

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2022-500023

(P2022-500023A)

(43) 公表日 令和4年1月4日(2022.1.4)

(51) Int.Cl.	F I	テーマコード (参考)
C 1 2 P 33/20 (2006.01)	C 1 2 P 33/20	2 B 0 3 0
C 1 2 N 5/10 (2006.01)	C 1 2 N 5/10 Z N A	4 B 0 5 0
C 1 2 P 13/00 (2006.01)	C 1 2 P 13/00	4 B 0 6 4
C 1 2 N 1/19 (2006.01)	C 1 2 N 1/19	4 B 0 6 5
C 1 2 N 1/13 (2006.01)	C 1 2 N 1/13	

審査請求 有 予備審査請求 未請求 (全 795 頁) 最終頁に続く

(21) 出願番号 特願2021-512835 (P2021-512835)  
 (86) (22) 出願日 令和1年9月5日 (2019.9.5)  
 (85) 翻訳文提出日 令和3年5月7日 (2021.5.7)  
 (86) 国際出願番号 PCT/IL2019/051000  
 (87) 国際公開番号 W02020/049572  
 (87) 国際公開日 令和2年3月12日 (2020.3.12)  
 (31) 優先権主張番号 16/123, 248  
 (32) 優先日 平成30年9月6日 (2018.9.6)  
 (33) 優先権主張国・地域又は機関 米国 (US)  
 (31) 優先権主張番号 268269  
 (32) 優先日 令和1年7月25日 (2019.7.25)  
 (33) 優先権主張国・地域又は機関 イスラエル (IL)

(71) 出願人 500018608  
 イエダ リサーチ アンド ディベロッ  
 メント カンパニー リミテッド  
 イスラエル国 7610002 レホボト  
 , ピーオー ボックス 95 アット  
 ワイツマン インスティテュート オブ  
 サイエンス  
 at the Weizmann Ins  
 titute of Science, P  
 O Box 95, 7610002 Re  
 hovot, Israel  
 (74) 代理人 110001379  
 特許業務法人 大島特許事務所

最終頁に続く

(54) 【発明の名称】 セルロース合成酵素様酵素及びその使用

(57) 【要約】

本開示は、セルロース合成酵素様 G (CSLG) 酵素の発現が増加または減少した遺伝子組換え細胞及び遺伝子組換え植物を提供する。本発明の遺伝子組換え細胞及び遺伝子組換え植物は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が、対応する非組換え細胞または非組換え植物と比較して増加または減少する。また本開示は、遺伝子組換え細胞内にステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法、植物またはその少なくとも一部の少なくとも1つの細胞における、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量を低減させる方法、並びに、植物またはその少なくとも一部の或る細胞におけるステロイド系サポニン、またはトリテルペノイド系サポニンの含有量を増加させる方法を提供する。

【選択図】 33 E

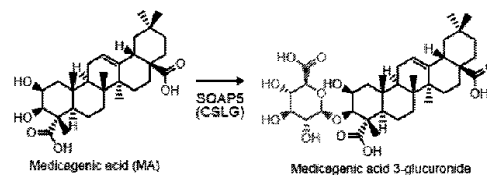


FIGURE 33E

## 【特許請求の範囲】

## 【請求項 1】

セルロース合成酵素様 G ( C S L G ) 酵素をコードする少なくとも 1 つの異種遺伝子の発現が、対応する非組換え細胞と比較して増加した遺伝子組換え細胞であって、

少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、

少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、

少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体

の含有量が、対応する非組換え細胞と比較して増加した、遺伝子組換え細胞。

10

## 【請求項 2】

請求項 1 に記載の遺伝子組換え細胞であって、

少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、

少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、

少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加し、

酵素をコードする少なくとも 1 つの追加の異種遺伝子をさらに発現し、

前記酵素は、サポニン - アミリン合成酵素、シトクロム P 4 5 0、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、UDP - グルコース 6 - デヒドロゲナーゼ 1、及びそれらの任意の組み合わせからなる群から選択される、遺伝子組換え細胞。

20

## 【請求項 3】

請求項 1 または 2 に記載の遺伝子組換え細胞であって、

前記少なくとも 1 つの異種遺伝子をコードする核酸配列は、

配列番号 3 0、3 2、3 4、3 6、3 8、4 0、6 5、8 0、9 3、9 5、9 7、9 9、1 0 1、1 0 3、または 1 0 5 のいずれか 1 つに記載の核酸配列であるか、または、

配列番号 3 0、3 2、3 4、3 6、3 8、4 0、6 5、8 0、9 3、9 5、9 7、9 9、1 0 1、1 0 3、または 1 0 5 のいずれか 1 つに記載の核酸配列に対して少なくとも 5

5 % の相同性及び少なくとも 5 5 % のカバレッジを有するそのホモログである、遺伝子組換え細胞。

30

## 【請求項 4】

請求項 1 または 2 に記載の遺伝子組換え細胞であって、

前記 C S L G 酵素のアミノ酸配列は、

配列番号 3 1、3 3、3 5、3 7、3 9、4 1、6 6、8 1、9 4、9 6、9 8、1 0 0、1 0 2、または 1 0 4 のいずれか 1 つに記載のアミノ酸配列であるか、または、

配列番号 3 1、3 3、3 5、3 7、3 9、4 1、6 6、8 1、9 4、9 6、9 8、1 0 0、1 0 2、または 1 0 4 のいずれか 1 つに記載のアミノ酸配列に対して少なくとも 5 5

% の相同性及び少なくとも 5 5 % のカバレッジを有するそのホモログである、遺伝子組換え細胞。

40

## 【請求項 5】

請求項 2 ~ 4 のいずれかに記載の遺伝子組換え細胞であって、

前記少なくとも 1 つの追加の異種遺伝子によりコードされる前記酵素、及び、該異種遺伝子をコードする核酸配列は、

( a ) サポニン - アミリン合成酵素、及び、配列番号 4 5 に記載の核酸配列、または、配列番号 4 5 に記載の核酸配列に対して少なくとも 8 0 % の相同性及び少なくとも 8 0 % のカバレッジを有するそのホモログ、

( b ) シトクロム P 4 5 0、及び、配列番号 4 6、5 1、または 5 3 のいずれか 1 つに記載の核酸配列、または、配列番号 4 6、5 1、または 5 3 のいずれか 1 つに記載の核酸

50

配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログ、

(c) グリコシルトランスフェラーゼ、及び、配列番号 55、57、59、または 61 のいずれか 1 つに記載の核酸配列、または、配列番号 55、57、59、または 61 のいずれか 1 つに記載の核酸配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログ、

(d) アセチルトランスフェラーゼ、及び、配列番号 63 に記載の核酸配列であるか、または、配列番号 63 に記載の核酸配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログ、

(e) UDP - グルコース 6 - デヒドロゲナーゼ 1、及び、配列番号 74 に記載の核酸配列であるか、または、配列番号 74 に記載の核酸配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログ、または、

(f) 上記の (a)、(b)、(c)、(d) 及び (e) の任意の組み合わせである、遺伝子組換え細胞。

【請求項 6】

請求項 2 ~ 4 のいずれかに記載の遺伝子組換え細胞であって、

前記少なくとも 1 つの追加の異種遺伝子によりコードされる前記酵素、及び、そのアミノ酸配列は、

(a) サポニン - アミリン合成酵素、及び、配列番号 48 に記載のアミノ酸配列、または配列番号 48 に記載のアミノ酸配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログ、

(b) シトクロム P 450、及び、配列番号 49、52、または 54 のいずれか 1 つに記載のアミノ酸配列、または、配列番号 49、52、または 54 のいずれか 1 つに記載のアミノ酸配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログ、

(c) グリコシルトランスフェラーゼ、及び、配列番号 56、58、60、または 62 のいずれか 1 つに記載のアミノ酸配列、または、配列番号 56、58、60、または 62 のいずれか 1 つに記載のアミノ酸配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログ、

(d) アセチルトランスフェラーゼ、及び、配列番号 64 に記載のアミノ酸配列、または配列番号 64 に記載のアミノ酸配列に対して少なくとも 80% の同一性及び 80% のカバレッジを有するそのホモログ、

(e) UDP - グルコース 6 - デヒドロゲナーゼ 1、及び、配列番号 75 に記載のアミノ酸配列、または配列番号 75 に記載のアミノ酸配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログ、または、

(f) 上記の (a)、(b)、(c)、(d) 及び (e) の任意の組み合わせである、遺伝子組換え細胞。

【請求項 7】

請求項 1 ~ 6 のいずれかに記載の遺伝子組換え細胞であって、

前記トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体は、甘味料、発泡剤、乳化剤、防腐剤、抗発癌剤、コレステロール低下剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、摂食阻害剤、抗真菌剤、またはそれらの任意の組み合わせを含む、遺伝子組換え細胞。

【請求項 8】

請求項 1 ~ 6 のいずれかに記載の遺伝子組換え細胞であって、

(a) 前記ステロイド系グリコアルカロイドは、エスクレオシド、デヒドロエスクレオシド、 - トマチン、トマチン、デヒドロトマチン、 - カコニン、 - ソラニン、 - ソラソニン、 - ソラマルギン、またはそれらの任意の組み合わせを含む、

(b) 前記ステロイド系サポニンは、ウットロシド B、トマトシド、またはそれらの任意の組み合わせを含む、

10

20

30

40

50

(c) 前記トリテルペノイド系サポニンは、メディカゲン酸 3 - O - グルクロニド (M A - 3 - G l c A) (化合物 6)、化合物 7、化合物 8、化合物 9、化合物 10、化合物 11、グリチルリチン (化合物 14)、グリチルレチン酸 3 - O - モノグルクロニド (化合物 15)、パヨゲニン (化合物 25)、パヨゲニン - ヘキサ A - ヘキサ - ヘキサ (化合物 31)、セルジャン酸 (化合物 26)、セルジャン酸 - ヘキサ A - ヘキサ (化合物 32)、ソヤサポゲノール A (化合物 29)、ソヤサポゲノール B (化合物 30)、ソヤサポゲノール A - ヘキサ A - ヘキサ - ペント (化合物 34)、ソヤサポニン V I (化合物 35)、ベタブルガロシド I V (化合物 33)、ヘデラゲニン - 3 G l c A、ギブソゲニン - 3 G l c A、ギブソゲン酸 - 3 G l c A、Q S - 2 1 アジュバント、またはそれらの任意の組み合わせを含む、

10

(d) 前記トリテルペノイド系サポニンの前記生合成中間体は、化合物 1、化合物 2、メディカゲン酸 (化合物 5)、オレアノール酸 (化合物 3)、オレアノール酸 - 3 G l c A、アウグスチン酸 (化合物 4)、アウグスチン酸 - 3 G l c A、グリチルレチン酸 (化合物 13)、またはそれらの任意の組み合わせを含む、または、

(e) 上記の (a)、(b)、(c)、及び (d) の任意の組み合わせである、遺伝子組換え細胞。

【請求項 9】

請求項 1 に記載の遺伝子組換え細胞であって、

フィトステロール、その誘導体、その代謝物、またはその生合成中間体、

コレステロール、その誘導体、その代謝物、またはその生合成中間体

20

フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、

コレスタノール、その誘導体、その代謝物、またはその生合成中間体、または、

フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体

の含有量が改変された、遺伝子組換え細胞。

【請求項 10】

請求項 1 ~ 9 のいずれかに記載の遺伝子組換え細胞であって、

前記細胞は、植物細胞、酵母細胞、または藻類細胞を含む、遺伝子組換え細胞。

【請求項 11】

請求項 10 に記載の遺伝子組換え細胞であって、

前記植物細胞は、イネ目、ナデシコ目、ナス目、マメ目、アオイ目、セリ目、フウチョウソウ目、クサスギカズラ目、ヤマノイモ目、またはユリ目の植物の細胞を含む、遺伝子組換え細胞。

30

【請求項 12】

請求項 11 に記載の遺伝子組換え細胞であって、

(a) 前記植物細胞がイネ目の植物の細胞を含む場合、前記植物は、イネ属、オオムギ属、カラスムギ属、及びコムギ属からなる群から選択される、

(b) 前記植物細胞がナデシコ目の植物の細胞を含む場合、前記植物は、ハウレンソウ属、アカザ属、ベータ属、レウム属、ドウカンソウ属、サボナリア属、及びジブソフィラ属からなる群から選択される、

(c) 前記植物細胞がナス目の植物の細胞を含む場合、前記植物は、ナス属、トウガラシ属、タバコ属、ヒヨス属、ダチュラ属、及びオオカミナスビ属からなる群から選択される、

40

(d) 前記植物細胞がマメ目の植物の細胞を含む場合、前記植物は、グリシリザ属、メディカゴ属、ダイズ属、ハス属、シセル属、インゲンマメ属、エンドウ属、ラッカセイ属、ルピナス属、及びアカシア属からなる群から選択される、

(e) 前記植物細胞がアオイ目の植物の細胞を含む場合、前記植物はカカオ属から選択される、

(f) 前記植物細胞がセリ目の植物の細胞を含む場合、前記植物は、ニンジン属、オランダミツバ属、パセリ属、オタネニンジン属、ミシマサイコ属、キツタ属、及びツボクサ属からなる群から選択される、または、

50

(g) 前記植物細胞がフウチョウソウ目の植物の細胞を含む場合、前記植物は、シロイヌナズナ属、アブラナ属、フウチョウボク属、及びパパイヤ属からなる群から選択される、遺伝子組換え細胞。

【請求項 13】

請求項 12 に記載の遺伝子組換え細胞であって、

(a) 前記植物細胞がナデシコ目の植物の細胞を含む場合、前記植物は、ホウレンソウ、ビートルート、及びキノアからなる群から選択される、

(b) 前記植物細胞がナス目の植物の細胞を含む場合、前記植物は、トマト、ワイルドトマト、ジャガイモ、ワイルドポテト、ナス、ペッパー、ピーマン、トウガラシ、チリペッパー、パプリカ、タバスコペッパー、グランドチェリー、タバコ、及びビタースイートからなる群から選択される、または、

(c) 前記植物細胞がマメ目の植物の細胞を含む場合、前記植物は、アルファルファ、大豆、ミヤコグサ、及び甘草からなる群から選択される、遺伝子組換え細胞。

【請求項 14】

請求項 13 に記載の遺伝子組換え細胞であって、

(a) 前記植物細胞は、

(i) - トマチン、その誘導体、その代謝物、またはその生合成中間体、トマチジン、その誘導体、その代謝物、またはその生合成中間体、デヒドロトマチン、その誘導体、その代謝物、またはその生合成中間体、または、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加したトマト植物の細胞であるか、または、

(ii) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、または、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変されたトマト植物の細胞である、

(b) 植物細胞は、

(i) - カコニン、その誘導体、その代謝物、またはその生合成中間体、 - ソラニン、その誘導体、その代謝物、またはその生合成中間体、または、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加したジャガイモ植物の細胞であるか、または、

(ii) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、または、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変されたジャガイモ植物の細胞である、または、

(c) 植物細胞は、

(i) - ソラソニン、その誘導体、その代謝物、またはその生合成中間体、 - ソラマルギン、その誘導体、その代謝物、またはその生合成中間体、または、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加したナス科植物の細胞であるか、または、

(ii) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変されたナス科植物の細胞である、遺伝子組換え細胞。

【請求項 15】

請求項 10 ~ 14 のいずれかに記載の遺伝子組換え細胞であって、

10

20

30

40

50

前記植物細胞は、葉の細胞、葉柄の細胞、植物の茎または幹の細胞、根の細胞、芽の細胞、塊茎の細胞、豆の細胞、穀物または穀粒の細胞、果実の細胞、木の実の細胞、マメの鞘の細胞、種子または種子の細胞、仮皮の細胞、包葉の細胞、または花の細胞を含む、遺伝子組換え細胞。

【請求項 16】

請求項 10 ~ 15 のいずれかに記載の遺伝子組換え細胞であって、

前記植物細胞は、植物またはその一部の細胞を含み、

前記植物の一部は、植物の葉、植物の葉柄、植物の茎または幹、植物の根、植物の芽、植物の塊茎、植物の豆、植物の穀物または穀粒、植物の果実、植物の木の实、植物のマメの鞘、植物の種子、植物の包葉、または植物の花を含む、遺伝子組換え細胞。

10

【請求項 17】

請求項 10 に記載の遺伝子組換え細胞であって、

前記酵母は、サッカロミセス属、シゾサッカロミセス属、ピチア属、ヤロウイア属、クリベロマイセス属、またはカンジダ属から選択される、遺伝子組換え細胞。

【請求項 18】

請求項 10 に記載の遺伝子組換え細胞であって、

前記藻類は、微細藻類、多細胞藻類、シアノバクテリア、珪藻類、緑藻植物（緑藻類）、紅藻植物（紅藻類）、褐藻植物（褐藻類）、ドナリエラ、クラミドモナス、またはヘマトコッカスから選択される、遺伝子組換え細胞。

20

【請求項 19】

セルロース合成酵素様 G (CSLG) 酵素をコードする少なくとも 1 つの CSLG 遺伝子の発現が、対応する非組換え植物における前記 CSLG 遺伝子の発現と比較して改変された少なくとも 1 つの細胞を含む遺伝子組換え植物であって、

少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、

少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、

少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体

の含有量が、対応する非組換え植物と比較して改変された、遺伝子組換え植物。

30

【請求項 20】

請求項 19 に記載の遺伝子組換え植物であって、

前記 CSLG 遺伝子は、内在性 CSLG 遺伝子を含む、遺伝子組換え植物。

【請求項 21】

請求項 20 に記載の遺伝子組換え植物であって、

前記内在性 CSLG 遺伝子は、突然変異を含み、

前記突然変異は、少なくとも 1 以上の点突然変異、挿入、欠失、またはそれらの任意の組み合わせを含み、

(a) 前記 CSLG 酵素は、安定性の向上、活性の増加、またはその両方を含み、前記含有量の改変は、少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が、対応する非組換え植物と比較して増加すること、または、

40

(b) 前記 CSLG 酵素は、安定性の低下、活性の低下、またはその両方を含み、前記含有量の改変は、少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が、対応する非組換え植物と比較して減少することである、遺伝子組換え植物。

50

## 【請求項 2 2】

請求項 2 0 に記載の遺伝子組換え植物であって、

( a ) 前記内在性 C S L G 遺伝子は、選択的にサイレンシングされるか、抑制されるか、または発現が低下し、前記含有量の改変は、少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が、対応する非組換え植物と比較して減少すること、または、

( b ) 前記内在性 C S L G 遺伝子は、過剰発現し、前記含有量の改変は、少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が、対応する非組換え植物と比較して増加することである、  
遺伝子組換え植物。

10

## 【請求項 2 3】

請求項 2 2 に記載の遺伝子組換え植物であって、

前記内在性 C S L G 遺伝子が選択的にサイレンシングされるか、抑制されるか、または発現が低下した場合、前記細胞は、前記 C S L G 遺伝子をコードするポリヌクレオチドを標的とする少なくとも 1 つのサイレンシング分子をさらに含み、

前記サイレンシング分子は、R N A 干渉分子またはアンチセンス分子から選択されるか、または、ウイルス誘導遺伝子サイレンシング系の構成要素である、  
遺伝子組換え植物。

20

## 【請求項 2 4】

請求項 1 9 に記載の遺伝子組換え植物であって、

前記 C S L G 遺伝子は、異種 C S L G 遺伝子を含み、

前記発現の改変は、前記異種遺伝子のデノボ発現を含む、  
遺伝子組換え植物。

## 【請求項 2 5】

請求項 2 4 に記載の遺伝子組換え植物であって、

前記発現の改変は、対応する非組換え細胞と比較して発現が増加することを含み、

前記含有量の改変は、少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して増加することを含む、  
遺伝子組換え植物。

30

## 【請求項 2 6】

請求項 1 9 ~ 2 5 のいずれかに記載の遺伝子組換え植物であって、

前記 C S L G 遺伝子をコードする核酸配列は、

配列番号 3 0、3 2、3 4、3 6、3 8、4 0、6 5、8 0、9 3、9 5、9 7、9 9、1 0 1、1 0 3、または 1 0 5 のいずれか 1 つに記載の核酸配列であるか、または、

配列番号 3 0、3 2、3 4、3 6、3 8、4 0、6 5、8 0、9 3、9 5、9 7、9 9、1 0 1、1 0 3、または 1 0 5 のいずれか 1 つに記載の核酸配列に対して少なくとも 5 5 % の相同性及び少なくとも 5 5 % のカバレッジを有するそのホモログである、  
遺伝子組換え植物。

40

## 【請求項 2 7】

請求項 1 9 ~ 2 5 のいずれかに記載の遺伝子組換え植物であって、

前記 C S L G 酵素のアミノ酸配列は、

配列番号 3 1、3 3、3 5、3 7、3 9、4 1、6 6、8 1、9 4、9 6、9 8、1 0 0、1 0 2、または 1 0 4 のいずれか 1 つに記載のアミノ酸配列であるか、または、

配列番号 3 1、3 3、3 5、3 7、3 9、4 1、6 6、8 1、9 4、9 6、9 8、1 0 0、1 0 2、または 1 0 4 のいずれか 1 つに記載のアミノ酸配列に対して少なくとも 5 5 % の相同性及び少なくとも 5 5 % のカバレッジを有するそのホモログである、  
遺伝子組換

50

え植物。

【請求項 28】

請求項 19 ~ 27 のいずれかに記載の遺伝子組換え植物であって、

前記少なくとも 1 つの CSLG 遺伝子の発現の改変は、

- (1) 前記核酸配列内に 1 以上の点突然変異を導入すること、
- (2) 前記核酸配列内に欠失を導入すること、
- (3) 前記核酸配列内に挿入を導入すること、または、
- (4) それらの任意の組み合わせを含み、

前記導入は、コード配列または非コード配列を突然変異させることを含む、遺伝子組換え植物。

10

【請求項 29】

請求項 28 に記載の遺伝子組換え植物であって、

前記少なくとも 1 つの CSLG 遺伝子の前記発現の改変は、前記少なくとも 1 つの CSLG 遺伝子を標的とするサイレンシング分子を導入することにより生じ、かつ、

(a) 前記 CSLG 遺伝子をコードする核酸配列が配列番号 30 に記載されている核酸配列である場合、サイレンシング分子は、配列番号 42 に記載の核酸配列またはその相補的配列であるか、または、配列番号 42 に記載の核酸配列またはその相補的配列に対して少なくとも 80 % の同一性及び少なくとも 80 % のカバレッジを有するそのホモログである、

20

(b) 前記 CSLG 遺伝子をコードする核酸配列が配列番号 34 に記載されている核酸配列である場合、サイレンシング分子は、配列番号 43 に記載の核酸配列またはその相補的配列であるか、または、配列番号 43 に記載の核酸配列またはその相補的配列に対して少なくとも 80 % の同一性及び少なくとも 80 % のカバレッジを有するそのホモログである、

(c) 前記 CSLG 遺伝子をコードする核酸配列が配列番号 38 に記載されている核酸配列である場合、サイレンシング分子は、配列番号 44 に記載の核酸配列またはその相補的配列であるか、または、配列番号 44 に記載の核酸配列またはその相補的配列に対して少なくとも 80 % の同一性及び少なくとも 80 % のカバレッジを有するそのホモログである、

30

(d) 前記 CSLG 遺伝子をコードする核酸配列が配列番号 65 または配列番号 93 に記載されている核酸配列である場合、サイレンシング分子は、配列番号 106 に記載の核酸配列またはその相補的配列であるか、または、配列番号 106 に記載の核酸配列またはその相補的配列に対して少なくとも 80 % の同一性及び少なくとも 80 % のカバレッジを有するそのホモログである、

(e) 前記 CSLG 遺伝子をコードする核酸配列が配列番号 95 に記載されている核酸配列である場合、サイレンシング分子は、配列番号 107 に記載の核酸配列またはその相補的配列であるか、または、配列番号 107 に記載の核酸配列またはその相補的配列に対して少なくとも 80 % の同一性及び少なくとも 80 % のカバレッジを有するそのホモログである、遺伝子組換え植物。

【請求項 30】

請求項 19 ~ 29 のいずれかに記載の遺伝子組換え植物であって、

前記トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体は、甘味料、発泡剤、乳化剤、防腐剤、抗発癌剤、コレステロール低下剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、摂食阻害剤、抗真菌剤、またはそれらの任意の組み合わせを含む、遺伝子組換え植物。

40

【請求項 31】

請求項 19 ~ 30 のいずれかに記載の遺伝子組換え植物であって、

(a) 前記ステロイド系グリコアルカロイドは、エスクレオシド、デヒドロエスクレオシド、 $\beta$ -トマチン、トマチン、デヒドロトマチン、 $\beta$ -カコニン、 $\beta$ -ソラニン、 $\beta$ -ソラソニン、 $\beta$ -ソラマルギン、またはそれらの任意の組み合わせを含む、

50

(b) 前記ステロイド系サポニン、ウットロシドB、トマトシド、またはそれらの任意の組み合わせを含む、

(c) 前記トリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GLCA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、パヨゲニン(化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサポゲノールA(化合物29)、ソヤサポゲノールB(化合物30)、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント(化合物34)、ソヤサポニンVI(化合物35)、ベタブルガロシドIV(化合物33)、ヘデラゲニン-3GLCA、ギブソゲニン-3GLCA、ギブソゲン酸-3GLCA、QS-21アジュバント、またはそれらの任意の組み合わせを含む、

(d) 前記トリテルペノイド系サポニンの前記生合成中間体は、化合物1、化合物2、メディカゲン酸(化合物5)、オレアノール酸(化合物3)、オレアノール酸-3GLCA、アウグスチン酸(化合物4)、アウグスチン酸-3GLCA、グリチルレチン酸(化合物13)、またはそれらの任意の組み合わせを含む、または、

(e) 上記の(a)、(b)、(c)、及び(d)の任意の組み合わせである、遺伝子組換え植物。

**【請求項32】**

請求項19に記載の遺伝子組換え植物であって、

フィトステロール、その誘導体、その代謝物、またはその生合成中間体、  
 コレステロール、その誘導体、その代謝物、またはその生合成中間体  
 フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、  
 コレスタノール、その誘導体、その代謝物、またはその生合成中間体、または、  
 フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体  
 の含有量が改変された、遺伝子組換え植物。

**【請求項33】**

請求項19～32のいずれかに記載の遺伝子組換え植物であって、

前記植物は、イネ目、ナデシコ目、ナス目、マメ目、アオイ目、セリ目、フウチョウソウ目、クサスギカズラ目、ヤマノイモ目、またはユリ目から選択される、遺伝子組換え植物。

**【請求項34】**

請求項33に記載の遺伝子組換え植物であって、

(a) 前記植物がイネ目の植物を含む場合、前記植物は、イネ属、オオムギ属、カラスムギ属、及びコムギ属からなる群から選択される、

(b) 前記植物がナデシコ目の植物を含む場合、前記植物は、ハウレンソウ属、アカザ属、ベータ属、レウム属、ドウカンソウ属、サポナリア属、及びジブソフィラ属からなる群から選択される、

(c) 前記植物がナス目の植物を含む場合、前記植物は、ナス属、トウガラシ属、タバコ属、ヒヨス属、ダチュラ属、及びオオカミナスビ属からなる群から選択される、

(d) 前記植物がマメ目の植物を含む場合、前記植物は、グリシリザ属、メディカゴ属、ダイズ属、ハス属、シセル属、インゲンマメ属、エンドウ属、ラッカセイ属、ルピナス属、及びアカシア属からなる群から選択される、

(e) 前記植物がアオイ目の植物を含む場合、前記植物はカカオ属から選択される、

(f) 前記植物細胞がセリ目の植物の細胞を含む場合、前記植物は、ニンジン属、オランダミツバ属、パセリ属、オタネニンジン属、ミシマサイコ属、キツタ属、及びツボクサ属からなる群から選択される、または、

(g) 前記植物がフウチョウソウ目の植物を含む場合、前記植物は、シロイヌナズナ属、アブラナ属、フウチョウボク属、及びパパイヤ属からなる群から選択される、遺伝子組換え植物。

10

20

30

40

50

## 【請求項 3 5】

請求項 3 4 に記載の遺伝子組換え植物であって、

( a ) 前記植物がナデシコ目の植物を含む場合、前記植物は、ホウレンソウ、ビートルート、及びキノアからなる群から選択される、

( b ) 前記植物がナス目の植物を含む場合、前記植物は、トマト、ワイルドトマト、ジャガイモ、ワイルドポテト、ナス、ペッパー、ピーマン、トウガラシ、チリペッパー、パプリカ、タバスコペッパー、グランドチェリー、タバコ、及びビタースイートからなる群から選択される、または、

( c ) 前記植物がマメ目の植物を含む場合、前記植物は、アルファルファ、大豆、ミヤコグサ、及び甘草からなる群から選択される、遺伝子組換え植物。

10

## 【請求項 3 6】

請求項 3 5 に記載の遺伝子組換え植物であって、

( a ) 前記植物は、

( i ) - トマチン、その誘導体、その代謝物、またはその生合成中間体、トマチジン、その誘導体、その代謝物、またはその生合成中間体、デヒドロトマチン、その誘導体、その代謝物、またはその生合成中間体、または、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加したトマト植物であるか、または、

( i i ) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、または、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変されたトマト植物である、

20

( b ) 前記植物は、

( i ) - カコニン、その誘導体、その代謝物、またはその生合成中間体、 - ソラニン、その誘導体、その代謝物、またはその生合成中間体、または、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加したジャガイモ植物であるか、または、

( i i ) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、または、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変されたジャガイモ植物である、または、

30

( c ) 前記植物は、

( i ) - ソラソニン、その誘導体、その代謝物、またはその生合成中間体、 - ソラマルギン、その誘導体、その代謝物、またはその生合成中間体、または、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加したナス科植物であるか、または、

( i i ) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、または、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変されたナス科植物である、遺伝子組換え植物。

40

## 【請求項 3 7】

請求項 2 0 ~ 3 6 のいずれかに記載の遺伝子組換え植物であって、

前記植物細胞は、葉の細胞、葉柄の細胞、植物の茎または幹の細胞、根の細胞、芽の細胞、塊茎の細胞、豆の細胞、穀物または穀粒の細胞、果実の細胞、木の実の細胞、マメの鞘の細胞、種子または種子の細胞、仮皮の細胞、包葉の細胞、または花の細胞を含む、遺伝子組換え植物。

50

## 【請求項 38】

遺伝子組換え細胞内にステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法であって、

(a) 前記細胞内に少なくとも1つの異種遺伝子を導入する導入ステップであって、前記少なくとも1つの異種遺伝子がセルロース合成酵素様 G (CSLG) 酵素をコードしており、前記少なくとも1つの異種遺伝子が任意選択でベクターに含まれている、該ステップと、

(b) 前記細胞内で前記少なくとも1つの異種遺伝子を発現させる発現ステップと、を含み、

前記細胞は、少なくとも1つのステロイド系アルカロイド、少なくとも1つのステロイド系サポニン、または少なくとも1つのトリテルペノイド系サポニンの含有量が、対応する非組換え細胞と比較して増加する、方法。

10

## 【請求項 39】

請求項 38 に記載の方法であって、

前記導入ステップは、

前記細胞内に少なくとも1つの追加の異種遺伝子を導入するステップであって、前記少なくとも1つの追加の異種遺伝子が、サポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、UDP - グルコース 6 - デヒドロゲナーゼ 1、及びそれらの任意の組み合わせからなる群から選択される酵素をコードしており、前記少なくとも1つの追加の異種遺伝子が任意選択でベクターに含まれている、該ステップと、

20

前記細胞内で前記少なくとも1つの異種遺伝子を発現させるステップと、をさらに含む、方法。

## 【請求項 40】

請求項 38 または 39 に記載の方法であって、

(a) 前記少なくとも1つの異種遺伝子が、プロモータ、転写終結配列、またはそれらの組み合わせに動作可能に連結されているか、

(b) 前記少なくとも1つの追加の異種遺伝子が、プロモータ、転写終結配列、またはそれらの組み合わせに動作可能に連結されているか、または、

(c) 上記の (a) と (b) の組み合わせである、方法。

30

## 【請求項 41】

請求項 38 ~ 40 のいずれかに記載の方法であって、

前記導入ステップは、

前記少なくとも1つの細胞を、

(a) 前記少なくとも1つの異種遺伝子、前記少なくとも1つの異種遺伝子をコードするポリヌクレオチド配列、または前記少なくとも1つの異種遺伝子を含むベクター、

(b) 前記少なくとも1つの追加の異種遺伝子、前記少なくとも1つの追加の異種遺伝子をコードするポリヌクレオチド配列、または前記少なくとも1つの追加の異種遺伝子を含むベクター、または、

(c) 上記の (a) と (b) の組み合わせを用いて形質転換することを含み、

40

前記発現は、一過性発現または構成的発現を含む、方法。

## 【請求項 42】

請求項 38 ~ 41 のいずれかに記載の方法であって、

前記少なくとも1つの異種 CSLG 遺伝子をコードする核酸配列は、

配列番号 30、32、34、36、38、40、65、80、93、95、97、99、101、103、または 105 のいずれか 1 つに記載の核酸配列であるか、または、

配列番号 30、32、34、36、38、40、65、80、93、95、97、99、101、103、または 105 のいずれか 1 つに記載の核酸配列に対して少なくとも 5

5% の相同性及び少なくとも 55% のカバレッジを有するそのホモログである、方法。

## 【請求項 43】

50

請求項 38 ~ 42 のいずれかに記載の方法であって、  
 前記コードされた C S L G 酵素のアミノ酸配列は、  
 配列番号 31、33、35、37、39、41、66、81、94、96、98、100、102、または 104 のいずれか 1 つに記載のアミノ酸配列であるか、または、  
 配列番号 31、33、35、37、39、41、66、81、94、96、98、100、102、または 104 のいずれか 1 つに記載のアミノ酸配列に対して少なくとも 55 % の相同性及び少なくとも 55 % のカバレッジを有するそのホモログである、方法。

【請求項 44】

請求項 38 ~ 43 のいずれかに記載の方法であって、  
 前記少なくとも 1 つの追加の異種遺伝子によりコードされる前記酵素、及び、該異種遺伝子をコードする核酸配列は、

(a) サポニン - アミリン合成酵素、及び、配列番号 45 に記載の核酸配列、または、配列番号 45 に記載の核酸配列に対して少なくとも 80 % の相同性及び少なくとも 80 % のカバレッジを有するそのホモログ、

(b) シトクロム P 450、及び、配列番号 46、51、または 53 のいずれか 1 つに記載の核酸配列、または、配列番号 46、51、または 53 のいずれか 1 つに記載の核酸配列に対して少なくとも 80 % の相同性及び少なくとも 80 % のカバレッジを有するそのホモログ、

(c) グリコシルトランスフェラーゼ、及び、配列番号 55、57、59、または 61 のいずれか 1 つに記載の核酸配列、または、配列番号 55、57、59、または 61 のいずれか 1 つに記載の核酸配列に対して少なくとも 80 % の相同性及び少なくとも 80 % のカバレッジを有するそのホモログ、

(d) アセチルトランスフェラーゼ、及び、配列番号 63 に記載の核酸配列であるか、または、配列番号 63 に記載の核酸配列に対して少なくとも 80 % の相同性及び少なくとも 80 % のカバレッジを有するそのホモログ、

(e) U D P - グルコース 6 - デヒドロゲナーゼ 1、及び、配列番号 74 に記載の核酸配列であるか、または、配列番号 74 に記載の核酸配列に対して少なくとも 80 % の相同性及び少なくとも 80 % のカバレッジを有するそのホモログ、または、

(f) 上記の (a)、(b)、(c)、(d) 及び (e) の任意の組み合わせである、方法。

【請求項 45】

請求項 38 ~ 43 のいずれかに記載の方法であって、  
 前記少なくとも 1 つの追加の異種遺伝子によりコードされる前記酵素、及び、そのアミノ酸配列は、

(a) サポニン - アミリン合成酵素、及び、配列番号 48 に記載のアミノ酸配列、または配列番号 48 に記載のアミノ酸配列に対して少なくとも 80 % の相同性及び少なくとも 80 % のカバレッジを有するそのホモログ、

(b) シトクロム P 450、及び、配列番号 49、52、または 54 のいずれか 1 つに記載のアミノ酸配列、または、配列番号 49、52、または 54 のいずれか 1 つに記載のアミノ酸配列に対して少なくとも 80 % の相同性及び少なくとも 80 % のカバレッジを有するそのホモログ、

(c) グリコシルトランスフェラーゼ、及び、配列番号 56、58、60、または 62 のいずれか 1 つに記載のアミノ酸配列、または、配列番号 56、58、60、または 62 のいずれか 1 つに記載のアミノ酸配列に対して少なくとも 80 % の相同性及び少なくとも 80 % のカバレッジを有するそのホモログ、

(d) アセチルトランスフェラーゼ、及び、配列番号 64 に記載のアミノ酸配列、または配列番号 64 に記載のアミノ酸配列に対して少なくとも 80 % の相同性及び 80 % のカバレッジを有するそのホモログ、

(e) U D P - グルコース 6 - デヒドロゲナーゼ 1、及び、配列番号 75 に記載のアミノ酸配列、または配列番号 75 に記載のアミノ酸配列に対して少なくとも 80 % の相同性

及び少なくとも 80% のカバレッジを有するそのホモログ、または、

(f) 上記の (a)、(b)、(c)、(d) 及び (e) の任意の組み合わせである、方法。

【請求項 46】

請求項 38 ~ 45 のいずれかに記載の方法であって、

(a) 前記トリテルペノイド系サポニンは、甘味料、発泡剤、乳化剤、防腐剤、抗発癌剤、コレステロール低下剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、摂食阻害剤、抗真菌剤、またはそれらの任意の組み合わせを含む、

(b) 前記ステロイド系グリコアルカロイドは、エスクレオシド、デヒドロエスクレオシド、  
- トマチン、トマチン、デヒドロトマチン、  
- カコニン、  
- ソラニン、  
- ソラソニン、  
- ソラマルギン、またはそれらの任意の組み合わせを含む、

(c) 前記ステロイド系サポニンは、ウットロシド B、トマトシド、またはそれらの任意の組み合わせを含む、

(d) 前記トリテルペノイド系サポニンは、メディカゲン酸 3-O-グルクロニド (M A - 3 - G l c A) (化合物 6)、化合物 7、化合物 8、化合物 9、化合物 10、化合物 11、グリチルリチン (化合物 14)、グリチルレチン酸 3-O-モノグルクロニド (化合物 15)、パヨゲニン (化合物 25)、パヨゲニン-ヘキサ A-ヘキサ-ヘキサ (化合物 31)、セルジャン酸 (化合物 26)、セルジャン酸-ヘキサ A-ヘキサ (化合物 32)、ソヤサボゲノール A (化合物 29)、ソヤサボゲノール B (化合物 30)、ソヤサボゲノール A-ヘキサ A-ヘキサ-ペント (化合物 34)、ソヤサポニン V I (化合物 35)、ベタブルガロシド I V (化合物 33)、ヘデラゲニン-3 G l c A、ギブソゲニン-3 G l c A、ギブソゲン酸-3 G l c A、Q S - 2 1 アジュバント、またはそれらの任意の組み合わせを含む、または、

(e) 上記の (a)、(b)、(c)、及び (d) の任意の組み合わせである、方法。

【請求項 47】

請求項 38 ~ 46 のいずれかに記載の方法であって、

フィトステロール、その誘導体、その代謝物、またはその生合成中間体、

コレステロール、その誘導体、その代謝物、またはその生合成中間体

フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、

コレスタノール、その誘導体、その代謝物、またはその生合成中間体、または、

フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体

の含有量を改変するステップをさらに含む、方法。

【請求項 48】

請求項 38 ~ 47 のいずれかに記載の方法であって、

前記細胞は、植物細胞、酵母細胞、または藻類細胞を含む、方法。

【請求項 49】

請求項 48 に記載の方法であって、

(a) 前記植物細胞は、イネ目、ナデシコ目、ナス目、マメ目、アオイ目、セリ目、フウチョウソウ目、クサスギカズラ目、ヤマノイモ目、またはユリ目から選択される植物の細胞を含む、

(b) 前記酵母は、サッカロミセス属、シゾサッカロミセス属、ピチア属、ヤロウイア属、クリベロマイセス属、またはカンジダ属から選択される、または、

(c) 前記藻類は、微細藻類、多細胞藻類、シアノバクテリア、珪藻類、緑藻植物 (緑藻類)、紅藻植物 (紅藻類)、褐藻植物 (褐藻類)、ドナリエラ、クラミドモナス、またはヘマトコッカスから選択される、方法。

【請求項 50】

請求項 49 に記載の方法であって、

(a) 前記植物細胞がイネ目の植物の細胞を含む場合、前記植物は、イネ属、オオムギ属、カラスムギ属、及びコムギ属からなる群から選択される、

(b) 前記植物細胞がナデシコ目の植物の細胞を含む場合、前記植物は、ハウレンソウ

属、アカザ属、ベータ属、レウム属、ドウカンソウ属、サボナリア属、及びジブソフィラ属からなる群から選択される、

(c) 前記植物細胞がナス目の植物の細胞を含む場合、前記植物は、ナス属、トウガラシ属、タバコ属、ヒヨス属、ダチュラ属、及びオオカミナスビ属からなる群から選択される、

(d) 前記植物細胞がマメ目の植物の細胞を含む場合、前記植物は、グリシリザ属、メディカゴ属、ダイズ属、ハス属、シセル属、インゲンマメ属、エンドウ属、ラッカセイ属、ルピナス属、及びアカシア属からなる群から選択される、

(e) 前記植物細胞がアオイ目の植物の細胞を含む場合、前記植物はカカオ属から選択される、

(f) 前記植物細胞がセリ目の植物の細胞を含む場合、前記植物は、ニンジン属、オランダミツバ属、パセリ属、オタネニンジン属、ミシマサイコ属、キツタ属、及びツボクサ属からなる群から選択される、または、

(g) 前記植物細胞がフウチョウソウ目の植物の細胞を含む場合、前記植物は、シロイヌナズナ属、アブラナ属、フウチョウボク属、及びパイヤ属からなる群から選択される、方法。

【請求項 5 1】

請求項 5 0 に記載の方法であって、

(a) 前記植物細胞がナデシコ目の植物の細胞を含む場合、前記植物は、ハウレンソウ、ビートルート、及びキノアからなる群から選択される、

(b) 前記植物細胞がナス目の植物の細胞を含む場合、前記植物は、トマト、ワイルドトマト、ジャガイモ、ワイルドポテト、ナス、ペッパー、ピーマン、トウガラシ、チリペッパー、パプリカ、タバスコペッパー、グランドチェリー、タバコ、及びビタースイートからなる群から選択される、または、

(c) 前記植物細胞がマメ目の植物の細胞を含む場合、前記植物は、アルファルファ、大豆、ミヤコグサ、及び甘草からなる群から選択される、方法。

【請求項 5 2】

植物またはその少なくとも一部の少なくとも 1 つの細胞における、少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を低減させる方法であって、

前記少なくとも 1 つの細胞の遺伝子を組み換える遺伝子組換えステップを含み、

前記遺伝子組換えステップは、

(a) セルロース合成酵素様 G (CSLG) 酵素をコードする少なくとも 1 つの遺伝子を標的とする少なくとも 1 つのサイレンシング分子を用いて前記少なくとも 1 つの植物細胞を形質転換するステップ、または、

(b) 前記 CSG 酵素をコードする前記少なくとも 1 つの遺伝子の核酸配列を突然変異させるステップであって、前記突然変異が、(1) 前記核酸配列内に 1 以上の点突然変異を導入すること、(2) 前記核酸配列内に欠失を導入すること、(3) 前記核酸配列内に挿入を導入すること、または、(4) それらの任意の組み合わせを含み、前記導入が、コード配列または非コード配列を突然変異させることを含む、該ステップと、を含み、

遺伝子が組み換えられた遺伝子組換え細胞における前記 CSG 酵素をコードする遺伝子の発現は、対応する非組換え植物の細胞と比較して減少し、

前記遺伝子組換え細胞を含む植物は、少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して減少する、方法。

【請求項 5 3】

10

20

30

40

50

請求項 5 2 に記載の方法であって、  
前記 C S L G 酵素をコードする前記少なくとも 1 つの遺伝子の前記核酸配列は、  
配列番号 3 0、3 2、3 4、3 6、3 8、4 0、6 5、8 0、9 3、9 5、9 7、9 9  
、1 0 1、1 0 3、または 1 0 5 のいずれか 1 つに記載の核酸配列であるか、または、  
配列番号 3 0、3 2、3 4、3 6、3 8、4 0、6 5、8 0、9 3、9 5、9 7、9 9  
、1 0 1、1 0 3、もしくは 1 0 5 のいずれか 1 つに記載の核酸配列に対して少なくとも  
5 5 % の同一性及び少なくとも 5 5 % のカバレッジを有するそのホモログである、方法。

【請求項 5 4】

請求項 5 2 または 5 3 に記載の方法であって、

( a ) 前記 C S L G 酵素をコードする前記少なくとも 1 つの遺伝子の前記核酸配列が配  
列番号 3 0 に記載されている核酸配列である場合、サイレンシング分子は、配列番号 4 2  
に記載の核酸配列またはその相補的配列であるか、または、配列番号 4 2 に記載の核酸配  
列またはその相補的配列に対して少なくとも 8 0 % の同一性及び少なくとも 8 0 % のカバ  
レッジを有するそのホモログである、

( b ) 前記 C S L G 酵素をコードする前記少なくとも 1 つの遺伝子の前記核酸配列が配  
列番号 3 4 に記載されている核酸配列である場合、サイレンシング分子は、配列番号 4 3  
に記載の核酸配列またはその相補的配列であるか、または、配列番号 4 3 に記載の核酸配  
列またはその相補的配列に対して少なくとも 8 0 % の同一性及び少なくとも 8 0 % のカバ  
レッジを有するそのホモログである、

( c ) 前記 C S L G 酵素をコードする前記少なくとも 1 つの遺伝子の前記核酸配列が配  
列番号 3 8 に記載されている核酸配列である場合、サイレンシング分子は、配列番号 4 4  
に記載の核酸配列またはその相補的配列であるか、または、配列番号 4 4 に記載の核酸配  
列またはその相補的配列に対して少なくとも 8 0 % の同一性及び少なくとも 8 0 % のカバ  
レッジを有するそのホモログである、

( d ) 前記 C S L G 酵素をコードする前記少なくとも 1 つの遺伝子の前記核酸配列が配  
列番号 6 5 または配列番号 9 3 に記載されている核酸配列である場合、サイレンシング分  
子は、配列番号 1 0 6 に記載の核酸配列またはその相補的配列であるか、または、配列番  
号 1 0 6 に記載の核酸配列またはその相補的配列に対して少なくとも 8 0 % の同一性及び  
少なくとも 8 0 % のカバレッジを有するそのホモログである、

( e ) 前記 C S L G 酵素をコードする前記少なくとも 1 つの遺伝子の前記核酸配列が配  
列番号 9 5 に記載されている核酸配列である場合、サイレンシング分子は、配列番号 1 0  
7 に記載の核酸配列またはその相補的配列であるか、または、配列番号 1 0 7 に記載の核  
酸配列またはその相補的配列に対して少なくとも 8 0 % の同一性及び少なくとも 8 0 % の  
カバレッジを有するそのホモログである、方法。

【請求項 5 5】

請求項 5 2 ~ 5 4 のいずれかに記載の方法であって、

( a ) 前記ステロイド系アルカロイド、前記ステロイド系サポニン、または前記トリテ  
ルペノイド系サポニンは、毒素または苦味化合物を含むか、ホルモン模倣特性を有するか  
、またはそれらの組み合わせであり、

( b ) 前記ステロイド系グリコアルカロイドは、エスクレオシド、デヒドロエスクレオ  
シド、  
- トマチン、トマチン、デヒドロトマチン、  
- カコニン、  
- ソラニン、  
- ソラソニン、  
- ソラマルギン、またはそれらの任意の組み合わせを含む、

( c ) 前記ステロイド系サポニンは、ウットロシド B、トマトシド、またはそれらの任  
意の組み合わせを含む、

( d ) 前記トリテルペノイド系サポニンは、メディカゲン酸 3 - O - グルクロニド ( M  
A - 3 - G l c A ) ( 化合物 6 )、化合物 7、化合物 8、化合物 9、化合物 1 0、化合物  
1 1、グリチルリチン ( 化合物 1 4 )、グリチルレチン酸 3 - O - モノグルクロニド ( 化  
合物 1 5 )、パヨゲニン ( 化合物 2 5 )、パヨゲニン - ヘキサ A - ヘキサ - ヘキサ ( 化  
合物 3 1 )、セルジャン酸 ( 化合物 2 6 )、セルジャン酸 - ヘキサ A - ヘキサ ( 化合物 3 2  
 )、ソヤサボゲノール A ( 化合物 2 9 )、ソヤサボゲノール B ( 化合物 3 0 )、ソヤサボ

10

20

30

40

50

ゲノール A - ヘキサ A - ヘキサ - ペント (化合物 34)、ソヤサポニン V I (化合物 35)、ベタブルガロシド I V (化合物 33)、ヘデラゲニン - 3 G l c A、ギブソゲニン - 3 G l c A、ギブソゲン酸 - 3 G l c A、Q S - 2 1 アジュバント、またはそれらの任意の組み合わせを含む、

(e) 前記トリテルペノイド系サポニンの前記生合成中間体は、化合物 1、化合物 2、メディカゲン酸 (化合物 5)、オレアノール酸 (化合物 3)、オレアノール酸 - 3 G l c A、アウグスチン酸 (化合物 4)、アウグスチン酸 - 3 G l c A、グリチルレチン酸 (化合物 13)、またはそれらの任意の組み合わせを含む、または、

(f) 上記の (a)、(b)、(c)、(d) 及び (e) の任意の組み合わせである、方法。

10

【請求項 56】

請求項 52 ~ 55 のいずれかに記載の方法であって、  
 フィトステロール、その誘導体、その代謝物、またはその生合成中間体、  
 コレステロール、その誘導体、その代謝物、またはその生合成中間体  
 フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、  
 コレスタノール、その誘導体、その代謝物、またはその生合成中間体、または、  
 フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体  
 の含有量を改変するステップをさらに含む、方法。

【請求項 57】

植物またはその少なくとも一部の少なくとも 1 つの細胞における、少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法であって、

20

前記少なくとも 1 つの細胞の遺伝子を組み換える遺伝子組換えステップを含み、  
 前記遺伝子組換えステップは、

(a) セルロース合成酵素様 G (CSLG) 酵素をコードする少なくとも 1 つの遺伝子の核酸配列を突然変異させるステップであって、前記突然変異が、(1) 前記核酸配列内に 1 以上の点突然変異を導入すること、(2) 前記核酸配列内に欠失を導入すること、(3) 前記核酸配列内に挿入を導入すること、または、(4) それらの任意の組み合わせを含み、前記導入が、コード配列または非コード配列を突然変異させることを含む、該ステップと、を含み、

30

(b) 前記 CSLG 酵素をコードする前記核酸配列を発現させるステップと、を含み、  
 遺伝子が組み換えられた遺伝子組換え細胞を含む植物は、少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物に比較して増加する、方法。

【請求項 58】

請求項 57 に記載の方法であって、

40

前記 CSLG 酵素をコードする前記少なくとも 1 つの遺伝子の前記核酸配列は、  
 配列番号 30、32、34、36、38、40、65、80、93、95、97、99、  
 101、103、または 105 のいずれか 1 つに記載の核酸配列であるか、または、  
 配列番号 30、32、34、36、38、40、65、80、93、95、97、99、  
 101、103、もしくは 105 のいずれか 1 つに記載の核酸配列に対して少なくとも 55% の相同性及び少なくとも 55% のカバレッジを有するそのホモログである、方法。

【請求項 59】

請求項 57 に記載の方法であって、

前記 CSLG 酵素のアミノ酸配列は、

配列番号 31、33、35、37、39、41、66、81、94、96、98、10

50

0、102、または104のいずれか1つに記載のアミノ酸配列であるか、または、  
配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列に対して少なくとも55%の相同性及び少なくとも55%のカバレッジを有するそのホモログである、方法。

【請求項60】

請求項57～59のいずれかに記載の方法であって、

(a) 前記ステロイド系アルカロイド、前記ステロイド系サポニン、または前記トリテルペノイド系サポニンは、甘味料、発泡剤、乳化剤、防腐剤、抗発癌剤、コレステロール低下剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、摂食阻害剤、抗真菌剤、またはそれらの任意の組み合わせを含む、

(b) 前記ステロイド系グリコアルカロイドは、エスクレオシド、デヒドロエスクレオシド、  
-トマチン、トマチン、デヒドロトマチン、  
-カコニン、  
-ソラニン、  
-ソラソニン、  
-ソラマルギン、またはそれらの任意の組み合わせを含む、

(c) 前記ステロイド系サポニンは、ウットロシドB、トマトシド、またはそれらの任意の組み合わせを含む、

(d) 前記トリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GlcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、パヨゲニン(化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサポゲノールA(化合物29)、ソヤサポゲノールB(化合物30)、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント(化合物34)、ソヤサポニンVI(化合物35)、ベタブルガロシドIV(化合物33)、ヘデラゲニン-3GlcA、ギブソゲニン-3GlcA、ギブソゲン酸-3GlcA、QS-21アジュバント、またはそれらの任意の組み合わせを含む、

(e) 前記トリテルペノイド系サポニンの前記生合成中間体は、化合物1、化合物2、メディカゲン酸(化合物5)、オレアノール酸(化合物3)、オレアノール酸-3GlcA、アウグスチン酸(化合物4)、アウグスチン酸-3GlcA、グリチルレチン酸(化合物13)、またはそれらの任意の組み合わせを含む、または、

(f) 上記の(a)、(b)、(c)、(d)及び(e)の任意の組み合わせである、方法。

【請求項61】

請求項57～60のいずれかに記載の方法であって、

(a) 前記遺伝子組換え細胞における、前記CSLG酵素をコードする前記少なくとも1つの遺伝子の発現は、対応する非組換え植物の細胞と比較して増加する、

(b) 前記遺伝子組換え細胞における、前記CSLG酵素の活性は、対応する非組換え植物細胞と比較して増加する、

(c) 前記遺伝子組換え細胞における、前記CSLG酵素の安定性は、対応する非組換え植物の細胞と比較して向上する、または、

(d) 上記の(a)、(b)、及び(c)の任意の組み合わせである、方法。

【請求項62】

請求項57～61のいずれかに記載の方法であって、

フィトステロール、その誘導体、その代謝物、またはその生合成中間体、

コレステロール、その誘導体、その代謝物、またはその生合成中間体

フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、

コレスタノール、その誘導体、その代謝物、またはその生合成中間体、または、

フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体

の含有量を改変するステップをさらに含む、方法。

【発明の詳細な説明】

【技術分野】

10

20

30

40

50

## 【0001】

(関連出願の相互参照)

本出願は、米国特許仮出願第61/831、164(2013年6月5日出願)に基づく優先権を主張する、PCT国際出願第PCT/IL2014/050497号(国際出願日は2014年6月2日)の国内段階出願として出願された、米国特許出願第14/895、059号(2015年12月1日出願)の継続部分出願である、米国特許出願第16/123、248号(2018年9月6日出願)に基づく優先権を主張するものである。上記の各出願の開示内容は、参照により本明細書中に援用される。また、本出願は、イスラエル国特許出願第268269号(2019年7月25日出願)に基づく優先権も主張するものである。上記の出願の開示内容は、参照により本明細書中に援用される。

10

## 【背景技術】

## 【0002】

植物界では、何十万種類もの低化合物が産生されており、その多くは属や科に固有のものである。30万種以上の植物の特異的代謝物(SM)の膨大な構造多様性は、それらのコア構造の様々な修飾により生じる。植物のSMは最も頻繁にグリコシル化され、このグリコシル化は、特異的代謝物(SM)の区画化、活性、溶解性、安定性、及び毒性に大きな影響を与える。特異的代謝物(SM)は、低分子量の窒素含有有機化合物であり、一般的には複素環構造を有する。

## 【0003】

サポニン-アルカロイド化合物の広範なグループは、植物界に広く分布し、サイトゾルのメバロン酸イソプレノイド生合成経路に由来する。サポニン及びステロイドアルカロイド(SA)は、植物が産生する特異的代謝物(SM)の2つの大きなクラスである。

20

## 【0004】

サポニンは、無数の植物種と100以上の科に見られる特異的代謝物(SM)の大きなグループであり、1以上のグリコシド部分で修飾された親油性トリテルペノイドまたはステロイド骨格(アグリコン)を表す。

## 【0005】

サポニンは、その名前(ラテン語で石鹸を意味する)が示唆しているように、両親媒性構造を有するため、水溶液中では石鹸のような性質を示す。サポニンのユニークな物理化学的特性は、植物界において、抗真菌性、抗菌性及び殺虫性を含む幅広い機能を提供する。サポニンの活性は、アグリコンの構造に依存するが、多くの場合、結合した糖にさらに依存する。例えば、オレアノール酸やメジカゲン酸(MA)由来のサポニンは、C-28位にカルボキシルが存在することに起因して溶血作用を示すが、アベナシンA-1は、殺菌性を示すためには、C-3位にL-アラビノースが結合していることを必要とする。さらに、コレステロール由来のステロイドサポニンは、植物界に広く存在し、その構造は非常に多様であり、C-5、6位が飽和しているもの(例えば、サラサボゲニン)もあるし、C-5、6位が不飽和のもの(例えば、ジオスゲニン)もある。

30

## 【0006】

多くの報告が、抗炎症及び免疫増強(アジュバント)から抗癌特性に至るまでの、ヒトの健康及び医療用途に対するサポニンの利点を強調している。糖活性酵素(CAZY)糖転移酵素1(GT1)スーパーファミリーに属するUDP糖転移酵素(UGT)メンバーは、サポニン及び他のすべての特異的代謝物(SM)クラスにおいて糖転移反応を行う。サポニンを装飾する最も一般的な糖類としては、D-グルコース、D-ガラクトース、L-アラビノース、D-グルクロン酸、D-キシロース、L-ラムノース、及びD-フコースが挙げられる。

40

## 【0007】

サボゲニンのC-3位に結合したグルクロン酸の存在は、ナデシコ目の種で特に一般的であるが、それに関与する酵素は依然として不明である。

## 【0008】

したがって、トリテルペノイドサポニンを含むサポニン類を産生するための代謝経路に

50

関与する酵素に関する知識、例えば、キレイン酸などにグルクロン酸を結合させるセルロース合成酵素様 G に関する知識に対する満たされていないニーズが依然として存在する。上記の知識が得られれば、自然に産生される産物の含有量を改変させるために、植物において鍵となる酵素を制御することができる。例えば、酵素を過剰に発現させたり、安定性を高めたりすることによって、天然の甘味料を大量に産生することができる。また、酵素を遺伝子レベルで下方制御することによって、植物産物（例えば、キノア）に苦味を加える化合物の産生を減少させることができる。さらに、高価値のサポニン（これに限定しないが、例えば、トリテルペノイドサポニン）を商業量で提供するために、異種系でこれらのサポニンを産生する方法に対する満たされていないニーズが依然として存在する。これらの高価値サポニンは、これに限定しないが、例えば、甘味料、発泡剤、乳化剤、防腐剤、抗発癌剤、コレステロール低下剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、摂食阻害剤、抗真菌剤、またはそれらの任意の組み合わせとして使用することができる。

10

#### 【0009】

加えて、アルカロイドなどの一部の特異的代謝物（SM）は、それらを産生する細胞にとって不可欠ではなく、生物の全体的な適応度に寄与するため、しばしば二次代謝産物と呼ばれる。アルカロイドは、一般的には複素環構造を有する、低分子量の窒素含有有機化合物であり、植物におけるそれらの生合成は、発生段階では、ストレス及び病原体に応答して厳密に制御されている。

20

#### 【0010】

例えば、ステロイドアルカロイド（SA）は、ナス目（Solanales order）の多数の植物に多く含まれていることから「ソラマメアルカロイド」とも呼ばれており、幅広い範囲の科の多数の植物に共通して存在することが分かっている。ステロイドアルカロイド（SA）は、多様な構造組成及び生物活性を有する。これらは低分子量の窒素含有コレステロール由来有機化合物であり、一般的には、C-27 コレスタン骨格と複素環窒素成分とから成る複素環構造を有する。

#### 【0011】

ナス属は、1350種と推定される顕花植物の中でも最大級の属であり、ナス科（ナス目）に属する種の約半分を代表する。ナス種としては、例えば、トマト（*Solanum lycopersicum*）、ジャガイモ（*Solanum tuberosum*）、ピタースイート（*Solanum dulcamara*）、ナス（*Solanum melogena*）などの食用植物が挙げられる、また、ナス科には、トウガラシ属（例えば、コショウ）などの様々な属が含まれる。ステロイドアルカロイド（SA）はまた、ユリ科の多くの種においても産生される。そのため、ステロイドアルカロイド（SA）は、広範な調査の対象となっている（Eich E. 2008. *Solanaceae and Convolvulaceae - secondary metabolites: biosynthesis, chemotaxonomy, biological and economic significance: a handbook*. Berlin: Springer）。

30

#### 【0012】

C-27 コレスタン骨格と複素環窒素成分とから成るステロイドアルカロイド（SA）は、サイトゾル内でコレステロールから合成されると考えられている。コレステロールからアルカミン SA への変換には、いくつかの水酸化、酸化、及びアミノ基転移反応（上記の Eich 2008）を必要とし、ほとんどの場合、さらにグリコシル化してステロイド系グリコアルカロイド（SGA）を形成する必要がある（Arnqvist L. et al. 2003. *Plant Physiology*. 131:1792-1799）。SA のグリコシル化により、付加されたオリゴ糖部分の成分がアルカミンステロイド骨格（アグリコン）の C-3 - （C-3 - ）のヒドロキシル基に直接共役したステロイド系グリコアルカロイド（SGA）が産生される。SGA の生合成は、D-グルコース、D-ガラクトース、L-ラムノース、D-キシロース、及び L-アラビノースなどの様々なオリゴ糖部分でアグリコンを修飾する UDP-糖転移酵素（UGT）をコードする遺伝子に依存し、最初の 2 つの単糖が主要な単位となる。

40

#### 【0013】

ステロイド系グリコアルカロイド（SGA）は、ナス科の数多くの植物、及び他の多く

50

の植物科によって産生される。S G Aの例としては、トマト (*Solanum lycopersicu*) の - トマチン及びデヒドロトマチン、ジャガイモ (*Solanum tuberosum*) の - カコニン及び - ソラニン、ナス (*Solanum melongena*) の - ソラマルギン及び - ソラソニンが挙げられる。また、S G Aは、トウガラシ属の様々な種類のコショウにも見られる。トマトでは100種類以上のS G Aが確認されており (Itkin et al., 2011, *Plant Cell* 23:4507-4525)、ジャガイモでは50種類以上のS G Aが確認されている (Shakya and Navarre, 2008, *J. Agric. Food Chem.* 56:6949-6958)。ナスにも、少なくとも1種類のS G Aが含まれている (Friedmann, 2006, *J. Agric. Food Chem.* 54:8655-8681)。

【0014】

ステロイドアルカロイド (S A) 及びステロイド系グリコアルカロイド (S G A) は、広範囲の病原体から植物を保護する役割を果たしており、フィトアンチシピン (抗菌性化合物) として知られている。S G Aは、細菌、真菌、卵菌、ウイルス、昆虫、及びより大型の動物などの、広範囲の病原体や捕食者に対する植物の耐性に寄与する。植物の食用部分に含まれるいくつかのS G Aは、その毒性作用に起因して、ヒトや他の哺乳類にとって抗栄養化合物と見なされる。例えば、S G Aの - カコニンと - ソラニンはジャガイモに含まれる主要な毒性物質であり、胃腸障害や神経障害を引き起こす可能性があり、高濃度の場合にはヒトに対して致死的な影響を与える可能性がある。毒性の機序としては、膜の破壊、及びアセチルコリンエステラーゼ活性の阻害 (Roddick J.G. 1989. *Phytochemistry* 28:2631-2634) がある。このような理由から、食用塊茎の新鮮重量において総S G A濃度が200 mg / kgを超えると、ヒトの消費に対して安全ではないと考えられる。

【0015】

ステロイド系アルカロイドの生合成経路を解明し、その産生を制御する試みが行われている。特許文献1には、ソラニジンUDP - グルコースグルコシルトランスフェラーゼ (S G T) という酵素をコードするジャガイモ由来のDNA配列が開示されている。また、特許文献1には、S G Tの産生を阻害することによって、ナス科植物、例えばジャガイモにおけるグリコアルカロイドのレベルを低減するための手段及び方法がさらに開示されている。

【0016】

同様に、特許文献2及び特許文献3には、UDP - グルコース : ソラニジングルコシルトランスフェラーゼ (S G T2) 及び - ソラニン / - カコニンラムノシルトランスフェラーゼ (S G T3) をコードするジャガイモ由来の核酸配列がそれぞれ開示されている。また、特許文献2及び特許文献3には、この核酸配列を含む組換えDNA分子、及びその使用、特に、S G T2 / S G T3の産生を阻害し、それによって、ジャガイモなどのナス科植物における主要なステロイド系グリコアルカロイド、 - カコニン、及び - ソラニンのレベルを低減するための該核酸配列及びアンチセンス構築物の使用も開示されている。

【0017】

近年、トマトのステロイドアルカロイドの代謝に関与すると考えられる3つの糖転移酵素が同定された (グリコアルカロイド代謝1 - 3 (G A M E 1 - 3))。より具体的には、G A M E 1発現が改変されると、トマト植物のS Aプロファイルが、繁殖部分及び栄養部分の両方で変化する。これらの遺伝子は、リコテトラオース部分を産生することにより、トマチジン ( - トマチンの前駆体) の代謝に部分的に関与していることが示唆されている (上記の「Itkin et al., 2011」)。

【0018】

特許文献4は、ステロール配糖体転移酵素をコードするDNA配列を含むトランスジェニック生物または細胞を産生する方法を開示している。このトランスジェニック生物には、細菌、真菌、植物、動物が含まれ、これらは、野生型の生物または細胞と比べて、ステロイドグリコシド、ステロイドアルカロイド、及び/またはステロールグリコシドの産生量の増加を示す。この合成化合物は、植物の保護のためだけでなく、製薬業界や食品業界においても有用である。

10

20

30

40

50

## 【 0 0 1 9 】

特許文献 5 には、ナス科の植物、例えばジャガイモ (*Solanum tuberosum*) に由来するグリコアルカロイド生合成酵素をコードする遺伝子が開示されている。また、特許文献 5 には、このタンパク質をコードする遺伝子を使用した新規生物の産生 / 検出方法も開示されている。

## 【 0 0 2 0 】

特許文献 6 及び特許文献 7 は、ステロイド系サポニン及びステロイド系アルカロイドの生合成におけるキー遺伝子、並びに、その遺伝子発現を改変し、ステロイド系サポニン及びステロイド系アルカロイドを産生するための手段及び方法に関するものである。

## 【 0 0 2 1 】

本発明の優先日後に公開された本発明の発明者の論文には、S G A の生合成に關与する 1 0 個の遺伝子の配列が記載されている。この 1 0 個の遺伝子のうちの 5 ~ 7 個の遺伝子は、第 7 染色体上にクラスターとして存在し、他の 2 個の遺伝子は、第 1 2 染色体上の重複ゲノム領域に隣接して存在することが分かっている。系統的機能解析の結果、コレステロールから始まり、トマトの S G A アグリコンに結合している四糖部分に至る、新規な S G A 生合成経路が提案された (Itkin M. et al., 2013 Science 341(6142):175-179)。

## 【 0 0 2 2 】

また、植物のコレステロール生合成経路は、植物ステロール代謝と重複することが分かっている (Sonawane et al. 2016. Nat. Plants 3: 16205)。例えば、コレステロール (3 - ) - コレスト - 5 - エン - 3 - オール) は、すべての動物細胞膜の必須構成成分としてすべての動物細胞で生合成される脂質分子の一種であるステロール (または修飾ステロイド) であり、これは、膜構造の完全性及び流動性に必須であり、動物細胞が細胞壁なしで機能することを可能にしている。また、これは、ステロイドホルモン、胆汁酸、ビタミン D、サポニン、S A、S G A の生合成の前駆体でもある。コレスタノール (5 - - コレスタン - 3 - - オール、コプロスタノール、5 - - コプロスタノール) は、糞便などの生体内に見られるコレステロール誘導体である。コレスタノールは、ほとんどの高等動物 (鳥類及び哺乳類) の腸内でのコレステロールの生物水素化による、コレステロールの変換により産生される、炭素数 2 7 のスタノールである。

## 【 0 0 2 3 】

植物はコレステロールをごく少量産生するが、植物ステロール (コレステロールやコレスタノールに類似する植物ステロールや植物スタノールを含む) も産生するため、腸管での再吸収においてコレステロールと競合し、その結果、コレステロールの再吸収を抑制する可能性がある。コレステロールは、医薬品、化粧品、及び他の用途の製造にしばしば使用される。植物性ステロールと植物性コレステロールとの両方を高レベルで産生することへの関心が高まっている。

## 【 0 0 2 4 】

例えば、トマト (例えば、*Solanum lycopersicum*、*Solanum pennellii*) では、 - トマチンとデヒドロトマチンとが主要な S G A であり、これらは主に緑色組織 (若葉、成熟葉、花芽、未成熟または成熟した緑色果実の皮や種子) に蓄積される。デヒドロトマチジン (すなわちトマチデノール) は、S G A 生合成において最初に産生される S A アグリコンであり、これは、さらに C - 5 位で水素化されてトマチジンを形成する。両アグリコンは、さらにグリコシル化され (四糖部分、すなわちリコテトローズ)、それぞれ、デヒドロトマチンと - トマチンを産生する。このように、S G A 経路は、トマチジン由来の S G A またはグリコシル化されたデヒドロトマチン誘導体のいずれかを形成するために、デヒドロトマチジンで分岐する。注目すべきは、デヒドロトマチジンとトマチジンとの構造の差異は、C - 5 位での二重結合の有無のみである。デヒドロトマチジンからトマチジンへの変換は、従来は、仮説的なヒドロゲナーゼによって触媒される単一の反応であると仮定されていた。トマトのほとんどの組織では、 - トマチンに対するデヒドロトマチンの割合は、約 2 . 5 ~ 1 0 % である。トマトの果実が成熟して赤くなると、 - トマチンとデヒドロトマチンのプール大部分が、エスクレオシド (主要 S G A ) とデヒドロエスクレ

10

20

30

40

50

オシド（マイナー SGA）にそれぞれ変換される。トマト果実が成熟して赤色期に達すると、 $\alpha$ -トマチンとデヒドロトマチンのプール全体が、エスクレオシド（主要 SGA）とデヒドロエスクレオシド（少量 SGA）にそれぞれ変換される。

#### 【0025】

栽培されているジャガイモ（例えば、*Solanum tuberosum*）では、 $\alpha$ -カコニンと $\beta$ -ソラニンが、同一のアグリコンであるソラニジン（C-5、6の二重結合が存在する）を共有する主要な SGA であり、それぞれカコトリオース部分とソラトリオース部位を有している。栽培されているジャガイモからはデミシジンやデミッシンが検出されないことから、これらの種は、ソラニジンをデミシジンに変換するヒドロゲナーゼ酵素を欠いていると示唆されている。活性ヒドロゲナーゼを含むと予測されるいくつかの野生ジャガイモ種（例えば、*S. demissum*、*S. chacoense*、*S. commersonii*）及びその体細胞雑種（*S. brevifolium*、*S. tuberosum*）は、デミシジンまたはそのグリコシル化型を産生し、デミシジンは、それらの主要な SGA の 1 つとなる。

10

#### 【0026】

ナス（例えば、*Solanum melongena*）では、 $\alpha$ -ソラマジンと $\beta$ -ソラソニンが、ソラソニンアグリコン（C-5、6の二重結合が存在する）由来の最も豊富な SGA である。一方、いくつかの野生のナス種、例えば *S. dulcamara*（ビタースイート）は、ソラソニンアグリコンから、ソラズルシジンまたはそのグリコシル化型、ソラズルシン A 及び $\beta$ -ソラズルシン（C-5、6二重結合が存在しない）を、主要な SGA として産生する。

#### 【0027】

多くのナス種（例えば、ナス）は、SGA に加えて、コレステロール由来の不飽和または飽和ステロイドサポニンも産生する。不飽和及び飽和ステロイドサポニンは、植物界に広く存在し、特に、単子葉植物、例えば、リュウゼツラン科（例えば、リュウゼツラン属やユッカ属）、アスパラガス科（例えば、アスパラガス）、ヤマノイモ科及びユリ科に多く含まれている。SGA と同様に、ステロイドサポニンの構造は非常に多様であり、C-5、6位が飽和しているものもあるし（例えば、サラサボゲニン）、不飽和のものもある（例えば、ジオスゲニン）。

20

#### 【0028】

すべての動物が産生する主要なステロールであるコレステロールは、SGA の生合成において重要な構成要素である。コレステロールから始まる SGA 生合成の中核に關与するトマト及びジャガイモの一連の遺伝子である GLYCOALKALOID METABOLISM (GAME) が、近年報告された。トマトの SGA 生合成経路は、2つの主要部分に分けられる。第1の主要部分では、GAME 6、GAME 8、GAME 11、GAME 4、及び GAME 12 酵素の作用により、コレステロールから SA アグリコンが形成されると考えられる。第2の主要部分では、UDP-糖転移酵素 (UGT) の作用により、SGA が産生される（トマトでは GAME 1、GAME 2、GAME 17、GAME 18、ジャガイモでは、ステロールアルカロイド・グリコシルト・ランスフェラーゼ (SGT 1)、SGT 2、SGT 3）

30

#### 【0029】

したがって、トリテルペノイドサポニンを含むサポニンを産生するための代謝経路に關与する酵素に関する知識、例えば、グルクロン酸をキレイン酸などに結合させるセルロース合成酵素様 G (CSLG) 酵素に関する知識に対する満たされていないニーズが依然として存在する。上記の知識が得られれば、自然に産生される産物の含有量を改変させるために、植物において鍵となる酵素を制御することができる。例えば、酵素を過剰に発現させたり、安定性を高めたりすることによって、天然の甘味料を大量に産生することができる。また、酵素を遺伝子レベルで下方制御することによって、植物産物（例えば、キノア）に苦味を加える化合物の産生を減少させることができる。さらに、高価値のサポニン（これに限定しないが、例えば、トリテルペノイドサポニン）を商業量で提供するために、異種系でこれらのサポニンを産生する方法に対する満たされていないニーズが依然として存在する。これらの高価値サポニンは、これに限定しないが、例えば、甘味料、発泡剤、

40

50

乳化剤、防腐剤、抗発癌剤、コレステロール低下剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、摂食阻害剤、抗真菌剤、またはそれらの任意の組み合わせとして使用することができる。

【0030】

加えて、食物量の増加及び品質の改善に対する要求が増加し続けている。栄養価の向上と抗栄養物質の除去は、両方とも強く要求されている。作物の栽培化の過程で、抗栄養物質の濃度は育種によって減少したが、多くの作物は、依然として、かなりの量の抗栄養物質、特にステロイド系グリコアルカロイドを含んでいる。

【0031】

もし、これらのSGAの合成を操作することができれば、従来の育種や遺伝子操作によって、SGAの濃度及び組成が改変された作物を開発することができ、それにより、広範囲の昆虫やその他の病原体に対する内在性の化学的バリアを植物に付与することが可能となる。

10

【0032】

加えて、植物ベースのコレステロールと、その逆に、植物コレステロールや他の植物ステロールの濃度を増加させた植物の両方に対する需要も存在する。

【0033】

したがって、有益な、特に治療効果を有する、サポニン、ステロイド系アルカロイド、及びステロイド系グリコアルカロイドの産生を制御するための手段及び方法が求められており、そのような手段及び方法を有することは非常に有益であると考えられる。

20

【0034】

以下に提供される開示は、いくつかの実施形態において、植物の特異的代謝物(SM)、より具体的にはサポニンがUGT1ファミリー酵素によって単にグリコシル化されるだけでないという知識を少なくとも記載することによって、上記の満たされていないニーズを満たす。一次細胞壁の生合成におけるその役割がよく知られているセルロース合成酵素(CESA)に関連するタンパク質が、サポニンのC-3位へのグルクロン酸の結合を可能にすると思われる。新しい型のトリテルペノイド糖転移酵素活性を有するセルロース合成酵素様(CSL)酵素に対する新機能分化を介した代謝進化は、その細胞内局在性の変化を伴い、それを小胞体(ER)に誘導した。また、このタンパク質の調節は、ステロイド系サポニンとステロイド系グリコアルカロイドの産生の厳密な制御をもたらしたが、遺伝子サイレンシングにより、コレステロールプールの蓄積をもたらされた。

30

【0035】

また、以下に提供される開示は、いくつかの実施形態において、コレステロールを、ステロイドサポニン、トリテルペノイドサポニン、ステロイドアルカロイド、及びステロイドグリコアルカロイドに変換する生合成経路におけるキー遺伝子及び酵素の知識を少なくとも記載することによって、上記の満たされていないニーズを満たす。

【先行技術文献】

【特許文献】

【0036】

【特許文献1】米国特許第5、959、180号明細書

40

【特許文献2】米国特許第7、375、259号明細書

【特許文献3】米国特許第7、439、419号明細書

【特許文献4】国際特許出願公開第WO 00/66716号

【特許文献5】米国特許出願公開第2012/0159676号明細書

【特許文献6】米国特許出願公開第2013/0167271号明細書

【特許文献7】国際出願公開第WO 2012/095843号

【特許文献8】米国特許出願公開第2015/0011008号明細書

【特許文献9】米国特許出願公開第2015/0004704号明細書

【特許文献10】米国特許出願公開第2013/0130389号明細書

【特許文献11】米国特許出願公開第2012/0094385号明細書

50

- 【特許文献 1 2】米国特許出願公開第 2 0 1 2 / 0 0 3 4 6 9 8 号明細書
- 【特許文献 1 3】米国特許出願公開第 2 0 1 1 / 0 3 0 0 6 3 3 号明細書
- 【特許文献 1 4】米国特許出願公開第 2 0 0 4 / 0 1 3 3 9 3 7 号明細書
- 【特許文献 1 5】米国特許出願公開第 2 0 0 9 / 0 2 6 4 3 2 0 号明細書
- 【特許文献 1 6】米国特許出願公開第 2 0 0 1 / 0 0 3 1 7 2 4 号明細書
- 【特許文献 1 7】米国特許出願公開第 2 0 0 3 / 0 0 4 9 7 8 5 号明細書
- 【特許文献 1 8】米国特許出願公開第 2 0 0 5 / 0 1 5 8 8 6 1 号明細書
- 【特許文献 1 9】米国特許出願公開第 2 0 0 7 / 0 2 6 4 7 1 6 号明細書
- 【特許文献 2 0】米国特許出願公開第 2 0 0 9 / 0 3 2 5 2 4 7 号明細書
- 【特許文献 2 1】米国特許出願公開第 2 0 1 0 / 0 1 9 0 2 2 3 号明細書
- 【特許文献 2 2】米国特許第 6、3 2 6、1 7 4 号明細書
- 【特許文献 2 3】米国特許第 4、9 8 7、0 7 1 号明細書

10

【発明の概要】

【課題を解決するための手段】

【0 0 3 7】

一態様によれば、本開示は、セルロース合成酵素様 G ( C S L G ) 酵素をコードする少なくとも 1 つの異種遺伝子の発現が、対応する非組換え細胞と比較して増加した遺伝子組換え細胞であって、少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え細胞と比較して増加した、遺伝子組換え細胞を提供する。

20

【0 0 3 8】

別の態様によれば、本開示は、セルロース合成酵素様 G ( C S L G ) 酵素をコードする少なくとも 1 つの C S L G 遺伝子の発現が、対応する非組換え植物における前記 C S L G 遺伝子の発現と比較して改変された少なくとも 1 つの細胞を含む遺伝子組換え植物であって、少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して改変された、遺伝子組換え植物を提供する。

30

【0 0 3 9】

さらに別の態様によれば、本開示は、遺伝子組換え細胞内にステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法であって、( a ) 前記細胞内に少なくとも 1 つの異種遺伝子を導入する導入ステップであって、前記少なくとも 1 つの異種遺伝子がセルロース合成酵素様 G ( C S L G ) 酵素をコードしており、前記少なくとも 1 つの異種遺伝子が任意選択でベクターに含まれている、該ステップと、( b ) 前記細胞内で前記少なくとも 1 つの異種遺伝子を発現させる発現ステップと、を含み、前記細胞は、少なくとも 1 つのステロイド系アルカロイド、少なくとも 1 つのステロイド系サポニン、または少なくとも 1 つのトリテルペノイド系サポニンの含有量が、対応する非組換え細胞と比較して増加する、方法を提供する。

40

【0 0 4 0】

さらに別の態様によれば、本開示は、植物またはその少なくとも一部の少なくとも 1 つの細胞における、少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を低減させる方法であって、前記少なくとも 1 つの細胞の遺伝子を組み換える遺伝子組換えステップを含み、前記遺伝子組換えステップは、( a ) セルロース合成酵素様 G ( C S L G ) 酵素をコードする少なくとも 1 つの遺伝子を標的とする少なくとも 1 つのサイレンシング分子を用いて前記

50

少なくとも1つの植物細胞を形質転換するステップ、または、(b)前記CSLG酵素をコードする前記少なくとも1つの遺伝子の核酸配列を突然変異させるステップであって、前記突然変異が、(1)前記核酸配列内に1以上の点突然変異を導入すること、(2)前記核酸配列内に欠失を導入すること、(3)前記核酸配列内に挿入を導入すること、または、(4)それらの任意の組み合わせを含み、前記導入が、コード配列または非コード配列を突然変異させることを含む、該ステップと、を含み、遺伝子が組み換えられた遺伝子組換え細胞における前記CSLG酵素をコードする遺伝子の発現は、対応する非組換え植物の細胞と比較して減少し、前記遺伝子組換え細胞を含む植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して減少する、方法を提供する。

10

20

30

40

50

#### 【0041】

さらに別の態様によれば、本開示は、植物またはその少なくとも一部の少なくとも1つの細胞における、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法であって、前記少なくとも1つの細胞の遺伝子を組み換える遺伝子組換えステップを含み、前記遺伝子組換えステップは、(a)セルロース合成酵素様G(CSLG)酵素をコードする少なくとも1つの遺伝子の核酸配列を突然変異させるステップであって、前記突然変異が、(1)前記核酸配列内に1以上の点突然変異を導入すること、(2)前記核酸配列内に欠失を導入すること、(3)前記核酸配列内に挿入を導入すること、または、(4)それらの任意の組み合わせを含み、前記導入が、コード配列または非コード配列を突然変異させることを含む、該ステップと、を含み、(b)前記CSLG酵素をコードする前記核酸配列を発現させるステップと、を含み、遺伝子が組み換えられた遺伝子組換え細胞を含む植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物にと比較して増加する、方法を提供する。

#### 【図面の簡単な説明】

#### 【0042】

ステロイド系グリコアルカロイド、ステロイド系サポニン、トリテルペノイド系サポニンを生産する方法及びその使用に関する主題は、本明細書の結論部分で特に指摘され、明確に主張されている。しかしながら、これらのステロイド系グリコアルカロイド、ステロイド系サポニン、及びトリテルペノイド系サポニンを産生する方法及びその使用は、その目的、特徴、及び利点とともに、組織及び操作方法の両方について、以下の詳細な説明を添付の図面とともに参照することによって最もよく理解されるであろう。

#### 【0043】

【図1】ナス科の植物のトリテルペノイド生合成経路において、コレステロールからトマチンに至るまでのステロイド系グリコアルカロイドの生合成経路を提案している。破線及び実線の矢印は、それぞれ、経路中の複数または単一の酵素反応を表す。

【図2】ナス科の植物のステロイド系アルカロイド関連遺伝子の共発現解析の結果を要約している。トマト(S1GAME1及びS1GAME4)及びジャガイモ(StSGT1及びStGAME4)の「ベイト」の共発現遺伝子の共有されたホモログ。実線( $r$ 値 $>0.8$ )及び破線( $r$ 値 $>0.63$ )で共発現遺伝子を結んでいる。\*は、トマトまたはジャガイモの7番染色体クラスターに位置している。St、Solanum tuberosum(ジャガイモ); S1、S. lycopersicum(トマト)である。遺伝子名の背景は、共発現が確認され

たバイトに対応する（上記の凡例）。SP、セリンプロテアーゼ；PI、プロテアーゼ阻害剤；UPL、ユビキチン・タンパク質リガーゼ；ELP、エクステンシン様タンパク質；PK、プロテインキナーゼ；SR、ステロールレダクターゼ；RL、受容体様。

【図3】トマト及びジャガイモの重複ゲノム領域で同定された遺伝子及びその共発現の概略図である。GAME1/SGT1（7番染色体）及びGAME4（12番染色体）をバイトとした共発現を、ジャガイモまたはトマトのいずれかで、ヒートマップの形態で示している（表5-8）。特定の遺伝子ファミリーを黒色の矢印で示し、他の遺伝子ファミリーの要素を白色の矢印で示している。

【図4A】トマトのGAME遺伝子の機能解析結果を示す図である。GAME8がサイレンシングされたトランスジェニック（RNAi）の葉には、野生型と比較して、22-（R）-ヒドロキシコレステロールが蓄積している。

【図4B】トマトのGAME遺伝子の機能解析結果を示す図である。GAME11がVIGSサイレンシングされた葉には、コレスタノール型のステロイドサポニン（STS）が蓄積している。

【図4C】トマトのGAME遺伝子の機能解析結果を示す図である。GAME12がVIGSサイレンシングされた葉には、STS（m/z = 753.4）が蓄積している。

【図4D】トマトのGAME遺伝子の機能解析結果を示す図である。ステロイド系アルカロイドのアグリコンであるトマチジンが、GAME1がサイレンシングされたトランスジェニック葉に蓄積している。

【図4E】トマトのGAME遺伝子の機能解析結果を示す図である。4つの組換えトマトGAMEグリコシルトランスフェラーゼの酵素活性アッセイの1つを示している。

【図4F】トマトのGAME遺伝子の機能解析結果を示す図である。4つの組換えトマトGAMEグリコシルトランスフェラーゼの酵素活性アッセイの1つを示している。

【図4G】トマトのGAME遺伝子の機能解析結果を示す図である。4つの組換えトマトGAMEグリコシルトランスフェラーゼの酵素活性アッセイの1つを示している。

【図4H】トマトのGAME遺伝子の機能解析結果を示す図である。4つの組換えトマトGAMEグリコシルトランスフェラーゼの酵素活性アッセイの1つを示している。

【図5A】野生型の植物と比較した、GAME9の発現が改変されたジャガイモ植物株の塊茎の皮におけるソラニン/カコニンレベルを示す図の1つであり、GAME9をサイレンシングした植物の塊茎におけるソラニンレベルを示す。

【図5B】野生型の植物と比較した、GAME9の発現が改変されたジャガイモ植物株の塊茎の皮におけるソラニン/カコニンレベルを示す図の1つであり、GAME9をサイレンシングした植物の塊茎におけるカコニンレベルを示す。

【図5C】野生型の植物と比較した、GAME9の発現が改変されたジャガイモ植物株の塊茎の皮におけるソラニン/カコニンレベルを示す図の1つであり、GAME9を過剰発現させた植物の塊茎におけるソラニンレベルを示す。

【図5D】野生型の植物と比較した、GAME9の発現が改変されたジャガイモ植物株の塊茎の皮におけるソラニン/カコニンレベルを示す図の1つであり、GAME9を過剰発現させた植物の塊茎におけるカコニンレベルを示す。

【図6】野生型の植物と比較した、GAME9をサイレンシング（RNAi）または過剰発現（OX）させたジャガイモ植物株の葉におけるソラニン/カコニンレベルを示す図である。

【図7】野生型の植物と比較した、GAME9がサイレンシングされたトマト植物株（RNAi、5871株）、または過剰発現されたトマト植物株（OX、5879株）の葉におけるトマチンレベルを示す図である。

【図8A】トマトにおけるGAME11ジオキシゲナーゼのサイレンシングの効果を示した図の1つであり、葉のトマチンレベル（m/z = 1034.5）を示している。

【図8B】トマトにおけるGAME11ジオキシゲナーゼのサイレンシングの効果を示した図の1つであり、葉のコレスタノール型ステロイド系サポニン（STS）（m/z = 1331.6、1333.6、1199.6、1201.6（主要サポニン））を示してい

10

20

30

40

50

る。

【図8C】トマトにおけるGAME 11ジオキシゲナーゼのサイレンシングの効果を示した図の1つであり、 $m/z = 1331.6$  (19.28分)のMS/MSスペクトルを示している。

【図8D】トマトにおけるGAME 11ジオキシゲナーゼのサイレンシングの効果を示した図の1つであり、19.28分に溶出したサポニンの断片化パターン、及びGAME 11でサイレンシングされた葉における蓄積を示している。図8CのMS/MSクロマトグラムの、アスタリスクが付された質量信号に対応している。

【図9A】GAME 18がサイレンシングされた成熟した緑色のトマトの果実から抽出された代謝物を示す図の1つであり、 - トマチン標準物質 ( $m/z = 740.5$ ) に対応する、新たに蓄積した化合物のピークを示している。

10

【図9B】GAME 18がサイレンシングされた成熟した緑色のトマトの果実から抽出された代謝物を示す図の1つであり、 - トマチン標準物質 ( $m/z = 740.5$ ) に対応する、新たに蓄積した化合物のピークを示している。

【図9C】GAME 18がサイレンシングされた成熟した緑色のトマトの果実から抽出された代謝物を示す図の1つであり、 - トマチン標準物質 ( $m/z = 740.5$ ) に対応する、新たに蓄積した化合物のピークを示している。

【図9D】GAME 18がサイレンシングされた成熟した緑色のトマトの果実から抽出された代謝物を示す図の1つであり、 - トマチンペントシド ( $m/z = 872.5$ ) に対応する、新たに蓄積された化合物のピークを示している。

20

【図9E】GAME 18がサイレンシングされた成熟した緑色のトマトの果実から抽出された代謝物を示す図の1つであり、 - トマチンペントシド ( $m/z = 872.5$ ) に対応する、新たに蓄積した化合物のピークを示している。

【図10A】トマトにおけるGAME 12トランスアミナーゼのサイレンシングの効果を示した図の1つであり、フラスタノール型のSTSを蓄積している。

【図10B】トマトにおけるGAME 12トランスアミナーゼのサイレンシングの効果を示した図である。GAME 12がサイレンシングされた葉には、STS ( $m/z = 753.4$ ) が蓄積されているが、WT (野生型) の葉にはわずかしが存在していない。

【図10C】トマトにおけるGAME 12トランスアミナーゼのサイレンシングの効果を示した図である。GAME 12がサイレンシングされた葉には、STS ( $m/z = 753.4$ ) が蓄積されているが、WT (野生型) の葉にはわずかしが存在していない。

30

【図10D】トマトにおけるGAME 12トランスアミナーゼのサイレンシングの効果を示した図の1つであり、19.71分における $m/z = 753.4$ のMS/MSスペクトル及び断片の解釈を示している。

【図11A】GAME 8をサイレンシングしたトマト植物の効果を示した図の1つであり、EI-GC/MSで得られたクロマトグラム (質量範囲172.5 - 173.5)、MSスペクトル、及び化合物の構造 (トリメチルシリル誘導体) を示している。GAME 8をサイレンシングした葉には、22 - (S) - コレステロール及び22 - (R) - コレステロールが蓄積されている。

【図11B】GAME 8をサイレンシングしたトマト植物の効果を示した図の1つであり、EI-GC/MSで得られたクロマトグラム (質量範囲172.5 - 173.5)、MSスペクトル、及び化合物の構造 (トリメチルシリル誘導体) を示している。22 - (R) - コレステロールの市販の標準物質を用いて推定同定を確認した。

40

【図11C】GAME 8をサイレンシングしたトマト植物の効果を示した図の1つであり、EI-GC/MSで得られたクロマトグラム (質量範囲172.5 - 173.5)、MSスペクトル、及び化合物の構造 (トリメチルシリル誘導体) を示している。22 - (S) - コレステロールの市販の標準物質を用いて推定同定を確認した。

【図11D】GAME 8をサイレンシングしたトマト植物の効果を示した図の1つであり、EI-GC/MSで得られたクロマトグラム (質量範囲172.5 - 173.5)、MSスペクトル、及び化合物の構造 (トリメチルシリル誘導体) を示している。GAME 8

50

がサイレンシングされた株には、WT (Q) と比較して、両異性体が蓄積されている。

【図 1 2】植物 CYP450 タンパク質ファミリーにおける GAME 遺伝子の系統樹の図である。枝上の番号は、各節を支持するブートストラップ反復の比率を示す。

【図 1 3】シクロアルテノール (Part I) から、コレステロール (Part III)、トマチジン (Part IIII)、 $\beta$ -トマチンを含むステロイド系グリコアルカロイドを経てリコペロシド/エスクレオシド (Part IV) に至る、ナス科植物における拡大された生合成経路の提案を示している。破線の矢印は、経路中の複数の酵素反応を表す。

【図 1 4 A】トマトにおける SGA 生合成の概要を示す図である。

【図 1 4 B】ジャガイモにおける SGA 生合成の概要を示す図である。

【図 1 4 C】ナスにおける SGA 生合成の概要を示す図である。

【図 1 4 D】トマチジンからトマチジン 3-O-グルクロニドの生成までの、ナス科植物における SGA 生合成経路の提案部分を示す図である。GAME15 セルロース合成酵素様 G 酵素が、トマチジンの水酸基の位置にグルクロン酸の付加を触媒している。

【図 1 5 A】LC-MS により測定した、野生型 (非形質転換) 及び GAME15-RNAi トマト株の葉における主要な SGA レベルを示す図である。#21、#22、#23 は、GAME15-RNAi を遺伝子導入した 3 株の別個のトマトである。数値は、3 つの生物学的複製の平均値  $\pm$  標準誤差を示す。アスタリスクは、スチューデントの t 検定によって算出した野生型試料に対する有意な変化を示している (\*P 値 < 0.05、\*\*P 値 < 0.01、\*\*\*P 値 < 0.001)。

【図 1 5 B】LC-MS により測定した、野生型 (非形質転換) 及び GAME15-RNAi トマト株の緑色の果実における主要な SGA レベルを示す図である。#21、#22、#23 は、GAME15-RNAi を遺伝子導入した 3 株の別個のトマトである。数値は、3 つの生物学的複製の平均値  $\pm$  標準誤差を示す。アスタリスクは、スチューデントの t 検定によって算出した野生型試料に対する有意な変化を示している (\*P 値 < 0.05、\*\*P 値 < 0.01、\*\*\*P 値 < 0.001)。

【図 1 5 C】LC-MS により測定した、野生型 (非形質転換) 及び GAME15-RNAi トマト株の赤い果実における主要な SGA レベルを示す図である。#21、#22、#23 は、GAME15-RNAi を遺伝子導入した 3 株の別個のトマトである。数値は、3 つの生物学的複製の平均値  $\pm$  標準誤差を示す。アスタリスクは、スチューデントの t 検定によって算出した野生型試料に対する有意な変化を示している (\*P 値 < 0.05、\*\*P 値 < 0.01、\*\*\*P 値 < 0.001)。

【図 1 6】LC-MS により測定した、GAME15-RNAi 株の葉における  $\beta$ -ソラニン及び  $\beta$ -カコニンのレベルを示す図である。#1、#2、#3 は、GAME15i を遺伝子導入した 3 株の別個のじゃがいもである。数値は、平均値  $\pm$  標準誤差を表す (n = 3)。スチューデントの t 検定を用いて、トランスジェニック株が野生型の植物に対して有意に異なるかどうかを評価した (\*P 値 < 0.05、\*\*P 値 < 0.01、\*\*\*P 値 < 0.001)。

【図 1 7】GAME15 をサイレンシングした植物から得られたトマトの葉のコレステロール含有量を示す図である。数値は、3 つの生物学的複製の平均値  $\pm$  標準誤差を示す。アスタリスクは、スチューデントの t 検定で算出した、野生型の (すなわち、形質変換されていない) 葉と比較した場合の 3 つの独立した導入遺伝子の葉 (#21、#22、#23) における有意な変化を示す (\*P 値 < 0.05、\*\*P 値 < 0.01、\*\*\*P 値 < 0.001)。試料調製時にはエピコレステロールを内部標準物質として用い、相対的なコレステロール値を、内部標準物質と比較した試料中のコレステロールのピーク面積の比で表した。分析は、GC-MS を用いて行った。

【図 1 8】ハウレンソウのサポニンの液体クロマトグラフィー・エレクトロスプレーイオン化・四重極/飛行時間型質量分析 (LC-ESI-QTOF-MS) スキャンの図である。上の図は、陰イオン化モードで取得したハウレンソウの葉の抽出物のトータルイオンクロマトグラム (TIC) を示している。赤色で表示されたクロマトグラムの断片は、試

10

20

30

40

50

料に含まれるサポニンを示している。下の図は、クロマトグラムの拡大図であり、主要なサポニンには番号が付されている。この番号は、実施例 15 で示した表 13 に記載されている様々な Y o s s o s i d e を表している。

【図 19 A】Y o s s o s i d e I V のタンデム質量分析 ( M S / M S ) の結果及び構造を示す図である。[ M - H ] 1 1 9 6 3 . 5 6 = - m / z 範囲 3 0 0 - 1 3 0 0 に由来する質量フラグメント。

【図 19 B】Y o s s o s i d e I V のタンデム質量分析 ( M S / M S ) の結果及び構造を示す図である。グルクロン酸残基の切断に由来する質量フラグメント、m / z 範囲 5 0 - 1 9 0 0 。

【図 19 C】Y o s s o s i d e I V のタンデム質量分析 ( M S / M S ) の結果及び構造を示す図である。Y o s s o s i d e I V の構造であり、矢印はフラグメントのパターンを示している。

【図 19 D】Y o s s o s i d e I V のタンデム質量分析 ( M S / M S ) の結果及び構造を示す図である。M A に結合したグルクロン酸のフラグメントに由来する M S / M S ピークの割り当てを示す表。

【図 19 E】Y o s s o s i d e V a のタンデム質量分析 ( M S / M S ) の結果及び構造を示す図である。[ M - H ] 1 3 0 5 . 5 7 1 9 0 = - m / z 範囲 3 0 0 - 1 4 0 0 に由来する質量フラグメント。

【図 19 F】Y o s s o s i d e V a のタンデム質量分析 ( M S / M S ) の結果及び構造を示す図である。グルクロン酸残基の切断に由来する質量フラグメント、m / z 範囲 5 0 - 1 9 0 0 。

【図 19 G】Y o s s o s i d e V a のタンデム質量分析 ( M S / M S ) の結果及び構造を示す図である。Y o s s o s i d e V a の構造であり、矢印は断片パターンを示す。アセチル基の位置 ( 赤で強調 ) は推定である。

【図 19 H】Y o s s o s i d e V のタンデム質量分析 ( M S / M S ) の結果及び構造を示す図である。[ M - H ] 1 3 0 5 . 5 7 3 7 = - m / z 範囲 3 0 0 - 1 4 0 0 に由来する質量フラグメント

【図 19 I】Y o s s o s i d e V のタンデム質量分析 ( M S / M S ) の結果及び構造を示す図である。グルクロン酸残基の切断に由来する質量フラグメント、m / z 範囲 5 0 - 1 9 0 0 。

【図 19 J】Y o s s o s i d e V のタンデム質量分析 ( M S / M S ) の結果及び構造を示す図である。Y o s s o s i d e V の構造であり、矢印はフラグメントパターンを示す。

【図 19 K】Y o s s o s i d e I X のタンデム質量分析 ( M S / M S ) の結果及び構造を示す図である。[ M - H ] 1 4 0 7 . 6 0 6 0 = - m / z 範囲 3 0 0 - 1 5 0 0

【図 19 L】Y o s s o s i d e I X のタンデム質量分析 ( M S / M S ) の結果及び構造を示す図である。グルクロン酸残基の切断に由来する質量断片、m / z 範囲 5 0 - 1 9 0 0 。

【図 19 M】Y o s s o s i d e I X のタンデム質量分析 ( M S / M S ) の結果及び構造を示す図である。Y o s s o s i d e I X の構造であり、矢印はフラグメンテーションパターンを示す。ペントースの構造及び結合場所は推定されている。

【図 19 N】Y o s s o s i d e X のタンデム質量分析 ( M S / M S ) の結果及び構造を示す図である。[ M - H ] 1 4 3 7 . 6 1 5 0 = - m / z 範囲 3 0 0 - 1 5 0 0 。

【図 19 O】Y o s s o s i d e X のタンデム質量分析 ( M S / M S ) の結果及び構造を示す図である。グルクロン酸残基の切断に由来する質量フラグメント、m / z 範囲 5 0 - 1 9 0 0 。

【図 19 P】Y o s s o s i d e X のタンデム質量分析 ( M S / M S ) の結果及び構造を示す図である。Y o s s o s i d e X の構造であり、矢印はフラグメント化パターンを示す。構造とペントースとの結合は推定である。

【図 20 A】共発現したサポニン - アミリン合成酵素 ( S O A P ) 遺伝子を含む、ホウ

10

20

30

40

50

レンソウのサポニンの生合成経路を示す図である。C - 20位（赤で強調）にグルクロン酸が結合したアセチルトリテルペノイドサポニンであるY o s s o s i d e Vの生成に至る完全な生合成経路。S O A P酵素の活性による各生成物は緑色で表示されている。S O A P酵素は、ハウレンソウのサポニン生産に関与する酵素である。

【図20B】共発現したサポニン - アミリン合成酵素（S O A P）遺伝子を含む、ハウレンソウのサポニンの生合成経路を示す図である。遺伝子共発現ネットワーク。赤色のドットはベイト（S O A P 1、S O A P 2、及びS o C Y P 7 1 6 A 2 6 8 v 2）を表し、緑色のドットで囲まれた円は全てのベイトと共発現している遺伝子を表している。拡大された注釈付きのドットは、ハウレンソウのウイルス誘導遺伝子サイレンシング（V I G S）によってサイレンシングされた遺伝子を表す。

10

【図20C】共発現したサポニン - アミリン合成酵素（S O A P）遺伝子を含む、ハウレンソウのサポニンの生合成経路を示す図である。S O A P 1 - 4を一過性発現させた植物からのオレアノール酸、アウグスチン酸、メディカゲン酸の抽出イオンクロマトグラム（E I C）[ m / z = 4 5 5 . 2 0 5（O A）、m / z = 4 7 1 . 2 0 5（A A）、m / z = 5 0 1 . 2 0 2（M A）、陰イオン化モード]を、S O A P 1 - 2 0、S O A P 1 - 2を単独で発現させた植物、及び対照（空のベクター（E V）を保有するアグロバクテリウム・リゾゲネスを浸透させた植物）と比較した。

【図21】Y o s s o s i d e Vの水素の重要な異種核多重結合（H M B C）の相関を示す図である。黒い矢印は、Y o s s o s i d e Vの6つのメチル基のH M B C相関を示す。赤い矢印は、他の重要な水素のH M B C相関を示す。数字は、メディカゲン酸及び糖部分の炭素原子に対応する。詳細は表1を参照。表1は、スペクトルデータを2つの部分に分けて表にまとめたものである。第1の部分はメディカゲン酸の原子（アグリコン部分）に対応し、第2の部分は糖部分の原子に対応する。Y o s s o s i d e V及びメディカゲン酸3 - O - グルクロニドの構造を、原子番号を付して図21及び図35Dにそれぞれ示している。表1において、a x（軸方向）及びe q（赤道方向）は、アグリコン及び糖の環構造におけるプロトンを示す。

20

【0044】

表1 . Y o s s o s i d e V ( 3 - O - [ - D - キシロピラノシル - ( 1 3 ) -  
- D - グルコピラノシル ] - 2 8 - O - [ - D - グルコピラノシル - ( 1 4 ) -  
- L - ラムノピラノシル - ( 1 2 ) - 4 - アセチル - - D - フコピラノシル ] - メ  
ディカゲン酸) 及びメディカゲン酸3 - O - グルクロニドについての<sup>1</sup>H及び<sup>13</sup>C N M  
Rスペクトルデータ

30

【0045】

【表 1 - 1】

Yossoside V					
アグリコン部分			糖部分		
原子番号	<sup>13</sup> Cシフト (ppm)	<sup>1</sup> Hシフト (ppm)	原子番号	<sup>13</sup> Cシフト (ppm)	<sup>1</sup> Hシフト (ppm)
1	44.77	2.08 (eq) 1.24 (ax)	$\beta$ -D-Fuc		
2	70.68	4.30	1	94.94	5.38 (ax)
3	86.47	4.05 (ax)	2	74.58	3.74 (ax)
4	53.01		3	74.57	3.88 (ax)
5	53.01	1.57 (ax)	4	75.13	5.06 (eq)
6	21.53	1.61	5	71.06	3.83 (ax)
		1.16	6	16.37	1.06
7	33.65	1.35 (eq) 1.49 (ax)	アセチル CO	172.69	
8	40.96		メチル	20.78	2.14
9	49.44	1.57 (ax)	$\alpha$ -L-Rha		
10	37.24		1	101.48	5.33 (eq)
11	24.68	1.93 (eq) 1.99 (ax)	2	71.66	3.90 (eq)
12	123.51	5.26	3	72.04	3.83 (ax)
13	144.57		4	84.52	3.53 (ax)
14	43.09		5	68.74	3.80 (ax)
15	28.96	1.18 (eq) 1.58 (ax)	6	18.28	1.32
16	23.92	1.60 (eq) 2.04 (ax)	$\beta$ -D-Glc		
17	48		1	106.4	4.47 (ax)
18	42.8	2.81	2	76.14	3.18 (ax)
19	47.22	1.13 1.71	3	78.28	3.34 (ax)
20	31.4		4	71.46	3.27 (ax)
21	34.77	1.22 (eq) 1.38 (ax)	5	78.09	3.25 (ax)
22	33.01	1.56 (eq) 1.75 (ax)	6a	62.85	3.66
23	182.07		6b	62.85	3.84
24	13.57	1.37 (ax)	$\beta$ -D-GlcA		
25	17.08	1.24 (ax)	1	104.79	4.39
26	17.66	0.78 (ax)	2	68.93	3.81
27	26.24	1.14 (ax)	3	70.94	3.81
28	177.9		4	74.31	3.43
29	33.45	0.89 (eq)	5	nd	nd
30	24.1	0.92 (ax)	6	nd	
			$\beta$ -D-Xyl		
			1	105.38	4.55 (ax)
			2	75.17	3.25 (ax)
			3	77.31	3.31
			4	70.89	3.48
			5a	66.85	3.87
			5b	66.85	3.2

10

20

30

40

【表 1 - 2】

メディカゲン酸3- $\beta$ -D-グルクロニド					
アグリコン部分			グルクロン酸		
原子番号	$^{13}\text{C}$ シフト (ppm)	$^1\text{H}$ シフト (ppm)	原子番号	$^{13}\text{C}$ シフト (ppm)	$^1\text{H}$ シフト (ppm)
1	43.72	1.24 (ax)	$\beta$ -D-GlcA		
		2.09 (eq)	1	103.14	4.41 (ax)
2	69	4.32	2	74.02	3.27 (ax)
3	87.14	4.07 (ax)	3	75.95	3.40 (ax)
4	52.97		4	74.52	3.60 (ax)
5	51.73	1.62 (ax)	5	72.09	3.39 (ax)
6	20.49	1.29	6	175.6	
		1.62			
7	32.44	1.24 (eq)			
		1.56 (ax)			
8	39.44				
9	48.18	1.60 (ax)			
10	36.12				
11	23.31	1.89 (eq)			
		2.02 (ax)			
12	121.05	5.22			
13	145.36				
14	41.9				
15	27.84	0.95 (eq)			
		1.87 (ax)			
16	23.2	1.57 (eq)			
		1.88 (ax)			
17	nd				
18	42.07	2.90			
19	46.73	1.08			
		1.65 (eq)			
20	30.43				
21	34.12	1.12 (eq)			
		1.34 (ax)			
22	32.94	1.49 (eq)			
		1.72 (ax)			
23	184.8				
24	13.46	1.36 (ax)			
25	15.8	1.26 (ax)			
26	16.97	0.88 (ax)			
27	25.15	1.14 (ax)			
28	nd				
29	32.56	0.87 (eq)			
30	22.81	0.95 (ax)			

10

20

30

40

## 【0047】

【図 2 2 A】アグリコン生合成の最初のステップに關与する遺伝子が、染色体上の近接した場所に位置していることを示す図である。オレアノール酸の生合成に關与する遺伝子が存在する染色体領域を模式的に示している。半透明の灰色のボックスには、SOAP1 (bAS) (配列番号 4 2 2)、CYP716A268 (SOAP2; 配列番号 4 6)、及びその切断された重複 CYP716A268v2 (配列番号 4 7) が含まれる。濃い灰色でマークされた遺伝子は、トリテルペノイド系サポニンの生合成に直接は關与していない。

【図 2 2 B】アグリコン生合成の最初のステップに關与する遺伝子が、染色体上の近接した場所に位置していることを示す図である。5つのハウレンソウのトランスクリプトーム

50

におけるSOAP1、SOAP2及びCYP716A268v2の発現パターン；SOWORLD - 長日期に生育した4週齢の植物の成熟葉；SOWOLL - 短日期に生育した8週齢の植物の成熟葉；SOWYLLD - 長日期に生育した4週齢の植物の若葉；SOWOLL - 短日期に生育した8週齢の植物の若葉；SOWFLD - 長日期に生育した4週齢の植物の花芽。詳細は対象と方法を参照。

【図23A】ハウレンソウにおけるサイレンシング(VIGS)によるSOAP1(bAS)及びSOAP2(CYP7123A2238)の機能的特徴を示すデータである。マグネシウム-ケラターゼサブユニットH(CHLH)をサイレンシングした植物で観察された萎黄病の表現型。CHLHは対照として、かつ、VIGS遺伝子サイレンシングのマーカーとして使用された。一般的に、この遺伝子はSOAP遺伝子やサポニン生合成には直接関係しないと言われている。EVは空のベクターを意味し、対照の葉を表す。

10

【図23B】ハウレンソウにおけるサイレンシング(VIGS)によるSOAP1(bAS)及びSOAP2(CYP7123A2238)の機能的特徴を示すデータである。サイレンシングされた植物におけるSOAP1及びSOAP2の発現のリアルタイムqPCR分析。数値(±SD)は3回の独立した生物学的実験の平均値。対照植物(CHLHサイレンシング)と比較して、統計的に有意な差を示した；\*\*\* $P < 0.001$ 。

【図23C】ハウレンソウにおけるサイレンシング(VIGS)によるSOAP1(bAS)及びSOAP2(CYP7123A2238)の機能的特徴を示すデータである。対照(CHLH)と比較した、SOAP1またはSOAP2をサイレンシングした植物から得られた - アミリン[m/z = 218、主要なフラグメント]のガスクロマトグラフィー質量分析(GC-MS)抽出イオンクロマトグラム(EIC)の整列である。 - アミリンは、SOAP2をサイレンシングした植物でのみ検出され、そのさらなる変換が阻害されている。

20

【図23D】ハウレンソウにおけるサイレンシング(VIGS)によるSOAP1(bAS)及びSOAP2(CYP7123A2238)の機能的特徴を示すデータである。対照(CHLH)と比較した、SOAP1またはSOAP2がサイレンシングされた植物からのYossoside IV[m/z = 12233.23、陰イオン化モード]の液体クロマトグラフィー質量分析(LC-MS)EICの整列である。Yossoside IVの含有量は、SOAP1及びSOAP2をサイレンシングした株では、それぞれ40%及び30%に減少した。

30

【図24A】ベンサミアナタバコにおける異種発現によるSOAP1(bAS；配列番号45)及びSOAP2(CYP2416A268；配列番号46)の機能的特徴を示すデータの図である。空のベクター(EV)単独及び基準の標準物質と比較した、SOAP1の一時的な発現による - アミリン[m/z = 218、主要なフラグメント]のGC-MS抽出イオンクロマトグラム(EIC)。インプラントでの実験で得られたクロマトグラムのみを縮尺通りに表示している。

【図24B】ベンサミアナタバコにおける異種発現によるSOAP1(bAS；配列番号45)及びSOAP2(CYP2416A268；配列番号46)の機能的特徴を示すデータの図である。空のベクター(EV)単独及び基準の標準物質と比較した、SOAP1及びSOAP2の一時的な発現によるオレアノール酸[m/z = 203、主要なフラグメント]のGC-MS EICの整列。インプラントでの実験で得られたクロマトグラムのみを縮尺に合わせている。GC-MS分析の詳細については「対象と方法」を参照。

40

【図25】シトクロムP450候補遺伝子が(VIGS)サイレンシングされたハウレンソウにおける選択されたサポニンの液体クロマトグラフィー質量分析(LC-MS)の分析結果を示すデータである。SOAP3(CYP72A655；配列番号51)、SOAP4(CYP72A654；配列番号53)、及び他の3つのCYP(p450シトクロム遺伝子)をサイレンシングした植物における、選択されたサポニン化合物の相対的な含有量(正規化されたピーク面積)を示している。受託番号については、表11を参照。数値(±SD)は3回の独立した生物学的実験の平均値を示す。対照植物(CHLHサイレンシング)と比較して統計的に有意な差を示した；\* $P < 0.05$ 、\*\* $P < 0.01$ 。

50

【図26A】シトクロムP450遺伝子のサイレンシング(VIGS)後にホウレンソウに蓄積するサポニンの特徴を示す図である。SOAP3(CYP72A655; 配列番号51)、SOAP4(CYP72A654; 配列番号53)、及び追加の3つのシトクロムP450遺伝子をサイレンシングしたホウレンソウの葉における、選択されたトリテルペノイド系サポニン化合物の相対的な含有量(正規化されたピーク面積)であり、受託番号は表11を参照。数値(±SD)は3回の独立した生物学的実験の平均値。対照植物(CHLHサイレンシング)と比較して、統計的に有意な差を示した; \* $P < 0.05$ 、\*\* $P < 0.01$ 。

【図26B】シトクロムP450遺伝子のサイレンシング(VIGS)後にホウレンソウに蓄積するサポニンの特徴を示す図である。主要なフラグメントをマークした、AA-アウグスチン酸のタンデム質量分析(MS/MS)スペクトル。

10

【図26C】シトクロムP450遺伝子のサイレンシング(VIGS)後にホウレンソウに蓄積するサポニンの特徴を示す図である。正確な質量測定、フラグメンテーション、及び文献に基づいたサポニン化合物の推定構造。詳細は表16を参照。

【図26D】シトクロムP450遺伝子のサイレンシング(VIGS)後にホウレンソウに蓄積するサポニンの特徴を示す図である。SOAP3(CYP72A655; 配列番号51)、SOAP4(CYP72A654; 配列番号53)、及び追加の3つのシトクロムP450遺伝子をサイレンシングしたホウレンソウの葉における、選択されたトリテルペノイド系サポニン化合物の相対的な含有量(正規化されたピーク面積)であり、受託番号は表11を参照。数値(±SD)は3回の独立した生物学的実験の平均値。対照植物(CHLHサイレンシング)と比較して、統計的に有意な差を示した; \* $P < 0.05$ 、\*\* $P < 0.01$ 。

20

【図26E】シトクロムP450遺伝子のサイレンシング(VIGS)後にホウレンソウに蓄積するサポニンの特徴を示す図である。主要なフラグメントをマークした、AA-アウグスチン酸のタンデム質量分析(MS/MS)スペクトル。

【図26F】シトクロムP450遺伝子のサイレンシング(VIGS)後にホウレンソウに蓄積するサポニンの特徴を示す図である。正確な質量測定、フラグメンテーション、及び文献に基づいたサポニン化合物の推定構造。詳細は表16を参照。

【図26G】シトクロムP450遺伝子のサイレンシング(VIGS)後にホウレンソウに蓄積するサポニンの特徴を示す図である。SOAP3(CYP72A655; 配列番号51)、SOAP4(CYP72A654; 配列番号53)、及び追加の3つのシトクロムP450遺伝子をサイレンシングしたホウレンソウの葉における、選択されたトリテルペノイド系サポニン化合物の相対的な含有量(正規化されたピーク面積)であり、受託番号は表11を参照。数値(±SD)は3回の独立した生物学的実験の平均値。対照植物(CHLHサイレンシング)と比較して、統計的に有意な差を示した; \* $P < 0.05$ 、\*\* $P < 0.01$ 。

30

【図26H】シトクロムP450遺伝子のサイレンシング(VIGS)後にホウレンソウに蓄積するサポニンの特徴を示す図である。主要なフラグメントをマークした、AA-アウグスチン酸のタンデム質量分析(MS/MS)スペクトル。

【図26I】シトクロムP450遺伝子のサイレンシング(VIGS)後にホウレンソウに蓄積するサポニンの特徴を示す図である。正確な質量測定、フラグメンテーション、及び文献に基づいたサポニン化合物の推定構造。詳細は表16を参照。

40

【図26J】シトクロムP450遺伝子のサイレンシング(VIGS)後にホウレンソウに蓄積するサポニンの特徴を示す図である。SOAP3(CYP72A655; 配列番号51)、SOAP4(CYP72A654; 配列番号53)、及び追加の3つのシトクロムP450遺伝子をサイレンシングしたホウレンソウの葉における、選択されたトリテルペノイド系サポニン化合物の相対的な含有量(正規化されたピーク面積)であり、受託番号は表11を参照。数値(±SD)は3回の独立した生物学的実験の平均値。対照植物(CHLHサイレンシング)と比較して、統計的に有意な差を示した; \* $P < 0.05$ 、\*\* $P < 0.01$ 。

50

【図26K】シトクロムP450遺伝子のサイレンシング(VIGS)後にホウレンソウに蓄積するサポニンの特徴を示す図である。主要なフラグメントをマークした、AA-アウグスチン酸のタンデム質量分析(MS/MS)スペクトル。

【図26L】シトクロムP450遺伝子のサイレンシング(VIGS)後にホウレンソウに蓄積するサポニンの特徴を示す図である。正確な質量測定、フラグメンテーション、及び文献に基づいたサポニン化合物の推定構造。詳細は表16を参照。

【図27】CYP P450をサイレンシングしたホウレンソウ植物におけるサポニンのアグリコンのGC-MS分析結果を示す図である。対照(CHLHサイレンシング)並びにオレアノール酸及びメディカゲン酸の基準の標準物質と比較した、SOAP3単独、SOAP4単独、及びCYP2をサイレンシングした植物に蓄積されたサポニンから放出されたトリテルペノイドアグリコン[m/z = 203及び262;すべてのアミリン型アグリコンに共通するフラグメント]のEICである。各EICの縮尺を、最も豊富なシグナルに対して正規化した。1-7の数字は検出された化合物の名前を表しており、それぞれ、1-オレアノール酸、2-ヘデラゲニン、3-アウグスチン酸、4-ギブソゲニン、5-パヨゲニン、6-不明、7-メディカゲン酸である。

10

【図28A】CYP P450をサイレンシングしたホウレンソウにおけるアグリコンの分析及び推定される特性を示す図である。FIGURE 28A)対照としてSOAP3、SOAP4、CYP2及びCHLHをサイレンシングした植物において検出されたサポニンの相対含有量。

【図28B】CYP P450をサイレンシングしたホウレンソウにおけるアグリコンの分析及び推定される特性を示す図である。特徴的なアグリコンの構造。1-7の数字は検出された化合物の名前を表しており、それぞれ、1-オレアノール酸、2-ヘデラゲニン、3-アウグスチン酸、4-ギブソゲニン、5-パヨゲニン、6-不明、7-メディカゲン酸である。

20

【図28C】CYP P450をサイレンシングしたホウレンソウにおけるアグリコンの分析及び推定される特性を示す図である。記載された化合物のフラグメンテーションと、オレアノール酸及びメディカゲン酸の基準の標準物質のフラグメンテーションとの比較。

【図29】グリコシルトランスフェラーゼ遺伝子がサイレンシングされた(VIGS)ホウレンソウにおける選択されたサポニンのLC-MS分析を示す図である。SOAP6-9(それぞれ、配列番号55、57、59、61)、SoGT1(配列番号117[遺伝子]及び配列番号118[ポリペプチド])、SoGT2(配列番号119[遺伝子]及び配列番号120[ポリペプチド])がサイレンシングされた植物における選択されたサポニン化合物の相対含有量(正規化されたピーク面積);受託番号については表11を参照。グリコシルトランスフェラーゼに使用した配列を以下に示す。

30

【0048】

ホウレンソウGT1の核酸遺伝子の配列は以下の通りである。atggccgacctcaaacaaagagcacaagcctacataattgctatgtttctcttggtttgcatatggctcattataaacccattcattccagctctctaacaaagctttctcccatggcattccaaaatctctttctctctcaataaccaggaaacattgatcgtatcaaaatctctcccttaatctctcacctcccaaacagctcattccccctcactattcccccaactgaaaggactctctccccaaactttgacagcagctctgaaagtaaacactcaaacctgctcagctctctcacaactagcacttgatcaaatgcagccccaaagtcaaaagctctattctctcaccagccacaagttaattctctttgattttgcatattcactggtttctctcagtagctcttgaaactaggcattcaaaagctgttcatttcaaatacattccccagctgttattcaaatcattatctcaactgtctctcaagaatgactgatccaaaataaacaccgacatttgaggacttgaaagaaacctctcaaggctattccc

40

50

a a a a c c t c a a c c g c c t c a g t g a a a c c t t c g a a g c t c a a g  
 a t t a c c t a t t c c t t t t c a a g a g t t t c g a t g g c g g a c c g t g  
 c c a t t t c g a a a a g a t a t t g g c a t t c a c a a a c a g c t g t g a t  
 g c t a t a c t t t a c a g g a c c t g c a a t g a a a t a g a a g g t c c a t  
 t c a t a g a t t a c t t c a a g a c c c a a a t a a a t a a a c c a c t g c t  
 t t t a g c t g g c c c a a g t g t t c c t c t a c c a c c c t c t g g t g a a  
 c t g g a t g a a a a a t g g g a g a t g t g g t t a g g t a a a t t t c c t g  
 a a a a g t c a g t c a t a t a c t g c a g c t t c g g a a g c g a g a c a t a  
 c t t g a a t g a t g c t c a g a t t c a g g a g c t t a c a c t t g g g t t g  
 g a g c t c a c t g g t c t g c c c t t t a t c t t g g t t t t g a a t t t t g  
 g a a c a a g t a a c a g c a c c g a t g c c c a c a a t a a g c t a g a a g c  
 a t c a t t a c c a g a a g g a t t t a g a g a g a a t c a a a g a c a g g  
 g g c g t t c t g c a t a ( 配列番号 117 ) 。

10

【0049】

ホウレンソウGT1のアミノ酸タンパク質の配列は以下の通りである。m a a s n k e  
 q s k l h i a m f p w f a y g h i n p f i q l s n k l s s h g i q i s f f s i p  
 g n i d r i k s s l n l s p p n q l i p l t i p p t e g l s p n f d s s s e v t  
 p q t a q l l t l a l d q m q p q v k a l f p h p q p q v i l f d f a y h w l p  
 s v a s e l g i k a v h f n t f p a v i n s y l t v p s r m t d p n k p p t f e  
 d l k n p p q g y p k t s t a s v k t f e a q d y l f l f k s f d g g p c h f e  
 k i l a f t n s c d a i l y r t c n e i e g p f i d y f k t q i n k p l l l a g  
 p s v p l p p s g e l d e k w e m w l g k f p e k s v i y c s f g s e t y l n d  
 a q i q e l t l g l e l t g l p f i l v l n f g t s n s t d a h n k l e a s l p  
 e g f r e r i k d r g v l h t g w v q q q n i l a h r s i g c f l t h a g f s s  
 v i e g i v n d c q l a f l p l k a d q f m i a k l f s g d l k a g v e v n r r  
 d e d g s f a k e d i f e a i k t i m v d t d k e p s r s i r e n h s n w r k f  
 l m n k e i e a s y i a n l a h e l k a l v q k a ( 配列番号 118 ) 。

20

【0050】

ホウレンソウGT2核酸遺伝子の配列は以下の通りである。a t g t g t g a c g a c  
 a a a a a a t c a t c t g t t t t g a g c a t a g c a t t t t a t c c g t g g t  
 t t g c t c t t g g t c a c c t t a c t t c a t t t c t c c g a t t a g c c a a  
 c a a a c t t g c a c a a a a t g g t c a c a a t g t g t c c t a t t t t a t c  
 c c a a c t a a t a c a t t a c c t a g a t t a c t t c c t c a c a a c c a t t  
 a c c c t g g c c a c c t t a c t t t c a t c c c c g t c a c c g t c c c a c c  
 c g t t g a c g g c c t c c c t c t c a g a g c c g a g a c c a c c a a c g a t  
 g t c c c c t c c t c g g c t a t a c a c c t t c t t a t g a c t g c c a t g g  
 a t t t g a c c c g t g a c a c t a t c g a g g c c c a t t t g g t t a g t a t  
 c a a a c c c g a t g t t g t t t t c t a c g a c t t t g c t t a t t g g a t t  
 c c c g a t c t a g c c c g a a a a c a c g g g t t c a a g t c a g t a c t c t  
 a c a t t a c a t c c t a t a t a g c a a g a t g t g c t t a t t t t g c c c c  
 c g a t t t g a a g t c g g g t c a t c a g t c c a c t g g g g c c g a a a t t  
 a t t g c g c c a c c a c c g g g t t t t c c g t c t c a g c a t t t c c g g a  
 t g c a a g c a c a c g a g g c t g a g a c t g t g g c a g a c g t a g g t a a  
 a g a g c a a g a t g g a t t a c a a g g t a t a a c t a t t t c t g a a a g g  
 a t g c g c a t t g c t t t t g g a a a a t g c g a c g c a a t t g g a g t a a  
 a g a g t t g t a a g g a g a t g g a a a a g g t g t a t a t t g a c t a c t g  
 t g a g a a g a t a t t t g g t a a g t c t g t a c t a c t a g c a g g t c c t  
 a t g g t c c c t a a a a c c c c a t c t t c c a a a c t t g a t g a a t a t t  
 t t g a t g g t t g g c t t a c g g g t t t t g g t g c t g c t a c t g t g a t  
 t t a t t g t g c a t t t g g g a g t g a a t g t g t t c t c g a a a t t a a c

30

40

50

c a a t t t c a a c a a c t t c t t c t t g g a c t a g a g c t c a c a g g a a  
 g g c c a t t t t t g g t g g c c a t g a a g c c g c c t a a g a a g t a t g a  
 a a c a a t a g a g t c g g c c t t a c c a g a a g g g t t t g a g a a g a g a  
 a c a a a a g g a a g g g g a a t c g t a c a t g a g g g t t g g g t g c a g c  
 a a c a a c t g a t a t ( 配列番号 1 1 9 ) 。

## 【 0 0 5 1 】

ホウレンソウ G T 2 アミノ酸遺伝子ポリペプチドは以下の通りである。m c d d k k s  
 s v l s i a f y p w f a l g h l t s f l r l a n k l a q n g h n v s y f i p t n  
 t l p r l l p h n h y p g h l t f i p v t v p p v d g l p l r a e t t n d v p s  
 s a i h l l m t a m d l t r d t i e a h l v s i k p d v v f y d f a y w i p d l  
 a r k h g f k s v l y i t s y i a r c a y f a p d l k s g h q s t g a e i i a p  
 p p g f p s q h f r m q a h e a e t v a d v g k e q d g l q g i t i s e r m r i  
 a f g k c d a i g v k s c k e m e k v y i d y c e k i f g k s v l l a g p m v p  
 k t p s s k l d e y f d g w l t g f g a a t v i y c a f g s e c v l e i n q f q  
 q l l l g l e l t g r p f l v a m k p p k k y e t i e s a l p e g f e k r t k g  
 r g i v h e g w v q q q l i l q h p s v g c f i t h c g v g s l s e a m v s k c  
 q v v l m p q a v d q f i n a r m m s l e l k i g v e v e k r e d d g l f t k e  
 a v h k a v s l v m e e e s e v a k e m r v s h d k w r e f l l q e g l e d s y  
 i s s f i q s l r q l t i g ( 配列番号 1 2 0 ) 。

## 【 0 0 5 2 】

数値 ( ± S D ) は 3 回の独立した生物学的実験の平均値を示す。対照植物 ( C H L H サイレンシング ) と比較して統計的に有意な差を示した ; \* P < 0 . 0 5 、 \*\* P < 0 . 0 1 、 \*\*\* P < 0 . 0 0 1 。

## 【 0 0 5 3 】

【 図 3 0 A 】サイレンシングされた ( V I G S ) グリコシルトランスフェラーゼを有するホウレンソウに蓄積されたサポニンの推定特性を示すデータである。S O A P 6 - S O A P 9 ( それぞれ、配列番号 5 5 、 5 7 、 5 9 、 6 1 ) 及び他の 2 つの G T をサイレンシングした植物における選択された化合物の相対含有量。数値 ( ± S D ) は 3 回の独立した生物学的実験の平均値。対照植物 ( C H L H サイレンシング ) と比較して統計的に有意な差を示した ; \*\* P < 0 . 0 1 、 \*\*\* P < 0 . 0 0 1 。

【 図 3 0 B 】サイレンシングされた ( V I G S ) グリコシルトランスフェラーゼを有するホウレンソウに蓄積されたサポニンの推定特性を示すデータである。主要なフラグメントを示した M S / M S スペクトル。

【 図 3 0 C 】サイレンシングされた ( V I G S ) グリコシルトランスフェラーゼを有するホウレンソウに蓄積されたサポニンの推定特性を示すデータである。蓄積されたサポニンの推定構造は、正確な質量測定及びフラグメンテーションパターンに基づいている。

【 図 3 1 A 】プラント ( ホウレンソウ ) 及びインビトロでの S O A P 1 0 アセチルトランスフェラーゼの特性を示す。V I G S でサイレンシングされた S O A P 1 0 ( 赤のトレース ) または対照としての C H L H ( 黒のトレース ) を有するホウレンソウの試料の L C - M S 分析。クロマトグラム青色及び黄色の領域には、それぞれ脱アセチル化及びアセチル化したサポニンが含まれている。ホウレンソウの葉で S O A P 1 0 をサイレンシングすると、アセチル化したサポニンの生成が減少し、脱アセチル化したサポニンの蓄積が増加した。

【 図 3 1 B 】プラント ( ホウレンソウ ) 及びインビトロでの S O A P 1 0 アセチルトランスフェラーゼの特性を示す。アシルトランスフェラーゼを V I G S でサイレンシングしたホウレンソウ、及び対照としての C H L H を有するホウレンソウにおける、選択した脱アセチル化サポニン ( Y o s s o s i d e I V 及び X I I ) 及びそのアセチル化対応物 ( Y o s s o s i d e V 及び V I I ) の相対含有量。縦軸は  $\log_{10}$  スケールである。

【 図 3 1 C 】プラント ( ホウレンソウ ) 及びインビトロでの S O A P 1 0 アセチルトランスフェラーゼの特性を示す。V I G S でサイレンシングされたホウレンソウ植物における

アセチル化サポニンに対する脱アセチル化サポニンの比率。SOAP10のサイレンシングにより、脱アセチル化サポニンの蓄積が増加した。

【図31D】プラント（ハウレンソウ）及びインビトロでのSOAP10アセチルトランスフェラーゼの特性を示す。インビトロでのSOAP10リコンビナントタンパク質の活性。基質（黒のトレース）-ハウレンソウ由来の脱アセチル化サポニンの精製画分、リコンビナントSOAP10をアセチル-CoA存在下で基質と共にインキュベートしたインビトロ酵素アッセイ（オレンジのトレース）、及び、比較のためのハウレンソウ葉の抽出物（緑のトレース）中の、脱アセチル化サポニン（ $m/z$  1101.51）及びアセチル化対応物（ $m/z$  1313.52）のEIC。点線は、クロマトグラムの領域を、左側及び右側の脱アセチル化サポニン及びアセチル化サポニンに、それぞれ分けている。「 $m/z = 42.01$ 」は、アセチル化によって生じた化合物の分子量の違いを示している。（SoAT1（配列番号109 [遺伝子] 及び配列番号113 [ポリペプチド]）、SoAT2（配列番号110 [遺伝子] 及び配列番号131 [ポリペプチド]）、SoAT3（配列番号111 [遺伝子] 及び配列番号115 [ポリペプチド]）、及びSoAT4（配列番号112 [遺伝子] 及び配列番号116 [ポリペプチド]）、機能的活性については表15を参照。）

10

【図31E】プラント（ハウレンソウ）及びインビトロでのSOAP10アセチルトランスフェラーゼの特性を示す。インビトロでのSOAP10リコンビナントタンパク質の活性。基質（黒のトレース）-ハウレンソウ由来の脱アセチル化サポニンの精製画分、リコンビナントSOAP10をアセチル-CoA存在下で基質と共にインキュベートしたインビトロ酵素アッセイ（オレンジのトレース）、及び、比較のためのハウレンソウ葉の抽出物（緑のトレース）中の、脱アセチル化サポニン（ $m/z$  1263.57）及びアセチル化対応物（ $m/z$  1305.58）のEIC。点線は、クロマトグラムの領域を、左側及び右側の脱アセチル化サポニン及びアセチル化サポニンに、それぞれ分けている。「 $m/z = 42.01$ 」は、アセチル化によって生じた化合物の分子量の違いを示している。（SoAT1（配列番号109 [遺伝子] 及び配列番号113 [ポリペプチド]）、SoAT2（配列番号110 [遺伝子] 及び配列番号131 [ポリペプチド]）、SoAT3（配列番号111 [遺伝子] 及び配列番号115 [ポリペプチド]）、及びSoAT4（配列番号112 [遺伝子] 及び配列番号116 [ポリペプチド]）、機能的活性については表15を参照。）

20

30

【図31F】プラント（ハウレンソウ）及びインビトロでのSOAP10アセチルトランスフェラーゼの特性を示す。インビトロでのSOAP10リコンビナントタンパク質の活性。基質（黒のトレース）-ハウレンソウ由来の脱アセチル化サポニンの精製画分、リコンビナントSOAP10をアセチル-CoA存在下で基質と共にインキュベートしたインビトロ酵素アッセイ（オレンジのトレース）、及び、比較のためのハウレンソウ葉の抽出物（緑のトレース）中の、脱アセチル化サポニン（ $m/z$  1395.61）及びアセチル化対応物（ $m/z$  3137.62）のEIC。点線は、クロマトグラムの領域を、左側及び右側の脱アセチル化サポニン及びアセチル化サポニンに、それぞれ分けている。「 $m/z = 42.01$ 」は、アセチル化によって生じた化合物の分子量の違いを示している。（SoAT1（配列番号109 [遺伝子] 及び配列番号113 [ポリペプチド]）、SoAT2（配列番号110 [遺伝子] 及び配列番号131 [ポリペプチド]）、SoAT3（配列番号111 [遺伝子] 及び配列番号115 [ポリペプチド]）、及びSoAT4（配列番号112 [遺伝子] 及び配列番号116 [ポリペプチド]）、機能的活性については表15を参照。）

40

【図32A】ベンサミアナタバコにおけるSOAP1-4及び6-10の発現のデータを示す図である。SOAP1-4単独、またはSOAP5（配列番号65または93 [遺伝子]、配列番号66 [ポリペプチド]）を含まない9つのSOAPすべてを一過性発現させた植物から得られたメディカゲン酸及びその（ポリ）グリコシル化誘導体のEIC [ $m/z = 501.32$  (MA);  $m/z = 663.38$  (MA + hex);  $m/z = 825.43$  (MA + 2 hex);  $m/z = 987.48$  (MA + 3 hex);  $m/z = 114$

50

9.53 (MA + 4hex) 陰イオン化モード] を、対照と比較した。挿入図は、クロマトグラム領域を20倍に拡大したものである。

【図3.2B】ベンサミアナタバコにおけるSOAP1-4及び6-10の発現のデータを示す図である。トリ-ヘキソシルメディカゲン酸のMSEスペクトルを示す。矢印はヘキソシル部分の消失を示している。

【図3.2C】ベンサミアナタバコにおけるSOAP1-4及び6-10の発現のデータを示す図である。ビス-ヘキソシルメディカゲン酸のMSEスペクトルを示す。矢印はヘキソシル部分の消失を示している。

【図3.2D】ベンサミアナタバコにおけるSOAP1-4及び6-10の発現のデータを示す図である。モノ-ヘキソシルメディカゲン酸のMSEスペクトルを示す。矢印はヘキソシル部分の消失を示している。

10

【図3.3A】セルロース合成酵素様G遺伝子(SOAP5; 配列番号65または93[遺伝子]、配列番号66[ポリペプチド])の技術的解析結果を示す図である。SOAP5がコードする酵素は、トリテルペノイドのアグリコンにグルクロン酸を結合させ、ハウレンソウのサポニン生合成経路において重要な役割を果たしている。ウイルス誘発性遺伝子サイレンシング(VIGs)後のハウレンソウの葉におけるSOAP5の相対発現の分析。提示された値は、4つの独立した生物学的複製から得られた。SOAP5の発現を、対照(マグネシウムケラターゼサブユニットHのみをサイレンシングした植物-CHLH)と比較した。

【図3.3B】セルロース合成酵素様G遺伝子(SOAP5; 配列番号65または93[遺伝子]、配列番号66[ポリペプチド])の技術的解析結果を示す図である。SOAP5がコードする酵素は、トリテルペノイドのアグリコンにグルクロン酸を結合させ、ハウレンソウのサポニン生合成経路において重要な役割を果たしている。SOAP5及び他のグリコシルトランスフェラーゼをサイレンシングしたハウレンソウの葉におけるメディカゲン酸の蓄積量を対照(CHLH)と比較した。有意差が示された; \*\*\* $P < 0.001$ 。

20

【図3.3C】セルロース合成酵素様G遺伝子(SOAP5; 配列番号65または93[遺伝子]、配列番号66[ポリペプチド])の技術的解析結果を示す図である。SOAP5がコードする酵素は、トリテルペノイドのアグリコンにグルクロン酸を結合させ、ハウレンソウのサポニン生合成経路において重要な役割を果たしている。SOAP1-5を一過性発現させた植物から得られたメディカゲン酸3-O-グルクロニド[m/z = 677.35 (MA-3-GlcA)、陰イオン化モード]のEICを、SOAP1-4のみを発現させた植物及びコントロール(空ベクター-EV)と比較した。

30

【図3.3D】セルロース合成酵素様G遺伝子(SOAP5; 配列番号65または93[遺伝子]、配列番号66[ポリペプチド])の技術的解析結果を示す図である。SOAP5がコードする酵素は、トリテルペノイドのアグリコンにグルクロン酸を結合させ、ハウレンソウのサポニン生合成経路において重要な役割を果たしている。SOAP1-5を発現する酵母、酵母対照(SOAP1-4 + SoUGD1; SOAP5単独及びEVで形質転換した酵母細胞)と比較した、SOAP1-5 + SoUGD1を発現する酵母からのMA-3-GlcA[m/z = 677.35、陰イオン化モード]のEIC。

40

【図3.3E】セルロース合成酵素様G遺伝子(SOAP5; 配列番号65または93[遺伝子]、配列番号66[ポリペプチド])の技術的解析結果を示す図である。SOAP5がコードする酵素は、トリテルペノイドのアグリコンにグルクロン酸を結合させ、ハウレンソウのサポニン生合成経路において重要な役割を果たしている。ハウレンソウにおけるSOAP5活性の模式図。SOAP5は、グルクロン酸(赤で強調)をC-3位のメディカゲン酸に結合させる。

【図3.3F】セルロース合成酵素様G遺伝子(SOAP5; 配列番号65または93[遺伝子]、配列番号66[ポリペプチド])の技術的解析結果を示す図である。SOAP5がコードする酵素は、トリテルペノイドのアグリコンにグルクロン酸を結合させ、ハウレンソウのサポニン生合成経路において重要な役割を果たしている。ベンサミアナタバコに

50

におけるハウレンソウサポニン生合成経路の再構成。10のSOAP遺伝子をすべて一過性発現させたベンサミアナタバコの葉の試料におけるYossoside V [m/z = 1305.57、陰イオン化モード]のEICを、対照植物 (EV) 及びハウレンソウの葉の抽出物と比較した。

【図34A】ハウレンソウでSOAP5をサイレンシング (VIGS) した結果を示すデータである。対照 (CHLHサイレンシング) 及び野生型 (WT) と比較した、SOAP5 (配列番号65または93) をサイレンシングした植物からのメディカゲン酸のEIC [m/z = 501.3、陰イオン化モード]。インプラント実験で得られたクロマトグラムのみを縮尺通りに表示している。

【図34B】ハウレンソウでSOAP5をサイレンシング (VIGS) した結果を示すデータである。植物に蓄積されたメディカゲン酸 [25V、陽イオンモード] のタンデム質量分析 (MS/MS) を、基準の標準物質と比較した。MA - メディカゲン酸。

【図35A】ベンサミアナタバコにおけるSOAP5 (配列番号65または93) の一時的な発現を示すデータの図である。SOAP1-5 (それぞれ、配列番号45、46、51、53及び65) を一過性発現させた植物から得たメディカゲン酸及びメディカゲン酸3-グルクロニドのEIC [m/z = 501.32 (MA); m/z = 677.35 (MA-3-GlcA)、陰イオン化モード] を、SOAP1-4 (それぞれ、配列番号45、46、51及び53) を単独で発現させた植物、SOAP5 (配列番号65または93) を単独で発現させた植物、及び対照 (空ベクター EV) と比較した。クロマトグラムは縮尺の通り。

【図35B】ベンサミアナタバコにおけるSOAP5 (配列番号65または93) の一時的な発現を示すデータの図である。フラグメントルートをマークしたMA-3-GlcAの構造。

【図35C】ベンサミアナタバコにおけるSOAP5 (配列番号65または93) の一時的な発現を示すデータの図である。MA-3-GlcAのMS/MS [45V、陰イオン化モード]、矢印はグルクロン酸部分の消失を示す。

【図35D】ベンサミアナタバコにおけるSOAP5 (配列番号65または93) の一時的な発現を示すデータの図である。メディカゲン酸3-O-グルクロニド (MA-3-GlcA) の水素の重要なHMBC相関を示している。黒い矢印は、MA-3-GlcAの6つのメチル基のHMBC相関を示す。赤い矢印は、MA-3-GlcAの他の (選択された) 水素のHMBC相関を表す。数字は、メディカゲン酸及びグルクロン酸部分の炭素原子に対応する。詳細は表1を参照。

【図36A】SOAP5 (配列番号66) の基質特異性を示す図である (表16参照)。SOAP1/2/5を一過性発現させた植物から得られたオレアノール酸3-GlcAのEIC [m/z = 631.38 (OA-3-GlcA)、陰イオン化モード] を、SOAP5を含まない同一のSOAPセットを発現させた植物、及び対照 (EV) と比較した。クロマトグラムは縮尺が異なる。

【図36B】SOAP5 (配列番号66) の基質特異性を示す図である (表16参照)。SOAP1/2/3/5を一過性発現させた植物から得られたアウグスチン酸3-GlcAのEIC [m/z = 647.38 (AA-3-GlcA)、陰イオン化モード] を、SOAP5を含まない同一のSOAPセットを発現させた植物、及び対照 (EV) と比較した。クロマトグラムは縮尺が異なる。

【図36C】SOAP5 (配列番号66) の基質特異性を示す図である (表16参照)。SOAP1/2/4/5を一過性発現させた植物から得られたギブソゲン酸3-GlcAのEIC [m/z = 661.35 (GA-3-GlcA)、陰イオン化モード] を、SOAP5を含まない同一のSOAPセットを発現させた植物、及び対照 (EV) と比較した。クロマトグラムは縮尺が異なる。

【図36D】SOAP5 (配列番号66) の基質特異性を示す図である (表16参照)。図36Aからのグルクロン酸誘導体のタンデム質量分析 (MS/MS) を示す [40V、陰イオン化モード]、矢印はグルクロン酸の消失を示す (m/z = 176.03; Glc

10

20

30

40

50

A - H<sub>2</sub>O)。

【図36E】SOAP5 (配列番号66)の基質特異性を示す図である(表16参照)。図36Bからのグルクロン酸誘導体のタンデム質量分析(MS/MS)を示す[40V、陰イオン化モード]、矢印はグルクロン酸の消失を示す( $m/z = 176.03$ ; GlcA - H<sub>2</sub>O)。

【図36F】SOAP5 (配列番号66)の基質特異性を示す図である(表16参照)。図36Cからのグルクロン酸誘導体のタンデム質量分析(MS/MS)を示す[40V、陰イオン化モード]、矢印はグルクロン酸の消失を示す( $m/z = 176.03$ ; GlcA - H<sub>2</sub>O)。

【図37A】ベンサミアナタバコにおけるSOAP活性の生成物を示す図である。クロマトグラムに示されたSOAP遺伝子の組み合わせを一過性発現させた植物から得たYossoside I ( $m/z = 823.42$ )のEIC。MS/MSスペクトルについては、図40Bを参照。

10

【図37B】ベンサミアナタバコにおけるSOAP活性の生成物を示す図である。クロマトグラムに示されたSOAP遺伝子の組み合わせを一過性発現させた植物から得たYossoside II ( $m/z = 969.27$ )のEIC。MS/MSスペクトルについては、図40Bを参照。

【図37C】ベンサミアナタバコにおけるSOAP活性の生成物を示す図である。クロマトグラムに示されたSOAP遺伝子の組み合わせを一過性発現させた植物から得たYossoside III ( $m/z = 1131.52$ ) Yossoside II ( $m/z = 969.27$ )のEIC。MS/MSスペクトルについては、図40Bを参照。

20

【図37D】ベンサミアナタバコにおけるSOAP活性の生成物を示す図である。クロマトグラムに示されたSOAP遺伝子の組み合わせを一過性発現させた植物から得たYossoside IV ( $m/z = 1263.57$ )のEIC。MS/MSスペクトルについては、図40Bを参照。

【図37E】ベンサミアナタバコにおけるSOAP活性の生成物を示す図である。クロマトグラムに示されたSOAP遺伝子の組み合わせを一過性発現させた植物から得たYossoside V ( $m/z = 1305.58$ )のEIC。MS/MSスペクトルについては、図40Bを参照。

【図38A】ベンサミアナタバコにおけるSOAP活性の主要な生成物の質量分析(MS)による同定を示す図である。MA-3-GlcAに対するSOAP6活性の生成物であるYossoside I (化合物7)のMS/MS。

30

【図38B】ベンサミアナタバコにおけるSOAP活性の主要な生成物の質量分析(MS)による同定を示す図である。Yossoside Iに対するSOAP7活性の生成物であるYossoside II (化合物8)のMS/MS。

【図38C】ベンサミアナタバコにおけるSOAP活性の主要な生成物の質量分析(MS)による同定を示す図である。Yossoside IIに対するSOAP8活性の生成物であるYossoside III (化合物9)のMS/MS。

【図38D】ベンサミアナタバコにおけるSOAP活性の主要な生成物の質量分析(MS)による同定を示す図である。Yossoside IIIに対するSOAP9活性の生成物であるYossoside IV (化合物10)のMS/MS。

40

【図38E】ベンサミアナタバコにおけるSOAP活性の主要な生成物の質量分析(MS)による同定を示す図である。Yossoside IVに対するSOAP10活性の生成物であるYossoside V (化合物11)のMS/MS。

【図39A】ベンサミアナタバコにおけるSOAP9活性の結果を示す図である。SOAP1-5+SOAP9を一過性発現させた植物から得られたメディカゲン酸3-GlcA-Xyl [  $m/z = 809.40$ 、陰イオン化モード]のEICを、SOAP1-5のみを発現させた植物及び対照(EV)と比較した。表16参照。

【図39B】ベンサミアナタバコにおけるSOAP9活性の結果を示す図である。809.4 (45V)のMS/MS、矢印はglcA-xyl部分の消失を示す( $m/z = 30$

50

8.07)。

【図39C】ベンサミアタバコにおけるSOAP9活性の結果を示す図である。SOAP9生成物(化合物10)の構造、矢印で示されたフラグメンテーションパターン。

【図40A】SOAP5を酵母で発現させた結果を示す図である。SOAP1-5を発現する植物、酵母対照、及びメディカゲン酸の基準の標準物質と比較した、SOAP1-5+UGD1(配列番号74[遺伝子]及び配列番号75[タンパク質])を発現する酵母からのメディカゲン酸及びメディカゲン酸3-GlcA[m/z=501.32(MA); m/z=677.35(MA-3-GlcA)、陰イオン化モード]のEIC。クロマトグラムは縮尺通り。

【図40B】SOAP5を酵母で発現させた結果を示す図である。酵母細胞から得られたメディカゲン酸3-GlcAのMS/MS[45V、陰イオン化モード]、矢印はグルクロン酸の消失を示す。

【図41A】酵母細胞におけるSOAP5の基質特異性を示した図である。SOAP1-5+UGD1を発現させた酵母から得られたオレアノール酸3-GlcA[m/z=631.38(OA-3-GlcA)、陰イオン化モード]のEICを、SOAP1-4+UGD1単独、SOAP1-5単独、SOAP1-4、SOAP1-2、SOAP5単独で発現させた細胞、及び対照(EV)と比較した。クロマトグラムは縮尺通り。

【図41B】酵母細胞におけるSOAP5の基質特異性を示した図である。SOAP1-5+UGD1を発現させた酵母から得られたギブソゲン酸3-GlcA[m/z=661.35(GA-3-GlcA)、陰イオン化モード]のEICを、SOAP1-4+UGD1単独、SOAP1-5単独、SOAP1-4、SOAP1-2、SOAP5単独で発現させた細胞、及び対照(EV)と比較した。クロマトグラムは縮尺通り。

【図41C】酵母細胞におけるSOAP5の基質特異性を示した図である。SOAP1-5+UGD1を発現させた酵母から得られたバヨゲニン3-GlcA[m/z=663.37(バヨゲニン-3-GlcA)、陰イオン化モード]のEICを、SOAP1-4+UGD1単独、SOAP1-5単独、SOAP1-4、SOAP1-2、SOAP5単独で発現させた細胞、及び対照(EV)と比較した。クロマトグラムは縮尺通り。

【図41D】酵母細胞におけるSOAP5の基質特異性を示した図である。図41Aのグルクロン酸誘導体のMS/MS[45V、陰イオン化モード]、矢印はグルクロン酸の消失を示す(m/z=176.03; GlcA-H<sub>2</sub>O)。

【図41E】酵母細胞におけるSOAP5の基質特異性を示した図である。図41Bのグルクロン酸誘導体のMS/MS[45V、陰イオン化モード]、矢印はグルクロン酸の消失を示す(m/z=176.03; GlcA-H<sub>2</sub>O)。

【図41F】酵母細胞におけるSOAP5の基質特異性を示した図である。図41Cのグルクロン酸誘導体のMS/MS[45V、陰イオン化モード]、矢印はグルクロン酸の消失を示す(m/z=176.03; GlcA-H<sub>2</sub>O)。

【図42A】セルロース合成酵素様G(CSLG)酵素のグルクロノシルトランスフェラーゼ活性が、系統的に離れた目に属する種の間で保存されていることを示す図である。系統樹は、ハウレンソウ由来のSOAP5と、ナデシコ目、アオイ科、セリ科、マメ目、及びナス目のCSLGタンパク質との密接な進化的関係を示している。系統樹は、セルロース合成酵素(CESA)に属するタンパク質と、セルロース合成酵素様酵素(CSLA、CSLB、CSLE及びCSLG)のいくつかのファミリーからなる。点線はCSLGタンパク質を含むクレードを囲み、吹き出しはGAME15やマメ目のCSLGを含むハウレンソウSOAP5に近縁のCSLGに対応している。

【図42B】セルロース合成酵素様G(CSLG)酵素のグルクロノシルトランスフェラーゼ活性が、系統的に離れた目に属する種の間で保存されていることを示す図である。メディカゲン酸は多くの種のCSLGによってグルクロン酸化される。SOAP1-4を単独で発現している植物及び対照(EV)と比較した、ハウレンソウ(SOAP5; 配列番号66)、キノア(CqCSL; 配列番号96)、ビートルート(BvCSL; 配列番号94)、アルファルファ(MsCSL; 配列番号98)、ダイズ(GmCSL; 配列番号

10

20

30

40

50

100)、ウラルカンゾウ (GuCSL; 配列番号 81 または配列番号 102)、ミヤコグサ (LjCSL; 配列番号 104) 由来の CSLG タンパク質に結合した SOAP1-4 を一過性発現させたベンサミアナタバコの葉の試料の MA-3-GlcA [m/z = 677.35、陰イオン化モード] の EIC。

【図 4 2 C】セルロース合成酵素様 G (CSLG) 酵素のグルクロノシルトランスフェラーゼ活性が、系統的に離れた目に属する種の間で保存されていることを示す図である。セルロース合成酵素様 G タンパク質は、複数の植物種においてトリテルペノイドサポニンの生合成の鍵となる酵素である。 - アミリン (中央) はシトクロム P450 (青丸) によって酸化され、CSLG (ピンクの丸) によって装飾された複数のアグリコン (1 - バヨゲニン、2 - セルジャン酸、3 - オレアノール酸、4 - メディカゲン酸、5 - グリチルレチン酸、6 - ソヤサポゲノール A、7 - ソヤサポゲノール B) を生成し、他のグリコシルトランスフェラーゼは、トリテルペノイドサポニン (8 - バヨゲニン - ヘキサ A - ヘキサ - ヘキサ (アルファルファ)、9 - セルジャン酸 (キノア (C. quinoa))、10 - beta vulgaroside IV (B. vulgaris)、11 - Yossoside V (ハウレンソウ)、12 - グリチルリチン (ウラルカンゾウ)、13 - ソヤサポゲノール A - ヘキサ A - ヘキサ - ペント (ミヤコグサ)、14 - ソヤサポニン VI (ダイズ)) を生成する。

10

【図 4 2 D】セルロース合成酵素様 G (CSLG) 酵素のグルクロノシルトランスフェラーゼ活性が、系統的に離れた目に属する種の間で保存されていることを示す図である。ウラルカンゾウにおけるグリチルリチンの生合成経路。強調された官能基は、特定の酵素の活性の結果である。

20

【図 4 2 E】セルロース合成酵素様 G (CSLG) 酵素のグルクロノシルトランスフェラーゼ活性が、系統的に離れた目に属する種の間で保存されていることを示す図である。4 つの遺伝子を発現している試料 (- UGT73P12; (配列番号 84))、及び対照 (bAS (配列番号 45) + CYP88D6 (配列番号 76) + CYP72A154 (配列番号 78)) と比較した、グリチルリチン生合成経路の 5 つの遺伝子 (bAS (配列番号 45) + CYP88D6 (配列番号 76) + CYP72A154 (配列番号 78) + GuCSL (配列番号 80 or 配列番号 103)、及び UGT73P12 (配列番号 84)) を一過性発現しているベンサミアナタバコの葉の試料から抽出したグリチルレチン酸 3-O-グルクロニド及びグリチルリチン [m/z 645.36、m/z 821.40 (負イオンモード)] の抽出イオンクロマトグラム (EIC) の測定結果。

30

【図 4 2 F】セルロース合成酵素様 G (CSLG) 酵素のグルクロノシルトランスフェラーゼ活性が、系統的に離れた目に属する種の間で保存されていることを示す図である。セルロース合成酵素様 (CSL) ポリペプチドの系統樹を、セルロース合成酵素様 G (CSLG) に特に重点を置いて提示している。青色の背景領域は、多くの植物目 (ナデシコ目、マメ目、セリ目、アオイ目、ナス目) の CSLG 酵素からなる。CSLG 酵素 (青色の背景) の中で、関連するサブクレードとしては、トリテルペノイドアグリコンに対するグルクロノシルトランスフェラーゼ活性が証明された CSLG 酵素 (太字) を含む赤色で描かれたサブクレード (大豆、甘草、ミヤコグサ、アルファルファ、レッドビート、キノア、ハウレンソウ由来の CSLG)、及び、ナス目 (ジャガイモ、トマト、ビタースイート) の CSLG からなる青色で描かれたサブクレードなどが含まれる; 太字の酵素はステロイド系アルカロイド及びステロイド系サポニンの生合成に関与している。

40

【図 4 3 A】キヌアにおけるサポニンの LC-ESI-QTOF-MS 分析を示す図である。キヌア (Chenopodium quinoa var. Read Head) の生後 1 ヶ月の植物の 80% メタノール抽出物の全イオンクロマトグラム (TIC)。半透明の四角は、サポニンを含むクロマトグラム領域を示す。推定される特徴については表 12 を参照。

【図 4 3 B】ビートルートにおけるサポニンの LC-ESI-QTOF-MS 分析を示す図である。ビートルート (Beta vulgaris var. Bohan) の生後 1 ヶ月のビーツの 80% メタノール抽出物の全イオンクロマトグラム (TIC)。半透明の四角は、サポニンを含むクロマトグラム領域を示す。推定される特徴については表 12 を参照。

50

【図43C】アルファルファにおけるサポニンのLC-ESI-QTOF-MS分析を示す図である。アルファルファ (*Medicago sativa*) の生後2週間の苗の80%メタノール抽出物の全イオンクロマトグラム (TIC)。半透明の四角は、サポニンを含むクロマトグラム領域を示す。推定される特徴については表12を参照。

【図44A】レッドビートにおけるCSLGのウイルス誘導遺伝子サイレンシング (VIGS) を示す図である。レッドビートにおけるウイルス誘導遺伝子サイレンシング; EV-空のベクター、CHLH-マグネシウムキラーゼサブユニットH、BVCSL-ビートルートセルロース合成酵素様G (配列番号95 [遺伝子]、VIGsサイレンシング配列 配列番号107)。

【図44B】レッドビートにおけるCSLGのウイルス誘導遺伝子サイレンシング (VIGS) を示す図である。サイレンシングされたBVCSLを有するレッドビート植物におけるトリテルペノイド系サポニン及びオレオノール酸の相対的な含有量 (ピーク面積) の分析。縦軸はlog10スケールである。数値は3回の独立した生物学的実験の平均値。対照植物 (CHLHサイレンシング単独) と比較して統計的に有意な差を示した; \* P < 0.05、\*\*\* P < 0.001。

【図44C】レッドビートにおけるCSLGのウイルス誘導遺伝子サイレンシング (VIGS) を示す図である。対照 (CHLH及びWT) 及び基準の標準物質と比較した、BVCSLがサイレンシングされた植物から得られたオレオノール酸 [m/z = 455.35、陰イオン化モード] のEIC (代謝物の同定については、表12のアルファルファを参照)。

【図45A】アルファルファの毛状根におけるCSLGのサイレンシングを示す図である。対照 (EV) と比較した、サイレンシングされたアルファルファセルロース合成酵素様 (MsCSL) を有する毛状根の培養物から得られた、アグリコンに結合したグルクロン酸を含む代表的なサポニンのEIC。

【図45B】アルファルファの毛状根におけるCSLGのサイレンシングを示す図である。サイレンシングされたMsCSLを有するアルファルファの毛状根におけるトリテルペノイド系サポニンの分析。縦軸はlog10スケール。数値は5回の独立した生物学的実験の平均値。対照 (EVのみで形質転換した根) と比較して統計的に有意な差を示す; \* P < 0.05、\*\*\* P < 0.01、\*\*\* P < 0.001。(asRNA-アンチセンスRNA) 使用したasRNAの配列を以下に記す。

【0054】

c t a c c c a c t c t t c c g c t t c a t a g a t a t a a t c c c c t c c a c  
a a t a g g a t g a c t t a g a g c a a g t a t a t a g a t a a t g a g g a a a  
a g t t g a c c a a a c a t t t c a t c a a a a t c c c t c a c a t t g a g t a  
g t c t c c a t a a a c c a c c a a a g a a g c a a a c a a t a t t t a t t g t  
g a g t a a c a c a a t c a c t g g t g c c a t g a g c a a t g c t g c a c c a  
t c a a a a t t a a a c t t a c c c t g c t c a t a t t t c t t t a g c t t a t  
c t t t g t c a a t t g c t t t g t c c g a t a a a g t g a a t t t t g c c t t  
g t t c a a c c c a a a c t t t t t t t c g c t g t t t c t a t a a t t g c g  
a a a a c g c a c c c a a t t g a t t t t a c a a t c c c a c t t c t t t g t t  
c a t t c c a c c a a a t c a t c g a g g a g c c a c c a g t t g a a a t t a c  
c t c a a t a t a a t g a t g a a t t t g a c t g g a t a g a t a c a a t a c t  
g t a a a c a c t a t a a a c c a a g g a t c t g t a a c c t t t g g a t a t a  
c a g g t a t t c c c t t c a a g a a g c a a a t c t g a g g a a c g a a t c c  
a t a t a a g a t g t a g g c t g t t g c a a a t t g g g t t g t g c t t a c c  
a a g t a a c a a a a a g t t a a a c a a t g a a t a g t g g g c a a t c t t g  
a a a g g c c a t a a g t g a a t g g g c t a t a t t t a g a g a t t g c a a g  
c a a c a a a a g t t c a g a t g a c c a c t t t a t c g g t t g a a t c a a t  
c c c t c t t t c a t g t c a g t t g g a g c a c a t c c t a a g a a a c a t g  
g t c t t t t a g g g t a a a g a t a a g t t g a t c t c c a t c c t c t a c a

10

20

30

40

50

a t g g a g g a g a t a g c c a g t a a c g g t a c t c t c a a g t t t t a t a  
 g c a t a c g a g a a t c c c a c c t c a t t a c c c c a a t t t g t g t t t c  
 t c t c a t a g g a a c a a g a a g c c a c t t c a c a t g c t t c t t g t a a  
 a a t t a c a t c t c t t g a a a t a t t c t g c t t c t t a g t t t g t t g a  
 c c a c g t a a c g c t t t t a g t g a t t c t a c a t a c a t g t t a g a c t  
 t g c c a a a g t t g t a t a g a g c a t c a a g t a g a t a g t c c c c t t t  
 t t g g t t t g g a c t t c c a a a g a g t a a t (配列番号108)。

【0055】

【図46A】GuCSLGによって触媒されるグリチルレチン酸3-O-モノグルクロニドの形成を示す図である。ウラルカンゾウ由来のタンパク質を発現するベンサミアナタバコから得られた試料のグリチルレチン酸3-O-モノグルクロニドのEIC [m/z = 645.36、陰イオン化モード]。bAS - - アミリン合成酵素 (配列番号48)、CYP88D6 - - アミリン11-オキシダーゼ (配列番号77)、CYP72A154-11-オキシ-b-アミリン30-オキシダーゼ (配列番号79)、GuUGAT (配列番号83) - グリチルレチン酸のグルクロニド化を行うことが報告されているグリコシルトランスフェラーゼ、UGT73P12 (配列番号85) - グリチルレチン酸3-O-モノグルクロニドにglcAを転移するグリコシルトランスフェラーゼ、GuCSL-セルロース合成酵素様G。 10

【図46B】GuCSLGによって触媒されるグリチルレチン酸3-O-モノグルクロニドの形成を示す図である。MS/MSは654.36 (グリチルレチン酸3-O-モノグルクロニドの異性体1; 55V)である。 20

【図46C】GuCSLGによって触媒されるグリチルレチン酸3-O-モノグルクロニドの形成を示す図である。MS/MS654.36 (異性体2; 55V)。

【図46D】GuCSLGによって触媒されるグリチルレチン酸3-O-モノグルクロニドの形成を示す図である。グリチルレチン酸3-O-モノグルクロニドの構造、及び矢印で示されたフラグメンテーションパターン。

【図46E】GuCSLGによって触媒されるグリチルレチン酸3-O-モノグルクロニドの形成を示す図である。検出されたイオンによるグルクロン酸のクロシングフラグメンテーションの命名法。

【図47A】GuCSLGを必要とするベンサミアナタバコにおけるグリチルリチン (グリチルレチン酸3-O-ジグルクロニド) の形成を示す図である。ウラルカンゾウ由来のタンパク質を発現させたベンサミアナタバコの試料のグリチルリチンのEIC (m/z = 821.40、陰イオン化モード)。bAS - b-アミリン合成酵素 [ - アミリン合成酵素] (配列番号48)、CYP88D6 - b-アミリン11オキシダーゼ (配列番号76 [遺伝子] 及び配列番号77 [ポリペプチド])、CYP72A154-11-オキシ-b-アミリン47-オキシダーゼ (配列番号78 [遺伝子] 及び配列番号79 [ポリペプチド])、GuUGAT (配列番号82 [遺伝子] 及び配列番号83 [ポリペプチド]) - グリチルレチン酸のグルクロン化を行うことが報告されているグリコシルトランスフェラーゼ、UGT73P12 - グリチルレチン酸3-O-モノグルクロニドにglcAを転移するグリコシルトランスフェラーゼ (配列番号84 [遺伝子] 及び配列番号85 [ポリペプチド])、GuCSL-セルロース合成酵素様G (配列番号80または配列番号103 [遺伝子] 及び配列番号81または配列番号102 [ポリペプチド])。 30 40

【図47B】GuCSLGを必要とするベンサミアナタバコにおけるグリチルリチン (グリチルレチン酸3-O-ジグルクロニド) の形成を示す図である。821.40 (異性体1) - bAS (配列番号45) + CYP88D6 (配列番号77) + CYP72A154 (配列番号79) + GuCSL (配列番号81または配列番号102) + UGT73P12 (配列番号85) の生成物のMS/MS (65V)。

【図47C】GuCSLGを必要とするベンサミアナタバコにおけるグリチルリチン (グリチルレチン酸3-O-ジグルクロニド) の形成を示す図である。821.40 (異性体2) - bAS (配列番号45) + CYP88D6 (配列番号77) + CYP72A154 50

(配列番号79) + GuCSL (配列番号81または配列番号102) + UGT73P12 (配列番号85)の生成物のMS/MS(65V)。

【図47D】GuCSLGを必要とするベンサミアナタバコにおけるグリチルリチン(グリチルレチン酸3-O-ジグルクロニド)の形成を示す図である。グリチルリチン基準標準物質、及び異性体1のMS/MS。

【図47E】GuCSLGを必要とするベンサミアナタバコにおけるグリチルリチン(グリチルレチン酸3-O-ジグルクロニド)の形成を示す図である。グリチルリチン基準標準物質、及び異性体2のMS/MS。

【図47F】GuCSLGを必要とするベンサミアナタバコにおけるグリチルリチン(グリチルレチン酸3-O-ジグルクロニド)の形成を示す図である。グリチルリチンの構造、及び矢印で示した断片化パターン。

【図48】SOAP5の構造を示す図である。シロイヌナズナのAtCESA1(配列番号67)、AtCESA3(配列番号68)、及びホウレンソウのそれらのオーソログ(SoCESA1(配列番号70)及びSoCESA3(配列番号71))のタンパク質配列を、SOAP5(配列番号66)及びシロイヌナズナのその対応物であるAtCSLG1(配列番号69)と共にアライメントした。

【図49A】SOAP5の予測される3D構造を示す図である。メディカゲン酸が酵素の活性部位に到達し(青-1、2、3)、脂質二重層の疎水性環境から離脱する(黄-4)ことを可能にするトンネルが予測されたSOAP5の3Dモデル。

【図49B】SOAP5の予測される立体構造を示す図である。メディカゲン酸が酵素の活性部位に到達し(青-1、2、3)、脂質二重層の疎水性環境から離脱する(黄-4)ことを可能にするトンネルが予測されたSOAP5の3Dモデル。

【図49C-1】SOAP5の予測される立体構造を示す図の1つである。図49Aで示したトンネルのプロファイル及び特性。

【図49C-2】SOAP5の予測される立体構造を示す図の1つである。図49Aで示したトンネルのプロファイル及び特性。

【図50A】SOAP5の部位特異的変異誘発データを示す図である。陰性のSOAP5、SOAP1-4単独、及び対照(EV)と比較した、SOAP5の変異体を有するSOAP1-4を一過性発現させた植物から得られたメディカゲン酸[m/z=501.32]及びメディカゲン酸3-GlcA[m/z=677.35]のEIC。

【図50B】SOAP5の部位特異的変異誘発データを示す図である。遊離MAに対するMA-3-GlcAの比率として表した、SOAP5の変異体を発現させた植物におけるMA及びMA-3-GlcAの蓄積。縦軸はlog<sub>10</sub>である。数値は3回の独立した生物学的実験の平均値。陰性のSOAP5を発現している植物と比較して統計的に有意な差があることを示している；\*\*\*P<0.001。

【図51-1】(A)-(F)からなる、SOAP5を含む、小胞体(ER)膜に共局在しているSOAPタンパク質を示す顕微鏡写真である。ベンサミアナタバコの葉で一次的に共発現させたSOAP5:RFPと、(A)ER:GFPマーカー、(B)ゴルジ:GFPマーカー、(C)SOAP1:GFP、(D)SOAP2:GFP、(E)SOAP3:GFP及び(F)SOAP4:GFPとの共焦点画像。各パネルには、GFP、RFP、及び両方のチャンネルを合成した画像が含まれている。スケールバーは50µm。

【図51-2】(G)及び(H)からなる、SOAP5を含む、小胞体(ER)膜に共局在しているSOAPタンパク質を示す顕微鏡写真である。(G)保存されたモチーフ(赤色)、置換されたアミノ酸(青色、矢印で示す)、及び6つの膜貫通ドメインを示したSOAP5のトポロジーモデル。(H)SOAP5を含む、小胞体(ER)膜に共局在しているSOAPタンパク質を示す顕微鏡写真である。示された構築物を発現させた細胞のFRITインデックス。値は平均値±SE。GFP及びRFPを発現させた陰性対照(NC)と比較した有意差を示す；\*P<0.05、\*\*P<0.01。

【図52-1】(A)-(D)からなる、トリテルペノイドアグリコンの形成に關与するSOAPが小胞体(ER)に共在して存在することを示す顕微鏡写真である。ベンサミア

10

20

30

40

50

ナタバコの葉に一過性共発現させた蛍光タンパク質 (GFP 及び RFP) と融合した SOAP の共焦点画像。(A) SOAP1 : GFP + SOAP2 : RFP。(B) SOAP1 : GFP + SOAP3 : RFP。(C) SOAP1 : GFP + SOAP4 : RFP。(D) SOAP2 : GFP + SOAP3 : RFP。パネルには、GFP、RFP、及び両方のチャンネルを合成した画像が含まれている。スケールバー - 50 ピコメートル。

【図52-2】(E) 及び (F) からなる、トリテルペノイドアグリコンの形成に關与する SOAP が小胞体 (ER) に共在して存在することを示す顕微鏡写真である。ベンサミアタバコの葉に一過性共発現させた蛍光タンパク質 (GFP 及び RFP) と融合した SOAP の共焦点画像。(E) SOAP2 - GFP + SOAP4 : RFP。(F) SOAP3 : GFP + SOAP4 : RFP。パネルには、GFP、RFP、及び両方のチャンネルを合成した画像が含まれている。スケールバー - 50 ピコメートル。

10

【図53】蛍光タンパク質 (FP) と融合したが、依然として活性を有する SOAP を示すデータである。陰性の SOAP5、SOAP1-4 単独、及び対照 (EV) と比較した、SOAP1 : eGFP、SOAP2 : eGFP、SOAP3 : mRFP、SOAP4 : eGFP、SOAP5 : mRFP を一過性発現させたベンサミアタバコ植物から得られたメディカゲン酸 [m/z = 501.32] 及びメディカゲン酸 3 - GlcA [m/z = 677.35] の EIC である。SOAP1 : GFP、SOAP2 : GFP、SOAP3 : RFP、SOAP4 : GFP、SOAP5 : RFP を、陰性の SOAP1-5、SOAP1-4 単独、対照 (EV) と比較した。

【図54A】蛍光タンパク質 (FP) と融合したが、依然として活性化している SOAP を示す図である。示された構築物を発現させた細胞の蛍光共鳴エネルギー移動 (FRET) 指数。値は平均値 ± SE。遊離 GFP 及び RFP を発現させた陰性対照 (NC) と比較した有意差をアスタリスクで示した ; \* P < 0.05、\*\* P < 0.01。

20

【図54B】蛍光タンパク質 (FP) と融合したが、まだ活性化している SOAP を示す図である。ER 内でのタンパク質の位置の可能性を示す模式図。

【発明を実施するための形態】

【0056】

以下の詳細な説明では、本発明の完全な理解を提供するために、多数の具体的な詳細を記載している。しかしながら、当業者であれば、これらの具体的な詳細がなくても本発明を実施することができることが理解できるであろう。他の例では、本発明を不明瞭にしないように、周知の方法、手順、及び構成要素は詳細に説明されていない。

30

【0057】

具体的には、セルロース合成酵素様 G (CSLG) 酵素をコードする少なくとも1つの異種遺伝子の発現が、対応する非組換え細胞と比較して増加した遺伝子組換え細胞であって、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え細胞と比較して増加した、遺伝子組換え細胞が開示される。いくつかの実施形態では、酵素をコードする少なくとも1つの追加の異種遺伝子をさらに発現し、前記酵素は、サポニン - アミリン合成酵素、シトクロム P450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、及び UDP - グルコース 6 - デヒドロゲナーゼ 1、及びそれらの任意の組み合わせからなる群から選択される。

40

【0058】

「CSLG」という用語は、Gファミリー由来のセルロース合成酵素様酵素を包含する。セルロース合成酵素様 (CSL) ファミリーには、例えば A、B、E、G などの、それぞれ異なる接尾辞を有するものがある。各ファミリーの接尾辞は、事前に特徴付けられたシロイヌナズナ由来のタンパク質に対する酵素の類似性に依りて基づいて付けられる。異なるファミリーの CSL 酵素は、異なる技術的特徴を有する。本明細書を通じて使用される、A、B、または E の接尾辞が付いていない用語「CSL」は、いくつかの実施形態で

50

は、「CSLG」という用語と、互換的に使用することができ、すべて同一の意味及び性質を有する。

【0059】

「CSLG」タンパク質は、いくつかの実施形態では、「GAME15」タンパク質及び「SOAP5」タンパク質を包含する。

【0060】

本明細書では、6つの検査済みCSLG遺伝子について説明する。いくつかの実施形態では、それらには、GAME15遺伝子と、それらがコードする酵素（トマト、野生トマト、ジャガイモ、野生ポテト、ナス、ピーマンの6種に由来する6つの酵素）が含まれる。これらは、グルクロン酸または他の糖部分を転移してステロイド系アルカロイドを産生するか、あるいは、グルクロン酸または他の糖部分をフロスタノール型サポニン（アグリコン）上に転移してステロイド系サポニンまたはその前駆体を産生するか、またはステロイド系アルカロイドまたはその前駆体を産生する（例えば、図1を参照）。

10

【0061】

本明細書では、グルクロン酸をトリテルペノイドアグリコン上に転移する7つの検査済みCSLG遺伝子（SOAP5遺伝子としても知られている）と、それがコードする酵素（ハウレンソウ、ウラルカンゾウ、レッドビート、キノア、アルファルファ、ダイズ、及びミヤコグサの7種に由来する7つの酵素）について説明する。「トリテルペノイドアグリコン」という用語は、いくつかの実施形態では、-アミリンから誘導される化合物群を包含する。したがって、本明細書に開示されるCSLG酵素は、ユニークな活性を提供する。

20

【0062】

当業者であれば、細胞、例えば、植物細胞、酵母細胞、藻類細胞、細菌細胞におけるステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンのうちの少なくとも1つの含有量を改変することは、内在性CSLG遺伝子を変異させることを含むか、CSLG遺伝子発現を調節することを含むか、あるいは、CSLG酵素を発現しない細胞に、または異種細胞で発現するCSLG酵素の活性を補うために、異種CSLGを添加することを含むことを理解できるであろう。いくつかの実施形態では、植物細胞は、植物またはその一部に含まれている。

【0063】

いくつかの実施形態では、少なくとも1つの遺伝子またはその任意の組み合わせの発現が改変される。この改変は、少なくとも1つの遺伝子を突然変異させることを含む。突然変異させることは、1以上の点突然変異の導入、ゲノム編集、細菌CRISPR/CAS系の使用、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、CSLG遺伝子またはポリヌクレオチドの発現は、例えば、欠失、挿入、または修飾によって、あるいは、CSLG遺伝子を標的とする少なくとも1つのサイレンシング分子の導入によって、サイレンシング、抑制、または低減される。

30

【0064】

少なくとも1つの遺伝子またはその組み合わせの発現を増加させることは、様々な手段によって達成することができ、そのすべてが本発明の範囲内に明示的に包含されることを理解されたい。いくつかの実施形態では、CSLGの発現の増強は、転写及び/または翻訳を増強する様々な分子（これに限定しないが、例えば、転写因子）を使用して、ゲノム及び/または転写レベルで影響を与えることができる。また、プロモータ、エンハンサーなどの追加を含む、少なくとも1つの遺伝子への突然変異の挿入も、その突然変異が遺伝子発現のアプレギュレーションをもたらす限り、使用することができる。他の実施形態では、アゴニストなどを使用してタンパク質レベルで発現を増加させる。

40

【0065】

いくつかの実施形態では、これらの遺伝子組換え細胞は、ステロイド系アルカロイドを産生する。いくつかの実施形態では、これらの遺伝子組換え細胞は、ステロイド系サポニンを産生する。いくつかの実施形態では、これらの遺伝子組換え細胞は、トリテルペノイ

50

ド系サポニンを産生する。いくつかの実施形態では、遺伝子組換え細胞は、それに対応する非組換え細胞と比較して、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、それに対応する非組換え細胞と比較して、少なくともステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、それに対応する非組換え細胞と比較して、少なくともトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加する。

【0066】

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの誘導体は、それぞれ、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンのグリコシル化誘導体を含む。

10

【0067】

いくつかの実施形態では、少なくとも1つの異種CSLG遺伝子をコードする核酸配列は、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、あるいは、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも55%の同一性を有し、かつ少なくとも55%のカバレッジを有するそのホモログであることが好ましい。いくつかの実施形態では、少なくとも1つの異種CSLG遺伝子をコードする核酸配列は、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、あるいは、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも55%の同一性を有し、かつ少なくとも80%のカバレッジを有するそのホモログであることが好ましい。

20

【0068】

いくつかの実施形態では、CSLG酵素のホモログは、アミノ酸配列に対して少なくとも55%、60%、65%、70%、75%、80%、85%、90%、95%、98%、または99%の同一性を有する。CSLG酵素のホモログをコードする遺伝子は、核酸配列に対して少なくとも55%、60%、65%、70%、75%、80%、85%、90%、95%、98%、または99%の同一性を有する。

30

【0069】

いくつかの実施形態では、CSLG酵素のホモログは、アミノ酸配列に対して少なくとも55%、60%、65%、70%、75%、80%、85%、90%、95%、98%、99%、または100%のカバレッジを有する。CSLG酵素のホモログをコードする遺伝子は、核酸配列に対して少なくとも55%、60%、65%、70%、75%、80%、85%、90%、95%、98%、99%、または100%のカバレッジを有する。

【0070】

いくつかの実施形態では、少なくとも1つの異種CSLG遺伝子をコードする核酸配列は、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、または配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも80%の同一性を有し、かつ少なくとも80%のカバレッジを有するそのホモログである。

40

【0071】

いくつかの実施形態では、コードされたCSLG酵素のアミノ酸配列は、配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列であるか、または配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または10

50

4のいずれか1つに記載のアミノ酸配列に対して少なくとも55%の相同性を有し、かつ少なくとも55%のカバレッジを有するそのホモログである。いくつかの実施形態では、コードされたCSLG酵素のアミノ酸配列は、配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列であるか、または配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列に対して少なくとも55%の相同性を有し、かつ少なくとも80%のカバレッジを有するそのホモログである。

【0072】

いくつかの実施形態では、コードされたCSLG酵素のアミノ酸配列は、配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列であるか、または配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列に対して少なくとも80%の相同性を有し、かつ少なくとも80%のカバレッジを有するそのホモログである。

10

【0073】

いくつかの実施形態では、少なくとも1つの追加の異種遺伝子をコードする核酸配列は、(a) - アミリン合成酵素、及び、配列番号45に記載の核酸配列であるか、または配列番号45に記載の核酸配列に対して少なくとも80%の相同性を有し、かつ少なくとも80%のカバレッジを有するそのホモログ、(b) シトクロムP450、及び、配列番号46、51、または53のいずれか1つに記載の核酸配列であるか、または配列番号46、51、または53のいずれか1つに記載の核酸配列に対して少なくとも80%の相同性を有し、かつ少なくとも80%のカバレッジを有するそのホモログ、(c) グリコシルトランスフェラーゼ、及び、配列番号55、57、59、または61のいずれか1つに記載の核酸配列であるか、または配列番号55、57、59、または61のいずれか1つに記載の核酸配列に対して少なくとも80%の相同性を有し、かつ少なくとも80%のカバレッジを有するそのホモログ、(d) アセチルトランスフェラーゼ、及び、配列番号63に記載の核酸配列であるか、または配列番号63に記載の核酸配列に対して少なくとも80%の相同性を有し、かつ少なくとも80%のカバレッジを有するそのホモログ、(e) UDP - グルコース6 - デヒドロゲナーゼ1、及び、配列番号74に記載の核酸配列であるか、または配列番号74に記載の核酸配列に対して少なくとも80%の相同性を有し、かつ少なくとも80%のカバレッジを有するそのホモログ、あるいは、(f) 上記の(a)、(b)、(c)、(d)、及び(e)の任意の組み合わせである。

20

30

【0074】

いくつかの実施形態では、コードされた少なくとも1つの追加の異種遺伝子のアミノ酸配列は、(a) - アミリン合成酵素、及び、配列番号48に記載のアミノ酸配列であるか、または配列番号48に記載のアミノ酸配列に対して少なくとも80%の相同性を有し、かつ少なくとも80%のカバレッジを有するそのホモログ、(b) シトクロムP450、及び、配列番号49、52、または54のいずれか1つに記載のアミノ酸配列であるか、または配列番号49、52、または54のいずれか1つに記載のアミノ酸配列に対して少なくとも80%の相同性を有し、かつ少なくとも80%のカバレッジを有するそのホモログ、(c) グリコシルトランスフェラーゼ、及び、配列番号56、58、60、または62のいずれか1つに記載のアミノ酸配列であるか、または配列番号56、58、60、または62のいずれか1つに記載のアミノ酸配列に対して少なくとも80%の相同性を有し、かつ少なくとも80%のカバレッジを有するそのホモログ、(d) アセチルトランスフェラーゼ、及び、配列番号64に記載のアミノ酸配列であるか、または配列番号64に記載のアミノ酸配列に対して少なくとも80%の相同性を有し、かつ少なくとも80%のカバレッジを有するそのホモログ、(e) UDP - グルコース6 - デヒドロゲナーゼ1、及び、配列番号75に記載のアミノ酸配列であるか、または配列番号75に記載のアミノ酸配列に対して少なくとも80%の相同性を有し、かつ少なくとも80%のカバレッジを

40

50

有するそのホモログ、あるいは、(f)上記の(a)、(b)、(c)、(d)、及び(e)の任意の組み合わせである。

【0075】

異種発現とは、ある遺伝子またはその一部を、本質的にはその遺伝子またはその遺伝子断片を持たない宿主細胞または宿主生物(例えば、以下に限定しないが、植物細胞、酵母細胞、藻類または植物など)において発現させることを含む。異種宿主への遺伝子の挿入は、当該技術分野で知られている任意の組換えDNA技術によって行われてもよく、例えば、酵母の場合は交配によって行われてもよい。

【0076】

トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体は、幅広い商業的用途があり、いくつかの実施形態では、甘味料、発泡剤、乳化剤、防腐剤、抗発癌剤、コレステロール低下剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、摂食阻害剤、抗真菌剤、またはそれらの任意の組み合わせを含む。

10

【0077】

いくつかの実施形態では、(a)ステロイド系グリコアルカロイドは、 $\beta$ -トマチン、トマチン、デヒドロトマチン、ヒドロキシトマチン、アセトキシトマチン、ジヒドロキシトマチン、 $\beta$ -カコニン、 $\beta$ -ソラニン、 $\beta$ -ソラソニン、 $\beta$ -ソラマルギン、それらの誘導体、関連化合物、またはそれらの任意の組み合わせを含む、(b)ステロイド系サポニンは、ウットロシドB、トマトシド、それらの誘導体、関連化合物、またはそれらの任意の組み合わせを含む、(c)トリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GlcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、パヨゲニン(化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサポゲノールA(化合物29)、ソヤサポゲノールB(化合物30)、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント(化合物34)、ソヤサポニンVI(化合物35)、ベタブルガロシドIV(化合物33)、ヘデラゲニン-3GlcA、ギブソゲンニン3GlcA、ギブソゲン酸-3GlcA、QS-21アジュバント、またはそれらの任意の組み合わせを含む、(d)トリテルペノイド系サポニンの生合成中間体は、化合物1、化合物2、メディカゲン酸(化合物5)、オレアノール酸(化合物3)

20

30

【0078】

いくつかの例示的な実施形態では、下流のステロイド系グリコアルカロイドは、エスクレオシドまたはデヒドロエスクレオシドからなる群から選択される。

【0079】

いくつかの実施形態では、遺伝子組換え細胞は、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変されている。

40

【0080】

いくつかの実施形態では、「含有量の改変」は、フィトステロール/フィトコレステロール/フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量の増加を含む。いくつかの実施形態では、「含有量の改変」は、フィトステロール、その誘導体、その代謝物、またはその生合成中間体の含有量の増加を含む。

【0081】

フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロー

50

ル、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が減少する。

【0082】

いくつかの実施形態では、遺伝子組換え細胞は、植物細胞、酵母細胞、または藻類細胞を含む。いくつかの実施形態では、遺伝子組換えされた植物細胞は、植物またはその一部の細胞を含む。いくつかの実施形態では、遺伝子組換え細胞は、酵母細胞である。いくつかの実施形態では、遺伝子組換え細胞は、藻類細胞である。

【0083】

いくつかの実施形態では、植物細胞は、イネ目、ナデシコ目、ナス目、マメ目、アオイ目、セリ目、フウチョウソウ目、クサスギカズラ目、ヤマノイモ目、またはユリ目の植物の細胞を含む。

【0084】

いくつかの実施形態では、(a) 植物細胞がイネ目の植物の細胞を含む場合、植物は、イネ属、オオムギ属、カラスムギ属、及びコムギ属からなる群から選択される、(b) 植物細胞がナデシコ目の植物の細胞を含む場合、植物は、ハウレンソウ属、アカザ属、ベータ属、レウム属、ドウカンソウ属、サポナリア属、及びジブソフィラ属からなる群から選択される、(c) 植物細胞がナス目の植物の細胞を含む場合、植物は、ナス属、トウガラシ属、タバコ属、ヒヨス属、ダチュラ属、及びオオカミナスビ属からなる群から選択される、(d) 植物細胞がマメ目の植物の細胞を含む場合、植物は、グリシリザ属、メディカゴ属、ダイズ属、ハス属、シセル属、インゲンマメ属、エンドウ属、ラッカセイ属、ルピナス属、及びアカシア属からなる群から選択される、(e) 植物細胞がアオイ目の植物の細胞を含む場合、植物はカカオ属から選択される、(f) 植物細胞がセリ目の植物の細胞を含む場合、植物は、ニンジン属、オランダミツバ属、パセリ属、オタネニンジン属、ミシマサイコ属、キツタ属、及びツボクサ属からなる群から選択される、あるいは、(g) 植物細胞がフウチョウソウ目の植物の細胞を含む場合、植物は、シロイヌナズナ属、アブラナ属、フウチョウボク属、及びパイヤ属からなる群から選択される。

【0085】

いくつかの実施形態では、(a) 植物細胞がナデシコ目の植物の細胞を含む場合、植物は、ハウレンソウ、ピーツ、及びキノアからなる群から選択される、(b) 植物細胞がナス目の植物の細胞を含む場合、植物は、トマト、野生型トマト、ジャガイモ、野生型ポテト、ナス、ペッパー、ピーマン、トウガラシ、チリペッパー、パプリカ、タバスコペッパー、グランドチェリー、タバコ、及びビタースイートからなる群から選択される、あるいは、(c) 植物細胞がマメ目の植物の細胞を含む場合、植物は、アルファルファ、大豆、ミヤコグサ、及び甘草からなる群から選択される。

【0086】

いくつかの実施形態では、植物細胞は、(i) - トマチン、その誘導体、その代謝物、またはその生合成中間体、トマチジン、その誘導体、その代謝物、またはその生合成中間体、あるいは、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加したトマト植物の細胞であるか、あるいは、(ii) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変されたトマト植物の細胞である。

【0087】

いくつかの実施形態では、植物細胞は、(i) - トマチン、その誘導体、その代謝物、またはその生合成中間体、トマチジン、その誘導体、その代謝物、またはその生合成中間体、あるいは、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中

10

20

30

40

50

間体の含有量が増加したトマト植物の細胞であるか、あるいは、( i i ) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が減少したトマト植物の細胞である。

【 0 0 8 8 】

いくつかの実施形態では、植物細胞は、( i ) - カコニン、その誘導体、その代謝物、またはその生合成中間体、 - ソラニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加したジャガイモ植物の細胞であるか、あるいは、( i i ) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変されたジャガイモ植物の細胞である。

10

【 0 0 8 9 】

いくつかの実施形態では、植物細胞は、( i ) - カコニン、その誘導体、その代謝物、またはその生合成中間体、 - ソラニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加したジャガイモ植物の細胞であるか、あるいは、( i i ) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が減少したジャガイモ植物の細胞である。

20

【 0 0 9 0 】

いくつかの実施形態では、植物細胞は、( i ) - ソラソニン、その誘導体、その代謝物、またはその生合成中間体、 - ソラマルギン、その誘導体、その代謝物、またはその生合成中間体、あるいは、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加したナス科植物の細胞であるか、あるいは、( i i ) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変されたナス科植物の細胞である。

30

【 0 0 9 1 】

いくつかの実施形態では、植物細胞は、( i ) - ソラソニン、その誘導体、その代謝物、またはその生合成中間体、 - ソラマルギン、その誘導体、その代謝物、またはその生合成中間体、あるいは、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加したナス科植物の細胞であるか、あるいは、( i i ) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が減少したナス科植物の細胞である。

40

【 0 0 9 2 】

いくつかの実施形態では、植物細胞は、葉の細胞、葉柄の細胞、植物の茎または幹の細胞、根の細胞、芽の細胞、塊茎の細胞、豆の細胞、穀物または穀粒の細胞、果実の細胞、木の実の細胞、マメの鞘の細胞、種子または種子の細胞、仮皮の細胞、包葉の細胞、また

50

は花の細胞を含む。いくつかの実施形態では、植物細胞は、植物またはその一部の細胞を含み、植物の一部は、植物の葉、植物の葉柄、植物の茎または幹、植物の根、植物の芽、植物の塊茎、植物の豆、植物の穀物または穀粒、植物の果実、植物の木の实、植物のマメの鞘、植物の種子、植物の包葉、または植物の花を含む。

【0093】

いくつかの実施形態では、酵母は、サッカロミセス属、シゾサッカロミセス属、ピチア属、ヤロウイア属、クリベロマイセス属、またはカンジダ属から選択される。

【0094】

いくつかの実施形態では、藻類は、微細藻類、多細胞藻類、シアノバクテリア、珪藻類、緑藻植物（緑藻類）、紅藻植物（紅藻類）、褐藻植物（褐藻類）、ドナリエラ、クラミドモナス、またはヘマトコッカスから選択される。

10

【0095】

具体的には、セルロース合成酵素様 G (CSLG) 酵素をコードする少なくとも1つのCSLG遺伝子の発現が、対応する非組換え植物におけるCSLG遺伝子の発現と比較して改変された少なくとも1つの細胞を含む遺伝子組換え植物が、本明細書に開示されている。この遺伝子組換え植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して改変されている。

20

【0096】

いくつかの実施形態では、少なくとも1つの遺伝子またはその任意の組み合わせの発現が改変され、発現の改変は、少なくとも1つの遺伝子を突然変異させることを含み、突然変異は、1以上の点突然変異の導入、ゲノム編集、細菌のCRISPR/CAS系の使用、またはそれらの組み合わせを含む。

【0097】

いくつかの実施形態では、少なくとも1つのCSLG酵素をコードする遺伝子の発現は、それぞれ対応する非組換え細胞または非組換え植物と比較して増加する。いくつかの実施形態では、遺伝子組換え細胞または遺伝子組換え植物は、CSLG遺伝子をコードするポリヌクレオチドを含み、このポリヌクレオチドの発現は、選択的に増加する。いくつかの実施形態では、遺伝子組換え植物は、少なくとも1つのCSLGからなる群から選択される少なくとも1つのタンパク質をコードする少なくとも1つの転写可能なポリヌクレオチドを含む、少なくとも1つの細胞を含む。

30

【0098】

いくつかの実施形態では、遺伝子組換え植物は、少なくとも1つのCSLG遺伝子をコードするポリヌクレオチドを標的とする少なくとも1つのサイレンシング分子を含む少なくとも1つの細胞を含む、形質転換された植物である。いくつかの実施形態では、形質転換された植物は、CSLG遺伝子をコードするポリヌクレオチドを含み、このポリヌクレオチドの発現が選択的にサイレンシングされるか、抑制されるか、または低下している。いくつかの実施形態では、形質転換された植物は、CSLG遺伝子をコードするポリヌクレオチドを含み、このポリヌクレオチドは、その発現をサイレンシングするか、抑制するか、または低下させるために、欠失、挿入、または修飾によって選択的に編集されており、または、遺伝子組換え植物は、遺伝子編集された植物の子孫である。いくつかの実施形態では、形質転換された植物は、CSLG遺伝子を標的とする少なくとも1つのサイレンシング分子を含む、少なくとも1つの細胞を含む。いくつかの実施形態では、サイレンシング分子は、RNA干渉分子及びアンチセンス分子から選択されるか、または、ウイルス誘導遺伝子サイレンシング系の構成要素である。いくつかの実施形態では、サイレンシング分子は、CSLG遺伝子の領域に実質的に相補的な核酸配列を有するポリヌクレオチドを含む。

40

【0099】

50

いくつかの実施形態では、形質転換された植物は、少なくとも1つのCSLG遺伝子を標的としたサイレンシング分子を含む複数の細胞を含む。追加の実施形態では、植物細胞の大部分がサイレンシング分子を含む。

【0100】

少なくとも1つの遺伝子またはその組み合わせの発現を阻害することは、様々な手段で達成され、それらはすべて本発明の範囲に明示的に包含されることを理解されたい。いくつかの実施形態では、CSLGの発現を阻害することは、以下に限定しないが、例えばアンチセンス、siRNA、リボザイム、またはDNAzyme分子を含む転写及び/または翻訳を阻害する様々な分子を用いて、ゲノム及び/または転写レベルで効果的に行うことができる。欠失、挿入、部位特異的な突然変異、ジンクフィンガーヌクレアーゼなどを

10

【0101】

いくつかの実施形態では、CSLG遺伝子は、内在性CSLG遺伝子を含む。

【0102】

いくつかの実施形態では、内在性CSLG遺伝子は、突然変異を含み、突然変異は、少なくとも1以上の点突然変異、挿入、欠失、またはそれらの任意の組み合わせを含み、(a)発現したCSLG酵素は、安定性の向上、活性の増加、またはその両方を含み、含有量の改変は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が、対応する非組換え植物と比較して増加すること、あるいは、(b)発現したCSLG酵素は、安定性の低下、活性の低下、またはその両方を含み、含有量の改変は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が、対応する非組換え植物と比較して減少することである。

20

【0103】

いくつかの実施形態では、(a)内在性CSLG遺伝子は、選択的にサイレンシングされるか、抑制されるか、または発現が低下し、含有量の改変は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が、対応する非組換え植物と比較して減少すること、あるいは、(b)内在性CSLG遺伝子は、過剰発現し、含有量の改変は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生

30

40

【0104】

いくつかの実施形態では、内在性CSLG遺伝子が選択的にサイレンシングされるか、抑制されるか、または発現が低下した場合、細胞は、CSLG遺伝子をコードするポリヌクレオチドを標的とする少なくとも1つのサイレンシング分子をさらに含み、サイレンシング分子は、RNA干渉分子またはアンチセンス分子から選択されるか、または、ウイルス誘導遺伝子サイレンシング系の構成要素である。

【0105】

いくつかの実施形態では、CSLG遺伝子は、異種CSLG遺伝子を含み、発現の改変は、異種遺伝子のデノボ発現を含む。

50

## 【0106】

いくつかの実施形態では、発現の改変は、対応する非組換え細胞と比較して発現が増加することを含み、含有量の改変は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して増加することを含む。

## 【0107】

いくつかの実施形態では、CSLG遺伝子をコードする核酸配列は、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、あるいは、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも55%の同一性を有し、かつ少なくとも55%のカバレッジを有するそのホモログである。いくつかの実施形態では、CSLG遺伝子をコードする核酸配列は、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、あるいは、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも55%の同一性を有し、かつ少なくとも80%のカバレッジを有するそのホモログである。

10

20

## 【0108】

いくつかの実施形態では、CSLG遺伝子をコードする核酸配列は、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、あるいは、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも80%の同一性を有し、かつ少なくとも80%のカバレッジを有するそのホモログである。

## 【0109】

いくつかの実施形態では、CSLG酵素のアミノ酸配列は、配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列であるか、あるいは、配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列に対して少なくとも55%の同一性を有し、かつ少なくとも55%のカバレッジを有するそのホモログである。いくつかの実施形態では、CSLG酵素のアミノ酸配列は、配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列であるか、あるいは、配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列に対して少なくとも55%の同一性を有し、かつ少なくとも80%のカバレッジを有するそのホモログである。

30

40

## 【0110】

いくつかの実施形態では、CSLG酵素のアミノ酸配列は、配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列であるか、あるいは、配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列に対して少なくとも80%の同一性を有し、かつ少なくとも80%のカバレッジを有するそのホモログである。

## 【0111】

いくつかの実施形態では、少なくとも1つのセルロース合成酵素様G(CSLG)遺伝子の発現の改変は、(1)核酸配列内に1以上の点突然変異を導入すること、(2)核酸

50

配列内に欠失を導入すること、(3) 核酸配列内に挿入を導入すること、あるいは、(4) それらの任意の組み合わせを含み、導入は、コード配列または非コード配列を突然変異させることを含む。

【0112】

いくつかの実施形態では、少なくとも1つのCSLG遺伝子の発現の改変は、少なくとも1つのCSLG遺伝子を標的とするサイレンシング分子を導入することにより生じ、かつ、(a) C S L G 遺伝子をコードする核酸配列が配列番号30に記載されている核酸配列である場合、サイレンシング分子は、配列番号42に記載の核酸配列またはその相補的配列であるか、または、配列番号42に記載の核酸配列またはその相補的配列に対して少なくとも80%の相同性を有し、かつ少なくとも80%のカバレッジを有するそのホモログである、(b) C S L G 遺伝子をコードする核酸配列が配列番号34に記載されている核酸配列である場合、サイレンシング分子は、配列番号43に記載の核酸配列またはその相補的配列であるか、または、配列番号43に記載の核酸配列またはその相補的配列に対して少なくとも80%の相同性を有し、かつ少なくとも80%のカバレッジを有するそのホモログである、(c) C S L G 遺伝子をコードする核酸配列が配列番号38：に記載されている核酸配列である場合、サイレンシング分子は、配列番号44に記載の核酸配列またはその相補的配列であるか、または、配列番号44に記載の核酸配列またはその相補的配列に対して少なくとも80%の相同性を有し、かつ少なくとも80%のカバレッジを有するそのホモログである、(d) C S L G 遺伝子をコードする核酸配列が配列番号65または配列番号93に記載されている核酸配列である場合、サイレンシング分子は、配列番号106に記載の核酸配列またはその相補的配列であるか、または、配列番号106に記載の核酸配列またはその相補的配列に対して少なくとも80%の相同性を有し、かつ少なくとも80%のカバレッジを有するそのホモログである、あるいは、(e) C S L G 遺伝子をコードする核酸配列が配列番号95に記載されている核酸配列である場合、サイレンシング分子は、配列番号107に記載の核酸配列またはその相補的配列であるか、または、配列番号107に記載の核酸配列またはその相補的配列に対して少なくとも80%の相同性を有し、かつ少なくとも80%のカバレッジを有するそのホモログである。

10

20

【0113】

トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体は、幅広い商業的用途があり、いくつかの実施形態では、甘味料、発泡剤、乳化剤、防腐剤、抗発癌剤、コレステロール低下剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、摂食阻害剤、抗真菌剤、またはそれらの任意の組み合わせを含む。

30

【0114】

いくつかの実施形態では、  
 (a) ステロイド系グリコアルカロイドは、 $\beta$ -トマチン、トマチン、デヒドロトマチン、ヒドロキシトマチン、アセトキシトマチン、ジヒドロキシトマチン、 $\beta$ -カコニン、 $\beta$ -ソラニン、 $\beta$ -ソラソニン、 $\beta$ -ソラマルギン、それらの誘導体、それらの経路中間体、またはそれらの任意の組み合わせを含む、(b) ステロイド系サポニンは、ウットロシドB、トマトシド、及びすべての関連化合物、それらの誘導体、それらの経路中間体、またはそれらの任意の組み合わせを含む、(c) トリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GlcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、バヨゲニン(化合物25)、バヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサポゲノールA(化合物29)、ソヤサポゲノールB(化合物30)、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント(化合物34)、ソヤサポニンVI(化合物35)、ベタブルガロシドIV(化合物33)、ヘデラゲニン-3GlcA、ジブソゲニン-3GlcA、ギブソゲン酸-3GlcA、QS-21アジュバント、またはそれらの任意の組み合わせを含む、(d) トリテルペノイド系サポニンの生合成中間体は、化合物1、化合物2、メディカゲン酸(化合物5)、オレアノール

40

50

酸（化合物 3）、オレアノール酸 - 3 G 1 c A、アウグスチン酸（化合物 4）、アウグスチン酸 - 3 G 1 c A、グリチルレチン酸（化合物 13）、またはそれらの任意の組み合わせを含む、あるいは、（e）上記の（a）、（b）、（c）、及び（d）の任意の組み合わせである。

【0115】

いくつかの実施形態では、遺伝子組換え植物は、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変されている。

10

【0116】

いくつかの実施形態では、遺伝子組換え植物は、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変され、（a）少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して増加する場合、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して減少するか、あるいは、（b）少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して減少する場合、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して増加する。

20

30

【0117】

ある実施形態では、ステロイド系サポニン、ステロイド系アラコイド、トリテルペノイド系サポニンのそれぞれが増加すると、他の成分の含有量が改変され得る。いくつかの実施形態では、改変は、含有量の増加を含む。いくつかの実施形態では、改変は、含有量の減少を含む。いくつかの実施形態では、改変は、一方の含有量の増加と、他方の含有量の減少とを含む。

40

【0118】

いくつかの実施形態では、植物は、イネ目、ナデシコ目、ナス目、マメ目、アオイ目、セリ目、フウチョウソウ目、クサスギカズラ目、ヤマノイモ目、またはユリ目から選択される。

【0119】

いくつかの実施形態では、（a）植物がイネ目の植物を含む場合、植物は、イネ属、オオムギ属、カラスムギ属、及びコムギ属からなる群から選択される、（b）植物がナデシコ目の植物を含む場合、植物は、ハウレンソウ属、アカザ属、ベータ属、レウム属、ドウ

50

カンソウ属、サボナリア属、及びジブソフィラ属からなる群から選択される、(c)植物がナス目の植物を含む場合、植物は、ナス属、トウガラシ属、タバコ属、ヒヨス属、ダチュラ属、及びオオカミナスビ属からなる群から選択される、(d)植物がマメ目の植物を含む場合、植物は、グリシリザ属、メディカゴ属、ダイズ属、ハス属、シセル属、インゲンマメ属、エンドウ属、ラッカセイ属、ルピナス属、及びアカシア属からなる群から選択される、(e)植物がアオイ目の植物を含む場合、植物は、カカオ属から選択される、(f)植物がセリ目の植物を含む場合、植物は、ニンジン属、オランダミツバ属、パセリ属、オタネニンジン属、ミシマサイコ属、キツタ属、及びツボクサ属からなる群から選択される、あるいは、(g)植物がフウチョウソウ目の植物を含む場合、植物は、シロイヌナズナ属、アブラナ属、フウチョウボク属、及びパパイア属からなる群から選択される。

10

## 【0120】

いくつかの実施形態では、(a)植物がナデシコ目の植物を含む場合、植物は、ホウレンソウ、ビートルート、及びキノアからなる群から選択される、(b)植物がナス目の植物を含む場合、植物は、トマト、野生型トマト、ジャガイモ、野生型ポテト、ナス、ペッパー、ピーマン、トウガラシ、チリペッパー、パプリカ、タバスコペッパー、グランドチェリー、タバコ、及びピタースイートからなる群から選択される、あるいは、(c)植物がマメ目の植物を含む場合、植物は、アルファルファ、大豆、ミヤコグサ、及び甘草からなる群から選択される。

## 【0121】

いくつかの実施形態では、植物は、(i) - トマチン、その誘導体、その代謝物、またはその生合成中間体、トマチジン、その誘導体、その代謝物、またはその生合成中間体、デヒドロトマチン、その誘導体、その代謝物、またはその生合成中間体、あるいは、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加したトマト植物であるか、あるいは、(ii) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が変更されたトマト植物である。

20

## 【0122】

いくつかの実施形態では、植物は、(i) - トマチン、その誘導体、その代謝物、またはその生合成中間体、トマチジン、その誘導体、その代謝物、またはその生合成中間体、デヒドロトマチン、その誘導体、その代謝物、またはその生合成中間体、あるいは、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加したトマト植物であるか、あるいは、(ii) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が減少したトマト植物である。

30

40

## 【0123】

いくつかの実施形態では、植物は、(i) - カコニン、その誘導体、その代謝物、またはその生合成中間体、 - ソラニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加したジャガイモ植物であるか、あるいは、(ii) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が変更されたジャガイモ植物である。

50

## 【 0 1 2 4 】

いくつかの実施形態では、植物は、( i ) - カコニン、その誘導体、その代謝物、またはその生合成中間体、 - ソラニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加したジャガイモ植物であるか、あるいは、( i i ) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が減少したジャガイモ植物である。

10

## 【 0 1 2 5 】

いくつかの実施形態では、植物は、( i ) - ソラソニン、その誘導体、その代謝物、またはその生合成中間体、 - ソラマルギン、その誘導体、その代謝物、またはその生合成中間体、あるいは、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加したナス科植物であるか、あるいは、( i i ) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変されたナス科植物である。

20

## 【 0 1 2 6 】

いくつかの実施形態では、植物は、( i ) - ソラソニン、その誘導体、その代謝物、またはその生合成中間体、 - ソラマルギン、その誘導体、その代謝物、またはその生合成中間体、あるいは、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加したナス科植物であるか、あるいは、( i i ) フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量が減少したナス科植物である。

30

## 【 0 1 2 7 】

いくつかの実施形態では、植物細胞は、葉の細胞、葉柄の細胞、植物の茎または幹の細胞、根の細胞、芽の細胞、塊茎の細胞、豆の細胞、穀物または穀粒の細胞、果実の細胞、木の実の細胞、マメの鞘の細胞、種子または種子の細胞、仮皮の細胞、包葉の細胞、または花の細胞を含む。

## 【 0 1 2 8 】

具体的には、本明細書に開示されているのは、遺伝子組換え細胞内にステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法であって、( a ) 細胞内に少なくとも1つの異種遺伝子を導入する導入ステップであって、少なくとも1つの異種遺伝子がセルロース合成酵素様 G ( C S L G ) 酵素をコードしており、少なくとも1つの異種遺伝子が任意選択でベクターに含まれている、該ステップと、( b ) 細胞内で少なくとも1つの異種遺伝子を発現させる発現ステップと、を含み、細胞は、少なくとも1つのステロイド系アルカロイド、少なくとも1つのステロイド系サポニン、または少なくとも1つのトリテルペノイド系サポニンの含有量が、対応する非組換え細胞と比較して増加する、方法である。

40

## 【 0 1 2 9 】

いくつかの実施形態では、導入ステップは、細胞内に少なくとも1つの追加の異種遺伝子を導入するステップであって、少なくとも1つの追加の異種遺伝子が、サポニン - アミリン合成酵素、シトクロム P 4 5 0、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、及び U D P - グルコース 6 - デヒドロゲナーゼ 1、及びそれらの任意の組み

50

合わせからなる群から選択される酵素をコードしており、少なくとも1つの追加の異種遺伝子が任意選択でベクターに含まれている、このステップと、細胞内で少なくとも1つの異種遺伝子を発現させるステップと、をさらに含む。

【0130】

いくつかの実施形態では、(a)少なくとも1つの異種遺伝子が、プロモータ、転写終結配列、またはそれらの組み合わせに動作可能に連結されているか、(b)少なくとも1つの追加の異種遺伝子が、プロモータ、転写終結配列、またはそれらの組み合わせに動作可能に連結されているか、あるいは、(c)上記の(a)と(b)の組み合わせである。

【0131】

いくつかの実施形態では、導入ステップは、少なくとも1つの細胞を、(a)少なくとも1つの異種遺伝子、少なくとも1つの異種遺伝子をコードするポリヌクレオチド配列、または少なくとも1つの異種遺伝子を含むベクター、(b)少なくとも1つの追加の異種遺伝子、少なくとも1つの追加の異種遺伝子をコードするポリヌクレオチド配列、または少なくとも1つの追加の異種遺伝子を含むベクター、あるいは、(c)上記の(a)と(b)の組み合わせを用いて形質転換することを含み、発現は、一過性発現または構成的発現を含む。

【0132】

具体的には、植物またはその少なくとも一部の少なくとも1つの細胞における、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を低減させる方法であって、少なくとも1つの細胞の遺伝子を組み換える遺伝子組換えステップを含み、遺伝子組換えステップは、(a)セルロース合成酵素様G(CSLG)酵素をコードする少なくとも1つの遺伝子を標的とする少なくとも1つのサイレンシング分子を用いて少なくとも1つの植物細胞を形質転換するステップ、あるいは、(b)CSLG酵素をコードする少なくとも1つの遺伝子の核酸配列を突然変異させるステップであって、突然変異が、(1)核酸配列内に1以上の点突然変異を導入すること、(2)核酸配列内に欠失を導入すること、(3)核酸配列内に挿入を導入すること、または、(4)それらの任意の組み合わせを含み、導入が、コード配列または非コード配列を突然変異させることを含む、該ステップと、を含み、遺伝子が組み換えられた遺伝子組換え細胞におけるCSLG酵素をコードする遺伝子の発現は、対応する非組換え植物の細胞と比較して減少し、遺伝子組換え細胞を含む植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して減少する、方法である。

【0133】

いくつかの実施形態では、(a)CSLG酵素をコードする少なくとも1つの遺伝子の核酸配列が配列番号30に記載されている核酸配列である場合、サイレンシング分子は、配列番号42に記載の核酸配列またはその相補的配列であるか、または、配列番号42に記載の核酸配列またはその相補的配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログである、(b)CSLG酵素をコードする少なくとも1つの遺伝子の核酸配列が配列番号34に記載されている核酸配列である場合、サイレンシング分子は、配列番号43に記載の核酸配列またはその相補的配列であるか、または、配列番号43に記載の核酸配列またはその相補的配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログである、(c)CSLG酵素をコードする少なくとも1つの遺伝子の核酸配列が配列番号38に記載されている核酸配列である場合、サイレンシング分子は、配列番号44に記載の核酸配列またはその相補的配列であるか、または、配列番号44に記載の核酸配列またはその相補的配列に対し

10

20

30

40

50

て少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログである、(d)CSLG酵素をコードする少なくとも1つの遺伝子の核酸配列が配列番号65または配列番号93に記載されている核酸配列である場合、サイレンシング分子は、配列番号106に記載の核酸配列またはその相補的配列であるか、または、配列番号106に記載の核酸配列またはその相補的配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログである、あるいは、(e)CSLG酵素をコードする少なくとも1つの遺伝子の核酸配列が配列番号95に記載されている核酸配列である場合、サイレンシング分子は、配列番号107に記載の核酸配列またはその相補的配列であるか、または配列番号107に記載の核酸配列またはその相補的配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログである。

10

**【0134】**

いくつかの実施形態では、(a)ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンは、毒素または苦味化合物を含むか、ホルモン模倣性を有するか、またはそれらの組み合わせであり、(b)ステロイド系グリコアルカロイドは、  
- トマチン、トマチン、  
- カコニン、  
- ソラニン、  
- ソラソニン、  
- ソラマルギン、またはそれらの任意の組み合わせを含む、(c)ステロイド系サポニンは、ウットロシドB、トマトシド、関連化合物(すべてステロイド系サポニンのファミリー)、経路中間体及び誘導体、またはそれらの任意の組み合わせを含む、(d)トリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GlcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、パヨゲニン(化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサボゲノールA(化合物29)、ソヤサボゲノールB(化合物30)、ソヤサボゲノールA-ヘキサA-ヘキサ-ペント(化合物34)、ソヤサポニンVI(化合物35)、ベタブルガロシドIV(化合物33)、ヘデラゲニン-3GlcA、ギブソゲンニン3GlcA、ギブソゲン酸-3GlcA、QS-21アジュバント、またはそれらの任意の組み合わせを含む、(e)トリテルペノイド系サポニンの生合成中間体は、化合物1、化合物2、メディカゲン酸(化合物5)、オレアノール酸(化合物3)、オレアノール酸-3GlcA、アウグスチン酸(化合物4)、アウグスチン酸-3GlcA、グリチルレチン酸(化合物13)、またはそれらの任意の組み合わせを含む、あるいは、(f)上記の(a)、(b)、(c)、(d)及び(e)の任意の組み合わせである。

20

30

**【0135】**

ある実施形態では、以下に限定しないが、例えば、キノアのように苦味を有するトリテルペノイド系サポニンや、有毒であるトリテルペノイド系サポニンなどのトリテルペノイド系サポニンの還元が望まれる。還元が望まれるトリテルペノイド系サポニンの他の例には、毒性を有するトリテルペノイド系サポニン、またはホルモンを模倣するトリテルペノイド系サポニンが含まれる。したがって、ある実施形態では、トリテルペノイド系サポニンの産生量の増加が有益であり、他の実施形態では、トリテルペノイド系サポニンの産生量の減少が有益である。

40

**【0136】**

いくつかの実施形態では、本方法は、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量を改変させるステップをさらに含む。

**【0137】**

いくつかの実施形態では、本方法は、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中

50

間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレスタノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレスタノール、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させるステップをさらに含む。

【0138】

いくつかの実施形態では、本方法は、遺伝子組換え植物から抽出されたフィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体を精製する精製ステップをさらに含む。いくつかの実施形態では、フィトステロールは、フィトコレステノールを含む。いくつかの実施形態では、遺伝子組換え植物から抽出されたフィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体コレステノール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体は、栄養剤または食品添加物、高付加価値のステロイド系化合物（例えば、プロビタミンD及び/またはジオスゲニン）、抗コレステロール剤（例えば、哺乳類または鳥類の腸内で食物コレステロールと競合する植物フィトステロール）、または化粧品を含む。

10

20

【0139】

具体的には、植物またはその少なくとも一部の少なくとも1つの細胞における、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法であって、少なくとも1つの細胞の遺伝子を組み換える遺伝子組換えステップを含み、遺伝子組換えステップは、(a)セルロース合成酵素様G(CSLG)酵素をコードする少なくとも1つの核酸配列を突然変異させるステップであって、突然変異が、(1)核酸配列内に1以上の点突然変異を導入すること、(2)核酸配列内に欠失を導入すること、(3)核酸配列内に挿入を導入すること、あるいは、(4)それらの任意の組み合わせを含み、導入が、コード配列または非コード配列を突然変異させることを含む、該ステップと、(b)CSLG酵素をコードする核酸を発現ステップと、を含み、遺伝子組換え細胞を含む植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して増加する、方法。

30

【0140】

いくつかの実施形態では、(a)ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンは、毒素または苦味化合物を含むか、ホルモン模倣特性を有するか、またはそれらの組み合わせであり、(b)ステロイド系グリコアルカロイドは、-トマチン、トマチン、-カコニン、-ソラニン、-ソラソニン、-ソラマルギン、またはそれらの任意の組み合わせを含む、(c)ステロイド系サポニンは、ウットロシドB、トマトシド、またはそれらの任意の組み合わせを含む、(d)トリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GlcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、パヨゲニン(化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサボゲノールA(化合物29)、ソヤサボゲノールB(化合物30)、ソヤサボゲノールA-ヘキサA

40

50

- ヘキサ - ペント (化合物 34)、ソヤサポニン VI (化合物 35)、ベタブルガロシド IV (化合物 33)、ヘデラゲニン - 3 Glc A、ギブソゲンニン 3 Glc A、ギブソゲン酸 - 3 Glc A、QS - 21 アジュバント、またはそれらの任意の組み合わせを含む、(e) トリテルペノイド系サポニンの生合成中間体は、化合物 1、化合物 2、メディカゲン酸 (化合物 5)、オレアノール酸 (化合物 3)、オレアノール酸 - 3 Glc A、アウグスチン酸 (化合物 4)、アウグスチン酸 - 3 Glc A、グリチルレチン酸 (化合物 13)、またはそれらの任意の組み合わせを含む、あるいは、(f) 上記の (a)、(b)、(c)、(d) 及び (e) の任意の組み合わせである。

【0141】

いくつかの実施形態では、(a) 遺伝子組換え細胞における、CSLG 酵素をコードする遺伝子の発現は、対応する非組換え植物の細胞と比較して増加する、(b) 遺伝子組換え細胞における、CSLG 酵素の活性は、対応する非組換え植物細胞と比較して増加する、(c) 遺伝子組換え細胞における、CSLG 酵素の安定性は、対応する非組換え植物の細胞と比較して向上する、あるいは、(d) 上記の (a)、(b)、及び (c) の任意の組み合わせである。

10

【0142】

いくつかの実施形態では、本方法は、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量を改変させるステップをさらに含む。

20

【0143】

いくつかの実施形態では、本方法は、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量を低減させるステップをさらに含む。

【0144】

本明細書に記載の各可能性は、本発明の個別の実施形態を表している。

30

【0145】

遺伝子組換え細胞・遺伝子組換え植物

【0146】

本明細書には、ステロイド系グリコアルカロイド/ステロイド系サポニン/トリテルペノイド系サポニンの代謝経路 (ステロイド系アルカロイド、ステロイド系サポニン、トリテルペノイド系サポニンの生合成経路) における重要な遺伝子の発現が改変された、遺伝子組換え細胞及び遺伝子組換え植物が開示されている。これらの遺伝子の発現を改変させると、ステロイド系アルカロイドのプロファイル、ステロイド系サポニンのプロファイル、またはトリテルペノイドのプロファイルも付随して改変される。非限定的な例では、ステロイド系アルカロイドの産生レベルが改変されることによって、病原菌に対する抵抗性が向上したステロイド系アルカロイドの含有量が増加した改良植物、または、植物の食用部分におけるこれらの二次化合物の含有量が減少した植物が得られ、その結果、抗栄養成分が減少または除去された改良作物を生産することができる。代替的に、または追加的に、別の非限定的な例では、本明細書に開示された遺伝子の発現の制御は、例えば栄養、化粧品、または製薬業界におけるさらなる使用のための、所望のステロイド系アルカロイドまたは植物由来のコレステロールの産生のために使用され得る。別の非限定的な例では、本明細書には、ナス目及び他の目の作物である植物を生産するための手段及び方法が開示されている。これらの植物は、食用部分に一般的に存在する有害な量のステロイド系アルカロイドを含まない。このように、本明細書に開示されている細胞及び植物は、栄養学的

40

50

及び商業的に重要な価値を有する。

【0147】

本明細書では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの生合成経路に關与する共発現遺伝子の配列が開示されている。本発明はさらに、この経路における重要な遺伝子を開示しており、その遺伝子の発現を改変されると、ステロイド系アルカロイドプロファイル、ステロイド系サポニンプロファイル、またはトリテルペノイドプロファイルが付随して改変される。一例として、ステロイド系アルカロイドの産生レベルを改変することによって、病原菌に対する抵抗性が向上したステロイド系アルカロイドの含有量が増加した改良植物、または、植物の食用部分におけるこれらの二次化合物の含有量が減少した植物が得られ、その結果、改良作物を生産することができる。代替的に、または追加的に、1つの非限定的な例では、本発明で明らかにされた遺伝子の発現の制御は、例えば栄養、化粧品、または製薬業界におけるさらなる使用のための、所望のステロイド系アルカロイド、植物由来のコレステロールまたは他のフィトステロールの産生のために使用され得る。したがって、本発明の細胞及び植物は、このように重要な栄養学的、化粧品的、薬学的、及び商業的価値を有する。

10

【0148】

本明細書では、遺伝子組換え細胞（例えば、植物細胞、酵母細胞、藻類細胞、細菌細胞、昆虫細胞）及び少なくとも1つの遺伝子組換え細胞を含む植物が開示されており、この遺伝子組換え細胞は、ステロイド系アルカロイド/ステロイド系サポニン/トリテルペノイド系サポニンの生合成経路内の酵素をコードする少なくとも1つの異種遺伝子を発現している。これらの酵素は、いくつかの実施形態では、セルロース合成酵素様G(CSLG)酵素、サポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、及びアシルトランスフェラーゼを含む。

20

【0149】

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、酵素をコードする少なくとも1つの異種遺伝子を発現し、酵素は、サポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、UDP - グルコース6 - デヒドロゲナーゼ1、及びセルロース合成酵素様G(CSLG)からなる群から選択される。図1は、コレステロールから始まるナス科の植物（ここではトマト）のトリテルペノイド生合成経路におけるステロイド系アルカロイド及びステロイド系サポニンの非限定的な例を示している。当業者は、この経路が、サポニンアグリコン及びグリコシル化ステロイド系アルカロイド生合成中間体及びグリコシル化サポニン生合成中間体の産生を包含し、これにより、ステロイド系アルカロイド（例えば、 - トマチン）またはステロイド系サポニン（例えば、ウットロシドB）が産生されることを認識するであろう。図20Aは、ハウレンソウにおけるトリテルペノイド系サポニン経路の非限定的な例を示している。当業者は、この経路がサポニンアグリコン及びグリコシル化サポニン生合成中間体の産生を包含し、これにより、トリテルペノイド系サポニン（Yossoside V（ヨッソシドV）；化合物11）が産生されることを認識するであろう。

30

【0150】

ステロイド系アルカロイド/ステロイド系サポニン/トリテルペノイド系サポニン生合成経路における遺伝子の少なくとも1つの発現を改変させると、いくつかの実施形態では、ステロイド系アルカロイドプロファイルまたはその経路の中間体、ステロイド系サポニンプロファイルまたはその経路の中間体、またはトリテルペノイド系サポニンプロファイルまたはその経路の中間体に、付随した改変が生じる。いくつかの実施形態では、遺伝子組換え細胞における発現は、対応する非組換え細胞と比較して増加している。いくつかの実施形態では、発現の増加は、細胞内に以前には存在しなかった酵素をコードする遺伝子からのデノボ発現を含む。いくつかの実施形態では、発現の増加は、細胞内に既に存在する遺伝子の発現の増加を含む。いくつかの実施形態では、発現の増加は、コードされた酵素の量の増加をもたらす。

40

【0151】

50

いくつかの実施形態では、ステロイド系アルカロイドのプロファイルの改変は、少なくとも1つのステロイド系アルカロイドまたはその生合成中間体の増加を含む。いくつかの実施形態では、ステロイド系アルカロイドのプロファイルの改変は、少なくとも1つのステロイド系アルカロイドまたはその生合成中間体の減少を含む。

【0152】

いくつかの実施形態では、ステロイド系サポニンのプロファイルの改変は、少なくとも1つのステロイド系サポニンまたはその生合成中間体の増加を含む。いくつかの実施形態では、ステロイド系サポニンのプロファイルの改変は、少なくとも1つのステロイド系サポニンまたはその生合成中間体の減少を含む。

【0153】

いくつかの実施形態では、トリテルペノイド系サポニンのプロファイルの改変は、少なくとも1つのトリテルペノイド系サポニンまたはその生合成中間体の増加を含む。いくつかの実施形態では、トリテルペノイド系サポニンのプロファイルの改変は、少なくとも1つのトリテルペノイド系サポニンまたはその生合成中間体の減少を含む。

【0154】

いくつかの実施形態では、ステロイド系アルカロイド/ステロイド系サポニン/トリテルペノイド系サポニン生合成経路の少なくとも1つの酵素をコードするヌクレオチド配列を、異種細胞（例えば、以下に限定しないが、植物細胞、藻類細胞、酵母細胞、昆虫細胞または細菌細胞など）または生物（例えば、以下に限定しないが、植物など）に導入することによって、少なくとも1つの酵素の発現が増加し、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの産生量が増加する。いくつかの実施形態では、ステロイド系アルカロイド/ステロイド系サポニン/トリテルペノイド系サポニン生合成経路の少なくとも1つの酵素をコードするヌクレオチド配列を、異種細胞または生物に導入することによって、異種細胞（例えば、以下に限定しないが、植物細胞、藻類細胞、酵母細胞、昆虫細胞または細菌細胞など）または生物（例えば、以下に限定しないが、植物など）が天然には産生していないステロイド系アルカロイド、ステロイド系サポニン、トリテルペノイド系サポニンが産生される。

【0155】

本明細書の記載を通して、遺伝子組換え細胞及びその使用について説明する。いくつかの実施形態では、遺伝子組換え細胞は、藻類細胞、酵母細胞、昆虫細胞または細菌細胞のいずれかから選択される。当業者であれば、本明細書に開示された遺伝子及び核酸配列を使用して、哺乳類細胞などの真核細胞、または細菌などの原核細胞を遺伝子組換えして、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法を知ることができるであろう。さらに、当業者であれば、本明細書に開示された遺伝子及び核酸配列を使用して、無細胞のインビトロシステムで少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生することができるであろう。いくつかの実施形態では、遺伝子組換え細胞は植物細胞であり、植物細胞は、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生するために使用される。いくつかの実施形態では、遺伝子組換え細胞は、植物細胞または酵母細胞であり、植物細胞または酵母細胞は、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生するために使用される。いくつかの実施形態では、遺伝子組換え細胞は酵母細胞であり、酵母細胞は、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生するために使用される。いくつかの実施形態では、遺伝子組換え細胞は藻類細胞であり、藻類細胞は、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生するために使用される。いくつかの実施形態では、遺伝子組換え細胞は昆虫細胞であり、昆虫細胞は、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生するために使用される。いくつかの実施形態では、遺伝子組換え細胞は細菌であり、細菌は、少なくとも

10

20

30

40

50

1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生するために使用される。いくつかの実施形態では、遺伝子組換え細胞は、植物またはその一部で構成される植物細胞であり、この植物細胞は、植物またはその一部において、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生するために使用される。いくつかの実施形態では、遺伝子組換え細胞は、植物またはその一部で構成される植物細胞であり、植物細胞は、植物またはその一部において、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの産生を減少させるために使用される。いくつかの実施形態では、遺伝子組換え細胞は、植物またはその一部で構成される植物細胞であり、植物細胞は、植物またはその一部において、少なくとも1つのステロイド性アルカロイド、ステロイド性サポニン、またはトリテルペノイド系サポニンの産生を増加させるために使用される。

10

**【0156】**

いくつかの実施形態では、異種発現は、1つの植物種からの少なくとも1つの異種植物遺伝子を、第2の植物種で発現させることを含む（例えば、下記の実施例23を参照）。いくつかの実施形態では、異種発現は、或る植物種からの少なくとも1つの異種植物遺伝子を酵母細胞で発現させることを含む（例えば、下記の実施例21を参照）。いくつかの実施形態では、異種発現は、或る植物種からの少なくとも1つの異種植物遺伝子を藻類細胞で発現させることを含む。いくつかの実施形態では、異種発現は、或る植物種からの少なくとも1つの異種植物遺伝子を、細菌細胞で発現させることを含む。いくつかの実施形態では、異種発現は、或る植物種からの少なくとも1つの異種植物遺伝子を昆虫細胞で発現させることを含む。いくつかの実施形態では、異種発現は、或る植物種からの少なくとも1つの異種植物遺伝子を、異なる種の植物の細胞で発現させることを含む。

20

**【0157】**

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの生合成経路から選択される酵素をコードする少なくとも1つの異種遺伝子を発現させることにより、その経路によって産生されるステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が改変されるか、少なくとも1つのその生合成中間体の含有量が改変されるか、あるいは、その組み合わせである。いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの生合成経路から選択された少なくとも2つの酵素をコードする1以上の異種遺伝子を発現させることにより、その経路によって産生されるステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が改変されるか、少なくとも1つのその生合成中間体の含有量が改変されるか、あるいは、その組み合わせである。いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの生合成経路から選択された少なくとも3つの酵素をコードする1以上の異種遺伝子を発現させることにより、その経路によって産生されるステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が改変されるか、少なくとも1つのその生合成中間体の含有量が改変されるか、あるいは、その組み合わせである。いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの生合成経路から選択された少なくとも4つの酵素をコードする1以上の異種遺伝子を発現させることにより、その経路によって産生されるステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が改変されるか、少なくとも1つのその生合成中間体の含有量が改変されるか、あるいは、その組み合わせである。いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの生合成経路から選択された少なくとも5つの酵素をコードする1以上の異種遺伝子を発現させることにより、その経路によって産生されるステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が改変されるか、少なくとも1つのその生合成中間体の含有量が改変されるか、あるいは

30

40

50

、その組み合わせである。いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの生合成経路から選択される少なくとも6つの酵素をコードする1以上の異種遺伝子を発現させることにより、その経路によって産生されるステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が改変されるか、少なくとも1つのその生合成中間体の含有量が改変されるか、あるいは、その組み合わせである。いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの生合成経路から選択された少なくとも7つの酵素をコードする1以上の異種遺伝子を発現させることにより、その経路によって産生されるステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が改変されるか、少なくとも1つのその生合成中間体の含有量が改変されるか、あるいは、その組み合わせである。いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの生合成経路から選択される少なくとも8つの酵素をコードする1以上の異種遺伝子を発現させることにより、その経路によって産生されるステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が改変されるか、少なくとも1つのその生合成中間体の含有量が改変されるか、あるいは、その組み合わせである。いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの生合成経路から選択される少なくとも9つの酵素をコードする1以上の異種遺伝子を発現させることにより、その経路によって産生されるステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が改変されるか、少なくとも1つのその生合成中間体の含有量が改変されるか、あるいは、その組み合わせである。いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの生合成経路から選択された少なくとも10の酵素をコードする2以上の異種遺伝子を発現させることにより、その経路によって産生されるステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が改変されるか、または少なくとも1つのその生合成中間体の含有量が改変されるか、あるいは、その組み合わせである。

【0158】

いくつかの実施形態では、セルロース合成酵素様G(CSLG)酵素をコードする少なくとも1つの異種遺伝子の発現が、対応する非組換え細胞と比較して増加した遺伝子組換え細胞が本明細書に開示されている。この遺伝子組換え細胞は、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え細胞と比較して増加している。いくつかの実施形態では、遺伝子組換え細胞は、酵素をコードする少なくとも1つの追加の異種遺伝子をさらに発現し、酵素は、サポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、UDP - グルコース6 - デヒドロゲナーゼ1、及びそれらの任意の組み合わせからなる群から選択される。

【0159】

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの誘導体は、それぞれステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンのグリコシル化誘導体を含む。

【0160】

いくつかの実施形態では、遺伝子組換え細胞は、異種CSLG及び異種サポニン - アミリン合成酵素を発現する。いくつかの実施形態では、遺伝子組換え細胞は、異種のCSLG及び少なくとも1つの異種シトクロムP450を発現する。いくつかの実施形態では、遺伝子組換え細胞は、異種CSLG及び1、2、または3つの異種シトクロムP450を発現する。一部の実施形態では、遺伝子組換え細胞は、異種CSLG、異種サポニン - アミリン合成酵素、及び1、2、または3つの異種シトクロムP450を発現する。いくつかの実施形態では、遺伝子組換え細胞は、異種のCSLG及び少なくとも1つの異種

のグリコシルトランスフェラーゼを発現する。いくつかの実施形態では、遺伝子組換え細胞は、異種CSLG及び少なくとも1、2、3、または4つの異種グリコシルトランスフェラーゼを発現する。いくつかの実施形態では、遺伝子組換え細胞は、異種CSLG、異種サポニン - アミリン合成酵素、及び少なくとも1、2、3、または4つの異種グリコシルトランスフェラーゼを発現する。いくつかの実施形態では、遺伝子組換え細胞は、異種CSLG、異種サポニン - アミリン合成酵素、少なくとも1、2、または3つの異種シトクロムP450、及び少なくとも1、2、3、または4つの異種グリコシルトランスフェラーゼを発現する。

#### 【0161】

いくつかの例では、必要な基質が利用できるように、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの経路に特に属さない追加の遺伝子を、同種または異種の遺伝子で発現させることが必要であり得る。いくつかの実施形態では、遺伝子組換え細胞は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの生合成経路の基質の産生に必要な酵素をコードする異種遺伝子が発現する。

10

#### 【0162】

いくつかの実施形態では、サポニン - アミリン合成酵素、シトクロムP450、C-2ヒドロキシラーゼ、C-23オキシダーゼ、グリコシルトランスフェラーゼ、UDP-グルコース6-デヒドロゲナーゼ1、アシルトランスフェラーゼ、またはセルロース合成酵素様G(CSLG)から選択される酵素をコードする異種遺伝子が、遺伝子組換え細胞内で発現される。いくつかの実施形態では、サポニン - アミリン合成酵素、シトクロムP450、C-2ヒドロキシラーゼ、C-23オキシダーゼ、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、またはセルロース合成酵素様G(CSLG)から選択される酵素をコードする異種遺伝子が、遺伝子組換え細胞内で発現される。いくつかの実施形態では、サポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、UDP-グリコシルトランスフェラーゼ、フコシルトランスフェラーゼ、キシロシルトランスフェラーゼ、アシルトランスフェラーゼ、BAHD-アシルトランスフェラーゼ、またはセルロース合成酵素様G(CSLG)から選択される酵素をコードする異種遺伝子が、遺伝子組換え細胞内で発現される。

20

#### 【0163】

いくつかの実施形態では、遺伝子組換え細胞は、CSLG活性を有する酵素をコードする少なくとも1つの異種遺伝子が発現する。いくつかの実施形態では、遺伝子組換え細胞は、サポニン - アミリン合成酵素活性を有する酵素をコードする少なくとも1つの異種遺伝子が発現する。いくつかの実施形態では、遺伝子組換え細胞は、シトクロムP450活性を有する酵素をコードする少なくとも1つの異種遺伝子が発現する。いくつかの実施形態では、遺伝子組換え細胞は、シトクロムP450活性を有する2つの酵素をコードする少なくとも2つの異種遺伝子が発現する。いくつかの実施形態では、遺伝子組換え細胞は、シトクロムP450活性を有する3つの酵素をコードする少なくとも3つの異種遺伝子が発現する。いくつかの実施形態では、シトクロムP450活性を有する酵素は、C-2ヒドロキシラーゼ活性を有する。いくつかの実施形態では、シトクロムP450活性を有する酵素は、C-23オキシダーゼ活性を有する。

30

40

#### 【0164】

いくつかの実施形態では、遺伝子組換え細胞は、サポニン - アミリン合成酵素活性、シトクロムP450活性、及びCSLG活性を有する酵素をコードする異種遺伝子が発現し、シトクロムP450酵素は、C-2ヒドロキシラーゼ活性を提供する酵素、C-23オキシダーゼ活性を提供する酵素、シトクロムP450活性、C-2ヒドロキシラーゼ活性及びC-23オキシダーゼ活性を提供する酵素の組み合わせ、を含んでいてもよい。

#### 【0165】

いくつかの実施形態では、遺伝子組換え細胞は、グリコシルトランスフェラーゼ酵素活性をコードする少なくとも1つの異種遺伝子が発現する。いくつかの実施形態では、グリ

50

コシルトランスフェラーゼ酵素は、UDP - グリコシルトランスフェラーゼ、フコシルトランスフェラーゼ、またはキシロシルトランスフェラーゼを含む。いくつかの実施形態では、遺伝子組換え細胞は、グリコシルトランスフェラーゼ酵素をコードする複数の異種遺伝子を発現する。いくつかの実施形態では、遺伝子組換え細胞は、UDP - グリコシルトランスフェラーゼ、フコシルトランスフェラーゼ、キシロシルトランスフェラーゼ、またはそれらの組み合わせから選択される、グリコシルトランスフェラーゼをコードする複数の異種遺伝子を発現する。いくつかの実施形態では、遺伝子組換え細胞は、当該技術分野で知られている既知のグリコシルトランスフェラーゼの組み合わせから選択される、グリコシルトランスフェラーゼをコードする複数の異種遺伝子を発現する。いくつかの実施形態では、遺伝子組換え細胞は、CSLG活性をコードする異種遺伝子、及び、少なくとも1つのグリコシルトランスフェラーゼ酵素活性をコードする異種遺伝子を発現する。いくつかの実施形態では、遺伝子組換え細胞は、CSLG活性をコードする異種遺伝子、及び、UDP - グリコシルトランスフェラーゼ、フコシルトランスフェラーゼ、キシロシルトランスフェラーゼ、またはそれらの組み合わせから選択される少なくとも1つのグリコシルトランスフェラーゼ酵素活性をコードする異種遺伝子を発現する。いくつかの実施形態では、遺伝子組換え細胞は、CSLG活性をコードする異種遺伝子、及び、UDP - グリコシルトランスフェラーゼ、フコシルトランスフェラーゼ、キシロシルトランスフェラーゼ、当該技術分野で知られている任意のグリコシルトランスフェラーゼ、またはそれらの組み合わせから選択される少なくとも1つのグリコシルトランスフェラーゼ酵素活性をコードする異種遺伝子を発現する。

10

20

**【0166】**

いくつかの実施形態では、遺伝子組換え細胞は、アシルトランスフェラーゼをコードする異種遺伝子を発現する。いくつかの実施形態では、遺伝子組換え細胞は、ベンジルアルコールアセチル、アントシアニン - O - ヒドロキシ - シンナモイル、アントラニレート - N - ヒドロキシ - シンナモイル / ベンゾイル、デアセチルピンドリンアセチルトランスフェラーゼ、(BAHD)アシルトランスフェラーゼをコードする異種遺伝子を発現する。いくつかの実施形態では、遺伝子組換え細胞は、CSLG活性を提供する酵素をコードする異種遺伝子、及び、アセチルトランスフェラーゼ活性をコードする酵素をコードする異種遺伝子を発現する。いくつかの実施形態では、遺伝子組換え細胞は、CSLG活性を提供する酵素をコードする異種遺伝子、及び、BAHDアセチルトランスフェラーゼ活性をコードする酵素をコードする異種遺伝子を発現する。

30

**【0167】**

いくつかの実施形態では、遺伝子組換え細胞は、トリテルペノイド系サポニン基質に対する酵素活性をコードする異種遺伝子を発現し、この酵素は、グリコシルトランスフェラーゼを含む。いくつかの実施形態では、遺伝子組換え細胞は、トリテルペノイド系サポニン基質に対する酵素活性をコードする異種遺伝子を発現し、この酵素は、UDP - グルコース6 - デヒドロゲナーゼ1を含む。いくつかの実施形態では、例えば、細胞が細菌細胞である場合、異種遺伝子は、スクアレインエポキシダーゼ、シトクロムP450レダクターゼ、UDP - グルコース6 - デヒドロゲナーゼ1、またはそれらの任意の組み合わせをコードする。

40

**【0168】**

いくつかの実施形態では、少なくとも1つのCSLG遺伝子の発現が、対応する非組換え細胞と比較して改変された少なくとも1つの細胞を含む遺伝子組換え植物が本明細書に開示されている。この遺伝子組換え植物は、少なくとも1つの細胞が異種CSLG遺伝子を含み、発現の改変が異種遺伝子のデノボ発現を含む。いくつかの実施形態では、少なくとも1つのCSLG遺伝子の発現が、対応する非組換え細胞と比較して改変された少なくとも1つの細胞を含む遺伝子組換え植物が本明細書に開示されている。この遺伝子組換え植物は、少なくとも1つの細胞はが異種CSLG遺伝子を含み、発現の改変が、機能的であってもよく非機能的であってもよい内在性遺伝子の存在下での異種遺伝子の発現を含む。

50

## 【0169】

いくつかの実施形態では、遺伝子組換え植物における異種CSLG遺伝子の発現は、対応する非組換え細胞と比較して増加し、その遺伝子組換え植物の細胞は、少なくとも1つのトリテルペノイド系サポニン、その代謝物、その誘導体、またはその生合成中間体の含有量が、対応する非組換え植物と比較して改変される。

## 【0170】

いくつかの実施形態では、本明細書に開示されているのは、トリテルペノイド系サポニン生合成経路に関与する酵素の配列である。この生合成経路は、サポニンアグリコン及びグリコシル化サポニン、及びグリコシル化トリテルペノイド系サポニンの産生を含む。いくつかの実施形態では、本明細書に開示されているのは、トリテルペノイド系サポニンの生合成経路に関与する酵素をコードする遺伝子の配列である。この生合成経路は、サポニンアグリコン及びグリコシル化サポニン、及びグリコシル化トリテルペノイド系サポニンを産生し、そのようなものとしては、これに限定しないが、例えば、3-O-[ -D-キシロピラノシル-(1 3)- -D-グルクロノピラノシル]-28-O-[ -D-グルコピラノシル-( 4)- -L-ラムノピラノシル-(1 2)-4-アセチル- -D-フコピラノシル]-メディカゲン酸(Yossoside V(ヨッソシドV); 化合物11)、3-O-[ -D-キシロピラノシル-(1 3)- -D-グルクロノピラノシル]-28-O-[ -D-グルコピラノシル-(1 4)- -L-ラムノピラノシル-(1 2)- -D-フコピラノシル]-メディカゲン酸(Yossoside IV(ヨッソシドIV); 化合物10)、3-O-[ -D-グルクロノピラノシル]-28-O-[ -D-グルコピラノシル-(1 4)- -L-ラムノピラノシル-(1 2)- -D-フコピラノシル]-メディカゲン酸(Yossoside III(ヨッソシドIII); 化合物9)、3-O-[ -D-グルクロノピラノシル]-28-O-[ -L-ラムノピラノシル-(1 2)- -D-フコピラノシル]-メディカゲン酸(Yossoside II(ヨッソシドII); 化合物8)、及び、3-O- -D-グルクロノピラノシル-28-O- -D-フコピラノシル-メディカゲン酸(Yossoside I(ヨッソシドI); 化合物7)(図20A参照)を含む。

## 【0171】

例えば、これに限定しないが、実施例において詳細に説明される化合物11のハウレンソウ・トリテルペノイド生合成経路で活性化する酵素が挙げられる。実施例20に提供される表16は、CSLG酵素がメディカゲン酸にグルクロン酸を付加してトリテルペノイド系サポニンであるメディカゲン酸3-グルクロニドを作る、ハウレンソウ生合成経路中の酵素のアミノ酸配列及び核酸配列を示す。実施例24において開示される表19は、シロイヌナズナから単離されたCSLGのアミノ酸配列を提供する。実施例において開示される表17は、他の植物種からのCSLGのホモログの核酸配列及びアミノ酸配列を提供する。

## 【0172】

いくつかの実施形態では、本明細書で開示される遺伝子組換え細胞は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる酵素をコードする少なくとも1つの異種遺伝子を発現する。いくつかの実施形態では、本明細書で開示される遺伝子組換え細胞は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる酵素をコードする異種遺伝子の配列を発現する。いくつかの実施形態では、本明細書で開示される遺伝子組換え細胞は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる2つの酵素をコードする少なくとも2つの異種遺伝子を発現する。いくつかの実施形態では、本明細書で開示される遺伝子組換え細胞は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる3つの酵素をコードする少なくとも3つの異種遺伝子を発現する。いくつかの実施形態では、本明細書で開示される遺伝子組換え細胞は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる4つの酵素をコードする少なくとも4つの異

10

20

30

40

50

種遺伝子を発現する。いくつかの実施形態では、本明細書で開示される遺伝子組換え細胞は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる5つの酵素をコードする少なくとも5つの異種遺伝子を発現する。いくつかの実施形態では、本明細書で開示される遺伝子組換え細胞は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる6つの酵素をコードする少なくとも6つの異種遺伝子を発現する。いくつかの実施形態では、本明細書で開示される遺伝子組換え細胞は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる7つの酵素をコードする少なくとも7つの異種遺伝子を発現する。いくつかの実施形態では、本明細書で開示される遺伝子組換え細胞は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる8つの酵素をコードする少なくとも8つの異種遺伝子を発現する。いくつかの実施形態では、本明細書で開示される遺伝子組換え細胞は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる9つの酵素をコードする少なくとも9つの異種遺伝子を発現する。いくつかの実施形態では、本明細書で開示される遺伝子組換え細胞は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる10の酵素をコードする少なくとも10の異種遺伝子を発現する。

10

20

30

40

50

**【0173】**

いくつかの実施形態では、少なくとも1つのセルロース合成酵素様G(CSLG)遺伝子の発現が、対応する非組換え植物におけるCSLG遺伝子の発現と比較して改変された少なくとも1つの細胞を含む遺伝子組換え植物が本明細書に開示されている。この遺伝子組換え植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して改変されている。いくつかの実施形態では、CSLG遺伝子は、内在性CSLG遺伝子を含む。いくつかの実施形態では、CSLG遺伝子は、異種CSLG遺伝子を含む。

**【0174】**

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの誘導体は、それぞれ、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンのグリコシル化誘導体を含む。

**【0175】**

いくつかの実施形態では、本明細書で開示される遺伝子組換え植物またはその一部は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる酵素をコードする少なくとも1つの異種遺伝子を発現する。いくつかの実施形態では、本明細書で開示される遺伝子組換え植物またはその一部は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる酵素をコードする異種遺伝子の配列を発現する。いくつかの実施形態では、本明細書で開示される遺伝子組換え植物またはその一部は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる2つの酵素をコードする少なくとも2つの異種遺伝子を発現する。いくつかの実施形態では、本明細書で開示される遺伝子組換え植物またはその一部は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる3つの酵素をコードする少なくとも3つの異種遺伝子を発現する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物またはその一部は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる4つの酵素をコードする少なくとも4つの異種遺伝子を発現する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物またはその一部は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる5つの酵素をコードする少なくとも5つの異種遺伝子を発現する。

いくつかの実施形態では、本明細書で開示される遺伝子組換え植物またはその一部は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる6つの酵素をコードする少なくとも6つの異種遺伝子を発現する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物またはその一部は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる7つの酵素をコードする少なくとも7つの異種遺伝子を発現する。いくつかの実施形態では、本明細書で開示される遺伝子組換え植物またはその一部は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる8つの酵素をコードする少なくとも8種類の異種遺伝子を発現する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物またはその一部は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる9つの酵素をコードする少なくとも9つの異種遺伝子を発現する。いくつかの実施形態では、本明細書で開示される遺伝子組換え植物またはその一部は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路に含まれる10の酵素をコードする少なくとも10の異種遺伝子を発現する。

10

20

30

40

50

**【0176】**

いくつかの実施形態では、異種遺伝子は、 - アミリン合成酵素、シトクロムP450、複数のシトクロムP450、グリコシルトランスフェラーゼ、複数のグリコシルトランスフェラーゼ、アシルトランスフェラーゼ、CSLG酵素、またはそれらの任意の組み合わせから選択される酵素を含む、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路中の酵素をコードする。いくつかの実施形態では、異種遺伝子は、 - アミリン合成酵素、シトクロムP450、複数のシトクロムP450、グリコシルトランスフェラーゼ、複数のグリコシルトランスフェラーゼ、アシルトランスフェラーゼ、CSLG酵素、またはそれらの任意の組み合わせから選択される酵素を含む、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路中の酵素をコードし、さらに、生合成経路の基質を提供するために必要な少なくとも1つの異種遺伝子を発現させる。いくつかの実施形態では、異種遺伝子は、 - アミリン合成酵素、シトクロムP450、複数のシトクロムP450、グリコシルトランスフェラーゼ、複数のグリコシルトランスフェラーゼ、アシルトランスフェラーゼ、CSLG酵素、またはそれらの任意の組み合わせを含むステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドの生合成経路中の酵素をコードし、さらに、これに限定しないが例えばUDP-グルコース6-デヒドロゲナーゼ1などの生合成経路の基質を提供するために必要な異種遺伝子を発現させる。

**【0177】**

いくつかの実施形態では、遺伝子組換え細胞は、ステロイド系アルカロイド/ステロイド系サポニン/トリテルペノイド系サポニン生合成経の酵素をコードする少なくとも1つの異種遺伝子を発現し、この少なくとも1つの異種遺伝子はCSLG酵素をコードする。

**【0178】**

いくつかの実施形態では、少なくとも1つの異種遺伝子が、C2-ヒドロキシラーゼまたはC-23オキシダーゼを含むシトクロムP450をコードするか、フルクトシルトランスフェラーゼ、キシロシルトランスフェラーゼ、またはUDP-グリコシルトランスフェラーゼを含むグリコシルトランスフェラーゼをコードするか、ベンジルアルコールアセチル、アントシアニン-O-ヒドロキシ-シンナモイル、アントラニレート-N-ヒドロキシ-シンナマイオール/ベンゾイル、またはデアセチルピンドリン(BAHD)アセチルトランスフェラーゼをコードするか、あるいは、それらの組み合わせである。

**【0179】**

いくつかの実施形態では、異種遺伝子または遺伝子の発現は、対応する非組換え細胞と比較して、発現の増加を含む。いくつかの実施形態では、遺伝子の発現の増加は、対応する非組換え細胞における酵素の含有量と比較して、その遺伝子によってコードされる酵素の含有量の増加をもたらす。当業者であれば、いくつかの実施形態では、発現量の増加は

、遺伝子の過剰発現を含み、その結果、対応する非組換え細胞と比較して、遺伝子によってコードされる酵素の存在量が増加することを理解できるであろう。いくつかの実施形態では、発現の増加は、細胞または生物、これに限定しないが、例えば、植物細胞、植物、酵母細胞、藻類細胞、昆虫細胞、または細菌細胞などでは自然に見られない遺伝子の発現、及びその結果としての酵素活性または活性の産生を含む。

【0180】

当業者であれば、いくつかの実施形態では、対応する非組換え細胞が、異種遺伝子を発現しないかまたは過剰に発現しない対照細胞を提供することを理解できるであろう。

【0181】

いくつかの実施形態では、異種細胞における遺伝子または遺伝子の発現は、遺伝子または遺伝子の一過性発現を含む。いくつかの実施形態では、異種細胞における遺伝子または遺伝子の発現は、遺伝子または遺伝子の構成的発現を含む。いくつかの実施形態では、異種細胞での遺伝子または遺伝子の発現は、遺伝子または遺伝子の一過性発現、遺伝子または遺伝子の構成的発現、またはそれらの組み合わせを含む。

10

【0182】

当業者であれば、一過性発現とは、核酸が真核細胞、例えば植物細胞や酵母細胞に導入された後に短時間で発現する、遺伝子の一時的な発現を意味することを理解できるであろう。逆に、構成的発現（安定発現とも呼ばれる）は、遺伝子の継続的な発現を包含する。

【0183】

当業者であれば、「遺伝子」という用語が、RNAまたはポリペプチドの産生に必要なコード化配列を含む核酸（例えば、DNAまたはRNA）配列を包含することを理解できるであろう。ポリペプチドは、完全長のコード配列またはその一部によってコードされる。遺伝子の「一部」は、その遺伝子の断片を意味する。断片の大きさは、数ヌクレオチドから、遺伝子配列全体から1ヌクレオチドを除いたものまで様々である。したがって、「遺伝子の少なくとも一部を含む核酸配列」は、その遺伝子の断片を含むか、または、遺伝子全体を含む。

20

【0184】

当業者であれば、「遺伝子」という用語は、任意選択で、構造遺伝子のコード領域を包含し、遺伝子が完全長のmRNAの長さに対応するように、5'末端及び3'末端の両方のコード領域に隣接して、いずれかの末端に約1kbの距離で位置する配列を含むことを理解できるであろう。コード領域の5'に位置し、mRNA上に存在する配列は、5'非翻訳配列と呼ばれる。また、コード領域の3'または下流に位置し、mRNA上に存在する配列は、3'非翻訳配列と呼ばれる。

30

【0185】

いくつかの実施形態では、遺伝子は、コード領域を含むDNA配列を含む。いくつかの実施形態では、遺伝子は、コード領域だけでなく、遺伝子のエクソン及び任意の介在イントロンを含む上流領域及び下流領域を有するDNA配列を含む。いくつかの実施形態では、上流領域及び下流領域は、非コード制御領域を含む。いくつかの実施形態では、上流領域及び下流領域は、これに限定しないが、例えば、プロモータ、エンハンサー、及びサイレンサーなどの制御配列を含む。制御配列の非限定的な例としては、これに限定しないが、AGGAボックス、TATAボックス、Inr、DPE、ZmUbi1、PvUbi1、PvUbi2、CaMV、35S、OsAct1、zE19、E8、TA29、A9、pDJ3S、B33、PAT1、alcA、G-box、ABRE、DRE、及びPCNAが挙げられる。

40

【0186】

調節領域は、いくつかの実施形態では、細胞、これに限定しないが、例えば、酵母、藻類、昆虫、細菌、または本明細書に記載されている植物細胞または植物内の特定の遺伝子の発現を増加または減少させることができる。調節領域は、いくつかの実施形態では、植物細胞内の特定の遺伝子の発現を増加または減少させることができる。調節領域は、いくつかの実施形態では、藻類細胞内の特定の遺伝子の発現を増加または減少させることがで

50

きる。制御領域は、いくつかの実施形態では、酵母細胞内の特定の遺伝子の発現を増加または減少させることができる。制御領域は、いくつかの実施形態では、昆虫細胞内の特定の遺伝子の発現を増加または減少させることができる。制御領域は、いくつかの実施形態では、細菌内の特定の遺伝子の発現を増加または減少させることができる。

【0187】

いくつかの実施形態では、遺伝子は、その遺伝子のエクソン及び任意の介在イントロンを含むその遺伝子のコード領域を含む。いくつかの実施形態では、遺伝子は、その制御配列を含む。いくつかの実施形態では、遺伝子は、遺伝子プロモータを含む。いくつかの実施形態では、遺伝子は、そのエンハンサー領域を含む。いくつかの実施形態では、遺伝子は、5´非コード配列を含む。いくつかの実施形態では、遺伝子は、3´非コード化配列を含む。

10

【0188】

いくつかの実施形態では、当業者であれば、DNAが、遺伝子のコード領域だけでなく、上流配列及び下流配列を有する遺伝子を含むことを理解できるであろう。他の実施形態では、DNAは、cDNA（相補的DNA）を含む。当業者であれば、cDNAは、逆転写酵素の作用によりRNAから逆転写された合成DNAを包含することを理解できるであろう。cDNAは、一本鎖または二本鎖であってよく、RNA配列の一部と実質的に同一の配列、またはRNA配列の一部と相補的な配列のいずれかまたは両方を有する鎖を含むことができる。さらに、cDNAは、上流調節配列及び下流調節配列を含むことができる。さらに他の実施形態では、DNAは、CDS（完全コード配列）を含む。当業者であれば、CDSが、完全長のタンパク質またはポリペプチドをコードするDNA配列を包含することを理解できるであろう。CDSは、通常、開始コドン（「ATG」）で始まり、最初のフレーム内停止コドン（「TAA」、「TAG」、「TGA」）で終わる（またはその1つ前で終わる）。当業者であれば、一実施形態では、cDNAがCDSを含むことを認識できるであろう。

20

【0189】

本明細書では、「ポリヌクレオチド」、「ポリヌクレオチド配列」、「核酸配列」、及び「単離されたポリヌクレオチド」という用語は、互換的に使用される。これらの用語は、ヌクレオチド配列などを包含する。ポリヌクレオチドは、一本鎖または二本鎖、直線状または分岐状であり、任意選択で、合成、非天然、または変化したヌクレオチド塩基を含む、RNAまたはDNAのポリマーまたはそれらのハイブリッドであってもよい。また、RNA/DNAハイブリッドも包含される。

30

【0190】

本明細書では、「相補的」または「その相補体」という用語を、相補的領域の全体にわたって別の指定されたポリヌクレオチドとワトソン&クリック塩基対を形成することができるポリヌクレオチドの配列を指すために使用している。この用語は、ポリヌクレオチドのペアにはそれらの配列のみに基づいて適用されるが、2つのポリヌクレオチドが実際に結合する特定の条件のセットには適用されない。

【0191】

本明細書で使用するとき、「構築物」という用語は、目的のポリヌクレオチドを含む、人工的に組み立てられたまたは単離された核酸分子を指す。一般に、構築物は、目的のポリヌクレオチドまたはポリヌクレオチド、場合によっては目的の遺伝子でもあるマーカー遺伝子、及び適切な調節配列を含むことができる。構築物に制御配列を含めることは任意選択であり、例えば、宿主細胞の制御配列を使用する状況では、そのような配列は必要とされないことを理解すべきである。制御要素としては、これに限定しないが、プロモータ、エンハンサー、複製起点、転写終結配列、ポリアデニル化シグナルなどが挙げられる。「構築物」という用語は、ベクターを含むが、これに限定されるものと見なすべきではない。

40

【0192】

「作動可能に連結される」という用語は、一方の機能が他方の機能によって制御される

50

ように、単一の核酸断片上の核酸配列が連結していることを指す。例えば、プロモータは、そのプロモータがそのコード配列の発現を調節することができる場合（すなわち、コード配列がプロモータの転写制御下にある場合）、コード配列と作動可能に連結される。コード配列は、センスまたはアンチセンスの方向で制御配列に作動可能に連結することができる。いくつかの実施形態では、「作動可能に連結された」及び「機能的に連結された」という用語は、互換的に使用することができ、すべて同一の意味及び性質を有する。

#### 【0193】

本明細書で使用するとき、「プロモータ要素」、「プロモータ」、「プロモータ配列」という用語は、一般にDNAポリマーのコード領域の5'末端に位置する（すなわち、先行する）DNA配列を指す。自然界で知られているほとんどのプロモータの位置は、転写領域よりも前にある。プロモータは、遺伝子の発現を活性化するスイッチとして機能する。遺伝子が活性化されると、その遺伝子は転写される、あるいは転写に関与すると言われる。転写は、遺伝子からmRNAを合成する。したがって、プロモータは、転写制御要素としての役割を果たすとともに、遺伝子のmRNAへの転写を開始するための場所を提供する。

10

#### 【0194】

異種遺伝子を含むポリヌクレオチド配列は、構成的発現または調節された（一過性の）発現を指示するプロモータの制御下で、例えば、植物細胞、藻類細胞、酵母細胞、昆虫細胞、または細菌において発現される。いくつかの実施形態では、植物細胞は、植物全体またはその一部である。調節された発現は、時間的または空間的に調節された発現、及び他の形態の誘導性または抑制性の発現を含む。時間的とは、例えば植物細胞培養物の特定の成長速度が得られた時点などの特定の時点で、発現が誘導されることを意味する（例えば、プロモータは、定常期または発達の特定の段階でのみ誘導される）。空間的とは、プロモータが、特定の器官、組織、または細胞でのみ活性化されることを意味する（例えば、植物の場合、プロモータは、根、葉、表皮、孔辺細胞などでのみ活性化される）。調節された発現の他の例は、植物細胞、酵母細胞、藻類細胞、昆虫細胞、または細菌に対して化学的または物理的な刺激を加えることによって、活性が誘導または抑制されたプロモータを含む。いくつかの実施形態では、プロモータは、誘導性プロモータである。いくつかの実施形態では、プロモータは、構成的プロモータである。いくつかの実施形態では、トリテルペノイド生合成経路の異種遺伝子（これに限定しないが、例えばCSLG）の調節された発現は、植物細胞に対して化学的または物理的刺

20

30

40

#### 【0195】

いくつかの実施形態では、発現は、環境、ホルモン、化学、及び/または発生のシグナルの制御下にある。植物細胞のそのようなプロモータには、（1）熱、（2）光、（3）アブジジン酸、ジャスモン酸メチルなどのホルモン、（4）創傷、または（5）サリチル酸、キトサン、金属などの化学物質によって制御されるプロモータが含まれる。構成的プロモータは、広範囲の条件下で、広範囲の細胞における発現を誘導する。植物細胞において異種ポリペプチドを発現させるのに有用な構成的植物プロモータの例としては、これに

50

限定しないが、単子葉植物を含むほとんどの植物組織において構成的な高レベルの発現を付与するカリフラワーモザイクウイルス (CaMV) 35Sプロモータ、ノバリン合成酵素プロモータ、及びオクトピン合成酵素プロモータが挙げられる。いくつかの実施形態では、プロモータは、誘導性プロモータである。本明細書で提供されるいくつかの実施形態では、誘導性植物プロモータとしては、これに限定しないが、GAL1、GAL10、PHO5、及びCUP1が挙げられる。いくつかの実施形態では、酵母プロモータは、構成的プロモータである。本明細書で提供されるいくつかの実施形態では、構成的な酵母プロモータとしては、PGK1、TDHS、ADH1、CYC1、ACT1、及びTEF1が挙げられる。いくつかの実施形態では、藻類プロモータは、誘導性プロモータである。いくつかの実施形態では、本明細書で提供される藻類プロモータは、光、pH、塩分濃度、温度、または栄養分の変更によって調節される。いくつかの実施形態では、誘導性藻類プロモータは、窒素レベル、窒素ストレス、または窒素飢餓に応答する。いくつかの実施形態では、藻類プロモータは、構成的プロモータである。いくつかの実施形態では、昆虫細胞に使用するための誘導性プロモータは、環境または栄養のシグナルに応答する。

10

20

30

40

50

#### 【0196】

発現カセットは、通常、「発現ベクター」と呼ばれる、少なくとも1つの異種遺伝子を含むDNAまたはRNA構築物中に提供され、これは、任意の遺伝的要素（例えば、プラスミド、染色体、ウイルスは）であり、細胞内でポリヌクレオチドを複製する自律的単位として作用するか（すなわち、それ自体の制御下で複製することができる）、または、宿主細胞染色体への挿入により複製可能となり、それに別のポリヌクレオチド断片を付着させて、付着した断片の複製及び/または発現をもたらすことができる。適切なベクターとしては、これに限定しないが、プラスミド、バクテリオファージ、コスミド、植物ウイルス、人工染色体などが挙げられる。発現カセットは、少なくとも1つの複製系も有するDNA構築物中に提供してもよい。複製系に加えて、多くの場合は、少なくとも1つのマーカーが存在し、これは、1以上の宿主に有用であるか、または個々の宿主に対して異なるマーカーが存在する。

#### 【0197】

別の実施形態では、選択遺伝子は、抗生物質や除草剤などの選択剤に対する耐性を付与する選択マーカーを提供する。形質転換される可能性のある細胞は、選択剤に暴露される。生残細胞の集団は、一般に、耐性が付与された遺伝子が組み込まれ、細胞生存を可能にするのに十分なレベルで発現している細胞である。細胞は、外因性DNAが安定して組み込まれていることを確認するために、さらに検査される。

#### 【0198】

有用な選択マーカー遺伝子としては、例えばカナマイシン (nptII)、ハイグロマイシンB (aphIV)、nptII、hpt、aadA、及びゲンタマイシン (aac3及びaac4)などの抗生物質に対する耐性、または、グルホシネート (barまたはpat)及びグリホサート (EPSPS)などの除草剤に対する耐性を付与したものが挙げられる。別の実施形態では、選択遺伝子は、代謝拮抗薬である。一実施形態では、代謝拮抗薬は、dhfである。

#### 【0199】

形質転換体を視覚的に同定する能力を提供するスクリーニング可能なマーカー、例えば、cat、lacZ、uidA、ルシフェラーゼ、緑色蛍光タンパク質 (GFP)などの着色タンパク質や蛍光タンパク質を発現する遺伝子、または、様々な発色基質が知られている - グルクロニダーゼやuidA遺伝子 (GUS)を発現する遺伝子など、を使用してもよい。また、スクリーニング可能なマーカーと選択可能なマーカーとの組み合わせは、形質転換細胞の同定に有用であると考えられる。

#### 【0200】

一実施形態では、選択遺伝子は、増殖のための基質となるか、または形質転換された組織の増殖及び分化を誘導する非毒性物質を条件とする陽性選択可能なマーカー遺伝子である。

## 【0201】

一実施形態では、選択剤への曝露を生き延びた植物細胞、またはスクリーニングアッセイで陽性とスコア付けされた細胞を再生培地中で培養し、植物に成長させてもよい。成長した植物体は、約85%の相対湿度、600ppmのCO<sub>2</sub>、及び25~250マイクロアインシュタインm<sup>2</sup>/sの光に制御された環境制御チャンバ内で、土壌を用いない植物成長用混合物に移して寒さに慣れさせた後に、成長のための温室または成長チャンバに移す。植物は、成長チャンバまたは温室のいずれかで成長させることが好ましい。植物は、初期組織に応じて、形質転換体が同定されてから約6週間~10ヶ月後に再生される。再生中、細胞は、約19~28の固形培地上で植物に成長する。再生した植物は、シュート及び根の発達段階に達した後、さらなる成長と検査のために温室に移される。当業者に知られている従来の植物育種法を用いて植物に受粉させ、種子を作る。

10

## 【0202】

形質転換された植物から子孫を回収し、外因性組換えポリヌクレオチドの発現を検査する。有用なアッセイとしては、例えばサザンプロット法、ノーザンプロット法、PCR法などの「分子生物学的」アッセイ、RNA（例えば、二本鎖RNA）またはタンパク質産物の存在を、例えば免疫学的手段（ELISA法、ウェスタンプロット法）または酵素機能などによって検出する「生化学的」アッセイ、葉または根のアッセイなどの植物の一部のアッセイ、再生植物の全体の用語型を分析することによるアッセイが挙げられる。

## 【0203】

当業者であれば、コードされた核酸配列を比較的無傷の状態を導入するための適切なベクターを選択することができるであろう。したがって、コードされた核酸が導入された宿主細胞、例えば植物プロトプラストを産生する任意のベクターで十分である。ベクターの選択、あるいはベクターの使用の有無は、通常、選択された形質転換の方法によって導かれる。

20

## 【0204】

植物細胞の場合、マーカーは、(a)抗生物質、毒素、重金属、特定の糖などの殺生物剤に対する保護をコードするか、(b)補助要求性宿主に原栄養を与えることで相補性を提供するか、または、(c)植物内での新規化合物の産生により可視用語型を提供する。いくつかの実施形態では、使用される遺伝子としては、ネオマイシンホスホトランスフェラーゼ(NPTII)、ハイグロマイシンホスホトランスフェラーゼ(HPT)、クロラムフェニコールアセチルトランスフェラーゼ(CAT)、ニトリラーゼ、及びゲンタマイシン耐性遺伝子が挙げられる。植物の宿主を選択するための適切なマーカーとしては、これに限定しないが、インジゴ産生を可能にする - グルクロニダーゼ、可視光産生を可能にするルシフェラーゼ、緑色蛍光タンパク質及びその変異体、カナマイシン耐性またはG418耐性を提供するNPTII、ハイグロマイシン耐性を提供するHPT、及び、グリホサート耐性を提供する変異ARO A遺伝子などが挙げられる。

30

## 【0205】

酵母細胞の場合、マーカーとしては、補助要求性宿主に原栄養を与えることにより相補性を提供するマーカー、これに限定しないが、例えば、必要な相補的活性をコードするHIS、LEU、TRP、及びURAマーカー遺伝子などが挙げられる。いくつかの実施形態では、酵母マーカーは、酵母に対して毒性を有する化合物に対する抵抗性を提供し、そのようなものとしては、これに限定しないが、例えば、G418に対する抵抗性を提供するkanMX4、ノーセオスリシンに対する抵抗性を提供するnatMX4、ハイグロマイシンBに対する抵抗性を提供するhphMX4、グルホシネートに対する抵抗性を提供するpatMX3、及び、ゼオシンに対する抵抗性を提供するZEOなどが挙げられる。

40

## 【0206】

プロモータ活性は、そのプロモータに作動可能に連結されたポリヌクレオチド配列、その相同体、その変異体、またはその断片の転写の程度を包含する。プロモータ活性は、産生されたRNA転写物の量を例えばノーザンプロット法によって測定することによって直

50

接的に測定してもよいし、または、例えばレポーター遺伝子がプロモータに連結されている場合などには、RNA転写物でコードされている産物を測定することによって間接的に測定してもよい。

【0207】

いくつかの実施形態では、本明細書で開示される、酵素をコードする少なくとも1つの異種遺伝子は、プロモータに作動可能に連結される。いくつかの実施形態では、プロモータへの作動可能な連結は、一過性発現をもたらす。いくつかの実施形態では、プロモータへの作動可能な連結は、構成的発現をもたらす。いくつかの実施形態では、或る異種遺伝子が構成的プロモータに作動可能に連結され、別の異種遺伝子が一過性発現を提供するプロモータに作動可能に連結されてもよい。いくつかの実施形態では、ポリヌクレオチド配列は、酵素をコードする少なくとも1つの異種遺伝子を含み、ポリヌクレオチドは、任意選択でベクター内に含まれる。

10

【0208】

本明細書で使用するとき、「エンハンサー」という用語は、プロモータの活性を刺激することができるDNA配列を指し、プロモータの生来の要素であってもよいし、プロモータのレベルや組織特異性を高めるために挿入された異種の要素であってもよい。

【0209】

本明細書で使用するとき、「発現」という用語は、機能的最終産物、例えば、mRNAまたはタンパク質の産生を意味する。いくつかの実施形態では、遺伝子組換え細胞は少なくとも1つの異種遺伝子を発現し、機能的最終産物はmRNAを含み、このmRNAが転写されて、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを含む機能的最終産物を産生する。

20

【0210】

「遺伝子組換え細胞」という用語は、人間によって遺伝子組換えされた細胞を指す。いくつかの実施形態では、遺伝子組換えは、酵素をコードする遺伝子を含む異種ポリヌクレオチドで、植物細胞、酵母細胞、藻類細胞、昆虫細胞、または細菌を形質転換することを含む。いくつかの実施形態では、遺伝子組換えは、トリテルペノイド系サポニン生合成経路において、活性な酵素をコードする遺伝子を含む異種ポリヌクレオチドで、植物細胞または酵母細胞を形質転換することを含む。本明細書で使用するとき、「遺伝子組換え細胞」及び「対応する未組換え細胞」は、それぞれ、遺伝子組換え細胞及び遺伝子組換えを欠く同一の種類 of 細胞を指す。

30

【0211】

いくつかの実施形態では、「遺伝子組換え細胞」及び「対応する非組換え細胞」は、それぞれ、同一の種類 of 遺伝子組換え植物細胞及び対応する未組換え植物細胞である。他の実施形態では、「遺伝子組換え細胞」及び「対応する未組換え細胞」は、それぞれ、同一の種類 of 遺伝子組換え酵母細胞及び対応する未組換えの酵母細胞である。

【0212】

「遺伝子組換え植物」という用語は、人間によって遺伝子組換えされた少なくとも1つの細胞を含む植物を指す。遺伝子組換えは、例えば、内在性ポリヌクレオチドまたは目的の遺伝子に対して突然変異、欠失、挿入、転移性因子などを導入することによる、内在性遺伝子の組み換えを含む。追加的または代替的に、遺伝子組換えは、異種ポリヌクレオチドで植物細胞を形質転換することを含む。本明細書で使用するとき、「遺伝子組換え植物」及び「対応する未組換え植物」は、それぞれ、少なくとも1つの遺伝子組換え細胞を含む植物、及び組換えを欠く同一の種類 of 植物を指す。

40

【0213】

いくつかの実施形態では、遺伝子組換え細胞の懸濁液及び遺伝子組換え細胞に由来する組織培養物も包含される。細胞懸濁液または組織培養物は、所望のステロイド系アルカロイド、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせ、ステロイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせ、あるいは、トリテルペノイド系サポニン、その誘導体、その代謝物、そ

50

の生合成中間体、またはそれらの任意の組み合わせの産生のために使用することができる。いくつかの実施形態では、遺伝子組換え細胞または組織培養物は、CSLG酵素の発現が改変された少なくとも1つの細胞を有する遺伝子組換え植物を再生するために使用される。

【0214】

当業者であれば、いくつかの実施形態では、遺伝子組換え植物の種子をさらに包含し、種子から成長した植物は、対応する未組換えの種子またはそれから成長した植物と比較して、少なくとも1つの細胞における少なくとも1つのCSLGの発現が改変されており、それによって、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が改変された少なくとも1つの細胞を有することを理解できるであろう。

10

【0215】

形質転換技術が組織培養の使用を必要とする場合、形質転換された細胞は、当業者によく知られている技術を用いて植物に再生される。再生された植物は、次いで、従来の育種技術を用いて、栽培し、同一のまたは異なる植物品種と交配させて種子を作り、その後、適切な条件下で選択される。

【0216】

当業者であれば、遺伝子組換え植物は、人間によって遺伝子組換えされた少なくとも1つの細胞を含む植物を包含することを理解できるであろう。いくつかの実施形態では、遺伝子組換えは、本明細書に記載された異種遺伝子を含む異種ポリヌクレオチドで、少なくとも1つの植物細胞、酵母細胞、藻類細胞、昆虫細胞、または細菌を形質転換することを含む。いくつかの実施形態では、遺伝子組換えは、それぞれが酵素をコードする2以上の異種遺伝子を含む異種ポリヌクレオチドで、少なくとも1つの植物細胞、酵母細胞、藻類細胞、昆虫細胞、または細菌を形質転換することを含む。いくつかの実施形態では、遺伝子組換えは、酵素をコードする複数の異種遺伝子をコードする複数の異種ポリヌクレオチドで、少なくとも1つの植物細胞、酵母細胞、藻類細胞、昆虫細胞、または細菌を形質転換することを含む。

20

【0217】

当業者であれば、少なくとも1つの植物細胞を異種ポリヌクレオチドまたは複数の異種ポリヌクレオチドで形質転換することを含む遺伝子組換え植物が、いくつかの実施形態では、「トランスジェニック植物」と呼ばれることを理解できるであろう。

30

【0218】

当業者であれば、「遺伝子組換え植物」と「対応する非組換え植物」との比較は、少なくとも1つの遺伝子組換え細胞を含む植物と、組み換えを欠いた同一の種類の植物との比較を包含することを理解できるであろう。

【0219】

当業者であれば、本明細書で開示される植物に関連して使用するとき、「トランスジェニック」という用語は、その細胞の1以上に少なくとも1つの異種の転写可能なポリヌクレオチドを含む植物を包含することを理解できるであろう。当業者であれば、本明細書で開示される酵母に関連して使用するとき、「トランスジェニック」という用語は、少なくとも1つの異種の転写可能なポリヌクレオチドを含む酵母を包含することを理解できるであろう。「トランスジェニック材料」という用語は、少なくとも1つの細胞に少なくとも1つの異種ポリヌクレオチドを含む少なくとも1つの細胞、複数の細胞または組織を含む植物またはその一部を広く包含する。したがって、「トランスジェニック植物」と「対応する非トランスジェニック植物」との比較、または「異種の転写可能なポリヌクレオチドを含む少なくとも1つの細胞を含む発現が改変された少なくとも1つの細胞を含む遺伝子組換え植物」と「対応する非組み換え植物」との比較は、「トランスジェニック植物」または「遺伝子組換え植物」と、異種の転写可能なポリヌクレオチドを欠いた同一の種類の植物との比較を包含する。当業者であれば、いくつかの実施形態では、「転写可能なポリヌクレオチド」が、RNAポリメラーゼによってRNA分子に転写されるポリヌクレオチ

40

50

ドを含むことを理解できるであろう。

【0220】

「形質転換体」または「形質転換細胞」という用語は、一次形質転換細胞及びその細胞に由来する培養物を移植回数に関係なく含む。すべての子孫は、意図的または不注意な突然変異に起因して、DNAの含有量が正確に同一ではない場合がある。最初に形質転換された細胞でスクリーニングされたのと同じの機能を有する突然変異体の子孫は、形質転換体の定義に含まれる。

【0221】

細胞の形質転換は、安定的であってもよいし、一過性であってもよい。「一過性の形質転換」または「一過性に形質転換された」という用語は、宿主細胞のゲノムへの外因性ポリヌクレオチドの組み込みを欠くときに、1以上の外因性ポリヌクレオチドを細胞に導入することを指す。対照的に、「安定的な形質転換」または「安定的に形質転換された」という用語は、1以上の外因性ポリヌクレオチドを細胞のゲノムに導入し、組み込むことを指す。「安定形質転換体」という用語は、1以上の外因性ポリヌクレオチドをゲノムまたはオルガネラDNAに安定的に組み込んだ細胞を指す。本明細書に開示された核酸、構築物、及び/またはベクターで形質転換された生物またはその細胞は、安定的な形質転換だけでなく、一過性の形質転換も可能であることを理解されたい。

10

【0222】

当業者であれば、「構築物」という用語は、目的のポリヌクレオチドを含む、人工的に組み立てられた、または単離された核酸分子を包含することを理解できるであろう。一般に、構築物は、目的のポリヌクレオチドまたはポリヌクレオチド、場合によっては目的の遺伝子でもあるマーカー遺伝子、及び適切な調節配列を含むことができる。構築物中に調節配列を含めることは任意選択であり、例えば、そのような配列は、宿主細胞の調節配列が使用される状況では必要とされないことを理解されたい。構築物という用語は、ベクターを含むが、これに限定されるものと見なすべきではない。

20

【0223】

追加的または代替的に、いくつかの実施形態では、遺伝子組換えは、少なくとも1つの植物細胞を異種ポリヌクレオチドまたは複数の異種ポリヌクレオチドで形質転換することを含む。当業者であれば、少なくとも1つの植物細胞を異種ポリヌクレオチドまたは複数の異種ポリヌクレオチドで形質転換することを含む遺伝子組換え植物が、いくつかの実施形態では、「トランスジェニック植物」と呼ばれることを理解できるであろう。

30

【0224】

ウイルスベクターは、形質転換抵抗性がより高い植物（例えば、大豆またはインゲンマメ）の形質転換に有用である。いくつかの実施形態では、ビーンポッドモットルウイルス（BPMV；コモウイルス属）ベクターなどのウイルスベクターは、外来遺伝子の発現及びウイルス誘導遺伝子サイレンシング（VIGS）に使用される（Zhang et al. (May 2010) Plant Physiol. 153: 52-65 ["Zhang 2010"]）。細胞の形質転換は、例えば、バイオオリステック法、または直接的なDNAラビング接種により行われる（Zhang 2010）。

【0225】

一実施形態では、植物の形質転換は、遺伝子銃または生物学的粒子送達システム（バイオオリステック法）を使用して外因性DNA（導入遺伝子）を細胞に送達することにより行われる（Rech et al. (2008) Nature Protocols 3(3): 410-418 ["Rech 2008"]）。いくつかの実施形態では、プラスミドを設計し、植物（例えば、大豆、豆、綿）の頂端分裂組織に、微粒子でコーティングしたDNAを衝突させ、その後、インビトロで培養し、トランスジェニック植物を選択する（Rech 2008）。他の実施形態では、未分化植物細胞のカルス、または未熟胚の群を、インビトロで、ゲル培地で培養する。いくつかの実施形態では、この細胞を、その後、オーキシンやジベレリンなどの一連の植物ホルモンで処理することにより植物が得られる。ステロイド系アルカロイド、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせ、ステロイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせ、あるいは、トリテル

40

50

ペノイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせは、上記のようなトランスジェニック植物または植物細胞を用いて産生することができる。

【0226】

タンパク質の「一過性発現」は、当技術分野で知られている様々な手段によって達成することができる。一実施形態では、タンパク質の一過性発現は、遺伝子組換えウイルスの使用によって達成される。いくつかの実施形態では、アグロインフィルトレーションを使用して、植物、分離した葉、または植物の別の部分における遺伝子の一過性発現を誘導する。アグロバクテリウム（例えば、アグロバクテリウム・ツメファシエンス）の懸濁液を、例えば、直接注入または真空濾過によって植物に導入するか、または、多孔質支持体上に固定化された植物細胞（植物細胞パック）と接触させる。バクテリアは、Tiプラスミド由来のT-DNAの導入を介して、植物細胞に目的の遺伝子を導入する。

10

【0227】

いくつかの実施形態では、藻類の形質転換の方法は、上述した方法のいずれかを含む。一実施形態では、藻類の形質転換は、ガラスビーズ支援型形質転換、粒子銃媒介型（バイオリスティック）形質転換、細胞壁を弱めるためのセルロース分解酵素による処理、または相同組換えを用いて達成される。

【0228】

別の態様では、本明細書に開示された核酸は、藻類に形質転換することができる。一実施形態では、藻類は、単細胞藻類である。別の実施形態では、藻類は、多細胞藻類である。一実施形態では、藻類は、シアノバクテリウム、珪藻、クラミドモナス、ドナリエラ、またはヘマトコッカスである。遺伝子は、藻類で過剰発現させることができる。ステロイド系アルカロイド、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせ、ステロイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせ、あるいは、トリテルペノイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせを、このようなトランスジェニック藻類を用いて産生することができる。藻類の形質転換のための方法は、当技術分野でよく知られており、特許文献8～特許文献14に完全に記載されている（これらの特許文献は、参照によりその全体が本明細書に組み込まれる）。さらに、微細藻類の培養は当技術分野でよく知られており、「Vuppaladadiyam et al., (2018) "Microalgae cultivation and metabolites production: a comprehensive review." Biofuels, Bioprod. Bioref. 12:304-324」に記載されている（この文献は、参照によりその全体が本明細書に組み込まれる）。

20

30

【0229】

別の態様では、本明細書に開示された核酸は、酵母に形質転換することができる。遺伝子は、酵母で過剰発現させることができる。酵母の形質転換のための方法は、当技術分野でよく知られており、特許文献15～特許文献21に完全に記載されている（これらの特許文献は、参照によりその全体が本明細書に組み込まれる）。ステロイド系アルカロイド、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせ、ステロイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせ、あるいは、トリテルペノイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせは、このようなトランスジェニック酵母を用いて産生することができる。

40

【0230】

別の実施形態では、本明細書に開示された核酸は、ウイルスに形質転換することができる。別の実施形態では、核酸は、ウイルスで過剰発現させることができる。ステロイド系アルカロイド、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせ、ステロイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせ、あるいは、トリテルペノイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせは、このようなウイルスの過

50

剰発現によって製造することができる。

【0231】

いくつかの実施形態では、細菌系では、発現させるポリペプチドの意図する用途に応じて、多数の発現ベクターを有利に選択することができる。一実施形態では、大量のポリペプチドが望まれる。一実施形態では、タンパク質産物の高レベルの発現を誘導するベクター、場合によっては疎水性シグナル配列との融合体として、発現した産物を、タンパク質産物が容易に精製される細菌のペリプラズムまたは培養液に誘導するベクターが望まれる。一実施形態では、特定の融合タンパク質は、ポリペプチドの回収を助けるために、特定の切断部位で操作される。一実施形態では、そのような操作に適応可能なベクターとしては、これに限定しないが、pETシリーズの大腸菌発現ベクター (Studier et al., *Methods in Enzymol.* 185:60-89 1990) が挙げられる。

10

【0232】

いくつかの実施形態では、ポリヌクレオチド配列は、少なくとも1つの異種遺伝子を含む。いくつかの実施形態では、ポリヌクレオチド配列は、酵素をコードする。いくつかの実施形態では、ポリヌクレオチド配列は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの生合成経路の酵素をコードする。いくつかの実施形態では、ポリヌクレオチド配列は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニン生合成経路の基質の産生に必要な酵素をコードする。

20

【0233】

いくつかの実施形態では、ポリヌクレオチド配列は、少なくとも2、3、4、5、6、7、8、9、10、またはそれ以上の個数の異種遺伝子を含む。いくつかの実施形態では、ポリヌクレオチド配列は、2、3、4、5、6、7、8、9、10、またはそれ以上の個数の酵素をコードし、各遺伝子は、単一の酵素をコードする。いくつかの実施形態では、ポリヌクレオチド配列は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニン生合成経路の2、3、4、5、6、7、8、9、10、またはそれ以上の個数の酵素をコードする。いくつかの実施形態では、ポリヌクレオチド配列は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニン経路の2、3、4、5、6、7、8、9、10、またはそれ以上の個数の酵素と、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニン生合成経路の基質の産生に必要な追加の酵素とをコードする。いくつかの実施形態では、ポリヌクレオチド配列は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニン経路の2、3、4、5、6、7、8、9、10、またはそれ以上の個数の酵素と、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニン生合成経路の基質の産生に必要な追加の複数の酵素とをコードする。

30

【0234】

いくつかの実施形態では、「ヌクレオチド配列」または「ポリヌクレオチド配列」という用語は、単一の連続したポリヌクレオチド配列を包含し、他の実施形態では、この用語は、複数の連続したヌクレオチド配列を包含する。

40

【0235】

本明細書に開示された共発現遺伝子アレイに基づいて、コレステロールから、例えば - トマチンへの経路が提案される (図1)。コレステロールは、GAME7によってC22位が水酸化された後、GAME8によってC26位が水酸化されることが、以前に記載されている (特許文献5)。22, 26-ジヒドロキシコレステロールは、次いで、C16で水酸化され、C22で酸化され、その後、GAME11及びGAME6によってEリングが閉じられ、フロスタノール型アグリコンが形成される。この反応順序は、GAME11をサイレンシングしたときに、C16及びヘミアセタールのEリングでの水酸化を欠くコレスタノール型サポニンが蓄積されるという本明細書中の知見からも裏付けられる (図8A - 図8D)。フロスタノール中間体は、GAME4によって酸化され、GAME12によって触媒されるアミノ基転移の基質である26-アルデヒドとなる。C22のアミ

50

ノ窒素が求核攻撃を受けるとトマチデノールが産生し、これが脱水素されてトマチジンとなる。トマチジンは、その後、GAME 1 によってT - Gal に変換される (Itkin et al., 2011 supra)。T - Gal はGAME 1 7 によってグリコシル化されて 1 - トマチンとなり、さらにGAME 1 8 によってグリコシル化されて 1 - トマチンとなり、最終的にGAME 2 によって 1 - トマチンに変換される (図 1)。

#### 【0236】

本明細書に記載されているように、生合成経路に関与する酵素及び/または他のタンパク質の発現を改変することによって、ステロイド系アルカロイド、ステロイド系グリコアルカロイド、ステロイド系サポニンのレベルを改変することができる。

#### 【0237】

ジャガイモ植物において、クラスタ化した酵素をコードする遺伝子と共発現した単一の遺伝子をサイレンシングすると、ステロイド系グリコアルカロイドである 1 - カコニンと 1 - ソラニンの含有量が有意に減少し、逆にこの遺伝子を過剰発現させると、これらの物質の含有量が有意に増加した (図 5 及び図 6)。この遺伝子は、AP 2 ドメインを含むコード配列を含むことが見出され、したがって、GAME 9 によってコードされた転写因子 (本明細書ではGAME 9 - 転写因子と称する) であると仮定された。

#### 【0238】

本発明によれば、GAME 9 転写因子をコードする遺伝子、2 - オキシグルタル酸依存性ジオキシゲナーゼをコードする遺伝子、ベーシック・ヘリックス・ループ・ヘリックス (BHLH) 転写因子をコードする遺伝子、またはそれらの任意の組み合わせからなる群から選択された少なくとも1つの遺伝子の発現が改変された少なくとも1つの細胞を含む遺伝子組換え植物であって、対応する未組換え植物と比較して、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量が改変された遺伝子組換え植物が提供される。本明細書中で、2 - オキシグルタル酸依存性ジオキシゲナーゼ (GAME 1 1) について例示したように、本発明の遺伝子の発現を操作することは、ステロイド系サポニン合成の操作をさらにもたらすことができる。

#### 【0239】

したがって、さらなる態様によれば、GAME 9 転写因子をコードする遺伝子、2 - オキシグルタル酸依存性ジオキシゲナーゼをコードする遺伝子、ベーシック・ヘリックス・ループ・ヘリックス (BHLH) 転写因子をコードする遺伝子、またはそれらの任意の組み合わせからなる群から選択された少なくとも1つの遺伝子の発現が改変された少なくとも1つの遺伝子の発現が改変された少なくとも1つの細胞を含む遺伝子組換え生物であって、未組換えまたは未編集の生物と比較して、ステロイド系サポニン及び/またはステロイド系アルカロイド、それらのグリコシル化誘導体または他の誘導体、それらの代謝物、またはそれらの生合成中間体から選択される少なくとも1つの化合物の含有量が改変された遺伝子組換え生物が提供される。

#### 【0240】

予期せぬことに、本発明は、ステロイド系アルカロイド生合成経路に関与する酵素及び/または転写因子の発現を調節することによって、ジャガイモ塊茎中のSGAレベルを大幅に低下させることができることが分かった。

#### 【0241】

いくつかの実施形態では、遺伝子組換え植物中の、GAME 9 転写因子をコードする遺伝子、2 - オキシグルタル酸依存性ジオキシゲナーゼをコードする遺伝子、BHLH 転写因子をコードする遺伝子、またはそれらの組み合わせからなる群から選択された少なくとも1つの遺伝子の発現が、対応する未組換えまたは未編集の植物における発現と比べて抑制され、これにより、遺伝子組換え植物は、対応する未組換え植物と比べて、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量が減少する。

#### 【0242】

化合物 1 1 のハウレンソウのトリテルペノイド生合成経路において活性な遺伝子及びそ

10

20

30

40

50

れをコードする酵素は、実施例において詳細に説明される。実施例 20 で提供される表 17 は、これらの酵素のアミノ酸配列及び核酸配列と、それらをコードする遺伝子を示す。実施例 24 で提供される表 19 は、シロイヌナズナから単離された追加の CSLG 酵素のアミノ酸配列を示す。

【0243】

いくつかの実施形態では、少なくとも 1 つの異種遺伝子のポリヌクレオチド配列は、配列番号 30、32、34、36、38、40、45、46、47、51、53、55、57、59、61、63、65、80、93、95、97、99、101、103、または 105 のいずれか 1 つに記載されている。いくつかの実施形態では、少なくとも 1 つの異種遺伝子のポリヌクレオチド配列は、配列番号 30、32、34、36、38、40、45、46、47、51、53、55、57、59、61、63、65、80、93、95、97、99、101、103、または 105 のいずれか 1 つに記載の核酸配列に対して少なくとも 55% の同一性を有する核酸配列を含む。

10

【0244】

いくつかの実施形態では、少なくとも 1 つの異種遺伝子のポリヌクレオチド配列は、配列番号 30、32、34、36、38、40、45、46、47、51、53、55、57、59、61、63、65、80、93、95、97、99、101、103、または 105 のいずれか 1 つに記載されている。いくつかの実施形態では、少なくとも 1 つの異種遺伝子のポリヌクレオチド配列は、配列番号 30、32、34、36、38、40、45、46、47、51、53、55、57、59、61、63、65、80、93、95、97、99、101、103、または 105 のいずれか 1 つに記載の核酸配列に対して少なくとも 80% の同一性を有する核酸配列を含む。

20

【0245】

いくつかの実施形態では、少なくとも 1 つの異種 CSLG 酵素をコードする核酸配列は、配列番号 30、32、34、36、38、40、65、80、93、95、97、99、101、103、または 105 のいずれか 1 つに記載されている。いくつかの実施形態では、少なくとも 1 つの異種 CSLG 酵素をコードする核酸配列は、配列番号 30、32、34、36、38、40、65、80、93、95、97、99、101、103、または 105 のいずれか 1 つに記載の核酸配列と少なくとも 55% の同一性及び少なくとも 55% のカバレッジを有するそのホモログである。いくつかの実施形態では、少なくとも 1 つの異種 CSLG 酵素をコードする核酸配列は、配列番号 30、32、34、36、38、40、65、80、93、95、97、99、101、103、または 105 のいずれか 1 つに記載の核酸配列と少なくとも 55% の同一性及び少なくとも 80% のカバレッジを有するそのホモログである。

30

【0246】

いくつかの実施形態では、少なくとも 1 つの異種 CSLG 酵素をコードする核酸配列は、配列番号 30、32、34、36、38、40、65、80、93、95、97、99、101、103、または 105 のいずれか 1 つに記載されている。いくつかの実施形態では、少なくとも 1 つの異種 CSLG 酵素をコードする核酸配列は、配列番号 30、32、34、36、38、40、65、80、93、95、97、99、101、103、または 105 のいずれか 1 つに記載の核酸配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログである。

40

【0247】

いくつかの実施形態では、少なくとも 1 つの異種遺伝子のポリヌクレオチド配列は、配列番号 30 に記載されている。いくつかの実施形態では、少なくとも 1 つの異種遺伝子のポリヌクレオチド配列は、配列番号 32 に記載されている。いくつかの実施形態では、少なくとも 1 つの異種遺伝子のポリヌクレオチド配列は、配列番号 34 に記載されている。

50







2、3、4、5、6、7、8、9、10、または11以上の少なくとも1つのステロイド系アルカロイド生合成酵素、ステロイド系サポニン生合成酵素、トリテルペノイド系サポニン生合成酵素、またはそれらの組み合わせをコードするポリヌクレオチド配列を含む。

【0252】

当業者であれば、配列の相同性は、共通の祖先からの系図に起因する配列の類似性を包含することを理解できるであろう。相同の生物学的構成要素（遺伝子、タンパク質、構造）はホモログと呼ばれる。ヌクレオチドまたはタンパク質の配列が関連している程度。2つの配列（DNA、RNA、またはアミノ酸）の間の類似性は、配列相同性の割合及び/または正しい置換の割合で表すことができる。

【0253】

当業者であれば、「ホモログ」という用語は、共通の祖先のDNAまたはポリペプチド（タンパク質）配列からの系図によって、それぞれ、第2の遺伝子またはポリペプチド（タンパク質）と関連している遺伝子またはポリペプチド（タンパク質）を包含することを理解できるであろう。したがって、いくつかの実施形態では、遺伝子のホモログは、その遺伝子と同様のヌクレオチド配列を含む。いくつかの実施形態では、遺伝子のホモログは、その遺伝子によってコードされるものと同じのポリペプチドを含む。いくつかの実施形態では、遺伝子のホモログは、その遺伝子によってコードされるものと同じの機能的特性を有するポリペプチドをコードする。いくつかの実施形態では、遺伝子のホモログは、その遺伝子によってコードされるポリペプチドと同様のアミノ酸配列を含むポリペプチドをコードする。一実施形態では、ポリペプチドのホモログは、そのポリペプチドと同様のアミノ酸配列を含む。いくつかの実施形態では、ポリペプチドのホモログは、そのポリペプチドと同一の機能的特性を有する。いくつかの実施形態では、ポリペプチドのホモログは、そのポリペプチドと同一のドメインを含む。いくつかの実施形態では、ポリペプチドのホモログは、そのポリペプチドと同様のドメインを含む。

【0254】

当業者であれば、パーセント相同性またはパーセント同一性は、これに限定しないが、例えばNCBI（アメリカ国立生物工学情報センター）のBLASTソフトウェアを使用して、デフォルトのパラメータを用いて決定できることを理解できるであろう。また、ホモログは、オーソログ、欠失、挿入、またはアミノ酸置換を含む置換変異体を指してもよい。いくつかの実施形態では、配列同一性または相同性は、BLAST、CLUSTALW、MUSCLE、及びHHpredなどの任意のタンパク質または核酸配列アライメントアルゴリズムを用いて決定することができる。

【0255】

当業者であれば、ホモログは、いくつかの実施形態では、同一の活性を有し得ることを理解できるであろう。いくつかの実施形態では、SOAP遺伝子のホモログは、対応するSOAP遺伝子と同一の活性を有する。いくつかの実施形態では、SOAPポリペプチドのホモログは、対応するSOAPポリペプチドと同一の酵素活性を有する。いくつかの実施形態では、SOAP5のホモログは、SOAP5と同一の酵素活性を有する。いくつかの実施形態では、GAME遺伝子のホモログは、対応するGAME遺伝子と同一の活性を有する。いくつかの実施形態では、GAMEポリペプチドのホモログは、対応するGAMEポリペプチドと同一の酵素活性を有する。いくつかの実施形態では、GAME15のホモログは、GAME15と同一の酵素活性を有する。いくつかの実施形態では、CSLG遺伝子のホモログは、セルロース合成酵素様G活性を有する酵素をコードする。いくつかの実施形態では、CSLG遺伝子のホモログは、グルクロン酸転移酵素活性を有する酵素をコードする。いくつかの実施形態では、CSLG遺伝子のホモログは、セルロース合成酵素様G活性及びグルクロン酸転移酵素活性を有する酵素をコードする。図42A及び42Fは、セルロース合成酵素様酵素の系統樹を示し、セルロース合成酵素様Gの下位の系統群は、本明細書に記載のSOAP5及びGAME15酵素を含む。

【0256】

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、少なくとも1つの異種CSLG遺伝子をコードする核酸配列を含み、その核酸配列は、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、あるいは、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも55%の同一性及び少なくとも55%のカバレッジを有するそのホモログである。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、少なくとも1つの異種CSLG遺伝子をコードする核酸配列を含み、その核酸配列は、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、あるいは、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも55%の同一性及び少なくとも80%のカバレッジを有するそのホモログである。

10

**【0257】**

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、少なくとも1つの異種CSLG遺伝子をコードする核酸配列を含み、その核酸配列は、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、あるいは、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログである。

20

**【0258】**

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、異種CSLG遺伝子に加えて、少なくとも1つの追加の異種遺伝子をコードする核酸配列を含み、追加の遺伝子は、(a) - アミリン合成酵素、及び、配列番号45に記載の核酸配列、または、配列番号45に記載の核酸配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログ、(b) シトクロムP450、及び、配列番号46、51、または53のいずれか1つに記載の核酸配列、または、配列番号46、51、または53のいずれか1つに記載の核酸配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログ、(c) グリコシルトランスフェラーゼ、及び、配列番号55、57、59、または61のいずれか1つに記載の核酸配列、または、配列番号55、57、59、または61のいずれか1つに記載の核酸配列に対して少なくとも80%の同一性及び80%のカバレッジを有するそのホモログ、(d) アセチルトランスフェラーゼ、及び、配列番号63に記載の核酸配列、または、配列番号63に記載の核酸配列に対して80%の同一性及び少なくとも80%のカバレッジを有するそのホモログ、(e) UDP-グルコース6-デヒドロゲナーゼ1、及び、配列番号74に記載の核酸配列、または、配列番号74に記載の核酸配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログ、あるいは、(f) 上記の(a)、(b)、(c)、(d)及び(e)の任意の組み合わせ、をコードする。

30

40

**【0259】**

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、少なくとも1つの異種CSLG遺伝子をコードする核酸配列を含み、その核酸配列は、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、あるいは、配列番号30、32、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも55%の同一性及び少なくとも55%のカバレッジを有するそのホモログである。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、少なくとも1つの異種CSLG遺伝子をコードする核酸配列を含み、その核酸配列は、配列番号30、32、34、36、38、40、65、8

50

0、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、あるいは、配列番号30、32、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも55%の相同性及び少なくとも80%のカバレッジを有するそのホモログである。

【0260】

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、少なくとも1つの異種CSLG遺伝子をコードする核酸配列を含み、その核酸配列は、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、あるいは、配列番号30、32、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログである。

10

【0261】

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、異種CSLG遺伝子に加えて、少なくとも1つの追加の異種遺伝子をコードする核酸配列を含み、追加の遺伝子は、(a) - アミリン合成酵素、及び、配列番号45に記載の核酸配列、または、配列番号45に記載の核酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログ、(b) シトクロムP450、及び、配列番号46、51、または53のいずれか1つに記載の核酸配列、または、配列番号46、51、または53のいずれか1つに記載の核酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログ、(c) グリコシルトランスフェラーゼ、及び、配列番号55、57、59、または61のいずれか1つに記載の核酸配列、または、配列番号55、57、59、または61のいずれか1つに記載の核酸配列に対して少なくとも80%の相同性及び80%のカバレッジを有するそのホモログ、(d) アセチルトランスフェラーゼ、及び、配列番号63に記載の核酸配列、または、配列番号63に記載の核酸配列に対して80%の相同性及び少なくとも80%のカバレッジを有するそのホモログ、(e) UDP - グルコース6 - デヒドロゲナーゼ1、及び、配列番号74に記載の核酸配列、または、配列番号74に記載の核酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログ、あるいは、(f) 上記の(a)、(b)、(c)、(d)及び(e)の任意の組み合わせ、をコードする。

20

30

【0262】

いくつかの実施形態では、本明細書に開示される酵素をコードする少なくとも1つの異種遺伝子を含むポリヌクレオチド配列は、配列番号30、32、34、36、38、40、45、46、47、51、53、55、57、57、59、61、63、65、80、93、95、97、99、101、103、105、またはそれらの組み合わせから選択される配列に対して少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、または少なくとも95%の相同性を有する核酸配列を含む。いくつかの実施形態では、本明細書に開示される酵素をコードする異種遺伝子の核酸配列は、配列番号30、32、34、36、38、40、45、46、47、51、53、55、57、59、61、63、65、80、93、95、97、99、101、103、105、またはそれらの組み合わせから選択される配列に対して少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、または少なくとも95%の相同性を有する核酸配列を含む。

40

【0263】

いくつかの実施形態では、本明細書に開示される酵素をコードする異種遺伝子の核酸配列は、配列番号30、32、34、36、38、40、45、46、47、51、53、55、57、59、61、63、65、80、93、95、97、99、101、103、105、またはそれらの組み合わせから選択されたる核酸配列を含む。

50

## 【0264】

いくつかの実施形態では、配列番号30、32、34、36、38、40、45、46、47、51、53、55、57、59、61、63、65、80、93、95、97、99、101、103、または105から選択される配列に対して少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、または少なくとも95%の相同性を有する核酸配列は、コドン最適化遺伝子配列を含む。

## 【0265】

いくつかの実施形態では、遺伝子を最適化することは、コドン使用量を宿主遺伝子のコドンバイアスに適合させることを含み、一実施形態では、宿主は植物である。いくつかの実施形態では、遺伝子を最適化することは、コドン使用量を宿主遺伝子のコドンバイアスに適合させることを含み、一実施形態では、宿主は酵母である。いくつかの実施形態では、遺伝子を最適化することは、コドン使用量を宿主遺伝子のコドンバイアスに適合させることを含み、一実施形態では、宿主は藻類である。コドン最適化遺伝子配列は、いくつかの実施形態では、本明細書に記載されているように、少なくとも1つの異種遺伝子の翻訳効率を向上させることによって、生体におけるタンパク質発現レベルを向上させる。いくつかの実施形態では、コドン最適化は、酵母における発現のためのコドン最適化を含む。いくつかの実施形態では、コドン最適化は、藻類における発現のためのコドン最適化を含む。いくつかの実施形態では、コドン最適化は、植物細胞における発現のためのコドン最適化を含む。

10

20

## 【0266】

いくつかの実施形態では、「最適化された」という用語は、所望の改変を包含し、一実施形態では遺伝子発現の改変であり、別の実施形態ではタンパク質発現の改変である。一実施形態では、最適化された遺伝子発現は、遺伝子発現の最適化された制御である。別の実施形態では、最適化された遺伝子発現は、遺伝子発現の増加である。この態様及び一実施形態では、野生型と比較して、遺伝子発現の2倍から1000倍の増加が企図される。別の実施形態では、遺伝子発現の2倍から500倍の増加、別の実施形態では、遺伝子発現の2倍から100倍の増加、別の実施形態では、遺伝子発現の2倍から50倍の増加、別の実施形態では、遺伝子発現の2倍から20倍の増加、別の実施形態では、遺伝子発現の2倍から10倍の増加、別の実施形態では、遺伝子発現の3倍から5倍の増加が企図される。

30

## 【0267】

別の実施形態では、最適化された遺伝子発現は、特定の環境条件下での遺伝子発現の増加であってもよい。別の実施形態では、最適化された遺伝子発現は、遺伝子発現の低下を含んでいてもよく、これは、一実施形態では特定の環境条件下でのみであってもよい。

## 【0268】

別の実施形態では、最適化された遺伝子発現は、遺伝子発現の持続時間の増加である。この態様及び一実施形態では、野生型と比較して、遺伝子発現の持続時間が2倍から1000倍に増加することが企図される。別の実施形態では、遺伝子発現の持続時間の2倍から500倍の増加、別の実施形態では、遺伝子発現の持続時間の2倍から100倍の増加、別の実施形態では、遺伝子発現の持続時間の2倍から50倍の増加、別の実施形態では、遺伝子発現の持続時間の2倍から20倍の増加、別の実施形態では、遺伝子発現の持続時間の2倍から10倍の増加、別の実施形態では、遺伝子発現の持続時間の3倍から5倍の増加が企図される。別の実施形態では、遺伝子発現の持続時間の増加は、非ベクター発現対照における遺伝子発現と比較されるか、または代替的に、野生型ベクター発現対照における遺伝子発現と比較される。

40

## 【0269】

当業者であれば、いくつかの実施形態では、酵素をコードする2以上の異種遺伝子が遺伝子組換え細胞で発現することを理解できるであろう。いくつかの実施形態では、細胞は、2以上の異種遺伝子をコードするポリヌクレオチド配列を含む。他の実施形態では、酵

50

素をコードする2つ以上の異種遺伝子が遺伝子組換え細胞で発現される場合、複数のポリヌクレオチド配列を使用することにより複数の異種遺伝子をコードしてもよく、ポリヌクレオチド配列は1以上の遺伝子を含んでいてもよい。いくつかの実施形態では、酵素をコードする2以上の異種遺伝子が遺伝子組換え細胞で発現される場合、複数のポリヌクレオチド配列を使用することにより複数の異種遺伝子をコードしてもよく、各ポリヌクレオチド配列は1つの単一の遺伝子を含む。いくつかの実施形態では、酵素をコードする2以上の異種遺伝子が遺伝子組換え細胞で発現される場合、複数のポリヌクレオチド配列を使用することにより複数の異種遺伝子をコードしてもよく、各ポリヌクレオチド配列は1以上の遺伝子を含む。

【0270】

全体を通して開示されているように、いくつかの実施形態では、遺伝子組換え細胞は、植物細胞を含む。いくつかの実施形態では、遺伝子組換え植物細胞は、植物またはその一部を含む。他の実施形態では、遺伝子組換え細胞は、酵母細胞を含む。他の実施形態では、遺伝子組換え細胞は、藻類細胞を含む。他の実施形態では、遺伝子組換え細胞は、昆虫細胞を含む。他の実施形態では、遺伝子組換え細胞は、細菌を含む。

【0271】

さらに、当業者であれば、いくつかの実施形態では、酵素をコードする2以上の異種遺伝子が、遺伝子組換え植物で発現することを理解できるであろう。いくつかの実施形態では、植物は、2以上の異種遺伝子をコードするポリヌクレオチド配列を含む。他の実施形態では、酵素をコードする2以上の異種遺伝子が遺伝子組換え植物で発現される場合、複数のポリヌクレオチド配列を使用することにより複数の異種遺伝子をコードしてもよく、ポリヌクレオチド配列は1以上の遺伝子を含んでいてもよい。いくつかの実施形態では、酵素をコードする2以上の異種遺伝子が遺伝子組換え植物で発現される場合、複数のポリヌクレオチド配列を使用することにより複数の異種遺伝子をコードしてもよく、各ポリヌクレオチド配列は1つの単一の遺伝子を含む。いくつかの実施形態では、酵素をコードする2以上の異種遺伝子が遺伝子組換え植物で発現される場合、複数のポリヌクレオチド配列を使用することにより複数の異種遺伝子をコードしてもよく、各ポリヌクレオチド配列は1以上の遺伝子を含む。

【0272】

いくつかの実施形態では、少なくとも1つの異種遺伝子によってコードされる酵素のアミノ酸配列は、ステロイド系アルカロイド生合成酵素、ステロイド系サポニン生合成酵素、及び/またはトリテルペノイド生合成酵素のアミノ酸配列を含む。

【0273】

いくつかの実施形態では、少なくとも1つの異種遺伝子によってコードされる酵素のアミノ酸配列は、配列番号31、33、35、37、39、41、48、49、50、52、54、56、58、60、62、64、81、94、96、98、100、102または104のいずれか1つに記載されている。いくつかの実施形態では、少なくとも1つの異種遺伝子によってコードされる酵素のアミノ酸配列は、配列番号31に記載されている。いくつかの実施形態では、少なくとも1つの異種遺伝子によってコードされる酵素のアミノ酸配列は、配列番号33に記載されている。いくつかの実施形態では、少なくとも1つの異種遺伝子によってコードされる酵素のアミノ酸配列は、配列番号35に記載されている。いくつかの実施形態では、少なくとも1つの異種遺伝子によってコードされる酵素のアミノ酸配列は、配列番号37に記載されている。いくつかの実施形態では、少なくとも1つの異種遺伝子によってコードされる酵素のアミノ酸配列は、配列番号39に記載されている。いくつかの実施形態では、少なくとも1つの異種遺伝子によってコードされる酵素のアミノ酸配列は、配列番号41に記載されている。いくつかの実施形態では、少なくとも1つの異種遺伝子によってコードされる酵素のアミノ酸配列は、配列番号48に記載されている。いくつかの実施形態では、少なくとも1つの異種遺伝子によってコードされる酵素のアミノ酸配列は、配列番号49に記載されている。いくつかの実施形態では、少なくとも1つの異種遺伝子によってコードされる酵素のアミノ酸配列は、配列番号50

10

20

30

40

50







の実施形態では、少なくとも1つの異種遺伝子によってコードされる酵素のアミノ酸配列は、配列番号104に記載されているアミノ酸配列に対して少なくとも80%の同一性を有する。

【0277】

いくつかの実施形態では、少なくとも1つの異種遺伝子によってコードされる酵素は、配列番号31、33、35、37、39、41、48、49、50、52、54、56、58、60、62、64、66、81、94、96、98、100、102、または104から選択される配列に対して少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、または少なくとも95%の同一性を有するアミノ酸配列を含む。いくつかの実施形態では、少なくとも1つの異種遺伝子によってコードされる酵素は、配列番号31、33、35、37、39、41、48、49、50、52、54、56、58、60、62、64、66、81、94、96、98、100、102または104から選択される配列に対して少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、または少なくとも95%の同一性を有するアミノ酸配列を含む。

10

【0278】

いくつかの実施形態では、少なくとも1つの異種遺伝子によってコードされる酵素は、配列番号31、33、35、37、39、41、48、49、50、52、54、56、58、60、62、64、66、81、94、96、98、100、102、または104から選択される配列に対して少なくとも80%、少なくとも85%、少なくとも90%、または少なくとも95%の同一性を有するアミノ酸配列を含む。いくつかの実施形態では、少なくとも1つの異種遺伝子によってコードされる酵素は、配列番号31、33、35、37、39、41、48、49、50、52、54、56、58、60、62、64、66、81、94、96、98、100、102、または104から選択される配列に対して少なくとも80%、少なくとも85%、または少なくとも95%の同一性を有するアミノ酸配列を含む。

20

【0279】

いくつかの実施形態では、コードされたCSLG酵素のアミノ酸配列は、配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列であるか、あるいは、配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列に対して少なくとも55%の同一性及び少なくとも55%のカバレッジを有するそのホモログである。いくつかの実施形態では、コードされたCSLG酵素のアミノ酸配列は、配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列であるか、あるいは、配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列に対して少なくとも55%の同一性及び少なくとも80%のカバレッジを有するそのホモログである。

30

40

【0280】

いくつかの実施形態では、コードされたCSLG酵素のアミノ酸配列は、配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列であるか、あるいは、配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログである。

【0281】

いくつかの実施形態では、コードされた少なくとも1つの追加の異種遺伝子のアミノ酸配列は - アミリン合成酵素をコードし、そのアミノ酸配列は、配列番号45に記載のア

50

ミノ酸配列、あるいは、配列番号45に記載のアミノ酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログである。いくつかの実施形態では、コードされた少なくとも1つの追加の異種遺伝子のアミノ酸配列はシトクロムP450をコードし、そのアミノ酸配列は、配列番号46、51、または53のいずれか1つに記載のアミノ酸配列であるか、あるいは、配列番号46、51、または53のいずれか1つに記載のアミノ酸配列と少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログである。いくつかの実施形態では、コードされた少なくとも1つの追加の異種遺伝子のアミノ酸配列はグリコシルトランスフェラーゼをコードし、そのアミノ酸配列は、配列番号55、57、59、または61のいずれか1つに記載のアミノ酸配列であるか、あるいは、配列番号55、57、59、または61のいずれか1つに記載のアミノ酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログである。いくつかの実施形態では、コードされた少なくとも1つの追加の異種遺伝子のアミノ酸配列はアセチルトランスフェラーゼをコードし、そのアミノ酸配列は、配列番号63に記載のアミノ酸配列、あるいは、配列番号63に記載のアミノ酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログである。いくつかの実施形態では、コードされた少なくとも1つの追加の異種遺伝子のアミノ酸配列は、UDP-グルコース6-デヒドロゲナーゼ1をコードし、配列番号74に記載のアミノ酸配列、あるいは、配列番号74に記載のアミノ酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログである。

10

20

**【0282】**

いくつかの実施形態では、遺伝子組換え植物細胞は、少なくとも1つの内在性CSLG遺伝子の発現が、対応する非組換え細胞と比較して改変されている。いくつかの実施形態では、遺伝子組換え植物細胞は、少なくとも1つの内在性CSLG遺伝子の発現が、対応する非組換え細胞と比較して改変され、その遺伝子組換え植物細胞は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え細胞と比較して改変されている。

**【0283】**

内在性CSLG遺伝子の発現を改変させると、ステロイド系アルカロイド生合成経路、ステロイド系サポニン生合成経路、及び/またはトリテルペノイド系サポニン生合成経路に影響を与え、いくつかの実施形態では、ステロイド系アルカロイドのプロファイルまたはその経路の中間体、ステロイド系サポニンのプロファイルまたはその経路の中間体、またはトリテルペノイド系サポニンのプロファイルまたはその経路の中間体に、それぞれ付随した改変をもたらす。

30

**【0284】**

いくつかの実施形態では、遺伝子組換え細胞における発現の改変は、対応する非組換え細胞と比較してCSLG遺伝子の発現が増加することを含む。いくつかの実施形態では、発現の増加は、コードされた酵素の量の増加をもたらす。いくつかの実施形態では、CSLG遺伝子の発現の増加は、少なくとも1つのステロイド系アルカロイドまたはその中間体の増加を含むステロイド系アルカロイドプロファイルの改変をもたらす。いくつかの実施形態では、CSLG遺伝子の発現の増加は、少なくとも1つのステロイド系サポニンまたはその中間体の増加を含むステロイド系サポニンプロファイルの改変をもたらす。いくつかの実施形態では、CSLG遺伝子の発現の増加は、少なくとも1つのトリテルペノイド系サポニンまたはその中間体の増加を含むトリテルペノイド系サポニンプロファイルの改変をもたらす。

40

**【0285】**

いくつかの実施形態では、遺伝子組換え細胞における発現の改変は、対応する非組換え細胞と比較してCSLG遺伝子の発現が低下することを含む。いくつかの実施形態では、発現の減少は、コードされた酵素の量の減少をもたらす。いくつかの実施形態では、CSLG遺伝子の発現の低下は、少なくとも1つのステロイド系アルカロイドまたはその中間

50

体の減少を含むステロイド系アルカロイドプロファイルの改変をもたらす。いくつかの実施形態では、CSLG遺伝子の発現の低下は、少なくとも1つのステロイド系サポニンまたはその中間体の減少を含むステロイド系サポニンプロファイルの改変をもたらす。いくつかの実施形態では、CSLG遺伝子の発現の低下は、少なくとも1つのトリテルペノイド系サポニンまたはその中間体の減少を含むトリテルペノイド系サポニンプロファイルの改変をもたらす。

【0286】

いくつかの実施形態では、内在性CSLG遺伝子の発現が改変された少なくとも1つの細胞は、植物細胞を含む。いくつかの実施形態では、内在性CSLG遺伝子の発現が改変された少なくとも1つの細胞は、植物またはその一部に含まれる植物細胞を含む。いくつかの実施形態では、内在性CSLG遺伝子の発現が改変された少なくとも1つの細胞は、CSLG遺伝子の突然変異の結果である。いくつかの実施形態では、突然変異は、少なくとも1以上の点突然変異、挿入、欠失、またはそれらの任意の組み合わせを含み、CSLG酵素は、安定性の向上、活性の増加、またはその両方を含み、含有量の改変は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が、対応する非組換え植物と比較して増加することである。いくつかの実施形態では、突然変異は、少なくとも1以上の点突然変異、挿入、欠失、またはそれらの任意の組み合わせを含み、CSLG酵素は、安定性の低下、活性の低下、またはその両方を含み、含有量の改変は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が、対応する非組換え植物と比較して減少することである。

10

20

【0287】

いくつかの実施形態では、内在性CSLG遺伝子の発現が改変され、改変は、CSLG遺伝子を突然変異させることを含み、突然変異は、1以上の点突然変異、挿入、欠失、またはゲノム編集の導入、または細菌のCRISPR/CAS系の使用、またはそれらの組み合わせを含む。いくつかの実施形態では、CSLG酵素をコードする遺伝子の発現が、対応する非組換え植物における発現と比較して増加し、これにより、遺伝子組換え植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して増加する。いくつかの実施形態では、CSLG酵素をコードする遺伝子の発現が、対応する非組換え植物における発現と比較して減少し、これにより、遺伝子改変植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して増加する。

30

【0288】

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの誘導体は、それぞれ、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンのグリコシル化誘導体を含む。

40

【0289】

CSLG遺伝子への、欠失、挿入、部位特異的変異、ジンクフィンガーヌクレアーゼなどを含む突然変異の挿入は、例えば、遺伝子発現の上方制御、発現したCSLG酵素の安定性の向上、CSLG酵素の活性の増加、遺伝子発現の下方制御、発現したCSLG酵素の安定性の低下、CSLG酵素の活性の低下、またはCSLG酵素の活性の消失などのために利用することができる。

【0290】

いくつかの実施形態では、突然変異したCSLG遺伝子は、CSLG遺伝子の核酸配列

50

に少なくとも1つの突然変異を有するポリヌクレオチドを含み、そのCSLG遺伝子の核酸配列は、配列番号31、33、35、37、39、41、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列またはその相補的配列である。

#### 【0291】

当業者であれば、遺伝子組換え植物は、人間によって遺伝子組換えされた少なくとも1つの細胞を含む植物を包含することを理解できるであろう。遺伝子組換えは、例えば、内在性ポリヌクレオチドまたは目的の遺伝子に対して突然変異、欠失、挿入、転移性因子などを導入することによる、内在性遺伝子の組み換えを含む。追加的または代替的に、いくつかの実施形態では、遺伝子組換えは、異種ポリヌクレオチドで植物細胞を形質転換することを含む。当業者であれば、少なくとも1つの植物細胞を異種ポリヌクレオチドまたは複数の異種ポリヌクレオチドで形質転換することを含む遺伝子組換え植物が、いくつかの実施形態では、同一の意味及び性質を有する「トランスジェニック植物」または「遺伝子組換え植物」と呼ばれることを理解できるであろう。

10

#### 【0292】

内在性CSLG遺伝子の発現を改変することは、いくつかの実施形態では、対応するタンパク質をコードする核酸分子に1以上の点突然変異を導入することによって達成することができる。突然変異は、例えば、部位特異的突然変異誘発を用いて導入することができる（例えば、「Wu Ed., 1993 Meth. In Enzymol. Vol. 217, San Diego: Academic Press; Higuchi, "Recombinant PCR" in Innis et al. Eds., 1990 PCR Protocols, San Diego: Academic Press, Inc」を参照）。このような突然変異誘発は、特定の所望のアミノ酸の挿入、欠失または置換を導入するために用いることができる。標的突然変異誘発のためのいくつかの技術は、ゲノム中の二本鎖切断(DSB)の標的誘導と、それに続く、変異性DNA修復に基づいている。これらの方法によってゲノム編集のために最も一般的に使用されるのは、ジンクフィンガーヌクレアーゼを含むカスタム設計されたヌクレアーゼ、メガヌクレアーゼを含むエンドヌクレアーゼ、及びキサントモナス由来の転写活性化因子様エフェクターヌクレアーゼ(TALEN)酵素である。

20

#### 【0293】

いくつかの実施形態では、CSLG遺伝子の発現を改変する場合、改変は、CSLG遺伝子を突然変異させることを含む。この突然変異は、CSLG遺伝子のコード領域、CSLG遺伝子の制御配列、またはそれらの組み合わせ中に存在する。

30

#### 【0294】

保存的変異体または非保存的変異体を産生するために、様々な種類の突然変異誘発を用いて、CSLG及びそれがコードする酵素を改変することができる。任意の利用可能な突然変異誘発方法を用いることができる。いくつかの実施形態では、突然変異誘発方法は、部位特異的突然変異誘発を含む。いくつかの実施形態では、突然変異誘発方法は、ランダム点突然変異誘発を含む。いくつかの実施形態では、突然変異誘発方法は、インビトロまたはインビボでの相同組換え(DNAシャッフリング)を含む。いくつかの実施形態では、突然変異誘発方法は、ウラシル含有テンプレートを用いた突然変異誘発を含む。いくつかの実施形態では、突然変異誘発方法は、オリゴヌクレオチド特異的突然変異誘発を含む。いくつかの実施形態では、突然変異誘発方法は、ホスホロチオエート修飾DNA突然変異誘発を含む。いくつかの実施形態では、突然変異誘発方法は、ギャップを有する二重鎖DNAを用いた突然変異誘発を含む。いくつかの実施形態では、突然変異誘発方法は、点ミスマッチ修復を含む。いくつかの実施形態では、突然変異誘発方法は、修復欠損宿主株を用いた突然変異誘発を含む。いくつかの実施形態では、突然変異誘発方法は、制限選択及び制限精製を含む。いくつかの実施形態では、突然変異誘発方法は、欠失突然変異誘発を含む。いくつかの実施形態では、突然変異誘発方法は、全遺伝子合成による突然変異誘発を含む。いくつかの実施形態では、突然変異誘発方法は、二本鎖切断修復を含む。いくつかの実施形態では、突然変異誘発方法は、キメラ構築物による突然変異誘発を含む。いくつかの実施形態では、突然変異誘発方法は、CRISPR/Casによる突然変異誘

40

50

発を含む。いくつかの実施形態では、突然変異誘発方法は、メガヌクレアーゼによる突然変異誘発を含む。いくつかの実施形態では、突然変異誘発方法は、ジンクフィンガーヌクレアーゼ (ZFN) による突然変異誘発を含む。いくつかの実施形態では、突然変異誘発方法は、転写活性化因子様エフェクターヌクレアーゼ (TALEN) による突然変異誘発を含む。いくつかの実施形態では、突然変異誘発方法は、当業者に知られている任意の他の突然変異誘発方法を含む。

【0295】

いくつかの実施形態では、突然変異誘発は、天然に存在する分子及び/または突然変異した分子に関する既知の情報によって誘導される。例として、この既知の情報は、配列、配列比較、物理的特性、結晶構造などを含む。いくつかの実施形態では、突然変異誘発は、本質的にランダムである。いくつかの実施形態では、突然変異誘発方法は、DNAシャッフリングである。

10

【0296】

当業者であれば、クラスタ化された規則的に間隔を置いて配置された短いパリンδροーム反復 (CRISPR) / CRISPR 関連タンパク質 (Cas) 系が、細菌 CRISPR / Cas 原核生物適応免疫系に基づくゲノム操作ツールを含むことを理解できるだろう。この RNA に基づく技術は、非常に特異的であり、カスタマイズ可能なスモールノンコーディング RNA によって導かれるゲノム DNA の標的切断を可能にし、それにより、非同相末端結合 (NHEJ) 及び相同性指向修復 (HDR) の両方のメカニズムによる遺伝子改変をもたらす (Belhaj K. et al., 2013. Plant Methods 2013, 9:39)。いくつかの実施形態では、CRISPR / Cas 系は、CRISPR / Cas 9 系を含む。

20

【0297】

いくつかの実施形態では、CRISPR / Cas 系は、当分野で既知の単一ガイド RNA (sgRNA) 及び/または Cas タンパク質を含む。いくつかの実施形態では、CRISPR / Cas 系は、予め選択された部位で切断するために新たに作製されたシングルガイド RNA (sgRNA) 及び/または Cas タンパク質を含む。当業者であれば、「シングルガイド RNA」、「sgRNA」、及び「gRNA」という用語が、互換的に使用することができ、すべて同一の性質及び意味を有することを理解できるであろう。sgRNA は、CRISPR RNA (crRNA) 及びトランスコードされた CRISPR RNA (tracrRNA) からなるキメラ RNA 分子を包含することができる。いくつかの実施形態では、crRNA は、GAME15 DNA の予め選択された領域に相補的であり、crRNA は、CRISPR 関連ポリペプチド (Cas) ヌクレアーゼタンパク質を予め選択された標的部位を「標的化」する。

30

【0298】

いくつかの実施形態では、相補的な crRNA 配列の長さは、19 ~ 22 ヌクレオチド長、例えば、標的部位に相補的な 19 ~ 22 個の連続ヌクレオチドである。別の実施形態では、DNA の領域に相補的な crRNA 配列の長さは、約 15 ~ 30 ヌクレオチド長である。別の実施形態では、DNA の領域に相補的な crRNA の配列の長さは、約 15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、または 30 ヌクレオチド長である。別の実施形態では、DNA の領域に相補的な crRNA 配列の長さは、20 ヌクレオチド長である。いくつかの実施形態では、crRNA は、sgRNA 分子の 5' 末端に位置する。別の実施形態では、crRNA は、予め選択された標的配列内で 100% の相補的を含む。別の実施形態では、crRNA は、予め選択された標的配列内で少なくとも 80% の相補的を含む。別の実施形態では、crRNA は、予め選択された標的配列内で少なくとも 85% の相補的を含む。別の実施形態では、crRNA は、予め選択された標的配列内で少なくとも 90% の相補的を含む。別の実施形態では、crRNA は、予め選択された標的配列内で少なくとも 95% の相補的を含む。別の実施形態では、crRNA は、予め選択された標的配列内で少なくとも 97% の相補的を含む。別の実施形態では、crRNA は、予め選択された標的配列内で少なくとも 99% の相補的を含む。別の実施形態では、tracrRNA は、100 ~ 300 ヌクレ

40

50

オチド長であり、Casヌクレアーゼ、例えば、CRISPR/Cas9複合体を形成するCas9タンパク質の結合部位を提供する。

【0299】

一実施形態では、突然変異誘発系は、CRISPR/Cas系を含む。別の実施形態では、CRISPR/Cas系は、CasヌクレアーゼとgRNA分子とを含み、gRNA分子は、予め選択された内在性標的部位内で結合し、それによって、Casヌクレアーゼが予め選択された内在性標的部位内のDNAを切断するように誘導する。

【0300】

いくつかの実施形態では、CRISPR/Cas系は、標的ポリヌクレオチドの領域、例えば予め選択された内在性標的部位に相補的または実質的に相補的なヌクレオチド配列を含むガイドRNA配列（「gRNA」または「sgRNA」）と、ヌクレアーゼ活性を有するタンパク質とを含む酵素系を含む。

10

【0301】

別の実施形態では、CRISPR-Cas系は、I型CRISPR-Cas系、II型CRISPR-Cas系、または、III型CRISPR-Cas系、またはそれらの誘導體を含む。別の実施形態では、CRISPR-Cas系は、自然発生のCRISPR-Cas系に由来する、操作された及び/またはプログラムされたヌクレアーゼ系を含む。別の実施形態では、CRISPR-Cas系は、操作された及び/または突然変異したCasタンパク質を含む。別の実施形態では、CRISPR-Cas系は、操作された及び/またはプログラムされたガイドRNAを含む。

20

【0302】

当業者であれば、ガイドRNAが、例えば予め選択された内在性標的部位などの標的DNA配列の領域に対して相補的または実質的に相補的な領域以外のヌクレオチド配列を含み得ることを理解できるであろう。別の実施形態では、ガイドRNAは、crRNAまたはその誘導體を含む。別の実施形態では、ガイドRNAは、crRNA:tracrRNAキメラを含む。

【0303】

別の実施形態では、gRNA分子は、少なくとも1つの相同染色体上の予め選択された内在性標的部位に相補的であり、かつそれに結合するドメインを含む。別の実施形態では、gRNA分子は、少なくとも1つの相同染色体上の多型対立遺伝子に相補的であり、かつそれに結合するドメインを含む。別の実施形態では、gRNA分子は、両方の相同染色体上の予め選択された内在性標的部位に相補的であり、かつそれに結合するドメインを含む。別の実施形態では、gRNA分子は、両方の相同染色体上の多型対立遺伝子に相補的であり、かつそれに結合するドメインを含む。

30

【0304】

Cas酵素は、DNAに二本鎖切断(DSB)を作ることができるRNAガイドのDNAエンドヌクレアーゼを含む。「Cas酵素」という用語は、「CRISPR関連エンドヌクレアーゼ」または「CRISPR関連ポリペプチド」という用語と互換的に使用することができ、すべての同一の性質及び意味を有する。一実施形態では、Cas酵素は、Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas6、Cas7、Cas8、Cas9、Cas10、C2c1、CasX、NgAgo、Cpf1、Csy1、Csy2、Csy3、Cse1、Cse2、Csc1、Csc2、Csa5、Csn2、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Csb1、Csb2、Csb3、Csx17、Csx14、Csx10、Csx16、CsaX、Csx3、Csx1、Csx15、Csf1、Csf2、Csf3、Csf4、それらのホモログ、及びそれらの改変版からなる群から選択される。別の実施形態では、Cas酵素は、Cas9を含む。別の実施形態では、Cas酵素は、Cas1を含む。別の実施形態では、Cas酵素は、Cas1Bを含む。別の実施形態では、Cas酵素は、Cas2を含む。別の実施形態では、Cas酵素は、Cas3を含む。別の実施形態では、Cas酵素は、Cas4を含む。別の実施形態では、Cas

40

50

酵素は、C a s 5 を含む。別の実施形態では、C a s 酵素は、C a s 6 を含む。別の実施形態では、C a s 酵素は、C a s 7 を含む。別の実施形態では、C a s 酵素は、C a s 8 を含む。別の実施形態では、C a s 酵素は、C a s 10 を含む。別の実施形態では、C a s 酵素は、C p f 1 を含む。別の実施形態では、C a s 酵素は、C s y 1 を含む。別の実施形態では、C a s 酵素は、C s y 2 を含む。別の実施形態では、C a s 酵素は、C s y 3 を含む。別の実施形態では、C a s 酵素は、C s e 1 を含む。別の実施形態では、C a s 酵素は、C s e 2 を含む。別の実施形態では、C a s 酵素は、C s c 1 を含む。別の実施形態では、C a s 酵素は、C s c 2 を含む。別の実施形態では、C a s 酵素は、C s a 5 を含む。別の実施形態では、C a s 酵素は、C s n 2 を含む。別の実施形態では、C a s 酵素は、C s m 2 を含む。別の実施形態では、C a s 酵素は、C s m 3 を含む。別の実施形態では、C a s 酵素は、C s m 4 を含む。別の実施形態では、C a s 酵素は、C s m 5 を含む。別の実施形態では、C a s 酵素は、C s m 6 を含む。別の実施形態では、C a s 酵素は、C m r 1 を含む。別の実施形態では、C a s 酵素は、C m r 3 を含む。別の実施形態では、C a s 酵素は、C m r 4 を含む。別の実施形態では、C a s 酵素は、C m r 5 を含む。別の実施形態では、C a s 酵素は、C m r 6 を含む。別の実施形態では、C a s 酵素は、C s b 1 を含む。別の実施形態では、C a s 酵素は、C s b 2 を含む。別の実施形態では、C a s 酵素は、C s b 3 を含む。別の実施形態では、C a s 酵素は、C s x 17 を含む。別の実施形態では、C a s 酵素は、C s x 14 を含む。別の実施形態では、C a s 酵素は、C s x 10 を含む。別の実施形態では、C a s 酵素は、C s x 16、C s a X を含む。別の実施形態では、C a s 酵素は、C s x 3 を含む。別の実施形態では、C a s 酵素は、C s x 1、C s x 15、C s f 1 を含む。別の実施形態では、C a s 酵素は、C s f 2 を含む。別の実施形態では、C a s 酵素は、C s f 3 を含む。別の実施形態では、C a s 酵素は、C s f 4 を含む。別の実施形態では、C a s 酵素は、C p f 1 を含む。別の実施形態では、C a s 酵素は、C 2 c l を含む。別の実施形態では、C a s 酵素は、C a s X を含む。別の実施形態では、C a s 酵素は、N g A g o を含む。別の実施形態では、C a s 酵素は、C a s のホモログである。別の実施形態では、C a s 酵素は、C a s オルソログである。別の実施形態では、C a s 酵素は、改変された C a s 酵素である。別の実施形態では、C a s 酵素は、当分野で知られている任意の C R I S P R 関連エンドヌクレアーゼである。

10

20

30

#### 【0305】

当業者であれば、「ジンクフィンガーヌクレアーゼ」または「ZFN」という用語は、すべて同一の意味と性質を有する互換性のあるものであり、ZFNは、DNAの二本鎖切断が可能で少なくとも1つのヌクレアーゼに作動可能に連結された少なくとも1つのジンクフィンガーDNA結合ドメインを含むキメラタンパク質分子を包含することを理解できるであろう。いくつかの実施形態では、ZFN系は、当分野で既知のZFNを含む。いくつかの実施形態では、ZFN系は、予め選択された部位を切断するために新たに作製されたZFNを含む。

#### 【0306】

いくつかの実施形態では、ZFNは、予め選択された内在性標的部位で二本鎖切断を生じさせる。いくつかの実施形態では、ZFNは、DNA結合ドメイン及びDNA切断ドメインを含み、DNA結合ドメインは、少なくとも1つの亜鉛フィンガーを含み、DNA切断ドメインに作動可能に連結されている。別の実施形態では、亜鉛フィンガーのDNA結合ドメインは、キメラタンパク質分子のN末端に位置し、DNA切断ドメインは、キメラタンパク質分子のC末端に位置する。別の実施形態では、亜鉛フィンガーDNA結合ドメインは、キメラタンパク質分子のC末端に位置し、DNA切断ドメインは、キメラタンパク質分子のN末端に位置する。別の実施形態では、ジンクフィンガー結合ドメインは、ジンクフィンガーヌクレアーゼにおける、例えば本明細書に開示されたような予め選択された内在性標的部位などの標的遺伝子座に結合することができる領域を包含する。別の実施形態では、ジンクフィンガーDNA結合ドメインは、少なくとも1つの相同染色体上の予め選択された内在性標的部位に結合するタンパク質ドメインを含む。別の実施形態では、

40

50

ジンクフィンガー-DNA結合ドメインは、少なくとも1つの相同染色体上の多型対立遺伝子に結合するタンパク質ドメインを含む。別の実施形態では、ジンクフィンガー-DNA結合ドメインは、両相同染色体上の予め選択された内在性標的部位に結合するタンパク質ドメインを含む。別の実施形態では、ジンクフィンガー-DNA結合ドメインは、両方の相同染色体上の多型対立遺伝子に結合するタンパク質ドメインを含む。

#### 【0307】

当業者であれば、「キメラタンパク質」という用語が、2以上のDNA断片を作動的に連結することによって作製されたDNA分子から発現されたタンパク質を説明するために使用されていることを理解できるだろう。DNA断片は、同一の種に由来するものであってもよいし、異なる種に由来するものであってもよい。また、DNA断片は、同一の遺伝子に由来するものであってもよいし、異なる遺伝子に由来するものであってもよい。当業者であれば、ZFNの「DNA切断ドメイン」という用語が、ジンクフィンガーヌクレアーゼにおける、ヌクレオチド鎖の核酸間の化学結合を破壊することができる領域を包含することを理解できるであろう。切断ドメインを含むタンパク質の例としては、制限酵素、トポイソメラーゼ、リコンビナーゼ、インテグラーゼ、DNaseなどが挙げられる。

#### 【0308】

当業者であれば、エンドヌクレアーゼには、制限エンドヌクレアーゼと同様に特定のDSB標的部位に結合して切断するメガヌクレアーゼ（ホーミングエンドヌクレアーゼ（HEase）とも呼ばれる）が含まれ、メガヌクレアーゼのDSB標的部位は一般的に約18bp以上であることを理解できるであろう。メガヌクレアーゼのドメイン、構造及び機能は、例えば、「Guhan and Muniyappa (2003) Crit Rev Biochem Mol Biol 38:199-248; Lucas et al., (2001) Nucleic Acids Res 29:960-9」、*「Jurica and Stoddard, (1999) Cell Mol Life Sci 55:1304-26」*、*「Stoddard, (2006) Q Rev Biophys 38:49-95」*、及び*「Moure et al., (2002) Nat Struct Biol 9:764」*を参照されたい。いくつかの例では、天然に存在する変異体、及び/または操作された誘導体メガヌダーゼが使用される。任意のメガヌダーゼを使用することができ、そのようなものとしては、これに限定しないが、1-Scel、1-Scell、1-Scelll、1-ScelV、1-SceV、1-SceVI、1-SceVII、1-CeuI、1-CeuAIIIP、1-Crel、1-Crepsb1P、1-Crepsb1lP、1-Crepsb1llP、1-Crepsb1VP、1-Tlil、1-Ppol、Pl-Pspl、F-Scel、F-Scell、F-Suvl、F-Tevl、F-Tevll、1-Amal、1-Anil、1-Chul、1-Cmoel、1-Cpal、1-Cpall、1-Csml、1-Cvul、1-CvuAIP、1-Ddil、1-Ddill、1-Dir1、1-Dmol、1-Hmul、1-Hmull、1-HsNIP、1-Llal、1-Msol、1-Naal、1-Nanl、1-NcllP、1-NgrIP、1-Nit1、1-Njal、1-Nsp236IP、1-Pakl、1-PbolP、1-PculP、1-PcuAI、1-PcuVI、1-PgrIP、1-PoblP、1-Por1、1-PorllP、1-PbpIP、1-SpBetallP、1-Scal、1-Sex1P、1-SnelP、1-Spoml、1-SpomCP、1-SpomlP、1-SpomllP、1-Squ1P、1-Ssp68031、1-SthPhiJP、1-SthPhiST3P、1-SthPhiSTe3bP、1-TdelP、1-Tevl、1-Tevll、1-Tevlll、1-UarAP、1-UarHGPAIP、1-UarHGPA13P、1-VinIP、1-ZbilP、Pl-Mtul、PI-MtuHIPII-MtuHIIP、Pl-Pful、Pl-Pfull、Pl-Pkol、Pl-Pkoll、PI-Rma43812IP、PI-SpBetallP、Pl-Scel、Pl-Tful、Pl-Tfull、Pl-Thyl、Pl-Tlil、Pl-Tlill、またはそれらの活性変異体や断片が挙げられる。TALエフェクターヌクレアーゼを用いて、植物または他の生物のゲノム内の特定の標的配列において二本鎖切断を行うことができる。TALエフェクターは、転写活性化因子（TAL）エフェクターまたはその機能的な部分を、エンドヌクレアーゼ（例えば、FokI）の触媒ドメインに融合させ

10

20

30

40

50

ることによって作製することができる。ユニークなモジュール型 T A L エフェクター D N A 結合ドメインは、任意の D N A 認識特異性を潜在的に有するタンパク質の設計を可能にする。したがって、T A L エフェクターのヌクレアーゼの D N A 結合ドメインは、特定の D N A 標的部 位を認識するように操作することができ、したがって、所望の標的配列において二本鎖切断を行うために使用することができる。「WO 2010/079430」、「Morbiter et al. (2010) PNAS 10.1073/pnas.1013133107」、「Scholze & Boch (2010) Virulence 1 :428-432」、「Christian et al. Genetics (2010) 186:757-761」、「Li et al. (2010) Nuc. Acids Res. (2010) doi:10.1093/nar/gkq704」、「Miller et al. (2011) Nature Biotechnology 29:143-148」(これらはすべて、参照により本明細書に組み込まれる)を参照されたい。

10

**【0309】**

いくつかの実施形態では、T A L E N 系は、T A L エフェクター D N A 結合ドメイン及び D N A 切断ドメインを含み、T A L エフェクター D N A 結合ドメインは、予め選択された内在性標的部 位内で結合し、それによって、D N A 切断ドメインを標的化して、予め選択された内在性標的部 位内で D N A を切断する。

**【0310】**

当業者であれば、「転写活性化因子様エフェクターヌクレアーゼ」、「T A L E N」、及び「T A L エフェクターヌクレアーゼ」という用語は、互換的に使用することができ、すべて同一の性質及び意味を有することを理解できるであろう。T A L E N は、その標的部 位、例えば本明細書に開示されたような予め選択された内在性標的部 位を認識して切断することができるヌクレアーゼを包含する。別の実施形態では、T A L E N は、T A L E ドメインとヌクレオチド切断ドメインとを含む融合タンパク質を含む。別の実施形態では、T A L E ドメインは、1 以上の T A L E 反復モジュールを介して、配列特異的な方法でヌクレオチドに結合するタンパク質ドメインを含む。当業者であれば、T A L E 反復モジュールは、植物の D N A 配列を認識する約 3 4 個の可変数のアミノ酸反復を含むことを認識できるであろう。さらに、反復モジュールは、新しい D N A 配列を標的とするために、単純な暗号にしたがって再配列させることができる。別の実施形態では、T A L E ドメインは、少なくとも 1 つの相同性染色体上の予め選択された内在性標的部 位に結合するタンパク質ドメインを含む。別の実施形態では、T A L E ドメインは、少なくとも 1 つの相同染色体上の多型対立遺伝子に結合するタンパク質ドメインを含む。別の実施形態では、T A L E ドメインは、両方の相同染色体上の予め選択された内在性標的部 位に結合するタンパク質ドメインを含む。別の実施形態では、T A L E ドメインは、両方の相同染色体上の多型対立遺伝子に結合するタンパク質ドメインを含む。

20

30

**【0311】**

一実施形態では、T A L E ドメインは、T A L E 反復モジュールの少なくとも 1 つを含む。別の実施形態では、T A L E ドメインは、1 ~ 3 0 個の T A L E 反復モジュールを含む。別の実施形態では、T A L E ドメインは、3 0 個を超える反復モジュールを含む。別の実施形態では、T A L E N 融合タンパク質は、N 末端ドメイン、1 以上の T A L E 反復モジュールとそれに続く半反復モジュール、リンカー、及びヌクレオチド切断ドメインを含む。

40

**【0312】**

エチルメチルスルホン酸 ( E M S ) などの薬剤を使用する化学的突然変異誘発法を用いて点変異の集団を得て、サイレンシングまたはダウンレギュレーションされる C S G L 遺伝子の突然変異体をスクリーニングすることができる。植物の場合は、自然集団からの遺伝子移入に基づく方法を用いることができる。培養種と野生種とを繰り返し交配し、野生ゲノムの所定の断片を含む植物を単離する。例えばトウモロコシやキングジョウなどの特定の種の植物は、天然のトランスポゾンを持している。これらのトランスポゾンは、トランスポゾンの配列内にトランスポザーゼが存在する自律的なものと、トランスポザーゼを有していない非自律的なものとがある。当業者であれば、トランスポゾンを「ジャンプ」させて、突然変異を生じさせることができる。あるいは、1 以上の所定の位置にランダム

50

なヌクレオチドを有する核酸配列を合成することにより、ランダムなアミノ酸置換を生じさせることもできる。

【0313】

いくつかの実施形態では、内在性CSLG遺伝子の発現は、それらの制御配列に1以上の点突然変異を導入することによって改変することができる。いくつかの実施形態では、異種CSLG遺伝子の発現は、それらの制御配列に1以上の点突然変異を導入することによって改変することができる。

【0314】

当業者であれば、「調節配列」は、コーディング配列の上流(5'非コーディング配列)、内部、または下流(3'非コーディング配列)に位置し、関連するコーディング配列の転写、RNAの処理または安定性、あるいは翻訳に影響を与えるヌクレオチド配列を指すことを理解できるであろう。いくつかの実施形態では、調節配列は、プロモータを含む。いくつかの実施形態では、調節配列は、翻訳リーダー配列を含む。いくつかの実施形態では、調節配列は、イントロンを含む。いくつかの実施形態では、調節配列は、ポリアデニル化認識配列を含む。いくつかの実施形態では、調節配列は、RNAプロセシング部位を含む。いくつかの実施形態では、調節配列は、エフェクター結合部位を含む。いくつかの実施形態では、制御配列は、ステムループ構造を含む。

10

【0315】

当業者であれば、「プロモータ」は、コーディング配列または機能性RNAの発現を制御することができるDNA配列を指すことを理解できるであろう。いくつかの実施形態では、コーディング配列は、プロモータ配列の3'に位置する。当業者であれば、異なるプロモータが、異なる組織または細胞型において、または異なる発生段階において、または異なる環境または生理学的条件に応答して、遺伝子の発現を指示することは、理解できるであろう。いくつかの実施形態では、プロモータは、構成的プロモータ、すなわち、ほとんどの細胞型において、ほとんどの時間で遺伝子を発現させるプロモータを含む。いくつかの実施形態では、プロモータは、制御されたプロモータ、すなわち、散発的な特定の刺激に応答して遺伝子を発現させるプロモータを含む。さらに、多くの場合、調節配列の正確な境界はまだ完全には定義されていないことが認識されている。

20

【0316】

当業者であれば、「3'非コード配列」または「転写ターミネーター」という用語が、コード配列の下流に位置するDNA配列を指すことを理解できるであろう。いくつかの実施形態では、3'非コード配列は、ポリアデニル化認識配列を含む。いくつかの実施形態では、3'非コード配列は、mRNAプロセシングに影響を与えることができる制御シグナルをコードする配列を含む。いくつかの実施形態では、3'非コード配列は、遺伝子発現に影響を与えることができる制御シグナルをコードする配列を含む。ポリアデニル化シグナルは、通常、mRNA前駆体の3'末端へのポリアデニル酸トラクトの付加に影響を与える。いくつかの実施形態では、3'非コード配列の変異は、遺伝子の転写に影響を与える。いくつかの実施形態では、3'非コード配列における変異は、RNAプロセシングに影響する。いくつかの実施形態では、3'非コード配列における変異は、遺伝子の安定性に影響を与える。いくつかの実施形態では、3'非コード配列における変異は、関連するコード化配列の翻訳に影響を与える。

30

40

【0317】

いくつかの実施形態では、CSLG酵素をコードする遺伝子の発現が、対応する非組換え植物における発現と比較して低減し、それによって、遺伝子組換え植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して低減する。

【0318】

50

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの誘導体は、それぞれ、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンのグリコシル化誘導体を含む。

【0319】

いくつかの実施形態では、遺伝子組換え植物は、CLSG酵素をコードするポリヌクレオチドを標的とする少なくとも1つのサイレンシング分子を含む少なくとも1つの細胞を含有するトランスジェニック植物である。いくつかの実施形態では、トランスジェニック植物は、CLSG酵素をコードするポリヌクレオチドを含み、そのポリヌクレオチドの発現は、選択的にサイレンシングされるか、抑制されるか、または発現が低減される。いくつかの実施形態では、トランスジェニック植物は、CSLG酵素をコードするポリヌクレオチドを含み、そのポリヌクレオチドは、その発現をサイレンシング、抑制、または低減するために、欠失、挿入、または修飾によって選択的に編集されているか、または、遺伝子組換え植物は、遺伝子編集された植物の子孫である。

10

【0320】

いくつかの実施形態では、トランスジェニック植物は、CSLG遺伝子を標的とする少なくとも1つのサイレンシング分子を含む少なくとも1つの細胞を含有する。

【0321】

いくつかの実施形態では、内在性CSLG遺伝子は、選択的にサイレンシングされるか、抑制されるか、または発現が低減され、改変された含有量は、対応する非組換え植物と比較した、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量の低減、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量の低減、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量の低減を含む。内在性CSLG遺伝子が、選択的にサイレンシングされるか、抑制されるか、または発現が低減される場合、上記の細胞は、CSLG遺伝子をコードするポリヌクレオチドを標的とする少なくとも1つのサイレンシング分子をさらに含み、そのサイレンシング分子は、RNA干渉分子またはアンチセンス分子から選択されるか、あるいは、そのサイレンシング分子は、ウイルス誘導遺伝子サイレンシング系の構成要素である。

20

【0322】

いくつかの実施形態では、トランスジェニック植物は、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載された核酸配列を標的とする少なくとも1つのサイレンシング分子を含む少なくとも1つの細胞を含有する。

30

【0323】

いくつかの実施形態では、トランスジェニック植物は、配列番号30に記載の核酸配列を標的とする少なくとも1つのサイレンシング分子を含む少なくとも1つの細胞を含有する。いくつかの実施形態では、トランスジェニック植物は、配列番号32に記載の核酸配列を標的とする少なくとも1つのサイレンシング分子を含む少なくとも1つの細胞を含有する。いくつかの実施形態では、トランスジェニック植物は、配列番号34に記載の核酸配列を標的とする少なくとも1つのサイレンシング分子を含む少なくとも1つの細胞を含有する。いくつかの実施形態では、トランスジェニック植物は、配列番号36に記載の核酸配列を標的とする少なくとも1つのサイレンシング分子を含む少なくとも1つの細胞を含有する。いくつかの実施形態では、トランスジェニック植物は、配列番号38に記載の核酸配列を標的とする少なくとも1つのサイレンシング分子を含む少なくとも1つの細胞を含有する。いくつかの実施形態では、トランスジェニック植物は、配列番号40に記載の核酸配列を標的とする少なくとも1つのサイレンシング分子を含む少なくとも1つの細胞を含有する。いくつかの実施形態では、トランスジェニック植物は、配列番号65に記載の核酸配列を標的とする少なくとも1つのサイレンシング分子を含む少なくとも1つの細胞を含有する。いくつかの実施形態では、トランスジェニック植物は、配列番号80に記載の核酸配列を標的とする少なくとも1つのサイレンシング分子を含む少なくとも1つ

40

50

の細胞を含有する。いくつかの実施形態では、トランスジェニック植物は、配列番号 93 に記載の核酸配列を標的とする少なくとも 1 つのサイレンシング分子を含む少なくとも 1 つの細胞を含有する。いくつかの実施形態では、トランスジェニック植物は、配列番号 95 に記載の核酸配列を標的とする少なくとも 1 つのサイレンシング分子を含む少なくとも 1 つの細胞を含有する。いくつかの実施形態では、トランスジェニック植物は、配列番号 97 に記載の核酸配列を標的とする少なくとも 1 つのサイレンシング分子を含む少なくとも 1 つの細胞を含有する。いくつかの実施形態では、トランスジェニック植物は、配列番号 99 に記載の核酸配列を標的とする少なくとも 1 つのサイレンシング分子を含む少なくとも 1 つの細胞を含有する。いくつかの実施形態では、トランスジェニック植物は、配列番号 101 に記載の核酸配列を標的とする少なくとも 1 つのサイレンシング分子を含む 10  
少なくとも 1 つの細胞を含有する。いくつかの実施形態では、トランスジェニック植物は、配列番号 103 に記載の核酸配列を標的とする少なくとも 1 つのサイレンシング分子を含む少なくとも 1 つの細胞を含有する。いくつかの実施形態では、トランスジェニック植物は、配列番号 105 に記載の核酸配列を標的とする少なくとも 1 つのサイレンシング分子を含む少なくとも 1 つの細胞を含有する。

#### 【0324】

いくつかの実施形態では、サイレンシング分子は RNA 干渉分子及びアンチセンス分子からなる群から選択されるか、または、サイレンシング分子はウイルス誘導遺伝子サイレンシング系の構成要素である。いくつかの実施形態では、サイレンシング分子は、配列番号 30、32、34、36、38、40、65、80、93、95、97、99、101  
20、103、または 105 のいずれか 1 つに記載された核酸配列、またはその相補的配列を有する CSLG 遺伝子の領域に対して実質的に相補的な核酸配列を有するポリヌクレオチドを含む。いくつかの実施形態では、植物または植物細胞における CSLG の発現をサイレンシング、抑制、または低減するアンチセンス分子は、配列番号 30、32、34、36、38、40、65、80、93、95、97、99、101、103、または 105  
のいずれか 1 つに記載の核酸配列、またはその相補的配列を有する CSLG 遺伝子の領域に対して実質的に相補的な核酸配列を有するポリヌクレオチドを含む。いくつかの実施形態では、植物または植物細胞における CSLG の発現をサイレンシング、抑制、または低減するアンチセンス分子は、配列番号 108 に記載のヌクレオチド配列を有する。いくつかの実施形態では、植物または植物細胞における CSLG の発現をサイレンシング、抑制  
30、または低減する VIG 分子は、配列番号 42、配列番号 43、配列番号 44、配列番号 106、または配列番号 107 に記載されたヌクレオチド配列を有する。いくつかの実施形態では、サイレンシング分子は、配列番号 30、32、34、36、38、40、65、80、93、95、97、99、101、103、または 105 のいずれか 1 つに記載の核酸配列、またはその相補的配列内の核酸配列を有する CSLG 断片を標的とする。いくつかの実施形態では、上記の植物細胞を含む遺伝子組換え植物は、トリテルペノイド系サポニン、その代謝物、その誘導體、またはその生合成中間体の含有量が減少する。

#### 【0325】

いくつかの実施形態では、遺伝子組換えまたは遺伝子編集された植物は、CSLG 遺伝子を標的とする少なくとも 1 つのサイレンシング分子を含む少なくとも 1 つの細胞を含有  
40する。いくつかの実施形態では、少なくとも 1 つのサイレンシング分子は、RNA 干渉分子及びアンチセンス分子からなる群から選択される。これらの実施形態では、遺伝子組換え植物は、ステロイド系アルカロイドまたはその生合成中間体、ステロイド系サポニンまたはその生合成中間体、あるいは、トリテルペノイド系サポニンまたはその生合成中間体の含有量が減少する。

#### 【0326】

CSLG 遺伝子を標的とするサイレンシング分子は、当業者に知られているようにして設計することができる。いくつかの実施形態では、サイレンシング分子は、CSLG 遺伝子の領域、これに限定しないが、例えば、トマト、野生型トマト、ジャガイモ、野生型ポテト、及びナスにおいて本明細書に開示されている GAME15 またはそのホモログ、に  
50

対して実質的に相補的な核酸配列を有するポリヌクレオチド含む。いくつかの実施形態では、サイレンシング分子は、CSLG遺伝子の領域、例えば、甘草、シロイヌナズナ、レッドビート、キノア、アルファルファ、大豆、またはミヤコグサにおいて本明細書に開示されているSOAP5またはそのホモログ、に対して実質的に相補的な核酸配列を有するポリヌクレオチド含む。

【0327】

いくつかの実施形態では、サイレンシング分子は、アンチセンスRNAである。

【0328】

いくつかの例示的な実施形態では、サイレンシング分子は、RNA干渉(RNAi)分子である。いくつかの実施形態では、サイレンシング分子は、二本鎖(ds)RNA分子である。いくつかの実施形態では、第1のポリヌクレオチド及び第2のポリヌクレオチドは、スパーサーによって分離されている。いくつかの実施形態では、スパーサー配列は、イントロンである。さらに他の実施形態では、第1のポリヌクレオチド及び第2のポリヌクレオチドの発現は、1つのプロモータに由来する。他の実施形態では、第1のポリヌクレオチド及び第2のポリヌクレオチドの発現は、2つのプロモータに由来する。2つのプロモータは、互いに同一であってもよいし、互いに異なってもよい。

10

【0329】

いくつかの実施形態では、「RNA干渉」または「RNAi」という用語は、低分子二本鎖RNAによって媒介される遺伝子発現のサイレンシングまたは低減を包含する。これは、動物や植物における、配列特異的な転写後の遺伝子サイレンシングのプロセスであり、その二重鎖領域における、サイレンシングされた遺伝子の配列と相同な阻害性RNA(iRNA)によって開始される。この遺伝子は、生物に対して内在性であっても外因性であってもよく、染色体に組み込まれていてもよく、または、ゲノムに組み込まれていないトランスフェクションベクターに存在していてもよい。遺伝子の発現は、完全にまたは部分的に阻害される。RNAiはまた、標的RNAの機能の全体または一部を阻害すると考えられる。

20

【0330】

一般的に、RNAi分子という用語は、センス配列とアンチセンス配列との両方を含む一本鎖または二本鎖のRNA分子を指す。例えば、RNA干渉分子は、自己相補的なセンス領域とアンチセンス領域とを含む二本鎖ポリヌクレオチド分子であってよく、アンチセンス領域は、標的核酸分子に対する相補性を有する。あるいは、RNAi分子は、自己相補的なセンス領域と、標的核酸分子に対する相補性を有するアンチセンス領域とを有する一本鎖ヘアピンポリヌクレオチドであってよい。または、RNAi分子は、2以上のループ構造、並びに、自己相補的なセンス領域と、標的核酸分子に対する相補性を有するアンチセンス領域とを含むステムを有する環状一本鎖ポリヌクレオチドであってもよい。この環状一本鎖ポリヌクレオチドは、インビボまたはインビトロで処理することにより、RNAiを媒介可能な活性分子を生成することができる。

30

【0331】

アンチセンス技術は、アンチセンスRNAまたはDNA分子が、標的センスDNAまたはRNA鎖と相互作用するプロセスである。センス鎖は、5'~3'のmRNA分子またはDNA分子である。センス鎖に対して相補的な鎖、すなわちミラー鎖は、アンチセンスと呼ばれる。アンチセンス鎖がセンスmRNA鎖と相互作用すると、その二重らせんは細胞にとって異物である認識され、分解され、その結果、タンパク質の産生量が減少したり、タンパク質が産生されなくなったりする。DNAはすでに二本鎖の分子であるが、アンチセンス技術を適用することにより、三本鎖を形成することができる。

40

【0332】

当業者であれば、本明細書で使用するとき、「相補的」または「その相補体」という用語は、相補的領域の全体にわたって別の特定のポリヌクレオチドとワトソン&クリック塩基対を形成することができるポリヌクレオチドの配列を包含するために使用されることを理解できるであろう。この用語は、2つのポリヌクレオチドが実際に結合する特定の条件

50

セットではなく、それらの配列のみに基づいてポリヌクレオチド対に適用される。

【0333】

C S L Gの細胞及び/または組織レベルのアンチセンス調節は、少なくとも1つのアンチセンス化合物、例えば、アンチセンスDNA、アンチセンスRNA、リボザイム、DNAザイム、ロック核酸(LNA)、アプタマーなどで、細胞または組織を形質転換することによって行うことができる。いくつかの実施形態では、分子は、化学的に修飾されている。他の実施形態では、アンチセンス分子は、アンチセンスDNA、またはアンチセンスDNA類似体である。

【0334】

RNAアンチセンス鎖には、触媒的なものであってもよいし、非触媒的なものであってもよい。触媒的なアンチセンス鎖(リボザイムとも呼ばれる)は、RNA分子を特定の配列で切断する。非触媒的なRNAアンチセンス鎖は、さらなるRNAプロセッシングを阻害する。

10

【0335】

C S L G遺伝子の細胞及び/または組織レベルのアンチセンス調節は、少なくとも1つのアンチセンス化合物、例えば、アンチセンスDNA、アンチセンスRNA、リボザイム、DNAザイム、ロック核酸(LNA)、VIG、アプタマーなどで、生物の細胞または組織を形質転換することによって行うことができる。いくつかの実施形態では、分子は、化学的に修飾されている。他の実施形態では、アンチセンス分子は、アンチセンスDNA、またはアンチセンスDNA類似体である。

20

【0336】

RNAiとは、特定の遺伝子産物を標的とし、その遺伝子の転写後のサイレンシングをもたらすために、相同性を有する二本鎖RNA(dsRNA)を導入することを指す。この現象は、シノラブディス・エレガンスにおいて、「Guo and Kemphues (1995, Cell, 81(4):611-620)」によって最初に報告され、その後、「Fire et al. (1998, Nature 391:806-811)」によって、インビトロRNA prepに存在するセンス鎖とアンチセンス鎖とがアニーリングして形成されたdsRNAの存在が、干渉活性が生じる原因であることが発見された。

【0337】

植物と動物の両方において、RNAiは、サイレンシングトリガーと相同なメッセンジャーRNAを破壊する配列特異的な多成分ヌクレアーゼであるRNA誘導サイレンシング複合体(RISC)によって媒介される。RISCは、二本鎖RNAトリガーに由来する短鎖RNA(約22ヌクレオチド)を含むことが知られている。この短鎖ヌクレオチドRNAの配列は、抑制される標的遺伝子と相同である。このように、短鎖ヌクレオチド配列は、多成分ヌクレアーゼであるRISCに特定のmRNAの破壊を指示するガイド配列として機能しているように思われる。

30

【0338】

RNAiを開始するために使用されるdsRNAは、天然ソースから単離してもよいし、または、DNAから転写するなどの既知の手段を用いて生成してもよい。標的配列に対するRNAi分子を生成するためのプラスミド及びベクターは、以下に例示するように、現在、容易に入手可能である。

40

【0339】

dsRNAは、2つの別個の鎖としてベクターから転写することができる。他の実施形態では、dsRNAを生成するために使用されるDNAの2つの鎖は、それぞれが少なくとも部分的に相補的な配列のDNA鎖と形成する、同一のまたは2つの異なる二重鎖に属し得る。このようにしてdsRNAが生成されると、転写されるDNA配列は2つのプロモータに挟まれ、一方のプロモータは一方の鎖の転写を制御し、他方のプロモータは相補的な他方の鎖の転写を制御する。これらの2つのプロモータは、互いに同一であってもよいし、互いに異なってもよい。あるいは、単一のプロモータが、自己相補的なセンス領域とアンチセンス領域とを有する一本鎖のヘアピンポリヌクレオチドの転写を誘導し、

50

それらをアニールして2本鎖RNAを生成することもできる。

#### 【0340】

阻害は、RNAの二重鎖領域に対応するヌクレオチド配列が遺伝的抑制の標的とされる点で、配列特異的である。標的遺伝子の一部と同一のヌクレオチド配列を含むRNA分子が、阻害のために好ましい。また、標的配列に相対する挿入、欠失、及び単一点突然変異を有するRNA配列も、阻害に有効であることが分かっている。したがって、配列相同性は、当分野で既知の配列比較及びアラインメントアルゴリズムによって最適化することができる（「Gribskov and Devereux, Sequence Analysis Primer, Stockton Press, 1991」を参照）。そして、例えば、デフォルトパラメータを用いて、BESTFITソフトウェアプログラムに実装されているSmith-Watermanアルゴリズム（例えば、  
「University of Wisconsin Genetic Computing Group」）によって、ヌクレオチド配列間の差異のパーセントを計算することができる。阻害RNAと標的遺伝子の部分とは、90%以上の配列相同性、またはさらには、100%の配列相同性があることが好ましい。あるいは、RNAの二本鎖領域は、標的遺伝子転写産物の一部とハイブリダイズすることができるヌクレオチド配列として機能的に定義してもよい。同一のヌクレオチド配列の長さは、少なくとも25、50、100、200、300または400塩基であり得る。使用することができるdsRNAの長さに上限はない。例えば、dsRNAの長さは、遺伝子の約21塩基対(bp)から遺伝子の全長またはそれ以上までの範囲をとることができる。

10

#### 【0341】

「RNA干渉」または「RNAi」は、低分子二本鎖RNAによって媒介される遺伝子発現のサイレンシングまたは低減を意味する。これは、動物や植物における、配列特異的な転写後の遺伝子サイレンシングのプロセスであり、その二重鎖領域における、サイレンシングされた遺伝子の配列と相同な阻害性RNA(iRNA)によって開始される。この遺伝子は、生物に対して内在性であっても外因性であってもよく、染色体に組み込まれていてもよく、または、ゲノムに組み込まれていないトランスフェクションベクターに存在していてもよい。遺伝子の発現は、完全にまたは部分的に阻害される。RNAiはまた、標的RNAの機能の全体または一部を阻害すると考えられる。

20

#### 【0342】

当業者であれば、RNAi分子という用語が、センス配列とアンチセンス配列との両方を含む一本鎖または二本鎖のRNA分子を指すことを理解できるであろう。例えば、RNA干渉分子は、自己相補的なセンス領域とアンチセンス領域とを含む二本鎖ポリヌクレオチド分子であってよく、アンチセンス領域は、標的核酸分子に対する相補性を有する。あるいは、RNAi分子は、自己相補的なセンス領域と、標的核酸分子に対する相補性を有するアンチセンス領域とを有する一本鎖ヘアピンポリヌクレオチドであってよい。または、RNAi分子は、2以上のループ構造、並びに、自己相補的なセンス領域と、標的核酸分子に対する相補性を有するアンチセンス領域とを含むステムを有する環状一本鎖ポリヌクレオチドであってよい。この環状一本鎖ポリヌクレオチドは、インビボまたはインビトロで処理することにより、RNAiを媒介可能な活性分子を生成することができる。

30

#### 【0343】

いくつかの実施形態では、RNA干渉(RNAi)を使用してCSLG遺伝子の発現をダウンレギュレートすることにより、植物またはその細胞またはその一部において、そのレベルを低減させることができる。植物と動物の両方において、RNAiは、サイレンシングトリガーと相同なメッセンジャーRNAを破壊する配列特異的な多成分ヌクレアーゼであるRNA誘導サイレンシング複合体(RISC)によって媒介される。RISCは、二本鎖RNAトリガーに由来する短鎖RNA(約22ヌクレオチド)を含むことが知られている。この短鎖ヌクレオチドRNAの配列は、抑制される標的遺伝子と相同である。このように、短鎖ヌクレオチド配列は、多成分ヌクレアーゼであるRISCに特定のmRNAの破壊を指示するガイド配列として機能しているように思われる。

40

#### 【0344】

50

RNAiを開始するために使用されるdsRNAは、天然ソースから単離してもよいし、または、DNAから転写するなどの既知の手段を用いて生成してもよい。標的配列に対するRNAi分子を生成するためのプラスミド及びベクターは、以下に例示するように、現在、容易に入手可能である。

【0345】

いくつかの例示的な実施形態では、dsRNAは、CSLGを標的とする。

【0346】

いくつかの実施形態では、RNA干渉(RNAi)を使用してCSLG遺伝子の発現をダウンレギュレートすることにより、植物またはその細胞またはその一部において、そのレベルを低減させることができる。植物と動物の両方において、RNAiは、サイレンシングトリガーと相同なメッセンジャーRNAを破壊する配列特異的な多成分ヌクレアーゼであるRNA誘導サイレンシング複合体(RISC)によって媒介される。RISCは、二本鎖RNAトリガーに由来する短鎖RNA(約22ヌクレオチド)を含むことが知られている。この短鎖ヌクレオチドRNAの配列は、抑制される標的遺伝子と相同である。このように、短鎖ヌクレオチド配列は、多成分ヌクレアーゼであるRISCに特定のmRNAの破壊を指示するガイド配列として機能しているように思われる。

10

【0347】

RNAiを開始するために使用されるdsRNAは、天然ソースから単離してもよいし、または、DNAから転写するなどの既知の手段を用いて生成してもよい。標的配列に対するRNAi分子を生成するためのプラスミド及びベクターは、以下に例示するように、現在、容易に入手可能である。

20

【0348】

dsRNAは、2つの別個の鎖としてベクターから転写することができる。他の実施形態では、dsRNAを生成するために使用されるDNAの2つの鎖は、それぞれが少なくとも部分的に相補的な配列のDNA鎖と形成する、同一のまたは2つの異なる二重鎖に属し得る。このようにしてdsRNAが生成されると、転写されるDNA配列は2つのプロモータに挟まれ、一方のプロモータは一方の鎖の転写を制御し、他方のプロモータは相補的な他方の鎖の転写を制御する。これらの2つのプロモータは、互いに同一であってもよし、互いに異なってもよい。あるいは、単一のプロモータが、自己相補的なセンス領域とアンチセンス領域とを有する一本鎖のヘアピンポリヌクレオチドの転写を誘導し、それらをアニールして二本鎖RNAを生成することもできる。

30

【0349】

阻害は、RNAの二重鎖領域に対応するヌクレオチド配列が遺伝的抑制の標的とされる点で、配列特異的である。標的遺伝子の一部と同一のヌクレオチド配列を含むRNA分子が、阻害のために好ましい。また、標的配列に相対する挿入、欠失、及び単一点突然変異を有するRNA配列も、阻害に有効であることが分かっている。したがって、配列相同性は、当分野で既知の配列比較及びアラインメントアルゴリズムによって最適化することができる(「Gribkov and Devereux, Sequence Analysis Primer, Stockton Press, 1991」を参照)。そして、例えば、デフォルトパラメータを用いて、BESTFITソフトウェアプログラムに実装されているSmith-Watermanアルゴリズム(例えば、「University of Wisconsin Genetic Computing Group」)によって、ヌクレオチド配列間の差異のパーセントを計算することができる。阻害RNAと標的遺伝子の部分とは、90%以上の配列相同性、またはさらには、100%の配列相同性があることが好ましい。あるいは、RNAの二本鎖領域は、標的遺伝子転写産物の一部とハイブリダイズすることができるヌクレオチド配列として機能的に定義してもよい。同一のヌクレオチド配列の長さは、少なくとも25、50、100、200、300または400塩基であり得る。使用することができるdsRNAの長さには上限はない。例えば、dsRNAの長さは、遺伝子の約21塩基対(bp)から遺伝子の全長またはそれ以上までの範囲をとることができる。

40

【0350】

50

C S L G の発現をダウンレギュレートすることができる別の薬剤は、共抑制分子である。共抑制 (co-suppression) とは、転写後のメカニズムであり、導入遺伝子と内在性遺伝子との両方がサイレンシングされることである。

【 0 3 5 1 】

C S L G の発現をダウンレギュレートすることができる別の薬剤は、C S L G の m R N A 転写産物または D N A 配列を特異的に切断することができる D N A 酵素分子である。D N A 酵素は、一本鎖と二本鎖の標的配列の両方を切断することができる一本鎖ポリヌクレオチドである。D N A 酵素の一般的なモデル (「 1 0 - 2 3 」モデル) が提案されている。「 1 0 - 2 3 」D N A 酵素は、1 5 個のデオキシリボヌクレオチドを含む触媒ドメインを有し、その両側に、それぞれ 7 ~ 9 個のデオキシリボヌクレオチドを含む 2 つの基質認識ドメインが隣接している。このタイプの D N A 酵素は、基質である R N A をプリン：ピリミジン結合部で効果的に切断することができる (D N A 酵素についてのレビューは、「Khachigian, L. M. (2002) *Curr Opin Mol Ther* 4, 119-121」を参照されたい)。

10

【 0 3 5 2 】

一本鎖及び二本鎖の標的切断部位を認識する、操作された人工 D N A 酵素の構築及び増幅の例は、特許文献 2 2 に開示されている。

【 0 3 5 3 】

「酵素核酸分子」または「酵素オリゴヌクレオチド」という用語は、基質結合領域において、特定の遺伝子標的に対して相補性を有し、かつ、C S L G の標的 R N A を特異的に切断して遺伝子をサイレンシングする酵素活性を有する核酸分子を指す。相補性領域は、酵素核酸分子の標的 R N A への十分なハイブリダイゼーション及びその後の切断を可能にする。酵素核酸という用語は、例えば、リボザイム、触媒 R N A、酵素 R N A、触媒 D N A、アプタザイムまたはアプタマー結合リボザイム、触媒オリゴヌクレオチド、ヌクレオザイム、D N A ザイム、R N A 酵素と互換的に使用される。本願に記載される特定の酵素核酸分子は、それに限定されるものではなく、本明細書に開示される酵素核酸分子は、1 以上の標的核酸領域に相補的な特定の基質結合部位を必要とし、その基質結合部位内またはその周囲に、分子に対して核酸切断及び/または連結活性を付与するヌクレオチド配列を有する。特許文献 2 3 には、そのような分子の例が開示されている。

20

【 0 3 5 4 】

いくつかの実施形態では、C S L G 遺伝子またはそれを標的とするサイレンシング分子は、その遺伝子またはそのサイレンシング分子の発現に必要なすべての要素を含む発現ベクターの一部を形成する。いくつかの実施形態では、発現は、構成的プロモータによって制御される。いくつかの実施形態では、構成的プロモータは、植物組織に特異的である。これらの実施形態では、組織特異的プロモータは、根特異的プロモータ、塊茎特異的プロモータ、葉特異的プロモータ、及び果実特異的プロモータからなる群から選択される。根特異的プロモータは、例えば、「Martinez, E. et al. 2003. *Curr. Biol.* 13:1435-1441」に記載されている。果実特異的プロモータは、とりわけ、「Estornell L.H et al. 2009」、*Plant Biotechnol. J.* 7:298-309、及び「Fernandez A. I. Et al. 2009 *Plant Physiol.* 151:1729-1740」に記載されている。塊茎特異的プロモータは、例えば、「Rocha-Sosa M, et al., 1989. *EMBO J.* 8:23-29」及び「McKibbin R.S. et al., 2006. *Plant Biotechnol J.* 4(4):409-18」に記載されている。葉特異的なプロモータは、例えば、「Yutao Yang, Guodong Yang, Shijuan Liu, Xingqi Guo and Chengchao Zheng. *Science in China Series C: Life Sciences.* 46: 651-660」に記載されている。

30

40

【 0 3 5 5 】

いくつかの実施形態では、発現ベクターは、3´非コード配列に制御要素をさらに含む。本明細書で使用するとき、「3´非コード配列」は、コード配列の下流に位置する D N A 配列を指し、ポリアデニル化認識配列、または、m R N A プロセッシングまたは遺伝子発現に影響を与える制御シグナルをコードする他の配列を含む。ポリアデニル化シグナルは、通常、m R N A 前駆体の 3´末端へのポリアデニル酸トラクトの付加に影響を与える。異なる 3´非コード配列の使用は、「Ingelbrecht I L et al. 1989. *Plant Cell* 1:671-

50

680」によって例示されている。

【0356】

当業者であれば、本明細書に記載された核酸配列及び形質転換ベクターの様々な構成要素が、核酸または核酸断片の発現をもたらすように、作動可能に連結されることを理解できるであろう。本明細書に開示される構築物及びベクターの構成要素を作動可能に連結するための技術は、当業者によく知られている。そのような技術としては、例えば、1以上の制限酵素部位を含む合成リンカーなどのリンカーの使用が挙げられる。

【0357】

変異CSLG遺伝子、遺伝子を標的とするサイレンシング分子の存在、及び/または遺伝子の過剰発現の検出は、当業者に既知の分子遺伝学の標準的な方法を用いて行われる。

10

【0358】

遺伝子またはサイレンシング分子の発現を測定するためには、核酸が発現する器官からcDNAまたはmRNAを取得する必要がある。その試料は、検出プロセスの前に、さらに処理してもよい。例えば、細胞または組織の試料中のポリヌクレオチドを、試料の他の構成要素から分離してもよいし、または、増幅してもよい。或る生物から取得したすべての試料は、任意の種類さらなる処理に供したものも含めて、その生物から取得したものと見なされる。

【0359】

遺伝子やサイレンシング分子の検出には、通常、候補となる改変された生物から取得したポリヌクレオチドの増幅を必要とする。DNAを増幅する方法は、当業者に知られている。DNAを増幅する方法は、当業者に知られている。最も一般的に使用されるDNA増幅方法は、PCR法である（ポリメラーゼ連鎖反応；例えば、「PCR Basics: from background to Bench, Springer Verlag, 2000; Eckert et al., 1991. PCR Methods and Applications 1:17」を参照）。別の適切な増幅方法としては、リガーゼ連鎖反応（LCR）、転写増幅及び自己持続型配列複製、及び核酸ベースの配列増幅（NASBA）が挙げられる。

20

【0360】

いくつかの実施形態では、CSLG遺伝子またはそのサイレンシング分子を含む核酸配列は、選択可能なマーカーをコードする核酸配列をさらに含む。いくつかの実施形態では、選択可能なマーカーは、抗生物質または除草剤に対する耐性を付与し、これらの実施形態では、トランスジェニック細胞または植物は、抗生物質または除草剤に対する耐性にしたがって選択される。

30

【0361】

いくつかの実施形態では、内在性CSLG遺伝子が選択的に過剰発現され、改変された含有量は、対応する非組換え植物と比較した、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量の増加を含む。

【0362】

CSLG遺伝子の過剰発現は、当業者に知られているような任意の方法によって生じさせることができる。いくつかの実施形態では、本明細書に開示されるのは、過剰発現され、少なくともCSLG酵素をコードする少なくとも1つの転写可能なポリヌクレオチドを含む少なくとも1つの細胞を含有する遺伝子組換え植物またはその一部であり、遺伝子組換え植物またはその一部は、対応する非遺伝子組換え植物と比較して、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加する。

40

【0363】

いくつかの実施形態では、本明細書に開示されるポリヌクレオチドは、植物細胞内でのそれらの発現を可能にするDNA構築物に組み込まれる。植物での使用に適したDNA構築物は、当業者に知られている。一実施形態では、DNA構築物は、プロモータ、エンハンサー、複製起点、転写終結配列、ポリアデニル化シグナルなどからなる群から選択される少なくとも1つの発現調節要素を含む。

50

## 【0364】

本明細書に開示されるDNA構築物は、達成すべき結果に応じて設計される。

## 【0365】

作物植物では、その食用部分（これに限定しないが、例えば、果実、種子、根、葉、塊茎など）において、有毒なステロイド系アルカロイドまたはステロイド系グリコアルカロイド、またはそれらの誘導體、代謝物、生合成中間体の含有量を減少させることが望まれている。一方、植物の非食用部分（これに限定しないが、例えば、非食用の根や葉）において、有毒なステロイド系グリコアルカロイドまたはその生合成中間体の含有量を豊富にすることは、広範囲の病原体に対する植物の抵抗性に寄与する。例示的な実施形態では、ステロイド系グリコアルカロイドを過剰に発現させた植物は、製薬産業においてそれらを産生するために使用することができる。

10

## 【0366】

作物植物では、ステロイド系サポニンの調節も望まれている。当業者であれば、いくつかの実施形態では、ステロイド系サポニンを調節することにより、ある化合物から、これに限定しないが、例えば、食品、化粧品、医薬品業界向けの、高価値の製品を作り出すことができることを理解できるであろう。いくつかの実施形態では、ステロイド系サポニンの調節により、病原体に対する植物の保護、これに限定しないが、例えば、細菌、昆虫、ウイルスに対する保護に役立つ化合物が植物中に産生される。いくつかの実施形態では、ステロイド系サポニンの調節は、植物から抗栄養化合物を除去することである。

20

## 【0367】

作物植物では、食用部分（これに限定しないが、例えば、果実や種子など）において、毒性、苦味、ホルモン模倣トリテルペノイド系サポニンやその生合成中間体を低減させることが望まれている。一方、甘味のあるトリテルペノイド系サポニンまたはその生合成中間体の含有量を豊富にすることは有益であり、このことは、天然甘味料の産生に利用することができる。QS-21などのトリテルペノイド系サポニンの含有量を増加させた植物は、製薬業界向けのアジュバントの産生に利用することができる。

## 【0368】

さらに別の実施形態では、本明細書に開示されるのは、少なくともCSLG遺伝子の発現が増強された遺伝子組換え植物または遺伝子編集植物であり、その遺伝子組換え植物または遺伝子編集植物は、対応する非組換え植物または非編集植物と比較して、少なくとも1つのステロイド系アルカロイド、少なくとも1つのステロイド系サポニン、または少なくとも1つのトリテルペノイド系サポニンの量が增加する。さらに別の実施形態では、本明細書に開示されるのは、少なくともCSLG遺伝子の発現が減少した遺伝子組換え植物または遺伝子編集植物であり、その遺伝子組換え植物または遺伝子編集植物は、対応する非組換え植物または非編集植物と比較して、少なくとも1つのステロイド系アルカロイドの量が減少し、かつ、少なくとも1つのステロイド系サポニンの量が減少し、少なくとも1つのトリテルペノイド系サポニンの量が減少する。

30

## 【0369】

本明細書及び後述の実施例では、ステロイド系アルカロイド、ステロイド系サポニン、及び/またはトリテルペノイド系サポニンの生合成経路に関与する酵素または他のタンパク質をコードする少なくとも1つの異種遺伝子（これに限定しないが、例えば、異種CSLG）を提供するか、または、トリテルペノイド合成経路の内在性遺伝子（これに限定しないが、例えば、CSLG遺伝子）を改変して、コードされている酵素の発現、活性または安定性を改変することによって、少なくとも1つのステロイド系アルカロイド、その誘導體、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導體、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導體、その代謝物、またはその生合成中間体のレベル（含有量）を改変することができる。

40

## 【0370】

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、または

50

トリテルペノイド系サポニンの誘導体は、それぞれ、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンのグリコシル化誘導体を含む。

【0371】

いくつかの実施形態では、植物細胞において、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が改変される。いくつかの実施形態では、藻類細胞において、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が改変される。いくつかの実施形態では、酵母において、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が改変される。いくつかの実施形態では、昆虫細胞において、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が改変される。いくつかの実施形態では、細菌において、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が改変される。いくつかの実施形態では、植物において、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が改変される。いくつかの実施形態では、植物の一部において、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が改変される。

10

20

30

【0372】

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの誘導体は、それぞれ、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンのグリコシル化誘導体を含む。

【0373】

例えば、いくつかの実施形態では、以下の実施例で例示するように、サポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、UDP - グルコース 6 - デヒドロゲナーゼ 1、またはそれらの任意の組み合わせの発現量を増加させた遺伝子組換え細胞（例えば、酵母細胞、植物細胞、藻類細胞、昆虫細胞、細菌）を含む遺伝子組換え植物またはその一部が使用される。他の実施形態では、実施例に示すように、サポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、及び UDP - グルコース 6 - デヒドロゲナーゼ 1 の発現を低減させるかまたはサイレンシングすると、対応するトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体のレベルが低下することが示されている。

40

【0374】

ステロイド系アルカロイド、その誘導体、その代謝物、及びその生合成中間体

50

## 【0375】

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が増加する。いくつかの実施形態では、本明細書に開示される少なくとも1つの遺伝子組換え細胞を含む遺伝子組換え植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が増加する。いくつかの実施形態では、インビトロ生合成系は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせを産生する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が増加する。いくつかの実施形態では、本明細書に開示される少なくとも1つの遺伝子組換え細胞を含む遺伝子組換え植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が減少する。いくつかの実施形態では、本明細書に開示される少なくとも1つの遺伝子組換え細胞を含む遺伝子組換え植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が増加し、かつ、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が減少する。

10

## 【0376】

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞においてステロイド系アルカロイドを産生する方法は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせを産生することを含む。いくつかの実施形態では、遺伝子組換え植物またはその一部においてステロイド系アルカロイドを産生する方法は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせを産生することを含む。いくつかの実施形態では、本明細書に開示されるのは、本明細書に記載されている少なくとも1つの遺伝子組換え細胞を含む植物またはその一部において、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量を減少させる方法である。いくつかの実施形態では、本明細書に開示されるのは、本明細書に記載されているような少なくとも1つの遺伝子組換え細胞を含む植物またはその一部において、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量を増加させる方法である。

20

30

## 【0377】

ステロイド系アルカロイドの生合成経路において、ステロイド系アルカロイドは、フロスタノール型サポニンアグリコンの形成後、ステロイド系サポニンから分岐する(図1)。酸化、転移、環形成などを少なくとも部分的に伴うさらなる段階を経て、ステロイド系アルカロイドが生合成される(図1)。

## 【0378】

ステロイド系グリコアルカロイドの例としては、これに限定しないが、 $\beta$ -トマチン、トマチン、デヒドロトマチン、 $\beta$ -シャオニン、 $\beta$ -ソラニン、 $\beta$ -ソラソニン、 $\beta$ -ソルマルギン、またはそれらの任意の組み合わせが挙げられる。

40

## 【0379】

いくつかの実施形態では、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させた遺伝子組換え細胞は、植物細胞、酵母細胞、藻類細胞、昆虫細胞、または細菌を含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させた遺伝子組換え細胞は、植物細胞を含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその

50

生合成中間体の含有量を増加させた遺伝子組換え細胞は、酵母細胞を含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させた遺伝子組換え細胞は、藻類細胞を含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させた遺伝子組換え細胞は、昆虫細胞を含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させた遺伝子組換え細胞は、細菌を含む。

【0380】

いくつかの例示的な実施形態では、下流のステロイド系グリコアルカロイドは、エスクレオシド及びデヒドロエスクレオシドからなる群から選択される。

10

【0381】

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、 - トマチン、トマチン、デヒドロトマチン、 - カコニン、 - ソラニン、 - ソラソニン、 - ソルマルギン、またはそれらの任意の組み合わせから選択される少なくとも1つのステロイド系アルカロイドの含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、 - トマチンの含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、トマチンの含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、デヒドロトマチンの含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、 - カコニンの含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、 - ソラニンの含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、 - ソラソニンの含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、 - ソルマルギンの含有量が増加する。

20

【0382】

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、 - トマチン、トマチン、デヒドロトマチン、 - カコニン、 - ソラニン、 - ソラソニン、 - ソルマルギン、またはそれらの任意の組み合わせから選択される少なくとも1つのステロイド系アルカロイドの含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、 - トマチンの含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、トマチンの含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、デヒドロトマチンの含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、 - カコニンの含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、 - ソラニンの含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、 - ソラソニンの含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、 - ソルマルギンの含有量が減少する。

30

【0383】

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、 - トマチン、トマチン、デヒドロトマチン、 - カコニン、 - ソラニン、 - ソラソニン、 - ソルマルギン、またはそれらの任意の組み合わせから選択される少なくとも1つステロイド系アルカロイドの含有量が増加し、かつ、 - トマチン、トマチン、デヒドロトマチン、 - カコニン、 - ソラニン、 - ソラソニン、 - ソルマルギン、またはそれらの任意の組み合わせから選択される少なくとも1つのステロイド系アルカロイドの含有量が減少する。

40

【0384】

いくつかの実施形態では、遺伝子組換え細胞においてステロイド系アルカロイドを産生する方法は、 - トマチン、トマチン、デヒドロトマチン、 - カコニン、 - ソラニン、 - ソラソニン、 - ソルマルギン、またはそれらの任意の組み合わせから選択される少なくとも1つのステロイド系アルカロイドを産生することを含む。いくつかの実施形態

50

では、遺伝子組換え細胞においてステロイド系アルカロイドを産生する方法は、  
 - トマチンを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてステロイド系アルカロイドを産生する方法は、トマチンを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてステロイド系アルカロイドを産生する方法は、デヒドロトマチンを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてステロイド系アルカロイドを産生する方法は、  
 - カコニンを生産することを含む。いくつかの実施形態では、遺伝子組換え細胞においてステロイド系アルカロイドを産生する方法は、  
 - ソラニンを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてステロイド系アルカロイドを産生する方法は、  
 - ソラソニンを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてステロイド系アルカロイドを産生する方法は、  
 - ソルマルギンを産生することを含む。

10

## 【0385】

いくつかの実施形態では、遺伝子組換え細胞において少なくとも1つのステロイド系アルカロイドを産生する方法は、植物細胞、酵母細胞、藻類細胞、昆虫細胞、または細菌において少なくとも1つのステロイド系アルカロイドを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞において少なくとも1つのステロイド系アルカロイドを産生する方法は、植物細胞において少なくとも1つのステロイド系アルカロイドを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞において少なくとも1つのステロイド系アルカロイドを産生する方法は、酵母細胞において少なくとも1つのステロイド系アルカロイドを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞において少なくとも1つのステロイド系アルカロイドを産生する方法は、藻類細胞において少なくとも1つのステロイド系アルカロイドを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞において少なくとも1つのステロイド系アルカロイドを産生する方法は、昆虫細胞において少なくとも1つのステロイド系アルカロイドを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞において少なくとも1つのステロイド系アルカロイドを産生する方法は、細菌において少なくとも1つのステロイド系アルカロイドを産生することを含む。いくつかの実施形態では、植物細胞は、植物またはその一部を含む。

20

## 【0386】

いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、  
 - トマチン、トマチン、デヒドロトマチン、  
 - カコニン、  
 - ソラニン、  
 - ソラソニン、  
 - ソルマルギン、またはそれらの任意の組み合わせから選択される少なくとも1つのステロイド系アルカロイドを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、  
 - トマチンを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、トマチンを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、デヒドロトマチンを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、  
 - カコニンを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、  
 - ソラニンを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、  
 - ソラソニンを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、  
 - ソルマルギンを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、コレステロールを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、  
 22 - ヒドロキシコレステロールを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、  
 22、26 - ジヒドロキシコレステロールを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法

30

40

50

は、フロスタノール型サポニンアグリコンを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、フロスタノール - 26 - アルデヒドを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、26 - アミノフロスタノールを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、トマチデノールを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、トマチジン産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、トマチジンガラクトシドを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、トマチン産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系アルカロイドを産生する方法は、トマチン産生することを含む。

10

**【0387】**

いくつかの実施形態では、インビトロ翻訳系においてステロイド系アルカロイドを産生する方法は、トマチン、デヒドロトマチン、トマチン、カコニン、ソラニン、ソラソニン、ソルマルギン、またはそれらの任意の組み合わせから選択される少なくとも1つのステロイド系アルカロイドを産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてステロイド系アルカロイドを産生する方法は、トマチン産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてステロイド系アルカロイドを産生する方法は、デヒドロトマチン産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてステロイド系アルカロイドを産生する方法は、カコニン産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてステロイド系アルカロイドを産生する方法は、ソラニンを産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてステロイド系アルカロイドを産生する方法は、ソラソニンを産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてステロイド系アルカロイドを産生する方法は、ソルマルギンを産生することを含む。

20

**【0388】**

いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を減少させる方法は、トマチン、デヒドロトマチン、トマチン、カコニン、ソラニン、ソラソニン、ソルマルギン、またはそれらの任意の組み合わせから選択される少なくとも1つのステロイド系アルカロイドの含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を減少させる方法は、トマチンの含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を減少させる方法は、トマチンの含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を減少させる方法は、デヒドロトマチンの含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を減少させる方法は、カコニンの含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を減少させる方法は、ソラニンの含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を減少させる方法は、ソラソニンの含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を減少させる方法は、ソルマルギンの含有量を減少させることを含む。

30

40

**【0389】**

いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を増加させる方法は、トマチン、デヒドロトマチン、トマチン、カコニン、ソラニン、ソラソニン、ソルマルギン、またはそれらの任意の組み合わせから選択され

50

る少なくとも1つのステロイド系アルカロイドの含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を増加させる方法は、  
- トマチンの含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を増加させる方法は、トマチンの含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を増加させる方法は、デヒドロトマチンの含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を増加させる方法は、  
- カコニンの含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を増加させる方法は、  
- ソラニンの含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を増加させる方法は、  
- ソラソニンの含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を増加させる方法は、  
- ソルマルギンの含有量を増加させることを含む。

10

20

30

40

50

**【0390】**

いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドと、少なくとも1つのステロイド系アルカロイドの生合成中間体との両方の含有量が改変される。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドと、少なくとも1つのステロイド系アルカロイドの生合成中間体との両方の含有量が増加する。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドと、少なくとも1つのステロイド系アルカロイドの生合成中間体との両方の含有量が減少する。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量が増加し、かつ、少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が減少する。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量が減少し、かつ、少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が増加する。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量を測定可能に改変することなく、少なくとも1つのステロイド系アルカロイドの含有量が改変される。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を測定可能に改変することなく、少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が改変される。

**【0391】**

「中間体」という用語は、いくつかの実施形態では、「生合成中間体」という用語と互換的に使用することができ、すべて同一の性質及び意味を有する。

**【0392】**

いくつかの実施形態では、ステロイド系アルカロイドの生合成中間体は、コレステロール、22-ヒドロキシコレステロール、22, 26-ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、フロスタノール-26-アルデヒド、26-アミノフロスタノール、トマチデノール、トマチジン、トマチジンガラクトシド、  
- トマチン、  
- 1-トマチン、またはそれらの任意の組み合わせを含む(図1参照)。

**【0393】**

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、コレステロール、22-ヒドロキシコレステロール、22, 26-ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、フロスタノール-26-アルデヒド、26-アミノフロスタノール、トマチデノール、トマチジン、トマチジンガラクトシド、  
- トマチン、  
- 1-トマチン、またはそれらの任意の組み合わせを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、コレステロールを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、22-ヒドロキシコレステロールを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、22, 26-ジヒドロキシコレステロールを含む少なくとも1つのステロイド系アルカロイドの生合成中間体

の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、フロスタノール型サボニンアグリコンを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、フロスタノール - 26 - アルデヒドを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、26 - アミノフロスタノールを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、トマチデノールを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、トマチジンを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、トマチジンガラクトシドを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、トマチンを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、トマチン - 1 - トマチンを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が増加する。

10

20

30

40

50

【0394】

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、コレステロール、22 - ヒドロキシコレステロール、22, 26 - ジヒドロキシコレステロール、フロスタノール型サボニンアグリコン、フロスタノール - 26 - アルデヒド、26 - アミノフロスタノール、トマチデノール、トマチジン、トマチジンガラクトシド、トマチン、トマチン - 1 - トマチン、またはそれらの任意の組み合わせを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え植物は、コレステロールを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え植物は、22 - ヒドロキシコレステロールを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え植物は、22, 26 - ジヒドロキシコレステロールを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え植物は、フロスタノール型サボニンアグリコンを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え植物は、フロスタノール - 26 - アルデヒドを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え植物は、26 - アミノフロスタノールを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え植物は、トマチデノールを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え植物は、トマチジンを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え植物は、トマチジンガラクトシドを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え植物は、トマチンを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え植物は、トマチン - 1 - トマチンを含む少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が減少する。

【0395】

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が増加し、かつ、少なくとも1つのステロイド系アルカロイドの生合成中間体の含有量が減少する。少なくとも1つのステロイド系アルカロイドの生合成中間体は、コレステロール、22 - ヒドロキシコレステロール、22, 26 - ジヒドロキシコレステロール、フロスタノール型サボニンアグリコン、フロスタノール - 26 - アルデヒド、26 - アミノフロスタノール、トマチデノール、トマチジン、トマチジンガラクトシド、トマチン、トマチン - 1 - トマチン、またはそれらの任



系アルカロイドの含有量を増加させる方法は、少なくともトマチデノールの含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を増加させる方法は、トマチジンの含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を増加させる方法は、少なくともトマチジンガラクトシドの含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を増加させる方法は、少なくともトマチンの含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドの含有量を増加させる方法は、少なくとも1 - トマインの含有量を増加させることを含む。

**【0398】**

いくつかの実施形態では、ステロイド系アルカロイドの生合成中間体は、コレステロール、22 - ヒドロキシコレステロール、22, 26 - ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、フロスタノール - 26 - アルデヒド、26 - アミノフロスタノール、トマチデノール、トマチジン、トマチジンガラクトシド、トマチン、1 - トマチン、またはそれらの任意の組み合わせを含む。

**【0399】**

予期せぬことに、本開示は、酵素（例えば、ステロイド系アルカロイド生合成経路の酵素）をコードする少なくとも1つの異種遺伝子を発現するように細胞または植物を遺伝的に改変することによって、細胞（例えば、植物細胞、酵母細胞、藻類細胞、昆虫細胞）、細菌、または植物において、ステロイド系アルカロイド、その誘導体、その代謝物、または生合成中間体の含有量を増加させることができることが分かった。他の実施形態では、本明細書に記載及び例示されているのは、植物細胞内の少なくとも1つの内在性遺伝子（これに限定しないが、例えば、CSLG遺伝子）を遺伝子的に改変して、その発現、活性、安定性、またはそれらの任意の組み合わせを調節する方法である。以下の実施例は、ステロイド系アルカロイド生合成経路の一部であることがこれまで知られていなかった酵素活性及び酵素を開示する。この知見がなければ、ステロイド系アルカロイド化合物の産生は不可能であった。

**【0400】**

いくつかの実施形態では、ステロイド系アルカロイド代謝経路は、ステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量が増加した細胞、あるいは、植物またはその一部においてこれらの化合物の含有量が増加した植物を生産することができる。いくつかの実施形態では、有用なステロイド系アルカロイドの含有量が増加した植物、これに限定しないが、例えば、作物植物が生産される。いくつかの実施形態では、本明細書に開示されるのは、ステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量が増加した細胞（これに限定しないが、例えば、植物細胞、酵母、または藻類細胞）または植物（これに限定しないが、作物植物）または植物の一部を生産するための手段及び方法である。代替的または追加的に、本明細書に開示される遺伝子の発現を制御することは、例えば製薬産業においてさらに使用する所望のステロイド系アルカロイドの生産のために、あるいは、食品または他の栄養補助食品（これに限定しないが、例えば、甘味料など）の生産のために使用することができる。いくつかの実施形態では、これらの高価値アルカロイドは、精製され、例えば、甘味料、発泡剤、乳化剤、防腐剤、抗発癌剤、コレステロール低下剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、摂食阻害剤、抗真菌剤、またはそれらの任意の組み合わせとして使用することができる。本明細書に開示される細胞及び植物は、栄養学的、薬学的、及び商業的に重要な価値を有する化合物を含有する。

**【0401】**

いくつかの実施形態では、ステロイド系アルカロイドの誘導体は、ステロイド系アルカロイドのグリコシル化誘導体を含む。

**【0402】**

いくつかの実施形態では、本明細書に記載の遺伝子組換え細胞、あるいは、少なくとも

10

20

30

40

50



含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドのまたはステロイド系アルカロイドの少なくとも5つの代謝産物の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドのまたはステロイド系アルカロイドの少なくとも6つの代謝産物の含有量が増加する。

【0406】

いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの少なくとも1つの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの少なくとも2つの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの少なくとも3つの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの少なくとも4つの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの少なくとも5つの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの少なくとも6つの生合成中間体の含有量が増加する。

10

20

【0407】

当業者であれば、ステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体に関する「含有量」及び「レベル」という用語は、対照細胞または対照植物と比較した、遺伝子組換え細胞または遺伝子組換え植物中の化合物の量、例えば、ステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の量を包含することを認識できるであろう。この文脈では、「含有量」及び「レベル」という用語は、互換的に使用することができ、すべて同一の意味及び性質を有する。

【0408】

いくつかの実施形態では、含有量が増加したステロイド系アルカロイドは、栄養剤、化粧品、医薬品、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、含有量が増加したステロイド系アルカロイドは、  
- トマチン、トマチン、デヒドロトマチン、  
- カコニン、  
- ソラニン、  
- ソラソニン、  
- ソルマルギン、またはそれらの任意の組み合わせから選択されるステロイド系アルカロイドを含む。いくつかの実施形態では、含有量が増加したステロイド系アルカロイドは、  
- トマチン、トマチン、デヒドロトマチン、  
- カコニン、  
- ソラニン、  
- ソラソニン、  
- ソルマルギン、またはそれらの任意の組み合わせから選択される少なくとも1、2、3、4、5、6、またはそれ以上のステロイド系アルカロイドを含む。

30

【0409】

いくつかの実施形態では、含有量が増加したステロイド系アルカロイドは、  
- トマチン、トマチン、デヒドロトマチン、  
- カコニン、  
- ソラニン、  
- ソラソニン、  
- ソルマルギン、またはそれらの任意の組み合わせから選択される少なくとも1、2、3、4、5、6、またはそれ以上のステロイド系アルカロイドを含む。

40

【0410】

いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において含有量が増加したステロイド系アルカロイドは、  
- トマチン、トマチン、デヒドロトマチン、  
- カコニン、  
- ソラニン、  
- ソラソニン、  
- ソルマルギン、またはそれらの任意の組み合わせから選択されるステロイド系アルカロイドを含む。

50

## 【0411】

いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、 $\beta$ -トマチンの含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、 $\beta$ -トマチンの含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、デヒドロトマチンの含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、 $\beta$ -カコニンの含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、 $\beta$ -ソラニンの含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、 $\beta$ -ソラソニンの含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、 $\beta$ -ソルマルギンの含有量が増加する。当業者であれば、いくつかの実施形態では、1つのステロイド系アルカロイドの含有量が増加する場合には、追加のステロイド系アルカロイド、または中間体、またはそれらの組み合わせも同一の時間の細胞内で増加し得ることを理解できるであろう。

10

20

## 【0412】

いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母細胞、藻類細胞）、または遺伝子組換え植物の一部において、ステロイド系アルカロイドの生合成中間体の含有量が増加する。生合成中間体は、コレステロール、22-ヒドロキシコレステロール、22,26-ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、フロスタノール-26-アルデヒド、26-アミノフロスタノール、トマチデノール、トマチジン、トマチジンガラクトシド、 $\beta$ -トマチン、 $\beta$ -1-トマチン、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ステロイド系アルカロイドの少なくとも2つの生合成中間体の含有量が増加する。生合成中間体は、コレステロール、22-ヒドロキシコレステロール、22,26-ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、フロスタノール-26-アルデヒド、26-アミノフロスタノール、トマチデノール、トマチジン、トマチジンガラクトシド、 $\beta$ -トマチン、 $\beta$ -1-トマチン、またはそれらの任意の組み合わせを含む。

30

## 【0413】

いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ステロイド系アルカロイドの生合成中間体の含有量が増加し、生合成中間体は、コレステロールを含む。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ステロイド系アルカロイドの生合成中間体の含有量が増加し、生合成中間体は、22-ヒドロキシコレステロールを含む。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ステロイド系アルカロイドの生合成中間体の含有量が増加し、生合成中間体は、22,26-ジヒドロキシコレステロールを含む。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ステロイド系アルカロイドの生合成中間体の含有量が増加し、生合成中間体は、フロスタノール型サポニンアグリコンを含む。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺

40

50

伝子組換え植物またはその一部において、ステロイド系アルカロイドの生合成中間体の含有量が増加し、生合成中間体は、フロスタノール-26-アルデヒドを含む。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ステロイド系アルカロイドの生合成中間体の含有量が増加し、生合成中間体は、26-アミノフロスタノールを含む。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ステロイド系アルカロイドの生合成中間体の含有量が増加し、生合成中間体は、トマチデノールを含む。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ステロイドアルキロイドの生合成中間体の含有量が増加し、生合成中間体は、トマチジンを含む。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ステロイド系アルカロイドの生合成中間体の含有量が増加し、生合成中間体は、トマチジンガラクトシドを含む。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ステロイド系アルカロイドの生合成中間体の含有量が増加し、生合成中間体は、トマチンを含む。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ステロイド系アルカロイドの生合成中間体の含有量が増加し、生合成中間体は、トマチンを含む。

10

20

## 【0414】

いくつかの実施形態では、少なくとも1つの遺伝子組換え細胞を含む植物またはその一部において含有量が減少したステロイド系アルカロイドは、苦味を有する化合物または毒素を含む。いくつかの実施形態では、含有量が減少したステロイド系アルカロイドは、トマチン、トマチン、デヒドロトマチン、カコニン、ソラニン、ソラソニン、ソルマルギン、またはそれらの任意の組み合わせから選択される少なくとも1、2、3、4、5、6、またはそれ以上のステロイド系アルカロイドを含む。

## 【0415】

いくつかの実施形態では、含有量が減少したステロイド系アルカロイドは、トマチン、トマチン、デヒドロトマチン、カコニン、ソラニン、ソラソニン、ソルマルギン、またはそれらの任意の組み合わせから選択される少なくとも1、2、3、4、5、6、またはそれ以上のステロイド系アルカロイドを含む。

30

## 【0416】

いくつかの実施形態では、少なくとも1つの遺伝子組換え細胞を含む遺伝子組換え植物またはその一部において含有量が減少したステロイド系アルカロイドは、トマチン、トマチン、デヒドロトマチン、カコニン、ソラニン、ソラソニン、ソルマルギン、またはそれらの任意の組み合わせから選択されるステロイド系アルカロイドを含む。

## 【0417】

いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、トマチンの含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、トマチンの含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、デヒドロトマチンの含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、カコニンの含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ソラニンの含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ソラソニンの含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物ま

40

50

たはその一部は、 - ソルマルギンの含有量が減少する。

【0418】

いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの生合成中間体の含有量が減少する。生合成中間体は、コレステロール、22 - ヒドロキシコレステロール、22, 26 - ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、フロスタノール - 26 - アルデヒド、26 - アミノフロスタノール、トマチデノール、トマチジン、トマチジンガラクトシド、 - トマチン、 - 1 - トマチン、またはそれらの任意の組み合わせを含む。

【0419】

いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの生合成中間体の含有量が減少し、生合成中間体は、コレステロールを含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの生合成中間体の含有量が減少し、生合成中間体は、22 - ヒドロキシコレステロールを含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの生合成中間体の含有量が減少し、生合成中間体は、22, 26 - ジヒドロキシコレステロールを含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの生合成中間体の含有量が減少し、生合成中間体は、フルソタノール型サポニンアグリコンを含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの生合成中間体の含有量が減少し、生合成中間体は、フロスタノール - 26 - アルデヒドを含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの生合成中間体の含有量が減少し、生合成中間体は、26 - アミノフロスタノールを含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの生合成中間体の含有量が減少し、生合成中間体は、トマチデノールを含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの生合成中間体の含有量が減少し、生合成中間体は、トマチジンを含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの生合成中間体の含有量が減少し、生合成中間体は、トマチジンガラクトシドを含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの生合成中間体の含有量が減少し、生合成中間体は、 - トマチンを含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系アルカロイドの生合成中間体の含有量が減少し、生合成中間体は、 - 1 - トマチンを含む。

【0420】

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量が増加し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変される。

【0421】

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量が増加し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間

10

20

30

40

50



導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体の含有量が増加する。

【0428】

いくつかの実施形態では、遺伝子組換え細胞または遺伝子組換え植物において、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体を産生する方法は、未組換え細胞または未組換え植物と比較して、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させ、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量を改変することを含む。

10

【0429】

いくつかの実施形態では、遺伝子組換え細胞または遺伝子組換え植物において、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体を産生する方法は、未組換え細胞または未組換え植物と比較して、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させ、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量を減少させることを含む。

20

【0430】

いくつかの実施形態では、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量を減少させる方法は、少なくとも1つの植物またはその一部の少なくとも1つの細胞を遺伝子組換えすることを含み、遺伝子組換えされた植物細胞は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量が減少し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変される。

30

【0431】

いくつかの実施形態では、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量を減少させる方法は、少なくとも1つの植物またはその一部の少なくとも1つの細胞を遺伝子組換えすることを含み、遺伝子組換えされた植物細胞は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量が減少し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が増加する。

40

【0432】

いくつかの実施形態では、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、少なくとも1つの植物またはその一部の少なくとも1つの細胞を遺伝子組換えすることを含み、遺伝子組換えされた植物細胞は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量が改変され、かつ、フィトステロール、その誘導体

50

、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が減少する。

【0433】

いくつかの実施形態では、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、少なくとも1つの植物またはその一部の少なくとも1つの細胞を遺伝子組換えすることを含み、遺伝子組換えされた植物細胞は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量が増加し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が減少する。

10

【0434】

ステロイド系サポニン、その誘導体、代謝物、及びその生合成中間体

【0435】

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が増加する。いくつかの実施形態では、本明細書に開示される少なくとも1つの遺伝子組換え細胞を含む遺伝子組換え植物は、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が増加する。いくつかの実施形態では、インビトロ生合成系は、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせを産生する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が増加する。いくつかの実施形態では、本明細書に開示される少なくとも1つの遺伝子組換え細胞を含む遺伝子組換え植物は、任意の組み合わせのうち、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え細胞を含む本明細書に開示される遺伝子組換え植物は、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が増加し、かつ、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が減少する。

20

30

【0436】

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞においてステロイド系サポニンを産生する方法は、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせを産生することを含む。いくつかの実施形態では、遺伝子組換え植物またはその一部においてステロイド系サポニンを産生する方法は、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせを産生することを含む。いくつかの実施形態では、本明細書に開示されるのは、本明細書に記載されている少なくとも1つの遺伝子組換え細胞を含む植物またはその一部において、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量を減少させる方法である。いくつかの実施形態では、本明細書に開示されるのは、本明細書に記載されているような少なくとも1つの遺伝子組換え細胞を含む植物またはその一部において、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物

40

50

、その生合成中間体、またはそれらの任意の組み合わせの含有量を増加させる方法である。

【0437】

ステロイド系サポニン、フロスタノール型のサポニンアグリコンの形成後、ステロイド系アルカロイドから分岐する(図1)。

【0438】

一般的に使用されているサポニンの命名法では、トリテルペノイド系サポニン(トリテルペンサポニンとも呼ばれる)とステロイド系サポニンを区別しているが、これは、そのアグリコンの構造と生化学的背景に基づいている。どちらのタイプのサポニンも、ステロール生合成の中心的な代謝産物である2,3-オキシドスカレンに由来すると考えられている。フィトステロール同化作用において、2,3-オキシドスカレンは、主にシクロアルテノールに環化される。「ステロイド系サポゲニン」は、シクロアルテノール形成の下流のフィトステロール経路の中間体に由来すると考えられている。

10

【0439】

ステロイド系サポニンの例としては、これに限定しないが、ウットロシドB、トマトリサイド、またはそれらの組み合わせが挙げられる。

【0440】

いくつかの実施形態では、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させた遺伝子組換え細胞は、植物細胞、酵母細胞、藻類細胞、昆虫細胞、または細菌を含む。いくつかの実施形態では、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させた遺伝子組換え細胞は、植物細胞を含む。いくつかの実施形態では、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させた遺伝子組換え細胞は、酵母細胞を含む。いくつかの実施形態では、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させた遺伝子組換え細胞は、藻類細胞を含む。いくつかの実施形態では、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させた遺伝子組換え細胞は、昆虫細胞を含む。いくつかの実施形態では、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させた遺伝子組換え細胞は、細菌を含む。

20

30

【0441】

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ウットロシドB、トマトリサイド、またはそれらの組み合わせから選択される少なくとも1つのステロイド系サポニンの含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ウットロシドBの含有量が増加する。

【0442】

いくつかの実施形態では、ステロイド系サポニンの誘導体は、ステロイド系サポニンのグリコシル化誘導体を含む。

【0443】

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ウットロシドB、トマトリサイド、またはそれらの組み合わせから選択される少なくとも1つのステロイド系サポニンの含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ウットロシドBの含有量が減少する。

40

【0444】

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、少なくとも1つのステロイド系サポニンの含有量が増加し、かつ、ウットロシドB、トマトリサイド、またはそれらの組み合わせから選択される少なくとも1つのステロイド系サポニンの含有量が減少する。

【0445】

いくつかの実施形態では、遺伝子組換え細胞においてステロイド系サポニンを産生する

50

方法は、ウットロシド B、トマトシド、またはそれらの組み合わせから選択される少なくとも 1 つのステロイド系サポニンを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてステロイド系サポニンを産生する方法は、ウットロシド B を産生することを含む。

【0446】

いくつかの実施形態では、遺伝子組換え細胞において少なくとも 1 つのステロイド系サポニンを産生する方法は、植物細胞、酵母細胞、藻類細胞、昆虫細胞、または細菌において少なくとも 1 つのステロイド系サポニンを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞において少なくとも 1 つのステロイド系サポニンを産生する方法は、植物細胞において少なくとも 1 つのステロイド系サポニンを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞において少なくとも 1 つのステロイド系サポニンを産生する方法は、酵母細胞において少なくとも 1 つのステロイド系サポニンを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞において少なくとも 1 つのステロイド系サポニンを産生する方法は、藻類細胞において少なくとも 1 つのステロイド系サポニンを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞において少なくとも 1 つのステロイド系サポニンを産生する方法は、昆虫細胞において少なくとも 1 つのステロイド系サポニンを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞において少なくとも 1 つのステロイド系サポニンを産生する方法は、細菌において少なくとも 1 つのステロイド系サポニンを産生することを含む。いくつかの実施形態では、植物細胞は、植物またはその一部を含む。

10

20

【0447】

いくつかの実施形態では、遺伝子組換え植物においてステロイド系サポニンを産生する方法は、ウットロシド B、トマトリサイド、またはそれらの組み合わせから選択される少なくとも 1 つのステロイド系サポニンを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてステロイド系サポニンを産生する方法は、ウットロシド B を産生することを含む。

【0448】

いくつかの実施形態では、インビトロ翻訳系においてステロイド系サポニンを産生する方法は、ウットロシド B、トマトシド、またはそれらの組み合わせから選択される少なくとも 1 つのステロイド系サポニンを産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてステロイド系サポニンを産生する方法は、ウットロシド B を産生することを含む。

30

【0449】

いくつかの実施形態では、少なくとも 1 つのステロイド系サポニンの含有量を減少させる方法は、ウットロシド B、トマトシド、またはそれらの組み合わせから選択される少なくとも 1 つのステロイド系サポニンの含有量を減少させることを含む。いくつかの実施形態では、少なくとも 1 つのステロイド系サポニンの含有量を減少させる方法は、ウットロシド B の含有量を減少させることを含む。

【0450】

いくつかの実施形態では、少なくとも 1 つのステロイド系サポニンの含有量を増加させる方法は、ウットロシド B、トマトシド、またはそれらの組み合わせから選択される少なくとも 1 つのステロイド系サポニンの含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのステロイド系サポニンは、ウットロシド B を含む。いくつかの実施形態では、少なくとも 1 つのステロイド系サポニンは、トマトシドを含む。

40

【0451】

いくつかの実施形態では、少なくとも 1 つのステロイド系サポニンと、少なくとも 1 つのステロイド系サポニンの生合成中間体との両方の含有量が改変される。いくつかの実施形態では、少なくとも 1 つのステロイド系サポニンと、少なくとも 1 つのステロイド系サポニンの生合成中間体との両方の含有量が増加する。いくつかの実施形態では、少なくとも 1 つのステロイド系サポニンと、少なくとも 1 つのステロイド系サポニンの生合成中間

50

体との両方の含有量が減少する。いくつかの実施形態では、少なくとも1つのステロイド系サポニンの含有量が増加し、かつ、少なくとも1つのステロイド系サポニンの生合成中間体の含有量が減少する。いくつかの実施形態では、少なくとも1つのステロイド系サポニンの含有量が減少し、かつ、少なくとも1つのステロイド系サポニンの生合成中間体の含有量が増加する。いくつかの実施形態では、ステロイド系サポニン中間体の含有量を測定可能に改変することなく、ステロイド系サポニンの含有量が改変される。いくつかの実施形態では、ステロイド系サポニンの含有量を測定可能に改変することなく、ステロイド系サポニン中間体の含有量が改変される。

【0452】

「中間体」という用語は、いくつかの実施形態では、「生合成中間体」という用語と互換的に使用することができ、すべて同一の性質及び意味を有する。

10

【0453】

いくつかの実施形態では、ステロイド系サポニンの生合成中間体は、コレステロール、22-ヒドロキシコレステロール、22, 26-ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、またはそれらの任意の組み合わせを含む。

【0454】

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、コレステロール、22-ヒドロキシコレステロール、22, 26-ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、またはそれらの任意の組み合わせを含む少なくとも1つのステロイド系サポニンの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、コレステロールを含む少なくとも1つのステロイド系サポニンの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、22-ヒドロキシコレステロールを含む少なくとも1つのステロイド系サポニンの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、22, 26-ジヒドロキシコレステロールを含む少なくとも1つのステロイド系サポニンの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、フロスタノール型サポニンアグリコンを含む少なくとも1つのステロイド系サポニンの生合成中間体の含有量が増加する。

20

【0455】

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、コレステロール、22-ヒドロキシコレステロール、22, 26-ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、またはそれらの任意の組み合わせを含む少なくとも1つのステロイド系サポニンの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え植物は、コレステロールを含む少なくとも1つのステロイド系サポニンの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え植物は、22-ヒドロキシコレステロールを含む少なくとも1つのステロイド系サポニンの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え植物は、22, 26-ジヒドロキシコレステロールを含む少なくとも1つのステロイド系サポニンの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え植物は、フロスタノール型サポニンアグリコンを含む少なくとも1つのステロイド系サポニンの生合成中間体の含有量が減少する。

30

40

【0456】

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、少なくとも1つのステロイド系サポニンの生合成中間体の含有量が増加し、かつ、少なくとも1つのステロイド系サポニンの生合成中間体の含有量が減少する。少なくとも1つのステロイド系サポニンの生合成中間体は、コレステロール、22-ヒドロキシコレステロール、22, 26-ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、またはそれらの任意の組み合わせから選択される。

【0457】

いくつかの実施形態では、少なくとも1つのステロイド系サポニンの含有量を減少させ

50

る方法は、コレステロール、22-ヒドロキシコレステロール、22,26-ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、またはそれらの任意の組み合わせから選択される少なくとも1つのステロイド系サポニンの生合成中間体の含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系サポニンの含有量を減少させる方法は、少なくともコレステロールの含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系サポニンの含有量を減少させる方法は、少なくとも22-ヒドロキシコレステロールの含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系サポニンの含有量を減少させる方法は、少なくとも22,26-ジヒドロキシコレステロールを減少させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系サポニンの含有量を減少させる方法は、少なくともフロスタノール型サポニンアグリコンを減少させることを含む。

10

**【0458】**

いくつかの実施形態では、少なくとも1つのステロイド系サポニンの含有量を増加させる方法は、コレステロール、22-ヒドロキシコレステロール、22,26-ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、またはそれらの任意の組み合わせから選択される少なくとも1つのステロイド系サポニンの生合成中間体の含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系サポニンの含有量を増加させる方法は、少なくともコレステロールの含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系サポニンの含有量を増加させる方法は、少なくとも22-ヒドロキシコレステロールの含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系サポニンの含有量を増加させる方法は、少なくとも22,26-ジヒドロキシコレステロールの含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系サポニンの含有量を増加させる方法は、少なくともフロスタノール型サポニンアグリコンの含有量を増加させることを含む。

20

**【0459】**

いくつかの実施形態では、ステロイド系サポニンの生合成中間体は、コレステロール、22-ヒドロキシコレステロール、22,26-ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、またはそれらの任意の組み合わせを含む。

30

**【0460】**

予期せぬことに、本開示は、酵素（例えば、ステロイド系サポニン生合成経路の酵素）をコードする少なくとも1つの異種遺伝子を発現するように細胞または植物を遺伝的に改変することによって、細胞（例えば、植物細胞、酵母細胞、藻類細胞、昆虫細胞）、細菌、または植物において、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させることができることが分かった。他の実施形態では、本明細書に記載及び例示されているのは、植物細胞内の少なくとも1つの内在性遺伝子（これに限定しないが、例えば、CSLG遺伝子）を遺伝子的に改変して、その発現、活性、安定性、またはそれらの任意の組み合わせを調節する方法である。以下の実施例は、ステロイド系サポニン生合成経路の一部であることがこれまで知られていなかった酵素活性及び酵素を開示する。この知見がなければ、ステロイド系サポニン化合物の産生は不可能であった。

40

**【0461】**

いくつかの実施形態では、ステロイド系サポニン代謝経路は、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加した細胞、あるいは、植物またはその一部においてこれらの化合物の含有量が増加した植物を生産することができる。いくつかの実施形態では、有用なステロイド系サポニンの含有量が増加した植物、これに限定しないが、例えば、作物植物が生産される。いくつかの実施形態では、本明細書に開示されるのは、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加した細胞（これに限定しないが、例えば、植物細胞、酵母、または藻類細胞）または植物（これに限定しないが、作物植物）または植物の一部を生産する

50

ための手段及び方法である。代替的または追加的に、本明細書に開示される遺伝子の発現を制御することは、例えば製薬産業においてさらに使用するための所望のステロイド系サポニンの生産のために、あるいは、食品または他の栄養補助食品（これに限定しないが、例えば、甘味料など）の生産のために使用することができる。いくつかの実施形態では、これらの高価値サポニンは、精製され、例えば、甘味料、発泡剤、乳化剤、防腐剤、抗発癌剤、コレステロール低下剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、摂食阻害剤、抗真菌剤、またはそれらの任意の組み合わせとして使用することができる。本明細書に開示される細胞及び植物は、栄養学的、薬学的、及び商業的に重要な価値を有する化合物を含有する。

**【0462】**

いくつかの実施形態では、本明細書に記載の遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、対応する非組換え細胞と比較して、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が改変されている。いくつかの実施形態では、含有量の改変は、含有量の増加を含む。いくつかの実施形態では、例えば、遺伝子組換え細胞または遺伝子組換え植物は、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が増加している。

10

**【0463】**

いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、少なくとも1つのステロイド系サポニンの含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、少なくとも2つのステロイド系サポニンの含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、少なくとも3つのステロイド系サポニンの含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、少なくとも4つのステロイド系サポニンの含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、少なくとも5つのステロイド系サポニンの含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、少なくとも6つのステロイド系サポニンの含有量が増加する。

20

30

**【0464】**

いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの少なくとも1つの誘導体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの少なくとも2つの誘導体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの少なくとも3つの誘導体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの少なくとも4つの誘導体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの少なくとも5つの誘導体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの少なくとも6つの誘導体の含有量が増加する。

40

**【0465】**

いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換

50

え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの少なくとも1つの代謝物の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの少なくとも2つの代謝物の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの少なくとも3つの代謝産物の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの少なくとも4つの代謝産物の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンのまたはステロイド系サポニンの少なくとも5つの代謝産物の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンのまたはステロイド系サポニンの少なくとも6つの代謝産物の含有量が増加する。

10

20

30

40

50

**【0466】**

いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの少なくとも1つの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの少なくとも2つの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの少なくとも3つの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの少なくとも4つの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの少なくとも5つの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの少なくとも6つの生合成中間体の含有量が増加する。

**【0467】**

当業者であれば、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体に関する「含有量」及び「レベル」という用語は、対照細胞または対照植物と比較した、遺伝子組換え細胞または遺伝子組換え植物中の化合物の量、例えば、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量を包含することを認識できるであろう。この文脈では、「含有量」及び「レベル」という用語は、互換的に使用することができ、すべて同一の意味及び性質を有する。

**【0468】**

いくつかの実施形態では、含有量が増加したステロイド系サポニンは、栄養剤、化粧品、医薬品、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、含有量が増加したステロイド系サポニンは、ウットロシドB、トマトシド、またはそれらの組み合わせから選択されるステロイド系サポニンを含む。いくつかの実施形態では、含有量が増加したステロイド系サポニンは、ウットロシドB、トマトシド、またはそれらの組み合わせから選択される少なくとも1、2、3、4、5、6、またはそれ以上のステロイド系サポニンを含む。

**【0469】**

いくつかの実施形態では、含有量が増加したステロイド系サポニンは、ウットロシドB、トマトシド、またはそれらの組み合わせから選択される少なくとも1、2、3、4、5、6、またはそれ以上のステロイド系サポニンを含む。

## 【0470】

いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において含有量が増加したステロイド系サポニン（例えば、ウットロシド B、トマトシド、またはそれらの組み合わせから選択されるステロイド系サポニン）を含む。

## 【0471】

いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ウットロシド B の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、トマトシドの含有量が増加する。

10

## 【0472】

いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ステロイド系サポニンの生合成中間体の含有量が増加し、生合成中間体は、コレステロール、22 - ヒドロキシコレステロール、22, 26 - ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ステロイド系サポニンの少なくとも 2 つの生合成中間体の含有量が増加し、生合成中間体は、コレステロール、22 - ヒドロキシコレステロール、22, 26 - ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、またはそれらの任意の組み合わせを含む。

20

## 【0473】

いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ステロイド系サポニンの生合成中間体の含有量が増加し、生合成中間体は、コレステロールを含む。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ステロイド系サポニンの生合成中間体の含有量が増加し、生合成中間体は、22 - ヒドロキシコレステロールを含む。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ステロイド系サポニンの生合成中間体の含有量が増加し、生合成中間体は、22, 26 - ジヒドロキシコレステロールを含む。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において、ステロイド系サポニンの生合成中間体の含有量が増加し、生合成中間体は、フロスタノール型サポニンアグリコンを含む。

30

## 【0474】

いくつかの実施形態では、少なくとも 1 つの遺伝子組換え細胞を含む植物またはその一部において含有量が減少したステロイド系サポニンは、苦味を有する化合物または毒素を含む。いくつかの実施形態では、含有量が減少したステロイド系サポニンは、ウットロシド B、トマトシド、またはそれらの組み合わせから選択されるステロイド系サポニンを含む。いくつかの実施形態では、含有量が減少したステロイド系サポニンは、ウットロシド B、トマトシド、またはそれらの組み合わせから選択される少なくとも 1、2、3、4、5、6、またはそれ以上のステロイド系サポニンを含む。いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌）、あるいは、遺伝子組換え植物またはその一部において含有量が増加したステロイド系サポニンは、ウットロシド B、トマトシド、またはそれらの組み合わせから選択されるステロイド系サポニンを含む。

40

## 【0475】

いくつかの実施形態では、遺伝子組換え細胞（例えば、植物細胞、酵母、藻類細胞、昆

50

虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において含有量が増加したステロイド系サポニン、ウットロシド B、トマトシド、またはそれらの組み合わせから選択されるステロイド系サポニンを含む。

【0476】

いくつかの実施形態では、少なくとも1つの遺伝子組換え細胞を含む遺伝子組換え植物またはその一部において含有量が減少したステロイド系サポニンは、ウットロシド B、トマトシド、またはそれらの組み合わせから選択されるステロイド系サポニンを含む。

【0477】

いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ウットロシド B の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、トマトシドの含有量が減少する。

10

【0478】

いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの生合成中間体の含有量が減少し、生合成中間体は、コレステロール、22-ヒドロキシコレステロール、22, 26-ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、またはそれらの任意の組み合わせを含む。

【0479】

いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの生合成中間体の含有量が減少し、生合成中間体は、コレステロールを含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの生合成中間体の含有量が減少し、生合成中間体は、22-ヒドロキシコレステロールを含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの生合成中間体の含有量が減少し、生合成中間体は、22, 26-ジヒドロキシコレステロールを含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ステロイド系サポニンの生合成中間体の含有量が減少し、生合成中間体は、フロスタノール型サポニンアグリコンを含む。

20

30

【0480】

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変される。

【0481】

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が減少する。

40

【0482】

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が減少し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、

50



テロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体の含有量を改変することを含む。

【0489】

いくつかの実施形態では、遺伝子組換え細胞または遺伝子組換え植物において、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体を産生する方法は、未組換え細胞または未組換え植物と比較して、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させ、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体の含有量を減少させることを含む。

10

【0490】

いくつかの実施形態では、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を減少させる方法は、少なくとも1つの植物またはその一部の少なくとも1つの細胞を遺伝子組換えすることを含み、遺伝子組換えされた植物細胞は、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が減少し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変される。

20

【0491】

いくつかの実施形態では、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を減少させる方法は、少なくとも1つの植物またはその一部の少なくとも1つの細胞を遺伝子組換えすることを含み、遺伝子組換えされた植物細胞は、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が減少し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体の含有量が増加する。

30

【0492】

いくつかの実施形態では、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、少なくとも1つの植物またはその一部の少なくとも1つの細胞を遺伝子組換えすることを含み、遺伝子組換えされた植物細胞は、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が改変され、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体の含有量が減少する。

40

【0493】

いくつかの実施形態では、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、少なくとも1つの植物ま

50

たはその一部の少なくとも1つの細胞を遺伝子組換えすることを含み、遺伝子組換えされた植物細胞は、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が減少する。

【0494】

トリテルペノイド系サポニン、その誘導体、その代謝物、及びその生合成中間体

10

【0495】

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、またはそれらの任意の組み合わせの含有量が増加する。いくつかの実施形態では、本明細書に開示される少なくとも1つの遺伝子組換え細胞を含む遺伝子組換え植物は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が増加する。いくつかの実施形態では、インビトロ生合成系は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせを産生する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が増加する。いくつかの実施形態では、本明細書に開示される少なくとも1つの遺伝子組換え細胞を含む遺伝子組換え植物は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え細胞を含む本明細書に開示される遺伝子組換え植物は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が増加し、かつ、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が減少する。

20

【0496】

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせを産生することを含む。いくつかの実施形態では、遺伝子組換え植物またはその一部において、トリテルペノイド系サポニンを産生する方法は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせを産生することを含む。いくつかの実施形態では、本明細書に開示されるのは、本明細書に記載されている少なくとも1つの遺伝子組換え細胞を含む植物またはその一部において、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量を減少させる方法である。いくつかの実施形態では、本明細書に開示されるのは、本明細書に記載されているような少なくとも1つの遺伝子組換え細胞を含む植物またはその一部において、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量を増加させる方法である。

30

40

【0497】

一般的に使用されているサポニンの命名法では、トリテルペノイド系サポニン（トリテルペンサポニンとも呼ばれる）とトリテルペノイド系サポニンとを区別しているが、これは、そのアグリコンの構造と生化学的背景に基づいている。どちらのタイプのサポニンも、ステロール生合成の中心的な代謝産物である2,3-オキシドスカレンに由来すると考えられている。フィトステロール同化作用において、2,3-オキシドスカレンは、主に

50

シクロアルテノールに環化される。「トリテルペノイド系サポニン」は2,3-オキシドスクアレンの代替環化によってフィトステロール経路から分岐し、「ステロイド系サポゲニン」は、シクロアルテノール形成の下流のフィトステロール経路の中間体に由来すると考えられている(図3A参照)。

【0498】

サポニンは、ダンマラン型サポニン、チルカラン型サポニン、ルパン型サポニン、オレアナン型サポニン、タラキサスタン型サポニン、ウルサン型サポニン、ホバン型サポニン、ククルピタン型サポニン、シクロアルタン型サポニン、ラノスタン型サポニン、及びステロイド型サポニンからなる群から選択することができる。アグリコン骨格のサポゲニン類も同様に分類することができ、ダンマラン型サポニン、チルカラン型サポニン、ルパン型サポニン、オレアナン型サポニン、タラキサスタン型サポニン、ウルサン型サポニン、ホバン型サポニン、ククルピタン型サポニン、シクロアルタン型サポニン、ラノスタン型サポニン、及びステロイド型サポニンからなる群から選択することができる。

10

【0499】

いくつかの実施形態では、本明細書に開示される「トリテルペノイド系サポゲニン」は、ダンマラン型サポゲニン、チルカラン型サポゲニン、ルパン型サポゲニン、オレアナン型サポゲニン、ウルサン型サポゲニン、及びホバン型サポゲニンからなる群から選択することができる。いくつかの実施形態では、本開示の方法によって産生されるトリテルペノイド系サポゲニンは、ダンマラン型サポゲニン、チルカラン型サポゲニン、ルパン型サポゲニン、オレアナン型サポゲニン、ウルサン型サポゲニン、またはホバン型サポゲニンである。

20

【0500】

トリテルペノイド系サポゲニンは、一般的に、四環式または五環式の骨格を有する。いくつかの実施形態では、サポゲニン構築ブロックは、それ自体が複数の修飾を有していてもよく、そのようなものとしては、これに限定しないが、例えば、アミリン、ルペオール、及びダンマレンジオールなどの前駆体サポゲニン骨格のヒドロキシル基、ケト基、アルデヒド基、カルボキシル基などの小さな官能基などが挙げられる。

【0501】

「トリテルペン」と「トリテルペノイド」という用語は、互換的に使用することができ、同一の意味及び性質を有する。トリテルペノイド系サポニンは、複数の機能を有しており、これに限定しないが、例えば、甘味料、発泡剤、乳化剤、防腐剤、抗発癌剤、コレステロール低下剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、摂食阻害剤、抗真菌剤、またはそれらの任意の組み合わせとして使用することができる。

30

【0502】

いくつかの実施形態では、本明細書に開示されるトリテルペノイド系サポゲニンは、天然に存在するトリテルペノイド系サポゲニンと構造的に関連する、天然ではない新たなトリテルペノイド化合物も包含する。これらの天然ではない新たなトリテルペノイド系サポゲニンは、合成真核生物宿主細胞、例えば、植物細胞または酵母細胞の遺伝子操作後に得られる新規化合物であってもよい。

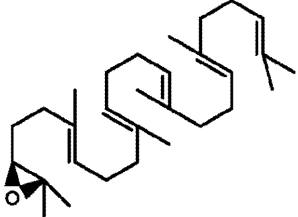
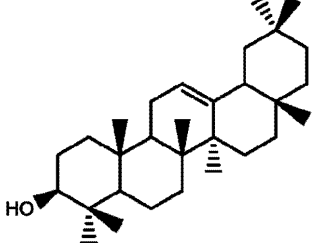
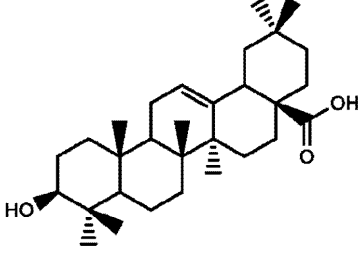
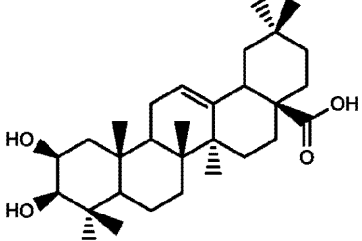
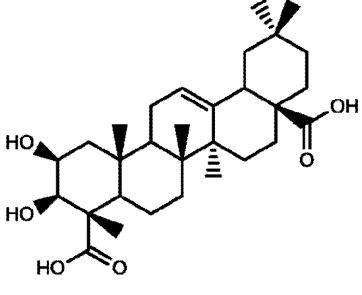
【0503】

トリテルペノイド系サポニン及びその生合成中間体の例を下記の表2に示す。

40

【0504】

【表 2 - 1】

化合物	構造式	名称
1		(S)-2,3-オキシドスクアレン
2		$\beta$ -アミリン
3		オレアノール酸 (OA)
4		アウグスチン酸 (AA)
5		メディカゲン酸 (MA)

10

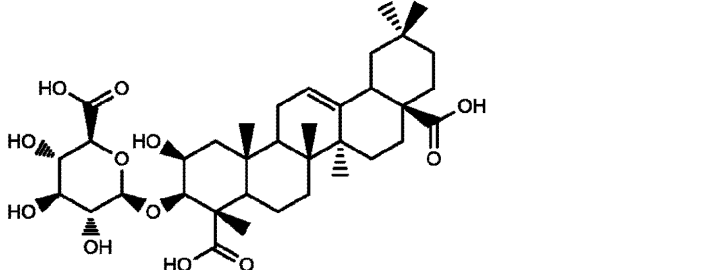
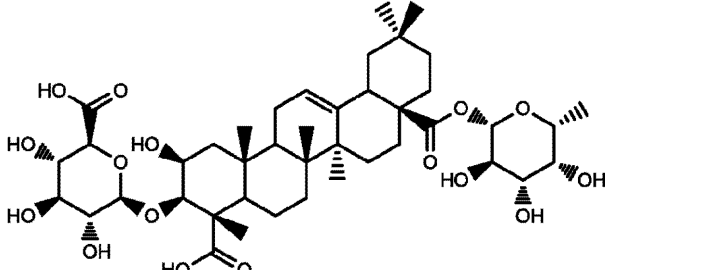
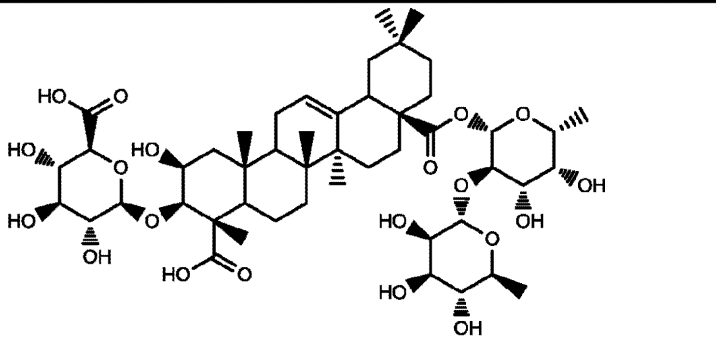
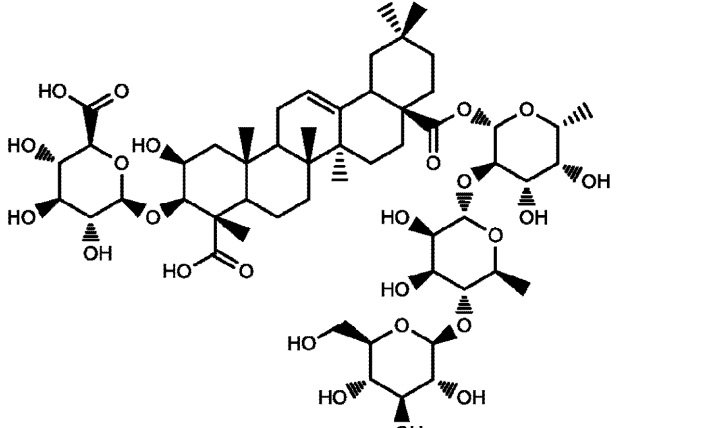
20

30

【 0 5 0 5 】

40

【表 2 - 2】

化合物	構造式	名称
6		メディカゲン酸 3-O- グルクロニド
7		Yossoside I 3-O-β-D- グルクロピラノシル -28-O-β-D- フコピラノシル- メディカゲン酸
8		Yossoside II 3-O- [β-D- グルクロノ ピラノシル] -28-O- [α-L- ラムピラノ シル-(1→2) -β-D- フコピラノシル] -メディカゲン酸
9		Yossoside III 3-O- [β-D- グルクロノ ピラノシル] -28-O- [β-D- グルコピラノ シル-(1→4) -α-L- ラムピラノシル -(1→2) -β-D- フコピラノシル] -メディカゲン酸

10

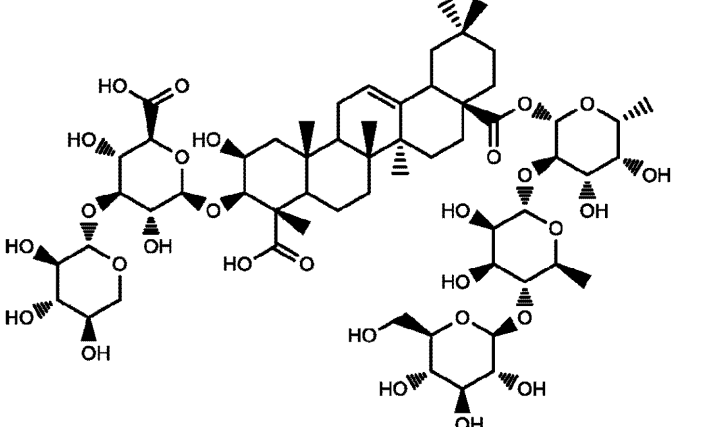
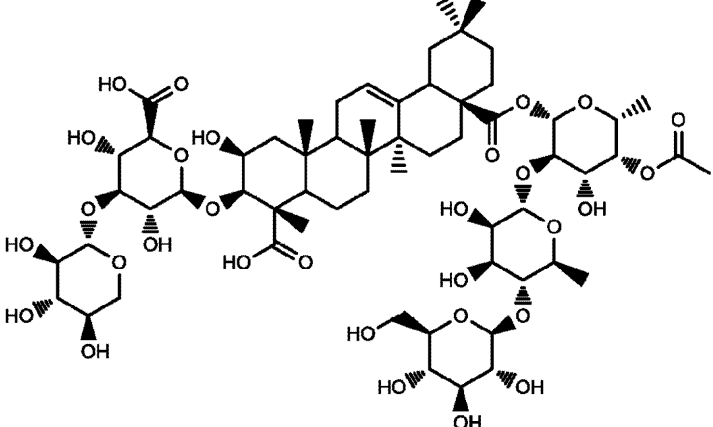
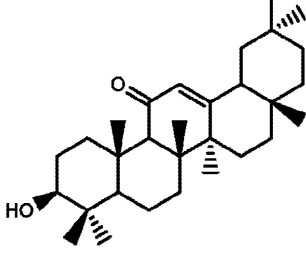
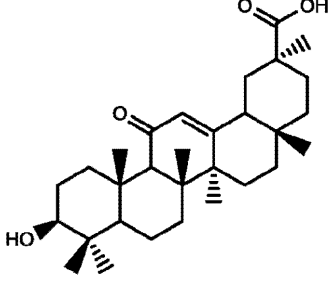
20

30

40

【 0 5 0 6 】

【表 2 - 3】

化合物	構造式	名称
10		YossosideIV 3-O- [β-D- キシロピラノシル -(1→3)-β-D- グルクロピラノシル] -28-O- [β-D-グルコ ピラノシル-(1→4) -α-L-ラムノ ピラノシル-(1→2) -β-D- フコピラノシル] -メディカゲン酸
11		YossosideV 3-O- [β-D- キシロピラノシル -(1→3) -β-D- グルクロノ ピラノシル] -28-O- [β-D- グルコピラノ シル-(1→4) -α-L- ラムノピラノシル -(1→2)-4- アセチル-β-D- フコピラノシル] -メディカゲン酸
12		11-オキソー β-アミリン
13		グリチルレチン酸

10

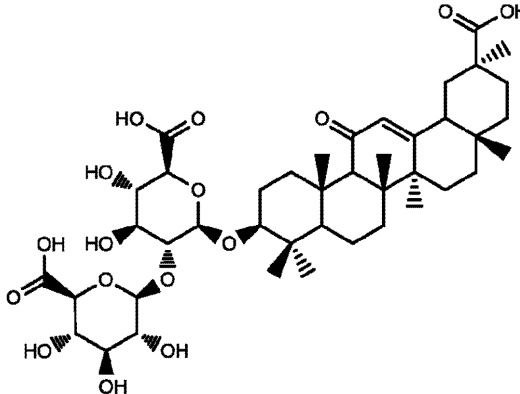
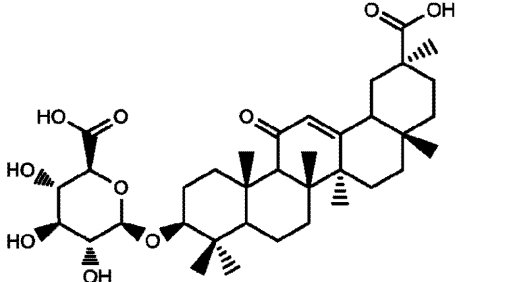
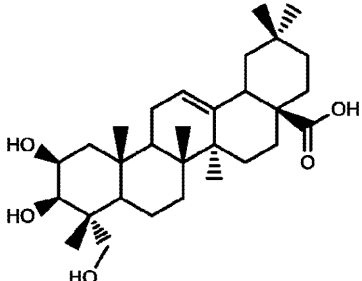
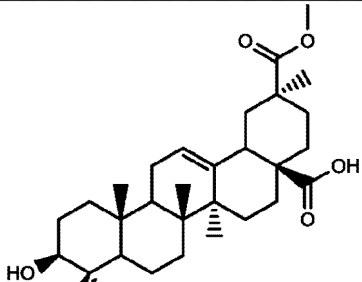
20

30

40

【 0 5 0 7 】

【表 2 - 4】

化合物	構造式	名称
14		グリチルリチン
15		グリチルレチン酸 3- $\text{O}$ - モノグルクロニド
16	not available	Yossoside Va
17	not available	Yossoside VI
18	not available	Yossoside VII
19	not available	Yossoside VIIa
20	not available	Yossoside VIII
21	not available	Yossoside IX
22	not available	Yossoside X
23	not available	Yossoside XI
24	not available	Yossoside XII
25		バヨゲニン
26		セルジャン酸

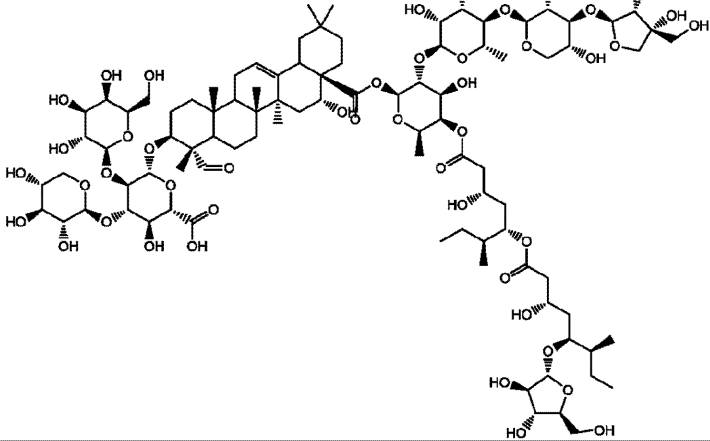
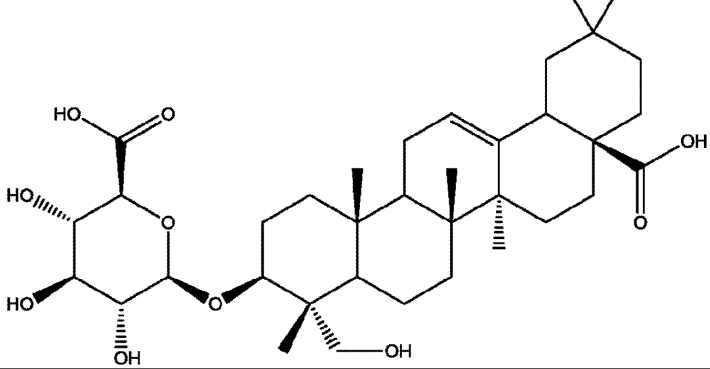
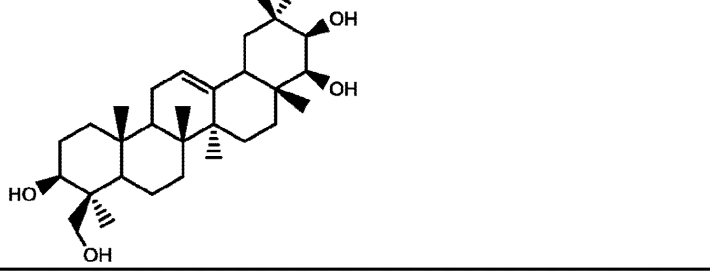
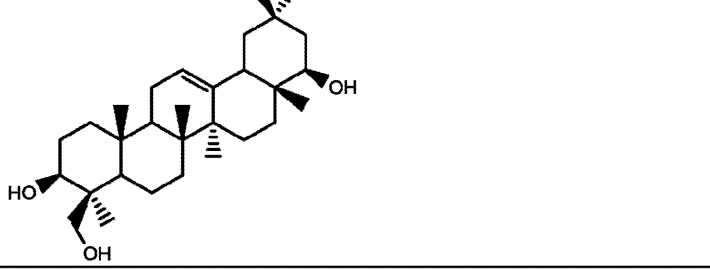
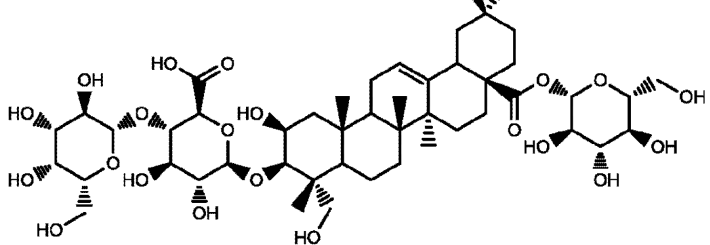
10

20

30

40

【表 2 - 5】

化合物	構造式	名称
27		QS-21
28		ヘデラゲニン -3GlcA
29		ソヤサポゲノールA
30		ソヤサポゲノールB
31		バヨゲニン -ヘキサ -ヘキサ -ヘキサ

10

20

30

40

【表 2 - 6】

化合物	構造式	名称
32		セルジャン酸 GlcA-glc
33		ベタブルガロシド IV
34		ソヤサポゲノールA -ヘキサ -ヘキサ -ペント
35		ソヤサポニンVI

10

20

30

40

## 【0510】

いくつかの実施形態では、トリテルペノイド系サポニンの誘導体は、トリテルペノイド系サポニンのグリコシル化誘導体を含む。

50

## 【0511】

いくつかの実施形態では、本明細書に開示されるトリテルペノイド系サポニンには、メディカゲン酸3-O-グルクロニド(MA-3-GlcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、バヨゲニン(化合物25)、バヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサポゲノールA(化合物29)、ソヤサポゲノールB(化合物30)、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント(化合物34)、ソヤサポニンVI(化合物35)、ベタブルガロシドIV(化合物33)、ヘデラゲニン-3GlcA、ギブソゲニン-3GlcA、ギブソゲン酸-3GlcA、QS-21アジュバント、またはそれらの任意の組み合わせを含む。

10

## 【0512】

いくつかの実施形態では、トリテルペノイド系サポニンには、メディカゲン酸3-O-グルクロニド(MA-3-GlcA)(化合物6)を含む。いくつかの実施形態では、トリテルペノイド系サポニンには、化合物7を含む。いくつかの実施形態では、トリテルペノイド系サポニンには、化合物8を含む。いくつかの実施形態では、トリテルペノイド系サポニンには、化合物9を含む。いくつかの実施形態では、トリテルペノイド系サポニンには、化合物10を含む。いくつかの実施形態では、トリテルペノイド系サポニンには、化合物11を含む。いくつかの実施形態では、トリテルペノイド系サポニンには、グリチルリチン(化合物14)を含む。いくつかの実施形態では、トリテルペノイド系サポニンには、グリチルレチン酸3-O-モノグルクロニド(化合物15)を含む。いくつかの実施形態では、トリテルペノイド系サポニンには、バヨゲニン(化合物25)を含む。いくつかの実施形態では、トリテルペノイド系サポニンには、バヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)を含む。いくつかの実施形態では、トリテルペノイド系サポニンには、セルジャン酸(化合物26)を含む。いくつかの実施形態では、トリテルペノイド系サポニンには、セルジャン酸-ヘキサA-ヘキサ(化合物32)を含む。いくつかの実施形態では、トリテルペノイド系サポニンには、ソヤサポゲノールA(化合物29)を含む。いくつかの実施形態では、トリテルペノイド系サポニンには、ソヤサポゲノールB(化合物30)を含む。いくつかの実施形態では、トリテルペノイド系サポニンには、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント(化合物34)を含む。いくつかの実施形態では、トリテルペノイド系サポニンには、ソヤサポニンVI(化合物35)を含む。いくつかの実施形態では、トリテルペノイド系サポニンには、ベタブルガロシドIV(化合物33)を含む。いくつかの実施形態では、トリテルペノイド系サポニンには、ヘデラゲニン-3GlcAを含む。いくつかの実施形態では、トリテルペノイド系サポニンには、ギブソゲニン-3GlcAを含む。いくつかの実施形態では、トリテルペノイド系サポニンには、ギブソゲン酸-3GlcAを含む。いくつかの実施形態では、トリテルペノイド系サポニンには、QS-21アジュバントを含む。

20

30

## 【0513】

トリテルペノイド系サポニンの含有量を改変することは、いくつかの実施形態では、これに限定しないが、例えば、甘味料であるグリチルリチン(化合物14)を産生する能力において、商業的に有益であり得る。いくつかの実施形態では、味料、発泡剤、乳化剤、防腐剤、抗発癌剤、コレステロール低下剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、摂食阻害剤、抗真菌剤、またはそれらの任意の組み合わせを含むトリテルペノイド系サポニンを産生することは、商業的に有益である。他の実施形態では、トリテルペノイド系サポニンの含有量を改変することは、これに限定しないが、例えば、植物のナデシコ目のキノア植物から産生されるキノアにおいて、苦味のあるトリテルペノイド系サポニンの含有量を減少されるかまたは除去するのに有益であると考えられる。他の実施形態では、トリテルペノイド系サポニンの含有量を改変することは、これに限定しないが、例えば、ホルモン模倣特性を有するトリテルペノイド系サポニンを減少されるかまたは除去するのに有益であり得る。

40

50

## 【0514】

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、メディカゲン酸 3-O-グルクロニド (MA-3-GlcA) (化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン (化合物14)、グリチルレチン酸 3-O-モノグルクロニド (化合物15)、パヨゲニン (化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ (化合物31)、セルジャン酸 (化合物26)、セルジャン酸-ヘキサA-ヘキサ (化合物32)、ソヤサポゲノールA (化合物29)、ソヤサポゲノールB (化合物30)、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント (化合物34)、ソヤサポニンVI (化合物35)、ベタブルガロシドIV (化合物33)、ヘデラゲニン-3GlcA、ギブソゲニン-3GlcA、ギブソゲン酸-3GlcA、QS-21アジュバント、またはそれらの任意の組み合わせから選択される少なくとも1つのトリテルペノイド系サポニンの含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、メディカゲン酸 3-O-グルクロニド (MA-3-GlcA) (化合物6) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、化合物7の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、化合物8の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、化合物9の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、化合物10の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、化合物11の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、グリチルリチン (化合物14) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、グリチルレチン酸 3-O-モノグルクロニド (化合物15) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、パヨゲニン (化合物25) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、パヨゲニン-ヘキサA-ヘキサ-ヘキサ (化合物31) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、セルジャン酸 (化合物26) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、セルジャン酸-ヘキサA-ヘキサ (化合物32) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、ソヤサポゲノールA (化合物29) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、ソヤサポゲノールB (化合物30) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント (化合物34) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、ソヤサポニンVI (化合物35) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、ベタブルガロシドIV (化合物33) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、ヘデラゲニン-3GlcAの含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、ギブソゲニン-3GlcAの含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、ギブソゲン酸-3GlcAの含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、QS-21アジュバントの含有量が増加する。

10

20

30

40

## 【0515】

いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させた遺伝子組換え細胞は、植物細胞、酵母細胞、藻類細胞、昆虫細胞、または細菌を含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させた遺伝子組換え細胞は、植物細胞を含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させた遺伝子組換え細胞は、酵母細胞を含む。いくつ

50

かの実施形態では、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させた遺伝子組換え細胞は、藻類細胞を含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させた遺伝子組換え細胞は、昆虫細胞を含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させた遺伝子組換え細胞は、細菌を含む。

【0516】

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、メディカゲン酸 3-O-グルクロニド (MA-3-GlcA) (化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン (化合物14)、グリチルレチン酸 3-O-モノグルクロニド (化合物15)、パヨゲニン (化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ (化合物31)、セルジャン酸 (化合物26)、セルジャン酸-ヘキサA-ヘキサ (化合物32)、ソヤサポゲノールA (化合物29)、ソヤサポゲノールB (化合物30)、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント (化合物34)、ソヤサポニンVI (化合物35)、ベタブルガロシドIV (化合物33)、ヘデラゲニン-3GlcA、ギブソゲニン-3GlcA、ギブソゲン酸-3GlcA、QS-21アジュバント、またはそれらの任意の組み合わせから選択される少なくとも1つのトリテルペノイド系サポニンの含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、メディカゲン酸 3-O-グルクロニド (MA-3-GlcA) (化合物6) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、化合物7の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、化合物8の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、化合物9の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、化合物10の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、化合物11の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、グリチルリチン (化合物14) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、グリチルレチン酸 3-O-モノグルクロニド (化合物15) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、パヨゲニン (化合物25) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、パヨゲニン-ヘキサA-ヘキサ-ヘキサ (化合物31) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、セルジャン酸 (化合物26) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、セルジャン酸-ヘキサA-ヘキサ (化合物32) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ソヤサポゲノールA (化合物29) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ソヤサポゲノールB (化合物30) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント (化合物34) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ソヤサポニンVI (化合物35) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ベタブルガロシドIV (化合物33) の含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ヘデラゲニン-3GlcAの含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ギブソゲニン-3GlcAの含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ギブソゲン酸-3GlcAの含有量が増加する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、QS-21アジュバントの含有量が増加する。

【0517】

10

20

30

40

50

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、メディカゲン酸 3-O-グルクロニド (MA-3-GlcA) (化合物 6)、化合物 7、化合物 8、化合物 9、化合物 10、化合物 11、グリチルリチン (化合物 14)、グリチルレチン酸 3-O-モノグルクロニド (化合物 15)、パヨゲニン (化合物 25)、パヨゲニン-ヘキサ A-ヘキサ-ヘキサ (化合物 31)、セルジャン酸 (化合物 26)、セルジャン酸-ヘキサ A-ヘキサ (化合物 32)、ソヤサポゲノール A (化合物 29)、ソヤサポゲノール B (化合物 30)、ソヤサポゲノール A-ヘキサ A-ヘキサ-ペント (化合物 34)、ソヤサポニン VI (化合物 35)、ベタブルガロシド IV (化合物 33)、ヘデラゲニン-3 GlcA、ギブソゲニン-3 GlcA、ギブソゲン酸-3 GlcA、QS-21 アジュバント、またはそれらの任意の組み合わせから選択される少なくとも 1 つのトリテルペノイド系サポニンの含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、メディカゲン酸 3-O-グルクロニド (MA-3-GlcA) (化合物 6) の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、化合物 7 の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、化合物 8 の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、化合物 9 の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、化合物 10 の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、化合物 11 の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、グリチルリチン (化合物 14) の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、グリチルレチン酸 3-O-モノグルクロニド (化合物 15) の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、パヨゲニン (化合物 25) の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、パヨゲニン-ヘキサ A-ヘキサ-ヘキサ (化合物 31) の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、セルジャン酸 (化合物 26) の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、セルジャン酸-ヘキサ A-ヘキサ (化合物 32) の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ソヤサポゲノール A (化合物 29) の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ソヤサポゲノール B (化合物 30) の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ソヤサポゲノール A-ヘキサ A-ヘキサ-ペント (化合物 34) の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ソヤサポニン VI (化合物 35) の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ベタブルガロシド IV (化合物 33) の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ヘデラゲニン-3 GlcA の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ギブソゲニン-3 GlcA の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、ギブソゲン酸-3 GlcA の含有量が減少する。いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、QS-21 アジュバントの含有量が減少する。

#### 【0518】

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、メディカゲン酸 3-O-グルクロニド (MA-3-GlcA) (化合物 6)、化合物 7、化合物 8、化合物 9、化合物 10、化合物 11、グリチルリチン (化合物 14)、グリチルレチン酸 3-O-モノグルクロニド (化合物 15)、パヨゲニン (化合物 25)、パヨゲニン-ヘキサ A-ヘキサ-ヘキサ (化合物 31)、セルジャン酸 (化合物 26)、セルジャン酸-ヘキサ A-ヘキサ (化合物 32)、ソヤサポゲノール A (化合物 29)、ソヤサポゲノール B (化合物 30)、ソヤサポゲノール A-ヘキサ A-ヘキサ-ペント (化合物 34)、ソヤサポニン VI (化合物 35)、ベタブルガロシド IV (化合物 33)、ヘデラゲニン-3

10

20

30

40

50

GlcA、ギブソゲニン - 3GlcA、ギブソゲン酸 - 3GlcA、QS - 21アジュバント、またはそれらの任意の組み合わせから選択される少なくとも1つのトリテルペノイド系サポニンの含有量が増加し、かつ、メディカゲン酸3 - O - グルクロニド (MA - 3 - GlcA) (化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン (化合物14)、グリチルレチン酸3 - O - モノグルクロニド (化合物15)、パヨゲニン (化合物25)、パヨゲニン - ヘキサA - ヘキサ - ヘキサ (化合物31)、セルジャン酸 (化合物26)、セルジャン酸 - ヘキサA - ヘキサ (化合物32)、ソヤサポゲノールA (化合物29)、ソヤサポゲノールB (化合物30)、ソヤサポゲノールA - ヘキサA - ヘキサ - ペント (化合物34)、ソヤサポニンVI (化合物35)、ベタブルガロシドIV (化合物33)、ヘデラゲニン - 3GlcA、ギブソゲニン - 3GlcA、ギブソゲン酸 - 3GlcA、QS - 21アジュバント、またはそれらの任意の組み合わせから選択される少なくとも1つのトリテルペノイド系サポニンの含有量が減少する。

#### 【0519】

いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、メディカゲン酸3 - O - グルクロニド (MA - 3 - GlcA) (化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン (化合物14)、グリチルレチン酸3 - O - モノグルクロニド (化合物15)、パヨゲニン (化合物25)、パヨゲニン - ヘキサA - ヘキサ - ヘキサ (化合物31)、セルジャン酸 (化合物26)、セルジャン酸 - ヘキサA - ヘキサ (化合物32)、ソヤサポゲノールA (化合物29)、ソヤサポゲノールB (化合物30)、ソヤサポゲノールA - ヘキサA - ヘキサ - ペント (化合物34)、ソヤサポニンVI (化合物35)、ベタブルガロシドIV (化合物33)、ヘデラゲニン - 3GlcA、ギブソゲニン - 3GlcA、ギブソゲン酸 - 3GlcA、QS - 21アジュバント、またはそれらの任意の組み合わせから選択される少なくとも1つのトリテルペノイド系サポニンを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、メディカゲン酸3 - O - グルクロニド (MA - 3 - GlcA) (化合物6)を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、化合物7を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、化合物8を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、化合物9を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、化合物10を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、化合物11を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、グリチルリチン (化合物14)を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、グリチルレチン酸3 - O - モノグルクロニド (化合物15)を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、パヨゲニン (化合物25)を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、パヨゲニン - ヘキサA - ヘキサ - ヘキサ (化合物31)を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、セルジャン酸 (化合物26)を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、セルジャン酸 - ヘキサA - ヘキサ (化合物32)を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、ソヤサポゲノールA (化合物29)を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、ソヤサポゲノールB (化合物30)を産生することを含む。いくつかの実施形態で

10

20

30

40

50

は、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、ソヤサポゲノール A - ヘキサ A - ヘキサ - ペント (化合物 34) を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、ソヤサポニン V I (化合物 35) を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、ベタブルガロシド I V (化合物 33) を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、ヘデラゲニン - 3 G l c A を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、ギブソゲニン - 3 G l c A を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、ギブソゲン酸 - 3 G l c A を産生することを含む。いくつかの実施形態では、遺伝子組換え細胞においてトリテルペノイド系サポニンを産生する方法は、Q S - 2 1 アジュバントを産生することを含む。

10

20

30

40

50

**【 0 5 2 0 】**

いくつかの実施形態では、遺伝子組換え細胞において少なくとも 1 つのトリテルペノイド系サポニンを産生する方法は、植物細胞、酵母細胞、藻類細胞、昆虫細胞、または細菌において少なくとも 1 つのトリテルペノイド系サポニンを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞において少なくとも 1 つのトリテルペノイド系サポニンを産生する方法は、植物細胞において少なくとも 1 つのトリテルペノイド系サポニンを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞において少なくとも 1 つのトリテルペノイド系サポニンを産生する方法は、酵母細胞において少なくとも 1 つのトリテルペノイド系サポニンを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞において少なくとも 1 つのトリテルペノイド系サポニンを産生する方法は、藻類細胞において少なくとも 1 つのトリテルペノイド系サポニンを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞において少なくとも 1 つのトリテルペノイド系サポニンを産生する方法は、昆虫細胞において少なくとも 1 つのトリテルペノイド系サポニンを産生することを含む。いくつかの実施形態では、遺伝子組換え細胞において少なくとも 1 つのトリテルペノイド系サポニンを産生する方法は、細菌において少なくとも 1 つのトリテルペノイド系サポニンを産生することを含む。いくつかの実施形態では、植物細胞は、植物またはその一部を含む。

**【 0 5 2 1 】**

いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、メディカゲン酸 3 - O - グルクロニド (M A - 3 - G l c A) (化合物 6)、化合物 7、化合物 8、化合物 9、化合物 10、化合物 11、グリチルリチン (化合物 14)、グリチルレチン酸 3 - O - モノグルクロニド (化合物 15)、パヨゲニン (化合物 25)、パヨゲニン - ヘキサ A - ヘキサ - ヘキサ (化合物 31)、セルジャン酸 (化合物 26)、セルジャン酸 - ヘキサ A - ヘキサ (化合物 32)、ソヤサポゲノール A (化合物 29)、ソヤサポゲノール B (化合物 30)、ソヤサポゲノール A - ヘキサ A - ヘキサ - ペント (化合物 34)、ソヤサポニン V I (化合物 35)、ベタブルガロシド I V (化合物 33)、ヘデラゲニン - 3 G l c A、ギブソゲニン - 3 G l c A、ギブソゲン酸 - 3 G l c A、Q S - 2 1 アジュバント、またはそれらの任意の組み合わせから選択される少なくとも 1 つのトリテルペノイド系サポニンを産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、メディカゲン酸 3 - O - グルクロニド (M A - 3 - G l c A) (化合物 6) を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、化合物 7 を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、化合物 8 を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、化合物 9 を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、化合物 10 を産生す

ることを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、化合物 11 を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、グリチルリチン（化合物 14）を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、グリチルレチン酸 3-O-モノグルクロニド（化合物 15）を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、パヨゲニン（化合物 25）を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、パヨゲニン-ヘキサ A-ヘキサ-ヘキサ（化合物 31）を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、セルジャン酸（化合物 26）を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、セルジャン酸-ヘキサ A-ヘキサ（化合物 32）を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、ソヤサポゲノール A（化合物 29）を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、ソヤサポゲノール B（化合物 30）を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、ソヤサポゲノール A-ヘキサ A-ヘキサ-ペント（化合物 34）を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、ソヤサポニン VI（化合物 35）を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、ベタブルガロシド IV（化合物 33）を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、ヘデラゲニン-3 Glc A を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、ギブソゲニン-3 Glc A を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、ギブソゲン酸-3 Glc A を産生することを含む。いくつかの実施形態では、遺伝子組換え植物においてトリテルペノイド系サポニンを産生する方法は、QS-21 アジュバントを産生することを含む。

#### 【0522】

いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、メディカゲン酸 3-O-グルクロニド（MA-3-Glc A）（化合物 6）、化合物 7、化合物 8、化合物 9、化合物 10、化合物 11、グリチルリチン（化合物 14）、グリチルレチン酸 3-O-モノグルクロニド（化合物 15）、パヨゲニン（化合物 25）、パヨゲニン-ヘキサ A-ヘキサ-ヘキサ（化合物 31）、セルジャン酸（化合物 26）、セルジャン酸-ヘキサ A-ヘキサ（化合物 32）、ソヤサポゲノール A（化合物 29）、ソヤサポゲノール B（化合物 30）、ソヤサポゲノール A-ヘキサ A-ヘキサ-ペント（化合物 34）、ソヤサポニン VI（化合物 35）、ベタブルガロシド IV（化合物 33）、ヘデラゲニン-3 Glc A、ギブソゲニン-3 Glc A、ギブソゲン酸-3 Glc A、QS-21 アジュバント、またはそれらの任意の組み合わせから選択される少なくとも 1 つのトリテルペノイド系サポニンを産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、メディカゲン酸 3-O-グルクロニド（MA-3-Glc A）（化合物 6）を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、化合物 7 を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、化合物 8 を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、化合物 9 を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、化合物 10 を産生す

10

20

30

40

50

ることを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、化合物 11 を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、グリチルリチン（化合物 14）を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、グリチルレチン酸 3-O-モノグルクロニド（化合物 15）を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、パヨゲニン（化合物 25）を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、パヨゲニン-ヘキサ A-ヘキサ-ヘキサ（化合物 31）を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、セルジャン酸（化合物 26）を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、セルジャン酸-ヘキサ A-ヘキサ（化合物 32）を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、ソヤサポゲノール A（化合物 29）を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、ソヤサポゲノール B（化合物 30）を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、ソヤサポゲノール A-ヘキサ A-ヘキサ-ペント（化合物 34）を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、ソヤサポニン VI（化合物 35）を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、ベタブルガロシド IV（化合物 33）を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、ヘデラゲニン-3 Glc A を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、ギブソゲニン-3 Glc A を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、ギブソゲン酸-3 Glc A を産生することを含む。いくつかの実施形態では、インビトロ翻訳系においてトリテルペノイド系サポニンを産生する方法は、QS-21 アジュバントを産生することを含む。

10

20

30

#### 【0523】

いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を減少させる方法は、メディカゲン酸 3-O-グルクロニド（MA-3-Glc A）（化合物 6）、化合物 7、化合物 8、化合物 9、化合物 10、化合物 11、グリチルリチン（化合物 14）、グリチルレチン酸 3-O-モノグルクロニド（化合物 15）、パヨゲニン（化合物 25）、パヨゲニン-ヘキサ A-ヘキサ-ヘキサ（化合物 31）、セルジャン酸（化合物 26）、セルジャン酸-ヘキサ A-ヘキサ（化合物 32）、ソヤサポゲノール A（化合物 29）、ソヤサポゲノール B（化合物 30）、ソヤサポゲノール A-ヘキサ A-ヘキサ-ペント（化合物 34）、ソヤサポニン VI（化合物 35）、ベタブルガロシド IV（化合物 33）、ヘデラゲニン-3 Glc A、ギブソゲニン-3 Glc A、ギブソゲン酸-3 Glc A、QS-21 アジュバント、またはそれらの任意の組み合わせから選択される少なくとも 1 つのトリテルペノイド系サポニンの含有量を減少させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を減少させる方法は、メディカゲン酸 3-O-グルクロニド（MA-3-Glc A）（化合物 6）の含有量を減少させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を減少させる方法は、化合物 7 の含有量を減少させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を減少させる方法は、化合物 8 の含有量を減少させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を減少させる方法は、化合物 9 の含有量を減少させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペ

40

50

ノイド系サポニンの含有量を減少させる方法は、化合物10の含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、化合物11の含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、グリチルリチン（化合物14）の含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、グリチルレチン酸3-O-モノグルクロニド（化合物15）の含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、パヨゲニン（化合物25）の含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、パヨゲニン-ヘキサA-ヘキサ-ヘキサ（化合物31）の含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、セルジャン酸（化合物26）の含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、セルジャン酸-ヘキサA-ヘキサ（化合物32）の含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、ソヤサポゲノールA（化合物29）の含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、ソヤサポゲノールB（化合物30）の含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント（化合物34）の含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、ソヤサポニンVI（化合物35）の含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、ベタブルガロシドIV（化合物33）の含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、ヘデラゲニン-3Gl c Aの含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、ギブソゲニン-3Gl c Aの含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、ギブソゲン酸-3Gl c Aの含有量を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、QS-21アジュバントの含有量を減少させることを含む。

10

20

30

#### 【0524】

いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を増加させる方法は、メディカゲン酸3-O-グルクロニド（MA-3-Gl c A）（化合物6）、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン（化合物14）、グリチルレチン酸3-O-モノグルクロニド（化合物15）、パヨゲニン（化合物25）、パヨゲニン-ヘキサA-ヘキサ-ヘキサ（化合物31）、セルジャン酸（化合物26）、セルジャン酸-ヘキサA-ヘキサ（化合物32）、ソヤサポゲノールA（化合物29）、ソヤサポゲノールB（化合物30）、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント（化合物34）、ソヤサポニンVI（化合物35）、ベタブルガロシドIV（化合物33）、ヘデラゲニン-3Gl c A、ギブソゲニン-3Gl c A、ギブソゲン酸-3Gl c A、QS-21アジュバント、またはそれらの任意の組み合わせから選択される少なくとも1つのトリテルペノイド系サポニンの含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を増加させる方法は、メディカゲン酸3-O-グルクロニド（MA-3-Gl c A）（化合物6）の含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を増加させる方法は、化合物7の含有量を増加させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を増

40

50

加させる方法は、化合物 8 の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、化合物 9 の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、化合物 10 の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、化合物 11 の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、グリチルリチン（化合物 14）の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、グリチルレチン酸 3 - O - モノグルクロニド（化合物 15）の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、パヨゲニン（化合物 25）の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、パヨゲニン - ヘキサ A - ヘキサ - ヘキサ（化合物 31）の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、セルジャン酸（化合物 26）の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、セルジャン酸 - ヘキサ A - ヘキサ（化合物 32）の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、ソヤサポゲノール A（化合物 29）の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、ソヤサポゲノール B（化合物 30）の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、ソヤサポゲノール A - ヘキサ A - ヘキサ - ペント（化合物 34）の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、ソヤサポニン VI（化合物 35）の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、ベタブルガロシド IV（化合物 33）の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、ヘデラゲニン - 3 Glc A の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、ギブソゲニン - 3 Glc A の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、ギブソゲン酸 - 3 Glc A の含有量を増加させることを含む。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量を増加させる方法は、QS - 21 アジュバントの含有量を増加させることを含む。

10

20

30

**【0525】**

いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンと、少なくとも 1 つのトリテルペノイド系サポニンの生合成中間体との両方の含有量が改変される。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンと、少なくとも 1 つのトリテルペノイド系サポニンの生合成中間体との両方の含有量が増加する。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンと、少なくとも 1 つのトリテルペノイド系サポニンの生合成中間体との両方の含有量が減少する。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量が増加し、かつ、少なくとも 1 つのトリテルペノイド系サポニンの生合成中間体の含有量が減少する。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンの含有量が減少し、かつ、少なくとも 1 つのトリテルペノイド系サポニンの生合成中間体の含有量が増加する。いくつかの実施形態では、トリテルペノイド系サポニン中間体の含有量を測定可能に改変することなく、トリテルペノイド系サポニンの含有量が改変される。いくつかの実施形態では、トリテルペノイド系サポニンの含有量を測定可能に改変することなく、トリテルペノ

40

50

イド系サポニン中間体の含有量が改変される。

【0526】

「中間体」という用語は、いくつかの実施形態では、「生合成中間体」という用語と互換的に使用することができ、すべて同一の性質及び意味を有する。

【0527】

いくつかの実施形態では、トリテルペノイド系サポニンの生合成中間体は、化合物1、化合物2、メディカゲン酸（化合物5）、オレアノール酸（化合物3）、オレアノール酸-3GlcA、アウグスチン酸（化合物4）、アウグスチン酸-3GlcA、グリチルレチン酸（化合物13）、またはそれらの任意の組み合わせを含む。

【0528】

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、化合物1、化合物2、メディカゲン酸（化合物5）、オレアノール酸（化合物3）、オレアノール酸-3GlcA、アウグスチン酸（化合物4）、アウグスチン酸-3GlcA、グリチルレチン酸（化合物13）、またはそれらの任意の組み合わせを含む少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、化合物1を含む少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、化合物2を含む少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、メディカゲン酸（化合物5）を含む少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、オレアノール酸（化合物3）を含む少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、オレアノール酸-3GlcAを含む少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、アウグスチン酸（化合物4）を含む少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、アウグスチン酸-3GlcAを含む少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞は、グリチルレチン酸（化合物13）を含む少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が増加する。

【0529】

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、化合物1、化合物2、メディカゲン酸（化合物5）、オレアノール酸（化合物3）、オレアノール酸-3GlcA、アウグスチン酸（化合物4）、アウグスチン酸-3GlcA、グリチルレチン酸（化合物13）、またはそれらの任意の組み合わせを含む少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え細胞は、化合物1を含む少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え細胞は、化合物2を含む少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え細胞は、メディカゲン酸（化合物5）を含む少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え細胞は、オレアノール酸（化合物3）を含む少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え細胞は、オレアノール酸-3GlcAを含む少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え細胞は、アウグスチン酸（化合物4）を含む少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え細胞は、アウグスチン酸-3GlcAを含む少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が減少する。いくつかの実施形態では、遺伝子組換え細胞は、グリチルレチン酸（化合物13）を含む少なくとも1つのトリテルペノイド系サポニ

10

20

30

40

50

ンの生合成中間体の含有量が減少する。

【0530】

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が増加し、かつ、少なくとも1つのトリテルペノイド系サポニンの生合成中間体の含有量が減少する。少なくとも1つのトリテルペノイド系サポニンの生合成中間体は、化合物1、化合物2、メディカゲン酸（化合物5）、オレアノール酸（化合物3）、オレアノール酸-3 Glc A、アウグスチン酸（化合物4）、アウグスチン酸-3 Glc A、グリチルレチン酸（化合物13）、またはそれらの任意の組み合わせから選択される。

【0531】

いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、化合物1、化合物2、メディカゲン酸（化合物5）、オレアノール酸（化合物3）、オレアノール酸-3 Glc A、アウグスチン酸（化合物4）、アウグスチン酸-3 Glc A、グリチルレチン酸（化合物13）、またはそれらの任意の組み合わせから選択される少なくとも1つのトリテルペノイド系サポニンの生合成中間体を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、少なくとも化合物1を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、少なくとも化合物2を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、少なくともメディカゲン酸（化合物5）を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、少なくともオレアノール酸（化合物3）を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、少なくともオレアノール酸-3 Glc Aを減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、少なくともアウグスチン酸（化合物4）を減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、少なくともアウグスチン酸-3 Glc Aを減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を減少させる方法は、少なくともグリチルレチン酸（化合物13）を減少させることを含む。

【0532】

いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を増加させる方法は、化合物1、化合物2、メディカゲン酸（化合物5）、オレアノール酸（化合物3）、オレアノール酸-3 Glc A、アウグスチン酸（化合物4）、アウグスチン酸-3 Glc A、グリチルレチン酸（化合物13）、またはそれらの任意の組み合わせから選択される少なくとも1つのトリテルペノイド系サポニンの生合成中間体を増加させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を増加させる方法は、少なくとも化合物1を増加させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を増加させる方法は、少なくとも化合物2を増加させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を増加させる方法は、少なくともメディカゲン酸（化合物5）を増加させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を増加させる方法は、少なくともオレアノール酸（化合物3）を増加させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を増加させる方法は、少なくともオレアノール酸-3 Glc Aを増加させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を増加させる方法は、少なくともアウグスチン酸（化合物4）を増加させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を増加させる方法は、少なくともアウグスチン酸-3 Glc Aを増加さ

10

20

30

40

50

せることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンの含有量を増加させる方法は、少なくともグリチルレチン酸（化合物13）を増加させることを含む。

#### 【0533】

いくつかの実施形態では、トリテルペノイド系サポニンの生合成中間体は、化合物1、化合物2、メディカゲン酸（化合物5）、オレアノール酸（化合物3）、オレアノール酸-3G1cA、アウグスチン酸（化合物4）、アウグスチン酸-3G1cA、グリチルレチン酸（化合物13）、またはそれらの任意の組み合わせを含む。

#### 【0534】

予期せぬことに、本開示は、酵素（例えばトリテルペノイド系サポニン生合成経路の酵素）をコードする少なくとも1つの異種遺伝子を発現するように細胞または植物を遺伝的に改変することによって、細胞（例えば、植物細胞、酵母細胞、藻類細胞、昆虫細胞）、細菌、または植物において、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させることができることが分かった。他の実施形態では、本明細書に記載及び例示されているのは、植物細胞内の少なくとも1つの内在性遺伝子（これに限定しないが、例えば、CSLG遺伝子）を遺伝子的に改変して、その発現、活性、安定性、またはそれらの任意の組み合わせを調節する方法である。以下の実施例では、トリテルペノイド系サポニン生合成経路の一部であることがこれまで知られていなかった酵素活性及び酵素を開示する。この知見がなければ、トリテルペノイド系サポニン化合物の産生は不可能であった。

10

20

#### 【0535】

いくつかの実施形態では、トリテルペノイド系サポニン代謝経路は、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加した細胞、あるいは、植物またはその一部においてこれらの化合物の含有量が増加した植物を生産することができる。いくつかの実施形態では、有用なトリテルペノイド系サポニンの含有量が増加した植物、これに限定しないが、例えば、作物植物が生産される。いくつかの実施形態では、本明細書に開示されるのは、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加した細胞（これに限定しないが、例えば、植物細胞、酵母、または藻類細胞）または植物（これに限定しないが、作物植物）または植物の一部を生産するための手段及び方法である。代替的または追加的に、本明細書に開示される遺伝子の発現を制御することは、例えば製薬産業においてさらに使用するための所望のトリテルペノイド系サポニンの生産のために、あるいは、食品または他の栄養補助食品（これに限定しないが、例えば、甘味料など）の生産のために使用することができる。いくつかの実施形態では、これらの高価値サポニンは、精製され、例えば、甘味料、発泡剤、乳化剤、防腐剤、抗発癌剤、コレステロール低下剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、摂食阻害剤、抗真菌剤、またはそれらの任意の組み合わせとして使用することができる。本明細書に開示される細胞及び植物は、栄養学的、薬学的、及び商業的に重要な価値を有する化合物を含有する。

30

#### 【0536】

いくつかの実施形態では、本明細書に記載の遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、対応する非組換え細胞と比較して、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が改変されている。いくつかの実施形態では、含有量の改変は、含有量の増加を含む。いくつかの実施形態では、例えば、遺伝子組換え細胞または遺伝子組換え植物は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、その生合成中間体、またはそれらの任意の組み合わせの含有量が増加している。

40

#### 【0537】

いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、少なくとも1つのトリテルペノイ

50



え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、トリテルペノイド系サポニンの少なくとも2つの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、トリテルペノイド系サポニンの少なくとも3つの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、トリテルペノイド系サポニンの少なくとも4つの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、トリテルペノイド系サポニンの少なくとも5つの生合成中間体の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞、あるいは、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、トリテルペノイド系サポニンの少なくとも6つの生合成中間体の含有量が増加する。

10

**【0541】**

当業者であれば、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体に関する「含有量」及び「レベル」という用語は、対照細胞または対照植物と比較した、遺伝子組換え細胞または遺伝子組換え植物中の化合物の量、例えば、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量を包含することを認識できるであろう。この文脈では、「含有量」及び「レベル」という用語は、互換的に使用することができ、すべて同一の意味及び性質を有する。

20

**【0542】**

いくつかの実施形態では、含有量が増加したトリテルペノイド系サポニンは、甘味料、発泡剤、乳化剤、防腐剤、抗発癌剤、コレステロール低下剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、摂食阻害剤、抗真菌剤、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、含有量が増加したトリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GLcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、パヨゲニン(化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサポゲノールA(化合物29)、ソヤサポゲノールB(化合物30)、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント(化合物34)、ソヤサポニンVI(化合物35)、ベタブルガロシドIV(化合物33)、ヘデラゲニン-3GLcA、ギブソゲニン-3GLcA、ギブソゲン酸-3GLcA、QS-21アジュバント、またはそれらの任意の組み合わせから選択されるトリテルペノイド系サポニンを含む。いくつかの実施形態では、含有量が増加したトリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GLcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、パヨゲニン(化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサポゲノールA(化合物29)、ソヤサポゲノールB(化合物30)、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント(化合物34)、ソヤサポニンVI(化合物35)、ベタブルガロシドIV(化合物33)、ヘデラゲニン-3GLcA、ギブソゲニン-3GLcA、ギブソゲン酸-3GLcA、QS-21アジュバント、またはそれらの任意の組み合わせから選択される少なくとも1、2、3、4、5、6、またはそれ以上のトリテルペノイド系サポニンを含む。

30

40

**【0543】**

いくつかの実施形態では、含有量が増加したトリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GLcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン

50

酸 3 - O - モノグルクロニド (化合物 15)、パヨゲニン (化合物 25)、パヨゲニン - ヘキサ A - ヘキサ - ヘキサ (化合物 31)、セルジャン酸 (化合物 26)、セルジャン酸 - ヘキサ A - ヘキサ (化合物 32)、ソヤサボゲノール A (化合物 29)、ソヤサボゲノール B (化合物 30)、ソヤサボゲノール A - ヘキサ A - ヘキサ - ペント (化合物 34)、ソヤサポニン V I (化合物 35)、ベタブルガロシド I V (化合物 33)、ヘデラゲニン - 3 G l c A、ギブソゲニン - 3 G l c A、ギブソゲン酸 - 3 G l c A、Q S - 2 1 アジュバント、またはそれらの任意の組み合わせから選択される少なくとも 1、2、3、4、5、6、またはそれ以上のトリテルペノイド系サポニンを含む。

【0544】

いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において含有量が増加したトリテルペノイド系サポニンは、メディカゲン酸 3 - O - グルクロニド (M A - 3 - G l c A) (化合物 6)、化合物 7、化合物 8、化合物 9、化合物 10、化合物 11、グリチルリチン (化合物 14)、グリチルレチン酸 3 - O - モノグルクロニド (化合物 15)、パヨゲニン (化合物 25)、パヨゲニン - ヘキサ A - ヘキサ - ヘキサ (化合物 31)、セルジャン酸 (化合物 26)、セルジャン酸 - ヘキサ A - ヘキサ (化合物 32)、ソヤサボゲノール A (化合物 29)、ソヤサボゲノール B (化合物 30)、ソヤサボゲノール A - ヘキサ A - ヘキサ - ペント (化合物 34)、ソヤサポニン V I (化合物 35)、ベタブルガロシド I V (化合物 33)、ヘデラゲニン - 3 G l c A、ギブソゲニン - 3 G l c A、ギブソゲン酸 - 3 G l c A、Q S - 2 1 アジュバント、またはそれらの任意の組み合わせから選択されるトリテルペノイド系サポニンを含む。

【0545】

いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、メディカゲン酸 3 - O - グルクロニド (M A - 3 - G l c A) (化合物 6) の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、化合物 7 の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、化合物 8 の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、化合物 9 の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、化合物 10 の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、化合物 11 の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、グリチルリチン (化合物 14) の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、グリチルレチン酸 3 - O - モノグルクロニド (化合物 15) の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、パヨゲニン (化合物 25) の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、パヨゲニン - ヘキサ A - ヘキサ - ヘキサ (化合物 31) の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、セルジャン酸 (化合物 26) の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部におい

10

20

30

40

50

て、セルジャン酸 - ヘキサ A - ヘキサ (化合物 32) の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、ソヤサポゲノール A (化合物 29) の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、ソヤサポゲノール B (化合物 30) の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、ソヤサポゲノール A - ヘキサ A - ヘキサ - ペント (化合物 34) の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、ソヤサポニン V I (化合物 35) の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、ベタブルガロシド I V (化合物 33) の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、ヘデラゲニン - 3 G l c A の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、ギブソゲニン - 3 G l c A の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、ギブソゲン酸 - 3 G l c A の含有量が増加する。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、Q S - 21 アジュバントの含有量が増加する。当業者であれば、いくつかの実施形態では、1つのトリテルペノイド系サポニンの含有量が増加する場合には、追加のトリテルペノイド系サポニン、または中間体、またはそれらの組み合わせも同一の時間の細胞内で増加し得ることを理解できるであろう。

10

20

**【0546】**

いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母細胞、藻類細胞)、または遺伝子組換え植物の一部において、トリテルペノイド系サポニンの生合成中間体の含有量が増加する。生合成中間体は、化合物 1、化合物 2、メディカゲン酸 (化合物 5)、オレアノール酸 (化合物 3)、オレアノール酸 - 3 G l c A、アウグスチン酸 (化合物 4)、アウグスチン酸 - 3 G l c A、グリチルレチン酸 (化合物 13)、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部は、トリテルペノイド系サポニンの少なくとも 2つの生合成中間体の含有量が増加する。生合成中間体は、化合物 1、化合物 2、メディカゲン酸 (化合物 5)、オレアノール酸 (化合物 3)、オレアノール酸 - 3 G l c A、アウグスチン酸 (化合物 4)、アウグスチン酸 - 3 G l c A、グリチルレチン酸 (化合物 13)、またはそれらの任意の組み合わせを含む。

30

**【0547】**

いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、トリテルペノイド系サポニンの生合成中間体の含有量が増加し、生合成中間体は、化合物 1を含む。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、トリテルペノイド系サポニンの生合成中間体の含有量が増加し、生合成中間体は、化合物 2を含む。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、トリテルペノイド系サポニンの生合成中間体の含有量が増加し、生合成中間体は、メディカゲン酸 (化合物 5)を含む。いくつかの実施形態では、遺伝子組換え細胞 (例えば、植物細胞、酵母、藻類細胞、昆

40

50

虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、トリテルペノイド系サポニンの生合成中間体の含有量が増加し、生合成中間体は、オレアノール酸(化合物3)を含む。いくつかの実施形態では、遺伝子組換え細胞(例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、トリテルペノイド系サポニンの生合成中間体の含有量が増加し、生合成中間体は、オレアノール酸-3GlcAを含む。いくつかの実施形態では、遺伝子組換え細胞(例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、トリテルペノイド系サポニンの生合成中間体の含有量が増加し、生合成中間体は、アウグスチン酸(化合物4)を含む。いくつかの実施形態では、遺伝子組換え細胞(例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、トリテルペノイド系サポニンの生合成中間体の含有量が増加し、生合成中間体は、アウグスチン酸-3GlcAを含む。いくつかの実施形態では、遺伝子組換え細胞(例えば、植物細胞、酵母、藻類細胞、昆虫細胞、細菌)、あるいは、遺伝子組換え植物またはその一部において、トリテルペノイド系サポニンの生合成中間体の含有量が増加し、生合成中間体は、グリチルレチン酸(化合物13)を含む。

10

## 【0548】

いくつかの実施形態では、少なくとも1つの遺伝子組換え細胞を含む植物またはその一部において含有量が減少したトリテルペノイド系サポニンは、苦味を有する化合物または毒素を含む。いくつかの実施形態では、含有量が減少したトリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GlcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、パヨゲニン(化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサポゲノールA(化合物29)、ソヤサポゲノールB(化合物30)、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント(化合物34)、ソヤサポニンVI(化合物35)、ベタブルガロシドIV(化合物33)、ヘデラゲニン-3GlcA、ギブソゲニン-3GlcA、ギブソゲン酸-3GlcA、QS-21アジュバント、またはそれらの任意の組み合わせから選択されるトリテルペノイド系サポニンを含む。いくつかの実施形態では、含有量が減少したトリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GlcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、パヨゲニン(化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサポゲノールA(化合物29)、ソヤサポゲノールB(化合物30)、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント(化合物34)、ソヤサポニンVI(化合物35)、ベタブルガロシドIV(化合物33)、ヘデラゲニン-3GlcA、ギブソゲニン-3GlcA、ギブソゲン酸-3GlcA、QS-21アジュバント、またはそれらの任意の組み合わせから選択される少なくとも1、2、3、4、5、6、またはそれ以上のトリテルペノイド系サポニンを含む。

20

30

## 【0549】

いくつかの実施形態では、含有量が減少したトリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GlcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、パヨゲニン(化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサポゲノールA(化合物29)、ソヤサポゲノールB(化合物30)、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント(化合物34)、ソヤサポニンVI(化合物35)、ベタブルガロシドIV(化合物33)、ヘデラゲニン-3GlcA、ギブソゲニン-3GlcA、ギブソゲン酸-3GlcA、QS-21アジュバント、またはそれらの任意の組み合わせから選択される少なくとも1、2、3、4

40

50

、 5、 6、 またはそれ以上のトリテルペノイド系サポニンを含む。

【 0 5 5 0 】

いくつかの実施形態では、少なくとも1つの遺伝子組換え細胞を含む遺伝子組換え植物またはその一部において含有量が減少したトリテルペノイド系サポニンは、メディカゲン酸 3 - O - グルクロニド ( M A - 3 - G l c A ) ( 化合物 6 )、化合物 7、化合物 8、化合物 9、化合物 10、化合物 11、グリチルリチン ( 化合物 14 )、グリチルレチン酸 3 - O - モノグルクロニド ( 化合物 15 )、バヨゲニン ( 化合物 25 )、バヨゲニン - ヘキサ A - ヘキサ - ヘキサ ( 化合物 31 )、セルジャン酸 ( 化合物 26 )、セルジャン酸 - ヘキサ A - ヘキサ ( 化合物 32 )、ソヤサポゲノール A ( 化合物 29 )、ソヤサポゲノール B ( 化合物 30 )、ソヤサポゲノール A - ヘキサ A - ヘキサ - ペント ( 化合物 34 )、ソヤサポニン V I ( 化合物 35 )、ベタブルガロシド I V ( 化合物 33 )、ヘデラゲニン - 3 G l c A、ギブソゲニン - 3 G l c A、ギブソゲン酸 - 3 G l c A、Q S - 2 1 アジュバント、またはそれらの任意の組み合わせから選択されるトリテルペノイド系サポニンを含む。

10

【 0 5 5 1 】

いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、メディカゲン酸 3 - O - グルクロニド ( M A - 3 - G l c A ) ( 化合物 6 ) の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、化合物 7 の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、化合物 8 の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、化合物 9 の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、化合物 10 の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、化合物 11 の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、グリチルリチン ( 化合物 14 ) の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、グリチルレチン酸 3 - O - モノグルクロニド ( 化合物 15 ) の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、バヨゲニン ( 化合物 25 ) の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、バヨゲニン - ヘキサ A - ヘキサ - ヘキサ ( 化合物 31 ) の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、セルジャン酸 ( 化合物 26 ) の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、セルジャン酸 - ヘキサ A - ヘキサ ( 化合物 32 ) の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ソヤサポゲノール A ( 化合物 29 ) の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ソヤサポゲノール B ( 化合物 30 ) の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ソヤサポゲノール A - ヘキサ A - ヘキサ - ペント ( 化合物 34 ) の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ソヤサポニン V I ( 化合物 35 ) の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ベタブルガロシド I V ( 化合物 33 ) の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ヘデラゲニン - 3 G l c A の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはそ

20

30

40

50

の一部は、ギブソゲニン - 3 G l c A の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、ギブソゲン酸 - 3 G l c A の含有量が減少する。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、Q S - 2 1 アジュバントの含有量が減少する。当業者であれば、いくつかの実施形態では、1つのトリテルペノイド系サポニンの含有量が減少する場合には、追加のトリテルペノイド系サポニン、または中間体、またはそれらの組み合わせも同一の時間の細胞内で減少し得ることを理解できるであろう。

**【 0 5 5 2 】**

いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部において、トリテルペノイド系サポニンの生合成中間体の含有量が減少する。生合成中間体は、化合物1、化合物2、メディカゲン酸（化合物5）、オレアノール酸（化合物3）、オレアノール酸 - 3 G l c A、アウグスチン酸（化合物4）、アウグスチン酸 - 3 G l c A、グリチルレチン酸（化合物13）、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部は、少なくとも2つのトリテルペノイド系サポニンの生合成中間体の含有量が減少する。生合成中間体は、化合物1、化合物2、メディカゲン酸（化合物5）、オレアノール酸（化合物3）、オレアノール酸 - 3 G l c A、アウグスチン酸（化合物4）、アウグスチン酸 - 3 G l c A、グリチルレチン酸（化合物13）、またはそれらの任意の組み合わせを含む。

10

20

**【 0 5 5 3 】**

いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部において、トリテルペノイド系サポニンの生合成中間体の含有量が減少し、生合成中間体は、化合物1を含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部において、トリテルペノイド系サポニンの生合成中間体の含有量が減少し、生合成中間体は、化合物2を含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部において、トリテルペノイド系サポニンの生合成中間体の含有量が減少し、生合成中間体は、メディカゲン酸（化合物5）を含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部において、トリテルペノイド系サポニンの生合成中間体の含有量が減少し、生合成中間体は、オレアノール酸（化合物3）を含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部において、トリテルペノイド系サポニンの生合成中間体の含有量が減少し、生合成中間体は、アウグスチン酸（化合物4）を含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部において、トリテルペノイド系サポニンの生合成中間体の含有量が減少し、生合成中間体は、アウグスチン酸 - 3 G l c A を含む。いくつかの実施形態では、少なくとも1つの遺伝子組換え植物細胞を含む遺伝子組換え植物またはその一部において、トリテルペノイド系サポニンの生合成中間体の含有量が減少し、生合成中間体は、グリチルレチン酸（化合物13）を含む。

30

40

**【 0 5 5 4 】**

いくつかの実施形態では、本明細書に開示される遺伝子組換え細胞は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その

50



いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が減少し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が増加する。

【0562】

いくつかの実施形態では、遺伝子組換え細胞または遺伝子組換え植物において、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体を産生する方法は、未組換え細胞または未組換え植物と比較して、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を改変し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量を減少させることを含む。

10

【0563】

いくつかの実施形態では、遺伝子組換え細胞または遺伝子組換え植物において、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体を産生する方法は、未組換え細胞または未組換え植物と比較して、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させ、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量を減少させることを含む。

20

【0564】

いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を減少させる方法は、少なくとも1つの植物またはその一部の少なくとも1つの細胞を遺伝子組換えすることを含み、遺伝子組換えされた植物細胞は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が減少し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変される。

30

【0565】

いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を減少させる方法は、少なくとも1つの植物またはその一部の少なくとも1つの細胞を遺伝子組換えすることを含み、遺伝子組換えされた植物細胞は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が減少し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が増加する。

40

【0566】

50

いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、少なくとも1つの植物またはその一部の少なくとも1つの細胞を遺伝子組換えすることを含み、遺伝子組換えされた植物細胞は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変される。

10

## 【0567】

いくつかの実施形態では、少なくとも1つのトリテルペ系ノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、少なくとも1つの植物またはその一部の少なくとも1つの細胞を遺伝子組換えすることを含み、遺伝子組換えされた植物細胞は、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が減少する。

20

## 【0568】

植物、植物の一部、及び細胞

## 【0569】

いくつかの実施形態では、酵素をコードする少なくとも1つの異種遺伝子を含む本明細書に記載の細胞または生物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、それぞれ対応する非組換え細胞または非トランスジェニック植物と比較して、増加している。他の実施形態では、ステロイド系アルカロイド経路、ステロイド系サポニン経路、またはトリテルペノイド合成経路の酵素をコードする少なくとも1つの遺伝子組換えされた内在性遺伝子を含む、本明細書に記載の少なくとも1つの遺伝子組換え細胞を含む遺伝子組換え植物またはその一部は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して、増加している。他の実施形態では、ステロイド系アルカロイド合成経路、ステロイド系サポニン合成経路、またはトリテルペノイド合成経路の酵素をコードする少なくとも1つの異種遺伝子を含む、本明細書に記載の少なくとも1つの遺伝子組換え細胞を含む遺伝子組換え植物またはその一部は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して、増加している。他の実施形態では、ステロイド系アルカロイド合成経路、ステロイド系サポニン合成経路、またはトリテルペノイド合成経路の酵素をコードする少なくとも1つの異種遺伝子を含む、本明細書に記載の少なくとも1つの遺伝子組換えされた内在性遺伝子を含む、本明細書に記載の少なくとも1つの遺伝子組換え細胞を含む遺伝子組換え植物またはその一部は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その

30

40

50

代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して、減少している。全体を通して記載されているように、当業者は、内在性遺伝子の組換えは、これに限定しないが、本明細書に詳細に記載されているように、遺伝子の発現を増加させること、遺伝子の発現を減少させること、遺伝子を突然変異させること、またはそれらの組み合わせを含むことを理解できるであろう。

【0570】

いくつかの実施形態では、CSLG酵素をコードする少なくとも1つの異種CSLG遺伝子を含む、本明細書に記載の細胞または生物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、それぞれ対応する非組換え細胞または非組換え植物と比較して、増加している。他の実施形態では、ステロイド系アルカロイド合成経路、ステロイド系サポニン合成経路、またはトリテルペノイド合成経路のCSLG酵素をコードする少なくとも1つの遺伝子組換えされた内在性CSLG遺伝子を含む、本明細書に記載の少なくとも1つの遺伝子組換え細胞を含む遺伝子組換え植物またはその一部は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して、増加している。他の実施形態では、ステロイド系アルカロイド合成経路、ステロイド系サポニン合成経路、またはトリテルペノイド合成経路の酵素をコードする少なくとも1つの異種CSLG遺伝子を含む、本明細書に記載の少なくとも1つの遺伝子組換え細胞を含む遺伝子組換え植物またはその一部は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して、増加している。他の実施形態では、ステロイド系アルカロイド合成経路、ステロイド系サポニン合成経路、またはトリテルペノイド合成経路の酵素をコードする少なくとも1つの異種CSLG遺伝子を含む、本明細書に記載の少なくとも1つの遺伝子組換え細胞を含む遺伝子組換え植物またはその一部は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して、減少している。

【0571】

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの誘導体は、それぞれ、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンのグリコシル化誘導体を含む。

【0572】

いくつかの実施形態では、例えば、少なくともサポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、CSLG酵素を通常は発現しない、及び/または、特定のステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、特定のステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、特定のトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体を通常は産生しない細胞、植物種、またはその一部は、サポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、CSLG酵素、あるいは、それらの任意の組み合わせを発現

10

20

30

40

50

するように遺伝子組み換えされる。いくつかの実施形態では、サポニン - アミリン合成酵素、シトクロム P 4 5 0、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、または CSLG 酵素を通常は発現しない、及び/または、特定のステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、特定のステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、特定のトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体を通常は産生しない細胞、植物種、または植物の一部は、ステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、それらの任意の組み合わせを産生するように遺伝子組み換えされる。他の実施形態では、所与の量のサポニン - アミリン合成酵素、シトクロム P 4 5 0、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、CSLG 酵素、またはそれらの任意の組み合わせを通常発現する細胞、植物種、または植物の一部は、サポニン - アミリン合成酵素、シトクロム P 4 5 0、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、CSLG 酵素、あるいは、それらの任意の組み合わせを過剰発現するように遺伝子組み換えされる。他の実施形態では、所与の量のサポニン - アミリン合成酵素、シトクロム P 4 5 0、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、CSLG 酵素、またはそれらの任意の組み合わせを通常発現する細胞、植物種、または植物の一部は、サポニン - アミリン合成酵素、シトクロム P 4 5 0、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、CSLG 酵素、あるいは、それらの組み合わせの発現を低下または抑制するように遺伝子組み換えされる。

10

20

### 【0573】

他の実施形態では、所与の量のステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体を通常産生する細胞、植物種、または植物の一部は、サポニン - アミリン合成酵素、シトクロム P 4 5 0、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、CSLG 酵素、あるいは、それらの組み合わせを過剰発現するように遺伝子組み換えされることにより、ステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、またはそれらの任意の組み合わせの産生量が増加する。

30

### 【0574】

他の実施形態では、所与の量のステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体を通常産生する細胞、植物種、または植物の一部は、サポニン - アミリン合成酵素、シトクロム P 4 5 0、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、CSLG 酵素、またはそれらの任意の組み合わせの発現を低下または抑制するように遺伝子組み換えされることにより、ステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、それらの任意の組み合わせの産生量が低減する。

40

### 【0575】

いくつかの実施形態では、遺伝子組換え植物の種子が本明細書に開示され、その種子から生育した植物は、対応する非組換え植物の種子から生育した植物と比較して、少なくとも

50

も1つのサポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG酵素を発現し、それにより、少なくとも1つのサポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG酵素の発現または過剰発現を含む、及び/または、ステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加する。

**【0576】**

本明細書に開示される細胞は、これに限定しないが、例えば、酵母、藻類、昆虫、細菌、及び植物細胞などの多数の品種を含む。いくつかの実施形態では、遺伝子組換え細胞は、植物細胞を含む。いくつかの実施形態では、遺伝子組換え細胞は、酵母細胞を含む。いくつかの実施形態では、遺伝子組換え細胞は、藻類細胞を含む。いくつかの実施形態では、遺伝子組換え細胞は、昆虫細胞を含む。いくつかの実施形態では、遺伝子組換え細胞は、細菌を含む。

10

**【0577】**

一般に、ほとんどの酵母細胞、藻類細胞、昆虫細胞、及び細菌は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを本質的には合成しない。いくつかの実施形態では、酵母細胞、藻類細胞、昆虫細胞、または細菌においてステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生するために、酵母細胞、藻類細胞、昆虫細胞、または細菌の遺伝子組換えは、トリテルペノイド系サポニン生合成経路の前駆体、基質、またはそれらの組み合わせを産生するために必要な酵素の導入を含む。例えば、出芽酵母を遺伝子組み換えすることにより、トリテルペノイド系サポニン生合成経路の基質に必要な前駆体を産生するために、異種UDP - グルコース6 - デヒドロゲナーゼIを発現させた下記の実施例14を参照されたい。いくつかの実施形態では、遺伝子組換え細胞は、ステロイド系アルカロイド生合成経路、ステロイド系サポニン生合成経路、またはトリテルペノイド生合成経路の前駆体、基質、またはそれらの組み合わせを産生するために、少なくとも1つの異種遺伝子を発現する。いくつかの実施形態では、遺伝子組換え細胞は、ステロイド系アルカロイド生合成経路、ステロイド系サポニン生合成経路、またはトリテルペノイド生合成経路の前駆体、基質、またはそれらの組み合わせを産生するために、少なくとも1つの異種遺伝子を発現し、ステロイド系アルカロイド生合成経路、ステロイド系サポニン生合成経路、またはトリテルペノイド生合成経路の酵素をコードする少なくとも1つの異種遺伝子を発現させる。

20

30

**【0578】**

いくつかの実施形態では、細胞は、微生物を含む。いくつかの実施形態では、遺伝子組換え細胞は、酵母細胞である。いくつかの実施形態では、酵母は、サッカロミセス目のものである。いくつかの実施形態では、酵母はサッカロミセス目のものであり、かつ、サッカロミセス属（例えば、出芽酵母）、シゾサッカロミセス属、ピチア属（例えば、ピチア酵母）、ハンセヌラ属（例えば、*Hansenula polymorpha*）、ヤロウイア属（例えば、*Yarrowia lipolytica*）、クリベロマイセス属（例えば、クリベロマイセス・ラクチス）、及びカンジダ属（例えば、カンジダ・アルピカンス、トルラ酵母）からなる群から選択される。いくつかの実施形態では、酵母は、サッカロミセス属、シゾサッカロミセス属、ピチア属、ヤロウイア属、クリベロマイセス属、またはカンジダ属から選択される。いくつかの実施形態では、酵母は、出芽酵母またはカンジダ・アルピカンスから選択される。

40

**【0579】**

いくつかの実施形態では、本明細書に記載のステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法において、遺伝子組換え酵母を利用することができる。

**【0580】**

50

いくつかの実施形態では、遺伝子組換え細胞は、藻類細胞である。いくつかの実施形態では、細胞は、これらに限定しないが、例えば緑藻植物（緑藻類）、紅藻植物（紅藻類）、または褐藻植物（褐藻類）などの様々な藻類のいずれかである。いくつかの実施形態では、緑藻植物は、クラミドモナス目またはクロレラ目のものである。いくつかの実施形態では、緑藻植物は、クラミドモナス目のものであり、クラミドモナス属（例えば、コナミドリムシ）及びドナリエラ属からなる群から選択される。いくつかの実施形態では、緑藻植物は、クロレラ目に属し、クロレラ属から選択される。

【0581】

いくつかの実施形態では、本明細書に記載のステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法において、遺伝子組換え藻類を利用することができる。

10

【0582】

いくつかの実施形態では、本明細書で開示される遺伝子組換え細胞は、細菌細胞である。一実施形態では、細菌細胞は、大腸菌細胞である。別の実施形態では、細胞は、*Acremonium rutilum*、ニホンコウジカビ、*Yarrowia lipolytica*、*Bacillus* sp. JPJ、*Brevundimonas* sp. SGJ、*E. herbicola*、シトロバクター・フロインディ、*Symbiobacterium*、または緑膿菌の細胞である。別の実施形態では、細菌細胞は、乳製品の発酵に關与する細菌である。一実施形態では、細菌は、乳酸連鎖球菌である。別の実施形態では、細菌はラクトバシラスである。一実施形態では、ラクトバシラスは、ラクトバチルス・ブルガリカスである。別の実施形態では、細菌は、ラクトコッカスまたはロイコノストックである。

20

【0583】

いくつかの実施形態では、本明細書に記載のステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法に、遺伝子組換え細菌を利用することができる。

【0584】

いくつかの実施形態では、本明細書で開示される遺伝子組換え植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変されている。

30

【0585】

いくつかの実施形態では、本明細書に開示される遺伝子組換え植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が減少している。

40

【0586】

いくつかの実施形態では、本明細書で開示される遺伝子組換え植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体

50

、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が減少し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が改変されている。

【0587】

いくつかの実施形態では、本明細書で開示される遺伝子組換え植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が減少し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が増加している。

10

【0588】

いくつかの実施形態では、遺伝子組換え細胞または遺伝子組換え植物において、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体を産生する方法は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を、非組換え細胞または非組換え植物と比較して、増加させることと、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量を改変させることと、を含む。

20

30

【0589】

いくつかの実施形態では、遺伝子組換え細胞または遺伝子組換え植物において、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体を産生する方法は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を、非組換え細胞または非組換え植物と比較して、増加させることと、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量を減少させることと、を含む。

40

【0590】

いくつかの実施形態では、少なくとも1つのステロイド系アルカロイド、その誘導体、

50



イド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、少なくとも1つの植物細胞、植物またはその一部の少なくとも1つの細胞の遺伝子を組み換える遺伝子組換えステップであって、遺伝子組換え植物細胞は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が増加し、かつ、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、あるいは、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量が減少している、該遺伝子組換えステップを含む。

10

**【0594】**

いくつかの実施形態では、遺伝子組換え細胞は、植物細胞である。植物細胞は、後述するように、広範囲の植物から得ることができる。植物細胞は、葉細胞、葉柄細胞、植物の茎または柄細胞、根細胞、芽細胞、塊茎細胞、豆細胞、穀物または穀粒細胞、果実細胞、ナッツ細胞、マメ細胞、種子または種子細胞、包葉細胞、カルス細胞、及び花細胞であってもよい。いくつかの実施形態では、遺伝子組換え細胞は、したがって、植物の葉、植物の葉柄、植物の茎または柄、植物の根、植物の芽、植物の塊茎、植物の豆、植物の穀物または穀粒、植物の果実、植物のナッツ、植物のマメ、植物の種子、植物の包葉、植物のカルス、または植物の花に含まれていてもよい。

20

**【0595】**

いくつかの実施形態では、遺伝子組換え植物細胞は、植物またはその一部とは独立して、培養で育てられる。植物及びその一部を培養する方法は、当該技術分野でよく知られており、例えば、「Ochoa-Villarreal M, Howat S, Hong S, et al. Plant cell culture strategies for the production of natural products. BMB Rep. 2016; 49(3):149-158」を参照されたい。この内容はその全体が本明細書に組み込まれている。

**【0596】**

いくつかの実施形態では、遺伝子組換え植物細胞は、植物またはその一部とは独立して、培養で育てられ、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生するために本明細書に提供される方法で利用される。遺伝子組換え細胞の懸濁液及び遺伝子組換え細胞に由来する組織培養物も、本明細書に記載の範囲に包含される。細胞懸濁液及び組織培養物は、所望のステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、所望のステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、所望のトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の産生のために使用することができ、これらはその後、細胞または成長培地から抽出される。あるいは、遺伝子組換え植物の細胞及び/または組織を用いて、サポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG酵素の発現または過剰発現を改変させたトランスジェニック植物を再生することができ、したがって、サポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、UDP-グルコース6-デヒドロゲナーゼ1、セルロース合成酵素様G(CSLG)は、それぞれ対応する非組換え細胞または非組換え植物と比較して、細胞内または植物内で発現または過剰発現する。したがって、ステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が改変される。

30

40

**【0597】**

いくつかの実施形態では、遺伝子組換え植物細胞は、植物に含まれる。いくつかの実施

50

形態では、遺伝子組換え植物細胞は、植物に含まれ、植物は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生するために本明細書に提供される方法に利用されている。いくつかの実施形態では、遺伝子組換え植物細胞は、植物組織に含まれる。いくつかの実施形態では、遺伝子組換え植物細胞は、植物組織に含まれ、植物組織は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生するために本明細書に提供される方法に利用されている。いくつかの実施形態では、遺伝子組換え植物細胞は、植物器官に含まれる。いくつかの実施形態では、遺伝子組換え植物細胞は、植物器官に含まれ、植物器官は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生するために本明細書に提供される方法に利用されている。いくつかの実施形態では、遺伝子組換え植物細胞は、植物の一部に含まれる。いくつかの実施形態では、遺伝子組換え植物細胞は、植物の一部に含まれ、植物の一部は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生するために本明細書に提供される方法に利用されている。

10

**【0598】**

いくつかの実施形態では、遺伝子組換え細胞は、葉細胞、葉柄細胞、植物の茎または柄細胞、根細胞、芽細胞、塊茎細胞、豆細胞、穀物または穀粒細胞、果実細胞、ナッツ細胞、マメ細胞、種子または種子細胞、包葉細胞、カルス細胞、または花細胞を含む植物細胞を含む。いくつかの実施形態では、遺伝子組換え細胞は、植物の葉、植物の葉柄、植物の茎または柄、植物の根、植物の芽、植物の塊茎、植物の豆、植物の穀物または穀粒、植物の果実、植物のナッツ、植物のマメ、植物の種子、植物の包葉、植物のカルス、または植物の花、またはそれらの組み合わせに含まれる植物細胞を含む。いくつかの実施形態では、遺伝子組換え植物細胞は、植物の一部に含まれ、植物の一部は、葉、葉柄、植物の茎または柄、根、芽、塊茎、豆、穀物または穀粒、果実、ナッツ、マメ科植物、種子または種、包葉、カルス、または花を含む。

20

**【0599】**

いくつかの実施形態では、本明細書に記載のステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法に、遺伝子組換え植物細胞を利用することができる。

**【0600】**

いくつかの実施形態では、遺伝子組換え細胞は、遺伝子組換え植物細胞（例えば、培養植物細胞）を含む。他の実施形態では、遺伝子組換え植物細胞は、植物または植物器官、植物組織、または植物の一部に含まれる。いくつかの実施形態では、「植物」という用語は、すべての同一の意味及び品質を有する「植物の一部」という用語と互換的に使用されてもよい。

30

**【0601】**

いくつかの実施形態では、遺伝子組換え植物細胞、植物、植物器官、植物組織、または植物の一部が、本明細書に記載のステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法に利用されてもよい。

**【0602】**

遺伝子組換え細胞、植物、植物の一部に含まれるステロイド系アルカロイド、ステロイド系サポニン、トリテルペノイド系サポニン、それらの誘導体、それらの代謝物、またはそれらの生合成中間体の含有量は、以下に例示するように、または当業者に知られているように、測定される。

40

**【0603】**

いくつかの実施形態では、子孫植物は、ステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、少なくとも1つの祖先植物と比較して、増加している。

50

## 【0604】

いくつかの実施形態では、本明細書で開示される植物、植物細胞、植物組織、または植物の一部は、ナデシコ目、ナス目、ナス目、アオイ目、セリ目、フウチョウソウ目、イネ目、シソ目、ウリ目、キク目、キントラノオ目、バラ目、ブナ目、ムクロジ目、ヤシ目、ツツジ目、リンドウ目、キンポウゲ目、ショウガ目、ユキノシタ目、ブドウ目、マツ目、ミズキ目、クサスギカズラ目、ヤマノイモ目、及びユリ目から選択される。いくつかの実施形態では、本明細書に開示される植物または植物細胞、植物組織または植物の一部は、ナデシコ目、ナス目、ナス目、アオイ目、セリ目、フウチョウソウ目、イネ目、シソ目、ウリ目、キク目、キントラノオ目、バラ目、ブナ目、ムクロジ目、ヤシ目、ツツジ目、リンドウ目、キンポウゲ目、ショウガ目、ユキノシタ目、ブドウ目、マツ目、ミズキ目、ク

10

## 【0605】

いくつかの実施形態では、本明細書に開示される植物、植物細胞、植物組織、または植物の一部は、イネ目、ナデシコ目、ナス目、マメ目、アオイ目、セリ目、フウチョウソウ目、クサスギカズラ目、ヤマノイモ目、またはユリ目の植物からの細胞を含む。

## 【0606】

いくつかの実施形態では、本明細書に開示される植物または植物細胞または植物組織または植物部分は、ナデシコ目に属し、ホウレンソウ属（例えば、ホウレンソウ）、アカザ属（例えば、キノア）、ベータ属（例えば、ビートルート）、レウム属（例えば、Rheum hybridum、ルバーブ、Rheum ribes）、ドウカンソウ属（例えば、ドウカンソウ）、サボ

20

## 【0607】

いくつかの実施形態では、本明細書に開示される植物または植物細胞または植物組織または植物部分は、ナス目に属し、タバコ属（例えば、ベンサミアナタバコ）、ナス属（例えば、トマト、野生型トマト）、ジャガイモ（*S. chacoense*、*S. tuberosum*）、ナス、ピターズスイート、トウガラシ属（例えば、トウガラシ）、ヒヨス属、ダチュラ属、及びオオカミナスビ属からなる群から選択される。いくつかの実施形態では、本明細書に開示される植物は、ナス科植物を含む。いくつかの実施形態では、ナス科植物は、トマト、野生型

30

## 【0608】

いくつかの実施形態では、植物はマメ目に属し、グリシリザ属（例えば、ウラルカンゾウ、スペインカンゾウ〔甘草〕）、メディカゴ属（例えば、アルファルファ、タルウマゴヤシ）、キラヤ属（例えば、キラヤ）、ダイズ属（例えば、ダイズ〔ダイズ〕）、ハス属（例えば、ミヤコグサ）、シセル属（例えば、ヒヨコマメ〔ヒヨコマメ、ガルバンゾマメ

40

## 【0609】

いくつかの実施形態では、植物はアオイ目に属し、カカオ属（例えば、カカオ）から選択される。

## 【0610】

いくつかの実施形態では、植物はセリ目に属し、ニンジン属（例えば、ノラニンジン）

50

、オランダミツバ属（例えば、セロリ）、パセリ属（例えば、パセリ）、オタネニンジン属（例えば、オタネニンジン）、ミシマサイコ属、キツタ属、及びツボクサ属（例えば、ツボクサ）からなる群から選択される。

【0611】

いくつかの実施形態では、植物はフウチョウソウ目に属し、シロイヌナズナ属（例えば、シロイヌナズナ）、アブラナ属（例えば、ヤセイカンラン [キャベツ]、カラシナ [ホワイトマスタード]、クロガラシ [ブラックマスタード]、セイヨウアブラナ）、フウチョウボク属（例えば、ケッパー [ケッパー]）、及びパパイヤ属（例えば、パパイヤ [パパイヤ]）からなる群から選択される。

【0612】

いくつかの実施形態では、植物はイネ目に属し、イネ属（例えば、イネ及びアフリカイネ [米]）、オオムギ属（例えば、オオムギ [大麦]）、カラスムギ属（例えば、オーツ麦 [オーツ麦]、アウエナストリゴサ）、及びコムギ属（例えば、スペルト小麦 [スペルト]）からなる群から選択される。

【0613】

いくつかの実施形態では、植物はシソ目に属し、アキギリ属（例えば、チア [チア]）、ゴマ属（例えば、ゴマ [ゴマ]）、及びオリーブ属（例えば、オリーブ [オリーブ]）からなる群から選択される。

【0614】

いくつかの実施形態では、植物はウリ目に属し、カボチャ属（例えば、これに限定しないが、ペポカボチャ、セイヨウカボチャ、*Cucurbita argyrosperma*、またはニホンカボチャを含むカボチャ）から選択される。

【0615】

いくつかの実施形態では、植物はキク目に属し、ヘリアンサス属（例えば、ヒマワリ [ヒマワリ]、*Helianthus verticillatus* [渦巻きヒマワリ]、クイモ [エルサレムアーティチョーク]）、ヨモギ属（例えば、クソニンジン）、*Galatella*（シオン）属（例えば、*Galatella sedifolia*）、タンポポ属（例えば、セイヨウタンポポ [タンポポ]）から選択される。

【0616】

いくつかの実施形態では、植物はキントラノオ目に属し、アマ属（例えば、アマニ [アマニ]）、オヒルギ属（例えば、オヒルギ）、トウダイグサ属（ミドリサンゴ）、トウゴマ属（トウゴマ [トウゴマ]）、メヒルギ属（メヒルギ）、及びヤエヤマヒルギ属（ヤエヤマヒルギ）からなる群から選択される。

【0617】

いくつかの実施形態では、植物はバラ目に属し、サクラ属（例えば、アーモンド [アーモンド]、*Prunus amygdalus*）及びアサ属（例えば、*Cannabis sativa*を含むアサ）からなる群から選択される。

【0618】

いくつかの実施形態では、植物はブナ目に属し、ハシバミ属（例えば、これに限定しないがセイヨウハシバミなどを含む、ヘーゼル/ヘーゼルナッツ/コブナッツ/フィルバートナッツ）、カバノキ属（例えば、シラカンバ [シルバーバーチ]、ヨーロッパダケカンバ [white or downy birch]、シラカンバ）、及びクルミ属（例えば、樫ぐるみ [ペルシヤグルミまたはセイヨウグルミ]、黒ぐるみ [黒ぐるみ]、バターナット [バターナット]）からなる群から選択される。

【0619】

いくつかの実施形態では、植物はムクロジ目に属し、カシューナッツ属（例えば、カシューナッツ [カシュー]）、カイノキ属（例えば、ピスタチオ [ピスタチオ]）、ミカン属（多数の種及び雑種）、トチノキ属（例えば、セイヨウトチノキ、トチノキ）、及びペガヌム属からなる群から選択される。

【0620】

10

20

30

40

50

いくつかの実施形態では、植物はヤシ目に属し、ココヤシ属（例えば、ココヤシ [ココナッツ]）から選択される。

【0621】

いくつかの実施形態では、植物はツツジ目に属し、イズセンリョウ属から選択される。

【0622】

いくつかの実施形態では、植物は、リンドウ目に属し、キョウチクトウ属（例えば、キョウチクトウ）、リンドウ属（例えば、リンドウマクロフィラ）、ニチニチソウ属（例えば、ニチニチソウ）、インドジャボク属、及びキナ属からなる群から選択される。

【0623】

いくつかの実施形態では、植物はキンボウゲ目に属し、クロタネソウ属（例えば、ニオイクロタネソウ）、ケシ属、ハナビシソウ属（例えば、ハナビシソウ）、オウレン属、メギ属、及びカラマツソウ属からなる群から選択される。

10

【0624】

いくつかの実施形態では、植物はショウガ目に属し、オオホザキアヤメ属（例えば、オオホザキアヤメ）から選択される。

【0625】

いくつかの実施形態では、植物はユキノシタ目に属し、カラニコエ属（例えば、コダカラベンケイ）から選択される。

【0626】

いくつかの実施形態では、植物はブドウ目に属し、ブドウ属（例えば、ヨーロッパブドウ [ブドウ]）から選択される。

20

【0627】

いくつかの実施形態では、植物はマツ目に属し、イチイ属（例えば、タイヘイヨウイチイ、ヨーロッパイチイ、イチイ、カナダイチイ、フロリダイチイ）から選択される。

【0628】

いくつかの実施形態では、植物は、ミズキ目に属し、カンレンボク属（例えば、カンレンボク）から選択される。いくつかの実施形態では、植物は、クサスギカズラ目に属し、リュウゼツラン属（例えば、アオノリュウゼツラン、アガベ・アテナータ、テキーラ）、クサスギカズラ属（例えば、アスパラガス）、またはイトラン属（例えば、イトラン）から選択される。いくつかの実施形態では、植物は、ヤマノイモ目に属し、Borderera属、ヤマノイモ属、Epipetrum属、Rajania属、ステノメリス属、またはティムス属から選択される。いくつかの実施形態では、植物は、ユリ目に属し、ユリ科（例えば、Clintonia borealis、ノモカリス・アベルタ、カロコルツスカタリナエ、Streptopus lariceolatusなど）から選択される。

30

【0629】

当業者であれば、植物の品種改良は、繁殖のために望ましい特性を有する植物を単に選択することから、遺伝学や染色体の知識を利用した方法、さらに複雑な分子技術まで、様々な手法で達成され得ることを理解できるであろう。

【0630】

当業者であれば、「ハイブリッド植物」という用語は、対象となる2つの植物を交配し、種子または組織によって繁殖させ、その後、植物を成長させることによって生成された植物を包含することを理解できるであろう。植物を有性交配する場合、受粉は、交配、自家受粉、または、形質転換されていない植物または別の形質転換された植物との戻し交配を含み得る。ハイブリッド植物は、第一世代及び第二世代の植物を含む。

40

【0631】

生物学的活性

【0632】

いくつかの実施形態では、遺伝子組換え細胞または遺伝子組換え植物は、少なくとも1つの異種遺伝子によってコードされる酵素と同一または類似の活性を有する内在性酵素を含む。他の実施形態では、遺伝子組換え細胞または遺伝子組換え植物は、少なくとも1つ

50

の異種遺伝子によってコードされる酵素と同一または類似の活性を有する内在性酵素を含んでいない。いくつかの実施形態では、遺伝子組換え植物細胞または遺伝子組換え植物は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニン合成酵素の活性が改変された内在性酵素を含む。

#### 【0633】

いくつかの実施形態では、異種遺伝子によってコードされる少なくとも1つのサポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、CSLG 酵素、またはそれらの組み合わせは、遺伝子組換え細胞または遺伝子組換え植物に存在する内在性のサポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLGタンパク質と比較して、改変されている。いくつかの実施形態では、少なくとも1つの内在性サポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、CSLG 酵素、またはそれらの組み合わせは、非遺伝子組換え細胞または非遺伝子組換え植物に存在する内在性のサポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLGタンパク質と比較して、改変されている。

10

#### 【0634】

当業者であれば、「生物学的活性」という用語が、アッセイによって測定できるタンパク質に関連する任意の活性を指すことを認識するであろう。いくつかの実施形態では、サポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG 酵素の生物活性は、ステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体を、細胞内及び植物内で生合成するために必要な酵素活性を有している。いくつかの実施形態では、サポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG 酵素の生物学的活性は、植物の少なくとも一部における、ステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量に影響を与える。いくつかの実施形態では、改変された生物学的活性は、酵素活性の増加を含む。いくつかの実施形態では、改変された生物学的活性は、酵素活性の低下を含む。いくつかの実施形態では、改変された生物学的活性は、ポリペプチドの安定性の向上を含む。いくつかの実施形態では、改変された生物学的活性は、ポリペプチドの安定性の低下を含む。いくつかの実施形態では、改変された生物学的活性は、酵素ポリペプチドの発現が増加して高レベルになることを含む。いくつかの実施形態では、改変された生物学的活性は、酵素ポリペプチドの発現の減少を含む。

20

30

40

#### 【0635】

いくつかの実施形態では、酵素の生物学的活性は、対照酵素と比較して、改変されている。

#### 【0636】

いくつかの実施形態では、改変された生物活性は、(a) サポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG 酵素の酵素活性の増加、サポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG 酵素の安定性の向上、サポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラー

50

ゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG酵素の酵素活性の低下、あるいは、サポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG酵素の安定性の低下を含む。

【0637】

いくつかの実施形態では、サポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG酵素の生物学的活性が増加する。いくつかの実施形態では、サポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG酵素の生物学的活性が低下する。いくつかの実施形態では、サポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG酵素は、安定性が向上している。いくつかの実施形態では、サポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG酵素は、安定性が低下している。

10

【0638】

いくつかの実施形態では、改変された生物学的活性は、遺伝子組換え細胞または遺伝子組換え植物に内在性酵素が存在する場合、その生物学的活性と比較して、少なくとも1つのサポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG酵素の酵素活性の増加、少なくとも1つのサポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG酵素の安定性の向上、少なくとも1つのサポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG酵素の酵素活性の低下、あるいは、少なくとも1つのサポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG酵素の安定性の低下を含む。

20

【0639】

トランスジェニック植物細胞、トランスジェニック酵母細胞、トランスジェニック藻類細胞、トランスジェニック昆虫細胞、及びトランスジェニック細菌

30

【0640】

少なくとも1つの異種遺伝子によってコードされる酵素をコードするポリヌクレオチドのクローニングは、当業者に知られている任意の方法によって行うことができる。この酵素は、少なくとも1つのサポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼまたはCSLG酵素を含む。本明細書に記載の少なくとも1つのサポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG酵素をコードするポリヌクレオチドのクローニングは、当業者に知られている任意の方法によって行うことができる。特定の実施形態では、CSLG酵素をコードするポリヌクレオチドのクローニングが作製される。様々なDNA構築物を使用することにより、所望の細胞または生物において所望の遺伝子を発現させることができる。いくつかの実施形態では、様々なDNA構築物を使用することにより、細胞内でCSLG遺伝子を発現させる。いくつかの実施形態では、様々なDNA構築物を使用することにより、植物細胞、酵母細胞、昆虫細胞、藻類細胞、または細菌においてCSLG遺伝子を発現させる。いくつかの実施形態では、様々なDNA構築物を使用することにより、植物細胞においてCSLG遺伝子を発現させる。いくつかの実施形態では、様々なDNA構築物を使用することにより、酵母細胞においてCSLG遺伝子を発現させる。いくつかの実施形態では、様々なDNA構築物を使用することにより

40

50

、植物細胞または酵母細胞においてCSLG遺伝子を発現させる。いくつかの実施形態では、様々なDNA構築物を使用することにより、藻類細胞においてCSLG遺伝子を発現させる。いくつかの実施形態では、様々なDNA構築物を使用することにより、昆虫細胞においてCSLG遺伝子を発現させる。いくつかの実施形態では、様々なDNA構築物を使用することにより、細菌においてCSLG遺伝子を発現させる。いくつかの実施形態では、様々なDNA構築物を使用することにより、細胞内で異種CSLG遺伝子を発現させる。いくつかの実施形態では、細胞が植物細胞である場合、CSLG遺伝子は異種遺伝子である。いくつかの実施形態では、細胞が植物細胞である場合、CSLG遺伝子は相同遺伝子であり、発現が改変（例えば、増加または減少）されるか、または、CSLG遺伝子が突然変異することによりCSLG酵素の活性が改変される。いくつかの実施形態では、細胞が植物細胞である場合、植物細胞は、植物またはその一部に含まれる。

10

**【0641】**

いくつかの実施形態では、遺伝子は、遺伝子の発現に必要なすべての要素と任意の調節要素を含む発現ベクターの一部を含む。

**【0642】**

いくつかの実施形態では、発現は、構成的プロモータによって制御される。他の実施形態では、発現は、一過性プロモータによって制御される。いくつかの実施形態では、構成的プロモータは、細胞、または、植物または植物組織に対して特異的である。いくつかの実施形態では、組織特異的プロモータは、根特異的プロモータ、塊茎特異的プロモータ、葉特異的プロモータ、及び果実特異的プロモータからなる群から選択される。根特異的プロモータは、例えば、「Martinez, E. et al. 2003. Curr. Biol. 13:1435-1441」に記載されている。果実特異的プロモータは、とりわけ、「Estornell L.H et al. 2009. Plant Biotechnol. J. 7:298-309」及び「Fernandez A. I. Et al. 2009 Plant Physiol. 151: 1729-1740」に記載されている。塊茎特異的プロモータは、例えば、「Rocha-Sosa M, et al., 1989. EMBO J. 8:23-29」及び「McKibbin R.S. et al., 2006. Plant Biotechnol J. 4(4):409-18」に記載されている。葉特異的プロモータ、例えば、「Yutao Yang, Guo dong Yang, Shijuan Liu, Xingqi Guo and Chengchao Zheng. Science in China Series C: Life Sciences. 46: 651-660」に記載されている。したがって、いくつかの実施形態では、プロモータは、酵母特異的プロモータを含む。したがって、いくつかの実施形態では、プロモータは、藻類特異的プロモータを含む。

20

30

**【0643】**

いくつかの実施形態では、発現ベクターは、3'非コード配列に制御要素をさらに含む。本明細書で使用するとき、「3'非コード配列」は、コード配列の下流に位置するDNA配列を指し、ポリアデニル化認識配列、及び、mRNAの処理または遺伝子発現に影響を与えることができる調節シグナルをコードする他の配列を含む。ポリアデニル化シグナルは、通常、mRNA前駆体の3'末端へのポリアデニル酸部分の付加に影響を与えることを特徴とする。植物の場合、異なる3'非コード配列の使用は、「Ingelbrecht I L et al. (1989. Plant Cell 1:671-680)」に例示されている。

**【0644】**

当業者であれば、本明細書に記載された核酸配列及び形質転換ベクターの様々な構成要素が、核酸または核酸断片の発現をもたらすように、作動可能に連結されることを理解できるであろう。本明細書に記載の構築物及びベクターの構成要素を作動的に連結する技術は、当業者にはよく知られている。そのような技術には、例えば1以上の制限酵素部位を含む合成リンカーなどのリンカーの使用が含まれる。

40

**【0645】**

当業者であれば、「作動可能に連結される」という用語は、一方の機能が他方の機能によって制御されるように、単一の核酸断片上の核酸配列が連結していることを指す。例えば、プロモータは、そのプロモータがそのコード配列の発現を調節することができる場合（すなわち、コード配列がプロモータの転写制御下にある場合）、コード配列と作動可能に連結される。コード配列は、センスまたはアンチセンスの方向で制御配列に作動可能に

50

連結することができる。

【0646】

本明細書で開示されている教示にしたがって植物を形質転換する方法は、当業者に知られている。本明細書で使用するとき、「形質転換」または「形質転換」という用語は、発現ベクターを含むDNA構築物などの外来DNAが受容細胞に入り、その受容細胞を、形質転換された、遺伝的に改変された、またはトランスジェニック細胞に改変するプロセスを表す。形質転換は、核酸配列が生物のゲノムに組み込まれ、安定的な遺伝性の形質を表す安定的なものであってもよいし、または、核酸配列が形質転換された細胞によって発現されるがゲノムに組み込まれず、一過性の形質を表す一過性のものであってもよい。いくつかの実施形態では、本明細書に開示される核酸配列は、細胞内に安定的に形質転換される。

10

【0647】

本発明の遺伝子組換え細胞または遺伝子組換え植物は、所望の少なくとも1つのサポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLGの含有量が改変し、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が改変される。いくつかの実施形態では、本明細書に記載されているように産生された遺伝子組換え細胞または遺伝子組換え植物は、一般的に、まず、遺伝子またはタンパク質の発現に基づいて選択される。少なくとも1つの異種遺伝子からコードされる酵素を発現する細胞または植物を同定し、次いで、少なくとも1つのサポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG、及び/または、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を分析する。

20

【0648】

いくつかの実施形態では、まず、所望の少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が改変された遺伝子組換え植物を、本明細書の教示にしたがって、遺伝子またはタンパク質の発現に基づいて選択する。次いで、遺伝子またはタンパク質の発現が増強または異常な植物について、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を分析する。いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの誘導体は、それぞれ、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンのグリコシル化誘導体を含む。

30

40

【0649】

少なくとも1つのサポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、またはCSLG遺伝子、及び/または、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の検出は、

50

当業者に知られている分子遺伝学の標準的な方法を用いて行われる。同様に、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の精製及び/または検出は、当業者に知られている分子遺伝学またはタンパク質化学の標準的な方法を用いて行われる。

**【0650】**

遺伝子を検出するためには、核酸を発現する細胞または植物からcDNAまたはmRNAを取得する必要がある。その試料は、検出プロセスの前に、さらに処理してもよい。例えば、細胞または組織の試料中のポリヌクレオチドを、試料の他の構成要素から分離してもよいし、または、増幅してもよい。或る生物から取得したすべての試料は、任意の種類のさらなる処理に供したものも含めて、その生物から取得したものと見なされる。

10

**【0651】**

遺伝子またはサイレンシング分子の発現を検出するためには、核酸を発現する器官からcDNAまたはmRNAを取得する必要がある。その試料は、検出プロセスの前に、さらに処理してもよい。例えば、細胞または組織の試料中のポリヌクレオチドを、試料の他の構成要素から分離してもよいし、または、増幅してもよい。或る生物から取得したすべての試料は、任意の種類のさらなる処理に供したものも含めて、その生物から取得したものと見なされる。

20

**【0652】**

いくつかの実施形態では、遺伝子の検出には、候補となる遺伝子組換え細胞、遺伝子組換え植物またはその一部から取得したポリヌクレオチドの増幅を必要とする。いくつかの実施形態では、植物の一部は、植物の器官、または植物の組織を含む。

**【0653】**

DNAを増幅する方法は、当業者に知られている。最も一般的に使用されるDNA増幅方法は、PCR法である（ポリメラーゼ連鎖反応；例えば、「PCR Basics: from background to Bench, Springer Verlag, 2000; Eckert et al., 1991. PCR Methods and Applications 1:17」を参照）。別の適切な増幅方法としては、リガーゼ連鎖反応（LCR）、転写増幅及び自己持続型配列複製、及び核酸ベースの配列増幅（NASBA）が挙げられる。

30

**【0654】**

いくつかの実施形態では、少なくとも1つのサポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、グルクロノシルトランスフェラーゼ、CSLG、またはそれらの任意の組み合わせを含む核酸配列は、いくつかの実施形態では、選択可能なマーカーをコードする核酸配列をさらに含む。いくつかの実施形態では、選択可能なマーカーは、抗生物質または除草剤に対する耐性を付与し、これらの実施形態では、トランスジェニック細胞または植物は、抗生物質または除草剤に対する耐性にしたがって選択される。

**【0655】**

育種

40

**【0656】**

いくつかの実施形態では、導入遺伝子編集による育種、導入遺伝子の使用、遺伝子または遺伝子の一過性発現の使用、分子マーカーの使用、またはそれらの任意の組合せを含む形質転換技術が、改変された発現を有する植物の育種に使用される。形質転換技術が組織培養の使用を必要とする場合、当業者によく知られた技術にしたがって、形質転換された細胞を植物に再生する。再生された植物は、その後、従来の育種技術を用いて同一または異なる植物品種と交配して種子を生産した後、適切な条件で選択される。

**【0657】**

ステロイド系アルカロイド及び/またはステロイド系サポニンの含有量は、以下に例示するように測定され、当業者には公知である。

50

## 【0658】

いくつかの実施形態では、子孫植物は、その祖先の植物の少なくとも1つと比較して、抗栄養成分または毒素の含有量の減少を含む。いくつかの実施形態では、子孫植物は、その祖先の植物の少なくとも1つと比較して、植物病原体、害虫、または捕食動物に対する改善された耐性を有する。

## 【0659】

いくつかの実施形態では、本明細書に開示される植物は、上述した目のいずれかの作物植物を含む。いくつかの実施形態では、本明細書に開示される植物は、イネ目、ナデシコ目、ナス目、マメ目、アオイ目、セリ目、フウチョウソウ目、クサスギカズラ目、ヤマノイモ目、または、ユリ目の作物植物を含む。

10

## 【0660】

いくつかの実施形態では、植物がイネ目の作物植物である場合、植物は、イネ属、オオムギ属、カラスムギ属、及びコムギ属からなる群から選択される。

## 【0661】

いくつかの実施形態では、植物がナデシコ目の作物植物である場合、植物は、ホウレンソウ属、アカザ属、ベータ属、レウム属、ドウカンソウ属、サボナリア属、及びジブソフィラ属からなる群から選択される。いくつかの実施形態では、ナデシコ目の植物は、ホウレンソウ、ビートルート、及びキノアからなる群から選択される。

## 【0662】

いくつかの実施形態では、植物がナス目の作物植物である場合、植物は、ナス属、トウガラシ属、タバコ属、ヒヨス属、ダチュラ属、及びオオカミナスビ属からなる群から選択される。いくつかの実施形態では、植物がナス目の作物植物である場合、植物は、トマト、野生型トマト、ジャガイモ、野生種ジャガイモ、トウガラシ、ナスビ、ズルカマラ、及びベンサミアナタバコからなる群から選択される。いくつかの実施形態では、ナス目の植物は、グランドチェリー、ナス、ジャガイモ、トマト、ペッパー、ピーマン、トウガラシ、チリペッパー、パプリカ、タバスコペッパー、タバコ、及びビタースイートからなる群から選択される。いくつかの実施形態では、ナス目の植物は、ナス科の植物を含む。

20

## 【0663】

いくつかの実施形態では、植物がマメ目の植物である場合、植物は、グリシリザ属、メディカゴ属、ダイズ属、ハス属、シセル属、インゲンマメ属、エンドウ属、ラッカセイ属、ルピナス属、及びアカシア属からなる群から選択される。いくつかの実施形態では、マメ目の植物は、アルファルファ、大豆、ミヤコグサ、及び甘草からなる群から選択される。

30

## 【0664】

いくつかの実施形態では、植物がアオイ目の植物である場合、植物は、カカオ属から選択される。

## 【0665】

いくつかの実施形態では、植物がセリ目の植物である場合、植物は、ニンジン属、オランダミツバ属、パセリ属、オタネニンジン属、ミシマサイコ属、キツタ属、及びツボクサ属からなる群から選択される。

40

## 【0666】

いくつかの実施形態では、植物がフウチョウソウ目の植物である場合、植物は、シロイヌナズナ属、アブラナ属、フウチョウボク属、及びパイヤ属からなる群から選択される。

## 【0667】

いくつかの実施形態では、植物は、本明細書に記載の任意の目、科、属、または種の植物である。

## 【0668】

当業者であれば、植物の品種改良は、繁殖のために望ましい特性を有する植物を単に選択することから、遺伝学や染色体の知識を利用した方法、さらに複雑な分子技術まで、様

50

々な手法で達成され得ることを理解できるであろう。

【0669】

当業者であれば、「ハイブリッド植物」という用語は、対象となる2つの植物を交配し、種子または組織によって繁殖させ、その後、植物を成長させることによって生成された植物を包含することを理解できるであろう。植物を有性交配する場合、受粉のステップは、交配、自家受粉、または、形質転換されていない植物若しくは別の形質転換された植物との戻し交配を含み得る。ハイブリッド植物は、第一世代及び第二世代の植物を含む。本明細書には、非限定的な例として、植物の抵抗性を向上させること、植物の抗栄養特性を低下させること、植物の毒素または苦味の化合物を減少させること、植物の医薬特性を向上させること、またはそれらの任意の組み合わせを含む、植物の形質を操作し改善するための方法が開示されている。

10

【0670】

バイオマーカー

【0671】

当業者であれば、「バイオマーカー」という用語は、生物において測定可能な物質を含み、この物質の存在が、対象の生物学的状態または状態を示すことを理解できるであろう。いくつかの実施形態では、バイオマーカーの存在は、対象の化合物または化合物群の存在を示す。いくつかの実施形態では、バイオマーカーの濃度は、対象の化合物または化合物群の濃度を示す。いくつかの実施形態では、バイオマーカーの濃度は、生物の表現型を示す。

20

【0672】

CSLG酵素は、多くの植物に含まれるステロイド系アルカロイド、ステロイド系サポニン、及びトリテルペノイド系サポニンの生合成に必須の役割を果たすことが、本明細書に開示されている。したがって、いくつかの実施形態では、CSLG遺伝子の発現レベルは、ステロイド系アルカロイド、その誘導体、その代謝物、または生合成中間体、ステロイド系サポニン、その誘導体、その代謝物、または生合成中間体、トリテルペノイド系サポニン、その誘導体、その代謝物、または生合成中間体、あるいは、それらの組み合わせを産生する植物の能力を示す。

【0673】

ステロイド系アルカロイド、ステロイド系サポニン、及びトリテルペノイド系サポニンの産生方法

30

【0674】

本明細書では、遺伝子組換え細胞、あるいは、遺伝子組換え植物細胞を含む植物または遺伝子組換え植物の一部を含む植物の一部においてステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法を提供する。この方法は、細胞内にポリヌクレオチド配列を導入する導入ステップであって、当該ポリヌクレオチド配列が、任意選択でベクターに含まれ、当該ポリヌクレオチド配列が、酵素をコードする少なくとも1つの異種遺伝子を含み、当該酵素が、サポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、UDP - グルコース6 - デヒドロゲナーゼI、及びセルロース合成酵素様G (CSLG) となる群から選択される、該ステップと、当該細胞内で少なくとも1つの異種遺伝子を発現させる発現ステップと、を含む。いくつかの実施形態では、導入されたポリヌクレオチド配列は、複数のポリヌクレオチド配列を含み、各ポリヌクレオチド配列は、酵素をコードする少なくとも1つの異種遺伝子を含み、当該酵素は、サポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、UDP - グルコース6 - デヒドロゲナーゼI、及びセルロース合成酵素様G (CSLG) となる群から選択される。いくつかの実施形態では、導入されたポリヌクレオチド配列は、2、3、4、5、6、7、8、9、10、または11以上のポリヌクレオチド配列を含み、各ポリヌクレオチド配列は、酵素をコードする少なくとも1つの異種遺伝子を含み、当該酵素は、サポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフ

40

50

ェラーゼ、アシルトランスフェラーゼ、UDP - グルコース 6 - デヒドロゲナーゼ I、及びセルロース合成酵素様 G (CSLG) からなる群から選択される。いくつかの実施形態では、導入されたポリヌクレオチド配列は、2、3、4、5、6、7、8、9、10、または 11 以上のポリヌクレオチド配列を含み、各ポリヌクレオチド配列は、酵素をコードする異種遺伝子を 1 つのみ含み、当該酵素は、サポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、UDP - グルコース 6 - デヒドロゲナーゼ I、及びセルロース合成酵素様 G (CSLG) からなる群から選択される。いくつかの実施形態では、導入されたポリヌクレオチド配列は、酵素をコードする 1、2、3、4、5、6、7、8、9、10、または 11 以上の異種遺伝子の任意の組み合わせをそれぞれ含む 2、3、4、5、6、7、8、9、10、または 11 以上のポリヌクレオチド配列を含み、当該酵素は、サポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、UDP - グルコース 6 - デヒドロゲナーゼ I、及びセルロース合成酵素様 G (CSLG) からなる群から選択される。当業者であれば、ポリヌクレオチド配列は、細胞に導入されて発現し、これにより、遺伝子組換え細胞においてコードされた酵素が発現することを理解できるであろう。

10

**【0675】**

いくつかの実施形態では、本明細書に開示されているのは、遺伝子組換え細胞内にステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法であって、(a) 細胞内に少なくとも 1 つの異種遺伝子を導入する導入ステップであって、少なくとも 1 つの異種遺伝子がセルロース合成酵素様 G (CSLG) 酵素をコードしており、少なくとも 1 つの異種遺伝子が任意選択でベクターに含まれている、該ステップと、(b) 細胞内で少なくとも 1 つの異種遺伝子を発現させる発現ステップと、を含み、細胞は、少なくとも 1 つのステロイド系アルカロイド、少なくとも 1 つのステロイド系サポニン、または少なくとも 1 つのトリテルペノイド系サポニンの含有量が、対応する非組換え細胞と比較して増加する。

20

**【0676】**

いくつかの実施形態では、導入ステップは、細胞内に少なくとも 1 つの追加の異種遺伝子を導入するステップであって、少なくとも 1 つの追加の異種遺伝子が、サポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、UDP - グルコース 6 - デヒドロゲナーゼ I、及びそれらの任意の組み合わせからなる群から選択される酵素をコードしており、少なくとも 1 つの追加の異種遺伝子が任意選択でベクター内に含まれている、該ステップと、細胞内で少なくとも 1 つの追加の異種遺伝子を発現させるステップと、をさらに含む。いくつかの実施形態では、(a) 少なくとも 1 つの異種遺伝子が、プロモータ、転写終結配列、またはそれらの組み合わせに動作可能に連結されているか、(b) 少なくとも 1 つの追加の異種遺伝子が、プロモータ、転写終結配列、またはそれらの組み合わせに動作可能に連結されているか、または、(c) 上記の (a) と (b) の組み合わせである。

30

**【0677】**

いくつかの実施形態では、導入ステップは、少なくとも 1 つの細胞を、(a) 少なくとも 1 つの異種遺伝子、少なくとも 1 つの異種遺伝子をコードするポリヌクレオチド配列、または少なくとも 1 つの異種遺伝子を含むベクター、(b) 少なくとも 1 つの追加の異種遺伝子、少なくとも 1 つの追加の異種遺伝子をコードするポリヌクレオチド配列、または少なくとも 1 つの追加の異種遺伝子を含むベクター、または、(c) 上記の (a) と (b) の組み合わせを用いて形質転換することを含み、発現は、一過性発現または構成的発現を含む。

40

**【0678】**

CSLG 酵素をコードするセルロース合成酵素様 G の核酸配列は、全体を通して詳細に記載され、実施例に例示されている。いくつかの実施形態では、少なくとも 1 つの異種 CSLG 遺伝子をコードする核酸配列は、30、32、34、36、38、40、65、8

50

0、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、または、30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも55%の相同性及び少なくとも55%のカバレッジを有するそのホモログであることが好ましい。CSLG酵素をコードするセルロース合成酵素様Gの核酸配列は、全体を通して詳細に記載され、実施例に例示されている。いくつかの実施形態では、少なくとも1つの異種CSLG遺伝子をコードする核酸配列は、30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、または、30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも55%の相同性及び少なくとも80%のカバレッジを有するそのホモログであることが好ましい。

10

**【0679】**

CSLG酵素をコードするセルロース合成酵素様Gの核酸配列は、全体を通して詳細に記載され、実施例に例示されている。いくつかの実施形態では、少なくとも1つの異種CSLG遺伝子をコードする核酸配列は、30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、または、30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログであることが好ましい。

20

**【0680】**

いくつかの実施形態では、トリテルペノイド系サポニンを産生する方法において、コードされるCSLG酵素のアミノ酸配列は、31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列であるか、または、31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列に対して少なくとも55%の相同性及び少なくとも55%のカバレッジを有するそのホモログである。いくつかの実施形態では、トリテルペノイド系サポニンを産生する方法において、コードされるCSLG酵素のアミノ酸配列は、31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列であるか、または、31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列に対して少なくとも55%の相同性及び少なくとも80%のカバレッジを有するそのホモログである。

30

**【0681】**

いくつかの実施形態では、トリテルペノイド系サポニンを産生する方法において、コードされるCSLG酵素のアミノ酸配列は、31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列であるか、または、31、33、35、37、39、41、66、81、94、96、98、100、102または104のいずれか1つに記載のアミノ酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログである。

40

**【0682】**

GAMEタンパク質であるGAME1及びGAME4を、本明細書の実施例に記載されているように、ジャガイモまたはトマトのベイトとして使用した。GAME1またはGAME4と共発現させた遺伝子を表5-8に示し、ステロイド系アルカロイド経路の酵素及びステロイド系サポニン経路の酵素の一部を、図1に示し、全体を通して詳細に説明し、実施例において例示した。図20Aに提示したトリテルペノイド系サポニン経路のSOAP酵素をそれぞれコードするSOAP核酸配列を、全体を通して詳細に説明し、実施例に

50

において例示した。

【0683】

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法において、少なくとも1つの追加の異種遺伝子をコードする核酸配列は、 - アミリン合成酵素をコードし、45に記載の核酸配列、または45に記載の核酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログ、シトクロムP450をコードし、46、51、または53のいずれか1つに記載の核酸配列、または46、51、または53のいずれか1つに記載の核酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログ、グリコシルトランスフェラーゼをコードし、55、57、59、61のいずれか1つに記載の核酸配列、または55、57、59、または61のいずれか1つに記載の核酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログ、アシルトランスフェラーゼをコードし、63に記載の核酸配列、または63に記載の核酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログ、UDP-グルコース6-デヒドロゲナーゼ1をコードし、74に記載の核酸配列、または74に記載の核酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログ、または、それらの任意の組み合わせである。

10

【0684】

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法において、コードされる少なくとも1つの追加の異種遺伝子のアミノ酸配列は、 - アミリン合成酵素をコードし、当該アミノ酸配列は、48に記載のアミノ酸配列、または48に記載のアミノ酸配列に対して少なくとも80%の相同性を有し、かつ少なくとも80%のカバレッジを有するそのホモログである、シトクロムP450をコードし、当該アミノ酸配列は、49、52、または54のいずれか1つに記載のアミノ酸配列、または49、52、または54のいずれか1つに記載のアミノ酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログである、グリコシルトランスフェラーゼをコードし、当該アミノ酸配列は、56、58、60、または62のいずれか1つに記載のアミノ酸配列、または56、58、60、または62のいずれか1つに記載のアミノ酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログである、アシルトランスフェラーゼをコードし、当該アミノ酸配列は、64に記載のアミノ酸配列、または64に記載のアミノ酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログである、UDP-グルコース6-デヒドロゲナーゼ1をコードし、当該アミノ酸配列は、75に記載のアミノ酸配列、または75に記載のアミノ酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログである、または、それらの任意の組み合わせである。

20

30

【0685】

いくつかの実施形態では、ステロイド系アルカロイドまたはステロイド系サポニンを産生する方法において、ステロイド系アルカロイドまたはステロイド系サポニンは、抗栄養剤、化粧品剤、医薬剤、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、ステロイド系アルカロイドまたはステロイド系サポニンを産生する方法において、ステロイド系アルカロイドまたはステロイド系サポニンは、抗栄養剤、化粧品剤、医薬剤、またはそれらの任意の組み合わせから選択される。

40

【0686】

いくつかの実施形態では、トリテルペノイド系サポニンを産生する方法において、トリテルペノイド系サポニンは、甘味料、発泡剤、乳化剤、防腐剤、抗癌剤、低コレステロール剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、防虫剤、抗真菌剤、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、トリテルペノイド系サポニンを産生する方法において、トリテルペノイド系サポニンは、甘味料、発泡

50

剤、乳化剤、防腐剤、抗癌剤、低コレステロール剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、防虫剤、抗真菌剤、またはそれらの任意の組み合わせから選択される。

【0687】

いくつかの実施形態では、ステロイド系アルカロイドを産生する方法において、ステロイド系アルカロイドは、 $\beta$ -トマチン、トマチン、 $\beta$ -カコニン、 $\beta$ -ソラニン、 $\beta$ -ソラソニン、 $\beta$ -ソルマルギン、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、ステロイド系アルカロイドを産生する方法において、ステロイド系アルカロイドは、 $\beta$ -トマチン、トマチン、 $\beta$ -カコニン、 $\beta$ -ソラニン、 $\beta$ -ソラソニン、 $\beta$ -ソルマルギン、またはそれらの任意の組み合わせからなる群から選択される。

10

【0688】

いくつかの実施形態では、ステロイド系サポニンを産生する方法において、ステロイド系サポニンは、ウットロシドB、トマトシド、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、ステロイド系アルカロイドを産生する方法において、ステロイド系アルカロイドは、ウットロシドB、トマトシド、またはそれらの組み合わせからなる群から選択される。

【0689】

いくつかの実施形態では、トリテルペノイド系サポニンを産生する方法において、トリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GlcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、パヨゲニン(化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサボゲノールA(化合物29)、ソヤサボゲノールB(化合物30)、ソヤサボゲノールA-ヘキサA-ヘキサ-ペント(化合物34)、ソヤサポニンVI(化合物35)、ベタブルガロシドIV(化合物33)、ヘデラゲニン-3GlcA、ギブソゲニン-3GlcA、ギブソゲン酸-3GlcA、QS-21アジュバント、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、トリテルペノイド系サポニンを産生する方法において、トリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GlcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、パヨゲニン(化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサボゲノールA(化合物29)、ソヤサボゲノールB(化合物30)、ソヤサボゲノールA-ヘキサA-ヘキサ-ペント(化合物34)、ソヤサポニンVI(化合物35)、ベタブルガロシドIV(化合物33)、ヘデラゲニン-3GlcA、ギブソゲニン-3GlcA、ギブソゲン酸-3GlcA、QS-21アジュバント、またはそれらの任意の組み合わせを含む。

20

30

【0690】

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの誘導体は、それぞれ、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンのグリコシル化誘導体を含む。

40

【0691】

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法において、細胞は、植物細胞、酵母細胞、藻類細胞、昆虫細胞、または細菌を含む。いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法において、細胞は、植物細胞、酵母細胞、または藻類細胞を含む。いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法において、細胞は、植物細胞または酵母細胞を含む。いくつかの実施形態で

50

は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法において、細胞は、植物細胞を含む。いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法において、細胞は、酵母細胞を含む。いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法において、細胞は、藻類細胞を含む。いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法において、細胞は、昆虫細胞を含む。いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法において、細胞は、細菌を含む。

10

**【0692】**

いくつかの実施形態では、少なくとも1つの異種遺伝子が、プロモータ、転写終結配列、またはそれらの組み合わせに動作可能に連結されている。

**【0693】**

いくつかの実施形態では、導入ステップは、少なくとも1つの細胞を、ポリヌクレオチド配列またはポリヌクレオチド配列を含むベクターで形質転換することを含み、発現は、一過性発現または構成的発現を含む。いくつかの実施形態では、ベクターは、発現ベクターを含む。いくつかの実施形態では、ベクターは、プラスミドベクターを含む。いくつかの実施形態では、ベクターは、宿主細胞のDNAに統合されている。いくつかの実施形態では、ベクターに含まれるポリヌクレオチドの一部は、宿主細胞のDNAに統合されている。いくつかの実施形態では、ベクターは、宿主細胞のDNAには統合せず、宿主細胞内で別の実体として複製している。

20

**【0694】**

いくつかの実施形態では、遺伝子組換え植物細胞を産生するために使用される発現ベクターは、アグロバクテリウムベクター及び非アグロバクテリウムベクターの両方を含む。アグロバクテリウムによって媒介される遺伝子導入は、DNAを植物の染色体に導入するアグロバクテリウム・ツメファシエンスの本質的な能力を利用するものであり、「G. Gheysen, G. Angenon, and M. Van Montagu, 1998, *Agrobacterium-mediated plant transformation: a scientifically intriguing story with significant applications in K. Lindsey (Ed.), Transgenic Plant Research, Harwood Academic Publishers, Amsterdam, pp.1-33; and in H. A. Stafford (2000), Botanical Review 66:99-118*」に詳細に記載されている。形質転換法の第2の群は、アグロバクテリウムによって媒介されない形質転換であり、これらの方法は、直接遺伝子導入法として知られている。概要は、「P. Barcelo and P. A. Lazzeri(1998)、*Direct gene transfer: chemical, electrical and physical methods in K. Lindsey (Ed.), Transgenic Plant Research, Harwood Academic Publishers, Amsterdam, pp.35-55*」に提供されている。方法としては、パーティクル・ガン・デリバリー、マイクロインジェクション、無傷の細胞のエレクトロポレーション、ポリエチレングリコールによって媒介されるプロトプラスト形質転換、プロトプラストのエレクトロポレーション、リポソームによって媒介される形質転換、シリコンウイスキーによって媒介される形質転換などが含まれる。適切な対照植物には、本明細書でトランスジェニック植物を生成するために使用される親系統の、遺伝的に改変されていない植物、すなわち非トランスジェニック植物が含まれる。

30

40

**【0695】**

遺伝子を導入した毛状根の培養物は、リゾビウム・リゾゲネスという病原菌によって形質転換することで得られる。毛状根の培養物の確立に用いられるプロトコルは、アグロバクテリウムによる感染に対する植物種の感受性と同様に様々である（「Toivounen et al. 1993; Vanhala et al. 1995」）。形質転換に用いられるアグロバクテリウム株は、根の形態や毛状根の培養物の二次代謝産物の蓄積の程度に大きな影響を与えることが知られている。プロトプラストなどを用いた系統的なクローンセレクションにより、高収量かつ安定した、単一細胞由来の毛根クローンを見出すことが可能である。これが可能であるの

50

は、毛根の培養物が、より大きな体細胞のバリエーションを有しているためである。形質転換のもう一つの可能性は、ウイルスベクターの使用である（「Turpen 1999」）。

【0696】

器官形成または胚形成のいずれかによってその後のクローン増殖が可能な任意の植物組織または植物細胞が、目的の発現ベクターで形質転換されてもよい。「器官形成」という用語は、分裂の中心から芽及び根が順次発生するプロセスを意味し、「胚形成」という用語は、体細胞からであろうと配偶子からであろうと、芽及び根が（連続してではなく）協働して発生するプロセスを意味する。選択される特定の組織は、形質転換される特定の種のために利用可能であり、かつ、形質転換される特定の種に最も適したクローン増殖システムによって異なる。例示的な組織ターゲットには、プロトプラスト、リーフディスク、花粉、胚、子葉、胚軸、大配偶体、カルス組織、既存の分裂組織（例えば、頂端分裂組織、腋芽、根部分裂組織）、及び誘導された分裂組織（例えば、子葉分裂組織、胚軸分裂組織）が含まれる。

10

【0697】

いくつかの実施形態では、上記で詳細に説明したような酵素をコードするポリヌクレオチド配列または少なくとも1つの異種遺伝子を含むベクターが、本明細書に開示されている。いくつかの実施形態では、本明細書に開示されているのは、上記で詳細に説明したような酵素をコードするポリヌクレオチド配列または少なくとも異種遺伝子を含む宿主細胞（これらに限定しないが、例えば植物細胞、酵母、または藻類細胞など）か、または、全体を通して記載したような酵素をコードするポリヌクレオチドまたは少なくとも1つの異種遺伝子を含むベクターである。

20

【0698】

いくつかの実施形態では、「ベクター」という用語は、操作可能に、または連続的に連結された別の核酸分子を輸送することができる核酸分子を包含する。ベクターは、これらに限定しないが、例えばファージ、ウイルス、プラスミド、ファージミド、コスミド、バクミド、または人工染色体などを含む任意の適切なタイプのものであってよい。特定のベクターは、そのベクターが導入された宿主細胞内で自律的に複製することができる（例えば、宿主細胞内で機能する複製の起点を有するベクターなど）。他のベクターは、宿主細胞への導入時に宿主細胞のゲノムに統合され、それによって宿主ゲノムとともに複製される。さらに、いくつかの実施形態では、ベクターは、対象の特定の遺伝子の発現を導くことができる。このようなベクターは、本明細書では「組換え発現ベクター」（または単に「発現ベクター」と呼ばれる。適切なベクターは、必要に応じて、かつ、特定の宿主生物（例えば、植物細胞、酵母細胞または藻類細胞）によって、プロモータ、エンハンサー、ターミネーター配列などの制御配列を有する。

30

【0699】

一般的に、本開示による組換えベクターは、少なくとも1つの「酵素をコードする異種遺伝子」または「酵素をコードする少なくとも1つの異種遺伝子を含む発現カセット」を含む。発現カセットは一般にDNA構築物であり、いくつかの実施形態では、（転写方向の5'から3'に向けて）プロモータ領域と、転写開始領域と動作可能に連結された、詳細に上述した本開示の酵素、ホモログ、変異体、またはその断片をコードする少なくとも1つの異種遺伝子を含むポリヌクレオチド配列と、RNAポリメラーゼの停止シグナル及びポリアデニル化シグナルを含む終結配列とを含む。これらの領域はすべて、形質転換される植物細胞、酵母細胞または藻類細胞などの生物学的細胞において機能可能であるべきということが理解される。好ましくはRNAポリメラーゼ結合部位を含む転写開始領域とポリアデニル化シグナルを含むプロモータ領域は、形質転換されるべき生物学的細胞に固有のものであってよく、その領域が生物学的細胞において機能している場合には、代替の供給源に由来するものであってよい。

40

【0700】

「組換え宿主細胞」（「発現宿主細胞」、「発現宿主システム」、「発現システム」、または単に「宿主細胞」という用語は、本明細書で使用される場合、組換えベクターが

50

導入された細胞を包含する。このような用語は、特定の対象細胞だけでなく、そのような細胞の子孫にも言及することが意図されていることを理解すべきである。突然変異や環境の影響のいずれかに起因して特定の改変が後の世代で発生するため、そのような子孫が親細胞と同一でない場合もあるが、本明細書で使用される「宿主細胞」という用語の範囲内に含まれる。組換え宿主細胞は、単離された細胞や培養された細胞株であってもよく、生体組織または生物の中に存在する細胞であってもよい。宿主細胞は、真菌、植物、または藻類由来のものであってもよい。

【0701】

いくつかの実施形態では、本明細書に開示されているのは、上述のベクターで形質転換されたトランスジェニック植物またはトランスジェニック植物由来の細胞である。

10

【0702】

ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、コードされた酵素は、サポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、UDP - グルコース 6 - デヒドロゲナーゼ I、またはセルロース合成酵素様 G (CSLG) を含む。ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、複数の酵素が異種遺伝子によってコードされ、コードされた酵素は、サポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、UDP - グルコース 6 - デヒドロゲナーゼ I、またはセルロース合成酵素様 G (CSLG) の組み合わせを含む。ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生するにおけるいくつかの実施形態では、1、2、3、4、5、6、7、8、9、10、または11以上の酵素が異種遺伝子によってコードされ、コードされた酵素は、サポニン - アミリン合成酵素、シトクロム P 450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、UDP - グルコース 6 - デヒドロゲナーゼ I、またはセルロース合成酵素様 G (CSLG) の組み合わせを含む。ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、コードされた酵素は、CSLGを含む。

20

【0703】

いくつかの実施形態では、シトクロム P 450 は、C 2 - ヒドロキシラーゼを含むか、または C - 23 オキシダーゼである。いくつかの実施形態では、グリコシルトランスフェラーゼは、フルクトシルトランスフェラーゼ、キシロシルトランスフェラーゼ、または UDP - グリコシルトランスフェラーゼ、またはそれらの組み合わせを含む。いくつかの実施形態では、アシルトランスフェラーゼは、ベンジルアルコールアセチル、アントシアニン - O - ヒドロキシ - シンナモイル、アントラニレート - N - ヒドロキシ - シンマイオール / ベンゾイル、デアセチルピンドリン (BAHD) アセチルトランスフェラーゼを含む。いくつかの実施形態では、トリテルペノイド系サポニンを産生する方法において、コードされたシトクロム P 450 は、C 2 - ヒドロキシラーゼを含むか、C - 23 オキシダーゼであるか、またはそれらの組み合わせである、グリコシルトランスフェラーゼは、フルクトシルトランスフェラーゼ、キシロシルトランスフェラーゼ、UDP - グリコシルトランスフェラーゼであるか、またはそれらの組み合わせを含む、アシルトランスフェラーゼは、ベンジルアルコールアセチル、アントシアニン - O - ヒドロキシ - シンナモイル、アントラニレート - N - ヒドロキシ - シンマイオール / ベンゾイル、デアセチルピンドリン (BAHD) アセチルトランスフェラーゼを含む、あるいは、コードされた酵素は、これらの酵素のいずれかの組み合わせを含み、追加の酵素が、ポリヌクレオチド配列に任意選択で含まれる。

30

40

【0704】

ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、少なくとも1つの異種遺伝子のポリヌクレオチド配列は、配列番号 30、32、34、36、38、40、45、46、47

50

、51、53、55、57、59、61、63、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列、あるいは、少なくとも1つの異種遺伝子のポリヌクレオチド配列は、配列番号30、32、34、36、38、40、45、46、47、51、53、55、57、59、61、63、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも55%の相同性を有する核酸配列を含む。複数の異種遺伝子が発現されるいくつかの実施形態では、当業者は、ポリヌクレオチド配列が、上記で詳細に説明したように、配列番号30、32、34、36、38、40、45、46、47、51、53、55、57、59、61、63、65、80、93、95、97、99、101、103、または105のいずれかに記載の核酸配列の組み合わせを含んでいてもよいことを理解できるであろう。複数の異種遺伝子が発現されるいくつかの実施形態では、当業者は、ポリヌクレオチド配列が、配列番号30、32、34、36、38、40、45、46、47、51、53、55、57、59、61、63、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも55%の相同性を有する核酸配列を含んでいてもよいことを理解できるであろう。

10

**【0705】**

ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、少なくとも1つの異種遺伝子のポリヌクレオチド配列は、配列番号30、32、34、36、38、40、45、46、47、51、53、55、57、59、61、63、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列、あるいは、少なくとも1つの異種遺伝子のポリヌクレオチド配列は、配列番号30、32、34、36、38、40、45、46、47、51、53、55、57、59、61、63、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも80%の相同性を有する核酸配列を含む。複数の異種遺伝子が発現されるいくつかの実施形態では、当業者は、ポリヌクレオチド配列が、上記で詳細に説明したように、配列番号30、32、34、36、38、40、45、46、47、51、53、55、57、59、61、63、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列の組み合わせを含んでいてもよいことを理解できるであろう。複数の異種遺伝子が発現されるいくつかの実施形態では、当業者は、ポリヌクレオチド配列が、配列番号30、32、34、36、38、40、45、46、47、51、53、55、57、59、61、63、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも80%の相同性を有する核酸配列を含んでいてもよいことを理解できるであろう。

20

30

**【0706】**

ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、少なくとも1つの異種遺伝子のポリヌクレオチド配列は、配列番号31、33、35、37、39、41、48、49、50、52、54、56、58、60、62、64、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列の酵素のアミノ酸配列をコードしている、あるいは、少なくとも1つの異種遺伝子のポリヌクレオチド配列は、配列番号31、33、35、37、39、41、48、49、50、52、54、56、58、60、62、64、81、94、96、98、100、102または104のいずれか1つに記載のアミノ酸配列に対して少なくとも55%の相同性を有するアミノ酸配列の酵素のアミノ酸配列をコードしている。

40

**【0707】**

ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、少なくとも1つの異種遺伝子のポリヌクレオチド配列は、配列番号31、33、35、37、39、41、48、49、50、52、54、56、58、60、62、64、81、94、96、98、100、10

50

2、または104のいずれか1つに記載のアミノ酸配列の酵素のアミノ酸配列をコードしている、あるいは、少なくとも1つの異種遺伝子のポリヌクレオチド配列は、配列番号31、33、35、37、39、41、48、49、50、52、54、56、58、60、62、64、81、94、96、98、100、102または104のいずれか1つに記載のアミノ酸配列に対して少なくとも80%の相同性を有するアミノ酸配列の酵素のアミノ酸配列をコードしている。

【0708】

ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、これに限定しないが、例えば植物細胞、酵母、または藻類細胞などの遺伝子組換え細胞は、少なくともステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が、対応する非組換え細胞と比較して、増加している。ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを製産する方法におけるいくつかの実施形態では、これに限定しないが、例えば植物細胞、酵母、藻類細胞、昆虫細胞、または細菌などの遺伝子組換え細胞は、少なくともステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が、対応する非組換え細胞と比較して、増加している。ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、遺伝子組換え細胞は、少なくともステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が、対応する非組換え細胞と比較して増加している植物細胞を含む。ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、遺伝子組換え細胞は、少なくともステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が、対応する非組換え細胞と比較して増加している酵母細胞を含む。ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、遺伝子組換え細胞は、少なくともステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が、対応する非組換え細胞と比較して増加している藻類細胞を含む。ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、遺伝子組換え細胞は、少なくともステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が、対応する非組換え細胞と比較して増加した植物細胞を含み、その植物細胞は、植物またはその一部を含む。

【0709】

いくつかの実施形態では、ステロイド系アルカロイドを産生する方法において、産生されるステロイド系アルカロイドは、抗栄養剤、化粧品剤、医薬剤、またはそれらの任意の組み合わせを含む、ステロイド系アルカロイドは、エスクレオシドまたはデヒドロエスクレオシドを含む、または、ステロイド系アルカロイドは、 $\beta$ -トマチン、トマチン、デヒドロトマチン、 $\beta$ -カコニン、 $\beta$ -ソラニン、 $\beta$ -ソラソニン、 $\beta$ -ソルマルギン、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、ステロイド系アルカロイドを産生する方法において、産生されるステロイド系アルカロイドは、 $\beta$ -トマチン、トマチン、デヒドロトマチン、 $\beta$ -カコニン、 $\beta$ -ソラニン、 $\beta$ -ソラソニン、 $\beta$ -ソルマルギン、またはそれらの任意の組み合わせを含む。

【0710】

いくつかの実施形態では、ステロイド系アルカロイドを産生する方法において、生成されるステロイド系アルカロイドは、 $\beta$ -トマチンを含む。いくつかの実施形態では、ステロイド系アルカロイドを産生する方法において、生成されるステロイド系アルカロイドは、トマチンを含む。いくつかの実施形態では、ステロイド系アルカロイドを産生する方法において、生成されるステロイド系アルカロイドは、デヒドロトマチンを含む。いくつかの実施形態では、ステロイド系アルカロイドを産生する方法において、生成されるステロイド系アルカロイドは、 $\beta$ -カコニンを含む。いくつかの実施形態では、ステロイド系ア

10

20

30

40

50

ルカロイドを産生する方法において、生成されるステロイド系アルカロイドは、 - ソラニンを含む。いくつかの実施形態では、ステロイド系アルカロイドを産生する方法において、生成されるステロイド系アルカロイドは、 - ソラソニンを含む。いくつかの実施形態では、ステロイド系アルカロイドを産生する方法において、生成されるステロイド系アルカロイドは、 - ソルマルギンを含む。

【0711】

いくつかの実施形態では、ステロイド系サポニンを産生する方法において、産生されるステロイド系サポニンは、抗栄養剤、化粧品剤、医薬剤、またはそれらの任意の組み合わせを含む、または、ステロイド系アルカロイドは、ウットロシドB、トマトシド、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、ステロイド系サポニンを産生する方法において、産生されるステロイド系サポニンは、ウットロシドB、トマトシド、またはそれらの任意の組み合わせを含む。

10

【0712】

いくつかの実施形態では、ステロイド系サポニンを産生する方法において、産生されるステロイド系サポニンは、ウットロシドBを含む。いくつかの実施形態では、ステロイド系サポニンを産生する方法において、産生されるステロイド系サポニンは、トマトシドを含む。

【0713】

トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、産生されるトリテルペノイド系サポニンは、甘味料、発泡剤、乳化剤、防腐剤、抗癌剤、低コレステロール剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、防虫剤、抗真菌剤、またはそれらの任意の組み合わせを含む、または、トリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GlcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、パヨゲニン(化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサポゲノールA(化合物29)、ソヤサポゲノールB(化合物30)、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント(化合物34)、ソヤサポニンVI(化合物35)、ベタブルガロシドIV(化合物33)、ヘデラゲニン-3GlcA、ギブソゲニン-3GlcA、ギブソゲン酸-3GlcA、QS-21アジュバント、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、トリテルペノイド系サポニンを産生する方法において、トリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GlcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、パヨゲニン(化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサポゲノールA(化合物29)、ソヤサポゲノールB(化合物30)、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント(化合物34)、ソヤサポニンVI(化合物35)、ベタブルガロシドIV(化合物33)、ヘデラゲニン-3GlcA、ギブソゲニン-3GlcA、ギブソゲン酸-3GlcA、QS-21アジュバント、またはそれらの任意の組み合わせを含む。

20

30

40

【0714】

トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GlcA)(化合物6)を含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、化合物7を含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、化合物8を含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、化合物9を含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態

50

では、製造されるトリテルペノイド系サポニン、化合物10を含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、化合物11を含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、グリチルリチン（化合物14）を含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、グリチルレチン酸3-O-モノグルクロニド（化合物15）を含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、バヨゲニン（化合物25）を含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、バヨゲニン-ヘキサA-ヘキサ-ヘキサ（化合物31）を含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、セルジャン酸（化合物26）を含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、セルジャン酸-ヘキサA-ヘキサ（化合物32）を含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、ソヤサポゲノールA（化合物29）を含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、ソヤサポゲノールB（化合物30）を含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント（化合物34）を含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、ソヤサポニンVI（化合物35）を含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、ベタブルガロシドIV（化合物33）を含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、ヘデラゲニン-3GlcA、ギブソゲニン-3GlcAを含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、ギブソゲン酸-3GlcAを含む。トリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、製造されるトリテルペノイド系サポニンは、QS-21アジュバント、またはそれらの任意の組み合わせを含む。

10

20

30

**【0715】**

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量を、それぞれ、非組換え細胞または非組換え植物と比較して、増加させることと、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、または、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量を、改

40

**【0716】**

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量を、それぞれ、非組換え細胞または非組換え植物と比較して、増加させることと、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その誘導体、その代謝物、またはその生合成中間体、フィトコレステロール、その誘導体、その代謝物、またはその生合成中間体、コレステノール、その誘導体、その代謝物、またはその生合成中間体、または、フィトコレステノール、その誘導体、その代謝物、またはその生合成中間体の含有量を、減

50

小さくすることと、を含む。

【0717】

ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、ポリヌクレオチド配列が導入される細胞は、植物細胞または酵母細胞を含む。あるいは、いくつかの実施形態では、ポリヌクレオチド配列が導入される細胞は、藻類細胞を含む。本明細書に記載されるように遺伝子組み換えされ得る植物細胞及び酵母細胞については、上記で詳細に開示されている。それらと同一の植物細胞及び酵母細胞は、いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法において使用されてもよい。ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法において、いくつかの実施形態では、ポリヌクレオチド配列が導入される細胞は、植物細胞、酵母細胞、藻類細胞、昆虫細胞または細菌を含む。

10

【0718】

ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法において、いくつかの実施形態では、細胞は、葉細胞、葉柄細胞、植物の茎または柄細胞、根細胞、芽細胞、塊茎細胞、豆細胞、穀物若しくは穀粒細胞、果実細胞、ナッツ細胞、マメ細胞、種子または種子細胞、包葉細胞、カルス細胞、または花細胞を含む植物細胞を含む。

20

【0719】

ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法において、いくつかの実施形態では、酵母細胞、藻類細胞、または植物細胞は、本明細書に記載の目、属、または種である。例えば、これに限定しないが、植物細胞は、イネ目、ナデシコ目、ナス目、マメ目、アオイ目、セリ目、フウチョウソウ目、クサスギカズラ目、ヤマノイモ目、またはユリ目の植物からの細胞を含み、または、酵母は、サッカロミセス属、シゾサッカロミセス属、ピチア属、ヤロウイア属、クリペロマイセス属、またはカンジダ属から選択される。

【0720】

ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生するいくつかの実施形態では、ナデシコ目の植物は、ハウレンソウ属、アカザ属、ベータ属、及びレウム属からなる群から選択される、ナス目の植物は、タバコ属、ナス属、及びトウガラシ属からなる群から選択される、マメ目の植物は、グリシリザ属、メデイカゴ属、キラヤ属、ダイズ属、及びハス属からなる群から選択される、セリ目の植物は、オタネニンジン属、ニンジン属、オランダミツバ属、及びパセリ属からなる群から選択される、イネ目の植物は、イネ属（例えば、イネ及びアフリカイネ〔米〕）、オオムギ属（例えば、オオムギ〔大麦〕）、カラスムギ属（例えば、オーツ麦〔オーツ麦〕、アウエナストリゴサ）、及びコムギ属（例えば、スペルト小麦〔スペルト〕）からなる群から選択される、または、フウチョウソウ目の植物は、シロイヌナズナ属（例えば、シロイヌナズナ）、アブラナ属（例えば、ヤセイカンラン〔キャベツ〕、カラシナ〔ホワイトマスタード〕、クロガラシ〔ブラックマスタード〕、セイヨウアブラナ）、フウチョウボク属（例えば、ケッパー〔ケッパー〕）、及びパパイア属（例えば、パパイア〔パパイア〕）からなる群から選択される。

30

40

【0721】

ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、ナス目の植物は、ベンサミアナタバコ、トマト、ジャガイモ、ナス、野生型トマト、野生型ジャガイモ、ズルカマラ、及びトウガラシからなる種の群から選択される、マメ目の植物は、ウラルカンゾウ、アルファルファ、タルウマゴヤシ、キラヤ、ダイズ、及びミヤコグサからなる種の群から選択される、アオイ目の植物は、カカオ属から選択される、セリ目の植物は、オタネニンジン、ノラニンジン、セロリ、及びパセリからなる種の群から選択される、植物は、カカオ種から選

50

扱われる、イネ目の植物は、イネ属（例えば、イネ及びアフリカイネ〔米〕）、オオムギ属（例えば、オオムギ〔大麦〕）、カラスムギ属（例えば、オーツ麦〔オーツ麦〕、アウエナストリゴサ）、及びコムギ属（例えば、スペルト小麦〔スペルト〕）からなる群から選択される、または、フウチョウソウ目の植物は、シロイヌナズナ属（例えば、シロイヌナズナ）、アブラナ属（例えば、ヤセイカンラン〔キャベツ〕、カラシナ〔ホワイトマスタード〕、クロガラシ〔ブラックマスタード〕、セイヨウアブラナ）、フウチョウボク属（例えば、ケッパー〔ケッパー〕）、及びパパイヤ属（例えば、パパイヤ〔パパイヤ〕）からなる属の群から選択される。

**【0722】**

ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法におけるいくつかの実施形態では、酵母は、サッカロミセス属、シゾサッカロミセス属、ピチア属、ヤロウイア属、クリベロマイセス属、またはカンジダ属から選択される。いくつかの実施形態では、酵母は、出芽酵母またはカンジダ・アルピカンスである。いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法において、当該方法は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを精製することをさらに含む。

10

**【0723】**

ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイドを精製する方法は、これらに限定しないが、例えばカラム液体クロマトグラフィー、分取中圧液体クロマトグラフィー、分取HPLCなど、当該技術分野で知られている標準的な精製技術を含む。これらの技術は、植物、その一部または植物細胞からステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを精製する実施形態で使用されてもよい。また、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを精製することなく、植物細胞または植物の一部をそのまま使用してもよい。

20

**【0724】**

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンが酵母細胞で産生されると、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンが細胞から分泌される。いくつかの実施形態では、分泌されたステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンは、酵母細胞の培地から精製される。

30

**【0725】**

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンは、当該技術分野で知られている方法を用いて、植物細胞、植物の一部、植物、藻類、酵母、昆虫細胞、または細菌の培養物から抽出される。

**【0726】**

いくつかの実施形態では、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法が、本明細書で提供される。この方法は、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する植物、植物の一部、コロニー、器官、組織、または細胞から、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを抽出することをさらに含む。葉、葉柄、植物の茎または柄、根、芽、塊茎、豆、穀物または穀粒、果実、ナッツ、マメ科植物、種子または種、包葉、カルス、あるいは花を含む。

40

**【0727】**

植物、植物の一部、植物細胞内のステロイド系アルカロイド、ステロイド系サポニン、トリテルペノイド系サポニンの含有量を改変する方法

**【0728】**

いくつかの実施形態では、遺伝子組換え植物、遺伝子組換え植物の一部または遺伝子組

50

換え植物細胞において、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量を改変する方法が、本明細書に開示されている。この方法は、植物、その一部または細胞におけるCSLG遺伝子を遺伝子組み換えすることを含み、遺伝子組み換えは、CSLG遺伝子の発現を増加させること、発現したCSLG酵素の活性を増加させること、発現したCSLG酵素の安定性を向上させること、またはそれらの任意の組み合わせを含む。この植物、植物の一部または植物細胞は、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が、対応する非組換え植物、植物の一部または植物細胞と比較して、増加している。他の実施形態では、遺伝子組換え植物、遺伝子組換え植物の一部または遺伝子組換え植物細胞において、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量を改変する方法が、本明細書に開示されている。この方法は、植物、その一部、または細胞におけるCSLG遺伝子を遺伝子組み換えすることを含み、遺伝子組み換えは、CSLG遺伝子の発現を減少または消失させること、発現したCSLG酵素の活性を低下させること、発現したCSLG酵素の安定性を低下させること、またはそれらの任意の組み合わせを含む。この植物、植物の一部または植物細胞は、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が、対応する非組換え植物、植物の一部、または植物細胞と比較して、減少している。

10

**【0729】**

いくつかの実施形態では、本明細書に詳細に記載されている遺伝子組換え植物は、本明細書に開示されている少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が改変されている。

20

**【0730】**

特定の実施形態では、植物またはその一部の少なくとも1つの細胞における、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を低減させる方法であって、少なくとも1つの植物細胞の遺伝子を組み換える遺伝子組換えステップを含み、遺伝子組換えステップは、(a)セルロース合成酵素様G(CSLG)酵素をコードする核酸遺伝子配列を標的とする少なくとも1つのサイレンシング分子を用いて少なくとも1つの植物細胞を形質転換するステップ、あるいは、(b)セルロース合成酵素様G(CSLG)酵素をコードする少なくとも1つの核酸配列を突然変異させるステップであって、変異誘発が、(1)核酸配列内に1以上の点突然変異を導入すること、(2)核酸配列内に欠失を導入すること、(3)核酸配列内に挿入を導入すること、または、(4)それらの任意の組み合わせを含み、導入が、コード配列または非コード配列を突然変異させることを含む、該ステップと、を含み、遺伝子組換え細胞におけるCSLG酵素をコードする遺伝子の発現は、対応する非組換え植物細胞と比較して減少し、遺伝子組換え細胞を含む植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1種のトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して減少する、該方法が、本明細書に開示されている。

30

40

**【0731】**

特定の実施形態では、植物またはその一部の少なくとも1つの細胞における、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法であって、少なくとも1つの植物細胞の遺伝子を組み換える遺伝子組換えステップを含み、遺伝子組換えステップは、(a)セルロース合成酵素様G(CSLG)酵素をコードする少なくとも1つの遺伝子の核酸配列を突然変異させるステップで

50

あって、変異誘発が、(1)核酸配列内に1以上の点突然変異を導入すること、(2)核酸配列内に欠失を導入すること、(3)核酸配列内に欠失を導入すること、または、(4)それらの任意の組み合わせを含み、導入が、コード配列または非コード配列を突然変異させることを含む、該ステップと、を含み、(b)CSLGをコードする核酸配列を発現させるステップと、を含み、遺伝子組換え細胞を含む植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して増加する、該方法が、本明細書に開示されている。

10

**【0732】**

ステロイド系アルカロイド生合成酵素、ステロイド系サポニン生合成酵素、またはトリテルペノイド系サポニン生合成酵素(例えばこれに限定しないがCSLG酵素など)をコードする少なくとも1つの核酸配列を突然変異させる方法は、上記の「Genetically Modified Cells & Genetically Modified Plants」のヘッダにおいて詳細に説明されている。同様に、ステロイド系アルカロイド生合成酵素、ステロイド系サポニン生合成酵素、またはトリテルペノイド系サポニン生合成酵素をコードする核酸配列、(例えば、CSLG酵素をコードする核酸配列)を標的とした少なくとも1つのサイレンシング分子で植物細胞を形質転換する方法は、上記の「Genetically Modified Cells & Genetically Modified Plants」のヘッダの下に詳細に記載されている。これらの方法はその全体が本明細書に組み込まれており、遺伝子組換え植物またはその一部に含まれる遺伝子組換え細胞を生産するために使用される方法は、特定の実施形態において、少なくとも1つのステロイド系アルカロイドの含有量を減少または増加させるために使用される方法と同一の方法であることを、当業者は理解できるであろう。これは特に、記載されている遺伝子組換え植物またはその一部は、ステロイド系アルカロイド、ステロイド系サポニン、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の少なくとも1つの含有量が改変されているためである。

20

**【0733】**

いくつかの実施形態では、少なくとも1つのサイレンシング分子がCSLG酵素をコードする核酸遺伝子配列を標的とする場合、CSLG遺伝子をコードする核酸配列は、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列、あるいは、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログである。いくつかの実施形態では、CSLG酵素をコードする核酸遺伝子配列が突然変異している場合、CSLG遺伝子をコードする核酸配列は、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列、あるいは、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログである。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを増加させる方法または減少させる方法において、CSLG酵素のアミノ酸配列は、配列番号31、33、35、37、39、41、66、81、94、96、98、100、102または104のいずれか1つに記載のアミノ酸配列、あるいは、配列番号31、33、35、37、39、41、66、81、94、96、98、100、102、または104のいずれか1つに記載のアミノ酸配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログである。

30

40

**【0734】**

50

いくつかの実施形態では、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの誘導体は、それぞれ、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンのグリコシル化誘導体を含む。

【0735】

いくつかの実施形態では、突然変異は、非コード領域における突然変異を含む。いくつかの実施形態では、突然変異は、コード領域における突然変異を含む。非コード領域における突然変異の非限定的な例は、CSLG遺伝子の発現を増加または減少させる突然変異である。いくつかの実施形態では、突然変異は、CSLG遺伝子の過剰発現を含み、遺伝子組換え植物細胞は、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量が増加している。過剰発現のための方法は、本明細書に詳細に記載されており、当業者は、CSLG遺伝子をコードする核酸配列を含む構築物に構成的または誘導的プロモータが組み込まれてもよいことを理解できるであろう。いくつかの実施形態では、突然変異は、CSLG遺伝子の発現の低下を含み、遺伝子組換え植物細胞は、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの減少を含む。発現を低下させるために核酸配列を特異的に標的化する方法は、全体を通して詳細に説明されており、当業者は、特定の実施形態において、発現の低下が、少なくとも1つのステロイド性アルカロイド、ステロイド性サポニン、またはトリテルペノイド系サポニンの含有量の低下をもたらすことを理解できるであろう。

10

【0736】

いくつかの実施形態では、CSLG酵素をコードする遺伝子のステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法において、CSLG酵素をコードする遺伝子の発現量が、遺伝子組換え植物細胞において、対応する非組換え植物細胞と比較して増加している、コードされたCSLG酵素の活性が、遺伝子組換え植物細胞において、対応する非組換え植物細胞と比較して増加している、コードされたCSLG酵素の安定性が、遺伝子組換え植物細胞において、対応する非組換え植物細胞と比較して増加している、または、それらの任意の組み合わせである。ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニン生合成経路の遺伝子（これに限定しないが、例えばCSLG遺伝子など）をコードする核酸を突然変異させる方法は、上記に詳細に記載されており、酵素の安定性、活性またはそれらの組み合わせを増加させるためのコード領域の突然変異、及び/または、過剰発現のためのCSLG遺伝子を発現する核酸配列の突然変異を含む。当業者であれば、特定の実施形態において、それらの突然変異の方法が本明細書でも使用され得ることを理解できるであろう。

20

30

【0737】

いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドを減少させる方法は、毒素、苦味化合物、またはそれらの組み合わせを含むステロイド系アルカロイドの含有量を、非組換え植物細胞を含む植物またはその一部におけるステロイド系アルカロイドの含有量と比較して、減少させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドを減少させる方法は、毒素、苦味化合物またはそれらの組み合わせを含むステロイド系アルカロイドの含有量を、非組換え植物細胞を含む植物またはその一部におけるステロイド系アルカロイドの含有量と比較して、消失させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドを減少させる方法は、毒素、苦味化合物、またはそれらの組み合わせを含むステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量を、非組換え植物細胞を含む植物またはその一部におけるステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量と比較して、減少させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体を減少させる方法は、毒素、苦味化合物、またはそれらの組み合わせを含

40

50

むステロイド性アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量を、非組換え植物細胞を含む植物またはその一部におけるステロイド系アルカロイドの含有量と比較して、消失させることを含む。

【0738】

いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドは、エスクレオシドまたはデヒドロエスクレオシドを含む。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドは、  
 - トマチン、トマチン、デヒドロトマチン、  
 - カコニン、  
 - ソラニン、  
 - ソラソニン、  
 - ソルマルギン、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、減少した少なくとも1つのステロイド系アルカロイドは、エスクレオシドまたはデヒドロエスクレオシドを含む。いくつかの実施形態では、減少した少なくとも1つのステロイド系アルカロイドは、  
 - トマチン、トマチン、デヒドロトマチン、  
 - カコニン、  
 - ソラニン、  
 - ソラソニン、  
 - ソルマルギン、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、減少した少なくとも1つのステロイド系アルカロイド生合成中間体は、コレステロール、22-ヒドロキシコレステロール、22, 26-ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、フロスタノール-26-アルデヒド、26アミノフロスタノール、トマチデノール、トマチジン、トマチジンガラクトシド、  
 - トマチン、または  
 - 1-トマチン、またはそれらの任意の組み合わせを含む。

10

【0739】

いくつかの実施形態では、減少した少なくとも1つのステロイド系アルカロイドは、エスクレオシドまたはデヒドロエスクレオシドが消失しているか、または、ほぼ消失している。いくつかの実施形態では、減少した少なくとも1つのステロイド系アルカロイドは、  
 - トマチン、トマチン、デヒドロトマチン、  
 - カコニン、  
 - ソラニン、  
 - ソラソニン、  
 - ソルマルギン、またはそれらの任意の組み合わせが消失しているか、または、ほぼ消失している。いくつかの実施形態では、減少した少なくとも1つのステロイド系アルカロイド生合成中間体は、コレステロール、22-ヒドロキシコレステロール、22, 26-ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、フロスタノール-26-アルデヒド、26アミノフロスタノール、トマチデノール、トマチジン、トマチジンガラクトシド、  
 - トマチン、  
 - 1-トマチン、またはそれらの任意の組み合わせが消失しているか、または、ほぼ消失している。いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドを減少させる方法は、少なくとも1つのステロイド系アルカロイドと少なくとも1つのステロイド系アルカロイド中間体との組み合わせを消失させることを含む。

20

30

【0740】

いくつかの実施形態では、少なくとも1つのステロイド系アルカロイドを増加させる方法は、毒素、苦味化合物、またはそれらの組み合わせを含むステロイド系アルカロイドを、非組換え植物細胞を含む植物またはその一部におけるステロイド系アルカロイドの含有量と比較して、増加させることを含む。いくつかの実施形態では、ステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、エスクレオシドまたはデヒドロエスクレオシドを含むステロイド系アルカロイドを増加させることを含む。いくつかの実施形態では、ステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、  
 - トマチン、トマチン、デヒドロトマチン、  
 - カコニン、  
 - ソラニン、  
 - ソラソニン、  
 - ソルマルギン、またはそれらの組み合わせを含むステロイド系アルカロイドを増加させることを含む。

40

【0741】

いくつかの実施形態では、ステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、コレステロール、22-ヒドロキシコレステロール、22, 26-ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、フロスタノール-26-アルデヒド、26-アミノフロスタノール、トマチデ

50

ノール、トマチジン、トマチジンガラクトシド、 $\beta$ -トマチン、 $\beta$ -1-トマチン、またはそれらの任意の組み合わせを含む少なくとも1つのステロイド系アルカロイド生合成中間体を増加させることを含む。いくつかの実施形態では、ステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、それらの組み合わせを増加させることを含む。

【0742】

いくつかの実施形態では、少なくとも1つのステロイド系サポニンを減少させる方法は、毒素、苦味化合物、またはそれらの組み合わせを含むステロイド系サポニンを、非組換え植物細胞を含む植物またはその一部におけるステロイド系サポニンの含有量と比較して、減少させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系サポニンを減少させる方法は、毒素、苦味化合物またはそれらの組み合わせを含むステロイド系サポニンを、非組換え植物細胞を含む植物またはその一部におけるステロイド系サポニンの含有量と比較して、消失させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系サポニンを減少させる方法は、毒素、苦味化合物、またはそれらの組み合わせを含むステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体を、非組換え植物細胞を含む植物またはその一部におけるステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量と比較して、減少させることを含む。いくつかの実施形態では、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体を減少させる方法は、毒素、苦味化合物またはそれらの組み合わせを含むステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体を、非組換え植物細胞を含む植物またはその一部におけるステロイド系サポニンの含有量と比較して、消失させることを含む。

10

20

【0743】

いくつかの実施形態では、少なくとも1つのステロイド系サポニンは、ウットロシドB、トマトシド、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、減少した少なくとも1つのステロイド系サポニンは、ウットロシドB、トマトシド、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、減少した少なくとも1つのステロイド系アルカロイド生合成中間体は、コレステロール、22-ヒドロキシコレステロール、22,26-ジヒドロキシコレステロール、フロスタノール型サポニンアグリコンを含む。

30

【0744】

いくつかの実施形態では、減少した少なくとも1つのステロイド系サポニンは、ウットロシドB、トマトシドまたはそれらの組み合わせが消失しているか、ほぼ消失している。いくつかの実施形態では、減少した少なくとも1つのステロイド系アルカロイド生合成中間体は、コレステロール、22-ヒドロキシコレステロール、22,26-ジヒドロキシコレステロール、フロスタノール型サポニンアグリコンが消失しているか、ほぼ消失している。いくつかの実施形態では、少なくとも1つのステロイド系サポニンを減少させる方法は、少なくとも1つのステロイド系サポニンと少なくとも1つのステロイド系サポニン生合成中間体との組み合わせを消失させることを含む。

40

【0745】

いくつかの実施形態では、少なくとも1つのステロイド系サポニンを増加させる方法は、毒素、苦味化合物またはそれらの組み合わせを含むステロイド系サポニンを、非組換え植物細胞を含む植物またはその一部におけるステロイド系サポニンの含有量と比較して、増加させることを含む。いくつかの実施形態では、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、ウットロシドB、トマトシド、またはそれらの組み合わせを含むステロイド系サポニンを増加させることを含む。

【0746】

いくつかの実施形態では、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、コレステロール、22-ヒドロキシコレス

50

テロール、22, 26 - ジヒドロキシコレステロール、フロスタノール型サポニンアグリコン、またはそれらの任意の組み合わせを含む少なくとも1つのステロイド系サポニン生合成中間体を増加させることを含む。いくつかの実施形態では、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、それらの組み合わせを増加させることを含む。

【0747】

いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンを減少させる方法は、毒素、苦味化合物またはそれらの組み合わせを含むトリテルペノイド系サポニンを、非組換え植物細胞を含む植物またはその一部におけるトリテルペノイド系サポニンの含有量と比較して、減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンを減少させる方法は、毒素、苦味化合物またはそれらの組み合わせを含むトリテルペノイド系サポニンを、非組換え植物細胞を含む植物またはその一部におけるトリテルペノイド系サポニンの含有量と比較して、消失させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンを減少させる方法は、毒素、苦味化合物、またはそれらの組み合わせを含むトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体を、非組換え植物細胞を含む植物またはその一部におけるトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量と比較して、減少させることを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体を減少させる方法は、毒素、苦味化合物、またはそれらの組み合わせを含むトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体を、非組換え植物細胞を含む植物またはその一部におけるトリテルペノイド系サポニンの含有量と比較して、消失させることを含む。

【0748】

いくつかの実施形態では、減少した少なくとも1つのトリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GlcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、パヨゲニン(化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサポゲノールA(化合物29)、ソヤサポゲノールB(化合物30)、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント(化合物34)、ソヤサポニンVI(化合物35)、ベタブルガロシドIV(化合物33)、ヘデラゲニン-3GlcA、ギブソゲニン-3GlcA、ギブソゲン酸-3GlcA、QS-21アジュバント、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、減少したトリテルペノイド系サポニン中間体は、化合物1、化合物2、薬原酸(化合物5)、オレアノール酸(化合物3)、オレアノール酸-3GlcA、オーガスチック酸(化合物4)、オーガスチック酸-3GlcA、またはグリチルレチン酸(化合物13)、またはそれらの任意の組み合わせを含む。いくつかの実施形態では、少なくとも1つのトリテルペノイド系サポニンを減少させる方法は、少なくとも1つのトリテルペノイド系サポニンと1つのトリテルペノイド系サポニン生合成中間体との組み合わせを減少させることを含む。

【0749】

いくつかの実施形態では、減少した少なくとも1つのトリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(MA-3-GlcA)(化合物6)、化合物7、化合物8、化合物9、化合物10、化合物11、グリチルリチン(化合物14)、グリチルレチン酸3-O-モノグルクロニド(化合物15)、パヨゲニン(化合物25)、パヨゲニン-ヘキサA-ヘキサ-ヘキサ(化合物31)、セルジャン酸(化合物26)、セルジャン酸-ヘキサA-ヘキサ(化合物32)、ソヤサポゲノールA(化合物29)、ソヤサポゲノールB(化合物30)、ソヤサポゲノールA-ヘキサA-ヘキサ-ペント(化合物34)、ソヤサポニンVI(化合物35)、ベタブルガロシドIV(化合物33)、ヘ

10

20

30

40

50

デラゲニン - 3 G l c A、ギブソゲニン - 3 G l c A、ギブソゲン酸 - 3 G l c A、Q S - 2 1 アジュバント、またはそれらの任意の組み合わせが消失しているか、または、ほぼ消失している。いくつかの実施形態では、減少したトリテルペノイド系サポニン中間体は、化合物 1、化合物 2、薬原酸（化合物 5）、オレアノール酸（化合物 3）、オレアノール酸 - 3 G l c A、オーガスチック酸（化合物 4）、オーガスチック酸 - 3 G l c A、またはグリチルレチン酸（化合物 1 3）、またはそれらの任意の組み合わせが消失しているか、または、ほぼ消失している。いくつかの実施形態では、少なくとも 1 つのトリテルペノイド系サポニンを減少させる方法は、少なくとも 1 つのトリテルペノイド系サポニンと 1 つのトリテルペノイド系サポニン生合成中間体との組み合わせを消失させることを含む。

10

**【 0 7 5 0 】**

いくつかの実施形態では、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、甘味料、発泡剤、乳化剤、防腐剤、抗癌剤、低コレステロール剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、防虫剤、抗真菌剤、またはそれらの任意の組み合わせを含むトリテルペノイド系サポニンを増加させることを含む。いくつかの実施形態では、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、メディカゲン酸 3 - O - グルクロニド（M A - 3 - G l c A）（化合物 6）、化合物 7、化合物 8、化合物 9、化合物 1 0、化合物 1 1、グリチルリチン（化合物 1 4）、グリチルレチン酸 3 - O - モノグルクロニド（化合物 1 5）、パヨゲニン（化合物 2 5）、パヨゲニン - ヘキサ A - ヘキサ - ヘキサ（化合物 3 1）、セルジャン酸（化合物 2 6）、セルジャン酸 - ヘキサ A - ヘキサ（化合物 3 2）、ソヤサポゲノール A（化合物 2 9）、ソヤサポゲノール B（化合物 3 0）、ソヤサポゲノール A - ヘキサ A - ヘキサ - ペント（化合物 3 4）、ソヤサポニン V I（化合物 3 5）、ベタブルガロシド I V（化合物 3 3）、ヘデラゲニン - 3 G l c A、ギブソゲニン - 3 G l c A、ギブソゲン酸 - 3 G l c A、Q S - 2 1 アジュバント、またはそれらの任意の組み合わせを含むトリテルペノイド系サポニンを増加させることを含む。いくつかの実施形態では、トリテルペノイド系サポニンの含有量を増加させる方法は、トリテルペノイド系サポニンであるグリチルリチン（化合物 1 4）を増加させることを含む。いくつかの実施形態では、トリテルペノイド系サポニンの含有量を増加させる方法は、トリテルペノイド系サポニンである Q S - 2 1 アジュバントを増加させることを含む。

20

30

**【 0 7 5 1 】**

いくつかの実施形態では、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、化合物 1、化合物 2、薬原酸（化合物 5）、オレアノール酸（化合物 3）、オレアノール酸 - 3 G l c A、オーガスチック酸（化合物 4）、オーガスチック酸 - 3 G l c A、またはグリチルレチン酸（化合物 1 3）、またはそれらの任意の組み合わせを含むトリテルペノイド系サポニン生合成中間体を増加させることを含む。いくつかの実施形態では、トリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、それらの組み合わせを増加させることを含む。

40

**【 0 7 5 2 】**

いくつかの実施形態では、少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法は、少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、あるいは、少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させることと、フィトステロール、その誘導体、その代謝物、またはその生合成中間体、コレステロール、その

50



の生合成中間体の含有量を改変する（増加または減少させる）方法において、当該方法は、少なくとも1つの植物細胞の遺伝子を組み換えることを含み、いくつかの実施形態では、当該植物細胞は、葉細胞、葉柄細胞、植物の茎または柄細胞、根細胞、芽細胞、塊茎細胞、豆細胞、穀物若しくは穀粒細胞、果実細胞、ナッツ細胞、マメ細胞、種子または種子細胞、包葉細胞、カルス細胞、または花細胞を含む。

【0757】

いくつかの実施形態では、ステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、ステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、またはトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を改変する（増加または減少させる）方法において、当該方法は、少なくとも1つの植物細胞を遺伝子的に改変することを含み、いくつかの実施形態では、植物細胞は、本明細書に詳細に記載されている目、属、または種を含む。例えば、ポ配列ス目、カリオフィラ目、ソラナレス目、ファバレス目、マルバレス目、アピ配列ス目、ブラシカレス目、アスパラガス目、ディオスコレ配列ス目、またはリリ配列ス目の植物からの細胞を含む植物細胞が挙げられるが、これに限定されない。

10

【0758】

ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量を改変する方法におけるいくつかの実施形態では、ナデシコ目の植物は、ホウレンソウ属、アカザ属、ベータ属、及びレウム属からなる群から選択される。ナス目の植物は、タバコ属、ナス属、トウガラシ属からなる群から選択される。マメ目の植物は、グリシリザ属、メディカゴ属、キラヤ属、ダイズ属、及びハス属からなる群から選択される。セリ目の植物は、オタネニンジン属、ニンジン属、オランダミツバ属、及びバセリ属からなる群から選択される。イネ目の植物は、イネ属（例えば、イネ及びアフリカイネ〔米〕）、オオムギ属（例えば、オオムギ〔大麦〕）、カラスムギ属（例えば、オーツ麦〔オーツ麦〕、アウエナストリゴサ）、及びコムギ属（例えば、スペルト小麦〔スペルト〕）からなる群から選択される、または、フウチョウソウ目の植物は、シロイヌナズナ属（例えば、シロイヌナズナ）、アブラナ属（例えば、ヤセイカンラン〔キャベツ〕、カラシナ〔ホワイトマスタード〕、クロガラシ〔ブラックマスタード〕、セイヨウアブラナ）、フウチョウボク属（例えば、ケッパー〔ケッパー〕）、及びバパイヤ属（例えば、バパイヤ〔バパイヤ〕）からなる群から選択される。

20

30

【0759】

ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量を改変する方法におけるいくつかの実施形態では、ナス目の植物は、ベンサミアナタバコ、トマト、ジャガイモ、ナス、野生型トマト、野生型ジャガイモ、ズルカマラ、及びトウガラシからなる種の群から選択される、マメ目の植物は、ウラルカンゾウ、アルファルファ、タルウマゴヤシ、キラヤ、ダイズ、及びミヤコグサからなる種の群から選択される、アオイ目の植物は、カカオ属から選択される、セリ目の植物は、オタネニンジン、ノラニンジン、セロリ、及びバセリからなる種の群から選択される、または、植物は、カカオ種から選択される。

【0760】

ステロイド系アルカロイド、ステロイド系サポニン、及びトリテルペノイド系サポニンのインビトロでの産生方法

40

【0761】

インビトロシステムは当該技術分野でよく知られている。いくつかの実施形態では、本明細書で開示されるトリテルペノイド系サポニンの産生方法において使用されるインビトロ翻訳システムは、「rabbit reticulocyte lysate」、「wheat germ extract」、または「E. coli cell-free system」を含む。無細胞システムを使用するための詳細な方法は、当該技術分野で知られ、自由に利用することができ、これに限定しないが、例えば以下のウェブサイト「<https://www.thermofisher.com/il/en/home/references/ambion-tech-support/large-scale-transcription/general-articles/the-basics-in-vitro-translation>。

50

html」で得られる情報を含む。

【0762】

いくつかの実施形態では、インビトロシステムにおいてステロイド系アルカロイドまたはステロイド系サポニンを産生する方法が、本明細書に開示されている。この方法は、(a) シトクロム P 4 5 0 ( G A M E 7 ) 遺伝子、シトクロム P 4 5 0 ( G A M E 8 ) 遺伝子、2 - オキソグルタル酸依存性ジオキシゲナーゼ ( G A M E 1 1 ) 遺伝子、シトクロム P 4 5 0 ( G A M E 6 ) 遺伝子、シトクロム P 4 5 0 ( G A M E 4 ) 遺伝子、トランスアミナーゼ ( G A M E 1 2 ) 遺伝子、トマチジン UDP - ガラクトシルトランスフェラーゼ ( G A M E 1 / S G T 1 ) 遺伝子、UDP - グリコシルトランスフェラーゼ ( G A M E 1 7 ) 遺伝子、UDP - グリコシルトランスフェラーゼ ( G A M E 1 8 ) 遺伝子、UDP - グリコシルトランスフェラーゼ / キシロシルトランスフェラーゼ ( G A M E 2 / S G T 3 ) 遺伝子、及びセルロース合成酵素様 G 遺伝子 ( C S L G 遺伝子または G A M E 1 5 遺伝子) をコードする核酸配列 ( この核酸配列は、任意選択でベクター内に含まれている ) を組み合わせること、( b ) 当該遺伝子を発現させること、及び、( c ) 発現した酵素の組み合わせを一緒にインキュベートすること、を含み、当該インキュベーションの生成物は、少なくとも1つのステロイド系アルカロイドまたはステロイド系サポニンを含む。

10

【0763】

いくつかの実施形態では、上記の核酸配列は、エチレン応答性エレメント結合因子 1 3 ( G A M E 9 ) 及び / または B H L H 転写因子をさらに含む。

【0764】

いくつかの実施形態では、インビトロシステムにおいてステロイド系アルカロイドまたはステロイド系サポニンを産生する方法が、本明細書に開示されている。この方法は、( d ) シトクロム P 4 5 0 ( G A M E 7 ) 遺伝子、シトクロム P 4 5 0 ( G A M E 8 ) 遺伝子、2 - オキソグルタル酸依存性ジオキシゲナーゼ ( G A M E 1 1 ) 遺伝子、シトクロム P 4 5 0 ( G A M E 6 ) 遺伝子、及びセルロース合成酵素様 G 遺伝子 ( C S L G 遺伝子または G A M E 1 5 遺伝子) をコードする核酸配列 ( この核酸配列は、任意選択でベクター内に含まれている ) を組み合わせること、( e ) 遺伝子を発現させること、及び、( f ) 発現した酵素の組み合わせを一緒にインキュベートすること、を含み、当該インキュベーションの生成物は、少なくとも1つのステロイド系アルカロイド、ステロイド系サポニンまたはそれらの組み合わせを含む。例えば、これに限定しないが、図 1、図 8 D、図 1 3、及び図 1 4 A - 1 4 D に示されている生合成経路及びそれによって得られる中間体及び産物を参照されたい。

20

30

【0765】

いくつかの実施形態では、生合成 G A M E 遺伝子及び / または C S L G 遺伝子は、コンストラクトを作製するために、本明細書で詳細に記載されているように、プロモータ、転写終結配列、またはそれらの組み合わせに対して動作可能に連結されている。いくつかの実施形態では、ステロイド系アルカロイドまたはステロイド系サポニンを産生する方法において、C S L G 遺伝子をコードする核酸配列は、配列番号 3 0、3 2、3 4、3 6、3 8、4 0、6 5、8 0、9 3、9 5、9 7、9 9、1 0 1、1 0 3、または 1 0 5 のいずれか 1 つに記載の核酸配列、あるいは、配列番号 3 0、3 2、3 4、3 6、3 8、4 0、6 5、8 0、9 3、9 5、9 7、9 9、または 1 0 1 のいずれか 1 つに記載の核酸配列に対して少なくとも 5 5 % の相同性及び少なくとも 5 5 % のカバレッジを有するそのホモログである。いくつかの実施形態では、ステロイド系アルカロイドまたはステロイド系サポニンを産生する方法において、C S L G 遺伝子をコードする核酸配列は、配列番号 3 0、3 2、3 4、3 6、3 8、4 0、6 5、8 0、9 3、9 5、9 7、9 9、1 0 1、1 0 3、または 1 0 5 のいずれか 1 つに記載の核酸配列、あるいは、3 0、3 2、3 4、3 6、3 8、4 0、6 5、8 0、9 3、9 5、9 7、9 9、または 1 0 1 のいずれか 1 つに記載の核酸配列に対して少なくとも 5 5 % の相同性及び少なくとも 8 0 % のカバレッジを有するそのホモログであることが好ましい。

40

【0766】

50

いくつかの実施形態では、生合成 G A M E 遺伝子及び/または C S L G 遺伝子は、コンストラクトを作製するために、本明細書で詳細に記載されているように、プロモータ、転写終結配列、またはそれらの組み合わせに対して動作可能に連結されている。いくつかの実施形態では、ステロイド系アルカロイドまたはステロイド系サポニンを産生する方法において、C S L G 遺伝子をコードする核酸配列は、配列番号 30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列、あるいは、配列番号30、32、34、36、38、40、65、80、93、95、97、99、または101のいずれか1つに記載の核酸配列に対して少なくとも80%の相同性及び少なくとも80%のカバレッジを有するそのホモログであることが好ましい。

10

**【0767】**

特定の実施形態では、ステロイド系アルカロイドを産生するインビトロの方法は、エスクレオシドまたはデヒドロエスクレオシドから選択されるステロイド系アルカロイドを産生することを含む。特定の実施形態では、ステロイド系アルカロイドを産生するインビトロの方法は、  
- トマチン、トマチン、  
- カコニン、  
- ソラニン、  
- ソラソニン、  
- ソルマルギン、またはそれらの任意の組み合わせから選択されるステロイド系アルカロイドを産生することを含む。

**【0768】**

特定の実施形態では、ステロイド系サポニンを産生するインビトロの方法は、ウットロシド B、トマトシドまたはそれらの任意の組み合わせから選択されるステロイド系サポニンを産生することを含む。

20

**【0769】**

いくつかの実施形態では、インビトロシステムにおいてトリテルペノイド系サポニンを産生する方法が、本明細書に開示されている。この方法は、(g) サポニン - アミリン合成酵素遺伝子 (S O A P 1)、シトクロム P 4 5 0 遺伝子 (S O A P 2)、C - 2 ヒドロキシラーゼを発現するシトクロム P 4 5 0 遺伝子 (S O A P 3)、C - 2 3 オキシダーゼを発現するシトクロム P 4 5 0 遺伝子 (S O A P 4)、UDP - グリコシルトランスフェラーゼまたはフコシルトランスフェラーゼを発現するグリコシルトランスフェラーゼ遺伝子 (S O A P 6)、UDP - グリコシルトランスフェラーゼを発現するグリコシルトランスフェラーゼ遺伝子 (S O A P 7)、UDP - グリコシルトランスフェラーゼを発現するグリコシルトランスフェラーゼ遺伝子 (S O A P 8)、UDP - グリコシルトランスフェラーゼまたはキシロシルトランスフェラーゼを発現するグリコシルトランスフェラーゼ遺伝子 (S O A P 9)、アシルトランスフェラーゼ遺伝子 (S O A P 10)、及びセルロース合成酵素様 G 遺伝子 (C S L G)、をコードする核酸配列 (この核酸配列は、任意選択でベクター内に含まれている) を組み合わせること、(h) 遺伝子を発現させること、及び、(i) 発現した酵素の組み合わせを一緒にインキュベートすること、を含み、当該インキュベーションの生成物は、少なくとも1つのトリテルペノイド系サポニンを含む。

30

**【0770】**

いくつかの実施形態では、生合成 S O A P 遺伝子及び/または C S L G 遺伝子は、コンストラクトを作製するために、本明細書で詳細に記載されているように、プロモータ、転写終結配列、またはそれらの組み合わせに対して動作可能に連結されている。いくつかの実施形態では、トリテルペノイド系サポニンを産生する方法において、C S L G 遺伝子をコードする核酸配列は、配列番号 30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列、あるいは、配列番号30、32、34、36、38、40、65、80、93、95、97、99、または101のいずれか1つに記載の核酸配列に対して少なくとも55%の相同性及び少なくとも55%のカバレッジを有するそのホモログである。いくつかの実施形態では、トリテルペノイド系サポニンを産生する方法において、C S L G 遺伝子をコードする核酸配列は、配列番号 30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列、あ

40

50

るいは、配列番号 30、32、34、36、38、40、65、80、93、95、97、99、または 101 のいずれか 1 つに記載の核酸配列に対して少なくとも 55% の相同性及び少なくとも 80% のカバレッジを有するそのホモログであることが好ましい。

【0771】

いくつかの実施形態では、生合成 S O A P 遺伝子及び/または C S L G 遺伝子は、コンストラクトを作製するために、本明細書で詳細に記載されているように、プロモータ、転写終結配列、またはそれらの組み合わせに対して動作可能に連結されている。いくつかの実施形態では、トリテルペノイド系サポニンを産生する方法において、C S L G 遺伝子をコードする核酸配列は、配列番号 30、32、34、36、38、40、65、80、93、95、97、99、101、103、または 105 のいずれか 1 つに記載の核酸配列、あるいは、配列番号 30、32、34、36、38、40、65、80、93、95、97、99、または 101 のいずれか 1 つに記載の核酸配列に対して少なくとも 80% の相同性及び少なくとも 80% のカバレッジを有するそのホモログであることが好ましい。

10

【0772】

いくつかの実施形態では、他のトリテルペノイド系サポニン生合成酵素をコードする核酸遺伝子配列は、(a) - アミリン合成酵素遺伝子をコードする核酸配列が、配列番号 45 に記載の核酸配列、または、配列番号 45 に記載の核酸配列に対して少なくとも 80% の相同性及び少なくとも 80% のカバレッジを有するそのホモログである、(b) シトクロム P 450 遺伝子をコードする核酸配列が、配列番号 46、51、または 53 のいずれか 1 つに記載の核酸配列、あるいは、配列番号 46、51、または 53 のいずれか 1 つに記載の核酸配列に対して少なくとも 80% の相同性及び少なくとも 80% のカバレッジを有するそのホモログである、(c) グリコシルトランスフェラーゼ遺伝子をコードする核酸配列が、配列番号 55、57、59、または 61 のいずれか 1 つに記載の核酸配列、あるいは、配列番号 55、57、59、または 61 のいずれか 1 つに記載の核酸配列に対して少なくとも 80% の相同性及び少なくとも 80% のカバレッジを有するそのホモログである、及び、(d) アシルトランスフェラーゼをコードする核酸配列が、配列番号 63 に記載の核酸配列、または、配列番号 63 に記載の核酸配列に対して少なくとも 80% の相同性及び少なくとも 80% のカバレッジを有するそのホモログである。

20

【0773】

特定の実施形態において、トリテルペノイド系サポニンを産生するインビトロの方法は、甘味料、発泡剤、乳化剤、防腐剤、抗癌剤、低コレステロール剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、防虫剤、抗真菌剤、またはそれらの任意の組み合わせから選択されるトリテルペノイド系サポニンを産生することを含む。

30

【0774】

特定の実施形態では、トリテルペノイド系サポニンを産生するインビトロの方法は、メディカゲン酸 3 - O - グルクロニド (M A - 3 - G l c A) (化合物 6)、化合物 7、化合物 8、化合物 9、化合物 10、化合物 11、グリチルリチン (化合物 14)、グリチルレチン酸 3 - O - モノグルクロニド (化合物 15)、パヨゲニン (化合物 25)、パヨゲニン - ヘキサ A - ヘキサ - ヘキサ (化合物 31)、セルジャン酸 (化合物 26)、セルジャン酸 - ヘキサ A - ヘキサ (化合物 32)、ソヤサポゲノール A (化合物 29)、ソヤサポゲノール B (化合物 30)、ソヤサポゲノール A - ヘキサ A - ヘキサ - ペント (化合物 34)、ソヤサポニン V I (化合物 35)、ベタブルガロシド I V (化合物 33)、ヘデラゲニン - 3 G l c A、ギブソゲニン - 3 G l c A、ギブソゲン酸 - 3 G l c A、Q S - 21 アジュバント、またはそれらの任意の組み合わせから選択されるトリテルペノイド系サポニンを産生することを含む。

40

【0775】

本明細書及び特許請求の範囲で使用するとき、単数形の「a」、「an」、及び「the」は、文脈が他の意味を明確に示さない限り、複数形も含むことを意図している。例えば、単数形の「分子」という用語は、分子の複数形も含む。

【0776】

50

本出願の全体を通じて、本発明の様々な実施形態を範囲形式で示すことがある。範囲形式での記載は単に利便性及び簡潔性のためであり、本発明の範囲に対する柔軟性のない限定と解釈されるべきではないことを理解されたい。したがって、範囲に関する記載は、可能な部分範囲のすべて、並びに、その範囲内に含まれる個々の数値を具体的に開示したものと見なされるべきである。例えば、1から6までという範囲の記述は、1から3まで、1から4まで、1から5まで、2から4まで、2から6まで、3から6まで、・・・、という部分的な範囲、並びに、その範囲内に含まれる個々の数（例えば、1、2、3、4、5、及び6）を具体的に開示したものと見なされるべきである。このことは、範囲の幅に関係なく適用される。

【0777】

本明細書において数値範囲を指定している場合は常に、指定された数値範囲内に含まれるすべての数値（分数または整数）を含むことを意味する。第1の指定数と第2の指定数と「の間の範囲」という用語、及び第1の指定数「から」第2の指定数「までの範囲」という用語は、本明細書では相互交換的に使用されており、第1の指定数及び第2の指定数、並びにこれらの間のすべての分数及び整数を含むことを意味する。

【0778】

さらに、当業者であれば、本明細の全体を通じて使用される「～を含む（comprising）」という用語は、本明細書に開示された遺伝子組換えまたは遺伝子編集植物、これらにおける遺伝子の発現を改変する方法、及び、これらにおけるSA及び/またはSGAの産生を改変する方法が、記載された要素を含むが、任意選択の他の要素を除外しないことを意味することを意図していることを理解できるであろう。したがって、「～からなる（consisting of）」は、他の要素の形跡以上のものを除外することを意味するものとする。当業者であれば、いくつかの実施形態では、「～を含む（comprising）」という用語が使用されているが、その用語は「～からなる（consisting of）」という用語に置き換えてもよく、このような置き換えは、具体的に記載されていない要素の包含範囲を狭めることになることを理解できるであろう。

【0779】

以下の例は、本発明のいくつかの実施形態をより完全に説明するために提示される。しかしながら、これらは、決して、本発明の広い範囲を限定するものとして解釈されるべきではない。また、これらは、決して、本明細書に開示された遺伝子組換え細胞及び遺伝子組換え植物、ステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生するための細胞及び植物の使用、あるいは、植物またはその一部におけるステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンの含有量を増加または減少させるための方法の広い範囲を制限するものとして解釈されるべきではない。当業者であれば、本明細書に開示された範囲から逸脱することなく、本明細書に開示された原理の多くの変形及び修正を容易に考案することができる。

【実施例】

【0780】

実施例1：実施例2～13の材料及び方法

【0781】

植物材料、処理、及び遺伝子導入植物の作製

【0782】

トマト（*Solanum lycopersicum*、栽培品種マイクロ・トム（Micro Tom））及びジャガイモ（*Solanum tuberosum*、栽培品種デジレー（Desiree））植物を、既報（前掲の「Itkin et al., 2001」）のように採取した。ジャガイモでは、緑色部分が乾燥し始めた際に、成熟塊茎（第3段階）を採取し、汚れを洗い落として乾燥し、4の完全暗所で保存した。

【0783】

対応するGAME9 DNA断片をpK7GWIWG2（II）バイナリーベクター及びpJCV52バイナリーベクターにそれぞれ導入することにより、GAME9サイレン

10

20

30

40

50

シング (RNAi) コンストラクト及び過剰発現 (OX) コンストラクトを作製した。Itkinら (前掲の「2011」) の既報に従い、トマト及びジャガイモにおけるGAME 9 のサイレンシング及び過剰発現の遺伝子導入系を作製し、組織抽出物を調製し、解析した。

【0784】

以下の表3に、本明細書記載のコンストラクトの作製に用いたオリゴヌクレオチドを示す。GAME 4サイレンシング (RNAi ; GAME 4i) コンストラクト、GAME 4過剰発現 (GAME 4oe) コンストラクト及びGAME 8サイレンシングコンストラクトは、既報 (前掲の「Itkin et al., 2001」 ; 「WO 2012/095843」) のように作製した。

【0785】

表3 : コンストラクトの作製に用いたオリゴヌクレオチド

【0786】

【表3】

名称	5'から3'方向への配列/ 説明	SEQ ID NO.
S107g0434 20 EcoRI Fw	AAAAAgaattcCGGATCTTCTCTCGAACTGGTCAA GAME11ウイルス誘導遺伝子サイレンシング(VIGS) コンストラクトを作製する	1
S107g0434 20 EcoRI Rv	AAAAAgaattcCACTTTCATTGCTTCATCCATTAGATCT GAME11 VIGSコンストラクトを作製する	2
S107g0435 00 EcoRI Fw	AAAAAgaattcCTTAGCTTATGGCCACATCACACCTT GAME18 VIGSコンストラクトを作製する	3
S107g043500 EcoRI Rv	AAAAAgaattcACTCAAGATTTGGTGAAGCTGTGGTT GAME18 VIGSコンストラクトを作製する	4
G8-Forward (AscI)	AAAAAGGCGCGCCAATCATAGAGAAGAAAGAAAGACG GAME8のRNAiを構築する	5
G8-Reverse (Not I)	AAAAAGCGGCCGCGACTCCTGCAGGAATTGTCATTCTC GAME8のRNAiを構築する	6
GAME9 RNAi NotI Fw	aaaaaGCGGCCGCATGAGTATTGTAATTGATGATGATGAA ATC GAME9のRNAiを構築する	7
GAME9 RNAi AscI Rv	aaaaGGCGGCCCGCCACACGCCACAGATGGTTCTT GAME9のRNAiを構築する	8
GAME9-Tom GW Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGAGTATT GTAATTGATGATGATGAAATC cDNAから遺伝子を取り出して過剰発現させる (トマトに好適)	9
GAME9-Tom GW Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTTTCATACTAC CTTCTGTCCTAAGCCT cDNAから遺伝子を取り出して過剰発現させる (トマトに好適)	10
GAME9-Pot GW Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGAATATT GCAATTGATGATGATGA cDNAから遺伝子を取り出して過剰発現させる (ジャガイモに好適)	11
GAME9-Pot GW Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTTTCATTGTAT CAACATTTGTAATTCACAC cDNAから遺伝子を取り出して過剰発現させる (ジャガイモに好適)	12

10

20

30

40

## 【0787】

共発現解析

## 【0788】

共発現解析においてトマトGAME1 (Solyc07g043490) 及びそのジャガイモのオルソログSGT1 (PGSC003DMG400011749) を「ベイト (baits)」として用い、共発現遺伝子 (各「ベイト」別) のリスト (r値 0.8で降順にソート) を作成した。次いで、2つの相同遺伝子を同定した (トマトにおけるSolyc12g006460 及びジャガイモにおけるPGSC0003DMG40002427

50

4)。これらは「ベイト」遺伝子と高い相関があった(両種において $r$ 値 $>0.9$ )。これらの遺伝子は、GLYCOALKALOID METABOLISM 4 (GAME 4、「WO 2012/095843」と同定した。GAME 4 遺伝子を「ベイト」として、前述の(GAME 1)共発現解析にさらに加えた。両種におけるGAME 1 (SGT 1)及びGAME 4の共発現のリストを用い、共発現相関ネットワークを作成した。解析は次のように行った。異なる組織、器官(果肉、皮、種子、根、葉、芽、花、花粉)、及び生育段階(合計25の試験)から得られたトマトRNAseqトランスクリプトームデータ(前掲の「Itkin et al., 2011」)、並びに異なる組織及び器官(合計40の試験)から得られたジャガイモRNAseqトランスクリプトームデータ(「US 2012/0159676」)を用いた。まず、Rスクリプトを用い、共発現解析(各種ごと)を行い、Perlスクリプトを用い、共発現遺伝子のリストをFASTAファイルとして作成した。最後に、BLASTallツール(「Camacho C. et al., 2009. BMC Bioinform 10:421」)を用い、これら2種間に共通のホモログを見出した。ホモログ類似性のtblastx基準としては、 $p$ 値 $>0.05$ 、最小25ヌクレオチド、及び少なくとも60パーセントの類似性を、各遺伝子についての全体的な相同性として設定した。共発現ネットワークは、Cytoscapeプログラム(「Shannon P. et al., 2003. Genome Res. 13:2498-2504」)によって可視化した。

10

【0789】

系統発生解析

【0790】

20

Muscleアルゴリズムを用いてタンパク質配列をアライメントし、系統樹を解析し、SeaView v4.3.5プログラムにより、PhyML 3.0による最尤法(「Exposito-Rodriguez M et al., 2008. BMC Plant Biol. 8:131」)を用い、以下の設定で可視化した。モデル - LG; 近似尤度比検定(aLRT)下平 - 長谷川様式(SH様式)の手順を統計的検定として用い、ブランチサポートを算出した(ブランチサポート - aLRT(SH様式)); 不変座位 - 最適化; 座位間の進化速度変化 - 最適化; 系統樹探索操作 - NNI及びSPRに最適; 開始系統樹 - BioNJ、樹形最適化。枝上の番号は、各節を支持するブートストラップ反復の比率を示す。この系統樹の作成に用いたタンパク質の受託番号及び器官名を以下の表4に示す。系統樹は図12に示す。

【0791】

30

表4: 系統樹の作成に用いた配列の受託番号

【0792】

【表 4 - 1】

図12に記載の名称	ラテン名及び一般名	受託番号
GuCYP88D6	ウラルカンゾウ	BAG68929.1
LjCYP88D4	ミヤコグサ	BAG68927.1
MtCYP88D3	タルウマゴヤシ	BAG68926.1
CmCYP88A2	セイヨウカボチャ	AF212991
AtCYP88A3	シロイヌナズナ	AAB71462.1
PsCYP88A7	エンドウ	AAO23064.1
ZmCYP88A1	トウモロコシ	NP_001105586.1
GmCYP88A26	ダイズ	XP_003516638.1
CaCYP89A35	トウガラシ	DQ114394
GmCYP89A36	ダイズ	DQ340245
ZmCYP89B17	トウモロコシ	CO465851.1
TmCYP89J1	一粒小麦	AY914081
SICYP88B1 (GAME4)	トマト	Solyc12g006460.1.1
SpimpCYP88B1 (GAME4)	カラントマト	contig_6356779
SpCYP88B1 (GAME4)	ソラナム・ピネリ	AW618484.1, BG135958.1
StCYP88B2 (GAME4)	ジャガイモ類Phureja	PGSC0003DMP400041994
StCYP88B1v2 (GAME4)	ジャガイモ類Tuberosum	PGSC0003DMP400041994
SICYP88C2	トマト	Solyc10g007860.2.1
SmCYP88B3 (GAME4)	ナス	FS071104, FS071103
OsCYP90A3	イネ	AC123526.1
SICYP90A5	トマト	Solyc06g051750.2.1
ScCYP90A8	オレンジ	DQ001728.1
ZcCYP90A11	ヒヤクニチソウ	BAE16977.1
PhCYP88C1	ペチュニア	AAZ39647.1
AaCYP90A13	クソニンジン	ABC94481.1
AtCYP710A1	シロイヌナズナ	AAC26690.1
SmCYP71A2	ナス	X71654.1
GmCYP93E1	ダイズ	AB231332
HICYP71C25	オオムギ	AY462228
NtCYP71D16	タバコ	AF166332
MeCYP71E7	キャッサバ	AY217351
TaCYP71F1	コムギ	AB036772
AoCYP71J1	アスパラガス	AB052131
MaCYP71N1v2	マレーヤマバショウ	AY062167
TaCYP72A6v1	コムギ	AF123604
ZmCYP72A16	トウモロコシ	AF465265
LeCYP72A51	トマト	Solyc10g051020.1.1
GmCYP72A61	ダイズ	DQ340241
MtCYP716A12	タルウマゴヤシ	ABC59076.1
StCYP716A13	ジャガイモ	PGSC0003DMP400013378
AaCYP716A14	クソニンジン	DQ363134
PsCYP716B2	シトカトウヒ	AY779543
SICYP718A6	トマト	Solyc07g055970.1.1
MtCYP718A8	タルウマゴヤシ	XP_003617455.1
PsCYP719B1	ケシ	EF451150

10

20

30

40

【表 4 - 2】

図12に記載の名称	ラテン名及び一般名	受託番号
StCYP72A186 (GAME7)	ジャガイモ	PGSC0003DMG402012386
StCYP72A186 (GAME7)	トマト	Solyc07g062520
StCYP72A188 (GAME6)	トマト	Solyc07g043460
StCYP72A188 (GAME6)	ジャガイモ	PGSC0003DMG400011750
GuCYP72A154	ウラルカンゾウ	BAL45206.1
MtCYP72A59	タルウマゴヤシ	ABC59078.1
NtCYP72A57	タバコ	ABC69414.1
NtCYP72A54	タバコ	ABC69417.1
CrCYP72A1	ニチニチソウ	gi461812
MtCYP72A63	タルウマゴヤシ	gi371940452
NpCYP72A2	ニコチアナ・ブルムバギニフォリア	AAB05376.3
StCYP734A7	トマト	Solyc03g120060.1.1
StCYP72A29	ジャガイモ	BAB86912.1
StSYP72a56	ジャガイモ	PGSC0003DMG400017325
StCYP72A208 (GAME8a)	ジャガイモ	PGSC0003DMG400026594
StCYP72A208 (GAME8b)	ジャガイモ	PGSC0003DMG400026586
StCYP72A208 (GAME8a)	トマト	TC243022
StCYP72A208 (GAME8b)	トマト	SGN-U578058

10

20

## 【0794】

代謝物の分析

## 【0795】

植物組織抽出物の調製並びにUPLC-qTOF-MS及びトマトの葉のフィトステロール含量による半極性化合物(ステロイド系アルカロイド及びステロイド系サポニンを含む)のプロファイリングを、既報(前掲の「Itkin et al., 2011」)のように行った。

## 【0796】

定量的リアルタイムPCRアッセイ

## 【0797】

RNAを単離し、定量的リアルタイムPCRを既報(前掲の「Itkin et al., 2011」)のように行った。また、ジャガイモ試料の内在性対照として、TIP41遺伝子(23)を用いた。オリゴヌクレオチドは、上記の表3に記載した。

## 【0798】

組換え酵素の産生

## 【0799】

GAME2、GAME17、及びGAME18をcDNAから増幅し、BamHI及びPstI(GAME2、GAME18)またはBamHI及びXhoI(GAME17)制限部位を用いてpACYCDUET-1にサブクローニングし、インサートをシーケンシングによって確認した。得られたプラスミド、pAC-GAME2/17/18を大腸菌BL21-DE3に形質転換した。GAME酵素を発現させるため、新鮮な終夜培養物を、クロラムフェニコール30µg/mLを含む2xYT培地25mL中で1:100に希釈し、A<sub>600nm</sub>が0.4となるまで、37及び250rpmでインキュベートし、GAME酵素を発現させた。次いで、IPTGを濃度が0.5mMとなるように添加し、インキュベーションを18及び250rpmで終夜継続した。翌日、遠心分離によって細胞を採取し、ペレットを、15%グリセロール、0.1mMEDTA及び5mM-メルカプトエタノールを含むpH=7.0の50mMトリスHCl2mL中に再懸濁した。細胞を超音波処理によって破壊した後、不溶性物質を遠心分離によって除去

30

40

50

し、可溶性分画を酵素の特性評価に用いた。タンパク質は、さらなる解析まで - 20 で保存した。

#### 【0800】

基質の調製

#### 【0801】

加水分解するため、 $\alpha$ -トマチン 35 mg を 1 N HCl 3 mL に溶解し、100 で15分間インキュベートした。次いで、溶液を氷上に置き、溶液のpHが9.0になるまでNH<sub>3</sub>を加えた。溶液を水飽和ブタノール4 mLで抽出した。ブタノール相を真空下で蒸発乾固させ、残存したペレットをメタノール1 mLに溶解し、さらなる使用まで - 20 で保存した。 $\alpha$ -トマチンの分解物を、Luna 5  $\mu$ m C18 (2) 100、LCカラム150 x 21.2 mm (Phenomenex、米国)により、25%アセトニトリル水溶液及び0.1%ギ酸によるアイソクラティック溶出を用いて分離した。化合物を3100質量検出器(Waters)で検出し、採取した。分画を凍結乾燥し、化合物の純度をLC-MSで確認した。生成物を同定するため、Waters 2996 PDA検出器に連結したWaters Alliance 2795 HPLCと、次いで正イオン化モードで稼働させたQTOF Ultima V4.00.00質量分析計(Waters, MS Technologies, 英国)とを用い、四重極飛行時間型質量分析法に連結した液体クロマトグラフィー(LC-QTOF-MS)を行った。C18プレカラム(2.0 x 4 mm; AJO-4286; Phenomenex、米国)に取り付けた分析用カラムLuna 3  $\mu$ m C18 (2) 100; 150 x 2.0 mm (Phenomenex、米国)を用いた。脱気した溶離液A [超純水:ギ酸(1000:1, v/v)]及び溶離液B [アセトニトリル:ギ酸(1000:1, v/v)]を流量0.19 mL/分で用いた。グラジエントを溶離液B 5%で開始し、45分間で溶離液Bを75%まで直線的に増加させ、その後カラムを洗浄し、次の注入前に15分間平衡化した。注入量は5  $\mu$ Lとした。この手順により、数ミリグラムの純粋な $\alpha$ -トマチン(トマチジン-ガラクトシド-グルコシド、T-Gal-Glu)及び1-トマチン(トマチジン-ガラクトシド-ジグルコシド、T-Gal-Glu-Glu)を得た。トマチジンガラクトシド(T-Gal)は、T-Gal-Gluによる強い混入のため、このようには精製できなかった。そこで、既報(前掲の「Itkin et al., 2011」)のように、トマチジン5 mgを、反応混合物1 mL中で、GAME 1及びUDP-ガラクトースと共にインキュベートした。T-Galを、固相抽出により、UDP-ガラクトースから精製した。Waters OASIS HLB 3 ccカラム(Waters Corp., Milford, MA)を100%メタノール6 mLでコンディショニングした後、超純水4 mLで洗浄した。10%メタノールを添加した反応混合物を充填し、次いでカートリッジを超純水4 mLで洗浄した。化合物を75%メタノール超純水溶液(v:v)1 mL及び100%メタノール0.4 mLで溶出した。完全に乾燥したペレットが得られるまで、高速真空濃縮器を用い、1つにまとめた溶出物から溶媒を除去した。

#### 【0802】

酵素アッセイ

#### 【0803】

基質T-Gal、1-及び $\alpha$ -トマチンを、1 mMとなるように、50% DMSOに溶解した。5 mM  $\beta$ -メルカプトエタノールを含むpH = 7.0の50 mM トリスHCl中で、最終反応体積100  $\mu$ L中、5  $\mu$ g/mLの酵素、8 mM UDP-キシロース、及び0.02 mM基質を用い、酵素アッセイを行った。攪拌しながら37 で2時間インキュベートした後、メタノール及び0.1%ギ酸300  $\mu$ Lを加えて反応を停止した後、短時間のボルテックスと15分間の超音波処理を行った。次いで、抽出物を13000 rpmで5分間遠心分離し、0.45  $\mu$ mフィルター(Minisart SRP4、Biotech GmbH、ドイツ)で濾過し、LC-MS(上記参照)によって分析した。生成物の量をLC-MSクロマトグラムにおけるピーク面積によって測定し、空のpACYCDUET-1を保有する大腸菌の酵素調製品の対照インキュベーションと比較し

た。検出に用いた質量は、 $1$ -トマチン ( $C_{50}H_{83}NO_{21}$ ;  $m/z = 1034.5$  ( $[M+H]^+$ ))、 $1$ -トマチン  $T-Gal-Glu-Glu$  ( $C_{45}H_{75}NO_{17}$ ;  $m/z = 902.51$  ( $[M+H]^+$ ))、 $2$ -トマチン ( $C_{44}H_{73}NO_{16}$ ;  $m/z = 872.50$  ( $[M+H]^+$ ))、 $2$ -トマチン  $T-Gal-Glu$  ( $C_{39}H_{65}NO_{12}$ ;  $m/z = 740.46$  ( $[M+H]^+$ ))、及び  $T-Gal$  ( $C_{33}H_{55}NO_7$ ;  $m/z = 578.41$  ( $[M+H]^+$ )) であった。

#### 【0804】

ウイルス誘導遺伝子サイレンシング (VIGS) 試験

#### 【0805】

既報 (「Orzaea D et al., 2009. Plant Physiol. 150:1122-1134」; 「Li R et al., 2006 J. Mass Spec. 41:1-22」) のようにして、GAME 遺伝子の断片を含むベクターを構築し、VIGS 試験を行った。空ベクター及びヘルパーベクター pTRV1 を含むアグロバクテリウムに感染させた植物を、対照として用いた。pTRV2\_\_DR\_\_GW ベクターの作製に用いたオリゴヌクレオチドは、上記の表 3 に示した。

10

#### 【0806】

ワイルドトマト種の遺伝子配列解析

#### 【0807】

3つのトマト野生種 (すなわち、ソラナム・ペネリ (*Solanum pennellii*)、カラントマト (*S. pimpinellifolium*)、及びソラナム・ケミエレウスキイ (*S. chmielewskii*) ) の部分的なリシーケンシングによって得られたゲノムデータ (Dr. Arnaud G. Bovy、未発表データ) を、トマトの 7 番及び 12 番染色体上の SGA 生合成遺伝子クラスターにアライメントする配列 (コンティグ) の有無について解析した。TopHat ツールキット (「Trapnell C. 2012. Nat. Protoc. 7:562-578」) を用い、野生種の読み取りを、リファレンスゲノムとして、トマトゲノム (ITAG 2.4) にマッピングした。マップされた読み取りを、IGV ゲノムブラウザ (「Robinson J T et al., 2011. Nat. Biotechnol. 29:24-26」) で可視化した。既存の栽培トマトの 7 番及び 12 番染色体の配列の遺伝子クラスターに対し、3つの野生種由来のコンティグの配列を組み立ててアライメントするため、CLCワークベンチ、CAP3 BWA 及び SAMtools ソフトウェアパッケージの組み合わせと、自家 Perl スクリプトとを用いた。

20

#### 【0808】

実施例 2 : SGA 生合成に関連する遺伝子

30

#### 【0809】

SGA 生合成に関連する遺伝子を見出すため、トマト及びジャガイモのトランスクリプトームデータを用いた共発現解析を行った。ジャガイモまたはトマトのいずれかにおける「ベイト」としての GAME1 / SGT1 (7 番染色体) 及び GAME4 (12 番染色体) の共発現を、以下の表 5 ~ 8 にヒートマップの形式で示す。GAME1 / SGT1 (7 番染色体) または GAME4 (12 番染色体) のいずれかと高度に共発現する遺伝子は、大きなフォント及び太字で示す。

#### 【0810】

表 5 : 受託番号、推定タンパク質、及び共発現 r 値 - トマト、7 番染色体

40

#### 【0811】

【表 5 - 1】

遺伝子名	推定タンパク質	トマトGAME1発現との 相関のr値
Solyc07g043310	アミノトランスフェラーゼ	-0.26
Solyc07g043320	未知タンパク質	0.12
Solyc07g043330	GRASファミリー転写因子	0.72
Solyc07g043340	未知タンパク質	
Solyc07g043350	未知タンパク質	
Solyc07g043360	60Sリボソームタンパク質L27	0.10
Solyc07g043370	トランスボゼース	
Solyc07g043380	未知タンパク質	
Solyc07g043390	セルロース合成酵素ファミリー タンパク質(GAME15)	0.92
Solyc07g043400	未知タンパク質	
Solyc07g043410	UDP-キシロースキシロ シルトランスフェラーゼ(GAME2)	
Solyc07g043420	2-オキシグルタル酸 依存性ジオキソゲナーゼ	0.79
Solyc07g043430	Gag-Polポリタンパク質	
Solyc07g043440	グルコシルトランスフェラーゼ様 タンパク質	
Solyc07g043450	ゼアチンO-グルコシルトランス フェラーゼ	
Solyc07g043460	シトクロムP450(GAME6)	0.91
Solyc07g043470	未知タンパク質	
Solyc07g043480	UDP-グルコースグルコ シルトランスフェラーゼ	0.88
Solyc07g043490	UDP-グルコシルトランスフェラーゼ ファミリー1タンパク質(GAME1)	1.00
Solyc07g043500	UDP-グルコシルトランスフェラーゼ	0.95
Solyc07g043510	システイン型ペプチダーゼ	-0.24
Solyc07g043520	トランスボゼース	
Solyc07g043530	未知タンパク質	
Solyc07g043540	未知タンパク質	
Solyc07g043550	UDP-アラビノース4-エピメラーゼ	0.70
Solyc07g043560	熱ショックタンパク質4	0.24
Solyc07g043570	アルドノケトレダクターゼ ファミリータンパク質	-0.09
Solyc07g043580	BHLH転写因子	0.43
Solyc07g043590	アミノキシナーゼファミリータンパク質	0.03
Solyc07g043600	ペントリコペプチドリピート 含有タンパク質	0.43
Solyc07g043610	オーキシン応答因子6	

10

20

30

40

【表 5 - 2】

遺伝子名	推定タンパク質	トマトGAME1発現との 相関のr値
Solyc07g043620	オーキシン応答因子6-1	0.65
Solyc07g043630	アシル-CoA合成酵素 /AMP-酸リガーゼII	
Solyc07g043640	アシル-CoA合成酵素 /AMP-酸リガーゼII	
Solyc07g043650	AMP- 依存性合成酵素及びリガーゼ	
Solyc07g043660	アシル-CoA合成酵素 /AMP-酸リガーゼII	-0.16
Solyc07g043670	ヒドロキシシナモイル CoAキナ酸トランスフェラーゼ2	
Solyc07g043680	エノイル-CoA-水和物	
Solyc07g043690	エノイル-CoA-水和物	
Solyc07g043700	アシルトランスフェラーゼ	

10

20

【0813】

表6：受託番号、推定タンパク質、及び共発現r値 - ジャガイモ、7番染色体

【0814】

【表 6】

遺伝子名	推定タンパク質	ジャガイモSGT1発現との 相関のr値
PGSC0003DMG400011754	$\gamma$ アミノ酪酸トランスアミナーゼ	-0.31
PGSC0003DMG400011753	Uro-接着因子A	-0.40
PGSC0003DMG400011742	DELLAタンパク質RGA	0.15
PGSC0003DMG400011741	60Sリボソームタンパク質L27	0.43
PGSC0003DMG400039612	機能未知の保存遺伝子	
PGSC0003DMG400011752	セルロース合成酵素 (GAME15)	0.90
PGSC0003DMG400011740	$\beta$ -ソラニンラムノシルトランス フェラーゼ (SGT3)	0.90
PGSC0003DMG400011751	2-オキシグルタル酸依存性 ジオキシゲナーゼ	0.87
PGSC0003DMG400011750	シトクロムP-450 (GAME6)	0.92
PGSC0003DMG400044993	未知タンパク質	
PGSC0003DMG400011749	ソラニンガラクトシルトランス フェラーゼ (SGT1)	1.00
PGSC0003DMG402015928	OTU様システインプロテアーゼ ファミリータンパク質	-0.24
PGSC0003DMG401015928	機能未知の保存タンパク質	-0.25
PGSC0003DMG400015927	UDP-アラビノース4-エピメラーゼ1	-0.21
PGSC0003DMG400015920	熱ショック70kDaタンパク質	-0.17
PGSC0003DMG402015926	アルド/ケトレダクターゼ	-0.05
PGSC0003DMG401015926	転写因子PIF5のアイソフォーム2	-0.33
PGSC0003DMG400015925	アミノキシダーゼ	0.11
PGSC0003DMG400015924	ペントトリコペプチドリピート含有 タンパク質	0.32
PGSC0003DMG400015919	ARF8	0.07
PGSC0003DMG400036440	AMP依存性リガーゼ	
PGSC0003DMG400015923	アシル:coAリガーゼ酢酸 -coA合成酵素	
PGSC0003DMG400015922	アシル:coAリガーゼ酢酸 -coA合成酵素	
PGSC0003DMG400044288	アシルトランスフェラーゼ	
PGSC0003DMG400015918	アシルトランスフェラーゼ	0.03

10

20

30

40

【0815】

表7：受託番号、推定タンパク質、及び共発現r値 - トマト、12番染色体

【0816】

【表7】

遺伝子名	推定タンパク質	トマトGAME4発現との 相関のr値
Solyc12g006530	シクロアルテノール合成酵素	0.08
Solyc12g006520	シクロアルテノール合成酵素	0.05
Solyc12g006510	シクロアルテノール合成酵素	-0.12
Solyc12g006500	リン酸輸送体タンパク質	0.15
Solyc12g006490	$\beta$ -1-3-ガラクトシル- $\alpha$ - グリコシル-糖タンパク質	0.03
Solyc12g006480	Nup205タンパク質	0.35
Solyc12g006470	$\gamma$ -アミノ酪酸アミノトランスフェラーゼ 様タンパク質	0.94
Solyc12g006460	シトクロムP450 (GAME4)	1.00
Solyc12g006450	$\gamma$ -アミノ酪酸アミノトランスフェラーゼ 様タンパク質	-0.13
Solyc12g006440	未知タンパク質	0.25
Solyc12g006430	UDP-グルクロノシルトランスフェラーゼ 1-182A1	
Solyc12g006420	トポイソメラーゼII関連タンパク質PAT1	0.08
Solyc12g006410	UDP-アラビノース4-エピメラーゼ	
Solyc12g006400	未知タンパク質	
Solyc12g006390	2-オキシグルタル酸依存性 ジオキシゲナーゼ	
Solyc12g006380	2-オキシグルタル酸依存性 ジオキシゲナーゼ	0.15
Solyc12g006370	アミノオキシダーゼファミリータンパク質	-0.16
Solyc12g006360	多剤耐性タンパク質mdtK	
Solyc12g006350	オーキシン応答因子6	0.35
Solyc12g006340	オーキシン応答因子6	0.47
Solyc12g006330	アシルトランスフェラーゼ様タンパク質	
Solyc12g006320	ATP依存性RNAヘリカーゼ	0.14
Solyc12g006310	小胞体-ゴルジ	0.25
Solyc12g006300	WDリピートタンパク質様	-0.03
Solyc12g006290	レティキュロンファミリータンパク質	0.19
Solyc12g006280	Myb様DNA結合タンパク質	

10

20

30

40

【0817】

表8：受託番号、推定タンパク質、及び共発現r値 - ジャガイモ、12番染色体

【0818】

【表 8】

遺伝子名	推定タンパク質	ジャガイモGAME4発現との 相関のr値
PGSC0003DMG400020034	$\beta$ -アミリン合成酵素	-0.13
PGSC0003DMG400024276	$\beta$ -アミリン合成酵素	-0.09
PGSC0003DMG400024277	機能未知の遺伝子	0.10
PGSC0003DMG400024278	フェニルアセトアルデヒド合成酵素	0.10
PGSC0003DMG400024279	機能未知の保存遺伝子	-0.16
PGSC0003DMG400024280	トリオースリン酸/リン酸輸送体、 非緑色色素体、葉緑体	-0.06
PGSC0003DMG400024271	アセチルグルコサミニルトランス フェラーゼ	-0.06
PGSC0003DMG400024273	耐性タンパク質PSH-RGH6	0.37
PGSC0003DMG400024281	$\gamma$ アミノ酪酸トランスアミナーゼ アイソフォーム2	0.94
PGSC0003DMG400024274	シトクロムP450モノオキシゲナーゼ GAME4	1.00
PGSC0003DMG400024275	$\gamma$ アミノ酪酸トランスアミナーゼ アイソフォーム3	0.37
PGSC0003DMG400024282	フォーチュン-1 (Fortune-1)	0.36
PGSC0003DMG400028806	UDP-グリコシルトランス フェラーゼ82A1様	-0.18
PGSC0003DMG401028807	トポイソメラーゼII関連タンパク質 PAT1	
PGSC0003DMG402028807	UDP-アラビノース4-エピメラーゼ	
PGSC0003DMG400028824	機能未知の遺伝子	
PGSC0003DMG400028808	2-オキシグルタル酸依存性 ジオキシゲナーゼ	-0.07
PGSC0003DMG400028809	2-オキシグルタル酸依存性 ジオキシゲナーゼ	0.61
PGSC0003DMG400028810	アミノオキシダーゼ	-0.04
PGSC0003DMG400028825	MATE輸送体	
PGSC0003DMG400028826	オーキシン応答因子6	
PGSC0003DMG400043090	インテグラーゼコアダメイン含有 タンパク質	
PGSC0003DMG400037700	WRKY転写因子27	
PGSC0003DMG400028811	アシルトランスフェラーゼ	
PGSC0003DMG400028812	DEADボックスATP依存性 RNAヘリカーゼ53	0.56
PGSC0003DMG400028814	WDリピートタンパク質	-0.10
PGSC0003DMG401028829	ポリガラクトソナーゼ	
PGSC0003DMG400028815	レティキュロンファミリータンパク質	0.08
PGSC0003DMG400028830	Myb様DNA結合タンパク質ドメイン、 SHAQKYFクラスファミリータンパク質	

10

20

30

40

各種由来の16の遺伝子が、GAME1/SGT1と共発現した(表9、図2)。これらの遺伝子のうちの1つは、以前GLYCOALKALOID METABOLISM 4 (GAME4)と呼ばれていたもので、シトクロムP450タンパク質の88Dサブファミリーのメンバーをコードする(図3)。GAME4及びGAME1/SGT1は、トマト及びジャガイモにおいて、極めて類似の発現プロファイルを示す(「WO 2010/095843」)。トマト及びジャガイモにおけるGAME1/SGT1及びGAME4遺伝子は、それらと共発現した遺伝子のいくつかと物理的に近くなるように、7番及び12番染色体に位置している(図2)。

#### 【0820】

GAME1/SGT1共発現遺伝子のクラスターは、7番染色体上の約200Kbpのゲノム領域に広がっている。GAME1と共に、トマトのクラスターは、7つの共発現遺伝子から構成されている。これらには、3つのUDP-グリコシルトランスフェラーゼ[GAME2(ジャガイモではSGT3と呼ぶ)、GAME17、及びGAME18]、72AサブファミリーのシトクロムP450(GAME6)、2-オキソグルタル酸依存性ジオキシゲナーゼ(GAME11)、及びセルロース合成酵素様タンパク質(GAME15)が含まれる。ジャガイモでは、GAME17及びGAME18 UDP-グリコシルトランスフェラーゼをコードするトマトの遺伝子のホモログがないため、このクラスターは5つの共発現遺伝子を含むものと考えられる。4つの組換えクラスター化トマトUDP-グリコシルトランスフェラーゼによって、酵素活性アッセイを行った。GAME17及びGAME18は、基質としてトマチジンガラクトシド(T-Gal)及び-トマチン(T-Gal-Glu)とそれぞれインキュベートした際、UDP-グルコシルトランスフェラーゼ活性を示した。一方、GAME2は、基質として1-トマチン(T-Gal-Glu-Glu)とインキュベートした際、UDP-キシロシルトランスフェラーゼ活性を示した(図4E、4F、及び4G)。GAME1は、トマトにおいて、トマチジンUDP-ガラクトシルトランスフェラーゼとして作用することが、既報において示されている(前掲の「Itkin et al., 2011」)。4つの組換えUGT酵素を単一の試験管内で、トマチジン及び全てのグリコシド供与体(UDP-ガラクトース、-グルコース及び-キシロース)とインキュベートすると、最終的なSGA生成物である-トマチンの蓄積が認められた(図4H)。

#### 【0821】

推定転写因子をコードする2つの遺伝子は、GAME1/SGT1及びGAME4と共発現した遺伝子の中から同定され(図4)、GAME9と呼ばれる1つの遺伝子は、トマトID Solyc01g090340及びジャガイモID PGSC0003DMG400025989によって同定された。この遺伝子はエチレン応答性要素結合因子13と呼ばれ、推定上のAP2ドメインを含む。もう一方の遺伝子は、トマトID Solyc03g046570及びジャガイモID PGSC0003DMG400012262によって同定されたBHLH-転写因子である。

#### 【0822】

表9：図2に示したジャガイモ及びトマトの両者における既知及び推定ステロイド系アルカロイド関連遺伝子と共発現させたホモログの詳細

#### 【0823】

トマト及びジャガイモの配列は、Sol Genomics Network(solgenomics.net)から入手した。共発現のr値 0.8。TCON番号：RNAseqデータの組み立てにおいて本発明者らによって付与されたコンティグ参照名。N/A：なし。

#### 【0824】

【表 9】

名称	トマトIDSolyc	ジャガイモの読み取り	トマトID
エクステンシン様タンパク質	Solyc01g006400	PGSC0003DMG400023230	TCONS_00007692
GAME 9	Solyc01g090340 [アミノ酸 SEQ ID NO: 13] [核酸 SEQ ID NO: 15]	PGSC0003DMG400025989 [アミノ酸 SEQ ID NO: 14] [核酸 SEQ ID NO: 16]	TCONS_00011729
$\delta$ (24)-ステロール レダクターゼ様	Solyc02g069490	PGSC0003DMG400021142	TCONS_00044548
BHLH転写因子	Solyc03g046570 [アミノ酸 SEQ ID NO:17] [核酸 SEQ ID NO: 18] [核酸 SEQ ID NO: 19]	PGSC0003DMG400012262 [核酸 SEQ ID NO: 20] [核酸 SEQ ID NO: 21]	TCONS_00055879
LRR受容体様 プロテインキナーゼ	Solyc05g009100	PGSC0003DMG400014576	TCONS_00101281
グリコシルトランスフェラーゼ	Solyc05g053120	PGSC0003DMG402027210	TCONS_00100675
セルロース合成酵素様 (GAME15)	Solyc07g043390	PGSC0003DMG400011752	TCONS_00135034
GAME6 (CYP72)	Solyc07g043460	PGSC0003DMG400011750	TCONS_00137734
GAME1 (ガラクト シルトランスフェラーゼ)	Solyc07g043490	PGSC0003DMG400011749	TCONS_00133014
GAME7 (CYP72)	Solyc07g062520 (GAME1 r値 0.66; GAME4 r値 0.71)	PGSC0003DMG402012386 (SGT1 r値 0.63; GAME4 r値 0.73)	TCONS_00132326
Srt/Thr プロテインキナーゼ6	Solyc08g066050	PGSC0003DMG400025461	TCONS_00151251
減数分裂 セリンプロテイナーゼ	Solyc08g077860	PGSC0003DMG401012339	TCONS_00149157
ステロールレダクターゼ	Solyc09g009040	PGSC0003DMG400002720	TCONS_00162820
ユビキチンタンパク質リガーゼ	Solyc10g008410	PGSC0003DMG400021683	TCONS_00183263
プロテイナーゼインヒビターII	Solyc11g020960	PGSC0003DMG402003479	TCONS_00194999
GAME4 (CYP88)	Solyc12g006460	PGSC0003DMG400024274	TCONS_00210154
$\gamma$ -アミノ酪酸アミノトランス フェラーゼ様タンパク質 (トランスアミナーゼ) (GAME12)	Solyc12g006470	PGSC0003DMG400024281	
$\beta$ -ソラニンラムノ シルトランスフェラーゼ (SGT3)	#N/A	PGSC0003DMG400011740	
2-オキソグルタル酸 依存性ジオキシゲナーゼ (GAME11)	Solyc07g043420 [アミノ酸 SEQ ID NO:22] [核酸 SEQ ID NO: 24] [核酸 SEQ ID NO: 25]	PGSC0003DMG400011751 [アミノ酸 SEQ ID NO:23] [核酸 SEQ ID NO: 26] [核酸 SEQ ID NO: 27]	
GAME18 (グリコシルトランスフェラーゼ)	Solyc07g043500	#N/A	
GAME17 (グリコシルトランスフェラーゼ)	Solyc07g043480	#N/A	

10

20

30

40

### 実施例 3 : GAME 9 転写因子の機能解析

#### 【0826】

対応する GAME 9 DNA断片を pK7GWIWG2 (II) バイナリーベクター及び pJCV52 バイナリーベクターにそれぞれ導入することにより、GAME 9サイレンシング (RNAi) コンストラクト及び過剰発現 (OX) コンストラクトを作製した。それぞれの GAME 9サイレンシングコンストラクト及び過剰発現コンストラクトによって形質転換した遺伝子導入トマト及びジャガイモ系を、既報 (前掲の「Itkin et al., 2011」) のように作製した。Itkinら (前掲の「2011」) の既報のように、組織抽出物を調製し、解析した。配列番号18は、GAME 9 RNAiサイレンシング分子の配列を示す。遺伝子導入及び野生型トマト植物及び/またはジャガイモ植物の葉及び/または塊茎から得た抽出物に対し、UPLC-TQ-MSを用いたステロイド系アルカロイドの代謝プロファイリングを行った。GAME 9をコードする遺伝子がサイレンシングされたジャガイモ系 (GAME 9-RNAi系) のジャガイモの塊茎の皮から得た抽出物では、  
- ソラニン及び - カコニンの低下が認められた (それぞれ図5A及び5B)。ジャガイモの GAME 9 - 過剰発現系の葉は、野生型に比べ、より高レベルの - ソラニン (図5C) 及び - カコニン (図5D) を含んでいた。同様の蓄積パターンがジャガイモの葉で認められ、RNAi系では - カコニン及び - ソラニンの量が低下し、GAME 9 - 転写因子を過剰発現する系では、これらのステロイド系アルカロイドの量が増加した (図6)。

10

#### 【0827】

トマトでは、GAME 9 - 転写因子 (5879と呼ばれる) を過剰反応する系の葉の抽出物が、野生型植物から得た葉の抽出物中の量に比べ、より高レベルの - トマチンを含んでいた。逆に、GAME 9 - 転写因子 (5871系) の発現をダウンレギュレーションさせたところ、 - トマチン含有量が著しく低下した。

20

#### 【0828】

### 実施例 4 : GAME 遺伝子の機能特性評価

#### 【0829】

### GAME 11サイレンシング植物

#### 【0830】

ウイルス誘導遺伝子サイレンシング (VIGS) は、植物の各種器官の遺伝子を全体的にサイレンシングすることが可能な、一般的に用いられる手法である (「Dinesh-Kumar S P et al., 2003. Methods Mol Biol 236:287-294」)。配列番号19は、GAME 11 RNAiサイレンシング分子の配列を示す。

30

#### 【0831】

クラスター中の推定ジオキシゲナーゼである GAME 11 を VIGS でサイレンシングしたものをを用いてトマトの葉を分析したところ、 - トマチンレベルが著しく低下し、いくつかのコレスタノール型ステロイド系サポニンが蓄積されていることが判明した。

#### 【0832】

トマトの GAME 11 ジオキシゲナーゼをサイレンシングすると、葉の - トマチンレベル ( $m/z = 1034.5$ ) が減少し (図8A)、コレスタノール型ステロイド系サポニン [すなわち、STSs;  $m/z = 1331.6$ 、 $1333.6$ 、 $1199.6$ 、 $1201.6$  (主要サポニン)] が蓄積する (図8B)。図8Cは、 $m/z = 1331.6$  ( $19.28$ 分) の MS/MS スペクトルを示す。図8Dは、GAME 11 をサイレンシングした葉に蓄積する、 $19.28$ 分で溶出するサポニンのフラグメンテーションパターンを示す。図8Cの MS/MS クロマトグラムにおいて、対応する質量シグナルをアスタリスクで示した。元素組成及びフラグメンテーションパターンによると、化合物は、(フロスタノール型サポニンに比べて) 水酸基1つとE環がなく、そのためE環の互変異性化及びマクラファティ転位の代わりに複数の水分子の消失を伴うフラグメンテーションをもたらす、コレスタノール型サポニンであることがわかる。

40

#### 【0833】

50

## GAME 18サイレンシング植物

## 【0834】

- トマチンの四糖類部分の形成におけるGAME 18の役割は、ウイルス誘導遺伝子サイレンシング(VIGS)アッセイにおいて、GAME 18をサイレンシングした果実では、対照試料には存在しなかった - トマチンが蓄積されたことから支持される(図9)。

## 【0835】

GAME 18をサイレンシングした成熟緑色果実から抽出した代謝物のうち、- トマチン標準物質( $m/z = 740.5$ ) (図9A~C)及び - トマチンペントシド( $m/z = 872.5$ ) (図9D~9E)に対応する、新たに蓄積した化合物のピークが検出された。

10

## 【0836】

## GAME 12サイレンシング植物

## 【0837】

トマトのGAME 12トランスアミナーゼをサイレンシングしたところ、フロスタノール型ステロイド系サポニンが蓄積した(図4D)。図10Aは、GAME 12をサイレンシングした葉においてSTS( $m/z = 753.4$ )が蓄積することを示す。一方、STSは、野生型の葉では少量のみが存在する(図10B)。図10Cは、 $19.28$ 分における $m/z = 753.4$ のMS/MSスペクトルを、フラグメントの説明と共に示す。MS/MS試験で認められた特徴的な質量フラグメントから判断した、 $19.71$ 分におけるSTSについて提案される構造を図10Dに示す。

20

## 【0838】

## GAME 7及びGAME 8の機能

## 【0839】

強固に共発現し、ゲノムの他の箇所位置する遺伝子についても、機能的な検討を行った。GAME 7及びGAME 8と呼ばれる2つの遺伝子は、シトクロムP450のCYP72サブファミリーに属する。GAME 7は両種(ジャガイモ及びトマト)で共発現したのに対し、StGAME 8a及びStGAME 8bはジャガイモでStSGT1及びStGAME 4と強く共発現した。現在、本発明者らは、GAME 7についてのSGA関連活性を実証できていないが、GAME 6については、SGAの代謝への関与が示唆されている(「US 20120159676」)。さらに、GAME 8をサイレンシングしたトマトの葉では、SGAの生合成経路(図1)において中間体として提案されている22-(R)-ヒドロキシコレステロールが蓄積された(図11A~11D)。GAME 8をサイレンシングした系では、野生型(図11D)に比べ、両異性体を蓄積する。(R)異性体がより多いことから、GAME 8の基質である可能性が高い。

30

## 【0840】

図12は、植物CYP450タンパク質ファミリーにおけるGAME遺伝子の系統樹を示す。枝上の番号は、各節を支持するブートストラップ反復の比率を示す。

## 【0841】

実施例5：ナス科植物で提案されている生合成経路

40

## 【0842】

トマトに関する図13(点線の矢印は経路における複数の酵素反応を表す)の概略図に示したように、ナス科植物では、拡張した生合成経路が提案されている。この経路は、便宜上4つの部分に分けることができる。パートIでは、一連の反応(例えば、SSR2、SMO3、SMO4によって触媒される)により、シクロアルテノールがコレステロールに変換される。副産物には、トリテルペノイド及びフィトステロールが含まれる。パートIIでは、一連の反応(例えば、GAME 11、GAME 6、GAME 4、GAME 12、GAME 25によって触媒される)により、コレステロールがトマチン(アグリコン)に変換される。副産物には、ステロイド系サポニン(例えば、ウットロシド(uttroside)B)が含まれる。パートIIIでは、一連の反応(例えば、GAME 1、GAME 17

50

、GAME 18、GAME 2によって触媒される)により、トマチンがステロイド系グリコアルカロイド(例えば、 $\beta$ -トマチン)に変換される。パートIVでは、一連の反応により、緑色のトマトのステロイド系グリコアルカロイド(例えば、 $\beta$ -トマチン)が赤色のトマトのリコペロシド(lycoperosides)及び/またはエスクレオシド(esculeosides)(例えば、エスクレオシドA)に変換される。

【0843】

実施例6：トマト、ジャガイモ、及びナスにおけるステロイド系グリコアルカロイド(SGA)生合成に關与する経路

【0844】

トマト、ジャガイモ、及びナスにおけるセルロース合成酵素様遺伝子(GAME 15)は、ステロイド系グリコアルカロイド(SGA)の生合成(図14A~14C)に關連することが確認されている。この遺伝子は、その他のSGA生合成遺伝子(例えば、GAME 4、GAME 12)及びSGA生合成の制御因子(例えば、GAME 9)とも、強く共発現されることが示されている。図14Dは、ナス科植物において提案された経路の非限定的な例を示す。例えば、これに限定されるものではないが、当業者であれば、GAME 15のグルクロン酸トランスフェラーゼ活性によるトマチジンからのトマチジン3-O-グルクロニドの産生が、この経路に包含されることを認識するであろう。特定の実施形態において、GAME 15酵素は、ナス科植物、例えば、これらに限定されるものではないが、ジャガイモ(*Solanum tuberosum*)、トマト(*Solanum lycopersicum*)、ズルカマラ(*Solanum dulcamara*)、及びナス(*Solanum melogena*)に存在するGAME 15を含む

【0845】

配列は以下のように同定された。

【0846】

セルロース合成酵素様\_\_トマト

【0847】

ATGAA A A A A A A C C A T G G A G C T C A A C A A A A G C A C T G T T C C A  
 C A A C C T A T C A C C A C C G T A T A C C G A C T C C A C A T G T T C A T C C  
 A C T C A A T A A T C A T G C T T G C A T T A A T A T A C T A C C G T G T A T C  
 T A A T T T G T T T A A A T T C G A A A A C A T T C T C A G T T T A C A A G C A  
 C T T G C T T G G G C G C T C A T C A C T T T T G G T G A A T T T A G T T T C A  
 T T C T C A A G T G G T T C T T C G G A C A A G G T A C T C G T T G G C G C C C  
 C G T T G A A C G A G A T G T T T T C C C T G A A A A C A T T A C T T G C A A A  
 G A T T C C G A T C T A C C G C C A A T T G A C G T A A T G G T A T T C A C T G  
 C C A A T C C T A A G A A A G A G C C A A T T G T A G A T G T C A T G A A C A C  
 T G T G A T A T C C G C A A T G G C T C T T G A T T A T C C C A C C G A T A A A  
 T T G G C T G T G T A T C T C G C T G A T G A T G G A G G A T G T C C A T T G T  
 C G T T G T A C G C C A T G G A A C A A G C G T G T T T G T T T G C A A A G C T  
 A T G G T T A C C T T T C T G T A G A A A C T A T G G A A T T A A A A C G A G A  
 T G C C C A A A A G C A T T T T T T T C T C C G T T A G G A G A T G A T G A C C  
 G T G T T C T T A A G A A T G A T G A T T T T G C T G C T G A A A T G A A A G A  
 A A T T A A A T T G A A A T A T G A A G A G T T C C A G C A G A A G G T G G A A  
 C A T G C T G G T G A A T C T G G A A A A A T C A A T G G T A A C G T A G T G C  
 C T G A T A G A G C T T C G C T T A T T A A G G T A A T A A A C G A G A G G G A  
 G A A C G A A A A G A G T G T G G A T G A T A T G A C G A A A A T G C C C T T G  
 C T A G T T T A T G T A T C C C G T G A A A G A A G A T T C A A C C G T C T T C  
 A T C A T T T C A A G G G T G G A T C T G C A A A T G C T C T A C T T C G A G T  
 T T C T G G A A T A A T G A G T A A T G C C C C C T A T G T A C T G G T G T T A  
 G A T T G T G A T T T C T T C T G T C A T G A T C C A A T A T C A G C T A G G A  
 A G G C A A T G T G T T T T C A T C T T G A T C C A A A G C T A T C A T C T G A

10  
20  
30  
40  
50

TTTAGCCTATGTTTCAGTTC CCTCAAGTCTTTTACAATGTC  
AGCAAGTCA GATATTTATGATGTC AAAATTAGACAGGCTT  
ACAAGACAATATGGCATGGAATGGATGGTATCCAAGGCC  
AGTGTTATCTGGGACTGGTTATTTTCTCAAGAGGAAAGCG  
TTATACACAAGTCCAGGAGTAAAAGAGGCGTATCTTAGTT  
CACCGGAAAAGCATTTTGGAAAGGAGTAAAAGGTTTCTTG  
TTCATTAGAGGAGAAAATGGTTATGTTAAGGCAGATAAA  
GTCATATCAGAAGATATCATAGAGGAAGCTAAGATGTTAG  
CTACTTGTGCATATGAGGATGGCACACATTGGGGTCAAGA  
GATTGGTTATTCATACGATTGTCATTTGGAGAGCACTTTT  
ACTGGTTATCTATTACACTGCAAAAGGGTGGACATCTACTT  
ATTTGTATCCAGACAGGCCATCTTTCTTGGGTTGTGCCCC  
AGTTGATATGCAAGGTTTCTCATCACAGCTCATCAAATGG  
GTTGCTGCACCTTACACAAGCTGGTTTATCACATCTCAATC  
CCATCACTTATGGTTTGGAGTAGTAGGATGAGGACTCTCCA  
ATGCATGTGCTATGCCTATTTGATGTATTTCACTCTTTAT  
TCTTGGGGAATGGTTATGTATGCTAGTGTTCCTTCTATTG  
GCCTTTTGTGGACTTCCAAGTCTATCCTGAGGTA CATGA  
TCCGTGGTTTGCAGTGTATGTGATTGCTTTCATATCGACA  
ATTTTGGGAGAATATGTCGGAGTCAATTCCAGAAGGGGGAT  
CAGTTAAAACGTGGTGGATGGAATACAGGGCATTGATGAT  
GATGGGAGTTAGCGCAATATGGTTAGGAGGATTGAAAGCT  
ATATATGACAAGATAGTCCGGAACACAAGGAGAGAAATTGT  
ATTTGTCCGGACAAGGCCAATTGACAAGGAAAAGCTCAAGAA  
ATACGAGAAGGGCAAAATTTGATTTCCAAGGAATAGGGATA  
CTTGCTCTGCCACTGATAGCATTTTCCGTGTTGAACCTCG  
TAGGCTTTCATTGTTGGAGCTAATCATGTCTTTATTACTAT  
GAAC TACGCAGGC GTGCTGGGCCAACTCCTCGTATCATCG  
TTCCTTCGTCTTTGTTGTCGTC ACTGTTGTCATTGATGTTG  
TATCTTTCTTAAAGGTTTCTTAA [配列番号30]。

10  
20  
30

【0848】

セルロース合成酵素様(トマト)

【0849】

MKKTME LNKSTVPQPITTVYRLHMF IHSIIMLAL IYYRV  
SNLKFENILSLQALAWALITFG EFSFILKWF FGQGTRWR  
PVERDVPENITCKDSDLPPIDVMVFTANPKKEPIVDVMN  
TVISAMALDYPTDKLAVYLADDGGCPLSLYAMEQACLF AK  
LWLPFCRNYGIKTRCPKAFFSPLGDDDRVLKND DFAAEMK  
EIKLKYE EFQQVEHAGESG KINGNVVPDRASLIKVINER  
ENEKSVDDMTKMPLLVYVSRE RRFNRLHHFKGGSANALLR  
VSGIMSNA PYVLVLD CDF FCHDPI SARKAMCFHLD PKLSS  
DLAYVQFPQVFYNVSKSDIYDVKIRQAYKT IWHGMDGIQG  
PVLSGTG YFLKRKALY TSPGVKEAYLSSPEKHFGRSKRFL  
ASLEEKNGYVKADKVI SEDIEEAKMLATCAYEDGTHWGQ  
EIGYSYDCHLESTFTGYLLHCKGWTSTYLYPDRPSFLGCA  
PVDMQGFS SQLIKWVAAL TQAGLSHLNPITYGLSSRMRTL  
QCMCYAYLMYFTLYSWGMVMYASVPSIGLLFDFQVYPEVH  
DPWF AVYVIAFISTILENMSSES IPEGGSVKTWMEYRALM  
MMGVSAIWL GGLKAIYDKIVGTQGEKLYLSDK AIDKEK LK  
KYEK GKFD FQGI GILALPLIAFSV LNLVGFIVGANHV FIT

40  
50

M N Y A G V L G Q L L V S S F F V F V V V T V V I D V V S F L K V S [ 配列番号 3  
1 ]。

【 0 8 5 0 】

セルロース合成酵素様 ( *solanum pennellii* )

【 0 8 5 1 】

A T G A A A A A A A C C A T G G A G C T C A A C A A A A G C A C T G T T C C A  
C A A C C T A T C A C C A C C G T A T A C C G A C T C C A C A T G T T C A T C C  
A C T C A A T A A T C A T G C T T G C A T T A A T A T A C T A C C G T G T A T C  
T A A T T T G T T T A A A T T C G A A A A C A T T C T C A G T T T A C A A G C A  
C T T G C T T G G C T A C T C A T C A C T T T T G G T G A A T T T A G T T T C A  
T T C T C A A G T G G T T C T T C G G A C A A G G A A C T C G T T G G C G C C C  
C G T T G A A C G A G A T G T T T T C C C T G A A A C A T T A C T T G C A A A  
G A T T C C G A T C T A C C G C C A A T T G A C G T A A T G G T G T T C A C T G  
C C A A T C C T A A G A A A G A G C C A A T T G T A G A T G T C A T G A A C A C  
T G T G A T A T C C G C A A T G G C T C T T G A T T A T C C C A C C G A T A A A  
T T G G C T G T G T A T C T G G C C G A T G A T G G A G G A T G T C C A T T G T  
C C T T G T A C G C C A T G G A A C A A G C A T G T T T G T T T G C A A A G C T  
A T G G T T A C C T T T C T G T A G A A A G T A T G G A A T T A A A A C G A G A  
T G C C C A A A A G C A T T T T T T T C T C C G T T A G G A G A T G A T G A C C  
G T G T T C T T A A G A A T G A T G A T T T T G C T G C T G A A A T G A A A G A  
A A T T A A A T T G A A A T A T G A A G A G T T C C A G C A G A A C G T G G A A  
C A T G C T G G T G A A T C T G G A A A A A T C A A T G G C A A C G T A G T G C  
C T G A C A G A G C T T C G C T T A T T A A G G T A A T A A A C G A G A G G G A  
G A A C G A A A A G A G T G T C G A T G A T T T A A C G A A A A T G C C C T T G  
C T A G T T T A T G T A T C C C G T G A A A G A A G A T T C A A C C G T C T T C  
A T C A T T T C A A G G G T G G A T C T G C A A A T G C T C T A C T T C G A G T  
T T C T G G A A T A A T G A G T A A T G C C C C C T A T G T A C T G G T G T T A  
G A T T G T G A T T T C T T C T G T C A T G A T C C G A T A T C A G C T A G G A  
A A G C A A T G T G T T T T C A T C T T G A T C C A A A G C T A T C A T C T G A  
T T T A G C C T A T G T T C A G T T C C C T C A A G T C T T T T A C A A T G T C  
A G C A A G T C C G A T A T T T A T G A T G T C A A A A T T A G A C A G G C T T  
A C A A G A C A A T A T G G C A T G G A A T G G A T G G T A T C C A A G G C C C  
A G T G T T A T C T G G A A C T G G T T A T T T T C T C A A G A G G A A G G C G  
T T A T A C A C A A G T C C A G G A G T A A A A G A G G C G T A T C T T A G T T  
C A C C G G A A A A G C A T T T T G G A A G G A G T A A A A A G T T C C T T G C  
T T C A T T A G A G G A G A A A A A T G G T T A T G T T A A G G C A G A T A A A  
G T C A T A T C A G A A G A T A T C A T A G A G G A A G C T A A G A T C T T A G  
C T A C T T G T G C A T A T G A G G A T G G C A C A C A T T G G G G T C A A G A  
G A T T G G T T A T T C A T A C G A T T G T C A T T T G G A G A G C A C T T T T  
A C T G G T T A T C T A T T A C A C T G C A A A G G G T G G A C A T C T A C T T  
A T T T G T A T C C A G A C A G G C C A T C T T T C T T G G G T T G T G C C C C  
A G T T G A T A T G C A A G G T T T C T C A T C A C A G C T C A T A A A A T G G  
G T T G C T G C A C T T A C A C A A G C T G G T C T A T C A C A T C T C A A T C  
C C A T C A C T T A T G G T T T G A G T A G T A G G A T G A G A A C T C T C C A  
A T G C A T G T G C T A T G C C T A T T T G A T G T A T T T C A C T C T T T A T  
T C T T G G G G A A T G G T T A T G T A T G C T A G T G T T C C T T C T A T T G  
G C C T T T T G T T T G G C T T C C A A G T C T A C C C T G A G G T A C A T G A  
T C C A T G G T T T G C A G T G T A T G T G A T T G C T T T C A T A T C G A C A  
A T T T T G G A G A A T A T G T C G G A G T C A A T T C C A G A A G G G G G A T  
C A G T T A A A A C G T G G T G G A T G G A A T A C A G G G C A T T G A T G A T

10

20

30

40

50

G A T G G G A G T T A G C G C A A T A T G G T T A G G A G G A T T G A A A G C T  
 A T A T A T G A C A A G A T A G T C G G A A C A C A A G G A G A G A A A T T G T  
 A T T T G T C G G A C A A G G C A A T T G A C A A G G A A A A G C T C A A G A A  
 A T A C G A G A A G G G C A A A T T T G A T T T C C A A G G A A T A G G G A T A  
 C T T G C T C T G C C A T T G A T A G C A T T T T C C G T G T T G A A C C T C G  
 T A G G C T T C A T T G T T G G A G C T A A T C A T G T C T T T A T T A C T A T  
 G A A C T A C G C A G G C G T G C T G G G C C A A C T C C T C G T A T C A T C A  
 T T C T T C G T C T T T G T T G T C G T C A C T G T T G T C A T T G A T G T T G  
 T A T C T T T C T T A A A G G T T T C T T A A [ 配列番号 3 2 ] 。

10

【 0 8 5 2 】

セルロース合成酵素様 ( solanum pennellii )

【 0 8 5 3 】

M K K T M E L N K S T V P Q P I T T V Y R L H M F I H S I I M L A L I Y Y R V  
 S N L F K F E N I L S L Q A L A W L L I T F G E F S F I L K W F F G Q G T R W R  
 P V E R D V F P E N I T C K D S D L P P I D V M V F T A N P K K E P I V D V M N  
 T V I S A M A L D Y P T D K L A V Y L A D D G G C P L S L Y A M E Q A C L F A K  
 L W L P F C R K Y G I K T R C P K A F F S P L G D D D R V L K N D D F A A E M K  
 E I K L K Y E E F Q Q N V E H A G E S G K I N G N V V P D R A S L I K V I N E R  
 E N E K S V D D L T K M P L L V Y V S R E R R F N R L H H F K G G S A N A L L R  
 V S G I M S N A P Y V L V L D C D F F C H D P I S A R K A M C F H L D P K L S S  
 D L A Y V Q F P Q V F Y N V S K S D I Y D V K I R Q A Y K T I W H G M D G I Q G  
 P V L S G T G Y F L K R K A L Y T S P G V K E A Y L S S P E K H F G R S K K F L  
 A S L E E K N G Y V K A D K V I S E D I I E E A K I L A T C A Y E D G T H W G Q  
 E I G Y S Y D C H L E S T F T G Y L L H C K G W T S T Y L Y P D R P S F L G C A  
 P V D M Q G F S S Q L I K W V A A L T Q A G L S H L N P I T Y G L S S R M R T L  
 Q C M C Y A Y L M Y F T L Y S W G M V M Y A S V P S I G L L F G F Q V Y P E V H  
 D P W F A V Y V I A F I S T I L E N M S E S I P E G G S V K T W W M E Y R A L M  
 M M G V S A I W L G G L K A I Y D K I V G T Q G E K L Y L S D K A I D K E K L K  
 K Y E K G K F D F Q G I G I L A L P L I A F S V L N L V G F I V G A N H V F I T  
 M N Y A G V L G Q L L V S S F F V F V V V T V V I D V V S F L K V S [ 配列番号 3  
 3 ] 。

20

30

【 0 8 5 4 】

セルロース合成酵素様 ( ジャガイモ )

【 0 8 5 5 】

A T G A A A A A A A C C A T G G A G C T C A A C A A A A G C A C T G T T C C A  
 C A A C C T A T C A C C A C C A T A T A C C G A C T C C A C A T G T T T A T C C  
 A C T C T A T A A T C A T G G T T G C A T T A A T A T A C T A C C G T G T A T C  
 T A A T T T G T T T A A A T T C G A A A A C A T T C T G A G T T T A C A A G C A  
 C T T G C T T G G G T A C T C A T C A C T T T T G G T G A A T T T A G T T T C A  
 T T C T C A A G T G G T T C T T C G G A C A A G G A A C T C G T T A T C G C C C  
 T G T T G A A A G A G A T G T T T T C C C T G A A A A C A T A A C T T G C A A A  
 G A T T C C G A T C T A C C A C C A A T T G A C G T A A T G G T A T T C A C T G  
 C C A A T C C T A A G A A A G A G C C A A T T G T G G A T G T C A T G A A C A C  
 T G T G A T A T C C G C A A T G G C T C T T G A T T A T C C T A C G G A T A A A  
 T T G G C T G T G T A T C T G G C T G A T G A T G G A G G A T G T C C T T T G T  
 C A T T G T A C G C C A T G G A A G A A G C A T G T G T G T T T G C A A A G C T  
 G T G G C T A C C T T T C T G T A G G A A G T A T G G A A T T A A A A C T A G A  
 T G C C C T A A A G C G T T T T T T T C T C C T T T A G G A G A T G A T G A A C  
 G T G T T C T T A A G A A T G A T G A T T T T G A T G C T G A A A T G A A A G A  
 A A T T A A A T T G A A A T A T G A A G A G T T C C A G C A G A A T G T G G A A

40

50

CGTGCTGGTGAATCTGGAAAAAATCAATGGTAACGTAGTGCC  
 CTGATAGAGCCTCGTTTATTAAAGGTAATAAACGACAGAAA  
 AGCGGAGAGCGAAAAGAGTGC CGATGATTTAACGAAAATG  
 CCCTTGCTAGTTTATGTATCCCGTGAAAAGAAGATTCAACC  
 GTCCTTCATCACTTCAAGGGTGGATCTGC AAAATGCTCTTCT  
 TCGAGTTTCTGGAAATAATGAGTAATGCCCCCTATAACTG  
 GTGTTAGATTGTGATTTCTTCTGT CATGATCCAATATCAG  
 CTAGGAAGGCAATGTGTTTTTCATCTTGATCCAAGCTATC  
 ATCTGATTTAGCTTATGTTTCAGTTC CCTCAAGTCTTTTAC  
 AATGTCAGCAAGTCCGATATTTATGATGTC AAAATTAGAC  
 AGGCTTACAAGACAATAATGGCATGGAATGGATGGTATCCA  
 AGGCCCAGTGTTATCAGGA ACTGGTTATTTTCTGAAGAGG  
 AAGGCGTTATACACGAGTCCAGGAGTAAAGGAGGAGTATC  
 TTAGTTTACCCGGA AAAAGCATTTTGGAAAGGAGTAAAAAGTT  
 CCTTGCTTCACTAGAGGAGAAAAAATGGTTATGTTAAGGCA  
 GAGAAAAGTCAATATCAGAA GATATCGTAGAGGAAGCTAAGA  
 CCTTAGCTACTTTGTGCATATGAGGATGGCACACATTGGGG  
 TCAAGAGATTGGTTATTCATACGATTGTCATTTGGAGAGC  
 ACTTTTACTGGTTATCTATTACACTGC AAAAGGGTGGAGAT  
 CGACTTATTTGTATCCAGACAGGCCATCTTCTTGGGTTG  
 TGCCCCAGTTGATATGCAAGGTTTCTCCTCACAGCTCATA  
 AAATGGGTTGCTGC ACTTACACAAGCTGGTTTATCACATC  
 TCAATCCCATCACTTATGGCTTTAGTAGCAGGATGAAAAC  
 TCTCCAATGCATGTGCTATGCCTATTTGATATAATTTCACT  
 CTTTATTTCTTGGGGAAATGGTTCTATAATGCTAGTGTTCCTT  
 CTATTGGCCTTTTGT TTGGCTTCCAAGTCTATCCCGATGT  
 ACATGATCCATGGTTTGCAGTGTATGTGATTGCTTTTCATA  
 TCGGCAATTTTGGAGAAATATGTCGGAGTCAATTCCTGATG  
 GGGGATCATTTA AATCTTGGTGGATGGAATACAGGGCACT  
 GATGATGATGGGAGTTAGTGC AATAATGGTTAGGAGGATTG  
 AAAGCTATATTAGACAGGATAATCGGAACAGAAAGGAGAGA  
 AATTGTATTTATCGGACAAGGCAATTGACAAAGGAAAAGCT  
 CAAGAAAATACGAGAAAGGGGA AATTTGATTTCCAAGGAATA  
 GGGATACTTGTGTACCATTTGATAGCATTTTCTTGTGTA  
 ACCTCGTAGGCTTCAATTGTTGGAGCTAATCATGTCTTTAT  
 TACTATGA ACTACGCAGGTGTGCTTGGCCA ACTCCTCGTA  
 TCATCCTTCTTCGTCTTTGTCTGGTCACTGTTGTCAATTG  
 ATGTCGTTTCTTTCTTA AAAAGGTTTCTTA A [ 配列番号 34 ]。

10

20

30

【0856】

セルロース合成酵素様 ( ジャガイモ )

40

【0857】

MELNKSTVPQPITTIYRLHMF IHSIIMVALIYYRVSNLF  
 KFENILSLQALAWVLITFG EFSFILKWF FGQGTRYR PVER  
 DVFPENITCKDSDLPPIDVMVFTANPKKEPIVDVMNTVIS  
 AMALDYPTDKLAVYLADDGGCPLSLYAMEEACVFAKLWLP  
 FCRKYGIKTRCPKAFFSPLGDDERVLKNDDFDAEMKEIKL  
 KYEEFQQNVERAGESGKINGNVVPDRASF I KVINDRKAES  
 EKSADDLTKMPLL VYVSRERRFNRLHHFKGGSANALLRVS  
 GIMSNAPIYILVLD CDF FCHDPI SARKAMCFHLDPKLSSDL  
 AYVQFPQVFY NVSKSDIYDVKIRQAYKTIWHGMDGIQGPV

50

L S G T G Y F L K R K A L Y T S P G V K E E Y L S S P E K H F G R S K K F L A S  
 L E E K N G Y V K A E K V I S E D I V E E A K T L A T C A Y E D G T H W G Q E I  
 G Y S Y D C H L E S T F T G Y L L H C K G W R S T Y L Y P D R P S F L G C A P V  
 D M Q G F S S Q L I K W V A A L T Q A G L S H L N P I T Y G F S S R M K T L Q C  
 M C Y A Y L I Y F T L Y S W G M V L Y A S V P S I G L L F G F Q V Y P D V H D P  
 W F A V Y V I A F I S A I L E N M S E S I P D G G S F K S W W M E Y R A L M M M  
 G V S A I W L G G L K A I L D R I I G T E G E K L Y L S D K A I D K E K L K K Y  
 E K G K F D F Q G I G I L A V P L I A F S L L N L V G F I V G A N H V F I T M N  
 Y A G V L G Q L L V S S F F V F V V V T V V I D V V S F L K V S [ 配列番号 3 5 ]

10

【 0 8 5 8 】

セルロース合成酵素様 ( solanum chacoense )

【 0 8 5 9 】

A T G A A A A A A A C C A T G G A G C T C A A C A A A A G C A C T G T T C C A  
 C A A C C T A T C A C C A C C A T A T A C C G A C T C C A C A T G T T C G T C C  
 A T T C T A T A A T C A T G G C T G C A T T A A T A T A C T A C C G T G T A T C  
 T A A T T T G T T T A A A T T C G A A A A C A T T C T G A G T T T A C A A G C A  
 C T T G C T T G G G T A C T C A T C A C T T T T G G T G A A T T T A G T T T C A  
 T T C T C A A G T G G T T C T T C G G A C A A G G A A C T C G T T G G C G C C C  
 T G T T G A A A G A G A T G T T T T C C C T G A A A A C A T A A C T T G C A A A  
 G A T T C C G A T C T A C C A C C A A T T G A C G T A A T G G T A T T C A C T G  
 C C A A T C C T A A G A A A G A G C C A A T T G T G G A T G T C A T G A A C A C  
 T G T G A T A T C C G C A A T G G C T C T A G A T T A T C C T A C G G A T A A A  
 T T G G C T G T G T A T C T G G C T G A T G A T G G A G G A T G T C C T T T G T  
 C A T T G T A C G C C A T G G A A G A A G C A T G T G T G T T T G C A A A G C T  
 G T G G C T A C C T T T C T G T A G G A A G T A T G G A A T T A A A A C C A G A  
 T G C C C T A A A G C G T T T T T T T C T C C T T T A G G A G A T G A T G A C C  
 G T G T T C T T A A G A A T G A T G A T T T T G A T G C T G A A A T G A A A G A  
 A A T T A A A T T G A A A T A T G A A G A G T T C C A G C A G A A T G T G G A A  
 C G T G C T G G T G A A T C T G G A A A A A T C A A T G G T A A C G T A G T G C  
 C T G A T A G A G C C T C G T T T A T T A A G G T A A T A A A C G A C A G A A A  
 A A C G G A G A G C G A A A A G A G T G C C G A T G A T T T A A C G A A A A T G  
 C C C T T G C T A G T T T A T G T A T C C C G T G A A A G A A G A T T C A A C C  
 G T C T T C A T C A C T T C A A G G G T G G A T C T G C A A A T G C T C T T C T  
 T C G A G T T T C T G G A A T A A T G A G T A A T G C C C C T A T A T A C T G  
 G T G T T A G A T T G T G A T T T C T T C T G T C A T G A T C C A A T A T C A G  
 C T A G G A A G G C A A T G T G T T T T C A T C T T G A T C C A A A G C T A T C  
 A T C T G A T T T A G C T T A T G T T C A G T T C C C T C A A G T C T T T T A C  
 A A T G T C A G C A A G T C C G A T A T T T A T G A T G T C A A A A T T A G A C  
 A G G C T T A C A A G A C A A T A T G G C A T G G A A T G G A T G G T A T C C A  
 A G G C C C A G T G T T A T C A G G A A C T G G T T A T T T T C T G A A G A G G  
 A A G G C G T T A T A C A C G A G T C C A G G A G T A A A G G A G G A G T A T C  
 T T A G T T C A C C G G A A A A G C A T T T T G G A A G G A G T A A A A A G T T  
 C C T T G C T T C A C T A G A G G A G A A A A A T G G T T A T G T T A A G G C A  
 G A G A A A G T C A T A T C A G A A G A T A T C G T A G A G G A A G C T A A G A  
 C C T T A G C T A C T T G T G C A T A T G A G G A T G G T A C A C A T T G G G G  
 T C A A G A G A T C G G T T A T T C A T A C G A T T G T C A T T T G G A G A G C  
 A C T T T T A C T G G T T A T C T A T T A C A C T G C A A A G G G T G G A C A T  
 C G A C T T A T T T G T A T C C A G A C A G G C C A T C T T T C T T G G G T T G  
 T G C T C C A G T T G A T A T G C A A G G T T T C T C C T C A C A G C T C A T A

20

30

40

50

A A A T G G G T T G C T G C A C T T A C A C A A G C T G G T T T A T C A C A T C  
 T C A A T C C C A T C A C T T A T G G C T T G A G T A G C A G G A T G A A A A C  
 T C T C C A A T G C A T G T G C T A T G C C T A T T T G A T A T A T T T C A C T  
 C T T T A T T C T T G G G G A A T G G T T C T A T A T G C T A G T A T T C C T T  
 C T A T T G G T C T T T T G T T T G G C T T C C A A G T C T A T C C G G A G G T  
 A C A T G A T C C A T G G T T T G C A G T G T A T G T G A T T G C T T T C A T A  
 T C G A C A A T T T T G G A G A A T A T G T C G G A G T C A A T T C C A G A A G  
 G G G G A T C A T T T A A A T C G T G G T G G A T G G A A T A C A G G G C A C T  
 G A T G A T G A T G G G A G T T A G T G C A A T A T G G T T A G G A G G A T T G  
 A A A G C T A T A T T A G A C A A G A T A A T C G G A A C A G A A G G A G A G A  
 A A T T G T A T T T G T C A G A C A A G G C A A T T G A C A A G G A A A A G C T  
 C A A G A A A T A C G A G A A G G G G A A A T T T G A T T T C C A A G G A A T A  
 G G G A T A C T T G C T G T A C C A T T G A T A G C A T T T T C C C T G T T G A  
 A C C T G G T A G G C T T C A T T G T T G G A G C T A A T C A T G T C T T T A T  
 T A C T A T G A A C T A C G C A G G T G T G C T T G G C C A A C T C C T C G T A  
 T C A T C C T T C T T C G T C T T T G T C G T G G T C A C T G T T G T C A T T G  
 A T G T C G T T T C T T T C T T A A A G G T T T C T T A A [ 配列番号 36 ]。

10

【 0 8 6 0 】

セルロース合成酵素様 (solanum chacoense)

【 0 8 6 1 】

20

M K K T M E L N K S T V P Q P I T T I Y R L H M F V H S I I M A A L I Y Y R V  
 S N L F K F E N I L S L Q A L A W V L I T F G E F S F I L K W F F G Q G T R W R  
 P V E R D V F P E N I T C K D S D L P P I D V M V F T A N P K K E P I V D V M N  
 T V I S A M A L D Y P T D K L A V Y L A D D G G C P L S L Y A M E E A C V F A K  
 L W L P F C R K Y G I K T R C P K A F F S P L G D D D R V L K N D D F D A E M K  
 E I K L K Y E E F Q Q N V E R A G E S G K I N G N V V P D R A S F I K V I N D R  
 K T E S E K S A D D L T K M P L L V Y V S R E R R F N R L H H F K G G S A N A L  
 L R V S G I M S N A P Y I L V L D C D F F C H D P I S A R K A M C F H L D P K L  
 S S D L A Y V Q F P Q V F Y N V S K S D I Y D V K I R Q A Y K T I W H G M D G I  
 Q G P V L S G T G Y F L K R K A L Y T S P G V K E E Y L S S P E K H F G R S K K  
 F L A S L E E K N G Y V K A E K V I S E D I V E E A K T L A T C A Y E D G T H W  
 G Q E I G Y S Y D C H L E S T F T G Y L L H C K G W T S T Y L Y P D R P S F L G  
 C A P V D M Q G F S S Q L I K W V A A L T Q A G L S H L N P I T Y G L S S R M K  
 T L Q C M C Y A Y L I Y F T L Y S W G M V L Y A S I P S I G L L F G F Q V Y P E  
 V H D P W F A V Y V I A F I S T I L E N M S E S I P E G G S F K S W W M E Y R A  
 L M M M G V S A I W L G G L K A I L D K I I G T E G E K L Y L S D K A I D K E K  
 L K K Y E K G K F D F Q G I G I L A V P L I A F S L L N L V G F I V G A N H V F  
 I T M N Y A G V L G Q L L V S S F F V F V V T V V I D V V S F L K V S [ 配列番号 37 ]。

30

【 0 8 6 2 】

セルロース合成酵素様 (ナス)

40

【 0 8 6 3 】

A T G A A A A A A C A A A T G G A G C T C A A C A G A A G T G T T G T A C C G  
 C A A C C T A T C A C C A C C A T T T A C C G T C T C C A C A T G T T T A T C C  
 A T G C C C T A A T C A T G C T A G C A C T A A T A T A C T A C C G T G T C T C  
 T A A T T T G G C C A A A T T C G A A A A C A T C C T C A G T T T A C A A G C A  
 C T T G C T T G G G C T C T T A T C A C G T T A G G T G A A C T T T G T T T C A  
 T A G T C A A G T G G T T C T T C G G A C A A G G G A C T C G T T G G C G T C C  
 T G T T G A T A G G G A T G T C T T C C C T G A A A A C A T C A C T T G T C C A  
 G A T T C C G A G C T A C C C C C A T T G A T G T C A T G G T T T T C A C T G

50

C A A A T C C T A A G A A A G A G C C A A T T G T G G A T G T C A T G A A C A C  
 T G T C A T A T C C G C A A T G G C T C T T G A T T A C C C G A C C G A C A A A  
 T T G G C C G T T T A T T T G T C T G A T G A T G G A G G A T G C C C C T T G A  
 C G T T G T A C G C A A T G G A G G A A G C T T G T T C C T T T G C C A A G T T  
 G T G G C T A C C T T T T T G T A G G A A G T A T G G A A T C A A A A C A A G G  
 T G C C C T A A G G C G T T T T T T T C T C C A T T A G G A G A A G A T G A C C  
 G T G T A T T G A A G A G T G A T G A C T T T G T T T C T G A A A T G A A A G A  
 A A T G A A G T C A A A A T A T G A A G A G T T C C A G C A G A A C G T G G A C  
 C G T G C T G G T G A A T C C G G A A A A A T C A A A G G T G A C G T A G T G C  
 C T G A T A G A C C C G C G T T T C T T A A G G T A C T A A A T G A C A G G A A 10  
 G A C G G A G A A C G A G A A G A G T G C A G A C G A T T T A A C T A A A A T G  
 C C T T T G C T A G T A T A C G T A T C C C G T G A A A G A A G A A C T C A C C  
 G T C G C C A T C A C T T C A A G G G T G G A T C T G C A A A T G C T C T T C T  
 T C G A G T T T C T G G G A T A A T C A G T A A T G C C C C C T A T A T A C T G  
 G T T T T A G A T T G T G A T T T C T T C T G T C A T G A T C C A A T A T C A G  
 C T C G G A A G G C A A T G T G T T T C C A T C T T G A T C C A A A A C T A T C  
 A C C T G A C T T A G C T T A C G T G C A G T T C C C T C A A G T G T T T T A C  
 A A T G T T A G C A A G T C C G A T A T T T A C G A C G T C A A A A T T A G A C  
 A G G C T T A C A A G A C A A T A T G G C A C G G G A T G G A T G G T A T C C A  
 A G G C C C A G T G T T A T C G G G A A C T G G T T A T T T T T T A A A A A A G 20  
 A A G G C G T T G T A C A C G A G T C C A G G T C T A A A A G A T G A G T A T C  
 T T A G T T C A C C G G A A A A G C A T T T C G G A A C G A G T A G A A A G T T  
 C A T T G C T T C A C T A G A G G A G A A T A A T T A T G T T A A G C A A G A G  
 A A A G T C A T A T C A G A A G A T A T C A T A G A G G A A G C T A A G A G A C  
 T G G C T A C T T G T G C A T A C G A G G A T G G C A C A C A T T G G G G T C A  
 A G A G G C A A A C A G G C C A T C T T T C T T G G G T T G T G C C C C A G T T  
 G A T A T G C A A G G T T T C T C C T C A C A G C T C A T A A A A T G G G T T G  
 C T G C A C T C A C A C A A G C A G G T C T A T C A C A T C T C A A T C C C A T  
 C A C T T A C G G C T T C A A G A G C A G A A T G A G A A C T C T C C A A G T C  
 T T G T G T T A T G C C T A T T T G A T G T A T T T C T C T C T T T A T T C T T 30  
 G G G G A A T G G T T C T A C A T G C T A G T G T T C C T T C T A T T G G C C T  
 T C T C T C T G G C A T T A A A A T C T A C C C G G A G G T G T A T G A T C C A  
 T G G T T T G T T G T G T A T G T G A T T G C T T T C A T A T C A A C A A T T T  
 T G G A G A A T A T G T C G G A A T C A A T T C C G G A A G G G G G A T C G G T  
 T A A A A C G T G G T G G A T G G A A T A C A G G G C A C T G A T G A T G A T G  
 G G A G T T A G T G C A A T A T G G C T A G G A G G A G T G A A A G C C A T A G  
 T A G A C A A G A T C A T C G G A A C G C A A G G A G A G A A A T T G T A T T T  
 G T C G G A C A A A G C A A T T G A C A A G G A A A A G C T C A A G A A A T A C  
 G A G A A G G G G A A A T T T G A T T T C C A A G G A A T A G G A A T A C T T G  
 C T G T A C C A T T G A T A A C A T T T T C T G T G T T G A A C C T G G T A G G 40  
 C T T C T T G G T T G G A A T T A A T C A A G T G T T G A T A A C G A T G A A G  
 T T C G C A G G C G T G C T G G G C C A A C T C C T C G T A T C A T C C T T C T  
 T C G T C T T T G T C G T C G T T A C T G T T G T C A T T G A T G T C G T A T C  
 T T T C T T A A A G G A T T C T T A A [ 配列番号 38 ]。

【 0 8 6 4 】  
 セルロース合成酵素様 (ナス)

【 0 8 6 5 】  
 M K K Q M E L N R S V V P Q P I T T I Y R L H M F I H A L I M L A L I Y Y R V  
 S N L A K F E N I L S L Q A L A W A L I T L G E L C F I V K W F F G Q G T R W R  
 P V D R D V F P E N I T C P D S E L P P I D V M V F T A N P K K E P I V D V M N 50

TVISAMALDYPTDKLAVYLSDDGGCPLTLYAMEEACSFALK  
 LWLPFCRKYGIKTRCPKAFFSPLGEDDRVLKSDDFVSEMK  
 EMKSKYEEFQQNVDRAGESGKIKGDVVPDRPAFLKVLNDR  
 KTEENEKSADDLTKMPLLVYVSRERRTHRRHHFKGGSANAL  
 LRVSGIISNAPYILVLDCCDFCHDPI SARKAMCFHLDPKL  
 SPDLAYVQFPQVFYNVSKSDIYDVKIRQAYKTIWHGMDGI  
 QGPVLSGTGYFLKKKALYTS PGLKDEYLS SPEKHFGT SRK  
 FIASLEENNYVKQEKVISEDIEEAKRLATCAYEDGTHWG  
 QEANRPSFLGCAPVDMQGFSSQLIKWVAAL TQAGLSHLNP  
 ITYGFKSRMRTLQVLCYAYLMYFSLYSWGMVLHASVPSIG  
 LLSGIKIYPEVYDPWFVVYVIAFISTILENMSSESIPEGGS  
 VKTWWMEYRALMMMGSVAIWLGGVKAIVDKIIGTQGEKLY  
 LSDKAIDKEKLLKYEKGFDFQGGIGILAVPLITFSVLNLV  
 GFLVGINQV LITMKFAGV LGLLVSSFFV FVVVTVVIVIDVV  
 SFLKDS [ 配列番号 39 ]。

10

【 0866 】

セルロース合成酵素様 ( トウガラシ ( capsicum annuum ) )

【 0867 】

ATGGAGCTCAACAGATGTACGGTGCAGCAACCTACCACT  
 GCCATATAACCGACTACACATGTTTCTCCACTCTCTAATCA  
 TGCTTGCAATTAGTATACTATCGTTTGTCTAATCTGTTT TA  
 CTTTCGAAAACGTCCTCACTTTACAAGCATT TGCATGGGGG  
 CTTATCACCTTAGGTGAAATTTGTTTCA TTGTCAAGTGGT  
 TCTTTGGTCAAGGGACTCGTTGGCGCCCGTTGT CAGGGA  
 AGTGTTCCTGGACAATA TACTTGCCAAGATTCCGAGCTG  
 CCCGCACTAGATGTGATGGTTTTC ACTGCCAATCCCAAGA  
 AAGAGCCAATTGTGGATGTCATGAACACTGTGATATCCGC  
 AATGGCTCTTGATTACCCGACGGATAAATTGGCTGTGTAT  
 CTGGCTGATGATGGAGGATGCCCTTGACGTTGTACGCCA  
 TGGAGGAGGCCTGTTCTTTTGCCAAGTTGTGGCTACCTTT  
 CTGTAGGAAGTATGGAATCAAAACAAGGTGCCCCAAAGCG  
 TTTTTTTCTCCATTAGGAGAAAGATGATCGTATCCTTAAGA  
 ACGATGACTTTGTAGCTGAAATGAAAGAAATTA AATTA  
 ATATGAGGAGTTCCAGCAGAAATGTA AACCTTGCTGGTGAA  
 TCCGGAAA AATCAAAAGGTGACGTAGTG CCTGATAGAGCCT  
 CGTTTATTAAGGTAATAAATGACAGGAAAATGGAGAACAA  
 GAAGAGTGCCGACGATATAACGAAAATGCCCTTTGCTAGTA  
 TACGTATCCCGTGAAAGAAAGATTTAACAGT CGTCATCACT  
 TCAAGGGTGGATCTGCAAAATGCTCTTCTTCGAGTTT CAGG  
 GATAATGAGTAATGCCCCCTATTTACTGGTCTTAGATTGT  
 GATTTCTTCTGT CATGATCCAACATCAGCTCGGAAGGCAA  
 TGTGTTTCCATCTTGATCCA AA ACTATCACCTTCTT AGC  
 TTATGTGCAGTTCCCTCAAGTGTTT TACAATGT CAGCAAG  
 TCCGATATAACGATGTCAAAATTAGACAGGCTTACAAGA  
 CAATATGGCACGGAATGGATGGTATCCAAGGCC CAGTGTT  
 ATCGGGAACTGGGTATTTTCTGAAGAGGAAAAGCGTTATAC  
 ACGAGTCCAGGTCTAAAGGATGAGTATCTTATTTCA CCGG  
 AAAAGCATTTTCGGATCAAGTAGAAAAGTTCA TTGCTTCTCT  
 AGAGGAGAACAAATGGTTATGTTAAGCAAGAGAAA ACTCATA  
 ACAGAAAGATATTATAGAGGAAGCGAAGACCTTGTC TACTT

20

30

40

50

GTGCATACGAGGATGGTACACGATGGGGCGAAGAGATCGG  
TTATACCTACAATTGCCATTTGGAGAGCACTTTTACCGGC  
TATCTTTTGC ACTGCAAAGGGTGGACATCAACATATTTGT  
ATCCAGAAAGGCCATCTTTCTTGGGTGTGCCCCAGTTGA  
TATGCAAGGATTCCTCCTCACAACTCACAA AATGGGTTGCT  
GCACTCACACAAGCTGGTCTATCACATCTCAATCCCATCA  
CTTACGGCATGAAGAGCAGGATTAAGACTATCCAATGCTT  
GTGCTATGCCTATTTGATGTATTTCTCTCTCTATTTCTTGG  
GGAATGGTTCTGCATGCTAGTGTTCCTTCTATTAGCCTTT  
TGCTTGGCATTC AAGTCTACCCCGAGGTCTATGATCCATG  
GTTTGCAGTGTATGTGCTTGGCTTTCATATCGACAATTTTG  
GAGAACATGTCAGAGTCAATTC CAGAAAGGC GGTT CAGTTA  
AAACTTTGGTGGATGGAATACAGGGC ACTGATGATGATGGG  
AGTTAGTGC AATATGGTTAGGAGGAGTGAAAGCTATAGTA  
GAAAAGATCATCGGA ACTCAAGGAGAGAAATTAATTTGT  
CGGACAAAGCAATTGACAAGGAAAAGCTCAAGAAATATGA  
GAAGGGGAAATTTGATTTCCAAGGGATAGGGATACTTGCT  
GTTCCATTGATAACATTCCTCAGCGTTGAATTTGGTAGGCT  
TCATGGTTGGAGCTAATCAAGTGATTCCTTACTATGAAGTT  
CGAAGCTTTGCTAGGCCCAACTCCTTGTGT CATCCTTCTTC  
GTCTTTGTGGTGGTCACCGTTGT CATAGATGTCCTATCTT  
TCTTAAAAGACTCTTAA [ 配列番号 40 ]。

10  
20  
30  
40

【 0 8 6 8 】

セルロース合成酵素様 ( トウガラシ ( capsicum annum ) )

【 0 8 6 9 】

MELNRCTVQQPTTAIYRLHMF LHS LIMLALVYYRLSNLF  
YFENVLTLQAF AWGLITLGEICFIVKWF FFGQGTRWRPVVR  
EVFLDNITCQDSELPALDVMVFTANPKKEPIVDVMNTVIS  
AMALDYPTDKLAVYLADDGGCPLTLYAMEEACSF AKLWLP  
FCRKYGIKTRCPKAF FSP LGEDDRILKNDDFVAEMKEIKL  
KYE EFQQNVNLAGESGKIKGDVVPDRASF IKVINDRKMEN  
KKSADDITKMP L LVYVSRERRFNSRHHFKGGSANALLRVS  
GIMSNA PYLLVLD CDF FCHDPTSARKAMCFHLDPKLSPSL  
AYVQFPQVFYNVSKSDIYDVKIRQAYKTIWHGMDGIQGPV  
LSGTGYFLKRKALYTSPGLKDEYLISPEKHFGSSRKFIAS  
LEENNGYVKQEKLITEDIIEEA KTLSTCAYEDGTRWGE EI  
GYTYNCHLESTFTGYL LHCKGWTSTYLYPERPSFLGCAPV  
DMQGFSSQLTKWVAAL TQAGLSHLNPITYGMKSRIKTIQC  
LCYAYLMYFSLYSWGMVLHASVPSISLLLG IQVYPEVYDP  
WFAVYVLA FISTILENMSESIPEGGSVKTWMEYRALMMM  
GVSAIWLGGVKAIVEKIIIGTQGEKLYLSDKAIDKEK LKKY  
EKGKFD FQGI GILAVPLITFSALNLVGF MVGANQVILTMK  
FEALLGQLLVSSFFV FVVTVVIDVLSFLKDS [ 配列番号 41 ]

30  
40  
50

。

【 0 8 7 0 】

それぞれの植物においてGAME 15をサイレンシングするため、以下の配列を作製した。

【 0 8 7 1 】

トマトにおいてGAME 15サイレンシングに用いた領域

【 0 8 7 2 】

50

GGCTCTTGATTATCCACCGATAAATTGGCTGTGTATCT  
CGCTGATGATGGAGGATGTC CATTGTCTGTACGCCATG  
GAACAAGCGTGT TTTGTTTGC AAAGCTATGGTTACCTTTCT  
GTAGAAACTATGGAATTA AAACGAGATGCCCAA AAGCATT  
TTTTTCTCCGTTAGGAGATGATGACCGTGTTCCTTAAGAAT  
GATGATTTTGTCTGCTGAAATGAAAGAAATTA AATTGAAAT  
ATGAAGAGTTCCAGCAGAAAGGTGGAACATGC [ 配列番号 4 2 ]。

【 0 8 7 3 】

ジャガイモにおいてGAME 15サイレンシングに用いた領域

【 0 8 7 4 】

10

GGCTCTTGATTATCCTACGGATAAATTGGCTGTGTATCT  
GGCTGATGATGGAGGATGTCCTTTGTCAATTGTACGCCATG  
GAAGAAGCATGTGTGT TTTGC AAAGCTGTGGCTACCTTTCT  
GTAGGAAGTATGGAATTA AAACCTAGATGCCCTAAAGCGTT  
TTTTTCTCCCTTTAGGAGATGATGAACGTGTTCCTTAAGAAT  
GATGATTTTGTCTGCTGAAATGAAAGAAATTA AATTGAAAT  
ATGAAGAGTTCCAGCAGAAATGTGGAACGTGCTGGTG [ 配列番  
号 4 3 ]。

【 0 8 7 5 】

ナスにおいてGAME 15サイレンシングに用いた領域

20

【 0 8 7 6 】

GGCTCTTGATTACCCGACCGACAAATTGGCCGTTTATTT  
GTCTGATGATGGAGGATGCCCTTGACGTTGTACGCAATG  
GAGGAAGCTTGTTCCTTTGCCAAGTTGTGGCTACCTTTTT  
GTAGGAAGTATGGAATCA AAACAAGGTGCCCTAAGGCGTT  
TTTTTCTCCATTAGGAGAAAGATGACCGTGTATTGAAGAGT  
GATGACTTTGTTTCTGAAATGAAAGAAATGAAGTCAAAAT  
ATGAAGAGTTCCAGCAGAAACGTGGACCGTGTCTGGTGAATC  
CGGAAAAATCA AAGGTGACGTAGTGCCTGATAGACCCGCG  
TTTCTTAAGGTACTAAATGACAGGAAGACGGAGAAACGAGA  
AGAGTGCAGACGATTTAACTAA AATGCCTTTGCTAGTATA  
CGTATCCC GTGAAAGAAAGAACTCACCGTCCCATCACTTC  
AAGGGTGG [ 配列番号 4 4 ]。

30

【 0 8 7 7 】

トマト及びジャガイモにおけるGAME 15 遺伝子に対するRNA i系を作製した。G  
AME 15 - RNA i 遺伝子導入トマト植物では、葉において - トマチン及び下流SG  
Aの著しい低下が見られ、GAME 15をサイレンシングした緑色果実において - トマ  
チンは検出されなかった。さらに、トマト果実の発育段階（例えば、催色期（breaker）  
及び赤色果実）中では、エスクレオシドまたは他のSGAは検出されなかった。また、G  
AME 15 - RNA i トマト植物の葉及び緑色果実において、SGAの前駆体であるコレ  
ステロールの15 ~ 20 倍の増加が認められた。ジャガイモでは、GAME 15をサイレ  
ンシングしたところ、 - カコニン及び - ソラニンが大きく低下し、これらの系統のコ  
レステロールの貯留が増加した。

40

【 0 8 7 8 】

実施例 7：GAME 15 - RNA i 遺伝子導入トマト、ジャガイモ及びナス植物の作製

【 0 8 7 9 】

選択した断片（サイレンシング配列、配列番号 4 2（トマト）、配列番号 4 3（ジャガ  
イモ）、及び配列番号 4 4（ナス））をpENTR/D-TOPO（Invitrogen）に（NotI及びAscIによって）導入し、Gateway LR Clonase  
II 酵素混合物（Invitrogen）を用い、この断片をpK7GWIWG2（

50

II) バイナリーベクターにさらにサブクローニングすることにより、GAME15-RNAiコンストラクトを作製した。既報(「Itkin et al. 2011. The Plant Cell 23:4507-25」; 「Sonawane et al. 2018. PNAS 115(23): E5418-E5428」)のようにして、ベクターをトマト、ジャガイモ及びナスに形質転換した。正のGAME25がダウンレギュレーションされた系を、LC-MS分析にさらに用いた。

【0880】

実施例8: GAME15をサイレンシングしたトマト植物はSGAプロファイルの激しい低下を示した

【0881】

SGA代謝におけるGAME15の正確な役割を解明するため、上記のトマトサイレンシング配列(配列番号42)を用い、GAME15-RNAi(GAME15i)遺伝子導入トマト系(#21、#22及び#23)を作製した。

10

【0882】

GAME15-RNAiの葉では、野生型トマトの葉に比べ、-トマチンの著しい低下が見られた(図15A)。さらに、引き続き、GAME15i果実のSGAプロファイルを、野生型のものと、異なる発育及び成熟の段階で比較した。トマトにおいて緑色から赤色果実に変化する間に、-トマチンはエスクレオシド及びリコペロシドに変換され、一方、デヒドロトマチンはデヒドロエスクレオシド及びデヒドロリコペロシドに変換される(図14A)。

20

【0883】

GAME15i緑色及び赤色果実は、SGA(例えば、-トマチンまたはエスクレオシドA)の痕跡を示さず、GAME15iサイレンシングにより、トマトの果実において、SGAが完全に喪失したことが示唆された(図15B及び15C)。

【0884】

実施例9: GAME15発現を変化させるとジャガイモにおけるSGAに大きな影響を及ぼす

【0885】

トマトと同様に、上記のジャガイモサイレンシング配列(配列番号43)を用い、ジャガイモ(#1、#2、及び#3)においてもGAME15iをサイレンシングし、ジャガイモSGA代謝に対する影響を解明した。

30

【0886】

ジャガイモにおいてGAME15をサイレンシングしたところ、ジャガイモの葉組織における主要SGAである-カコニン(影付きバー)及び-ソラニン(白抜きバー)が、野生型ジャガイモの葉組織に比べ、大幅に低下した(図16)。

【0887】

実施例10: GAME15をサイレンシングしたトマトの葉におけるコレステロールの高蓄積

【0888】

コレステロールは、SGAの生合成において重要な前駆体としての役割を果たす(「Sonawane et al., 2016, Nat. Plants 3: 16205」)。GAME15iをサイレンシングしたトマト植物においてSGAが著しく低下し、その後完全な喪失が認められたことから、これらの植物におけるコレステロールレベルを調べた。GAME15iをサイレンシングしたトマト植物の葉において、野生型トマト植物の葉に比べ、コレステロール(SGA前駆体)の約15~20倍の増加が認められた(図17)。

40

【0889】

実施例11: GAME15発現を変化させ、ナスにおけるSGAへの影響を観察する

【0890】

ジャガイモと同様に、上記のナスサイレンシング配列(配列番号44)を用い、ナスにおいてもGAME15iをサイレンシングし、ナスSGA代謝に対する影響を解明した。

【0891】

50

ナスにおいてGAME 15をサイレンシングする影響を、 - ソラニン及び/または - ソラマルギンのレベルの低下について、野生型ナスと比較した(図14C)。

【0892】

実施例12: トマト、ジャガイモ、及びナスにおけるGAME 15の過剰発現

【0893】

さらに、トマト、ジャガイモ、及び/若しくはナス植物の遺伝子を組み換えるか、または遺伝子を編集し、GAME 15を過剰発現させた。

【0894】

トマト植物において - トマチン並びにエスクレオシド及び/またはリコペロシド(図14A)の産生を増加させるため、トマト植物の遺伝子を組み換えるか、または遺伝子を編集し、GAME 15を過剰発現させた。

10

【0895】

ジャガイモ植物において - ソラニン及び/または - カコニン(図14B)の産生を増加させるため、ジャガイモ植物の遺伝子を組み換えるか、または遺伝子を編集し、GAME 15を過剰発現させた。

【0896】

ナス植物において - ソラソニン及び/または - ソラマルギン(図14B)の産生を増加させるため、ナス植物の遺伝子を組み換えるか、または遺伝子を編集し、GAME 15を過剰発現させた。

【0897】

20

実施例13: SGAのレベル及び組成が変更された植物及び作物

【0898】

前述の内容に基づき、古典的な育種または遺伝子工学(例えば、遺伝子組換えまたは遺伝子導入植物、及び遺伝子編集植物など)によって、様々な昆虫及び他の病原体に対する障壁を植物に付与し、かつ/または抗栄養化合物(例えば、ジャガイモからカコニン及び/またはソラニン)を除去し、SGAのレベル及び組成を変更した、ナス科植物(例えば、トマト、ジャガイモ、ナス、及び/またはペッパー植物)及び/または作物が作製される。

【0899】

さらに、高コレステロールまたは高フィトステロールトマト系は、高価値のステロイド系化合物(例えば、プロビタミンD及び/またはジオスゲニン)を合成生物学ツールなどによって工学的に作り出すのに用いられる。

30

【0900】

また、高フィトステロール(例えば、フィトコレステロール)系は、化粧品に用いられる成分を産生するのに用いられる。

【0901】

他の例では、ナス科植物(例えば、トマト、ジャガイモ、ナス、及び/またはペッパー植物)及び/または作物は、SGAのレベルを高め、かつ/またはフィトステロールのレベルを下げて作製される。

【0902】

40

実施例14: 実施例14~25の材料及び方法

【0903】

サポニンの分析の液体クロマトグラフィー - 質量分析(LC-MS)及びタンデム質量分析(MS/MS)パラメータ

【0904】

試料から得た4つの生物学的複製物( $n = 4$ )を、代謝分析に用いた。簡潔に説明すると、凍結粉末化した植物組織100mgを、内部標準物質(ポナステロンA、 $C = 1.5 \mu\text{g/mL}$ )を混合した80%メタノール300 $\mu\text{L}$ で抽出し、短時間ボルテックスした後、室温で20分間超音波処理した。抽出物を14000 $\times g$ で10分間遠心分離し、0.22 $\mu\text{m}$ のフィルターで濾過した。試料を、SYNAPT-G2 qTOF検出器(タ

50

ンデム四重極 / 飛行時間型質量分析計、W a t e r s ) に連結したU P L C ( W a t e r s A c q u i t y ) で構成された高分解能U P L C / q T O Fシステムを用いて分析した。代謝物の分離は、1 0 0 m m × 内径 2 . 1 m m 、 1 . 7 μ m のU P L C B E H C 1 8 カラム ( W a t e r s A c q u i t y ) で行った。移動相の組成は 0 . 1 % ギ酸を含むアセトニトリル : 水 ( 5 : 9 5 、 v / v ; 移動相 A ) 及び 0 . 1 % ギ酸を含むアセトニトリル ( 移動相 B ) であった。流量は 0 . 3 m L / 分であり、カラム温度は 3 5 に維持した。

**【 0 9 0 5 】**

以下の直線グラジエントを用い、ホウレンソウ、ビートルート ( *Beta vulgaris* ) 中のトリテルペノイド系サポニン、酵母細胞中及び S O A P 1 0 を用いて *i n v i t r o* で産生されたトリテルペノイドの分析を行った。移動相 A を 5 分間で 1 0 0 % から 8 5 % とし、移動相 A を 2 分間で 8 5 % から 7 5 % とし、移動相 A を 7 5 % で 3 分間維持し、グラジエントを継続し、移動相 A を 1 0 分間で 6 5 % とし、移動相 A を 2 分間で 6 5 % から 4 0 % とし、移動相 A を 4 0 % から 1 分間で移動相 B を 1 0 0 % とした後、移動相 B を 1 0 0 % で 3 . 5 分間し、その後 0 . 5 分以内に初期状態 ( 移動相 A 1 0 0 % ) に戻し、移動相 A 1 0 0 % で 1 分間コンディショニングした。

10

**【 0 9 0 6 】**

キノア ( *Chenopodium quinoa* ) 及びアルファルファ ( *Medicago sativa* ) のトリテルペノイド系サポニンを分析するため、以下の 4 0 分間のグラジエントを用いた。移動相 A を 2 2 分間で 1 0 0 % から 7 2 % とし、移動相 A を 1 4 分間で 7 2 % から 0 % とし、移動相 B を 1 0 0 % で 2 分間維持した後、 0 . 5 分以内に初期状態 ( 移動相 A 1 0 0 % ) に戻し、移動相 A 1 0 0 % で 1 . 5 分間コンディショニングした。

20

**【 0 9 0 7 】**

以下の設定を用いた。キャピラリー 2 k V ; コーン 2 7 V ; ソース温度を 1 4 0 、脱溶媒和を 4 5 0 、脱溶媒ガス流量を 8 0 0 L / 時間に設定した。アルゴンを衝突ガスとして用いた。エレクトロスプレーイオン化 ( E S I ) を、陰イオン化モード、m / z 範囲 5 0 ~ 1 6 0 0 D a で用いた。ギ酸ナトリウムを用いて質量分析 ( M S ) システムを較正し、L e u - エンケファリンをロックマスとして用いた。M a s s L y n x ソフトウェアバージョン 4 . 1 ( W a t e r s ) を用い、装置を制御し、正確な質量及び元素組成を算出した。また、1 5 種の標準代謝物の混合物を 1 0 個の試料のそれぞれの後に注入し、品質対照として用いた。データの取得は、試料中の全ての検出可能な成分由来の正確な質量前駆体及びフラグメントイオンの情報を記録するエネルギー勾配による M S <sup>E</sup> モードで行った。M S <sup>E</sup> モードは、2 つの機能をすばやく切り替える。1 つ目は低エネルギーの正確な質量前駆体イオンスペクトルを取得することであり、2 つ目は高エネルギーの正確な質量フラグメントイオンスペクトルを取得することである。衝突エネルギーは、低エネルギーの機能については 4 e V 、高エネルギーの機能については 1 5 ~ 5 0 e V の勾配に設定した。

30

**【 0 9 0 8 】**

ベンサミアナタバコ ( *N. benthamiana* ) 中のウラルカンゾウ ( *Glycyrrhiza uralensis* ) 酵素の生成物の分離には、別のグラジエントを用いた。移動相 A を 7 5 % 、 5 分間で開始した後、移動相 A を 1 5 分間で 7 5 % から 5 0 % とし、移動相 A を 2 分間で 5 0 % から 3 0 % とし、移動相 A を 3 0 % から移動相 B を 1 0 0 % とし、移動相 B を 1 0 0 % で 3 . 5 分間維持した後、 0 . 5 分以内に初期状態 ( 移動相 A 7 5 % ) に戻し、移動相 A 7 5 % で 1 分間コンディショニングした。流量は 0 . 3 m L / 分であり、カラム温度は 3 5 に維持した。試料は、I m p a c t H D U H R - Q q T O F ( B r u k e r ) に連結したU P L C ( 1 2 9 0 I n f i n i t y I I 、 A g i l e n t ) で構成された高分解能U P L C / q T O Fシステムを用いて分析した。代謝物の分離は、1 0 0 m m × 内径 2 . 1 m m 、 1 . 7 μ m のU P L C B E H C 1 8 カラム ( W a t e r s A c q u i t y ) で行った。移動相の組成は 0 . 1 % ギ酸を含むアセトニトリル : 水 ( 5 : 9 5 、 v / v ; 移動相 A ) 及び 0 . 1 % ギ酸を含むアセトニトリル ( 移動相 B ) であった。E S I を

40

50

、陰イオン化モード、 $m/z$  範囲 50 ~ 1600 Da において、以下のパラメータで用いた。乾燥ガス 200 °C、8 L / 分；ネブライザ：2 Bar；キャピラリー：4.2 kV。

#### 【0909】

代謝物は、標準化合物の保持時間及び質量フラグメントを比較することによって同定した。対応する標準物質がない場合は、保持時間、元素組成及びフラグメンテーションを文献に記載のもの（表14）と比較することによって推定して同定した。ハウレンソウ中のトリテルペノイド系サポニンの相対定量を、Target Lynx (Waters) プログラムを用いて行った。各試料中のサポニンのピーク面積は、試料の取り扱いに起因するばらつきを低減するため、内標準物質のピーク面積に対して標準化した。

10

#### 【0910】

ハウレンソウトリテルペノイド系アグリコンのガスクロマトグラフ - 質量 (GC - MS) 分析

#### 【0911】

粉末化した凍結組織 (200 mg) を、80% MeOH 600  $\mu$ L で3回抽出した。採取した分画を高速真空 (25、78 mbar、終夜 (O/N)) で蒸発させた。乾燥残渣を、2 M HCl 50% MeOH 溶液 500  $\mu$ L に溶解し、65 °C で5時間加熱した。加水分解した試料を、高速真空 (50、78 mbar、1.5時間) で蒸発乾固した。乾燥残渣をトルエン : MeOH (3 : 2 ; v/v) 200  $\mu$ L に溶解し、トリメチルシリルジアゾメタン (TMSCHN<sub>2</sub>) 75  $\mu$ L を加えた。メチル化混合物を室温で40分間インキュベートして蒸発乾固し、MSTFA 80  $\mu$ L に再懸濁し、室温で20分間、65 °C で10分間インキュベートした。試料をガラスインサートに移し、1  $\mu$ L を、前述 (31) したようにして、GC - MS に注入した。GC - MS システムは、COMBI PAL オートサンプラー (CTC Analytics) と、プログラム昇温気化 (PTV) 注入器を備えた微量 GC 超ガスクロマトグラフと、DSQ 四重極質量分析計 (Thermo Electron) とから構成されていた。GC は、30 m x 0.25 mm x 0.25  $\mu$ m の Zebron ZB - 5 ms MS カラム (Phenomenex) で行った。PTV スプリット法を、以下のように行った。試料は、一定温度スプリットレスモードで分析した。PTV 注入口温度は 280 °C に設定した。分析物は、以下のクロマトグラフィー条件を用いて分離した。ヘリウムをキャリアガスとして流量 1.2 mL / 分で用いた。170 °C で開始した温度グラジエントをこの温度で1.5分間維持し、280 °C まで 37 °C / 分で昇温した後、300 °C まで 1.5 °C / 分で昇温し、300 °C で 5.0分間維持した。溶離液を、イオン化電圧 70 eV の電子衝撃モードでフラグメント化した。再構築したイオンクロマトグラム及び質量スペクトルを、Xcalibur ソフトウェア (ThermoFinnigan) を用いて評価した。化合物は、保持時間指標及び質量スペクトルを、同じ装置で分析した正式な標準物質に対して生成されたもの、及び文献 (「E. Biazzi et al., CYP72A67 Catalyzes a Key Oxidative Step in Medicago truncatula Hemolytic Saponin Biosynthesis. Molecular Plant. 8, 1493-1506 (2015)」) で報告されたものと比較することによって同定した。

20

30

#### 【0912】

NMR 分析用のサポニンの精製

40

#### 【0913】

システムは、既報 (「B. Khakimov, L. H. Tseng, M. Godejohann, S. Bak, S. B. Engels, Screening for Triterpenoid Saponins in Plants Using Hyphenated Analytical Platforms. Molecules. 21, pii: E1614 (2016)」) のように、クォータナリポンプ、オートサンプラー、ダイオードアレイ検出器、Bruker / Spark Prospekt II LC - SPE システム (Spark) を備えた Agilent 1290 Infinity II UPLC システムと、Bruker NMR MS Interface (BNMI - HP) を介して連結された Impact HD UHR - QqTOF 質量分析計 (Bruker) とから構成された。MS スペクトルは、ネガティブモード

50

で、 $m/z$  50 ~ 1700 において取得した。較正は、各クロマトグラフィー測定を開始時及び終了時に注入される 10 mM トリフルオロ酢酸ナトリウム (NaTFA、Sigma-Aldrich) により、自動的に行なった。分離は、XBridge LC カラム (BEH C18, 粒径 5  $\mu$ m、及び 250 mm  $\times$  4.6 mm; Waters) で行った。クロマトグラフィー条件は以下の通りであった。流量 0.9 mL/min で、移動相 A (5% ACN + 0.1% FA) 74% 及び移動相 B (100% ACN + 0.1% FA) 26% の溶媒組成を、60 分で移動相 A 70% とする直線グラジエントを開始した後、62 分で移動相 B を 100% とする別の直線グラジエントを行い、移動相 B 100% を 3 分間維持した後、65.5 分で移動相 A を 74% とする直線グラジエントを行い、4.5 分間維持して平衡化する。サポニンは、予め設定した時間幅で SPE カートリッジに採取し、トラッピングは、強度の閾値によって開始した。ヨッソシド (Yossoside) I V ( $m/z$  1263.58、閾値 200000、時間幅 18 ~ 23 分)、ヨッソシド X I I ( $m/z$  1233.57、閾値 58000、時間幅 27 ~ 32 分)、ヨッソシド V a ( $m/z$  1305.59、閾値 200000、時間幅 32 ~ 40 分)、ヨッソシド X ( $m/z$  1437.64、閾値 150000、時間幅 40 ~ 48 分) 及びヨッソシド V ( $m/z$  1305.59、閾値 200000、時間幅 54 ~ 65 分)。このトラッピング法では、カートリッジへの分析物の保持を高めるため、溶離液が SPE カートリッジを通過する前に、補充流量 2.5 mL/分の水を溶離液に加えた。トラッピングには、GP 樹脂を充填した、10 mm  $\times$  2 mm の SPE カートリッジを用いた。各カートリッジには同じ化合物を 5 回ロードし、1つのサポニンのトラッピングに 5 ~ 10 個のカートリッジを用いた。NMR 測定の前に、SPE カートリッジを窒素流で乾燥させ、各カートリッジの分画を重水素化メタノール 150  $\mu$ L で溶出させた。試料は 96 ウェルプレートに溶出させた。同じ化合物を含む溶離液をプールし、窒素流下で乾燥し、凍結乾燥し、D<sub>2</sub>O 200  $\mu$ L に再懸濁し、再び凍結乾燥して微量の H<sub>2</sub>O を除去した。乾燥化合物を、3-プロピオン-2,2,3,3-d<sub>4</sub> 酸ナトリウム塩 (TMSP、<sup>1</sup>H 及び <sup>13</sup>C スペクトルの内部化学シフト基準として用いた) を 0.01% 添加した 90% MeOD-d<sub>4</sub> と 10% D<sub>2</sub>O との混液 60  $\mu$ L に溶解し、1.7 mm の NMR 試験管に移した。

【0914】

NMR 法

【0915】

NMR スペクトルは、シールドグラジエントコイルを備えた 1.7 mm の TXI-z CryoProbe 及び 5 mm の TCI-xyz CryoProbe を備えた Bruker AVANCE NEO-600 NMR 分光計で記録した。全てのスペクトルを 293 K で取得した。ヨッソシド V の NMR データは 1.7 mm の CryoProbe 及び 5 mm の CryoProbe の両方で記録し、メディカゲン酸 3-O-グルクロニド (MA-3-GlcA) のスペクトルは 5 mm の CryoProbe で取得した (5 mm のプローブ中に 1.7 mm の試験管を保持するための、特殊な NMR スピナータービンを用いた)。

【0916】

<sup>1</sup>H 及び <sup>13</sup>C の化学シフトの帰属は、別々の 1D 及び 2D NMR 手法に基づき、全てパルス磁場勾配選択を用いた。

【0917】

アグリコン単位の構造及び立体化学は、1D <sup>1</sup>H 及び <sup>13</sup>C、2D 等核 COSY 並びに 2D 異核 HSQC 及び HMB C スペクトルによって決定した。単糖単位の同定には、帰属を完結し、個々の糖単位内の <sup>1</sup>H シグナル間の結合性を詳細に描写するため、2D 等核 TOCSY スペクトルも記録する必要があった。HMB C スペクトルにおいて認められた、2 ~ 3 結合分離した <sup>1</sup>H と <sup>13</sup>C との間の相関関係を、ヨッソシド V については図 21 に、MA-3-GlcA については図 35 D に示す。

【0918】

<sup>1</sup>H 1D NMR スペクトルは、16 k のデータポイント及び 2.5 秒のリサイクル

10

20

30

40

50

遅延を用いて取得した。 $^1\text{H}$   $^13\text{C}$  スペクトル編集 DEPTQ NMR スペクトルは、65 k のデータポイント及び3秒のリサイクル遅延を用いて取得した。

【0919】

2D  $^1\text{H}$  -  $^1\text{H}$  COSY、TOCSY及びROESYスペクトルは、16384 - 8192 ( $t_2$ )  $\times$  400 - 512 ( $t_1$ ) のデータポイントを用いて取得した。2D TOCSYスペクトルは、等方性混合時間100 ~ 300 m秒を用いて取得した。本試験で用いたT-ROESY試験は、ROESY試験におけるTOCSYの移行を効果的に抑制するTOCSY-less ROESYであった。2D T-ROESYスペクトルは、スピロックパルス100 ~ 400 m秒を用いて記録した。

【0920】

2D  $^1\text{H}$  -  $^13\text{C}$  HSQC及びHMBCスペクトルは、4096 ( $t_2$ )  $\times$  400 - 512 ( $t_1$ ) のデータポイントを用いて記録した。多重度編集HSQCにより、正相関を生じさせるメチル基及びメチン基を、負のピークとして現れるメチレン基から区別することができる。遠隔カップリングを進展させるためのHMBC遅延を設定し、 $J_{\text{H},\text{C}} = 8\text{ Hz}$  の遠隔カップリングを観察した。

10

【0921】

化学シフトの帰属は、HMBC結合性(図21及び図35D)及びTOCSYの相関(データ提示なし)由来の情報の組み合わせに基づいて行った。

【0922】

ヨッソシドV及びMA-3-GlcAの $^1\text{H}$ 及び $^13\text{C}$ の化学シフト値は、表1に示す。予想通り、これら2つの化合物のアグリコン単位の化学シフト値はよく一致し、同様のHMBC相関パターンが認められる(図21及び図35D)。アキシアルまたはエクアトリアルへのプロトンの帰属は、観察された隣接するJカップリングに基づいて行い、大きな値( $> 10\text{ Hz}$ )は、アキシアルプロトンを示し、ROESYスペクトルで認められた相関によってさらに支持される。

20

【0923】

アグリコン領域のHMBCの相関は、他のトリテルペノイド系サポニン(「H. Schrodter et al., A triterpene saponin from *Herniaria glabra*. *Phytochemistry*. 34,1609-13 (1993)」; 「E. P. Mazzola et al., Utility of coupled-HSQC experiments in the intact structural elucidation of three complex saponins from *Blighia sapida*. *Carbohydr Res*. 346, 759-68 (2011)」)において認められたものとよく一致している。MA-3-GlcAにおけるメディカゲン酸とグルクロン酸との間の結合部位は、GlcAのアノマー水素とアグリコンのC3との間の強いHMBC結合性、並びにMA-H3/GlcAのアノマーC1との間に認められた相補的な相関(図18D中の赤矢印)によって支持された。同様に、糖単位間の結合部位は、1つの糖のアノマー水素/炭素間、並びにFuc-Rha-Glc間のグリカン間の矢印によって示されるような、直接結合したグリカンのプロトン/炭素間に認められた、相補的な3結合HMBC結合性(図21)に基づいて決定した。GlcAとアグリコンとの間に先に認められた相関がないこと、並びにGlcA内のJカップリングがないことで示されるように、ヨッソシドV(GlcAに結合したXyl)におけるキシロースの付加により、GlcAの立体配座空間は制限されていると考えられる。しかしながら、化学シフトの予測は、ヨッソシドVにおけるGlcAの帰属をさらに支持するものである。

30

40

【0924】

RNA-Seqライブラリの作成及びデータ解析

【0925】

RNA-Seqライブラリは、若干の変更を加え、既報(「S. Zhong et al., High-throughput illumina strand-specific RNA sequencing library preparation *Cold Spring Harb Protoc*. 8, 940-9 (2011)」)のように作成した。簡潔に説明すると、全RNA 5  $\mu\text{g}$ を、Dynabeads Oligo(dT)<sub>25</sub>(Invitrogen)を用いたポリ(A)RNAの取得に用い、94 で5分間断片化し、溶出させた。第1鎖cD

50

NAは、ランダムプライマー及びdNTPを用いた逆転写酵素SuperScript II (Invitrogen)を用いて合成し、一方、第2鎖cDNAは、dUTPを用いたDNAポリメラーゼI (Enzymatics)を用いて生成した。末端修復(Enzymatics)、Klenow 3'-5' (Enzymatics)によるdA-ターミナル及びアダプターライゲーション(Quick T4 DNA Ligase、NEB)の後、dUTP含有第2鎖をウラシルDNAグリコシラーゼ(Enzymatics)によって消化した。得られた第1鎖アダプターライゲーションcDNAを14サイクルのPCR濃縮(NEB Next High-Fidelity PCR Master Mix、NEB)に用いた。インデックス付きライブラリをプールし、the Crown Institute for Genomics, The Nancy and Stephen Grand Israel National Center for Personalized Medicine, Weizmann Institute of Scienceにおいて、HiSeq2500 (Illumina)の単一ラン上で配列決定した(ペアエンド、125bp)。FastQC (「S. W. Wingett, S. Andrews, FastQ Screen: A tool for multi-genome mapping and quality control. Version 2. F1000Res. 7, 1338 (2018)」)を用い、生データの品質を評価した。Tophat v2.0.13 (「C. Trapnell, L. Pachter, S. L. Salzberg, TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 25, 1105-11 (2009)」)を用い、生の読み取りを、ホウレンソウリファレンスゲノムのバージョン1.0.1 (The Beta vulgaris resource (<http://bvseq.boku.ac.at/Genome/Download/index.shtml>))に対してアライメントした。HTSeq (「S. Anders S, P. T. Pyl, W. Huber, HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics. 31, 166-9 (2015)」)を用いて定量を行った後、DESeq2 (「M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014)」)を用いて標準化及び差次的遺伝子発現解析を行った。

【0926】

共発現解析

【0927】

CoExpNetVizソフトウェア(「O. Tzfadia, T. Diels, S. De Meyer, K. Vandepoele, A. Aharoni, Y. Van de Peer, CoExpNetViz: Comparative Co-Expression Networks Construction and Visualization Tool. Front Plant Sci. 6, 1194 (2016)」)を用いて共発現解析を行った。簡潔に説明すると、低いパーセンタイル順位の相関閾値を5、高いパーセンタイル順位の相関閾値を95に設定した共発現解析において、ホウレンソウbAS (SOAP1)、CYP716A268 (SOAP2)及びCYP716A268v2 (SOAP2様)を「ベイト」として用いた。共発現した遺伝子を、各ベイトの相関係数( $r > 0.9$ ; PCC)に従い、さらにフィルタリングした(表14)。解析は、異なる組織及び生育段階(SO4WORLD - 長日に発育した4週齢の植物の成熟葉; SO8WOLSL - 短日に発育した8週齢の植物の成熟葉; SO4WYLLD - 長日に発育した4週齢の植物の若葉; SO8WOLSL - 短日に発育した8週齢の植物の若葉; SO4WFLD - 長日に発育した4週齢の植物の花芽)のホウレンソウRNA-Seqトランスクリプトームデータを用いて行った。共発現ネットワークは、Cytoscapeソフトウェア(「P. Shannon P et al., Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498-504 (2003)」)によって可視化した。

【0928】

クローニング

【0929】

Phusion High-Fidelity DNA Polymerase (New England Biolabs)を製造業者の使用説明書に従い、全てのPCR増

10

20

30

40

50

幅ステップに用いた。クローニング用の制限酵素及びT4リガーゼは、New England Biolabsから購入した。オリゴヌクレオチドプライマーは、Sigma-Aldrichから購入した。アガロースゲルから切り出したDNAを、Gel/PCR抽出キット(Hy-Labs)を用いて精製した。他の異種宿主に形質転換する前に、大腸菌TOP10細胞(Invitrogen)をプラスミドの単離に用いた。AccuPrep(商品名)Plasmid Mini Extraction Kit(Bioneer)を用いて、プラスミドDNAを大腸菌培養物から単離した。クローニングに用いたプライマーのリストについては、表11を参照されたい。

#### 【0930】

VIGSのコンストラクトの作製

10

#### 【0931】

ホウレンソウ中の候補遺伝子をサイレンシングするため、全ての遺伝子配列(200~400bp)を、ホウレンソウの葉のcDNAテンプレートから増幅した。精製した増幅産物を、SacIによって消化された組換えpTRV2ベクター(EcoRIによって挿入された約200bpのマグネシウムキラターゼサブユニットH)に挿入した。

#### 【0932】

GoldenBraidシステム及びベンサミアナタバコ(*N. benthamiana*)における発現

#### 【0933】

同じアグロバクテリウム株由来の複数の遺伝子をベンサミアナタバコにおいて一過性に発現させるため、GoldenBraid 3.0(GB)システム(「M. Vazquez-Villar et al., GB3.0: a platform for plant bio-design that connects functional DNA elements with associated biological data. *Nucleic Acids Res.* 45, 2196-2209 (2017)」)を用い、複数のコンストラクトを作製した。最初のステップで、BsaI及びBsmBI制限部位を取り出し、それらをpUPD2 GoldenBraidエンターベクターに挿入することにより、全てのCDSと、トマトユビキチンプロモータ及びターミネータとを馴化させた。次のステップで、遺伝子をターミネータ及びプロモータと共に、pDGB3 1またはpDGB3 2 GBベクターにサブクローニングした。第1レベルのベクターにより、ベンサミアナタバコにおいて単一遺伝子が一過性に発現する。続いて、pDGB3 1:SOAP1をpDGB3 2:SOAP2と組み合わせてpDGB3 1:SOAP1+SOAP2とし、pDGB3 1:SOAP3をpDGB3 2:SOAP4と組み合わせてpDGB3 2:SOAP3+SOAP4とした。オメガベクターの組み合わせにより、ベンサミアナタバコにおいてメディカゲン酸の産生をもたらす4つの遺伝子が発現する。他の遺伝子、セルロース合成酵素様G(SOAP5)、グリコシルトランスフェラーゼ(SOAP6-9)及びアシルトランスフェラーゼ(SOAP10)を、アルファベクターにおいて用いた。キノア(*C. quinoa*)、ビートルート(*B. vulgaris*)、ミヤコグサ(*L. japonicus*)、ダイズ(*Glycine max*(*G. max*))、アルファルファ(*M. sativa*)及びウラルカンゾウ(*G. uralensis*)を、同じ手法でクローニングした。cDNAテンプレートを作製するため、2週齢の実生(ミヤコグサ、ダイズ、アルファルファ及びウラルカンゾウ)、成熟葉(ビートルート)または花芽の周りの葉(キノア)から抽出したRNAを用いた。

20

30

40

#### 【0934】

酵母系(pESC)へのクローニング

#### 【0935】

S.セレビスエ(*S. cerevisiae*)におけるSOAP1(bAS)、CYP(SOAP2-4)及びCSLG(SOAP5)の発現のため、目的の配列を保有する、クローニングしたpDGB3ベクターから配列を増幅し、精製した増幅産物を、一連のpESC(Amp<sup>R</sup>)プラスミドに挿入し、1つのプラスミドから2つの遺伝子を同時に発現させた。NotI/SacI制限酵素を用いてSOAP1を、BamHI/SalI制限酵素を用いてSOAP2を、それぞれpESC-HISプラスミドに挿入した。NotI/Sa

50

c I 制限酵素を用いて S O A P 3 を、B a m H I / X h o I 制限酵素を用いて S O A P 4 を、それぞれ p E S C - L E U プラスミドに挿入した。N o t I / S a c I 制限酵素を用いて S O A P 5 を、S a l I / X h o I 制限酵素を用いて S o U G D 1 を、それぞれ p E S C - L E U プラスミドに挿入した。

【0936】

p E T - 2 8 b へのクローニング及び大腸菌における発現

【0937】

N d e I / N o t I 制限部位を用い、S O A P 1 0 を p E T - 2 8 b ( K a n <sup>R</sup> ) ベクターにクローニングし、大腸菌 B L 2 1 ( D E 3 ) 細胞において発現させた。

【0938】

S O A P 1 0 による I n v i t r o 酵素アッセイ

【0939】

p E T - 2 8 b : S O A P 1 0 で形質転換した細菌細胞を、L B 培地において 3 7 °C で培養した。培養物が A 6 0 0 = 0 . 6 に達したとき、イソプロピル - 1 - チオ - β - D - ガラクトピラノシド ( I P T G ) 2 0 0 μ M により、1 5 °C で 2 4 時間、タンパク質発現を誘導した。1 0 % グリセロール、0 . 1 % T r i t o n X - 1 0 0、1 m M P M S F 及びリゾチーム 1 0 0 m g / m L を含む 5 0 m M トリス H C l p H 8 . 0 中で、超音波処理によって細菌細胞を溶解した。

【0940】

S O A P 1 0 酵素アッセイにおける基質として、ハウレンソウデスアセチルサポニンの分画を用いた。2 0 0 μ L の管中で、細胞溶解物 6 μ L を、6 . 1 8 m M アセチルコエンザイム A ( A c e t y l - C o A ) 2 μ L 及びサポニン基質 ( 1 0 μ g ) 2 m L と混合した。反応を 2 8 °C で 2 . 5 時間行った。S O A P 1 0 と混合しない細胞溶解物を、対照として用いた。メタノール 2 0 μ L を添加して反応を停止した後、短時間ボルテックスした。最後に、上記のように、抽出物を 1 4 0 0 0 × g で 1 0 分間遠心分離し、L C - M S によって分析した。

【0941】

V I G S の手順

【0942】

ウイルス誘導遺伝子サイレンシングを、既報 ( 「M. Senthil-Kumar, K. S. Mysore, Tobacco rattle virus-based virus-induced gene silencing in Nicotiana benthamiana. Nat Protoc. 9, 1549-62 (2014)」 ) のように行った。簡潔に説明すると、サイレンシングのための配列を保有する p T R V 1 コンストラクト及び p T R V 2 コンストラクトを、アグロバクテリウム・ツメファシエンス ( Agrobacterium tumefaciens ) ( G V 3 1 0 1 ) の電気穿孔細胞に形質転換した。形質転換細胞を、カナマイシン 5 0 μ g / m L 及びゲンタマイシン 5 0 μ g / m L を含む L B プレートにおいて 2 8 °C で培養した。抗生物質を添加した L B 培地 1 0 m L にシングルコロニーを接種し、2 8 °C で終夜培養した。細胞を 3 0 0 0 × g で 1 0 分間遠心分離し、上清を除去した。ペレットを浸透媒体 ( 1 0 0 m M M E S 緩衝液、2 m M N a <sub>3</sub> P O <sub>4</sub> · 1 2 H <sub>2</sub> O、1 0 0 μ M アセトシリゴン ) 5 m L に再懸濁し、再び遠心分離した。ペレットを浸透媒体 1 0 m L に再び再懸濁し、室温で 2 時間インキュベートした。1 m L の無針シリンジを用い、アグロバクテリウムの懸濁液 ( 各株について O D 6 0 0 = 0 . 3 ) を 1 週齢のハウレンソウ子葉 ( 植物ごとに 1 つ ) の下側に浸透させた。植物を 1 6 時間の光周期、2 3 °C で生育させた。2 週間後に、マグネシウムキラターゼサブユニット H ( C H C L ) のサイレンシングの最初の兆候が見られた。組織の採取の前に、植物をさらに 2 週間生育させた。葉を採取し、液体窒素中で凍結し、後の加工処理用に - 8 0 °C で保存した。生物学的複製物は、全て異なる植物から得た 4 枚の葉で構成した。以下の表 1 0 は、赤カブ及びハウレンソウにおいて C S L G のサイレンシングに用いたハウレンソウ及び赤カブの核酸配列を示す。

【0943】

表 1 0 : V I G 配列

10

20

30

40

50

【 0 9 4 4 】

【 表 1 0 】

名称	SEQ ID NO:	核酸配列
SoCSLG (SOAP5) VIGS ハウレンソウ	106	TTCCGTAGCCATCTTCTCGCTCTTCTACTACCGTTTCACTTCCTTCTTCA ACTCCGACATCTCCATACTTGCTTACTCCTTACTCACCACCGCCGA CTCTTCTTAACCTTTCTATGGGCTTTTACTCAGGCTTTCCGGTGGCGTCCCG TAATGAGGGAAGTCTCCGGGTACGAATCCATCAAACCCGAACA ACTACCGGGTTTGGATGTCTTCATTGTCCTGCTGACCCGACAAAGGAGCCAGT TCTGGAGGTGATGAACTCCGTGATATCATCCATGGCGTTGGATTATCCG GTTGATAGACTGGCGGTTTACTTGTCCGGATGACGGTGGTTCTCCGTTGT CGAAGG
BvCSLG VIGS_ 赤カブ	107	CCTTCGGCTCCTCCACAAAATTCATTGCGTCAGTTAGTTCAA ACTCCAA GCAAAAATATGGCCTTGAAGGAAATGACAAGAGACGACTTGT TAGAAGA GGCTAAAAATTTGGCTACTTGTGCATATGAATCAAACACTGA ATGGGG TAACAAGATTGGATATTCGTATGAGTGTGTTGTTGGAGAGT ACATTTACC GGATATCTCTTACATTGCAAAGGATGGATTTCCGTGTATCTTT ACCCAA AAAGACCCTGCTTCTTAGGATGCACGACGATTGACATGAAAG ATGCCA TGGTTCAACTAATGAAATGGACCTCTGGATTACTAGGAGTTGG CATATC AAAGTTTAGCCCTCTAACTTATGCCTTTTCGAGGATGTCTAT ATTACAA AGCATGTGCTACGGTTACTTCACATTTTCAGCCCTTTTCGGAG TTTCGT

10

20

【 0 9 4 5 】

SOAP5の部位特異的突然変異誘発

【 0 9 4 6 】

QuickChange Site-Directed Mutagenesis Kit (Agilent)を用い、製造業者の使用説明書に従って、SOAP5の部位特異的突然変異誘発を行った。

【 0 9 4 7 】

遺伝子発現解析 - qPCR

【 0 9 4 8 】

各VIGSサイレンシング系について3つ/4つの生物学的複製物(n=3/4)を用い、遺伝子発現解析を行った。トリゾール法(Sigma-Aldrich)により、RNAの単離を行った。高容量cDNA逆転写キット(Applied Biosystems)を用い、DNase I(Sigma-Aldrich)処理したRNAを逆転写した。Primer Blastソフトウェア(NCBI)によって遺伝子特異的オリゴヌクレオチドを設計した。内在性対照として、翻訳伸長因子アルファ1(EF1-)遺伝子を用いた。使用したオリゴヌクレオチドを表11に示す。

30

【 0 9 4 9 】

表11：使用したプライマーのリスト

【 0 9 5 0 】

【表 1 1 - 1】

A	B	C	D	E	F	G
遺伝子名	注釈	種	遺伝子座	目的	プライマ名	配列
オキシドスクアレシクラーゼ						
1						
2						
3	β-アミリン合成酵素, bAS	ホウレンソウ	Sp_107620_kqnh	VIGS	Sp107620_vF	CTAGAGCTCATGTTCAAGAAGTTAT ACCC
4					Sp107620_vR	TTCGAGTCCCTCCTTTTGTGTAGT AA
5				ベンサミアナタバコ における発現	Sp107620_GB_1F	GCGCCGTCGCTCGAATGTGGAG GTTGAAGGTTGG
6					Sp107620_GB_1R	GCGCCGTCGCTGCTCATATGIG ATCTCCT
7					Sp107620_GB_2F	GCGCCGTCGACAGCCTCGACGA CATTAAAG
8					Sp107620_GB_2R	GCGCCGTCGCTCAAAGCTCAAGT AGAGTTGATAGAAGTAA
9				酵母における発現	Sp107620_NoII_F	AAAAGGCCCGCATGTGGAGTTG AAGTTTGA
10					Sp107620_SacI_R	AAAAGAGCTCTCAAGTAGAGTTGAT AGAAGGTAATG
11				共局在化	Sp107620_GB_1F	GCGCCGTCGCTCGAATGTGGAG GTTGAAGGTTGG
12					Sp107620_GB_CT_R	GCGCCGTCGCTCAGCAACCAGT AGAGTTGATAGAAGGTAATGA
13	シトクロムP450					
14	CYP715A56	ホウレンソウ	Sp_046690_sjiw	VIGS	Sp046690_vF	CTAGAGTCCCTTCGTCGTTGTTGTC GTCC
15					Sp046690_vR	TTCGAGTCCCTATGACGAGCCCAA GCAT
16	CYP72A656	ホウレンソウ	Sp_148230_dgra	VIGS	Sp148230_vF	CTAGAGTCAAGCTACTTTTGGCGTG CTCA
17					Sp148230_vR	TTCGAGTCTCTGGCCCTCATAGTT CGGA
18	CYP72A685	ホウレンソウ	Sp_085350_pxjs	VIGS	Sp085350_vF	CTAGAGTCTTGTGGGTTGCTTGT CCCT
19					Sp085350_vR	TTCGAGTCCGGTCTTTGTTGGCAGC AACC
20	CYP716A268	ホウレンソウ	Sp_107660_kiag	VIGS	Sp107660_vF	CTAGAGTCTGGTGGGGAGTCATT TGAA
21					Sp107660_vR	TTCGAGTCCCAAGCCAGCCAGAAG GTGTA

【 0 9 5 1 】

【表 1 1 - 2】

A	B	C	D	E	F	G
1 遺伝子名	注釈	種	遺伝子座	目的	プライマ名	配列
22				ベンサミアナタバコ における発現	Sp107660_GB_1F	GCGCCGTCGCTCGAATGGAAC CTTCTTATGTGG
23					Sp107660_GB_1R	GCGCCGTCGCTCAAAGCTTAAGC AGCTACAGTCGAGG
24				酵母における発現	Sp107660_BamHI_ F	AAAAGGATCCATGGAACCTCTCTTA TGTGTGG
25					Sp107660_Sall_R	AAAAGTCGACTTAAGCAGCTACAGT TCGAGGATG
26				共局在化	Sp107660_GB_1F	GCGCCGTCGCTCGAATGGAAC CTTCTTATGTGG
27					Sp107660_GB_CT R	GCGCCGTCGCTCAGCAACCAGC AGCTACAGTCGAGGAT
28 SOAP3	CYP72A655	ホウレンソウ	Sp_085340_meek	VIGS	Sp085340_vF	CTAGAGCTCGTTGCATTGCTTGTC CTGG
29					Sp085340_vR	TTGAGCTCTGATTGCAGAAGCGT AGGG
30				ベンサミアナタバコ における発現	Sp085340_GB_1F	GCGCCGTCGCTCGAATGATAGAA ATCGGGTATATGTAAA
31					Sp085340_GB_1R	GCGCCGTCGCTCAAAGCTTAGTC CTGAGCTTATGATAAT
32				酵母における発現	Sp085340_NotI_F	AAAAGGGCCGCATGATAGAAATCG GGTATATTGTAAAATG
33					Sp085340_SacI_R	AAAAGAGCTCTTAGTCCCTGAGCTT ATGTATAATGAG
34				共局在化	Sp085340_GB_1F	GCGCCGTCGCTCGAATGATAGAA ATCGGGTATATGTAAA
35					Sp085340_GB_CT R	GCGCCGTCGCTCAGCAACCAGTC CTGAGCTTATGATAATG
36 SOAP4	CYP72A654	ホウレンソウ	Sp_040350_wdny	VIGS	Sp040350_vF	CTAGAGCTCTTCAAAGAGCGCGAG TGGT
37					Sp040350_vR	TTGAGCTGGCCTGCACCAGAATTG TTGA
38				ベンサミアナタバコ における発現	Sp040350_GB_1F	GCGCCGTCGCTCGAATGATTTCA AAGAGCGCGAGT
39					Sp040350_GB_1R	GCGCCGTCGTTGTCTCTTGACCA GCCAAG
40					Sp040350_GB_2F	GCGCCGTCGCAACCTCAGTGG CCCTAAC

【 0 9 5 2 】

10

20

30

40

【表 1 1 - 3】

A	B	C	D	E	F	G
1 遺伝子名	注釈	種	遺伝子座	目的	プライマ名	配列
41					Sp040350_GB_2R	GC GCCG TCTCGCTCAAAGCTTAAAA TCGATGTAATAATAATGTGGG
42				酵母における発現	Sp040350_BamHI_F	AAAAGGATCCATGATTTCAAAGAGC GCGAGT
43					Sp040350_XhoI_R	AAACTCGAGTTAAATCGATGTAA AATAATGTGGG
44				共局在化	Sp040350_GB_1F	GGCCGCTCGCTCGAATGATTTCA AAGAGCGCGAGT
45					Sp040350_GB_CT_R	GGCCGCTCGCTCAGCAACCAAA CGATGTAATAATAATGTGGG
46	GuCYP88D6	ウラルカンゾウ	Glyur000561s0002345 1.1	ペンサミアナタバコ における発現	GuCYP88D6_GB_1F	GGCCGCTCGCTCGAATGGAAGT ACATTGGGTTTGC
47					GuCYP88D6_GB_1R	GC GCCG TCTCGCTCAAAGCCTAAG CACATGATACCTTATCAC
48	GuCYP72A15	ウラルカンゾウ	Glyur000890s0001907 1.1	ペンサミアナタバコ における発現	GuCYP72A154_GB_1F	GC GCCG TCTCGCTCGAATGGATGC ATCTTCCACACC
49					GuCYP72A154_GB_1R	GC GCCG TCTCGAAAGACCAATGGAT TTTGATTGTG
50					GuCYP72A154_GB_2F	GC GCCG TCTCGCTTCCAAAGATGA TGCTGCA
51					GuCYP72A154_GB_2R	GC GCCG TCTCGTTGCTCTTGCCCT GCCAGG
52					GuCYP72A154_GB_3F	GC GCCG TCTCGACAACCGCAGCTTT GCTGGC
53					GuCYP72A154_GB_3R	GC GCCG TCTCGCTCAAAGCTTACAG TTTATGCAGAATGATGGG
54				リコシトランスフェラーゼ		
55	セルロース 合成酵素様 G2, CSLG2	ホウレンソウ	Sp_076690_ejcm/Spo 12715	VIGS	Sp076690_vF	CTAGAGCTTCCGTAGCCATCTTC TCGC
56					Sp076690_vR	TTCGAGCTCCCTTCGACAACGGAGA ACCA
57				ペンサミアナタバコ における発現	Sp076690_GB_1F	GC GCCG TCTCGCTCGAATGGCAAC TTCTCACATTCGC
58					Sp076690_GB_1R	GC GCCG TCTCGGAAGACGGTTAAC AATGGCTC
59					Sp076690_GB_2F	GC GCCG TCTCGCTCACATCTTCCT CCATTCC

【 0 9 5 3 】

【表 1 1 - 4】

A	B	C	D	E	F	G
1 遺伝子名	注釈	種	遺伝子座	目的	プライマ名	配列
60					Sp076690_GB_2R	GCGCCGTCGCGCCTCTTTCCCTG GCTACA
61					Sp076690_GB_3F	GCGCCGTCGAGGCCTGGTCGTC CTCATCG
62					Sp076690_GB_3R	GCGCCGTCGCTCAAAGCTTATAA CCATCCCITAACAACAGG
63				酵母における発現	Sp012715F_NotI_F	AAAAGCGCCGCATGGCAACTTCTC ACATTCGCAA
64					Sp012715R_SacI_R	AAAAGAGCTTTATAACCATCCCTTA ACAACAGGGTAG
65				共局在化	Sp076690_GB_1F	GCGCCGTCGCTCGAATGGCAAC TTCTCACATTCCG
66					Sp076690_GB_CT_R	GCGCCGTCGCTCACGAACCTAAG CATCCCTTAACAACAGG
67				突然変異誘発	Sp012715_M1_F	CGTATGAGTGTGTTGGAGGATAC ATTCACCTGGATATATG
68					Sp012715_M1_R	CATATCCAGTGAATGTATCCTCCA ACAAGCACTCATAG
69					Sp012715_M2_F	CTACGGTTCAACTAATAAGATGGAC CTCCTCATTACTTGG
70					Sp012715_M2_R	CCAAGTAATGAGGAGGTCATCTTA TTAGTTGAACCCGTAG
71 SoGT1	UGT79M1	ホウレンソウ	Sp_049940_mmut	VIGS	Sp049940_vF	CTAGAGCTCAGCTTCCCTCCCATGG CATC
72					Sp049940_vR	TTCGAGCTCACTTGTGGCTGGAGTT GAGG
73 SoGT2	UGT79N1	ホウレンソウ	Sp_148240_fnwi	VIGS	Sp148240_vF	CTAGAGCTCCCCTCCTCGGCTATAC ACCT
74					Sp148240_vR	TTCGAGCTCCCCTAGTCTGCCACAG TCTC
75 SOAP6	UGT74BB2	ホウレンソウ	Sp_170930_hjgg	VIGS	Sp170930vF	CTAGAGCTCCCGAAGCCTCACTCCA ACTT
76					Sp170930vR	TTCGAGCTCTCAGAACTCGGGGATT GTGC
77				ペンサミアナタバコ における発現	Sp170930_GB_1F	GCGCCGTCGCTCGAATGACGGG AAAAGGAAGAACG
78					Sp170930_GB_1R	GCGCCGTCGCTCAAAGCTTAGGA GGACGCAAGCCAGT

【 0 9 5 4 】

10

20

30

40

【表 1 1 - 5】

A	B	C	D	E	F	G
1 遺伝子名	注釈	種	遺伝子座	目的	プライマ名	配列
79 SOAP6v2	UGT74BB1	ホウレンソウ	Sp_170920_oudh	VIGS	Sp170920_vF	CTAGAGCTCTCGTGGCTCGTATGGG AAAG
80					Sp170920_vR	TTCGAGCTCAGGAACAGATTCAGCC GCAA
81 SOAP7	UGT79K1	ホウレンソウ	Sp_020820_yeau	VIGS	Sp020820_vF	CTAGAGCTCTAGCCACAAAGCTGGG GATG
82					Sp020820_vR	TTCGAGCTCGGGAGGAATTCACACT CGGG
83				ペンサミアナタバコ における発現	Sp020820_GB_1F	GCGCCGCTCGCTCGAATGGGTAA AACAGTAGCAGCT
84					Sp020820_GB_1R	GCGCCGCTCGCTCAAAGCTTAATT TGCAGTAAGAAAGCGTTTC
85 SOAP8	UGT79L2	ホウレンソウ	Sp_113700_suxh	VIGS	Sp113700_vF	CTAGAGCTCGTGATGTTCCCATGGC TTGC
86					Sp113700_vR	TTCGAGCTCAAGTGGGCTTATGG GTTG
87				ペンサミアナタバコ における発現	Sp113700_GB_1F	GCGCCGCTCGCTCGAATGGGTGG AGAGAAAGAGTTG
88					Sp113700_GB_1R	GCGCCGCTCGCTCAAAGCTAAG GAACAAGGGCTTGTA
89 SOAP8v2	UGT79L1	ホウレンソウ	Sp_072870_funr	VIGS	Sp072870_vF	CTAGAGCTCGGCTTGCCTTGGACA CTTG
90					Sp072870_vR	TTCGAGCTCAAGACAAGGTCGGGCT TCTG
91 SOAP9	UGT73BS1	ホウレンソウ	Sp_170320_dmqj	VIGS	Sp170320_vF	CTAGAGCTCGTGTGCTGAGGTT GTTG
92					Sp170320_vR	TTCGAGCTCGTGCAGGTACCCGGA TAGG
93				ペンサミアナタバコ における発現	Sp170320_GB_1F	GCGCCGCTCGCTCGAATGGAGCT TTCAAACCCTAGC
94					Sp170320_GB_1R	GCGCCGCTCGGCGACGAACCACC TTCCTCA
95					Sp170320_GB_2F	GCGCCGCTCGTGGGTAACAATTT AAGTGATTG
96					Sp170320_GB_2R	GCGCCGCTCGCTCAAAGCTTATTC TGAGTTTGGACACTG
97 BvCSLG	CSLG2	ビートルート	XM_010673823.2	VIGS	BvCSL_vF	AAAAGAGCTCCCTTCGGCTCCTCCA CAAAA

【 0 9 5 5 】

10

20

30

40

【表 1 1 - 6】

A	B	C	D	E	F	G
1 遺伝子名	注釈	種	遺伝子座	目的	プライマ名	配列
98					BvCSL_vR	AAAAGAGCTCACGAAACTCCGAAAA GGGCT
99				ペンサミアナタハコ における発現	BvCSL_GB_1F	GGCCGTCGCTCGAATGCTTCT CTCCACATTGC
100					BvCSL_GB_1R	GGCCGTCGCGCTCTTTCCCTG GATACG
101					BvCSL_GB_2F	GGCCGTCGAGGGCCAAATCGTC CTCATCG
102					BvCSL_GB_2R	GGCCGTCGCGCATCTTGTCAAT TCCTTCAAG
103					BvCSL_GB_3F	GGCCGTCGCGATGACTTGTAGA AGAGGCT
104					BvCSL_GB_3R	GGCCGTCGCTCAAAGCTCAATC ACGTCTTTCTTACTTT
105 CqCSL	CSLG2	アルファアラブ	XM_021866098.1	ペンサミアナタハコ における発現	CqCSL_GB_1F	GGCCGTCGCTCGAATGGCGGC AACACACATTTG
106					CqCSL_GB_1R	GGCCGTCGCGCTCTTTCCCTG GCTACA
107					CqCSL_GB_2F	GGCCGTCGAGGGCCAGGTCATC CTCATCG
108					CqCSL_GB_2R	GGCCGTCGCTCAAAGCTTATTC TTTCTTCTAAGTTTGTCTG
109 MsCSL	CSLG2	ダイズ	MSAD_299835	ペンサミアナタハコ における発現	MsCSL_GB_1F	GGCCGTCGCTCGAATGGCAAC CTTCACATTTAC
110					MsCSL_GB_1R	GGCCGTCGCGCATCTCTTGAATA TTCTGCTTCT
111					MsCSL_GB_2F	GGCCGTCGCGATGTAATTTTACA AGAAGCATGTG
112					MsCSL_GB_2R	GGCCGTCGCTCAAAGCCTACC CACTCTCCGCTTCA
113				アルファアラブ毛状根に おけるサイレンシング	MsCSL_asGB_1F	GGCCGTCGCTCGAATGCTACC CACTCTCCGCTTCA
114					MsCSL_asGB_1R	GGCCGTCGCTCAAAGCATGGC AACCTTCACTTTCAC
115 GmCSL	CSLG2	ウラルカンゾウ	NM_001365113.1	ペンサミアナタハコ における発現	GmCSL_GB_1F	GGCCGTCGCTCGAATGGCGAC CTTCCACACAGA
116					GmCSL_GB_1R	GGCCGTCGCGGATCTCTTCCAAA TTCGTCA

【 0 9 5 6 】

10

20

30

40

【表 1 1 - 7】

A	B	C	D	E	F	G
遺伝子名	注釈	種	遺伝子座	目的	プライマ名	配列
1						
117					GmCSL_GB_2F	GCGCCGTCGATCCTAAAAATCGT TCCATTGIGT
118					GmCSL_GB_2R	GCGCCGTCGCTCAAAGCCTATTG CACCTTGCTTTTCAIG
119	GuCSL	ウラルカンゾウ	Glyur003152s0003749 1.1	ベンサミアナタバコ における発現	GuCSL_GB_1F	GCGCCGTCGCTCGAATGGCAAG CTTACCCTTCA
120					GuCSL_GB_1R	GCGCCGTCGATACCAGAGTGTTG ATCACCT
121					GuCSL_GB_2F	GCGCCGTCGGTATCTGCCCTTGC CATGGA
122					GuCSL_GB_2R	GCGCCGTCGAGATCTCGCTCTGA CCTGAA
123					GuCSL_GB_3F	GCGCCGTCGATCTCATCAAGGCT AAATACGAG
124					GuCSL_GB_3R	GCGCCGTCGCTCAAAGCCTATCC ACTTTGCTTTTCAIG
125	GuUGAT	ウラルカンゾウ	KT759000.1	ベンサミアナタバコ における発現	GuUGAT_GB_1F	GCGCCGTCGCTCGAATGACCAT GGTAACGAGAAT
126					GuUGAT_GB_1R	GCGCCGTCGGGCGGACGATCCTCC TTCCTCC
127					GuUGAT_GB_2F	GCGCCGTCGTCGCACAACGATTT TAACTCTTTA
128					GuUGAT_GB_2R	GCGCCGTCGCTCAAAGCTTAATG GGCACGCGACCTCA
129	GuUGT73P1	ウラルカンゾウ	Scaffold00629 (LC314779)	ベンサミアナタバコ における発現	GuUGT73P12_GB_1F	GCGCCGTCGCTCGAATGGACTC CTTTGGGGTTGA
130					GuUGT73P12_GB_1R	GCGCCGTCGGGCTCATTTCAACG GAACAGT
131					GuUGT73P12_GB_2F	GCGCCGTCGAGGGCCAGATTTCAT AGTCACT
132					GuUGT73P12_GB_2R	GCGCCGTCGCTCAAAGCTTAAGC CACTGCCCTCCATTAA
133						
アシルトランスフェラーゼ						
134	SoAT1	ホウレンソウ	Sp_074630_vgho	VIGS	Spo04549_vF	CTAGAGCTCAGGGCTTGCTATCGAT CCAG

10

20

30

40

【表 1 1 - 8】

	A	B	C	D	E	F	G
1	遺伝子名	注釈	種	遺伝子座	目的	プライマ名	配列
135						Spo04549_vR	TTCGAGCTCCATGGCCTTGAGTCTC AGCA
136	SoAT2	O-アシルトランス フェラーゼ WSD1様	ホウレンソウ	Sp_123780_pgiiy	VIGS	Spo21561_vF Spo21561_vR	CTAGAGTCCCGGAAATGAAGGTCT GGGT TTCGAGCTCAGGGGGTTCGAGCATT TGTC
137							
138	SoAT3	アシル-CoA ーステロールO- アシルトランス フェラーゼ1	ホウレンソウ	Sp_149180_nwmy	VIGS	Spo15788_vF Spo15788_vR	CTAGAGTCTGATAAAGGGCCCACTT TGCG TTCGAGTCTCCAGTTCGAGCCCT AACAA
139							
140	SoAT4	マロニル -CoA: アントシアニン5 -O-グルコシド -6''-マロニル トランスフェラーゼ 様	ホウレンソウ	Sp_198340_focw	VIGS	Spo13090_vF Spo13090_vR	CTAGAGTCAAGCCAAAGCAAAAGG CACC TTCGAGTCTCGGTTTCCCCCACC AAAA
141							
142	SOAP10	サルタリジノール 7-O- アセチルトランス フェラーゼ	ホウレンソウ	Sp_125800_kzws	VIGS	Spo02253_vF Spo02253_vR	CTAGAGTCCAACAACCTTCCAAGG CGGC TTCGAGTCTCTTGAAGCCCATTTGC TGCT
143							
144					大腸菌における発現	Spo02253_NdeI_F Spo02253_NotI_R	GAAACAATGGGGGAAGTCAACCAT GAAGAAG CTTTGCGGCCGCCCTAATTAGGAGTA GCAAAAGCAAGG
145							
146					パンサミアナタバコ における発現	Spo02253_GB_1F Spo02253_GB_1R	GCGCCGTCGCTCGAATGGGGGA AGTCAACCATGA GCGCCGTCGCTGTCTCAAATTTG GTGGGT
147							
148						Spo02253_GB_2F	GCGCCGTCGACAGTGACAGGGT TTATTTGG

【 0 9 5 8 】

10

20

30

40

【表 11 - 9】

	A	B	C	D	E	F	G
1	遺伝子名	注釈	種	遺伝子座	目的	プライマ名	配列
149						Spo02253_GB_2R	GCGCCGTCGCGCTCCCATCGTT GGGAATT
150						Spo02253_GB_3F	GCGCCGTCGAGGCCCCCTAAGG TCAGGAA
151						Spo02253_GB_3R	GCGCCGTCGCTCAAAGCCTAATT AGGAGTAGCAAAAGCAAG
152					その他		
153	SoCHLH	マグネシウム キラターゼ サブユニットH	ホウレンソウ	Sp_082000_pixp	VIGS	SoCHLH_vF	ACTAGAAATCCAAGTGGGGATGAGT GATGCTTG
154						SoCHLH_vR	GTTCGAATTCCTAGCAGTGTGATG ATGGAACTC
155	BvCHLH	マグネシウム キラターゼ サブユニットH	ビートルート	XM_010674548.2	VIGS	BvCHLH_vF	AAAAGAATTCGGAGGCAAGAGGGG CTAAAG
156						BvCHLH_vR	AAAAGAATTCCTTCCCTTCAAGTCC GCCC
157	SoUGD1	UDP- グルコース6- デヒドロゲナーゼ1	ホウレンソウ	Sp_189830_pzca	酵母における発現	SoUGD1_Sall_F	AAAAGTCGACATGGTGAAGAAACTG AAGATTTGCTG
158						SoUGD1_XhoI_R	AAAAGTCGAGTTAAGCTACGGCAGG CATGTC

10

20

30

40

【0959】

ベンサムアナタバコ (*N. benthamiana*) における一過性発現及び候補遺伝子のスクリーニング

【0960】

50

p D B G 3 コンストラクト及び p D B G 3 コンストラクトを、アグロバクテリウム・ツメファシエンス (*Agrobacterium tumefaciens*) (G V 3 1 0 1) の電気穿孔細胞に形質転換した。形質転換細胞を、カナマイシン (p D B G 3 ) 5 0  $\mu$  g / m L またはスベクチノマイシン (p D B G 3 ) 2 0 0  $\mu$  g / m L、及びゲンタマイシン 5 0  $\mu$  g / m L を含む L B プレートにおいて 2 8  $^{\circ}$  C で培養した。抗生物質を添加した L B 培地 1 0 m L にシングルコロニーを接種し、2 8  $^{\circ}$  C で終夜培養した。細胞を 3 0 0 0  $\times$  g で 1 0 分間遠心分離し、上清を除去した。ペレットを浸透媒体 ( 1 0 0 m M M E S 緩衝液、2 m M N a <sub>3</sub> P O <sub>4</sub>  $\cdot$  1 2 H <sub>2</sub> O、1 0 0  $\mu$  M アセトシリンゴン ) 5 m L に再懸濁し、再び遠心分離した。ペレットを浸透媒体 1 0 m L に再び再懸濁し、室温で 2 時間インキュベートした。1 m L の無針シリンジを用い、アグロバクテリウムの懸濁液 ( 各株について O D 6 0 0 = 0 . 3 ) をベンサミアナタバコの葉の下側に浸透させた。浸透前に、植物を 1 6 時間の光サイクル下で 4 ~ 5 週間生育させた。浸透から 4 日後に葉を採取し、液体窒素中で凍結し、後の加工処理用に - 8 0  $^{\circ}$  C で保存した。生物学的複製物は、全て異なるタバコ植物から得た数枚の葉で構成した。

【 0 9 6 1 】

S . セレピシエ W A T 1 1 における S O A P 遺伝子の発現及び代謝物の抽出

【 0 9 6 2 】

Y e a s t m a k e r ( 商品名 ) 酵素形質転換システム ( C l o n t e c h ) を用い、p E S C コンストラクトをサッカロミセス・セレピシエ (*Saccharomyces cerevisiae*) W A T 1 1 に形質転換した。1 つの遺伝子 ( S O A P 5 )、2 つの遺伝子 ( S O A P 1 ~ 2 )、4 つの遺伝子 ( S O A P 1 ~ 4 )、5 つの遺伝子 ( S O A P 1 ~ 5 または S O A P 1 ~ 4 + S o U G D 1 ) 及び 6 つの遺伝子 ( S O A P 1 ~ 5 + S o U G D 1 ) を発現させる p E S C ベクターの様々な組み合わせによって、酵母細胞を形質転換した。形質転換した酵母を、適切なアミノ酸及び 2 % グルコースを添加した S D 最小培地で培養した。コロニーをスクリーニングし、導入遺伝子の存在を P C R によって確認した。遺伝子発現を誘導するため、2 % ガラクトースを添加した最小培地に形質転換細胞を移し、3 0  $^{\circ}$  C で 2 4 時間培養した。培養物を 7 0 0  $\times$  g で 1 0 分間遠心分離し、ペレットを H <sub>2</sub> O 1 m L に再懸濁し、2 m L の E p p e n d o r f 管に移し、8 0 0 0  $\times$  g で 1 分間遠心分離した。細胞ペレットを秤量し、同量の水と 2 倍量のガラスビーズ ( 直径 5 0 0  $\mu$  m ) とを加え、1 分間のボルテックスを 5 回行い、その都度、ボルテックスの間に細胞を氷上に 1 分間保持した。溶解細胞をメタノール 5 0 0  $\mu$  L と混合し、1 4 0 0 0 r p m で 5 分間遠心分離し、澄明な上清を採取し、S p e e d V a c 中で終夜乾燥した。乾燥残渣を 8 0 % メタノール 1 5 0  $\mu$  L に溶解し、0 . 2 2  $\mu$  m のフィルターで濾過し、L C - M S で分析した。

【 0 9 6 3 】

細胞内局在化及び共焦点顕微鏡分析

【 0 9 6 4 】

G o l d e n B r a i d システム ( 前掲の「M. Vazquez-Vilar et al., (2017)」) を用い、蛍光タンパク質 ( F P ) との融合において検討したタンパク質を発現させるプラスミドを作製した。S O A P 1 : G F P、- G F P または - R F P と融合した S O A P 2 ~ 4、及び S O A P 5 : R F P を用いた。細胞コンパートメントマーカーとして、A B R C ( E R g k C D 3 - 9 5 5 及び G g k C D 3 9 6 3 ) から得たプラスミドを用いた (「B. K. Nelson, X. Cai, A. Nebenfuhr, A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. Plant J . 51, 1126-36 (2007)」)。

【 0 9 6 5 】

ハウレンソウセルロース合成酵素様 G の細胞内局在化試験のため、ベンサミアナタバコ表皮細胞において、E R またはゴルジマーカーと共に、S O A P 5 : R F P を一過性に発現させた。細菌の吸光度 A 6 0 0 n m として 0 . 1 5 を、各アグロバクテリウム株の浸透に用いた。浸透の 7 2 時間後に、葉片 ( 直径約 0 . 4 c m ) を採取し、共焦点顕微鏡法により、以下のパラメータを用いて蛍光を分析した。蛍光は N i k o n E c l i p s e

10

20

30

40

50

A1顕微鏡で観察し、488nmのレーザーで励起してGFP（ERまたはゴルジマーカ）の画像を取得し、561nmで赤色蛍光タンパク質（RFP）シグナルの画像を取得した。シグナル-ノイズ比を高めるため、各走査画素を4回サンプリングし、平均した。同じパラメータを用い、SOAP遺伝子の共局在化を確認した。

【0966】

蛍光共鳴エネルギー移動（FRET）分析

【0967】

FRET Analyzerプラグイン（「M. Hachet-Haas et al., FRET and colocalization analyzer a method to validate measurements of sensitized emission FRET acquired by confocal microscopy and available as an ImageJ Plug-in. *Microsc Res Tech.* 69, 941-56 (2006)」）Fiji/ImageJを用い、平均FRETインデックス（すなわち、閾値を適用してバックグラウンドノイズを除去した後の各画素においてFRETに起因する、受容体、mRFP融合タンパク質、強度の平均）を算出した。平均FRETインデックスは、近接性/相互作用について試験対象となる各タンパク質の組み合わせに対し、3つの独立した画像について算出した。GFP及びRFPを発現する細胞のみを用い、供与体及び受容体のブリードスルーを算出した。

10

【0968】

系統発生解析

【0969】

NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) の非冗長タンパク質データベースにおいてBlastpオプションを用い、様々な問い合わせ配列によって相同性検索を行った。Muscleアルゴリズムを用いてタンパク質配列をアライメントし、RAxML高速ブートストラッピングとそれに続くML検索により、系統樹を推測した（「A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 30, 1312-3 (2014)」）。最終的な系統樹の尤度を、座位間の進化速度の不均質性のGAMMAモデル下で評価し、最適化した。GAMMAモデルパラメータは、最大0.1000000000対数尤度単位の精度で評価した。RAxMLの高速ブートストラップオプションを用い、1000個のブートストラップの複製を行った。用いた置換行列はDAYHOFFであった。iTOL（「I. Letunic, P. Bork, Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44, W242-5 (2016)」）によって系統発生樹を可視化した。

20

30

【0970】

アルファルファ（*M. sativa*）毛状根におけるMsCSLのサイレンシング

【0971】

GoldenBraidシステムを用い、35Sプロモータ下でMsCSLを標的とするアンチセンスRNAと、35S:KanRカセットとを有するベクターpDBG3-2を作製した。アルファルファの形質転換及び遺伝子導入毛状根の作製を、既報（「A. Boisson-Dernier, M. Chabaud, F. Garcia, G. Becard, C. Rosenberg, D. G. Barker, *Agrobacterium rhizogenes*-transformed roots of *Medicago truncatula* for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. *Mol Plant Microbe Interact.* 14, 695-700 (2001)」）のように行った。Arqua1の代わりに、アグロバクテリウム・リゾゲネス（*A. rhizogenes*）のATCC15834株を用いた。遺伝子導入根を採取し、粉碎し、代謝物を上記のように抽出した。

40

【0972】

テンプレートベースのタンパク質構造モデリング

【0973】

RaptorX（「M. Kallberg et al., Template-based protein structure modeling using the RaptorX web server. *Nature Protocols.* 7, 1511-1522 (2012)」）を用い、サイクリック-ジ-GMP及びUDPとの複合体におけるロドバクター・スフェロイデ

50

ス (Rhodobacter sphaeroides) の結晶構造をテンプレート (PDB エントリ: 4P00) として用い (「J. L. Morgan, J. T. McNamara, J. Zimmer, Mechanism of activation of bacterial cellulose synthase by cyclic di-GMP. Nat Struct Mol Biol. 21, 489-96 (2014)」)、SOAP5 の相同性モデルを作製した。最良のモデルを、チャネル、トンネル及び細孔の位置の特定並びに特性評価を自動的に行うための汎用ツールキットである MOLE2.5 (「L. Pravda et al., MOLEonline: a web-based tool for analyzing channels, tunnels and pores (2018 update). Nucleic Acids Res. 46, W368-W373 (2018)」) のテンプレートとして用いた。SOAP5 の膜貫通トポロジーを表すスキームを、Protter (「U. Omasits, C. H. Ahrens, S. Muller, B. Wollscheid, Protter: interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics. 30, 884-6 (2014)」) で作成した。

10

【0974】

実施例 15 : ホウレンソウにおけるトリテルペノイド系サポニン生合成経路

【0975】

目的: ナデシコ目 (Caryophyllales) の一種であるホウレンソウ (Spinacia oleracea) におけるトリテルペノイド系サポニン生合成経路、特に、トリテルペノイド系サポニンのグリコシル化に關与する酵素を見出すこと。

【0976】

方法: 液体クロマトグラフィー - 質量分析法を、ホウレンソウの葉におけるサポニンの分析に用いた。上記の材料及び方法を参照されたい。

20

【0977】

結果:

【0978】

ホウレンソウの葉を分析したところ (LC-MS を使用)、アグリコンとしてメディカゲン酸を含む 20 を超える分子を含む複合型トリテルペノイド系サポニン混合物 (図 18、図 19A ~ 19P、表 12) が明らかとなり、それらのおよそ半分がアセチル化されていた。以下の表 13 は、図 18 の下方の走査によって同定された各化合物についての詳細を示す。以下の表 12 では、以下の用語を用いる。試料 - 試料が由来する試験 / 遺伝子型の種類; 保持時間 - 分単位の保持時間; 推定名 - 推定代謝物の同定; 分子式 - 代謝物またはその FA 付加物の分子式; 理論  $m/z$  - イオン  $[M-H]^-$ 、 $[M+H]^+$  について算出した理論的モノアイソトピック質量; 実測  $m/z$  - 実測された質量;  $m/z$  誤差 (ppm) - 理論値と実測  $m/z$  値との ppm 単位の差; MS / MS フラグメント - イオン  $[M-H]^-$ 、 $[M+H]^+$  から得たフラグメント; MS / MS CE (eV) - フラグメンテーションに用いた衝突エネルギー; UV / Vis - UV / Vis 吸収極大。UV / Vis スペクトル (200 ~ 600 nm) は、材料及び方法に記載した LC 条件において、Acquity 2996 PDA 検出器を備えた UPLC (Waters, Acquity) 装置で取得した。(S) - 標準化合物によって確認された同定; (NMR) - NMR によって確認された同定; MA - メディカゲン酸; AA - アウグスチン酸 (2-ヒドロキシオレアノール酸); B - バヨゲニン; PG - ポリガラゲニン; OA - オレアノール酸; G - ギブソゲニン; GA - ギブソゲン酸; H - ヘデラゲニン; GhA - グリチルレチン酸; hexA - ヘキサロン酸; hex - ヘキソース; dhex - デオキシヘキソース; pent - ペントース; GlcA - グルクロン酸; Ac - アセチル。

30

40

【0979】

表 12 : 質量分析に基づく試験した植物中のトリテルペノイド系サポニンの同定

【0980】

【表 1 2 - 1】

Section A

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
試料	ベンザミナナノコ で検出した遊分子 の組み合わせ	保持時間	推定名	分子式	理論 m/z [M+H] <sup>+</sup>	実測 m/z [M+H] <sup>+</sup>	m/z 誤差 (ppm)	ES(+) 理論 m/z [M+H] <sup>+</sup>	ES(+) 実測 m/z [M+H] <sup>+</sup>	誤差 (ppm)	MS/MS ES(+) フラグメント	MS/MS ES(+) フラグメント	ES(+) CE (eV)	
1	ホウレンソウ	-	13.7		MA-dhex+thex+pent+hexA	C53 H82 O24	1101.5118	1101.5118	-0.2	-	-	-	-	-
2	ホウレンソウ	-	12.8	Yosossida VI	MA-dhex+thex+hexA	C54 H84 O25	1131.5223	1131.5223	-1.9	-	-	-	-	-
3	ホウレンソウ	-	19.4	Yosossida XII	MA-dhex+thex+Ac+pent+hexA	C55 H84 O25	1143.5223	1143.5223	3.7	-	-	-	-	-
4	ホウレンソウ	-	14.7	Yosossida XII	MA-dhex+thex+pent+pent+hexA	C56 H90 O28	1233.5534	1233.5534	1.1	-	-	-	-	-
5	ホウレンソウ	-	12.8	Yosossida IV	MA-dhex+thex+hex+pent+hexA	C59 H92 O29	1263.5713	1263.5713	0.6	-	-	-	-	-
7	ホウレンソウ	-	20.84	Yosossida VII	MA-dhex+thex+Ac+pent+pent+hexA	C60 H92 O29	1275.5646	1275.5646	-0.5	-	-	-	-	-
8	ホウレンソウ	-	15.8	Yosossida Va	MA-dhex+thex+Ac+hex+pent+hexA	C61 H94 O30	1305.5732	1305.5732	1.6	-	-	-	-	-
9	ホウレンソウ	-	18.6	Yosossida V (NMI)	MA-dhex+thex+Ac+hex+pent+hexA	C61 H94 O30	1305.5732	1305.5732	2.2	-	-	-	-	-
10	ホウレンソウ	-	14.2	Yosossida VIII	MA-dhex+thex+pent+pent+pent+hexA	C63 H98 O32	1365.5981	1365.5981	1.3	-	-	-	-	-
11	ホウレンソウ	-	12.4	Yosossida XI	MA-dhex+thex+hex+pent+pent+hexA	C64 H100 O33	1395.6069	1395.6069	-0.1	-	-	-	-	-
12	ホウレンソウ	-	18.9	Yosossida IX	MA-dhex+thex+Ac+pent+pent+pent+hexA	C65 H100 O33	1407.6107	1407.6107	2.7	-	-	-	-	-
13	ホウレンソウ	-	16.7	Yosossida X	MA-dhex+thex+Ac+hex+pent+pent+hexA	C66 H102 O34	1437.6174	1437.6174	0.6	-	-	-	-	-
14		-	14											
15	ホウレンソウ	-	16.6		AA-hex+hexA	C42 H68 O15	809.4323	809.4323	1.7	-	-	-	-	-
16	ホウレンソウ	-	10.3		B-hex+hexA	C42 H68 O16	825.4273	825.4273	0.6	-	-	-	-	-
17	ホウレンソウ	-	15.4		AA-hex+pent+hexA	C47 H74 O19	941.4746	941.4746	-1.0	-	-	-	-	-
18	ホウレンソウ	-	15.4		AA-hex+hexA+3-oxopyruvic acid and glycolic acid isomer 1	C47 H70 O21	969.4331	969.4326	-0.5	-	-	-	-	-
19	ホウレンソウ	-	16.2		AA-hex+hexA+3-oxopyruvic acid and glycolic acid isomer 2	C47 H70 O21	969.4331	969.4327	-0.4	-	-	-	-	-
20	ホウレンソウ	-	13.1		AA-hex+hexA	C48 H76 O20	971.4866	971.4866	1.4	-	-	-	-	-
21	ホウレンソウ	-	14.4		AA-hex+pent+pent+hexA	C52 H82 O23	1073.5169	1073.5216	4.4	-	-	-	-	-
22	ホウレンソウ	-	11.6		PG-hex+pent+pent+hexA	C52 H80 O24	1087.4961	1087.4960	2.7	-	-	-	-	-
23	ホウレンソウ	-	18.7	Yosossida VIIa	MA-dhex+thex+Ac+pent+pent+hexA	C56 H92 O29	1275.5646	1275.5675	2.9	-	-	-	-	-
24	ホウレンソウ	-	17.2		OA-hex+hexA+3-oxopyruvic acid and glycolic acid (Betavulgaroside I) isomer 1	C47 H70 O20	953.4382	953.4366	-1.7	-	-	-	-	-
25	ホウレンソウ	-	18.4		OA-hex+hexA+3-oxopyruvic acid and glycolic acid (Betavulgaroside I) isomer 2	C47 H70 O20	953.4382	953.4369	-1.4	-	-	-	-	-
26	ホウレンソウ	-	14.6		G-hex+hexA+3-oxopyruvic acid and glycolic acid (Basilasaponin B)	C47 H68 O21	967.4175	967.4183	0.8	-	-	-	-	-
27	ホウレンソウ	-	16.1		GA-hex+pent+thex+hexA	C59 H92 O28	1247.5697	1247.5685	-0.2	-	-	-	-	-
28	ホウレンソウ	-	21.2		GA-hex+pent+thex+hexA+Ac+hexA	C61 H94 O29	1289.5803	1289.5781	-1.7	-	-	-	-	-
29		-												
30	ホウレンソウ	-	10.74		MA-hex+hexA+3-oxopyruvic acid and glycolic acid isomer 1	C47 H68 O23	999.4073	999.4097	2.4	-	-	-	-	-
31	ホウレンソウ	-	11.90		MA-hex+hexA+3-oxopyruvic acid and glycolic acid isomer 2	C47 H68 O23	999.4081	999.4087	1.6	-	-	-	-	-
32	ホウレンソウ	-	15.30		MA-dhex+hexA+3-oxopyruvic acid and glycolic acid (Betavulgaroside II)	C47 H68 O22	993.4124	993.4155	3.2	-	-	-	-	-
33	ホウレンソウ	-	15.31		MA-hexA+thex+pent (Betavulgaroside III)	C47 H72 O20	955.4539	955.4543	0.4	-	-	-	-	-
34	ホウレンソウ	-	11.00		MA+pent+thex+hexA	C47 H72 O21	971.4488	971.4506	1.9	-	-	-	-	-
35	ホウレンソウ	-	13.40		MA-dhex+thex+hexA	C48 H74 O20	969.4695	969.4661	-3.5	-	-	-	-	-
36	ホウレンソウ	-	17.10		OA-hexA+hex+pent	C47 H74 O18	925.4797	925.4789	-0.9	-	-	-	-	-
37	ホウレンソウ	-	12.50		AA-H-hexA+thex	C42 H68 O15	809.4323	809.4335	1.5	-	-	-	-	-
38	ホウレンソウ	-	14.62		G-hexA+thex	C42 H64 O16	807.4167	807.4156	-1.4	-	-	-	-	-
39	ホウレンソウ	-	12.90		GA/PG-hex+hexA	C42 H64 O16	823.4116	823.4129	1.6	-	-	-	-	-
40	ホウレンソウ	-	10.00		MA-dhex+pent+pent+hex+hexA	C58 H90 O29	1249.5490	1249.5537	3.8	-	-	-	-	-
41	ホウレンソウ	-	8.50		MA-dhex+thex+hex+hexA	C54 H84 O26	1147.5173	1147.5226	4.6	-	-	-	-	-
42	ホウレンソウ	-	9.80		MA-dhex+pent+hex+hexA	C53 H82 O25	1117.5067	1117.5078	1.0	-	-	-	-	-
43	ホウレンソウ	-	11.00		MA-hex+hexA	C42 H64 O17	839.4065	839.4077	1.4	-	-	-	-	-
44	ホウレンソウ	-	21.00		MA+pent+hexA	C41 H62 O16	809.3960	809.3938	-2.7	-	-	-	-	-
45	ホウレンソウ	-	15.25		PG/GA-hex+hexA	C42 H64 O16	823.4116	823.4132	1.9	-	-	-	-	-
46	ホウレンソウ	-	13.20		MA-dhex+thex+pent+pent+hexA	C58 H90 O28	1233.5640	1233.5499	-3.3	-	-	-	-	-
47	ホウレンソウ	-	17.60		MA-dhex+thex+Ac+pent+pent+hexA	C60 H92 O29	1275.5646	1275.5627	-1.5	-	-	-	-	-
48	ホウレンソウ	-	11.15		MA-hex+thex	C42 H68 O16	825.4273	825.4283	1.2	-	-	-	-	-
49	ホウレンソウ	-	22.73	Medicagenic acid (S)	C30 H46 O6	503.3373	501.3216	-1.6	-	-	-	-	-	-
50		-												25
51	アルファアルファ7	-	21.61	Sapontin 1	B-hex+Hex+hexA	C46 H76 O21	997.4801	997.4788	-1.3	-	-	-	-	-
52	アルファアルファ7	-	25.63	Sapontin 2	B-hex+hexA	C42 H68 O16	825.4273	825.4273	0.0	-	-	-	-	-
53	アルファアルファ7	-	23.04	Sapontin 3	Soyasapogenol A+dhex+hex+hexA	C48 H78 O19	957.5059	957.5059	-0.2	-	-	-	-	-
54	アルファアルファ7	-	24.11	Sapontin 4	GA/PG+hex+hexA+maltolyl	C45 H66 O19	909.4120	909.4120	0.0	-	-	-	-	-

【表 1 2 - 2】

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	試料	ベンザミナタハコ で発現した遺伝子 の組み合わせ	保持時間	推定名	アグリコン+糖	分子式	理論 m/z (M+H) <sup>+</sup>	ES(+) 実測 m/z (M+H) <sup>+</sup>	m/z 誤差 (ppm)	ES(+) 理論 m/z (M+H) <sup>+</sup>	ES(+) 実測 m/z (M+H) <sup>+</sup>	m/z 誤差 (ppm)	MISMS ES(+) フラグメント	MISMS ES(+) CE (eV)
1														
55	アルファルブ7	M s C S L が サイレンシング された アルファルブ7 由来の 遺伝子	25.25	Sapontin 5	B-2-hex+hex+hexA	C48 H78 O20	-	971.4852	-	971.4852	971.4849	-0.3	-	-
56	アルファルブ7		22.84	Sapontin 6	H-hex+hex+dlhex+hexA	C54 H86 O24	-	1117.5431	-	1117.5420	1117.5420	-1.0	-	-
57	アルファルブ7		22.92	Sapontin 7	G-A-PG-hex+hexA	C42 H64 O16	-	823.4116	-	823.4078	823.4078	-4.6	-	-
58	アルファルブ7		28.14	Sapontin 8	Soyasapogenin B-dhex+hex+hex+hexA	C54 H88 O23	-	1103.5638	-	1103.5638	1103.5635	-0.3	-	-
60	ピートリート		14.95	B-Sapontin 1 (beta-ujferoside V)	O-hex+hex+hex+AcI	C53 H82 O25	-	1117.5067	-	1117.5060	1117.5060	-0.6	-	-
61	ピートリート		15.24	B-Sapontin 2	O-hex+hex+hex+pent	C53 H84 O23	-	1087.5225	-	1087.5216	1087.5216	-0.9	-	-
62	ピートリート		15.67	B-Sapontin 3 (beta-ujferoside X)	O-hex+hex+hex+pent+Ad	C52 H80 O24	-	1087.4961	-	1087.4978	1087.4978	1.6	-	-
63	ピートリート		16.95	B-Sapontin 4	O-hex+pent+hexA	C47 H74 O18	-	925.4773	-	925.4782	925.4782	-1.6	-	-
64	ピートリート		17.54	B-Sapontin 5 (beta-ujferoside III)	O-hex+hex+hex+Ad	C47 H72 O20	-	955.4539	-	955.4536	955.4536	-0.3	-	-
65														
66	ベンザミナタハコ における発現	SOAP1-4+SOAP5	20.98	medicagenic acid 3- O-glucuronides (NMR)	MA+GlcA	C36 H54 O12	-	677.3543	-	677.3522	677.3522	-2.2	-	-
67	ベンザミナタハコ における発現	SOAP1-5+SOAP6	14.96	Yessoside 1	MA+GlcA+Fu	C42 H64 O16	-	823.4116	-	823.4104	823.4104	-1.5	-	-
68	ベンザミナタハコ における発現	SOAP1-6+SOAP7	13.21	Yessoside 2	MA+GlcA+Fu+Rha	C48 H74 O20	-	999.4695	-	999.4693	999.4693	-0.2	-	-
69	ベンザミナタハコ における発現	SOAP1-7+SOAP8	12.08	Yessoside 3	MA+GlcA+Fu+Rha+Glc	C54 H84 O25	-	1131.5200	-	1131.5199	1131.5199	-0.1	-	-
70	ベンザミナタハコ における発現	SOAP1-8+SOAP9	12.37	Yessoside 4	MA+GlcA+Fu+Rha+Glc+Xyl	C59 H92 O29	-	1263.5646	-	1263.5649	1263.5649	0.2	-	-
71	ベンザミナタハコ における発現	SOAP1-9+SOAP10	18.18	Yessoside 5	MA+GlcA+Fu+Ac+Rha+Glc+Xyl	C61 H94 O30	-	1305.5752	-	1305.5736	1305.5736	-1.2	-	-
72														
73	ベンザミナタハコ における発現	BAS-CYP72A154+C YP88D6+GlcSL	17.2	glycyrrhetic acid 3- O-glucuronides	GMA-GlcA	C36 H54 O10	-	645.3644	-	645.3644	645.3644	0.1	-	-
74	ベンザミナタハコ における発現	BAS-CYP72A154+C YP88D6+GlcSL	19.0	glycyrrhetic acid 3- O-glucuronides	GMA-GlcA	C36 H54 O10	-	645.3644	-	645.3644	645.3635	1.2	-	-
75														
76	ベンザミナタハコ における発現	BAS-CYP72A154+C YP88D6+SOAP5	17.2	glycyrrhetic acid 3- O-glucuronides	GMA-GlcA	C36 H54 O10	-	645.3644	-	645.3644	645.3648	-0.6	-	-
77	ベンザミナタハコ における発現	BAS-CYP72A154+C YP88D6+SOAP5	19.0	glycyrrhetic acid 3- O-glucuronides	GMA-GlcA	C36 H54 O10	-	645.3644	-	645.3644	645.3659	-2.3	-	-
78														
79	ベンザミナタハコ における発現	BAS-CYP72A154+C YP88D6+GlcSL+UG T73P12	11.8	glycyrrizin isomer 1 (S)	GMA-GlcA-GlcA	C42 H62 O16	-	821.3965	-	821.3965	821.3963	0.3	-	-
80	ベンザミナタハコ における発現	BAS-CYP72A154+C YP88D6+GlcSL+UG T73P12	13.4	glycyrrizin isomer 2 (S)	GMA-GlcA-GlcA	C42 H62 O16	-	821.3965	-	821.3965	821.3960	0.6	-	-
81														
82	ベンザミナタハコ における発現	SOAP1-4+GlcSL	14.2	medicagenic acid 3- O-glucuronides (S)	MA+GlcA	C36 H54 O12	-	677.3543	-	677.3543	677.3535	1.0	-	-

【 0 9 8 2 】

10

20

30

40

【表 1 2 - 3】

Section B

A	B	P	Q	R	S	T
		MS/MS ES(+): フラガメント	MS/MS ES(+): CE (eV)	UV/Vis	他の種で検出	リアレンジ
1	ホウレンソウ	959.4370 [M+pent-H] <sup>+</sup> ; 955.4099 [M+hex-H] <sup>+</sup> ; 793.4330 [M+hexA-pent-H] <sup>+</sup> ; 501.3204 [M+hexA-pent-dhex-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
2	ホウレンソウ	955.4599 [M+hexA-H] <sup>+</sup> ; 823.4120 [M+hex-dhex-H] <sup>+</sup> ; 677.3559 [M+hex-dhex-H] <sup>+</sup> ; 501.3264 [M+hex-dhex-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
3	ホウレンソウ	1011.4806 [M+pent-H] <sup>+</sup> ; 855.4438 [M+pent-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
4	ホウレンソウ	925.4760 [M+hexA-pent-H] <sup>+</sup> ; 501.3236 [M+hexA-pent-dhex-H] <sup>+</sup> ;	70	-	-	-
5	ホウレンソウ	955.4693 [M+hexA-pent-H] <sup>+</sup> ; 809.3924 [M+hexA-pent-hexA-H] <sup>+</sup> ; 823.4101 [M+pent-pent-dhex-H] <sup>+</sup> ;	70	-	-	-
6	ホウレンソウ	1143.5160 [M+pent-H] <sup>+</sup> ; 1011.4427 [M+pent-pent-H] <sup>+</sup> ; 967.4788 [M+pent-hexA-H] <sup>+</sup> ; 823.4101 [M+pent-pent-dhex-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
7	ホウレンソウ	997.4946 [M+hexA-pent-H] <sup>+</sup> ; 955.4602 [M+hexA-pent-hexA-H] <sup>+</sup> ; 501.3232 [M+hexA-pent-dhex-H] <sup>+</sup> ;	70	-	-	-
8	ホウレンソウ	997.4946 [M+hexA-pent-H] <sup>+</sup> ; 955.4678 [M+hexA-pent-hexA-H] <sup>+</sup> ; 501.3210 [M+hexA-pent-dhex-H] <sup>+</sup> ;	70	-	-	-
9	ホウレンソウ	1067.5228 [M+hexA-pent-H] <sup>+</sup> ; 925.4878 [M+hexA-pent-pent-H] <sup>+</sup> ; 647.3782 [M+hexA-pent-pent-dhex-H] <sup>+</sup> ;	70	-	-	-
10	ホウレンソウ	1263.5773 [M+pent-H] <sup>+</sup> ; 1087.5523 [M+hexA-pent-H] <sup>+</sup> ; 955.4677 [M+hexA-pent-pent-H] <sup>+</sup> ; 909.3959 [M+hexA-pent-pent-dhex-H] <sup>+</sup> ;	70	-	-	-
11	ホウレンソウ	955.4677 [M+hexA-pent-H] <sup>+</sup> ;	70	-	-	-
12	ホウレンソウ	1099.5310 [M+hexA-pent-H] <sup>+</sup> ; 1057.5170 [M+hexA-pent-hexA-H] <sup>+</sup> ; 501.3190 [M+hexA-pent-pent-dhex-H] <sup>+</sup> ;	70	-	-	-
13	ホウレンソウ	1129.5430 [M+pent-hexA-H] <sup>+</sup> ; 1087.5220 [M+pent-hexA-H] <sup>+</sup> ; 501.3220 [M+hexA-pent-pent-hex-dhex-H] <sup>+</sup> ;	70	-	-	-
14	ホウレンソウ	847.3809 [M+hex-H] <sup>+</sup> ; 603.3903 [M+hex-CO2-H] <sup>+</sup> ; 585.3865 [M+hex-CO2-H2O-H] <sup>+</sup> ; 471.3473 [M+hex-hexA-H] <sup>+</sup> ;	60	-	-	-
15	ホウレンソウ	663.3741 [M+hex-H] <sup>+</sup> ; 619.3843 [M+hex-CO2-H] <sup>+</sup> ; 601.3750 [M+hex-CO2-H2O-H] <sup>+</sup> ; 487.3431 [M+pent-hex-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
16	ホウレンソウ	809.4302 [M+pent-H] <sup>+</sup> ; 647.3781 [M+pent-hex-H] <sup>+</sup> ; 603.4377 [M+pent-hex-CO2-H] <sup>+</sup> ; 585.3978 [M+pent-hex-CO2-H2O-H] <sup>+</sup> ; 471.3474 [M+pent-hex-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
17	ホウレンソウ	809.4344 [M+oxypyruvic and glycolic acid-H] <sup>+</sup> ; 647.3788 [M+hex-oxypyruvic and glycolic acid-H] <sup>+</sup> ; 585.3795 [M+hex-oxypyruvic and glycolic acid-CO2-H2O-H] <sup>+</sup> ; 471.3475 [M+hex-oxypyruvic and glycolic acid-hexA-H] <sup>+</sup> ;	60	-	-	-
18	ホウレンソウ	809.4326 [M+oxypyruvic and glycolic acid-H] <sup>+</sup> ; 647.3788 [M+hex-oxypyruvic and glycolic acid-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
19	ホウレンソウ	809.4267 [M+hex-H] <sup>+</sup> ; 417.3479 [M+hex-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
20	ホウレンソウ	911.4585 [M+hex-H] <sup>+</sup> ; 647.3839 [M+hex-pent-pent-H] <sup>+</sup> ; 471.3416 [M+hex-pent-pent-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
21	ホウレンソウ	955.4252 [M+pent-H] <sup>+</sup> ; 925.4351 [M+hex-H] <sup>+</sup> ; 617.3639 [M+hex-H] <sup>+</sup> ; 485.3267 [M+hex-pent-pent-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
22	ホウレンソウ	1143.5077 [M+pent-H] <sup>+</sup> ; 1011.4468 [M+pent-pent-H] <sup>+</sup> ; 677.3535 [M+pent-pent-dhex-H] <sup>+</sup> ; 501.3269 [M+pent-pent-dhex-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
23	ホウレンソウ	799.4954 [M+oxypyruvic and glycolic acid-H] <sup>+</sup> ; 631.3893 [M+oxypyruvic and glycolic acid-hexA-H] <sup>+</sup> ;	60	-	ピーチル/ツルムラサキ	(54)
24	ホウレンソウ	H2O-H] <sup>+</sup> ; 455.3514 [M+oxypyruvic and glycolic acid-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	ピーチル/ツルムラサキ	(54)
25	ホウレンソウ	799.4974 [M+oxypyruvic and glycolic acid-H] <sup>+</sup> ; 631.3844 [M+oxypyruvic and glycolic acid-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	ツルムラサキ	(55)
26	ホウレンソウ	807.4161 [M+oxypyruvic and glycolic acid-H] <sup>+</sup> ; 645.3807 [M+oxypyruvic and glycolic acid-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
27	ホウレンソウ	H2O-H] <sup>+</sup> ; 459.3307 [M+oxypyruvic and glycolic acid-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
28	ホウレンソウ	1115.5173 [M+pent-H] <sup>+</sup> ; 969.4326 [M+hex-pent-H] <sup>+</sup> ; 939.4197 [M+hex-dhex-H] <sup>+</sup> ; 823.4352 [M+hex-pent-dhex-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
29	ホウレンソウ	959.4322 [M+hex-pent-hexA-H] <sup>+</sup> ; 939.4189 [M+hex-dhex-hexA-H] <sup>+</sup> ; 823.4349 [M+hex-pent-dhex-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
30	ホウレンソウ	955.4204 [M+CO2-H] <sup>+</sup> ; 897.4118 [M+CO2-C2H2O2-H] <sup>+</sup> ; 839.4088 [M+CO2-C2H2O2-H] <sup>+</sup> ; 677.3566 [M+oxypyruvic and glycolic acid-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
31	ホウレンソウ	955.4320 [M+CO2-H] <sup>+</sup> ; 897.4130 [M+CO2-C2H2O2-H] <sup>+</sup> ; 839.4070 [M+CO2-C2H2O2-H] <sup>+</sup> ; 677.3503 [M+oxypyruvic and glycolic acid-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
32	ホウレンソウ	939.4240 [M+CO2-H] <sup>+</sup> ; 881.4178 [M+CO2-C2H2O2-H] <sup>+</sup> ; 823.4127 [M+CO2-C2H2O2-H] <sup>+</sup> ; 677.3565 [M+oxypyruvic and glycolic acid-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
33	ホウレンソウ	647.3743 [M+hexA-pent-H] <sup>+</sup> ; 501.3240 [M+hexA-pent-dhex-H] <sup>+</sup> ;	50	-	ピーチル	(56)
34	ホウレンソウ	663.3754 [M+hex-H] <sup>+</sup> ; 617.3639 [M+hex-hexA-H] <sup>+</sup> ;	50	-	-	-
35	ホウレンソウ	823.4380 [M+hex-H] <sup>+</sup> ; 677.3533 [M+hex-hexA-H] <sup>+</sup> ; 501.3225 [M+hex-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
36	ホウレンソウ	799.4944 [M+hex-H] <sup>+</sup> ; 651.3826 [M+pent-hex-H] <sup>+</sup> ; 599.3984 [M+pent-hex-CO2-H2O-H] <sup>+</sup> ; 455.3519 [M+pent-hex-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
37	ホウレンソウ	647.3774 [M+hex-H] <sup>+</sup> ; 471.3475 [M+hex-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
38	ホウレンソウ	645.3636 [M+hex-H] <sup>+</sup> ; 469.3323 [M+hex-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
39	ホウレンソウ	661.3593 [M+hex-H] <sup>+</sup> ; 647.3801 [M+hexA-H] <sup>+</sup> ; 617.3719 [M+hex-CO2-H] <sup>+</sup> ; 599.3574 [M+hex-CO2-H2O-H] <sup>+</sup> ; 485.3271 [M+hex-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
40	ホウレンソウ	1117.5071 [M+pent-H] <sup>+</sup> ; 985.4347 [M+pent-pent-H] <sup>+</sup> ; 809.4266 [M+pent-hexA-H] <sup>+</sup> ; 501.3194 [M+hexA-hex-pent-pent-dhex-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
41	ホウレンソウ	995.4606 [M+hex-H] <sup>+</sup> ; 823.3881 [M+hex-H] <sup>+</sup> ; 617.3605 [M+hex-dhex-H] <sup>+</sup> ; 501.3212 [M+hex-dhex-hexA-H] <sup>+</sup> ;	70	-	-	-
42	ホウレンソウ	995.4552 [M+pent-H] <sup>+</sup> ; 955.4471 [M+hex-H] <sup>+</sup> ; 677.3557 [M+pent-hex-dhex-H] <sup>+</sup> ; 501.3250 [M+pent-hex-dhex-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
43	ホウレンソウ	677.3486 [M+hex-H] <sup>+</sup> ; 501.3224 [M+hex-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
44	ホウレンソウ	501.3207 [M+hexA-pent-H] <sup>+</sup> ; 325.0773 [M+hex-pent-H2O-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
45	ホウレンソウ	661.3582 [M+hex-H] <sup>+</sup> ; 647.3805 [M+hexA-H] <sup>+</sup> ; 617.3690 [M+hex-CO2-H] <sup>+</sup> ; 599.3590 [M+hex-CO2-H2O-H] <sup>+</sup> ; 485.3292 [M+hex-hexA-H] <sup>+</sup> ;	50	-	-	-
46	ホウレンソウ	925.4778 [M+hexA-pent-H] <sup>+</sup> ; 501.3233 [M+hexA-pent-dhex-hexA-H] <sup>+</sup> ;	70	-	-	-
47	ホウレンソウ	1143.5116 [M+pent-H] <sup>+</sup> ; 1101.5002 [M+pent-hexA-H] <sup>+</sup> ; 967.4825 [M+pent-hexA-H] <sup>+</sup> ; 939.4345 [M+pent-hexA-pent-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
48	ホウレンソウ	501.3212 [M+hex-hexA-pent-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
49	ホウレンソウ	663.3759 [M+hex-H] <sup>+</sup> ; 501.3222 [M+hex-hexA-H] <sup>+</sup> ;	ramp 1.5-50 eV	-	-	-
50	ホウレンソウ			-	-	-
51	アルブアルブ	825.4302 [M+hex-H] <sup>+</sup> ; 663.3773 [M+hex-hex-H] <sup>+</sup> ; 487.3424 [M+hex-hexA-H] <sup>+</sup> ;	45	-	タルウムコヤン	(Main text reference 14)
52	アルブアルブ	663.3765 [M+hex-H] <sup>+</sup> ; 487.3432 [M+hex-hexA-H] <sup>+</sup> ;	45	-	タルウムコヤン	(Main text reference 14)
53	アルブアルブ	939.4989 [M+H2O-H] <sup>+</sup> ; 895.5098 [M+CO2-H2O-H] <sup>+</sup> ; 811.4192 [M+hex-H] <sup>+</sup> ; 767.4695 [M+hex-CO2-H2O-H] <sup>+</sup> ; 651.3887 [M+hex-hex-H2O-H] <sup>+</sup> ; 613.3731 [M+hex-hex-H2O-H2O-H] <sup>+</sup> ; 473.3655 [M+hex-dhex-hexA-H] <sup>+</sup> ;	45	-	タルウムコヤン	(Main text reference 14)
54	アルブアルブ	959.3806 [M+malonyl-hex-CO2-H2O-H] <sup>+</sup> ; 823.4179 [M+malonyl-H] <sup>+</sup> ; 805.3989 [M+malonyl-H2O-H] <sup>+</sup> ; 703.3777 [M+hex-CO2-H] <sup>+</sup> ; 643.3525 [M+malonyl-hex-H2O-H] <sup>+</sup> ;	45	-	タルウムコヤン	(Main text reference 14)

【 0 9 8 3 】

【表 1 2 - 4】

A	B	P	Q	R	S	T
	試料	MS MS ES(-) 試料	MS MS ES(-) CE (eV)	UV/Vis	他の種で検出	リファレンス
1						
55	アルファルプ	953,4714 [M+H2O-H] <sup>+</sup> ; 909,4915 [M-CO2-H2O-H] <sup>+</sup> ; 825,4289 [M+hex-H] <sup>+</sup> ; 763,4302 [M+hex-CO2-H2O-H] <sup>+</sup> ; 645,3669 [M+hex-H2O-H] <sup>+</sup> ; 601,3768 [M+hex-Hex+H2O-CO2-H] <sup>+</sup> ; 555,3728 [M+hex-Hex+H2O-H] <sup>+</sup> ; 487,3447 [M+hex-Hex-H2O-H] <sup>+</sup> ; 1089,5315 [M+H2O-H] <sup>+</sup> ; 955,4952 [M+hex-H] <sup>+</sup> ; 863,5005 [M+hex-CO2-H2O-H] <sup>+</sup> ; 761,4280 [M+hex-Hex+H2O-H] <sup>+</sup> ; 747,4388 [M+hex-Hex-CO2-H2O-H] <sup>+</sup> ; 728,4260 [M+hex-Hex+H2O-H2O-CO2-H] <sup>+</sup> ; 629,3726 [M+hex-Hex+H2O-H] <sup>+</sup> ; 685,3828 [M+hex-Hex+CO2-H] <sup>+</sup> ; 539,3792 [M+hex-Hex-H2O-H] <sup>+</sup> ; 471,3504 [M+hex-Hex+H2O-H] <sup>+</sup>	90	-	タルウマコヤシ	(Main text reference 14)
56	アルファルプ	471,3504 [M+hex-Hex+H2O-H] <sup>+</sup>	80	-	タルウマコヤシ	(Main text reference 14)
57	アルファルプ	661,3585 [M+hex-H] <sup>+</sup> ; 617,3689 [M+hex-CO2-H] <sup>+</sup> ; 599,3587 [M+hex-CO2-H2O-H] <sup>+</sup> ; 485,3268 [M+hex-Hex+H] <sup>+</sup> ; 1085,5536 [M+H2O-H] <sup>+</sup> ; 1041,5638 [M+H2O-CO2-H] <sup>+</sup> ; 957,5061 [M+hex-H] <sup>+</sup> ; 895,5071 [M+hex+H2O-CO2-H] <sup>+</sup> ; 795,4518 [M+hex-Hex+H] <sup>+</sup> ; 777,4427 [M+hex-Hex+H2O-H] <sup>+</sup> ; 759,4297 [M+hex-Hex-H2O-H] <sup>+</sup> ; 457,3981 [M+hex-Hex-H2O-H] <sup>+</sup>	45	-	タルウマコヤシ	(Main text reference 14)
58	アルファルプ	1085,5536 [M+H2O-H] <sup>+</sup> ; 759,4297 [M+hex-Hex-H2O-H] <sup>+</sup> ; 457,3981 [M+hex-Hex-H2O-H] <sup>+</sup>	45	-	タルウマコヤシ	(Main text reference 14)
59						
60	ビートル	997,5342 [M-C3H4O5-H] <sup>+</sup> ; 955,5225 [M+Ac-H] <sup>+</sup> ; 835,4744 [M-C3H4O5-Hex-H] <sup>+</sup> ; 793,4632 [M+Ac-Hex-H] <sup>+</sup> ; 731,4694 [M+Ac-Hex-CO2-H2O-H] <sup>+</sup> ; 631,4028 [M+Ac-Hex+H] <sup>+</sup> ; 613,3919 [M+Ac-Hex-Hex+H2O-H] <sup>+</sup> ; 485,3650 [M+Ac-Hex-Hex+H2O-H] <sup>+</sup>	60	-	ビートル	(57)
61	ビートル	953,5191 [M+pent-H] <sup>+</sup> ; 925,5099 [M+hex-H] <sup>+</sup> ; 793,4635 [M+hex-pent-H] <sup>+</sup> ; 731,4694 [M+hex-CO2-pent-H2O-H] <sup>+</sup> ; 613,3912 [M+hex-pent-Hex+H2O-H] <sup>+</sup> ; 558,4006 [M+hex-pent-Hex-H2O-CO2-H] <sup>+</sup> ; 551,3894 [M+hex-pent-Hex-H2O-CO2-H2O-H] <sup>+</sup> ; 455,3650 [M+hex-pent-Hex-Hex+H] <sup>+</sup>	60	-	ビートル	(57)
62	ビートル	997,5213 [M-C3H4O5-H] <sup>+</sup> ; 955,5092 [M+Ac-H] <sup>+</sup> ; 805,4618 [M-C3H4O5-Hex-H] <sup>+</sup> ; 763,4607 [M+hex-Ac+H] <sup>+</sup> ; 743,4690 [M-C3H4O5-Hex+H2O-CO2-H] <sup>+</sup> ; 593,3999 [M-C3H4O5-Hex+H2O-CO2-pent-H2O-H] <sup>+</sup> ; 455,3639 [M+Ac-Hex-pent-Hex+H] <sup>+</sup> ; 763,4516 [M+hex-H] <sup>+</sup> ; 701,4487 [M+hex-CO2-H2O-H] <sup>+</sup> ; 631,4028 [M+hex-pent-H] <sup>+</sup> ; 569,4016 [M+hex-CO2-H2O-pent-H] <sup>+</sup> ; 551,3904 [M+hex-CO2-H2O-pent-H2O-H] <sup>+</sup> ; 455,3654 [M+hex-pent-Hex+H] <sup>+</sup>	80	-	ビートル	(57)
63	ビートル	833,4793 [M-C3H4O5-H] <sup>+</sup> ; 793,4630 [M+Ac+H] <sup>+</sup> ; 673,4187 [M-C3H4O5-Hex+H] <sup>+</sup> ; 631,4045 [M+Ac-Hex+H] <sup>+</sup> ; 569,4015 [M+Ac-Hex-CO2-H2O-H] <sup>+</sup> ; 455,3655 [M+hex-Hex+Ac+H] <sup>+</sup>	80	-	ビートル	(57)
64	ビートル	455,3655 [M+hex-Hex+Ac+H] <sup>+</sup>	80	-	ビートル	(57)
65	ペンシツカハに おける処理					
66	ペンシツカハに おける処理	501,3206 [M-GlcA-H] <sup>+</sup> ; 483,3096 [M-GlcA+H2O-H] <sup>+</sup> ; 439,3184 [M-GlcA+H2O-CO2-H] <sup>+</sup> ; 193,0346 [GlcA-H] <sup>+</sup> ; 175,0238 [GlcA+H2O-H] <sup>+</sup>	45	-		
67	ペンシツカハに おける処理	647,3776 [M+hexA-H] <sup>+</sup> ; 501,3237 [M+hexA-H] <sup>+</sup>	45	-		
68	ペンシツカハに おける処理	793,4233 [M+hexA-H] <sup>+</sup> ; 501,3219 [M+hexA-Hex+H] <sup>+</sup> ; 409,3182 [M+hexA-Hex+H2O-H] <sup>+</sup>	50	-		
69	ペンシツカハに おける処理	955,4602 [M+hexA-H] <sup>+</sup> ; 677,3536 [M+hex-Hex+H] <sup>+</sup> ; 501,3195 [M+hexA-Hex+H2O-H] <sup>+</sup>	55	-		
70	ペンシツカハに おける処理	955,4648 [M+hexA-pent-H] <sup>+</sup> ; 793,4109 [M+hexA-pent-Hex+H] <sup>+</sup> ; 501,3018 [M+hexA-pent-Hex+H2O-H] <sup>+</sup> ; 483,3283 [M+hexA-pent-Hex-Hex+H] <sup>+</sup>	70	-		
71	ペンシツカハに おける処理	997,4995 [M+hexA-pent-H] <sup>+</sup> ; 955,4683 [M+hexA-pent-Hex+H] <sup>+</sup> ; 501,3227 [M+hexA-pent-Hex+H2O-H] <sup>+</sup> ; 439,3179 [M+hexA-pent-Hex-Hex+H] <sup>+</sup>	70	-		
72						
73	ペンシツカハに おける処理	883,3628 [M-CO2-H2O-H] <sup>+</sup> ; 569,3469 [M-C2O3H4-H] <sup>+</sup> ; 523,3415 [M-C3O4H4-H2O-H] <sup>+</sup> ; 469,3315 [M-GlcA-H] <sup>+</sup> ; 455,3419 [M-GlcA-CO2-H] <sup>+</sup>	55	-	ウラルカンゾウ	(58)
74	ペンシツカハに おける処理	883,3635 [M-CO2-H2O-H] <sup>+</sup> ; 569,3470 [M-C2O3H4-H] <sup>+</sup> ; 523,3423 [M-C3O4H4-H2O-H] <sup>+</sup> ; 469,3318 [M-GlcA-H] <sup>+</sup> ; 455,3419 [M-GlcA-CO2-H] <sup>+</sup>	55	-	ウラルカンゾウ	(58)
75						
76	ペンシツカハに おける処理	883,3631 [M-CO2-H2O-H] <sup>+</sup> ; 569,3468 [M-C2O3H4-H] <sup>+</sup> ; 523,3418 [M-C3O4H4-H2O-H] <sup>+</sup> ; 469,3317 [M-GlcA-H] <sup>+</sup> ; 455,3422 [M-GlcA-CO2-H] <sup>+</sup>	55	-	ウラルカンゾウ	(58)
77	ペンシツカハに おける処理	883,3633 [M-CO2-H2O-H] <sup>+</sup> ; 569,3470 [M-C2O3H4-H] <sup>+</sup> ; 523,3421 [M-C3O4H4-H2O-H] <sup>+</sup> ; 469,3320 [M-GlcA-H] <sup>+</sup> ; 455,3415 [M-GlcA-CO2-H] <sup>+</sup>	55	-	ウラルカンゾウ	(58)
78						
79	ペンシツカハに おける処理	645,3636 [M-GlcA-H] <sup>+</sup> ; 469,3303 [M-GlcA-GlcA-H] <sup>+</sup> ; 351,0574 [GlcA-H] <sup>+</sup> ; 193,0357 [GlcA-H] <sup>+</sup> ; 175,0238 [GlcA+H2O-H] <sup>+</sup>	65	-	ウラルカンゾウ	(58)
80	ペンシツカハに おける処理	645,3615 [M-GlcA-H] <sup>+</sup> ; 469,3323 [M-GlcA-GlcA-H] <sup>+</sup> ; 351,0577 [GlcA-H] <sup>+</sup> ; 193,0354 [GlcA-H] <sup>+</sup> ; 175,0250 [GlcA+H2O-H] <sup>+</sup>	65	-	ウラルカンゾウ	(58)
81						
82	ペンシツカハに おける処理	501,3212 [M-GlcA-H] <sup>+</sup> ; 483,3084 [M-GlcA+H2O-H] <sup>+</sup> ; 439,3176 [M-GlcA+H2O-CO2-H] <sup>+</sup> ; 193,0334 [GlcA-H] <sup>+</sup> ; 175,0342 [GlcA+H2O-H] <sup>+</sup>	45	-		

10

20

30

40

【 0 9 8 4 】

表 1 3 : ホウレンソウ中のサポニン組成の要約

【 0 9 8 5 】

【表 1 3】

No.	名称	分子式	m/z 実測値 [M-H]	m/z 算出値 [M-H]	$\Delta m/z$ ppm
1	Yossoside XI	C84H100O33	1395.6068	1395.6069	-0.1
2	Yossoside VI	C47H88O30	1131.5261	1131.5282	-1.9
3	Yossoside IV	C52H98O34	1263.5713	1263.5705	0.6
4	Yossoside VIII	C63H98O32	1365.5981	1365.5963	1.3
5	Yossoside XII	C58H90O28	1233.5554	1233.5540	1.1
6	Yossoside Va	C61H94O30	1305.5773	1305.5752	1.6
7	Yossoside X	C66H102O34	1437.6160	1437.6151	0.6
8	Yossoside VII	C60H92O29	1275.5675	1275.5646	2.3
9	Yossoside V	C61H94O30	1305.5781	1305.5752	2.2
10	Yossoside IX	C85H100O33	1407.6107	1407.6069	2.7
11	Yossoside VIIa	C60H92O29	1275.5640	1275.5646	-0.5

10

## 【0986】

20

質量分析法フラグメンテーション分析並びに一連の1D及び2D-NMR試験によって、最も豊富なハウレンソウサポニン（ここではヨッソシドVと呼ぶ）の構造が確立した。このサポニンは、C-3位のヒドロキシに結合したグルクロン酸及びキシロースと、C-28のカルボキシルに結合したアセチルフコース、ラムノース及びグルコースとを有するメディカゲン酸（アグリコン）を含む（図20A~20C、図21）。

## 【0987】

要約：ハウレンソウにおいて11種のトリテルペノイド系サポニンが同定され、ヨッソシドVが最も豊富であった。

## 【0988】

実施例16：ハウレンソウトランスクリプトームデータを用いたサポニン生合成遺伝子の同定

30

## 【0989】

目的：トリテルペノイド系サポニン生合成経路におけるトリテルペノイド系アグリコン生合成遺伝子を同定すること。

## 【0990】

方法：上記の材料及び方法を参照されたい。

## 【0991】

結果：異なる含有量のサポニンを発現する5種のハウレンソウ組織からトランスクリプトームデータを作成し、高サポニン含有試料及び低サポニン含有試料間で異なる発現量、並びに既知のトリテルペノイド系アグリコン生合成遺伝子（「E. Biazzi et al., CYP72A67 Catalyzes a Key Oxidative Step in Medicago truncatula Hemolytic Saponin Biosynthesis. Molecular Plant. 8, 1493-1506 (2015)」）に対する相同性に基づき、遺伝子候補を選択した。

40

## 【0992】

候補のうち、サポニン - アミリン合成酵素（SOBAS; SOAP1と呼ぶ; Sp\_107620\_kpnh; 配列番号45 [遺伝子配列]、配列番号48 [ポリペプチド配列]）、シトクロムP450 CYP716A268 (SOAP2; Sp\_107660\_kiqg; 配列番号46 [遺伝子配列]、配列番号49 [ポリペプチド配列])、及びCYP716A268v2 (SOAP2様; Sp\_107670\_ptqx; 配列番号47 [遺伝子配列]、配列番号6 [ポリペプチド配列])を含む3つの遺伝子位置を調べた

50

ところ、代謝遺伝子クラスターの形態における、それらの密接な物理的遺伝子位置が明らかとなった（図 2 2 A 及び 2 2 B）。

【0993】

ハウレンソウにおける SOAP 1 及び SOAP 2 のウイルス誘導遺伝子サイレンシング（VIGS）アッセイの結果、サポニン含有量が低下し（図 2 3 A ~ 2 3 D）、その一方でペンサミアナタバコにおいては、それらを過剰発現させたところ、 $\alpha$ -アミリン及びオレアノール酸が産生した（図 2 4 A 及び 2 4 B）。

【0994】

概要：これらの結果は、これらの遺伝子がハウレンソウサポニン生合成に関与していることを示す有力な証拠をもたらすものである。

10

【0995】

実施例 17：さらなるサポニン生合成遺伝子の同定のための共発現解析及びコードされた酵素の特性評価

【0996】

目的：トリテルペノイド系サポニン生合成経路において、さらなるトリテルペノイド系アグリコン生合成遺伝子を同定すること。

【0997】

方法：上記の材料及び方法を参照されたい。

【0998】

結果：SOAP 1、SOAP 2 及び SOAP 2 様ヌクレオチド（配列番号 45 ~ 47 のヌクレオチド配列が発現された）を、後に共発現解析（ $r > 0.9$ ；ピアソン相関係数 - PCC）においてベイトとして用い、VIGS を用いた機能特性評価のため、さらに、p 450 シトクロムをコードする 5 つの遺伝子、グリコシルトランスフェラーゼをコードする 8 つの遺伝子、及びアシルトランスフェラーゼをコードする 5 つの遺伝子を同定した（図 2 0 A 及び 2 0 B；表 1 4）。SOAP 1、Soap 2、及び SOAP 2 様ヌクレオチドは、それぞれ配列番号 48、配列番号 49、及び配列番号 50 のポリペプチド配列をコードする。これらの試験により、CYP72A655（SOAP 3；Sp\_\_085340\_\_mek；配列番号 51 [ヌクレオチド]、配列番号 52 [ポリペプチド]）及び CYP72A654（SOAP 4；Sp\_\_040350\_\_wdny；配列番号 53 [ヌクレオチド]、配列番号 54 [ポリペプチド]）が、ヨソシドアグリコンの生合成に関与していることが明らかとなった（図 2 5）。以下の表 1 4 において、用いられる用語には以下のものがある：共発現した遺伝子（coexpressed gene）- 共発現相関係数が 0.9 超である遺伝子；アノテーション（annotation）- シロイヌナズナ属プロテオーム（TAIR 10）に対するブラスト解析に基づく遺伝子機能予測；及びアミノ酸配列（aa seq）- 共発現した遺伝子のアミノ酸配列

20

30

【0999】

表 1 4：3 種のベイト SOAP 1、SOAP 2 及び SOAP 2 様によるハウレンソウ遺伝子の共発現解析

【1000】

【表 1 4 - 1】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
3	sp_000460_nyyz	0.96371	0.9376205	Pentatricopeptide repeat (PPR) superfamily protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion; CONTAINS InterPro DOMAIN/s: Pentatricopeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: Tetratricopeptide repeat (TPR)-like superfamily protein (TAIR:AT5G0110.1); Has 32626 Blast hits to 11011 proteins in 270 species: Archae - 3; Bacteria - 29; Metazoa - 476; Fungi - 329; Plants - 30874; Viruses - 0; Other Eukaryotes - 915 (source: NCBI BLINK).	MVSFLRHSIPAVTNLFLSIGKPTKRLLSHRLVADD PKHTVASSVCDLSLRKGLNWDTLRSRYGVSVDLDD LIERVLLFEKPEIEAKRALSFHWAHNNKLGHG LKHYCIIIQILARARLVNDARSLLLESILMRGENEV NCSLRFVVDMLLNTYKVTNSCSFVFDLMVQSY SKLRMYSAAFVCCYLDERGFSLSLVTFNTLLHVI QKSPDSDNLWEVYGHMIOKRINPNDATVKIMINA LKAGELOKYMNIGRVRHGKGCSPMVTWNTSLVF RMLGEGRIEKGVLVLRMLQNLKLLDDVSYSLV YARLRSGEINLAREVYEGMVLSSGFLPNSFVYTLF IEGYCEKGMVEEALNLMDEMVMNGLKPYDDTYN SILVACARIGASVDSFRLCEEEMTEKGLLPSTSAFN EMVKGKLCESGNMKQSNMDMLTYLMEKGYLPDET TELHLIDGYGREKNIQEALKLYEMEFRRSISPLPI LTLSLIRTLCKCRKLEEAADKYLAIMKRRSLSLTVCIY KPLISAHLENGNKERAHHLYNEMIQRGVKNQOPM PPDFSEDLDS
4	sp_000760_grtg	0.9403550	0.9993832	HTR12; FUNCTIONS IN: DNA binding; INVOLVED IN: double fertilization forming a zygote and endosperm; LOCATED IN: chromosome, centromeric region, nucleus; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAIN/s: Histone H3 (InterPro:IPR000164), Histone-fold (InterPro:IPR009072), Histone core (InterPro:IPR007125); BEST Arabidopsis thaliana protein match is: male-gamete-specific histone H3 (TAIR:AT1G19890.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	MRTKHPAAKKSSSKSPRSKAPSTPSKSNQSKST PIRSPKSKSGRKSVAAPRIKKPYRHRPGVTALREI RKYOKSVDLLIPSAPFVRTVKEISHQFSPMVTWR KVEAMKALQQAEDFVCLFEDAMLCAFHAKRVT LMKKDIELARRIGGERRR

10

20

30

40

【表 1 4 - 2】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_ktgg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
5	sp_000910_wmxj	0.985934	0.983733	gamma-glutamyl hydrolase 2 (GGH2); FUNCTIONS IN: omega peptidase activity; INVOLVED IN: tetrahydrofolylpolyglutamate metabolic process; LOCATED IN: vacuole; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Peptidase C26, gamma-glutamyl hydrolase (InterPro:IPR015527), Peptidase C26 (InterPro:IPR011697); BEST Arabidopsis thaliana protein match is: gamma-glutamyl hydrolase 1 (TAIR:AT1G78660.2); Has 406 Blast hits to 402 proteins in 115 species: Archae - 3; Bacteria - 51; Metazoa - 191; Fungi - 0; Plants - 77; Viruses - 0; Other Eukaryotes - 84 (source: NCBI BLink).	MFPSNTPSSTSHLSNLLWFFLWIPLITSLSLSQL TSAEPNNMIIMLTSSDYLPARKSCVPDPDLNFR PVIGVSHPGDGGASGRKNGTNSVYIAASVYKVF ESAGARVPLVYTDPIDVFKKLELVNGVLFVGGW AKNGTYFEVWQIFKVKLEKNDAGKFFPCYAICL GFELLTMVSKDNINLENFSASYQASTLQFVDNIV SGTVFORPEPELLRKLSTDCLVQNHKFAISPQT LOANHLLSSFFDVIITSTDSDNKVVYVSTVQAHQY PITGFQWHPEKNAFEWGYPMIPHSEDAIQVTOH VANFFISEARRSPNRPVVDVNLNLLNLYNSPTYC GKKNGYDEVIYIFG
6	sp_002040_oxuo	0.973201	0.973201	purple acid phosphatase 29 (PAP29); FUNCTIONS IN: protein serine/threonine phosphatase activity, acid phosphatase activity; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Metallophosphoesterase (InterPro:IPR004843), Phosphoesterase A12946880 (InterPro:IPR011230); BEST Arabidopsis thaliana protein match is: purple acid phosphatase 14 (TAIR:AT2G46880.1); Has 636 Blast hits to 631 proteins in 208 species: Archae - 0; Bacteria - 277; Metazoa - 0; Fungi - 208; Plants - 117; Viruses - 0; Other Eukaryotes - 34 (source: NCBI BLink). DNVVRNY-binding protein Kin17, conserved region; FUNCTIONS IN: zinc ion binding; INVOLVED IN: biological_process unknown; LOCATED IN: intracellular; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINs: Zinc finger, C2H2-like (InterPro:IPR015880), DNA/RNA-binding protein Kin17, conserved region (InterPro:IPR019447), KOW (InterPro:IPR005824); BEST Arabidopsis thaliana protein match is: DNA/RNA-binding protein Kin17, conserved region (TAIR:AT5G51795.1); Has 1885 Blast hits to 1561 proteins in 251 species: Archae - 4; Bacteria - 25; Metazoa - 715; Fungi - 242; Plants - 160; Viruses - 13; Other Eukaryotes - 726 (source: NCBI BLink).	MHYANGKOTPCSDVLPQLQTCSDLNITTDYLRR FIQENPDLVFTGDNVHSTDAYNFRSMDEVFA PAIESNIPWAILGNHDVNVLQIGIPRQAVMEYIV GSKPFENKSVLNLYLDDTGDYTYSTFTFYLGYE WICPQQSWFLRTSKQLQETYSAPQPQPQQG PAPGLVFFHIPIPEFNNISKSDMVGVKQENVICPG VNSGFFDTMLAAGDIKAAAFVGHGHTNDYCGKLC DIQLCYDAGFGYHGYGIAGWPRRARVINVKLDT QDGNWGPVNSIVTWKRLHDDRLTIIDPEVLWTS AGRDOSKSVFKPIMSE
7	sp_005360_psus	0.963032	0.963032		MEGKSGLKALSNRMKSKGLQKLMHYCEMCKQK CRDANGFKGCHCMSDGHLRQMQIFGENPKKFIEL FSGEFEDAFLELMRRSHRFSRVAATVYNEYISD RTHVHMNSTKWLTLTEFVMIHLGRTGKCRVDHTE KGLFLYSARSENLFYERRKNRKLKSLVAEEVK QENEIVKIQIRAAASVAVPETEVRTPLEIQINNI MLCSEPEMGNKKISFAFGFGSSSTSVTSLKPRE DSSMSFGLLEDVPAELGLSKKKTIGDVGDSGGC VS

10

20

30

40

【表 1 4 - 3】

	A	B	C	D	E	F
1						
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
8	sp_008840_kejm	3.98035	3.981213	3.97675	receptor-like protein kinase 1 (RLK1); FUNCTIONS IN: in 6 functions; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Curculin-like (mannose-binding) lectin (InterPro:IPR001480), Protein Kinase, ATP binding site (InterPro:IPR017441), Serine/threonine-protein kinase domain (InterPro:IPR002290), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Serine/threonine-protein Kinase, active site (InterPro:IPR008271), Protein kinase-like domain (InterPro:IPR011009), Protein Kinase, catalytic domain (InterPro:IPR000719), Tyrosine-protein kinase, catalytic domain (InterPro:IPR020635); BEST Arabidopsis thaliana protein match is: lectin protein kinase family protein (TAIR:AT1G34300.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).	MASPPSCILFSFFTLFLLOSIAAQQGGQITVGSLS TAINGSLLLSPSGDFAFGFHQLPNKKNLFLAIW YAKIPDTIWHANDGNPVPQENICMTDQNPPEL GSGICGFNSICSLGDASDDRRPTCTCPGYYSFR DPNDKYSSCKPDFLDGCEGYAERSLKGYNLKL VPTTDWPFNDYERLDYPYNEEDCKTSCLEKCFCA AVIFKEPNTKD-HGCWKKRAPLSNGRKDITVLT TWIKGVNINISNDPLNPFVSDPPSKAKNKVKSLT KVLGGQRKTVTEYKNIDCFYQELIEATNGFKE ELGRGAFGVVYKGMINGGGLSTSVAVKLLDRISG DTEKKEKTEVNVVIGQTHHKNLVQLVGFQKEDDQ RLLVYEMSNGLAAAYLFDGLRPSWKDRKVTQ GTARGLLYLHEECSTQILHCDINFPQNILLDEHQNP RISDFGLAKILVNLQHTHTNTGIRGTGKYVVAPEWF RNKFPVTAKVDVYSFGVLLLEICCRESYCMELIAE KEPF

10

20

30

40

【表 1 4 - 4】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
9	sp_008960_semr	0.9884122	0.8823473	receptor-like protein kinase 1 (RLK1); FUNCTIONS IN: in 6 functions; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Curculin-like (mannose-binding) lectin (InterPro:IPR011480), Protein kinase, ATP binding site (InterPro:IPR017441), Serine/threonine-protein kinase domain (InterPro:IPR022290), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Serine/threonine-protein kinase, active site (InterPro:IPR008271), Protein kinase-like domain (InterPro:IPR011009), Protein kinase, catalytic domain (InterPro:IPR000719), Tyrosine-protein kinase, catalytic domain (InterPro:IPR020635); BEST Arabidopsis thaliana protein match is: lectin protein kinase family protein (TAIR:AT1G34300.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).	MAKPFTCFSEFFATLIFLPLPIYAQSSGKIIVGQSL LSAATNAPSPWSPGDFAFGHQLPTNSNLLLA IWAYKIPDTVWYANDGNPNVREGSKVTLTADKGL VLSDPQASSLLWSTSDDFSNNGVSYGFMST GNLVKNSSNDPWWQSFNHPDTLLPTQIMEIN GVVNSRLENNFRKGRFQLRIPGNLNLVNTDR ASGNYGAYISGTNERPNTGHRLIYNESGYMYV LRTDGYDILLQNKNTPTKSYQYRATLNFQVGL MWYSIPKTSISNGW/SMDQALPNDVCTSIGASGP QLSGGICGFNSICVGLGDDKRPCTCKCPGYSLLDS NDTYGSKPDKADGCVEYAQRSMKNEYRLERL PSTNWPFSDEKLGPNVEDDCKMSCLTDFCAA VAFQDSNSPICWKKKAPLSYGRQDTKVRETTWI KVGNVFNSNDP_LNTFVPLFPKAKNKVKSLDKFL LGGSVFNLVLSAISLGLIYLHKKP_LKEFEQGQR KKLREYNSNVHCFYSYQELTDATKGFTEELGRGA FGVYRGMISTVYGPICVAVKLDLRISHDADKEF RTEVNGIGQTHHRLVRLVGFCKEGDQRLLYVE YMSNGTLANYLFGELRPSWVARIQITLGIARGLLY LHEECSTQIHCIDIKFQMLLDDNHNARIADFLAK LLVNLQHTHTAIRGTGYVAPWFRNKPTVTKV DYYVSGVLLLEICCRRSVCMFEIEEGAILTDWA FDCYQSNLTDSLNVNDMEALNDWQRLKRFNMS LWCIQEDFSLRPTMKMMVMQMLEGLAEVPPKPCP TMLSIITTTNTKPLG
10	sp_009290_wckf	0.9239123	0.8515365	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G25265.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLink).	MMMMNGSKIMGRGGMVLVVALFVGFCAIHYSSTL SLTNTQLKVLGKYGPAPQVVELVETVPDPKQDA VSGPTRMKFHVAVTSDSISYKWKQCRIMYIYWK RVDFLPGSDMGGFTRVLSHGQPDNLNLEEIPSFV VDFLPEGVDKGYVILNRPVAFVQWLEKATIEEY ILMGEFDHLFVNLPLNLAQGENPAAYGYFFIRPD LNEKIVRFKYPNGSISNVDPVIGSSPVIITKSLKEIA PTWMMNISLRINKNDPEADKTFDWWQEMYYAIAASA LHGKHLHEHFVGPMDGNVEISYIHYTYGCD YNNKGGELLYGQKTDWRFDFKRSYQDAPPPKPLTL PPLGAPETVRLIQMVNEATAMIPWF

10

20

30

40

【表 1 4 - 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
				SUPPRESSOR OF ACTIN 9 (SAC9); FUNCTIONS IN: inositol or phosphatidylinositol phosphatase activity; INVOLVED IN: response to high light intensity, phosphatidylinositol metabolic process, response to cold, phosphoinositide-mediated signaling, response to osmotic stress; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: WW/Rsp5/WWP (InterPro:IPR01202), Synaptojanin, N-terminal (InterPro:IPR002013); BEST Arabidopsis thaliana protein match is: Phosphoinositide phosphatase family protein (TAIR:AT1G17340.1); Has 3998 Blast hits to 2752 proteins in 237 species; Archae - 0; Bacteria - 0; Metazoa - 2445; Fungi - 904; Plants - 298; Viruses - 0; Other Eukaryotes - 351 (source: NCBI BLINK).	
11	sp_010800_skmm	0.945733	0.9841832	0.964053	MEAMKLEIERLRNLSTAERYRALLAVSTDPASIN PNLLDDLQWIRFCRITSLAMVQGAALDEKIKTSI GLEKNDVDFWNVSGFRQSCYGGVCEVHAE NGGGGANIYREIISRFLKVFVYVFLWEKRLSNML CWKGSISACKL
12	sp_010890_uhjp	0.963735	0.9245354	0.968321	NKVQVPEVQTEVYHVAVPDPQEKYVPEFVLS TEKVNDDQFPQGVSSLSLAIENDKGVHIVAVGE VAIKFSAEYRKDIPKIKKIPKIKWQKIECPQPPG SKESGYVGRYMIETIALRQMFIREKTMFLFLFRG CYAAVPAADSDVPAAGAAVAGEFVRCCC

10

20

30

40

【表 1 4 - 6】

A	B	C	D	E	F	
1						
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	
13	sp_011580_zgon			Leucine-rich repeat transmembrane protein kinase; FUNCTIONS IN: kinase activity, INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: plasma membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Serine/threonine-protein kinase domain (InterPro:IPR002290), Leucine-rich repeat (InterPro:IPR001611), Serine-threonine/tyrosine-protein kinase (InterPro:IPR001245), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase active site (InterPro:IPR008271), Protein kinase catalytic domain (InterPro:IPR000719), Malectin/receptor-like protein kinase (InterPro:IPR021720), Tyrosine-protein kinase, Arabidopsis thaliana protein match is: Leucine-rich repeat transmembrane protein kinase (TAIR:AT1G53430.1); Has 190921 Blast hits to 136069 proteins in 4848 species: Archae - 114; Bacteria - 18757; Metazoa - 52699; Fungi - 11122; Plants - 65531; Viruses - 458; Other Eukaryotes - 22240 (source: NCBI BLINK).		MAPFFFTSLFFQLTSLFIFFTFNFGSDAQLIPQIEV KTLLEISQKLENTYWKVTQNSCSDGGRSFNISIGV NILSNVTCDSFNSGKICHVTIHLKGLNLITGSLP EEFGDLTQLQEIDLSRNYVNGTIPKLSKIP LVTLS LLGNRNIGPIPIEIQISTLEELELEDNQLGGPLPK SLGNLKSRRMVLAAFFNGTIPETFFVNLKNTLTEL RLSGNVLGKIPDFLGNLTNLSIDIQGTSLEGPPI STSSLTNLILRISDINTRSFRITFPNLNMANLQY LVLKNGITGSIPIDVIGTLANIKLMDLSYNQLSGQI PQSMENLRLENLFTNNSFSGTLPSWILNSRHN FDVSYNFTGSPFASCQQSNVNLVAYSYSKASN SVNWCLRRDLPCSNKPOYYSLFINGCGGTLKSFQ GNVYDEDSFGSPSTFFADGEKWAYSSSTGAFLY DDNGNFLAMETSNHNVGTGYQTARLAPLSLKYGG LCLLQGSYKVKLHFAEIMFSDGQNFNGIGRRFFD VAIQGDVYKLDNFIESEAGGVGKAVVKEFDVNT GSTLEILYVWAGRGTGSPDRGYVGLISGIATV PNFKVNOGLSAGANGVILASCALVLLVVLWKR GYVGGKSENEELRRLGTGYFTLRQIKTATDDFD HRNKIGEGGFGVPVKGVLDPDGKVIKVLSSKSK QGNREFVNEIGMISALQHPNLVLYGCCIEGKELL LVYEMENNSLARALFGKENCKLHLDWQTRKRI CLGITRGLAYLHEESRLKIVHRDIKATNLLDKDL NAKISDFGLAKLDEDETHISTRIAGTMPKEEFVY LLDWAYVLLQEQGNLLELDPTLESSYPKDEALRL LNIALCTNPSPSLRPAMSSVTSMIEGQTPICAPI VKRSGANDLRFKAFERLSQDSQTHVSGSSMSS QVQRISIDGLWSDTSVSMQSKDGITEMSVSTSR LLPDDVKLE
14	sp_013360_axec			SANT, DNA-binding (InterPro:IPR001005), Molecular chaperone, heat shock protein, Hsp40, DnaJ (InterPro:IPR015609), Homeodomain-like (InterPro:IPR009057); BEST Arabidopsis thaliana protein match is: RAD-like 5 (TAIR:AT1G19510.1); Has 590 Blast hits to 589 proteins in 66 species: Archae - 0; Bacteria - 0; Metazoa - 121; Fungi - 0; Plants - 466; Viruses - 0; Other Eukaryotes - 3 (source: NCBI BLINK).	MSSNGWTPKONKAFERALKDYDKDTPDRWNNI AKEVGGKTVVEEVKRYQDILEDIQHESGRVFPF YVNWKM	

10

20

30

40

【表 1 4 - 7】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 ELMOICED-12 family protein; INVOLVED IN: phagocytosis; LOCATED IN: cytoskeleton; CONTAINS InterPro DOMAINs: Engulfment/cell motility, ELMO (InterPro:IPR006816); BEST Arabidopsis thaliana protein match is: ELMO/CED-12 family protein (TAIR:AT2G44770.1); Has 873 Blast hits to 873 proteins in 155 species: Archae - 0; Bacteria - 0; Metazoa - 490; Fungi - 72; Plants - 180; Viruses - 0; Other Eukaryotes - 131 (source: NCBI BLink).	aa seq MKIYFPPSRSDHQEALKVLWYTTYDQPLHGLIS EQWKEMGWGGRDPSTDFRAGGFISENLFFAK TFSTFQHLVKKLGKGRATWEYFVAVGVNITFM LIQMLDLDAPKGRTRFVRAFLQMLSDNEWAFDLL YCVAFVYMDKQWLEKKNATYMEFNILKSTRAQL EKELLMDEVGLIEDMPFSFSLLC
15	sp_013520_aqcr	0.931229	0.937526		MFNHRHLPLISFNSPKPLSFLSLSTSAIKALPS SPHHFLSPNSIDFLP_LSSLSLQSLHQTQKSHALAV FGFLPSSVSLCASLILNYAKFGNPFVSKSLFEKSV FYCKTPFLWNTLLRGYSNSGVYDVYVYVYHMIRT HVRPDDHTFPVVIKACADHCENKKGREIHGFSIK VGFENLVYVGNLTLSPYSSCCDLVGAQNVFDEM PDRDVSWNMTIGVFSANGWSNEAVEVFLQMKI WSEFLPNAVSVSVLPACAEELKDEVAVRVIHGYY VKVGLINQVRVGNALVDYGGYKCGELCSKQVFN EIEKNGVSWNTVISSFAHAHKCYSDAL YMFMLMID SKEKPNGITLSSLLPILVELECFDAGKEVHGLSIKT GLDSDVFINSLIDMYAKARNFTKASIFQNTSVKN VVTWAMVANLAQNSQELAAIGLVYRKMQSFGE PNAVTFNVLFPACARIGALLSGKEIHGRCIRKGIS PDFLNSLIDMYAKCGLVDLAGNVFNVSNRDQIS YNTLIVGYSQTTECLKSLTLFKEMGLIGMEMDTIS FYGAISACANSAAALPKGKEIHAFTRRLFHTHSFIS NSLLDLYMKCGKIGLAQINIFRAHKKDYASWNT MILGYMVGDFETAISMIFEAMFREDGVKHDVSFI AVLSACSHGGLVEKGRKYFEEMKSLIVKNNMH YTCMIDLLGRAGLMEEAVELEIKELPFEPDANIWG ALLGASRIHGNIYAKLAENLFLVLPKPEHSGYYVL LSNMLAEAGRWDIETRVRELIMKSRGVKQKDPAY SWVQNRDQVRAFLVGES
16	sp_016480_idkw	0.936789	0.948324	Pentatricopeptide repeat (PPR) superfamily protein; CONTAINS InterPro DOMAINs: Penitricopeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: Tetratricopeptide repeat (TPP)-like superfamily protein (TAIR:AT3G57430.1); Has 47794 Blast hits to 13382 proteins in 245 species: Archae - 0; Bacteria - 5; Metazoa - 44; Fungi - 66; Plants - 47218; Viruses - 0; Other Eukaryotes - 461 (source: NCBI BLink).	

10

20

30

40

【表 1 4 - 8】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
17	sp_017870_ceny	0.985764	0.921744	0.958512	Phosphoglycerate mutase family protein; FUNCTIONS IN: catalytic activity, INVOLVED IN: metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Histidine phosphatase superfamily, clade-1 (InterPro:IPR013078). Phosphoglycerate/bisphosphoglycerate mutase, active site (InterPro:IPR001345); BEST Arabidopsis thaliana protein match is: Phosphoglycerate mutase family protein (TAIR:AT5G04120.1); Has 16927 Blast hits to 16736 proteins in 2503 species: Archaea - 87; Bacteria - 12791; Metazoa - 870; Fungi - 486; Plants - 272; Viruses - 0; Other Eukaryotes - 2421 (source: NCBI BLINK).	MAASDRVSMCEVKNNDTIEGAAEPVDGTCIEIV VRHGETEWNALRFKMQMDIDLNEVGRQOAVT VAHRLSRETNISCIYSSDLKRAFETAETIASRCGG LEVWTDKDLRERHIGDLOGLTFREAPWVNPJAFE ALKSPENDQEIPGGESRQVYKFRCTSSLQRIAE KHRGARVAVTHGAVIEMLYKRAIPGGRAGEIWN TSISIFQLFEGDKWSIKLWNDKSHLRETEYLEAS GGGDASC
18	sp_020820_yeau	0.977932	0.943386	0.956003	UDP-glucose:flavonoid 3-O-glucosyltransferase (UF3GT); FUNCTIONS IN: transferase activity, transferring glycosyl groups; INVOLVED IN: N-terminal protein myristoylation, response to sucrose stimulus; LOCATED IN: cellular_component unknown; EXPRESSED IN: 14 plant structures; EXPRESSED DURING: LP.04 four leaves visible, 4 anthesis, C globular stage, E expanded cotyledon stage, D bilateral stage; CONTAINS InterPro DOMAIN/s: UDP-glucuronosyl/UDP-glucosyltransferase (InterPro:IPR002213); BEST Arabidopsis thaliana protein match is: UDP-Glucosyltransferase superfamily protein (TAIR:AT5G54010.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archaea - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 985; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MGKTVAASEQ-HIVMIPWFAYGHILPYFELSINKLA EKGHKILTVPNKVKLDLEPKIRHPSLSLHAFVTP HIEPLPGTETCSVDYFIELOHHLAVAMDRARPEV ESISAIIDPKDILFYDNAYWVPEIATKLGMSV FYQACALSITRIKQTPSASASASAKLFTLPKW VLTPKVLGDARANYGEGITYYQRVKKALSSCDAI ALRTCREEGESSDILAAGYKPKVFLTGPVLPFEV FLPPLDNSWAEWLAKFGPKSVVLCGFGSQYVPD KAQLQEMALALEDTGLPFLMSVKPPTTECATIEEA LPEGFSERKERGVHGGVWQQQLQILAHPSVGC FICHCYGSMMWEGLLSDNQLVLLPQLPQLMMA QMLAEKLVGMVDRREDDGWSRKNLCOAVK SVMDPHSEFAALLKNNI-HANFRDKLLTNGFMANY LEVFDQDLKRFLLTAN

10

20

30

40

【表 1 4 - 9】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa_seq WWWKTEGIP TGVWVFWVAFALERVYGRVRFDET SRKEGPKFLRAKVIPTRRPLIPGFYFYEYSQDNF KVVHLRYEGVFNFCQNCCKGKIGHRKPNQVQEE LVKXIEVRRHLKGCGLNFEFMEVTDAPPLYSKKV MGLDDIEKFRTRLLNLRWGOPTATDSDNNDS SDSYSPSPDPNFSQRSTGNSATSATNKTNHT GKKRKGDNQADGGGNSKRRRRANGEASDLR SIRERRGLANLQKLPSQNILMGWCMAVSVE NGSISKAVAGLVTPSQPCKNRVSHYHQSSNSL FNASGHTINTSCNLESASLLTIPQLQPTWVSS PRLLVLPVFSIKWSVPYPSFIMTAPTVPFSPFW GFFLEFVLYHSPVFLSSPVYSLEVELEKLPVLM ISTTCSIDSLNLQHCYLYPNTFSLNHLKPYPIQ TILVPHDILLIPIPLLLAMDTPPMSTSLVIAS KAVVLSQODRIFTRKNTLLDPFKPKSTSSAYSWS YNTNIQFSIPKHLVSAQLSPKFNNSAGFSSTPI ETLIHACSONLQALSVLSPVANGDTSNNVSTPP SSPASNFTAPSSLESAEVESMIKLDYALDSD STTMNSWDVEPEYSHVEVKMYKKRAGMRQH EDGPRKKWCPSLDDEIRVAIVEGKIYPEENPII LKRUEGGQTSNGHGANSSNSPWWKKNQLR YEGRGFRFLSKDDDLWEKLWVAEEGEDINT NIKTSINILEPGSKLFVKQKLOQLNKTRKWCLD YKKEMIGITWKNFEEELAIQGNHNIESSIGEVEER KKCREKAFFQWQYWRQSKSWDAFGDQSTA FFYKSVGRSCRNDIRAKDDDDGVWTKKEEIQN EYKYNNLFAQSPYVAGHEDINMKFWLEDIPSJIT AHINALMQFFSEEDIKVAFAFKPLKSPGPDGVP PVFIQKYWATAILGIFRLAPSKKRTLDLHOVSGIK PSRMTLLGPPSSGKTTLLALAGKLDNNLKVSG KITYCGHELKEFVKPRTSAVIKQNDIHLGEITVRET LDFSGRCLGVSRYLLKELLKQKQAGINPQVDE DTMARKKTVNLADTLMLMGHQMDFPSGGRYIHT MSAGEISRTPPGEAKSSHHEATYDFEFWMEETE DVGPSNHLVDSLEESQEDSDGGVEEAEAEPRR VAVVRGNDQEMELKCSRWQLEPHFAGGSDDE LPALTKGEWDQIVVHFQVESAVQGGKLNLPKS WLPSCQFKDHVFLASAGLDHGLDQAFAMSKFP PVEKQINLTKWELTQMVEOKLKGIVGGPPAKFE FRCGCISPRRICKQLSGGIFPLVQVDSDDGFILS SLLKVTGEGYKVPVLSVKYINKTKVANKFLVDKIR ELRMVTEVOS
2	分析に用いたベイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	pleiotropic drug resistance 11 (PDR1); FUNCTIONS IN: ATPase activity, coupled to transmembrane movement of substances; INVOLVED IN: drug transmembrane transport; LOCATED IN: membrane; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: 4 anthesis, C globular stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAINs: ATPase, AAA+ type, core (InterPro:IPR003593), ABC transporter-like (InterPro:IPR003439), Plant PDR ABC transporter associated (InterPro:IPR013581), ABC-2 type transporter (InterPro:IPR013525); BEST Arabidopsis thaliana protein match is: pleiotropic drug resistance 6 (TAIR:AT2G36380.1); Has 357151 Blast hits to 273757 proteins in 3981 species: Archae - 7090; Bacteria - 287486; Metazoa - 9208; Fungi - 6204; Plants - 5777; Viruses - 4; Other Eukaryotes - 41382 (source: NCBI BLINK).
19	sp_020990_kogi	0.952213	0.987957		
20	sp_021210_xeed	0.924177	0.9791		GAMMA-IRRADIATION AND MITOMYCIN C INDUCED 1 (GMI1); BEST Arabidopsis thaliana protein match is: defective in meristem silencing 3 (TAIR:AT3G49250.1); Has 114 Blast hits to 102 proteins in 24 species: Archae - 0; Bacteria - 7; Metazoa - 8; Fungi - 4; Plants - 91; Viruses - 0; Other Eukaryotes - 4 (source: NCBI BLINK).

【表 1 4 - 1 0】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたペアとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
21	sp_023090_letn	0.951575	0.987891	0.99926	AGAMOUS-like 20 (AGL20); FUNCTIONS IN: sequence-specific DNA binding transcription factor activity; INVOLVED IN: flower development, response to cold, positive regulation of flower development, maintenance of inflorescence meristem identity; LOCATED IN: nucleus; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAINs: Transcription factor, MADS-box (InterPro:IPR002100), Transcription factor, K-box (InterPro:IPR002487); BEST Arabidopsis thaliana protein match is: AGAMOUS-like 42 (TAIR:AT5G62165.3); Has 7364 Blast hits to 7361 proteins in 926 species: Archaea - 0; Bacteria - 16; Melazoa - 673; Fungi - 316; Plants - 6262; Viruses - 0; Other Eukaryotes - 97 (source: NCBI BLink).	MVRGKTQMKRIENDSSRQVTFSKRRNGLLKKAF ELSVLCAEVALJIFSPTGKLYEFASSRCCSTSRTI DRYQKMKELRGVSKVSDRNWLOFKEDANELE KKLSLSKFRLLGDSLESGIDELQDLKLDK SLSTIRAKKNALFKVQIDKLYEEEEKLLKENARLKE KLNLSLVPVLSVRQPTDDDDQRNNVQDDVETELFI GLPESRNSRCLP MRWSSARAHVYVDMTSVYDGGGWWAGQVSSAG QVSAIG
22	sp_026540_smp	0.954322	0.944192	0.979128	uncharacterized protein	
23	sp_028080_hieu	0.982014	0.984432	0.999831	NB-ARC domain-containing disease resistance protein; FUNCTIONS IN: protein binding; INVOLVED IN: apoptosis, defense response; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAINs: Leucine-rich repeat-containing protein (InterPro:IPR015766), NB-ARC (InterPro:IPR002182), Leucine-rich repeat (InterPro:IPR01611), Disease resistance protein (InterPro:IPR000767); BEST Arabidopsis thaliana protein match is: LRR and NB-ARC domains-containing disease resistance protein (TAIR:AT3G14460.1); Has 25652 Blast hits to 19792 proteins in 981 species: Archaea - 38; Bacteria - 1467; Melazoa - 4997; Fungi - 304; Plants - 18213; Viruses - 0; Other Eukaryotes - 633 (source: NCBI BLink).	MAEAVLPLAEAAKPMIGLVSRVTAIEQLCRDF KRDISELQKIQISSTVLMNTSATNTRSSSSSN TMELWVRDVMVNHVEDLDVLAIEKQRQVVM KRTFLWPLKQVYFRKISRWIKRINTITGEALIEA QTIQLKDLVHSTSPPTATNFITNEGAAAAAENR LOQLRRYVDISLIVGRDSDVSVVVKLCNPSKDC YNTSNCLSVAVVYGGIGGKTVLAKKGGPGLL KHVKYLRVLDLQDMGLEEVPEVAVAGLKLKLSL AANPIKVPFVTELYNLSLQNLKCDLLKEVPGG LSNLVLRHLMIKILNLPAGMIGQLRCLQTLPLPW VNEGQFISELGDLPHYVSGSLAIHGLEHNSKQD AESASISEKSRVKKRLRLWSDDDWIKHKEVLD GLKPKSNLRYLIICGYGGEIFPFWLMRMDVIKIIL VGCRRCKLPTLHLPFLKLELJEDMESLECIQEE LHNTAILTEKSSMVFYPSLKLRLIGLHKLTWE VEPSSGGWTLFPLLEELMAYCYHELTTPTNTPFS LKQTLVTNITSGRPLSNITNKSSVKVCLVQLDWI TDITDPEYLAESCSSLQTLKIGCVNLSLLENL SMLISQVLEIKRCDKLYIPNLSNLSGLRELMIHF NLALTSVPKGLQSCNSLERLWFACCPKVETLPLD TGLTRLONLRSRMTKMITQPPDWLYTLPLCLRL GIGSYWELDEVPMDFLVKIVSLESJHLFGWEKL QSLPEQLOQFTTLKELGICDCDNMEYLPDWVGN LSSRLYLSPTRCRNLRKLPMTMDAMKRLTKLSFLKI EDCPELEAVQFNGSEYHKHSHIEIKIVPTY

10

20

30

40

【表 1 4 - 1 1】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	分析に用いたバइटとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-lite)	注釈	aa seq
24	sp_032370_ufri	0.993628	0.993533	0.953132	unknown protein; Has 109 Blast hits to 109 proteins in 44 species: Archae - 0; Bacteria - 2; Metazoa - 61; Fungi - 8; Plants - 36; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLINK).	MVKGLQQNLPAADVTLQDLERHCLASDGSFV SKSAYYDLQLAREEMCRERLRYLEAMAIYCEAIA MVEEYQQALSVANFRDVQGLYPQLGSKNXPQY ETLEHRLVWAEAAQKLRPLJSKDGGEIHEEDIEKW SIMSRSLDSTSNLTMSSSTNSTNYGTSTPSTV ANVTPSNSTDMSDFGIGVGNRYLGITSTYLVWQ TQLHQGTAEQMSLFRFEIARLRVKCDQADAFV MDDVDGSSGYQNVSSRLPERVKIIEIEERDEEAL KEDLYSADRFKFAEYVNLQILGVLIKVKDKLQ NQHKYDELQKTWQIKRHEAMGAKLRVLDHILLVA TYTQETIPALHKIRNYLVEATEEASTAYNKAVTRL REYQGVDPHFDTIARQYHDIVKVKLENMOWTIHQV EMDLKYLDPDHSST
25	sp_033550_oqaz	0.9939321	0.9933137	0.9939538	RESISTANT TO P. SYRINGAE 5 (RPS5); CONTAINS InterPro DOMAINs: Leucine-rich repeat-containing protein (InterPro:IPR015766), NB-ARC (InterPro:IPR002182), Disease resistance protein (InterPro:IPR000767); BEST Arabidopsis thaliana protein match is: LRR and NB-ARC domains-containing disease resistance protein (TAIR:AT1G12280.1).	MLANEVITLGAIFGLTISIAGSTACFYNSYKTSRI NL.SILKVKVQNLLEIQELQOKLNGIDVLKHPRE WISRRREFLETWNISDLESRLNENCVGGQLPNFC SRFGALKSIEESNHGLDVLDSGKNVSREIKVKG SIQPEVNVGNARKEIKRIRWEQVSOAETAGVCI HGMTGVGKTAFAAAHNKALKELTAFENVIWVWV ENSGSLKHQEEELASKLQVHLPGTTV'NKELRSA LMQGRFLLVLDSSMWQAFSLHDIGIPEPVAGSKL IVTSKTYSYCRKITKNIKREIYEIKPLLEAEAWDL FOGEGVNVNLESETVSAQRAVKVNDGLPLAIK LFAETLSEVHDDYGGIDAAWKDELFSLSRSTSL LEGYLIKPELIDYMWEGLLGPISSLGESMRIG RQTLNELFAHLLLESVKSEGESLKMNLVRRHV AMDASKSGNHFFIKGOSLGEFFSAYDWPANT ERASLVQNLRALQNSPDCNKLSTLLQDNPLNL VHHEFFSKIRNLRYLDLSRTSSSLPKFSGLTN LHALLRSCRNLKLLPSSLNHLRVIDISETPLEO VPDGLNLINLRDLKTKIEVFPATVIGKITEM EELLTAENGGYVWGSNLIFFQWPGACIEELDL KKMVVQLAFINPEVFNRFVQVDDQQRAIQLR SKFYVGGYSGDGDIGDINSVVMGDCSIRVPRGT SELYIMKHSOELLKSLKFDGCMRDLTIVDYSRFDA LTYLFTLEMLGSLKSKICVKCHKMKMSVQIPAA EANPTSIITPKLTKFLYDLQNLSEVYDGKVLSCPS LHGFEVFNCKLKLKPKVLVGRQNCIIEIYGEKEWI DKVMEDGSPHLKYTEALVPAELSTFSLRFRSE SVFSRDTHEDSGRSLGATIPRWHRHLRHPFRFGS AVFSPADTLHIASTTSDAISEVSVLPHAEAGTS QRGNRTRSSRSD

10

20

30

40

【表 1 4 - 1 2】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
26	sp_037920_hhky	0.953715	0.83387	unknown protein; CONTAINS InterPro DOMAINs: Protein of unknown function DUF2439 (InterPro:IPR018838); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:ATI G43722.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MDDVKKWTVTYTKHVQKRVYQDGSLLHLSNN KVMLYDESEQLLDSKFLTKDEAVKPGESLTFASF LVDIGDFDNCNPTSNANPKPRDKTQMPQREKS RHTFSSMKPNTAGRKTSLVSVSPSQIREFKKN EELKYGAHSPKSPGTSSFKWEALYTTOLTOKA KKYHDGFRMIYCGSQGRQIMLYDSSKNOIAKRF MKKDEVIDTGGSLTFDSHLVDIGDPMGTNGDDA DYKQERNNAKRSKSTLLDHVGSNGSVIGKRR DDACSTKRVNSDSSGDFDKINQKKNFASSSL RDANQILSLRKPTQDTSIVAEAPMQCPTTSS EPVQIHADSOAFDSSLOVDQCPVIAKVKHNEK AVDGGQSSTFSENGCSYLERLCTQGGIELRSNEF ESPPKSQPIVGSVDHGRPIVICSSKRYVESIDNRS VYDSETHLQPSSTTKLPETDPENKEKTASTKDM SKRFFGSLMRMKTKKSVKNEKLNKIMECP'SFDLGI
27	sp_040350_wcny	0.973027	0.984336	cytochrome P450, family 72, subfamily A, polypeptide 15 (CYP72A15); FUNCTIONS IN: electron carrier activity, monooxygenase activity, iron ion binding, oxygen binding, heme binding; INVOLVED IN: response to cyclopentenone; LOCATED IN: endomembrane system; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Cytochrome P450 (InterPro:IPR001128), Cytochrome P450, E-class, group 1 (InterPro:IPR002401); BEST Arabidopsis thaliana protein match is: cytochrome P450, family 72, subfamily A, polypeptide 13 (TAIR:AT3G14660.1); Has 31875 Blast hits to 31741 proteins in 1631 species: Archae - 65; Bacteria - 4943; Metazoa - 11298; Fungi - 6149; Plants - 8001; Viruses - 3; Other Eukaryotes - 1416 (source: NCBI BLINK).	MISKSASGVAIASLGCIIILYMMWKLKGLWLTPK KLEKCLKQQLVGNYSYKFLIGDMKESKLRNEAL QKPIPTHYDYNRIQPHIHOILNNSGAGKNIYTWL GPVPTILIQPELIKDAFNRMNFQKPRLNPYTQM LSTGLPNYEGQKWAKHRKLLNPAFQLDKLMIH TFETCVTDLTKWLEKLVCKTGSSEVDIWPYLTLL TGDGIARAAGSSFFEDGRRIFELLTKQDVISLLK YSYIPGFKYMPIKGNRKMKEADNEIKPLTNIINRR RKAMEAGEAPKDDLLGMLESNANEARQVNEE SGSSKRKSDLTMSFPEMIDACKOFFLAGQETTSSV ALTWMLLAKHQDWQTRARQEVLFATFGMNTP DFDGIHNLKIVTMILYEVLFVLPVPAVSRVRHD RETKLGDVLPQGVGVSVFSLHAHILNPEIWGDDA KEFKPDRFAEAGIAKATGNNSYFFFGWGPFCIG QNFALVEAKMALCMLQRFSDLSPSYIHATPSLI SLOPQHGHAHILHFRF VLDQQPGSLECGYVIRFVYDIFTKHRDSQDLTT DYSST
28	sp_044310_sjto	0.985125	0.976974	uncharacterized protein	

10

20

30

40

【表 1 4 - 1 3】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
29	sp_049910_uyku	0.976375	0.961653	ER lumen protein retaining receptor family protein; FUNCTIONS IN: ER retention sequence binding, receptor activity; INVOLVED IN: protein retention in ER lumen, protein transport; LOCATED IN: integral to membrane; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: ER lumen protein retaining receptor (InterPro:IPR000133); BEST Arabidopsis thaliana protein match is: ER lumen protein retaining receptor family protein (TAIR:AT4G38790.1); Has 889 Blast hits to 889 proteins in 227 species: Archae - 0; Bacteria - 0; Metazoa - 314; Fungi - 182; Plants - 232; Viruses - 0; Other Eukaryotes - 161 (source: NCBI BLINK). UDP-Glycosyltransferase superfamily protein; FUNCTIONS IN: transferase activity, transferring glycosyl groups; INVOLVED IN: metabolic process; LOCATED IN: endomembrane system; EXPRESSED IN: 6 plant structures; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage, 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: UDP-glucuronyl/UDP-glucosyltransferase (InterPro:IPR002213); BEST Arabidopsis thaliana protein match is: UDP-Glycosyltransferase superfamily protein (TAIR:AT5G53990.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MGDKRATPTAVKQLSRWIRSRVROQVGMASM AAPILLIMKVFVKNYTYFYVAQLAHSHGILVLYK LTIKTKCSGLSKTQELTSIFLMRIICHFMIRDIYI VLDSITLGSSTLWVIYMIYKLNKNTYLQDLDMIPFY YVLVPAPVIAFYGNPVGISGLFYMYSKVIEAVSV LPQLRIQNAKVVPEFTAHYVVFALGVARFFSCAN WILLIDTKGTMLTSGISGHLYLLIILLAEVYQTFILA DFCYYYVYKSVVEGQSMRMLPFSMV
30	sp_049940_mmul	0.933433	0.933907		MAASNKEQSKLHAMFPWFAYGHINFIQLSNKL SSHGIQISFISIPGNIDRIKSSLNLSPPNQLIPLIP PTEGLSPNFDSSSEVTPQATQLLTLALDQMOPQ VKALFFHPQPQVILDFAYHWPVASELGIKAV HFNTFFAVINSYLTVPSRMTDPNKPPPTFEDLKNP PQGYPKTSTASVKTFEAQDYLFKFSDFGGPCH FEKILAFNSCDAILYRTCNEIEGFIDYFKTQINKP LLLAGPSVPLPPSGELDEKWEWMLGKFKPEKSVI YCSFGSETYLNDAQIQELTLGLELTGLPFLVLF GTSNSTDAHNKLEASLPEGFRERIKDRGVLHTG WVQQQNILAHRSIGCFLTHAGFSVIEGIVNDQQ LAFLLKADQFMIAKLFGSGLKAGVEVNRDEDG SFAKEDIFAIKTIMVDTKPEPSRSIRENHSNWRK FLMNKEIASYIANLAHELKALVQKA

10

20

30

40

【表 1 4 - 1 4】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
31	sp_051630_sset	0.854572	0.822075	Major facilitator superfamily protein; FUNCTIONS IN: carbohydrate transmembrane transporter activity, sugar:hydrogen symporter activity; INVOLVED IN: transport, transmembrane transport; LOCATED IN: integral to membrane, membrane; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 7 growth stages; CONTAINS InterPro DOMAINs: Major facilitator superfamily (InterPro:IPR020846), General substrate transporter (InterPro:IPR005828), Sugar/inositol transporter (InterPro:IPR003663), Major facilitator superfamily, general substrate transporter (InterPro:IPR016196); BEST Arabidopsis thaliana protein match is: Major facilitator superfamily protein (TAIR:AT2G48020.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MKEEKVQLLLNTDNNDKLVDEAGHHGLFPLPE ANDTSLAVAAAAAVTPMLLFSVFIAVLGSFSGFV VGYTSPVQAEIMKELDLSTAQFSLFGSLLTVGGA LGSWCGKLDYIGRRRVLQLTSLVFIGWFAISV SQGAWLLDLGRWMLGFTGGISGFAIPVYVAEFT PKNLRGGFVILHVMITVGVSAAFGLIGLVTWRTL ALIGIPSLYQVIGFIFIPSPRWLMMMGKMKFEFA ALQCFRGGVDITQEAADIRNSAEALDQIEKYSIL QLFHRKYAYHIVGIGLSALGALVGLNGLIFYAGSI FESAGFVKVGTVALASFQLPSWVLGVILMDKCG RRPLIMISAAGLFIGCSLTGLAFLFEDHHLFGGISP SIALIGILYVAFYPIGVGGGTCTIIEIYPLNIKSGAG SIAAYVSSLSSWIVSYAFNFSMEWSSSGTFFVLA GVCVFTLIFVAKLVPETKGTLEEVHKTMSDVWQ
32	sp_052880_qnfy	0.875383	0.846745	XH/XS domain-containing protein; CONTAINS InterPro DOMAINs: Domain of unknown function XS (InterPro:IPR005380), Domain of unknown function XH (InterPro:IPR005379), Domain of unknown function, putative Zinc finger, XS/XH (InterPro:IPR005381); BEST Arabidopsis thaliana protein match is: XH/XS domain-containing protein (TAIR:AT1G15910.1); Has 52398 Blast hits to 32322 proteins in 2074 species: Archae - 615; Bacteria - 6626; Metazoa - 25839; Fungi - 3724; Plants - 1961; Viruses - 202; Other Eukaryotes - 13431 (source: NCBI BLINK).	MENGNQDELLKTLRNLVVKLARELDYKNFTLMEL EEKNAKDLNWMKEKEKLIERSAAISVLLSENV DKDLLLDEKSARVIALTEEKQMLRETTKSMNTL KEESERQINEKSAVINTLKAQNAADKERLLEEKEW LLDERLTCMNTLVKEKERLVNEKSAVIRTLMIQNA EKEMLNEEKSAAISLMAEKEQGLEALQKELKKT ENMKLNKELQCDLESCCKGELERLKEKKNLQME TNNTELEWVEYCSKPCGKELNVQPREPESRVQ TYEEINCMPLOMQNPNEDLRTVILFRKDLGKAE LLESMEMENRILMVKELKSNHELQVARKAAATEGL LAMHFRPMNIGIKMGVNLKPFPRDACSRRFRS GNWDEKFGNWEKESVMLCSSLWQREISDPGWQ PFKQETVNGKLTVIDENDEKLRELKEEWEAEAY KAVKEALVEVNEYNASGRYPVSELWNHKKQRKA TLEEVIQYIKQWRTRTNKKNKMAV

10

20

30

40

【表 1 4 - 1 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
33	sp_056760_rown	0.952633	0.925079	0.9464273	<p>P-loop containing nucleoside triphosphate hydrolases superfamily protein; FUNCTIONS IN: nucleobase, nucleoside, nucleotide kinase activity, nucleotide kinase activity, ATP binding; INVOLVED IN: nucleobase, nucleoside, nucleotide and nucleic acid metabolic process, anaerobic respiration, nucleotide metabolic process; LOCATED IN: mitochondrion; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Adenylate kinase (InterPro:IPR000850); BEST Arabidopsis thaliana protein match is: adenosine kinase (TAIR:AT2G37250.1); Has 9469 Blast hits to 9393 proteins in 2915 species: Archae - 95; Bacteria - 5951; Metazoa - 891; Fungi - 419; Plants - 430; Viruses - 0; Other Eukaryotes - 1683 (source: NCBI BLINK).</p> <p>Phosphoglycerate mutase family protein; LOCATED IN: chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Histidine phosphatase superfamily, clade-1 (InterPro:IPR013078); Has 1391 Blast hits to 1391 proteins in 390 species: Archae - 1; Bacteria - 792; Metazoa - 0; Fungi - 0; Plants - 44; Viruses - 0; Other Eukaryotes - 554 (source: NCBI BLINK).</p> <p>YELLOW-LEAF-SPECIFIC GENE 8 (YLS8); CONTAINS InterPro DOMAIN/s: Thioredoxin fold (InterPro:IPR012335), mRNA splicing factor, thioredoxin-like U5 snRNP (InterPro:IPR004123), Thioredoxin-like fold (InterPro:IPR012336); BEST Arabidopsis thaliana protein match is: mRNA splicing factor, thioredoxin-like U5 snRNP (TAIR:AT3G24730.1); Has 564 Blast hits to 564 proteins in 220 species: Archae - 0; Bacteria - 0; Metazoa - 221; Fungi - 146; Plants - 112; Viruses - 0; Other Eukaryotes - 85 (source: NCBI BLINK).</p>
34	sp_050610_zonc	0.976479	0.885918	0.966873	<p>MGNKGWEEETASMTGTVIEIKTCNAALLEGTG KSWEEMLSFLLPFALVPLPKFTSGKAKMGVHAN SFMAPG</p>
35	sp_061210_ojgc	0.976639	0.862438	0.9598827	<p>MSYLLPHLHSGWAVDQAILAEERLVIIRFGHDW DETCMQMDVLSVAETIKNFVAVLYLVDITEVPDF NTMYELVDPSTIMFFFRNKHMIDLTGTNNKINW ALKDKQEFIDIVETVYRGARKGRGLVIAPKDYSTK YRY</p>

10

20

30

40

【表 1 4 - 1 6】

A	B	C	D	E	F	
1						
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	
36	sp_064650_ufan	0.922775	0.867018	0.945318	receptor-like kinase in in flowers 3 (RKF3); FUNCTIONS IN: receptor signaling protein serine/threonine kinase activity, kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: plasma membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase- like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine- protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: Protein kinase superfamily protein (TAIR:AT1G11050.1); Has 112833 Blast hits to 111726 proteins in 4356 species: Archae - 97; Bacteria - 12546; Metazoa - 41926; Fungi - 9138; Plants - 32648; Viruses - 383; Other Eukaryotes - 16095 (source: NCBI BLINK).	MCHLHLLFLSLQLLLSSAQNDTVPSTSSSSP CPLNFTLSTLRDADVPKSNNDVANGCGYAEGL RLVSLQSTGTSFPVNTSESCWTSYDSVVD SFLPGKIRSSCGFRRTQWISGEGCMNISTRSEFER LVQSAALDVAVSNQSLNGLSPGCAAC.TSLTTL QASYLTSNPFNGNISDCTNLP LIYAGALINQYGP SDFGTAKCLFLDFSSQNRNRRKPKILLISLTVV GFICISIFVGGYFFSFLKKNRAKMAEKRRINRV GSGAGPNLSSIGGSTHLMKFGYDEIKRATKNFHR DNLIGKGGYGVYKGLLPDNTTEIATVRRKNCSA AGDAFTHEVEIASIRHVNVLVLRGYCVATTNLE GHQRIIVCDLMHNGSLHDHLYGSKLWPFNPE GFHSLSTRVAGTMGVVAPEYALYGOLTERSDVF SFGWLELLSGKAILSDAGNPSLSPLVTDWA WSLVRKHGGLDVEEGMPELGPESVMEKYVLT VLCSDHPQLYARPTMDQVVKILETENMQVPAIPER PIFNALEIEDIERSASSSSGSGQMSTSGGGFQSYYS MGHYDFGSHNGLEEVEDN
37	sp_056190_ewgc	0.972328	0.847288	0.9263287	COBRA (COB); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to salt stress, multidimensional cell growth, cellulose microfibril organization; LOCATED IN: in 6 components; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Glycosyl-phosphatidylinositol-anchored, plant (InterPro:IPR006918), COBRA-like (InterPro:IPR017391); BEST Arabidopsis thaliana protein match is: COBRA-like protein 1 precursor (TAIR:AT3G02210.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MESCCRSRISKLTSTLFIALLSCSFFTSSEAY DALDPNGNITIKWDVINWTPDGYLAVTWFNFOQ YRHQITPQWTLGWTWAKKDVIVSMVGAQTTEQ GDGSKFGNIPHCCKDPTWDLPLPGTFPNQOJA NCCRGGVLTAWGQDTANAVAGAFQISVGGAGTS NKTVKLPRNFTLKAPGPGYTCGFAKVVVPTQFT TADKRRRTQALMTWNVCTYSQFLAQKPTCCV SLSSFYNDTIVPCPTCCGQCNITIPGSCVDGD SPYLASASAGGPKASNTPLVQCTSHMCPVRIHW HVKNYKYDYWFAKVTITNFNYRMNYSWNLVVQ HPNFDNLQTSFSFMYKGLTPYDGLNDTAMLWGV KFVYDFLNQAGPFGNQVSELLFQKEKSSFSFEK GFAPRRYVYFNGDNCVMPPPDAYPWLPNASSR PIFLLQPVMTLAAALFLLAYA MLSQFSSMIEGAIKGRRLKTSMMIMIGLGRQVW LPKYVMLMLQKTEGLRQ
38	sp_070160_ugjc	0.952033	0.853398	0.977037	uncharacterized protein	

10

20

30

40

【表 14 - 17】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたペアとの相関係数 sp_107620_kqnh sp_107660_klqg (SOAP2)	sp_107620_kqnh (SOAP2)	sp_107670_ptqx (SOAP2- <i>like</i> )	注釈	aa seq
39	sp_074730_rixn	0.993346	0.945731	0.952812	mevalonate kinase (MK); FUNCTIONS IN: mevalonate kinase activity; INVOLVED IN: isoprenoid biosynthetic process, metabolic process, phosphorylation; LOCATED IN: cytoplasm; EXPRESSED IN: 30 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAIN/s: Mevalonate kinase (InterPro:IPR006205), GHMP kinase, ATP-binding, conserved site (InterPro:IPR006203), Mevalonate:galactokinase (InterPro:IPR006206), Ribosomal protein S5 domain 2-type fold (InterPro:IPR020568), GHMP kinase (InterPro:IPR006204), Ribosomal protein S5 domain 2-type fold, subgroup (InterPro:IPR014721), GHMP kinase, C-terminal (InterPro:IPR013750).	MAVTRAPGKILLAGEHAVVHGSTAVAAASINLCTS VSLFSPGDMLEKLEKDMLEFESWPIGKIKEVIPD FDTPSPSTPLPCSSSETLKSIAISLYEEQNIPEAKITL ASGYSALFLWLYTATQGYKPAATVTVTSELPLGAGL GSSAAFCVALSGALLALSECCELDYSQQGWLTL RNSELDLVNKWAFEGEKIHRPSPGIDNTVSTYG NMIQFKSGEITRLKSNTPLEKMLITNTKVGKRNTKAL VAGYSERTLRYPEAMTCYFKAVDSISKELSNIIQT SASNDIAVTEKEDKVEELMEMINOGLLQGMGVSH ASLETYLRITLTKYKLRSLTGAGGGGCVLTLPLTL LSSIVVDKVVVAELESCEGFCQFTAGIGGNGLEVCV SDNAQI
40	sp_074810_agcn	0.977297	0.935336	0.966715	Arabidopsis protein of unknown function (DUF241); LOCATED IN: chloroplast; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF241, plant (InterPro:IPR004320); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR-AT4G01590.2); Has 1807 Blast hits to 1807 proteins in 277 species: Archaea - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MAAASHHPVRSISLPARSHPTSLKIEETINKLKSRI YQSDSTSCYTSGSDSIQLGLNLVELYARVQELV SSLSSQHGKLVVEALESSVSLDTCISARDLISGF KEQVCDLQSAALRRKSSSGTEKEINIMYMSFRKVS KEYSRCLKVLKQAEAKYGITAPLLDNDPSSLNLN RLRRETYATIVFKSLFLFLSTPTKPKWLSISKL VTKNSISKEAVTQLNEVSGVDSAIQNESKYDVML LRGKIQALDSSLQILEGEVGVCLFRCLLQNRVSLLN ILTQSI

10

20

30

40

【表 1 4 - 1 8】

A	B	C	D	E	F	
1						
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	
41	sp_076690_ejcm	0.963713	0.923413	0.953234	<p>cellulose synthase like G1 (CSLG1); FUNCTIONS IN: cellulose synthase activity, transferase activity, transferring glycosyl groups, transferase activity; INVOLVED IN: cellulose biosynthetic process; LOCATED IN: membrane; EXPRESSED IN: hypocotyl, sepal, male gametophyte, flower, carpel; EXPRESSED DURING: 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Cellulose synthase (InterPro:IPR005150); BEST Arabidopsis thaliana protein match is: cellulose synthase like G3 (TAIR:AT4G23990.1); Has 3059 Blast hits to 2371 proteins in 475 species: Archae - 2; Bacteria - 903; Metazoa - 2; Fungi - 10; Plants - 2111; Viruses - 0; Other Eukaryotes - 31 (source: NCBI/BLINK).</p>	<p>MATSHIRNVQLTRAVNRLHIFLHVAIFSLFYRF TSFFNSDISLAYSLTTAEFLFLWAFTQAFRW RPVMREVSGYESIKPEQLPGLDVFIVTADPTKEP VLEVMNSVISSMALDYPVDRLAYLSDDGGSPLS KEAIKAYEFALKWIPFCNKYNVKTRCPOAFFSPL ADGERLDWNESEFMADQLELQTKYEAFRNYVEKE SGDNTKCTAVHDRPFCVEIHDKONGESDVKM PLLVYAREKRPGRPHRFKAGALNALLRVSSLMS NAPYLLVDCDMYCHDPTSAQSCMCFHLDTNMA SSLAYVQPIQIFYNVSKNDIYDGOARSAHMTKW KGMDFLRGYPVINGTYLKRKALFKPNNEDY LNSQPEKAFSGSTKLIALLRENSKONLAIKETED ELYQEARNLATCTYEANTLWSEYGYSEVCELE STFTGYMLHCRGWKSYLYPKRPFCLGCTTDM KDATVQLIKWTSLLGIALSKSPLT-AMSSMSIL QSMCYAITFTGLFAAPLYVGVVLPISLLKGFPI PKVSDPWILPFVLFVSSHQHLQHYLEVLESDKSAIQ WWNEVRIWMMKSVTACFLGLETEAMKKGQVTA TFRLLTNKVVKEKMDKYEKEREDFSGAAMLMPV NILVLLNMVVFIGGLMRVINNSYDQMFQALF-S FFVLLLSYPVVKGWL</p>
42	sp_079260_kmfz	0.981037	0.944412	0.974061	<p>Regulator of chromosome condensation (RCC1) family protein; CONTAINS InterPro DOMAIN/s: Regulator of chromosome condensation/beta-lactamase-inhibitor protein II (InterPro:IPR009091), Regulator of chromosome condensation, RCC1 (InterPro:IPR000408); BEST Arabidopsis thaliana protein match is: Regulator of chromosome condensation (RCC1) family protein (TAIR:AT3G53830.1); Has 20178 Blast hits to 6212 proteins in 476 species: Archae - 78; Bacteria - 2696; Metazoa - 7040; Fungi - 1181; Plants - 2975; Viruses - 2; Other Eukaryotes - 6206 (source: NCBI/BLINK).</p>	<p>MNGEGGGGESEVQMEIEMREERERKRYVMW GYLPGALPQRTPLQSPTLVK-PPSSGNFWKDVQ GGGCGFAMASIDTQKLTWGSTDDDLGQSVYVTS KHGETPEAPLPTEAPIVRAAAGWAHCYVATDR GEVYTWGWKECVPSGKVFVGDSSSGGSEKDAL EKQSSLFPEQVSPRGSKNANGIGPGAVLENKGA GEETTKRRRVSSSKQTVESSTASDETLTALPCLV TLNPGVRISTV/AGGRHRTLVLASDVGGVWGW GYGEGQLGLGSRIRMVSSPHVPCLESSYRKH RHSGFSQGLNSDGHGFRVPGTYIKGIACGGRH SAAITDAGAILAFGWGLYGVQVNRFDREKSRVDV QDEAHSSQRSDDDYFLPTLCSLCCGQGDVLS PTCVSLLGRIRIEAVALWHTVCGISAGEDVYAFG GNQFGQLGTGSDQAEATLPRLLDAPSIENVATIV SCGARHSSITGDGKVFVSWGNWYKYGQLGDDAI DRNIPSOVKLDGSAVAKNACVWHTLLLAESPT MAASTSLPSTLTLFLLILLNFTYSSGSSMAIG SSRNLQAAASVHKQTDNTKQNIQIPNCSEIVAKSHC SSNPKRCWCRSDVLDSTCSTTSDAWRFRFSQIFI CD</p>
43	sp_079290_hwwi	0.995356	0.959503	0.961348	<p>uncharacterized protein</p>	

10

20

30

40

【表 1 4 - 1 9】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
44	sp_079580_zhwh	0.3572886	0.381127	Major facilitator superfamily protein; FUNCTIONS IN: transporter activity; INVOLVED IN: oligopeptide transport; LOCATED IN: membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINS: Oligopeptide transporter (InterPro:IPR00109), Major facilitator superfamily, general substrate transporter (InterPro:IPR016196); BEST Arabidopsis thaliana protein match is: Major facilitator superfamily protein (TAIR:AT3G16180.1); Has 5212 Blast hits to 5001 proteins in 870 species: Archae - 0; Bacteria - 1843; Metazoa - 575; Fungi - 389; Plants - 2143; Viruses - 0; Other Eukaryotes - 262 (source: NCBI BLINK).	MREPLLTSNGGLKTLPIIANEAFEKAASFGLLP NMILYLMRGYGMELATGSIFFWNAATNFTPVIG AFLADSYVGRFWMIGFSGIFSFLGMIFLWLTAIP GATPCDDSSGECKSWNVFQLTLLVSFAFLMSIG AGGIRASLAFGADQLLKDINFSDAGILERFLNWL YYSACCGAVVGVLFVYQDHFAGWVKVGFVSV VIMLLSALSFSSASTLYIKFKAKAGSLIAFAQVILW SWKRNLFSEDELYYQSKGSRLLIPSDKLRFLN KACVGNAAEELTPDGVVAIDPWNLCTVNQVFEVK ALIKVPMWSTGIMMSVTTTQTFPPVLOAGSMR HITSNFQVAGSFGFTVTSIILWIAIYDRIVPLATK ITGKPGQLSVRRMRMGVGLLSFMSVMVCAIEGVR RETAIKEGFSDPKAVTDMASAYWLLFHHWISGM AEALFLVQNEFFYSEFPKSMSSVATTISGLSTS VASLFAJLIMSSINSITSKGGKESWSSNINRAHF DYYYWLLAGLSFLNLYLLOSRAYGPPSPYDQDR FEDNEEIEEG
45	sp_080440_xyfo	0.3510733	0.3502385	FUSED (FU); CONTAINS InterPro DOMAIN/s: Protein kinase, ATP-binding site (InterPro:IPR017441), Serine/threonine-protein kinase domain (InterPro:IPR002290), Armadillo-like helical (InterPro:IPR011989), Armadillo (InterPro:IPR000225), 26S proteasome non-ATPase regulatory subunit 5 (InterPro:IPR019538), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271), Protein kinase, catalytic domain (InterPro:IPR000719), Armadillo-type fold (InterPro:IPR016024); BEST Arabidopsis thaliana protein match is: serine/threonine protein kinase 2 (TAIR:AT3G08720.2); Has 138334 Blast hits to 135590 proteins in 4834 species: Archae - 164; Bacteria - 15793; Metazoa - 51363; Fungi - 13983; Plants - 33508; Viruses - 544; Other Eukaryotes - 22979 (source: NCBI BLINK).	MGIASNYHMEPVGEGSFGKYVYKRRKCTGOV VAMKFIKNGKTEKDLGRLQEQEIEILKXKHENIE MLDSFETAEEFCWTEFAQGEISKLEDHEVLPE AQVQAIKQLVSALHYLHNSRIVHRDMKPNILIC SGSYVKLCDFGFARAMSTKTVLLHSFKGTPLYM APELVREQPYNHTVDLSLGVILYELFVGLPPFS ADSFYSLFRLIVKDPVKYPENMSLEFKTKLGLLN KIPQORLTWPTLLEHPIGSHLLKDLPTAKAA LKGKTPVILHPDTONFESGOALDRFENNSSTLE GAKTIGQEMCALELILEPVMMLSKGTHNCCCPSL VQADIRDCDGLPDVAGGSDKVVETITRAFVNSEPV QDAMWYLCQCKDEVLLATIQLLHRCCHMSGIL SNVLCGFPRSLPAAATISGEKTFYSEIFSIQCG SSLNKLPTSEAIKGFYTNISGLALHSCLLATI AQSFKSAGNSAALITTSSEIQOSRSLSSACLFS FNDGAVTSLQPHCASAMAFASIIUESGASVNSSI MLERHVNCINDISGAHLKLLKSWCWPGGGGLN GVIPEIDAAIDELAFPPVATENSLGYPSYVTSMYS GIPLNVSSPAANLYMEEGDNKKEIKENMKYVEIL IENLYKIERADILQRLKELCCDDCAIRAKVGNVAV GNMCLHSSYFYHLLGKYGIISILLIDRCTDSDLOTR NFACFVAGNAAFHDDLVEELRRSIPQLTNLLIST EKDDIKVNAAGTLYSLVRHSNKLGDDEMSGAIQ ALLKVVSDCESKALNSLKMDAIDASPLKNALCAL AKMCTCYPPCRNVIFSSSEFQVIAQLRQSPAPKIA KFARDICQATGTRLEAYT

10

20

30

40

【表 1 4 - 2 0】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_kitg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	sp_107620_kqnh (SOAP1)	sp_107660_kitg (SOAP2)	sp_107670_ptqx (SOAP2-like)	cytochrome P450, family 71, subfamily A, polypeptide 21 (CYP71A21); FUNCTIONS IN: electron carrier activity, monooxygenase activity, iron ion binding, oxygen binding, heme binding; INVOLVED IN: oxidation reduction; CONTAINS InterPro DOMAINs: Cytochrome P450 (InterPro:IPR001128), Cytochrome P450, conserved site (InterPro:IPR017972), Cytochrome P450 E-class, group I (InterPro:IPR002401); BEST Arabidopsis thaliana protein match is: cytochrome P450, family 71, subfamily A, polypeptide 22 (TAIR:AT3G48310.1); Has 32159 Blast hits to 31925 proteins in 1622 species: Archaea - 46; Bacteria - 2978; Metazoa - 11709; Fungi - 6921; Plants - 9489; Viruses - 3; Other Eukaryotes - 1013 (source: NCBI/Link).	MSSLFLQLNLFFHPLSLVIFVYFYKSLSAKPA TAKNLPSPKLPILGNLHQLGLPHRSLOSLSQR YGKLLFIYIGSKPSLVSSADAAQEIIMKTHDMIFS NRPKSLAASRLFYGCKDVAFAFSPYGEYWRQMRCI CVTQLLSNRTKVSFRKIRREEAALMVESIERCGS SVNLFREIMTFNDVICRASFGMKYDHEGGFNL KKLLDKVAMELLGDFCVGDFIPWLSWIDRVSGLEG RLEKVAKPLDAFLEKVIQEHQIRQSTKIHNNELDK VKNVQNFVDILLEVQSENREALPLDLSIKAVILQPS HNPTLKFWSIAPGKISLPRQLQSDRYALK
46	sp_084440_pkxa	0.928x7	0.947395	0.973x5	Leucine-rich repeat protein kinase family protein; FUNCTIONS IN: protein serine/threonine kinase activity, kinase activity, ATP binding; INVOLVED IN: transmembrane receptor protein tyrosine kinase signaling pathway, protein amino acid phosphorylation; LOCATED IN: endomembrane system; EXPRESSED IN: shoot apex, embryo, leaf whorl, sepal, flower; EXPRESSED DURING: petal differentiation and expansion stage, D bilateral stage; CONTAINS InterPro DOMAIN/s: Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase domain (InterPro:IPR002290), Leucine-rich repeat-containing N-terminal domain, type 2 (InterPro:IPR013210), Tyrosine-protein kinase, catalytic domain (InterPro:IPR020635), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009); BEST Arabidopsis thaliana protein match is: Leucine-rich repeat protein kinase family protein (TAIR:AT5G35390.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archaea - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI/Link).
47	sp_085210_mwkt	0.970198	0.945105	0.985154	MTVIFILLTKYRSPKYGTGKKTARDIETAPHEKAK LKQSNHGDTYPNNQARALERATSYKKSQNGKLC FLRNDROKFEHQDQLLKASAEVLLGSGSFGSSYKA AIMGGPTFVVKFRQMNVMNGKEEFQELIRKLGRL RHPNLLPLVAFFHFKDEKLIISDFVENGSLASHLH GRHTPDQPGLRWSMRLKIKGVARGLAYLYRELP ELSLPHGHKSSNVLLDEKFEPLLSDYAFVPLNK EHGEQVMAAYKSPFTQTNAVTKKTDVWCGLILI LEIL.TGKFPANY.LKHGGGNDLAWVDSVWKEE WSSEMFDMEMFRASRSCKGEMLLKIKIGMWCCCE WNEDKRWSLRDAVEKIEELKEMDDDDNLDADFPSSG FSYSARFSFSSDSSYSRIASITTEQV

10

20

30

40

【表 1 4 - 2 1】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
48	sp_085340_meek	0.925743	0.9422	cytochrome P450, family 72, subfamily A, polypeptide 15 (CYP72A15); FUNCTIONS IN: electron carrier activity, monooxygenase activity, iron ion binding, oxygen binding, heme binding; INVOLVED IN: response to cyclopentenone; LOCATED IN: endomembrane system; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Cytochrome P450 (InterPro:IPR001128), Cytochrome P450, E-class, group 1 (InterPro:IPR02401); BEST Arabidopsis thaliana protein match is: cytochrome P450, family 72, subfamily A, polypeptide 13 (TAIR:AT3G14660.1); Has 31875 Blast hits to 31741 proteins in 1631 species: Archaea - 65; Bacteria - 4943; Metazoa - 11298; Fungi - 6149; Plants - 8001; Viruses - 3; Other Eukaryotes - 1416 (source: NCBI BLINK).	MIEIGYIVKVICLVIVRWVWVIVNWNVWVWFTPKRL EKFLRKOGLDGNYSYRFLLDGLKDMSKMRKEARQ KPIPTFDHFFHRLILPFHNNHFNKYGESFFSWMGPI PVNVAEQEVKGVFTRIKFQKAKLNPLVALLY PGLVSAEGDKWVKHRKLNINPAFHMEKLMHPA FGASVLDVMVNWKEKIVSKTIGSSEVDVWPFVSSL TADAISRAAFSSYDEGRKIFELVLEQTEITLRLLO SVYIPGWMYVPTKTRRRMKTVNSEIQNLLTGIVK RKKAMEAGEAAKDDLLGILLESNYKDTENVLSNK KKLSMTFOELIDECKLFYLAGQESTSVLLAWTVIL LKGHTWQARAREEVVATFGKNEPDEFGLNHLK IVTMLNEVLRLLPYVYCTITRKNRNFHDYQLGNTV PRGAMVTMSAYRIQRDPKIWGGDDAKEFNPRFES EGYAKATKGNIAFFPFGWGRICIGQNFALIEAK MAYSMVLQRFSEFELSPSYTHAPTITLITLQPOOGA HLIHKLRD MEDDDLDQDFINMCGNIANNDEFMSNDELLYE FQTAADIVAEPPVPTVINLVRVFFGLIYVMKEGI GVYILFILFGSIF
49	sp_085760_qcir	0.935753	0.928625	uncharacterized protein	
50	sp_087690_iwaxz	0.928753	0.963424	lactoylglycylthione lyase family protein / glyoxalase I family protein; FUNCTIONS IN: calmodulin binding, lactoylglycylthione lyase activity; INVOLVED IN: response to cadmium ion, carbohydrate metabolic process; LOCATED IN: chloroplast; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Glyoxalase I (InterPro:IPR018146), Glyoxalase I, conserved site (InterPro:IPR004361), Glyoxalase I, bleomycin resistance protein/dioxygenase (InterPro:IPR004360); BEST Arabidopsis thaliana protein match is: Glyoxalase/Bleomycin resistance protein/Dioxygenase superfamily protein (TAIR:AT1G67280.1); Has 35933 Blast hits to 34131 proteins in 2444 species: Archaea - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	MESASETKESASNNPG-LHSTPDDATKGYIMQQT MFRKIDPKLSLDFYSRVLGNLSLLKRLDSDPTMKFSI YFLGYEDVASAPSPVDRTVWTLSSKATIELTHY WGTESDPEFKGYHNGNSEPRGFHGIGITVDDTC KACERFEQLGVFEFKKPPDDGKMKGITFIKDPDGY WIEIFDLKTIRSVAAAS

10

20

30

40

【表 1 4 - 2 2】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
51	sp_088530_wpzu			<p>polygalacturonase inhibiting protein 1 (PGIP1);            FUNCTIONS IN: polygalacturonase inhibitor activity;            INVOLVED IN: signal transduction, defense response;            LOCATED IN: cell wall, plant-type cell wall;            EXPRESSED IN: 20 plant structures; EXPRESSED            DURING: 13 growth stages; CONTAINS InterPro            DOMAIN/s: Leucine-rich repeat-containing N-terminal            domain, type 2 (InterPro:IPR013210), Leucine-rich            repeat (InterPro:IPR001611); BEST Arabidopsis            thaliana protein match is: polygalacturonase inhibiting            protein 2 (TAIR:AT5G06870.1); Has 1807 Blast hits to            1807 proteins in 277 species: Archae - 0; Bacteria - 0;            Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0;            Other Eukaryotes - 339 (source: NCBI BLINK).            non race-specific disease resistance 1 (NDR1);            CONTAINS InterPro DOMAIN/s: Late embryogenesis            abundant protein, group 2 (InterPro:IPR04864);            BEST Arabidopsis thaliana protein match is: Late            embryogenesis abundant (LEA) hydroxyproline-rich            glycoprotein family (TAIR:AT3G20590.1); Has 471            Blast hits to 471 proteins in 25 species: Archae - 0;            Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 471;            Viruses - 0; Other Eukaryotes - 0 (source: NCBI            BLINK).</p>	<p>MPSNRDEMSLFSVFILLLLSSSTISADRGHPDDKK            ALFRIKAFNNAYIFASWTHDTCDCCEWYLVKCCDD            TTHRIYSLAVSSDDKVSQHPIDAVGDLPLYLESILTF            LRLPKLTGPIQPSILKLNLFWLSHNNLTGPIPD            FLGYLPKLTYNLSVKNLTGHPSSLSKHLALGALF            LERNYLSGPIPEFSGFGKNNPFDYLLKAKNQLSG            PIPKSLGGINFLLDISRNMLTGDASVFFGFRNKL            QTVDISRNKFSFDLTKVGLPKTLNLTLELSHNNIYG            NLPPTLAKLPMLQRFNVSYNRLCGKIPTGGDLNR            FDKYSYAHNKCLCGSPLPPCK</p>
52	sp_089620_qnin				<p>MGVEEGGCCRCFCFSGLTALFLWLNKTTN            PTCSENFVDVFTLNKTANTTFRNNHTIRYDLKLN            KENKDKGIYDNLNTFYKPNLSFVTLGNATFNP            FYGGHGKNTHRISVIDARGVKVENATATMAEFR            VEMNTSVRFKIFWTKRHRDLHADLLVNEQGIL            VKKKKKKGIKFKSGVWNSKVCNFYLLGILGIFTYIL            LW</p>

10

20

30

40

【表 1 4 - 2 3】

A	B	C	D	E	F
1	分析に用いたペイトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa_seq MECVVQFLCTEAVKPTWIEVYVSRVTAERFSRFGFE KDLKLRKIQISNIVLHMTPTNTNINSRDPVEL WVRNVKVIHNVENLMDVLSYEKQHVVMKFRLL RTSFNGKTKLLFWPKRVYFRLLKIGKWSIRNTN IEEALKEAETISQRLGLTHSPSLPAATNIITSAEN RLQQLRRNVDMISQIVGRSDVDVTVVVKLQNPFSH DFYVSSNRLSLIAIVGIGGIGKTVVAKVVFEDKR VMDNFERIWWVAVAFVSSEGLNKMLEYMSN NQENWADTRSSLEICIGGSHESMILVTRATVA ETMHAFVHSLPESEDDSWSLFNRAVAFAMSP EYISILETTGKVVQKCKGVPKAIKISGGMLQKRS HPYEWKIERSEIWDLLEGGDTGILPSLMLSXKLN SSSFLCKCFSCATLPKSSHMAFVSXKHELLLE ADNQVTAISDVHRHAIYEHKELKTDWCWDEKVAI ENLRTLAYKNSPGGLLKHVKYLFVLDLEDMLK EVPFCIAGLKCKFLGLVNNPIKEFEFETKLFNLQ SLHFNECDQIKVEVFGGLSNLYNLRHLDPGLKLP RGVIGQLTCLQTLPIGVREGGQISELGGPLPHIS DKLEIQGLLEIKSKEDAKSASISEKSRVKKLRFW NTDNDDDWVHHEEVLGDKPNRNLRLDIQGYG GENFPFWLMPFMDMIKIELTDCRRCRKLPTLGH LHSEKLEKIKNMESELCIGVELYCYTYTVGASSIS SSSSNSSEVKYFPSLKQLQLLLEKLEKTEWIEPS SGEWNIFLLEELQMEDCPKELKTLKYKFSLEHL TVKNITSGLLSKILSNKGSVTCIKIVQIEGITEID LPEELAEYCTSLKTLRHRCQNLRLPESLSKLISL EVLQIKVCPNLVSLPSLSSLETLELEFYNNVGLTS VPEGLQCKSLERLWVDCCPKVELPDLTGLSH LCDLNIRMTKMINOPPDWLYRPLCLTKLTIGGYL GRNEFPDMFLWKITSLEALNLCGWGKLOSLPE QIQHLLTKRLGIYDFKNLEYLPDWLGNLSSLDAL STTRCENLRKLPTEAMKRLTKLSYLTILECPLE MARRTRGPKGKGKHEEGEREGEREAGLGG STPLRTPKGGGGGLQRRYVGEGRERKGNAN REGMGVSRGRIKVMVSDGGGSGPGRIRAEVGFDD DGGGVVS MKTEIGDITSSPEOPSLCYSSVOKSSSNVGVCRTL FSNIGR
53	sp_090460_plyf	0.322237	0.353825	NB-ARC domain-containing disease resistance protein; FUNCTIONS IN: protein binding; INVOLVED IN: apoptosis, defense response; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAINs: Leucine-rich repeat-containing protein (InterPro:IPR015766), NB-ARC (InterPro:IPR002182), Leucine-rich repeat (InterPro:IPR001611), Disease resistance protein (InterPro:IPR000787); BEST Arabidopsis thaliana protein match is: LRR and NB-ARC domains-containing disease resistance protein (TAIR:AT3G14460.1); Has 25652 Blast hits to 19792 proteins in 681 species: Archae - 38; Bacteria - 1467; Metazoa - 4997; Fungi - 304; Plants - 18213; Viruses - 0; Other Eukaryotes - 633 (source: NCBI BLink).	
54	sp_090690_mjc	0.322237	0.342304	uncharacterized protein	
55	sp_095820_trmg	0.322237	0.333948	uncharacterized protein	

10

20

30

40

【表 1 4 - 2 4】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
56	sp_096930_munp 0.965354	0.931383	0.975453	uncharacterized protein	MGLPRFVILGRPTGGVTLRYVVERGAQHRRLL HLSSPEVISERNFTKLEVVVRAPPTIGRDMVHIRSC YNNKFLVLESEGSNVFAEAEDEPEVDRSIWSTL FRVTVNNRLLSLLQSSNAQHSQVLSNSSLGV RNSPFAYWFDVTNMTLAILPEVIAFKGDNNRYL NSRHRTQNRLLKYESNTLLDNGVHHQTQLRNG NVRFRSIFGRFRWRKSNPEWIFGSDDDRTGSDI NTVFRPKLDGNKISLLNLGNRYCMMRRVGGGQP KLVDLTLASSSTVGPLTVMVEEAVRQREIYNIHY NMDESRYNVVAFESGNTQDNTYNTLPLPVKR VGFKYQVTRTFNSSHWSKVAVYNSKIKVTVPK FVEGEATLEASYDGKVEWSELNETITQTKDLFIP VPPMTKSIVTLILGRGSDVPFSYEQRDILPNGRE VHKHMEDGLFSGLNTVEVFTHTVREISLSTLSESE LEEARRLSVDEEAAAPALSSKL
57	sp_097340_nzpp 0.975407	0.953393	0.975906	Class-II DAHP synthetase family protein; FUNCTIONS IN: 3-deoxy-7-phosphoheptulonate synthase activity; INVOLVED IN: aromatic amino acid family biosynthetic process; LOCATED IN: membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINS: DAHP synthetase, class II (InterPro:IPR002480); BEST Arabidopsis thaliana protein match is: 3-deoxy-D-arabino- heptulosonate 7-phosphate synthase 1 (TAIR:AT4G39980.1); Has 3811 Blast hits to 3796 proteins in 685 species: Archae - 0; Bacteria - 1237; Metazoa - 0; Fungi - 110; Plants - 171; Viruses - 0; Other Eukaryotes - 2293 (source: NCBI BLink).	MALSGNTILPTKTFLLPSTTTAKPTTFSLPLKPR HISAVOSSDKNHPITSSKPTIPSAASAATTAISSIA KXSSSEQQWTLSEWRSKKALQLPEYDQKELD AVLKTIEDFPPIVAFEGARILEERLTDALGNFVL MGGDCAEFKEFNANNIRDFTRVLLQMGAVLMY GGQMPVVKVGRMAGQFAPKDFEVRDGVKLL PSYRGDNINGDAFDEKSRIPDQPMIRAYCQSA TLNLLRAFATGGYAAQWVQWNLGFAERSEQ ADKYQELAHRVDEALGFMAAAGLTDHPLMQST EFWTSHEC.LLPYEQSLTRKDSSTSGRYDCCSAH MVMVGERTRQLDGAHIEFLRGIANPLGIKYSDKM DPNELIKLIDILNAENKPGRIITIRMGAEINMRVKL PHLIRAVRRAGQIVTWSDPMHGNITKAPCGLKT RPFDAIRSEVKAFDFVHEQEGSYPGGVHLEMTG QNVTECIGGGQTVTFDDLSSRYHTHCDPRLNAS QALELAFIADRLRNRMRKMSOASLTS
58	sp_098440_zepn 0.965385	0.989339	0.946939	RNA-binding (RRM/RBD/RNP motifs) family protein; FUNCTIONS IN: nucleotide binding, nucleic acid binding; INVOLVED IN: biological process unknown; LOCATED IN: cellular component unknown; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAINS: RNA recognition motif, RNP-1 (InterPro:IPR00504), Nucleotide-binding, alpha-beta plait (InterPro:IPR012677); BEST Arabidopsis thaliana protein match is: penicilliricopeptide (PPR) repeat-containing protein (TAIR:AT5G27300.1).	MPRYKDEPPTVRYVTVCGDESKYLIVRNVAAALGCG DELRLQFAAYGDEECKPMDAEDCEAFTDVMYWI KFHMISNARFAKRLDESFLGNRLQVSYATEYE TLDSTKEKLEGRARKEVLRNLKSLKSGRYGDEM KSKGVETHPLSHGAGVSSQVSSQLDLKGGWYA KEMNQYSEVKHSSITRYSSDQDYFPPSQSMNETV KRVREKLDKIESSSSSNTQSGSASKKIRVDNRRR

10

20

30

40

【表 1 4 - 2 5】

A	B	C	D	E	F	
1						
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	Polypeptide cyclase/dehydrase and lipid transport superfamily protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to biotic stimulus, defense response; LOCATED IN: cellular_component unknown; EXPRESSED IN: 8 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: Bet v 1 allergen (InterPro:IPR000916); BEST Arabidopsis thaliana protein match is: Polyketide cyclase/dehydrase and lipid transport superfamily protein (TAIR:AT5G28000.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).	aa seq	
59	sp_1099260_aa0e	0.9743275	0.9531295	0.9612833	MSSLKRKLEGEIEIRAAAGDVYHDIYKSRPHHLSH ASPHFVOSCDLHEGELGKPGSIWVWYKLGKGLKAL TMKCVIEIIEEKKLVQRVVEGDSLEEYKLVG TCHVIAKDSSETSIVRWMLEYEKMHAGIPEPSALL DACLEAAKDADDHHGIKK	
60	sp_102570_knxq	0.9388267	0.9233259	0.9602318	Aluminium induced protein with YGL and LRDR motifs; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cytosol, nucleus, plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; BEST Arabidopsis thaliana protein match is: Aluminium induced protein with YGL and LRDR motifs (TAIR:AT5G43830.1); Has 1324 Blast hits to 1324 proteins in 483 species: Archae - 0; Bacteria - 791; Metazoa - 46; Fungi - 0; Plants - 333; Viruses - 0; Other Eukaryotes - 154 (source: NCBI BLink).	MLAVFNKVEAKQPEELKNAESLGNVKEDALIQQF SSVHPNALSVHLGDSGFIASFSSDKQNP LLPRLFA VVDDIFLFGHIEINIPSLKQYGLTKTANEVSIIE AYRSLDRGSSPADHALRNIEGKFARFYDSSIKT TFFSVDTDGVSPLFWGTDSEDLHLVLSDIPEVMQ NACGRSFAPFPKGGFFTSAGGLKSYEHPRKVLK AVPWVDSSGOACGSGTFMVVTEESNKENKQLGIKS GMFPRVGSADADWSKKI
61	sp_104090_cxdg	0.9707727	0.9352415	0.9414324	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G02370.2); Has 122 Blast hits to 122 proteins in 54 species: Archae - 0; Bacteria - 0; Metazoa - 74; Fungi - 4; Plants - 41; Viruses - 0; Other Eukaryotes - 3 (source: NCBI BLink).	MGGRREAEELNSLGGLSDSSEYSSQDSADDEW CSEAGSETKQCRSVLWACWKKDETMGAEYMN HKGNLWKTGINRNGKMYLSJEETLFIQIGALLV DDKDDTTFLSEVYPKVAEEKSGCSWEGFKYQHL KSLGYIVGRVGPWSQKSTKNCTMWHEDTSETD NVTQIIFRTSIVQPLNLSLNEANPLSEDELPNKF KKTFPEESKEVHLPVYEVYLPNSOFKKTCPGDP DFVVYPTSGHPFVSFDIKELERQSNGLPIKIVGHV DEGRVLSFSFDKVELPVL
62	sp_107890_mqkw	0.9455231	0.9452723	0.9280399	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G22758.1); Has 131 Blast hits to 131 proteins in 17 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 131; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink).	MSPENTYGHLIKAVLEVEFEKERRRRLNENHPSC FELHYSPFSLQNIYSCQIKIQKIGFEFVSRRSIKF GSTESDADT

10

20

30

40

【表 1 4 - 2 6】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたペイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MIRFGQLKMSIEIGIYKQPVTIYLHCSVWDQVA EELSPIKRLGAPRVAAMLRRXIGKMFYKTLCKCK FRAGVCPVYVYTNCFNFAHGLEELRKPFPNWEIVA AHEEERGVTVEQREEHQIPTISTSLRNESPRTY KGRHCKFYVEEGCPYGENCTLHDEQSRARES VAISVNPTVGGGTAYGANGPTQKPSNWKTRICN KWETT GSCPFSGKCHFAHGLAELHKFSGLDGEG RDSKMPAETKPSGMMTKDTDLTSPNLLTLPRQEY HHVGIPISSSSVYVQAIVGQRAAHKWKGPDKISR IYGDWIDDIE
63	sp_107930_fsnu	0.944987	0.943236	0.938656	Zinc finger (CCH-H-type) family protein; CONTAINS InterPro DOMAINs: Zinc finger, CCHH-type (InterPro:IPR000571); BEST Arabidopsis thaliana protein match is: Zinc finger C-x8-C-x5-C-x3-H type family protein (TAIR:AT2G35430.1); Has 13421 Blast hits to 5509 proteins in 634 species: Archae - 13; Bacteria - 2784; Metazoa - 4906; Fungi - 574; Plants - 3348; Viruses - 263; Other Eukaryotes - 1533 (source: NCBI BLink).
64	sp_112710_rsf	0.934036	0.93387	0.930249	unknown protein; INVOLVED IN: vegetative to reproductive phase transition of meristem; LOCATED IN: endomembrane system; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G40640.1); Has 103 Blast hits to 103 proteins in 14 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 103; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink). TINY2 (TINY2); FUNCTIONS IN: DNA binding, sequence-specific DNA binding transcription factor activity; INVOLVED IN: regulation of transcription, DNA-dependent; LOCATED IN: nucleus; EXPRESSED IN: shoot apex, embryo, hypocotyl, leaf; EXPRESSED DURING: LP, 10 ten leaves visible, D bilateral stage; CONTAINS InterPro DOMAIN/s: DNA-binding, integrase-type (InterPro:IPR016177), Pathogenesis-related transcriptional factor/ERF, DNA-binding (InterPro:IPR001471); BEST Arabidopsis thaliana protein match is: Integrase-type DNA-binding superfamily protein (TAIR:AT5G25610.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).
65	sp_112720_jcmg	0.967237	0.940704	0.93263	MADGHERTEPETSQNDTTSSTSSQTPSGLPNNK FKRNSKDPVYRGMRAWKGVWVSEIREPRKK SRWLGTFNTPMAARAHDAVALCKIGNSAILNF PELAKSLFRFASNSPRDVQAAAALAAMVSPSP SQSPAVSSSSSEVNGPEEELSQIVLPSLEDEGA DSAESMSELVLYDSVEEVAWYSSVWTDMMMLRD FLYESTADISLADYGVGVYTY

【 1 0 2 6】

10

20

30

40

【表 1 4 - 2 7】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_krig (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
66	sp_113700_suxh			UDP-Glycosyltransferase superfamily protein; FUNCTIONS IN: transferase activity, transferring glycosyl groups; INVOLVED IN: response to salt stress, N-terminal protein myristoylation; LOCATED IN: cellular component unknown; EXPRESSED IN: guard cell, CONTAINS InterPro DOMAIN/s: UDP-glucuronosyl/UDP-glucosyltransferase (InterPro:IPR002213); BEST Arabidopsis thaliana protein match is: UDP-Glycosyltransferase superfamily protein (TAIR:AT4G27570.1); Has 5043 Blast hits to 5008 proteins in 241 species: Archae - 0; Bacteria - 56; Metazoa - 380; Fungi - 15; Plants - 4575; Viruses - 3; Other Eukaryotes - 14 (source: NCBI BLINK).	MGGEKLRIVMPWLAFGHFIPYLH-SNKLAEKG HKITLLPNKARLQLESNLHPSLITFHSHITVPPLET LPYGTETADISLDQHGEISISMDRTRPEVESFLS THKFDLVLVDMAHWPEIAAKVGIKSYVNVVCA IAVSHVRPSP_LPKGTAAHVPLPLSSVPKWLSLQ HGSSTPYFEGEGITLLERSVISLSSADAIAIRTCREI EGVYCDRVAATFNKPVLVTSHALPDLLELSPLE TRWAEWLARFEPGVSIFCCCLGSSQHVLDAPQLQE LALGLEMTGLPFLMAVKPPVGGTSLSEEVLPFGFN DRVSGRGVHGGWVQQQQAHPQLSGLGCFVTLG GSSSMWEGLVSEQLVLLPQLADQTLYAKLMAD ELKVGKVEREENGWMTKRSLCEAIKSVMDSDS DISHVVRKNHAKYRSMILISPGFISGYYIDNFIKDLQA LVP
67	sp_114580_dhwu			Heavy metal transport/detoxification superfamily protein; FUNCTIONS IN: metal ion binding; INVOLVED IN: metal ion transport; LOCATED IN: cellular component unknown; CONTAINS InterPro DOMAIN/s: Heavy metal transport/cetoxification protein (InterPro:IPR006121); BEST Arabidopsis thaliana protein match is: Heavy metal transport/detoxification superfamily protein (TAIR:AT3G05920.1); Has 700 Blast hits to 667 proteins in 32 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 700; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).	MDDNKGILTVGVEVEPVIKAIKKGIGQTPEIVSV GPPKPORRRGPFYNGSSCRYSYDPQKQGWDTA MYPCS_LPCNFSCKECS

10

20

30

40

【表 1 4 - 2 8】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
68	sp_115130_uqnn	0.977524	0.8477397	0.823588	<p>starch branching enzyme 2.2 (SBE2.2); FUNCTIONS IN: 1,4-alpha-glucan branching enzyme activity; INVOLVED IN: amylopectin biosynthetic process, starch metabolic process; LOCATED IN: chloroplast, chloroplast stroma; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINs: Glycoside hydrolase, family 13, N-terminal (InterPro:IPR004193), Immunoglobulin E-set (InterPro:IPR014756), Glycoside hydrolase, catalytic core (InterPro:IPR017853), Alpha-amylase, C-terminal all beta (InterPro:IPR006048), Immunoglobulin-like fold (InterPro:IPR013783), Glycosyl hydrolase, family 13, all-beta (InterPro:IPR013780), Glycoside hydrolase, subgroup, catalytic core (InterPro:IPR013781), Glycosyl hydrolase, family 13, catalytic domain (InterPro:IPR006047); BEST Arabidopsis thaliana protein match is: starch branching enzyme 2.1 (TAIR:AT2G36390.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archaea - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).</p>	<p>MVYISGIRLPSTVPSIHINKTSCFHGDIRRNSFS  LF_LNKKSSASLSRKIYAKKSSYSDSDSPSATLADS  QKILVPSGEAEGSLSDLESLTSDPNLNVPEGFHDF  DVQTMKEGSEADKHGSAVATRSBDKDVQNF  VTRRTVPPGNGQKIYEIDPLKDKHRTHLDRYS  MYRKLREINKYEGGLEEFSSRGYEKMGFNRSAE  GITYREWAPGAKAASLVGDFNNWNSNADVMTR  NEFGVWEIFLNNADGSPAIPHGSRVKIHMETPS  GIKDSIPAWIKFSVQAPGEIPYNGIYDPPPEEKYI  FKHPQKRRPKSLRIVESHIGMSSTEPVINTYANFR  DDVLPRIKRLGYNVQIMAIQIEHSYYASFGYHVT  NFFAPSSRFQTPEDLKSIDRAHELGIWVLMDIVH  SHASNNVLDGLNMFDTGDSHYFHSARGYHWM  WDSRLFNYSWEVLYRLLSNARWWLDEYKFDG  FRFDGVTSMMYTHHGLQVSFTGNYYEYFGFATD  VDVAVNYMLVNDMIHGLFPEAIAIGEDVSGMPTF  CLPTQDGGVGFYRLHMAIAKWIQLLQDDED  WRMGDIHTLTNRRWDENCISYAESHQALVGD  KTIAFWLMDKMDYDFMAVDPRPSTPLIDRGIALHK  MIRLITMGLGEGYLNFMGNFEGHPPEWIDFPRG  DQHLPTGKVVPGNNNSYDKCRRRFDLGDANYLR  YRGLQEFQDQAMQHLNENYGFMTSDHQYISRKSE  DDRIIVFERGNLVEVFNHFWNSYFYDQVGCCLKP  GKYKIVLDSDDPLFGGFNRIDHSAEHFTFEGWYD  NRPFSFLYAPCRAAVVVALVEDEAAHALVEE  DSEFGHALVEDDPEPAMQALPVKDETKPAEE</p>

10

20

30

40

【表 1 4 - 2 9】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
69	sp_115320_inco	0.343688	0.922527	0.3993188	<p>Leucine-rich repeat protein kinase family protein; FUNCTIONS IN: protein serine/threonine kinase activity, kinase activity, ATP binding; INVOLVED IN: transmembrane receptor protein tyrosine kinase signaling pathway, protein amino acid phosphorylation; LOCATED IN: plasma membrane; EXPRESSED IN: 13 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Leucine-rich repeat-containing N-terminal domain, type 2 (InterPro:IPR013210), Leucine-rich repeat (InterPro:IPR001611), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: Leucine-rich repeat protein kinase family protein (TAIR:AT3G47580.1); Has 213556 Blast hits to 133070 proteins in 4776 species: Archaea - 194; Bacteria - 22980; Metazoa - 68117; Fungi - 10236; Plants - 86850; Viruses - 328; Other Eukaryotes - 24851 (source: NCBI/Link).</p>	<p>MKQPPLLVTQAAVILLVVIAMQPHPLRVANAISN ETDYMALLAIKSQLIDHPNGVLSWVNSLHCL WEGVICGRKHKRVTLKLFGRGLSGTSPFIGNL SELIWILDNSLSLHSGSIPPEIGRLRLQSLYLPNNT LSGGIPASLSSCINLQNLAINLEGELEPMEIGA LSNLKWLHLHKNKLSGPIKLIINLTSLVRIKGIHT NAFTGPIPNIPKLTNLTSLSELSNQLSGIIPSPISN LSSLOILDLSMNQLSGIIPSPISNLSLQILYLAYNQ LRGSIPIHALTLPQLOYLQLAYNYFSGPLPISHN LTSLSRSELSNYYFKGTISNLHFGHLYNLQRLVISN NNLSGDINFTIKLVNCSKLESLESDVGNNOFTGLVPK VLANLSTNLEIWIENQVOTGGIPVGLGNLNLRLIL SMKSSGLTGTVPDLGELQNLLEIDLDSFNRLTGEI PSSFGNLSHLNLFYENKLGQRPVSSLGNCAN LLYLYSYNQFNGLPIELFAKTTNFQIQLWLDQNH FQGPLTMEISKQINLEVFSMSNNKFFGSPIDVFS LPALQELVMEIDFFHGLIPFSFASLSKLLKLDLSQ NNLSGVPVEYFATFPLIVLNLNHNNEFGRVPTKG KFRMSRALKLTITASAFVGLVMAICWILYLKG VFANSSATSVGNKNCGGSSKHLHPCRVENER KFRMSRALKLTITASAFVGLVMAICWILYLKG QSKKRKSTSSDALVKERFLKYSYDMLLKATDGF SQENLLSGSTFGSVFKGILDKGTAVKVLNLQOR GDSKSFMAECEALRYRHRNLVGLITACSSIDFKR NDFKALVYEFMPNGLSKDKWLYKSGYLSLQRLVNI AIDVAHALNYLQCEETPIVHCDLKPNNILLDNDM VAHVGDGFLAKVLGLOPLHPNQSSSIVRGTGTY AAPEYGLGGEASPEADLYSYGILLLELMTKRRPT DNMFKEDNLFHMFKAALPHQVLIHVDSTLMEDE SDEPKRGPKPHMELRKTTECMVSVKIGVACSS HLPRDRMKISEAISELQKARNILLSPKHKNRNLPRG ESD</p>
70	sp_118860_auyc	0.3938639	0.352687	0.263043	<p>uncharacterized protein</p>	<p>MHTMKPTVSNKAQPEGSIANARIMEECLAFISRY LNGIETKFRMSRNDMDMQESLLFLKSVFKKG NPLGKRTFKQLSFLDWKQAGLYLVMNCQEVQPF IGEYTTIHGPKPSLDWFQSQWIKLYQERDPRVT AELLSLRGPKSVKSYQGYVYNGYRFRHTTKRQ RRRKTQNSGVVWVKDEESEKDFGFLKEVIEL EYDAPNNGDKEPIVWLFRCVWFDVYREGGRIK DKFGGISLNFKKFLGTNEPFAMASQVGGVYVRF DDEAEIKYNEFKLHEKEVGETGADNVSDVAYE KVLGYRSGYARGLKGHPVLSKDEKKGRALELV TVVQLQENNTMRAEFEEHHAETLRKENEAKKK QEELEQLKIMEIKIGLESLT</p>

10

20

30

40

【表 1 4 - 3 0】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
				HXXXD-type acyl-transferase family protein; FUNCTIONS IN: transferase activity, transferring acyl groups other than amino-acyl groups; transferase activity; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: root; CONTAINS InterPro DOMAINS: Transferase (InterPro:IPR003480); BEST Arabidopsis thaliana protein match is: HXXXD-type acyl-transferase family protein (TAIR:AT3G30280.1); Has 2415 Blast hits to 2403 proteins in 134 species; Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 58; Plants - 2353; Viruses - 0; Other Eukaryotes - 4 (source: NCBI BLink).	MGEVNHVEVEIIEIETIKPSSLLPPKTPPKTTLS HLDQAAPLYYYPLLLYYTNTTTTTTPTSQIRVDITST LKTSLSKTLDFHPIAGRCVDDSTICCNHQGPPIE TKVSNILDMVNSPEKMKLLIKFLPHAEFQDVRP VSDNLHAFQVNVFRCGGVIIGSYV LHKLLDGI GTFKKNWSTIANDERVKDDDLVQDFEATKAFP PRATAPMLPRNOQLPKAAEKNPNPVKLVTKS FVFDIVSLKMMFMMAKSELVPKTKFETVTFGIW EGTLSTLRNSGVEVEHTSLIPVIRPRMSPPPLPR GSMGNLLKNAKAGANTSSNGLDLVKHEHSSLS QTTQKINPPPPPPPPPTTTTTHSSLSQTTQKI NTPPPTTTTHSSLSQTTQKINTTTTAEVILTKRK VDNPVTQNRGNYLFTSWKIKGLDEADDFGFGKP VWVFPDGRPKVRNMFILTDYRHPETGVEGIAA WITLEEKQMQCLKSNPEFLAFATPN
71	sp_125800_kzws	0.982317	0.982318		MSLEFTNGFQSNDDRLKEVKAFDAMGKVKGL IDSGVKIPKIFIRP.LDESEDSKTPCVKLVQVPTINL EGIDRKDRYNEIVKEVLDASETWGFQVNVHGIS VELLDEMLQRYRMFHEQNTVEVKEFYRETLLKQ VYVNSNYDLYTSRSANWKD TLVNAVYYGHLDPK ELPEICRDVMDYTNHVMKLSDILLELSVALGLG PDHLREMEFKGWSLVNHYYPACFAPD.LTLAGS KHSDSAFLTILLQDQIGGLQVLHENQVWVWQPKP AAFVNVIGDILQMVSNKLRVHHRVIANLIGPRIS VAFFLRGVLSEKLYGPIKELLSKESPPVREFTL SDFYSYFFSRPLNESGFYFKV
72	sp_126640_aupr	0.932137	0.921087	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein; FUNCTIONS IN: rRNA methyltransferase activity; INVOLVED IN: rRNA processing; LOCATED IN: cytoplasm; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: rRNA small subunit methyltransferase G (InterPro:IPR003682); Has 6840 Blast hits to 6840 proteins in 2435 species; Archae - 0; Bacteria - 5096; Metazoa - 2; Fungi - 0; Plants - 38; Viruses - 0; Other Eukaryotes - 1704 (source: NCBI BLink).	MLCLRSKLIILSPFSFRFTFIEHLPLYNPKPSLNFLLNT RKSLNTAASVKSSHFFETLNP HQKHQISLYVDALIQ WNOQMNLTAATERSEVEMERHIEDSLAILPPIINES YTSHCDSCESSKLVDTGVTGAGLPGJLAIACPG WKVTTLESNMKRCGLFLEHAIDVTGVSNOVVRD RAENFGQTLFRESFDI AVARVAELRYLAEYCIIP LVRKGGLFVAAGKGDHDFQEEIRNAERAMNLLGAS LLHQHLVESHGPFPGQRTAIVCFKDRDLTSSRKYPRS PGTPAKMPL
73	sp_129380_nkfi	0.931436	0.936147		

10

20

30

40

【表 1 4 - 3 1】

1	A	B	C	D	E	F
		分析に用いたバイトとの相関係数			注釈	aa seq
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	Leucine-rich repeat transmembrane protein kinase; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: plasma membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Maleic/retinoid receptor-like protein kinase (InterPro:IPR021720), Protein kinase, catalytic domain (InterPro:IPR000719), Leucine-rich repeat (InterPro:IPR001611), Serine-threonine/tyrosine-protein kinase (InterPro:IPR001245), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: Leucine-rich repeat transmembrane protein kinase (TAIR:AT1G53420.1); Has 185540 Blast hits to Bacteria - 17971; Metazoa - 51218; Fungi - 10628; Plants - 83318; Viruses - 446; Other Eukaryotes - 21837 (source: NCBI BLINK).	MESSFHTVIVFWVLFSLFGLSSELQCRRTAAKEV EALREIAKELGKDWNFNEDPCSNHTSWFSRGG PARQTYNQVTCNCVSSRWHLTNISLKGQDLSG SLPQSLVKPLVYKTDLSRNLLSGMLPDEWATMK LEFVSLMANQLSGQIPKFMGNITLLYLIESNSF SGTIPAEELGKLVNHLFLDNNLNRLGHLPIELTKL ASLTELRVSSNNFSGQIPNYFHHWQLEKLEQIQG SGFDGPIPFSSILLTNLIELRISDLNLEGSEFFSLE NLGLSRMLKSKLYGTIPSIQIAMSRLKLDLDS FNKLSGIPNELQFLNKLQNVYLSNMLSGPIPS WITNKEDSSHIDLSYNHFDESSAPSSCNKETLNL YRFSFRGENLESSVCLKNFKCKARYSLHINCQGA ETTVAKTYQKDDERGGAAKFIIPRAEWGFFSSS GEFWFSQLPDYYIAKNSAIRLKDNLNLMRAFLS ATSLTYASCLGKGNVTVLHFCEIVFANNSYSSL GRRIFDIQDKLVKRNFNIVKEANGTKIVKIVKVFV HIPVSDSLEIRFYAGKSGTGIPVRGTGYPLYSAL SVQSDFKPPLDKKQIIRLSGVVAILLCTTLGGFIL WRKHFRSDIASREQVLAGLDLQTLGFLTFRQIKAA TNNFSDSKIGEDNSQLKLNWLRQKICIDVARG LTFLHEESTIKIVHRDIKATNILLDGLNNAKISDFGL ARLDEGESTHISTRVAGTM
74	sp_135830_r9sg	0.968137	0.981157	0.937834	Serinc-domain containing serine and sphingolipid biosynthesis protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: membrane; CONTAINS InterPro DOMAIN/s: TMS membrane protein/tumour differentially expressed protein (InterPro:IPR005016); BEST Arabidopsis thaliana protein match is: Serinc-domain containing serine and sphingolipid biosynthesis protein (TAIR:AT3G24460.1); Has 814 Blast hits to 786 proteins in 182 species: Archaea - 0; Bacteria - 0; Metazoa - 443; Fungi - 138; Plants - 166; Viruses - 0; Other Eukaryotes - 67 (source: NCBI BLINK).	MAGERTDLRVSEAVYLEELSRTPQKKIDVSEKR KSLKRVYYGIIFLTTFNFAVWLVRYDIQRIPIENH CKLGLXVRDYIQRVIPENHFLRTCGVGGHDCIQTI GVLRIISFGCFFIFLMLFTTLTNKLOEYRNAWHS GEFARIGAGVFLALQLSVIQIFIAWNNYWPDPV KFKQSCSLGLFMSTVFYVASICGVWALYLYVYPRS SCTLNFIFFWTAVLLVMMMLTJHVKVNRGLLSSG IMAAVYVFLCWSAIRSEPADKCSPOKQVTHGH DWTWTFEFGICAVIMATFSTGIDSESFOFRKDEV EEEDDIPYKYGFFHFLVFLSGAMYFAMLFINWDLN SSTRTWSIDVGVWASTWVKINIEWFAATVYMWKLI SPVVRQAKVDEGAIQDQSC MNPFGVPSINTEITKAMSRYDKDSITQEDRAREAT RLIHAEDKNLSALKHALKLSYDNGVSTLCLLYN ATGDTLLELVQDQKDDWLVGYIYKEEPPSSFENGQ WLAFLHAHPTSQAYGSEAAARYRGNKWKGDVVR DFLVAWSTPWSQNSAYTEVREKDHFPQYWD YKVSKLENANKITKDESDEYCASTVSGIGTSPEF JAVLKHKFSPLP
75	sp_140780_kjml	0.944493	0.925374	0.9543363	Serinc-domain containing serine and sphingolipid biosynthesis protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: membrane; CONTAINS InterPro DOMAIN/s: TMS membrane protein/tumour differentially expressed protein (InterPro:IPR005016); BEST Arabidopsis thaliana protein match is: Serinc-domain containing serine and sphingolipid biosynthesis protein (TAIR:AT3G24460.1); Has 814 Blast hits to 786 proteins in 182 species: Archaea - 0; Bacteria - 0; Metazoa - 443; Fungi - 138; Plants - 166; Viruses - 0; Other Eukaryotes - 67 (source: NCBI BLINK).	MAGERTDLRVSEAVYLEELSRTPQKKIDVSEKR KSLKRVYYGIIFLTTFNFAVWLVRYDIQRIPIENH CKLGLXVRDYIQRVIPENHFLRTCGVGGHDCIQTI GVLRIISFGCFFIFLMLFTTLTNKLOEYRNAWHS GEFARIGAGVFLALQLSVIQIFIAWNNYWPDPV KFKQSCSLGLFMSTVFYVASICGVWALYLYVYPRS SCTLNFIFFWTAVLLVMMMLTJHVKVNRGLLSSG IMAAVYVFLCWSAIRSEPADKCSPOKQVTHGH DWTWTFEFGICAVIMATFSTGIDSESFOFRKDEV EEEDDIPYKYGFFHFLVFLSGAMYFAMLFINWDLN SSTRTWSIDVGVWASTWVKINIEWFAATVYMWKLI SPVVRQAKVDEGAIQDQSC MNPFGVPSINTEITKAMSRYDKDSITQEDRAREAT RLIHAEDKNLSALKHALKLSYDNGVSTLCLLYN ATGDTLLELVQDQKDDWLVGYIYKEEPPSSFENGQ WLAFLHAHPTSQAYGSEAAARYRGNKWKGDVVR DFLVAWSTPWSQNSAYTEVREKDHFPQYWD YKVSKLENANKITKDESDEYCASTVSGIGTSPEF JAVLKHKFSPLP
76	sp_146060_hxqp	0.958415	0.928288	0.937738	uncharacterized protein	MAGERTDLRVSEAVYLEELSRTPQKKIDVSEKR KSLKRVYYGIIFLTTFNFAVWLVRYDIQRIPIENH CKLGLXVRDYIQRVIPENHFLRTCGVGGHDCIQTI GVLRIISFGCFFIFLMLFTTLTNKLOEYRNAWHS GEFARIGAGVFLALQLSVIQIFIAWNNYWPDPV KFKQSCSLGLFMSTVFYVASICGVWALYLYVYPRS SCTLNFIFFWTAVLLVMMMLTJHVKVNRGLLSSG IMAAVYVFLCWSAIRSEPADKCSPOKQVTHGH DWTWTFEFGICAVIMATFSTGIDSESFOFRKDEV EEEDDIPYKYGFFHFLVFLSGAMYFAMLFINWDLN SSTRTWSIDVGVWASTWVKINIEWFAATVYMWKLI SPVVRQAKVDEGAIQDQSC MNPFGVPSINTEITKAMSRYDKDSITQEDRAREAT RLIHAEDKNLSALKHALKLSYDNGVSTLCLLYN ATGDTLLELVQDQKDDWLVGYIYKEEPPSSFENGQ WLAFLHAHPTSQAYGSEAAARYRGNKWKGDVVR DFLVAWSTPWSQNSAYTEVREKDHFPQYWD YKVSKLENANKITKDESDEYCASTVSGIGTSPEF JAVLKHKFSPLP

10

20

30

40

【表 1 4 - 3 2】

A	B	C	D	E	F
1	分析に用いたバイトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
77	sp_146920_mcah	0.388811	0.533543	Prolyl oligopeptidase family protein; FUNCTIONS IN: serine-type peptidase activity, serine-type endopeptidase activity; INVOLVED IN: proteolysis; LOCATED IN: chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN's: Peptidase S9, prolyl oligopeptidase, catalytic domain (InterPro:IPR01375), Peptidase S9A, oligopeptidase, N-terminal beta-propeller (InterPro:IPR004106), Peptidase S9A, prolyl oligopeptidase (InterPro:IPR002470); BEST Arabidopsis thaliana protein match is: Prolyl oligopeptidase family protein (TAIR:AT1G20380.1); Has 8127 Blast hits to 8050 proteins in 1368 species: Archaea - 105; Bacteria - 3923; Metazoa - 339; Fungi - 43; Plants - 244; Viruses - 0; Other Eukaryotes - 3473 (source: NCBI BLINK).	MLSRGPVIGSNFQLCAGFLLPRTTRTSLSSSL SPPNKYSSNYFCCHLHKHKPTLLNSSHFLSPQT FSSMASASHKPLHYPLSRDDSVDDYFGVQITD PYRWLEDPDSEETKEFVENQVKLDSVLGECET RDKLKEKMTLEFDYPRFYPPKRCCKYFYFHNA GLOPQSVLYVQDSLDGKPEVLLDPNQLSEDTV ALNTYVSSEDAKYFYGLSASGSDWVTIKVMRIE DKGVEPDSLSWVFGSGITWTDNKGFFYSRYP PNEGQISDAGTETNSLNQELCYHVLGTDQSEDI LCWRDPENPKHSFGAEVTDGKYLKLLSTSECE VVKVYYLDLTLNPIEGFRGGKLLPFTKLIDT FDAAYDVVANDDTVFIFRTNKDAPRYKLVFVDLK SFGIWDVIPQSEKDVLESAYAVNGNHLLVCYLS DVKHVLVLRDLQTSLLHLPLEIGSYDDISARRK DSVVFRRFTSFLSPGIYHCDLKSDVPELKIIFREIV VPKFRDFEFQVDQVFSKDKGTIPMFIVTKKGI LDGSHPCLLYGYGGFNISITPSFGVSRVLLKHLG AVFCIANIRGGEGYGEWFKAGSLSKKQNCFFDD FISAGEFLVSYGYTQPKKLCIEGSSNGLLVGACI NQRPIFGCALAHVGMMDMLRFRHKFTIGHAWTS DYGCSDKKEEFHWLKYSPLNVRPWRDLSSDR TAQYPTMLLTADHDDRVVPLHSLKLLATLQHL CTSLNSPQKNPVGRIECKAGHGAGRPTOKMID EAADRYGFMAKMMVNASWTD
78	sp_147470_tnri	0.988207	0.944002	uncharacterized protein	MANPFGSYVDEKIDEMERYIGTKTQEDRARE AMHLINEDDKNSKAEAYAESYKEYYGGVSTMC MVMYATGDTLTYTNDNDWYGFISRTPTYTEIGN QWAAAFQHVHNTGASGSEAAVYVRGNKDKGHE RDFMLSWSTPWGPWYKKNKAYCEMGGYDSFQS RWDIYDKLNNSGYSDHVRDRDGVKIDVDTATGG APIFIHATIKIPFSS

10

20

30

40

【表 1 4 - 3 3】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
79	sp_148210_yuyx	0.983042	0.938131	PDI-like 1-1 (PDIL1-1); FUNCTIONS IN: protein disulfide isomerase activity; INVOLVED IN: in 6 processes; LOCATED IN: in 8 components; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAIN(s): Disulphide isomerase (InterPro:IPR005788), Thioredoxin fold (InterPro:IPR012335), Thioredoxin domain (InterPro:IPR013766), Thioredoxin, conserved site (InterPro:IPR017937), Thioredoxin-like subdomain (InterPro:IPR006652), Protein disulphide isomerase (InterPro:IPR005792), Thioredoxin-like (InterPro:IPR017936), Thioredoxin-like fold (InterPro:IPR012336); BEST Arabidopsis thaliana protein match is: PDI-like 1-2 (TAIR:AT1G77510.1); Has 31573 Blast hits to 17834 proteins in 2924 species: Archae - 337; Bacteria - 15428; Metazoa - 5451; Fungi - 1644; Plants - 2422; Viruses - 29; Other Eukaryotes - 6262 (source: NCBI/BLINK).	MALKIGLLEVLVLLATIGGIQVEAEKEEYVITLDHNSFTDVSXKHDFFILVEFYAPWCGYCKQLAPEYKAAATLSKHDPPIITLAKVDLNEDTNKEHFKSKFEIRAFPTLKLINGGKIVHDYKNPPQSADDIVGQVKKLLGVPVSAEIKSVNDASKIIDQKVKVIAGVFPKFEGEEYGNFKLAELKRSDFEFGHTLDKAVLPRGDSVKVGVVFLKLPFDLDFADTKDFDLAALAKFVQEEEDVPSVIPFTMDPSNTYLRFKAFSGENTKIMMVLDRKSENAALAKSRFSDAAILLKGNKIKYFLVGDLESKDLKLFKLEDSSHAPLVFIHKTEAQKYLKTNVEPDQLAPWLKFKLDGKLEPFLRSQPIPEENNEPVKVVVLDLSQDMVFNSEKEVEFLFYAPWCGHCKSLAPILDEVAISYEKADANIMIAKFDASANDAPKDTFNVKGFTTMTFFITKSEIVEYNGGRAKEDIVGFIEKHRITTEKNKSEAGIKSEPEKENNHVDETVASKDEL
80	sp_148230_dgra	0.993452	0.976422	cytochrome P450, family 72, subfamily A, polypeptide 15 (CYP72A15); FUNCTIONS IN: electron carrier activity, monooxygenase activity, iron ion binding, oxygen binding, heme binding; INVOLVED IN: response to cyclopentenone; LOCATED IN: endomembrane system; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN(s): Cytochrome P450 (InterPro:IPR001128), Cytochrome P450, E-class, group 1 (InterPro:IPR002401); BEST Arabidopsis thaliana protein match is: cytochrome P450, family 72, subfamily A, polypeptide 13 (TAIR:AT3G14660.1); Has 31875 Blast hits to 31741 proteins in 1631 species: Archae - 65; Bacteria - 4943; Metazoa - 11298; Fungi - 6149; Plants - 8001; Viruses - 3; Other Eukaryotes - 1416 (source: NCBI/BLINK).	MKESSTLRAQALNKPIPFTHDYQRIQPFTHQLINNSCAGAGKNYTWLGPVPTLITQPELIKDAFNRMNFKPRLNPTQILSTGLFNYEGQKWKAKHRKLNPAFQLDKLMPAFETCVTDTLNKKWEKLVSKTGSSEVDYVPHFTLLTGDGIARAAFGSSFEDGRFIDLLKEQKDLVISLLKYSYIPGFKYLPTKGNKMKETENEIKPLLNLHHRKAMEAGEAPKDDLLGMLESNANEARQVNESESRRQRSDITMSFHEMIDACKLFFLAGQETTSVALTWAMILLAKHQDYQTRAREEVLATFGTKTDFDGVHNRKLVITVTHEVLRLYPIIPATRRRAHERETKGLDLVPOQGVGVSFSLHAHLNPEIWGDDAKEFKPERFEGEIAKATKGNNSYFPFGWGRPRICIGQAFALIQVKMALSMLQRFSELSFSYIHAPITSLALQPOHGHAHVHLRHL

10

20

30

40

【表 1 4 - 3 4】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiag (SOAP2)	sp_107670_ptax (SOAP2-like)	注釈	aa seq
81	sp_148240_fnwi 0.943312	0.863363	0.9573	UDP-Glycosyltransferase superfamily protein; FUNCTIONS IN: transferase activity, transferring glycosyl groups; INVOLVED IN: N-terminal protein myristoylation, metabolic process; LOCATED IN: cellular component unknown; CONTAINS InterPro DOMAIN/s: UDP-glucuronosyl/UDP- glucosyltransferase (InterPro:IPR002213); BEST Arabidopsis thaliana protein match is: UDP- Glycosyltransferase superfamily protein (TAIR:AT4G27560.1); Has 4908 Blast hits to 4874 proteins in 244 species: Archae - 0; Bacteria - 41; Metazoa - 215; Fungi - 17; Plants - 4624; Viruses - 3; Other Eukaryotes - 8 (source: NCBI Link).	MCDKSSVLSIAFYWPWFALGHLTSFLRLANKLA QNGHNVSYPITNTLPRLLPHNHYPGHLTRIPVTY PPVDGLPLRAEITNDYVSSAHLMLTAMDLTRDTI EAHLVSIKPDVVFYDFAYWIPDLARKHGFKSVLYI TSYIARCAYFAPDLKSGHQSTGAEIAPPGFPSQ HFRMQAHEAETVADVKGQDGLQGITISERMRIA FGKCDAGVKSCKEMEKVYIDYCEKIFGKSVLLAG PMVPKTPSSKLDYFDGWLTFGGAATVIYCAFQ SECVLEINQFQILLGLELTGRPELVAMKPPKPY ETIESALPEGFEKRTKGRGIVHEGWVQQQLQIH PSVGCFFIT-HCGVGSLSSEAMVSKCOVLMPOAVD QFINARMMISLEIKIGIVEKREDDGLFTKEAVHK AVSLVMEEESEVAKEMRYSHDKWREFLLQEGLE DSYISSFIQSLRQLTIG
82	sp_152910_apyy 0.993371	0.973337	0.954123	serine carboxypeptidase-like 18 (scpl18); FUNCTIONS IN: serine-type carboxypeptidase activity; INVOLVED IN: proteolysis; LOCATED IN: endomembrane system; EXPRESSED IN: root; CONTAINS InterPro DOMAIN/s: Peptidase S10, serine carboxypeptidase (InterPro:IPR001563); BEST Arabidopsis thaliana protein match is: serine carboxypeptidase-like 7 (TAIR:AT3G10450.1); Has 3653 Blast hits to 3603 proteins in 397 species: Archae - 0; Bacteria - 261; Metazoa - 616; Fungi - 829; Plants - 1537; Viruses - 0; Other Eukaryotes - 410 (source: NCBI Link).	FFGNLPPYLETYVRYVGENDEVQLFYFIKSERD PEKDPMLWLTTGGPGCSAL-SGLVLEIVSRCLIN STMSSGRKTDGHRWFWSCVGGCWSLTAHRR RENEKEMERCPNLNFTSACNWHKSFYTFQLNPS SWTKVASIIFLDSVPGTGFYSYATNPEGYTDDITS SRHIYQFLRKWLMDHDGFKDHPFLFISGDSYCGKI VPVVQEISNGNRAGFEPAMNLQGYVLGNPYTRK EDDNKSKFYAHRVLSLLSDDLYESATRINCNGTY VDVDVDDKDNKCTRILQAIISYNLDPLFDAHVLPK CISHQTWCQVYFSCSSDMA

10

20

30

40

【表 1 4 - 3 5】

I.	A	B	C	D	E	F
		分析に用いたベイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiag (SOAP2)	sp_107670_ptax (SOAP2-like)	注釈	aa seq
83	sp_157560_hiix	0.943223	0.922333	0.942034	<p>purple acid phosphatase 26 (PAP26); FUNCTIONS IN: protein serine/threonine phosphatase activity, acid phosphatase activity; INVOLVED IN: phosphate ion homeostasis; LOCATED IN: vacuole, plant-type cell wall; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAIN/s: Purple acid phosphatase, N-terminal (InterPro:IPR015914), Metallophosphoesterase (InterPro:IPR004843), Purple acid phosphatase-like, N-terminal (InterPro:IPR008963); BEST Arabidopsis thaliana protein match is: purple acid phosphatase 12 (TAIR:AT2G27190.1); Has 1843 Blast hits to 1825 proteins in 387 species: Archae - 3; Bacteria - 583; Metazoa - 186; Fungi - 81; Plants - 765; Viruses - 0; Other Eukaryotes - 225 (source: NCBI BLINK).</p>	<p>MGSADRAQRKEEIGLNGADNYGPIWTVVTKRPS TKLKGQGRSANSSYKGANSKINEYRGNKSTPVI NVHQFTGSGNARYMKGONITQLERINSSANLDFQ KNNNDGDFKGYASAKSKGLEEFELGHVPCALTH LVGGSDSPVSSLPFOINSAPVQSNPTPTSHSSST LSPNLHNSVSSVFNNDQPNISISISIGKCOOSIS TNADHISSDNLPAORSSLELDORAKAKHAESS GGERRSEPTVQNPQSVPTTSDNDEPMFIDGGE KSGEGNCMHAGEYSIGSAEERDNSTQEANASKG GSEYLTSHASFTSSYRLDPPSTDIPLDSEFFAAP EAYNAPQGVHITQGDYDGKAVIISVWVTPYLHGSH TVYGTGKGYTSKAKATNTSYSNFYKSGQIYH CLVDGLENTKYRYRIGSGFSPNREFWFQTPPEV GFDPSPYFGLIGLDLQTMNSLSTLQHYMKGEGQ TVLFVGDLYADRHIDEAPHEGGLRWDTWGRLV EPSTAYQPWWTGCGNHDIDYFPKLLGGEFMQF RERYSPYKASESTPLWYAIRRGPAAHVIVLASY SGFVYTAQYRWLDNELKVKDRKKTPLWLVIMHV PLYSNTHAYMEGESMRVFEQWFVYDYKVDVW FAGHVHSYERSHRVSNLKNISNGACSPVLDESA PFYVTIGDGNSEGIAGEYISQPHYSAFREPSF GHALDIKNTKTHAYYYWHRNQDGNKVTADSFVLH NQYWTNYV</p>
84	sp_158080_oxzh	0.932227	0.935311	0.956383	<p>monogalactosyldiacylglycerol synthase 2 (MGD2); CONTAINS InterPro DOMAIN/s: Monogalactosyldiacylglycerol synthase (InterPro:IPR009695), Glycosyl transferase, family 28, C-terminal (InterPro:IPR007235); BEST Arabidopsis thaliana protein match is: monogalactosyldiacylglycerol synthase type C (TAIR:AT2G11810.1); Has 1572 Blast hits to 1572 proteins in 585 species: Archae - 0; Bacteria - 1385; Metazoa - 0; Fungi - 0; Plants - 105; Viruses - 0; Other Eukaryotes - 82 (source: NCBI BLINK).</p>	<p>MVMSVIAPKRSSLSVFSRVGGYYSKKRCSSQ FYPSSSGISSSSSMVIDEMEDDEDGSTMEMW QLGADRTRKNVILIMSDTGGGHRASAEAIRDAFNL EFGDEYNIHKDVKWEYTGWPLNDMERQYKFW KHVQLWSLAFHSTSPKWIHSVYVLAIAAFYAKEY EAGLMEYKPDIIISVHPLMQHPLWVLRWQGLQK KVIFVTITLNSCHRTWTFHPAVNRCYCPSEEEVA KRASLDGLEDSQIQVGLPVRPSPFARAVL_SKDDL RVELEMDPLPAVLLMGGGEGMPVQKTAKALA ESLFDKLEKPIGOLIIICGRNKS_LAST LESKEWKI PVKVRGFEKQMEKWMGACDCHITKAGPGTIAEA MIRGLPILNDYIPGQEKGNVYVYVENGAGVFTRS PKETARVAEWFSFKGDELKRMSENALKLAKPEA VFEIVRDLHELADQRGPLANLPYMLTSSFTSLI</p>
85	sp_166530_hghx	0.970423	0.969423	0.930762	<p>F-box family protein; CONTAINS InterPro DOMAIN/s: F-box domain, cyclin-like (InterPro:IPR01810), F-box domain, Skp2-like (InterPro:IPR022384); Has 68 Blast hits to 68 proteins in 20 species: Archae - 0; Bacteria - 0; Metazoa - 4; Fungi - 0; Plants - 59; Viruses - 0; Other Eukaryotes - 5 (source: NCBI BLINK).</p>	<p>MLVCKQWKDIAEDDYLWKKLCARRWPSLCKRP SPSTMKYKLYGTFHKRQRRLDLPPLRSLFNDL DFYIDIWAEQGSIFSEVSGPVLHAGIKSRPIGID IVRCKDGPYRMLPVEPRFSPVLGQTASYSVVL VERKDFNKACINKLSFDYVDRTSFLAIAFDYLDL APIHFFISGIRAWISLFTDGDENRDRVADVFGEIMD FCDAASNETEYVLLDMLDWK</p>

10

20

30

40

【 1 0 3 5】

【表 1 4 - 3 6】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
86	sp_168890_mkyl 0.955238	0.93422	0.975875	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G14850.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	MVVRMNDIKMPGDRHFRIDILEIKLQISRKLGN LRAEKYFGLDRFFGLCKSEFDKLCIVAVGRE NISLHNLKISLNACLSKSPQKRSKLETPIAEK VANGFQRSCLQSLCRDALPPSPRRGRIPGFRDR KSKDRQSPGPNKSNNTVEAKDLGHKTLQEQS ATELLS.GSRPIEVNSVEDGEEVDQAASPGIHSR SPVTPPLGIPWNGKGNRKYVLSLSSSGNLESC HGNGELPNTSCLRKRLEQKLEKEGLGISQDCVNL LNDGLDLYLKRLIKPSLELANSRNGNKRQKVGGE DVHVHGLNGLWPVRYLQSPVSSSVTILDFRFA MELNPEVLGEDWPTQLKLEKCLREFESEYS
87	sp_168980_dmck 0.991194	0.981698	0.993833	FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: cultured cell; CONTAINS InterPro DOMAIN/s: Golgi apparatus membrane protein TVP15 (InterPro:IPR013714); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MRNAEGEISEQRNPPSPSPHQSRRTSNTK PDAKLRNRNAPFLIVCKCFVITALSALCIANVLS AFRSFKNGSDVDFGFCYAYALAFVVAETEW SLIFRFWKVLEWPGRSMLQIFVAVMTRAFSDFI GKQRDIVLQSIASVLLSCGVVYVLSGILCIGVIK RSLEEKEISRDCGAVKLEELEKRRHELEALLIEEE NV
88	sp_169540_akgo 0.953251	0.921448	0.938737	EARLY LIGHT-INDUCIBLE PROTEIN 2 (ELIP2); BEST Arabidopsis thaliana protein match is: Chlorophyll A-B binding family protein (TAIR:AT3G22840.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MAMSMQVQALLSNSVGFSTGSSSKAYPVGFNP RIAKRVNFGVKMAEEKVGGEPKLDGSLPKPNP TVLSSPTSIPTRPQPKVSTNFTDLFAFSGPAPE RINGRLAMIGFVAIAVELSNGTDLFSQISGSGSQ WFLGTSILFSAASLPLQGSRPEAKTEGLMNAV AELWNGRIAMILGLVALAFTEFVYTGSAVY
89	sp_170320_dmqi 0.982212	0.988954	0.998831	UDP-glucosyl transferase 73B5 (UGT73B5); FUNCTIONS IN: quercetin 3-O-glucosyltransferase activity, UDP-glucosyltransferase activity, UDP- glucosyltransferase activity, transferase activity, transferring glycosyl groups; INVOLVED IN: response to other organism; LOCATED IN: endomembrane system; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage, LP 04 four leaves visible, 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: UDP- glucuronosyl/UDP-glucosyltransferase (InterPro:IPR002213); BEST Arabidopsis thaliana protein match is: UDP-glucosyltransferase 73B4 (TAIR:AT2G15490.3).	MELSNPTTPTLNATQPLRGYFPLITVPSHSNLV DIAKLFSSRGVHVTLTTHHTSLRFKQSIHDWGFK IDLHIVDFPFREVLPEGVENYSDATPEGASQLF QAFMMLQKPMEDAIRAAKPDFIVSDRYHWSVD LARELAIRLIF-HVRCYFALCAAEEVAKFAPHEKV ESDSDLFLPDLPTIHMTRLQLPEWIQTRNMFY VLNERMDEADRECYGVVNSCYELERAYADFVY SNLGRRAWTCIGYPVHCVKYGGKKGDDSKKHS CFEWDKMGEGEVYVSGTSCFSPAQISELAT ALEMSGHPFWVVRNNGEKLLPDGFEERITEQDK GVLIKDWAPQVKILEHPAVGGFLTHCGWNSVTE SLAAGVPMVTWPLGAEQFFNEKLSVGLKVGVE VGSEKWSRGIVPNTDMIEKDKIERAIKELMSKEP EAEEERQKVKELSKAPRNAVEEGGSSRNLLSDLI EKLQRLKANEISVSTNSE

10

20

30

40

【表 1 4 - 3 7】

A	B	C	D	E	F	
	分析に用いたペイトとの相関係数					
1	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)			
2	共発現遺伝子			注釈	aa seq	
90	sp_170920_oudh	0.9953733	0.9441829	0.9855877	<p>Uridine diphosphate glycosyltransferase 74E2 (UGT74E2); CONTAINS InterPro DOMAIN(s): UDP-glucuronosyl/UDP-glucosyltransferase (InterPro:IPR002213); BEST Arabidopsis thaliana protein match is: UDP-Glycosyltransferase superfamily protein (TAIR:AT1G05675.1); Has 8397 Blast hits to 8309 proteins in 513 species: Archae - 0; Bacteria - 605; Metazoa - 2499; Fungi - 31; Plants - 5061; Viruses - 123; Other Eukaryotes - 78 (source: NCBI BLink).</p>	<p>MLVYDAYFPWALDVAKDHILAAAYFVSSCAYVA SFYPMFLFEFGSDQHPVVAESVPGLLVELLPS REEMERYAPKACACSSDDKPDTKKPLHPVYQ MIVVSSITT_LHLADFLVFNSTDLHLEHQVVKWMTNL WRVKTVGLLPSPAYLDKRIESDVNYCVKPYKPNN EACMSWLNAAKQVASVYVYVFGSVAKLSVEQISEI AKALKQIPSSFLWIVREAEQEKLPNDFITETSEKG LVMISWCPQLDVLAAHEAVGCFHCGWNSVIEAT SFGVPMLGMPQFMDHFLDAHFLKLVWVGVIRVK ADEKNFVTCDEIKRGLKIIYGERGINKIKENA TKW KELAKEAVGEGGSSDNSTGEIKWLASS</p>
91	sp_170930_hjgc	0.9953733	0.9441829	0.9855877	<p>Uridine diphosphate glycosyltransferase 74E2 (UGT74E2); CONTAINS InterPro DOMAIN(s): UDP-glucuronosyl/UDP-glucosyltransferase (InterPro:IPR002213); BEST Arabidopsis thaliana protein match is: UDP-Glycosyltransferase superfamily protein (TAIR:AT1G05675.1); Has 8397 Blast hits to 8309 proteins in 513 species: Archae - 0; Bacteria - 605; Metazoa - 2499; Fungi - 31; Plants - 5061; Viruses - 123; Other Eukaryotes - 78 (source: NCBI BLink).</p> <p>cysteine-rich RLK (RECEPTOR-like protein kinase) 25 (CRK25); FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: endomembrane system; EXPRESSED IN: root; CONTAINS InterPro DOMAIN(s): Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Protein of unknown function DUF26 (InterPro:IPR002902), Serine-threonine/tyrosine-protein kinase (InterPro:IPR001245), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: cysteine-rich RLK (RECEPTOR-like protein kinase) 10 (TAIR:AT4G23180.1); Has 124291 Blast hits to 122724 proteins in 4390 species: Archae - 112; Bacteria - 13863; Metazoa - 45275; Fungi - 10764; Plants - 35367; Viruses - 449; Other Eukaryotes - 18461 (source: NCBI BLink).</p>	<p>MTGKGRTMEVIMMPPFHQGHLPMLQFAKRF A WKGAGSRITLATTSTAGNMNTSKNNNNNDY DELTVESIYDDTDDSQLKFMGRMGKFKSEASLQL GRLLTTSDINNKCMILVYDAYLPWALDYGKDHNI QAAAFVQACAYMASFYPMFLFEFGSDDDQHPV AAKAAEVPSSLVELPSREEMERYAPKCAQSPS SDDKPNYKSLHPVYRMVYVSSITLHLADFV_LIN SFDHLEHQLDVLAAHEAVGCFHCGWNSVIEATNF GVPMLGMPQFMDQFLDAHFMKLVWVGVIRAKA DEKNFVTCDEIKRGLKIIYGERGINKIKENA KW KDLAKEAVGEGGSSDNSTGEIKWLASS</p>
92	sp_171760_xdak	0.9953733	0.9441829	0.9855877	<p>cysteine-rich RLK (RECEPTOR-like protein kinase) 25 (CRK25); FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: endomembrane system; EXPRESSED IN: root; CONTAINS InterPro DOMAIN(s): Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Protein of unknown function DUF26 (InterPro:IPR002902), Serine-threonine/tyrosine-protein kinase (InterPro:IPR001245), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: cysteine-rich RLK (RECEPTOR-like protein kinase) 10 (TAIR:AT4G23180.1); Has 124291 Blast hits to 122724 proteins in 4390 species: Archae - 112; Bacteria - 13863; Metazoa - 45275; Fungi - 10764; Plants - 35367; Viruses - 449; Other Eukaryotes - 18461 (source: NCBI BLink).</p>	<p>MLVLLVWFVINASNAVPLTDYIWIIDCYNSTTP PNSSYETNLNHLHLLSTLAPQTSNSKHGFFTKLNT TQNDQHPVYGLASCRGDINPDCRDCVROAST QVIRACPEQTOAVFVMEQCMLRYSNRFSHLLQ ESNFRFTNHNVTNNVNDWIKLLNDLTIETSK AASDCKGNKFKATKRVWVFSQKLYTLAQCNP D LTVADCAECFSLARSQYDGMARQGLVLPSCN IRYQVKQFYGENSNYSITPLTSLPPPPHHHIST TYTGKKKQTTNVAIVAVPILGVLSISFSTFYFK KAKKRRSLDLKNVLDLTLTAESLMFELELEAATN NFAFIRKLGQGGFGSVYK</p>

10

20

30

40

【 1 0 3 7】

【表 1 4 - 3 8】

A	B	C	D	E	F
1	分析に用いたペイトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
93	sp_172300_mjoa	0.962317	0.943339	<p>RNA-dependent RNA polymerase family protein; CONTAINS InterPro DOMAINs: RNA-dependent RNA polymerase, eukaryotic-type (InterPro:IPR007855); BEST Arabidopsis thaliana protein match is: RNA-dependent RNA polymerase family protein (TAIR:AT2G19910.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLink).</p>	<p>MASVPELPPLEAVEEQLGICHEKSLPADAV  ARRYLSEVGEQESFRILNIIRNSGEVRKLSGFVHL  VKNRSRSPPLASPTISPOGNTTSGKRRKMES  DSPGGLSSVRRVCFINSPTKIVIRVDPPKSELDQ  STVEEVCLNSPTKIEEKHLKSETDRVLDLEKL  EFFRKTLLNYGCAKLEDAVYGLSADKILSYQSL  GMRDFENEIWNDFGRNICOCOKERINKPGYLKA  ATLLQKTVEDENVLMVKFMEQETKSLDTESSIMQ  SPAFNKVSEEGIFVGRKCYRFFVFKDGGKEEKKK  NPTSSPKVCYFVCTEGFASLLKOPPHLLKINMSV  QKARGLFMHKATNLSNYMARFSLLSKTTILEL  NYDVLRVVEVDYPCRDGNGDVIYDKDKAMIFT  DGTGFISEDLAVRCPHNCFNQYNNDENVOAML  RDPGVKIFCYGSLGSLAEAFRLPHNLGLTEKLSL  PGVPLLQCFRMFYKGAIKGTVLVNRKLPNTIQ  VRRSMLKIKHBEKLSGAPTYNSFEVATSNKPKD  AAYSKNLIALHYGGVPKDVFLSMLWEAITKPLSV  LYKTKAAVRGNADVYSITSESEFCSLPFCHWQM  KVYKLFEGNSRLCLSVGLSLRSTSVIPLFGLTT</p>
94	sp_174820_kxmh	0.9462312	0.9443283	<p>YELLOW STRIPE like 3 (YSL3); CONTAINS InterPro DOMAINs: Oligopeptide transporter OPT superfamily (InterPro:IPR004813); BEST Arabidopsis thaliana protein match is: YELLOW STRIPE like 2 (TAIR:AT5G24380.1).</p>	<p>MRSVLEIREGDFEEDKVCDEIKRVPWQK  QITIRGIASILIGVYVVMKLCITGLVPLNLSVA  ALLAFVLRWTSLQKLGIVTFPFRQENTIVQT  CAVACYSIAGGGFGSYLLGLNRKTYEQAGVDT  QGNAPISFKEPVGWMTGFLFVSSFVGLLALVPL  RKIMIDYKLYPSGTATANLINGFHTSRGENEPR  VEVHGFVKCLISFLWFFQWYSYGGNQYGFPPQ  FPKFGKAWENSFYDFDNMTYIGAGMISSHLVNL  SLLLGAVLSWGMWPLIGDMKGWFPSSLPSS  MKSINGYKVFISIALIGDGLYNFFKTLFFIRSIHD  RLKDKLSSCPDSSKQALDLDLVRNEVFMREGPI  WLACFGYTFEISIIVIPKVPPELKWYVVVAVM  LAPLSFCNAYGTGLTDINMAYNYGKVALFVLSA  VAGKDSGAVVGLVGGCLIKSIVSISSDLMHDFKT  GHLTLTSPRSMILLSQAIGTAVGCVVAPLTFFLFYN  AFDVGNDGEYKAPYALYRNMAILGVQGSALP  QHCLQCYAFFGFAVANLVRDTSPPKIKGWVPL  PMAMAVFLVGGYFAIDMCVGLVVFVFAWHFLNS  QKASLMVPYVAVSGLICDGLWILPSSILALVKRNP  SIR</p>

10

20

30

40

【表 1 4 - 3 9】

A	B	C	D	E	F	
1	sp_107620_kqnh (SOAP1)	分析に用いたベイトとの相関係数	sp_107670_ptqx (SOAP2-like)		aa seq	
2	共発現遺伝子			注釈		
95	sp_175360_idcc 0.952331	0.958437	0.982454	Glycosyl hydrolases family 31 protein; FUNCTIONS IN: hydrolase activity, hydrolyzing O-glycosyl compounds; INVOLVED IN: carbohydrate metabolic process; LOCATED IN: apoplast, vacuole, plant-type cell wall; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Glycoside hydrolase, family 31 (InterPro:IPR00322); Glycoside hydrolase, catalytic core (InterPro:IPR017853); BEST Arabidopsis thaliana protein match is: alpha-xylosidase 1 (TAIR:AT1G68560.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).		MKKKIPSALGILLVLLQYLAVAGISTSENDEPGVI GYGYKVKVSDSGTRRSLLTALLQLVKNSVYVG PDQLLSITASLESNDRLRVRITDAKHRRWEPDNI LHRHQPPPSPHLSLLYRTLSSPTTNRKILLS HPNSDLTFSLINTTFFGFTISRKSTHDLVLDATPD PTNPTFLIFKQYLLHLSLPGTRAHYGLGEHS KPTFQLAHNQTLTMMWAADIPSSNDYNLYGSHF FYMDVRSSPVAGSTHGVLLFNSNGMDVEYTGNR ITYKVIIGIDLYFFAGFSPAQVVEQFTRVIGRPAP MPYAFGGQCRGYHYHDVYELQSVVAGYAKAKI PLEVMWTDIDYMDAYKDFLDFVNFPLDKMKKF VNNLHKNGQKYYVILDGPGISTNKTYETVIRGMKH DYFLKRNKPYLGSVWPGVPYFPDFLKPFSALTF WTDEIKRFLNLLPVDGLWIDMNEISNFISSPPIPG STLNDPYPKINNSGYMLPIINKTIPTTAMHYGDPE YNVHNLFYGLARVTRAALIKLTKRPFVLSRSTF SGSGKTYAHWTGDNAAATWDLVYSPISMLDFFGL FGIPMVGADICGFLNGTTEELCRRWIQLAGAYPF SHDHSSLGTTYQELRYWESVAASARKVGLGRLYT LLPYFYTLMEQAQLNGIPIARPLFFSFDDDIKTYGI SSQFLLGKGVMSVPLKPGVSVYAYFPRGNWF DLFDYTRSVTASTGRVYVTLAPPDHNINHIQEGNI LAMQGMATTTQAARKTPFHLLVWMSDCGASFGI LFLDDGVEVTMGVNRGKWTFVKFIAASAKQTCH TSDVYVSGEFASQKVVIDKVTILGLRKRKTKINGY TVQTVGAVTRKGDKSKLKSSTPDRKGEFVAVIEISGL NLLGREFEKLVLH
96	sp_177730_zato 0.98274	0.936322	0.992142	Putative membrane lipoprotein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: male gametophyte, flower, pollen tube; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage, 4 anthesis; BEST Arabidopsis thaliana protein match is: Putative membrane lipoprotein (TAIR:AT5G17590.1); Has 36 Blast hits to 36 proteins in 9 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 36; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink).		MAVPEDSEAGGWYIGISLLFFITVSGAIFLVLVM IQPTSIHSWYVPIGGALVCFWFVFWFLTLRYIV SRACFRMSCFCGTSGVDHDFVIPSIVSAGVWINA SIGTNMGVWDGNVPIVGDLENGFNENGENEN GNRASGVRLQPGSPRRVHFGGVVLGEDDDGG GGAVGNOQESPKLSLSSSSSSSSSSSSSSSSSS KLCMAS

10

20

30

40

【表 1 4 - 4 0】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたペイトとの相関係数			注釈	aa seq
		sp_107620_kqnh (SOAP1)	sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	Zinc finger C-x8-C-x5-C-x3-H-type family protein; FUNCTIONS IN: zinc ion binding, nucleic acid binding; INVOLVED IN: biological process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; CONTAINS InterPro DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, CCHH-type (InterPro:IPR000571); Has 904 Blast hits to 850 proteins in 161 species: Archae - 9; Bacteria - 40; Metazoa - 431; Fungi - 54; Plants - 83; Viruses - 2; Other Eukaryotes - 285 (source: NCBI BLINK).	MMVETKPKFKTKLCLVYGRGRCARETCTTFAHGDS ELRRFSASLPDRRIDRDNDLRSKLDRIKRSPLRRC SPERDARGSHAYKRISPHSPKRRKLRITHEYLEG PNGFSEGSKNEEGSADSRDKMIMHVSGFSRKML EDELKIQIYVDVLFNOKSQMEIELEERYQEADS LSSKIODLEAQLAGEKDKSRVDSKIKLVRRAHV HHTRLQDEVKRSQLEQLVGLVQQLVDRDATPFSAIE KLVNINHEEKVLELKDSVPSVRKQPHSSMEASQ GLKSARGNGVRVLCSSRNPLVRIDCQQAIEIRNK RKDVAGPYTESNPRMRNASK-HFDPLDKVKKSL ESSLVAPPTSMAAHAVDDHIELMEAEENPEVLGTP TNFGSSLLFLLPPLPTMSQNNYFQDEGNDANVD VDGHEEDIEEYVDIV
97	sp_178160_fplg	0.987438	0.98747	0.990205	low-molecular-weight, cysteine-rich 69 (LCR69); FUNCTIONS IN: peptidase inhibitor activity; INVOLVED IN: defense response; LOCATED IN: plasma membrane, plant-type cell wall; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Gamma thionin (InterPro:IPR008176), Knottin (InterPro:IPR003614), Gamma Purothionin (InterPro:IPR008177); BEST Arabidopsis thaliana protein match is: low-molecular-weight cysteine-rich 68 (TAIR:AT2G02130.1); Has 400 Blast hits to 400 proteins in 82 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 400; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).	MERSRFRVSVLLMLVLVSTEIGTKVAEARICES ASYRFKIGCYFRSNCANVCNKNEGFPGGRCRGRFR RRCLCYKHCG
98	sp_181290_kwoe	0.98247	0.98857	0.978233	60S acidic ribosomal protein family; FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translational elongation, response to cold; LOCATED IN: cytosol, cytosolic ribosome, ribosome, nucleus; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ribosomal protein 60S (InterPro:IPR001813); BEST Arabidopsis thaliana protein match is: 60S acidic ribosomal protein family (TAIR:AT2G27110.4); Has 1851 Blast hits to 1850 proteins in 328 species: Archae - 0; Bacteria - 0; Metazoa - 772; Fungi - 418; Plants - 417; Viruses - 0; Other Eukaryotes - 244 (source: NCBI BLINK).	MKIVAACLSSLGGKACPTLDIIKKNLRSRGGPFSR HFFHNGSLIDEDI
99	sp_183370_xrwc	0.963244	0.950138	0.977528		

10

20

30

40

【表 1 4 - 4 1】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh sp_107660_kiqg (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
100	sp_185520_rkzu	0.926219	0.952985	0.975341	ZIP metal ion transporter family; FUNCTIONS IN: metal ion transmembrane transporter activity; INVOLVED IN: metal ion transport, transmembrane transport; LOCATED IN: endomembrane system, membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Zinc/iron permease (InterPro:IPR003689); BEST Arabidopsis thaliana protein match is: ZIP metal ion transporter family (TAIR:AT3G08650.2); Has 3662 Blast hits to 3630 proteins in 1201 species: Archae - 124; Bacteria - 2163; Metazoa - 678; Fungi - 116; Plants - 147; Viruses - 0; Other Eukaryotes - 434 (source: NCBI BLink).	MDSQVLVALALSLVGLLSTL GALFVVLSDAPNL KMLGLLQGFAGLMLISISFFDLAHNAINSIGFLK NIWFFSGVIFAIINFIPEPSLAPTSNVKSKRRAN GEDKGDLMKKHRRQVMYSGITIGISLHNFPF GMAYFLGSMKGLRVGNLALALAIHNIPEGVAVA LPVYFATQSKWQAFKIASLSGLAELPLGVIVAYLF PISLPEILDGLGAVGGVMAFLTLHEMLPLAFD YAGQKQAVKAVFFGMAFMSASLYFLEISLPADLS L
101	sp_187550_ghsy	0.966134	0.945535	0.943033	ATP binding; nucleic acid binding; helicases; FUNCTIONS IN: helicase activity, nucleic acid binding, ATP binding; INVOLVED IN: biological_process unknown; EXPRESSED IN: embryo; EXPRESSED DURING: D bilateral stage; CONTAINS InterPro DOMAINs: Pentatricopeptide repeat (InterPro:IPR002885), DNA/RNA helicase, C-terminal (InterPro:IPR001650); BEST Arabidopsis thaliana protein match is: Tetrairicopeptide repeat (TPR)-like superfamily protein (TAIR:AT1G12300.1); Has 84744 Blast hits to 32036 proteins in 2584 species: Archae - 123; Bacteria - 11324; Metazoa - 2691; Fungi - 2451; Plants - 63793; Viruses - 4; Other Eukaryotes - 4418 (source: NCBI BLink).	MANKLGWVWOKASAFINGGNMTHRLSTLSINYLL SASSYFHSHTNFHFFVENEYEVDSSELFESVRFQQ COMGFRKLDLDAVSLCNRMICMNPSPIMLFTLILS SMTKIKPRVFLNYSYTVIHLNSHLEWLGVSRNLYSF NILVNCYC-HLGHVNFSGSILGRIUKVGYCPDIVTFS TLINGLVQNGLEEQATELLDKIVKLGFPNIVTYG ALIKGLCRIGNYANALNLLVTMFSGFFLCKPDAV MYSTIIDLKDKLLTOAFDLFTEMKSGDQDPVYF TYCTLVRGMCJAGKWDKACGMNEMLESNIDPD VETYNMLVDMCCCKSRKIREAEAILGUMTKRQDAP STVYSSLLDGYCSCGRLLDDASEVMDLMVKYGC MPDVVAFNTLLKGCCKSKKLDKALDLYQNMHLR GLAPNVFNTLMDALCKDNRIHCASQLKEMEA AGLGPDAVYSSLLDQLCKSQIDEALEMVRQFM GIGKLPDIVIYSLIDGLCGVGRYQEARDFIFSLLG KGLKPNQYTYNALIKGLCKKGLMINDAIDLKTMF DNGGSPNDCNVDYTIIRGLLSANDNEKALEFVFIK SKGFTLNAHTNSLNFNASLSDHDHDSVKCCQGSFAE VSVEPEA

10

20

30

40

【表 1 4 - 4 2】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
102	sp_190300_zfwy	0.3857	0.077325	0.993502	<p>squalene synthase 1 (SQS1); FUNCTIONS IN: farnesyl-diphosphate farnesyltransferase activity; INVOLVED IN: sterol biosynthetic process; LOCATED IN: endoplasmic reticulum, plasma membrane, endoplasmic reticulum membrane; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Squalene/phytoene synthase, conserved site (InterPro:IPR019845), Terpenoid synthase (InterPro:IPR008949), Farnesyl-diphosphate farnesyltransferase (InterPro:IPR006449), Squalene/phytoene synthase (InterPro:IPR002060); BEST Arabidopsis thaliana protein match is: squalene synthase 2 (TAIR:AT4G34650.1); Has 1469 Blast hits to 1468 proteins in 521 species: Archae - 28; Bacteria - 581; Metazoa - 114; Fungi - 174; Plants - 415; Viruses - 0; Other Eukaryotes - 157 (source: NCBI BLink).</p>	<p>MGSLGAILKHPDEFYPLLLKLMAYKEAEKQIPPEL HWGFCYSMLHKVSRSFALVIQQLGTELRNAVCV FYLLRALDTEVDDTSTIATDVKLPILRDFHKHIYDR DWHFACGKTKHYKVLMDFFHYVSTAFLELDRGYQ VAIEITRRMGAGMAKFCICNEVETIGDYDEYCHY AGLVGLGLSKLFHNTGLEDLASDDLNSMGLFLQ KTNIRDYLEDINEIKGRMFWPREIWSKYVKNLE DLKYEENSEKAVQCLNDMVTNALFHVEDCLKYM SDLRDPAIFRFAIPQIMSIGT LAICYNLQVFRGV VKMRRGLTAVIDQTNTPDPYGYGAFYDFACMMIK PKVDKRFDPNATKTLRISIAIQKICMDSGTLNKRKL HVVKNKSAYTPIMMVVFFIVLAVMIFARLSANRQN NYLSAVMAFL</p>
103	sp_191200_fgi	0.978481	0.943583	0.960787	uncharacterized protein	<p>MILLSVYCNLLSKFGILAEVSLISQTHIRGVVCLVI LVPSKILVHITWLVNWRVYFKLFFVCVLAEVWYT NIMKWLF</p>

10

20

30

40

【表 1 4 - 4 3】

A	B	C	D	E	F
1	分析に用いたバイトとの相関係数	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	共発現遺伝子				MRNLKSLQISSTLLRSSVTHKPLSSSLTSLSSSE PLUSODPISESSLPNQSPISNDFNEGYSIPISQNH QNPFTNFHSHINGLLSSKDPSPSAMEYFKWAQKOP GFGEGVDSFCLLHNFMGLEGQQCAVKELVSTQ NFPDAKTVIRGLNDCSNRFGFDLDSRIYNYLLCG FIKTRIEDSISCVEMLENKIDLWVSYLNDLIYEL GIORKMIKEKELYDKVFCQGNVNGGTVRIMMK ACLKDGPFSEADRYFRRLARDKGVELDPAAFSCAI VAVSQKPDNSAACELLMEMKEKGWVPSGTYT QTIACVKORNLMEALRLKDEMYSCGKPMNLVVA TTLMKGYCVSGDLESALGLFAKLDEYELSPNIT FSVLIDGCCINGNMEKASELYDQMIKIEGIEFSVYN VNSMIKGFLLARLEEKAYEMFNKAESGLVDDSS LNSLLSCLCKDGNMDEAQKFWGMALCQGIENL QNYNLLILGYCRKGNVAMAHSMYVEMIGKGLNP NVFTFTLNGHFKKGEVDKALNLFDEMDDQDIDC TDHTYNTMIGGLCKHGKTSAPNMLENFIKEANF VPSGMTYNSIVGGLIKEGLDLSALSVNEMGENG VSPDICTYTFIDAFRKSNNFDLAKMONEMISKG LELDITAYGTLIDGCCRRDMESACKLNFELCKS GLSPNIYNNMISGYRNINNMVVALSFQKMWTE QGVLCDLETTYTLIDGLLKEDKLSALEFFDDMW EKGRPDKITYTVLVKGLGKGLKLENARNFKEEM EEKGVTNLSIYNALVAGYFREGNLPFAFKLYDE MLERGLTPDKMTDFLNMNANFKPGYQSSRTSSD MPIFCKEALA
104	sp_191910_jamq	0.942882	0.9349537	LATERAL ORGAN JUNCTION (LOJ); CONTAINS InterPro DOMAINs: Pentatricopeptide repeat (InterPro:IPR002888); BEST Arabidopsis thaliana protein match is: Pentatricopeptide repeat (PPR) superfamily protein (TAIR:AT3G54980.1); Has 63662 Blast hits to 14403 proteins in 306 species: Archae - 7; Bacteria - 57; Metazoa - 980; Fungi - 1296; Plants - 59204; Viruses - 0; Other Eukaryotes - 2118 (source: NCBI BLink).	MGLSVSILSAWHEILSQKFLTVSDGLLSSPVKR FPVSFRFTESFQKSTTDQSSPNFISETNHCOSG ETGAKDHHKHPENSTSKTRISETNHCQSTE MGAKKEHPNKEKIVLEKSPSFNTLVQEYDPTRT SSVVVVTGSNGLVHKPLPAITLPEPAVLFSRPVY SELDAAVKVKQVYSYRTRNLADCAVVVEEL WWKALDFASLKRSSVSFFNNDKQETAVAKWAR AKTRVAVKVGKLSQNEKAGKALRHWLEAIDPR HRYGHNLIHYDWFQSESSOPFFYWLIDGDGKE INLEKPRTKLQQCKIKYLGPIERIEYEVEDGK LTYKNGELVTVGECKWIFVLSRSLYIGQKK KGFQHSFLAGAAATTAAGRJIADKGVLEAIWPY SGHYLPTTEENKFEFISLEDNHDVLDLTVNKKCAYN EEGSKYFPEDESSEPEKTTDIDESIPAEVKTKE ESEFPFARRLSCKWVVTGNGPRIGCVRDYPMDL QSRFALEQVNLSPRVTPGSGFPIPSRPFSPKVRVS PRLAYMGLSPRTQMAAA
105	sp_194050_pwuc	0.941858	0.938504	embryo sac development arrest; 39 (EDA39); CONTAINS InterPro DOMAINs: IQ calmodulin-binding region (InterPro:IPR000048); BEST Arabidopsis thaliana protein match is: calmodulin-binding family protein (TAIR:AT2G26190.1); Has 379 Blast hits to 276 proteins in 60 species: Archae - 0; Bacteria - 22; Metazoa - 1; Fungi - 104; Plants - 246; Viruses - 0; Other Eukaryotes - 6 (source: NCBI BLink).	

10

20

30

40

【 1 0 4 3】

【表 1 4 - 4 4】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたペアトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
106	sp_199780_ehtzg	0.953374	0.972588	0.998753	alpha/beta-Hydrolases superfamily protein; FUNCTIONS IN: hydrolase activity; LOCATED IN: endomembrane system; EXPRESSED IN: shoot apex, hypocotyl, root, leaf; CONTAINS InterPro DOMAIN/s: Alpha/beta hydrolase fold-1 (InterPro:IPR000073); BEST Arabidopsis thaliana protein match is: alpha/beta-Hydrolases superfamily protein (TAIR:AT4G36610.1); Has 8797 Blast hits to 8792 proteins in 1517 species: Archae - 89; Bacteria - 6094; Metazoa - 449; Fungi - 67; Plants - 445; Viruses - 0; Other Eukaryotes - 1653 (source: NCBI BLink).	MVNIVSCYMSMLHKIMKLAGMHPHKEVEIEPQTVI HFVYNNNDNDNNAKAKKRRRRPALVLIHGF ADGILTWQFQVLAIRKKYSLYIPDLFFGGSH TSKAKITPRFHADCFKGLRKLGVEEKCVVVGFS YGGFVGFQMAEHPELVDLMVVTGSLALSESI SSKSLERIGFSSWPDYLMPNSVEGVKVLDDIGSY SFPWMPKVFYKHYLE
107	sp_199960_garc	0.983334	0.987591	0.997239	Chloroplast-targeted copper chaperone protein; FUNCTIONS IN: metal ion binding; INVOLVED IN: metal ion transport; LOCATED IN: cellular_component unknown; EXPRESSED IN: 12 plant_structures; EXPRESSED DURING: 7 growth stages; CONTAINS InterPro DOMAIN/s: Heavy metal transport/detoxification protein (InterPro:IPR006121), Copper chaperone, chloroplast-targeted, predicted (InterPro:IPR016578); BEST Arabidopsis thaliana protein match is: Chloroplast-targeted copper chaperone protein (TAIR:AT3G53530.1); Has 1086 Blast hits to 1014 proteins in 44 species: Archae - 0; Bacteria - 10; Metazoa - 2; Fungi - 4; Plants - 1070; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink).	MPIINIKIIGIDFSCASPASTAICSDQNFMVKN TTKSLNRLNSYLNHYKNTLNQHQSSSPRACPIS ELPFTPLSSFRKTRKSPSPSVKHSDLTPRR KSTADIADVRRRYDKCSNNSILSPHSHSSRY LLGESKLLDLSLSEHQNTVNYKSSSDALPLVPS TPLLVSREAHETKSFSGFLDVQKGNKSLTIVNPS DORLIKSSKYDYDGHQNRHKSIMGYDTNARNDLSL VLKSSSAAARSDQVVVLRVSLHCKGEGEGLR KHLKMEGYTFSIDLETKKVTVIGNVTFLSVLSI SRVYKNAQFWPFAASAPSSSSSSSSSTSDIDFN
108	sp_202730_niqq	0.943332	0.943677	0.974819	high mobility group B3 (HMG3); FUNCTIONS IN: chromatin binding, structural constituent of chromatin, DNA binding, sequence-specific DNA binding transcription factor activity; INVOLVED IN: chromatin assembly or disassembly; LOCATED IN: chromatin; EXPRESSED IN: 24 plant_structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: High mobility group, superfamily (InterPro:IPR009071), High mobility group, HMG1/HMG2 (InterPro:IPR000910); BEST Arabidopsis thaliana protein match is: high mobility group B2 (TAIR:AT1G20693.2); Has 5742 Blast hits to 4932 proteins in 507 species: Archae - 0; Bacteria - 0; Metazoa - 4274; Fungi - 383; Plants - 553; Viruses - 0; Other Eukaryotes - 532 (source: NCBI BLink).	MKGGKGAETTTTRRGGDAKLSVKKGAPAKKG KPAKDPNKPRPPSAFFVMEEFRTVYKQKHPN NKSVAVVGKAGGKWKSLTDAEKAPYQKAEKR KADYEKNMTTYNNKLAEGKDAEEEEESDKSRSE VNDDEDEDGSGEEDDD

10

20

30

40

【 1 0 4 4】

【表 1 4 - 4 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	<p>UDP-glucosyl transferase 74DT (UGT74DT);</p> <p>FUNCTIONS IN: transferase activity, transferring hexosyl groups, UDP-glycosyltransferase activity, transferase activity, transferring glycosyl groups;</p> <p>INVOLVED IN: metabolic process; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: UDP-glucuronosyl/UDP-glucosyltransferase (InterPro:IPR002213); BEST Arabidopsis thaliana protein match is: UDP-Glycosyltransferase superfamily protein (TAIR:AT1G05675.1); Has 8199 Blast hits to 8111 proteins in 470 species: Archae - 0; Bacteria - 484; Metazoa - 2414; Fungi - 25; Plants - 5092; Viruses - 117; Other Eukaryotes - 67 (source: NCBI BLINK).</p> <p>uncharacterized protein</p> <p>basic helix-loop-helix (bHLH) DNA-binding superfamily protein; FUNCTIONS IN: DNA binding, sequence-specific DNA binding transcription factor activity;</p> <p>INVOLVED IN: regulation of transcription; LOCATED IN: nucleus; CONTAINS InterPro DOMAIN/s: Helix-loop-helix DNA-binding domain (InterPro:IPR001092), Helix-loop-helix DNA-binding superfamily helix-loop-helix (bHLH) DNA-binding superfamily protein (TAIR:AT2G22750.2); Has 2910 Blast hits to 2902 proteins in 174 species: Archae - 4; Bacteria - 2; Metazoa - 77; Fungi - 51; Plants - 2768; Viruses - 0; Other Eukaryotes - 8 (source: NCBI BLINK).</p> <p>S-adenosylmethionine carrier 1 (SAMC1); CONTAINS InterPro DOMAIN/s: Mitochondrial carrier protein (InterPro:IPR02067), Mitochondrial substrate carrier (InterPro:IPR001993), Mitochondrial substrate/solute carrier (InterPro:IPR018108); BEST Arabidopsis thaliana protein match is: S-adenosylmethionine carrier 2 (TAIR:AT1G34055.1).</p> <p>Peptidase family C54 protein; CONTAINS InterPro DOMAIN/s: Peptidase C54 (InterPro:IPR005078); BEST Arabidopsis thaliana protein match is: Peptidase family C54 protein (TAIR:AT3G59950.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).</p>	aa seq
109	sp_213800_kcljw	0.983405 0.933635	0.941979 0.927053	0.977783 0.978378	MTNLWRVRTVGPLLPSSYLDKRIKSDVDYGINFY KPNNEACMNWLNKGAASVYVYVSGSVANLSVE QIAEIAKALKKIPSSFLWIVREAEQEKLPEDFITET SEKGLVMSWCTQLDYLAEAVGCFITHCGWIN LSRALAVLASASLGSRREREFLRHVNIKE
110	sp_214020_fsci				
111	sp_214780_pqju	0.959688	0.96324	0.983005	MDDKEPTPEITAPDIKVRVFNNDNVLITIVHEIPHGL VTKIMREISNLHLFIANMNSMPIGNISQVSIVAEVS PKPY
112	sp_000230_zzhd	#N/A	0.947647	0.930005	TAGAVGGMBDASLRVPTTELVKORMQTGQFVSAP DVRLLIVAKEGFKGLPIV/SASRFAY
113	sp_000570_tqfy	#N/A	0.9603447	0.9283638	AHVPDQVKNSEVANIADKDDIEAQFALSIDPSLEIG FYCRDI

10

20

30

40

【 1 0 4 5】

【表 1 4 - 4 6】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	alpha/beta-hydrolases superfamily protein; FUNCTIONS IN: hydrolase activity; INVOLVED IN: metabolic process; LOCATED IN: cellular component unknown; EXPRESSED IN: stem, hypocotyl, root, flower; EXPRESSED DURING: petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Lipase, GDHG, active site (InterPro:IPR002168), Alpha/beta hydrolase fold-3 (InterPro:IPR013094); BEST Arabidopsis thaliana protein match is: carboxylesterase 13 (TAIR:AT3G48700.1); Has 8596 Blast hits to 8578 proteins in 1414 species: Archae - 104; Bacteria - 4918; Metazoa - 486; Fungi - 808; Plants - 1365; Viruses - 3; Other Eukaryotes - 912 (source: NCBI BLINK).	aa seq
114	#N/A	0.558442	0.924077	uncharacterized protein	RNARSMLLAAEAHEADAEMILIPNQSPRKADDV TNJAVDTFHCFPIFSSEINEKRNWRSPLLQIQA IAIERFSLGARQWQLDLRPPKPPFEIKHHEVAI VKYNNLPSRRVNLKEYLDRVFFYDAVMNEEASN DKGLRLKLSQVNDIFIFILLDE
115	#N/A	0.346883	0.971385	uncharacterized protein	MSSTSQTDINHQRRLIDPAFPKFFFTQAPRLQ NFFKSHLTGLRKAGORKDVHELOODIRKEOKSS TDLEADLQKQLWFRANPHHWGSGPPPEINVTVPK GSLCNLNVFEDVGLPPDAVYVIVIDPDSKRVFRNI KEVISRKL_VDEGSRQVVEVEQAIAWNLWLSG TISVHVLVDQNRDHTMRFKQVNTGFMKKFEGC WRVEPVFVDEKMCSPKPKTSLSDYRSCTGGKG RIGSRVSLQQLQPAIVPPPIISWYLRGITSKTSE TLMNVLLEEVARIKGISQVAGSNGDEKSVSEITNE PLDIDEVRDIKRWALRRRRHALHYRKK
116	#N/A	0.924332	0.368824	Domain of unknown function (DUF220); CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF220 (InterPro:IPR003863); BEST Arabidopsis thaliana protein match is: Domain of unknown function (DUF220) (TAIR:AT1G23560.1); Has 170 Blast hits to 157 proteins in 17 species: Archae - 0; Bacteria - 4; Metazoa - 0; Fungi - 0; Plants - 157; Viruses - 0; Other Eukaryotes - 9 (source: NCBI BLINK).	
117	#N/A	0.943522	0.337843		

10

20

30

40

【表 1 4 - 4 7】

A	B	C	D	E	F
1		分析に用いたペイトとの相関係数			
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
118	sp_025850_ngzk #N/A	0.327054	0.342337	Pectin lyase-like superfamily protein; FUNCTIONS IN: polygalacturonase activity; INVOLVED IN: carbohydrate metabolic process; EXPRESSED IN: stem, root, stamen, seed; EXPRESSED DURING: 4 anthesis, E expanded cotyledon stage; CONTAINS InterPro DOMAINs: Pectin lyase fold/virulence factor (InterPro:IPR011050), Glycoside hydrolase, family 28 (InterPro:IPR000743), Pectin lyase fold (InterPro:IPR012334), Parallel beta-helix repeat (InterPro:IPR006626); BEST Arabidopsis thaliana protein match is: Pectin lyase-like superfamily protein (TAIR:AT4G01890.1); Has 5413 Blast hits to 5130 proteins in 651 species: Archae - 6; Bacteria - 1539; Metazoa - 261; Fungi - 1343; Plants - 1850; Viruses - 100; Other Eukaryotes - 314 (source: NCBI BLink).	MPLHVQNPNDQYSQISIPSPSYSPNIFDSFSDDDHHS VGYNPAPPSPDDDDHGVSSTDAHRRRRQTH HKFSGHKAAGHRHQAAPAPSHRQPAQPTGVF DVRSFGAGDGLVDDTLAFKSAWDAACGGDNA KVVAPKGFSLVQPLIFNGPKSGVVFQLEGTVV PDGPDAPKDSKQKRWLVFYRIHSLTQGGGLI DGRGDKWNLPCPKPHKNGTSTHGPCDPSVA MKFSGNTLTVRGLRLINSPQFNIRIDGQDVNV DLSIKSPANPNTDGIHVENTNDVTIYNSIANGD DCISIGAGSHNLDIRNVTGCGPHGISIGSLGIHQY KACVSNITVSNCIKHSDNGIRIKTWQGGSSVSG INFEGHMDTVRNPIIDQYCYCDASAKCSNQTSAVSI KDITYSNIKGYDTRSPVHACSDPAPCTNITIS GLELLPAPVTGMDLVTDHKLQVLSMEPFYCNAY GNNIGEVTNPPVYCLLDGMPYTSDDVEQC
119	sp_031470_htwz #N/A	0.324122	0.3303735	Eukaryotic aspartyl protease family protein; FUNCTIONS IN: aspartic-type endopeptidase activity; INVOLVED IN: response to salt stress; LOCATED IN: cell wall, plasma membrane, membrane, plant-type cell wall; EXPRESSED IN: 6 plant structures; CONTAINS InterPro DOMAINs: Peptidase aspartic (InterPro:IPR021109), Peptidase aspartic, catalytic (InterPro:IPR009007), Peptidase A1 (InterPro:IPR001461); BEST Arabidopsis thaliana protein match is: Eukaryotic aspartyl protease family protein (TAIR:AT1G03230.1); Has 1398 Blast hits to 1391 proteins in 46 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 1396; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLink).	MASYNLHHPPIFYLLIYTSISISISISISITTSALFL PKALVLTWKDPSTLYQVTKLIQRTQPQLSLDLGT PISWVICDNRYISSTHPAYCGSPQCTLANSKLC RTCPPAKPGKCNNTCTTSLNPFNTIIISPFGG TELADVIRPTDDGNSPFGQLVTVPNHLFVCAPKS LLOGLAKVSGVAGLGRSKISLPSQVSSYFFRPR KFTLCLPPLTESGVVFFGVDVYFNFPFGGAASAN LSYTLVANPFRPSDYFVIGVQGIKVGEDVPSVN KSKLSINSGNGGTMISLSHPYTIMETSYKAITIF NEQLTVVYKRVNVPVAFQFCYDSENTIGFSPGPM VQDIDLVLQSKVQWTAIPLNSMVRMNDKAYCLG ILDGGNPNPSTSIIGGFQIEDNLLQFDLCKSMIGFG SSLIYNRMRCNSNFNFTSNVI
120	sp_044330_oxia #N/A	0.323236	0.347231	Arabidopsis protein of unknown function (DUF241); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: root; CONTAINS InterPro DOMAINs: Protein of unknown function DUF241, plant (InterPro:IPR004320); BEST Arabidopsis thaliana protein match is: Arabidopsis protein of unknown function (DUF241) (TAIR:AT2G17070.1); Has 750 Blast hits to 734 proteins in 87 species: Archae - 2; Bacteria - 37; Metazoa - 79; Fungi - 28; Plants - 534; Viruses - 2; Other Eukaryotes - 68 (source: NCBI BLink).	MASYHTRSFSPNSHPVADQLDAQLSRLRSSQ AASTSPLTSKLSGLNDLYKCYEEFLQPNQKQTV SQSQGNVVEQVLDGSLRLLDIVSTSRDLALSK ERLQDICSVLRRCRCSGEIDITNEVSEYLRKTRKASK KIHKCLKINAVDKTNESIXXGEIDITNEVSEYLRK TRKASKKIHKCLKINAVDRTNESIAIESMLKDVQT LTVDFKSLSYIGGSQKSSWSFSKLVRRQGSBKE EAASISVFEAVDATLSQKNKANVDMSQLKLESEI QEIDEVLESFRHLIKTRATLLNLSN

10

20

30

40

【 1 0 4 7】

【表 1 4 - 4 8】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
121	sp_044550_qdkt #N/A	0.9228834	0.9228833	Kunitz family trypsin and protease inhibitor protein; FUNCTIONS IN: endopeptidase inhibitor activity; INVOLVED IN: biological_process unknown; LOCATED IN: apoplast, cell wall; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Proteinase inhibitor I3, Kunitz legume (InterPro:IPR002160), Kunitz inhibitor ST1-like (InterPro:IPR011065); BEST Arabidopsis thaliana protein match is: kunitz trypsin inhibitor 1 (TAIR:AT1G73260.1); Has 1019 Blast hits to 1019 proteins in 106 species: Archaea - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 1018; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLINK).	MAHLLLPAAATLLTLLFSPPTATAKDSVLDVD GNPVEAGSQYVLTSGWGGAGRGGLTAASLGG PCFLYVQYRLWVQKSPVTFYPSNPQKHITQ SSDVNIAFAPNPICRNOGVWQLGYNGISFYITTN GVIGKQEDVDFANWFKIQAGGLPGAYKIVYCF AVPSTGPEVPDPSRRRTGPEPDTYCEQLDAT VPVPPMGLSLLTPVIDNPLPFFGFFVKVNTTAN HSY
122	sp_062500_hayx #N/A	0.9328821	0.9328839	UDP-Glycosyltransferase superfamily protein; FUNCTIONS IN: transferase activity, transferring glycosyl groups; INVOLVED IN: lipid glycosylation, biosynthetic process, carbohydrate metabolic process; LOCATED IN: chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Glycosyl transferase, family 28 (InterPro:IPR004276), Glycosyl transferase, family 28, C-terminal (InterPro:IPR007235); Has 7866 Blast hits to 7866 proteins in 2456 species: Archaea - 0; Bacteria - 5567; Metazoa - 4; Fungi - 2; Plants - 34; Viruses - 0; Other Eukaryotes - 2259 (source: NCBI BLINK).	MVVVATISQLSLOFPPKHTQFKPIYPPKPRSQVF CCLSIDPPTLESSNVEERKQKQLRVVFAAGGAG GQIFPAVAIADEIRIINPDAQILFIGTDAGMESTAVP ASGYEFAIPATPLSRPILSVNNLSIPLKWKVKSMA ASWKIMKDFKPEIVGTGGYVSPFICLIAIKGLR MVIQEQSSIPGLANWFLSLVADLVFVAYVNSVDR FPTKIKKIVVSGNPVRLIKKIFVSKAVARTYFFPR LAKVSESEAVVLLVGGSGANGSMNIALFNIFYQ MLEORPNLIHIWETGVESFNEMESLVRNHPRLVL APFLRMDMGGYGAADLVVSRAGSMTCTELLVTG KPAILPSPNDAEGHQFQNASLMADLAGSRILTED ELDSITLRNAITDIDNDVLMASMSDRALQAAKPN AGEEIAERVIALYESSVYKA
123	sp_062670_phih #N/A	0.9433333	0.9233361	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G34560.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archaea - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MNPLNINHIKVEKAKIKRYQKLRKITRFRVFEII LILAVISRLSLOPLALFKLSNECFYSLSIFVSPKF VFIGNAIVTLFVNSGRLSNNECSPKTKTTPPYQR KNKNMMKDDCKSKCKQGITTEENAMMAMIFPPS EQQQVIVYQFQSKSEKMIKREKVEKCCGLRR SNTKISLKGCGCSNDVKECHYPEDHMSNEEFQ KTENFIARQQKFLREEEFSALKE MASYEASQQVTKFMFEGNTSRRRMSAAMVATM
124	sp_063080_tirx #N/A	0.9383382	0.9333373	uncharacterized protein	EMRLELISFGELDGRNTNGGGG

10

20

30

40

【表 1 4 - 4 9】

A	B	C	D	E	F
1					
2	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
125	#N/A	0.3485618	0.5243354	<p>Pentatricopeptide repeat (PPR) superfamily protein; FUNCTIONS IN: binding; INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion; CONTAINS InterPro DOMAINs: Pentatricopeptide repeat (InterPro:IPR002885), Tetatricopeptide-like helical (InterPro:IPR011990); BEST Arabidopsis thaliana protein match is: Tetatricopeptide repeat (TPP)-like superfamily protein (TAIR:AT4G39530.1); Has 56701 Blast hits to 14050 proteins in 283 species: Archae - 2; Bacteria - 22; Metazoa - 148; Fungi - 141; Plants - 55528; Viruses - 0; Other Eukaryotes - 860 (source: NCBI BLINK).</p>	<p>MKRELFSTPNIHLSIIKHLIAQNLPQALKNLSLFL SSPFLTDQTYALFLKSGHVLDTFLATSIITHFSKS GDFSRNGLLDAHGVDVAFNAMLGYAQFHQ SSPSFALYNKLTGYLSPDKYTLSSLIKSCDNLK IRVSHV/CVKYGFNLNGFVSVGLIHKYGGGDVE SAEKCFEELDLSDVYVWTAMINGVYVWNGFEKGG RQIFMAMRGLGLEINEFSLTAVV GALREVREGE QIHGFSFKGLIFDGS LHSNVVMNMYKCGHRL DALKVDEIPNPVYVSWTSRIGASRDGGEALEAL KFMLLNNVAVNELTVNVLTIIEPRLLIAGRQVH SLCHKAGYLTLSVNNALISMYGKCREMDDARLV FDELDYDPIISWNSLIGGYSDNGLLSAVLCIFHQ MRDLGVQHDLYTLASV LGAVTGPQFVLDVMQIH SHIIKSGLVSDSMISCLISSYKGGKSCLDHAKKVL SSTQEVNIGHLOVMAATAV TAGYPADAVECFKS AMNLWSEMGTTLSLIFKACSSLTNLGGQKAHC LALKAGLSDRFRVETAVDFYCKCGSIADARNIVF DYSSHNLAAWNAMIMFGAQHCYNEVLSLFR MTELGIKPDEITYLGILOSCCHAGLVNEANIHLNS MSELHGLLPCLEHYATMIDLGRVGLLKEARLVIE KLPTVDAQVWQSLLSACHNSGNLEIGEVAAKEL LKLQPENDSAYILLSNLYARAHWDVAVRRLRRE MKDKFIYKEPGSSWIEAGSVYHFLAEDTSHPEN EKIYFYLQSIACQMLP_LSEVTEDSFVCMDDGW</p>
126	#N/A	0.922632	0.933353	<p>uncharacterized protein</p> <p>HSP20-like chaperones superfamily protein; CONTAINS InterPro DOMAINs: Heat shock protein Hsp20 (InterPro:IPR002068), HSP20-like chaperone (InterPro:IPR008978); BEST Arabidopsis thaliana protein match is: HSP20-like chaperones superfamily protein (TAIR:AT2G27140.1); Has 638 Blast hits to 638 proteins in 88 species: Archae - 0; Bacteria - 6; Metazoa - 0; Fungi - 34; Plants - 583; Viruses - 0; Other Eukaryotes - 15 (source: NCBI BLINK).</p>	<p>PAAQHFNLYKAVGSSMDLYYVYRWNLQLTQ NSIDMAVLRKRAEDVESLLEDSKKLDAASDLE AANKRSLDIEPKLQAEAEADGLDQQLASINASL PGVKSAAKAVDKLQSDWFDNDILTLRHNGGW CAAHRRVVCHVKLLDEDGWQVQFEDGYEELKYYK</p> <p>MAIKTKSSTPGSPSYVDFHPACDFLKAEGIETLIH LPDFKKEQLRVQVWKEGVVIVSGERATNEDGSK RSRFVKETIPEGCNDINDIRAKFINGRLNVTMPKKI LITPPPPPPPPQLIQEQFOPRSPSSSSPPPPKPEA VTDQKPKPEAASVSDDQTPKGTIGGTETSCSY HGQPKSINRMFELGSLVQISRLQGRKNVALGFGIA VVTMVAIGAFVASKLGSDDTTSSPSSLYEDSGLM</p>
127	#N/A	0.9223957	0.9333831		

10

20

30

40

【表 1 4 - 5 0】

A	B	C	D	E	F
1					
2	分析に用いたペアトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
128	sp_116810_wrcx #N/A	0.990343	0.993261	annexin 4 (ANNAT4); FUNCTIONS IN: calcium-dependent; phospholipid binding, calcium ion binding; INVOLVED IN: in 6 processes; LOCATED IN: cell surface; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAINs: Annexin like protein (InterPro:IPR015472), Annexin repeat (InterPro:IPR018502), Annexin (InterPro:IPR001464); BEST Arabidopsis thaliana protein match is: annexin 8 (TAIR:AT5G12380.1); Has 2371 Blast hits to 2094 proteins in 191 species: Archae - 0; Bacteria - 0; Metazoa - 1728; Fungi - 90; Plants - 415; Viruses - 0; Other Eukaryotes - 138 (source: NCBI BLINK).	MADVSHLELOAVTKAFSGLGVDEKLSISLGIAGKWOH DHKNSFRKASPSHFADKERHFEKWNQOVELI KHEALRFKASLVWTHMPWERDARLLKEALIKG PEGYNLVEIACSRSSSEELLGARKAYHSLFDSSE EDVASHVDGHRKLLALLVSAVRYEGSRVHDGN AKSEAEVLHRAINNISDDEELVRILCTRSLKHS AVFKHYKQLFGKDLSEEFHGNLSLKAALCCLVIPE VYFSEVLDAAMPREGATDFTEQGLTRVIVSOADG NMEEIKAAAYHNLFGAHLSHKIEETCLGNFKDFLLT LVARGE
129	sp_120600_gjhz #N/A	0.921325	0.963121	hydroperoxide lyase 1 (HPL1); FUNCTIONS IN: electron carrier activity, monooxygenase activity, iron ion binding, heme binding; INVOLVED IN: fatty acid metabolic process, response to wounding; LOCATED IN: chloroplast envelope; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Cytochrome P450 (InterPro:IPR001128), Cytochrome P450, E-class, group IV (InterPro:IPR002403); BEST Arabidopsis thaliana protein match is: allene oxide synthase (TAIR:AT5G42650.1); Has 6712 Blast hits to 6702 proteins in 615 species: Archae - 25; Bacteria - 392; Metazoa - 3265; Fungi - 548; Plants - 2246; Viruses - 0; Other Eukaryotes - 236 (source: NCBI BLINK).	MMMTISATMSSTPNPIPTLLTPPSPPPANLPLR PVPGSYGLPVLGPIGDRNLNYFWFGQDSFFKKK MEKYKSSVFRITNVPCCFFFFAGVNPVIALLDVP SFSYLFDMIDIIDKKNVLVGDMPSTKFTGDMRVG VYLDITTEEKHTKKNFSTDLIKKSGWVSSLN TEIMWETVEESLNKDGSTKIPTLQKGFNLSQV MLRANPSNYSELCENGHIMDKWLAIQLLPIIN MFQPLEEIFHSWSYPYLYKGDYKLVNFVAKE AREVVEVTGVEFGLSECEAIHNLFLMLGFNAFGG FSRFLPALVNNLRENKGGIIDLKKEVRGKCNLSL TSLTFKSVQDMPIMTSFVYETLRLOQPPVPLQFGR ARKDFELTRHESRFEIKKGLCGYQTLVMDPK IFDDPDTFVSEFRMGETGGELLNYLFWVSNQPQT GKSDASNKQCAASDYVPLTACLFLAHLFLRYDSI TLDSSSGAITKVKQAT

10

20

30

40

【表 1 4 - 5 1】

A	B	C	D	E	F
1	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
130 sp_131330_pwyu	#N/A	3_357237	6_5228531	<p>disproportionating enzyme 2 (DPE2); FUNCTIONS IN: 4-alpha-glucanotransferase activity, heteroglycan binding; INVOLVED IN: polysaccharide metabolic process, circadian rhythm, maltose catabolic process, maltose metabolic process, starch catabolic process; LOCATED IN: cytosol; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Immunoglobulin-like fold (InterPro:IPR013783), Carbohydrate-binding-like fold (InterPro:IPR013784), Glycoside hydrolase, family 77 (InterPro:IPR003385), Glycoside hydrolase, catalytic core (InterPro:IPR017853), Glycoside hydrolase, carbohydrate-binding (InterPro:IPR002044), Glycoside hydrolase, subgroup, catalytic core (InterPro:IPR013781); BEST Arabidopsis thaliana protein match is: disproportionating enzyme (TAIR:AT5G64860.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).</p>	<p>MVNLGLFSSSENGVKSVSVMFRIPYFTQWGQSLL VCGSPSLGSWNRRRGLLSPTSQSGSELIWGGT IEVSSGCECEYSYVVDDDDKNILRWEMGKRRKLL LPSSIQDGEIVAFHDLWQAGSDTLFSSSAFKDVIF RNSSSLTVRPLGLVQNKLGLEDVSVNFKICCP NLDEGTSYVYIGSPQKLGQWKTEGGLKLSYAGD SVWQADCVIGKEDFPKYRYGKYDKAGKFAIEIG TNRKLLESLSQARYIFLSDGMMREMPWFRGSG MAIPMFSVRSEEDLGVGEFLDLKLVVDWAVASG FHLVQLLPINDTSVHGMMWDSYPYSSLSVFALH PLYLRVQALSENPEEIKLEIVKAKERLDVKDYDE ATMATKLSIAKKIFSLKDLTKSSSFQKFFLENEE WLKPYAFCFLRDFEETSQHSQWGRFSDSKEK LEKLVSDSPHYMIVCFHYIYQYHLHLQLSEAAEH ARKKGWLVKGDLPVGNDRNSVDTWVYPNLFRMN TSTGAPPDYFDKNGQNWGFPTYNWEEEMSKDNY GWIWRGLTQMAKYFTAFRIDHILGFFRIWELPE HAYTGLVGFRRSPILSQEELKEKIGWDFDRLSR PYIRQEFLEEKFGSSWTVAALFNECQKGHYEF KEDCNTEKKAISKDKDSTGKSLLLGSEAEIRINLF DLLRNIVLRDPEDTSKYPFRNLEDITTSFTDLDD HSKNVLRRLYYDYFFHROENLWRONAMKTLPLV LNSDDMLACGEDLGMPSQVHPVMQELGLGLRI QRMPNEPDLFEGIPSQYGYMTVCAPSCHDCSTM RAWWEEDEERRRRFCKTVMGSKLPPSKCPEI AYFIIRQHVESPSMWAIFPLQDLALKNLEYTTTTPA TEETINDTNPKHYWFRFHSHTVLEMLKLDQELKT AIKELVRSRSGRSYPHDDDETVAVSNKGNPEPPAV ASKVIQPNGASQEEEEEAVAVAVL</p>

10

20

30

40

【表 1 4 - 5 2】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq  MYSNSSHLKLGCKPAERKALLDFKSSVTFDLKLVSSWEGEDCCSWSRVKCNPTTGHIVELRLNGNYSYLFEARKIMESSLVYLLLEINHLLEYLDVSANNFNSAHDYFVNDFNLSKIPKFMGSMKQLRDCAFTEGIEVNIWLSVFNLCNLSLDLSSNKIQAKLLEGRENISRTICTHFDIEMLDLSWNSISGRIPYWLGEEFRSLKSLSGNNLSGIEPTSLGMLSSMKHLDISNPLEAAVLSESHFANLSSLTDLFMDQTLTLKLSYYLPPFQLQNFYASYCKINGPFFQWLKTKKNISNLYLSDAGIHGALPKWFHNMHSLSYVSLSNNOITGCPFPVNFEGISLSSNLSLSCQYLTMNSRNTVYSSOGVDLSDNSISGQLEHLYHDMPNLSTLVLSRLINGSIPESMCHFTSLQILDHKNRSLGTPHCLGESSELEFVHLAFNNLSGDIPCFRLLTHVENEWSSYLSFVSLNDNMLSGQIPSCLANLSTLEYLDVGEKLSLNGEIPMFITGLNPLVLRILRNKNGLEGSIQEYDYSIFYADMNQVNEVIKIERAYTTLLKYMVNDLSCNLLIGSPEGITNLTYLRNLSYNHLSGQIPMSIGCLKSLESJDLRSRNLQGTPIPIISMGAIFLESLLSYNNLSGQIPTGNQLQTLNDPLSYAGNPYLCGDPLAKKCKSDNDSPNGSSGQDNNEDVSEDKHETMWFYLVVMSGFATGFWCVVGTFLFKKSWRHAFRRRVEFVQDWLYVAVVVRVPKLMRRFI
131	sp_133890_wxdc #N/A	0.331238S	0.213303S	disease resistance family protein / LRR family protein; INVOLVED IN: signal transduction, defense response to fungus, defense response; LOCATED IN: cell wall; EXPRESSED IN: 10 plant structures; EXPRESSED DURING: 7 growth stages; CONTAINS InterPro DOMAIN/s: Leucine-rich repeat-containing N-terminal domain, type 2 (InterPro:IPR013210), Leucine-rich repeat (InterPro:IPR001611); BEST Arabidopsis thaliana protein match is: Leucine-rich repeat transmembrane protein kinase (TAIR:AT4G20140.1); Has 135121 Blast hits to 32840 proteins in 1181 species: Archaea - 56; Bacteria - 8201; Metazoa - 34800; Fungi - 1569; Plants - 79859; Viruses - 2; Other Eukaryotes - 10634 (source: NCBI BLINK).	
132	sp_149180_nwmy #N/A	0.331481S	0.2983272	acyl-CoA sterol acyltransferase 1 (ASAT1); FUNCTIONS IN: acyltransferase activity; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 10 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: Wax synthase (InterPro:IPR017088); BEST Arabidopsis thaliana protein match is: MBOAT (membrane bound O-acyl transferase) family protein (TAIR:AT5G55350.1); Has 801 Blast hits to 791 proteins in 235 species: Archaea - 0; Bacteria - 473; Metazoa - 0; Fungi - 36; Plants - 238; Viruses - 0; Other Eukaryotes - 54 (source: NCBI BLINK).	MVTRILRPTVYVYPTLEYSTKVVGRTWATLPAMVS TFVFSAIMHELIFYYLGRNWPTEVTFWFFLLHGL CLCIVEIVAKKLVGGKWRIPRWRISGPGPATVLFVWGT GFWLFPLLLKAGLDTRPFQFEAAVAKFVRSLLKA ALTF

10

20

30

40

【表 1 4 - 5 3】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	分析に用いたペイトとの相関係数 sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 glutamate receptor 2.3 (TAIR:3); FUNCTIONS IN: intracellular ligand-gated ion channel activity; INVOLVED IN: cellular calcium ion homeostasis, response to light stimulus; LOCATED IN: endomembrane system, membrane; CONTAINS InterPro DOMAINs: Extracellular solute-binding protein, family 3 (InterPro:IPR001638), Ionotropic glutamate receptor (InterPro:IPR001320), Extracellular ligand-binding receptor (InterPro:IPR001828), Glutamate receptor-related (InterPro:IPR015683), Ionotropic glutamate-like receptor, plant (InterPro:IPR017103); BEST Arabidopsis thaliana protein match is: glutamate receptor 2.2 (TAIR:AT2G24720.1); Has 5318 Blast hits to 5220 proteins in 711 species: Archae - 48; Bacteria - 1305; Metazoa - 3164; Fungi - 0; Plants - 641; Viruses - 0; Other Eukaryotes - 160 (source: NCBI BLINK).	aa seq
133	sp_155800_psd0	#N/A	0.924890	0.936107	MMKKKCCIFVMMNMFVFIPIVLYLGGQNTTETT PRELRIAPVREFKDWVSEHDPNTGEQISGFAL EYFKAVVNGKLLSPIPKYKPIFVSDGVTVMNGTFF DMLRAVFNQ
134	sp_156240_sfrz	#N/A	0.921897	0.973544	MSSIGSNESQIHPPTTEEEASCNAMCLATWVSPS MVFKTVLELNVDIYFTGPEAQLSAEEIVAKLPAT NPHAAVLDRLRLLASFTILTCTQQVGGDGGPK RLYGLGPVQCFLAKNDDEVSAAAFANLMDKVI MDTWKHLKEAVLVGGSAFRNSNGTSFFDYLGTD SRFYKVFYEGMDHSTIVMKLLDTYQGFKGISTI IDVGGATGLTKMILDKYPSIKAINFDLPHIKDAPS YPGVEHVGGMDYSPKGDALFIMKWTFSQGTD EKCVKLLKNCHEALPEDGKVICEYVVPFLPKTSY QJHTAFIWDSIMLFTGGKSRITKKEFEALGKEAG FEDFQLVCFACDTWVMEYLLKKKSDN

10

20

30

40

【表 1 4 - 5 4】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたペアとの相関係数 sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	alpha/beta-Hydrolases superfamily protein; FUNCTIONS IN: hydrolase activity; INVOLVED IN: metabolic process; LOCATED IN: cellular component unknown; EXPRESSED IN: stem, hypocotyl, root, flower; EXPRESSED DURING: petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Lipase, GDXG, active site (InterPro:IPR002168), Alpha/beta hydrolase fold-3 (InterPro:IPR013094); BEST Arabidopsis thaliana protein match is: carboxylesterase 13 (TAIR:AT3G48700.1); Has 8596 Blast hits to 8578 proteins in 1414 species: Archae - 104; Bacteria - 4918; Metazoa - 486; Fungi - 808; Plants - 1365; Viruses - 3; Other Eukaryotes - 912 (source: NCBI BLINK).	aa seq MNSRAKEVKVELPGFIRVYEDDSVERFYGSPQV PPTLEDPETGVSSKDVISADKNISARIYLPKIITNN HPILVYFHGGDFCVESAFSLHHQRYMILLASQAG VLVISVEYRLAPEHP_LPIAYEDCWITALNWTSLD DPWLTKHGNLDLFIIGDSSGANIVHNLAMKAG TQSLDHGVKILGAFSSHYSYFLDEPSSRLDQVMA YKIWDFVYPNALGGSNSPMPINPMVGDGAPSLAGL GCSKVLVVAELEGGQRNSEIRYNAVKDSGFQG SVLELFTKGEDHCFHFIFNPDTHNSKLLFQRLTSFL RN MQAFHCKRVKSLIHTVDNKNKTEKTRLLVHGKIQ VVPRAAATAGSRGRSNMQLYPYCCCHREAVFQ EANHPGEEGLQLHLMGIVTVIDRCQGNSSTLN YDDGHHQKKTQIFSWITLIENCHHHHPHPCCSA EQENSIPKINRGLQQRVIAIVTDHPGHAFRQKADL TKIFEIVL
135	sp_178980_gxiz #N/A	0.95192	0.928271	uncharacterized protein	
136	sp_180410_nknp #N/A	0.924565	0.926107	Low temperature and salt responsive protein family; INVOLVED IN: response to cold, hyperosmotic salinity response; LOCATED IN: endomembrane system, integral to membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Uncharacterised protein family UPF0057 (InterPro:IPR000612); BEST Arabidopsis thaliana protein match is: Low temperature and salt responsive protein family (TAIR:AT2G24040.1); Has 1255 Blast hits to 1255 proteins in 413 species: Archae - 0; Bacteria - 524; Metazoa - 40; Fungi - 292; Plants - 377; Viruses - 0; Other Eukaryotes - 22 (source: NCBI BLINK).	
137	sp_191450_ttk #N/A	0.961747	0.961233		

10

20

30

40

【表 1 4 - 5 5】

I	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたバइटとの相関係数 sp_107620_kqnh sp_107660_kiqg (SOAIP1) (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	<p>MFYSYLFQFITALCTWIHPHLLTSASSFOQTSDDPPIS KPGCQDICGNVTIPYFPGVGSNCYHTLPYAVTCN TSLSHPRFLFLNKYNLEIEIQINIIHKFHLQDQNTTVN TPLQRICNTRDANISSIDLSGTPYSFSSDINSVYVFG GGSSALVLDNRNNILGGCTSVGNNSHMIDDCI GFGCCRFDLPSSEFYRYNIFNFGGNCNTNAMLVN MSSPIISKAVDPSLFEVAVVPTVLEWTPGNITNPG DDPNIDCCGSADGYDCYCYNGTSGNPLYPYGC QVPSECKEKLGCIGPENHYVCVRNNSSEDKAYN FLVLLVLLGGIVGLMLCTGNISCIYLCIKRRKHKM MKKEFFQONGGFLKQQLNSQNSGNINDTKIFTA QELREATNYSKQNILGGGGFVYKGLPNDQV VAIKKSMVSDQSQVEQFNEVILSQTNHRNVVK LLGCCLEIQVPLLYEFVSNGLFEHHSDDARWL SWENCLRISTEAANAIAIYLSAASMPIIHRDIKSSN ILLDDNFTAKIADFGVSRVLPIDQTOVTLVQGTG GYLDPVFLTSRLSEKSDVYVSGVWLVLLTKRK PISWETGMEDRNAMIFISSVKNQNFDFILEPQLV KEASEEQIISIGNLVEKCLSVKGEDRFPMSKDVAR VLEEVANQSNNEIHEMSSEQRFPVQPAIEQDSSR QRFK</p> <p>MRPTTIPTATAQLGVPSVYGYAPAFGSARGFGT RGFWFPPTLVSPLVLSQDPITPGSSAVTSYVL GFWNPSSDCSVPTKSMTTLQTEVQSPACKPGV SPLRSRKRTRDMEKELVAFKECCQGNKDEGP ELIEFDGLCLSKYQMYRMVAMDEYAGIKYKVAS KIYTRWKAELGAGRGHRSVMLPEPFGF MSPLSPFQPHHRTAPLFCPLFPSPVTDVAVV INLWGGFTSSLPISRLLSLVPAM</p>
138	sp_192500_nzmu	#N/A	0.3375216	0.9231387	THO5; CONTAINS InterPro DOMAINs: THO complex, subunit 5 (InterPro:IPR019163); BEST Arabidopsis thaliana protein match is: THO complex, subunit 5 (TAIR:AT1G45233.2); Has 35333 Blast hits to 34131 proteins in 2444 species: Archaea - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	
139	sp_195330_fxln	#N/A	0.346982	0.933845	THO5; CONTAINS InterPro DOMAINs: THO complex, subunit 5 (InterPro:IPR019163); BEST Arabidopsis thaliana protein match is: THO complex, subunit 5 (TAIR:AT1G45233.2); Has 35333 Blast hits to 34131 proteins in 2444 species: Archaea - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	
140	sp_205140_yqdr	#N/A	0.290339	0.923881	uncharacterized protein	

10

20

30

40

【表 1 4 - 5 6】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
141	#N/A	0.344185	0.92315	S-locus lectin protein kinase family protein; FUNCTIONS IN: protein serine/threonine kinase activity, sugar binding, protein kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation, recognition of pollen; LOCATED IN: endomembrane system; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Curculin-like (mannose-binding) lectin (InterPro:IPR001480), Protein kinase, ATP binding site (InterPro:IPR017441), Apple-like (InterPro:IPR003609), PAN-1 domain (InterPro:IPR003014), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271), Protein kinase, catalytic domain (InterPro:IPR000719), S-locus glycoprotein (InterPro:IPR000858), EGF-like (InterPro:IPR006210); BEST Arabidopsis italiana protein match is: receptor-like protein kinase 4 (TAIR:AT4G0340.1); Has 116366 Blast hits to 114909 proteins in 4392 species: Archae - 99; Bacteria - 12770; Metazoa - 42761; Fungi - 9470; Plants - 34242; Viruses - 402; Other Eukaryotes - 16622 (source: NCBI BLINK).	MLDDMGTFASDQLKFSALDVGFGIKRRLTMDF DGNLRIYSLNKLGLWNITWTAVAKPCDIHGLCG RNGICEYTPKPKCSCPPSYEPTDISDWSKGRCP RFHRSQSDSEFIEMSHVDYGYDLNSTITTPASYYDD CREIC:LGDCRCOAFNYRLTGEQMCFTKSALFNG VKSVSFFASISINQRRKFWYLYSFAIVIGTVEVILLI AGWWFLFGKHSLSGSLDGYRAISNQFRSFSYN ELKTATRNFKVELGRGGFVAVYKGLADEREIAV KKLENVQGEVEEFWAEVSTIGKINHMNLARMWG FCSGRKHLLVVELVENGLDRHLFSTKTVLWG KERGGHSSSELTRIRGTGYMAPEWAINLPITAK VDVYSGVVLELIKIRLSSWVVDQNSSEEEIPE LVRFRLLAKRKIQNGEDSWVEDLVDPRLEGKFS RNQAAMMIEVGL-SCLDDDDRNKRPTMESVAQALA ECEDEL
142	#N/A	0.3551735	0.9237028	uncharacterized protein	

10

20

30

40

【表 1 4 - 5 7】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
143	sp_211990_qexc	#N/A	0.9344:1.4 0.9353:2.5	ANTHOCYANIN11 (ATAN11); CONTAINS InterPro DOMAIN/s: WD40 repeat-like-containing domain (InterPro:IPRO11046), WD40 repeat 2 (InterPro:IPRO19782), WD40 repeat, conserved site (InterPro:IPRO19775), WD40-repeat-containing domain (InterPro:IPRO17986), WD40 repeat-containing domain (InterPro:IPRO15943), WD40 repeat, subgroup (InterPro:IPRO19781); BEST Arabidopsis thaliana protein match is: Transducin/WD40 repeat-like superfamily protein (TAIR:AT3G26640.1); Has 4731 Blast hits to 4446 proteins in 351 species: Archaea - 6; Bacteria - 291; Metazoa - 1848; Fungi - 1200; Plants - 767; Viruses - 0; Other Eukaryotes - 619 (source: NCBI BLLink).	MGASSEPMNPEPFSDEQHKQRFSEIYTYEAPWH IYSMNWSVRRDKNYPLAIAISLLEHAQNRVEIVQL HDSTGEIKSDPFLSDHPYPTTKAIFIPDKDCSKP DLLATSSDYLRWQINSDSNSSKIELKSLINGNK NSEFCGLTDFDWNENAEPRRIGTSSIDTCTIWDI ERETVDTOIAHDKEVYDIAWGGAGVFASVSAD GSVRYFDLRDKEHSTIYESSSEPDTPPLVRLGWNK QDPRYMATIMDSPKVWVLDIRFPTLPVVELQRHQ ASANAIAWAPHSSCHICTAGDDMQALIWDLSSM GQPVYEGGLDPLAYTAGAEIEQLQWSSSQPDWV AIAFSNKLQILRV
144	sp_001340_gcdp	#N/A	0.9336:3.5 0.9331:2.7	uncharacterized protein	MTDSLPRFVILGRPTGGVYLYVYVEKGAQHRRL HLLSPEVISDRNFTKLEVVRAPPTIGLDMVHIRSC YNNKFLVLGSEGNVFAEAEDEPEVDRSIWSC TL FRATFFNRLSLLQSSNAQHSVQVSNSS-LGIV HSPFAYWFDVTNMTLAILPEYIAFKGDNNRYLN SRFRTKNRLKFESDNLDDTGVVHHQTHTLRNGNV RFWSIHFGKFWKSHSEWVFGDSDDKTGGDIN TVFRPIKLDGNKIALNLGNNGYCMRRVQRGQPR LIDTLIASASSYGPLAVVEVEAVRQREIYNIHYN MDESRIYVVPFESGNVTQDTYNTLPLPVIKYG FKYQVTRTFDSSHWSWKVAYNSKINVTIPKFV EGEATLDASIEYKVEWSELNETTITQTKDLIYV PPRTKSVVTLILGRGSDVDPFSEYEQROILPNGRE VHKHMEGLFSGNLTEVITHYREISLSTLSESEL EEAPRLSEAEAAAAALSSKL

10

20

30

40

【表 1 4 - 5 8】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数 sp_107620_kqnh (SOAP1)		sp_107670_ptqx (SOAP2-like)	注釈 UDP-glycosyltransferase superfamily protein; FUNCTIONS IN: transferase activity, transferring glycosyl groups; INVOLVED IN: metabolic process; LOCATED IN: endomembrane system; EXPRESSED IN: 6 plant structures; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage, 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN's: UDP-glucuronosyl/UDP-glycosyltransferase (InterPro:IPR002213); BEST Arabidopsis thaliana protein match is: UDP-Glycosyltransferase superfamily protein (TAIR:AT5G53990.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).	aa seq MEYSKKEGNNLHIGMPWFYFAYGHINPIQLSNK LSSHGIQISFFSIPGNINRIKSSNLNLTPPNQLPLTIP PIEGLSPNFDSSSEVTHETSQLSLALDQMPPQV KSHLLQLQPOIKVILFDYAYHWLPSLASELGIKSV HFNTPFLVNSYLTVPSSRMTDPNKIPTVEEFKPK PHGYPKTSTAASVKTFEAEAYNFKKCFNGCHYE KLLGFTNRCDAIYRTOCNEVEVTSDELDTKWRIW LSKFPEKSVIYCSFGSETSLNNAIQIELTLGLELT GLPFILVNFSGTNTIDAQNKLEASLPDGFREIRIK NRGVLHTGWVQQQNILAHVSGCFLTHAGFSSV IEGIVNDCHSVN
145	sp_002570_ltmw	0.9483338	#N/A	0.9533485	FASCICLIN-like arabinogalactan protein 8 (FLA8); LOCATED IN: anchored to plasma membrane, apoplast, plasma membrane, anchored to membrane, plant-type cell wall; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN's: FAS1 domain (InterPro:IPR000782); BEST Arabidopsis thaliana protein match is: FASCICLIN-like arabinogalactan protein 10 (TAIR:AT3G60900.1); Has 14155 Blast hits to 6880 proteins in 856 species: Archae - 79; Bacteria - 4506; Metazoa - 1333; Fungi - 788; Plants - 1999; Viruses - 868; Other Eukaryotes - 4582 (source: NCBI BLink).	MASAKTTTVFVSLAIFAVSYAHNITEILSSDPSSL QFSSFLTQTKLDDEINRSQTTV_LALDNAMATLT AKHPIISVWKNLLSHLLLDYDPPKHLHSIDNGTTLT TTLYQTTGSASSQEGSVNITDLKGGSVGFGSGA PGSKLDSTYTKSVKQIPYINISVLEISAPIAPALLTA ASSDVNVTGVLEKAGCKTFASLISSSGV_LKVYQS QMDKGITFAPNDEAFKAKGVDPDLTKLSNAEIVSL LOFHALLSSYSPIGTLKTTKGSPLTASGVGKYDLS PSAAGDSVTLHTGVDSSRIASITLLDAPVWYVAVD NVLPLVALFGMSPSPAPSPDSDSPSPAPAPAD SPSPAGTFMSSPPAPSSSESPADSPADSPAGTS ADENTKNGNERRVPAAMFIAVGLSVVSGVSLVL
146	sp_003130_hrrf	0.9487302	#N/A	0.9561284		

10

20

30

40

【表 1 4 - 5 9】

	A	B	C	D	E	F
1		分析に用いたペアとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_ktag (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2					cytochrome P450, family 78, subfamily A, polypeptide 6 (CYP78A6); FUNCTIONS IN: electron carrier activity, monooxygenase activity, iron ion binding, oxygen binding, heme binding; INVOLVED IN: oxidation reduction; LOCATED IN: endomembrane system; EXPRESSED IN: 13 plant structures; EXPRESSED DURING: C. globular stage, 4 leaf senescence stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAINs: Cytochrome P450 (InterPro:IPR01128), Cytochrome P450, conserved site (InterPro:IPR017972), Cytochrome P450, E-class, group 1 (InterPro:IPR002401); BEST Arabidopsis thaliana protein match is: cytochrome p450 78a9 (TAIR:AT3G61880.2); Has 32637 Blast hits to 32519 proteins in 1695 species: Archae - 48; Bacteria - 3588; Metazoa - 11617; Fungi - 7095; Plants - 9209; Viruses - 3; Other Eukaryotes - 1077 (source: NCBI BLINK).	MSTNINIDTWVFALEFSKCNHFSTFNLFISLFIIVAV WLTVALHWSYGGPAWGGKHLRARMIFHWAKK PIPGPRLPIIGSMLMTGLAHRKISDMAKTCNAK RLMAFSLGETRIVTCHHDVAKDILHSSNFADRP KESAYSIMFNRAIGFAPYGVYVWRTLRVASTHM FSPKQIKASEGRYVIA SAMITGYFGTGTGPPQPR DVLRRASLSNMMSVFGQKCSLVEGPENNKE SEELREMYEGYDLLGMLNWDHLSWLSDFDF QKIRFRCSQLPKVNRVFGRIITEHRANVPNSP DRVDVLLNLPSSDKLSDSDMISVLWEMIFRGTDT VAVLIEWLARVYLPDVSQSRVHDELNDVYGRSR AVYSESDITSLVLTAVIKELVLRVHPPGLLSWARL AIGDITVDGHHYPAGTTAMVDMWVAITRDFEWA EPDTFNDRFLGSGAASEISLGSDLRLAPFGSG RRACP GKALGLATV/SFWVASLMHEVEWVSSAA DATVDLSEVLKLSCEMAHPLTVEVRRRRASVM
147	sp_005150_thof	0.323743	#N/A	0.3273124		MGNISIQDQTKMSAAKSSLRGCASKFPQPLSSE DLQEAQTLGDTVYGVNHDKLLQDIYRSSEGDK DVLRQVFRWDLAPYGDVFGHGFQAHKRVHD KTSYDLHSYVQGRGPLDSRKTSDFAFISTVSN SWYPPVATGDGNSKTVEVYRYEYAPGGIWTSL TLGKNDKRFKRLDEVVFGGASQYVRSQAQLFR LTGNPDGTTTIERANNVLLINNFPSPQSHPLKLLN IGNPVSdyTDSNRTPLSQTYVGSNNNNNTNII NYYANDVANLDVYMSAALRSSTNEAYVLOQL VTLNLYDPGGTHDHDVDFGFSVSEMFKSLEYTP FGEHGIDGALKSRDDEAFIFCANLCLVINYAPDTR NDTVKGMVTYVEMFPFKGTVFENGIDSACEIK SAEDEAYFFKGNLCAHINYGSNPQLITQGGIQFI PLKGTVEFGIGVDASASHAEVAYEYFKYNNY GQINWVEDGVEVLYGVFVWMDNAPVIHSFVFPQ KNRGLDIRSNP GALPSPDEYDYDEL
148	sp_008780_frur	0.323382	#N/A	0.323303	uncharacterized protein	

10

20

30

40

【表 1 4 - 6 0】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 carboxylesterase 20 (CXE20); INVOLVED IN: metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: root, flower; EXPRESSED DURING: petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Alpha.beta hydrolase fold-3 (InterPro:IPR013094); BEST Arabidopsis thaliana protein match is: alpha.beta-Hydrolases superfamily protein (TAIR:AT2G45600.1); Has 9350 Blast hits to 9331 proteins in 1590 species: Archae - 113; Bacteria - 5436; Metazoa - 449; Fungi - 960; Plants - 1345; Viruses - 3; Other Eukaryotes - 1044 (source: NCBI BLink).	aa seq MKNPIKLGTFIFLVTILHNNHLAKAQPPITINPYEF LHIPNPDGSIIRIKEFFFTVPNTTSLAISKDVPL NPYKKTWVRLYLPSGLSGPPGYLNNLPIVFAHG GAFVLLSTATPDIDLFLSNTASKLHLVSVSEYRL APEHRLPAAYDDVLEALYWVKDKDEWVGKYG DVSRCIMMGESVGGNIVNVGLRASVLTTRVLKPL VIRGLVLIQIPFYGGNLRGKYDPNIEVVADLLW NLSPLIGANWNPYFNPTFGGGSSNLDIIRRLG WRVAFAGCDGDFDKQKLVFEFFKGRGLDVIY FGRGYYHGVEFVDPSESKIFEFVRYVYSSYI MACEMKAIKRYGELLFTAINATISNGADAHIDEF PAQQHQILPISNILFEAFIPNAARAGGSLTCLGE GEHCWTECCPGLTKFLGNTSFCGWCPSAGDP CGALDHCCDGYSCDGFSSGTCH
149	sp_014050_fdgk (SOAP1)	#N/A	#N/A	uncharacterized protein	HGYSNSWADQWDSNNDQSSSLNKNKNSNN NGVKSQKSMTKKLSKKFGESFNKTEASANGVE KTKESVSKGMRKAKDITTSQSSISWVYQKFRSSY K
150	sp_015660_nxer (SOAP1)	#N/A	#N/A	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G19020.1); Has 84 Blast hits to 84 proteins in 15 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 84; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink).	MPERNVSVWTALLSGYSQNGFFKDALMVFEDM HHSVYKGNQFTYGSALRGCTKLMCLFSGKQIQG CLQKGRFSSNLFYQSALVDLHSHKCGSINDASIIQF TMPYRDLVCWNTMIDGYAVQGFVNDAFQLPFLM MKGETIPDCFTFGLLRACQGNLRSRVQIHGL MIQHGYLEHLLNRLLDIAYAKCGSTKHAYTVFK SMNNVYIFIVELWKTTSCTSKVFSYFDVALGN ALIDMYKCGELKEAIRAFDQMEEKVVISWSSLID GYAKHGYGHEAMLLYRKMEHEGLKPNRVDVTLISL FACSHAGLIREGWEFCFNMMVDKHNIAIPEKHYH CMDVLAGGGQLEEFNLQKMNIIQPNASHWGA ILGASVHGDIQGLKAAKLLFYLDTKNSYVAL ASIYAAASGLWNSDWSWGTYKFMEEENLKKSPGGSF DLCEEKRTALLOS
151	sp_017800_nkhn (SOAP1)	#N/A	#N/A	Tetrairicopeptide repeat (TPR)-like superfamily protein; FUNCTIONS IN: binding; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 14 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: Tetrairicopeptide-helical (InterPro:IPR011990), Pentairicopeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: Tetrairicopeptide repeat (TPR)-like superfamily protein (TAIR:AT2G27610.1); Has 41688 Blast hits to 13542 proteins in 246 species: Archae - 0; Bacteria - 10; Metazoa - 149; Fungi - 146; Plants - 40815; Viruses - 0; Other Eukaryotes - 568 (source: NCBI BLink).	
152	sp_022280_xyzq (SOAP1)	#N/A	#N/A		

10

20

30

40

【表 1 4 - 6 1】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
153	sp_027880_kudg	0.923046	#N/A	0.83443	basic helix-loop-helix (bHLH) DNA-binding superfamily protein; FUNCTIONS IN: transcription regulator activity, sequence-specific DNA binding transcription factor activity; INVOLVED IN: regulation of transcription; LOCATED IN: nucleus; CONTAINS InterPro DOMAIN's: Helix-loop-helix DNA-binding domain (InterPro:IPR001092), Helix-loop-helix DNA-binding (InterPro:IPR011598); BEST Arabidopsis thaliana protein match is: sequence-specific DNA binding transcription factors;transcription regulators (TAIR:AT5G50010.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLink)	MHNIRKCSNRRGVAPLENQGVVNGQHNVSPA SAFGGSLPSSAKMPPRVHDIIEFQSEVCPQNFII VDEIDANQSRIMYHPGIAHNLSTPNMDFHETPNG EITQENFGNFFEGGIVSPDMEDSDYIDLMLSLDE NQQSEEEVSTAGTQGGYDGGSSSPESFCNYTTK CRKATYPCSVQRASTGSSSHNEKNKFKEARKMVK VLRGMVPGGDMNTAAVFDAAVHHLKGLGKV QRLGLSNHKHFP
154	sp_032700_tnie	0.933038	#N/A	0.7244362	Integrase-type DNA-binding superfamily protein; FUNCTIONS IN: DNA binding, sequence-specific DNA binding transcription factor activity; INVOLVED IN: regulation of transcription, DNA-dependent; LOCATED IN: nucleus; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN's: DNA-binding, integrase-type (InterPro:IPR016177), Pathogenesis-related transcriptional factor/ERF, DNA-binding (InterPro:IPR001471); BEST Arabidopsis thaliana protein match is: Integrase-type DNA-binding superfamily protein (TAIR:AT5G18450.1); Has 5751 Blast hits to 5635 proteins in 247 species: Archae - 0; Bacteria - 2; Metazoa - 0; Fungi - 5730; Viruses - 2; Other Eukaryotes - 17 (source: NCBI BLink).	MDNSKSTLSKPWKKGPARGGPQONATCMYR GVRQRTWGWKVAEIREPKKRSRLWLGSFSTAE EAAALAYDEAARRLYGPDAYLNLPHLRCSFNPLNK SQRFWRFPKKNFIMFPASSGSSSTPSLLNVAQ PSLHVHQRLQOEKKNKSEPKVVKSKRVLVEQVS SNEHINWEKPKQDLNLEFLQMGVLEGETTYTND KYDVAVNHDETAKEIFIEEMNSEWNLVELRED HQQAESRSRFGDYTMDLITLPTTIWDF

10

20

30

40

【表 1 4 - 6 2】

I	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
155	sp_038870_ruye	0.3439733	#N/A	0.3324831	cysteine-rich RLK (RECEPTOR-like protein kinase) 5 (CRK5); FUNCTIONS IN: kinase activity; INVOLVED IN: defense response to bacterium; programmed cell death, response to salicylic acid stimulus; LOCATED IN: endomembrane system; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAIN's: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Protein of unknown function DUF26 (InterPro:IPR02902), Serine-threonine/tyrosine-protein kinase (InterPro:IPR01245), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: cysteine-rich RLK (RECEPTOR-like protein kinase) 4 (TAIR:AT3G45860.1); Has 121629 Blast hits to 120099 proteins in 4447 species; Archae - 100; Bacteria - 13629; Metazoa - 44864; Fungi - 10560; Plants - 34243; Viruses - 400; Other Eukaryotes - 17833 (source: NCBI Link).	KSKADRRSHILNISLGVVAILVLLVALLIYFLCTKRRITNDSRRRENHITNFKNVGESFMSVESLRIIDVAILEIATNINFEEDNKLGGAGFGGYMAPEYIEYGEISTKS DVYSFGVLVLEISGGKISDRQSELAENLLTLAWMHWSEGLPLELVDLSTLGGDSYSSNEIRCLHLALLC VQGDVEDRPSMLSIIVFYINCSSYLPTPQQPAFLI SRRKGLSLGLGSLNDFDKSTKSGSSSGTSGQHS VKSTSRSNNSFSGIESTMY
156	sp_041880_wyef	0.966235	#N/A	0.971423	uncharacterized protein	MARVSNLVVVPAILAITFMLSTATVLGKVTTSE EMEMPVGYKMMLEEMALARKVQSNMYDQPAAV RRSQAAMKGLMDVLVSSSSEKAMQCVAAAGGYC NLVFGPRCCTNDWACVPPGLVGGVCVA

10

20

30

40

【表 1 4 - 6 3】

A	B	C	D	E	F
1					
2	sp_107620_kqmh (SOAPT)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
157	sp_046690_sjiw 0.9428554	#N/A	0.9443533	cytochrome P450, family 715, subfamily A, polypeptide 1 (CYP715A1); FUNCTIONS IN: electron carrier activity, monooxygenase activity, iron ion binding, oxygen binding, heme binding; INVOLVED IN: response to salt stress; LOCATED IN: endomembrane system; EXPRESSED IN: 8 plant structures; EXPRESSED DURING: 4 anthesis, C globular stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Cytochrome P450 (InterPro:IPR001128), Cytochrome P450, E-class, group I (InterPro:IPR002401), Cytochrome P450, conserved site (InterPro:IPR017972); BEST Arabidopsis thaliana protein match is: cytochrome P450, family 735, subfamily A, polypeptide 1 (TAIR:AT5G38450.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MSLVHFVYGFIFISLFLWLKLNACVISPVRTLRKL RNSGLGGPTSPFLGNLSDMKOMLIKAKLSLSSSS LLSSSESISHDIHSSLLPYFAQWKKLHGKYVYV WMGVPEVYIADPEFLKEMSGKVFVSRNWGKPN VFRSDRVPMLGEGGLSMIEGNAWARHRHITPAF SSTNLKGMVDSMVESTIKMIKNWSTLLGNSSKLE LELDVERFPVLGGETMAKTNFGIHNSEVENTLL EKLREIQISLFKHTRYVGYPGKFLSMRQTLKAR KLKGEIDGILLINARREIIGVEPQHEHDLRLLE GEDHFDGQLTKLEIDCKTFFGG-HETVALSMT WTSM.LLAMYKAWQDELREEIKEVIGDKDDIDFTM LARLKKMGWVFNESQRLYSPSPNLQOARSDIQ VGNVVPNGTNIWIDVGMHDDPDLWGNVDNEF KPERFDGDL.YGGCKNKMGYLPFGFGGRMCIG INYALMEYKIVMISLILRRFSFVAPSYHHFFTYFFS FRPGSGVPLVVKPL
158	sp_053590_purn 0.9403215	#N/A	0.9372205	AP2/B3-like transcriptional factor family protein; CONTAINS InterPro DOMAIN/s: Transcriptional factor B3 (InterPro:IPR003340); BEST Arabidopsis thaliana protein match is: AP2/B3-like transcriptional factor family protein (TAIR:AT3G11580.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MNFSLNLPSSDTLWYPTTMEHPKSLPFFYH SSLSSKHDDVDSNDSNTNIAATNAIDRSETTS SSPSSADQEYMFKEKPLTPSDVGLNRLVPIKQHA EKYFP LAGSSTVDSSTSSVGDKSTGGVGGGVL LGFEDESGKCKWFRYSYWNSSQSYVLTGWSR YKKEKLDAGDVL.FQRHRELESDRFFIGWRRRG SAPLSQSNPVRVYNNNNNNNNNNNNNGNYGVVQ SQSYNGAPQHQIASFPSYRDPDVLHTGTETVRQK GRKSGGGNSVRLFGVNLCCIDGCEQEEDEEA QTSDSSSQGGQVFGVGPYAYNYFDYVYNNHY PSSHMDMNYPQDVNQMRM
159	sp_056310_xuqe 0.945208	#N/A	0.9338837	late embryogenesis abundant protein, putative / LEA protein, putative; INVOLVED IN: embryo development ending in seed dormancy; EXPRESSED IN: seed; CONTAINS InterPro DOMAIN/s: Late embryogenesis abundant protein, group 4 (InterPro:IPR004238); BEST Arabidopsis thaliana protein match is: embryonic cell protein 63 (TAIR:AT2G36640.1); Has 58153 Blast hits to 33178 proteins in 2866 species: Archae - 714; Bacteria - 20429; Metazoa - 14816; Fungi - 4800; Plants - 4772; Viruses - 420; Other Eukaryotes - 12202 (source: NCBI BLINK).	MMERQRFMGKMLSLVIVAVVVAACNNSVSAH EVEEIEKESGESWTGWAKIEQGLGKQQGEE LSSGAGEAKEKVTDKASMTKDAAMDKAKEYGRE ASNKASNVKDRVQQQAEATAADKAQEAETVKAD AQEAKEAARAKQAQAKEMAKDEAQLHLETTTRNK AQESKTKMKNVEETKEAARVAQARAREMAKDE AKLESESTGNKAQEAKEKERMNTNKVEQTKTEATRK AQEAKEMAKDAKAAVEEHSWVWDKASEGYDA LKKKASDALQAKEMAQDMTSARYIINDDDDDDD DEL

10

20

30

40

【表 1 4 - 6 4】

I	A	B	C	D	E	F
		分析に用いたペイトとの相関係数			注釈	aa seq
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_piqx (SOAP2-like)		Phosphorylase superfamily protein; FUNCTIONS IN: catalytic activity; INVOLVED IN: nucleoside metabolic process; LOCATED IN: endomembrane system; CONTAINS InterPro DOMAIN/s: Nucleoside phosphorylase (InterPro:IPR000845), Nucleoside phosphorylase, family 1 (InterPro:IPR018017); BEST Arabidopsis thaliana protein match is: Phosphorylase superfamily protein (TAIR:AT4G24350.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI Link).	MGMRYELTVILMVLVHQSQVLRSTHPLHEVVDR INHHGGPYIGLV_LAFPAEQVPLNSGLFVFNRSRN PFIQLSGRKNFKIKGVDVIVYVLTGEQTYNAAITV QSLLQTFNIKGVHYGIAGCTNTSMSVGDYGVPH SFAFTSSWKWLDGFSSEETSEKELKFGAYNHPQT GENLLRIQIFSPVQLFTNGKAMEEIIFFWVDAH WYNLATQLQRWKLQCCQLNETACLKEPKVVFGL RGSTANVFMNSLAFAKFLYEESISTADEESAIV VHASLSNGVPCVVFGRGVSVDVAGGEKALALAGLSS LAAINSFKVAAKFVELIGRKKATF
160	sp_064490_imro	#N/A	#N/A	0.963217		MMMANVTERKYEDFPVSDLLSEKEDTLLLYLP GFRKEQVKQLTTSRILRISGERPTGDNKWKRFN KEVHTSNVDMKEMTAKFEGGILYRQPKLITPAE PAATAAPAEPAAPAEQ.LDQQGGPTGGNMEKKE EGEMDRKEKDEISAESNEERKGVWMMNELREQ VGVRIQLFKELSKPEKRLRLAVYVLLLLVFCQITK MFKSFTKVEEQTVLEFNHQEF
161	sp_066380_dtgp	#N/A	#N/A	0.943253	HSP20-like chaperones superfamily protein; FUNCTIONS IN: molecular function unknown; INVOLVED IN: response to heat; LOCATED IN: membrane; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Heat shock protein Hsp20 (InterPro:IPR002068), HSP20-like chaperone (InterPro:IPR008978); BEST Arabidopsis thaliana protein match is: HSP20-like chaperones superfamily protein (TAIR:AT5G20970.1); Has 712 Blast hits to 708 proteins in 96 species: Archae - 0; Bacteria - 37; Metazoa - 7; Fungi - 2; Plants - 649; Viruses - 0; Other Eukaryotes - 17 (source: NCBI Link).	MGTTVNLILFPLSLELLTLQLPQNAENFKLANN YTISATNPELAASNGGKCNFGKWVFDKSKYPL YDSLNCPEIDLEFNCGSRPDKLFQKYRWQPFQ NLPRFNGLYFLRWKRGKIMFVGDLSLNFQFESL ACMIHSWPNTKTSYIRGGALPSLTFOEFGVTIM LYRTPFLDIVREKVGRLKLDSSISGNAWKEMDM LIFNSWHWHTHTGRAQPDWDYIQVGNKLYRDMN RWIAFYKGLTTWARWVDLNVNPSKTKVFFYGV PSHDFGSDWDEPSSKCSSETEPFFATRYPAGTP QEWAVVEKVMRSRIRKPVYWLDTTLSQYRKDGH PSKYGGHGKDCSHWCLPGLPDTWNLQLLYAALF Q
162	sp_068240_kqid	#N/A	#N/A	0.963211	TRICHOME BIREFRINGENCE-LIKE 39 (TBL39); INVOLVED IN: biological process unknown; LOCATED IN: vacuole; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF231, plant (InterPro:IPR004253); BEST Arabidopsis thaliana protein match is: Plant protein of unknown function (DUF828) (TAIR:AT2G3110.2); Has 1331 Blast hits to 1312 proteins in 27 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 2; Plants - 1329; Viruses - 0; Other Eukaryotes - 0 (source: NCBI Link).	

10

20

30

40

【表 1 4 - 6 5】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2				<p>HXXXD-type acyl-transferase family protein;            FUNCTIONS IN: transferase activity, transferring acyl groups other than amino-acyl groups, transferase activity; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Transferase (InterPro:IPR003480); BEST Arabidopsis thaliana protein match is: HXXXD-type acyl-transferase family protein (TAIR:AT5G07860.1); Has 2450 Blast hits to 2444 proteins in 162 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 125; Plants - 2319; Viruses - 0; Other Eukaryotes - 6 (source: NCBI BLINK).</p>	<p>MAKSEQETMAKSEQETSIKLVSECFVKPKYEIKS            AKOPYHLGPMDLVMLTIDPIQKGLVFTIKNSPLFL            SSESNDNIEIIRTKVVSRLLEKHKLSLALVHFYPL            AGRFTTQKQPEHNTSLVFDCKNGKPGARFIHATS            LDFTISDILSPVDVSVHSFFDLGEKHVNYDCHTK            ALLSIQVTELLDGVFVIGFMSHSVVDGTSFIHFVN            TLSEIFKSDDFTTISRAPILNYRPPCDIPIKFFFLDV            EGFICRAYNPGLRERIFHFSLNSMLRLKAMANQ            ECGTQNVLSSFOALTAVVWRSITRVRNLPKDEQ            TTCFMAMGSRTRLNPPLSDDYFGNFMISTKFAK            AEELGNSLGVAVMNLRKIIMSTDEKSILATYKAL            ADSPVIPRETIPTIPGPHGMRVIRIIGSSRFDMYGP            FGLGRALAAARMGYGNKDDGKITANPGCCGCGGS            VDLEICLRPHIMASLEVDQEFMGEFVS</p>
163	sp_074630_ygho 0.924372	#N/A	0.923833	<p>CD2-binding protein-related; FUNCTIONS IN:            molecular_function unknown; INVOLVED IN:            biological_process unknown; LOCATED IN:            cellular_component unknown; EXPRESSED IN: 25            plant structures; EXPRESSED DURING: 15 growth            stages; BEST Arabidopsis thaliana protein match is:            suppressor of abi3-5 (TAIR:AT3G54230.2); Has            30201 Blast hits to 17322 proteins in 780 species:            Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi -            3422; Plants - 5037; Viruses - 0; Other Eukaryotes -            2996 (source: NCBI BLINK).</p>	<p>MEGFSSSRRKRFIDDDDSNKNKQKRVRFPKK            KKLKTGEAPEPVAVKDEAEVAGFVDLSDPRTAA            KERAKRRGQMTTQLFTEDGEDVEVEVAHISAAEV            QYQENESFVDDGVQLEPNLEREREVEEYFDEAG            NFVEYVPEKEEKDAWLDLSLEVYPKLANKVSTYV            NSDDEEQEMSADDTGKMKRRIADVLEPGETVLQ            ALKRLKGSTDKRVKMPAEKVVFDQLTEDAMRL            MENDYNYVHEEREVFRVREAEYEALEARARNND            VSGIEDDGLDMFAEDNINVAANSEPNNDTSTPS            VGTSTQLAVENGSDSGSTWAGQSDYVYDEAS            GYYYSSSLGYYYDPSTALFCSASSGQWYKYNEA            TGTYYDEYQSETEAPSVAS</p>
164	sp_084110_jhrq 0.933339	#N/A	0.923648		

10

20

30

40

【表 1 4 - 6 6】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
165	sp_094130_zpgz	#N/A	0:328013	Leucine-rich repeat receptor-like protein kinase family protein; FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: plasma membrane, membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Leucine-rich repeat-containing N-terminal domain, type 2 (InterPro:IPR013210), Tyrosine-protein kinase, active site (InterPro:IPR008266), Leucine-rich repeat like domain (InterPro:IPR01611), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), BEST Arabidopsis thaliana protein match is: Protein kinase family protein with leucine-rich repeat domain (TAIR:AT1G35710.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI Link).	MSFKALFLFFCILLQFSFISFAGAIQAEAEALVE WKQS_HLSLNOSELSW/S/PRDALTKNSNVTTP CHWLVGCEGNSVTR_NLTVNGLTGTIPPEVGY MISLTELRLNINLITGPIPPSIGNLRGLKTLVSYGN KLSGSLPKFEDNLTNLTLCFLSNNTLSGSLPEKIC QGGILDFCASNRFSGTVPKGLKNCSTLRLRL DRNNFYGNITEDFAAYPVVDYIDLSYNNFEGEISS DWAKCKNMTSLKISDNNITGKIPPELGKATKLHYL DLSNNQLVGEIPKELGNLKSFLNLTLSNNKLSGNI PREVGNLDDLASLDLGGANGLNGTIPEIGNCPKMI YLNLSRNLGNGVWPQIGSLVSLQVLLDLSRNSL SGEIPQLGNLNDLEELDLSHNNFSGQIPSTFDQ MQSLISVDVSYNNLEGLPYGKGFVMSPLVSFME NKALCGNITGIPCPRALNFSKPKGKNAVALITAPI LVGVFVIVAVAVALLKGGRRGKGEKNRKRVG SQRENNMFSWSDGKLVYEDIKQATKGFDFVKYCI GEGGHGTVYKAVLSTGQIVAVKLLKTIQNPSEFN AKNFEAEIALSKIRHNMVKNLQNGFCSCHAQHSLV YEYLERGSLGKLLGIEKEARELDWEKRVIIVIGIA NALCYMHYDCSPPIHRDLSSNNVLLDRDYEPRV SDFGTARLLSLDSCRNIWLHCSW
166	sp_092350_fyt	#N/A	0:328124	Peroxidase superfamily protein; FUNCTIONS IN: peroxidase activity, heme binding; INVOLVED IN: oxidation reduction, response to oxidative stress; LOCATED IN: cell wall, membrane, plant-type cell wall; EXPRESSED IN: root, leaf; CONTAINS InterPro DOMAIN/s: Haem peroxidase (InterPro:IPR010255), Plant peroxidase (InterPro:IPR000823), Peroxidases haem-ligand binding site (InterPro:IPR019793), Peroxidase, active site (InterPro:IPR019794), Haem peroxidase, plant/fungal/bacterial (InterPro:IPR02016); BEST Arabidopsis thaliana protein match is: Peroxidase superfamily protein (TAIR:AT3G03670.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI Link).	MKGSFLTLILFLGLIMTRVCHCYGGYGTQLQ QRFYEGKGGNDVSVIFGIHKIASEPDTVSD LVRLSFHDFVRGCDGSLFDGDNTEQKAVINCG IGGIDTINDLDAVEKVCVGVVSCDILVIGARAAI SLAGGKYEVETGRRDGVVSIKSEAOAIPPTM PIQAIALFAQKGLTKDDFVLLGGHTVGTAHCH SPKDRLYNRYTTQKPDSTISTITLLQNLOKTCPLN DETDNEAFLDQTPDSNFKIDNAYYKQILAQNGVL EIDQNLAASSPQTKDMVNTLAQGPDTFQNLFGRA MVKMARIGVLTGTQGIKSCSGSINT

10

20

30

40

【表 1 4 - 6 7】

A	B	C	D	E	F	
1						
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	
167	sp_097540_mwnc	#N/A	0.9342338	0.9342338	TCSRRKTSMDTLAELGWLSDFFHEQKGEFAKAI KCLEAICQSPVSELPVIEVKTRLRRIATLLLRHSHNV NHAKSLERSQLLKSIPSCFELKCRAYSLLSQCY HLVGAIPTQKQILNKGLELTKSLNGFAAKLWYC NFSSQLSTALIEGDYQNSLATLEQGYKCAEMC YPELQMFATSKLHVHLMQWDDVDMVERALAR CDEVWNSFELNKTQOCLGLLFYNELLHIFYRFRI CDYKNAHQHDKLDAAMKADLQKTEHVQKLTSEL NTINQTLSDPAMKQRDCSMLIEKRALLKQQLVTT TSNSVAQGYTDSVVEYERRLSADKLELAPYID GEWLPKNAVYALVDLTVVMFGRPKGHFCKTKR IQSGTQVIVKVELSKLGITDGVREVDLQHSAIWMA GVYLLMLMQLENKAVDLTRADFFIAEQALVQM RGWYIRFPTILQACESIEMLRGQYSHSVGCFDE AAFHFVEAKLTKESKIQAMCQVYAAVSYICGDA ESSOQALDLPVYVWVMSDFIGAREKTSVLFQYV LLLMKQNLQEARIRLATGLQTHSNLGNLQLVS QYLTVLGSLALALHDTVQARELLRSALTAKKLNLD VPTQIVLVSLSALYQELGERGEMENAEYQK KIDDLHKRFADAYASHHTALINKDKIEVHQLQES DSKRAMASQSMKVDLPIPEISGLASSTPVKSSSR LVDLDSGRRGKRKI	Tetraicopeptide repeat (TPP)-like superfamily protein; Has 214 Blast hits to 140 proteins in 56 species: Archaea - 0; Bacteria - 4; Metazoa - 151; Fungi - 0; Plants - 55; Viruses - 0; Other Eukaryotes - 4 (source: NCBI BLink).
168	sp_101870_mdrns	#N/A	0.9386392	0.9386392	MAEQSNFLAEKRYAVVTGANKGIGLEICRQLASQ GVVILTSRDGKKGLEALNDLKSIGSSDNLHYHQL DVTDPMSITALAGFNSKFGKLDLVNNAIGGGFII DYDAIKGIFRNINYDEMIMTQYELAKECELETTY GVKRTTEALLQLELSDSPRINVSSTGMLKNIP NERIRGVLDGAEALTEEKVEAILNELLTDFKDGFS KEKEWPSRMAAYTLSSKAALNAYARILAKKYPSIIIS CVCPGVTKTDMNGNLGQLTYVEEGAASPVRAVLM PHGSPSGLFYARSEVSSVE MFSKFAVFLMFLATSGVFDIARYQPRSSGIMT EELLEAEVKALASLREVERAVIASKLKQNLMMRT TEDNCKISGTCISGFGPPEQCCSGACVPHPIIKMF VCC	NAD(P)-binding Rossmann-fold superfamily protein; FUNCTIONS IN: oxidoreductase activity, (-)-menthol dehydrogenase activity, (+)-neomenthol dehydrogenase activity; INVOLVED IN: response to karrikin, defense response; LOCATED IN: plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: NAD(P)-binding domain (InterPro:IPR016040), Glucose/ribitol dehydrogenase (InterPro:IPR002347), Short-chain dehydrogenase/reductase SDR (InterPro:IPR002198); BEST Arabidopsis thaliana protein match is: NAD(P)-binding Rossmann-fold superfamily protein (TAIR:AT2G24190.1); Has 78247 Blast hits to 78183 proteins in 3234 species: Archaea - 733; Bacteria - 53633; Metazoa - 4843; Fungi - 3816; Plants - 2340; Viruses - 0; Other Eukaryotes - 12882 (source: NCBI BLink).
169	sp_114650_ukiw	#N/A	0.9348233	0.9348233	uncharacterized protein	

10

20

30

40

【表 1 4 - 6 8】

I	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh sp_107660_kiqg (SOAP1)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	
170	sp_118730_ygz	0.221048	#N/A	0.237327	basic helix-loop-helix (bHLH) DNA-binding superfamily protein; FUNCTIONS IN: DNA binding, sequence-specific DNA binding transcription factor activity; INVOLVED IN: regulation of transcription; LOCATED IN: nucleus; CONTAINS InterPro DOMAIN/s: Helix-loop-helix DNA-binding domain (InterPro:IPR01092), Helix-loop-helix DNA-binding (InterPro:IPR011598); BEST Arabidopsis thaliana protein match is: basic helix-loop-helix (bHLH) DNA-binding superfamily protein (TAIR:AT2G22750.2); Has 2910 Blast hits to 2902 proteins in 174 species: Archae - 4; Bacteria - 2; Metazoa - 77; Fungi - 51; Plants - 2768; Viruses - 0; Other Eukaryotes - 8 (source: NCBI BLINK).	MEKQALSTMEELLSSAALHGEDLCYPLLGKSNITW YPESQWELEYDLVDQFITSPPSYTNSDSSEIISHD DYEFYDNHNDGNDNASTVKMLMIENHNKVV SEKQAVNGKRRKRKESQVODHILAERKRRET LAQFLSSTLIPGLKIDKTSILRGAIKHMKHLQER VVLENLGAQQKNLKSLLVVKKCOLTVDDDDN NDDNGFSIIDDGSSNDAYLNNNSPEIQVKTDD DTLLLRVHCEQQKDIMTKLFDVVVKKHMSVAHCA VIPFTNLADITVITVOMEDGFSMDVKDFVRSIRYILH PAS
171	sp_118740_hmni	0.962863	#N/A	0.333224	basic helix-loop-helix (bHLH) DNA-binding superfamily protein; FUNCTIONS IN: DNA binding, sequence-specific DNA binding transcription factor activity; INVOLVED IN: regulation of transcription; LOCATED IN: nucleus; CONTAINS InterPro DOMAIN/s: Helix-loop-helix DNA-binding domain (InterPro:IPR01092), Helix-loop-helix DNA-binding (InterPro:IPR011598); BEST Arabidopsis thaliana protein match is: basic helix-loop-helix (bHLH) DNA-binding superfamily protein (TAIR:AT2G22750.2); Has 2910 Blast hits to 2902 proteins in 174 species: Archae - 4; Bacteria - 2; Metazoa - 77; Fungi - 51; Plants - 2768; Viruses - 0; Other Eukaryotes - 8 (source: NCBI BLINK).	MEKQAFSTMEELLSSAALHGEDLCYPLLGKSNITW YPESQWELEYDLVDQFITSPPSYTNSDSSEIISHN DFNDDNNNDGDDNASTVKMLKTTPTGVQOE PDRQSENNKEVTKQVNVGGGGGKRRKRQE SQVQDHILAERKRRET LAQFLSSTLIPGLKIDK TTI LGEAIKHMKHLQEKVNALET LAQQKTLKSM VVVKYQLTVANDNDNDNDGSRNAINFNNSI PEIQVKTDDTLLKVVHCEQQEKIMTKLFDVVVKK HMSVWNCALIPFKNLAODITVTKMKEGFDLDVKG FVRSIRSLHPTS
172	sp_123950_hmts	0.958147	#N/A	0.332423	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MTVMARAADVTSKKGTTFFSFFQYSGLKFLGRFN PFMRYGLPLISLTVFGLGLHLLQGGKDKISKVKD DKEWEIIEERKALSRTPGVNAYNPKNISLEELKA FQEKVDINSYDYKRIPLPNEGKSAR

10

20

30

40

【表 1 4 - 6 9】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
173	sp_128030_ztbg	#N/A	0.943525	fructose-bisphosphate aldolase 2 (FBA2); FUNCTIONS IN: fructose-bisphosphate aldolase activity, catalytic activity; INVOLVED IN: response to cadmium ion, pentose-phosphate shunt, response to abscisic acid stimulus; LOCATED IN: in 7 components; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAINs: Aldolase-type TIM barrel (InterPro:IPR013785), Fructose-bisphosphate aldolase, class-I (InterPro:IPR000741), BEST Arabidopsis thaliana protein match is: fructose-bisphosphate aldolase 1 (TAIR:AT2G21330.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MATATATTTTSLKSSSAEGLDKSDWIKGQS IRHVAVKCNFVNSLSVKASSYADELVKTKSI ASPGRGILAMDESATCGKRLDSIGLENTEANRQ AYRTLLVSAPGLGQYVSGAILFEETLYQSTSDGK KMVDVLVEQNIMPQIKVDKGLVPLAGSNNSWC QGLDGLATRSAAVYQGGARFAKWRVTYVSPNGP SALAVKEAAWGLARYASISQDNGLVPIPEILLD GEHGIDRTFEVANKVWAEVYFYLTENNVMFEGIL LKPSMVTPGAECRDRASPOQVAGYTLRLLRQRI PPAVFGMFLSGQSEVEATLNLNAMNQANPW HVSFSYARALQNTCLKTWGGRPENVAQAQETLL LFAKSNLSLAQLGKYTEEGESEEAQKGMFVKGYT Y
174	sp_133220_yrmt	#N/A	0.944883	hydroxy methylglutaryl CoA reductase 1 (HMG1); CONTAINS InterPro DOMAINs: Hydroxymethylglutaryl-CoA reductase, class I, catalytic (InterPro:IPR004554), Hydroxymethylglutaryl-CoA reductase, class I/II, substrate-binding (InterPro:IPR009029), Hydroxymethylglutaryl-CoA reductase, class I/II, NAD/NADP-binding (InterPro:IPR009023), Hydroxymethylglutaryl-CoA reductase, class I/II, catalytic (InterPro:IPR002202); BEST Arabidopsis thaliana protein match is: 3-hydroxy-3-methylglutaryl-CoA reductase 2 (TAIR:AT2G17370.1); Has 2212 Blast hits to 2210 proteins in 935 species: Archae - 202; Bacteria - 1018; Metazoa - 225; Fungi - 225; Plants - 266; Viruses - 1; Other Eukaryotes - 275 (source: NCBI BLINK).	MDVRRRQPPPPSALKRKTMMAGDESPKASDALP LPLYITNGVFLTLFFAVVYFLLQWRWREKIRNSTPL HWVSLSEISAINAFASFIYLLGFFGIDFVQSFISPR SSYDDDEEVIDDATNEQQILKEDKRISGGIPIPLPE KKLEVLHQLPKTLDLTPDEEEVYQSVVSGSTPSYS LESKLGDCFRAAKIRREALQRMITGRSIVGLPLDG FDYDSILGCCCEMPYGFQIPYGIAGPLLLDGIIEFS VPMATTEGCLVASTNRGCKAIYLSGGAASVILKD GMTRAPVVRFGSARRAADLKFLEDPANFETISL VFNKSSRFARLQEVKCAIAGKNLYVRFTRCTGDA MGNMNVSKGVQNVLEFQSDFFDMVDVIGISGNY CSDKPAAVNWIEGRGKSVVCEAVIGEEVWVKVL KTDVASLVELNMLKNTLGSVAGSLGGFNAHAA NIVSAYIATGODPAQNVESQOCITMMEAVNGGR D.LHVSVMPSIEVGTGGTQLASQSAACLINLLG VKGANREAAAGGNARLLATVAVGSYLAGELSLMS AIAAGQLVRSHMKYNRSSRDMSKYVAVSS

10

20

30

40

【表 1 4 - 7 0】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたペイトとの相関係数 sp_107620_kqnh (SOAP1) 0.9369335	sp_107660_kiqg (SOAP2) #N/A	sp_107670_ptqx (SOAP2-like) 0.943135	注釈	aa seq
175	sp_133830_rrk	0.9369335	#N/A	0.943135	<p>ARAY; FUNCTIONS IN: GTP binding; INVOLVED IN: protein transport, small GTPase mediated signal transduction; LOCATED IN: endosome, plasma membrane; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Ras GTPase (InterPro:IPR001806), Small GTP-binding protein (InterPro:IPR005225), Small GTPase (InterPro:IPR020851), Ras (InterPro:IPR013753), Ras small GTPase, Rab type (InterPro:IPR003579), Rab5-related (InterPro:IPR015599); BEST Arabidopsis thaliana protein match is: RAB homolog 1 (TAIR:AT5G45130.1); Has 26416 Blast hits to 26384 proteins in 719 species: Archaea - 27; Bacteria - 133; Metazoa - 13862; Fungi - 3690; Plants - 2969; Viruses - 20; Other Eukaryotes - 5715 (source: NCBI BLINK).</p>	<p>MAGSAGKVVYKAKLVLLGDMGAGKTNLALRFVVG QFFDCQEPVGTATYFTQTLVYKEATINFDVWDTA GOERYHSMAPLYYRNATAALLYDISDKHSFVRA KKLVQELLTEETPIVYLSLVGNKADLETEREVATE EAETLSQIENGLFFMETSAKTSNINELFYIAKRV AEVCPQQPSDIDLQNEIQNRRFSCCSG</p>
176	sp_134000_joqi	0.973699	#N/A	0.945602	<p>EMPFINDLICHER IM DUNKELROTEN LICHT 1 (EID1); FUNCTIONS IN: ubiquitin-protein ligase activity; INVOLVED IN: photoperiodism, flowering, red, far-red light: phototransduction, leaf development; regulation of photomorphogenesis; LOCATED IN: nucleus; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; Has 121 Blast hits to 91 proteins in 15 species: Archaea - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 121; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).</p>	<p>MAETTFKLSDDVTLKIFSLLEDDPRNWARLSCV CRKFNISITCCKSKKSLYSPFISADLLPVSSLDG WASLYKIAVCGPGLQHAGVLFDSNHFGLDRELG PTEEYLHQTOQTRTETETNTLLSSPPSSKSPA TTVTEVGEAVVHDDDDNNNNINIEFGGWSLYD DLFYDTDNSETATPMEVTTGVADLDSKVVTT NPILIEEDDFESKFVWTKLENDEGVKRRKVNVS LORSHLASGVWNLSREQGNKLLASRRDDCLYI SDWPGCVHIEEKRSYMLFRGVFKNFKRTRVWRT INDGNRSKIVLNCAFCSCNEVWDLHSAFCLKRGF GFHDDGEPVVRAYVYCENGHVSGAWTDLPLYT</p>
177	sp_136880_away	0.982293	#N/A	0.926318	<p>Adenine nucleotide alpha hydrolases-like superfamily protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to stress; LOCATED IN: vacuole; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: UspA (InterPro:IPR006016), Rossmann-like alpha/beta/alpha sandwich fold (InterPro:IPR014729), Universal stress protein A (InterPro:IPR006015); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G49050.1); Has 2830 Blast hits to 2747 proteins in 619 species: Archaea - 112; Bacteria - 1833; Metazoa - 129; Fungi - 32; Plants - 667; Viruses - 0; Other Eukaryotes - 57 (source: NCBI BLINK).</p>	<p>MAMSSTEKAMVIVGIDESDEFHALEWTLDKFFFT PYAPNFFNVLLVYAKPTSTTITGLTGPAAQVL SFVEADLKRVAARIIIKAKDLCHSKSVHDVSVKVI GBARSVLCGAVEKKNINASMLVVGSHGVGALKRRTF LGSVSDYCAHNAHCSVMVIVKPKQTND</p>

10

20

30

40

【表 1 4 - 7 1】

A	B	C	D	E	F
1					
2	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MSNLGRKWSLLELFGKSRKNSGPIKPKKQVYTKLRR NSNNTVFSKQFYGKSLRSCNTNPKNFMATRFRFSY VARAVLAADPASQQADKYNDLGDIELQVSVSPM SSGCSKISKEIVTNSHDDNMVHLHWGIRDSKEK WALP-SRRPDGTRVHKDRALRTPFKSGSGGCSLT IEIDPAVQAIIEFLINEARNIWFKNNQNFHVRLLP RKSEHAQVPEDLVQVQAYLRWERMGKMYTP EQEKVEYAARAELMDEVARGVSIQELRERLKK GSSNASEDKGPPSKSESRIPIDDLVSIOYSLRWE KAGKPNYSDEQSREFEAEKELQRELDKGSV DDIRKIAKGEIKTRVTKVEKORNYSSERIQRKK RDLKQVISKHTAEPYDENVSTKVTDPKTLSAVELF SEAKQAQDGDLLRKKMFKLADKQLLVLYTKLDG KTKVCLATDFKEPLTHWALSAAKGNESQFPQS ILPPGSSLLDKAAETQFVSVSSDDSSQQFFLQSL EIEIEGDDFVGMFVYVSSDNNWMDKNGSDFYVDI RVKPKQKADGNGTAKSLEKIASLESEAKQSF MHRFNKADLLDEAVDAGELGLAGILVWMRFMA TRQLIWNKNYKPREISKAGDRLTDNLQNVFRS QPQHQELIRIMIMSTVGRGEGDVGGQRIIDEILVI QRNNDCKGGIMEEWHQKLNHTSPDDVVICQAL IDYIKSDFDISVYWKTLNENGISKERLLSYDRAHS EPHFRSEKQEGLLRDLGHYMRTLKAVHSGADLE AAISNMGYKTEGGQGMVGNIDPLPLGAPGPF DLLQFILVHVEDKNVEPLLEALLEARQELRLLFK SHSRLLKDLIFLDIALDSSVRTVIERGYEELNNAQ EKIMYFISMVVENLALSSDDNEDLVYCIKGWKRA LDMSSNDKDWALYAKSVLDRTRLALASKALELY QHILQPSADYLGSLLEVDQDQALSFTEEMIRAGSA ASVSALLNRLDPVLRRETANLGSWQIISPVYEAQVY VVVDELLSVQNKTYEQPTILVANHVGRGEEIIPD GTVAVLTDFDMPDVLSHVSVRARNRKGVCFATCFD
178	sp_138570_dixk	#N/A	0.37341	STARCH EXCESS 1 (SEX1); FUNCTIONS IN: protein binding, alpha-glucan, water dikinase activity; INVOLVED IN: cold acclimation, response to symbiotic fungus, response to trehalose stimulus, circadian rhythm, starch catabolic process; LOCATED IN: mitochondrion, chloroplast stroma, chloroplast, chloroplast envelope; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAINs: Pyruvate phosphate dikinase, PEP/pyruvate-binding (InterPro:IPR002192); BEST Arabidopsis thaliana protein match is: phosphoglucan, water dikinase (TAIR:AT4G24450.1); Has 2623 Blast hits to 2596 proteins in 1116 species: Archaea - 196; Bacteria - 1989; Metazoa - 24; Fungi - 8; Plants - 221; Viruses - 0; Other Eukaryotes - 185 (source: NCBI BLLink).	

10

20

30

40

【表 1 4 - 7 2】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
179	sp_141040_qmwp	0.956432	#N/A	0.934272	<p>Integrase-type DNA-binding superfamily protein; FUNCTIONS IN: DNA binding, sequence-specific DNA binding transcription factor activity; INVOLVED IN: regulation of transcription, DNA-dependent; LOCATED IN: nucleus, chloroplast; EXPRESSED IN: 10 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: DNA-binding, integrase-type (InterPro:IPR016177), Pathogenesis-related transcriptional factor/ERF, DNA-binding (InterPro:IPR001471); BEST Arabidopsis thaliana protein match is: Integrase-type DNA-binding superfamily protein (TAIR:AT5G25810.1); Has 5404 Blast hits to 5361 proteins in 232 species: Archaea - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 5399; Viruses - 0; Other Eukaryotes - 5 (source: NCBI BLINK).</p>	MEHLPQCSSEFEDRAISPGSPSSGGPQSSKSSQGS GTRHPVYHGVRKRRWGWVSEIREPRKTRIW LGSFPMEMAARAYDVAHCLGKCKAQLNFP IDFLPRPTSCAPKDIQAAAANAANAIVEGKVS AAADHFDHDDVEYDYYDDFWSEIELPELMDT GQGGNFMDCYTSSSLGEFYF
180	sp_151150_kxyd	0.932439	#N/A	0.963753	<p>thioredoxin 2 (TRX2); FUNCTIONS IN: oxidoreductase activity, acting on sulfur group of donors, disulfide as acceptor; INVOLVED IN: N-terminal protein myristoylation, sulfate assimilation; LOCATED IN: cytosol, plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Thioredoxin fold (InterPro:IPR012335), Thioredoxin, core (InterPro:IPR015467), Thioredoxin domain (InterPro:IPR013766), Thioredoxin, conserved site (InterPro:IPR017937), Thioredoxin-like subdomain (InterPro:IPR006662), Thioredoxin-like (InterPro:IPR017936), Thioredoxin-like fold (InterPro:IPR012336); BEST Arabidopsis thaliana protein match is: thioredoxin H-type 7 (TAIR:AT1G59730.1); Has 19714 Blast hits to 19583 proteins in 2950 species: Archaea - 262; Bacteria - 11021; Metazoa - 1828; Fungi - 880; Plants - 1790; Viruses - 14; Other Eukaryotes - 3919 (source: NCBI BLINK).</p>	MGARLSTIPENRGLYSATAGPGSSYPSGYNKPA ASDRYVNSFPAENKAPAAAPVCAAPFPFPAN SNP'SNNYPAATTTKSSVAAGCGGDVAGGGRG GAFVNVKPAANTKEVIVFHSSTKWKEYFEASKQS NKLWVYFTATWCGPCRYMEPSIKEIAAKHTDIDL VKIDVDELFNVSREYGVQAMPTFLFMKKGKQIDK VVGARKEELQRKVEKHA

10

20

30

40

【 1 0 7 2】

【表 1 4 - 7 3】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	共発現遺伝子			O-methyltransferase 1 (OMT1); FUNCTIONS IN: myricetin 3-O-methyltransferase activity, quercetin 3-O-methyltransferase activity, caffeate O-methyltransferase activity; INVOLVED IN: lignin biosynthetic process, flavonol biosynthetic process; LOCATED IN: cytosol, nucleus, plasma membrane, cytoplasm; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Winged helix-turn-helix transcription repressor DNA-binding (InterPro:IPR011991), Plant methyltransferase dimerisation (InterPro:IPR012967), O-methyltransferase, family 2 (InterPro:IPR001077), O-methyltransferase, COMT, eukaryota (InterPro:IPR016461); BEST Arabidopsis thaliana protein match is: O-methyltransferase family protein (TAIR:AT1G77520.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archaea - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MGNVTENAELQGLSEVYKHFVFAFTESLVLRSV ELRIADIHSNGQPMISLSQIASKIDSSPSINITNLER IMRMLYRKIKIFSSSAEPTALYGLTSTSKWLHDQE LSLAPVLLTFTHTMTTAWYQISPSIKENGTAFEM AHGESFWDMASSNPEFNKLFNDGMSAGTKPVL DSJIEGYKDGFSKLEGLVDVGGALGVYVSKIVEV YPHIKGINFDKPNAFANAPQYPGVIHVGGDMFKE VPSADNVFKISVLIDWPDEECITILKIKCRKVAEK KGGKVIIEIILHPDRYEMFDDVAISLDLMMTNFNG GKARTENEWMELLLKVKVGTDFINIIPLPSIMSIVCV
181	sp_151320_gtwi	#N/A	0.938231	cysteine-rich RLK (RECEPTOR-like protein kinase) 26 (CRK26); FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: endomembrane system; EXPRESSED IN: stem, root; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Serine/threonine-protein kinase domain (InterPro:IPF002290), Protein of unknown function DUF26 (InterPro:IPR002902), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271), Protein kinase, catalytic domain (InterPro:IPF000719), Tyrosine-protein kinase, catalytic domain (InterPro:IPF020635); BEST Arabidopsis thaliana protein match is: cysteine-rich RLK (RECEPTOR-like protein kinase) 29 (TAIR:AT4G21410.1); Has 124218 Blast hits to 122701 proteins in 4579 species: Archaea - 106; Bacteria - 13898; Metazoa - 45472; Fungi - 10958; Plants - 35042; Viruses - 435; Other Eukaryotes - 18307 (source: NCBI BLINK).	MGKTLFIFASLMMVSLWQLRIVATGINFSDAACS SVANYTINSVYQKLNLLFSDLSSKSTQKFKYNST IGDDPKVYGLYQCREDVSLICRCQIKKATQKIV QVCPKKEAIVWYLECMRLRYKSSIFSLDMSVH DYKHSSGINSYVQSPVLAEMDKVIDQAAALST SSGHFATRARRWSSVYVYCLAQCTFDLSGRDC NNCLHIALRIMFIANFNNSVAIGIFFPSCQLRYGRV QFYDDLVPKPPAPTSSSSPITMTNP
182	sp_151640_xhpu	#N/A	0.952219		

10

20

30

40

【表 1 4 - 7 4】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
183	sp_163870_noif 0.928419	#N/A	0.942328	IQ-domain 33 (Iq33); CONTAINS InterPro DOMAIN/s: IQ calmodulin-binding region (InterPro:IPR000048); BEST Arabidopsis thaliana protein match is: IQ-domain 3 (TAIR:AT3G52290.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archaea - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MGITGELVRSVFSKSPRSIKIKTDEINMKFNGVEKK RWKTSVKYLCGDELFDSDVAEEDSASVRSRSS EATVMQMLQDDGYSTODIESTKYDNPDRHNSF SSNFDEDTAA TIQSAFRGFKVRSRLRETTDFON DFEEGKSPGRESMATSIIEVQTGNSIGVSIQEE DTSASTKLHHLKSSAYKPKEDWINDSTLDSTKS KMRIONFLAEMTRRERALAYAFSQQLRICKSKKH TKSDPDMGWSWLERWMTAQIEVPLVEEEDY TKQLTENEAVSRTRKARPLRKKFLDLSFEKKE GSNEIPVQVD TSTNSNGEDKHGYDSRPNRLKNP TSSSRKTIVPTHYQKESKYSKKNRSKETEENIT ROKARQSESSSICNDGASWSPPDPDHSVSGHN SKSDP
184	sp_171000_pxii 0.928416	#N/A	0.939333	cytochrome P450, family 96, subfamily A, polypeptide 10 (CYP96A10); FUNCTIONS IN: electron carrier activity, monooxygenase activity, iron ion binding, heme binding; INVOLVED IN: oxication reduction; LOCATED IN: endomembrane system; EXPRESSED IN: 6 plant structures; EXPRESSED DURING: 4 anthesis, C globular stage; CONTAINS InterPro DOMAIN/s: Cytochrome P450 (InterPro:IPR001128), Cytochrome P450, E-class, group I (InterPro:IPR002401), Cytochrome P450, conserved site (InterPro:IPR017972); BEST Arabidopsis thaliana protein match is: cytochrome P450, family 96, subfamily A, polypeptide 9 (TAIR:AT4G39480.1); Has 28273 Blast hits to 28182 proteins in 1460 species: Archaea - 44; Bacteria - 2431; Metazoa - 10638; Fungi - 6078; Plants - 7933; Viruses - 3; Other Eukaryotes - 1145 (source: NCBI BLINK).	MAGGDTSVAFTWFFYLLVKYPQVTAKIREELDA ILITQSNYKDDDFVKNFSDNNKKLVYLHAALES RMYPPLFNHKTQSVKSDVLPESHKVGPNTEIYFD MYAMGRMKSLSWGDCCNEFKPERWITKQGNFKL EFSNKLAFAGAGPRICQGRSMTFIQMKVIAAIIPK YDIVAVEGHQVQVDPVSVLQMKHGFVKVIVRPPSI PLHDS
185	sp_171320_gmgq 0.856094	#N/A	0.838674	KDSB; FUNCTIONS IN: 3-deoxy-manno-octulosonate cytidyltransferase activity, nucleotidyltransferase activity; INVOLVED IN: lipopolysaccharide biosynthetic process; LOCATED IN: mitochondrion; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: 3-deoxy-D-manno-octulosonate cytidyltransferase (InterPro:IPR004528), Acylneuraminate cytidyltransferase (InterPro:IPR003329); Has 7496 Blast hits to 7495 proteins in 1549 species: Archaea - 38; Bacteria - 4007; Metazoa - 8; Fungi - 0; Plants - 48; Viruses - 0; Other Eukaryotes - 3395 (source: NCBI BLINK).	MNPTSICTSSSSSSSSSTKSWSVHGLVLGAAA AAAIGAHTYFYLRRSRKFRSRVIGIIPAFASSRF QGGKPLHILGKPMIORTYEOAKLASMLDQVVVAT DQNKIACCRGFGADVIMTSESCRNGAERCSEA LQKLNKYYDIWVNIQGGDEPLIEPDIIDGIVKALQSA PDAVFSTAVTPLKAGDAFDTRKVKVVDNLGYAI YFSRGLIPIYKSGMGNPDPFYLLHLGICSDAEFL KIYPELSTPLQLEEDLEQLKYLENGYKLVKIVE HEAHGVDVPEDEVKEIQVMRERNLS

10

20

30

40

【表 1 4 - 7 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
186	sp_174510_ypax	#N/A	0.933773	LACHESIS (LIS); CONTAINS InterPro DOMAINs: WD40 repeat 2 (InterPro:IPR019782), WD40 repeat, conserved site (InterPro:IPR019775), Pre-mRNA processing factor 4 (PRP4) like (InterPro:IPR014906), WD40 repeat (InterPro:IPR001680), Splicing factor motif (InterPro:IPR003648), G-protein beta WD-40 repeat, region (InterPro:IPR020472), WD40 repeat-like-containing domain (InterPro:IPR011046), WD40-repeat-containing domain (InterPro:IPR017986), WD40/YVTN repeat-like-containing domain (InterPro:IPR015943), WD40 repeat, subgroup (InterPro:IPR019781); BEST Arabidopsis thaliana protein match is: Transducin/WD40 repeat-like superfamily protein (TAIR:AT2G05720.1); Has 96331 Blast hits to 36444 proteins in 930 species: Archaea - 78; Bacteria - 11869; Metazoa - 38307; Fungi - 21059; Plants - 12275; Viruses - 15; Other Eukaryotes - 12728 (source: NCBI Blink).	LLHMIQLRYETWHSIALSSLFCXILALEGHIKPVY GVDFSPNGYHLATGSEDNTRWRDLRKKSLYII PAHSLKLSQVKFEPQEGYLLVTASDYDTTKWVSS RDFKLIRLTLSGHEARVTSLDVTGGEIQGCRWFVC GLFVVCVWLEGGFVVVSSGLCLYGFDFEWFVRDG FMVVGAVYGVGRVVS
187	sp_177240_yfer	#N/A	0.934187	Mannose-binding lectin superfamily protein; CONTAINS InterPro DOMAINs: Mannose-binding lectin (InterPro:IPR001229); BEST Arabidopsis thaliana protein match is: Mannose-binding lectin superfamily protein (TAIR:AT1G73040.1); Has 3242 Blast hits to 933 proteins in 57 species: Archaea - 0; Bacteria - 14; Metazoa - 135; Fungi - 2; Plants - 3082; Viruses - 0; Other Eukaryotes - 9 (source: NCBI Blink).	MANQLTNWSELEGEKTKVMIRHGFVVDGIGF ETVKNSNGIPLSSGFEFITYISGSYGMYYRETCT TMKIHNTLRPDYGYGPAQGAEGVTDFFSLPL NSSIVGFFGSGVYLASIGINAEERTMTIPYGPYGN SESSPNWSIEINEGORFSKVRISHGYVDGIGFDI TDQSGKTTPTQLFGSGGSPSEIDLGVNETITOI SGSYGKYFGKDVATIMIKIHNTKRPDGGYGPYGRA QEATDVKPFSTPLPLKGVVFSFQGHYGGVYLSIG ISDKYKTYVEKYGPYGRA
188	sp_178750_skdf	#N/A	0.933768	RNA-dependent RNA polymerase family protein; CONTAINS InterPro DOMAINs: RNA-dependent RNA polymerase, eukaryotic-type (InterPro:IPR007855); BEST Arabidopsis thaliana protein match is: RNA-dependent RNA polymerase family protein (TAIR:AT2G19920.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archaea - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI Blink).	DQQLSGPVLVYRNPFAHFGDHHVWNRVYEAL KDFVGNAKYGVFFPSKGLRSMADFIAGGDYDGD MYWCRNPELLKFFKPFSEPWKSNQPNATKQIS LPSADDVLEMKLFRLYRGVRCOPFNAMGIAEAC WVYMDRLLTLGDCCADEKATLRRKIDQIDYH DALDAPKKGAAQVEVPDDLKLRYPHYLKGGEHR TYRKSILGEIHDVDQLSAPIKFWKLECFEADIPG TYKVLWKNHYDKYRPFEMQMALNESRNADFFVYQ KYRKELYGGVDDLLDSEKAKWKDIRLEALTIYHVC YDYAKAAEDPKKCGFAWKVAGDALWIKYI-LQSTG EKPILFAPSVVRQMMR

10

20

30

40

【表 1 4 - 7 6】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
189	sp_179790_koyu	#N/A	0.923228	<p>wall associated kinase-like 2 (WAKL2); FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: integral to membrane; EXPRESSED IN: stem, cotyledon, leaf whorl, leaf; EXPRESSED DURING: LP.06 six leaves visible, LP.04 four leaves visible, LP.10 ten leaves visible, LP.08 eight leaves visible, LP.12 twelve leaves visible; CONTAINS InterPro DOMAIN/s: Wall-associated kinase (InterPro:IPR013695), EGF-like calcium-binding, conserved site (InterPro:IPR018097), Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: wall associated kinase-like 4 (TAIR:AT1G16150.1); Has 122569 Blast hits to 120935 proteins in 4599 species: Archae - 125; Bacteria - 14250; Metazoa - 44921; Fungi - 10430; Plants - 34238; Viruses - 502; Other Eukaryotes - 18093 (source: NCBI BLINK).</p>	<p>MHLSWLLLLLFSVSAATLAGTDALITPKNMITAAK TSKPGCQQKCGNLTIYPFGIGKQCYFSGSFGIC NTSYNPPRTFPVIAWGWDFIDITETQMIVIRNSGV ASSCHDSQKLNSSSSTGLDDMWGSPFTVSSSTAN KLTVIGCNDYAVFGGSDDAKYORAGCIAMCN NRAEPVAGSCPGLGCCQTSPLGWNWQGSFISIT LFSVEDGQLTTIGKSLSDQSCSYAFYAKSGSFIFG GASDLVTNNTDDVNRIMEQVPMVLDWNVQLKI SCKQVKNSTSNPYLCOQSNTKCVDTAEGGYRCS CLPGYEGNPLYQPGGTDINECVSGPNPCVMTG VNTPGSYKSCARGYVGDGLKSGTGTGCFYNKRA ANLIRITLGLSSLGVIIVLLVAVWFSTIARRKIM KQKVYFKONGGLLQEQMATDSDSVAAIKMFTI DELEKATDHFNDRIILRGAGALTYLHSSSIPFI HRDIKSSNILLDDTYRAKLSDFGTSKTVGIDQTHV TTPQVMGTFGYMDPEYHLLGQYTEKSDVYVSGV TVLLELLTGQAVRSALAEKESLASWFLSHMKNP RLLDIDSRVLOEDSKEEFLLIAEIAEKCLNQDRNS RPTMKEVLEIEIAYLSHLPOANEQNWPKSEQIF SGSISDGWSSSTAFLYENSSLSAEVSLFNRSR IMATSKAFVLVILFAATLLIISDDNVSARELYHPNST PQTHASTAGYDGGSSPHHSYDDDDPTQEPIFPPT HKKKKTHG</p>
190	sp_193630_uhio	#N/A	0.9234147	<p>uncharacterized protein</p> <p>polygalacturonase inhibiting protein 2 (PGIP2); FUNCTIONS IN: polygalacturonase inhibitor activity; INVOLVED IN: response to salt stress, signal transduction, defense response; LOCATED IN: cell wall, plant-type cell wall; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Leucine-rich repeat-containing N-terminal domain, type 2 (InterPro:IPR013210), Leucine-rich repeat (InterPro:IPR001611); BEST Arabidopsis thaliana protein match is: polygalacturonase inhibiting protein 1 (TAIR:AT5G06860.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).</p>	<p>MAQAHLIIVSLLLVLALFVLDLSYSPITDLCHPHDK KVLRLKHNHLGNPSSLSWDPNTDCAGWTGIQCN SQGHVIRFAINQALDIHGPIPPFLDRPALTSLYFD GIPNLSGYPIYIGKLTLDLDFSIGTNTVGTGPIPEFL GQLTNLNLLYSNKFTGPIPNFLGLLKSLYLDIS SNLLTGPIVALAQ.TNIEILSLSNNKLSGNIPIDFIA QKLSIKILELNYSLSGPIPSLSGLFLPNISSIGLAG NQLSGPIPKSLGRGNLNFIRLSQNKLTGDAASFLFE KANKVMILEVANLLKDFTKVDLAPLSLTMNIG HNMIYGLPKRFLSFLSTDSVDVSYNQLCGPIPNIG RRYKRVDPPIFAHNKCLCGGPFPGCK</p>
191	sp_197810_urfr	#N/A	0.923883		

10

20

30

40

【表 1 4 - 7 7】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
192	sp_198340_focw	#N/A	0.943252	HXXXD-type acyl-l transferase family protein; FUNCTIONS IN: transferase activity, transferring acyl groups other than amino-acyl groups, transferase activity; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAINs: Transferase (InterPro:IPR003480); BEST Arabidopsis thaliana protein match is: HXXXD-type acyl-l transferase family protein (TAIR:AT1G03495.1); Has 1873 Blast hits to 1858 proteins in 109 species: Archaea - 0; Bacteria - 0; Metazoa - 0; Fungi - 18; Plants - 1853; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLINK).	MAPPSSSTTGGNGSSFAVNIMASFYISQQPS TTNSHSIPLTFDFDIPWILQYPLQLPFFFLPSTPQ SSSSSSSSDFDHNLYLEFSSTILPRLKHSLSAL QYVFPFSGKLTHTTTPNNLVFSTSDSDSVELTV SLCDADFNGLCSFLPRSTHLFQQLVPSLPNISS NLTTFPAPLLAIQITFFFTSSPGFSIGFASHPVLS QRTFSNELYSWASFSEKFDNINISLAP'SFPVSDRS VILDPDRLEPLLLEQWLGLLESKPTMSTKMKLRPP PAYVRGSLRSTFVMGSPDIANATQWLQTOCEKL NRSYPVLLSPYVVTCAFIVTFCFLRVRQNSAVTK AKAKGTMVYFGIAGGITRPFYRYPVPAKYLGNVGF GRAAQRRELLKEGEGMLAAADAIGLTIKKDKD VLGGAEKWIYEWQTLMESEDIHVGVGSPKVNLY ETDFWVWGPKKIEIESTDVTRAISSLTQSRDMIKRG IEIGLTLPNSIMDDFESSIFTQGLLVFQN
193	sp_199120_jzrw	#N/A	0.921698	Protein kinase superfamily protein; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; EXPRESSED IN: petal, flower; EXPRESSED DURING: 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAINs: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: Protein kinase superfamily protein (TAIR:AT2G18890.1); Has 113420 Blast hits to 112131 proteins in 3664 species: Archaea - 114; Bacteria - 13552; Metazoa - 41580; Fungi - 9353; Plants - 32095; Viruses - 456; Other Eukaryotes - 16270 (source: NCBI BLINK).	MIPSTAPARIIVATSLEVDVKKLLSWAINILSHPN DTIALHVLVGKEPKKLVKPSREYKFRRAKSFVL SAMGEFAKTCQCKQ

10

20

30

40

【表 1 4 - 7 8】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
194 sp_199710_jns	0.948838	#N/A	0.952514	purine permease 1 (PUF1); CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF6, transmembrane (InterPro:IPR000620), Protein of unknown function DUF250 (InterPro:IPR004853); BEST Arabidopsis thaliana protein match is: purine permease 3 (TAIR:AT1G28220.1); Has 617 Blast hits to 607 proteins in 89 species: Archae - 4; Bacteria - 97; Metazoa - 10; Fungi - 14; Plants - 377; Viruses - 0; Other Eukaryotes - 115 (source: NCBI/Link). zinc ion binding;nucleic acid binding;zinc ion binding; FUNCTIONS IN: zinc ion binding, nucleic acid binding; INVOLVED IN: biological process unknown; LOCATED IN: intracellular; CONTAINS InterPro DOMAIN/s: Zinc finger, C2H2-like (InterPro:IPR015880), Zinc finger, C2H2-type (InterPro:IPR007087); Has 287 Blast hits to 281 proteins in 115 species: Archae - 0; Bacteria - 0; Metazoa - 116; Fungi - 82; Plants - 70; Viruses - 0; Other Eukaryotes - 19 (source: NCBI/Link).	MYAAHSFITAFYHRDGTVNRRLVNLADIRSAAS STPIGGVTSFDHTIGIDPLNKLWFRIFFPANSPSS HSLPVIYFHGGGFVSPATYAYDTFCRKVSGD VGAVISVYRLAPEHKWPAQYDDGDFDLKFLD DDEMRKIECFANADLGRCLAGDSAGANLAH HVAVRGGARFKTVRVRGLISIQIPFYGGEERTDS ELRTKQWPVLTLSQDFFWKAFLPNGSDRDHPG SSVFGPKSKDVSGLSEFPPTLVVGGDLDRDW GMRYDGLKRSRGGKVELVQYSNVGHFFYSAPEL PEYKLVSKIGEFVENQLRLOGIGISGGPLVMRLY YIHGGKSIWSSWLETGGWPLMIPLLVAYFRRL RRRRGGADPTQAAKLIPIINTNTFFFAAIVGLVAGA ADYCYAYGVKHIPVSTSSLLATQLAFTAVFAFVLV NQKFTYPSINSIALLTFAAILAVHNSDKPAGESN GOYLLGFLLTASSALFALMLPMIELTYKRVKQSV VVINYSLVMEIQLVLSFSATVMCTVGMMLASGDYK EYAREMQDYKHGKAMYFVVLCSAIVWQFFYLGA AGVIFYGSSLLSGVIISVALPITEIAAVFFYHETFOV DKGVSLLSLLSLWGFISYFYGEAKEIKKQKRRQRN NTQSIDVES
195 sp_200090_ziji	0.840358	#N/A	0.827312		MAMQVEEGKSGSEFFYWKPVHRKFAFDSPFFS SGDVERQLLDKQVALVLTEDMQQLDMEDHER EIFCPIVCGGARLNSLEGFDHYHSCHTASCVC HKSYPTORLLSHLSEVHDSYFKAKVSKGFSMYE CLVETCGLKFKSYKARHQHLVDKHKFPITYEYFK KAHLKKQRQKYGRKPKALCKEETSNAMEPLDNEA DDSMEDLTSAISQLSTSDSYVPSISFGHRRSRGL AFVPRSVQRERRRFGYSPSTMN

10

20

30

40

【表 1 4 - 7 9】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
196	sp_213110_prhn	0.935336	#N/A	0.968252	vesicle-associated membrane protein 727 (VAMP727); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: vacuole organization, protein targeting to vacuole; LOCATED IN: endosome, SNARE complex, membrane; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Longin (InterPro:IPR010908), Longin-like (InterPro:IPR011012), Synaptobrevin (InterPro:IPR001388); BEST Arabidopsis thaliana protein match is: vesicle-associated membrane protein 726 (TAIR:AT1G04760.1); Has 2397 Blast hits to 2395 proteins in 262 species: Archae - 0; Bacteria - 0; Metazoa - 970; Fungi - 430; Plants - 618; Viruses - 0; Other Eukaryotes - 379 (source: NCBI BLINK).	ILNMSFRLHAMTDSNPQLQEMMQNSVLCQLVFP EMMQVLDCCRIRPKIFSKYAFNCYDLNLVYVKFF WDTSCYFVLYYKQADSFORQGRQLRRKMWFGN IQLKLTVGGSIILFVIVVWLENPQLSDL
197	sp_005860_fgmk	#N/A	#N/A	0.938522	integrase-type DNA-binding superfamily protein; CONTAINS InterPro DOMAIN/s: DNA-binding, integrase-type (InterPro:IPR016177), Pathogenesis-related transcriptional factor/ERF, DNA-binding (InterPro:IPR001471); BEST Arabidopsis thaliana protein match is: Integrase-type DNA-binding superfamily protein (TAIR:AT1G15360.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MARPOORFRGVRORHWGWSVSEIRHPIKTRI WLGTFETAEDAARAYDEAARLMCGPRARTNFPY NPNAPQTSSSKLLSATLTKLHRCYMASLQVTKN SAMQAGQORGPQFSGTGNVVGPKFGETCAW SPEYGSTDGLVLPNPEPNWTKMVENAGLEPOC HFTMLEDDHIEQMIEELLDYGSIEIAI

10

20

30

40

【表 1 4 - 8 0】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq  MGTDKANDVTKNGIASKIPAEAHPLSEESTEIASN INYHAQFSPHFSPKQPEQAIFYATAESVDRLI QQWNETYVHFHEANPKQTYLQSMEYLQGRALT NAVGNLNIQDAYADALSKLGHQLEIVEOEKDA LNGGLRLASCFLDSMATLNLPAWGYGLRYRH GLFKLISKEGGSEIPEWLEKFSWVWV FPVRFQVQVNPDGSRKWVDGEIQALAYDVI PGYKNTISLRLWEAKASAAADFNLFQNDGQYE SSLLHSHRAQIQICAVLYPGDATEGGKLLRLKQOY FLCSASLQDIIFRRKFKRDKGKTPLKWSEFFSKYA VQLNDHTPTLAPELMLRLMDDEGLSWDEAWGV TTRAIAYTNHTVPEALEKWSSVMWKLPRHM EITEIDKRFVELIHSTRPDLKEDKYNMCLDNNPQ KAVVRMANLCVVSHTYNGVAQLHSDLTSLDFA DYVSIWPTKFNKNGTNGITPRRWLFCNPELSDIIT KWLLKTDWVYTNLRLVLRQFADNEDLQSEWA SAKMASKRRLAQYILEITGVSDPNLSLFDIQKRIH EYKQRLNILGAIYMYKLLKEMSPEDRSKTTARTI MIGGKAFATYTNKRIVKLVNDVGAVNTDPDVM NYLKVVPYNNVSVABKLPISGSELSQHSIATGM EASGTSNMKFAKNGCLIGLTDGAVNVEIREEIGEE NFFLFGATDEVPKLRKDRENGLFRPDPFRFEAK DFIRSGAFSGYDYNPLDLSLEGNSSYGRGDYFLV GHDFFSYMDAQAKVDEAYKNQKWLKMSILSTA GSGKFSDDRTIAQYAKEIWIIEECPVPSNIKE
198	sp_008600_oszt	#N/A	#N/A	alpha-glucan phosphorylase 2 (PHS2); FUNCTIONS IN: phosphorylase activity, transferase activity, transferring glycosyl groups; INVOLVED IN: response to cadmium ion, response to water deprivation; LOCATED IN: cytosol, chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAINs: Glycosyl transferase, family 35 (InterPro:IPR000811), Glycogen/starch/alpha-glucan phosphorylase (InterPro:IPR011833); BEST Arabidopsis thaliana protein match is: Glycosyl transferase, family 35 (TAIR:AT3G29320.1); Has 5792 Blast hits to 5741 proteins in 1741 species: Archae - 74; Bacteria - 3953; Metazoa - 555; Fungi - 138; Plants - 232; Viruses - 2; Other Eukaryotes - 838 (source: NCBI BLINK).	Phosphatidate cytidyltransferase family protein; FUNCTIONS IN: phosphatidate cytidyltransferase activity, transferase activity, transferring phosphorus- containing groups; INVOLVED IN: phospholipid biosynthetic process; LOCATED IN: membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Phosphatidate cytidyltransferase (InterPro:IPR000374); BEST Arabidopsis thaliana protein match is: vitamin E pathway gene 5 (TAIR:AT5G04490.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).
199	sp_016880_prex	#N/A	#N/A		MLTPRRHESYNAFSSVKHFSHLSLSPSELTSPSIN FSGELFRFRANTSHSLRKRKTVAANIMLPDNP VISDICATVITAAIARGVYLGWELTAKYDVDFSKLN RKLVLHISLGLVFMCLWPLFSSGYRGAFACLVPFG LNAIRVLLIGLIGYKDEATVKSMSRFRGDYRELLKG PFYASAIISFACLVFWRTSPIGIAAICNLGADGM ADVGRRIKDKLFPYKNKS YAGSIAMATAGFYA SYGYMHYLSAFGFMQENWGLVFRLLWLSVSSV VESLPISTQLDDNLTVPASFSTGFLFF

10

20

30

40

【表 1 4 - 8 1】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたペイトとの相関係数 sp_107660_kriq (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈  ammonium transporter 1:1 (AMT1:1); FUNCTIONS IN: ammonium transmembrane transporter activity; INVOLVED IN: ammonium transport, transport, protein polymerization, response to karrikin; LOCATED IN: nucleus, plasma membrane, membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Ammonium transporter (InterPro:IPR001905), Ammonium transporter, conserved site (InterPro:IPR018047); BEST Arabidopsis thaliana protein match is: ammonium transporter 1;3 (TAIR:AT3G24300.1); Has 11619 Blast hits to 11599 proteins in 2072 species: Archae - 224; Bacteria - 4688; Metazoa - 462; Fungi - 431; Plants - 505; Viruses - 0; Other Eukaryotes - 5309 (source: NCBI BLINK).	aa seq
200	sp_018590_jidd #N/A	#N/A	0.642533		MAPSWEESVASSINTYLLFSAYLVFVQMQLGFAM LCAGSVRAKNALNIMLTNVVDVAVVGSLSYLYLGF AFAFGTGSEFPNFIGTQFFALKDVPDQKDYDSFF LYQWFAIAVAGITSGSIAERTQFSAYLIFSFLSG FVYVVAHWVWSSHGWLSAGSATLFLSGAIDF AGSVVHLVGGIAGLWGSIEGPRVGRDFAFGKP VPIKGHNATLVVLTGTLWFGVWFGFNGSFGSEIL VPTYTTDQGNWTAIGRTAVTTTLAGSMSGLVTL FGRRLIGHWDAMDVNGVGLGGFVAITSGGSVW EPWAAICGFVAAWYVLIQNLMLALHLHFDLPPLEAA QLHGGCGAWGLIFTGLFAKEEFIIQAYDSGKQDA SRSHGLLLGGWGLFGSQVLELIVVGVWYSVTM GPLFYVLRHKJLLRISIDDEVAGLIDISSHGGYAYH NQDEDERPRRYKYDYLQMHN
201	sp_026640_raed #N/A	#N/A	0.642533	LOCATED IN: plasma membrane; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINs: Protein of unknown function DUF3754 (InterPro:IPR022227); BEST Arabidopsis thaliana protein match is: aminopeptidases (TAIR:AT5G13940.1); Has 308 Blast hits to 303 proteins in 113 species: Archae - 0; Bacteria - 164; Metazoa - 10; Fungi - 0; Plants - 111; Viruses - 0; Other Eukaryotes - 23 (source: NCBI BLINK).	MGKKKKEFIRLDRESVIPISKRLIHLTLQLFESM SEREEFRFCRVEYTIRAWYLVKFEEMQLFNIL FEPVAGAHLKERNLSEDEIDELEQKFLVLLFQM MKNKSNFKIVTDQEFEEVAASGSYCLNLIQIKVDEDK LDSTLLRSYFSKHPQDHLPHYFADQYMIFFRGGGL DOMTGYFIQWKVDALISRTVQGFLLRLLGLKKIFS RKGRLWYVVRNQLLEGIVTECEDEYDDNDGDE HSLLIERIRVQNMDSIGNLFSKTTIQTPEITFEKIIVY RLHSTKDRAYVVKHKHPMADMEIVLPEKKSPPG LTPSDWVKLLVSAAGLTVITQLCQTKANIKVLA TVCSAVAYCAKIYFTFOANLKYQSLVQGSMYD KQLDSGRGTLHLCDVDVQIQEVEKVEVAVFFVLIKY GKLSRODLQKGEELLSDQNEKCNKFNFDIDDAHK LEKLSLVYKDEMEQYSSVDVGVKANEFRITTEIV TIACQGDQ

10

20

30

40

【表 1 4 - 8 2】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	myo-inositol monophosphatase like 2 (IMPL2); FUNCTIONS IN: histidinol-phosphatase activity, 3'(2'),5-bisphosphate nucleotidase activity, L- galactose-1-phosphate phosphatase activity, inositol or phosphatidylinositol phosphatase activity, inositol- 1(or 4)-monophosphatase activity; INVOLVED IN: sulfur metabolic process, histidine biosynthetic process; LOCATED IN: chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Histidinol- phosphate phosphatase, putative, inositol monophosphatase (InterPro:IPR011809), Inositol monophosphatase (InterPro:IPR000760), Inositol monophosphatase, metal-binding site (InterPro:IPR020593); BEST Arabidopsis thaliana protein match is: Inositol monophosphatase family protein (TAIR:AT3G02870.1); Has 15134 Blast hits to 15130 proteins in 2267 species: Archae - 157; Bacteria - 8374; Metazoa - 557; Fungi - 304; Plants - 238; Viruses - 0; Other Eukaryotes - 5504 (source: NCBI BLINK).	aa seq
202	sp_027050_wigk #N/A	#N/A	{333}471	GDSL-motif lipase 5 (GLIP5); FUNCTIONS IN: lipase activity, carboxylesterase activity; INVOLVED IN: lipid metabolic process; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Lipase, GDSL (InterPro:IPR001087); BEST Arabidopsis thaliana protein match is: GDSL lipase 1 (TAIR:AT5G40990.1); Has 3342 Blast hits to 3298 proteins in 168 species: Archae - 0; Bacteria - 231; Metazoa - 0; Fungi - 23; Plants - 3083; Viruses - 0; Other Eukaryotes - 5 (source: NCBI BLINK).	MLESLSLIYELLISFSPNINCLISSNHKHHHPPTS LFFGDSILDVGNINININTSLDQANFPWPGYVTF HHPTGRFSDGRVISDFIAERANLPSIPPYLQPGK WRFYDGVNFAFASAGGALVGTFRGYVIDLHMQVK HYKTVEKW_LKNKLGNVKAKTKLERSVYLFSIGTN DYMSLFLTNSSTHLTNHNKQVYVSMVINITSVYN EYKTGGRKFGFLNLPVGCFFPALRMLNPGQND ECQTEVSSYIKLHNKKAISIALKDLAKDLPFKFALF DYVTSTLQRVNHPSKFGYKEGKTACCGTGRFRG VSSCGRRRPVKEYELCKNINDHIFWDSLSHYTEKA SKQITDEMWVSNINYIHGYSYTIKDLFLHLP
203	sp_031320_wiin #N/A	#N/A	{333}312		

10

20

30

40

【表 1 4 - 8 3】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
204	#N/A	#N/A	0_323362	tubulin alpha-5 (TUA5); FUNCTIONS IN: structural constituent of cytoskeleton; INVOLVED IN: response to cadmium ion, microtubule-based process; LOCATED IN: tubulin complex, cytosol, cell wall, membrane; EXPRESSED IN: 6 plant structures; CONTAINS InterPro DOMAIN's: Alpha_tubulin (InterPro:IPR002452), Tubulin (InterPro:IPR000217), Tubulin/FisZ, GTPase domain (InterPro:IPR003008), Tubulin/FisZ, N-terminal (InterPro:IPR019746), Tubulin/FisZ, C-terminal (InterPro:IPR008280), Beta tubulin, autoregulation binding site (InterPro:IPR013838), Tubulin, conserved site (InterPro:IPR017975), Tubulin/FisZ, 2-layer sandwich domain (InterPro:IPR018316); BEST Arabidopsis thaliana protein match is: tubulin alpha-3 (TAIR:AT5G19770.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archaea - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MREISIHIGAGIQVGNQSCWELCYLEHGIQPDGL MPSDSTIGVGHDAFNTEFFSETSAAGKHVPRAFVDF LEPTVIDEVRTGYRQLFHPPEQLISGKEDAANFV ARGHYTVGKEIVDLGLDRIRKLDNCTGLQGLFLV FNAVGGTGGSLGSLLELRLSVLDYDYGKSKLGF YPSQVSTAVVEPYNSVLSLTHALLEHTDVSLLD NEAYDICRKSLEIERPSTYNLRLISQISSLTTSL RFDGAINVDITFQTNLVPYPRIFHMLSSYAPVISA EKAYHEQLSIPETNAVFEFSSMMMAKCDPRHGKY MACCLMYRGDVPKDVNAAVATIKTKRTVQFVD WCPTGFKCGINQPPPTVVPGGD_LAKVQRAVCGMI SNNTAVAEVFSRIDHKFDLMYAKRAFVHWYVGM GMEEGEFSEAREDLAALAKDYEEVGAEGVDDEE EEEEY
205	#N/A	#N/A	0_323341:37	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein; FUNCTIONS IN: oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, 2-oxoglutarate as one donor, and incorporation of one atom each of oxygen into both donors, oxidoreductase activity, iron ion binding; INVOLVED IN: oxidation reduction; LOCATED IN: cellular_component unknown; EXPRESSED IN: 7 plant structures; EXPRESSED DURING: 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN's: Isopenicillin N synthase (InterPro:IPR002283), Oxoglutarate/iron-dependent oxygenase (InterPro:IPR005123); BEST Arabidopsis thaliana protein match is: 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein (TAIR:AT3G47190.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archaea - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MSESNIIENHSYPIFRSPFNDRTRKLDLKDDEQGE GAESTGTLDPPLPLPLIDFQCLDDEKLDTACKE WGIIFRLENHGVPSLLRQLQEHAKTLFRLPFEAK QGLAKRCPIKVFYGTALTSSTSSGVALLRSDALNM DWVEGLNVPSQLHQHLQHIHHDPLINDFRDLLE DYRGLHTRIASTLFEAMINLKLDOQVKPCLDLSE TGLRYYRPRTHINEHVKGLHEHTDSSISILSED QVGGIQLLKKDKWLKVEPIPTNLVNLGDMMQV LSDDEYKSVKHKVELNKNKERVSYGYVFFPEENT LINSKYPFSSNQFRAQVQDQDKLHGTKIGLQNF KLPHTSQTV

10

20

30

40

【表 1 4 - 8 4】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
206 sp_034930_fifr	#N/A	#N/A	0.923381	<p>SUPPRESSOR OF SALICYLIC ACID INSENSITIVITY 1 (SSI1); FUNCTIONS IN: transferase activity; transferring glycosyl groups, starch synthase activity; INVOLVED IN: amylopectin biosynthetic process; LOCATED IN: chloroplast; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Glycoген/starch synthases, ADP-glucose type (InterPro:IPR011835), Starch synthase, catalytic domain (InterPro:IPR013534), Glycosyl transferase, group 1 (InterPro:IPR001296); BEST Arabidopsis thaliana protein match is: starch synthase 2 (TAIR:AT3G01180.1).</p>	<p>MESLSQSFHVKGLNYSKFNLANPSEFRINQ VGFDFHLRQRCKRLCFVRCQSFSGKNEGVS SADGSSAIEEORNEGHFQOLIGPVKADASSIVS FNLYSQSGDGKSNVFLDLEKPNENDKEEKISSGS PSSFVFERGVKESVEKILEETIVEGEDKTVAKS KLSWVFTAAPYSKTGGLGDVCGSLPIALAER GHRVWVIAPRYHGTAAADKYAGAFDANCRIKVN CFGGEQEVAFHQRDGDVWVFDHPSYHRPG NPYGDSFGAFGDNQFRFTLLCHAACEAPLVLP GGYTYGEKGLFLVNDWHAGLVPVLLAANYRPG VYKDARSILVHNLHQGVPEPAVYDNLGLPPQW YGFLEWVFEWARAHELDKGEAVNLIKGAVTA DRILTVSQGYSEITTEGGYGLHELLTSRKVIN GIINGINSEWDPSNDAHIAAPYSIDDLGKAECK AALQKELGLPVKPDCLPGLIGFRLDYKGDIIQAA MPELLGDDVQFVMLGTGVAKYESWWMRETEARY KDKFRGWGFSVPISHRITAGCDILLMPFRFEPCLINQLYAMRYGTPVYVHSTGGLRDTVETYNPYAV PGDSSGTGTGWGFAPLAKDSMLLSALRNLKTY REYKGSWLGLMKRGMQDFTWDSAAAHHYEQV FEWAFTDPPYC</p>
207 sp_035120_edcp	#N/A	#N/A	0.83438	<p>PR5-like receptor kinase (PR5K); FUNCTIONS IN: transmembrane receptor protein serine/threonine kinase activity, kinase activity; INVOLVED IN: response to fungus; LOCATED IN: endomembrane system; EXPRESSED IN: 6 plant structures; EXPRESSED DURING: petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Thaumatin, conserved site (InterPro:IPR017949), Serine-threonine/tyrosine-like domain (InterPro:IPR001245), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271), Thaumatin, pathogenesis-related (InterPro:IPR01938); BEST Arabidopsis thaliana protein match is: receptor serine/threonine kinase, putative (TAIR:AT1G70250.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).</p>	<p>MILIAFIIIIYAKORRRKRKFSLLWTKTMMMNQNV AFELKSYGSLAPKRYTYAEIRKMTNMFEDKLEGGE FGAVYKGLPDNSRFVAVKLNKLGKNGDDLNN EYVSMGRTNHVVVTLGLGFCFEGNKRALIFEFFP NGSLKFEIVTADHSLTSETLFSIAVGIAGLDYL HRCGNTRLHFHDIKPHNILLDEELRPKISDFGLARL CFPNRKESLISMSGARGTIGYIAPEVFCRNFQGGV SHKSDVYSYGMMLDMVCGGRNINIVSEAKGSSEV YFPEWIYSQLEIEDELENDIVVHGMIKDDDEEKEVR RN</p>

10

20

30

40

【表 1 4 - 8 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
				<p>DEFECTIVELY ORGANIZED TRIBUTARIES 3 (DOT3); FUNCTIONS IN: signal transducer activity; INVOLVED IN: in 6 processes; LOCATED IN: cellular component unknown; EXPRESSED IN: leaf primordium; CONTAINS InterPro DOMAIN/s: NPH3 (InterPro:IPR004249), BTB/POZ fold (InterPro:IPR011333), BTB/POZ-like (InterPro:IPR000210); BEST Arabidopsis thaliana protein match is: Phototropic-responsive NPH3 family protein (TAIR:AT5G64330.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 796; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).</p>	<p>MKKPVPAAKAEIQENNSDAADQITLPTENNVD YTFQKKDHSWFVTSDIPTD.LLIQIENITFHVKYA LLTKSGYGRTEFQPTSSGLENNIKLNDFFGGPVT FKEVLKFCYGPVPLSANNVAPLRCASEFLEMTE EFEDGNLISKSEAFITVLLSSWRDSVVLKACEN LSPLAENLQVRRCCDSISRKAFQDKNEIIDDKA WWFDDVATLRIDHFGRIIAAMKPKGKFPDLGSCI KHYAEKWLPMDGEVEGRRHWYGNKNELOISI CGANHEDDIGNKDKKIIIESLISILPQKEAVSCK FLLWMLKRAMICYATPALISELEKRVGSLVDAN VQDLLIPRCTNKDRGGLNSPEERTVHSIDTVQRIL EYFLMHQKKTEDLDVSKLLDNYLAEIATDPNL SISKFQVLAELPANTRYCHDGLYRAIDTYLKTHP SLPEQERRRCLRIMNCQRLSLDACMHVAGNDRL PLKVIQVLFSEQVKMRTAIOQSKEQTEGNTREND ESWSSAKKEIMNLKTELEKVKERMSELORDYSE LOREYIMQINKPRFNSSWIFRWKMKMSISLTFEK AEEIEIGEDQETPNQEGSTPNRIIRLSIT</p>
208	#N/A	#N/A	0.948337		
				<p>Protein kinase superfamily protein; FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation, cell wall macromolecule catabolic process; LOCATED IN: plasma membrane; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Peptidoglycan-binding lysin domain (InterPro:IPR018392), Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009); BEST Arabidopsis thaliana protein match is: protein kinase family protein / peptidoglycan-binding LysM domain-containing protein (TAIR:AT2G23770.1); Has 102109 Blast hits to 101070 proteins in 4036 species: Archae - 73; Bacteria - 10993; Metazoa - 38718; Fungi - 7726; Plants - 30221; Viruses - 349; Other Eukaryotes - 14029 (source: NCBI BLINK).</p>	<p>MFPKISTKFSQISHCLLLLFLINTHHSNSQPPTN GYTCTTTNTSNNIYPCQTYLYRKPPEYLDLA SIADLFNVSRLSISKSPSINSPFSLLPDQPLFIPIT CGGNIPVRNQTLSLFAFARMYTFKPDOTFYFVST HVGNGLTTYSEVFNPELDVNLQIGSDAYFFIF CKCPNKTQGNVNTCLVSYVFPQDDTLDSIAELF KYTKQSIADVNANNLGAFTVFPVSKLVTLHQP PGVNTQAGGSNTNQTTAMNGNSDDHKKAV VGLGVGLGVSV.LLVLSGLLIYREIERYRKKLKER ELEVISRPENGGMGDQRRRIKQLEQNLLADVSD TLDKYKVEIDELRKATGGFGHFKFIQGSVYKGGI DGKDFAKKMKWNAYEELKIQKVNHNGLVRLLEG FSIDPEDASCYLVYEIENGSLNSWLHENKRKLN WKARLRSIDVANGLOYIHEHTFPQVVKDKIKSSN ILLDSNMRRAKIANFGLARSGCNAITMHWGTQGYI APEYLTGQVSPKMDVDFSGVILLESLSGREVIDE EGKVLWVSAHEIWECEGDEGKARKLMEWIDIKFL QKESCSVDALTNIMTVAVACVKNKDPKRPMSMVD VYVALCKSDLLFCDDISGNEFFSSSPITAR</p>
209	#N/A	#N/A	0.938313		

10

20

30

40

【表 1 4 - 8 6】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 Villin-like 1 (VLNT); CONTAINS InterPro DOMAIN/s: Gelsolin (InterPro:IPR007122), Villin headpiece (InterPro:IPR003128), Gelsolin domain (InterPro:IPR007123); BEST Arabidopsis thaliana protein match is: villin 4 (TAIR:AT4G30160.1); Has 3095 Blast hits to 1746 proteins in 186 species: Archae - 0; Bacteria - 6; Metazoa - 2441; Fungi - 102; Plants - 219; Viruses - 0; Other Eukaryotes - 327 (source: NCBI BLINK).	aa seq
210	sp_049620_ccgu	#N/A	#N/A	0.953742		MLEHDVHYWLGKDAINEVESLMVSDKALELDLAL GSCTMQYREKQGQESAKFLSYFKPCVPCEGVY TSGLERSNRGTYHITLLA
211	sp_054040_cmic	#N/A	#N/A	0.923313	PAP/OAS1 substrate-binding domain superfamily; FUNCTIONS IN: nucleotidyltransferase activity; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; CONTAINS InterPro DOMAIN/s: Nucleotidyl transferase domain (InterPro:IPR002934), PAP/25A-associated (InterPro:IPR002058); BEST Arabidopsis thaliana protein match is: Nucleotidyltransferase family protein (TAIR:AT2G40520.4); Has 391 Blast hits to 365 proteins in 105 species: Archae - 0; Bacteria - 0; Metazoa - 57; Fungi - 102; Plants - 177; Viruses - 0; Other Eukaryotes - 55 (source: NCBI BLINK).	MDTQDLRSPLPSSFSSSSSSSDVIPFTIGATRW QVGDRTAQQLAVIQPNAESENRRKEIEYLOFLIE DALGIKVPFPGSVPLKTYLSHGIDILTAVSHFSTS QDLAADICCLLENEKKNANVQDLQVKDVICVPA QVRVVKTYGGICVDISFNQTAGLCAFYLEQVD RFIGKDLFKRSVILIKAWCYYESRLLGGFHGLIS TYALETMVLHVINLPHSSLPCPLAVLYKFLYYST FDWNSYCVDINGLVIPLPEIVVQSPQDQDDVLL TDTFLRSSREALSSMGVGRTRNDSAFSVKFLNLV DPLKEDNNLGRSVCKGNFFRIKSALTYGAKQLGG VLMPLQNIQNGLEKFFKDTLERNKGKRRADVIV PVDEYGSNTSTGNSGHDHLLSGIHYGLWFFHDYG MMYPVLPETSENSRWVSSQWVDMGQNVQLDH SMYNNVQSPSPSSVNRVTSFPQTNAASVKEKW KSRGTGTYPNLEHLMRRNIVLEARKNEHECEKE QGGKGGKEVESVDSVDPACTPREANLSSDEGANTD VFNFSIEDFFVLPGSKPITIKVHQLLEAPEPEQVKE MSVSLDLSLEFDNKHSSRSQSLTEAGKQSDSDV SAILDLVDPTTAVVLEDSQESSKEEKMKWPKLEL ADDFEFPPLSRPEPEPVITLNKKEFFPLRASLKSO KKGKWKKDKF

10

20

30

40

【表 1 4 - 8 7】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_ktgg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
212	sp_055020_skr0	#N/A	#N/A	Concanavalin A-like lectin protein kinase family protein; FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: endomembrane system; CONTAINS InterPro DOMAIN/s: Legume lectin, beta chain (InterPro:IPR001220), Protein kinase, ATP binding site (InterPro:IPR017441), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Concanavalin A-like lectin/glucanase, subgroup (InterPro:IPR013320), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271), Protein kinase, catalytic domain (InterPro:IPR000719), Concanavalin A-like lectin/glucanase (InterPro:IPR008985); BEST Arabidopsis thaliana protein match is: Concanavalin A-like lectin protein kinase family protein (TAIR:AT3G45330.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MFQVMAETCSYLHKKIALIFFFVALTFAQDEKQF MYNGFQGAKLHLDGLANIHPNGLLOMTNSLQG SGHVFYQPLKFNTRSFSFTTFVAINTDYPGHG GHGIAFVSSSMNFDHAPAEYLGIFINISNGNINS NHVIAVELDTYQNAVFKDIDGNHVGIDINGLESVG SASVAYYSDEERKNMSLQSGGTPMQIWIYDYG AKMLMNVTVAPLKHHPKPKKYLSTPVNLSPALLD PVVYGFSSSTGLTATEHYISGWSWSQIGQAQDL DLTKLPSFKIRSQRRKKLVFVLIIVLSLLEMLIA GVAFVWRKYEELKEPWEQEYISHRFSKDL TATKGFKDELLGFGFGKVKYKELISITGKPTTAT DVFSFGIFLEVACGRRPNSLQEFREDFVLVDW VYDCWKKGVILETSDPTLEGNHQVEEMELVCLK GLLCSHPKPEGRPSMRNVVQFLHGDASLPDIPS DYDHQSNYSYFGGWSRISLPSLSSLGNISRQG CDLSEALFFRFSNIL
213	sp_057480_jiorn	#N/A	#N/A	Glycosyl hydrolase superfamily protein; FUNCTIONS IN: cation binding, hydrolase activity, hydrolyzing O-glycosyl compounds, catalytic activity; INVOLVED IN: defense response to fungus, incompatible interaction, response to salt stress; LOCATED IN: cell wall, plasma membrane; EXPRESSED IN: 11 plant structures; EXPRESSED DURING: LP.06 six leaves visible, LP.04 four leaves visible, 4 anthesis, petal differentiation and expansion stage, LP.08 eight leaves visible; CONTAINS InterPro DOMAIN/s: Glycoside hydrolase, catalytic core (InterPro:IPR017853), Glycoside hydrolase, family 17 (InterPro:IPR000490), Glycoside hydrolase, subgroup, catalytic core (InterPro:IPR013781); BEST Arabidopsis thaliana protein match is: beta-1,3-glucanase 1 (TAIR:AT3G57270.1); Has 2169 Blast hits to 2154 proteins in 141 species: Archae - 0; Bacteria - 6; Metazoa - 5; Fungi - 22; Plants - 212; Viruses - 0; Other Eukaryotes - 15 (source: NCBI BLINK).	AQGVYCYGMLGDNLPTEKDVVHLYKSNIGIRAMR LYSPLNATLQALGGSGIKMLDVPNETVPSLASS DPSSAMQWVTYVVPFASSIKYIVVGNIEHPSDK QASSVLPAHQNLNALKQNKLGDLIKVSTSIDITL IVNSYPPSDGQFNDTSSYTAPIKFLSTNKSPLL VNVYTFAYVSNQRDISLNYALFTSPGTVVTDPNN NKNYKLFDMAMDSVYAAALAKAGAPNTAVVNSE TGWPSDGGDAATVNNAGTYRNLINLHVEQGTPL RFGQAIETYLFAMFDEDEKKGIGANTERNFGLFTPT KHPKYGQLNFS

10

20

30

40

【表 1 4 - 8 8】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq  MASEQPASTSGRCPGKEPVTDDIDMATGEYSDDG DSIFEDSAGKDVDEHRGQDGDGDEAGVSGRNDTD VGNDSIGDDDDPADVIVAEVLADEAKEQPYFTD DSTDVHRREELPPLDSLPRTKTLDRDFGFDENW RPVCTQPPPKGYDYDITISTYSTRASSRRAVTV KNFVGTGSKTKSSASISKTAQPSNMQSNPTGR GGKSRKRSLTSGRSDPNLKSAAAGDDVSAAP SIEILDIEGGGGENEPLQHCATSAAVTGGNSM PPVSRSENRSGDSRLNYHLPLRTGGWIEDTPFKI PPAMKSWFGSPGGEPTDQFYPNIDLCCGESVQ SSIDLYYRWNQLQAQNSIDIAELKRAEAAEA ALKDSAKKLETASSDLEAVNKRLDEVEPKLKAET ERADDLTOQLIDLNRGLPGVYKKTAAKRAVDKLLQ SD MENPHFLSLTLVDLSLSSTAPLSISANADHPLR RSDSCCCGVAALLLHRVSPSPSPSILLRLTPPP WWSLLPPLLPHKWWVTLSL MQYSPCLLVSPFINPFLISLVLFVFLSFLVILISLQ LCYFYPKHRLQLPPSGMGWVYLGTELKLYTEN PNSFFSNRQKRYGDFKTHILGCPVCMISSPEAAK IVLVTKAHLFKPTYPPSKEKIMIGPEALFFHQGDYH SKLKLQSSFLPCTLKASVSDIENIVLNFPLTWE NNTINTLHEMKKYAFDVAISAFAGDNKELEIEGKN LYHSLKGYNSMPLNFPFGTPFHKSMKARKLLNET LRKLIARRRNNIDGGLLGVLLGSRREQKTGKNN ELSDSQIADNMGVFAAHDTTASVLTWLLKYLHD NHLLQAVTKEQGGIKORLEIANRGLTWEDTRN MPLTSRVIQETLRTASILSFTIFREAVHDFEFGYLI PKGWKVLPLFRSIIHCAEFFHPPEKFDPSRFEVP PRPNTYMPFGNGVHSCPGSELAKLEMLIFLHLHT TTYRWEVVGDEDDGIQYGFPPVPRGFFPIMVHPQ K
214	sp_063280_hapr	#N/A	0.2443544	uncharacterized protein	
215	sp_071600_ifwj	#N/A	0.9353238	uncharacterized protein	
216	sp_075610_odsj	#N/A	0.2333853	cytochrome P450, family 707, subfamily A, polypeptide 2 (CYP707A2); FUNCTIONS IN: oxygen binding, (+)-abscisic acid 8-hydroxylase activity; INVOLVED IN: in 7 processes; LOCATED IN: endomembrane system; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAINs: Cytochrome P450 (InterPro:IPR001128), Cytochrome P450, E-class, group 1 (InterPro:IPR002401), Cytochrome P450, conserved site (InterPro:IPR017972); BEST Arabidopsis thaliana protein match is: cytochrome P450, family 707, subfamily A, polypeptide 3 (TAIR:AT5G45340.1).	

10

20

30

40

【表 1 4 - 8 9】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
217	sp_076640_oyhh #N/A	#N/A	0.528333	Protein kinase superfamily protein; FUNCTIONS IN: protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: endomembrane system; EXPRESSED IN: 7 plant structures; EXPRESSED DURING: C globular stage, F mature embryo stage, petal differentiation and expansion stage, E expanded cotyledon stage, D bilateral stage; CONTAINS InterPro DOMAIN/s: Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009); BEST Arabidopsis thaliana protein match is: Protein kinase superfamily protein (TAIR:AT5G61570.2); Has 1807 Blast hits to 1807 proteins in 277 species: Archaea - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MKQTQFLKFLIPFTFIILIIYLKRNKPEKNQP SETGDPDELGNTINKNKNEEEEEELMCFQGG EDLTVDPILDAPGEVIGKSSYGTLYKANLYRNIN NNNSILLRFLRPACSGTIREVPVVAHGLVHSH LVPLKAFYMGPRGEKLLVHPFYAHFNLAQFKDS NNESHKWDVVISIALGLAKGLAYLHTGLHKPIHG NLKSKNVLGSDFPYISDYGLHLVLPNTAGQEM LETSAAQGYKAPELMKMKDVSESTDIYSLGILLE LLTGKEPVITANSDDSEQLYLPNVRRAILDHRIND MYHPQIVSNQTKQRHKTEDCALKLVLQAMSCC SASPSSLRFEASEVARKVEEIGKEQALQHKRT
218	sp_077030_crpq #N/A	#N/A	0.241405	Tetratricopeptide repeat (TPP)-like superfamily protein; FUNCTIONS IN: binding; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Pentatricopeptide repeat (InterPro:IPR002885), Tetratricopeptide-like helical (InterPro:IPR011990); BEST Arabidopsis thaliana protein match is: Pentatricopeptide repeat (PPR) superfamily protein (TAIR:AT1G73710.1); Has 56373 Blast hits to 14756 proteins in 361 species: Archaea - 8; Bacteria - 288; Metazoa - 580; Fungi - 796; Plants - 52865; Viruses - 0; Other Eukaryotes - 1836 (source: NCBI BLINK).	MFSKAQCFDINCFFKSPCHLHTSFNIFKFSVPIET HDSKPKRDSNVLEIPNKGYCKFRFEKQANNS CGVIDKSSLSKNPHSGVDELGRLOSGNSRNQQ KKINGEVYEFAGSGNFPFAKVKVHTRCSTKAWASYG GCIPSLALDSISDLDEAFRPWEERLSNERSIIL KEQSSWERAMEIFEWFKKGCYELNWHYHNIMLR ILGKARKWVELESWDEMKMKRIAPINSTYGTLLID VYSGGLFREEALVWLRMNEQGMPEDEVMTGI VVMQYKKEGEFRKAEDFFKMMSSDKSKEKNT FOGHKMGAMKVKSGDQLSLSLYTYNTLIDTYGK SGQLQEMSQTFQKIMDDGVIPDTITFTNTMIHYGN IGRLEEVNKLVKMAELNRSPTRTYNIILSYLAK RDDINLAFSFAFMKQDYLPQDPVSYRRTLLYAFSI ROMVFEAEMLISEMDKRGLEIDEFTQSALTRMYI DAGMLEKSWSWFERFHLEGNMTTECYSANIDAY GERGYVKEAEVFCQEKKNKIGVLEFNMAIKAY GIGSYDKACQLFDSMGKLGVPFDRCSYLSLQSL AAADLPHIAKPYLLKMQEAGLVNDCVPCYALUSF AKLGQLELAKSLYKEMKGYNVRPDDVLYGVLINAS AESGNVSEAMEYDALKKADLNDVINDVNSLILKY TKVGYLEEAIVETRYLLRSFEKGDVYSSNCHIDL SKFSMVTAEAEVDFDVRERGEANEFFSFAMMLC MYKRVGRLEEAEMAEKMQELGLLTELSTYNSV LALFAMVGRLEKAVSIFKMSLEIEPDDSDTSYKSL GFVLVKSQVPKQAVAKLEVARKKDKTKSGLQAVV SALTSSVCRDDYGNYSDD

10

20

30

40

【表 1 4 - 9 0】

	A	B	C	D	E	F
1		分析に用いたペアとの相関係数				
2	sp_107620_kqnh	sp_107660_klqg	sp_107670_ptqx	sp_107670_ptqx	注釈	aa seq
		(SOAP1)	(SOAP2)	(SOAP2-like)	basic helix-loop-helix (bHLH) DNA-binding superfamily protein; FUNCTIONS IN: DNA binding, sequence-specific DNA binding transcription factor activity; INVOLVED IN: regulation of transcription; LOCATED IN: nucleus; CONTAINS InterPro DOMAINs: Helix-loop-helix DNA-binding domain (InterPro:IPR010992), Helix-loop-helix DNA-binding (InterPro:IPR011598); BEST Arabidopsis thaliana protein match is: basic helix-loop-helix (bHLH) DNA-binding superfamily protein (TAIR:AT2G22750.2); Has 2910 Blast hits to Metazoa - 77; Fungi - 51; Plants - 2768; Viruses - 0; Other Eukaryotes - 8 (source: NCBI BLINK).	MEIQVPSLNNSHNEYNQSETEAASFDEQLTA ALLGEDFFPLPAKNCCTYPDQSHWDFEYDLVD QFITPPSYSSNSSASYDYNNNNNYDDVNDNTM VSPGVQEPGRGSKRKEVITEKQAVNGGTRKRK RHPQVQDHLAERKRREVLGHEFSLSTMIPGL KKDIDKTSILGEAIKMKHLQEKVNVLEKVSACQKS ARSSMVVKKYQLGVENDDSDDTGNGSSIVSD NSASDGGGGCGSNGSDSSNNAMN.LNNSIPEIE VKITDETLKLVYCEITONGVMAKLLKVVENHLLS VTNCSLLPFDFKLALDITVAIQIQEGFDMNVKDFVR SIRSVLRPAGI
219	sp_079960_puqw	#N/A	#N/A	0.343343		SNGVINGLQRNGRFKEAIDILCEKNLVKEAVQLLN HIQHPASAIYSLLLIICLQPSLNEGRRVHDIKA SGFVPGVVISNLFMLDMYVRCGSLSDAQKVFDEM PERDLCSWNTMIKGYGKTGKVELARKLFDMPQ RDGFSWTAMISCYVQHRRPKEALNLYRKMQSC GSWDSNKFTISSVIAASASLPSLLSGKETHGRIM RLGLECDSAVWSALSDMYAKCGCIEGARHIFDKT VDRDVVSWTTMIDGYFEDGRWKDGFGLSDML RSRVFPNEFTFAGVLNACADRALENVKGQAHGH MIRIGVDSDFSFAASALVHMYSKCGNISSAKNAFE GQCTPDLVSWTSLIVFAGNGHTEESLELFLML RSQDTEPDHVTFFVGLSACTHAGLVEKGEYFHSI KDRHGLTPIMEHYSCLLDLLSRAGRFKEAEEYIMN MPMKPKDFLWASLLGGCRIHKHVELAELAAKALF EIEPENAATYAT_LANIYATAGRWEEVAVKVRKMDM NRGVYVKKPGLSIEWVRRVHRFLYGDTS_HLSYN EIHAFLEISIKMKQEGYVPTDFVLHDVVEEQK EQNLSYHSEKLAVALFAGVISTPPTGTPKPKVFNLRTC VDCHTAIKIFRITRKRKIVRDSNRHFFHFFEDGSCSC GDYVW
220	sp_080330_ahq	#N/A	#N/A	0.343343	Pentatricopeptide repeat (PPR) superfamily protein; CONTAINS InterPro DOMAINs: Pentatricopeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: pentatricopeptide (PPR) repeat-containing protein (TAIR:AT2G22070.1); Has 43848 Blast hits to 14238 proteins in 253 species: Archae - 0; Bacteria - 12; Metazoa - 20; Fungi - 118; Plants - 43065; Viruses - 0; Other Eukaryotes - 633 (source: NCBI BLINK).	

10

20

30

40

【表 1 4 - 9 1】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
221	sp_092890_sdyj #N/A	#N/A	0.932692	MATE efflux family protein; FUNCTIONS IN: antiporter activity, drug transmembrane transporter activity; INVOLVED IN: drug transmembrane transport, transmembrane transport; LOCATED IN: chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Multi antimicrobial extrusion protein MatE (InterPro:IPR002528); BEST Arabidopsis thaliana protein match is: MATE efflux family protein (TAIR:AT4G39030.1); Has 7799 Blast hits to 7783 proteins in 1733 species: Archaea - 152; Bacteria - 6456; Metazoa - 30; Fungi - 49; Plants - 387; Viruses - 0; Other Eukaryotes - 725 (source: NCBI BLINK).	MQKAIWSSPYQFKLPINQNSLSSHFRPKTS PKFPNLCISKFRPNRPNLRKNCCISPSQEVGEN LETHEPESEISGFEINPENQNLQFPKPTNEVDDF ANQSTWDQIFEIHKFSGPATGLWICGPLMSLIDTA VIGQSSLELAALGPATVLCDDYGVFVFMFLSIATS NMVATSLAQKDEDDQVQHQISMLLFMGLASGVLL LLLTCKFGEWALTFVGVHNVHLVSAANSYQIQR ALAWPAILVGSVAOSASLGMKDSMGLPKALLIAS ALNMGDILVLCFTFLGYGIAGAAWATAASQVAAAY AMVEALNRKGYNGYSFVSPVLSLSELOEIFGIAAPY FLTMVSKVTYGLLIYFATSMGTHVVAHAHGVLLNI FGMCSWGEPLSQTAGSFMPELLYGANRSLYK AFSLKSLLIIGVSVGLLGTAAATAVFPVLPNFETS DTVIGEMRNWVICFFLALSVTPCILPCEGTLLAG RDLKIFISIMSGCFAGALVLLFVRSQGLGLAGC WSALVGFQWARFISISLRLLSFGVLFSEDLDRF RLADLKVA
222	sp_094570_sqpg #N/A	#N/A	0.982062	CLAVATA3/ESP-RELATED 26 (CLE26); Has 30201 Blast hits to 17322 proteins in 780 species: Archaea - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MTTTTTTSQDKHNKHEVTVGRGKLSARSKV DINYSKRKVPNGPDPHINRRAGNSRRRPPGQA  MAFHSSCSGGSNLPEFFQQLNLRSPVPTKP VHTLILSTNWTGVTITPPSFTFPGWNINGTIYSGI RNGGELYKISTFTLRDFLMNP LEVPOSLWNLYMW VPGDDGPKTLRPAHPSLQDQVDSMSKVAGIVLFR WWTQRLAACSVDR_LVIPAELOQLSDPEVPLPGK LYRLITEDWVRAQYELSCNNVTFTCAHSDKKV NRQISLSCAWGWKKGSEKLMGGLKLRDNNVYEM YSPYLERFVTKDNISFDARVETNSLPLKLNRLMVM LWNSRGVARWFRDFDQLMDDHSPDIVLTD TRYSKRMAEEIIMDIPEFDSVSVVEALFGSGGAIIL WNSLHYTFIRIGKEQRIALHGVIQCAESYD
223	sp_095520_qhxy #N/A	#N/A	0.931195	uncharacterized protein	

10

20

30

40

【表 1 4 - 9 2】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
224	sp_095840_iuti	#N/A	#N/A	0.3324136	<p>Pentatricopeptide repeat (PPR) superfamily protein; CONTAINS InterPro DOMAIN/s: Pentatricopeptide repeat (InterPro:IPR02888); BEST Arabidopsis thaliana protein match is: Tetrairicopeptide repeat (TPR)-like superfamily protein (TAIR:AT3G29230.1); Has 37071 Blast hits to 13818 proteins in 252 species: Archae - 0; Bacteria - 21; Metazoa - 136; Fungi - 146; Plants - 36104; Viruses - 0; Other Eukaryotes - 664 (source: NCBI BLink).</p>	<p>MMNTKLLNLSALFSSYKLSYDEVCSILQRCTKAK          ASYSAKHIGMLISGTDVSRMSLDKILGYASC          GNLRYAHLDFKMPNPNFVAFNWMISVMSFHGN          YQEAIYFSLMOKMGTLPNKYTLFVLKSCVGLL          DLNKGGRGVHAVMNMKMGFECDALVCNALVDMYC          KCGNTQYARRLFDNRMSKDIASWTSMISGYCNT          GKLINEALFEQMKLEGFEPNDFTNWNAITGYAQ          SGRDRGALFLFSRMRNREGLCCD_LITWNAI_LSGFS          RSKHPFEALKLIFREMLIAGIRPNHVTVTGLLPVCG          LIGSPQGRREHLGLICRLSLVNVVYVASALIDMYS          KSGSVKAAKNVFEGLVFRVALVWV/IIGCLGKNG          VVFESLRFEEVMEQEGQFQNEVTLVCVLSACSH          GGLVEKAMEIFRSMKKEYQVEANKEHYACVVDL          LSRAGEVEEAYELIKRMPMEATEESIFGAFNGCNI          HGRDRLAEKMADSLKAAWKPGALVTLNSIHA          AGEEWQGVQNVRLKMLKIGVYKPKGFSL</p>
225	sp_101590_daje	#N/A	#N/A	0.9549323	<p>Ankyrin repeat family protein; CONTAINS InterPro DOMAIN/s: Ankyrin repeat-containing domain (InterPro:IPR020683), Ankyrin repeat (InterPro:IPR002110); BEST Arabidopsis thaliana protein match is: Ankyrin repeat family protein (TAIR:AT5G35810.1); Has 10613 Blast hits to 6237 proteins in 350 species: Archae - 27; Bacteria - 619; Metazoa - 5684; Fungi - 682; Plants - 1506; Viruses - 31; Other Eukaryotes - 2164 (source: NCBI BLink).</p> <p>Polyketide cyclase/dehydrase and lipid transport superfamily protein; CONTAINS InterPro DOMAIN/s: Lipid-binding START (InterPro:IPR002913); Has 197 Blast hits to 196 proteins in 72 species: Archae - 0; Bacteria - 22; Metazoa - 96; Fungi - 2; Plants - 31; Viruses - 0; Other Eukaryotes - 46 (source: NCBI BLink).</p>	<p>MSGQPPSGSQVSSQQASSSSTTHDVTLSIIPSP          TSSATVPPPPRPPPPPLSAAVISSSRADFLKFGI          PLYKAALGNWPAAEQILOQSPDWIRHPKAGGE          RVLHIAAARTRHFVKELVKLMIEGNIDDLALSND          AKNTALCLATSGVWEIAKVMVDNMGALPNIRGS          QDMTPQLQMAVLLGRRDMWVYLGIVTDNNQLTDV          DRIALTSSINTNLFVYVNHILKEHPELALLRDGKR          ETALHALARKPVRSVDYHEGSIWKTLYLGLGGKL          AQGKFPQSSLEQLVQRLWDEVIKQEDKNISRLI          GSPWRLLFAAKCGNVEFLTTLIRTYPLDIWKVDE          ENRSIFHAIINRHEEIKLIYELGAIKDLIAYDKDGN          DNNMLHLGALAPSHRLNCGVGAALQMGRELL          WFKAVEKVSSEYAEAEAKNSYKTPRALFTEHRE          GLFAKGEVWMMKQTAQCSVVAALATVVFAMF          TLPGGIDSKGTFVFNRRRLFVFAIGNALSLSFSSC          AAILMFLSITSRYAHEHDFLKSPLKMLGLTGMFI          SVITMMAAFTVTSFITFREGVKSVMWIIPIPIALFAA          VPVALFAFQQYPLLLLDIYKSTYRSWSLFLQPTKPK          LFPYSYS</p>
226	sp_114190_irms	#N/A	#N/A	0.945233		<p>IYQQWDPDLVEARYIKDLEENLSIIRLRFGBASRP          LFKNREFVYVERRETMDGTLVAVASLKPKEIAG          LYPKONKAIRGILLLOSQGWVWVEKLEDHSCMVTYV          VQ</p>

10

20

30

40

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたベクトルとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
227	sp_116620_oodk #N/A	#N/A	0.3347734	RNA binding; RNA-directed DNA polymerases; FUNCTIONS IN: RNA binding, RNA-directed DNA polymerase activity; INVOLVED IN: RNA-dependent DNA replication; LOCATED IN: cellular component unknown; CONTAINS InterPro DOMAIN/s: RNA-directed DNA polymerase (reverse transcriptase) (InterPro:IPR000477), Protein of unknown function DUF1204 (InterPro:IPR009596), RNA-directed DNA polymerase (reverse transcriptase), related (InterPro:IPR015706); BEST Arabidopsis thaliana protein match is: Protein of unknown function (DUF1204) (TAIR:AT4G20700.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archaea - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 365; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MWASRPNMSSSVSVEADDADLYFDGHGVSQPGLLSSADSPFQAQFLSRPCGILLSGGPTFPDPFCVQDNICSGSPNPNPGRRRPPGMEAGPSN WVEYPPDYHLSPVPDFGSPSPSLSESFPVTPATDSGYFPDSSNADDFAGNNGDSSDAESEVQVQVLLAAPPADGVLAVDTSPVYSPASVDSSTRLQPLSPESVSAFAVSNVDHPESVTQFRPIGLCNVIYKCIKCLTHRLQGMSSLVSETONAFVPGRLMSNDCLLAHELISFVNNCRSKSKCYAAIKLDMINKAYDRVLWDFLKTLSAFGFFPYWVHIHQCVSTVSYQVLVNGEPTSSFRPCGGLRQGDPLSPYLFVLCMENRRAGPSIAPEARSRSRLQKLVAEDELPKQSI LQTOASMVETDPSWPEPKLPEVRDKSPFRAYC KRSKLDIRSLQVQAOQCRSSRRPKLSSSMKLI AEDPVSEKQMMISFGSGCRGFLQKVLTKSSVS SSSPVS KALRAVPPSSAVLKPKAWEVAASQGPSR V
228	sp_118780_oaqp #N/A	#N/A	0.35719	uncharacterized protein	MEERAGDPSPPQIQSNLSALSQRFTPQQLLTASAVMKAMAE_LSSRRTEGNQIVVTGNRPGAMPVFN LYETLEHEWLPAQPEPILVQSENPGRKRKRHSR RRRSTRRESVSEQEGSFPAGRESTPYHEEYI VQSPQP
229	sp_119720_znuk #N/A	#N/A	0.323636	uncharacterized protein	MVASLPLCAAILSNKNKYLISVYKGGKTKHLRFSGEEIVSRSTKFAIERSKEDDENTELVHLRCCCHNNKYLVKKEAEGWIAAAAEKPEEDESWSSTLFEPIVESHGTYGRTVLRFKHVQSQLYASLSKDDDDDDDDDDSLSKLQRLQEIYEDVNSDDVVTIIDWGTLVILKRVAFKSSSDGKYLVDQQRSSAHPCI RSSVDKDDARVIHEVYTNPDGTLSEIKSSDADGFWTSHEGDEIVPEFADNTSEOKLATMEIKKGIQFTVPIKISSNEIALRGLP_LNKFCIDGTSKELYCKA NSILRETKVVEEVVISREVSNIIEYHLLDARIYDEA TVLSTNIHNTSENAFTEYSFRFYDKRISRWTN GGSLGLVPVNIKSAIGLIPTENGVISGETGYE WGESKTEKEMKTYNTSVPMQTQATIRLVSTGTCDDVPFYSVREHVHDGKTVVYTMDDGLYTGIFNYKLOQTOMETKPI

10

20

30

40

【表 1 4 - 9 4】

1	A	B	C	D	E	F
		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
					<p>Leucine-rich repeat protein kinase family protein; FUNCTIONS IN: protein serine/threonine kinase activity, kinase activity, ATP binding; INVOLVED IN: transmembrane receptor protein tyrosine kinase signaling pathway, protein amino acid phosphorylation; LOCATED IN: plasma membrane; EXPRESSED IN: 13 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Leucine-rich repeat-containing N-terminal domain, type 2 (InterPro:IPR013210), Leucine-rich repeat (InterPro:IPR01611), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: Leucine-rich repeat protein kinase family protein (TAIR:AT3G47580.1); Has 213556 Blast hits to 133070 proteins in 4776 species: Archae - 194; Bacteria - 22980; Metazoa - 68117; Fungi - 10236; Plants - 86850; Viruses - 328; Other Eukaryotes - 24851 (source: NCBI BLINK).</p>	
230	sp_120350_rpma	#N/A	#N/A	0.9281533		<p>MSCSNSFFYWIIFSLSVFFCFLEFSIASDITNDKIAL LEIKAKISDDPLGVTSSWNETLHFCEWYGVTCG GRHORVTVLDQSSKLTGTLSPYLGNSLFLRILD RNNSFSSVLPPEIGRLHRLQTLQLSNNSIGGEIPS SISNCYSLTGYLILNKVGEIPTVYGSLSHLOFLK LTRNLTGSIPISSIGNLSSLSFVSGNINLVGKIPD SLGKLNLTSLCDNKFSGELPFSIFNLSLLTYL CLYQNDFEGNLPDIGNTLPLQWFSIFNFRFSG HIPTSISNASNLQVLDLSKNLGRQVPSLHKLVNL IHLVYSNLSLFGQADLNFVSSLANATNLQWLE IDQNKFGVFFKICKFSSITKLVSDNINIAEIPNC IENLAEMQMFYSKNALSSVIPHSIGKLNQLETLN LGDNKLTVIPPSIGNLTKLMTDLSNRRLEGQIP PALENCRSLIGLFLSNLNLSGSIQSLYNIPLSIA LYVDGNHLTGSLSSEEVGQLKLEALDVSQNMLS GEIPSSLGSCLSELLGMDENNFQGPINALQKL TGLVELGLSYNNLSGKIPTFLASLNQLNLSYNNL EGEVPVIGGVFNATNLSIIGNKRICGGISALKPHC NVSPTVQRKLNKPKLILVILSTILGVTLLVLLV LSYTCLEKTKETTTTNDSEKNRNFYSQTLFKA TNGFSSENLIGTSSGVVYKGLIEDGSTVAIKTFN LEYRGASKSFLVECSVLRISIRHRNLLKVTYCSSI DYQGDKFKALVYEMWNGSLDNWLHPHIEATRGV EDTNDSDRLNLRQRLDIADYAFSLEYLHHPHGG ASVHCDLKFNSVLLDENMVAHVGDGFLAKFLMK DTDSHAKGNSIGVGTGTGTYTPEYAGNEVSTR GDVYVSGILLLEMTFGKRPNTNEMFKGGLSLHCFV KEALPEHVIKLDHVAEDINFEETDNSMMLEALIS ILGIALCCSSEVPRERLNMINDVTAKLSSIRNRLGG ITCSNKRGEFKQGHGIDNV</p>

10

20

30

40

【表 1 4 - 9 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1) 分析に用いたバイトとの相関係数	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	<p>注釈</p> <p>Oxidoreductase family protein; FUNCTIONS IN: oxidoreductase activity, binding, catalytic activity; INVOLVED IN: oxidation reduction, metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Oxidoreductase, N-terminal (InterPro:IPR000688), Oxidoreductase, C-terminal (InterPro:IPR004104), NAD(P)-binding domain (InterPro:IPR016040); BEST Arabidopsis thaliana protein match is: Glycerinaldehyde-3-phosphate dehydrogenase-like family protein (TAIR:AT1G34200.1); Has 17148 Blast hits to 17147 proteins in 2160 species: Archaea - 344; Bacteria - 12203; Metazoa - 250; Fungi - 605; Plants - 99; Viruses - 0; Other Eukaryotes - 3647 (source: NCBI BLink).</p>	aa seq
231	sp_121650_zwed #N/A	#N/A	#N/A		MENRRKVKVHAIFTLADDYFLKNDIRVKPDLDAL GALGDVGCYRIRALWAANYELPKFVTLRNP NDDGVLSCDAALHWEDGKIATFCSFLSLTM DITVIGTNGTLHYNDFLIPYKASEVSFKTASKWF TDKALEWTHKSEQVSTDLPQEVLMVKEFSG VDGKIFRDLAPEMKWPTISRKTQLVIDAVKVSIDK GLOAVEVVY
232	sp_123780_pgiv #N/A	#N/A	#N/A	<p>O-acyltransferase (WSD1-like) family protein; CONTAINS InterPro DOMAIN/s: O-acyltransferase, WSD1, C-terminal (InterPro:IPR009721), O- acyltransferase, WSD1, N-terminal (InterPro:IPR004255); BEST Arabidopsis thaliana protein match is: O-acyltransferase (WSD1-like) family protein (TAIR:AT2G38995.2); Has 1220 Blast hits to 1207 proteins in 180 species: Archaea - 4; Bacteria - 957; Metazoa - 16; Fungi - 0; Plants - 224; Viruses - 0; Other Eukaryotes - 19 (source: NCBI BLink).</p>	MTPNLQIVTNGGKPEDEAEPVSPSTGQYFNSKV LSVCVLAILEIDVPIDDSCVIPQLRDVFLPMNPRFS SIMISDNKDYKQWKRVEVNLQDHVVVPSVPDGL SVESDYKYFDEYLTKITVDPLPQDRPLWELHVIKY PTSKAAGHFHWKLLHHALGDGYTLMGVLLSGVNR ADDP/SLPLTFPSTRSSSLVTNKMNIISWVPRTF SAIYNGVYVYFVWVSLKSTCKADDKTPIRSGNEGL GFHPMKSTIELSLDQIKFIKTKLGATVNDILAGIIFL GVRKYMQATDTEQGNSESTALVLFNTRNIGYIM TAEQMKKAGMKIWNQGFALHIAIPLQINDKCSN PLDYVYEARQKISRKSSPSSVYLTAGCCLLRKC KGPEAAAEFIQSSTTNKASILMTNMIIGPIERTVMD HPVKGMHFVMTGSPQDRSRD

10

20

30

40

【表 1 4 - 9 6】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 Oxidoreductase family protein; FUNCTIONS IN: oxidoreductase activity, binding, catalytic activity; INVOLVED IN: oxidation reduction, metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Oxidoreductase, N-terminal (InterPro:IPR000683), Oxidoreductase, C-terminal (InterPro:IPR004104), NAD(P)-binding domain (InterPro:IPR016040); BEST Arabidopsis thaliana protein match is: Glyceroldehyde-3-phosphate dehydrogenase-like family protein (TAIR:AT1G34200.1); Has 17148 Blast hits to 17147 proteins in 2160 species: Archaea - 344; Bacteria - 12203; Metazoa - 250; Fungi - 605; Plants - 99; Viruses - 0; Other Eukaryotes - 3647 (source: NCBI BLink).	aa seq
233	sp_127590_adcq #N/A	#N/A	0.938873	uncharacterized protein	MAESGSNAIRFGILGCAEISRKVARAISLSPNSPP HAIGRSRLEKAQTFATTNSDFPSTKYVYGSYDEVI DDPEVDVAVYVPLPSTLHLOWVRAAEKKKHVIVE KPPALDVKELDVMNLNACKANGVQYMDGTMWMIH HPRTKIMTDFLADSORFGHLKA MLAASAVIDAKYEARTTASIMTEILLKAEVKAMSS LREVERAVASKLKNVMMTEDKCSPSGNICSGF GPPEQCCSGACVPHPIRIFVCG
234	sp_129340_kdia #N/A	#N/A	0.933815	Ankyrin repeat family protein; CONTAINS InterPro DOMAIN/s: Ankyrin repeat-containing domain (InterPro:IPR020683), Ankyrin repeat (InterPro:IPR002110); BEST Arabidopsis thaliana protein match is: Ankyrin repeat family protein (TAIR:AT5G54620.1); Has 21854 Blast hits to 10883 proteins in 482 species: Archaea - 19; Bacteria - 1460; Metazoa - 11922; Fungi - 1177; Plants - 2352; Viruses - 52; Other Eukaryotes - 4872 (source: NCBI BLink).	MEKERDKSEVTVLYEAALTGDVPSMLNLLQ DQFILERGIEKSCHEMQSPLHVAVNMGHIQFVRE ILRRKPOFAEVVDRLRRSSPLHIACAIKGRLEVINI LVEEKPNMCFADQDGRTPIHVAVVNGNVEVME VLLKEKHGAAWERTMGGETVHLCKVKNQLDAL KSLVNTDDPELLNFADSDNGDTALHLAVAYKESE MVECLLAQNKIAKDVRIKKEKTIVDIHKETNKGSE ADKIDRLLKYYKAPAKYGLKRQTKENKWLLEEQ NTALMVAASIAATMAFQYGMNPPGGVWQDNIN GHEAGTSIMSYSYSDGADYVNLVSNITIGLISSLSV ILLISGLPCKRFFVFLRITLWISVSAATAATYLLAI GYVTPNVINKASLVNDAMDYSVEVWLWLVAVILL GHGVILFKLIGHDRRQKIREFYKTHFKVPEAVP KRVLESQITCPV
235	sp_131200_inri #N/A	#N/A	0.938873		

10

20

30

40

【表 1 4 - 9 7】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_ktdg (SOAP2)	sp_107670_ptqx (SOAP2-like)	<p>注釈</p> <p>C2H2 and C2HC zinc fingers superfamily protein; FUNCTIONS IN: sequence-specific DNA binding transcription factor activity; INVOLVED IN: response to chitin, regulation of transcription; LOCATED IN: intracellular; EXPRESSED IN: root; CONTAINS InterPro DOMAIN/s: Zinc finger, C2H2-like (InterPro:IPR015880), Zinc finger, C2H2-type (InterPro:IPR007087); BEST Arabidopsis thaliana protein match is: C2H2 and C2HC zinc fingers superfamily protein (TAIR:AT1G34370.1); Has 32454 Blast hits to 15503 proteins in 208 species: Archae - 0; Bacteria - 0; Metazoa - 30756; Fungi - 38; Plants - 696; Viruses - 0; Other Eukaryotes - 964 (source: NCBI BLINK).</p>	aa seq
236	sp_134450_jzcs	#N/A	#N/A	0.923734	<p>LRR and NB-ARC domains-containing disease resistance protein; FUNCTIONS IN: ATP binding; INVOLVED IN: apoptosis, defense response; CONTAINS InterPro DOMAIN/s: NB-ARC (InterPro:IPR002182), Leucine-rich repeat (InterPro:IPR001611), Disease resistance protein (InterPro:IPR00767); BEST Arabidopsis thaliana protein match is: NB-ARC domain-containing disease resistance protein (TAIR:AT3G14470.1); Has 38623 Blast hits to 22788 proteins in 858 species: Archae - 54; Bacteria - 2838; Metazoa - 8237; Fungi - 658; Plants - 25379; Viruses - 2; Other Eukaryotes - 1455 (source: NCBI BLINK).</p>	<p>MPLNLRGELKIQFRESYSYDATNAEEGSYLLSK KHIKALVIKAKYETPIVDERLLEVLQPHHDLIKI YGYKSIRLPGWVAIAAANLPRVLCLELRFERLQH LPQLSQLHLKILQLDNMPGVEYEMDDKGNIPCF PLLEELIWNFPKLGWWRPEVGVETDGVIPGL HKLTFGGPNLTSFPGIPHVKEKFRANKFNPYW</p>
238	sp_155920_gjdz	#N/A	#N/A	0.9953747	<p>uncharacterized protein</p>	<p>MAKFVYVILKVIIGSIL-SNPIESLCSKLPKPCIESI VSFNINQSIISIIETGFNEGALQMEANLLDQCC GRFTVETISRDKQCFCSSISIQEHLRSNNITILLS ECSTSDGMLLFYDDDLCCDIQAPSSAPVFLPFSY VLSHTQLSPTSTPIPTPRNKIIIAIIVTVTLW LIALCIGFFSMKPKKEKPNVAQSHHQSGKTEK</p>

10

20

30

40

【表 1 4 - 9 8】

A	B	C	D	E	F
1	分析に用いたバイトとの相関係数	sp_107660_kicqg (SOAP2)	sp_107670_ptqtx (SOAP2-like)	注釈	aa seq
2	共発現遺伝子			<p>NAD(P)-binding Rossmann-fold superfamily protein; FUNCTIONS IN: oxidoreductase activity, binding, catalytic activity; INVOLVED IN: oxidation reduction, metabolic process; EXPRESSED IN: 16 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: Short-chain dehydrogenase/reductase, conserved site (InterPro:IPR020904), NAD(P)-binding domain (InterPro:IPR016040), Glucose/ribitol dehydrogenase (InterPro:IPR002347), Short-chain dehydrogenase/reductase SDR (InterPro:IPR002198); BEST Arabidopsis thaliana protein match is: senescence-associated gene 13 (TAIR:AT2G29350.1); Has 121057 Blast hits to 120745 proteins in 3594 species: Archae - 991; Bacteria - 79410; Metazoa - 5524; Fungi - 6125; Plants - 2737; Viruses - 5; Other Eukaryotes - 26265 (source: NCBI BLINK).</p>	
239	sp_156640_yxdd	#N/A	#N/A		

10

20

30

40

【表 1 4 - 9 9】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh   sp_107660_kiqg (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MFFLGS DILTYLYAKQIKKVPDTFSSSTDHLYLSSFK HLLKEEIRAEFSAGVESIAQSPACEISKYRISKHHK PPKSLYYEIMKDIRVEKNESEHNYMLQSGDLIVL TDMRPRMEDLIREKDNLF.LFAFVASSNVDEPWI QIFSSKEFEPKLTKKHGRFFAFILMNMTTSFRI WSALNPDPKSSSMGLIQKVLQYDSTVDDDCSLC VNEKPFNVKGLNVFDLJSTLGLDESOQKAVLNTIS MRNCSHQSNVNLWGPFGTGKTKTVASSLFVL LKLKCRTLTCAPTNIAVQVANRLVRIFLENGEL DTYGLDVLFGHEERMKINDHDELDMDVFLDYRK EVLGECLEPLHGWKGTLESMSKSFLEDPQTRYND YLQQHEVEDEENGSKRIDGVKKNKYSMDMLKT EDLMIFEEFFKRFVSLADRLVYCKLKNMYTHLPTT SLPLNVAKQIMIRLVHVLKTTTNRKVKVSDHKNLM MKREEILCILNYLSDQFPKLFKGSVHEFCLSNAR LLFTASASIKYSKNVEWIDEAAQLKECESMPL QVPGLNVAVLIGDDRQLPAMVQSKALGDLNFRG SLFQRLAKIGKIKKHLNIIQYRMHPSISLFPNKEFY ENRIIDAPKVEEISYRSFLKGAMYGSYFNFVSR GKEDFEKGHSRPNIEEAVVDHIVAKLFGGLDFK KOKLTVGVISPYKGGVCLIQEKIGNKYANYKENFA VNVRSVDGFGGGEEEDVHISTVRCNGNGSVGFSL NHORTNVALTRARYCLWIVGNWTTLAKSGSWW KKVAIDAKTRGCFYDADDDLLHSHKQAKTGS�AKIN QLSVSKPDPFFGSLNLGKAKWKVFFADAFKSSITSI SSKKCKNAQKLLKKIADGWRLOSNTEAEKCV ATCGIAPELLELYKVKEQLYLWVSDVIQEEESKYT QAIKVVWDLVLPASKIPKLAKELDILFSKYTSEFIMNR CKYKSIQEGDLVLPNSWQLHLSVLSAENLSDQFSV MNIIRDRQESSTGSLRDNMKGAGSGISSSSRRRRR RAGTA MESLRSFLCFKTSQSSVVFPTTTTTFTKLPSPRF WHKAKAGYAVFGAKQAGKRGKGLVLSLQKDELCE GYADIETMWMKMIHSTCSKDIHSDGNQKRPRYLVF CFKPS
240	sp_159110_dayj	#N/A	#N/A	0.242818	P-loop containing nucleoside triphosphate hydrolases superfamily protein; BEST Arabidopsis thaliana protein match is: P-loop containing nucleoside triphosphate hydrolases superfamily protein (TAIR:AT1G65780.1); Has 6257 Blast hits to 5371 proteins in 1056 species: Archae - 197; Bacteria - 1787; Metazoa - 1336; Fungi - 1200; Plants - 672; Viruses - 106; Other Eukaryotes - 959 (source: NCBI BLink).	
241	sp_151290_shyy	#N/A	#N/A	0.948223	uncharacterized protein	

10

20

30

40

【表 1 4 - 1 0 0】

I	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
242	sp_170040_wkpp	#N/A	#N/A	3.323332	Major facilitator superfamily protein; FUNCTIONS IN: transporter activity; INVOLVED IN: oligopeptide transport; LOCATED IN: plasma membrane, membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Oligopeptide transporter (InterPro:IPR000109), Major facilitator superfamily, general substrate transporter (InterPro:IPR016196); BEST Arabidopsis thaliana protein match is: Major facilitator superfamily protein (TAIR:AT1G52190.1); Has 5584 Blast hits to 5329 proteins in 963 species: Archae - 0; Bacteria - 2098; Metazoa - 610; Fungi - 414; Plants - 2170; Viruses - 0; Other Eukaryotes - 292 (source: NCBI BLINK).	MAIEEEKMAFLLDHQNSEHENPKPRTKGGIITM PFIIVNEAFEKVSYSYGLLPNMIIFYLMNVYHLEAAK GSIIILWSASSNAMAIFGGFLSDSYFGRFRVIAF GTFCSSLGTTMIWSTSMIPWMKSPCAMSSQDC DSPTSLHYLVLFSGFGLISVAGCVRPCSLAFGA DQVNDKENPNERNVLDSEFNWYYASTGLSTVIAL TFVYIQDKFGVQIGFVGVPLMVISLATLFGSPL YVHVKARDNLFIGLQVPAFAAFQRRKISFRAGAE VRYCYGNGSEATTPSDHIRCLNFRACVINPQDF NPDGSASNPWTCTVEOVERLKSIVRILPMWTT GFMMILVALNNSYTTLQAKTLDRHFSSSQIPAG TFTVFSIVTITFWAFYDRAVVP_LMAKYFNMPRG LGPFVRMIGILLSCVSLLVAGVSVESRRKVAIDS GLEDKPNEVLEMSAMWLVPLV_LGLAEGFNVAV GQIEFYAHLYKGMVAVSSLYGSLLVQVVDVDFSSV GGKYSWLSSNLNKGHDVYDYYWLLAFCLVNFLY FLCCRMVYGTSESHSRFRRYESVEHEESA
243	sp_170860_dkxg	#N/A	#N/A	3.332144	Leucine-rich repeat protein kinase family protein; FUNCTIONS IN: protein serine/threonine kinase activity, kinase activity, ATP binding; INVOLVED IN: transmembrane receptor protein tyrosine kinase signaling pathway, protein amino acid phosphorylation; LOCATED IN: plasma membrane; EXPRESSED IN: 13 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Leucine-rich repeat-containing N-terminal domain, type 2 (InterPro:IPR013210), Leucine-rich repeat (InterPro:IPR001611), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: Leucine-rich repeat protein kinase family protein (TAIR:AT3G47580.1); Has 213556 Blast hits to 133070 proteins in 4776 species: Archae - 194; Bacteria - 22980; Metazoa - 68117; Fungi - 10236; Plants - 86850; Viruses - 328; Other Eukaryotes - 24851 (source: NCBI BLINK).	MEISNQVSLVKLDVSSNKFSGEIPDVFNNLPTLQ KLYMAHNFHGLIPQSFASLKSLLLEDLSHNRLS GPVPRYLSFFPLIYNLNVNFECEVPTKGFIFANA SAISLVGNRLCGGISELHARCVRKAPKTRMY YTLKLVIFIHITIGILATGSCWLVLRHRYKREPVS LDAIGKFSKVSYHMLIKATNGFSMGNLLGSGT FGSVFQGNLDGKTVAVKVNLQQRGGNKSFWMA ECEAMRNIRHNLVGIITAGLAQVWISKEIKSKLL SHALNYLHHECGNSVHCDLKPNNILLHDMVAR VSDFGAKVLRPLHPNCGSSVAVRGTIGYVAPE YGLGAESPADYISYGIILLLELMTSKRPTDNMF MEENLHMYSKDALHDRVLVKIVDSTLLEHENEEVE SSEIKTQVVLQNRRECITLVKIGVACSNQLPRDR MKISDVIIELQEARKILLISKQRFQRTFTITQM

10

20

30

40

【 1 1 0 0】

【表 1 4 - 1 0 1】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_ktgg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2					MRLSMKSLVLPYDLPMSSTPYVPVPGHQ QCSITKRTQELTFGSQQSINRQSLNQTCLTPVWR QSRANRWKEYMGVRDWDGLDPLDLRDEI LRYGSEIATYRGFEFNPSSSSYKSKKXFF EQGLPTPGYRLRHLRATSGIQVPGWASSWS PMQTSWIGYVAVCTNKAMIKQLGRREVWIALFGT ATCLEWLENLRATLAPVPGISVENDVGPWNV DSPPMVESGFLSYTSGNDWMPSLQDAIRIEIGRI LDLYQDKEPLSITITGSLGAALATLAAYDIKTTFD RAPLVTVISFGGPRVGNQSFSLVDKQGTILRV VNFSDLITKLPGFVTDNEDFVIDTSPATKKIGI KNSPRTSHNNKNNKNNRNPRTTTSVSDSTRF GKILSHFGKGIKSSSKSRSLGVTKSGSSKMKLM GNILMNWVQDFRVDTRWVYVYAEVGRLELLEN RESAEQLQYYLSNMDVAKCHDLKTYLNLVEDC PLKKNFVRRSLNKLQRCRTRNAHEKDNINNSG VLAL MEDGFNRNVNDFVT
244	sp_172940_kgac	#N/A	0.3323243	DEFECTIVE ANTER DEHISCENCE 1 (DAD1); CONTAINS InterPro DOMAIN/s: Lipase, class 3 (InterPro:IPR002921); BEST Arabidopsis thaliana protein match is: alpha/beta-hydrolases superfamily protein (TAIR:AT4G16820.1); Has 1715 Blast hits to 1709 proteins in 324 species: Archae - 0; Bacteria - 331; Metazoa - 67; Fungi - 376; Plants - 633; Viruses - 9; Other Eukaryotes - 299 (source: NCBI BLINK).	
245	sp_173100_ziaf	#N/A	0.3246318	uncharacterized protein	
				Leucine-rich repeat transmembrane protein kinase; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: endomembrane system; EXPRESSED DURING: 15 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Serine/threonine-protein kinase domain (InterPro:IPR022290), Serine- threonine/tyrosine-protein kinase (InterPro:IPR001245), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271), Protein kinase, catalytic domain (InterPro:IPR000719), Malecin/receptor-like protein kinase (InterPro:IPR021720), Tyrosine-protein kinase, catalytic domain (InterPro:IPR020635); BEST Arabidopsis thaliana protein match is: Leucine-rich repeat transmembrane protein kinase (TAIR:AT1G29740.1); Has 108967 Blast hits to 107923 proteins in 3166 species: Archae - 87; Bacteria - 12310; Metazoa - 39616; Fungi - 9390; Plants - 31491; Viruses - 414; Other Eukaryotes - 15659 (source: NCBI BLINK).	MELVDPMLGTDYNTTEQVLRMIKVALLCITNPSPAV RPVMSEVLRMLEGQMLVKELTLPGLCNNDWFY EASKRYQDSTSDSETQTLRFPNARNNANSLSDL SPIPTNHPKPPVIEPVQT
246	sp_176550_tgfx	#N/A	0.324244		

10

20

30

40

【表 1 4 - 1 0 2】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kfkg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
247	sp_183980_pscg #N/A	#N/A	0.657374	NAD(P)-binding Rossmann-fold superfamily protein; FUNCTIONS IN: oxidoreductase activity, binding, catalytic activity; INVOLVED IN: oxidation reduction, metabolic process; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: NAD(P)-binding domain (InterPro:IPR016040), Glucose/fructol dehydrogenase (InterPro:IPR002347), Short-chain dehydrogenase/reductase SDR (InterPro:IPR002198); BEST Arabidopsis thaliana protein match is: NAD(P)-binding Rossmann-fold superfamily protein (TAIR:AT5G51030.1); Has 50499 Blast hits to 50443 proteins in 2986 species: Archae - 509; Bacteria - 33822; Metazoa - 3124; Fungi - 3071; Plants - 1791; Viruses - 1; Other Eukaryotes - 8181 (source: NCBI BLINK).	MMKMKQFNLIISRKIPTLFYSTASFNRASVSLV QGASRGIGLEFVKQLVDRDKKGVHLATCRNPDQ STGLELNKQFSDRLNLLKIDVTIDENTIEAATTSV RETYGYNLLVNTSGILSIPNLLHPETTLAKVQKS ALMLTYEVNAVGPLLMIKHMWPLLKEGGKGTM RDAAVVANLSARVSGSDNLLGGVHWSRASKTA LNQLTKTVSVELARKKDPVICLLHPGTVDLDSR PFQRNVPODKLFTREFSVSKLLSIIDGAKIHDNGK FFAWDGOEIPW
248	sp_187280_gxza #N/A	#N/A	0.633431	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G17580.1); Has 40 Blast hits to 40 proteins in 11 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 40; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).	MAASWVKTTTSPFKKACTFVKHQPPKDKKSKQE NENTVMALHGEVMACTYEDVQVMWMSMLDKSKP GACNMTS MGDLTKWVSNISSEALELTPAVSSYCCYRYPEN LRKVNEDAYRPLLVSIGFSYYPGDPRLVAMHQKLL RYLQSFQRYYTCYNLDYYVQVIREWEVWARQCY VESLDTTRLTSPQFIEMLLYDAAFLIELFLRNHFHE FIEENDRIFNKPRMMLLEVTRDIRLEENQLPFFILKQ LYDLVQPSQNKQVSLFDITYKFTLTKNEAVPAR IIDADVKHLVDILRLCYLPCSVRGOHPRTPNPKFEF ASSVSELSEAGVKFVASPSENLLDIRFVKGVLEV PREVYVDDTDESLEFRNIMVFEQCHYFDOSYIDYFA FMDSLINTPKDVGILVQSGVQIENWLGNEEVANL FSNIFKQTRLKGSDFYVSGICKDLNAYAGTPWNR WKALKIHDYFNHPWAAISVYVAVMLLTLVQVVT GRK
249	sp_188860_zcgg #N/A	#N/A	0.634132	unknown protein; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF247, plant (InterPro:IPR004158), Protein of unknown function DUF862, eukaryotic (InterPro:IPR008580); BEST Arabidopsis thaliana protein match is: Plant protein of unknown function (DUF247) (TAIR:AT5G11290.1); Has 1967 Blast hits to 1844 proteins in 183 species: Archae - 6; Bacteria - 6; Metazoa - 223; Fungi - 83; Plants - 1477; Viruses - 0; Other Eukaryotes - 178 (source: NCBI BLINK).	

10

20

30

40

【表 1 4 - 1 0 3】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈  2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein; FUNCTIONS IN: oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, 2-oxoglutarate as one donor, and incorporation of one atom each of oxygen into both donors. oxidoreductase activity; INVOLVED IN: biosynthetic process; LOCATED IN: cellular component unknown; EXPRESSED IN: hypocoely, root; CONTAINS InterPro DOMAIN/s: Oxoglutarate/iron-dependent oxygenase (InterPro:IPR005123); BEST Arabidopsis thaliana protein match is: 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein (TAIR:AT3G60290.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archaea - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	aa seq
250	#N/A	#N/A	3333332	Plant protein of unknown function (DUF936); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: culture cell; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF936; Plant (InterPro:IPR010341); BEST Arabidopsis thaliana protein match is: Plant protein of unknown function (DUF936) (TAIR:AT3G14170.1); Has 718 Blast hits to 418 proteins in 99 species: Archaea - 0; Bacteria - 149; Metazoa - 50; Fungi - 25; Plants - 280; Viruses - 0; Other Eukaryotes - 214 (source: NCBI BLINK).	
251	sp_192960_cgkm	#N/A	3333328		

10

20

30

40

【表 1 4 - 1 0 4】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
252	sp_193940_axoc #N/A	#N/A	0.1273931	FUNCTIONS IN: molecular_function unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Autophagy-related protein 27 (InterPro:IPR018939); Has 138 Blast hits to 138 proteins in 57 species: Archae - 0; Bacteria - 0; Metazoa - 32; Fungi - 62; Plants - 33; Viruses - 0; Other Eukaryotes - 11 (source: NCBI BLINK).	MGSHYSSKLTLLSILVGNFSSVFAATCEYGT RERNKLYSLSLTPVKNFRHRHGLVSEDFGYKVA NDTIWIFQLCDGMIFNDPPRCADCWCDCGGPSR CGMGSALVAKYTGYYACTTIGREINFDINLTD RKTPDKGVTITMSSYTDGVCNLSVSVICDSKGIK EPESLKLKLGACHYATELRHPSPGCPKIAVHATGW GWFRLLTLILCSVGAYLLIGGVYRFFHLGIRGLDV IPNLDWISLPLYSIKSVFMSLVQRFRGPTYYHTSS YSSVNF
253	sp_195460_jgcs #N/A	#N/A	0.541838	uncharacterized protein	MYSSQVKPFGGLKINATDLGAPVPAATKSVCCG S
254	sp_209810_wuzr #N/A	#N/A	0.1277437	beta glucosidase 11 (BGLU11); FUNCTIONS IN: cation binding, hydrolase activity, hydrolyzing O-glycosyl compounds, catalytic activity; INVOLVED IN: carbohydrate metabolic process; LOCATED IN: endomembrane system; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Glycoside hydrolase, family 1 (InterPro:IPR001360), Glycoside hydrolase, family 1, active site (InterPro:IPR018120), Glycoside hydrolase, catalytic core (InterPro:IPR017853), Glycoside hydrolase, subgroup, catalytic core (InterPro:IPR013781); BEST Arabidopsis thaliana protein match is: beta glucosidase 10 (TAIR:AT4G27830.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	IEPHVTLHIDVPOALVDDYESFLNRRIEDFTAYV DVCREFGDRVRYWTTLNEPFIYVGGYDIGNYP PGRCSPPFGNCSRGNSETPEYIAHLLLAHGS VKLYKEYQAKQLGFGVMNLLSYHFSPITNTRENI IATRRTYDFLIGWFMHPLTYGDYPEIMKKNVGSRI PVFTEKESNLVKGSFDFIGINYYSIMETKDYSTGL SKYPRDRFFADIGVKWT CMI
255	sp_211820_gkmm #N/A	#N/A	0.544334	uncharacterized protein	MGLONNSRKEGPEQETNFSENLWKYFAA
256	sp_212470_xagd #N/A	#N/A	0.5951177	BCL-2-associated athanogene 3 (BAG3); INVOLVED IN: regulation of apoptosis, apoptosis; LOCATED IN: cellular_component unknown; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: Apoptosis regulator, Bcl-2 protein, BAG (InterPro:IPR003103), Ubiquitin supergroup (InterPro:IPR019955); BEST Arabidopsis thaliana protein match is: BCL-2-associated athanogene 2 (TAIR:AT5G62100.2); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MGGLEWEMRPPGGFMVQKRNLDNFRNYSIPIPTI KVLVKNSSYHEIHINSHASFGELKRLRAQVTGLH PQDKLIFKEKERESRFLDYSGVKDGSRIILVOD INCNRFIESRRNSNIERNAMKSVAEYSIEVDKLAG QVTSMESVYKGGRRVADKELLTIELLMAQMIKLD GIADGDVKLRKTKQVVRIOKYIDSIDRMKQNKV ASRYGGGQVGLVWSQQQQQRPFIHQQQQRECKNF SNGMLTTSKWEKY

10

20

30

40

【 1 1 0 4】

【表 1 4 - 1 0 5】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
257	sp_002450_dsmi	0.832284	0.92306	#N/A	FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; CONTAINS InterPro DOMAINs: Metal-dependent phosphohydrolase, HD subdomain (InterPro:IPR006674), Ppx/GppA phosphatase (InterPro:IPR003695).	MVIPLSQNLFASVDMGNTNFKLIIHLEPSSAKFT TIHNSDAVALGRNSSPSSSSPPLISPDQSF RAIESLKFOEILRSHNVTRTRMVGTSAIRAENR SIFLNEVRENLGDFHVTYVLSGVEEARLYLVQLQ FLHVYDKRVLTDIGGGSTEFVIGDHGEWYADSL RLGHLTLTONFGESDLYGLRGYVRELEGSLVD KVKILGFEVAVGSSGTTRAVENAIWCSYGDDRE KGGMRVLEEGSRRLWVFSREELRDLVERLCT VFESEKRVRRDEFFRRSEFIVAGAVLLEEIMEL LGISEMQVSKYALGEGVVAEMVSGEFPEFDLNA NVRWKSIMRFLSMRFNGNDRMRNGVQCSAIKDI FEGRLKYTKVFEDHFTDDVCLDEKDFEYLAACM LHNIGLYLGGKGYHKNSYIVMEHSOLDGYTSEEI QLIAQLARHHRKKFPNFEYTSLOGFFPREKFGILCS ILRLSVIQKHKLKLEIKFDDSESVLVLVKKGD QPILLDSKSLAKDIETKTRKEAEFFEGVHFHQKLV VVMEDGSESSARQ
258	sp_033270_jhzc	0.905510	0.952288	#N/A	REVERSAL OF THE DET PHENOTYPE 3 (TED3); FUNCTIONS IN: zinc ion binding; INVOLVED IN: protein import into peroxisome matrix, fatty acid beta-oxidation, peroxisome organization, photomorphogenesis; LOCATED IN: cytosol, peroxisome; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAINs: Zinc finger, RING-type, conserved site (InterPro:IPR017907), Pex N-terminal (InterPro:IPR006845), Zinc finger, RING-type (InterPro:IPR001841); Has 30201 Blast hits to 17322 proteins in 780 species: Archaea - 12; Bacteria - 1396; Metazoa - 17336; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MDRQTLASSSSPSSISEVSPDAWITTYKTLPL QWNSPFLSHQIPIITISRVNCFDAARLDVEMSAIML KEQLVKVFLMKPGILFOYEPEDAFLEFLIWRFS IWDKPTFGNALMNLRYRDERVMEYRKGVRTGL EGPGLTAAQKMWYCIATVGGQYVTRLQSFSAF RRWGDSEORTLQRAWLLLOHVEGFYKAAFS NLLLFLLTGKYRSLIERALRALRYVGSFNMNRVSV FEYMRQLVWNEFESEMLLLPLINSPLKSLFR PFSKDKASESASNESSPCICQSGSPSIFIALPCRH RYCYGLRTRCAATPSFRCARCSQPVTAMQRLG GSVDEKSSKK
259	sp_039080_ztzh	0.931236	0.943488	#N/A	cyclin-related; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; Has 5668 Blast hits to 3895 proteins in 309 species: Archaea - 0; Bacteria - 250; Metazoa - 3167; Fungi - 631; Plants - 477; Viruses - 7; Other Eukaryotes - 1136 (source: NCBI BLINK).	MSDSKGVPRERSPLKSEVGDKDGDKKMHM QLQDALLESPLASDSKQEVKASSMLDSKNG KGETRGHSNTTAVPRRSYHQHDDRRRTAARDGS RSGRRSTSEGWKDSRGGEDFAAHRSAPSADP KYVEGKSRADKRGWIDSRGDRASHRSGPS DPKVRREGKSGAEGGDSQVWRHDFRIFAEAGDAQ PSKKRRPFREEKLPVKSEDDVNAVVPERSKVSH PTGCTARWEQKAHDRRLERNEHDRAYLRERS FSNIRGVENTGFPSPDRFNGVSNNGKFIGRENF GGRKGYHSSGGVKEWKWKHDLYDEANRSPTPKK EEDPVAKYEALLSL

10

20

30

40

【 1 1 0 5】

【表 1 4 - 1 0 6】

A	B	C	D	E	F
1	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiag (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2				Inositol 1,3,4-trisphosphate 5/6-kinase family protein; FUNCTIONS IN: inositol or phosphatidylinositol kinase activity, catalytic activity; INVOLVED IN: inositol and derivative phosphorylation; LOCATED IN: intracellular; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAINs: Inositol-tetrakisphosphate 1-kinase (InterPro:IPR017427), ATP-grasp fold (InterPro:IPR011761), Inositol 1,3,4-trisphosphate 56-kinase (InterPro:IPR008656); BEST Arabidopsis thaliana protein match is: Inositol 1,3,4-trisphosphate 5/6-kinase family protein (TAIR:AT4G08170.2); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MAEQNNEIISVGYALPSKKYDVSFIKPSLVSVAKTR GIDFIPIDQPLLOQGPFFHCVIHKVYGGDDWNREL LDFKARNPNVVIDSPAAIQRLLHNRISMLDVKNL KPASSDAALTPSVGTPNQILLSEEIMMRSLSHQ MSVVFNNGLKVKREFAPIVLQEFVNHGGVLFKI YVAGENVRCVQRKSLPFDNDSDDHPFFLQFEFEEG FMPFAQISNLATTEMQOEIDNIEMPPLODFVNNLAK GLREGTINLNFNLDLRDSSVSGGNRYLIDINYYFP GFAKLPGEEMLTDFFLDVIWHFSSSAVEKA
260	sp_052000_anyg	0.945173	#N/A		MPCSRKQLKLLYNVTATKFFLFFLYLFLSFLIAK LFTYFRTEEEYEDDISYEDDEDEDEEEEAASE EEEEERYSHMEEEEEGDHLVADIVNDGEG LIFVANYGDDDYERNNDHSIVHSNSYETLFDQPV EDFNDEAEGCFENSALTSSSELFDRGGDYTH HGGDYGGDNYDINEDYASRFESNHLYDRSFEI EEEEIHKNDHENPHQCEAVYHDMQPNSPSYI EVKVVAPSGNHDVAGEFDESKTDVVKQDSNFT RDEKFLIFAPPKSEAKLTLGLEMENEGIFEDTYT GSTSKSSEWRSSIRDSGTDPPFSSSSRRSCP WESVAVYQKYDEEMTFDRISAQKLHETESLRSI QVPRRSQRIVHKLGNKKEKSSGGDDIYKNTN PYNELEVAVYVAQICLITWALNWNVYETFRRTKQ KSKTELODDDHGCPANIAQQFQFQVLVQRYIE NEFYQRGKRPEIYARLSAAPNLLQVPEYHDYE GEEKELVGGRIASALFLOILEDISIQTFMNFLLKADR ENHCQVITALKRPFPGSIDPTVLNVRKCKNKKKK LKLKEVERTRKCFKRRRLTREEEMEILMSLIDLK VSRVLRMTEITQEQLHWCFOKMAKPTILEGLKLR DSSPLFFPAH
261	sp_075130_figr	0.928387	#N/A	FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation; LOCATED IN: endomembrane system, ribosome, intracellular; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: 4 anthesis, petal differentiation and expansion stage, D bilateral stage; CONTAINS InterPro DOMAINs: Ribosomal protein L34e (InterPro:IPR008195), Protein of unknown function DJF1666 (InterPro:IPR012870); BEST Arabidopsis thaliana protein match is: Protein of unknown function (DJF1666) (TAIR:AT3G20260.1); Has 1252 Blast hits to 1091 proteins in 184 species: Archae - 4; Bacteria - 49; Metazoa - 367; Fungi - 151; Plants - 117; Viruses - 41; Other Eukaryotes - 523 (source: NCBI BLINK).	

10

20

30

40

【表 1 4 - 1 0 7】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
				hydroxyproline-rich glycoprotein family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: stem; BEST Arabidopsis thaliana protein match is: hydroxyproline-rich glycoprotein family protein (TAIR:AT3G26910.1); Has 1387 Blast hits to 1097 proteins in 177 species: Archae - 4; Bacteria - 37; Metazoa - 593; Fungi - 281; Plants - 163; Viruses - 19; Other Eukaryotes - 290 (source: NCBI BLink).	MKLKDFFERNVKDSHEDSFIREKRIIGSQAPLFLQ KFKPAEKIRLOQTSSSRKFKHSYALPKPGDTRISA RSENRVLSTKPNPLLAENYHSPSLVPNTYEKLV TEMHNLHA
262	sp_077670_jsex	G.944376 0.823694	#N/A	Mannose-binding lectin superfamily protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; CONTAINS InterPro DOMAIN/s: Mannose-binding lectin (InterPro:IPR001229); BEST Arabidopsis thaliana protein match is: Mannose-binding lectin superfamily protein (TAIR:AT5G38550.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).	MAEKPATSNFVEGTKNAAGKAVTAVTTQTEKISK SAGEMSEKTKNSVGDALNKAGEFVKPAPPPPEE SLMDKAKGLFK
263	sp_080320_atmm	G.928251 0.835044	#N/A	galacturonosyltransferase-like 2 (GATL2); FUNCTIONS IN: transferase activity, transferring hexosyl groups, polygalacturonate 4-alpha- galacturonosyltransferase activity, transferase activity, transferring glycosyl groups; INVOLVED IN: carbohydrate biosynthetic process; LOCATED IN: endomembrane system; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: Glycosyl transferase, family 8 (InterPro:IPR002495); BEST Arabidopsis thaliana protein match is: Nucleotide-diphospho-sugar transferases superfamily protein (TAIR:AT1G19300.1); Has 1885 Blast hits to 1878 proteins in 434 species: Archae - 2; Bacteria - 935; Metazoa - 151; Fungi - 0; Plants - 716; Viruses - 0; Other Eukaryotes - 81 (source: NCBI BLink).	MPPTRPFLVGVVWVFLITIVSATATATIDGSSQ KQRFREAPQFYNSEECPKVSTEFEDDDHQDYSL PSSALDPDSVLCVCSDEAVHVMTLDLTYIRGSM AIIQSVLQSSCPQNIYHFVISATFRSNASELHTTIS SSFPYLRFDLYPFDDSAVAGLSTIRSALDCPLN YARPYLANLLPLCVGRVYLDSDVWVDDISTLY NFNLVDDKVLAAPEYCNANFTTYFTFWSNPFL SVTFADRACYFNTGMVID.NRWRAGDYTRRI EEWMELOKRMRYELGSLPPFLLVFGGNIPVDH RWNOHGLGGDNFRGLCRDLHPGPVSLLLHWISGK GKPPWARLDANRPPCLDALWAPYDILLQTPFALDS
264	sp_103930_jfx	G.944376 0.823694	#N/A		

10

20

30

40

【表 1 4 - 1 0 8】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
265	sp_112680_fwka	0.3410773	0.9448332	#N/A	Carbohydrate-binding X8 domain superfamily protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: X8 (InterPro:IPR012946); BEST Arabidopsis thaliana protein match is: Carbohydrate-binding X8 domain superfamily protein (TAIR:AT2G04910.2); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MAWNLLPLLLLLLLLLLVVNPISGDDNDGEDDDPG QQWCVADVQTPDNELESAINWACGQQDVKCSM IQENQPCYQPNITTKDHADVFVNSYFCQNHKNKGG NCYFNAAIITEKDPSPHDNCHFEYLPSN
266	sp_116800_mxdp	0.3363373	0.9239332	#N/A	uncharacterized protein	MLFHTGSKNKQQTIGEKKVNEKVALETRVFKD NSFTRGGEVKEAGEVLIVLHDEYKGRAFTINCKIE AFGKIMQRFDRRRRWVNMFGGLYHLCGKH LPRNFVYWLMTRVDPVNVQIFSAPNGIDFFMSKN QIRWIFGPIKGVIPKTENDADEEMKAKAQFLGA YGKTWESDHPNIGAVISDAIPVNPFIQRLGEGT WEDKDEEFTMLIAAVEVWLCPTQCGRSPS MLYACSLGMOAREYDWCNLYVDFIERVKVFCR DFHTFGWTGCVGGCSMYLVIFYLRLNRMPLO WGFPRLKALDINDLLKARQADLLPSGDYKIGV CFSYPLTISANM

10

20

30

40

【表 1 4 - 1 0 9】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
267	sp_117040_thpe	0.93267	0.93267	#N/A	Zinc finger (C3HC4-type RING finger) family protein; FUNCTIONS IN: ubiquitin-protein ligase activity, zinc ion binding; INVOLVED IN: N-terminal protein myristoylation; LOCATED IN: plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, RING-type (InterPro:IPR001841), Zinc finger, C3HC4 RING-type (InterPro:IPR018957), von Willebrand factor, type A (InterPro:IPR002035); BEST Arabidopsis thaliana protein match is: Zinc finger (C3HC4-type RING finger) family protein (TAIR:AT2G38970.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).	MRRRWQRILNRNASDFAVETTNNINDAHREFTV KYFISDQKCSICESLKEGEDRALYISECSHAFHY NCILYNALHGNNICPTCNLSLWKNL PFLNASPTPY PNKIHIPNPPPPPNFSPKFPPIRHLPYADDDPLS SSTASSFFNNNHIPKPPPPPPKFPPIRPLPFAD DDPLPSSPATTASSFFPQLSESDTTL YPTNNVHI HIPKPPPPKFPPIRIRISFADDNPLPSSPVTTIA SSSFHEVALRESDTLL YPTNNHINHIFRPPPPPT FPHYSYQVLSSTFTTTTTTATTTASSATPLQQQ GANMTVKAFFEYPALGKSEGSNFALVGIQAPP LSPREADSTRAPIDLIADVSGSMTGNKITLL KKAMVFNINLGPDRLSITFTSKARRVTPLWRM TDRKNDSHAVNSIVVEEATNIIAGLKMAYQVLD QRREKNPVSSILLSDGQDNENEPGKHFRLLPY CIRSDTSSNSAHHQTDVIPVHTFGFGSDHDA TTMHAIADISRGTFYSIRIQQIDAFACQIGLLSV VAQLEIQVETASVKYRIQCIPSGKRRNIGYSKE HGVIY-DNVYAEVQKFLVLSVPPAELGTLTQL LEVKCLYMDPLSKONKVSFTVKEIYRDPVEDLS ETDROVCLEVSKEKHRITVAKGIAQAQAMAERG QLHEAQSSLQNIQITLETVRDSSGFGDAAHCEMLTE EVKVKQIMMASKGVYEGTGRGFTKSMMLCHGY QRAATQSSKAPAGRSVAMSYGTNNMKMVMVSK SRRMFRGGDDDDQDSIP
268	sp_117750_hpak	0.93267	0.93267	#N/A	NB-ARC domain-containing disease resistance protein; FUNCTIONS IN: protein binding; INVOLVED IN: apoptosis, defense response; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Leucine-rich repeat-containing protein (InterPro:IPR015766), NB-ARC (InterPro:IPR002182), Leucine-rich repeat (InterPro:IPR001611), Disease resistance protein (InterPro:IPR000767); BEST Arabidopsis thaliana protein match is: LRR and NB-ARC domains-containing disease resistance protein (TAIR:AT3G14460.1); Has 25652 Blast hits to 19792 proteins in 581 species: Archae - 38; Bacteria - 1467; Metazoa - 4997; Fungi - 304; Plants - 18213; Viruses - 0; Other Eukaryotes - 633 (source: NCBI BLink).	MAEAILPMAKILJGKIASGFSDLVTVYVTEIIGA KSAKDKLIADKMEICAVLEDEEEEFNKANKI WLRDLKSVYDIDLLDEVAFDALQRRVNOGYFG RQLRYLSSSNPLISRFDSLHKIESLRKKLEIVAK KNEFDLTKRPVKAKVEIRDPDFDECYVYEPAAIG RNEAKQDIISKL TALGDASELSVFSVGMGGIGKT ALAKLVHHDVNFDFLKWCVSDKFIHPKIEDILK AGTGNTPNQDMSMLVKKLGNLLNGKKYFLVLD DVWVEIDEWRKLNLLAVGKSGSSILTTTFLAN VALMAETMESYNLDCLSHDHVCWVSIFKQAFKEG EDGRYPLNLAIGWCIVYKGRGPIYVCKLGNLLR KLTSLPNSIKYLOKLOIKNCEELDFDKGLEGLQ SLRLLYISNLKLRPNDCISAAAASLOHLSVENCL GLVELPYWLRHFASLRITIEIYCRNLLSLPKTISHL SLOELKIGCPHLSKFCGVPGGEDYSLVQHVPK FELDEVCFLLTK

10

20

30

40

【 1 1 0 9】

【表 1 4 - 1 1 0】

I	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
269	sp_142090_cuqd	0.936333	0.933483	#N/A	PEANUT 1 (PNT1); FUNCTIONS IN: mannosyltransferase activity, transferase activity, transferring glycosyl groups, alpha-1,4-mannosyltransferase activity; INVOLVED IN: GPI anchor biosynthetic process, cellulose biosynthetic process, plant-type cell wall biogenesis, embryo development, cell division; LOCATED IN: integral to membrane, endoplasmic reticulum membrane; CONTAINS InterPro DOMAINs: Mannosyltransferase, DXD (InterPro:IPR007704); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2995 (source: NCBI BLINK).	MVINFLVLFSTILRVLLLYGDWQDTHMEVRYT DVDVLFSDAANLIFSGESPFKRTTYRYSPLAFL LIPNALHRSWGGKLYLFAADLLYGLIRITLKLKRGV PEKICTYVMAWLNFPTFTTIGTRGNCEPIVCAVI LWVLTICMRGKYVQAADVYGLVHFRIPIYIYALPI VLVLDRCQVIKSGQKPMLVKWSHOGNSNQC.LG DTPKSSVSLASIFSGTRILFGLMSGAVFLGFT GLFYLYGWWEYLHEALLYHLTRTDPRHMFYIF HIYLHYEQELSLKLAFLPOFAVQLVLMFNAL DLPFCLVQTVAFVAFNKVITAGYFVWFCLIPVIL PWSKMKLKWEGVICILVWMIQLNWLMLWGYLL EFKGNVLLPLWGSALLFMAANCYVLIIRCHRY SSIFQHLNSSKERLKVK
270	sp_176320_kfpu	0.921984	0.921985	#N/A	Tetratricopeptide repeat (TPP)-like superfamily protein; CONTAINS InterPro DOMAINs: Pentatricopeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: Tetratricopeptide repeat (TPP)-like superfamily protein (TAIR:AT4G21065.1); Has 36578 Blast hits to 13720 proteins in 224 species: Archae - 0; Bacteria - 18; Metazoa - 87; Fungi - 88; Plants - 35844; Viruses - 0; Other Eukaryotes - 541 (source: NCBI BLINK).	MSKEGVPPDRYTLPTIKAVGQAYSLKTSQQLHG VAICHGVEDDEFVESGLNMYVYKAGEFRDAYKVF DGSVPRNLGWSNIAIAGLAQSERVYRDVDFVLR MRRSGFEPDGMTMSVSWACGGDLRLALQLH KCVFQVWKKEEKVMMNSVIDMYGKCGRMDL ANCVFLEMNEQNSVSWTSMIVGYATHGLAKEAM EFFHAMIEACVYPNHVTFVGYLSACVHGGKVDE GKRCFQMMRNLYGIEPNVKHYGCMVDLLGRAGL LEEARNMVETPMKPNFNGAIWGGCLLAASERYNDV KMGWVYAKLFLELEPNWINDGVVYVLSNIYASRDM WEEVEKREIMKORRLGKLGQYSSVTL.SIIN
271	sp_177950_zogd	0.934038	0.937583	#N/A	sodium/calcium exchanger family protein / calcium-binding EF hand family protein; FUNCTIONS IN: calcium ion binding; INVOLVED IN: transmembrane transport; LOCATED IN: vacuolar membrane, plasma membrane, vacuole, membrane, plant-type vacuole; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAINs: EF-Hand 1, calcium-binding site (InterPro:IPR018247), Sodium/calcium exchanger membrane region (InterPro:IPR004837), EF-HAND 2 (InterPro:IPR018249), EF-hand-like domain (InterPro:IPR011992), Calcium-binding EF-hand (InterPro:IPR002048); BEST Arabidopsis thaliana protein match is: Calcium-binding EF-hand family protein (TAIR:AT1G2920.1); Has 622 Blast hits to 603 proteins in 166 species: Archae - 12; Bacteria - 69; Metazoa - 8; Fungi - 194; Plants - 285; Viruses - 0; Other Eukaryotes - 54 (source: NCBI BLINK).	MGMLAGSTIMLLTVIWGCCVYVKGKCDIONSRLIN GIDTKGFHLQDSGVSTDIWTSYAARIMASVVPFL VVQLPINALNSTSGRHYSILISLVSLLLIGYCVYVQ VSPQWQERRLDYVYKHKRVMITGFLKHLHETHATG KFYNRDGTINRDVVEKLFKAIKIDGNDGNHLSGSEL RAFVVMQIFEGMGEDYFLDKVMKEFDTRDPDD RIDQVFMQISKFLALVRGEKAIDHNDADTKYL SHYDEQATKEHFLVDDFLDEEKEEIEENRKAATIK AILLGLGTYVAFADPLVDSVHSFSEATNIPPEFI SFVFLPTINSGKAVSVISFASKKRSVSLAFCE LYGVEVTMGNVGLTVLLTVYARGLTWDFSSVEL IVFLVCTLMGALTCFRITTFPLWTALTALLCPPFSLL LYVLDNFFGWS

10

20

30

40

【 1 1 1 0】

【表 1 4 - 1 1 1】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
272	sp_187200_zfmi	0.36375%	#N/A	<p>plantacyanin (ARPN); FUNCTIONS IN: electron carrier activity, copper ion binding; INVOLVED IN: pollination, anther development; LOCATED IN: middle lamella-containing extracellular matrix, apoplast; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: Plastocyanin-like (InterPro:IPR003245), Cupredoxin (InterPro:IPR008972); BEST Arabidopsis thaliana protein match is: Cupredoxin superfamily protein (TAIR:AT5G26330.1); Has 1174 Blast hits to 1133 proteins in 56 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 1174; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink).</p>	<p>MGQGRGSARSALVGLAILLAIIVQPTLAAVYNI GWSFNVNGARGKSFVAGDVLVFKYIKGQHNVA VNGRGYASCSAPRGARTYSSGGDRIKLTRGQNY FICSPFGHCGGGMKAIAK</p>
273	sp_198600_lqjw	0.46765%	#N/A	<p>cyclin-dependent kinase B2;2 (CDKB2;2); CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase domain (InterPro:IPR002290), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: cyclin-dependent kinase B2;1 (TAIR:AT1G76540.1); Has 109034 Blast hits to 107608 proteins in 3051 species: Archae - 93; Bacteria - 11427; Metazoa - 41458; Fungi - 12141; Plants - 24398; Viruses - 437; Other Eukaryotes - 19080 (source: NCBI BLink).</p>	<p>MEATATGGEKIALSAMDFAFEKLEKVGEGTYGKV YRAREKATGKIVALKKTRLHEHDEGEVPPPTLREV SLLRMLSRDPHVVRLLDVKQGQNKKEGKTVLYLV FEYMDTDLKKYIRSFRTNGDTPPIIKSLMYQLC KGVAFCHGHGVIHRDLKPHNLLMDRKTTLIKIADL GLARAFVIKKYTHEILTLWYRAPEVLLGATHYS CGVDMWSVGCIFAEILLNNOALFCGDSLEQQLLHI FRLLGTFNEQIWPGVTKLDWHEYPQWSPKDIS SAYPGLDKEGDLKLLSKMLIYEPSKRISAKKAMEH PYFDDLDKTHL</p>
274	sp_206500_lysz	0.37617%	#N/A	<p>2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein; CONTAINS InterPro DOMAIN/s: Oxoglutarate/iron-dependent oxygenase (InterPro:IPR005123); BEST Arabidopsis thaliana protein match is: 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein (TAIR:AT1G49390.1); Has 8498 Blast hits to 8450 proteins in 980 species: Archae - 0; Bacteria - 1051; Metazoa - 121; Fungi - 1005; Plants - 4926; Viruses - 0; Other Eukaryotes - 1395 (source: NCBI BLink).</p>	<p>MAASSVSPMKPELPPKPTQELSKITGDQLPDRYI HQFPETAADHSAIKYMDNPIDLNLNSAVSTHQPE ELOKLRSSLSWGGFQLRNHGLTSSLLHQIROV EKEFFGLPVEYKQKHSRTLHWVEGYGNDTVSDE QKYCWNDRLVLKVPVDQRNFKFWPEDLPKFR ETLDEYTKVEYRRVLEMLKASKSLNLDENSEVRE CEGDEGIYMLARFNLYPPCSTPDHVVGLKPHSD ASAVTILLQDNEVEGLQVQKDNQWFKVPIVPDTL FINVGDQLEIMSNGLIKSAVHKVVDKIKERERTSVAI ACFPYIEKIGPISLIDKERPQLYKVKNYISLYLQ YSPATERAITAMKL</p>

10

20

30

40

【表 1 4 - 1 1 2】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 Cupredoxin superfamily protein; FUNCTIONS IN: electron carrier activity, copper ion binding; CONTAINS InterPro DOMAINS: Plastocyanin-like (InterPro:IPR003245), Cupredoxin (InterPro:IPR008972); BEST Arabidopsis thaliana protein match is: Cupredoxin superfamily protein (TAIR:AT5G26330.1); Has 1434 Blast hits to 1387 proteins in 64 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 1434; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).	aa seq MGITQTSKSLYNILNLMKLLIISLVFTIAKSDVIA VGGENRWSDGFNYVSWAQNTLFTVGDVLEFSY TRGSHNVLDVTSATFRSDNQSTGNVAMYSGGDD HVQLNEAKQYWFICTVREHCRRLGMKFGINVTTS LNRTSNSTLNDPSSGNKHNSGGFYLLAVWIFSMF L
275	sp_215910_xzds 0.9436482	0.965138	#N/A	beta glucosidase 16 (BGLU16); FUNCTIONS IN: cation binding, hydrolase activity, hydrolyzing O-glycosyl compounds, catalytic activity; INVOLVED IN: carbohydrate metabolic process; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINS: Glycoside hydrolase; family 1 (InterPro:IPR001360), Glycoside hydrolase; family 1, active site (InterPro:IPR018120), Glycoside hydrolase, catalytic core (InterPro:IPR017853), Glycoside hydrolase, subgroup, catalytic core (InterPro:IPR013781); BEST Arabidopsis thaliana protein match is: beta glucosidase 13 (TAIR:AT5G44640.1); Has 11475 Blast hits to 11098 proteins in 1475 species: Archae - 144; Bacteria - 7936; Metazoa - 716; Fungi - 200; Plants - 1457; Viruses - 0; Other Eukaryotes - 1020 (source: NCBI BLINK).	MAVLHSAFPQALHSSVT_LKNPNSLFSIRNTSK YYLPTKISCSLIPNVIKAAKTVVNPVISVDGLQPA TFTPPKDDLIKEDFPFGFGASTSALQTEGRGD EGGRG;PSTWDFMIKADKGHKAVDSYNYLKEDVK LLKETGLNTYRFSIAWPRILPKGTIESGINQEGIDF YNRFIDELIANGITPVLTFHFDLPVSLQSEFYGF RPNVDHFKEFADLCEKFGDRVYKWTINEPYV FGSYGYKMGFLPNALKDDPLDGLTYIATHHILAHAA AKLYKEKYQKSGGKIGISLPLCKWYMPHASADP KDFVATKTQLDFTLWYVMDPLTRGDYPESMKKS VKGLPVFTKAEKELIKGAFDYDYGINYYTSRYVEHT CVVDAMVGGTITIKYSEQKVDDEGNLVGWQAQGG EDIVYRDPDGLKNLLLVKTEYDNPVYVTEANGVP DKFTITSSIEDLIAGQNTTELEAALYDDYRIGHTE HLVAVRGAIRGANVKGIVVWVSLMDNMEVYGAGY ETRFGLNFTDYNNDNHKRYPKSAKAWFANFLKEE TTNNGSLKKN
276	sp_014560_mxdg #N/A	0.947381	#N/A	VPS9B; CONTAINS InterPro DOMAINS: Vacuolar sorting protein 9, subgroup (InterPro:IPR013995), Vacuolar sorting protein 9 (InterPro:IPR003123); BEST Arabidopsis thaliana protein match is: Vacuolar sorting protein 9 (VPS9) domain (TAIR:AT3G19770.1); Has 1393 Blast hits to 1376 proteins in 227 species: Archae - 0; Bacteria - 0; Metazoa - 508; Fungi - 217; Plants - 509; Viruses - 0; Other Eukaryotes - 159 (source: NCBI BLINK).	MSNLFQTFQYISQTPQNHTSNQTHAHFPIPHL NSPFFKRRTRITTCGLRIGPKRPLWRSSFLSTE AIQAVOSLKLAKSPVKLQQVFNTRITRLLKADLLG ALSELQFNHLDLALVKVDFVKEVETADLSLYD MILLGKNIKQAEEELDELKKEGLQPDVRAYTE LIGAVFQVDMVEMKAMETYNEMKASGFTPKLTL MIMVRNLEKAGEQELAAASVKRDFEYVDEPERFL QQVEKNYSYSTD
277	sp_017220_zcfr #N/A	0.943837	#N/A	uncharacterized protein	MILVGDDEEDGEDLKGSLRHMRLTTDQKDEEDG WALASVFGFSGVEEYVARRYEQNRVTEFQIEI
278	sp_023330_reqs #N/A	0.952084	#N/A		

10

20

30

40

【表 1 4 - 1 1 3】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
279	sp_031030_kwll	#N/A	0.945223	#N/A	embryo defective 2739 (emb2739); CONTAINS InterPro DOMAINs: Integrator complex, subunit 3 (InterPro:IPR019333); Has 179 Blast hits to 170 proteins in 63 species: Archae - 0; Bacteria - 0; Metazoa - 131; Fungi - 0; Plants - 37; Viruses - 0; Other Eukaryotes - 11 (source: NCBI BLink).	MKYLHAIVNDGYCLFHILISQIVDGLYSKMLEFPKVK QLIWSVMIDVSGVGFDSLLVSLLRQIVGGDFDS EGNLWLSFEIWNLFSTKWDSDILDEEPLVLSALYT YLRLLADHYRLAQSPLVLRKEIGFCVRLMREK FHLGLRIGDLVRLQLDLYHIFEFRAIWKLLNPN NSFQTSFTDISHIYGMKTSRYSRYLLRITPEMETQ LRFLLTHVKLGSORRHQAWFRKFLFAPERETLV SDIVRFICSHHPSNEIQSDIMPRWAVIGVLLKL CSKKYFEANVKTLYFDWLVFFDEKVDNVMNVEP AMLLMNSIPRYDITNSLLEFLLVVENYDMDRK NVIYNGVSSAISMLVSKGVIGSLDIFSSCKLLSPFV KEMLQRFMSKVQSGNPSPEMPPEYLSSEMPPO TSPNPPYLEIQTHQSAKLRACPRNAELSAQHSYLV KPTSAASLCEGQVDNLNLRKIREFMISKEGQ KLVENILRYINPANEGKFGGIGSEYLACELKKE ELAGYNLFSVPSLNSDDNADDEVQSVYTAIIR TFISQHQSRILMLFWWSRNGCLVGPRLLSYASR LAHEAHEAGYASYTTDINYTKMPLLNKHIEGYM SLTTSQNDHVNAGTTPNPEYDKKLSKMDSAFA SYRFSHAKDLSKSEVGIKLLSMDVISCGLW KSKFFKDLFCYFNSLNDPISLSDDEDVILLVTQLED TDLLYMQMDLGLKRVTFIGEDYKVLQLIRRSYW KQVDQHMVWGLIRSELTVSEFFVEKLLGCFEFT DFDSKTSVGVGGHLCSSCAPTEIVGAVMLL PNDKFPGFAGAIKSHWTVTNASMFCNSVADF.LN KLGKNGKSPLDWDGFLNQSGLLFLLEFYKKAQG ERGVNMLEKALILSNLEIRPGETTPVMDTG
280	sp_034290_khnd	#N/A	0.93505	#N/A	ROOT UV-B SENSITIVE 3 (RUS3); CONTAINS InterPro DOMAINs: Protein of unknown function DJF647 (InterPro:IPR006968); BEST Arabidopsis thaliana protein match is: Protein of unknown function, DJF647 (TAIR:AT3G45890.1); Has 450 Blast hits to 446 proteins in 132 species: Archae - 0; Bacteria - 0; Metazoa - 113; Fungi - 75; Plants - 194; Viruses - 0; Other Eukaryotes - 68 (source: NCBI BLink).	MAEQILVEEWAGSSSKLSITATISSTPNISIRFRD SRFKLVRQILSAFVPEGFPSSVTPDYVPFQIWD SLQGLSTYIRMMSTQALLSAIGVGEKSATVIGAT QWFLRDLTGMGLGILFFYQGSNLDNSNAKMWR LVADLMDLMDLMDLSPFSAFYFILCLGSLSR SFTGVASGATRAALTQHAFALQDQNAADISAKESQ ETVATMGALGMFLAHVTRGHSMAMWFSFLCL TMVHIYANKAVRCLSALNGERSLLOHFMK TGQVLSPEKVSMEEHVLPWLWNPWSKKFKGLH SRINLGVKSSLNHAAMTDLFFCAGSYKAKYL LLEKNGMISIIHKDSTPYDILOSYHALVWANI.SNK SKSGHVESFSWMDKYEDFIKKSQGWKTERLLS HSYVWFRASWFIGQSDKID

10

20

30

40

【表 1 4 - 1 1 4】

1	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
281	sp_038050_rktw	#N/A	0.843312	#N/A	<p>ARM repeat superfamily protein; FUNCTIONS IN: binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: Armadillo-type fold (InterPro:IPR016024).</p>	<p>MAPESNLEASPVINTLQWLEHCHENGVEVNDAE EERFQLHADENGPTHPSAPPDEFIDISTTVDP YIISLRKLLPADERQGGDLLGLSVWNNSTQRSDA DQVVDGASSPKRNSGLNINLNGDLETMESOPVGL EQPSRSAGPDDNDTGLPKKTAREEVWEECGCV LWDLAANKHEAEYMQVNLVLEVLANLLTSDSVR ITEISLGIIGNLACHVEYVMDKITSTNKLETTVDQVF SDDPPCLGEVCHRLTLGLGSGSVVWARALQLD HILSRVLWVAGNTLNPVLEIKSIGLFLAEMENQHV VPILLPTFMNLGLPDLVSLLAVELGLKLTSDRTSE RNVFGSDNVTYLLEHLHCLKTFSSLFTPKVQGGC TDFVMYPALDSILRAVEALSVDYDYSERISSKKEY VQMAELIKIPEKIEISSCVTASVLLANILADSPDM AFKLSDFLLQSLDLDFPFTSEEFARNALW/SIL ARLLVRAQEAEMSSSRFSQFVLMVDRSDMLE ALLESQWKEKSPKRDNLTTSEAKFDSSTAIAIKGVK ILNQWNLVKDEYSVVCTDEDSVRSKSYPPDRDQV KLLDCCLKHAS</p>
282	sp_040250_geuu	#N/A	0.2835887	#N/A	<p>plant VAP homolog 12 (PVA12); FUNCTIONS IN: structural molecule activity; LOCATED IN: cytosol, endoplasmic reticulum, nucleus, plasma membrane; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: PapD-like (InterPro:IPR008962), Major sperm protein (InterPro:IPR000535), Vesicle-associated membrane protein (InterPro:IPR016763); BEST Arabidopsis thaliana protein match is: vesicle associated protein (TAIR:AT3G60600.1); Has 1106 Blast hits to 1079 proteins in 208 species: Archae - 0; Bacteria - 0; Metazoa - 441; Fungi - 140; Plants - 434; Viruses - 0; Other Eukaryotes - 91 (source: NCBI BLink).</p>	<p>MNSGEFLSIQPOELQFPFELNKKQISCTMQLTNLT DSYVAFKVKTTNPKKYSVRPNVTGIVLPHSNHDL VTMQAQKEAPADMQRDKFLLQSVIASPETTKD ITSDMFNKEGGLVDEWKLKVSYSVPPPPSAV HEGSDGVSVDNGHMDTFSAAATHVETEDDSR QVNAISRLEEKNSAIQQNNKHLHQEMDHFRRREI SKKRATIPFMVYIIVGIVIGIMLGYLYKRI</p>
283	sp_048010_ucep	#N/A	0.5333393	#N/A	<p>S-adenosyl-L-methionine-dependent methyltransferases superfamily protein; CONTAINS InterPro DOMAIN/s: SAM dependent carboxyl methyltransferase (InterPro:IPR005299); BEST Arabidopsis thaliana protein match is: S-adenosyl-L-methionine-dependent methyltransferases superfamily protein (TAIR:AT3G11480.1); Has 909 Blast hits to 891 proteins in 125 species: Archae - 0; Bacteria - 69; Metazoa - 9; Fungi - 5; Plants - 730; Viruses - 0; Other Eukaryotes - 96 (source: NCBI BLink).</p>	<p>MEIKNYLHMVYVNGELSYSONSSVOARVWLIN QSLLLENAIQSLITSKDSGLNIADLGCYGNVPLGL VSFLVEIMQKSKKELLKLDNDDQVPLQIYMNND LPSNDFNLFKEMMNLVLLKXKDDDDGVPCLFLM GAPGSYDRLPFNKSLHLVHANYSLHWLSKVPP RIYDDOORMSINKINIYSETSPAKYGRIRYRDOFEQ DFTLFLKRSRKEVIANGCMLTLRGRSSSTNPLT WTTTFEFKFTKTLASLYSEGLIKEEKLDTDFPCY CATKEQLSIAKKEGSFEVETSKTMTWVDVAPEIE DKWERAQVITKIIRAFSESLSYSHFGEDISTLLYD KLVHFANQHLVDDQLAQNHAVTVLLRRKK</p>

10

20

30

40

【表 1 4 - 1 1 5】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
284	sp_048090_axht	#N/A	0.923115	#N/A	<p>sinapoylgucose 1 (SNG1); CONTAINS InterPro DOMAIN/s: Peptidase S10, serine carboxypeptidase (InterPro:IPR001563); BEST Arabidopsis thaliana protein match is: serine carboxypeptidase-like 9 (TAIR:AT2G23010.2); Has 2871 Blast hits to 2808 proteins in 264 species: Archae - 0; Bacteria - 35; Metazoa - 597; Fungi - 474; Plants - 1434; Viruses - 0; Other Eukaryotes - 331 (source: NCBI BLINK).</p> <p>DNA binding; FUNCTIONS IN: DNA binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: Zinc finger, BED-type predicted (InterPro:IPR003656), Protein of unknown function DUF659 (InterPro:IPR007021); BEST Arabidopsis thaliana protein match is: tAT transposon superfamily (TAIR:AT3G22220.2); Has 518 Blast hits to 485 proteins in 17 species: Archae - 0; Bacteria - 8; Metazoa - 8; Fungi - 0; Plants - 505; Viruses - 0; Other Eukaryotes - 5 (source: NCBI BLINK).</p>	<p>ESYYLLENWANDIQTOAALHIQKGTKSSWIRCN MTLAYAYNVSESSFGYHQNFQDFIRALIY</p>
285	sp_062340_nrhw	#N/A	0.94717	#N/A		<p>MSLGNFDLLVSCPGLDPKDPARKYGRPDNNR NNFTCNFGGKVTGGAYRLKQHLVGGFRNVTKC YSCPEHYKKEEYKFFMIKKQQAQVGGQMLPRVTQ SDVDGCNDDEEDCETLNSKGGSSVKATSKRPRL KGFMDMYFTPIPEDAISGRKNKNGRQRGSFSY VPSRSVQSGADPIASDTPGRRDLQLQGTSPSHFR FRNAVYVRRRRRFR</p>

10

20

30

40

【表 1 4 - 1 1 6】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MWRKFCIQAVYQIENITIKNTLSLKKVWIKQVKNRS NFSQVLDEIKPHISNDFAIQLQKTSYSSPWQR VLPFDTPETISYLKAKRSAERFQVLOFAVCPTLRA SKLTAYPNFHLFPRDELKIGMPSYSFACQTSYL TAMARLGFDFNSCYDGISYSIRAQESSMKHLPG NTMLRAHASQSSPAVSADSIYIQRITRVTWIK ACKDSKGLNEDPMVRAALRKMVIMGEHGSRPCLDL DVCSEKRVQLVLEVMKEYSDDTVPLLIIPAKRGGT CQLRIVLTSSEQDKILFOKELQDLLEEFOSKNVRG FFREVIDLISAGKPVIAHNSLDDFTFIHSHKFLSTLP PTIDFKHLSLSSFFPYIIDMGHVTKEFNRSRNMNN IPAASSFLRSRFFAPIDMDIPIKGTPEMSEKING QNVLOLSHLFAKLNTIIRGQHHVISEDKILADNF ANVFDYSSPLPENSDDDVRISSGSLRKLSCDN MVFIFGFRSGLSTGELKTLSSGSHAEAFSDVQ LVDKSCAVALQOPGLAGAFYKAMESGGGCHL GGLEKLLSEGLFAADYGTYLACNMSKWDSNLA DSLDQVLMASADQFFVDSHGRSPYEIGSFNQRK QLFSISPTTRKIRGRNGNGLMGAVCATHNFQVW VVGAGIIGLIARDLLLRSLNSVAVVDAAPVCSGAT GAGOGYLWMVHKTPGTGEVWDLAMFRSRKLWQL LADNLQDQGSDDPLEALGWKKTGSLIGKTEAQL NQLKSRVQQLSMAGIRAEYLSNNDLAIREPELAIN QQSGAALPDDCCQLDAHRMTAFIEKDNRLFLSQ GRYAEFYHDPVAVGFIRSNKGKIDGKTKQKNTLY GRKATIAAGCWTGSLMDKLLQNSDIVLHVPIQPR KGFLLALEFNFLRVTHGLMEAGYVDHOKSVISA SDKVDSEALSISMTVTRDASGNLLIGSSRQFAG FNTELDTFIVDRIVERACDFPSSLKEFSVKNLSE QMNVRIGLRPMYMPDGKPVIGVPGITNLFASGHE GGGLSMALGTAEMVTDMLDNLPLQDLQCKPFAVE GROC MENLRISQISTQRIQYSVEPDIQDATMEDYLNNC MENNALAQEREEESTNEIEAEQHIYDSEEPD QDIVGTLIGYTAETVEELLVRINEKHATSLNLSNTE DGS TIETSATAAPETSITTRAPETNSAPTGTTHATT TSDTSVQQAATSEPATNTEPMMVLDPERCT TKGRNKRPRGPFVKKKKKGGKTAAPPTADFGTITPN LRLF
286	sp_082780_spxw	3 3653432	#N/A	Polynucleotidyl transferase, ribonuclease H-like superfamily protein; FUNCTIONS IN: ribonuclease activity, nucleic acid binding; INVOLVED IN: RNA modification; LOCATED IN: nucleus; CONTAINS InterPro DOMAINs: Ribonuclease CAF1 (InterPro:IPR006941), Polynucleotidyl transferase, ribonuclease H fold (InterPro:IPR012337); BEST Arabidopsis thaliana protein match is: Polynucleotidyl transferase, ribonuclease H-like superfamily protein (TAIR:AT1G55870.1); Has 339 Blast hits to 335 proteins in 120 species: Archae - 0; Bacteria - 10; Metazoa - 154; Fungi - 41; Plants - 69; Viruses - 0; Other Eukaryotes - 65 (source: NCBI/Link).	#N/A
287	sp_063090_jhxo	3 842252	#N/A	uncharacterized protein	#N/A

10

20

30

40

【 1 1 1 6】

【表 1 4 - 1 1 7】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
288	sp_065250_adcs	#N/A	0.323737	#N/A	<p>EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Cactin protein, cactus-binding domain, C-terminal (InterPro:IPR019134), Cactin, central region (InterPro:IPR018816); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G36815.2); Has 11711 Blast hits to 7382 proteins in 452 species: Archae - 31; Bacteria - 352; Metazoa - 6006; Fungi - 1138; Plants - 599; Viruses - 33; Other Eukaryotes - 3552 (source: NCBI BLINK).</p>	<p>MARHERSKSSRRREKPPSSRRKKQRRHSDNDSA VSOSETDYDSSRRDRSRRRRRSNSHRRRESS PETSSESGDRRKRKKKVEEIKKEYLAKKAQKKA MRVAKLKTHTVSGYSNESNPFGDANLNEKFVW RKKIERDVCDGVPMEIEFTVKAEEKKQKDRMSEIE KVKKRREERALEKAKQREEMILLORERSRAEAD DFEKKEEFHFDQSKVRSRIRLREGRTKPIDVLT KYLDDSEDFDVELDEPYMVKFGLTVKEIEELHDDI KLHLGLDFAITPTHEVYWEALMVYGEWELGESR RKDAIDRARVRGEPPELLAEERGMHSSIEDV RNFLEKSYKDLVELQAIQIESEMRSRGAKVVEY WEAVVKRLHIYKAKACIKIHKALLRKHLLRLEO PLOGENVKSEEEKIEEIEEVEEVKVCLDARAFSP EPIITHEEAQSEAEADGSGFSPQLLHSEDTTEEAIDP EEDRVILERKRLAVIEEQRAKEAMASKPRAQPED DLEVRARKVMGVVEDGDVAFGAGAEINLDSQVY WWHDKYRPRPKPKYFNVRHTGYEWMKYNQTHY DHNDPPPKIVOGYKFNIFYPDLVDKIKKAPTYYIEKD GDSSETCHIRFHAGPPYEDIAFRVKNKEWEYSHKK GYKCTFERGILLHVFNFKRYRFR</p>

10

20

30

40

【表 1 4 - 1 1 8】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MSRKNIDGPFSSINRFLLENRQFRKFNVKINLQPKL TLISSGTLAVLYVWLSILTFVGNNDKNGFSSSSWEK AEVFMFRIVQALTQVAIFVLVVHEQRHFHATCHPMS LRFYWIACFVVLVFSISGLTMLSAGEELDSFM RTDDMFSLVSLPFCFYLLIVASQTSIGVIVKEEN GLSGRTEVNETVDDPNVSDYAKANLFSKTMWL WMNPVINKGYKNPKIEDVPSLPPDHLAERMSKL FESQWPKSGEHTKHPVAVTLKCFWKNIVYTG LAVIKLGMVYVGVLIQSFVNFSTGMSSPYEGY YLVTLLIAKAAEVLSAHFNFHTQRVGLLIRSTLIT SLYRKLGLKSCSRQAHHGIGQVNYMVAQAQOL GDLMNQLHSIWMPLQLFAALALYFNMVGSAlA TFIGVILGTFVIRTKCNVFNQYQVMRNFDRJK ATNEMINMRVYKQAWENYFHKRIQSYRELEYG WLSKFSYTAANMGLVWSTFVALSVLAFGVAIV MGMYLDAGTVTVTIMKIQLEPVRNFPQSLIQIS QAMISLGRLDGYSLELDQGLEEREEGCQDDIA VEVNGSFSWEDEATEQTLKGLNFQVKKGELSA IVGTVGGKSSLLAAILGEMHKISGKVKVCGTTAY VAQTSWIONATIQENILFGLPMNKEKYEKIRVCS LEKDLIEMFEFGDKTEIGERGINLGGQKQRIQLAR AVYQDCDYLLDDIFSAVDAHTGSDIFKECVRGA LKDKTLLVTHQVDFLHNVDDQIFVMRDGMIVQSG KYNEILEAGLDFGALVAAYENSMELVETSSHATE DDTPKSPISENKTIERYKSEKALTKTDSQKGTSK LIEDEEREITGRVSLVLYKYSYEAFAFGWWGVSAV VIVSLFVIVSLLGSDYWLAYETSEERFTPSLFIIV YAVIGIISCLLVARAFVITFLGLKTAQGFNFQILDS ILHAPMSFFDTTSPGRILNRASDTQTNIDLTPFL GMTVMYFNLSILVTVCVYAWPTVWLLVPSIWLIN LWYRGGYIISRELTRLGSIKSPYIHFHSETYSG VMTIFRCFRROALFCKQNIKIEKVDLSLKMDFYNGA MAEENNTSESHEKGFHLEPEYDIDGRGSIYKIL TLSMKRRFSPQVVGQAQTPSPGAGEVRDLRLRL VEIDNESRKLKLVFSLLEKDESLESYVKKRIEDP KVSRMWLLVEQVLACQEGDTNFQVIMGDDNSI EDGNITTYDVKSIEKDEIGEERKNERFRKVRNITS GHQPVPVGVTTITANGK
289	sp_068740_kuop	0.9302393	#N/A	multidrug resistance-associated protein 10 (MRP10); FUNCTIONS IN: ATPase activity, coupled to transmembrane movement of substances; INVOLVED IN: transport, transmembrane transport; LOCATED IN: vacuolar membrane, plasma membrane, vacuole, plant-type vacuole; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: ATPase, AAA+ type, core (InterPro:IPR003593), ABC transporter-like (InterPro:IPR003439), ABC transporter, transmembrane domain, type 1 (InterPro:IPR011527), ABC transporter integral membrane type 1 (InterPro:IPR017940), ABC transporter, transmembrane domain (InterPro:IPR01140), ABC transporter, conserved site (InterPro:IPR017871); BEST Arabidopsis thaliana protein match is: multidrug resistance-associated protein 4 (TAIR:AT2G47800.1); Has 708251 Blast hits to 364353 proteins in 4024 species: Archae - 13144; Bacteria - 571430; Metazoa - 13503; Fungi - 8752; Plants - 6721; Viruses - 16; Other Eukaryotes - 94685 (source: NCBI/Link).	#N/A
290	sp_074060_fdfx	0.922536	#N/A	uncharacterized protein	#N/A

10

20

30

40

【表 1 4 - 1 1 9】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1) 分析に用いたバइटとの相関係数	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MGSFLLAFQVETRRLLWGLVFTITFCQCECPK GNSLWLSFRGDEIPEKKVLSFFGDNISTYSASL SNMSSPSTISVISEDNRNNSVRIHLGNITTECGSNT SSNEAPKDFNLAPSEPPVAPFLSHDVASNLST EFSAPPLLIHVSENKTYISVNPGLGNTTERRSAF NEAPKEDFALPPGHNGDSNLAPVESHVDVGRPNL SMFSTPPPPPPPPPPPPPIHVSAKPKSKMF DMKVLTLSDMSEELLQNYLSPPTMVPWSSKVD EELVLAKQLISNAPPLKDVLYAPLYRNASMFHRS YELMEOMLKYYKKEGDRIRHRPPLEGYASEG WFMKHEANRHFVTKRPEAKHLFLYLFSSRQLE ETLYPNSSHKNLIQHLEDYVDMISSKYAFWNR TEGTDHFLVACHDWAPAEKRMGNCIRALCNA DKKEGKFGKDVSLPETVYRKARNPLRDIGGKPP SKFKTLAFYAGNMHGYLRFLLQYWGNNIDPDM QIYGSLPRSKDNRNYNYMKSSKYCICPRGYEVN SPRWEAIFFECEVPIISDNFVPPFFETLNWESFA VFAEKDIPNLKILVSIPEKTFRFRHMRVKKPGL CLAQENMGKLLKAFHVETRRLLWLTGLVFTIV LFOCIELPKGNLSWSMFSPGEVSEQKVNFFGKN STSYLVSNMSSPSAVNSKIVSEEDGNFSGRYDE GRNVVPISENPGLGNTTKSNSSFTESLKEYFAL HKGNSNLALIKPLGAPLVLSHDVAPNFSMEFGTS APPAPPPLVHVSANKNRLTYFEKEVLSLSDMSEA LIQNFASPTMIPRWSKKADEELVLAKQLIESAPP LKDGLYAALYRNASMFQRSYELMEQMLKYYKKE GDTRIFHRPPLLGIYASEGWFMKHMEANRHFVY KKPKEAHLFLYLFSSLQLEETLYVPKSHDRKPLIQ HLEDYADMISAKYPFWFNRTEGTDHFLVACHDWV PSETKROMNCIRAINSDVKEGKFGKDVSLPE TNVFKSQNSLRDVGKPPSKRKTlafyagntlh GYLRFILLQYWGNNITDPDMQIHGTLPRSKDNRN
291	sp_079500_yq1d	#N/A	3323335	#N/A	Exostosin family protein; FUNCTIONS IN: catalytic activity; LOCATED IN: membrane; EXPRESSED IN: stem, root, stamen; EXPRESSED DURING: 4 anthesis; CONTAINS InterPro DOMAIN/s: Exostosin-like (InterPro:IPR004263); BEST Arabidopsis thaliana protein match is: Exostosin family protein (TAIR:AT4G32790.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archaea - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI/BLINK).	

10

20

30

40

【表 1 4 - 1 2 0】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたペイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
292	sp_082830_szmt	#N/A	0.931234	#N/A	DIE2/ALG10 family; FUNCTIONS IN: transferase activity, transferring hexosyl groups; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system, integral to membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Alpha-1, 2 glucosyltransferase Alg10 (InterPro:IPR016900), Glycosyltransferase_ALG10 (InterPro:IPR007006); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MGRTVVALIVSLWMPVSVILVNRVPHSHYMDEIFHI PQVQYCKGNFRSDWPMITTPGLYYLSLAHIAS LFPGMWVQRISSESDLCSTAILRNSVNGALAVCSI LYEITOLKPAIDRRKAALNTVLSLPLHWFFTY LYTDVAVTLVFLAVYLTSLKKNYFLSALMGALAL CIRQTNIIWMLFVACDGDVIDMLLVNRQSVGIQKSD ISIAEKNRISPKDNFSAKPGVRRRKMKGTDQVDFP PTVTGGNEATL SCSHL SGLDFEFQYIIQAWHLK WELMVSFTPELVLVLFVAFVSVYVWNGSVVVGAKI WELMVSFTPELVLVLFVAFVSVYVWNGSVVVGAKI AHAVSPHFQAQTMVFAMVSALFSAPFHFTLGHAT ALLRSFFKIRPLSFLFGVILTAGVLSVHFFSIAHP YLLADNRHPYFYWRKVINAHQMLKMYFLVPFYVY SWLSLHLGKARSRIWVFLAYFLACAAVLPVPLV EFRYYTIPFFIFMUNSGIEIEIKWLLIAILVYVNSF TMYMFLFRSFEWSHESGVQRFFW
293	sp_087510_gkqg	#N/A	0.929332	#N/A	respiratory burst oxidase protein F (RBOH F); FUNCTIONS IN: NAD(P)H oxidase activity; INVOLVED IN: in 10 processes; LOCATED IN: plasma membrane, integral to plasma membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ferredoxin reductase-type FAD-binding domain (InterPro:IPR017927), Cytochrome b245, heavy chain (InterPro:IPR000778), EF-Hand 1, calcium-binding site (InterPro:IPR018247), EF-hand-like domain (InterPro:IPR011992), Ferric reductase-like transmembrane component, N-terminal (InterPro:IPR013130), NADPH oxidase Respiratory burst (InterPro:IPR013623), Ferric reductase, NAD binding (InterPro:IPR013121), EF-HAND 2 (InterPro:IPR018249), FAD-binding 8 (InterPro:IPR013112), Riboflavin synthase-like beta-barrel (InterPro:IPR017938); BEST Arabidopsis thaliana protein match is: Riboflavin synthase-like superfamily protein (TAIR:AT4G11230.1); Has 2419 Blast hits to 2261 proteins in 348 species: Archae - 6; Bacteria - 265; Metazoa - 742; Fungi - 676; Plants - 525; Viruses - 0; Other Eukaryotes - 205 (source: NCBI BLINK).	MRGYARHELRRWGSNDNIRTPAMAMSTGSSPGS DFTTSGYSTPRSTATAEEFVEVTLDIEDDETLILRS VEPAAITPPPTMNAESSSSRPNSLASMAGME IPTASTSRTPAIRRTTSSSSRIRQFSQELKAEISKA KQFSQAQLTKRFSGKFLNYSNMKGSADVDPAL AARNRRQLAQLNTRTSGAHRALKGLRFISNKG SASCEVESNEFEKAKDGLYRTDFAKOIGMEKS EEFALEMFDALSRRLKVEKISDELLEYEWSQI TDQSFDRLQIFFDMVDKNEGDQLTEEEVEIIML SASANKLSRLKEQAEYAAIMEELDPEGLGYEL WQLETLLOKDTYLSYSOALSYSQALSQNLQGL RYKTPMQKMRMTMNFYVDNWKRIWVLSWILL MTGLFIWKFQYRNKAAFQVMGYCLTTAKGME TLKLNVAALILFPVCRNTIWLRLAYLVPFDDN INFHTKTAASIVLGIHLVGNHLACDFVRLERSPPQ DYALYLIDFGKTKPTYGDLVRGVEGIGLMMVLM AIAFTLATRYFRGGDKLPKPFDRDLTGYNFAWYS HLLIIVYICLIHIGIQLYLVHKWYKTTTMMYLAVP VLLYAGERTLRFRRSGLYPVRLKVAIYPGNVLT QMTKPPQFRYKSGQYMFVQCPAVSPEFWHPFSI TSAPGDDFLSIIHRLKGDWTEELQRIFKAACETPI PGRSGLLRADETTKKS_LPKLLIDGYPGAPADYR KYDVLVLLVGLGIGATPFSILKLNNTVKKKEEDEE KAADLYVPDLESGSPKQKPFRTTNAYFYWVT REQOSDFWFKGVMNEVAELDQRGVEMHNYLT SVYEEGDARFALITMVAALNHAKNKNGVDIVSGTKV RTHFAKPNWKKVLSKIGSKHANARIGVYCGAPV LAKELKVLNENEYNOVRTGTTKFEFHEKHF

10

20

30

40

【 1 1 2 0】

【表 1 4 - 1 2 1】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたバइटとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
294	sp_093090_ycwu	#N/A	#N/A	AAR2 protein family; CONTAINS InterPro DOMAIN/s: AAR2 (InterPro:IPR007946); Has 319 Blast hits to 307 proteins in 161 species: Archae - 0; Bacteria - 3; Metazoa - 130; Fungi - 91; Plants - 45; Viruses - 0; Other Eukaryotes - 50 (source: NCBI/BLINK).	MEPEKALELVKQGSTLLFLDVPQFTLIGIDTQVYS AGPEFKMIPPGVHFVYSSSTKDGKEFSPIIG FFIETSEVVRKWNLEERLVKVPSEEEERYA QAVKQFEDRHLGPTLSDYGDWKRLSNYTKD VMERIEPIGGEITVACESSLVGNPKTITMEVALAE QLKNSKFHREAESSEKGCYYTPVPWVWIKQKGIS GAELTSVNLDKTHILELILMKECGNEDRFLGEL QFAFIALMGOSLDGFLQWKSLSLLFGCTQAPF HTFRSLFVKVKVYQKLYGFQORDHARSCGIEK ESSLLDDSWLTSDFLHLLCKEPELLVDEASVL DGDLLTWTRKLRLELLENTLWFEFQKSDMDRLY FEEDDEYAPVVVEMLNESN
295	sp_096790_omah	#N/A	#N/A	uncharacterized protein LEAFYPETIOLE (LEF); CONTAINS InterPro DOMAIN/s: DNA-binding, integrase-type (InterPro:IPR016177), Pathogenesis-related transcriptional factor/ERF, DNA-binding (InterPro:IPR001471); BEST Arabidopsis thaliana protein match is: Integrase-type DNA-binding superfamily protein (TAIR:AT1G28160.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MAAGPRKRPQLASKNNPTRDRTRTTGFGGGPG GYGSKGASEEGVNHSTRPNQQRPPPTPMRSK EALRVPIRRGGFWFAGFGCRRYTGLWTSNQL PTATTSRRRCDDPFRMAASNLVGEYWEIKVKS EQILQYDLLLLYKLVVDGKLNWRSSHSDEEEEK SRVITKVKVYVHAKEKVKVKGAKPPPPGARKS MRIATNEPKLIHSPTYVEISODDEDEEGEQEEDI EIKDKEQAKVPQPKDYVNPRLVHVPPPTVWPI GSSGFGGSSVSSSSSSFGSOKDPSMAELKEKV GQLKERMARWLEIQEGTDARLGLLAQYALVS VQEMGSLQQLNTQVTRRSSTLQNLLEH
296	sp_103010_ortt	#N/A	#N/A	WRKY41; FUNCTIONS IN: sequence-specific DNA binding transcription factor activity; INVOLVED IN: regulation of transcription, DNA-dependent, regulation of transcription; CONTAINS InterPro DOMAIN/s: DNA- binding WRKY (InterPro:IPR003657); BEST Arabidopsis thaliana protein match is: WRKY family transcription factor (TAIR:AT4G23810.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MSKTKSTKTTSPQQQDVGNPPIKFLGVRRRPWG RYAAEIRDPSKERHWLGLTFDTAEAAALAYDRAS RSMRGPRTNFVYSDMPHGSSVTSLSLSPDNQN QHONQONHQNQOPEFSLSPGNDYDYKAVS NAQVSSSSSSSYOENPTIFTYNNGADDSVINTN YNNNCFYANSNDIELPLPDQFTGELAQLSWPS PAVGGNSSEYLGLDNNAGVYVADVPPFAFIAY HGDLQGLVRTFASOGETPQVIDVDADPFSKLTAD PNNSTQIMFGSIPATISMDTKAGSANAKDTNPA DGSKGAGTSSTNE
297	sp_110710_zzug	#N/A	#N/A		MEEVMGDFRSARLINELQGMFAKELNRIMNP ENKGCSDDDDDQVALNLSLEKMLSSFDKTTTI AKLLHIQLPHSPSEPHSLTISSPISENSHPDSDPP SKRRRVMMRWTKIVKMSSEKRGVEDGFSWRKY GKRDILGAKFPFGYRCRTHRTKGGCAATKQVQR SDGDESIQIMYRGEHTCFQAGAGAFDDIIPAQM KSEQQPPSPKQEYVYVGLSPEGLKVESEQEP ENRAQYFRSLSTPANLQVQVQVPPPIESYLSMYSYS DFDSTTTSVSNYLVLPYNHGLTVETSGTTEFFE AVSGPNVSTNSPMMDDDFEFPFIEELLDFGGNFSFE L

【 1 1 2 1】

10

20

30

40

【表 1 4 - 1 2 2】

	A	B	C	D	E	F
1						
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	分析に用いたバइटとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
298	sp_118630_coxg	#N/A	0.35365	#N/A	nuclear factor Y, subunit B11 (NF-YB11); FUNCTIONS IN: sequence-specific DNA binding transcription factor activity; INVOLVED IN: regulation of transcription; LOCATED IN: intracellular; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Transcription factor CBF/NF-Y/archaeal histone (InterPro:IPR003958), Histone-fold (InterPro:IPR009072); BEST Arabidopsis thaliana protein match is: nuclear factor Y, subunit B1 (TAIR:AT2G38880.7); Has 148542 Blast hits to 60498 proteins in 2540 species: Archae - 708; Bacteria - 15552; Metazoa - 38589; Fungi - 16830; Plants - 7103; Viruses - 1761; Other Eukaryotes - 68019 (source: NCBI Link).	MGKKEKATAVATEAEELPKTIVRRVYVKEKLSQISTD GGDVTLHKEALIAFSESARIFIHVLSATAHDCINDS NFGIINAEVDVLALEEVEFSEFVEPLRAALDEFKK KAAAGKFRAGSAKAKEASKKRKTDEETATKNVKAE VTATENGEEEDTATENGKDEETATENGEEEDTA TENGKDEETATENGEEEDTGTENGEEETATEN EEEEETATENGEEEDTATENGKEEENEDEEED EESENNDDNDDDDDDSPDS

10

20

30

40

【 1 1 2 2】

【表 1 4 - 1 2 3】

A	B	C	D	E	F
1	分析に用いたバイトとの相関係数	sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	sp_107620_kqnh (SOAP1)	sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
299	sp_120540_rwnz	0.9283133	#N/A	<p>Pentatricopeptide repeat (PPR) superfamily protein;  CONTAINS InterPro DOMAIN/s: Pentatricopeptide  repeat (InterPro:IPR002885); BEST Arabidopsis  thaliana protein match is: Pentatricopeptide repeat  (PPR) superfamily protein (TAIR:AT5G55840.1); Has  34111 Blast hits to 7189 proteins in 218 species:  Archae - 2; Bacteria - 9; Metazoa - 117; Fungi - 126;  Plants - 33142; Viruses - 0; Other Eukaryotes - 715  (source: NCBI BLINK).</p>	MLPRRATSLTSTVINGVCSSTSSFFPKYTRQPLSA FLSNSSSNKSNTHFNAPTITAPLVGNLVSFKYS HFWEHRIDRYTANSNLNQLLNISVVCPECRKFL RVSSWEAENWLEVLGQGLGSGNLRIDVEVVESL WEIFRRVSVRNGGFKHLPRSYEVNASLVNVGM FREVESLSTAEATRGILLANEEISGGIVEGINLGE LGKAILMYDTMRELCLAPLSLSCYRSLLGLLMKED KINQAIRVYEDMVDVGFGEIGDADAKNLSLIGFLC WEGEIREARKLKKIVGGWKFSTEILSEIADVYY EKKDFEDLLSFFIEMKCAPDVLVGNRISSLCSNF GIRADKFMMELENLGFVNPETITFGYICWSCLD GRIRDACYLSEILSRGLKPDHSHYNALSTMFKE GIQDHAREILNEMIDTGKPNLATYRVLLAGYSMS RQFNEVQLMMFEMVNHGLVLSPTDNPLSKAFAI LGLDPLTVKVRDNDVKNFRTEFEVDALGNGLYLD TDVDFEKTLLTGVLQESLIPDFNLLVMRECDVKN VKKALSMYDEMLQWQDELSLPSALVRLVCRS TCHVKAIPDFDKVFLFRKLDHETLNFVEALGK NRFVDQCRLLHEMYRRHLLPIFSATYASLVSLC QFGKSDLYGWSLWAGKDNWSPWDDDFIILENL CRRRMLKEALELFDMMMLKVNPGGLGTQICDVFLEK LSLNGFTVGVDILVVELKRNKFNELDCSAYSHLKG YCLEKRFSEALIFDKIRSDSSALSMDALLLIPOL CKVNRLEAVALKDAGSKDHPYTLFSLHSALIGC CRAGKVAEADGLLLGMLLNMMHPKDEIFNALLHA HCHNSSLTKTEKIGIMMKRIGICLSVSSYRDLYLL CKKGMFYSAQGLKSLVVVRNIPFVSVYVNIILFRL FQTGKHYLDVLDVLELEESGQQLDEVGNFJHJG FSACKDMKRSKMFNDMINKGLRPSNRS.LRAVH YLCRNGQVONALELNQIESRRWILCFPSVHYAM EGLFLDGRIGQAEERFYSQIMEKGGSTPETIYNNLI RNFRCRYGRLETAVNILLDQILKKNKSIDASSYDCII

10

20

30

40

【表 1 4 - 1 2 4】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1) 分析に用いたバイトとの相関係数	sp_107660_kiqg (SOAP2) sp_107670_ptqx (SOAP2-like)		注釈	aa seq
300	sp_121370_awkj	0.348093	#N/A	<p>lectin protein kinase family protein; FUNCTIONS IN: in 6 functions; INVOLVED IN: protein amino acid phosphorylation, recognition of pollen; LOCATED IN: plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Curculin-like (mannose-binding) lectin (InterPro:IPR001480), Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), S-locus glycoprotein (InterPro:IPR000858), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: S-locus lectin protein kinase family protein (TAIR:AT2G19130.1); Has 120498 Blast hits to 118775 proteins in 4615 species: Archaea - 117; Bacteria - 12995; Metazoa - 44949; Fungi - 10067; Plants - 34445; Viruses - 432; Other Eukaryotes - 17493 (source: NCBI BLink).</p>	MKPFCISKPTFRIFRFFRFFLSSAGADIPFLGSLSA SNPTQTWSSPNNSFSLHFLSLNSSSFSLSITYDS LPWISAVADAAATFSLADGNLRLLSGAGNTL WQSNSTNSGVSHATLDDSGNFCLKNRSNFVW STFDHPTDITVPSQNFVGMVLSGLYSFLNRT GNLTQWKNKDVYLSGLINATVSTNLSPRLGLQ SIGILSYDDNLGSMVIMYGSYAESGNILRFLKL DSDGNLRIYSSKGGSTSVRWAVNDCCQVFG YCGHMGICSYNDTSPVCGPSENFDPVPNDSS KGCKLKRIRDCPGSVAMQLMDHTKFLTYPPENL NDYFVIGISACFGNCLSGGSCVGTILADGGQ CYKTSFVSGYQSAALPSTSYKVCSPVEPNPP PSSEEGGGGGKGLRWVAVVLTLLVLI EGTLWWWCCRKSPKFGGLSAQYALLEYASGAP VQFSYRDLQOATLNFKELGAGGFGVYKGVLA NKTIVAKQLEEGEKGQFRMEVATISSHTHLNL VRLIGFCSEGHRLLYVEFMKNGSLDSFLFFATE NDKNGRVLWENRFSIALGTARGITYLHEECRDC IVHCDIKPENILLDDNYSKAVSDFGLAKLVNPRDH RRTLTSVRGTRGYLAPEWLANLPITSKSDVYSFG MVLLEISGRNFEVSDQDRKFKFNWAYEEFDK GIVKTLDRKIADSDVNDMDQVTRAVKVSFWCIG QPSQRPMPGRVVMLEGMEIEIKPPKPVHVEG SISGTLNMSISSALSTFATSAHGGPSSSSSFQ TPVIEELKELWGMAFPITAMNCLVYRAVAVSFLFL GRGLSLEAGGALSIGFTNITGYSVLVGLASGLEP VCSOAYGSONWDLSSLHRMIFLFIPIISFLW INLGPIMIFLGDQOEVTAAMAAYCIFSPLDLLTNT FLOPLRVYLRSQOVTKPMWCTLLAVIFHVPLNY WLVVMVLGWYKGVAVSSVVTNLINMAVMLGGYV CFKEKIQNQMNMKMMKMMWNNWGLKGVGVEG LKPLLRALIPISCLGICLEWVWVYEVITLSGYLPNP
301	sp_126610_yyhw	0.322693	#N/A	<p>2A6; FUNCTIONS IN: oxidoreductase activity; LOCATED IN: endomembrane system; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAINs: Oxoglutarate/iron-dependent oxygenase (InterPro:IPR005123); BEST Arabidopsis thaliana protein match is: 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein (TAIR:AT1G03400.1); Has 8330 Blast hits to 8293 proteins in 997 species: Archaea - 0; Bacteria - 1128; Metazoa - 114; Fungi - 883; Plants - 4883; Viruses - 0; Other Eukaryotes - 1322 (source: NCBI BLink).</p>	MDEYSNGSQPSYDRVKELKAFDDTKTGKGLVD SGVKTLPKFIRPSEELKEOKTPOVNLQVPIVDF DGKIKDRHRYSDIVKEVLDASEKWWGFFQVNH GISSELLDNMRSYRHFHEQDAEVRKLYTRDRA KQVLYQSNFDLYTSRYANWRPDLTVDTTCSGHL DPNELPEICRDMVVEYFNQVLLSDVLELLSMA LGLGPDRLKEMGCNKGWTAVALNYPACDPDKLT LGSKHSOSTFTTLLQDQIGLQVLENQWVNV RPIPGALVNIQDMLQIMNSDKLSVYHRVTANE VGPRISSAFFLKGILSSPKLYGPIKDLLSEENPAKY REFTISEFQGNFYSRPDLDFGQFYFK

10

20

30

40

【 1 1 2 4】

【表 1 4 - 1 2 5】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2					MANIIISQLLAARDGDVKFLTEALATQPADYFL VRSKNKGDGEHYNFHIVLNEQSEFLKEALGRL PVSTVHLLCNQASSDFNYNFHCAALRGNLAIVN LLVKYKSSSSSSSSSSSSSSSSSSSSSSSSSS TVDVNGKTPFHVALDRGRDDCALEIMSLDKELC NMVNDKGISPLFLAAQQGCAOVGVKILTSGHSYN TSGEDGSTPLDVVFNCSSEEVCKLLLDKHPKIMVK HYDENGRTLLHKWVMGKGLWPIQLLNREKPN MWRKFEINLLCTTEKLTNNINPLHAAONHNEAVV QVVQLLYEAYIDAYEQGEVLOPSPWKWENKDG DTPLMVSLNKHEELALYFLTVDMENLVHATKSL YYAVLNGONDAENILITVHPTSFDLMLKGGNG RNVLHIANCTERTGDILVEKLPWLINELDEYGVKS PLDIAAEEVGTAWLJELLLKKDPSLSANAPFAWIEA CKFGHKDAHFIDNSTDFRTLCCRKDSPLHHIQ LKSYYKYKEFLDIPIQEMKNMPDFTGATPLHRA LERRDIDLVEALLSIDGVFRNIKTNRKSATEILKE LCKEDYKWSMCRRIQINPWLRTEFISTGMANC MRYSGRSRSGTTPSAGEMRGTLSSVAALLATL FAAGFTVPGGLNEDDGEALAKAAFLVFLADTY AMCCSMLVLFCLIMSMVGDQDKSLILDRSMVILV QSLYGLTVAFTMGYATISHKSLWAAIIVIMCSIVA ISSNRAILDVLDKLIPISEVERERRR MIQRFFSAQLLGNVELIPSSNVLGIARAAATRDGVAK TMAASAARTRITLLELSTGGRLGSSSAGLN
302	#N/A	3384377	#N/A	ankyrin repeat family protein; CONTAINS InterPro DOMAIN/s: Ankyrin repeat-containing domain (InterPro:IPR020683); Ankyrin repeat (InterPro:IPR002110); BEST Arabidopsis thaliana protein match is: Ankyrin repeat family protein (TAIR:AT4G03500.1); Has 38664 Blast hits to 17395 proteins in 590 species: Archaea - 39; Bacteria - 2692; Metazoa - 19825; Fungi - 3120; Plants - 3302; Viruses - 203; Other Eukaryotes - 9483 (source: NCBI/BLINK).	
303	#N/A	3324983	#N/A	uncharacterized protein	MPASAVCQSTLYPSYCKDVLQSAKNTANVVEYG RFSFKRSVQAARFLKVVNKYLHDSRKLKLTPSA VAALQCCQLAEINLVDFLADSFNAAVNTQENKTL ALTTVEDVQTLFSLGILTNQOTCLDGLKSAESAWS FDRGITSPLSTDTKLYSIALSLFKAWIPKPKKGS KPPRGGKHPGFFGNGGRLPGLKMSDNARAVF EASLSRRHERLLGSDNEGDQVLYDVVTVSKDG SGNFSINDIAAIAAPNKTEASDGYLYLVYSAGVY EENVSIDKKRYIMMIGDGINQITIGNRSVADKVF STFGSATFVGTGTFVGMNMTIRNTAGAINHQAV ALRSGSDLFAFYNCSEFEGYQDTLYAHSRQFQYR ECDIQYDFIFGNNANVFCNCPNYPRLDNQFN AITAQRTPNQNTGISIENCIVIRASDDLANST TKTYLGRFWKEYSRVYLESFIDNVDPDSGWRE YHVMNATDASSFTVSTFVFGDEWLPTQGTGVPYTS GLL
304	#N/A	3341187	#N/A	ATPMEPCRB; FUNCTIONS IN: pectinesterase activity; INVOLVED IN: response to salt stress; LOCATED IN: endomembrane system, cell wall; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Pectinesterase, active site (InterPro:IPR018040); Pectin lyase fold/virulence factor (InterPro:IPR011050); Pectinesterase, catalytic (InterPro:IPR000070); Pectinesterase inhibitor (InterPro:IPR006501); Pectin lyase fold (InterPro:IPR012334); BEST Arabidopsis thaliana protein match is: Plant invertase/pectin methyltransferase inhibitor superfamily (TAIR:AT1G02810.1); Has 2755 Blast hits to 2700 proteins in 321 species: Archaea - 8; Bacteria - 597; Metazoa - 1; Fungi - 199; Plants - 1923; Viruses - 0; Other Eukaryotes - 27 (source: NCBI/BLINK).	

10

20

30

40

【 1 1 2 5】

【表 1 4 - 1 2 6】

I	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 Adaptin ear-binding coat-associated protein 1 NECAP-1; FUNCTIONS IN: molecular function unknown; INVOLVED IN: endocytosis; LOCATED IN: membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Adaptin ear-binding coat-associated protein 1 NECAP-1 (InterPro:IPR012466); BEST Arabidopsis thaliana protein match is: non-intrinsic ABC protein 4 (TAIR:AT1G03900.1); Has 463 Blast hits to 463 Metazoa - 232; Fungi - 69; Plants - 94; Viruses - 0; Other Eukaryotes - 66 (source: NCBI BLINK).	aa seq
305	sp_156390_kchz	#N/A	0.3371545	#N/A		MRADEWDVNVKWAWEGLVKVSKGEECHIRLEDK STGELYARAFLRKGEPPHVTVIDSSRYLLRIIE KTGE
306	sp_167450_zihe	#N/A	0.336471	#N/A	uncharacterized protein	MIHNRGRTFSGSRVMDQEPYVTFQVRNLYLEPIPL SDFSNASRRGTFRVYASMFEEPAATVQHGFTFR FGRDNATYEVGSSRQPVYHKKHQSSLCFSAPVY GNSQSLPVVSDDFNVVSISSANAAMNDPSPNSGIA ATAVISSAMPADRLNALDQRVVEGKQVVFSEGA FSDIPSPQRSFIDSPPLQVSTYQPRMFSESLYLR SVNVIIPPLMIPKASPVLFSSVVMKQVNVVSSL KDPPTPTGTKSTGLSPTASLA
307	sp_168240_fyem	#N/A	0.326835	#N/A	MLP-like protein 423 (MLP423); INVOLVED IN: response to biotic stimulus, defense response; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Bet v I allergen (InterPro:IPR000916); BEST Arabidopsis thaliana protein match is: MLP-like protein 28 (TAIR:AT1G70830.3).	MGVFTYTLADITSPVAAPRLFQALAIIDNHNLMIPKV VPHLVKSVFEVFGDSTTVGCVKQINLPDEAPFKY VKNRVDEIDASNLYLKYTCIEGDVFPETVEYVYVE DKYEPSEAGTRCKMVAHYHLKGDNVMMKDDDLV KAKEGIQKIFKAVEDHLIANNPQVYA MLRLARYVLOQSSQLGRVSCSPRIWFSTMFQ TEGQSQVTSNDFGKENSNDHSLFLGQKITRKEK TRFLNALDLEDSKDAYGTLDAAVVAWEQNFPIA ALKNVLLALEKEQWHRVQIKWMLSKGQGT VGTYQQLIRALDMDHRAEEAHNIWMKKIGNLHS VPWQTCHELMISVYHRNNIMLEHLVLFRRLEAFH RKPPEKSVQKVADAYELGLVDDKERLLKYADL LSKNRDESKMKSINPVRKSRKKTSSGSRSDVANG EKA
308	sp_171070_eimc	#N/A	0.322811	#N/A	BEST Arabidopsis thaliana protein match is: Pentapeptide repeat (PPR) superfamily protein (TAIR:AT4G21190.1); Has 111 Blast hits to 111 proteins in 15 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 109; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLINK).	

10

20

30

40

【表 1 4 - 1 2 7】

A	B	C	D	E	F
共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 DOWNY MILDREW RESISTANT 6 (DMR6); FUNCTIONS IN: oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, 2-oxoglutarate as one donor, and incorporation of one atom each of oxygen into both donors, oxidoreductase activity, INVOLVED IN: response to fungus, response to bacterium, flavonoid biosynthetic process; LOCATED IN: cellular component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Oxoglutarate/iron-dependent oxygenase (InterPro:IPR005123); BEST Arabidopsis thaliana protein match is: 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein (TAIR:AT4G10490.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).	aa seq  MEKFSVSDGKTLPENYIFPEQRPGKLLPSTS NYVTPVIDLGLKLDGPDHSIVVHQLQACQELGAF QVFNHGVARRLMDDARQMLDEFFKLPADKSAK ITSDPNSKFKMYTSTVNYDEEKIHLWRDAFQOVC TPLEECIHPWPAKPAKYREIIEEYTRKVNELGLRIL DLICEGLGIEAGYFANEQSQHQHIELVNHYHPDDEWV GVDPPEAFVYHGNIMQIASNGKLLNGDHRVAVTN AYEARESAAALFFMPTKDCGFIEPAKALITAEENPLY RGFRYHEYLAAHNTARRLKDPDLLLLKPFRI
sp_133850_nuza	#N/A	0.331253	#N/A	polygalacturonase inhibiting protein 1 (PGIP1); FUNCTIONS IN: polygalacturonase inhibitor activity; INVOLVED IN: signal transduction, defense response; LOCATED IN: cell wall, plant-type cell wall; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Leucine-rich repeat-containing N-terminal domain, type 2 (InterPro:IPR013210), Leucine-rich repeat (InterPro:IPR001611); BEST Arabidopsis thaliana protein match is: polygalacturonase inhibiting protein 2 (TAIR:AT5G06870.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).	MSNHNKISILSLLLLSSTISAEEKCHPDHDKKTLFKI KQAFHNAYLFAFWTHDTCCEWYLVKCDVTVN RIISLVSFRDDDEVAGSIPNTVGDLPYLEYLGFTNL PKLTGSIPOQSISKLNKLSLWISHTNLGTGSIPIDFVG RLKTLTYINLSVNLKLTGPIPPSLSSLHLKLSALFLEN NQLKGTIPNSFGGFKNNPFDYLLKAKNQLSGPLP KYLGAHFVSVLDSRNMFMTGDSILFGVVKGLQT LDSIRNKVFDLTKVGLPNLNLNLSLSSHMMIYGS PPTIAKVSYLQQFN MNMIDKGSRRKTKLTWKVVECPRSQQVWVWVESG YYVMKYMDDQTVSLVSKKERDF
sp_202780_mzjw	#N/A	0.321763	#N/A	uncharacterized protein	
sp_204930_xign	#N/A	0.356533	#N/A		

10

20

30

40

【 1 1 2 7】

【表 1 4 - 1 2 8】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh sp_107660_kiqg (SOAP1)	sp_107670_ptqx (SOAP2-like)		注釈 actin binding; FUNCTIONS IN: actin binding; INVOLVED IN: cellular component organization, actin cytoskeleton organization; EXPRESSED IN: shoot apex, stem; CONTAINS InterPro DOMAIN/s: Actin-binding FH2/DRF autoregulatory (InterPro:IPR003104), C2 calcium/lipid-binding domain, CaLB (InterPro:IPR008973), Tersin phosphatase, C2 domain (InterPro:IPR014020), Actin-binding FH2 (InterPro:IPR015425); BEST Arabidopsis thaliana protein match is: Actin-binding FH2 protein (TAIR:AT5G07770.1).	aa seq  MLKSKIKSFPCFLVGSMAALRPLPLPPFPAPS FLAAPLNPNFTGDASGQLHRSPLLSSLORRCA VTNAPTFLHLAPYLVLGRQSERIHNLAFAHATALE LVTDFKSLNTVNSACDEVRNSGNLKEIMQLLNL GNLLNSGTARGDYVQHKNFSTSEFTPYHSSAKGT HHIRHLWGYGL  MVVIDCHVIGIFVSVLLGFLVLLFSIFRDVKSKSXS ANKKDGRRNSVGQTEVARSSDDVCNVSDGPD VIIVGAGVAGAACTLAKDGRRVHVIERDLSEPD RIVGELLQPGGYLKLIELGLEDCVGNIDAQRIVGY GLFKDGNKLSYPLEQFHSDVSGRSFHNGRFIQ RMRKASTLPNVRLEQGTVTSLEENGMVKGIN YKNKNGEVMKAYAPLTVCDGQFNSLRSLCSP QVDVPSHFVGLLENCQLPFANHGHLVILDFSPV LFYPISTEVRCILDVDPAGDKLPSIGNGEMAKYL KTVVAPQVAEIRDSFVAAYDKGNIPTMPNRSMP AIPQPTPGALLMGDAFNMIRHPLTGGGMTVALSDI VVLRLDLKPLQDLNDAASLCSYLECFYTLRKPVA STINTLAGALYKVFCAASPEAHKEMRDACDFDYL LGGVFSNGFPIALLSGLNPRPMSLVAHFFFAVAVYG VGRMLIPFSPYKRVCLGTRLLSAGAGIIFIMKAE GFRQMFPLATVPAYHRRHHPVK
312	sp_004340_zdis	G927736	#N/A	#N/A	XF1; FUNCTIONS IN: squalene monoxygenase activity; INVOLVED IN: response to water deprivation, sterol biosynthetic process; LOCATED IN: endomembrane system, integral to membrane; EXPRESSED IN: 29 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Squalene epoxidase (InterPro:IPR013698); BEST Arabidopsis thaliana protein match is: squalene epoxidase 2 (TAIR:AT2G22830.1); Has 1994 Blast hits to 1990 proteins in 731 species: Archae - 43; Bacteria - 1249; Metazoa - 112; Fungi - 225; Plants - 178; Viruses - 0; Other Eukaryotes - 187 (source: NCBI BLINK).	
313	sp_008240_esex	G952739	#N/A	#N/A		

10

20

30

40

【表 1 4 - 1 2 9】

A	B	C	D	E	F
1	分析に用いたバイトとの相関係数 sp_107620_kqnh sp_107660_kiqg (SOAP1)	sp_107670_ptqx (SOAP2-like)		注釈	aa seq MWAFAWNTTSPFFCSSTTSTVAVVAVT TDDSDSFC LRTGSSPRTQIIRDEESDFIAHPTRTSLSMLIQIR AQESEFASPTRRPSATGEIRCSCHPFAENS NGGNAGGKPYRNKSSNGPARRKSLIAFRKAE KDEEKRALEKIQKVEEDKAERRRKLGLPPKSKG SQQHVSNCLAGGTHESIHCIDEREALLQFKQGI QFDRGGLSSWQEQASQHHLECCQWEGVHCN NHTGHVTLALRGVAGDQPCLEGTVSTSLKIK HLKYLDSLNDFFGGEVPSFIGFTNLQHLNLSK AGFSGEIPHQLGNLFLASLDLSCDFGNSYGCY ADSLKWLHLTLRDLISLGSIDLSTQSDDWYQVA NNLTSVLHMERCGLSPSILNSISYMNSSSFSV IDLNNFNDSIFKWLFLNKGIGSTLVHLDLSDN KIPGPIPHNGKMLSLSYLDLSSNLLQGPISAFW NMRSISYSLATNQLGEPIDTISNLKHLHLISG NAFQGRGNFPKSLGSLQNLQDIFFNNDLDEL STIQALSQCVRKSLVLELSONRIWGSIPDIISTF SLLDGLLNTNRLNGTISPGIQQLMLEILDLSYNS LKDTISHNLNLNLSLSDLSDNPDVYHISPDW VPPFQLTSLILESKLGFSPFKWLVQTNLSYIGI SNAGISDPTISFWNSLSPSNLVYLNMSYNMISG VLPNLSVTLKPIVDLSSNILSGAIPSFRLNATKL YLNNRFRDLSPPFCPKQKTSLEFLNANLFGS EIPDCWYFDQLLVHLENNNFRKLPASIGALT LYTLHLRNNSLSGELPKLQNLCTSLVILDGYNLT TGKIPQTIQHSFENLRLSLQNNFVGLPLSLCQ LSQLQILDVSNHISGILPKLYNLAMTGTEDGLP DLVYDNYVNYYYSEYAFRLMWRKKEQRFQSRM GOVIADISDNELEGHIPEGISSLTGLIFLNSRNLH SGNIASKVGLVSLFDFLSDNHLSGEIPFTLSEV ASLEILDLSNNLSGKIPVGTQLQSFDSSSYMGN PGLGGLPLPKCSGDQPPANNVNVQEDDAFSEQNED
314	sp_011450_qrpi 2,752,349	#N/A	#N/A	disease resistance family protein / LRR family protein; INVOLVED IN: signal transduction, defense response to fungus, defense response; LOCATED IN: cell wall; EXPRESSED IN: 10 plant structures; EXPRESSED DURING: 7 growth stages; CONTAINS InterPro DOMAIN/s: Leucine-rich repeat-containing N-terminal domain, type 2 (InterPro:IPR013210), Leucine-rich repeat (InterPro:IPR001611); BEST Arabidopsis thaliana protein match is: Leucine-rich repeat transmembrane protein kinase (TAIR:AT4G20140.1); Has 135121 Blast hits to 32840 proteins in 1181 species: Archae - 56; Bacteria - 8201; Metazoa - 34800; Fungi - 1569; Plants - 79859; Viruses - 2; Other Eukaryotes - 10634 (source: NCBI BLINK).	

10

20

30

40

【表 1 4 - 1 3 0】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2					<p>MGHQENGDSNNALKFIEDVTSRCDEYQCQVLA EILSQNAEVEYLKRYGLPKGCVDREKFKSKVPVV MYEELKSDIQRIANGDFSPILCARPFTQLFSSSGT SGGEPNLIPTADQEIPIRRHQYKVLPLLSQCFP GMESGKLLCGFYQRHEVVRVDSHFASLWVHIF LRDHYSELCDMDSSGT.LSSKISDMALRAHMAEH VVKQPEPELAKLVYNAKGENWEGILKKIWPNIK YLDIITGSMAGYIPIINFYSGGVPVIVSLYKTCSEGA FGNINLPICDPYEVSYTLIPNMMAYFEFLPINTSSS DREGSSDQLVLDADVKGHEYELVWVTTYDGLYR LRVEDVLSPTGFYNSTPQKFRVRRNNVLLSVHAE KTTESELQNTMEKYVDYLNQFNFTIRVVDYTSYAD VKTIPIGHYVYCELSITSDATNKLGLDGMMEKC CLTIEDSMSYVYRYRVHNPAPGPLEIRVVKKGTF DKMRDFAISRGASLQYKVPKPRVYKFIPLLELLNS GVISTHFSPLPHCSL</p>
315	sp_011710_wkjp	#N/A	#N/A	<p>BRU6; CONTAINS InterPro DOMAIN/s: GH3 auxin-responsive promoter (InterPro:IPR004993); BEST Arabidopsis thaliana protein match is: Auxin-responsive GH3 family protein (TAIR:AT1G59950.1); Has 1707 Blast hits to 1543 proteins in 239 species: Archae - 1; Bacteria - 597; Metazoa - 55; Fungi - 2; Plants - 674; Viruses - 0; Other Eukaryotes - 378 (source: NCBI BLINK).</p>	<p>MGEDEDSLIVNLPIDLLAPDRVSTAATLRQACIE VGFYVIVNHGVYDKLLGRLEESKFFDLPEEKE MKLDRKEIRNGYTPYSEKLDTSINTKGDSSKESFYI GPIHDELINLQWVPEELLNWRATMEAYHKQI LNAGKSLSLAPALNLNENFFENVGALDQMPV LRLHYPEGDNNCNDVLDGASAHSDYGMITLLV TDGVPGLQVCREKDRPQRIWEDVPHVVEGAIWNI GDITERWTNCLFRSTLHRVVPKGRERFSAALFLD PPHNFVSVECLETCSSSESNPPRFPPIRYNDYKLER YRVTYGA</p>
316	sp_011890_wuzw	#N/A	#N/A	<p>2-oxoglutarate (2OG) and Fe(l)-dependent oxygenase superfamily protein; FUNCTIONS IN: oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, 2-oxoglutarate as one donor, and incorporation of one atom each of oxygen into both donors, oxidoreductase activity, iron ion binding; INVOLVED IN: oxidation reduction, flavonoid biosynthetic process; LOCATED IN: peroxisome; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: T3 growth stages; CONTAINS InterPro DOMAIN/s: Isopenicillin N synthase (InterPro:IPR002283), Oxoglutarate/iron-dependent oxygenase (InterPro:IPR005123); BEST Arabidopsis thaliana protein match is: 2-oxoglutarate (2OG) and Fe(l)-dependent oxygenase superfamily protein (TAIR:AT4G16765.1); Has 6583 Blast hits to 6566 proteins in 839 species: Archae - 0; Bacteria - 934; Metazoa - 52; Fungi - 818; Plants - 3817; Viruses - 0; Other Eukaryotes - 962 (source: NCBI BLINK).</p>	

10

20

30

40

【 1 1 3 0】

【表 1 4 - 1 3 1】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	共発現遺伝子			GDSL-like Lipase/Acylhydrolase superfamily protein; FUNCTIONS IN: hydrolase activity, acting on ester bonds, carboxylesterase activity; INVOLVED IN: lipid metabolic process; LOCATED IN: endomembrane system; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Lipase, GDSL (InterPro:IPR01087); BEST Arabidopsis thaliana protein match is: GDSL-like Lipase/Acylhydrolase superfamily protein (TAIR:AT5G45670.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLink).	MMRGAVAVLTLIIITVRCPEQVPCYFIFGDLSLSD VGNNDLVTKAKANYPPYGDIDFGAIPTRFNSNG RVFVDRITELLEFDNYLFPSSAKGEEILKGVNYAS GAAGIRYETGGHGERISMYDQIMHNETVAKIK EIGKTSAPVKLEECLYVNIIGSNDYINNYFKPGYY YISSLLYNPDEFAEELINQYAANLKLHSLGARAKIA VFGGLGICGSEGEIYFTGTNGSLCVDKINAQVQPF NQKLVNMVDDLTMELSDLQLTYIGSAASTDTGSA IADAPCKVRRDDFCIEFEACEGRNYYIFWDGF HTTEAVNVLTEIAYSALDTKHHHPVDIKTLVSIDR ESMINQVKNSSLSYSSA
317	sp_012520_wxgw	#N/A	#N/A	Lung seven transmembrane receptor family protein; LOCATED IN: endomembrane system, integral to membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Transmembrane receptor, eukaryota (InterPro:IPR009637); BEST Arabidopsis thaliana protein match is: Lung seven transmembrane receptor family protein (TAIR:AT5G18520.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).	MALHSATLTLIFVALFSLSTAEIRTTKVKSDNR PIPIDFEGFTHGRLELNVSGITLSTRNKLDSLK GFFLCTRDSWLHVLEQLEEEISECALDHLVKVY YTFNSLPTDPSFGSLYYETDSDQYTLVFANCY PENVQISMNIRSAMYNLDGKTRTRDYLVSAGRTL PRIFYFFSFIYGIAGVFWTYLYKKRLTVFGIHFFM LGVIMKALNLLCEAEDEKSYIKRTGSAHGWDLV YIFSLKGLTLTLVIGTGWISFLKPYLQDREKKV LMVIFLQVAVNIAGVVIDETSPNQDQWVTWKQV FLLVDVICCAVLPPIVWISIKNLREAAATDQKAAV NLMKLTFRHYIYVICYIFTRVWVVALETITTSYR YLWTSVVAALATLAFYVFTGYKFKPEAHNPYFA IDDEEEAAVEQLKLEDEFEL
318	sp_013260_kgkl	#N/A	#N/A	uncharacterized protein	MEDDDQPPTFEELLLLEELLESQQLSDDLLDVID IPTMDFLSMEQOMNQRFRDDEASQAEVWYLSDD DNNVDYEAHINHYDEHHSDIISSEYFQNMQPSET SNTNRNHGKTFELSDNQRVAVAKMLLANKD MTLPWGVAKKLANEYDTEVPPDLLTSIRQYMQ EGOININEFRQRVNEIMGVREDPFYQNGRHRQ
319	sp_014380_pjhe	#N/A	#N/A		

10

20

30

40

【表 1 4 - 1 3 2】

I	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
320	sp_017120_eipt	0.923119	#N/A	#N/A	F-box family protein; CONTAINS InterPro DOMAIN/s: F-box domain, cyclin-like (InterPro:IPR001810), F-box domain, Skp2-like (InterPro:IPR022364), F-box associated domain, type 3 (InterPro:IPR013187), F-box associated interaction domain (InterPro:IPR017451); BEST Arabidopsis thaliana protein match is: F-box and associated interaction domains-containing protein (TAIR:AT5G52610.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MSDIPPEVINKILERLPVKSLLRFKCVSKTWYSLIS NRNLEFAINNLKHLSSNDNTLLAIFNDNWLSQS NEFDIKDPFVTRFDLGPSEYDSEYCFNARIV GSCNGLVAVLDGCLCLMVSJKTELFRDITVWV MKEYGVKESWAPLFSLSSEIDIFKPLPMPLIAFSKD EFLVRQDVGWVLRITLPLIYADSNLSQSLMAYAL VLRTLHLYADNLTVPSPGVGDALGCHCCVGSLLV VFNQVGPQNRQ
321	sp_017920_yraq	0.923119	#N/A	#N/A	uncharacterized protein	MGLKHLVSRPPTPSPSPSDSMTRNKNGYYVDR VSVQRPTTYMDSVANSMPMERSTPSFHVERSTP SIQAESEIMFNYSNDEFKSVKGVQPRDIWITD EKGFYYKLNFEFCQLRKGGSILVRFELGDIKQ DSLCPVGIKTRAVNKKLSNIVTMVREHFVIFDG LVVDKAVRRAGKSWKNHRYLILKQYDFPVNKTLL DQNYDSIPDGISSSSWINDLVNYYWFSLKGTKKKEK GEYRELTFYKSVYAKEDGRFKEGTISRQFMEDA NNKVOENLASSSTSCKSKEIENE
322	sp_020030_fqha	0.923119	#N/A	#N/A	uncharacterized protein	MAREITVVAHNFHFLTEKSKVWVSTRIQFFYK TTVTIAPLFKASLGP
323	sp_020300_mcup	0.923335	#N/A	#N/A	Phototropic-responsive NPH3 family protein; FUNCTIONS IN: signal transducer activity; INVOLVED IN: response to light stimulus; LOCATED IN: plasma membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: NPH3 (InterPro:IPR004249), BTB/POZ fold (InterPro:IPR011333), BTB/POZ-like (InterPro:IPR000210); BEST Arabidopsis thaliana protein match is: Phototropic-responsive NPH3 family protein (TAIR:AT3G50840.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MAEKSGTSTPSKQAWFCCTTGLPSDIIVEVDN MTFHLKHFPLMSKSKLHDLIREQETKNLPTNK RRQDTQNDVDESKEEEDDEEDENLEGGQCH VSLTDFPGGSEIFTAAKFCYGVKIDLTAAKVAPL RCAGEVLEMTTEEYSEDNLISKTERYFANTVLLKNIK DSVKTLKSCFKLPPLENLGVDRCDIASSSKASLY DPTLFGWPVNDVVEESRKKKKKNAAPPSKVSD AWVEDLAQLSSHLFRVISTMRMKEVNPVIESC LIHYAKRCFFGISRSNRKPTATAPSMMASEAEQR QLLETISNLPENHTRSPQAAATSFLLGLRTANI LNASETCSALEKKIGAGLNQATLDDLLIPYSYL TETLYDVCIERILSHFLANHEEREIEGEGDHTL GNVRTPAIMHVKGKLDIGLYSEIASDANLKPQKFIQ LALSPPKEKARLYDDGVYRAVDVYLKAHPWVLEE ERDKVSGVMDCQKLTLEACTHAQNERLPLRAV VQVLFQEQLQRAHJAGTLITADIAPISESTPPSLLG HDEEEERDRGGGGGGGGGGGGTWEKAIKENQ VLRFLDMSMRTRVYELERECSSMMKVKIEKFDKD KVKGVVNSPKRWRSTMS

10

20

30

40

【表 1 4 - 1 3 3】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
324	sp_022650_pgkf	#N/A	#N/A	amino acid permease 4 (AAP4); CONTAINS InterPro DOMAIN/s: Amino acid transporter, transmembrane (InterPro:IPR013057); BEST Arabidopsis thaliana protein match is: amino acid permease 2 (TAIR:AT5G09220.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).	MNFDDAGRIRGTWWTATAHAHTAVIGSGVLAVP WSVAQMGWIFGFLSLFGFAVITYITARLLCDCYR TPDPVKGRNHSYMDAVRATLGKKEVLIKAFQAQ YSLWGTVMVGYTITAAATSMLSVKRNTCYHTKGA EADCCPSSNFFLIFGAIHIVLSQLPSLEKVTFLSV IAATMSFCYSFIALYLCVAKFVSNPVRGTLTGVE IGINGGGLSLSKWLQSSQALGNIAFSYTYAQLLI EIEDTLKSPPAENKSMKRATLYSIALTTTFVYVSLG CMGYAFAGNEAPGNVLTAFHEPFWLVDLANIGVV IHLTAAQVFAQPIFASYEKWKLAIRWPESPLIHT YTIQVPRFKALSYQFTLCRLLRPLVVVVTVLTAM MFPFFNAVLGLLGSIAPFFLTVYLPVRMYMITQVK LNKGALKWCMQLQGLSLLALFVSLISAVGVSADIA DRLKNTKLFNANI
325	sp_026020_gong	#N/A	#N/A	basic helix-loop-helix (bHLH) DNA-binding superfamily protein; FUNCTIONS IN: DNA binding, sequence-specific DNA binding transcription factor activity; INVOLVED IN: regulation of transcription; LOCATED IN: nucleus; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Helix-loop-helix DNA-binding domain (InterPro:IPR001092), Helix-loop-helix DNA-binding (InterPro:IPR011598); BEST Arabidopsis thaliana protein match is: basic helix-loop-helix (bHLH) DNA-binding superfamily protein (TAIR:AT3G56770.1); Has 2073 Blast hits to 2066 proteins in 145 species: Archae - 0; Bacteria - 2; Metazoa - 255; Fungi - 6; Plants - 1802; Viruses - 0; Other Eukaryotes - 8 (source: NCBI BLink).	MPHFHFNATNTSSSMHLQELPANCYALEFQEPF EGFAEFEDKASADSSHNHROAEKRRDRINGQ LARLRRLIPKSEKMDKAALLERVVEQVKDLKRKA SEISKFITPSEIDITIECHRKENHPNHNNNNNIN NNSNSNIMFMKATLCCEDRPDLIPELIRELKDNL TVIGADIGCLGGRMKSLLLYYKLDNSNDGVVVCNV SAIKOSLMGVLTRIIISWSSASTSSFRVTSKRQRRF FSS
326	sp_026820_nizic	#N/A	#N/A	Ankyrin repeat family protein; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; CONTAINS InterPro DOMAIN/s: Ankyrin repeat-containing domain (InterPro:IPR020683), Ankyrin repeat (InterPro:IPR002110); BEST Arabidopsis thaliana protein match is: Ankyrin repeat family protein (TAIR:AT1G34050.1); Has 20807 Blast hits to 11728 proteins in 551 species: Archae - 16; Bacteria - 1568; Metazoa - 11429; Fungi - 1235; Plants - 2476; Viruses - 139; Other Eukaryotes - 3944 (source: NCBI BLink).	TEETFRKEDIESMKAKEASHWLKDGLTPLFR AFQTNLSTATELCTISSEAGTCDANGQTFWH LLVHFPSGYSYFVFRLLQNEVLRDLLKLDKDKKGNTP LHLAIEGNRYDLVKDFLDWIWERERELLPKGKDQW LVNLLQIQNKAGKTPADLIRESFSLPPNVSHF

10

20

30

40

【 1 1 3 3】

【表 1 4 - 1 3 4】

	A	B	C	D	E	F
1						
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈  indole-3-acetic acid inducible 32 (IAA32); FUNCTIONS IN: sequence-specific DNA binding transcription factor activity; INVOLVED IN: response to auxin stimulus, embryo development ending in seed dormancy, pollen development; LOCATED IN: nucleus; CONTAINS InterPro DOMAIN/s: Aux/IAA- ARF-dimerisation (InterPro:IPR011525), AUX/IAA protein (InterPro:IPR003311); BEST Arabidopsis thaliana protein match is: indole-3-acetic acid inducible 34 (TAIR:AT1G15050.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archaea - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	aa seq
327	sp_027800_hwal	0.4923119	#N/A	#N/A	uncharacterized protein	MESNMASFLNLTSAIDLGLSLNTHIHTOACDPVDQ RRQGMQLHSTIIPGGYNNHNHMEMWPYLKSPFK SKSNPDDCGEDLEGVESKERWPYVKYNMIDGVIV GRKICLDHLYGYPALQLEDIMFGGFTFYGVRLF QRGSEFMLFYKDREENWRSAGDVPWKDFVDCA KRLFRVWRKRVLLS
328	sp_032290_xsaw	0.9754432	#N/A	#N/A	uncharacterized protein	MGTDSVAHYPSVFSDDCSYGGDGGSELDGVA VPPVGMYSYRRDMVDVDFSVPAEEVMEVLSDYVF R
329	sp_037830_teft	0.453119	#N/A	#N/A	uncharacterized protein	MDEA
330	sp_039740_fjzg	0.4923119	#N/A	#N/A	uncharacterized protein	MVSNAMAILRLETRRRRRGGGYPYRSGEGRSRI RVGDDLIGN
331	sp_040000_zixd	0.356624	#N/A	#N/A	Eukaryotic aspartyl protease family protein; FUNCTIONS IN: aspartic-type endopeptidase activity; INVOLVED IN: response to salt stress; LOCATED IN: cell wall, plasma membrane, membrane, plant-type cell wall; EXPRESSED IN: 6 plant structures; CONTAINS InterPro DOMAIN/s: Peptidase aspartic (InterPro:IPR021109), Peptidase aspartic, catalytic (InterPro:IPR009007), Peptidase A1 (InterPro:IPR001461); BEST Arabidopsis thaliana protein match is: Eukaryotic aspartyl protease family protein (TAIR:AT1G03230.1); Has 1398 Blast hits to 1391 proteins in 46 species: Archaea - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 1396; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLINK).	MASGPFLLSFLFSAFHSFNANLPLVVKDNT KSQYITFLVGRTRPSTRVYLVIDLANQWSWFAC DLYGYNTHSHEVPCNSPKCERYAPGTFCSGNAC GALAYTPWSNGLFAGDLYEDTVSLYKQTQONS AQLSKVYLPKLPFTCADKGPLKGLSPRTKGVVTL ARVPVALHSQISSSTFKVARKFALCLPSSISLGYN GAIYFGGSPDYLSKSLITTLPLVFNPNSTAPYSEGD ASVEYFIKVRSIKIDGTPIRFRSSLLSFDIKEGVGG TKVSTMDPYTLVHSEIYKTLVNVFVTKATAMKMT RVASVSPFGTCFSSKSVVDNAKTGPAVPIIDIMD GKSPLLHNKNRWRIYGANSMVKVSDTVMCLGF VDGGSKPTTSIVIGGKQMEDNLVEIDLETSLKLGVT SSLLRMGTNCSKFKGI

10

20

30

40

【 1 1 3 4】

【表 1 4 - 1 3 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
332	sp_040030_zdic 0.922371	#N/A	#N/A	Eukaryotic aspartyl protease family protein; INVOLVED IN: aspartic-type endopeptidase activity; cell wall, plasma membrane, membrane, plant-type cell wall; EXPRESSED IN: 6 plant structures; CONTAINS InterPro DOMAINs: Peptidase aspartic (InterPro:IPR021109), Peptidase aspartic, catalytic (InterPro:IPR009007), Peptidase A1 (InterPro:IPR001461); BEST Arabidopsis thaliana protein match is: Eukaryotic aspartyl protease family protein (TAIR:AT1G03230.1); Has 1398 Blast hits to 1391 proteins in 46 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 1396; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLink).	MASSNRRTFSLFLVMSLSALSFHFSNAHFVLPVVK DNTNSKYSYTSLVGNHHSPRVNVLVDLSNKWS WFTCDFGYNITTSHTGVPCHTPKCDRYAPLAPCS DGCVVDAAYTPSSNFIYSNNLLEDTTKVYKIQOQN SLTLQSNVNLNFPFTCGDNVPLKGLSPHTKGV TLARVHVSLSHQISSTFKVARKFALCLPSSPTSS GYNGAIFGGSPNYLSKSLITPLVINPNSSTPIYS PGDSSVEYFIKVKSEIETPIRFRSSLLSPDIKEAV GGTKISITMDPYTLHLSEIYKTLINVFITKATVMKMI RVASVAPFGACFSSKSVVNDAKTGPKYPIIDIMD GKSPLFARWRIYGANSVMVKVSDTMCLGVFDGG LKPTTSVIGGKQMEDNLVEIDLET3KLGVTSSLL HLGTSCSKFKGI
333	sp_041370_tnfi 0.922326	#N/A	#N/A	Expressed protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLink).	MAVVTQEPILCRDLRLNLMRHLNIRGSSNNIN NNNNSNRSSPSTPSSGTARNSDGHASSLDLFFS PRSRDKLHCRPIDDVIVETEHTKGTLLDRVSLIEHR LLKLCMQLEQEMDAEKQDKHSQDENKSRSPQT HKKSGLKQFVKSVCVKGINKHKPNK
334	sp_042210_agke 0.934436	#N/A	#N/A	GATA transcription factor 1 (GATA1); CONTAINS InterPro DOMAINs: Zinc finger, NHR/GATA-type (InterPro:IPR013088), Zinc finger, GATA-type (InterPro:IPR000679); BEST Arabidopsis thaliana protein match is: GATA transcription factor 9 (TAIR:AT4G32890.1); Has 1635 Blast hits to 1587 proteins in 195 species: Archae - 0; Bacteria - 2; Metazoa - 169; Fungi - 549; Plants - 842; Viruses - 0; Other Eukaryotes - 73 (source: NCBI BLink).	MESFDKAARFMDLDFSSDIGEEESDDEIIPGNS SKSHHFLHHRHLNHRHRNRQSPFAEFVEEEL WISNKDAPMVEESLVLDLPNHPRIASSYNTNHS PISVLDNSSSSSANSLSLTHGSLNNNINSNG GRGNMMLINLSSDFKVPKARSKHGRKRKRD MLGSQFKCFVYEKKGARKLASSVSASSKSSSTL GRKQLHCGSEKTPQWRAPGSPGPKTLONACGVR YKSGRLCDEYRPAASSPTFSSDLSHSNHRKIVEMR RQKMPLTGSGSNSLEIG MTPFGSRSRSGFGRGGGGGGGGGGGGGAGA GGGRARRSDGWVPPRRIKNSV_LHFNNGTCT SCASQIPRQQQQQQQQQQQISLTMPSISISYV HNVSAPPYSIQGGNPPRLRFLCKELTYSDVKK RNMKILPRVIVASKERNEVLNGECSVPNNGDLGM GLMTSSYSPSACNYYVNLQLSDFCVDEMEEVA TPRTPYVGNATWTATIVSFDLPLPLQTFNLDFPI DNMSWT MFKMGKPKIAYGFIILFSLSSISISIDGACSPSEDN NKYPVIGILSHPRRNG
335	sp_047220_sodq 0.923119	#N/A	#N/A	uncharacterized protein	
336	sp_048110_xhda 0.940733	#N/A	#N/A	uncharacterized protein	

10

20

30

40

【 1 1 3 5】

【表 1 4 - 1 3 6】

I	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたバइटとの相関係数 sp_107620_kqnh sp_107660_kiqg (SOAP1) (SOAP2)	sp_107670_ptqx (SOAP2-like)		注釈	aa seq
337	sp_048300_qhsr	0.923112	#N/A	#N/A	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein; FUNCTIONS IN: lipid binding; INVOLVED IN: lipid transport; LOCATED IN: chloroplast thylakoid membrane; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Bifunctional inhibitor/plant lipid transfer protein/seed storage (InterPro:IPR016140), Plant lipid inhibitor (InterPro:IPR003612), Plant lipid transfer protein/hydrophobic protein, helical domain (InterPro:IPR013770); BEST Arabidopsis thaliana protein match is: Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein (TAIR:AT5G46890.1); Has 782 Blast hits to 776 proteins in 62 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 782; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink).	MASKAFASLTLCLNLLFFLLVSVSYAPNCPSS GHHGGHHGGHHGGHHGGSPSTPTTPMPPAQA KGFIDALKLGVCANVLGDLINVVVGTPPKTPCQPL IQGLVDLEAAICLCTAIKANILGINLDIHLSLTLLNY CGKGVSPSGFCQPY MEGKQKCRITFCPKGAWRPLPILPYWDRKRLDD RVKELDSLKVVSSNSAPSATEATKCAIMSHSPEN FPGKTNVVELEKVVSVQPCSSACSNTAQDIYE VPCSYIDSSNIKESYNEPILLSPEAKSSKSDS KDKIDWEKPLLLSPGSSNESPKLDRKTDKPLLF KHPTLGGVADFELOQFNRLKQLEDQRNGKLDL AWGEIWKSDVKELLISPEAKSSKNTNDSKDKDW EKKPLLSPGSSNESPKLDRKTDKPLLFHPTLG GVADFELOQFNRLKQLEDQRVTRGHETLKKL KKKN
338	sp_048730_m/ei	0.923112	#N/A	#N/A	uncharacterized protein	MKEAGNLLFKQGDIEDALEKYGAGIILGCFQFEE NDDRAEFFDLANCIILLNSAACFSRKRFEFQVGI CSVILEFNPNNKASYRRAMAALIELGRSLAYPD LLMASHNDPRNTEVRKKLEEVKOSREKEVSGKH SLGDVPIGLGLGSPPRKSKGELVKEHTIDQVO LHEKHEIMSEKIGCDVRLGLNSTVSEPVLSIMKDK HLSESSVWEENLMSAVLEGAEEEMMEDVLSVSN SENSKVPENYRVNRLRSGSSLSISKDYQLLL KGGTIQYFNSKVASPMTIRVIGGGDNTMKNRYE CHEKRLQDSFRVDMIPQHSSVSGLNEELNQRE ERLRLDGHYLNLFSSNKESESEMSVTTENYHHIG DASEYRDDACFKSTASIQQDEKDFPGVKVRRKD GRGIGRKATRKQVDKKENAIVTHSYVPSFSIKVEK KRKVVTPSDCIQRRPKKYLVALTNSIESYASSF ESCASSVESSICSSDKSCNNKVICVSEHMANSQE
339	sp_048870_qkod	0.923112	#N/A	#N/A	Tetrairicopeptide repeat (TPR)-like superfamily protein; FUNCTIONS IN: peptidyl-prolyl cis-trans isomerase activity, binding; INVOLVED IN: protein folding; CONTAINS InterPro DOMAIN/s: Tetrairicopeptide TPR-1 (InterPro:IPR001440), Tetrairicopeptide-like helical (InterPro:IPR011990), Tetrairicopeptide repeat-containing (InterPro:IPR013026), Tetrairicopeptide repeat (InterPro:IPR019734), Peptidyl-prolyl cis-trans isomerase, FKBP-type (InterPro:IPR011791); BEST Arabidopsis thaliana protein match is: rotamase FKBP 1 (TAIR:AT3G25230.1); Has 3730 Blast hits to 3559 proteins in 279 species: Archae - 0; Bacteria - 41; Metazoa - 2106; Fungi - 376; Plants - 567; Viruses - 0; Other Eukaryotes - 640 (source: NCBI BLink).	

10

20

30

40

【 1 1 3 6】

【表 1 4 - 1 3 7】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
340	sp_048940_weag	#N/A	#N/A	Heavy metal transport/detoxification superfamily protein; FUNCTIONS IN: metal ion transport; LOCATED IN: INVOLVED IN: metal ion transport; LOCATED IN: cellular component unknown; CONTAINS InterPro DOMAIN/s: Heavy metal transport/detoxification protein (InterPro:IPR006121); BEST Arabidopsis thaliana protein match is: Heavy metal transport/detoxification superfamily protein (TAIR:AT4G10465.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLink).	MPKSSRPLSLQTVELKVRMCCTGCEERVVKDAILK LRGVDSEIVLSMEKVTVVGVYDRNKVVKAVRRA GKRAEFWPYPNPPLYFTSANNYFKDMAHDYKET YNYRHGYNVGDKKHGSPLISSRGDDDKVSNMIFND DNVHACTLIM
341	sp_050470_ptpo	#N/A	#N/A	ent-kaurenoic acid hydroxylase 2 (KAO2); CONTAINS InterPro DOMAIN/s: Cytochrome P450, B-class (InterPro:IPR002397), Cytochrome P450 (InterPro:IPR001128), Cytochrome P450, conserved site (InterPro:IPR017972); BEST Arabidopsis thaliana protein match is: cytochrome P450, family 88, subfamily A, polypeptide 3 (TAIR:AT1G05160.1).	IMFGKPCVLTTPESCCKVLSDDATFGPGPEPESITK LMGKKAFFHGITQEEHKRLRLLTTAALSGQEALSI YLKHIEITINSLKEWAGMSRPIELTEMRKIAFKV IMFVMVGDSTDFHIMEALEKEYTTLNHGLKSMAI DLPGFAYHNAKARKNLQKIFQGVVDRKRIEETM ANKQEEKKDMLDMLMGIEDENKSKLSDEEIVDLII MFLNAGHESGHAAWAVVLLQDHPPEVFQKAK VINTVFGNQVDVEM
342	sp_053050_hrne	#N/A	#N/A	nuclear factor Y, subunit C3 (NF-YC3); FUNCTIONS IN: DNA binding, sequence-specific DNA binding transcription factor activity; INVOLVED IN: regulation of transcription, DNA-dependent; LOCATED IN: nucleus, intracellular; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Transcription factor CBF/NF-Y/archaeal histone (InterPro:IPR003958), Histone-fold (InterPro:IPR009072); BEST Arabidopsis thaliana protein match is: nuclear factor Y, subunit C9 (TAIR:AT1G08970.2); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLink).	MDSHWSLNHSDNKFVDTTTRYINQQQQQQQ CQQQKQKPHIDSASGGAGGSSPLQQQLRGFW QSQIQESKQDVFDRNHNPLARIKKIMKSDDNVR MISADAPVLLSRACQMFQDLSLRAWNSTQEGK RKTLCQKTDISAALSTTDLDFLVDILPREKPLISAT MLPPVNSASSSLGVICDHQ.LSQHQGYSVGEGTG LTMGLQMHQLVQDQDDQLDGGCNCSSNYIHRG WPNVLSHHSAWPPP
343	sp_053090_gkfc	#N/A	#N/A	uncharacterized protein	MDGAKNAEIVKTNVTEQTKIDTKIKRTAGEIM AETTKTHVGDALEKAGDGVVKEPPEKESLMDKA KGLFK

10

20

30

40

A	B	C	D	E	F
1	分析に用いたバイトとの相関係数	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	共発現遺伝子	sp_107620_kqnh (SOAP1)		<p>embryo defective 1075 (emb1075); FUNCTIONS IN: pyridoxal phosphate binding, carboxy-lyase activity, catalytic activity; INVOLVED IN: cellular amino acid metabolic process, embryo development ending in seed dormancy; LOCATED IN: cellular_component unknown; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINS: Pyridoxal phosphate-dependent transferase, major domain (InterPro:IPR015424), Pyridoxal phosphate-dependent decarboxylase (InterPro:IPR002129), Pyridoxal-phosphate binding site (InterPro:IPR021115), Pyridoxal phosphate-dependent transferase, major region, subdomain 1 (InterPro:IPR015421); BEST Arabidopsis thaliana protein match is: glutamate decarboxylase 2 (TAIR:AT1G65960.1); Has 3259 Blast hits to 3251 proteins in 995 species: Archae - 187; Bacteria - 1878; Metazoa - 518; Fungi - 211; Plants - 249; Viruses - 11; Other Eukaryotes - 205 (source: NCBI BLink).</p>	<p>MAGCHALPQEIHQDFLAVAHITSNNGDTSGEARKH LIDNNGEREIVRQRFRVHTTDLKIQEOPGAANGNI AKKSCDVNGKREIVTRTDLEKEQPGAANGNI GEKEAYMEDLVKHETWMDTKYHLYGYPGNLD FNFDLGRLLQQLNNCGDPFIEGNYGMHTKD FEVAVLDWFASLWEIEKQYWGYYVTNNGGTEGN FHGLLIGREMFDPDGLYTSKESHYSVFKAAHMYR MDCIKVDLTVSGEIDYDGLKTRLENKDKPAIILT IGTTYKGGVDNIDLVIKTLKEGGIPRNFYIHODGA LFLGLMLPFINRAPQVTFKPKIDSVTISGHKFGCPI TCGQIVRKHQHNVLAKNVEYASRDITMGSFNGQ APIFLWYALNRKGYNGFRKEVQKCMQNRARYLRD RLIEVGSVMLNELSNVTVFERPKNEEFQIHWQL SCQGNIAHVVMNPNTVDKLDMFVDEMVEKRLS QPFCVAFEIGKENCVCSLH</p>
344	sp_054370_dsqe	#N/A	#N/A		<p>MRSFQPYQHNNHLLYYRPPSTQPLLPPTPSHY HNPPFRSHQPPTLPPKPHNYHRQLPTLPNYH NVLISTSPVDHSTLSPAEELRVYSSPFFYSPVS PNYCGAYYPHCITTTCSPPDDHNAAYSSWSPKY SLPSPVFPSPSPSPSPSPSPVRYEESSSKSPVE SPSPSPVYSPYSPYSPYSPYSPYSPYSPYSPYSPY</p>
345	sp_055670_rnri	#N/A	#N/A	uncharacterized protein	<p>MYTFKKQATTAFSIGVGDNTSTRVGEVLKPIQ GOKNNSKGLTVDDSDVNDGAFDFP_LFTYPSAT RNNIEGWELSPVPVHEENMDDELSEDEDHIQ NGNEFDNHQNEACEDILENVHVDREISQTIQD VNVVEGSSDFDEKNNMOKKRGKRNCKEAKLKP DEKLEITFYNNRAIGTNHKVFAFHGLGIIVRDTNCP VVRVKWDDIGDKKEKHEHMMWSAVTVRI</p>
346	sp_058510_kipf	#N/A	#N/A	uncharacterized protein	

10

20

30

40

【表 1 4 - 1 3 9】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
347	sp_058560_ccgr	0.923119	#N/A	#N/A	uncharacterized protein	KEGWCSSVKANNSTNSGGWSTGKANVDGVS NSWRSQKEEKQPKSNDRSVGDGNFDDAKEE GLSSWDAAKVKERTQSSWGQPSKNSAQSSA DHWGFDKSNQPKSSGWGSEDTNAGKDSEKQ DSWKGKSNPFGGWNKKQPEGGRGWSNNTG EWSRKTQNCNQNQCNCNQNQRPRGPNHDS PRVALTATRKRMDPEFTISTPTDRTPITSTPTATP TDRTPTVSTPTISTPNISTPTDRTPITSTPTDRTPS SAWGKPAQETDNIIGWKKNNPQRDSENILEGXFY FNYPCLNPNYFNPNYCLNPNYINPTDRTPITSTPTD RTPTISTPTNRTSSAWGKPAQETDNIIGWKKNN QRDSENLEGTSGWNDKLOKENRFSKOSQPA SSKDWDSSTGNVTTGNTGFGVEKNEKFPWDVAS NVSVKSTWGQTGGNSWKKNEQDEKDDGDPQG LPWGKSHKSSDSWTSGGQNGHPVSGVXGVS EKQETLFLGTT
348	sp_062440_ojur	0.932333	#N/A	#N/A	glutathione S-transferase TAU 8 (GSTU8); FUNCTIONS IN: glutathione transferase activity; INVOLVED IN: response to calcium ion, toxin catabolic process; LOCATED IN: cytoplasm; EXPRESSED IN: 7 plant structures; EXPRESSED DURING: LP_06 six leaves visible, LP_04 four leaves visible, 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Thioredoxin fold (InterPro:IPR012335), Glutathione S-transferase, C-terminal (InterPro:IPR004046), Glutathione S- transferase, C-terminal-like (InterPro:IPR010987), Glutathione S-transferase/chloride channel, C- terminal (InterPro:IPR017933), Glutathione S- transferase, N-terminal (InterPro:IPR004045), Thioredoxin-like fold (InterPro:IPR012336); BEST Arabidopsis thaliana protein match is: glutathione S- transferase tau 7 (TAIR:AT2G29420.1); Has 7045 Blast hits to 7029 proteins in 1100 species: Archaea - 0; Bacteria - 3564; Metazoa - 831; Fungi - 183; Plants - 2012; Viruses - 0; Other Eukaryotes - 455 (source: NCBI BLINK).	MALKLGIPYDYQEEEPYGNKSPPELLKSNPVYKKI PVLJHMGKPIAESQVILEYIDQVWDGYSILPTDPY QRSLARFWAKFIDDKAMPSLWNAMWAKEDDEQE KTLKEAKENLEMLEKEIEGKFKFFGGENIGFVDIVA NFIFGFWLGTIQQVLGKELTKETYPHLSNWAEEY VNCDIIKQNVPPKERLFTLFLKHEAGF

10

20

30

40

【表 1 4 - 1 4 0】

A	B	C	D	E	F
1					
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh sp_107660_klqg (SOAP1) (SOAP2)-like	sp_107670_ptqx (SOAP2)-like	注釈	aa seq
349	sp_063480_wpxr	#N/A	#N/A	Protein of unknown function (DUF1442); CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF1442 (InterPro:IPR009902); BEST Arabidopsis thaliana protein match is: Protein of unknown function (DUF1442) (TAIR:AT2G45360.1); Has 92 Blast hits to 92 proteins in 12 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 92; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).	MAHWSAENATKAYLRTIKMGRKGRKEPDVAEFIS ALAGGNNAKLMEICGSIAGPTLLGLVAAAKQTS GRYVCILPGKELQASKQSLRSYLNHVEFVIGDP LTLTNRYSADFVTVDCNINGHEAIFKTKGQKG ALVWYNAQHLYPSYNELGGHFLPIGEGLLSVK PKHHDIRKSKWVVKVDKCTGEEHVFRITCPQOK VIQA
350	sp_063840_txwp	#N/A	#N/A	uncharacterized protein	MRQLMEGETRKRKITNICGEDEEEDEDIKMEKFF ALIRSANDIRDQLLDAEGSSKERAKKAVATAEKK AEVAVAGRAAQRRNEKGFASFATATAAAVDADA VVPNKAAAAAATAAAAAAADSAAARISRSAAAA AAEENSKKAADQKNFKRGRIFLDLNSL MAGSSEMKGEEKLSFEQAISLDEIIMNARKSFC AEDMVNHSVAENKAEVAGHGLVIGNLSNLTG GPLTVLDSKLVAGYIKDYPNPLLVGTSILQVQV GVSNDGLKAAVLYTGKNIACVDCGWLAWADSQ GKGRVYAECCGPISKFNCLNWERIEKKLDDPCC PAINPYKADEQGTGTSVIVGLSGVEWPLTYETRD LLAGIISG
351	sp_064960_mipu	#N/A	#N/A	uncharacterized protein	MGTEESADCLLGESISNGVVYDSSYSARKEKLI IDTDPGIDDCMAILMAFQTOEVDIIGLTTIFGNVQT ENATRNALLCEIAERPDVPAEAGSPGLKGGMP RVADFHGTDLGNINLLPKARIEKSAAOFLV DKVSEHPGEVSIKALGPLNLALAIKDDSSAFASKA KRIVVLGGAFFALGNVNPAAEANIFGDPEAADVV FTSGANITVGINITQVKLTDSDLHGLRESKGRH TRLLTDMCKFYRDWHVKSQGVYGFIFLHDPVSEV AVVRPDLFTYKKGVRVETQGGICAGYTLMDQGL KKWNSNPWTGYSPISSVAVTNNVEEVLMYKS QLTQS
352	sp_065850_iki	#N/A	#N/A	uridine-ribosyltransferase 1 (URH1); CONTAINS InterPro DOMAIN/s: Inosine/uridine-preferring nucleoside hydrolase (InterPro:IPR001910); BEST Arabidopsis thaliana protein match is: uridine-ribosyltransferase 2 (TAIR:AT1G05620.1); Has 6361 Blast hits to 6297 proteins in 1364 species: Archae - 75; Bacteria - 4675; Metazoa - 179; Fungi - 233; Plants - 187; Viruses - 0; Other Eukaryotes - 1012 (source: NCBI BLINK).	MRREKARTGEIGKEKPEAGYGAEGPIPIQIEG PILEGAIPIPIPIPIPIPIPIPIPIPIPIPIPIPIPI VFVFTKT
353	sp_070260_rqas	#N/A	#N/A	uncharacterized protein	

10

20

30

40

【 1 1 4 0】

【表 1 4 - 1 4 1】

1	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
354	sp_070580_zelq	0.971336	#N/A	#N/A	RNase H family protein; FUNCTIONS IN: ribonuclease H activity, nucleic acid binding; INVOLVED IN: biological_process unknown; EXPRESSED IN: 11 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: Ribosomal protein L9/RNase H1, N-terminal (InterPro:IPR009027), Polynucleotidyl transferase, ribonuclease H fold (InterPro:IPR012337), Ribonuclease H (InterPro:IPR002156); BEST Arabidopsis thaliana protein match is: RNase H family protein (TAIR:AT5G51080.2); Has 3412 Blast hits to 3412 proteins in 738 species: Archae - 71; Bacteria - 1533; Metazoa - 7; Fungi - 2; Plants - 1270; Viruses - 0; Other Eukaryotes - 529 (source: NCBI BLINK).	MQQVFKLCROSSIPSPFQPAKLAFTIFHAAKALAVATSEAAANNQDSTSSANDAVSWISLDVSKKH SRSLFNAAYVPLJLSENDAAVAEATGSGKPKHHARI NSGFGAHVHLQKWSINQLVKSISLSSSTSVFL CLGVQVWYSSKSNSSGKQKLPKAVMEEEEKGA YFVWRKGDVGVYKSLSECOAGVSSVSDLSATV YKGNLSPKDAEEYLASRGLHNLALYTIKASDLKDC PFGKIVPCPEKTCILHFDGSSKGNPIRAGAGAVL RTLGNLICRIRQQLGNATSSVAEYRAFILGLKKA LEMGCTRIQAQQGDNKIVCMQGSYVLLNPSYGND RLSNIYPI
355	sp_070600_cqhx	0.933333	#N/A	#N/A	Xanthine/uracil permease family protein; FUNCTIONS IN: transmembrane transporter activity; INVOLVED IN: transport, transmembrane transport; LOCATED IN: membrane; CONTAINS InterPro DOMAIN/s: Xanthine/uracil/vitamin C permease (InterPro:IPR006043); BEST Arabidopsis thaliana protein match is: Xanthine/uracil/vitamin C permease (TAIR:AT5G25420.1); Has 8520 Blast hits to 8482 proteins in 1844 species: Archae - 65; Bacteria - 6911; Metazoa - 346; Fungi - 119; Plants - 452; Viruses - 1; Other Eukaryotes - 626 (source: NCBI BLINK). terpene synthase 14 (TPST4); CONTAINS InterPro DOMAIN/s: Terpene synthase, metal-binding domain (InterPro:IPR005630), Terpenoid synthase (InterPro:IPR008949), Terpenoid cyclases/protein prenyltransferase alpha-alpha toroid (InterPro:IPR008930), Terpene synthase-like (InterPro:IPR001906); BEST Arabidopsis thaliana protein match is: terpene synthase 02 (TAIR:AT4G16730.1); Has 1701 Blast hits to 1676 proteins in 178 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 1696; Viruses - 0; Other Eukaryotes - 5 (source: NCBI BLINK).	MMSGNVGCGNAGGGNNTVAQEVKPHVME TLSGIQYCNTPPLFEIILGFQHSFLTGIITVISS LTPQMGGSNHDKAKVQSMFMFVSGTSTLLOTL VGTRLPVAIGASHAFVPMTTIVLTRLSCINQPOF LDQTEKTMEMQGALLVSSVQIILGFSGIWRIF VRYLSPLSMVLITFTGLGLYSLGFPMADLDCVEIG LPHILLIFSOYMSGFLKTSRLIFDRYAIFFTAIWI YALILTASGVHNHRPPTQLSCRTDRAGLMEAD WYVYLRPFQWGWKPTVNMREATPVFAGLVALIES TGTFFIAARYGSATPVPPSVISRGTFWLGIGTLIN AVFGAVAGPTASVESAGLLGLTRIGSRVRIQISAG FVILFFSIFGKFTAFFASVFPFSMAATYCILOYVSS AGLNHLOFCNLNSFRTLFIILGFSSFMASVFPQQFR DYLLTNSGGQLPTRSKWFFDTLATIFVSSPAVAIM VAILDNTISEEALFNDSGKRWVERFIQFTKDIR NAEFYSLPCELGNCFPSL
356	sp_070640_izrr	0.933333	#N/A	#N/A		MVSIVFATTNTPITPTTKITYYKESFINRHHSS INTKEYRLHEGVYRKRIGIMTSKTRGYDIFNLLKDK DGLQLKOEIYKEDVKGLL SVCECSILGETNDISEQ LLRTLNNMEEOSKASITDNTLRYPPYHKNLSILM VQNVLDEFDHYVSKFLHKSNNVKEWLNVQVQLA RVDFGNARMTRYEVQVQISIVKKNHNFQKELSF ARIDPFKWWMMW/SYGILOGPNM/SEDRIELAKAISF IYMIDDFDNGTYSDELLELFTFTEANGWEYTDSDIN QLPNCMKYCFKALLETTFNDFSSKILTKHGWNPE QCLRKLQVQLA

【 1 1 4 1 1】

10

20

30

40

【表 1 4 - 1 4 2】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP)	分析に用いたバイトとの相関係数	sp_107670_ptqx (SOAP2-like)	sp_107660_kicg (SOAP2)	aa seq
				<p>cysteine-rich RLK (RECEPTOR-like protein kinase) 8 (CRK8); FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Reverse transcriptase, RNA-dependent DNA polymerase (InterPro:IPR013103), Serine/threonine-protein kinase domain (InterPro:IPR002290), Protein of unknown function DUF26 (InterPro:IPR002902), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Serine/threonine-protein kinase, active site (InterPro:IPR008271), Protein kinase-like domain (InterPro:IPR011009), Protein kinase, catalytic domain (InterPro:IPR000719), Tyrosine-protein kinase, catalytic domain (InterPro:IPR020635); BEST Arabidopsis thaliana protein match is: cysteine-rich RLK (RECEPTOR-like protein kinase) 6 (TAIR:AT4G23140.1); Has 131284 Blast hits to 128961 proteins in 4748 species: Archae - 114; Bacteria - 13787; Metazoa - 45525; Fungi - 11866; Plants - 40839; Viruses - 427; Other Eukaryotes - 18726 (source: NCBI BLINK).</p>	<p>MTSIRTYLGLAASLDLKLQDLVKVTAFLHGDLEHEE IYMDQPGGFDEKGDKLVFKLKLKSLYGLKQAPR QWYRKFDSFMTSNGYKCTTADPCVYFRKFPFG NWYRLLVYDQDMVLIYGDANVQIQLKAFKSLFDMK DLGFSKILGMEIARDRKGKVLWSQEKYIEWVL ERFNMKHKAPVSTPLAT-HFKLSKRAEGRNALGP SMHLMV</p>
357	sp_071950_ikaf	#N/A	#N/A	<p>ATBFRUCT1; FUNCTIONS IN: hydrolase activity, hydrolyzing O-glycosyl compounds, beta-fructofuranosidase activity; INVOLVED IN: response to karrikin, response to wounding; LOCATED IN: cell wall; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Glycoside hydrolase, family 32 (InterPro:IPR001362), Glycoside hydrolase, family 32, active site (InterPro:IPR018053), Glycosyl hydrolases family 32, N-terminal (InterPro:IPR013148), Glycosyl hydrolase family 32, C-terminal (InterPro:IPR013189), Concavalin A-like lectin/glucanase (InterPro:IPR008985); BEST Arabidopsis thaliana protein match is: beta-fructofuranosidase 5 (TAIR:AT1G55120.1); Has 4230 Blast hits to 4174 proteins in 1237 species: Archae - 18; Bacteria - 2583; Metazoa - 93; Fungi - 290; Plants - 1044; Viruses - 0; Other Eukaryotes - 202 (source: NCBI BLINK).</p>	<p>MALCYSTKLVILLGGSLLLGYVIELQASHHVQN ATLDQSLISADQPYRTAYHFQPVQINWMDPCG PVYKGYHLFYQYMPIGNSPDGPRVWGHSTSF DMVNSFQPLILEPQMSVNMNGSYSGSTTIQGY IPAILFTGITPQNEQVQDLAFPKDISDPYLKEWVL VPENPVMSPTLENNINASSFRDPTTAWRLSDGN WRVLIGSKIDRLGLALLFRSMDFINWVQADEPLY SYDOTGMWVCEPDPFFVYINGPFIQVDTSAIGPTV KHVLKVSLLDDTKFDIYAIGTYDITTKDAYIPDAGFL NDSSLRYDYGYKYASKFYDNLKMKORILLSWINE SSTVEDDIKKGWSGIHAIPRVRWLDKSRKLIQW PIAEIERLVNVPVILPTQVIKGGSTIEIPRITAAQAD VEISFRVDMPIERFDASWTNPQLLCSQKGPV SGGGLGPFGLLTLASTGLEEYTAFFRIFKAYDTNK FIVLMCSQSRSSLNPTDKTTYGTGFVNDPIRE GLSLRVLIHSHVNESFANKGNKVNITARVYPTLAIYD KTRLVYAFNYGVNNDITWLTAWSMRKANIA</p>
358	sp_072040_guid	#N/A	#N/A		

10

20

30

40

【 1 1 4 2】

【表 1 4 - 1 4 3】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 UDP-Glycosyltransferase superfamily protein; FUNCTIONS IN: transferase activity, transferring glycosyl groups; INVOLVED IN: response to salt stress, N-terminal protein myristoylation; LOCATED IN: cellular component unknown; EXPRESSED IN: guard cell; CONTAINS InterPro DOMAIN/s: UDP-glycuronosyl/UDP-glycosyltransferase (InterPro:IPR002213); BEST Arabidopsis thaliana protein match is: UDP-Glycosyltransferase superfamily protein (TAIR:AT4G27570.1); Has 5043 Blast hits to 5008 proteins in 241 species: Archae - 0; Bacteria - 56; Metazoa - 380; Fungi - 15; Plants - 4575; Viruses - 3; Other Eukaryotes - 14 (source: NCBI BLink).	aa seq MDRTRPEVESILSTQKPDLVFYDMAH-WIPDIASK LGIKSVSYNSVCAISMHVPSVLPKAAATSQVA LSSSYVGSVPQWFRNDHNSVYFGDEITLLER SIISLSSADAIAIRTCREIEGVYCDRVAATAENKPV VTSHALPDLLEQLSPLETRWADW-LARFGPRSVIF CCLGSQHYLVDAALQLELALGLEMTGLPELMAVK PPVGYASLKELLPEGYNDRVSGRGVWHGGWVQ QQQILAHPSLGCVFTHC.GSSSSMWEGLVSETQLV LLPQLPDDTLNAKLMADLKVGVKVERQONGWV SKQGLCEAINFRVMEDESDISHVVRKNHAKYTSML ISPDFITGYIDNFIKDLQALIS
359	sp_072870_funtr	#N/A	#N/A		KKRRNQQLKPPSSARGPRFSATAPLVDSPLLA DSSFLTDSFLADSPLOEQEESGTLDSIGASLG RLRKKDPSLTAEMIAALWCLHFLQTLTFLMGHGYI #N/A
360	sp_073010_xeas	#N/A	#N/A	uncharacterized protein	MLFKEERKMGNTSMVYVNFVYHGRKFIHFKPR GYEGGLSETIISRNRSLTHMHKLEIKLGCKKEY TIWYVEDMKRFHEDGIAIKDPDDDFYNLYELTKSNL VHIVVQHRDVEEEGKGLQNERWRRLTQQRLPK NLG
361	sp_073891z2584	#N/A	#N/A	uncharacterized protein	MLVLLVFTDQVQSLNFSFPPFNPNTPNITYRND STPTNGAIQLTINRAVAGLSGVSGRASYNHPVRL WDQSSNRRTNFTTHFSFDIQVIPDPNYYGDGIAF FIAPFDDSANAPADSFSGGCLGIGRGODESGNST TEYPVAVAFEDTHNAWDFGESGAPSNDRHVGIN VNSVFSRTNFRLESLNKRNRTTGNAWITYDSSSTNN LSVFLSYDANFPVNGESNLSYVIDLSSILPFRVRY GFSSTGANYELHNLISWNFNSTLEDRIILPAPR QSNKAALVGGIVAGVSALLILLGLVAFMWWRR KRRDDMKFDVDDDFDEETGPRFRFYKELSOATN NFNEEGKLGEGGEGVYRGFMVMSREIAVKKI SRGSKQKKEYVSEVKAISRHLRHNKLVQLLQVWC HROGHLVLYEYMPNGSLDLSHLYTGKRVLPWP DERYRQAQGLASAVLYLHEEWEGCVLHRDIKSSN VMLDSNFNAKLGDFGLARLVHNLGSGQTTEAS KYVMIEWVWRLYEGEGLLEAVDKTQVGEFIOQI ECLMVGLWCC.HPEYTFRPSIRGVINLVNLESPL PNLPMKFPVYVYCSPTTNVKNFVSSSSSSSSSS SYSSSGFTGSSTGTSSSSQSVVSGTAPLLSQN LTV
362	sp_074000_slue	#N/A	#N/A	Concanavalin A-like lectin protein kinase family protein; FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: endomembrane system; EXPRESSED IN: root; CONTAINS InterPro DOMAIN/s: Legume lectin, beta chain (InterPro:IPR001220), Protein kinase, ATP binding site (InterPro:IPR017441), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Concanavalin A-like lectin/glucanase, subgroup (InterPro:IPR013320), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271), Protein kinase, catalytic domain (InterPro:IPR000719), Concanavalin A-like lectin/glucanase (InterPro:IPR008988); BEST Arabidopsis thaliana protein match is: Concanavalin A-like lectin protein kinase family protein (TAIR:AT5G65600.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).	
363	sp_074680_doa1	#N/A	#N/A		

10

20

30

40

【 1 1 4 3】

【表 1 4 - 1 4 4】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
364 sp_076670_xjri	Q_5226473	#N/A	#N/A	<p>cellulose synthase like G2 (CSLG2); FUNCTIONS IN: cellulose synthase activity, transferase activity, transferring glycosyl groups, transferase activity; INVOLVED IN: cellulose biosynthetic process, polysaccharide biosynthetic process; LOCATED IN: membrane; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAINs: Cellulose synthase (InterPro:IPR005150); BEST Arabidopsis thaliana protein match is: cellulose synthase like G1 (TAIR:AT4G24010.1); Has 2950 Blast hits to 2333 proteins in 442 species: Archae - 3; Bacteria - 792; Metazoa - 2; Fungi - 20; Plants - 2101; Viruses - 0; Other Eukaryotes - 32 (source: NCBI BLINK).</p>	<p>MENPFHEFHVKLRATLNITYSLFYFLALTFIFYR SAAFFLTKATGVLFAELVFSFLWAIKQPMYWRP ITRTVFPEKLPKDEKLPGLIDVLICTADPNKEPTFEV INTVISAMSLDYP AHKLSVY LDDAGASVTLNAM KQAWVFASWVLPFCRKYHLKTIQOAYFQNP DQEESCCQTNMDFVQHSMLLEEKYEAFKQV RSNKLSDVDNINNTSRDHPPLVQVINENCIEDG AVTLNPEAGEMPLLVYVSRKRPSPHCFKAGA LNALLRVSSIMNSPYTLILDQDMYCNDISARQA MCFRIDPKISSLGWVQFPQKFNHISDITDYDSQL RVTWVYWPAGDGLQGPCIGTNYIKRVLVLSY NITDDGINLKEVRSSLGFSNELTNSVVQNDKLN KTVKDKVEVSMVQEAHFLASCTYEKGTMWGQ KYGFCYGTVEDAKTSIRLQNKGWRSYLNPTRP QFLGSATTSLNEMLVQGIWYGGLLDCLSKYC PLYRPSRMQFLQKMYYSWVLLALDFIPVLCFAI IPICFLYGIPIVYKLPDGPFAVAFVSTQLKHL SVDVVYSGSVSAWVNEQRTWMAKCLSCYL YG TIVCITLGLHEATFTLTDKLGDDTTNWDYDKGIYD FRTSKMFLPVTAVILNCCFIFGGVARLIVIRNWD AMFGQMLFSLVLMGFPVLEGLVLRKDNARISV SASMWSTLISLILLGLYLIRH</p>
365 sp_078000_hnxy	Q_5257123	#N/A	#N/A	<p>Protein kinase superfamily protein; FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: chloroplast; EXPRESSED IN: shoot apex, embryo, root, flower, seed; EXPRESSED DURING: F mature embryo stage, petal differentiation and expansion stage, E expanded cotyledon stage, D bilateral stage; CONTAINS InterPro DOMAINs: Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase domain (InterPro:IPR02290), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: D6 protein kinase like 2 (TAIR:AT5G47750.1); Has 111692 Blast hits to 81240 proteins in 2732 species: Archae - 30; Bacteria - 14475; Metazoa - 44429; Fungi - 12685; Plants - 18495; Viruses - 325; Other Eukaryotes - 21253 (source: NCBI BLINK).</p>	<p>MNKPRKORSLETESMNSLFQDLTVKDSVSVSL SVLSPCSSTTASGSDVGGSSKPPNIGCLLKLVE NGESEKASASSLSNSSNGVDNTNDASSFRSSC PNKPHKGNDIRWEAIOCVKGGKDGELGLSHFRL LKKLGFDIGVYLAELRGMGCLFAMKVMKMG LAGRKKLORAQTEKDLGLLDHPFLPTLYSHFETD KFSCLLMEFCNGGDLHILROROPNKHFSEQASR FYASEVLLALEYLHVMGWVYRDLKPNVLRVED GHIMLSDFDLSLRCCVNPTLYKGSYOPNCTNSY CIEPRCIEFACKLPVCVEPSCLOPSCFRPLFNK SSKVKREKFNFTNSSLILIAEPTNARMSRMSGT HEYLAPEIRGDGHGSADVWTFGIFLYELLHGK TPFKGNQNETLFWVGGSLKFFPEGSPASFAAK DLIRGLLYKDPQKRLGFKRGATEIKQHPFFESVN WALIRGSTPEIPKPTDLTYNHTSKTSSEPPSEK QGASDSDRSRSGPYLDFFEF</p>

10

20

30

40

【 1 1 4 4】

【表 1 4 - 1 4 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kicq (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
366	sp_078820_rkx 0.923119	#N/A	#N/A	Aminotransferase-like, plant mobile domain family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN's: Aminotransferase-like, plant mobile domain (InterPro:IPR019557); BEST Arabidopsis thaliana protein match is: Aminotransferase-like, plant mobile domain family protein (TAIR:AT2G25010.1); Has 691 Blast hits to 683 proteins in 18 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 691; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink).	TFLSGPTDNFDPWIGMVEDVSTLGDYCWGDL GYATLVGQMSLAVRSDPWRHRHLITLAGVPRLI ELWAFEHLPLWLPAPKGGORPLEYPAGRRWGWK KLTARPPDPTVWDLIRDGNPEHVWVTPWLSFR GTYASVRDSYALGQMRVLFVGRDDPVWYLGER
367	sp_079440_pti 0.923119	#N/A	#N/A	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G05858.1); Has 48 Blast hits to 48 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 48; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink).	MGSSSKKRASTRTNHLDFTPRRKVFPSEDE RGYVVAEPGIDRKASAFIAKFHEARRFEAQTMAI PNSQTS
368	sp_080820_syp 0.923119	#N/A	#N/A	uncharacterized protein	#N/A
369	sp_080861z2824 0.923119	#N/A	#N/A	uncharacterized protein	#N/A
370	sp_081120_kzde 0.923119	#N/A	#N/A	uncharacterized protein	#N/A
371	sp_081530_scku 0.923127	#N/A	#N/A	fatty acid reductase 1 (FAR1); FUNCTIONS IN: oxidoreductase activity, acting on the C-H group of donors, fatty-acyl-CoA reductase (alcohol-forming) activity, long-chain-fatty-acyl-CoA reductase activity; INVOLVED IN: response to salt stress, microsporogenesis, response to wounding, suberin biosynthetic process; LOCATED IN: chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN's: Male sterility (InterPro:IPR004262), NAD(P)-binding domain (InterPro:IPR016040), Male sterility, NAD-binding (InterPro:IPR013120); BEST Arabidopsis thaliana protein match is: fatty acid reductase 4 (TAIR:AT3G44540.1); Has 2592 Blast hits to 2555 proteins in 497 species: Archae - 6; Bacteria - 759; Metazoa - 1002; Fungi - 307; Plants - 285; Viruses - 0; Other Eukaryotes - 233 (source: NCBI BLink).	MCKCRSHGTFFLGLLOSSHLNICYHQDLHQRP LRFGRSRHRFCSDRCCALLIGAWHLPRRPFIGHAL QRHPFSGLYDSAGELLHTP  MDQDSILKFLDNKAILVTGGAGFLAKIFVEKVLRS OPNVNKYLLLRAPDNNNSASSRLQNEVRYIFA

【 1 1 4 5】

10

20

30

40

【表 1 4 - 1 4 6】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
372	sp_082770_zfhh	0.923363	#N/A	#N/A	<p>N-MYC downregulated-like 2 (NDL2); CONTAINS InterPro DOMAINs: Pollen specific protein SF21 (InterPro:IPR015511), Ndr (InterPro:IPR004142); BEST Arabidopsis thaliana protein match is: N-MYC downregulated-like 1 (TAIR:AT5G56750.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archaea - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLink).</p> <p>Pyridoxal phosphate (PLP)-dependent transferases superfamily protein; FUNCTIONS IN: pyridoxal phosphate binding, carboxy-lyase activity, catalytic activity, tyrosine decarboxylase activity, INVOLVED IN: response to wounding, cellular amino acid metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Aromatic-L-amino-acid decarboxylase (InterPro:IPR010977), Pyridoxal phosphate-dependent transferase, major domain (InterPro:IPR015424), Pyridoxal phosphate-dependent decarboxylase (InterPro:IPR002129), Pyridoxal-phosphate binding site (InterPro:IPR021115), Pyridoxal phosphate-dependent transferase, major region, subdomain 1 (InterPro:IPR015421), Pyridoxal phosphate-dependent transferase, major region, subdomain 2 (InterPro:IPR015422); BEST Arabidopsis thaliana protein match is: L-tyrosine decarboxylase (TAIR:AT4G28680.1); Has 5452 Blast hits to 5431 proteins in 1792 species: Archaea - 99; Bacteria - 1888; Metazoa - 2300; Fungi - 284; Plants - 262; Viruses - 5; Other Eukaryotes - 634 (source: NCBI BLink).</p>	<p>MGSDSDSISVFIESISIDSKKEYVYKRTIGSGPYSVAV YGDQDKPALVYTPDLALNYASCFQGLFLCPEALS LLVHNFQVHISFPFGHEWGAAPISYEYEASTPSVD DLADQIVEILDYFGLGEVCMCMGVTAGAYLTLFA TKYKQVRLGLLISPLCRAPSWTEWLYNKVMSNL LYFYGMCVVKEMLLKRYFSKELRGNLSLVPESDI VOACRKLDDRESSTNMVHFEAINRRTDISEGLK KLCRSLICVGEDSSFSHSEAVHMATKLDRRFSAL VEVOACGCVLVEEQPQAMLIPLEYFLMGYGFYR QSYKSLSPRSPSPSCISPELLSPESMGLKLPKIP TRISLEY</p>
373	sp_083530_gyrt	0.9036339	#N/A	#N/A		<p>MGSLDIKSLESTTQMNMNPNLDPEEFRKQGY MVIDFLAEYKYSIEKYPVRSQVEAGYLRRKLPVSA PNDPVSMTILADIOTDIVPGLTHWQSPNYFAYE PSSGSTAGLLGETLASEGFNVVGNWISSPASTEL ESIVMDWLAEMLSLPKSFSTFGSGGGVMMGT CEALTTITARDSDLKIGRHNINKLVVYGSQDTH CSFFKSAKIAGISIPANFRITKTKANAFSLRGDSL RATIOADVNAGLIPFVAVTVGTTSTAADVAVGPI CDVAKEYNMVHVDAAYAGNGCIGPELRFHMIDG VEKADSFNAHKWFLLTDCCLFVKDPNLRV CLSTNPEYLKKNKATDNTQVVDYKDWQLT.SRRF RSLKMMWVIRSYGVDLFRNLRSHVDMAYFEK MVVKDTRFEIVVPRSFALVCFRLSASAVLECSGQ KFLFEETLNEINTELESVNSDGIYMTSHSLVDGIF MIRFAVGATLTIHHVTVYAWKVVQVEHATALLAR SMIAC</p>

10

20

30

40

【 1 1 4 6】

【表 1 4 - 1 4 7】

A	B	C	D	E	F	
1						
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	
				cytochrome P450, family 710, subfamily A, polypeptide 1 (CYP710A1); FUNCTIONS IN: C-22 sterol desaturase activity, oxygen binding; INVOLVED IN: oxidation reduction; LOCATED IN: endomembrane system; EXPRESSED IN: 14 plant structures; EXPRESSED DURING: LP.04 four leaves visible, 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Cytochrome P450 (InterPro:IPR01128), Cytochrome P450 conserved site (InterPro:IPR01792), Cytochrome P450, E-class, group 1 (InterPro:IPR002401); BEST Arabidopsis thaliana protein match is: cytochrome P450, family 710, subfamily A, polypeptide 2 (TAIR:AT2G34490.1); Has 25674 Blast hits to 25623 proteins in 1435 species: Archae - 46; Bacteria - 2777; Metazoa - 10394; Fungi - 4411; Plants - 7239; Viruses - 0; Other Eukaryotes - 807 (source: NCBI BLink).		MSYNSMDSWAILLAAPFVATLVALVILIEQIYLYKK KSVAPGPTLVLPFLGNVTLIRNPEKFRWDDQADY AKSSPLGISVNYIIGKFXAKSSPLGISVNYIIGKFMV FTRDSEMSHKIFANVPDADFQWVGHFPGKGLFG EHNMINLYGQEHKDLRRRIAPNFTPKALATYSDIQ QLVILKHLTRWVENINDPKVALRVLARDMNLTSQ TVFVGPYLDEEKREGKVDYNYFNVGVVALPIDIP GFAFRNARLAVERLVERLTRCVESKMKVENE EPTCLVDYWMQGMVNEERELQQPLQFFPSPPP PPPNCSNHQIGAHMDFLFAAGDASTALLWA VTLLDHPVLLKVVREEVESWIKVESNKLITPEQL VSMKYTQAVAREVVFYRAPATLVPHMVHQDFQL AENFTIPKGSIVFPVYESSFQGFSEPRDFPDR FYLEEREDQVYKRNLYLVFGGAHQCVGQRYAL NLLVLFAMFTSLIDFKRHRTEGCDLNFCTICPK DDCLVLSLKRCAANFPALSLP
374	sp_085530_zifo Q35373	#N/A	#N/A	uncharacterized protein	#N/A	
375	sp_086041z3059 Q323119	#N/A	#N/A	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; Has 20719 Blast hits to 6096 proteins in 607 species: Archae - 22; Bacteria - 3243; Metazoa - 4364; Fungi - 2270; Plants - 237; Viruses - 128; Other Eukaryotes - 10455 (source: NCBI BLink).	MAPLPLFFRFTFYLLFASSLFASSYSHRIFLSRR NLLDNTNSKKNQTKLIKPSSSSSSKNQTISSNSN SSKKNQTKLAKPIKTSKKNVGPAPNTVGPKNV STKSSDPLKSKSITKNKTKSSPSPSTPTTKLPSSDT KSDQNKQKQKQKLSKSPPTWMSQSDDDND TDYDFMAEFRDLSNKFHQTIFPDMHFRFSTTKMY ITKANKIEITQSFKPYVSKQYHSSASTAISFAIIVP FLVSLLFNRTAYFSLQKLLIFIQIYLSIYFSLCFSS LITGLEPLRFFYATSQSTVWVWLQVLTGLVWMYL LMLLMYLLVVFSTESGQKFLGQFLAQFTVGFVAVG FHYMYTVHRVVLHLPKTNWVHGIYATCFELVI CLLARADRTTKAYHDEEGVEEGKIN	
376	sp_086530_hkee Q323177	#N/A	#N/A	uncharacterized protein	#N/A	
377	sp_087591z3096 Q323119	#N/A	#N/A	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G56260.2); Has 477 Blast hits to 219 proteins in 41 species: Archae - 0; Bacteria - 4; Metazoa - 91; Fungi - 61; Plants - 144; Viruses - 0; Other Eukaryotes - 177 (source: NCBI BLink).	MYQIKEDDKLFSKLVTKESMANPFRVYGGN SVGVFPVWEAKPGTPKHTKHIDY'SKEEENKQY LPPLTPPCYYLKQPNKLPNTNGNSKRSSSRIFLG NLFPFKSMHRRQISTTSSSSSTLSCPSSMLV SDASTMRSSDESRLNLSLSLSLSRSGSGSGSG SGSGTRRRRTLSSPGLFDFLVNGDQEDVEEED MRSPRSVFCFGLGCSRRIRSC	
378	sp_087890_qxaw Q323175	#N/A	#N/A			

10

20

30

40

【表 1 4 - 1 4 8】

A	B	C	D	E	F
1					
2	分析に用いたペイトとの相関係数 sp_107620_kqmh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
379	sp_088320_effk	#N/A	#N/A	MUTS-homologue 5 (MSH5); FUNCTIONS IN: damaged DNA binding, mismatched DNA binding, ATP binding; INVOLVED IN: mismatch repair, chiasma assembly, reciprocal meiotic recombination; LOCATED IN: condensin nuclear chromosome; EXPRESSED IN: male gametophyte, flower, pollen tube; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage, L00 in/inflorescence just visible; CONTAINS InterPro DOMAIN/s: DNA mismatch repair protein MutS, core (InterPro:IPR007696), DNA mismatch repair protein MutS, C-terminal (InterPro:IPR000432); BEST Arabidopsis thaliana protein match is: MUTS homolog 2 (TAIR:AT3G18524.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MGRRLRNWFLRPILDAEILNSRLDAISFFLCSEE LMVSVRGTLKSIKIDVPYILKIKFNSSPSSVSTSDW AAFLKLSLLHLNKFVGVISENLRGOANHFLCDI VDKAASCITTELSEYVELVIGIIVDTREKERGYGTIV KEGFCCELDLQRIYEELPELLEEVSSLELARLPN LHHGECVPCIVYIHIQIGYLLWKFEFKLDECTLQNL EDLEFAFSDGENEPEKRFYRTSKTRELDNLLGDI YHKILDMERVITRDLISHLVLLFSTHLLKAMNFAAEL DCLLSALVAQONNYYRPNLTTDAIDIQNGRHVL QEMTVDTFIPNDTKILDEGKINIITGNYSKGSYIK QVALIVFLSHIGSFVPAADAAYVGLTDRIFCAGMSK LMAAEQSTFMIDLQQVGMMLRQATSRSLCLLDE FGKGLTDEGIGLGGTVVYFAGNDNTPKVFVCT HLTEIFHENLLPKSEKINFYTMVSLRPEEESWPAE DIVFLYRYPGQAHLSYGLHCAMLGAVADNVVKR AAEILDATGKGNHVDRLSNENIQADQQQYQKALE KLLAFDARNGNLDLLFEDMFCTS
380	sp_089871z3160	#N/A	#N/A	uncharacterized protein	#N/A
381	sp_090991z3216	#N/A	#N/A	uncharacterized protein	#N/A
382	sp_092191z3260	#N/A	#N/A	uncharacterized protein	#N/A
383	sp_092771z3313	#N/A	#N/A	uncharacterized protein	#N/A
384	sp_094840_irsr	#N/A	#N/A	Protein of unknown function, DUF584; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF584 (InterPro:IPR007608); BEST Arabidopsis thaliana protein match is: Protein of unknown function, DUF584 (TAIR:AT4G21970.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MENNRYNWROYRGEFEEDVWAVIKENGETT NNHLRNPRESISRRLPFAARMIPGVQKHQHIDDNS NNESELVOSSAPVNIPIVDSWYKSPVKMKMYGG CSSSFDYEDDNDYDLVLEDDNNEEYAGDGRLLPH EWLAKKLARSQISSFSVCEGVGRTLKGRDLSRV RNAILSKTGF
385	sp_096961z3475	#N/A	#N/A	uncharacterized protein	#N/A
386	sp_098111z3526	#N/A	#N/A	uncharacterized protein	#N/A

10

20

30

40

【 1 1 4 8】

【表 1 4 - 1 4 9】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
387	sp_099160_fatq	0.923119	#N/A	#N/A	Peroxidase superfamily protein; FUNCTIONS IN: protein binding, peroxidase activity; INVOLVED IN: response to oxidative stress, oxidation reduction; LOCATED IN: apoplast, cell wall; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: LP.04 four leaves visible, 4 anthesis, C globular stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Haem peroxidase (InterPro:IPR010255), Plant peroxidase (InterPro:IPR000823), Peroxidases heme-ligand binding site (InterPro:IPR019793), Haem peroxidase, plant/fungal/bacterial (InterPro:IPR02016), Peroxidase, active site (InterPro:IPR019794); BEST Arabidopsis thaliana protein match is: Peroxidase superfamily protein (TAIR:AT5G58400.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	GCDGSIILLNDTANFTGKTAIPNNVSRGYDVIDT IKSQVENICPGIVSCADIVAAARDSWALGGPS WYPLGRDRDSTTASITLANSIDIPAPFNITELLRS FSNKGFSQEMVALSGAHTIGQARCTTFPTRIN ETNINSTFATSLQANCPRNGDDNLSPLDRTSST TFDNAYFTNLLSLEGLVHTDQELFNNGTGTITDS FVSSYSSTSTFTDFSDAMVKMGVNPVGELEE PCLVMLPC
388	sp_099861z3606	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
389	sp_101740_azug	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
390	sp_103201z3740	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
391	sp_104821z3793	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
392	sp_105691z3827	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
393	sp_106341z3848	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
394	sp_106671z3860	0.923119	#N/A	#N/A	uncharacterized protein	#N/A

10

20

30

40

【表 1 4 - 1 5 0】

I	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh sp_107660_kiqg (SOAP1) (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	MASTYLLHPTTISIPFKQLNLQSLQPRFFLNTQTR TIVYDFRIRASDGVENADSYLNWKNVAVNREK KEIEFNIAQKLPIDGGVDEETPDDLEKKSFEFQ KILEYSKEERDKVQRMQVIDRAAAIAAARALLKG RKADDDVANSLSSESVRVSDNSGVLSPETENGG IDSGSSYKQSRISAVGPPGDFWSTPTPPEDD LNRSDVSKIQKAPKTNPHTNTVIEKERSADYLSI PLESTYSENLIQLPPFQSLMEVHKVEALNSSQ EATVMQEEIGSQFSAHAAEAHNLHNVEEKPFY GINPEGSRVWKTGVETRTDGVVCRWTLNRGV TADGVVEEKEYWEAADGLDYKELGSEKSGRD SMGNVWREYWKETMWQEHGLLHMEKTAADKW GKNAGKEEWQEKWLEKYDASGKADKWAHKWC SIDPNTQVEIGHAHHWHERWGETYDRQGGSIKY TDKWAERCEGDDWVKWGDKWDENFDQNGHG VKOGETWVEGKYGDQWNRWRTWGERHNGSGWL HKYKSSSGE-HWDTHAQEDTWYERFP:HFQFMH CFENSPQLREYKPKPSDI
395	sp_109480_ctfd	0.955038	#N/A	#N/A	unknown protein; EXPRESSED IN: 16 plant structures; EXPRESSED DURING: 10 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G42430.2).	#N/A
396	sp_110171z4101	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
397	sp_110320_uirmq	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
398	sp_111040_mkxq	0.923119	#N/A	#N/A	KOBITO (KOB1); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G57200.1); Has 94 Blast hits to 94 proteins in 32 species: Arohae - 0; Bacteria - 14; Metazoa - 0; Fungi - 0; Plants - 75; Viruses - 0; Other Eukaryotes - 5 (source: NCBI BLink).	#N/A
399	sp_111070_gdwh	0.942754	#N/A	#N/A	uncharacterized protein	#N/A
400	sp_113330_keon	0.933387	#N/A	#N/A	uncharacterized protein	#N/A
401	sp_113751z4321	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
402	sp_114251z4332	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
403	sp_114381z4337	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
404	sp_114891z4370	0.923119	#N/A	#N/A	uncharacterized protein	#N/A

【 1 1 5 0】

10

20

30

40

【表 1 4 - 1 5 1】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAPI)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq #N/A
405	sp_117601z4468 0.923112	#N/A	#N/A	uncharacterized protein	#N/A
406	sp_117630_eqqm 0.923112	#N/A	#N/A	NAC domain containing protein 41 (NAC041); FUNCTIONS IN: sequence-specific DNA binding transcription factor activity; INVOLVED IN: multicellular organismal development, regulation of transcription; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: No apical meristem (NAM) protein (InterPro:IPR003441); BEST Arabidopsis thaliana protein match is: NAC domain containing protein 83 (TAIR:AT5G13180.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archaea - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2998 (source: NCBI Link).	MMKHNFIIEGVSRLLPPGFRFQPTDEELLFQYLRC KIFGLPLASVIPEVHFNLDPWELPGDSKEERY VFTKRGAKHKNGDRNHRTRATKATNSGYWEETG VEEHMSPPNQNLGMPKIGVKKTLVFNIGKSPNY TSPDWMHEYSLVYKHNHQTGSKMEDWVYHVF LKQRVEENDDAIMLDYDNNNGHYNL.TSNIDRYPC CSSSSSSNSFAEVSSYDSSQSDSESSYNFL
407	sp_121070_qtw 0.923112	#N/A	#N/A	uncharacterized protein	#N/A
408	sp_121851z4625 0.923112	#N/A	#N/A	uncharacterized protein	#N/A
409	sp_123481z4690 0.923112	#N/A	#N/A	uncharacterized protein	#N/A
410	sp_124020_ikzx 0.923112	#N/A	#N/A	receptor like protein 2 (RLP2); CONTAINS InterPro DOMAINs: Leucine-rich repeat-containing N-terminal domain, type 2 (InterPro:IPR013210), Leucine-rich repeat (InterPro:IPR001611); BEST Arabidopsis thaliana protein match is: Leucine-rich receptor-like protein kinase family protein (TAIR:AT1G72300.1); Has 90767 Blast hits to 27052 proteins in 1087 species: Archaea - 40; Bacteria - 7835; Metazoa - 18696; Fungi - 1120; Plants - 55519; Viruses - 2; Other Eukaryotes - 7565 (source: NCBI Link).	MOIRNVPALQAINSFHLFILLLYSSFTSYLQACNP NDQQLAFNSLSSTLSPLNWSSDSDCCSSWE GIGCDDNGRVTILWPLRGLGEGISSLGLTLRLS KLNLSCHNLTGFLPDGLFSLKSLKLEIFDISSNFLYG NLSASFTSDGSAFASTIQIVIDVSSNKFHGGIEPYW FFQGSNLSNFASNNDFDTHIPSAICVTCPLRTL DFSENTSGQFTVGLGGCLKLRVLRAGFNYSLG VLPQDIYRSLIELSLPVNNGIGTGKRISRLANL RIINLYSNINNFAGTVPQIEIGLKLLEQLHLHNNLNG SIPASLMNCTNLKILRVNLFQGGKIPDSIGNLPSL FYLDLNMNLRGFEFFLQLSRLPALISDRVAOKVN RNYLQCPVFEISQKQLHVLKLDLKNVFSGRISSPKF ANLNLNEMLDLSONNLTKIPTSLOTLNFLSKFN SYNNLEGAIPGGQFDFTTESSYISNPGLCGRV.LQ SSCSAQPVQRGGYQDADKSKKHILLEGFTTGF IFSVLAIQAYYHRRVLFYSKKR

10

20

30

40

【 1 1 5 1】

【表 1 4 - 1 5 2】

A	B	C	D	E	F
1					
2	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
411	0.923119	#N/A	#N/A	uncharacterized protein	MRRRTTIAELSLTESFMAKAEEDSGIMDPVCEMEE GSEESGYGVERESGLGLGLEGIMGSVDSDEG GFETEVGVNIRGVISGDNAGESYGSILGSAIV ESLNWVVLCEAEFV
412	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
413	0.923119	#N/A	#N/A	Fumarylacetoacetate (FAA) hydrolase family; FUNCTIONS IN: copper ion binding, catalytic activity; INVOLVED IN: metabolic process; LOCATED IN: mitochondrion; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Fumarylacetoacetase, C-terminal-like (InterPro:IPR002529), Fumarylacetoacetase, C- terminal-related (InterPro:IPR011234); BEST Arabidopsis thaliana protein match is: Fumarylacetoacetate (FAA) hydrolase family (TAIR:AT3G16700.1); Has 10822 Blast hits to 10670 proteins in 1907 species: Archae - 249; Bacteria - 6730; Metazoa - 306; Fungi - 435; Plants - 75; Viruses - 0; Other Eukaryotes - 3027 (source: NCBI BLINK).	ATGLPWLLGKGLDTFTSSEMIPVSAVDPHPNLE LWFDVDGESKQKSTKDMIFEIPFLISHIRSFMTL MEGDVILTGTSLV
414	0.942311	#N/A	#N/A	Transmembrane amino acid transporter family protein; CONTAINS InterPro DOMAIN/s: Amino acid transporter, transmembrane (InterPro:IPR013057); BEST Arabidopsis thaliana protein match is: Transmembrane amino acid transporter family protein (TAIR:AT3G28960.1); Has 5045 Blast hits to 4985 proteins in 317 species: Archae - 15; Bacteria - 153; Metazoa - 1796; Fungi - 1055; Plants - 1275; Viruses - 9; Other Eukaryotes - 742 (source: NCBI BLINK).	NGVLVRYLAAGGWWSLGLLFTISAAATYTSLL VKRCMEVDPRIKSYTHIGEYAFGEIGKIIVSILYTD LYVVTTFULLEGDNLNLPKLSLGIPIDGKSSFII LVVALLPVLFDNLSILAYISATGVFASILILVSV WVGVDGVDGHHVDESGFFKWKGIPTAISLYML CYSSHVPFALYTSMQKKHDFSKVLLICFTLSTFI YVSMALVGLMFGSSVESQITLNLPTDKVSSQIAI YTLITPIAKYALIKPIVISTEGWFSNEYQNSKCF KVTIRIVLASQVVIALLPYFGYFMSLTGALLCST ASLTMPCILCYLKITSTNNSRPFQIQILLIWCIVFLS IAILIFGTYTSLVDILAETH
415	0.923119	#N/A	#N/A	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein; CONTAINS InterPro DOMAIN/s: Oxoglutarate/iron-dependent oxygenase (InterPro:IPR005123); BEST Arabidopsis thaliana protein match is: 2-oxoglutarate (2OG) and Fe(II)- dependent oxygenase superfamily protein (TAIR:AT5G9540.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MGTMSVDITKSFQFAYDRYKELKAFDDTKSGVK GIVDSSAQOTVPIRIFRPTDELSEDLNHPFSVKRQVP VISFNGDGTKERDVIQEYSNHVKNLSDVILRLLSV ALGKPEHLKETECQEGEQTLLTNHYYPACPAPEL TMGTSKHSDFSLTLLQDQIGGLQVFNENQWV DIQPIPGAFVYVNIQDMLQISNDKLSKNHRVIANL VGPRIYSVGLFLRG

10

20

30

40

【 1 1 5 2】

【表 1 4 - 1 5 3】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
416	sp_126880_azjn	#N/A	#N/A	alpha carbonic anhydrase 7 (ACA7); FUNCTIONS IN: carbonate dehydratase activity, zinc ion binding; INVOLVED IN: one-carbon metabolic process; LOCATED IN: endomembrane system; CONTAINS InterPro DOMAINs: Carbonic anhydrase, alpha-class, catalytic domain (InterPro:IPR01148); Carbonic anhydrase, CAH1-like (InterPro:IPR018340); Carbonic anhydrase, alpha-class, conserved site (InterPro:IPR018338); BEST Arabidopsis thaliana protein match is: alpha carbonic anhydrase 5 (TAIR:AT1G08065.1); Has 3434 Blast hits to 3405 proteins in 557 species: Archae - 0; Bacteria - 714; Metazoa - 2114; Fungi - 83; Plants - 330; Viruses - 6; Other Eukaryotes - 187 (source: NCBI BLINK).	MEKSIILVFSRALLSLLVLSFATSTIAQEVEDESEF NYIRGDVIRGPNWGNIKDEWEKCKTKGMQSPFI IFPGLPERIPESDRVIRSYQTGQTTLKNRGHDKL RYALEVHFVHQSLDDKIAVIGVYDYGNPDPFLS SIENELKLSNKNSSSEIIMSGIIVNPSDVRMEGSR YRYMGSLLTPPCDEGVIVTETKVGVSLEQV KLLRDVANDDARENARPLQPLNWRRRVSLYDTRL RIEDNDTLITPYVSSM
417	sp_127681z4859	#N/A	#N/A	uncharacterized protein	#N/A
418	sp_128150_hfid	#N/A	#N/A	uncharacterized protein	#N/A
419	sp_129111z4909	#N/A	#N/A	uncharacterized protein	#N/A
420	sp_129520_zhzz	#N/A	#N/A	CRINKLY4 related 3 (CCR3); FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAINs: Protein kinase, ATP binding site (InterPro:IPR017441), Regulator of chromosome condensation/beta-lactamase-inhibitor protein II (InterPro:IPR009091), Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: CRINKLY4 related 4 (TAIR:AT5G47850.1); Has 117697 Blast hits to 115463 proteins in 4296 species: Archae - 105; Bacteria - 13273; Metazoa - 42642; Fungi - 9844; Plants - 33569; Viruses - 607; Other Eukaryotes - 17657 (source: NCBI BLINK).	MKKKASCFPLERCQFFLLVLLFFHITMVSFGFS SSTIAASHGNTVCGILAGKYQGQCFCQKGTFS VLPNISYEAISSGGQDFFCGLSSGGFDLFCWDTIF PSFNHFKRIYHNTFRFLDLDLVGGDDQVAVQAR TGNIRCREARLGLTPTLPRVEETAFLALTS KGFSCGIVKDERNLVLCWGESGGDFIEREFGNM SMLSVAKSHVCGIAFSGLLICKGINDSGQLNVP LERPFHSELALGRYSCALORSNGSLICWGNR GNINTSVITDQKVQNLFSGISISGVYICGLRTDD FTVICWGLSGKLNITSVLSFPMVPGRCVRES QLCGVLPSESLCSGSDTVICRSCVIELPLPKPKI KLEAGNVPPIPSPTAISPPSYPISSPPSKETLR PLWAYEVFVGVGILSGMFAFYFYLYLKRPGSSKD GKKVLDVARDKQHDVSGSSNTGDSAFIASNASS APOSRSSTICLSSMTLGRARSWDYSSKHVEKS EVFTLFEAAATKNFLENRIGRGSFGTYKGL SNIIGCEVAIKREEICTRERGNAFESELVLLTRVH HKHLVSLVGFCEEQNERLLVVEYMSNGALHDHL HTKNTKSNVSNWIKIRINALDAARGIEYLHNY AVPPIIHRDIKSNILLDANVWVDFGSLKGP ENQEGSMSIKPVGTGVYIDPEYVYLVKLVTKSDV YFGVWVLELLTGKAVFKEGKESGPTGWEYA GPFVSVGEVWVLDKRVEIPVMINESEAVELLAYT AMMCVNLEGKERPTMSDVSNLERALTLCKGSS GTFSPFPNFYSFS

10

20

30

40

【 1 1 5 3】

【表 1 4 - 1 5 4】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MFCLVLLSSMYACGYSEKSHPPINPGLTGLG TPGRDGKPGKDGKDGAPGQGRGTGGRGTP GGRTGPTSGGTPGTPGTPGTPGTPR PPNPTHLSALVSTKTRSNIGGPGAPSPVGTGPA PATSGTPTGTCGKPGSPGAPATPGTPTP GTPGSRSTPSPGIPGSPGIPGVPPTGAPGTH CM
421	sp_131590_nhhw 0.964736	#N/A	#N/A	uncharacterized protein	#N/A
422	sp_133801_z511 0.923115	#N/A	#N/A	uncharacterized protein	#N/A
423	sp_134850_udxp 0.939374	#N/A	#N/A	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G01990.1); Has 125 Blast hits to 123 proteins in 14 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 125; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink).	MAYTKCCSSLGCWDLHNNL_GFLNADILSSSYD DMENNNNNNNNNNAASSMLQASDYDNYSS SFSNDLLSTGYLEDALFEFSVRKRRRVMMVSN SNSNSNNNTDFGVNNDYNDHQFATHHHFDDLS CWSNLFANNYDMQSEDEQELSHATLLNHHGHMHC SLGEITEKPIKEMYSVVKIEETNLLLEEAIRR TPKLMDSSSYSKHNSFNKSTSFAYNSDTNST DSSISVVTDEEEKRKNNNNNNNNNKTTTTT TTRKMVYVYFGLVKPGGEGDITLKDINRRILMPPT RPLKHPYVGFACRPVSPNGPGLS
424	sp_135570_aqh 0.929715	#N/A	#N/A	Ankyrin repeat family protein; CONTAINS InterPro DOMAIN/s: Ankyrin repeat-containing domain (InterPro:IPR020683), Ankyrin repeat (InterPro:IPR002110); BEST Arabidopsis thaliana protein match is: Ankyrin repeat family protein (TAIR:AT4G09460.1); Has 38634 Blast hits to 17459 proteins in 653 species: Archae - 49; Bacteria - 2374; Metazoa - 20314; Fungi - 3164; Plants - 3215; Viruses - 213; Other Eukaryotes - 9305 (source: NCBI BLink).	MARIKQPKNGGAALWIKFVPGEDMSTVITLVEK WSHYTCWPDANGSTPLHQAASLNTPYNLQVIRT ILYHWFGSAEYVGDASGKTLHLHMINRMPYQEAT NLLKFEIYALRNYQDCCGNTPLHIAAINLDINMV RVLLESSTKLSIKNMEGNSAASLIQHNLLQMLR KRNMTCEESEAVDRADITFLFRMSKSGMDFLL SQDSKGRNVFPHILMQIKNGSRILHDDDFVFIQOVL GNFPLICQTDLDKGDTPHVLVVRNHPDPTTC

10

20

30

40

【 1 1 5 4】

【表 1 4 - 1 5 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
425	sp_136640_nneo	#N/A	#N/A	Protein kinase family protein; FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation; CONTAINS InterPro DOMAIN(s): Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: Protein kinase superfamily protein (TAIR:AT3G53840.1); Has 114140 Blast hits to 113200 proteins in 4569 species: Archae - 121; Bacteria - 12456; Metazoa - 42740; Fungi - 9504; Plants - 32338; Viruses - 412; Other Eukaryotes - 16569 (source: NCBI BLINK).	MTHTKLVKLLIQLVVWLLSVTKQKAVATLWC PDCGKTPVYPLSTSQSGDQLYKVRCTTGKLV LDTLNGSSYIITISPKIQQLMISSQEFYRNSCTRA DLRSGGWLNQTLPFSDTSNTILKFNCSKSTVRL SSMDCTSSPCYKMKRCRIGLHCCSYTTGGSK TAYRVDIAEASCTAYESFVKLNLPVVRKWPOSSG IEIQWETPQVICTLPLDCHGLANSTCLPDRAVA GQKRCFCNKGFWDPINGTCQSLTFAACGTVII IVFLVLDQHKHKHREAHKRLVKQREDVLSHNN HGKAALFSGKEMKNATKNFSENLLGVGGFGE VFKGNLDDGTAAIKRAKLGNTLGMDOVLNEVCI LCQVNRSLVRLGCLLDLQDPVLVYVYVWNGTL YEHFHGRYSNEWGTLWQRRLLIAKQATADGLAY LHFSVPPHFHRDKSSNILLDEKLNKAVSDFGLSR LVDTATHISTCAQGTGLYLDPEYQNYQLTDRS DVYSFGVLELLTSKKAIDFNREDDVNLAVVYTK LADREKLFDAVDAGLTTQASNLELETMKALGLLA VACVDEQRKRPMSPREVTEELYIISIDKETSLL
426	sp_138800_opgi	#N/A	#N/A	Protein kinase superfamily protein; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN(s): Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase domain (InterPro:IPR002290), Tyrosine-protein kinase, catalytic domain (InterPro:IPR020635), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: Protein kinase superfamily protein (TAIR:AT3G62220.1).	IEWELGKGSYGRVLWGALKWACSSNKTGCKHT TRYEFLLAQVSMVSRLLKHESFVKLLVLMSCYAAVL LFCSTRSVPADLHLFVW MARFNGDESEETPEKDKDYDALNEKQGGPEPDI ANDEYLRRIDFEELADAAALGVVANHPTEGSS SKKGGKRGSGKARGFTKGLGKTGPMFLQFD SLLRPSGKWRKLYGQQIAALHHYITASLHCCPAV HCTACSRFVQCVTVGOOCSAATGQOCSAATGQ QCSEAHVRKTDQVVVQQPMH
427	sp_139610_wxoe	#N/A	#N/A	uncharacterized protein	

10

20

30

40

【表 1 4 - 1 5 6】

1	A	B	C	D	E	F
		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
428	sp_140380_lizax	0.986533	#N/A	#N/A	nucleolar protein gar2-related; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G57780.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archaea - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLink).	MRESEKKVKQNNPPKPSVRTPRENKSVQEN VRKPLNAKESHTKASRAKPPSNISITSDKNVGAEP SEVYENVVHYVDDAHRSEDASQGPKGDEDDVL EDSSSEFGKKGDDSDYESLKGSTSSQDDIQGVD DLKIDGASNVSNKIPNKGSLKVSDTQEKSNKGVK HSHSLPREPMKSDGSKVVPKDDSSSECSEGETDK PIEVTNQEICSGSDDDTVCAKGNMTEQERAAALK QKAEIMENRIEKLEDELRYWASLEISLSYVPEHGS SAHKVHTPARRLSRLYHACKHWTRSKRATIARN TVSGLILVSKTCGNDVARLTFWLSNTVW_LREISQ AFGISGKSPARFSAALNGSTKGSSEGNAYPGKLG GHSSGKQVNIKISQLFEDWEETSTFTSALEKIESW IFSRVVEVWVWQALTPNMQKPVEDLIRQKISGKL LGFSLGDDQQGGTFSIELWKNFAQGAERLCPVR AAGHECGCLPVLARMVMEQCVRRLDVAMFNAIL RESAHQIPTDPVSDPIVDSRVLPIPAGDFSGGGA QLKNSVGNWSRWFSDKLDMEEDSLPIDDTEYTS DDHNIHDKTKGFFLLNELSDLLMLPKDMLMDRSI RYEVCPSVNLGVKIKRILCYFTPDEFPCDPAVPGEVL EALNAESFVERKVGGEAFGSSFFYGAPVVSFFPP SSGDVSEKVAETGGKPGVERNIMCIVQKKGYNMS DEELDDIGSPLTSFIEQSSPRILQINGNAGLKSLEQ TAGCIVNSRYELLREVVWSL

10

20

30

40

【表 1 4 - 1 5 7】

A	B	C	D	E	F
1					
2	分析に用いたペイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
429	0.923119	#N/A	#N/A	cation/H+ exchanger 20 (CHX20); CONTAINS InterPro DOMAIN/s: Cation/H+ exchanger (InterPro:IPR006153); BEST Arabidopsis thaliana protein match is: cation/H+ exchanger 18 (TAIR:AT5G41610.1); Has 10038 Blast hits to 10006 proteins in 2135 species: Archaea - 380; Bacteria - 7898; Metazoa - 11; Fungi - 167; Plants - 678; Viruses - 0; Other Eukaryotes - 904 (source: NCBI BLink).	MVSFSLCSTPLOGTEEFRQRKQSPVQVPRIV NNTNANQRLLMDIKLMKGRVEFLKHANKVNEE HEKMKYEMSMIAEGRSKTEIKNDNVYLLSLCDV ISVVGILLGPSALGRSHQYMDKVPKWSPTILE SVASIGLLFFFLVGLDLGSIIRSRGKEAFSIALA GISLPFIGGAAVAVILRRITIEGADHAGYAPFLVFM GVLSITAPFVLAARLAEKLLTKVGETAMAAAFA NDVAAWILLALAVLAGTGGEGGHVSPVLSLW VLLAGVGFVAFMLAIKPLMGWMMKRCSAEQGV VDELVITLAVLVAGFITDFIGHSIFGAFIFLIIP KEGEFAGKLLERIEDVSGLLPLVFASSGLKTDV TKIQGAVAWGLLVIAIACFGKIFGTFYVAKLCR MKGREALGILMNTKGLVELIVNIGKEKKVIND EMFAILVLMALFTTITPLVMAIFNPAKHGSDIET QKRKLQDPTSIHKDDTSSTSKSQQFRILACAHG PGNVSGLINLIESTRSSAVKSNLKYIMHLVELTE RSSIMMVHRLRRNGLPFIDPTRRRSRVVEHDRV ASAFETYAQLGHVMVRPVTTISALSTMDDVHRV AEGKRVMMILPFHKMMWRKDEESGEMKVVENLG HGWRGVNQRMLKNPSSIAVFLDRGFGGGRVT QNESTSQRQLGCVVFFGGPDDREALELGARM MDHPERIVTVIRFIETQGMDSNVKLMPSPEKCS DSNTYFSTVAMNTEEKAENEDEVVEEFRNKWE GTMFIEIQOANNIEEAVLTLGRNGEYDLIVGKG RFPSTMVAELADRAQAEHAELGPIGDILSSSNGIV SSVLVVOGHTMVHVEEMPVSRVVDGEDNHINDA TNQV
430	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
431	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
432	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
433	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
434	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
435	0.935417	#N/A	#N/A	Protein of unknown function (DUF803); CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF803 (InterPro:IPR008521); BEST Arabidopsis thaliana protein match is: Protein of unknown function (DUF803) (TAIR:AT1G34470.1); Has 1265 Blast hits to 1244 proteins in 219 species: Archaea - 0; Bacteria - 92; Metazoa - 420; Fungi - 372; Plants - 263; Viruses - 0; Other Eukaryotes - 118 (source: NCBI BLink).	MASSSDNSDPDKGMSADNIKGLVMALSSSFFIGA SFVYKKGKKGAGAGVVRAGVGGYSYLVEPLWW IGMITMIVGEIANFAAYAFAPAILVPLGLALSISAV LAHVILKEKLHIFMGLGCVLVGSSTTVLVHAPVE REIESVIEVWNLATEPGLFVYAAALIMTTVFLIFHVI PEYQGTHIMVYIGVCSLVGSLVSMVSKALGIAYKL TLSGINQLIYQWTWAFSMIVAAGIITOMNYLNKAL TDFNTAVVSPYYVMFTSFTILASVIMFKDWNQRON PTQIVTEMCGEFTILSGTFLHLKTKDLDSDGVSSCL PVKLSKHIGENGFEDESEGIPLRRQDSVSRST

10

20

30

40

【 1 1 5 7】

【表 1 4 - 1 5 8】

A	B	C	D	E	F
1	分析に用いたバイトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
436	sp_155690_sqcd	#N/A	#N/A	<p>Glycosyl hydrolase superfamily protein; FUNCTIONS IN: cation binding, hydrolase activity, hydrolyzing O-glycosyl compounds, catalytic activity; INVOLVED IN: carbohydrate metabolic process; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN's: Glycoside hydrolase, catalytic core (InterPro:IPR017853), Glycoside hydrolase, family 5 (InterPro:IPR001547), Glycoside hydrolase, subgroup, catalytic core (InterPro:IPR013781); BEST Arabidopsis thaliana protein match is: Glycosyl hydrolase superfamily protein (TAIR:AT4G28320.1); Has 827 Blast hits to 815 proteins in 205 species: Archae - 18; Bacteria - 244; Metazoa - 38; Fungi - 187; Plants - 310; Viruses - 0; Other Eukaryotes - 30 (source: NCBI/BLINK).</p>	<p>MKKTNNSEWREKYAYPIGIATISTILYLNFNDQ LNFEEPIFWQKMGVGVNSTNFIHGHDDHHQ DDRRVYVNGWVWYLMWVWDPSPRSKVKM LKRGAOMGLNVCRTWAFSDGGPNLQILPGVF NERVLKGLDYVIVEAKNGIRLIFLSVNLNLFAGG KAQYVWQADAGINISSDTSFFADSTIKGYKA YKAIVARKNSLSGVYIIEEPAIFAWELMNEPRCD SDCSSILQAWITEMVAYIKSLDQKHLVITGLEGFY GPKTAERKSVNPGEWAASLGTDFMQNSAIDGID FASVHAYPDSWMPHANTTAKLDFLOMMLDSSHIS DGNMLKLPILFTEVGYPIKDGKGGISYGDTSLKT VYDSIFSAKKGQAGAGALIWQLLVEGMEEYGD RFSLVAWEQFSTYKMMVEQSCCKLSNINTGKRKS AGELNGKDFPCSGIKYIP</p>
437	sp_155810_mczy	#N/A	#N/A	<p>curculin-like (mannose-binding) lectin family protein / PAN domain-containing protein; FUNCTIONS IN: sugar binding, protein kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation, recognition of pollen; LOCATED IN: endomembrane system; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN's: Curculin-like (mannose-binding) lectin (InterPro:IPR001480), PAN-2 domain (InterPro:IPR013227), Apple-like (InterPro:IPR003609), Protein kinase, catalytic domain (InterPro:IPR000719), S-locus glycoprotein (InterPro:IPR000858), Serine-threonine/tyrosine-protein kinase (InterPro:IPR001245), Protein kinase-like domain (InterPro:IPR011009); BEST Arabidopsis thaliana protein match is: lectin protein kinase family protein (TAIR:AT1G34300.1); Has 27641 Blast hits to 27042 proteins in 566 species: Archae - 0; Bacteria - 139; Metazoa - 3989; Fungi - 187; Plants - 22616; Viruses - 64; Other Eukaryotes - 666 (source: NCBI/BLINK).</p>	<p>IMVTRCIFVGTSAITMLCQFLVSLVFSVAVSEISLGS KLTVSENQLWVSSNGDFALGFTHSDNVNQYSL GIRFHSGSIPVDKQSVWVAGAEITVSNQSYFQL MEEGGLVLDPSNNGVWWTNTSGKSVASASL QDDGNLVLIDGNKDMWQSFDFPDTLLPGQV LSVHKTLRALSGNLVSSHYSLYLNDGGQLQFLRW ETGVLYWRGSPNSLSKASLSSNGSFLQDDRR SEAFWVFAEDHNDNWKFRFLRLDVGDLNLRYS WIDESVSWKSVWQAVENQCDVFATCGDQGGIC SSRANSLGKCPFGSAMENSKCLAPYRQQSC SSRAMVEYKHTSLYGYPPSDTIITSLDFCQRLCE KDFLCTAVTYLNDGSRARCOIKKTRFITGSSDPSLS SISFVKCTDPAFVLPNKSPATAASPSLPSLSSKS NQSSRFPIRVIDAALVSFCLFCMLQIAGVFWVC KRRRTAVRKASPPSHGLISAGLVRWVTTYGEIKEL TGNFKDQLGKSVYKSVLQNNQPAVVKRLESQIE ERRFRAMVSMGSIHKNLVKVEGYCCESGNRFL VYELLKNGSLEDHMDPASCCKLSWKNRMLICL DVAKAISYLHTGCRDFVSHGKLCGNVLDENLE AKVTEFLGKLDGIETYAEKIDVDFGKMWLYLIC GRQEEVCEWSYKMWVSGYGRVWVVKRIEGEV DFEELERGLRIAFWCLQVDERMRPMSMGEVNVL EGLTIVDPPPPFPFSDSRNLQEDILAESELY</p>

10

20

30

40

【表 1 4 - 1 5 9】

A	B	C	D	E	F	
1						
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	
438	sp_157070_pyrh	#N/A	#N/A	bidirectional amino acid transporter 1 (BAT1); FUNCTIONS IN: arginine transmembrane transporter activity, L-lysine transmembrane transporter activity, L-alanine transmembrane transporter activity, L-glutamate transmembrane transporter activity; INVOLVED IN: amino acid transport, transport, transmembrane transport; LOCATED IN: integral to membrane, membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Amino acid permease subfamily (InterPro:IPR004756), Amino acid/polyamine transporter 1 (InterPro:IPR002293), Amino acid permease domain (InterPro:IPR004841); Has 15465 Blast hits to 15422 proteins in 1894 species: Archae - 296; Bacteria - 11987; Metazoa - 438; Fungi - 2169; Plants - 254; Viruses - 0; Other Eukaryotes - 321 (source: NCBI BLink).		MDSIDVDRPQLLGSERSISDDNRLRQLQQLGYKQEL INFSLSAIANFAVFTSIISVLTGLTFMFMNGLAYGG PVTMYGWPVIGIMTMMVGLSMAEICSAPFTSG GLYFWSAKLGNRWPFPASWLTGWFNIVGQW AVTTSVDYSLAQMIQVILLSTGGKNGGGYEASKY VMGFHGGLLHMLNLSLHISWLSFFGQFAAW NVAGVFLMILVPLVAKAKASAEFVTHFHTDNG EGKSVYFVGLLMSQYTLGSDASAHMTEETKN ADKNGPKGISAVISIIVGWYGVITFAVDIPYLL SSENDAGGYAIAQVFLAFKTRYGTGIGILCLI VAVAVFFCGMSSMTSNRMAYAFSRDGMAMPMS SYWQKVNKQEIPIINAYWMSAFISFCMALTSLGSL VAFAEMYSIATIGLYISYALPIFFRYTIARKSVFPGP FNLGRYGVIGWIAVITWITISILFSLPVAYPINTE TLNYTPVAVGGLLILTVSYWVFSAKSWFKGLKN VTV
439	sp_160031z6573	#N/A	#N/A	uncharacterized protein	#N/A	
440	sp_161020_ustu	#N/A	#N/A	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein; CONTAINS InterPro DOMAIN/s: Oxoglutarate/iron-dependent oxygenase (InterPro:IPR005123); BEST Arabidopsis thaliana protein match is: 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein (TAIR:AT3G61400.1); Has 8339 Blast hits to 8299 proteins in 1003 species: Archae - 0; Bacteria - 1128; Metazoa - 122; Fungi - 867; Plants - 4923; Viruses - 0; Other Eukaryotes - 1299 (source: NCBI BLink).		MYSSKSSEIYDRKGDIAKAFDESKLGVKGLVD AGVTLLPRIFWQEATAPRESEVYVYTTSDNPHV NIPVIDLSGVWEDGDLEKFRFVVKAVSEACEK WGFQAVNHGVPTHYLDEMINGAKRFHELDVES KKEYVMVRGNSGRKFSYSNMYLYTKVYTNW RDTTIAGMEFMPDPQDFPFCREITIEYAKHMKK VALTFLELISEAMGVESNYLKDMDCCQEFLLHT YPPCPQPELAIGINQHADNDIVTLQLQDEIGGLQ VLHDNQWHDIPHIPGAVIINTGDLLQLSNDKYKS VIHRVQSKVGVPRISVPCFFRPKTNPRILAPIKEL LSPENFAYRPTTTTGEYIAHYHSTGQDYGVKPKAL DHFKLAN
441	sp_163000_inwc	#N/A	#N/A	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G34090.1); Has 73 Blast hits to 73 proteins in 21 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 65; Viruses - 0; Other Eukaryotes - 8 (source: NCBI BLink).		KRGFYFAWRQTWHFPPCNTFVKQLRCYDFES VTVANSTIWDQYIHRMLVEEDACSPKDRNRELL HYSANAGKLYREGSFIDSQLPNLDYHLRKVGVF PDILERKINKHLEKGNRVYALITGEFYAKKHQHPG FGRPVAFNAHILMRVGRDLEAKDAARGALKSPW WTLGYQVEVAEIAKWKDEIQCYVKKKLKLTDEGK REDLSRQKELAQIALDQAAFLDLASVEGNWDDF VRDISECYKVAAGLHDMARFIAYKS MPGAKFWKYNLAAPAPPRYKVPQGRPKANKR RLEAGEKGGKGGKAKGDQPTQAEHPKPKPKON KCGKCGGLGHNRITCKSTGPKSEKPNKGGGR PTSNTPWVAEQRRKKEERVHKAAVATAACAGSS SQPQSQVTSQQSQP-HQH
442	sp_163030_xjsc	#N/A	#N/A	uncharacterized protein		

【 1 1 5 9】

10

20

30

40

【表 1 4 - 1 6 0】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
443	sp_164920_emoj 0.323119	#N/A	#N/A	uncharacterized protein	EEKRGEAGEDEAGKQGTWEYGGVSGWSPSGA GVVVEQWWEVWVWVAVVWVNGGCVVVP
444	sp_165380_awih 0.959029	#N/A	#N/A	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: root; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G17350.1); Has 261 Blast hits to 261 proteins in 13 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 261; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).	MFCERFPDRCDJTTCDFPASYLCRHFPECAPNNKA LRALSCSYYPKATESLRMMLKHCDTYTTSVHWK SVGGIGTSOPYDQVPEYGVYRVDIFEIPLTASCLR CQDRAKGGTGGFDTSNODFVCLCEERNVTY CKDHQVERHGGHTKAVAGTTVASIAGVGVIGAG V
445	sp_166571z7270 0.323119	#N/A	#N/A	uncharacterized protein	#N/A
446	sp_166761z7281 0.323119	#N/A	#N/A	uncharacterized protein	#N/A
447	sp_166881z7285 0.323119	#N/A	#N/A	uncharacterized protein	#N/A
448	sp_168211z7384 0.323119	#N/A	#N/A	uncharacterized protein	#N/A
449	sp_170411z7481 0.323119	#N/A	#N/A	uncharacterized protein	#N/A
450	sp_171160_fxje 0.323375	#N/A	#N/A	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein; FUNCTIONS IN: lipid binding; INVOLVED IN: lipid transport; LOCATED IN: endomembrane system; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Bifunctional inhibitor/plant lipid transfer protein/seed storage (InterPro:IPR016140); Plant lipid transfer protein/seed storage/trypsin-alpha amylase inhibitor (InterPro:IPR003612); Plant lipid transfer protein/hydrophobic protein, helical domain (InterPro:IPR013770); BEST Arabidopsis thaliana protein match is: Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein (TAIR:AT1G48750.1); Has 308 Blast hits to 307 proteins in 31 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 308; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).	MATKVATLLLAAVVAVMLLVETPATEAVTCSPV ELTPCAPVIMYNOAPSNAACCEKMGEQKPCLCGY AKNPTLKSINSPGAKRFMASTCKVDVKG

10

20

30

40

【表 1 4 - 1 6 1】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 Peroxidase superfamily protein; FUNCTIONS IN: peroxidase activity, heme binding; INVOLVED IN: response to oxidative stress, oxidation reduction; LOCATED IN: cell wall, vacuole, membrane, plant-type cell wall; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Haem peroxidase (InterPro:IPR010255), Plant peroxidase (InterPro:IPR000823), Peroxidases haem-ligand binding site (InterPro:IPR019793), Haem peroxidase, plant/fungal/bacterial (InterPro:IPR019794), BEST Peroxidase, active site (InterPro:IPR019794); BEST Arabidopsis thaliana protein match is: Peroxidase superfamily protein (TAIR:AT2G18150.1); Has 4523 Blast hits to 4498 proteins in 279 species: Archae - 0; Bacteria - 4; Metazoa - 1; Fungi - 207; Plants - 4264; Viruses - 0; Other Eukaryotes - 47 (source: NCBI BLINK).	aa seq MGSSKFLPFLQLVLCSSLLLSISNGSIVPVVPLS YTFYDSNCPGLEIIVRSHLWVFNINITQAAGLLR LHFHDCFOGCDGSLVLDGSGTSGPSEQDAPPNL TLRAAFTIIDLRLALVHNACGPIVSCADIVALAAR ESVFLPNGPYQIPLGRRDSYVFATRPVTLANLP PPTFNITSLINFAFAAKNLTTDLVLSGGHTIGRG HCASFANRLYPTQDSTLQTFANLKLCTPNTT DNSTNLDIRSPNLFDNKYFYDLMNRRQGLFTSDQD LYTDSRTFNIVTSAVFNQSLFFKEFVDFDGLKMGQ LGVLTGTQGEIIRTNCSATNVNINIDLASV
451	sp_171340_ufrj	#N/A	#N/A	expansin-like A2 (EXLA2); INVOLVED IN: response to cyclopentenone, plant-type cell wall organization, unidimensional cell growth, plant-type cell wall loosening; LOCATED IN: endomembrane system, extracellular region; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Banwin-related endoglucanase (InterPro:IPR009009), Pollen allergen, N-terminal (InterPro:IPR014734), Rare lipoprotein A (InterPro:IPR005132), Major pollen allergen Lol pl (InterPro:IPR005795), Expansin/Lol pl (InterPro:IPR007118), Expansin 45, endoglucanase-like (InterPro:IPR007112), Pollen allergen/expansin, C-terminal (InterPro:IPR007117); BEST Arabidopsis thaliana protein match is: expansin-like A1 (TAIR:AT3G45970.1); Has 1877 Blast hits to 1874 proteins in 122 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 1869; Viruses - 0; Other Eukaryotes - 8 (source: NCBI BLINK).	MSLISLMLFFFFFFFAYASACDQCVLKASVFTKA SSLSSGCGYGSLLALFNNGGHVAAGVPSLYKDG AGCGACFQIRCKNSGICNKAGTKVIMSDVKNNSN STDFVLSRRALPAMALHGKDDQQLKLGWVDEYK RIPCXYGQNLAVRVEESSKNPGYLAIKVLYQGG QTEIVADIRVGSNNWIFMSRKYGVAVWETSQYP GPGPLQLRFVVTSGFDGKNVVAENVLPANWKT GVTYDSKVQIDDIALLEGCTPCADSHWN
452	sp_171440_eirj	#N/A	#N/A		

10

20

30

40

【表 1 4 - 1 6 2】

	A	B	C	D	E	F
1						
2	共染現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
453	sp_172180_sdki	0:9243119	#N/A	#N/A	SIGNAL PEPTIDE PEPTIDASE-LIKE 2 (SPPL2); FUNCTIONS IN: peptidase activity, aspartic-type endopeptidase activity; INVOLVED IN: proteolysis; LOCATED IN: endosome, plasma membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protease-associated PA (InterPro:IPR003137), Peptidase A22, presenilin signal peptide (InterPro:IPR006639), Peptidase A22B, signal peptide peptidase (InterPro:IPR007369); BEST Arabidopsis thaliana protein match is: SIGNAL PEPTIDE PEPTIDASE-LIKE 4 (TAIR:AT1G01650.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	VYATEVGTVLLSCAFFYDIFWVFKWFFHESVM IMVSVVYVLFNSEGGCFSLQL
454	sp_173760_jwjm	0:953119	#N/A	#N/A	multidrug resistance-associated protein 14 (MRP14); FUNCTIONS IN: ATPase activity, coupled to transmembrane movement of substances; INVOLVED IN: transport, transmembrane transport; LOCATED IN: plasma membrane; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: ATPase, AAA+ type, core (InterPro:IPR003593), ABC transporter-like (InterPro:IPR003439), ABC transporter integral membrane type 1 (InterPro:IPR017940), ABC transporter, transmembrane domain, type 1 (InterPro:IPR011527), ABC transporter, transmembrane domain (InterPro:IPR001140), ABC transporter, conserved site (InterPro:IPR017871); BEST Arabidopsis thaliana protein match is: multidrug resistance-associated protein 6 (TAIR:AT3G21250.2); Has 673798 Blast hits to 360971 proteins in 4075 species: Archae - 12279; Bacteria - 544381; Metazoa - 13093; Fungi - 8967; Plants - 6941; Viruses - 21; Other Eukaryotes - 88116 (source: NCBI BLINK).	MRLNGITKSMVANHLAESLAGAITIRAFQEEERFC SKNFDLIDKKNARFFNFNFAATEW_LIQRIEILSAVVL ATAALCMVLLPNGTSSGGFIMGALSYGLGLNLGM VFSIQNQCILANQISVERLYQYMDIPSEALEVIND SRPPQNWPVSGKVEYNLQIRYRIDTPLVVKGINC VIEGDKVGIIVGRTSGSGKTLIGALFRFLVEPVDG KIIVDGVDCISGLNDLRSHFGIIPQDPTLFNGSVR FNLDPLCQHTDEEIVEV
455	sp_174161z7883	0:9243119	#N/A	#N/A	uncharacterized protein	#N/A

10

20

30

40

【表 1 4 - 1 6 3】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
456	sp_176310_fnxx 0.9583037	#N/A	#N/A	Polyketide cyclase/dehydrase and lipid transport superfamily protein; CONTAINS InterPro DOMAIN/s: Polyketide cyclase/dehydrase (InterPro:IPR019587); BEST Arabidopsis thaliana protein match is: Polyketide cyclase/dehydrase and lipid transport superfamily protein (TAIR:AT4G32870.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI/BLINK).	MESVSSHKKWKQKARVELAGYKAEVWPPFLAD FOGLNKWFPTLSTCIRVEGVEGQPGCVRYCAGF KTHIEHANENHVNWTKCKLLSIDPDNMTFSYIID GNVGFSSYYSKVRVLLQVDGCVIEWDYEVEPVE GWKPKDLDEFIIGKGLKVMNAKRMEALKGRAMIN MIECCGGSWLEFASGLRGGVWVWDEGLRCCGG AAWLMRRRRELKMLLCLVYVHFKECGVG
457	sp_177510_rknn 0.923119	#N/A	#N/A	uncharacterized protein	
458	sp_181790_jczu 0.922226	#N/A	#N/A	aldehyde dehydrogenase 2B7 (ALDH2B7); FUNCTIONS IN: 3-chloroallyl aldehyde dehydrogenase activity, aldehyde dehydrogenase (NAD) activity; INVOLVED IN: oxidation reduction, metabolic process; LOCATED IN: mitochondrion; EXPRESSED IN: 13 plant structures; EXPRESSED DURING: LP 06 six leaves visible, 4 anthesis, LP 10 ten leaves visible, 4 leaf senescence stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Aldehyde/histidinol dehydrogenase (InterPro:IPR016161), Aldehyde dehydrogenase (InterPro:IPR015590), Aldehyde dehydrogenase, N-terminal (InterPro:IPR016162), Aldehyde dehydrogenase, conserved site (InterPro:IPR016160); BEST Arabidopsis thaliana protein match is: aldehyde dehydrogenase 2B4 (TAIR:AT3G48000.1); Has 60835 Blast hits to 60447 proteins in 2994 species: Archae - 474; Bacteria - 34339; Metazoa - 2649; Fungi - 2135; Plants - 1675; Viruses - 0; Other Eukaryotes - 19563 (source: NCBI BLINK).	MNVDKLAFTGSRRTGKLVLELASKSNLKPFTLEL GGNSPFIICENADIHQAMETTCFALSFQPGTMLL CWVLYLCRP
459	sp_182041z8496 0.923119	#N/A	#N/A	uncharacterized protein	#N/A
460	sp_182310_umxi 0.923119	#N/A	#N/A	uncharacterized protein	#N/A
461	sp_182991z8566 0.923119	#N/A	#N/A	uncharacterized protein	#N/A
462	sp_183491z8707 0.923119	#N/A	#N/A	uncharacterized protein	#N/A

【 1 1 6 3】

10

20

30

40

【表 1 4 - 1 6 4】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたペイトとの相関係数 sp_107660_kicg sp_107670_ptqx (SOAP2-like)		注釈	aa seq
463	sp_183550_aazg	#N/A	#N/A	Squamosa promoter-binding protein-like (SBP domain) transcription factor family protein; FUNCTIONS IN: DNA binding, sequence-specific DNA binding transcription factor activity; INVOLVED IN: regulation of transcription; LOCATED IN: nucleus; CONTAINS InterPro DOMAINs: Transcription factor, SBP-box (InterPro:IPR004333); BEST Arabidopsis thaliana protein match is: Squamosa promoter-binding protein-like (SBP domain) transcription factor family protein (TAIR:AT5G50670.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MEWDESGWNGVSKLEIEDNISYHHHLSPLAA GNITTAAIRYARGTSSKTPVIVAGKEORFCQCS RFQSLFEFDEVKRSYRKRLDGHNRRRRKRQPEA FYMFNCAAGFLSSHTADGMVQFCPCQSHTVNW PTMSGPPDLYLRNQEITYEPLLVGVGQIGNRTV KVSADNSGFSLSLLSRYSSDDICVESAMQPTTMS STGQGGGSCSTSLHNNNSIVQFSCSHGLEDMIE TSASLSRTHNITLQSGHDEFDRPSSRIFFPGW E
464	sp_183911z8741	#N/A	#N/A	uncharacterized protein	#N/A
465	sp_184101z8837	#N/A	#N/A	uncharacterized protein	#N/A
466	sp_185291z8971	#N/A	#N/A	uncharacterized protein	#N/A
467	sp_185301z8976	#N/A	#N/A	uncharacterized protein	#N/A
468	sp_186321z9043	#N/A	#N/A	uncharacterized protein	#N/A
469	sp_187700_qkcc	#N/A	#N/A	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G73940.1); Has 58 Blast hits to 58 proteins in 12 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 58; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).	MVFRTRAILRMTSKKPKMKPIELKGGQPEQTQT ITRAIFDILKEGGPLTIAQTWERLQEVGLKDLTSKT QMKTYMKWMKERQKLMICNHNHVGSHKQFLYTT WFTKENIKPAKPGSYDSKE
470	sp_190361z9640	#N/A	#N/A	uncharacterized protein	#N/A
471	sp_190601z9675	#N/A	#N/A	uncharacterized protein	#N/A
472	sp_192281z9788	#N/A	#N/A	uncharacterized protein	#N/A
473	sp_192751z9929	#N/A	#N/A	uncharacterized protein	#N/A
474	sp_193180_pccw	#N/A	#N/A	uncharacterized protein	#N/A
475	sp_193311z9971	#N/A	#N/A	uncharacterized protein	#N/A
476	sp_193621z9987	#N/A	#N/A	uncharacterized protein	#N/A

10

20

30

40

【 1 1 6 4】

【表 1 4 - 1 6 5】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
477	sp_194260_zfku	0.923119	#N/A	#N/A	Peroxidase superfamily protein; FUNCTIONS IN: peroxidase activity, heme binding; INVOLVED IN: oxidation reduction, response to oxidative stress; LOCATED IN: endomembrane system; EXPRESSED IN: root; CONTAINS InterPro DOMAIN/s: Haem peroxidase (InterPro:IPR000823), Peroxidases haem-ligand binding site (InterPro:IPR019793), Haem peroxidase, plant/fungal/bacterial (InterPro:IPR02016), Peroxidase, active site (InterPro:IPR019794); BEST Arabidopsis thaliana protein match is: Peroxidase superfamily protein (TAIR:AT1G49570.1); Has 4658 Blast hits to 4636 proteins in 311 species: Archae - 0; Bacteria - 0; Metazoa - 8; Fungi - 376; Plants - 4203; Viruses - 0; Other Eukaryotes - 71 (source: NCBI BLINK).	MNSNKYFSALFVIITLGLISAVSCODLSCNYYDKD CENFEIIVAKKVVWVWFKKDKTIAASLRLHFHDC AVRFGDASILLDYPGSEKWAATSSKTLRGEVIDD IKATLEKCKPNLVSQADILTAARADATVLVKGPYV SNEYGRKDGJLSVAQEAADDLPNGREDVTTLIEFY QALGLNMLDMVLSGAHTIGRSTCGAIKINRLSRY DENGKLSHTIMAPTLYLNYLRKCKNDNTYVDLDA RTPYTFDNEYKNIKEMKMGLLYSQALYTDHRTS EYIATLANDDGGSPFYLFSSSMVKLGNILGDAQD DGEGEVRQHC SRLNSGY
478	sp_194421z10073	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
479	sp_194650_fick	0.923119	#N/A	#N/A	ATMP4K ALPHA1; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: response to salt stress, hyperosmotic response, response to wounding; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase domain (InterPro:IPR002290), Tyrosine-protein kinase, catalytic domain (InterPro:IPR020635), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009); BEST Arabidopsis thaliana protein match is: Protein kinase superfamily protein (TAIR:AT3G15220.1); Has 132200 Blast hits to 130116 proteins in 4312 species: Archae - 185; Bacteria - 15420; Metazoa - 49247; Fungi - 12380; Plants - 32481; Viruses - 583; Other Eukaryotes - 21904 (source: NCBI BLINK).	MADVAAFIEAAGSRFGDLELIGRSGFDVYKGFDF KELNKLVAIKVIDLEESDIEDIQEISVLSQCRS QYITEYYGSLVHOTKLWIMEYVWAGGSVADLLQS GPPLEISISACILRDLHAIIDYLHAEQKIHHRDIKAAAN ILLTENGDKVAVDFGVSAQLTKTISRRK
480	sp_194691z10097	0.923119	#N/A	#N/A	uncharacterized protein	#N/A
481	sp_195441z10201	0.923119	#N/A	#N/A	uncharacterized protein	#N/A

10

20

30

40

【表 1 4 - 1 6 6】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
482	sp_197700_ghsd	0.988837	#N/A	#N/A	Major facilitator superfamily protein; FUNCTIONS IN: carbohydrate transmembrane transporter activity, sugar:hydrogen symporter activity; INVOLVED IN: transport, transmembrane transport; LOCATED IN: integral to membrane, membrane; EXPRESSED IN: leaf whorl, sepal, flower; EXPRESSED DURING: petal differentiation and expansion stage; CONTAINS InterPro:IPR005829), Major facilitator superfamily (InterPro:IPR020846), Sugar/inositol transporter (InterPro:IPR003663), General substrate transporter (InterPro:IPR005828), Major facilitator superfamily, general substrate transporter (InterPro:IPR016196); BEST Arabidopsis thaliana protein match is: Major facilitator superfamily protein (TAIR:AT4G04750.1); Has 23023 Blast hits to 22536 proteins in 1742 species: Archae - 399; Bacteria - 9001; Metazoa - 4095; Fungi - 6023; Plants - 2554; Viruses - 0; Other Eukaryotes - 951 (source: NCBI BLink).	MPTNSDRFDGDWKKIKALISSKTVIQVGYTTPTQ TSIVKDLNLSEAEWIDPRLP #N/A
483	sp_198561z10674	0.923113	#N/A	#N/A	uncharacterized protein	#N/A
484	sp_198890_xshx	0.723222	#N/A	#N/A	Cornichon family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: intracellular signaling pathway; LOCATED IN: endomembrane system, membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Cornichon (InterPro:IPR003377); BEST Arabidopsis thaliana protein match is: Cornichon family protein (TAIR:AT1G12390.1); Has 605 Blast hits to 605 proteins in 175 species: Archae - 0; Bacteria - 0; Metazoa - 316; Fungi - 159; Plants - 93; Viruses - 0; Other Eukaryotes - 37 (source: NCBI BLink).	MHWDLILW_LISLISNIVLLVLLIYQLICLSDLEADYM NPYETAANINSILPEFGLQAFCALFLVVGHFFM FLVTLPAVYHARLFARGEHLVDVTEIFRALSVEK KHRUKLGLYLLFFFLVFRFRLVIAIYNSLADDEEEALH GFWVF #N/A
485	sp_199031z10717	0.923113	#N/A	#N/A	uncharacterized protein	#N/A

10

20

30

40

【表 1 4 - 1 6 7】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
486	sp_199840_kgne	#N/A	#N/A	NIMA-related kinase 2 (NEK2); FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: cellular component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase domain (InterPro:IPR002290), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR009271); BEST Arabidopsis thaliana protein match is: NIMA-related kinase 3 (TAIR:AT5G28290.1); Has 124376 Blast hits to 122221 proteins in 4182 species: Archae - 113; Bacteria - 14280; Metazoa - 45892; Fungi - 12262; Plants - 30619; Viruses - 487; Other Eukaryotes - 20723 (source: NCBI BLINK).	LSIFVKKVHPVIESYADYGLRSLAVATOVCAAGP MSMLQLHETLYLFCELDMTKYFCYGLSVGSTE ENKRESTVAPLPAKYSGAFRGLKRLRNPKLFR PSADELLRHPLLPQYVYVEKIHLHTRNPMCDIADDN WDDTDSNCIDKATLPEAEDEKIDIPSMSEAEIIDS FYDDRIKPTLSDLERRLAKSKAODPFRRAAATK OVPSKASNDVRNPNPLSRAKTFTGTSKRRRPELSKN KQSVSFDASM
487	sp_200701z11025	#N/A	#N/A	uncharacterized protein	#N/A
488	sp_200741z11027	#N/A	#N/A	uncharacterized protein	#N/A
489	sp_200981z11050	#N/A	#N/A	uncharacterized protein	#N/A
490	sp_203671z11605	#N/A	#N/A	uncharacterized protein	#N/A
491	sp_203861z11635	#N/A	#N/A	uncharacterized protein	#N/A
492	sp_204020_zoaq	#N/A	#N/A	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: anchored to plasma membrane, plasma membrane, anchored to membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G30700.1); Has 65 Blast hits to 65 proteins in 12 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 65; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).	ACFFNFIAAESIMGITATDCFTAFAPYLANVCCP QLHATLVILIGSSKKTMLFALNETLSKQCLSDVN QVLASQGANDNLQQICKVWESNIEGSLVKDV REFEKTIDTSLKLMASC
493	sp_204651z11817	#N/A	#N/A	uncharacterized protein	#N/A
494	sp_204990_emnw	#N/A	#N/A	uncharacterized protein	MKPKRLVFCSSRRPSSSFQSPFPKLDKAERVIS SSTYQHW

10

20

30

40

【 1 1 6 7】

【表 1 4 - 1 6 8】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
495	sp_206970_eizs	#N/A	#N/A	Pentatricopeptide repeat (PPR) superfamily protein; CONTAINS InterPro DOMAINs: Pentatricopeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: pentatricopeptide (PPR) repeat-containing protein (TAIR:AT2G22070.1); Has 46650 Blast hits to 13875 proteins in 226 species: Archaea - 0; Bacteria - 4; Metazoa - 74; Fungi - 58; Plants - 45878; Viruses - 0; Other Eukaryotes - 646 (source: NCBI BLINK).	YARNRFRKALEMFCMHVSEGIQVYDKYILTTIVSV CADYGLLSFGEQIHGLVHKIGHKFDIVFCSSLVDM YSKCGCLLENQVDFDKTTTRNTILWTAINGNAFH GFGREAVKLFKLMDDIQFNEVTFVAVL.TACSH AGLVDEGRNYFTVMXGIEHFTCMWDLFGRAGYL EEIKNFYENNL-SHLSVWVRAFLSSCRLLHKNYDL GKWVYSEKLLLOLEPL
496	sp_207541z12186	#N/A	#N/A	uncharacterized protein	#N/A
497	sp_208330_izqx	#N/A	#N/A	Poly(nucleotidyl) transferase, ribonuclease H-like superfamily protein; FUNCTIONS IN: nucleic acid binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAINs: Poly(nucleotidyl) transferase, ribonuclease H fold (InterPro:IPR012337); BEST Arabidopsis thaliana protein match is: thioredoxin family protein (TAIR:ATI.G52990.1); Has 414 Blast hits to 414 proteins in 36 species: Archaea - 14; Bacteria - 19; Metazoa - 0; Fungi - 5; Plants - 373; Viruses - 0; Other Eukaryotes - 3 (source: NCBI BLINK).	DGRVVCPTLQAASFTKLVADYIEYSQRVFKPRPN FTRTSPARLSSPFGYKIVNTDANINSEAGVGLG ALLRDDSAPVAAKYKRLKARWSVEQAEVEAVK YGMQLANRLGFTKVLMECDAVTVYQAINNNYK YSPLLL.YYEDIAARMKLSFDTFQCSHXQVQGGCLP PHQVTLR
498	sp_208441z13982	#N/A	#N/A	uncharacterized protein	#N/A
499	sp_208921z15390	#N/A	#N/A	uncharacterized protein	#N/A
500	sp_209531z16277	#N/A	#N/A	uncharacterized protein	#N/A
501	sp_210441z17064	#N/A	#N/A	uncharacterized protein	#N/A
502	sp_211480_shin	#N/A	#N/A	uncharacterized protein	MSRAAPRTPALRTEPSAEHISIMPSTAAPKLGVE FRNFKLLAEAGPSKSKSVSPMSRAAPRTP ALRTEPSAEHISIMPSTAAPK.LGVEFRNFKLLAE AQPSSKSKSVS
503	sp_211851z17798	#N/A	#N/A	uncharacterized protein	#N/A
504	sp_213401z18122	#N/A	#N/A	uncharacterized protein	#N/A
505	sp_213691z18153	#N/A	#N/A	uncharacterized protein	#N/A
506	sp_213691z18154	#N/A	#N/A	uncharacterized protein	#N/A
507	sp_213941z18231	#N/A	#N/A	uncharacterized protein	#N/A
508	sp_213980_ack	#N/A	#N/A	uncharacterized protein	MTHKDATTINLVQWYANDNFAQRKSKSICNSTLS

【 1 1 6 8】

10

20

30

40

【表 1 4 - 1 6 9】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
509	sp_214250_arhi	#N/A	#N/A	O-Glycosyl hydrolases family 17 protein; FUNCTIONS IN: cation binding, hydrolase activity, hydrolyzing O-glycosyl compounds, catalytic activity; INVOLVED IN: carbohydrate metabolic process; LOCATED IN: anchored to membrane; EXPRESSED DURING: L mature pollen structures; EXPRESSED DURING: L mature pollen stage, LP_06 six leaves visible, M germinated pollen stage, LP_04 four leaves visible, 4 anthesis; CONTAINS InterPro DOMAIN/s: X8 (InterPro:IPR012946), Glycoside hydrolase, catalytic core (InterPro:IPR017853), Glycoside hydrolase, family 17 (InterPro:IPR000490), Glycoside hydrolase, subgroup, catalytic core (InterPro:IPR013781); BEST Arabidopsis thaliana protein match is: O-Glycosyl hydrolases family 17 protein (TAIR:AT5G18220.1); Has 2719 Blast hits to 2656 proteins in 173 species: Archae - 0; Bacteria - 8; Metazoa - 5; Fungi - 96; Plants - 2598; Viruses - 0; Other Eukaryotes - 12 (source: NCBI BLINK). RNA-dependent RNA polymerase family protein; FUNCTIONS IN: RNA-directed RNA polymerase activity, INVOLVED IN: posttranscriptional gene silencing; LOCATED IN: cellular_component unknown; EXPRESSED IN: 14 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: RNA-dependent RNA polymerase, eukaryotic-type (InterPro:IPR007855); BEST Arabidopsis thaliana protein match is: RNA-dependent RNA polymerase family protein (TAIR:AT2G19910.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	TYADYTPLYGSSCHGLDANENISYAFNYYQTK YQKLEACNFRGLAKLTQNIQSTGNCLTIQIASSG LRTAAPAMFLNVLVLIYLF
510	sp_214270_ditz	#N/A	#N/A	uncharacterized protein	#N/A
511	sp_215551z18525	#N/A	#N/A	uncharacterized protein	#N/A
512	sp_215871z18557	#N/A	#N/A	uncharacterized protein	#N/A

10

20

30

40

【 1 1 6 9】

【表 14 - 170】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)		注釈	aa seq
513	sp_216290_qcaa	0.923333	#N/A	#N/A	FAD-binding Berberine family protein; FUNCTIONS IN: oxidoreductase activity, FAD binding, catalytic activity; LOCATED IN: endomembrane system; EXPRESSED IN: embryo, synergid; EXPRESSED DURING: C glaucular stage; CONTAINS InterPro DOMAIN/s: FAD-binding, type 2 (InterPro:IPR016166), Berberine/berberine-like (InterPro:IPR012951), FAD linked oxidase, N-terminal (InterPro:IPR006094); BEST Arabidopsis thaliana protein match is: FAD-binding Berberine family protein (TAIR:AT1G26420.1); Has 4385 Blast hits to 4167 proteins in 621 species: Archae - 36; Bacteria - 1918; Metazoa - 13; Fungi - 1639; Plants - 666; Viruses - 0; Other Eukaryotes - 113 (source: NCBI BLINK).	ATFLSIWIGLVSPILSVVVTNNNNIHLDAFIQCLE NESHVSDGYPFSEFILTPNNSDFQVHNRFTNLL RYEQPTRKPLLVAPKSELINPAVAVTCARNGL KMKIRSGHDFEGISFTASFPFILLDMFNLNKVDI DIDTEVAWQSGATLGCYIYIAEKSPHAFAPAG CPVSGAGGHFSGGGYGLRKYGLSDNIVDAR VYDVNGRILDRSMGEDFFWAILGGGGSGFVIL SWKLNLRVVPKVTNVNVDVKNKPGVNHMLYK WQNIHKLPNDLLIRVESSYRTPENYENPTIYA SFSMFLGDRLENLTKMKNKSPFELGLTRKDCCKEM SWIDSTMFWHRIPDGTREILLNYDIKHSAYRKIK GDFVKKPIPEAALPQIWEKMLQGTQVYMEWTPY GGRMDEIDESAIPPYHRKGNLFMIFYDLLWIEDG VDEFYLNKARELHQFMGQFVSDSPRESLVNYVD LDIGSSWEKYFKANFPKLVMMVKTVDPTDFFSHE QSIPTLTKLENKDFKRGYSNLLRLPMS
514	sp_000180_aycw	0.819452	-0.65531	-0.855357	flavodoxin family protein / radical SAM domain-containing protein; FUNCTIONS IN: iron-sulfur cluster binding, oxidoreductase activity, FMN binding, catalytic activity; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 7 growth stages; CONTAINS InterPro DOMAIN/s: Flavodoxin/nitric oxide synthase (InterPro:IPR008254), Radical SAM (InterPro:IPR007197), Wycosine base formation (InterPro:IPR013917); BEST Arabidopsis thaliana protein match is: P450 reductase 2 (TAIR:AT4G30210.2); Has 3108 Blast hits to 3089 proteins in 883 species: Archae - 191; Bacteria - 976; Metazoa - 668; Fungi - 593; Plants - 205; Viruses - 0; Other Eukaryotes - 475 (source: NCBI BLINK).	METTPSLACANKCVFCWRHHTNPVGSWQWIK MDDPVEIVNTALDLHTKMIKMGKVPGLVEERLS EGLSPRHCSALSLVGEPIPEINTLVDELHRRRIS TFLVTNAGFPEKIKLKPITQLYVSDAAATRDSLK AYDRPLFGDFWERFLESLSLNEKHQRTVYRLTL VKGWNAEDVDGYFKLVSLGQDFIEIKGVTYCGS TATSKLTMENVPWHADVKAFSEAVLRSEGEYE VACEHEHSCCVLAKVKDFKINGQWFTWIDYEKF HDLVASGEPFKSSDYMAFTPPWAVYGAVEGGF DPOQARFRKERHNSKLSQSK
515	sp_000360_ejow	#N/A	-0.861523	#N/A	E3 Ubiquitin ligase family protein; FUNCTIONS IN: zinc ion binding; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, RING-type (InterPro:IPR001841), E3 Ubiquitin ligase (InterPro:IPR022170); BEST Arabidopsis thaliana protein match is: E3 Ubiquitin ligase family protein (TAIR:AT1G59560.1); Has 1960 Blast hits to 1959 proteins in 225 species: Archae - 0; Bacteria - 0; Metazoa - 1182; Fungi - 34; Plants - 350; Viruses - 140; Other Eukaryotes - 254 (source: NCBI BLINK).	MSAGDQAAVAIAQQLLAGDGACLGALGYAV RSLTKFAATSSALRKIGDAPSVQVSDLRFLIKEDS GGDGEGERNSEVLVVRGLVASKSAVKGWSMKS LWPDLVSHESGDRGVVLFQSKQGIYNEWRGF FGWYPLDRAVLKAWREOQTTSTRTPFVLVDG RQQTQSEYVVDLDSNHPHLLPTIYQHMQPIEA SPYTFLOALFGHEYPVGLDEEKILPLGKEVTA VSPYTFLOALFGHEYPVGLDEEKILPLGKEVTA GICSSKSGALEIKSCRDLPCLSELKDKDILVQLS FRTKILLWSGIFIGSVSIGILGYAIVRNWNKWKQW RYQRRAEQNHPPAAAEEDNSPSESEVEDVED VGDVPDGLQCVICLMRRRRAFAVPCGHLVCCR R CAFSVERDLAPKCPVCRQSGIRGVSRIYDA

10

20

30

40

【 1 1 7 0 】

【表 1 4 - 1 7 1】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)		
2	共発現遺伝子			注釈	aa seq
516	sp_000640_ugxj	#N/A	-0.3223388	<p>Tetrapeptide repeat (TPR)-like superfamily protein; CONTAINS InterPro DOMAINs: Pentapeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: Tetrapeptide repeat (TPR)-like superfamily protein (TAIR:AT5G39710.1); Has 28218 Blast hits to 10356 proteins in 245 species: Archae - 1; Bacteria - 15; Metazoa - 229; Fungi - 251; Plants - 27094; Viruses - 0; Other Eukaryotes - 628 (source: NCBI BLink).</p>	<p>MSSTFRRALFPRATTTTFFSSISSAHAHSPSP STSSVATAVSILTHHRKSRWSALLSLFPHGFTF PQTSQILLGRINPHLSLRFLLTTRHSLSSLHSY ATHILARGRHKTRAKTLIKSAIREFDLLGADDSQ VFSGKCKVFRALAEYRVCDSPAFFVFDLILSLI EMNKLDPCDIVRMLHSRAIYINASTCIDLIVSASKI RGCFYDYVREYVFGYDSDGFVDHDKKVNCKLV PNVHIFNELMLSFRHNGLLDMVEEVVIEVMKSN EPNSYSYSLMAAYCEDNDMEKAEKVVWEELRGK GLNPDIVGNTLIGGYCRIGEVKRGEEELYKEMGL RGLNGSVTYEHLIGGYCLCGDYDSAMLLYNDM RRKGFNPNSLIVDSLVRILCGKRSVSEGLEFFRVA MKNADFVYSESSYQLLIEGLCDEGMTKEALTQA QMAGKGFKPNQIYDAFIDGYMKIGNTDEAENLR KEMMDK</p>
517	sp_000790_mejd	#N/A	#N/A	<p>beta-amylase 3 (BMY3); FUNCTIONS IN: cation binding, beta-amylase activity, catalytic activity; INVOLVED IN: cellulose biosynthetic process, carbohydrate metabolic process, polysaccharide catabolic process; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Glycoside hydrolase, family 14 (InterPro:IPR01554), Glycoside hydrolase, catalytic core (InterPro:IPR017853), Glycoside hydrolase, subgroup, catalytic core (InterPro:IPR013781); BEST Arabidopsis thaliana protein match is: chloroplast beta-amylase (TAIR:AT4G17090.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).</p>	<p>MLIPKRVSFENSHRVGFTLKAQQDAVAIEEK ANSILLTSCSVGGVRLVGLPLDVTQYNTVNH ERAMTAGLALKLLGVGVELPWWVGIVKDNM GIYDWSGYLAIEMVQNLGLKLRISCFCHASQEP KISLPDWSRVSGEADPNIFMDRSGRQYKDCLS LAVDDLPIVFDSTSPQVQNFLESNFNAIFSPMGS TITGITYGLPGDGLRYPFHDATTTALNNQLGV GEFQCYDKYMLSHLKYAETGKPMWGLGGPH DVPKDESPNANFFKDDGGSWETPYGDFFLS WYSHQLVEHGNRILSLASTVFEIPTTISGKLPV HSWYKTRSPCELTAGFYNTVNRDGYEVEMIF AKNSCRILLPGLDLSDAHQQNGSLLVQIKDSCSTKQ GVGMVGENSSYCRVNNKFEIKSNLSEEKAVS SFTYQRMGADFFSPKHFPIFTAFVRAVNLQDL DVDDMPDITTTTTRAVETIHRSEALKLQTA</p>

10

20

30

40

【表 1 4 - 1 7 2】

A	B	C	D	E	F
1	分析に用いたペイトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
518	sp_000880_wxck	-0.873383	#N/A	P-loop containing nucleoside triphosphate hydrolases superfamily protein; FUNCTIONS IN: DNA helicase activity, nucleoside-triphosphatase activity, nucleotide binding, ATP binding; LOCATED IN: nucleolus, chloroplast; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: TIP49, C-terminal (InterPro:IPR010339), ATPase, AAA+ type, core (InterPro:IPR003593); BEST Arabidopsis thaliana protein match is: P-loop containing nucleoside triphosphate hydrolases superfamily protein (TAIR:AT3G49830.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink). Ribosomal RNA: adenine dimethylase rrmjy protein; FUNCTIONS IN: rRNA methyltransferase activity, rRNA (adenine-N6,-)-dimethyltransferase activity, rRNA (adenine) methyltransferase activity; INVOLVED IN: rRNA modification, rRNA processing; LOCATED IN: cellular component unknown; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ribosomal RNA adenine methylase transferase, N-terminal (InterPro:IPR020598), Ribosomal RNA adenine dimethylase (InterPro:IPR011530), Ribosomal RNA adenine methylase transferase, conserved site (InterPro:IPR020596), Ribosomal RNA adenine methylase transferase (InterPro:IPR001737); BEST Arabidopsis thaliana protein match is: Ribosomal RNA adenine dimethylase family protein (TAIR:AT5G66360.2); Has 9871 Blast hits to 9860 proteins in 3038 species: Archae - 214; Bacteria - 6519; Metazoa - 271; Fungi - 146; Plants - 120; Viruses - 0; Other Eukaryotes - 2601 (source: NCBI BLink).	MAGKIRKEKQKGSNHFQGGIPFHKSKGQHLKN PMLVESIVQKAGIKPTDIILEIGPGTGNLTKKLEA GKKVIAVELDPRMVLKRRRFQGGTFPSTRLEVIQ GDVLLKCELPFFDICVANIPYQISSPLVFKLLDINQK QNFRCAIMFQREFAMRLVAQPQDKLYCRLAVNT QLLARVFKLLKVGKNNFRPPKVDSSVVRIEPRR PLPIPVGKLEWNGLLSLCFSRKNTKLSIFGQK KYLEMLEKNYRTLESGLQKDNDSIIDMSILNED VSMDDVWNGKADDEDDQMEVEEEDGEDDADSD NNEMQINDSGFKGKVGVLKQYDYADKRSSKLTQ NEFHLLSVFNQAGIHFTSTK MSSLPNSFSTVEKEKEKEKEKEKEKEKEKE KEEKEKSKSKSSNRLTKEEQFQAKNQTFPPL ERMEINRGFINTTSLVEGPPRSISLVIGFSINIDIWF RGDWGDKLDNDALCQWLQSWDAMDPEVARIT RGPFWKFNFSDFDILSDDEAHYKHLLAGYDQPL SAHFEGPDGSCITMYAILLHVAQWVQVEVIDV NVVPTAVTYTVHGNIFASYGHFESDNDSDIKDYC KTTLLKTGDLESVKVERWSKIKLSRSIVDTLGGM MTELLFG
519	sp_000960_ucxn	-0.888376	#N/A	uncharacterized protein	
520	sp_002390_ttfp	-0.881844	-0.853079		

10

20

30

40

【 1 1 7 2】

【表 1 4 - 1 7 3】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa_seq
521 sp_003590_tmqe	#N/A	#N/A	#N/A	-0.834473	Protein of unknown function (DJF1664); CONTAINS InterPro DOMAINs: Protein of unknown function DJF1664 (InterPro:IPR012458); BEST Arabidopsis thaliana protein match is: Protein of unknown function (DJF1664) (TAIR:AT2G02730.2); Has 199 Blast hits to 190 proteins in 29 species: Archae - 0; Bacteria - 14; Metazoa - 2; Fungi - 2; Plants - 161; Viruses - 0; Other Eukaryotes - 20 (source: NCBI/BLINK).	MAMQAGLGVSRLLLLVGGYTGIVLVKNGKLSDI IGELQGLVKLENSDDSANADSEIAKQVRKLAQE VRNIASSRQITVYVNGSGGTFSSLLPAAASLQAL AYGYMWWKGLSFDLMYYTKRNMATAVANLTA NLDNITDRLSKVKLHLTQRIENLDGKVDQREMS KQIKDEVCKTRSEIDNVNSQLYALKDVTNSLQKD MESEFAKQMLTLSGVKYLCDYVDGRTSKAPKIS QHLQEQVKHGKIPRAQLSSSNALGLQGLKDAESV NVTRSIKEIE
522 sp_003620_widu	#N/A	#N/A	-0.902736	-0.859728	Cystathionine beta-synthase (CBS) protein; CONTAINS InterPro DOMAINs: Cystathionine beta-synthase, core (InterPro:IPR000644); BEST Arabidopsis thaliana protein match is: SNF1-related protein kinase regulatory subunit gamma 1 (TAIR:AT3G48530.1); Has 879 Blast hits to 799 proteins in 237 species: Archae - 117; Bacteria - 170; Metazoa - 187; Fungi - 63; Plants - 180; Viruses - 0; Other Eukaryotes - 162 (source: NCBI/BLINK).	MDTATTVGYDAYFETVQIQSRKLLPISLQETL TSA FASIPVSSFPKVPGGKVEISAETSVSDTYKVLSE CNILSAPVTDMEKANSNDWKERYIGIIDYSAILW VLOGAELAAVALSATSAAGYGAGAGITLGAVA LGATGPVAVAGLAAAAGVAAVAGGLAVDKGMG KDAPTAASELGGDDFYKVIQEEPPKSTKYREVVK SFRSTPFLPVESTESNMLTVLLLKRYMRNVPIE PGKAFMKNFTQSAVVRGLEGCGLDWFDCIAA HPISDLGLPMSRDOVYSIDSESELILEAFKMKDH HIGGLPVVEGPKRKVGNLSIRDIRYLLLOPNLFSN YRHLTVNDFMKTVTSMCQEAQRIMPPITCKSDTT LGDVHLSLASKSVHRIYVVAEEDEVIGVITLDRVIS CFIYEPPNYDEYLGYSAEVLLDR

10

20

30

40

【表 1 4 - 1 7 4】

1	A	B	C	D	E	F
		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MTSEGGSAQVAAVATADPFNSIEYRAEVNQGRAAF QYPSAAGATPVAVPIHGGDNSSLGNGTVQNV SYYNHQPATVQDATNVPSTSTSTASGTDNA SHMYTSTYSSSTNPGYASSEYTNYSYGOQ QSTEPSGYQQQSTQPSGYQSQSTQPSGYQQQST TOPSGYQQQSTQPSGYQQQSTQPSGYQQQST QSSGYQQQSTQPSGYQQQSTQPSGYQQQSTQ PSGYQQQSTQPSGYQQQSSOPYTOATGYQNS GSPYQISSFQNTGTYGTTYPSTYYPNGDYQ TTGAYNGSSYSOTNLWSQGNANYPHQYESY SNDTSSAYTSSNVAATTDYQQYQQWADYNS QTEVTCAPGTEENTASATAALSCPIGVSSGYAN PVSQPATYTLAWKPEIDPSALSSSHQSTSTAGT YNDYWKHAQAYQNHASNLQSNFQPLDVSH SHPVQDQQQLATQPLNSLYTEVQQTYPQP LOPHTPADTLRVSNLQIPTNPRIASNLALAMPKI KDTVSGYTAKPAYSVSLPKPTNKVLSSSASDS NPKFGIFPAALRGYVERALARCKDEKQFAASQE VMKEIITKATADGTLYSRDWDEPLFLPDPDTSTLN IEQSAAPVLSLAFKRSRFRKRSRWEPLPEDKP IEKLTSTYQDSARPVWANYNGRDHKSVAKSER KDATFSRFSTVDQSTSRSAHFPNKRQRFNSDL SATENGDASSDSDKEQSLTAHYSGAILTANSPEE KKFRSRRFRFEKNGGORAENRFRPKNAGVGN YLYSRRATALVLSKTFEDGGGRAVEDIDWDALTI KGTGCOEIEKRYLRLTSDPDSVTRPEEVLEKALH MVGSSQKNLYKCDOLKSRQDLTVQIRNELTV KYVETHGRLELLEFGDLPEFNQSOQLKSLYGEGL AGCHMEFSAYNLLSISLNNRKLDSLSMSRLSD ETKHEAVKHALSVRSVAVTSGNYVLEFFRYLKYKSD NLNLTCLMDLCVEKMRFEAVKCMSRTRYPYVSY VAQVLFASAPEGEGGDEGDVGLLEECMEWLK
523	sp_003800_wscm	-0.307295	#N/A	#N/A	SAC3/GANP/Nim1/mis3/eIF-3 p25 family; CONTAINS InterPro DOMAINs: SAC3/GANP/Nim1/mis3/eIF-3 p25 (InterPro:IPR005062); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	

10

20

30

40

【 1 1 7 4】

【表 1 4 - 1 7 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたペアとの相関係数 sp_107670_ptqx (SOAP2-like)		注釈	aa seq
524	sp_004140_rkiz #N/A	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	CAB UNDEREXPRESSED 1 (CUE1); FUNCTIONS IN: antiporter activity; INVOLVED IN: transport; LOCATED IN: mitochondrion, chloroplast, plastid, plastid inner membrane, chloroplast envelope; EXPRESSED IN: callus; CONTAINS InterPro DOMAINs: Protein of unknown function DJF6, transmembrane (InterPro:IPR000620), Tpt phosphate/phosphoenolpyruvate translocator (InterPro:IPR004696), Protein of unknown function DJF250 (InterPro:IPR004853); BEST Arabidopsis thaliana protein match is: phosphoenolpyruvate (pep)/phosphate translocator 2 (TAIR:AT3G01550.1); Has 2791 Blast hits to 2789 proteins in 409 species: Archae - 9; Bacteria - 295; Metazoa - 571; Fungi - 421; Plants - 1194; Viruses - 0; Other Eukaryotes - 301 (source: NCBI BLINK).	MQSSAISPSLPHLKLQNPFSQHRRLPPLNSFS VRFSLSSAITSVPRLTIVRSSSEEDSFPSPRRSS RFDGLFSSPSLNFSGKRSSQDDFAVKATSSSLPE SADSDEAPQTMSTQLLEIGILFVLWVINYFYNIYNIK QVEFDF
525	sp_004220_jsph	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	unknown protein; Has 1524 Blast hits to 1298 proteins in 225 species: Archae - 9; Bacteria - 84; Metazoa - 474; Fungi - 184; Plants - 98; Viruses - 17; Other Eukaryotes - 658 (source: NCBI BLINK).	MGKKFKDKKNSATFQLLSRDSSDPVYIEEPEND RVFVRVNNNEVSVNGIDEEDHVESADDYGGSEV FRGPTLTDVRREILELGFDDGGYNYLTLREIR NTGGGSAYFQNPKAELHQLPHDKVAYDASRVRI KEASEDTSPTLYNVAEKTVA/RVQKAVDPEIAA LLEDNESDFGSDIDLEEDFWQANHSEETGDVY EDTTLVWVNFQEEVINGSDTLSCDQITEVASNS TVVEWQSVNEKSRTORLLDQDFMLESQEYGS SDDYDNYVAEQEISLAERLKNVLDREVDDLEF NENYKAPAEILRRCAEYAEKYENEDGNDKVVIV EESSESEKWDCEFTVSTYSLNDNHPGRIGVPG RIRQKILTEITISGVLKSGSKIIISLGGKARLPVDYL PHGKKKESKETTQVVAKTEQPKRKPFGQETKEEK KERKAAVKEERREARRLKKDLKELYRDEAKLAQ KYVAAITGPSSIHLLG

10

20

30

40

【表 1 4 - 1 7 6】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
				glucose-6-phosphate dehydrogenase 2 (G6PD2); CONTAINS InterPro DOMAINs: Glucose-6-phosphate dehydrogenase, C-terminal (InterPro:IPR022675), Glucose-6-phosphate dehydrogenase, active site (InterPro:IPR019796), NAD(P)-binding domain (InterPro:IPR016040), Glucose-6-phosphate dehydrogenase (InterPro:IPR001282), Glucose-6-phosphate dehydrogenase, NAD-binding (InterPro:IPR022674); BEST Arabidopsis thaliana protein match is: glucose-6-phosphate dehydrogenase 3 (TAIR:AT1G24280.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MATFSPSHCASSSSFFYSSSSKSRNGRLGMV KISSISKQSAPIKISYSKQQQSGRNPKSNVAFMQD GAVATKIIPVEGENVSSVETLKDGILSATSSKESN DLSGFDVRKDESTVITVVGASGDGLAKKIFPFLF ALYEDCLPKHFTIFGYARSKMTDDELRTMWSKT LTCRIDKRENCGEKMEEFKFCFFYQSGQYDSRE NFVELDKLQRQREAGRIANRLFYLSIPPNIIDAVR CASLSASSANGWTRVIVEKPFGRDSESSAALTN GLKQYLKEDQIFRIDHYLGKELVENLSVLRFSNLIF EPLWSRQYRINVLIFSEDFGTEGRGGYFDHYGI IRDIMQNHLLQILALFAMETPVSLDAEDIRNEKVK VLRSMRPLQLDDVVGQYKQSQTRGGVTPYGYTE DKTVPKDSLPTFAAAAFLDNARWIDGVPELMDKA GKALHNRSAEIRVQFRHVPGNLYARNFGSNLDQ SMNELVIRVQFDEAIYLNKKNKVPGLGMRDLRSTL NLHYAARYAKEIPDAYERLLLDAVEGERLFRIRD ELDAAWALFTPLLKELOKKNVPEYYPYGSRGCP VGAHYLAARYNVRWGDGSGSEP
526	sp_004890_dmip	#N/A	#N/A	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 12 growth stages; Has 16 Blast hits to 16 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 16; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).	MEFONLIKDKKFWFASVLIWAAGLOGHMMW/L QRQDAFKQKFGDLNESDDQKNLN
527	sp_005030_fqaq	#N/A	#N/A	unknown protein; Has 20 Blast hits to 20 proteins in 10 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 20; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).	MEEDWKEGVKRGCGERLQNALYDCHRRMPAG PARNSGCRHLNRAFTDCLVAVICPSEIEAVRSLC SSGGTAMKFSQCQCEARLSLVC--SQHQDQDHP
528	sp_005450_plpz	#N/A	#N/A		

10

20

30

40

【表 1 4 - 1 7 7】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	<p>pescadillo-related; FUNCTIONS IN: transcription coactivator activity; INVOLVED IN: cell proliferation; LOCATED IN: nucleolus, intracellular; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Pescadillo, N-terminal (InterPro:IPR010613), BRCT (InterPro:IPR001357); Has 503 Blast hits to 494 proteins in 230 species: Archae - 0; Bacteria - 0; Metazoa - 173; Fungi - 168; Plants - 48; Viruses - 0; Other Eukaryotes - 114 (source: NCBI BLINK).</p>	<p>MPKHYPAGKKKEGNAKYVYTRSQAVKGLQVTL PLFRKMCWKGIFPRDPKPKKIKGNHQTYYHTKDIS FLHHDPLVDIHRDIRAQDKKIKKAYSKNKPRAVR LEEHPYNDRLRIERYPRFDALGDLDDCLTMV HLFAALPTLQNEKVEVHIHSRRRLSHEWQAYIS RTHKLRKTFISVIGYQADVGDGQKITWIAPHSIQ QVLTEDVDNFNMLTFLEYETLLAFVNFKLYHSIN VKYPPMLDPRLEALADLYALSRYIASRSKEQDN HELASQVDESQTLAQLOHQQLAPSEPGALMHLIE EAETVDEADEETKECRSLFKDTKFFLGREVPREP LLFVITAFGGTISWEGDGPFEESDPSITHQIVDR PTGNHMLFSRAYIQQWYVFSVNAIRILLPTEDYV VGRVPPHLSPFVDNEAEGYIPDYAQIKQLQDA ARIMPVEVGNHQDNTQNVLDGVDADRIEANKA AEMIKKMSAQKQFHDELKMEVEGVSFATGTIK NVKVTDPDESDEKQISEHPDMSEMTMSRRKRG LYESIRINKERKAKYVDVIESRKRNIKESQK</p>
529	sp_005830_ihwq	#N/A	#N/A	<p>ARM repeat superfamily protein; FUNCTIONS IN: protein transport activity, binding; INVOLVED IN: intracellular protein transport, protein import into nucleus, docking; LOCATED IN: nucleus, chloroplast, nuclear pore, cytoplasm; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Importin-beta, N-terminal (InterPro:IPR001494), Armadillo-like helical (InterPro:IPR011989), Armadillo-type fold (InterPro:IPR016024); BEST Arabidopsis thaliana protein match is: ARM repeat superfamily protein (TAIR:AT3G08947.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).</p>	<p>MVMEITQLLSAQSYDGGQVKRHAEDSLKQFQDQ NLPNFYLSLAGEASEEKPADSRKLAGLILKNELD AKEQHRKFFELVQKWVSLDAGLKNQIKTFLQSLT SPVHDARSTASQVAVKAGIELPKQEWPELIGSL LGNHQLPHVHQATLETGLYLCSEVSPDWDQD HYNKILTAVVGGMNETNNEVRLAATHALYNAL GFAHANFNNDMERDYIMRVVCEATMSPDLKIRQ AAFECLVAISSTYYEKLAPYIQDIFSITAKAVREDE EPVALQAVEFWSSICDEEIDILEEYGGDFGDSFI PCFYFIKQALPALVPMLETLKQDEQDQDEGA WNIAMAGTCLGLVARTVGDIVPLVMPFIEENIT KPDWRQREAAATYAFSGILEGSPSPDKLTAINVVAL NFMLTALTKDPNSHVKDTTAWTLGRIFEFLHGSAL METPIIPANCQQIVTVLQAMKDDTPNVAEKACG ALYFLAQGGYEDVASSSPLTPYFQEIQCILLTVTH REDAGEARLRTAAAYETLNEVVRCSSTDEAPLVLQ LVPVIMMELHKTLESNLSPPDGKQKQGLGCG CLQVIQKLGSSSEPTRAVLQYADQIMQLFLGVFA CFRSATVHEEAMLGIGALAYTGPDFAKYMNFEYK YLEMGLQNFGEYQVCAVTVGVVGDLCRALEDKIL PYCDGIMTQLLLSSNQLHRSVKPFIFSCFQGDIA LAIGENFEKYLVMYMPMLQSAADLSAHTAGADD EILEYTNLLRINGILEAYSIGLQGFKNKTKTQLLPH ARHLSFDSIYQEKDMDDDAVMKTAVGLVGLDADT LASSGLSIQQFPTCRELLDECLSSDDHHSIKQSAE WAKLAVSRVSI</p>
530	sp_006460_eyhq	#N/A	#N/A	<p>-0.875%026</p>	

10

20

30

40

【表 1 4 - 1 7 8】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
531	sp_006590_fjix	#N/A	-0.89387	#N/A	uncharacterized protein	EYLTFSIKFLKWGDFTNFKYREELKFKVHMSTFFS ASRAATLGRAARRRGERARDADACTKVAGRAAPKE RE
532	sp_006900_ahhc	#N/A	-0.807164	-0.460338	MINICHROMOSOME MAINTENANCE 3 (MCM3); FUNCTIONS IN: nucleoside-triphosphatase activity, DNA-dependent ATPase activity, DNA binding, nucleoside binding, ATP binding; INVOLVED IN: cell proliferation, DNA-dependent DNA replication initiation, DNA unwinding involved in replication; LOCATED IN: nuclear chromatin; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAINS: Nucleic acid- binding, OB-fold-like (InterPro:IPR016027), Nucleic acid-binding, OB-fold (InterPro:IPR012340), ATPase, AAA+ type, core (InterPro:IPR003593), DNA- dependent ATPase MCM (InterPro:IPR001208), DNA- dependent ATPase MCM, conserved site (InterPro:IPR018525), MCM protein 3 (InterPro:IPR008046); BEST Arabidopsis thaliana protein match is: Minichromosome maintenance (MCM2/3/5) family protein (TAIR:AT4G02060.2); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MWIKKAVILLMLGGTEKNLKNKNGTHLRGDINMMMV GDPFVAKSOLLREIMNIAALAISTTSRGSSGVGLT AVVTSDEMGKEG

10

20

30

40

【表 1 4 - 1 7 9】

I	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数		sp_107670_ptqx		
2	共発現遺伝子	sp_107660_kiqg (SOAP2)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
533	sp_007210_sudc	C_893727	#N/A	#N/A	Heat shock protein DnaJ with tetratricopeptide repeat; FUNCTIONS IN: binding, heat shock protein binding; INVOLVED IN: protein folding; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAIN/s: Tetratricopeptide TPR-1 (InterPro:IPR001440), Tetratricopeptide-like helical (InterPro:IPR011990), Heat shock protein DnaJ, N-terminal (InterPro:IPR001623), Tetratricopeptide repeat-containing (InterPro:IPR013026), Tetratricopeptide repeat (InterPro:IPR019734), Heat shock protein DnaJ, conserved site (InterPro:IPR018253); BEST Arabidopsis thaliana protein match is: Heat shock protein DnaJ with tetratricopeptide repeat (TAIR:AT5G12430.1); Has 24543 Blast hits to 21405 proteins in 3009 species: Archae - 131; Bacteria - 8314; Metazoa - 5854; Fungi - 2610; Plants - 2808; Viruses - 9; Other Eukaryotes - 4817 (source: NCBI BLink).	MPPFYVYVDSRRPQSDSINFPNGSTTHFNDQSS FNTGSSSTSPAVGRGSSSTRAKPRLVKLRKSS KSSLSSEADLDDGGINPFRSVSSINTDASNCITN SSSDSFVSRFRNVGFVGVNSIDNNINNNIKV NDGGKVAEASVFGFRDNLHKQEFSENGTE RDNNCTENAVFRFGATAKDOIGVSNVKCGENG EFELLDELQKIDKIDKERCCTAEDKVDLNNVDLTK GGDRFFVFKTGGVSYDECGSSVPLRSDVDA TN LSNGKGNINESKMGAFYRYVDDSAFVFGSSG NFSKMGNSSSSQSSVPAENDGARIKEDFLSPF FSSLDADAGNAKSEACSSSSERKVEFSSGGLP EGSGVSEFEFLPQFDAPLFSANFTSGLGRKPE SRGLKDKLRKTKGKSKKPTKLPVQVQPLSS EGVTQGEASECYSPMDFSPYEDTTNNDVPLTG DHETYQTKVTTGNLSLDVHTDSEINDSDTIKSDN HRGNEAQFSSDKASRDEENITFTSPTSQNH VSSPKRLLKKYRMPASISSNASRVQSSLDSS PLLSKLOGAGGYFFASASKDQHEAAGVEHDC EFSAAEKACEKWRPARGNEAYEKGELSKAEDFYT WGVNCPSPDASGAAFRPLVRCYSNRAATRMAC GRVREAIQDCLTAVKLDPTFHRAQIRAGNCYLL GEVEEAMQCFKYIGDGNVLCIDRKAVISADGI CKSQKVLQSTSRSAELLKLTSEAAADALOHIAE ALAIINSEKLLKMKQAQALCTLSRHDEAIALCKTL DFAKKNFTAIFGNVEYVDGSSNDRLWRYRILSKC QFHLGNLEVALDLDLEKQENKTEGIDPLSATIREL LKQKSAGNSAFHSGNHEAVEHYTSAIMSSMDA RPFAAICFNRAAANOALGQITDAISDCSLAIALD ENYAKAVRRATLHEMIRDYKHAADLQRLSSFL CKPVSVEKSKQSGGKNGKELRQVKQRLSMEEE AKRDTPLDLYLILGVKKSQDVTSEIKKAYRKAALRH HPDKAAGSRVKSQGDVGEGRHFKWIDIAELVHKDS

10

20

30

40

【表 1 4 - 1 8 0】

1	A	B	C	D	E	F
		分析に用いたバイトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)		注釈	aa seq
534	sp_007320_xjcx	#N/A	#N/A	-0.6526237	<p>eukaryotic translation initiation factor 3C (EIF3C);  FUNCTIONS IN: translational initiation factor activity;  INVOLVED IN: translational initiation; LOCATED IN:  cytosol, nucleus; EXPRESSED IN: 25 plant  structures; EXPRESSED DURING: 15 growth stages;  CONTAINS InterPro DOMAINS: Winged helix-turn-  helix transcription repressor DNA-binding  (InterPro:IPR011991), Proteasome component (PC)  domain (InterPro:IPR000717), Eukaryotic translation  initiation factor 3 subunit 8, N-terminal  (InterPro:IPR008905); BEST Arabidopsis thaliana  protein match is: eukaryotic translation initiation factor  3 subunit C2 (TAIR:AT3G22860.1).</p>	MSSRFWDQGNSDTEEEESDYDDVDEPVADEAG AKAKYGGDRYKEDNGNDSDDSDHHRVPIPGK DKRFQEMTATVEQMKNAIKINDWVSLQESYDKIN KQLEKVMRVAETGKVPNYKALVLEDFLNELLS NKEAKKMWSTSNKALNKMKNLKNKKNQYED MITKRENPESDEKDEGDDSSGTEFTESEGED EDTDKEDDKGEGDDVEDDVENIGGWQTPGKNN RLDILLKNPSEITWDTVDKVKYKEVAGRGKKG GKVELVERLFTLTKIAKTPAQKLEILFSWSAQFDI NPSLSGHMPINWKKCVQNMFMVILDILTQYPNLLV DQSVPEENETQKQPDHKGITIRWGNLVAFLERI DVEFFKSLQCIDPYTREFVERLRDEPLFLVLAQNV QEYLEKSGNLKAAASKIALKPFVEFYYPQEVYDA MRKLAEQSEEGNDEAAEVEFKDAGLGRGPPLEFV VTPELVPRKSTFPERCRALMDSLVSLLYQFGDER TKARAMLCDIYHAIMDEFFIARDLLMSHLQDTI QHMDISTQILFNRAMAQLGLCAFRIGLSEGHSC LSELYSAGRVRELLAQGFSQSRHYHEKTPQEMK ERFQOMPYHIHINLELEAVHLTCAMLLEVPYMA ANTFDSQRIISKTFRRLEASERQTLRGPENY RDHAMAATRGLTQGDQKAYDVITSLDVMWLLK NKESVFEMLKSKIEAALRTYLFYAYVSYSSLSVD QLSKMFEISDPEIHSYKMMVMNEELRASWDQP TQCIIIFHEVHTRLQALAFQLTEKLSFFAENNERA MEAKLGGSLDGLPLRRRDPDYAQVTSTGGG GGGGRWHDNSYNGRSGRSGYGGKSMNFG QFSGGGYGGHSGRGGGGGGGGYGRARGQHNS SAPMVSLLNRGAR

10

20

30

40

【表 1 4 - 1 8 1】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh sp_107660_kiqg (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
535	sp_007710_asekk	-0.382829	-0.381548	-0.3839248	adenine nucleotide transporter 1 (ADNT1); FUNCTIONS IN: binding, ADP transmembrane transporter activity, ATP transmembrane transporter activity, AMP transmembrane transporter activity; INVOLVED IN: in 6 processes; LOCATED IN: mitochondrion, mitochondrial inner membrane, plasma membrane, plastid, membrane; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Mitochondrial carrier protein (InterPro:IPR002067), Mitochondrial substrate carrier (InterPro:IPR001993), Mitochondrial substrate/solute carrier (InterPro:IPR018108); BEST Arabidopsis thaliana protein match is: Mitochondrial substrate carrier family protein (TAIR:AT5G51050.1); Has 27969 Blast hits to 14393 proteins in 468 species: Archae - 0; Bacteria - 2; Metazoa - 12077; Fungi - 8125; Plants - 4770; Viruses - 0; Other Eukaryotes - 3015 (source: NCBI BLink).	MASEDVKQRGEAAAVTTIVLAEAAKAAATEGVKSK PKNALLSISKSLAAGGIAGGLSRTAVAFPLERLKL QVQNPHSIKYNGTIQGLKYIWQTEGFRGLFKGNG TNCARIVPNSAVKFFSYEQASKGILWL YQQQTGN ENAE LTPVLR LGAGACAGIAMSATYPMDMVVRGR LTVQTAQSPQQYRGYHAYHALTTVLRREGARALYKG WLPVSGVWPVYVGLNFAVAYESLKDWLKEKAFGL VQDNELSVTTRLGCGAVAGTIGQTVA YPLD VVR RRMQMVGWKDASTVITGDGQKIALEYTG MIDAF FKTVRYEGFRALYKGLV P NSVKVVP SIAFV TYE VVKDLLGVEMRISD
536	sp_008030_ekix	#N/A	-0.387438	#N/A	basic leucine zipper 42 (bZIP42); FUNCTIONS IN: DNA binding, protein homodimerization activity, protein heterodimerization activity, sequence-specific DNA binding transcription factor activity; INVOLVED IN: regulation of: transcription, DNA-dependent; EXPRESSED IN: inflorescence meristem, flower; EXPRESSED DURING: petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Basic-leucine zipper (bZIP) transcription factor (InterPro:IPR004827), bZIP transcription factor, bZIP-1 (InterPro:IPR011616); BEST Arabidopsis thaliana protein match is: basic leucine zipper 43 (TAIR:AT5G38800.1); Has 1669 Blast hits to 1669 proteins in 135 species: Archae - 0; Bacteria - 2; Metazoa - 133; Fungi - 19; Plants - 1463; Viruses - 0; Other Eukaryotes - 52 (source: NCBI BLink).	MSSNSTSESEEMHHQQGKMMIDRRQRMMIS NRESARRSRMRQRHLDLWSQVRLRNENNML IEKLNHVSESHETILQENVR LKQEASDLRQLNGM QLATTFNGLRDLEDIPCTTAHLRAESNSTHSITTS STDLLM

10

20

30

40

【 1 1 8 1】

【表 1 4 - 1 8 2】

I	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)		注釈	aa seq
537	sp_008300_wwdn	-0.333942	-0.859241	#N/A	Ring/U-Box superfamily protein; INVOLVED IN: biological_process ion binding; LOCATED IN: chloroplast; CONTAINS unknown; DOMAIN/s: Zinc finger, RING-type InterPro:IPR001841; BEST Arabidopsis thaliana (InterPro:IPR001841); BEST Arabidopsis thaliana protein match is: RING/U-box superfamily protein (TAIR:AT5G04460.3); Has 608 Blast hits to 493 proteins in 135 species: Archae - 2; Bacteria - 207; Metazoa - 113; Fungi - 27; Plants - 119; Viruses - 0; Other Eukaryotes - 140 (source: NCBI BLINK).	MAIAGLHNVPVISSILRESQSSARIGNPGSISTR ASSIRQMWLELEDCVNRVRDRGRRLRQQERN DGLSNCGVGEGRGLEDSESENGSVINSESLT GPLNEOYDRQSLASEQSDFG.LVERARVRFIFQ EWKNTGVANQATNVPHRNRGVRSQWLGETER ERVRVREWIQMTSQPREDDQVSEVSPRIEVRD GQVDHINESQVENNRRPIRRLCGRQALLDLAK KEQERRQELQGLVETQVPSRFAHRNRIQS.LLRIR CFQNMRSAESRRPSSSTAESELGLLRQRTVSGL RGGFLRLDNVQVQASDLGSSLENSPHSFRDD QTEANSANEVVDGAHVPNNPANGFLDGADEQN CPANEVLSGAHEQNPNANENLDVCEQRNVEGD ANVNPRESGNTNFQELEVVQVETRLVASEDEE NISETATSGPADGFSDEGYLLMNSQEPSRSGSE EFGEIEGTHELSRHFNTHREEINVLNQLDNEVHE SEEVNDQDFSIQVEDWPEQVTTYNESGWQGNV LSNQWRDGDQAEATNRQHVGDGEEHFDGALESEV GEONDAQEDYWHESGSLTPREWLGMPSTSG VTSGRDLTYFSDDDNVHHMELRELVNRVRVSS LLHSDFRSLDQLIQSVVERQTNVPEDEWELHEM SLPPTTQNDQDLSTGHSESEPIHVEETQRTLP RPPPLPOAHWHGHANW'SRHESRQRPGTEWEII NDLRIDMARLQORLNNMQRMLEASMDMQLLELQ RSVRQEVSAALNRSATPSELFENVLPKDSFKWD CVRKGICCSNTNIDSLLYRCGHMCTCTCKADLL VQGDGKCPMCOAPVIEEGLYRNLKLSKLGQSIS LQMPVLTKSPVFLSL

10

20

30

40

【 1 1 8 2】

【表 1 4 - 1 8 3】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
538	sp_008750_cisf	-0.342787	-0.380373	-0.354787	<p>Nucleolar GTP-binding protein; FUNCTIONS IN: GTP binding, nucleotide binding; LOCATED IN: membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: GTP 1/OBG (InterPro:IPR006073), Nucleolar GTP-binding 1 (InterPro:IPR010674), NOG, C-terminal (InterPro:IPR012973); BEST Arabidopsis thaliana protein match is: Nucleolar GTP-binding protein (TAIR:AT1G10300.1); Has 9379 Blast hits to 9179 proteins in 2035 species: Archae - 370; Bacteria - 5195; Metazoa - 1143; Fungi - 469; Plants - 325; Viruses - 0; Other Eukaryotes - 1877 (source: NCBI BLink).</p>	<pre> DEFPRLDDIHPFYGDLLHVLVYKNDHYKLAGQVN TARNLISKISKDYVKLLKYGDSLRYRKCKLVAALG RMCTVLRISPSLAYLEQVRQHMARLPSIDPNTR TLICGYPNVGVKSSFMNKITRADVDVQPYAFTTKS LFVGHDTYKYLRYQVIDTPGILDRPFEDRNIEMIC SITALAHLRAAILFFDISGSGYTIQAQAALFHSIK SLFMNKPLIIVCNKTDLLAFDNIPEDDKLVMDMK AEALKTIVGOGGGEATNEEEGVLLTMSLTTEEGVM AVKNAACDRLLDQRVELMKSKKINDCLNRFHVA MPKPRDQKERPAICIQGVLEAKAKAEKAEKAEKE HRITERQLEDENGGAGYVSASLKKHYLANDEWK EDIMPEFLDGHNVYDFVDPDILHRLEELEREGL RLADEAEEDFEMDGEELTVEEKEALAEIRRKKSIL IQEHLRLLKSTAGSRVAVPRKFDKDKRKFITERMG RQLTKIGIDPAAILRARSRSRQKQRQRSPEVGG EGDVMMDVQPNKIKLRALSRRRSRSLRFP GEVVPGEKFSIQKVKAINLAKKSSKKRNKNAH KGEGRVPIPNLKPKHLYSGKRGIGKTDRR </pre>

10

20

30

40

【表 1 4 - 1 8 4】

	A	B	C	D	E	F	
1							
2	共発現遺伝子	分析に用いたバइटとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	
539	sp_009050_gfng	#N/A	#N/A	-0.864246	ubiquitin-specific protease 8 (UBP8); FUNCTIONS IN: ubiquitin-specific protease activity, ubiquitin thiolesterase activity; INVOLVED IN: ubiquitin-dependent protein catabolic process; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Peptidase C19, ubiquitin carboxyl-terminal hydrolase 2, conserved site (InterPro:IPR018200), Peptidase C19, ubiquitin carboxyl-terminal hydrolase 2 (InterPro:IPR001394); BEST Arabidopsis thaliana protein match is: ubiquitin-specific protease 5 (TAIR:AT2G40930.1); Has 10748 Blast hits to 7392 proteins in 262 species: Archaea - 0; Bacteria - 6; Metazoa - 5718; Fungi - 1753; Plants - 1297; Viruses - 10; Other Eukaryotes - 1964 (source: NCBI BLINK).		MPSSSSSTLTQQRFRKLLYAFFSLJFLLSRL FDSLJFSLSSAMDNLFPDDFLSVGNDAITTSNN DDGDDDDGLLYPQSRRRRLHSSPPSSSSTS TLY_RSDRLFVLYCYKWKDAREGCLGSKGEIKG DRYTTAELSDDENKRESDDDEFVNLKREVDD DDNGGDDKGGKGFSGREYALITESMWFGLKW HSDYCTRIKDMGCSLAAEDNEQDYFPIQIRLYYM QQTDSL VYKISLMDNPIQLYKSAQYDFVQSNLVC IWFDFSGMTQFFERGGALFPNGCPGQAPAEVLL ELQVYGFSGFNAGEKTEIAMQSSSSGSDSCG SCSIKMDGTVQSSSSPFLSLQSCTHRYGETGVL GLTGLNLGNTCFMNSAIQCLAHTPKLIDYFLGDF KKEINYENPLGMNGELALSFGDLRKLWAPGRM PVAPRLFKSTLENFAPQFCGYNQHQHDSQEFALL DGLHEDLNRVKHKPYQLNDTEGRPDVKVADEH WEIQARNDLSIIDL YHGQFRSTLVCPICKRVSVT FDFPMYLSLPLPSTTMTMTVTLLSTDGSMPPSP LTVTVPKSGRCKDLIQALGTAVLQDDETLFVAE VYNHCLHSLEVPDSLALIRDCDYLVAYRLPKES ESSLLVYFMHQRCERCFQFGRMRRRSWSPFGIPL VARIPNISDGSQIHDTKKLLRPLNMDDEDYDDI ENSDSELEFPNSDINSWSSDDEMGDDLHVPSD FQFHITDERGMLRGRRIKVGKVLDSVASNKVYV LVSWSKPYLRRYDTCRLLPEVFKHVFAAMKP QESVSLYKCLEGFLKEEPLGPEDMWMYCPSCKEH RQASKKLDLWRLPELVVHLKRFYSRFLKNKLE TYVDFPLEELDLTGFLAHESGPRLNRYRLYALSN HYGSMGGGHYATVYVYHGKGWYEFNDSFVELVR EEDVKTSAAAYVLYRRISDT
540	sp_009140_ckki	#N/A	-0.887473	-0.937478	Polynucleotidyl transferase, ribonuclease H-like superfamily protein; FUNCTIONS IN: exonuclease activity; LOCATED IN: intracellular, chloroplast; EXPRESSED IN: 14 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: Exonuclease, RNase T/DNA polymerase III (InterPro:IPR013520); Has 576 Blast hits to 576 proteins in 123 species: Archaea - 0; Bacteria - 44; Metazoa - 336; Fungi - 16; Plants - 80; Viruses - 0; Other Eukaryotes - 100 (source: NCBI BLINK).	MOKGLEASLGCVQSKSYLYGPTYPRFSFGQOT MELKYTNGEKOANVDSNNGAVVTHDKLAVQSEY CGRPVRQESYLLWSALTPNSQKDVHNSQSHLN DGCYFANMGNFYQTTHWLPQVPMQMWPOQSY PDYVFDQFYVVIDFEATCDKERPNPPOEIEFF SVLNSRTGQLEGCFQMYRPTHNLSDFCCKS LTGIQSQVDKGYLLSDALLMHDKWLDEKGIKNT NFAVWSDWDCRVMLESECFRKRIRKPLYFNR WINLPIFNKIFGSGRRNLKQAVELAGLTWEGRA HCGLLDDAKNTALLSHLMQLGKLSITNSLKWQT ADHP SAVQVQQTVTITTFESNLTYQPQNYKTAPS VTAIFQPANLLKEPYVYCYGLTSTKRVNQNPG PKHGCVFFGCGNWTNSNGPCCQCFKFWWSAS DSGKPY	

【 1 1 8 4】

10

20

30

40

【表 1 4 - 1 8 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	BED zinc finger, hAT family dimerisation domain; CONTAINS InterPro DOMAINs: HAT dimerisation (InterPro:IPR008906), Zinc finger, BED-type predicted (InterPro:IPR003656); BEST Arabidopsis thaliana protein match is: BED zinc finger, hAT family dimerisation domain (TAIR:AT1G18560.1); Has 1233 Blast hits to 1128 proteins in 97 species: Archae - 0; Bacteria - 4; Metazoa - 232; Fungi - 219; Plants - 755; Viruses - 0; Other Eukaryotes - 23 (source: NCBI BLINK).	注釈 注釈
541	sp_010180_udyc	-0.839057	-0.881247	LOW EXPRESSION OF OSMOTICALLY RESPONSIVE GENES 4 (LOS4); FUNCTIONS IN: RNA helicase activity, RNA-dependent ATPase activity, ATP-dependent helicase activity, INVOLVED IN: poly(A)+ mRNA export from nucleus, response to cold, response to heat, response to abscisic acid stimulus; LOCATED IN: nuclear envelope, plasma membrane, cytoplasm; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: RNA helicase, DEAD-box type, Q motif (InterPro:IPR014014), DNA/RNA helicase, DEAD/DEAH box type, N-terminal (InterPro:IPR011545), DEAD-like helicase, N-terminal (InterPro:IPR014001), DNA/RNA helicase, C-terminal (InterPro:IPR001650), Helicase, superfamily 1/2, ATP-binding domain (InterPro:IPR014021); BEST Arabidopsis thaliana protein match is: eukaryotic initiation factor 4A-III (TAIR:AT3G19760.1); Has 40925 Blast hits to 40635 proteins in 3033 species: Archae - 702; Bacteria - 21081; Metazoa - 5720; Fungi - 4311; Plants - 2419; Viruses - 28; Other Eukaryotes - 6664 (source: NCBI BLINK).	MADSTEATVEAPGSSSTPSSSVPESSAAASESC SAPEIKPAEYKRSWGDVEDEEFTDQLSIKESKDL DEPDSNIKAVITGDTPTYSAKKFEDLNSEELTK GLYVEMRFRERPSKVQSVTLPMILTPPYRNLIJQA HNGSGKTTCTFVGLMLSRVDPKQQAQALCICPT RELAIQNLVYKIKMGKYTGISAERAVPENLDNPA HTPISKRPVTSQVIIGTGTGTTKIKWMSFRKLSMSS MKILVFDEADHMAEAGGFKDLSLRIMNDIKRSAP SCQVLLFSATSERVYKAFVKSAIQDYNQLFVKKE ELSLQSVKQKRVKPCPELEKQVYVREKILELQCK VGGTIIVRTRNSASYLHKALVDLGLYEVTTIQGAL THDDRDKIVKEFKDGLTKYLIJSTDLVLRGFDQSQ VNLVWNYDLPVKHIDPNKYDYPNKAEPDCEVYLH RVGRAGRFGRKGVFNLCDDIDNKVMGKIENHF GCEVIEVPDYRDNDEFEKALKDAGLL
542	sp_010730_piy	#N/A	-0.823335	Nucleotide-diphospho-sugar transferases superfamily protein; CONTAINS InterPro DOMAINs: Glycosyl transferase, family 2 (InterPro:IPR001173); BEST Arabidopsis thaliana protein match is: Nucleotide-diphospho-sugar transferases superfamily protein (TAIR:AT2G39630.1); Has 22705 Blast hits to 22656 proteins in 2589 species: Archae - 933; Bacteria - 16103; Metazoa - 288; Fungi - 260; Plants - 98; Viruses - 23; Other Eukaryotes - 5000 (source: NCBI BLINK).	MAEKOEREKSKFSIIVPTNYERLNIALVYLKPKHL RDIDFEIVDDGSDPGTQIEIVKQLQQVYGGDRILL AARPKLGLGTAYAHGLKHASGNFVIMDADLSH HPKYLPFSFIKQMETGASIVTGRYVSGGVYHG WNLMRKLTFRGANVLAHTILWPGYSDLTGSRFL YRKSVDLIRSCVSKGYVFOEMIVRASKRGLH IIEVPTFVDRVYVYVSSKLGSGSEIVEYLKGLVYLLFT T
543	sp_011570_rfc	-0.931667	-0.875437		

10

20

30

40

【 1 1 8 5】

【表 1 4 - 1 8 6】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
544	sp_011640_gwnu -0.8833331	#N/A	-0.885722	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G34310.3); Has 84 Blast hits to 78 proteins in 24 species: Archae - 2; Bacteria - 4; Metazoa - 9; Fungi - 8; Plants - 55; Viruses - 0; Other Eukaryotes - 6 (source: NCBI BLink).	MPSGAKRKAAKKAKQKSHINTDGGSHLQIGDEA KSHDERSDGGEGERESSEERWVSKNEEGEQH ATEFSKESSEENYDDMRVNVEDDVRKTKNGKH KSSSSSSSSSEDEWIEKEKSVVSEVQSSSL PEASRENPIVETVYASVEESIPVEESVKQVLSLP QETNTRSESADVENSMPIDDIELGMKQNLDEVLS SADGKDKVSSGMTEFLQEKIEKKVIGVEEGEPAN LSKSLHNTGIDAEASLPPSSGVTSADACCEDNHV KESETHKYTEEQPLVSSAPLPVVRTSLMNCCLLF EVLSSGASR
545	sp_012380_tgua -0.987327	-0.982107	-0.984449	Single-stranded nucleic acid binding R3H protein; FUNCTIONS IN: nucleic acid binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Single- stranded nucleic acid binding R3H (InterPro:IPR001374); BEST Arabidopsis thaliana protein match is: Single-stranded nucleic acid binding R3H protein (TAIR:AT2G40960.1); Has 389 Blast hits to 389 proteins in 86 species: Archae - 0; Bacteria - 0; Metazoa - 196; Fungi - 53; Plants - 134; Viruses - 0; Other Eukaryotes - 6 (source: NCBI BLink).	MDGNAVEDPVAPESWEVADVEESMKRLLSA SSSAKDSATPSSSSNCIAADDLTASGPLNLAFS SGAENAGVGLSDDAINQVDFLREALQMPRE LSILREQDVEQIQDPTQQQLEFQQLPTSYLRLA AHRVAGHYFLQSMVLLDNLSPDGGSGRIIVQKTS TECFRPRILADIPVSLPQEDSGVVKVAIKQRPNK HSSSSVGNSSGSKANSKVSVEERKEENRFRAPAR IFNSNSQNGSPGLKPEIEAKKAENLQLPSSGPPK VDDKPAAGSDLSGSRVAVVVDTSGSSRSARS RTEKDPISRYPNSRVAIFRDRDIRKDPDYDRN YDRYLQRFDPFGFHAGPYTVQPMYSPAVNYNT EFPQLGSTPRPOISADPPRPLQHLPLWASPSN PAGLGYGHPEPTMMAAYSPNHVMPRSTSAIYLHT IQYPCQRPGMPFIHPHEHIHQSYQQSQQQQHD ATFGFARPR
546	sp_012530_qgin -0.918451	-0.9810354	-0.918459	LOCATED IN: mitochondrion; EXPRESSED IN: shoot apex, embryo, flower, seed; EXPRESSED DURING: petal differentiation and expansion stage, E expanded cotyledon stage, D bilateral stage; BEST Arabidopsis thaliana protein match is: PIF / Ping-Pong family of plant transposases (TAIR:AT3G55350.1).	MAASPKPSTKRRTTKRKTATTATITLSPLTAAI TTALTSSHFFHHELHLLFQSOSVLOSSLSSVF LSLKLQPHQLALPLSSPSSCCWFQRFLLSSS NDEDPRWVDFRMSKPTFYRLNSLSNSISSVLP SSISPDFAIGVAICRLAHNWSFVFKRFLGSSSD ACRVYAVCKVIGQLGGLYGGFGSELRIWVGF ERISLNCGGVLGVEKFFIEGELLGGNGFLMQAL VDSQGRFLDVSAGWPSFTDPSILRQTRFLIJSVD ESKELNGPEVIDGLMPPQYIMGSDSCYCPCLPW LLTPFARRGEAEDSSSSKKEFNEAKHRATDLVR MAFGRVRANWKLDRKWKEEFAEFLPFIITACVL NNFLMKHGELLSHGDLCGVREODIPVFDGEVDE SAAPIRDLALHLRSVRTQDN

10

20

30

40

【表 1 4 - 1 8 7】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
547	sp_012600_uczs	#N/A	-0.854913	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family; CONTAINS InterPro DOMAIN/s: Late embryogenesis abundant protein, group 2 (InterPro:IPR004864); BEST Arabidopsis thaliana protein match is: hydroxyproline-rich glycoprotein family protein (TAIR:AT3G52460.1); Has 868 Blast hits to 868 proteins in 31 species: Archae - 2; Bacteria - 0; Metazoa - 2; Fungi - 0; Plants - 862; Viruses - 0; Other Eukaryotes - 2 (sources: NCBI BLink).	MTDPSRATGYPAPPQNGYPPPPPPQSQSNYYY PNHNPNYNNAPQAYYPAAPPRATFFRRFFAA LIAALIFGVILLVWLVLRPIRPSFTVDSVLSLNV SSSPQISGTWLSLHFSVSNPNHKLVRVYDEIAR LYYRSSLLTRNQLAPFDQDTKNITTMDAKLVASS AYESELVNRKIGDRSTGMVNFVQV/LALVRFKA GGWRAPRRILRVFCDNLVSFSSNSTMGTLSSGG PKKFCVSGI
548	sp_012680_xqaa	#N/A	-0.865343	Intron maturase, type II family protein; FUNCTIONS IN: RNA binding, RNA-directed DNA polymerase activity, INVOLVED IN: RNA-dependent DNA replication, RNA splicing; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Intron maturase, type II (InterPro:IPR000442), RNA-directed DNA polymerase (reverse transcriptase) (InterPro:IPR000477); BEST Arabidopsis thaliana protein match is: Intron maturase, type II family protein (TAIR:AT1G30010.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).	MVLIIRITGSGILPLLRRHSYRLNGFALSRSFSGF HNNWRKPDPPDPLSLMKEDGVAVCSHWIEN FREPEKSASNLTSYLKRFELWVLAQKVVYADET GSYLPRHSIQKPALEDLRLNVAVLDRFKWGAR LNFYIRSPDKTEYESLSKRKIKITLITTTQAPYQD RIVQEVLFMWLEPVYVESRFSQSYAFRPGRTAHT VLRVIRSFAGYLWYLGKDLSTLDGMKVGLVINA LMFDVDRKVIDLIRSAVLMVPIETKIDGEEKQKK KKRKYOKKKVLADDEPKPDPYWLDSFFGFAPKE AEKVPFCWGHCGILSPLLANILDELHWMMEGKIK EFYHPSKSDVIWNSPEGEAEQGNISWPEFVPTS GPDKTRKMDYIRYGGHILIGIRPRADAVALRKOL VEFCDDQYLLKLDNDSLPIEHITKGMFLDHLVLCR RVVYPTLRYTATGGKIISEKGVGTLTSSVTSUKQC IRQFRKMNLLKGDREPDPPQCFRMMFHATQAHN SQMNKLLCTMVEVYRFADNRKKNVNFVSIIRGS LAKLYAAKYKLRSAKVKISDRTLSRP_LKEKKG SPEYQNMRLRMGLAESIEGLKYTRMSLVLPETDYP FFGNWRPDHEKALLEFLSLDDPKTLEDQORNCIKE QGLISPDQYISMLVWNYKRNKASLSDQLSLLNSGG KNIQSDQLEMESSNEDCVSEDEDEDEEDKERIN MAHV

10

20

30

40

【表 1 4 - 1 8 8】

A	B	C	D	E	F
1	分析に用いたペイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2				PHD finger family protein; FUNCTIONS IN: DNA binding, zinc ion binding; INVOLVED IN: regulation of transcription, DNA-dependent; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: DDT domain (InterPro:IPR004022), Zinc finger, PHD-type, conserved site (InterPro:IPR019786), EF-Hand 1, calcium-binding site (InterPro:IPR018247), Zinc finger, PHD-type (InterPro:IPR001965), DDT domain superfamily (InterPro:IPR018501), DDT domain, subgroup (InterPro:IPR018500), Zinc finger, FYVE/PHD-type (InterPro:IPR011011), Zinc finger, PHD-finger (InterPro:IPR019787); BEST Arabidopsis thaliana protein match is: metalloendopeptidases;zinc ion binding;DNA binding (TAIR:AT5G35210.2); Has 3640 Blast hits to 3265 proteins in 204 species: Archae - 0; Bacteria - 0; Metazoa - 2369; Fungi - 398; Plants - 598; Viruses - 0; Other Eukaryotes - 275 (source: NCBI/BLINK).	MESGGKSRGFRFRKTRFRGDEKNGENVDGEGE VQDMVRDNDNDDNEGEDVKNAAGNSRKRGRPP KRRAVEIRGDSVWGRYVVKDFEGSGVFLGKIVSY DSGLRYVYEDGCEDELDSDELREYLIGEERQKF VGDLLRRKQKLDENVADLLEKKQKSDENVSPKT VVKSKSERKRNVRVVDVEVEKGNVAEPIIMDK QDCGNMLNDTEKVEVPISSGRFDEEVDGSEVDD DDVSDSSDSCEYGGGWETRSEIEIPAIPPELP QSSGSIGVPEEHVPHLFSVYGFRLRSFIRLFLSPF TLDDLGGVCTQNTLLDAIHVGLLRALRWHLE AQATDGSSELASKCLRCDWGLLDLITWPVYLVHY LMVMGHLNRKWKRFYDVLQKDYCTLSIGTKLIM VLQILCDDALDAAETRAEIDNRESEFVGVDSDGM SSFSENGPRKYPRSAKTSIAKERENLILEAKNN TFPGSDSLTVKGNQDTSKSDNIDEDENGDECLR CGMDGVLLCCDGGCPSAYHSRCIGVYKLSIPTGD WFCPECKISKLGPPITMKTSLRGAEVFGVDFYAQ LFLGTGNHLLVYSLKTSPPFVRYNCNDIPQVWT ALCSSEYHAVSYSGICQAILKYWELPQVLLPPPI CQTAMVLEKDDSDSVLLSAETGNQVLNKVEPEN YASGVGESSLNYSVGGPNEKSSSTHGYDDVLIG APVQNDIAPRMNGMNESSTMSNKLKKNAGLSG SSPQGLADSSVMGISSCTSGYKSSSAGQRNSV CLPGNVYLPKYRNEAGVRAVQQWKDNGNCLYI GTSFKYHYHYNHGNFAATAAANFAGLSDEN RLSEVHSDPRKVMASAKVTLQAKAFSSATSRRF WPSTEKLVLEIPRERCWSGLHCKAPVSSRKAACL LNQAALSATRAAMKFCGSLKLVKNVEGGLYGITA YTLYLEESMRGLIVGPFHSAVYQKQWRKKEVEQA STCREMKSLLEENLNCINALSSEWTKLVDSLIM DELSVQSACTAVSNQKRGVPRKVKHLATAERIV EDSNNSLNDLNRWRSGLSKLYFORATLPCPVLR MSPDVLSSSTQDIGKSPKDKKMYTKRKKDPLA SLNKSCKEKLKALNQLLRFILCKKRE
549	#N/A	#N/A	-0.8932336		
550	#N/A	-0.8133394	-0.8930734	uncharacterized protein	

10

20

30

40

【表 1 4 - 1 8 9】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
551	sp_013730_tmpz	#N/A	#0.857703	Penitricopeptide repeat (PPR) superfamily protein; CONTAINS InterPro DOMAIN/s: Penitricopeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: Tetratricopeptide repeat (TPR)-like superfamily protein (TAIR:AT5G01110.1); Has 56384 Blast hits to 14280 proteins in 305 species: Archae - 4; Bacteria - 41; Metazoa - 1079; Fungi - 1056; Plants - 52130; Viruses - 0; Other Eukaryotes - 2074 (source: NCBI BLINK).	MKQLHQLRRQSNFSSSAVITNTCKPPQPPPPPP PPLTQNLQTSILNSQWHEVLAHTLTPAVISIT FFNLFSPKTVVSKFLDQINPHLDADSNLALVIV SLLPSPKSAVQLAKRIINRLVTSDDFFTKLGSR EKLNVEISICYDYWIRGLCQLKKAEEAFKCFKMM KDLGVLPKVETCNMMLSLFLRFNYTQMAWVLYA EMFKLISSNTHTFNIMINVLCKEGLKKAIEFVE CMEIAGFKPNVATYNTLINGYCSRGKIETARKILC EMKVRGIYDSTYTYGTIINGLCKEGRVDESVRFF DKMVDSSLVPTAVNFNALIDGYCNKGDLEKAFG YRDEMVKKGVAPTLPTYNMLIHCLLMIEGKDDTAY QLIKEMEEKGIKDDVTYNILINAYCQSGNWKAF ELLDEMLVEEIQTRVYSSLIYVLSKSRMKEAT DLFEKTRKGLTDLVMFNTLINGHCYNGKMEKA LLV-SEMDRRKVNPDVWYNTLMQGLCRVGVKE EARLVDKMGKRGISPDHVSFNTLISGYSKRGDM KDAFKVDEMLDIGNPTLLTYNALIRGLCKNKEG ALAQELLKEMVGRGITPNDSTYFYLVDSIGDVEN FLENSRPQQVKIQCKKNQVKGLGPYS
552	sp_013900_tyut	#N/A	#0.89372537	UDP-Glycosyltransferase superfamily protein; FUNCTIONS IN: transferase activity, transferring hexosyl groups, UDP-glycosyltransferase activity; INVOLVED IN: metabolic process; LOCATED IN: endomembrane system; EXPRESSED IN: 8 plant structures; EXPRESSED DURING: 4 anthesis, LP.02 two leaves visible, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: UDP-glucuronosyl/UDP-glycosyltransferase (InterPro:IPR002213); BEST Arabidopsis thaliana protein match is: UDP-glycosyltransferase 76C1 (TAIR:AT5G05870.1); Has 7583 Blast hits to 7530 proteins in 440 species: Archae - 0; Bacteria - 352; Metazoa - 2017; Fungi - 25; Plants - 5047; Viruses - 86; Other Eukaryotes - 56 (source: NCBI BLINK).	MGLLAIATATATASMEKQVIKLGCRYLLEFPYPY VGHTTFMFLGNALYSKGFSTYVIQTRLNSSDQT TFPNFTFHYEEALWSEALASVDPDPWEAVTLIN KSFPLPFRDLSQILEADASKDIEPVACLISDPLWD FAGTIADELNLRALRFTGGLLAFVLYDSVFLRQ KGYPIQSESEEEAVSEIPLKVKDLPPEAIHDTL AVVWKEAKTYRLGYICNTFKELEGSFLDLFGRSLP GTPIFPIGLHKKYVPTVMGEKLEDEEHSITWLNQ APNSVLVYSEFGTIAAISKEQLEVAWGLAKSQOP FLWVVRPKMINGLVKDDNDIFNPFLETVSGRGY TYTWAQQLVLAHPSVGGGFWTHCGWNSSTIESIC QGVPMICLPFFGDQKVMARHVSDILRGLHLEKGI EQONIEAIRRLMLGKEGEMRERVAALKEEAQS CLMEGSSFKALDQLTNHILSFCSAEIPKKHRR

10

20

30

40

【 1 1 8 9】

【表 14 - 190】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
553	sp_013960_ecsx	-0.85508285	-0.9170823	-0.9422959	<p>P-loop containing nucleoside triphosphate hydrolases superfamily protein; FUNCTIONS IN: helicase activity, nucleic acid binding, ATP binding, ATP-dependent helicase activity; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: DNA/RNA helicase, DEAD/DEAH box type, N-terminal (InterPro:IPR011545), RNA helicase, DEAD-box type, Q motif (InterPro:IPR014014), RNA helicase, ATP-dependent, DEAD-box, conserved site (InterPro:IPR000629), DEAD-like helicase, N-terminal (InterPro:IPR014001), DNA/RNA helicase, C-terminal (InterPro:IPR001650), Helicase, superfamily 1/2, ATP-binding domain (InterPro:IPR014021); BEST Arabidopsis thaliana protein match is: DEA(D/H)-box RNA helicase family protein (TAIR:AT5G60990.1); Has 60083 Blast hits to 51461 proteins in 3186 species: Archae - 589; Bacteria - 20345; Metazoa - 13239; Fungi - 7387; Plants - 3446; Viruses - 273; Other Eukaryotes - 14804 (source: NCBI BLINK).</p>	<p>MFHSPVYGGASSNHLFSDDTDFGFVSLLEI  DESYYNLAIPESREDNVKEKSGTSHKAKTRKRV  TIGAECDVGVIEVDKYLEQGSVKEKKNKKRRKT  KQVQTOPKDGDEIEADKEMEVEEYVNOKKRKL  KQVQQTETNTEDGEQLVDENEFNAWHELRLH  PDLMKAYKLGFKQPTPIQKACIPAAAHQGKDVIG  AAETGSGKTLAFGPIQLRLLDEREKGEKLMSED  AEAVEILAPRGFIRALIITPRELALQVTDHLKAAA  TSTNFRVAIVGGMSTEKQERLLKERFEIVVGT  GRLWELMSGGEKHLVELHLSFFVLDDEADRMIE  NGHFQELRSIIDLLPGTSGVTEKHFQDVQNCITIS  NFQRRKQRTVFSATIALSRDFRKKLKRGSLSK  HSATDGLSSLEALSERAFAGMKPDAVFDLTS  TSNLEESFLECEEDKDAYLYILCVHGGGQTVF  CSSTALRHVAALLRLNLSVCLVHGEMQQRARL  KALDRFRFNKNGILVADSGAARGLDIPGVRTVIH  QLPFAAEYVHRSGRTARGNTDGSLLLVSSNE  RSKFASLCKSFQKDTLQRFPIVNSYMPEVRKRLS  LACQIDKISRDRSKEKANKSWFKRNAESIGLDVE  DNDSEERVNCRCQKKTSSLOMKLQQLNLTLL  SRFLQPKFSFHRFLAGAGVSPLLQHQLEALAQ  KCINDKSSGNKKNVVIGQDCVPEPLQALRRAGP  EVRFLDLNALAEKQQDMEKLLRRKRKDAKFKRQEQ  KRQEKRRRKEAIGS</p>

10

20

30

40

【表 1 4 - 1 9 1】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
554	sp_014140_gqgg	-0.326774	-0.961377	0.36412	Outer arm dynein light chain 1 protein; INVOLVED IN: biological_process unknown; LOCATED IN: vacuole; CONTAINS InterPro DOMAIN's: Leucine-rich repeat, typical subtype (InterPro:IPR003591), Leucine-rich repeat (InterPro:IPR001611); BEST Arabidopsis thaliana protein match is: Outer arm dynein light chain 1 protein (TAIR:AT4G03260.2); Has 24271 Blast hits to 14439 proteins in 747 species: Archae - 6; Bacteria - 9514; Metazoa - 8699; Fungi - 1002; Plants - 3127; Viruses - 44; Other Eukaryotes - 1879 (source: NCBI BLink).	MAVLFKCFKFIGRKKKVDKDAVSDNGVRTLQVK VEHPMKPSENWDDLKSTFSFVVPFDIPEISTCE VKLMNSDSPVEGDQEEVAYEGEDEHDETSIMSR DNDSNLQARDNIGTPRNLSPHQIGSSPCAEEED DVAKNNEIIDTGRNGHVSDFGMAREEKWASPR LSRSCSNIESREVLKTTQSFKQLQELSNITNEEA RSPCCSHSSALTHRSADKVMKHKHSYQILFPRS KKLWKLFLWHRNLCKPEIAKTRVSSLLASV NQGGYSSDTMEPTRYKQKVESTGSESEQ SSYNKSSHEEDNNGLEGRFSGLWPNQWYAF SGESSSLKRVDDWVKDFEFOFSALEEGNVVD DDDEQIVFPSSPKVDDSVVKNPTHFPRHNSKINL TEELLHANSVIQSLNSSSTVAHVSGMGLKVIPIH FSSLRSLNNSNSIVQITQGSLLPKGLHVLDSRINKI TTIEGLRELTRLRVLDLSYNRISRIGHGLFSCSLK ELYLAGNKISEVEGLHRLKLTVLDLDFNKITTTKA LGQLVANYSLLALNILGNPILSNLSSDQLPRAVC GLLSKLAYLNKQPVNQKAREVSMDIKSSRS NSWNARRKSVKKTGGQGSMLSGSVRSNASVRQ RTRNKAKREGNQQRSLKIRITTTTTTS

10

20

30

40

【表 1 4 - 1 9 2】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kriq (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
555	sp_014440_dsja	#N/A	#N/A	59.97053%	<p>Cation efflux family protein; FUNCTIONS IN: cation transmembrane transporter activity, efflux transmembrane transporter activity; INVOLVED IN: cation transport, transmembrane transport; LOCATED IN: membrane; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Cation efflux protein (InterPro:IPR002524); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G38380.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).</p>	<p>MEDSHHHHRPQRLSVPPRTNVLTLRHP SAYP  LTPPTPSKTRLSRSTSTINDKSSIPFLLFLSLR  SLYSVLPFLRSSPFSFLPFSFLVLSLLSFLTLFSFS  LYSKHPRLPFLSLSLSSQSVRHLHLFKSLVLMV  FLRFQALRYCGTAAMILAELSGNMAFRWKDQ  NWDRDGHRSRVRGFFALFSGFLLSYSWDRMDC  FPLSYNLGKIQNSLMVVVDGNCRLRIVPMLLPFLS  GFLGCRYERDLMNWGTIRQLGRKRVQLISLFFTTT  MLFFPAINMLVFEAEGGVSIMNLGWSLANTVLF  GVLLNEWYNDKRVNPSDEKEFLITFLCTLVLE  LVYFPELSLWGLICGFLWVAIQQLNWARSRVY  ELGAESETFLSMWRPIRHVLSERKSRKIALFLM  INTGYMVVEFVAGFMSNSLGLISDACHMLFDCAA  LAIGLYASYISRLPANSQFNRYGRGFEVLSGYVN  AVLLLVGALIVLESLELRIDPQEIISTNSLLAVSIGG  LLVNVVGLIFFEHEEHHHAHGSCSHSDSHSHT  HDDHDHVTVMHHDGVLVYGESKNGSHDMHGES  HQHDHVHSHGDHDSHTHGHHDHSHHHGHND  HDSHGHHDHDSHAHHDHDSHAHHDHDSHAH  HDSHSDQNHSCGHHDQHHHQGHGHHDHND  NEVCKSIPTVPGVDSCSHSHHLDKSTESSHKKP  RHQHDHNMEGIFLHVADTLGSGVGVVISTLLIKY  KGWL VADPACSFIFSVLIVSVIPLLRNSAEVLLOR  VPRAHEQDLKAAVTDVAMKIKGVGQIQLHVWSLT  NTDVISTLHLRISEIDQAAATKARVMHVLEDAGVK  DLTLQVEGVKSL</p>

10

20

30

40

【 1 1 9 2】

【表 1 4 - 1 9 3】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2					<p>WGRKKKLVDDDETSVATLSEESQVTEEDKVVYF TGGKKKGGKNSKGAQSRDEEVEVAFVA GKKSSKGGVKGAFELLDGDDNDQKDGSD DEEDVSVSFTGKKSSKSGKGGKASFALLG DDGDEDEVKGGSEGEDDVPVSVFTGKKKPSK GSKKGGSKTSFALLDDEDNDGDNVQSEDEAP SMFVGGKPKSGKSGSSLSAALLEEDGEE DDSHSEAPSVSKTAYDEADAVVAVFTGKKKS SKKGGSSALTALGDDITTKMETKIANNAEVDDYL PQEKVSAVVPETSKNKKKKGGRTAEQQDLD KILAELEGPAVSKPEPTPAPSQVEKADTQLEPV GLTEAAGEKEGEDEGIVESAAKKKKKKEKE KKAASASSAEIKVEKEDVKNKAVDKKLP KHVREMQEALARRKMEERQKQEEERLRKEE EERLRQELLRQAEELKKKEKEKELRLKKEA GLFLAKQKEEQKREIMKNQFLATAGNKLPSAD NGAPTRPIYTTKSKSNRQTNVKAQTLDEME TNQAVGDSSELVETEKVDEVCTIDVEEKSEVADE IEDVVAEEDNEDDEDWDAKSWDESLSLPATR AFEDEETDSDGQSLVKVEKQAAQSVRDTGRPA ATANEAHSSQATKGGTSLKQKHSDSEAVGKKR SKETEVEQKPKPSDQAKNGEGDLRSPICIMG HVDGTGKLLDCIRGTNVQEGEAGGITQQIGATY FPAENIRERTRELKADAKLVPGLLVIDTPGHESF TNLRSRGSGLCDIALVDMHGLEPQIESLNLRL MRNTEFIVALNKVDRLYGWKTCRNAPIVKAMKQ QSKDYQNEFNVRTIGVITQFKEGGLNTELYYKKN DMGETYNIPTSAISGEGIPDMLLVLWQWSQKTM IEKLFTRDEICTVLEKVVVEGHGTTIDVYLVNGV LHEGDQIWCVMQGPICATIFALLTPHPMKELRV KGTYLHKKQIKAAQGIKISAOGLHEAHPGTALYV GPNDDVEEIKLAMMEDMESVLSRIDKSGEGVCV</p>
556	sp_014450_yfzl	-0.831725	-0.713539	#N/A	

10

20

30

40

【表 1 4 - 1 9 4】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptax (SOAP2-like)	注釈	aa seq
2	共発現遺伝子				<pre> MDDSVVVVVDDGGGGVWFVKKKRRSSSKFSLQ NWVGGFSRQVNGKGRSSNVKYSNISKPGKD SVQRGVSSHSSSKEHEKDTCLDKCVVGGDTG RPKTPPKIITTTSSQSSSTANENVTEKSDIVQRI KWGDLEDDALDCGTAAVSEIRFGEIGDDDLAMC CTSRNVCDLSACNLSKDYVNVNNSFEVLEEASD MCYSKPPFPVDESLEENCKEVEIVASTDVGMLST TTKVIDANCGLDHHKELVNDVENVHEDHACGS SNDVFCQSSIEEPAKIVELSAPILTPGVCDLGFSS VKDEDNADGMLIARDMIDISVPNLAEAEPENCNA ASVPADCGSGIPNSSSDNLLNPTMINREPEG GESKERFRERLWCFLENLRAVDELYLLCELEC DLGQVKEAMVVLQEAESDFRELNIRVEEFELTK SCPPTSGAVMSIKGDRHRPHALSWEVRRMTTSP HRAEILSSSEAFKVKQEEFEREKERASGEVKKRHK RLPTMDEHKRENIPDTSELETKSRRKSGISDAAR LNSTRERRSTEGRRSKAPVQNSRVALVTSASD PTTTKLPFRGSSAAFVGGKCKEQLGSLAEMEK LFEGDVFRRHNVAGDKKEKERNQSPWRPVDAAW KEKRNWEEILAPRRLLSRVSHSPCMSRRSIERA RVLHDKLMSPEKKKKTAVDMKKEAEEKHARAMR IRSELENERVQKLORTSEKLNRVSEWHAVRSMK LREGMYARHORSESRHEAHIAQVVKRAGDESIIK VNEVRFITSLNEENKFLRCKLHDSMRRAEKF QVMKTKQKEDMARFEAVIERKLIIEAEKLRQRLAE TQRRKEEAQLRREERKASSAAFEARTMEQLRR KEERAKAQEEAELLAQRLAERLSESDORRKYV LEQIRERASMDFRDQSSPLMFRSLNKEGGQRLT PPRNDDNQTINPSSSESGFNNGIGMQQSLKR RIKKIROKLMALKHDLPENPAEAENVGIGYRTAVAT APAKIGRWLQDLQRLRQERKEGAASIGLITAEIMI KYLDDGKELELOSSRQAGLLDFISSALPASHISKPE </pre>
557	sp_015320_cmxs	-0.873892	-0.85537%	uncharacterized protein	-0.919379

10

20

30

40

【表 1 4 - 1 9 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
558	sp_015800_tqrp	-0.351483	-0.397113	multidrug resistance-associated protein 14 (MRP14); FUNCTIONS IN: ATPase activity, coupled to transmembrane movement of substances; INVOLVED IN: transport, transmembrane transport; LOCATED IN: plasma membrane; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: ATPase, AAA+ type, core (InterPro:IPR003593), ABC transporter-like (InterPro:IPR003439), ABC transporter integral membrane type 1 (InterPro:IPR017940), ABC transporter, transmembrane domain, type 1 (InterPro:IPR011527), ABC transporter, transmembrane domain (InterPro:IPR01140), ABC transporter, conserved site (InterPro:IPR017871); BEST Arabidopsis thaliana protein match is: multidrug resistance-associated protein 6 (TAIR:AT3G21250.2); Has 673798 Blast hits to 360971 proteins in 4075 species: Archae - 12279; Bacteria - 544381; Metazoa - 13093; Fungi - 8967; Plants - 6941; Viruses - 21; Other Eukaryotes - 88116 (source: NCBI Link).	MLSSFELFANSHCDNVNCGDRSRVCFHFMVSEG KIVRASYPHELLMTSEAFQNLVNYKDYDTTASQNSKE VLRDRTAKTVTSKILQEDTKENLNYSVDLQIKREERE MGDAGIQPHLQYLKYGNALYFPLSNFMQWLQH LSSPSCWNRFSGLQRLRSMIQHLWEEYSVGLPLV YLKNCQIYYSDDHTLAIKRIGYTLGIMAGSNCSASL KDTQHSVPGGTLASQLTTHNSQPLVCI
559	sp_016290_hnar	#N/A	#N/A	minichromosome maintenance 8 (MCM8); FUNCTIONS IN: nucleoside-triphosphatase activity, DNA-dependent ATPase activity, DNA binding, nucleoside binding, ATP binding; INVOLVED IN: DNA-dependent DNA replication initiation, DNA replication; LOCATED IN: cellular component unknown; EXPRESSED IN: shoot apex, embryo, flower, seed; EXPRESSED DURING: F mature embryo stage, petal differentiation and expansion stage, E expanded cotyledon stage, D bilateral stage; CONTAINS InterPro DOMAIN/s: Nucleic acid-binding, OB-fold-like (InterPro:IPR016027), Nucleic acid-binding, OB-fold (InterPro:IPR012340), ATPase, AAA+ type, core (InterPro:IPR003593), DNA-dependent ATPase MCM (InterPro:IPR001208); BEST Arabidopsis thaliana protein match is: Minichromosome maintenance (MCM2/3/5) family protein (TAIR:AT2G16440.1); Has 4027 Blast hits to 3651 proteins in 539 species: Archae - 371; Bacteria - 134; Metazoa - 1249; Fungi - 930; Plants - 423; Viruses - 3; Other Eukaryotes - 917 (source: NCBI Link).	MDGGGESIVMTMMQRPFAAEPCCDLLHIWPIY FPQIELSFQDWRYLISKLVCFFATPRGNHLVSQV KHEDDTSLSLQNFQEFKICDLECFYVNLLEDYPK EALLCMGAAIHKVTSFOEKSLESHGKINIRLHNY PESLJALRNLKAAYIDRLVSVRGTAVKVSTVPLV MQMTFKICQCGADIIRFFPDGKYSPPATCMRCC KGKTFVPIRPTAQSIDFQKIRLQELLPENHEEGR VPHRTVECELTELDVVCIPGDVVTVTGIHQINNY MDLGGKSKSHOGLYYLYLEAVSVTNSKSLSP EDSEYTTADARATEMSDLFSFSPRDLERFVKFNE EHGPDVFRMIOQICPSYNGHELKAGITLALFGG AAKHSMDQNKVPVRGDIHVIVGDPGLGKSQLQ VAAAVSPRGYVCGNATTNAGLTVAVVKDRMTS DYAFEAGAVVLADRGCLCIDEFDKMSAEHQALLE AMEQCVSVAKAGLVASLSARTSILAAANPVGG HYNRKTYNENLKMSSPLSRFDLVFILLDRPDE HMDKRLSEHIMARKQPDGGQGVFILGSPDILSLL EGHLSSEDRVTLT

10

20

30

40

【 1 1 9 5】

【表 1 4 - 1 9 6】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2				related to ubiquitin 1 (RUB1); INVOLVED IN: protein modification process, response to auxin stimulus, protein neddylation, ethylene biosynthetic process, embryo development; LOCATED IN: plasma membrane; EXPRESSED IN: 6 plant structures; EXPRESSED DURING: L mature pollen stage; M germinated pollen stage; CONTAINS InterPro DOMAINs: Ubiquitin subgroup (InterPro:IPR019956), Ubiquitin conserved site (InterPro:IPR019954), Ubiquitin (InterPro:IPR000626), Ubiquitin supergroup (InterPro:IPR019955); BEST Arabidopsis thaliana protein match is: ubiquitin 7 (TAIR:AT2G35635.1); Has 16064 Blast hits to 7349 proteins in 721 species: Archae - 0; Bacteria - 36; Metazoa - 7330; Fungi - 1879; Plants - 3570; Viruses - 331; Other Eukaryotes - 2918 (source: NCBI BLINK).	
560	sp_016390_trim -0.8371232	#N/A	-0.83552		MQIFVKTLTKTKITLEVESDITDNNKAKIKQKKEGI PPDQQRLLIFAGKQLEDGRTTLADYNIQKESLHLV LRLRGGTMIKVKTLTGKEIEIDIEPNDDTIDRIKERVE EKEGIPPVQQRLLIYAGKQLADDDKTKADYNIIEGGS VLHLVLLALRGGACASASA
561	sp_016410_mtuo -0.8372843	-0.8362457	-0.8357624	DGD1 SUPPRESSOR 1 (DGS1); INVOLVED IN: galactolipid biosynthetic process, hydrogen peroxide biosynthetic process; LOCATED IN: mitochondrial outer membrane, mitochondrion, protein complex; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Nuclear control of ATP synthase 2 (InterPro:IPR019946); Has 223 Blast hits to 223 proteins in 116 species: Archae - 0; Bacteria - 2; Metazoa - 1; Fungi - 125; Plants - 59; Viruses - 0; Other Eukaryotes - 36 (source: NCBI BLINK).	MEVASPSQEDAESRNFRTLTLYSNIWTRFSNS PFIGISLRCRQFGRAKSRKQKLCPLPLPSKSV SSJL TEASSIYDVL EDVFEHLSNMHNVOKNLRY WQAKAEGSNAQKAYFMIFERGPRALFDETFEYV RNFIAEDSSFOILCHSASGFISERVAVLDSLOSAL ATFLAQYMMIDKQGEDLVDPQKALPSSLVTIN GLFSDLENTIAQVHASMISLIDSCAGENKSSALVF NKMTEVQLEQPOWSDSEIRDAVDIINQIQKLD YISLVARHOKPRMTRYWIRYTCGAVGLSVCSL WIVRHSRLVSPDIDNWFSDAKESTVSWFTEHV EHPILSIRDELFEFRKRRHRGVMEEVEVHLTANSL RRMLVAFCEQANGKLPENATDLDLQIVMGRY EKEMMHPVKSVMFGGDLARMLIQVQKLDIETA MLELNQILKANEINFAIALPFAVIVSMLIVRS WIMQDKGAEGRGRLARRQRRLLIVEIEKRIMHYQ SCMDMGQAQNAQCSFGLVLYTLDRLYRAVERH AKMTGEWSCLRQDIIDLAKPDLQIITYKLTITSRME RVYDCLLPSLKG

10

20

30

40

【表 1 4 - 1 9 7】

A	B	C	D	E	F
1	sp_107620_kqnh	sp_107660_kiqg	sp_107670_ptqx		
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
562	sp_016450_fxnk	#N/A	#N/A	<p>molybdopterin biosynthesis MoeE family protein;            FUNCTIONS IN: molecular_function unknown;            INVOLVED IN: Mo-molybdopterin cofactor            biosynthetic process; LOCATED IN:            cellular_component unknown; EXPRESSED IN: 22            plant structures; EXPRESSED DURING: 13 growth            stages; CONTAINS InterPro DOMAIN/s:            Molybdopterin biosynthesis MoeE            (InterPro:IPR003448); Has 30201 Blast hits to 17322            proteins in 780 species: Archae - 12; Bacteria - 1396;            Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses            - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).</p>	<p>MSEEDKTLIELEENPIDIKYINFRAPKCGAIAT            FSGTTRDTFDGKAVVELRYEAYVPMIRCIKSICT            SAFETWNLIAVAHCLGPVPVGGTSVFIATSSV            HRVDALDACQYLDELKATVPIWKKEVYANGEVW            KENSEFLERRTEIGVPVGETSVFIATSSVHRVDA            LDACQYLDELKATVPIWKKEVYANGEVWKENSE            FLERRSELSENFRRPLNRSSEKKGCCGCTKVRV            DEERSNKKSSENTGINQ</p>
563	sp_016990_fuiw	#N/A	#N/A	<p>YODA (YDA); FUNCTIONS IN: protein            serine/threonine kinase activity, protein kinase            activity, kinase activity, ATP binding; INVOLVED IN:            stomatal complex morphogenesis, embryo            development ending in seed dormancy; LOCATED            IN: cellular_component unknown; EXPRESSED IN:            22 plant structures; EXPRESSED DURING: 13 growth            stages; CONTAINS InterPro DOMAIN/s: Protein            kinase, ATP binding site (InterPro:IPR017441),            Protein kinase, catalytic domain            (InterPro:IPR000719), Serine/threonine-protein kinase            domain (InterPro:IPR002290), Serine/threonine-            protein kinase-like domain (InterPro:IPR017442),            Protein kinase-like domain (InterPro:IPR011009);            BEST Arabidopsis thaliana protein match is: mitogen-            activated protein kinase kinase kinase 3            (TAIR:AT1G53570.1); Has 133707 Blast hits to            131404 proteins in 4998 species: Archae - 148;            Bacteria - 15191; Metazoa - 50613; Fungi - 12908;            Plants - 32496; Viruses - 572; Other Eukaryotes -            21779 (source: NCBI BLINK).</p>	<p>MPSWWGKSSSKDAKKKTGDETILDAIRKFRNS            SDVKLDSRSGASRRHCIDTVSERGAQSAESRS            PPSKVKSHVGRCCFAERPHQAQLPLGLHPA            NWSRTDSGISLSSKPRTEKSKSSLFLPLRPSCI            RSRADPAELSDVWTVNSVSSASSIESDDLTESRL            RSPQAFDLDIGARTAAAGSPSSVVLKDGATKAAQ            NDYVEAQKPVSLPLSNQSTTASPRRRRLGNHV            PNLLVPPHGFACSDPSSMSSPSRSPRLTPNSE            QILNAAFWASKPYPDVTVLLGGHCHCSSPGSGQNS            GHNSMGGDISGQLFWQQPSRGSPEYSPIPSPR            MTSFGPSRRVQSGAVTPIHPRAAGSGSDMHAS            WGDGKQOSHLPLPVVTIPSPSPFHGNSAAT            SPVSPRPARAENLPSPTSRRWKKGRLLGRGTFG            HVYLGFNSETGEMCAMKEVTLFSDDPKSKESAK            QLQOEIALSRLRHPNINQYQYGSSETVGDRLYYLE            FVSGGSYKLLQEYQGLGELAIRSYTQQLSGLAF            LHGKNTVHRDIKGANLLVDPNGRVKLDLDFGMAK            HITGQSCPLSKFGSPYWMAPVIVKNSGNCNLAV            DIWSLQCTVLEMATTKPPWVSYEGVAAFMKIGN            SKELPAIPDHLSEDEKSFVKLCIQRDPOORLSAA            QLLEHPVKNATAPLVRPIPSPPEADTPIINILNGT            RSLGVGVERNLSSLDSERLAVHSSRVSPARPRS            SEYIIPRNISCPVSPIGSPLLNPRSPQHLNGMMS            PPSIPSSGTTSGSSTPLTGYSGAIPFFHSQSQQS            LYVQEGFGLPKSPNSVFNAPCYHDPNPNILR            GMSQSGPHAITLVSPKNGHNHVKQVERERSCO            GGDPGNGRTILVDRVSKQLLREHVKKQNPVSEF            SPSRSLPFSRTHHGL</p>

10

20

30

40

【表 1 4 - 1 9 8】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
564	sp_017400_qppq	-0.88430%	-0.93337%	general regulatory factor 9 (GRF9); FUNCTIONS IN: protein phosphorylated amino acid binding, calcium ion binding; LOCATED IN: chloroplast stroma, nucleus, plasma membrane, cytoplasm; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: 14-3-3 protein (InterPro:IPR000308); BEST Arabidopsis thaliana protein match is: general regulatory factor 12 (TAIR:AT1G26480.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archaea - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI/BLINK).	MASTKERESFYIAKLAEOAERYDEMVMKKNLA KLDVELTVEERNL SVGKYNVVGARRASWRILSS IEQKEEGKNEQWVKRIEYRQKVELELTNICSDI MTVINEHLIPSA TAGESTVYFYKMKGYRYRLAE FKFGNERKEAADQSMKAYESATTTAAEALPPTH PIRLGLALNFSVYFYELINSPERACHLAKQAFDEAI SELSLINEDSYKDSLTLMQLLRDNLTLWTSDIPEE GDFQKADGAALKAGDDAE MITSLECPAFSPFPMVNTTEIGAVLAVMRRNRVWG TRFTADDQQDHSLSLKLTKLKKLFTSHAHWHDI DPVYTLQPF.LSVQSEETGAPITSVALS SVYKILTL DIFDMETINIQDTMHLVDVAVTSCHEFVTDPASEE VYLMKILQVLLACVKCAAAVALSNQHVCNVDTCF RVVQOTS AKGELLQRIARHSMHEL VRSIFSRLEI LCHELDSGGNLINHEIQVNSDPIFVAPQPDSDK NTPTD SKL TETCSTSSSSAPFVDCPSNGTTTTMDT VQYESHLMNQGFVPCWELHFLC SLINVKEHA DMNPKNSLAFDEEDVPLFALGLINS AIELGGSSF QRHPRLKIQDNLFCNLMOFGMSMSPILSMVC SIVNL YQHLRTKLQLEAFFSSVLSLAQSRYG ASYQQQEVAMEALYDFCQRQRTFIMEMYANYDCD LTSRNIFEELVLLSKSAFPVNTPLSTMHVLALEG LVAVIQGMAERTADERYVDLGNITSITL EYKPFWL FKNGYEEFGHWVFLRQTKYIKORLMAGADHF NRDPKEGLEFLOSTHLLPEKLDQAQSVACFLRYTT GLDKNLVGDFLGNHEEFVQVINEFAQTDFDKG MNMMDTALRIFLETFRLPGEAQIQRVLEAFSARY YEQSPHILANKDAALLSYSHLLNTDQHNSOVKK KMTTEEDFIRNRLINDGNLPRELLELYSSICEN EIRITPEOGSSGCEIRLSRWIDLMORSQNTSPYIA AAKYPLDDDMFAIVAGPAVAASVIFDHAEDVD YQLCIDGFSAAKIAAYLHDSVLDVDDIVLSCKFTT LLNPSYLEEFLAYSNDRTRARISITTTFTIANRYG DCIGTGWONJINCLIRFEYEGFLSAFVCAADADEP ALSALPEHVAKPLANTLSSPHIQISVAPKPKPSGLMG RFSOLL SFETEPEISOPSEELAAARLCTOEIVOKG QINNIFTESKLOADSLMHLVRAIVGAEQIQAEITG YPEDEANATFCLELLIATLNNRDRISLLWQSVYE HISNIVQSTMPCALVEKAVFGLLRICQRLLPYKE NLIDELLSQLV.LKLDARVADAYCEQITQEVGQL
565	sp_018230_yayi	#N/A	-0.3653%	GNOM (GN); FUNCTIONS IN: protein homodimerization activity, GTP:GDP antiporter activity; INVOLVED IN: in 13 processes; LOCATED IN: intracellular; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: SEC7-like (InterPro:IPR000904); BEST Arabidopsis thaliana protein match is: GNOM-like 1 (TAIR:AT5G39500.1).	

10

20

30

40

【表 1 4 - 1 9 9】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈  PELOTA (PEL1); FUNCTIONS IN: translation release factor activity; INVOLVED IN: meiosis, translational termination, translation; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: eRF1 domain 2 (InterPro:IPR005141), eRF1 domain 3 (InterPro:IPR005142), eRF1 domain 1 (InterPro:IPR005140). Probable translation factor pelota (InterPro:IPR004405); BEST Arabidopsis thaliana protein match is: Eukaryotic release factor 1 (eRF1) family protein (TAIR:AT3G58390.1); Has 776 Blast hits to 774 proteins in 323 species: Archae - 255; Bacteria - 0; Metazoa - 140; Fungi - 159; Plants - 51; Viruses - 1; Other Eukaryotes - 170 (source: NCBI BLink).	aa seq  MKIVRRDYVHGFGPSIKMIPDEADDMMIAYNLIA EGDSVLAVTVRVKVMREAAASGGRAERVKLRLEI VVEAVEYKEGSLVRIRGKNTLEND-HVKIGAFHT LELELHRPVLVKDIWDLALDMLYQAADPAASA DLVLIMQEGLAHLLLVGKSVTLTRSRVSIIPRK HGPAIAGYESALKFFFNVLQAFKLYYDFNVVRC AVIASPGFTKQDFYRHLLEAEERKQLRSIENKSR VLVHTTSGYKHSLSREVLDPNVWSMIKDTKAARE VFRALDFNMLSNDFPARACYGPKHVEVAHERMA IQTLITDEQFRNADIATRKYYVNLIGSIKASGGSV HIFSSMHVSGEQLAQLTGVAAILRFPPLDLDIEM  MALKLNKKGWHTGSLRNIEVWVWKAQKNEAE QRKLEELRQIIEERDEFRLVLEQOAGLTPRQE RLEFLYDLSGLAVGKPFSSDGFQALGSPFSADNSA SASASAPVLASAPAPSSAQAPVPGALFEKQPQ SANDAWRKLHSDPLLMIRQEQEAMARIKINPV KNAMIRKSVESKKHDESHDGKHGKVKSKHKKH SSSSKHHSESESDSHEERERRRKVKQQRKEHH SESESDSHEERERRRKVKQQRKEHHSDTESDEE RERRRKTKNKKKEHSDSDSDSDSDSDSDSDSDSD HFNYDGKSDRINKDSDRIOGQYENSKHSEYES KHTGLPPRQDYHESKHGLPMDQYHESKHHTGLP PRQDYHESKHHTALPPRQESHRRRGRAPQLSEEE RAARLREMQMDAETHHEEQRWKRLKADENDAR EASQAGNSRGGKSLDSVQKSVYGAEGKGGSTI EESVRRRSHFLOGRSSADEGNAFRR
566	sp_018890_pzwn	-0.927159	-0.935854	CBF1-interacting co-repressor CIR, N-terminal;Pre-mRNA splicing factor; CONTAINS InterPro DOMAIN/s: CBF1-interacting co-repressor CIR, N-terminal (InterPro:IPR019339), Pre-mRNA splicing factor (InterPro:IPR022209); BEST Arabidopsis thaliana protein match is: CBF1-interacting co-repressor CIR, N-terminal;Pre-mRNA splicing factor (TAIR:AT2G44195.1); Has 59661 Blast hits to 33745 proteins in 1751 species: Archae - 130; Bacteria - 4372; Metazoa - 27826; Fungi - 6422; Plants - 3866; Viruses - 275; Other Eukaryotes - 16770 (source: NCBI BLink).	
567	sp_019490_aoeg	-0.837335	#N/A	phosphoprotein phosphatase inhibitors; FUNCTIONS IN: phosphoprotein phosphatase inhibitor activity; INVOLVED IN: regulation of signal transduction, regulation of phosphoprotein phosphatase activity; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Protein phosphatase inhibitor 2 (IPP-2) (InterPro:IPR007062); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLink).	
568	sp_019500_ixjs	-0.844236	-0.886328		

10

20

30

40

【表 1 4 - 2 0 0】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたペアとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
569	sp_020010_coop	-0.3314	-0.334255	#N/A	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 16 growth stages; Has 42 Blast hits to 42 proteins in 17 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 41; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLINK).	MSSTNAAKSWLQSLRRYKAPWQITGFCASPEY KSAVPRATEYRLFPCPATIPQEAIVPSSNPDTVFDI KYFPRDQRNRPAIRRTVLKADVEKMMAEKTF QPSDYPIPYLTATIEEDMINTHGGGYQ
570	sp_020070_xujw	#N/A	#N/A	-0.630330	SMAD/FHA domain-containing protein ; CONTAINS InterPro DOMAIN/s: SMAD/FHA domain (InterPro:IPR008984), Forkhead-associated (FHA) domain (InterPro:IPR00253); BEST Arabidopsis thaliana protein match is: SMAD/FHA domain- containing protein (TAIR:AT3G07260.1); Has 459 Blast hits to 458 proteins in 156 species: Archae - 0; Bacteria - 0; Metazoa - 122; Fungi - 202; Plants - 94; Viruses - 0; Other Eukaryotes - 41 (source: NCBI BLINK).	MAHTAVAGGGGSDVEAGFAKLOGEDEFYYMQ TYSIILGRNSKSTYDYLASLGGGMINSRHHARI FYDFORRRFALEV LKNGCLVEGV LHLPGNIPVK LDSQDLLQIGDKFYL LPRNLLGGAAPRRHHVY NTP LPPAVPARNSGYGGGGSRKGRMKEFY AAPAAGEEYEEEEGGRRMCGEDGGGLSKMIR RYGDGSGGGHGDYGYGGSGGGKGLSGTLE KKSDFRSRVDPEADNQQLMQLLEEKDWTVSATL LSDMCGPDWMAMEKHLHTELMEQYGHVWHHS RVFRYLTSEDWTFPEKGPWYGLMLLRKYPE HFVINTRSKGRVTLFEVSLVTLTT
571	sp_020340_amsq	-0.225039	#N/A	-0.301533	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MSGREDDSDAPEEYV TALGQQDEEIRKALQLS KSRVRRGKERRKLAARLTPROKSKDEVPDGD VEMQVDFVDEDESOKSVNTGMLPSDVVNLLAA REKOVFLSDSEDEKEKELKPKKRAKRAFSGVETV ILKELPPHCVQNSLEFLKRRKMQVPRSNVAVLNN SKQALRLSSSGLLK

10

20

30

40

【表 1 4 - 2 0 1】

A	B	C	D	E	F
1	分析に用いたペアとの相関係数	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	sp_107620_kqnh (SOAP1)			protein kinase family protein / peptidoglycan-binding LysM contain-containing protein; FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation, cell wall macromolecule catabolic process; LOCATED IN: endomembrane system; EXPRESSED DURING: 22 plant stages; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Peptidoglycan-binding Lysin subgroup (InterPro:IPR002482), Serine/threonine-protein kinase domain (InterPro:IPR002290), Serine-threonine/tyrosine-protein kinase (InterPro:IPR01245), Serine/threonine-protein kinase, active site (InterPro:IPR008271), Protein kinase-like domain (InterPro:IPR011009), Protein kinase, catalytic domain (InterPro:IPR000719), Tyrosine-protein kinase, catalytic domain (InterPro:IPR020635); BEST Arabidopsis thaliana protein match is: chitin elicitor receptor kinase 1 (TAIR:AT3G21630.1); Has 117626 Blast hits to 116415 proteins in 4493 species: Archae - 110; Bacteria - 13527; Metazoa - 43649; Fungi - 10139; Plants - 32288; Viruses - 471; Other Eukaryotes - 17442 (source: NCBI BLink).	
572	sp_020530_msca	-0.884377	-0.8603	-0.917146	MDHEGRGLDRVHSRTHKNSLCSLRKSSNILLDG AFRAKISDLGLSKLVGTTSDDEEASATRVVGTGY LAPGS
573	sp_020760_mzey	-0.816596	-0.944275	-0.846535	MALRRALGWSGDGELMRDCKPQSRILMRQTAGIF TYGGALGFVWLCRLHYGTGNKCLI

10

20

30

40

【表 1 4 - 2 0 2】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
574	sp_020910_ppkh -0.859436	-0.859436	#N/A	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G14750.1); Has 5478 Blast hits to 4354 proteins in 533 species: Archae - 87; Bacteria - 653; Metazoa - 2554; Fungi - 380; Plants - 418; Viruses - 16; Other Eukaryotes - 1370 (source: NCBI BLINK).	MGSKGRLLPPHMRHPLHGPGLVHPEPFGAGIRP PHCGFPHGDMLEPPPEILEQKASQHVEMERLAT ENQRLAATHGSSRQQLAAAOQELQMLETOIRAS KSEREQQTRGLMDRISKMETELKAVDRLLKDLK QARTEAQTLEVRQELITKAQQLNSELORAHVDV QQIPAMSELNHLRQEFHQYRTTYEHEKKVYRD HIESLQAMDKEYKMADEVAKLRTELNNTSNVDK RSVYGSAAAGYGDIDASAHPNSAGSSYEDGYGM QQVHHQFPASGAATTASTPAAATGGGVVASSGG TPTYAPQSGTASGRVGYGNPGEYAGRGLTYD POGSQVYNAQRASGYDGYNRGNVYDMQRPP YDFQRAAYEMPWANNYDAHSRGAAPPPPPS CGOAPVNNMAYGSAAPPPAARTTAGNEMQTRG GNTARR
575	sp_022320_njtt -0.913134	-0.922964	-0.9228365	Pentatricopeptide repeat (PPR) superfamily protein; CONTAINS InterPro DOMAINs: Pentatricopeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: Pentatricopeptide repeat (PPR) superfamily protein (TAIR:AT4G17616.1); Has 9904 Blast hits to 5133 proteins in 176 species: Archae - 0; Bacteria - 34; Metazoa - 111; Fungi - 31; Plants - 9502; Viruses - 0; Other Eukaryotes - 226 (source: NCBI BLINK).	MVKMLRIKLCQILSSWLGTRYVISSVLSNYR EINPSQPAISSFVNGSRLNPSLLYQGNVANRTKF FSTTAGTILVOARDPAKLSLDIHNAVNDHRYDDA WRYFEQHLEMEGFPKRSVNVNKLLAGYAETLDSQ LLERAFGLVERAFEEGKHTLFEKVTLIYAFALAR CSQPVPASVLRKLRIMEEYPPVTAWSAVLAYS LSSGAYLAAELILEIGYLFQDNRIDPRKKSAPLI AMKPNLTAISVALAGCLVSGTTRIAEQLLDMMPRI GVKTDATLILMAHYERNRREELKLRHIDEA HDLDIQFRQYYNCLLNCHLKFGLDSDASRMVLE MLRKAEANNSLAAATLRFEPANSYSPKIVSNITE LSISDKGNGHFMSFEDEFCDRDNFICLEGETKEV LISLSLQKQTELMITERGILQPTTEKTYVKLAKAF LEACKTKDLAEFLIKVEKEDSPFVEDSPLVDVIN SCISLGLDQAHDLDEMCFAGIKISSVYSSLLK AYCKEKLSEVMSLLREVRKAGIQLDASCYDALL QSRVLRDQTQGLNLFKEMKETKIPRAGHQEYVE VQARSCKASSEAGFMAKLLQEIKEGETVDCGVH DWNVNHFFCKKRLMQDAEKALKKMRSLGHTPN AQTFFHSMVTGYAAGIKGYEIVTELVGEMKSLAAA TPMKFDQELDSVLYTFVRGGFFFRANEVEMNM EKGNMFDIKYRYRTLFLKYHKTLYKGTKPKFETE AGSKKRDATAFAFKKVVGLF
576	sp_022610_ktnf #N/A	-0.821137	-0.85511	uncharacterized protein	MADLYDPVWVPTDPRGAMTLSFSRKRKNGD KVESKSKGNGDGKDKVSSKRGKYSNDDVFE MLKGYPVSLERHMMGDSSDSE

10

20

30

40

【表 1 4 - 2 0 3】

A	B	C	D	E	F
1	分析に用いたサイトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
577	sp_022780_iqwn	-0.354134	-0.323485	dicer-like 1 (DCL1); CONTAINS InterPro DOMAIN/s: Restriction endonuclease type I, R subunit/Type III, Fes subunit (InterPro:IPR006935), Double-stranded RNA-binding (InterPro:IPR001159), Argonaute/Dicer protein, PAZ (InterPro:IPR003100), Ribonuclease III (InterPro:IPR000999), Double-stranded RNA-binding-like (InterPro:IPR014720), DEAD-like helicase, C-terminal (InterPro:IPR014001), DNA/RNA helicase, C-terminal (InterPro:IPR001650), Helicase, superfamily 1/2, ATP-binding domain (InterPro:IPR014021), Dicer double-stranded RNA-binding fold (InterPro:IPR005034); BEST Arabidopsis thaliana protein match is: dicer-like 3 (TAIR:AT3G43920.2); Has 21958 Blast hits to 17420 proteins in 2982 species: Archaea - 328; Bacteria - 11461; Metazoa - 3615; Fungi - 1668; Plants - 1373; Viruses - 45; Other Eukaryotes - 3468 (source: NCBI BLINK).	MIEGINKVSDSRFINVWLDVACEEODDIIICFEIIGD FGFEFGSCNQPVEVAGIDPSPFFGGIDHILDSIKNGAG LPIAADNNDGMDIITPPP TTPKDMQVENLVNCGN VEVQRSSDNKLNIGVDNGSNMNDNDNDNDNMN MNINGDGKVVVDVRTSNGDSKRPSAREDCNV DRDRDSEERGGKRTLSSGFRDERGYSSRVQS YGESNGRKYREFDE TGRDRDRDRDRDLGGH RREYNNINGRRRESRDRERDRREGKGYWEREGP GMKEIYRVGSGYEPARNRDANKVEFGKNOECNV RVGKNEEGNKEKLEERARHYQLEVLQCAKOK NTIAFLETGAGKTLAVLLIKSICNRLQQQNKVLA VFLPKVPLYYQAAEVIREFQOTGYSVGHYCGEMG QDFWDTRRWQREFETKQVLMTAGILLNLRHSI IKMEYIHLILDECHHAVKHPYSLVMSEFYHTTP KEKRPATFGMTASPVNLSKVSSQVDCAIKIRNLET KLDSSVCTIKDRKELEKHPMPSEIWEYDKAASL LYLHEQITQMEVTVVEEAAARAS TRSKWQFMGAR DAGAKEELRQVYGVSERTESDGAANLQKLRVAV NYALGELGQWCAYKYAHNLFVALQNDERANYQL DYKVFQESYLSKVWALLQCLSEGAAFNKDASSTE EVNGFHQDNGSVDEIEEGELPDCHVWSSGGEHVD VVIAGAAVADGKVTPKVQSLVKVLMKYQHTDDFR AIIFVERVVAALVLPKVFALPFLSNIKFCASLGHN NNOEMRTSOMQDTIARFRDGRVTLVATVSAEE GLDIRQCNVIRFDLAKTVLAIYQSRGRARKPGS DYLLMVERGNMMSHETFLRNARNSEETLRKEAIER TDISHLKG TAGLVSADNVPGSVYQVASTGATVSL NSAVGLHFYCSQLPSDDRYSLHPEFIMERHEKPG GSSTESCRHLHLPNAPENLEGLPCSSMRLAQQ AVCLAAGKHLHEMGAFDMLLPDKGSGGDEGEKI EPNDEGEPLPGTARHREFYPEGIASILRGEWILS GRDGCNSKIFHLVYVYVWKEHVHVGSSKDPFLTQ

10

20

30

40

【表 1 4 - 2 0 4】

	A	B	C	D	E	F
1						
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
578	sp_023010_ecqs	-0.81722%	-0.99253%	-0.32274%	Major Facilitator Superfamily with SPX (SYG1/Pho81/XPR1) contain-containing protein; FUNCTIONS IN: molecular function unknown; INVOLVED IN: transmembrane transport; LOCATED IN: cellular component unknown; CONTAINS InterPro DOMAIN/s: SPX, N-terminal (InterPro:IPR004331), Major facilitator superfamily (InterPro:IPR011701), Major facilitator superfamily, general substrate transporter (InterPro:IPR016196); BEST Arabidopsis thaliana protein match is: Major Facilitator Superfamily with SPX (SYG1/Pho81/XPR1) domain-containing protein (TAIR:AT4G11810.1).	MVAFGKIKERQLOEWQGYINYINKLMKKVKQY GOHIETGLRRQVLDKDFSRMLDIQIEKIVLFLLE QQGLLAGRLAKGECQHNSPDOOSDIATISQLREA YREVGLDLLKLLYEVEINAIGLRKILKFKDFKRFGR FTDYVKTADHPYSYQLQQVFKHVGIGAVGAI RNLGLQERQGSYASIQDPALPLQDPVIDSIKSA VDRLGYSYTNFLHFLAQHALIMQDEILPSPVEEELV DKSYHFDLSILNLANITLYMINTYIIVPTADDSMS LGAATVCGIVGAMAVAQVSSVYFSAWSNRSY FRPLVFSSIVLYGNIMYALAFDFNSWILLIGRVF CGLSARAVNRRYISDCVPLKIRMQASAGFVSAS ALGMACGPALAGLLQTNKLFNLTFNQDTLFGW VMAYAWLLYLVLWIAFKEPSHEDDYVEQQSNT ETQONDVMERGLQOPLITEKIERSEEKERNGD DDGDACEDESEEAPEEESRKPVTSFASAYRLLT PSIKYQLLIYFMUKYAMEILLSVSWVTTYFGWT TSSVSIFLACLGLTLPVNIWVGSYISINMFQERQIL LASQIMVCIGILLSFNFFVKYTPQVYVCSGJMFVS AEVLEGVNLISLRVMSSRLSRGTYNGGLLSTEA GTLARVADGTTLIGLLGTSSLLNLTLLPSFLICY GCIATCYTYSNLS

10

20

30

40

【表 1 4 - 2 0 5】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	分析に用いたパイロットとの相関係数 sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq VVVAVVFCGLVNVCFYFRFRSRCSJWIGTELLCRAG LKEEDLFLAVPPGFESFKAKKVKQADSEPCSV SSNLSSQESVWVNSEPGIGGGKISRSLRHPS MNYGRYNQCSDEESDCEKLDQNFPTKPLPKG VIRGCTEQNCQKVIARWHPAEACRADLEEVV FYPSSEFKDTRIEKIRPOAELYGICRIIPPLWK PPCLLKEKDIWEHAKFVTRVQRVDKLONRNASR KGSKVHSHSRKRRRCGRMDVEFGSGNGKTSE PGEAEFGPEPPEFTLDEFQRYADTFKVQYFRS ENCTDSGVENWEPVSDNIEGEGYWRVVEKPTKEI EVLYGADLETGVFGSGFKFSDQSCSTSDSYV KSGWNLNLPRLAGSVLSYENADISGVQVPWLY VGMCFSSFCQWVEDHHLVSLNYMHFGAPKLWY GVPGSDAKLLEEMKRLPDLFDEQPDLLHKLVT QLSPSILKSAGVYVFCVQNAgefVLTFRPRAYHA GFNSGFNCAEAVNVAPVDWLP-HQNAVELYRE QGRKTTISHDKLLGAAAREAVKANWELSLRNT MDNLRWKDVCGRDGLAKSLRARVDTENVRREF LCKFSKALKMENSFDATTERECSICFFDLHLSAA GCRGSPDRFACLNHAKQLCSCAWDGKFFLFRYD VSELNLLADALEGKLSAIYRWARLDLGLALSSCS KDNQGLTGLSGAVSEQPGENIICKTNSQPIVTLK NNEVIAKTPAVLASSVGMKSLVEKGGKFPDVAL EATKVAPILSRQSKRKSQKQIGKDGEEVFNKR FTGMPEQLKLSHGPEDLICLSDDEDEGSNKTYI DANKGTIGQDSPNVPEIEGSDHKLNLCDLANAAL AAEANLISDGHKDDCMSCSTSPKEETNVKGET VAVSDEHNSCHQSGISGSSIEKKDPDASTPDSFL QNAACVFNVGGIGKFRDES VQVRLDDARILPQS PSCSQSSVDKYPKQKGPRIAKVRRRINCVELLE FGVIVPGLNLSWTSRAIFPKGFKSRVRYINISDPTN MCNYISRLDGGGLNGFLFMVLEHCPSEVFVQTL
2	sp_107620_kqnh (SOAP1)	分析に用いたパイロットとの相関係数 sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq VVVAVVFCGLVNVCFYFRFRSRCSJWIGTELLCRAG LKEEDLFLAVPPGFESFKAKKVKQADSEPCSV SSNLSSQESVWVNSEPGIGGGKISRSLRHPS MNYGRYNQCSDEESDCEKLDQNFPTKPLPKG VIRGCTEQNCQKVIARWHPAEACRADLEEVV FYPSSEFKDTRIEKIRPOAELYGICRIIPPLWK PPCLLKEKDIWEHAKFVTRVQRVDKLONRNASR KGSKVHSHSRKRRRCGRMDVEFGSGNGKTSE PGEAEFGPEPPEFTLDEFQRYADTFKVQYFRS ENCTDSGVENWEPVSDNIEGEGYWRVVEKPTKEI EVLYGADLETGVFGSGFKFSDQSCSTSDSYV KSGWNLNLPRLAGSVLSYENADISGVQVPWLY VGMCFSSFCQWVEDHHLVSLNYMHFGAPKLWY GVPGSDAKLLEEMKRLPDLFDEQPDLLHKLVT QLSPSILKSAGVYVFCVQNAgefVLTFRPRAYHA GFNSGFNCAEAVNVAPVDWLP-HQNAVELYRE QGRKTTISHDKLLGAAAREAVKANWELSLRNT MDNLRWKDVCGRDGLAKSLRARVDTENVRREF LCKFSKALKMENSFDATTERECSICFFDLHLSAA GCRGSPDRFACLNHAKQLCSCAWDGKFFLFRYD VSELNLLADALEGKLSAIYRWARLDLGLALSSCS KDNQGLTGLSGAVSEQPGENIICKTNSQPIVTLK NNEVIAKTPAVLASSVGMKSLVEKGGKFPDVAL EATKVAPILSRQSKRKSQKQIGKDGEEVFNKR FTGMPEQLKLSHGPEDLICLSDDEDEGSNKTYI DANKGTIGQDSPNVPEIEGSDHKLNLCDLANAAL AAEANLISDGHKDDCMSCSTSPKEETNVKGET VAVSDEHNSCHQSGISGSSIEKKDPDASTPDSFL QNAACVFNVGGIGKFRDES VQVRLDDARILPQS PSCSQSSVDKYPKQKGPRIAKVRRRINCVELLE FGVIVPGLNLSWTSRAIFPKGFKSRVRYINISDPTN MCNYISRLDGGGLNGFLFMVLEHCPSEVFVQTL
579	sp_023490_loj	#N/A	#N/A	PKDM7D; FUNCTIONS IN: sequence-specific DNA binding; transcription factor activity; INVOLVED IN: regulation of transcription; LOCATED IN: nucleus; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Transcription factor jumonji/asparlyl beta-hydroxylase (InterPro:IPR003347), FY-rich, C-terminal (InterPro:IPR003889), FY-rich, N-terminal (InterPro:IPR003888), Transcription factor jumonji, JmjN (InterPro:IPR003349), Zinc finger, C5HC2-type (InterPro:IPR004198), FY-rich, C-terminal subgroup (InterPro:IPR018516), Transcription factor jumonji (InterPro:IPR013129), FY-rich, N-terminal subgroup (InterPro:IPR018518); BEST Arabidopsis thaliana protein match is: Transcription factor jumonji (jmi) family protein / zinc finger (C5HC2 type) family protein (TAIR:AT1G30810.2).	-0.8683171
580	sp_024110_wuro	#N/A	#N/A	RNA-binding (RRM/RBD/RNP motifs) family protein; FUNCTIONS IN: TATA-binding protein binding, nucleic acid binding; INVOLVED IN: biological process unknown; LOCATED IN: cellular component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: RNA recognition motif, RNP-1 (InterPro:IPR000504); Has 35333 Blast hits to 34131 proteins in 2444 species: Archaea - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	-0.393107

10

20

30

40

【 1 2 0 5】

【表 1 4 - 2 0 6】

	A	B	C	D	E	F
1						
2	共発露遺伝子	分析に用いたサイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 Transducin/WD40 repeat-like superfamily protein; CONTAINS InterPro DOMAIN/s: WD40 repeat-like- containing domain (InterPro:IPR011046), WD40 repeat (InterPro:IPR001680), WD40/YVTN repeat-like- containing domain (InterPro:IPR015943), WD40 repeat, subgroup (InterPro:IPR019781); Has 672 Blast hits to 646 proteins in 166 species: Archae - 0; Bacteria - 97; Metazoa - 271; Fungi - 156; Plants - 85; Viruses - 0; Other Eukaryotes - 63 (source: NCBI BLINK).	aa seq
581	sp_024490_etgk	#N/A	-0.836827	#N/A		MGVFAVDTFDLADVEVWLSLDDSSIVVWDSLLEYK VLIYTPDGRC
582	sp_024650_kuyy	-0.834326	#N/A	-0.654436	uncharacterized protein	MOFSKAGLDVSVSDMGTGTRKKVKGSCSESEVDAS CCSPFEVAAPKPVLIETVATNGFTYVPSHGPDE QHSVDDLGGTEPEELIEVAIGDSCVMNLDGNS SHIQCPDTKKSVAAEYSISAGLYGCLLILVQC HRISLHLRCHASYRRYFKNLCCCKGKVD

10

20

30

40

【 1 2 0 6】

【表 1 4 - 2 0 7】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	<p>calcium-dependent protein kinase 1 (CDPK1);            FUNCTIONS IN: calmodulin-dependent protein kinase activity, protein kinase activity, kinase activity;            INVOLVED IN: response to water deprivation, response to salt stress, protein amino acid phosphorylation, N-terminal protein myristoylation, abscisic acid mediated signalling pathway; LOCATED IN: chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages;            CONTAINS InterPro DOMAINs: Protein kinase, ATP binding site (InterPro:IPR017441), EF-Hand 1, calcium-binding site (InterPro:IPR018247), Serine/threonine-protein kinase domain (InterPro:IPR002290), EF-hand-like domain (InterPro:IPR011992), Calcium-binding EF-hand (InterPro:IPR002048), EF-hand (InterPro:IPR018248), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271), Protein kinase, catalytic domain (InterPro:IPR000719), EF-HAND 2 (InterPro:IPR018249), Calcium-dependent protein kinase (InterPro:IPR020642), Calcium/calmodulin-dependent protein kinase-like (InterPro:IPR020636); BEST Arabidopsis thaliana (TAIR:AT1G74740.1); Has 125370 Blast hits to 119826 proteins in 3847 species: Archae - 176; Bacteria - 14154; Metazoa - 46310; Fungi - 16874; Plants - 24900; Viruses - 438; Other Eukaryotes - 22518 (source: NCBI BLINK).</p>	aa seq
583	sp_024780_ydyr	#N/A	#N/A		<p>MGNACIRPEELPKKTPQTKKKKTRPNPYSINQ            PQPAPRVYVKDIIPAVGPRISDKYVIGRELGRGEF            GITYLCTDRETCDRLAVKSKRKRKRLTAIDVEDVFR            REVAIMSTLVPHPNIVRLKAVYEDNEAVHLVMEI            CEGGELFDRIVARGHYSERAAATVAKTVAEYVK            MCHDNGVMHRDLKPNFLFANKKENSPLKAIKF            GLSVFFRPGFERFEIVGSPYYMAPEVLRNRYGQ            EVDVWSAGVILYLLCGVPPFWAETEQQGVALAIL            RGYLDFKREPWPQISDSAKSLVKQMLDPPDKKR            LTAQQVLEHPWLQNAKASNVPLGDIVRTRLKQ            FAGMNRFKRRLRVAIEHLTYEEVEIIRDMLMD            TDRDGKTYEELRTGLKKGVSQGLGEPEIKLLMEV            ADADGNGYLDYGEFVAVTHLQKMNDEHFFHRA            FMFYFDKDGSGYIDLDELRLDALSDSAETDADVLT            EIMREYDTKDKGRISYEFEVAMMKAGTDWRKAS            RQYSRERFKSLSLNLMKDGSLQLHDGMTGEAVV            V</p>

10

20

30

40

【表 1 4 - 2 0 8】

A	B	C	D	E	F
1	分析に用いたペイトとの相関係数 sp_107620_kqnh sp_107670_ptqx	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	共発現遺伝子			<p>BR11 suppressor 1 (BSU1)-like 3 (BSL3);          FUNCTIONS IN: hydrolase activity, manganese ion binding, protein serine/threonine phosphatase activity, phosphoprotein phosphatase activity, iron ion binding;          INVOLVED IN: biological_process unknown;          LOCATED IN: cytosol, plasma membrane;          EXPRESSED IN: guard cell, cultured cell; CONTAINS InterPro DOMAIN/s: Serine/threonine_protein_phosphatase, BSU1 (InterPro:IPR012391), Metallophosphoesterase (InterPro:IPR004843), Galactose oxidase/kelch, beta-propeller (InterPro:IPR011043), Kelch repeat type 2 (InterPro:IPR011498), Kelch-type beta propeller (InterPro:IPR015915), Serine/threonine-specific protein phosphatase/bis(5-nucleosyl)-tetraphosphatase (InterPro:IPR006186); BEST Arabidopsis thaliana protein match is: BR11 suppressor 1 (BSU1)-like 2 (TAIR:AT1G08420.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).</p>	<p>MEVDSMSSMENDQDPSSSTQNINSSPSSPPPL          PSSSSPLSSVQTTPASSSPMEKDVQQQLRQQ          QQGQIVGPRCAPSYSVNAIEKKEKEDGPGPRCGH          TLTAVPAVGEESTGYIGPRILIFGGATALEGNS          AAAAGTPSSAGSAGIRLAGATADIHCHYDVLTKNWS          RIAPIGEPTPRAAAHVATAVGTMVVIGGIGPAGL          SAEDLHVLDLTOQRPRWHRVVVQGGPGPRYG          HVMALVQORYLMICGGNDGKRPLADVWALDAAA          KPYEWRKLEPEGEGPPCMYATASARSDGLLLL          CGGRDANSVPLASAYGLAKHRDGRWWEWAIAPG          VSPSPRYQHAAVFVNRHLVHSGGALGGRRMVE          DSSSVAVLDTAAGVWCDTKSVVTSVRTGRYSAD          AAGGDASVELTRRCRHAHAAVAGDLFIYGGLRGG          VLLDILLVAEDLAAEITTTAASIAAAAAAASNPP          GRPLGRYGTDDRLRESDESPVDPGAVVLGNPV          APPVNGDIYDINHENAMAQGRRLSKGVEYLVE          ASAAEAESATLAAKARQVNGEVELPDRGDR          GAETPSGKPIPTMIKPDSSAPISITPPPGVRLHHR          AVVVAETGGALGGMVRSLSIDQFENEGRRVSY          GTPENATAARKLLDRQMSINSVPKVVWVHLLKPR          GWKPPVRRQFFLDCNDIADLDCSAERIFASEPSV          LQLRAPIKIFGDLHGQFGDMLRFLFDEYGGSPSTAG          DIAYIDYLFGLDYVDRGQHSLETTILLSLKVEYPH          NVHLIRGNHEAADINALFGFRIECIERMGERDGIW          AWHRFNRLFNWLPLAALIEKIKIIMHGGIGRSINH          VEQIENIQRPITMEAGSIVLMDLLWSDPTENDSVE          GLRPNARGPLVTFGPDPRVMEFCNNNDLQIVR          AH-ECVMDGFERFAQGHILTLFSATNYCGTANNA          GAILYLGRDLVWPKLIHPLPPAISSPETSFDRIHE          DTWMQELNANRPPPTTRGRPQVANDRGLAWI</p>
584	sp_024870_skya	#N/A	0.873168		

10

20

30

40

【表 1 4 - 2 0 9】

I	A	B	C	D	E	F
1		分析に用いたペアとの相関係数	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	共発現遺伝子				<p>Nucleotidyltransferase family protein; FUNCTIONS IN: nucleic acid binding, nucleotidyltransferase activity; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED DURING: 4 anthesis, F mature embryo stage, petal differentiation and expansion stage, E expanded cotyledon stage, D bilateral stage; CONTAINS InterPro:IPR002934, PAP/25A-associated (InterPro:IPR002058); BEST Arabidopsis thaliana protein match is: Nucleotidyltransferase family protein (TAIR:AT4G00060.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).</p>	<p>MAVTKYSETLSPFLPSSAADTTATAGNDEYDD  EUVFRNEISLSSLEPTSTDAAAVDYFSLDVNDFSD  KVIQTNPSPEVVAERDKAEQIPERTLESGWF  RADCRFKSPMLQLHKEILDGDFLSPTPEEAASR  NSAIERSVDIRIWIWPCRYVEVFGSRYRTGLYLPSS  DIDVILESNIKTPQIGLOALARALSOKGIKKMQV  IGKARVPIKFEKVTVDQFDFIDVDNGPKAAEYI  QDAISRLLPPLPCLLKVFLQOQRELNVEVSGVGG  SYALLAWLIALRNVNDCQSPENLQILLVKFFE  FYAHLNTWVGVCNCRGRTFYLKSRRKGFQSK  GRPLFISIEDPQTPGNDIGKNSFYQIRSFAMMA  YATLNTKTIIMNLGNRSLGTIIRPDPVLLERKGG  SKGELTFNSLLPGAGEALDSLDEKQDILCNWRV  DNDYEEEEELPRGGNAESSGASSSGKRRKASK  SKSYGKVKENGSGFGKNNYEENGSRKERKRDKK  RSKSSG</p>
585	sp_025100_pcos	-0.884173	#N/A	#N/A	uncharacterized protein	MFAFRRRQSOQLNRRILSPVVDLRSPLTTRHSS LATRRAARRSQSGDFTVIWCDWFRLLAWSGC
586	sp_025110_lqaf	-0.885334	-0.942932	-0.830352	<p>ATCOAE; FUNCTIONS IN: ATP binding, dephospho-CoA kinase activity; INVOLVED IN: coenzyme A biosynthetic process; LOCATED IN: peroxisome, chloroplast, vacuole; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro:IPR001977; Has 7074 Blast hits to 4684; Metazoa - 251; Fungi - 135; Plants - 58; Viruses - 0; Other Eukaryotes - 1945 (source: NCBI BLINK).</p>	<p>MRIVGLTGGIASKGKTVSKLKAQGIPIVDIVAR  DALKKGTTGGWKKVDAFGDGLQPDGEVDRPNL  GRIVFSNPKRQLNRLNLLAPYSSGIFQEIILKWLK  GHKVVLDVPLLFEAKMDKWTNPVIVVWVDRET  QLNRLVSRDRITEEENKINKDAQMPLDSSKDKADI  VIDNTGSLDNLNHLQNLVRYVNRPLTWSEFVLS  RQGAFAFAFVYVGVFVAKNIFKRL</p>
587	sp_025140_mazi	-0.883776	#N/A	-0.824136		

10

20

30

40

【表 1 4 - 2 1 0】

A	B	C	D	E	F	
1	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数	sp_107670_ptqx (SOAP2-like)			
2	sp_107660_kiqg (SOAP2)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	
588	sp_025650_iaes	-0.892419	-0.244133	-0.953382	<p>Homeodomain-like transcriptional regulator;            FUNCTIONS IN: sequence-specific DNA binding            transcription factor activity; INVOLVED IN: regulation            of transcription, DNA-dependent, regulation of            transcription; LOCATED IN: nucleus; EXPRESSED            IN: 22 plant structures; EXPRESSED DURING: 13            growth stages; CONTAINS InterPro DOMAIN/s: DDT            domain superfamily (InterPro:IPR018501), Homeobox            (InterPro:IPR001356), Homeodomain-like            (InterPro:IPR009057), DDT domain,            subgroup (InterPro:IPR018500), DDT domain            (InterPro:IPR004022), Homeodomain-related            (InterPro:IPR012287); BEST Arabidopsis thaliana            protein match is: homeobox-1 (TAIR:AT1G28420.1);            Has 67503 Blast hits to 37811 proteins in 2117            species: Archae - 251; Bacteria - 7621; Metazoa -            32181; Fungi - 4940; Plants - 2680; Viruses - 356;            Other Eukaryotes - 19475 (source: NCBI BLINK).</p>	<pre> MEGGSSDGEKPKFPESGGGGGGGGGGGGGEPKTK KRKMKTASQLEVLEKTYQVATYPAEALRAELSVQ LGLTDFRLQMMWFCHRRLLKDRKPPQGGGGGG GGGGSGKRSKKREASPAATAATGGPNVAVAP LLPPAPVPVGLRDEMLGAEELGNEHGGSGSS PYGHMDPRGYVGRPGVAVPRISGHGRYYEPS QSWAELRAIAFVAQLGEPRLRDDGPAALGMFDP PPGAFGAPITTTAPQKPTGRTYEDRADAKPVK GSNRPHEVQFLPEQPTVRSSEYERVAQSHYYG SPSDGPGTKTPPLSSGRSYIHGYEEVPPVYVGP NQVSSLNLMSSQKVGSLMPSTSGDYETIARKSP YNNMGMESPFSAHPTSGLDNALISSERRATQEE DAGRLRERKRNEEARVAKEIEAHEKMKRKELE QDILRKRREQLRKMEDRERKREKLLKTRIVEKM KQREERYQREQKREMERERKFLKTRIVEKM KQKEELRREREARQAANDRAVARRLAKESTE LIEDERLEMEIAASRKSLSLSLSDSETLQNLDSF REMLRVFPKSVQLKPKFPATRPWMGSEENVGN LLMWRFLVTFADVLGLWPFTLDEFYQAFHDYD PRLLGEHVSLRTRIKDIEDVARTPSTGLVGNQNS VANPVGHQVQEGAYAWGDFIRSWQCINLPLT WPEVLRQLALSAGFGPKLKRNTEQAYSREEYE GNDGIDISNLTGAAVESAVAIMRERGYSNPRRS RHRLTPGTVKFAAFHVLSEGGNGLNILEVADRI QKSGLRDLTTSKTPEASIAAALSRTDKLFERTAPS TYCVRSFPRKPADGEAILSSAREKIREFKGFVD GYEADVDGMEADDEKEDSESDIADPEIDDL GTELTPNDVNPFLYDGNSSVPKVESDNGSRDLH SHWDKTPNGLADAGNVLSSAHLESFKELNDNG CFAEESMDVGGICQETNSDTHEDTDIDESCFGE PWVQGLMEGEYSNLSVEERLNSLVALLGVATEG NTRIVLEERLEAATLKKQMWAEAGLDKRRLIKE           </pre>

10

20

30

40

【 1 2 1 0】

【表 1 4 - 2 1 1】

A	B	C	D	E	F
1	分析に用いたペイトとの相関係数 sp_107620_kqnh	sp_107660_klog (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	共発現遺伝子				MEETGLRVPFRSTYMLVFPQLLALCKDEDFGVWRQ AIVSGTKSFCSSILOELAQCFHRHGKVERWAEEM WMHIEFKDAVLSLVFEPAFVGTTRLLALKLETCL LLLTSDASETERLTSGGKELRFNFVSWLASGHPVL DAAKLISEANRLFTSLDMLQSSSSLPGLTITVN CLASIAKRPEHYGOVLPDLLDFHNSSEIRGGHA ASIQSVRSALLGFLRCTHPVIESRDRLYOELRL MSAGDAVDQVLRQVQDKQMRNRRSRRDARVQK EEHSSSTQRPVLSGDSIKRKLPLVDNEEPDNVFEWA SKFMRYTSNGKSPDLVQIHGSGDDNPVANGISS EAPVVDNSVNIPIEQMIAVIAALLAEGERASESLEL LISQIHPDLLADIVITNMKHLPKTAPLAKQGLPSN SQSSSSVSAAGGADIMSLPVHSSLSAHPAAV AVSSSLPFGSANLQVARRDRPRDPRRLDPR RVAVPAEATLLPNAEEDTHVSSVFEFGPASNRS SLTSALNTEPYSVPDAEVDGKSRITLLVPEKEIMP KEEEADESKEFLSSSEFNVLVNGAHSFDPPTVNES SITLTSKMEPTDAASALTQFSEEDDSDPPASNTS ACEDTCEDLPEVPSVYLTEKQRIKVLAVERIV KSYKQLKQGFESKTRMSLLARLVSWMGDDLV MMVQQDIVLDYRDQQGHDLVNHLVYHLHLITIE VEQSSGAADLYEKFLGLVAKSLDTPASDKSFS RLGGEAPMLTDSVTRLENLNCYPDGSMDIRGKDIR DDRTVTQGLGAVWSLILGRPLNRQACLSVLKCAV HSEDDTRAKNRLVSNKLYQLNVYSETIEQYAKN MLL-SAVDCTTDADASESICSGREVGVAEETSVTG PQLIDFGTSHAVSQAVASMSFFEAORLISLYFSLC TKKPGLLQIVFDVYARAGKIVKQAIHRHPIRLRAL GSSYSQLLRIISDPPMGSENLLMLVQLQISEGTPP PPDLISTVKHLVETKLDATILIPMLSS-SKNEVLP FRLVDLPLDKFQQALAHILQGSASHTGPALTPAE VLVAIHDIIVPEKEIKALKKVTDCSACACFEQRTVFT
589	sp_025820_mrtis	-0.864325	-0.951793	-0.921175	CONTAINS InterPro DOMAIN/s: Symplekin_riht junction protein C-terminal (InterPro:IPR022075); BEST Arabidopsis thaliana protein match is: HEAT repeat-containing protein (TAIR:AT5G01400.1); Has 593 Blast hits to 297 proteins in 111 species: Archaea - 0; Bacteria - 3; Metazoa - 139; Fungi - 49; Plants - 57; Viruses - 0; Other Eukaryotes - 345 (source: NCBI BLink).

10

20

30

40

【 1 2 1 1】

【表 1 4 - 2 1 2】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたペアとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
590	sp_026180_qgyj	#N/A	#N/A	-0.8502834	FAR1-related sequence 5 (FRS5); FUNCTIONS IN: zinc ion binding; INVOLVED IN: response to red or far red light; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: MULE transposase, conserved domain related (InterPro:IPR004330), Zinc finger, PMZ-type (InterPro:IPR006564), Zinc finger, SWIM-type (InterPro:IPR007527); BEST Arabidopsis thaliana protein match is: FAR1-related sequence 3 (TAIR:AT2G27110.2); Has 1793 Blast hits to 1580 proteins in 47 species: Archae - 2; Bacteria - 0; Metazoa - 4; Fungi - 136; Plants - 1646; Viruses - 0; Other Eukaryotes - 5 (source: NCBI BLINK).	MENLRISVIQSTQTQVVPVPEPDIQDATMEDVLYNC IENNALAQEREEEEAMEAEEDDIISDSEEPDQDIV GTOMGYTAETVEELLGFYQKHASAVGFSIRKGN TRFKVGRTRIVLEKTYVCSAAGVTNNGKNNKTKVQ TWPVVPKKERKPRQVSIITRTOCRACLRVKMNS EGRYEVVNHVIMHNDLTRSQWYHLHRSEKQIT EEKREAIETMQKSGNSYLVKLLCSV
591	sp_026370_zdww	#N/A	-0.874031	#N/A	unknown protein; Has 19 Blast hits to 19 proteins in 8 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 19; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).	MAISPVKVEVLSLELFRELLIAISQSEPENRVSSKSF YEYCNLSLNGAKYKKNEDKTEELRSELISLISCSASSAD NVIMPSMSMQEAEAS

10

20

30

40

【表 1 4 - 2 1 3】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
592	sp_026920_oruh	#N/A	-0.673345	<p>FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Nop14-like protein (InterPro:IPR007276); Has 69842 Blast hits to 35213 proteins in 1572 species: Archaea - 363; Bacteria - 20593; Metazoa - 20851; Fungi - 8010; Plants - 2912; Viruses - 517; Other Eukaryotes - 16596 (source: NCBI BLINK).</p>	<p>MAKSETMKNKKNKSKKAGVKKVSMKQKSAKPD NPFETIWSRRKFDVIGKRRKGEERRIGQARTRAI DKRKKTLLEVEOSGKASVFDNRHIGEQNDALG EFDKAILRSQFEROMKIKKTSKYNLSDEEEDLDI YDGLRDDYDVEHGSDEADAGDSSSILRHGRK SYDANIAEEENFLKSKKEVMEIIEIQKSKFFKEQKA REKEENEKLVQLDKDFESLVHNSALSSLTQPNK MNAKALVNPNFSEKELGKVEVMSVSKVSKODEP DYEKLVSEMAFEMRARPDRRTKTAEEIAEEKE RLERLEQRQDRMRATDDAEEEDDSESDGKP FTNKHSHVSGDDLGDSTFYDDKSLNKKGWDEI LERNDTEDEDEDESRNSDDSEEDDEEGGTE DEEDDGKAETLKWELSDDEVIGTDLLEEDKS DEEGGEEDNKTHSEIEDKNERLANPVRKVKKK TALLDLGMAQRSSLKQELPFKINAPGTMEELHAL LDDHSNEDIEAIRRIRICNAIALAENRKKMQVFY GLLLQYFATIPNKVPLNFQLLLLVKPLMEMSSEI PFFSAICARERIORTRTKFSGDIKDPGNSCWPSM KTLTLLRLWSMIFPCSDFRHVVMTFASLLMCEYL MRCPLVSVRDAAGSFLCSVLLVTKNSKKFCPEA LTFKDLKDLAATVPRCPKDSQMRTQDHPKPLWIG ESIEDIKPLDFLKMNLPKDSSFFNSHDFRASVLVS VVETLRGVNIYGELKAFPEIFSPIKALMIEVADQK NMPNPLQEKIRDVVELIKKAKDEYDEHRQPLEMR KQKPVPIKLLNPKFEENFIKGRDYDPRIRAEENRK MKKLLRKEARGAVSELKDNHFLQAVKEKEKQR VDEERAETYGKARAFLOEQEHNFKSGQLGKGRK RRR</p>
593	sp_026950_rcux	#N/A	#N/A	uncharacterized protein	MVESTSNSTRDLDRLLRSGHLVGNFPEPSAE

10

20

30

40

【 1 2 1 3】

【表 1 4 - 2 1 4】

A	B	C	D	E	F
1	分析に用いたペアトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2				注釈	aa seq
594	sp_027370_yznu	#N/A	-0.846304	HIT zinc finger; PAPA-1-like conserved region; CONTAINS InterPro DOMAINs: Zinc finger, HIT-type (InterPro:IPR007529), PAPA-1-like conserved region (InterPro:IPR006880); BEST Arabidopsis thaliana protein match is: PAPA-1-like family protein / zinc finger (HIT type) family protein (TAIR:AT3G06660.1); Has 3160 Blast hits to 2461 proteins in 330 species: Archae - 7; Bacteria - 184; Metazoa - 1235; Fungi - 390; Plantis - 202; Viruses - 26; Other Eukaryotes - 1116 (source: NCBI BLINK).	MEDLNSSREFEIGNAVKRRSHRRPRPESQPLS DGRDVSSLSSTPPSDDVKPGSSDENSGGDHS SKFKELSLSOCASRFPSPGVEGENPVRKSKVY DIGLELYGSSNHRKRSSEGLAPGNWRSTSRRE KESPEPOKTMALNKKADEGTGSPQDRLNDRNK VKVKLVGGVTRTIDANSKSSRSLDATVLOKSA VKDSEIDNSPDKGGKGGLOGIPWKDFSKVGF LVREESMKGKGGKNSVQKGDSDSARSKRS SKKRAIDGDDDDDEIRYLEKVNKAKYFTGGE DDEESSRKHRRLSKFSNGTIPQNSEVASSRFLG KDGKRRSPENEFEGEEELSGDGELEGGKKKK ARKEPVVIVEPEKRELTTRRQKALQSGRDSAP GGSAVEFPDGLPPAPSRKQKEKLESEVDOQLKY EAAQRKMHNEKIARENEAEAIRRLGQDSNRKK REDKKKRVYEELAQEREAEDELLRASNSIRLVMGP NGTTVTFNEMLPSPFFDSKPCSYPPPRESCAGP SCTNPKYRDSKTKLPLCSLQCYKAVQGNVPTQ AATC
595	sp_027940_ssgd	-0.944634	-0.85951	ARGONAUTE 4 (AGO4); CONTAINS InterPro DOMAINs: Domain of unknown function DUF1785 (InterPro:IPR014811), Slem cell self-renewal protein Pwi (InterPro:IPR003165), Argonaute/Dicer protein, PAZ (InterPro:IPR003100), Polynucleotidyl transferase, ribonuclease H fold (InterPro:IPR012337); BEST Arabidopsis thaliana protein match is: Argonaute family protein (TAIR:AT5G21150.1).	MDKDDNRRVFRPMSRPFQGMKGRKISLCTNHF RVELTNSSGYFFHYSVSICNKDGHVVEGKIGRR IMDKIHETYLELAEKDFSYDGESSLFTWGLP KAMDFVYLDKISSKFRGNDNSINGSRSPQTGDK KROKLFQTRTFRAIKYGVKIPAGASIDMLRGK YSEEVHEVIKVDIVLQRAAKQDCLVRSQFFHD DQRNFTDIGGVFACKGLHSSFGPTQNGLTLNM DSSRTTILKPGVLDLVSQNVKNPSCIDWSKA RRMLNLRKIFPYDRESRITGLSEYVCKQRRFPL GDEGRGDESQKERVTYDYVVKHKHILRYSG DFPCINIGTKSKLYPIELCSLLPLQRYTKLSIPQ RLKFIESIROKPLDKMTMLTEAWKTMNYAEPML QFCGVSSISSTQVEGRVLPAPKLVGRGEEIIPR AGCVNLNRKMLAEPMSLSHWAVVSRFRDASI VKDRLAKFGEIMGIFMEPPICIEENGRRFRQAPPA IVERMFEQILDMLKPLRLRSLKNTKTPDFICLL PNKDSIYGPWKRTLIEFGISNOCISSKLDMDQY FNSLLKINAKLGLHSLRLAVEWNCRLVSKKPTM IFGMDVYHGHYGVPSIASVSGRWCWPFSLARYRA SVRFSQPKTEMLDLSFKPPGNAEKDEGMIRELL VDFYSSGKVKPFSQIIFRDRGVSESEFKQV.LNVEF HQILEACKRLDSDWNPFTIIAOKRHHTKFFOFG QPHGNVPPGTVDVSKVCHPHYDYFVYMCPHAAFS GTARPIHYHLLDEIGPDSDELQELVHLSLYVYQKC TKAISVTPRIYAHLAASKTSHFLOYATSEIFSGHT EAGVSDLPKHLKDVCKSMFFCC

10

20

30

40

【表 1 4 - 2 1 5】

	A	B	C	D	E	F
1		分析に用いたペアトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
596	sp_027960_taqj	#N/A	-0.836131	-0.852306	Galactose mutarotase-like superfamily protein; FUNCTIONS IN: isomerase activity, carbohydrate binding, aldose 1-epimerase activity, catalytic activity; INVOLVED IN: galactose metabolic process, carbohydrate metabolic process; LOCATED IN: cellular component unknown; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Glycoside hydrolase-type carbohydrate-binding (InterPro:IPR011013), Aldose 1-epimerase (InterPro:IPR008183), Glycoside hydrolase-type carbohydrate-binding, subgroup (InterPro:IPR014718); BEST Arabidopsis thaliana protein match is: Galactose mutarotase-like superfamily protein (TAIR:AT3G61610.1); Has 1923 Blast hits to 1920 proteins in 763 species: Archaee - 0; Bacteria - 1269; Metazoa - 39; Fungi - 135; Plants - 272; Viruses - 0; Other Eukaryotes - 208 (source: NCBI BLINK).	SSTLKALMALTRLFFAGFVAVSLPRFLDSLPRMMMLLFTSFKLMYLYGAHVTSWKNEHKQELLFVSNKAIFKPPKPIRGGISICFFQFDTLGYVGAHGAYARNRVWSDPNPPLPMTSPTNSYVDLLKPKSEEDLKIWPHSYEFFRLISLGFPGSELLTSRVNRNIDSKPFSFTFAYHTYFSVDVSEVRVEGLETLDYLDNNTKRRERFTEQGDATITEGEVDKVVYLSVPSKIALDHERKRTYVIRKDGLLDAVVWNPWDKKAKALADFGDNEYKHMLCVEAAAIEPITLKPCEEWTGREGISTVPSYYCSGQLDPQKVLMLGN
597	sp_028180_arxc	-0.933476	-0.885677	-0.923612	3'-5' exonuclease domain-containing protein / K homology domain-containing protein / KH domain-containing protein; FUNCTIONS IN: 3'-5' exonuclease activity, RNA binding, nucleic acid binding, and nucleic acid metabolic process; LOCATED IN: intracellular; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: K Homology (InterPro:IPR004087), K Homology, type 1, subgroup (InterPro:IPR018111), K Homology, type 1 (InterPro:IPR004088), Polynucleotidyl transferase, ribonuclease H fold (InterPro:IPR012337), 3&apos;-5&apos;; exonuclease (InterPro:IPR002562); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT2G25920.1); Has 1067 Blast hits to 1067 proteins in 347 species: Archaee - 0; Bacteria - 413; Metazoa - 196; Fungi - 59; Plants - 125; Viruses - 0; Other Eukaryotes - 274 (source: NCBI BLINK).	MADDPSSHHHVLPLPSNLGGLVSDDAPLVIPIHVTEASQLPVEFLPCPETKLVIGFDCEGVDLCRYGTLICIMQLAFPFDALYVDAIEGGDMLEACKPALESRYITKVIHCKRDRSEALYFQFGIKLQNVMDTQIAYSLEEQEGRRPVPDDYISFVALLADPRVCGGISYLEKEEVRVLLRQDPKFWKYRPLSELMIFRAAADVYRFLLYYHKMMEKLNASHLWNLALRGELYCRFCFISDDNLTWPALPPIPDLPVDTQAPPEEILSVLDVPPGKMGVRIGRKGASILSIKESCNAEILIGGSKGPPDKVFIIGRVKQVRAEAMIFGRMLDM

10

20

30

40

【 1 2 1 5】

【表 1 4 - 2 1 6】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 ATBET12; CONTAINS InterPro DOMAINs: Target SNARE coiled-coil domain (InterPro:IPR000727); BEST Arabidopsis thaliana protein match is: BET1P/SFT1P-like protein 14A (TAIR:AT3G59170.1); Has 338 Blast hits to 337 proteins in 95 species: Archae - 0; Bacteria - 0; Metazoa - 144; Fungi - 35; Plants - 128; Viruses - 0; Other Eukaryotes - 31 (source: NCBI BLINK).	aa seq
598	sp_028670_sawg	#N/A	-0.947435	FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: root; CONTAINS InterPro DOMAINs: Protein of unknown function DUF642 (InterPro:IPR006946); Galactose-binding domain-like (InterPro:IPR008979); BEST Arabidopsis thaliana protein match is: Protein of unknown function; DUF642 (TAIR:AT2G41800.1); Has 306 Blast hits to 298 proteins in 21 species: Archae - 0; Bacteria - 4; Metazoa - 0; Fungi - 0; Plants - 302; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).	MSYRRETRASKIGLEFDGYNLSLEEGGLRASTSYSS DISEHENDKALDTLKDRVFFLQKLSGDIQDEVLNH NRVLDRMGNEMDASRGILSGTRDRFKMWERKS GRAMCKLYGFLVAFFLLYYIIRTLLYQSEETAKEW LAIEAEEGSLDIPCEYEVETL
599	sp_028720_fiuc	-0.924039	-0.943612	RNA-binding KH domain-containing protein; FUNCTIONS IN: RNA binding, nucleic acid binding; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: K Homology, type 1, subgroup (InterPro:IPR018111); K Homology (InterPro:IPR004087); K Homology, type 1 (InterPro:IPR004088); BEST Arabidopsis thaliana protein match is: RNA-binding KH domain-containing protein (TAIR:AT5G15270.1); Has 6723 Blast hits to 3876 proteins in 329 species: Archae - 3; Bacteria - 266; Metazoa - 4280; Fungi - 575; Plants - 1237; Viruses - 20; Other Eukaryotes - 342 (source: NCBI BLINK).	MASFNYLCVILLMMINSRYAAQLAPVSVISTD GLLPNGNEQGNPNSDLKKITVYKYSLPLKWKVK GLVRYVSGGHKRVGFTFNVTRGVHAVSLGQDITF IRVYGSNTADAFAPSKPLYLSDKADTFAWGFQT AASTAYLFSNPTGHPKYPACGPLIQAVAIKELPP PPTTTIDNLIQNGNFEIGPYTLENFTTGYVYHPYT LNPVSPFGWIIESRYTIRYINSKDFFIPSGSSAVE LLVGGENAIAQVIGITPKMYNLYTYIGDANNG CIGMEVQASADKETIKVYEQKGGVTHASF RFQAVSLSFTRISFWSVFYHNTYVDDPNRFCCGPV LDNRYVVLDP
600	sp_028900_izk	#N/A	-0.862435		MSLSLTPAKRQHRSHENSGKFKQKASLUNST KQPRKTSSTGGVFRVLCPSRRIESVIGKENNILS QHETGAKYRIDDSIPGDEKLFIMGLDKEAES SGOHKNDNEDTKSSKEQGSSEHDHNVNDADKE TPTNAEDLSKKNSSVQNALLLVFNQFYEGERE TEEPNEESNAAVTFRLLVLADQVGCLLGKGG SIIQMSAESGAQIRVLPDRKLPQCASSTDEVVQI TGLADAVKKAIRTVSNQILENALHDNDSFPSNTP GRTSQSHNHVPRPESYFLHSLPBGSPGGGFR DSADFFHSGVPSLFPKFHENFMFGRMDDVPDVTF RLLCPDERVGGVIGKGAIRTLQHETGCEIKVLE GFSGDGEERMIIVSGPAHPDDDRISAVQDVALRVQN RVAKAVPDSKEKGMMAKLLVFSNQIGCLLKGKG SIIGEMRKYTGAYRILGKDOMPKSASANEVWQI NGEPEVVDALLOITRLOHFFRDAFPVSNVHPS NLAFSDRAPPFPSPFMGREVEPRGFMFKLGPSE RNTDYIGLPPHGGFFHPDEHPFLHDIHRHGAP NLISERSWHFQNLIEGGPMLLDHGGPPPHRMS GFGGGSQLAVNTTVDVVVPSVIVPIIYGEDGGC LRQIRQISDAKITITDPKOGAPETVIISGTPEQTHA AGSLQIAQFVMLENESG

10

20

30

40

【 1 2 1 6】

【表 1 4 - 2 1 7】

	A	B	C	D	E	F
1.		分析に用いたペイトとの相関係数		sp_107670_ptqx (SOAP2-like)	注釈	aa seq
601 sp_029530_dujr		#N/A	-0.878437	#N/A	<p>FUNCTIONS IN: molecular_function unknown;  INVOLVED IN: biological_process unknown;  LOCATED IN: chloroplast; CONTAINS InierPro  DOMAIN/s: Defecis-in-morphology protein 1-like,  mitochondrial (InierPro:IPRO19190); BEST  Arabidopsis thaliana, protein match is: unknown  protein (TAIR:AT3G57110.1); Has 1807 Blast hits to  1807 proteins in 277 species: Archae - 0; Bacteria - 0;  Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0;  Other Eukaryotes - 339 (source: NCBI BLINK).</p>	MFOGATEDGKTMVYKYGPWLGASPHKFRSSRLS FSEREKEKMLDRLNTHAPRVARHGGFNDPNA IRLEPPGVAPKLSFSSPRSSRPTSPPCARESMLV AVKMGSQLVLRPRELNAGNVNLSNVDVGV LKKGMEVTEGHKKTSGVDVADAGKPCQCGMM PVVQSGKEVTSSELIKEATITVDKGSFVGTGPK PRKWNKLRKRGDTSCSVDGVKEDMSGSPDYCIT SSEIQENDAEIKIPVEIVSDEEMAFIEAASFASRR SSLSVLCSTSFQSIASSTRSFSSIVGNDVEDLG KIQSPQRKKNKVSLSLKRFRKNRGLSVTLTST EWCEKQMEFVLTIGRPEKTNAMKAGITRHAALLEE EVSKKVEVRVQTLIEDIWAIVKFINFICANQLLFDG LTREIPIVGFVQGVWVWVGVDELRLMPVTGNRKL TLVDTKTRVQRLPTEPQRRNGRLQLMFKRLL DNMISNFPSEOFYVHFSDPHNLSDEVKMNAA DSGFFAETLGEVRYVSNVCSVLPVDDQLLRY ELQADNSLIGEDEFNYDEDWLKAQMHSCLEFWQ GEREAKYYPAAEKWKCCQYCKFATVCPQTDTE SQSSVQGS

10

20

30

40

【表 1 4 - 2 1 8】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたペアトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kriqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq WATERSGGAGGRFRFRKPKYKNSQATPYDRPTAL RPTQLVGGGEREKNSWLSKLVDPASRLTAGA HKIFSSVFRKRLLPPPPPPQHSPPGERTSEIRDVQP EITTKVSSSGTKEPLVYKGNDSGGVSESSGVAD LEYLQKQKTFTRSEIDHLSALLSRVTDNLAGEK GRSELDTSDDKVAHSRREILTNLPAIENGGERVL PAVISTPVRSKVVVEDVVASPAELAKAYMGSRPT KASSVLGPKSDMAREDL SAPYNPFKQKSPAV SVIQRPTSEIGASANGYMTPRYGRSAIYSMGRT PYSRHPTDIRKSAEMAFDTHGASSSSSHNITTE TQQLSGSKKQVLRSSALESDIGFVGGPIRRIQ KPNLLHHKILSLPASGVPHIKHGYGALTEATEVPS SIDWSSNHRSSKALAEANGVHSTSNGVVPSQSI ETAQIFQHLDFSSKGGKSPKKEIMGGSSAAK MTTDMREGLAHKSMSTESKVLQTIQDSRKEKG LMDSSNAHDHSLKTSRVEGGTQRIVIPSVEDG AEAGSSKALEYMNHRSSFDQQQKRSFQMSA HEDYLELDDVDVHTNGVASAPMIKGVKCNLAESS AAFSDNPOSAAVSAVKPIAPEIKPSLNSLGSNKEE SLPDGSAQAFSGFTIPTGLTSSSTTSKQAAV SSQFTAEINKFSSNSAEVVPISFYFSSVVESTAA KDIKPELSSISAEANKADELDTAKGNLHFGDV SKKMEGSSVSSGKSLFFETQPNGATLTNVPST VSVSRVTSPPVPISSSNDQRLAPVSVSSPSSLSI SNGNAVSTLSDTSAGTVINTATTGGMPIFGSNTS QFSTSIAPAVPFGSNLQFQSPSPVFTFAAATSAS ESVNSGAKLSDSAFCNATSGQFSAKPTETASV GNINIFGFSASTLKSTSETPPVLOAAIFSSQASS AVASTTSIAGLSSQSAFSSASHVFGFSAPSSAYL STGANPFTTGSAGSASSVFLSSVTAPVSSSNIT SNIFGLQSSKPAVTGSTFGASGFGFLKASEASA TNSNSTPIVFGATSSSSAAKNSTPLVFGAASSSA
602	sp_029560_sait	#N/A	-0.871825	#N/A	BEST Arabidopsis thaliana protein match is: nucleoporin-related (TAIR:AT5G20200.1); Has 61042 Blast hits to 31782 proteins in 2093 species: Archae - 202; Bacteria - 16480; Metazoa - 16017; Fungi - 12552; Plants - 1653; Viruses - 629; Other Eukaryotes - 13509 (source: NCBI BLINK).	

10

20

30

40

【表 1 4 - 2 1 9】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたペアとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
603	sp_029850_esmd	-0.930492	-0.932716	-0.930118	maternal effect embryo arrest 12 (MEE12); CONTAINS InterPro DOMAINs: Transcription initiation factor Rim7 (InterPro:IPR021752); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G01335.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLink).	MAETDPSKIMCSNDYLGLYVDDDDGRYYCTNCN SOAVDIIDQANADEFLDKGNNGASINYNAANKRSA PLYSADPDRDYLGPTEPADFGKTPKGYKDE YYNTIRMRHILGVQLMIQFQCEALVEKGVPLIC GIAASWLRFRVATTRVDDDDWADKTVLDSETVSP SVFNFRFMIFQGEKRYGVGKHCINEPRNMHNER MMWWSRS_LKKTIPLSCSLAVSYLACHIAKEPVL TDIKWSIEGKLPYFGAFPKEIAESIEIEKYDHHVSI CPLKTSTMFRIPIQLPLHKLLEMAASIAQSVGLNLP PVNFIYRIAHRYLKLQSLPVEKILPYACRIQEWLMP PDLWLSASEEKLPSSRVCSVMSILVVAIRILYNLNGF GKWESTLSAPTFFSSRDDDTASSGKSNANKGAE KASKNLGGESGANLSSVHKHFEFDSTELLRLERR YDLSDKFEYSKDLPTYLYQYKDVVFAVGEPLK DPTFIEELWDFYEKNREDIKTSEAFETESNQKRS RDNDKIFSTPSKDNKKHRDEGITSSSSADSGFQC IYNSPKWDQAGADNTQQSSQADQAPATVEKGEA SLEDIALKYLKIDMEEHRCFYIPPRVEIKRRDHLHY SRKKADGSLAFVAHADYYILLRACAKVAQVDIRIM HVSVLSLERRLSWIEKWIDNSFLEKCSSTGSGVESE DEDEGLEDEYVNLNLSLNI

10

20

30

40

【表 1 4 - 2 2 0】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2				Regulator of chromosome condensation (RCC1) family with FYVE zinc finger domain; FUNCTIONS IN: chromatin binding, zinc ion binding, Ran GTPase binding; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINS: Regulator of chromosome condensation, RCC1 (InterPro:IPR000408), Disease resistance/zinc finger/chromosome condensation-like region (InterPro:IPR013591), Pleckstrin homology (InterPro:IPR01849), Zinc finger, FYVE-type (InterPro:IPR000306), Regulator of chromosome condensation/beta-lactamase-inhibitor protein II (InterPro:IPR009091), Zinc finger, FYVE-related (InterPro:IPR017455), Pleckstrin homology-type (InterPro:IPR011993), Zinc finger, FYVE/PHD-type (InterPro:IPR011011); BEST Arabidopsis thaliana protein match is: Regulator of chromosome condensation (RCC1) family with FYVE zinc finger domain (TAIR:AT5G12350.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	WTTDLTRNGPYVERDIEQATVLRKKGAYLKYTRRR CKPKFCPRLSNDESVLIWFSGKEEKHLKLSHVS RIISGRTPIQRYPPEKEYQSFLYNDRSLDLI CKDKDEAEVWFGKALISRSHQKWRTESRSD SIPSEANSRPTYTRRSPHLSPFASGDGLKQDSS DHLRLQSPYGGPPKNALDTAFSDVMFTAPKGF FPSDSATMSVHSISSGSDSVYNHMKGMPDTR VLSASVSSSQSGSLDEGDALGDVFIWGEGTG DGVLGGVRRLLGSCFNKIDSLPKALESVAVLD VONIACGGRHAALVTKQGEIFSWGEEITGGRLLG GYDADVSHPKLIDALNTNIELVACGEHHTCAVTL SGELYTWGDGAFNFGLLGHGNDVSHWVVKRVN GPLEGHVSYSICGPWHTAVVTSAGQLTFDGDGT FGVLGHGDRKSVSMPREVDLSKGLRTVRAACGV WHTAAVVEVMVGGSSNCSSGKFTWGDGDK GRLGHGDKAKLVPTCVAAALVEPTFCQVACGHSI TVALLTSGHYVTMGSPYVYQVGNPHSDGKLPTR VEGKILKCFVEEIIACGAYHVAALTRTEVYTWGK GANGLRHGDTEDRNGPTLVGALKDKQVRSIAC GTNFTAALCLHKWVSGVDQSMCGCRLPFNFKR KRHNCYNCGLLFCHSCSSKSKFASMAPNPNKP YRVCDNCLAKLRKAVETGASTRSSHARRGIWVQ GSESENEREDKLDGSRGQLARYSSWDSSEGRS KKNKKYEYDSSRVSPIPINGGQQWGANISRSFNP MFGSSKKFFSASVPSGSRVSRATSPISRPSPPR STTPPTLSGLMSPKVTVDDAKRANDGLSQEVIR LRAQVETLTRKAQVQVELEERTNKQLKEAIAIG EETAKCKAAKEVIKSLTAQLKEMAEERLPVGTATRN NKSPSFPLSSSTPTFADISNSEEIOfNDQIIPSLD SFGFNGLVLSNGSATPSSRRSSVQSRVNSEATRIN GNRTGEPRPENEWEQDEPGVYITFVSLPGG IKDLKRVRFRRKRFSEKQAEQWVAENRARIYEH
604	sp_031040_frad	-0.85:0.188	-0.34:0.138	#N/A	

10

20

30

40

【 1 2 2 0】

【表 1 4 - 2 2 1】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたペプチドの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq WVNSTGSGFVNSSEWVWYKTEFSSPSTINLGSSEI PTFSSISGLSKPRLMKSRKSNQAKSMGSSSETRRI DPGFNFRPVLNDNLAGVSSRLESEDDNGAFMNS SNSSKHLNLKVTSKVEKVGDELDIEIRKLRINQNI ARQSKGVSNDLPSLELPDKIKNLNLGGSGRENSR VSGVVVDNVSSGFGTSMESSELPIDLERLNLKDTG EMGNPNVVGASGNSFVFGSSTRIMKSDTPSI ASTLPNRIKLNLIADPKNSGTTKEMINONNRT ANRTSEGEVDTMILDEMAKLEILGDSQSSDHPR TEQVGSMSKEDSRIPCQPKSRMQINEHDEHFIGP QVAETQSDKGFYDASTSSFSAGVHFQPVRH EFOAPCTDGAEKKEGFTSSKQDGLRIPHVELTP DSKGNLFTGSKVQTFSTKGLVKDAKSKGRKA KVKEPKSVLLSFANNFTSKENMMQPEPSESYSY MDVSPYRETEYDEQLSREHSLTSENQTSLSGSQT SMVSNDAIDEDLIRATESMDINESNAVKHDEADD QVQVDCVSAVESFMSASENLDENGENEVESN SEECGYFAASFASHGESSFTFAASSNQDQAS AAMRFNKIKIPMKGGSDYIPTRASVLSDAQFSPI ARVSVLPSQVVQSDVLPRTSSGPARAAQEAACEK AMEGNVIGDAKKVTSLPRTSSGPARAAQEAACEK WRLRGNQAYANGDLEKAEYYSQGLSCIPQHET SPDCLRALVLCYSNRRAAARMSRGRMREALRDCL VAAEIDPKFRVQLRAANCYLALGEIGDASLHFKK LLQSTDVCVERKIIIEASDGLQKQAKASEFMDRS SGLLQRTLVDAESALGVIAEALISPYSEKLLSEK AEALLMLRYEEVIQLCGQTLEAKNSP_LLSAD GQSYLLEASEQLLKSYSFRWRCCDMFKSNFYL GKLEEAVEFLGKQECWMSVMVAKFGNETLESFIP QLSTARELINKSAGNGAFAGKHKKEAVEHYTA ALSFNVESRPFSAVCFANRAAAYQALGOITDAIA DCSLAIALDENY_LKAISRRATLFEIMIRDYEQAAKD MLKICDEAAXALREMHRLKSRKRWTMFKGSDLLM CGI
605	sp_031110_aqmg	#N/A	#N/A	Heat shock protein DnaJ with tetrairicopeptide repeat; FUNCTIONS IN: heat shock protein binding, binding; INVOLVED IN: protein folding; LOCATED IN: chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Tetrairicopeptide TPR-1 (InterPro:IPR001440), Tetrairicopeptide-like helical (InterPro:IPR011990), Heat shock protein DnaJ, N-terminal (InterPro:IPR01623), Tetrairicopeptide repeat-containing (InterPro:IPR013026), Tetrairicopeptide repeat (InterPro:IPR019734), Heat shock protein DnaJ, conserved site (InterPro:IPR018253); BEST Arabidopsis thaliana protein match is: Heat shock protein DnaJ with tetrairicopeptide repeat (TAIR:AT2G41520.1); Has 27844 Blast hits to 24585 proteins in 3132 species: Archae - 222; Bacteria - 9738; Metazoa - 5941; Fungi - 3019; Plants - 3250; Viruses - 11; Other Eukaryotes - 5663 (source: NCBI BLINK).	#N/A
606	sp_031200_qoqj	#N/A	#N/A	uncharacterized protein	

10

20

30

40

【 1 2 2 1】

【表 1 4 - 2 2 2】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたバइटとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
607	sp_031440_awod	-0.976364	-0.881136	HIS triad family protein 3 (HIT3); FUNCTIONS IN: protein kinase C binding, adenylylsulfatase activity, zinc ion binding; INVOLVED IN: sulfur metabolic process, purine ribonucleotide metabolic process; LOCATED IN: peroxisome, plasma membrane; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Histidine triad-like motif (InterPro:IPR011146), Histidine triad, conserved site (InterPro:IPR019808), Histidine triad (HIT) protein (InterPro:IPR001310), Histidine triad motif (InterPro:IPR011151); BEST Arabidopsis thaliana protein match is: HISTIDINE TRIAD NUCLEOTIDE-BINDING 2 (TAIR:AT1G31160.1); Has 7957 Blast hits to 7956 proteins in 2592 species: Archae - 148; Bacteria - 4849; Metazoa - 343; Fungi - 139; Plants - 127; Viruses - 0; Other Eukaryotes - 2351 (source: NCBI BLINK).	MASEKEAALAAASAPADSPITFDKINKEIPANIVF EDEQVLAFRDINPOAPTHIIPKVKDGLTGLSQAE ERHYEILGRLLHTAKLVAKOEGLEDGFRVINDGP KGCOSVYHIVHLLGGROMNWPAG
608	sp_031490_gamj	-0.849352	#N/A	uncharacterized protein	MFYSVPYDMSLVNGQRDCLCSRPAASYGDDQAA QHHELALPTRD
609	sp_032880_gxac	-0.975230	-0.990071	uncharacterized protein	MTKQTAVGRVVAAGVAQCKTGWERETWLLTKK CRGCGKSPR
610	sp_033170_gmxx	-0.989226	-0.852767	ATUBC2-1; CONTAINS InterPro DOMAIN/s: Ubiquitin-conjugating enzyme/RWD-like (InterPro:IPR016135), Ubiquitin-conjugating enzyme, E2 (InterPro:IPR000608); BEST Arabidopsis thaliana protein match is: ubiquitin-conjugating enzyme 18 (TAIR:AT5G42990.1); Has 9046 Blast hits to 9044 proteins in 385 species: Archae - 0; Bacteria - 0; Metazoa - 3956; Fungi - 2002; Plants - 1701; Viruses - 20; Other Eukaryotes - 1367 (source: NCBI BLINK).	MTSSSTQSRKV_SKIAGNRLQKELAEWQLNPPP GLKHKIDNLRWVIDVSGAPGTLYANETYQLQV DFPEHYPMEAQPVIFYPAPLPHIYSNGHICLDI LYDSWSPAMTVNSICISILSMLSSSTSKQRPPDN DRYVKNCRNGRSPKETRWWWFHDDDKA

10

20

30

40

【表 1 4 - 2 2 3】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
611	sp_033250_jpic	-0.592946	-0.975135	phytochrome-associated protein 2 (PAP2); FUNCTIONS IN: sequence-specific DNA binding transcription factor activity; INVOLVED IN: response to auxin stimulus, regulation of translation; LOCATED IN: intracellular, nucleus; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINS: Aux/IAA-ARF- dimerisation (InterPro:IPR011525), AUX/IAA protein (InterPro:IPR003311); BEST Arabidopsis thaliana protein match is: indoleacetic acid-induced protein 8 (TAIR:AT2G22670.4); Has 2237 Blast hits to 2234 proteins in 81 species: Archae - 0; Bacteria - 0; Metazoa - 1; Fungi - 0; Plants - 2235; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink).	MSRPFQHDYIGFSEVSSMENSVKSPNFSLKTELS LGLPGSEPERKPGSYGVTFGKDFEGKSQNGY SICSPKSSASGAKRGFSDAIDGCEKWWFSINGNS DPNLSKDGAKGNKFTCVENNNVSPSPKPVQDK QPQLSAANELASAPAAKTQVVGWPPPIRSFRKNT LASSPKNNGHTEEKSVCLYKVKVSMDDGAPYLK VDLQTYRNYMELSSALEKMFSCFITGGYCSNGH GKQDADLQNAEYLLTYEDKDGDMWMLVGDVPWR MFTENCRKLRIMKGTGAIGLAPNAAEKSNTA
612	sp_033500_ihik	-0.870285	#N/A	ACCUMULATION OF PHOTOSYSTEM ONE 4 (APO4); INVOLVED IN: biological process unknown; LOCATED IN: mitochondrion; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINS: Protein of unknown function DUF794, plant (InterPro:IPR008512); BEST Arabidopsis thaliana protein match is: Arabidopsis thaliana protein of unknown function (DUF794) (TAIR:AT5G57930.2); Has 163 Blast hits to 121 proteins in 13 species: Archaea - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 163; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink).	MALFRHRNWCKLVDNIEINEIGRRWYATNTKVD WDKLRPMIKRIENRARGYPIKSMIPVAEQVLOA RSDLIRGVSALLKVLVWVSKYCEVHVGVGSHLI RTCCGHRRLNKQDHWIRASLNDILVPVETFHL DNMVQDVIRHEQRFDFERVPAIVELCYOAGVIPH D.LGLQSLDLYSNSTVDFALLTGPDSLASDLRTIG NRTLRAWETLRHGVOKLLLVYSKVKYCYSEVH VGFSGHKARLCLGVFKYQNFNFRGNHFWKAGVDD LVPPKIVWCRRPQDPPVLLHEDRNYGHPAVV D.LCMKAGTVPVVKYLMMKAGGLRGRGSPDKVL G

10

20

30

40

【 1 2 2 3】

【表 1 4 - 2 2 4】

A	B	C	D	E	F
1					
2	分析に用いたペアとの相関係数 sp_107620_kqnh sp_107660_kiqg (SOAP1) (SOAP2)	sp_107670_ptqx (SOAP2-like)		注釈	aa seq
613 sp_033650_qdon	-0.847811	#N/A	-0.8327	BR11 suppressor 1 (BSU1)-like 1 (BSL1); FUNCTIONS IN: protein binding, protein serine/threonine phosphatase activity; LOCATED IN: plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN's: Serine/threonine protein phosphatase, BSU1 (InterPro:IPR012391), Metallophosphoesterase (InterPro:IPR004843), Galactose oxidase/kelch, beta-propeller (InterPro:IPR01043), Kelch-type beta propeller (InterPro:IPR015915), Serine/threonine-specific tetraphosphatase/bis(5-nucleosyl)- tetraphosphatase (InterPro:IPR006186); BEST Arabidopsis thaliana protein match is: Serine/threonine protein phosphatase family protein (TAIR:ATI03445.1); Has 11450 Blast hits to 9180 proteins in 548 species: Archae - 80; Bacteria - 322; Metazoa - 3883; Fungi - 1977; Plants - 2720; Viruses - 6; Other Eukaryotes - 2462 (source: NCBI BLINK).	MASKPWLQAPSYTQLESFWDSDDEDAPGPRCG HTLTAANKSQGPRLLIFGGATAIEGGASSAP GIRLAGVTNSVHSYDYLTKKWTRIRPAGDPPSPR AAHAAAAGTMMVFGGIGGIPAGHSTDDLYLDL TNDKYKWRVWVQGGGPRYGHVMDLVAQR YLVTVSGNDGKRVLSDSVWLDTAQKPYVWQRL NPEGDRPSARMYATASARSDGMFLLCGGRDST GAPLADAYGLLHRNGQEWTLAPGVSPSPRY QHAAVFYGARLHVTGGALRGGAIEGEAAVYL DTAAGVWLDRSGLVTTSKANKQGTENDPSLELM RRQRHASASVGRVRYMYGGLRGDYLDDFLVAE NTVSQDDNSLPLASERASSLTDPRINLSLTMP ATPVTGDRHEKTLPGGVSMEEKSVDKLVEASAA EVEAVRTWQAAQAAASSPPDESPLSQSSPIT DTPSDGSDTDGDVFLHPRAVVAVAKETVGNLGG MVRQLSLDQFENESRRMIPVADQSTLGGKFTRS KSPQGLLHKKVISNLLRPNWKPVPVNRKFFLDSYE VGEICHAEEQIFMHEPSVLQKAPIKVFDDLHGQ FGDLMLRFLDEYGFSTAGDITYDYLFLGDVYDR GQHSLETTLLALKIEYSPVHLIRGHNHEAADINA LFGFRLECIERMENDGIWAWTRFNQLFNYLPLA ALIEKKIIMHGIGRISQIAVEQIEKLERPITMDAG SIVLMDLLWSDPTENDSIEGLRPNARGPGLVTFG PDRVTEFKKINKLQHIRAHECVMDGFERFAQGG LITLFSATNYCGTANNAGAILVGRGLVVPKLIHP LPPPMQSPETSPERAMDDTWMQELNIQRPTPT RGRPQPDHDRNSLAYI
614 sp_034220_totz	#N/A	-0.845671	-0.832888	RING/U-box superfamily protein; FUNCTIONS IN: zinc ion binding; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN's: Zinc finger, RING- type (InterPro:IPR001841); BEST Arabidopsis thaliana protein match is: RING/U-box superfamily protein (TAIR:ATI1G17145.1).	MGNKLGKRRGVDEKYTRPQGLYQIKDQVYKKL FKLILESKLAPCYGGDECDGGLLEECPICLYYP TLNRSRCMKSICTECFLQMKNPSTRPTQCPF CKTSNYAVEYRGGKSKEERGLEQIEEQKVIKAI RMRQEMEEEEERQHKKEETSSSERDPRTENID CSTSGGVEEETAPSDSNGTQMLRQPHTRHG RDDDFDLLEDIMVMAIWLSIQENGRAEYLSLN DEVPEQYTAGDHYMSPVAPVAGSSSSPSSGGLA CAIAALAEFGQSTGESSTHYSEPVSRDLMLPSS RSSKRMKTKQRHAECSKQICSSSMISEGCEV SPALEVAEVTGYSINSDSSEGEVILPPLPPPM YRWNKRGNQIRKYSEFERASFKNLHDDRWA EDWFEFHGSEVAEGTYYASSDESNEIDQAAVL PPPPPPPLPPYVVGSGGSDHGEYLLPPP YVTDGFGQSPIPGPIVPENFEEQMLAMAVSLAE ATARTTAGGVAVWH

10

20

30

40

【 1 2 2 4】

【表 1 4 - 2 2 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
				DC1 domain-containing protein; FUNCTIONS IN: oxidoreductase activity, antioxidant activity; INVOLVED IN: response to calcium ion, pollen tube growth, pollen tube guidance; LOCATED IN: cellular component unknown; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Thioredoxin fold (InterPro:IPR012335), Alkyl hydroperoxide reductase/ Thiol specific antioxidant/ Mal allergen (InterPro:IPR017936), Thioredoxin-like fold (InterPro:IPR012336), C1-like (InterPro:IPR011424), Thioredoxin, conserved site (InterPro:IPR017937); BEST Arabidopsis thaliana protein match is: protein kinase C-like zinc finger protein (TAIR:AT4G31240.2); Has 6688 Blast hits to 3903 proteins in 794 species: Archae - 4; Bacteria - 4185; Metazoa - 634; Fungi - 4; Plants - 553; Viruses - 0; Other Eukaryotes - 1308 (source: NCBI BLINK).	MDSARKRSADSATLVKKTGDIVDDFWPLKLD PFENL.SRFQENKGCYNRFFYSFLRVSKDGEYGD DIKIMVSELPDMEIGDPDVLSSLKFKGQSNLYV RNNDNDQVELNKDTHFGKYVMVCWFHPIHCL DGDALFIQSAANICYEMFDLDDFEIVLGLKMNPK YNNORVFSHLSAFPSPSCLVVPFDDPKRSDYICC YLEVRIRCEFLVDGNSIALFSGEPDFEMGERYS ASRLCSNYSAKDLEISLPIKTRLSIINNTVVRIN EEGVQEEMSFLRLNEKVVGLYLWTEGIRCIIDE SFIDTLDVMNKGCRDILVADALQHHMQQTWEHC GCIPGQLVIVGNTEYVDLYGRDILGKYGSVAYPF TRGLINSVGLKRELTLESLLTSESNLYLKDGF TQINVASLQPRDILLYFGSEDSGLIKSLQNLVSEV KLENPNFEIVYVRRNNTSNYCNMMDIGLPLVRP FEADHSSSLGKKLFPVVPVDTFVFRFGNDGHC SVGGNNL.REDYCSVGGNSLRFDVIFMLERLIGIME DYIRYCC
615	sp_034910_oips	-0.847785	-0.833142	-0.841185	MEAAQLSSSWQFSCDLEVNFESEKAYIVYAALN VDKELOPKVKRVMISVSDGKLAHYHQVAEARFL RASFSAFVDVLRLATRTIEEFSPGMEI MENKINDDREGGVNSLSSEPNMSEALCICCAL WMEKIGSVMEGMKVRMKW
616	sp_034990_pexn	-0.945491	-0.922928	-0.937575	
617	sp_035630_neliu	-0.824825	#N/A	#N/A	MGASLSNSAEGGANGGNGPSLGDPEACIACV FMYLTPPEICNLARLNRAPFRGAASDNDVWESKLP PNYQDLILLSPERYQGFSGKDFISLLARSIFED DTKQVCVDKVTGRVCLSVSARAMSITGIEDRRY WSWLPTSESRFQVIYV.LQQVWVFEIDGVVKKFPL PPDYI.LTFRHLGKFKSKRLGRVSNFEHTHGWD IKPVRFELTTWDNQHATSECYLDDETEPDMKEN LKRGCVIEYKVGFEIVRSDSVPEVFRFSMKQIDC THSKGGLCVDSISIVPSDLRERRRRLLSN
618	sp_035650_rkws	#N/A	-0.87747	-0.868647	

10

20

30

40

【 1 2 2 5】

【表 1 4 - 2 2 6】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq WMTETVYGSKVVWVDFEYVWIDGVEYTKTSGDE LHKVTNGKSVITEVSKVYKDEEAPAGGVDMDIT KLAYLHEPGVLSNLAMPYSINEIYTYGNILIAVNP FQRLPHLYDTHMMGOYKAGFGELESPHVFAIAD ASYRAMINEGKSNLSVSGESGAGKTEITKMLMR YLAYLGGRRSGVEGRTVEQQVLESNPVLEAFGNA KTLRNNSRFKGFVEIOFDKHGKISGAAIRTYLL ERSRVQVSDPERNYHCFYLLCAAPQEVDRKYK LDSPDKFHYLNQTNCKYKLDGVDDEEYLA TRRA MDIVGISAEDEQAEIFRVVAAIHLGNVWFAKGTVEV DSSVLKDEKSRFHLAMTAEALLRCDIKSLEDALIKR VMVTPVEVITRTLDAHSAQISRDFAKTYSRLLFD WLVDKINNSIGODPNKSIIGVLDIYGFESFKLNSF EQFCINFTEKQLQHFHQVFKMEQEEYTKEEIN WSYIEFVNDQVLDLMIKPPGGVIALDEACMFP KSTHETFAQKLYQTYKHKFKPKLSRTSFTIC HYAGDVVYQADQFLDKNDKYVIAEHQDILLQASK CSFVASLFPPLPEETSSTKSKFSSIGSRFKLQOQL METLNSLSTPHYRVCYKPNLLKPAIFENVNINQL RCGGVLEAIRISCAGYPTRRTFEFLRFVGLAPE VLDDGKDDKIVCQKILDKMGLKYQIGTKVFLR AGQMAELDTRRAEVLGNAAKRIQRHGRTRYARKE FISYRNAIQIOSWCRGVLACKVHEELCROASAIR IQKNFRCYKARQSYLRVRLSAVTVQTGLRAMHA RNMFRYKKTETKAAIIQTRQRCHMAYSYYQNLQK AALVTCGGVRRRVRARKELRQLKMAARETGHLK ELKDKLEKRVEDLTWRLQFEGQLRKNLEEKGGQ EIAKLOEALQAMKQVEEANARITKEREAAKAVV EDAPPVVIKEVPMVQDTEKVESLTIENL KALLE KEKEAAEVTRKSVSDTEERNAELVKKLEGSDDTI NQLKHFVORMEEKLSNAESESQVLRQOQLEMT QKLSAESEIQDMAEKLKAESELOALRQOQALTM  MEERERVLNQKKTEQEEHEQEKDIEPKQOQLW SWGAGTEGQLGTGNLKDDELTPQLLNLNSLSLF APISILSCGGAHVIALTSGGKVLWGRGSSGQLG HDQIMPNCSFKLVALSLSFNHSHVSAGWNHSGF VSDSGCLFTCGDGSFQGLGHGDFKSYSSPIEVG FFNSKHVEQVACGMFHSVLVSKGNSGDKVMYV FGSGKRGQLGISEDKIKLVSPVVCVSGFENAEV DILANGDQSAALSADGRLTYWGRGFSSSSNVCI PQSPISAFSEFEQAAAGWNHALVLTSEGEVFMGLG GSHHGVLSEFPQKMRHNHNSDEPSLSKVCGLV GKVVVHIAAGAEHSAIVTENGQVYTWGWEHG QLGLDSTSDHTAPQLVLPEDRSLSIRVYCGSGFT FALKTHT
619	sp_035830_cyaw	#N/A	#N/A	-0.8882331	XIK; FUNCTIONS IN: motor activity; INVOLVED IN: in 11 processes; LOCATED IN: endomembrane system; myosin complex; EXPRESSED IN: 6 plant structures; EXPRESSED DURING: seedling growth; CONTAINS InterPro DOMAIN/s: Dil domain (InterPro:IPR018444), Dilute (InterPro:IPR002710), Myosin, N-terminal, SH3-like (InterPro:IPR004009), Myosin head, motor domain (InterPro:IPR001609), IQ calmodulin-binding region (InterPro:IPR000048); BEST Arabidopsis thaliana protein match is: Myosin family protein with Dil domain (TAIR:AT1G54560.1); Has 40570 Blast hits to 26875 proteins in 2480 species: Archae - 466; Bacteria - 5527; Metazoa - 21427; Fungi - 3168; Plants - 2202; Viruses - 135; Other Eukaryotes - 7645 (source: NCBI BLINK).	
620	sp_037040_wcpn	#N/A	-0.8882331	#N/A	Regulator of chromosome condensation (RCC1) family protein; CONTAINS InterPro DOMAIN/s: Regulator of chromosome condensation/beta-lactamase-inhibitor protein II (InterPro:IPR009091), Regulator of chromosome condensation, RCC1 (InterPro:IPR000408); BEST Arabidopsis thaliana protein match is: Regulator of chromosome condensation (RCC1) family with FYVE zinc finger domain (TAIR:AT5G19420.1); Has 19815 Blast hits to 6041 proteins in 472 species: Archae - 85; Bacteria - 2553; Metazoa - 7293; Fungi - 979; Plants - 2476; Viruses - 2; Other Eukaryotes - 6427 (source: NCBI BLINK).	

10

20

30

40

【表 1 4 - 2 2 7】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
621	sp_037140_nteo	-0.386322	#N/A	-0.380943	uncharacterized protein	MSQLTQEKGDVADTWLNELYQQRVYLHQIQD ETKASSKAVKPSFVRSALASKRKKEDEAKAPSK KPKLKITKVKVKEKETVEKFHDFHDPKAKELTLA VAGSKGAINVLSAFDEKVTTPAGKEREREIVSTNG GGTSETPAKQLVKAAPLGDPEVVGDKVPSFVF DRVPRQTGEASLEQKWDEEVGEPFSTKRFDNK SOAKIVEGIMNVVPKYREFVQGGDDAQLGAIQA TLVDLLIKMDGAKKWHREVQRDLKQVTKAAFV QIEEKINEQILADITIGELDKALAAKAEKLEAQ SSMEARVVELEKDLAARADVAFLSADAKSAYIK GEEDTKKATHMAWDFVT
622	sp_037170_lkrj	-0.386322	-0.847304	#N/A	C2H2 and C2HC zinc fingers superfamily protein; FUNCTIONS IN: zinc ion binding, nucleic acid binding; LOCATED IN: nucleus; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, U1-C type (InterPro:IPR013085), Zinc finger, U1-type (InterPro:IPR003604), Zinc finger, C2H2-type matrix (InterPro:IPR006930); Has 30201 Blast hits to 17322 proteins in 780 species; Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MPRYCYDYCDTYLTHDSPSVRQKHNEGYKHKAN VRIYYQQFEEQQTOSLIDQRIKEHLSGGGAFQIG AAYAARWRPGRFPPLLPAPLPIPGSLPPPPAQMMA PGIRLPVFRPPPPGPPGFPVYGAPPQQIAGQPGA PGPPGPPGAPVQANG
623	sp_037510_fxuf	-0.386322	-0.951935	-0.833447	Nucleotide-sugar transporter family protein; FUNCTIONS IN: nucleotide-sugar transmembrane transporter activity, sugar:hydrogen symporter activity; INVOLVED IN: carbohydrate transport, nucleotide- sugar transport; LOCATED IN: integral to membrane, Golgi membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Nucleotide-sugar transporter (InterPro:IPR007271), UDP/CMP-sugar transporter (InterPro:IPR021189); BEST Arabidopsis thaliana protein match is: Nucleotide-sugar transporter family protein (TAIR:AT4G35335.1); Has 30201 Blast hits to 17322 proteins in 780 species; Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	REVVTSSVLACEVAKVICAFVLIIFNGSLPKLKD WSLLRALTFSGLPAAYALONSLQISYKLNLSLT FSLNLQTKLFFTAFFGYMILRQKQSIQIQIGALFLI LAALLSICEGTSKAKSRNPEEVLFGVIVLIAS VLSGLASSLQWASQVKKHSSYLMTEIENSIFGSL CLLTSLSQSPDGTAMRKHGFFQGWTFPLTLVPVV ANALGGILVGLVTTVYAGGVRKGFVIVSALLVTALL QFICDGIPPSPYCLVALPLVITSVSIYQKYPYRVKK KEA

10

20

30

40

【 1 2 2 7】

【表 1 4 - 2 2 8】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたペアとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
624	sp_037550_dcd	-0.881733	-0.88328	#N/A	Hypoxia-responsive family protein; CONTAINS InterPro DOMAINs: Hypoxia induced protein, conserved region (InterPro:IPR007667); BEST Arabidopsis italiana protein match is: Hypoxia-responsive family protein (TAIR:AT5G27760.1); Has 144 Blast hits to 144 proteins in 38 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 45; Plants - 99; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink).	MAKESKTTMESVRSWVWFQHKLRVTGCLWLSGIN GSIAYNWSRPQMKTSVRLIHARKEELNYFCRF  MEAVVSELMASPLPGSLDRERKDVCGLSVVS PGDDIECDNMDGLLASSALLFORSRPRGRKK RVALADRGWGFTSEIKKFSYNLVEFVNHNGGL LSGRNLKGLSLNVMIKGGDGNMNHQCKAVA QLNGCVHKQEKVMKFAADCRDGVHNSAVDRSR LANWMSKEREKMKVASHLGEEDQHTETGVFS RISELIGVDKLLFGGEDGDGWGELLSWERRMD IDIQKDTHSCGVRMLLAIKDFAKGYESLRITDLEE ARRLLLTQDILSPFNLERDRIKELVNGANRKKL
625	sp_037790_xmku	#N/A	-0.885965	#N/A	uncharacterized protein	MATTDATPKFHINHELRLSSLEEQQQNLTMDF SPSHTHDGGLLFWQFMVAGSIAGCAEHMAMFPV DTIKTRMQALGSCPIKTAGVRTALRSLQSDGPA GLYRGIAMGLGAGPAHAYFYSMEFFKKSFSG GNPNHVAHAMSGVCATMSDAVLTPIMDVMKQ RLQLSGNGGAYRGVWDCVKLVKNEGFAFY ASRYTTVMNPFTAVHFATYEAAKKGLTEISPD ADDERLVVHATAGAAAGALAAVITPLDVVKTL QCQGVGCDRFSSSGSRDVLKTIKDKGYRGLM RGWAPRMIFHAPAAACWCSTYEAAKTFFQELND QKNSSSVT
626	sp_038000_ftpx	-0.987983	-0.984474	-0.983318	Mitochondrial substrate carrier family protein; FUNCTIONS IN: bincing; INVOLVED IN: transport, mitochondrial transport, transmembrane transport; LOCATED IN: mitochondrial inner membrane, membrane; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINs: Mitochondrial carrier protein (InterPro:IPR002067), Mitochondrial substrate carrier (InterPro:IPR01993), Mitochondrial substrate/solute carrier (InterPro:IPR018108); BEST Arabidopsis italiana protein match is: Mitochondrial substrate carrier family protein (TAIR:AT2G30160.1); Has 27351 Blast hits to 14015 proteins in 457 species: Archae - 0; Bacteria - 6; Metazoa - 11451; Fungi - 7735; Plants - 5292; Viruses - 0; Other Eukaryotes - 2867 (source: NCBI BLink).	

10

20

30

40

【 1 2 2 8】

【表 1 4 - 2 2 9】

	A	B	C	D	E	F
1						
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
627	sp_038170_jmqo	-0.87026	#N/A	-0.884585	<p>Prefoldin chaperone subunit family protein;            FUNCTIONS IN: unfolded protein binding; INVOLVED IN: protein folding; LOCATED IN: prefoldin complex; EXPRESSED IN: pollen tube; CONTAINS InterPro DOMAIN/s: Prefoldin alpha-like (InterPro:IPR004127), Prefoldin (InterPro:IPR009053); Has 429 Blast hits to 422 proteins in 142 species: Archae - 27; Bacteria - 8; Metazoa - 215; Fungi - 17; Plants - 37; Viruses - 4; Other Eukaryotes - 121 (source: NCBI BLINK).</p> <p>alfin-like 6 (AL6); CONTAINS InterPro DOMAIN/s: Zinc finger, PHD-type, conserved site (InterPro:IPR019786), Zinc finger, PHD-type (InterPro:IPR001965), Protein of unknown function DUF3594 (InterPro:IPR021993), Zinc finger, FYVE/PHD-type (InterPro:IPR011011), Zinc finger, PHD-finger (InterPro:IPR019787); BEST Arabidopsis thaliana protein match is: alfin-like 7 (TAIR:AT1G14510.1); Has 1820 Blast hits to 1777 proteins in 193 species: Archae - 0; Bacteria - 0; Metazoa - 798; Fungi - 350; Plants - 563; Viruses - 0; Other Eukaryotes - 109 (source: NCBI BLINK).</p>	<p>MATARGKGTVTSLSSTFPAAEEARKAADLVNHTIN            EKFSLDNLRSEFVADNTNLINLVQRLPDELHHHIQ            VPGKAAFFPGRVLRVHTNEF_VLLGEGYFADRTAK            QTVDILKRRGQTLSEHYSDSVKAVIKDKAEASFFY            RTAEAAEGLVREIREELESDDESEGEKEDAAAS            EPEVIGISSAAVLDDDEYARIMSRLEDELEKEELE            EEEEDTEDHRDVTNGEDFNNSQSEVQHHTK            VLSDSISLPNADAKSIHTYLNLPNHKDDCSRGK            DILQKFVERPMEEDNVPVWKTSPFSESRSKSEV            YVPAPKPEVGNQLSRQVKTDLSKAFATGSIERTD            NLQISPDREPINQPSNSQSPKPVSRFKMORRGV</p>
628	sp_039020_guqe	#N/A	-0.895212	#N/A		<p>MEGMPONTPRTYDEVFNDFKGRVGLIKALTTD            VDKFYQQCDPEKENLCLYGLPNETWEVNLVVEE            VPPELPEALGINFARDGMQEKDWLSLVAVHSD            SWLLAVAFYFGARFGFTKNERRKLFQMINDMPT            VFELINGNAKQPKENPGPHNSSKNKNSGVKPSR            QPENHSPKMKMPTPPKAEASGGEEDDEEQG            ATLCGACGDNVYASDEFWICDVCWEKWFHGKCV            RITPAKAEHIKQYKQPGCSTKRARY            MRNIRKLFSSFTLLAEIQWYTLPLPLYQYYK            KQOQKYTTVQITTNKGVKGLSKLCKNMMAKAT            QSAAEKAVRKLQENFDFVDADINYYTTCMHORKK            RRIAEIKL TREEMNFIDLKRYYYFOECAENRSLKEQ            LVKLLNFRDKGSG</p>
629	sp_039160_anky	#N/A	-0.866056	#N/A	uncharacterized protein	<p>MTYRKRVRATKTDAGSASKSLPSRTSSLSSEVA            SQPTKKTATQSQSPQVPSKVFPGFPFGTTTS            MVAPPCSSTAKSAQQAIAWKTPPTGSAASNSLEAS            TKNKIPPKITSIKSIQIPNIPPSKPLKYFSTSSGQIPP            NMPHPVPPSSHHTKTSAGDSQSQQKSVQEPKEN            STKSLSLREYPDWFRKNGAKMI</p>
630	sp_039300_htyn	#N/A	-0.863056	#N/A	uncharacterized protein	

10

20

30

40

【 1 2 2 9】

【表 1 4 - 2 3 0】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_ktgg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
631	sp_039570_npk	-0.856776	-0.927384	<p>Trimeric LpxA-like enzyme: FUNCTIONS IN: binding, transferase activity, translation initiation factor activity, nucleotidyltransferase activity; INVOLVED IN: biosynthetic process, regulation of translational initiation; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Trimeric LpxA-like (InterPro:IPR011004), eIF4-gamma/eIF5/eIF2-epsilon (InterPro:IPR003307), Armadillo-type fold (InterPro:IPR016024), Nucleotidyl transferase (InterPro:IPR005835); BEST Arabidopsis thaliana protein match is: Trimeric LpxA-like enzyme (TAIR:AT4G18300.1); Has 8121 Blast hits to 7961 proteins in 1548 species: Archae - 574; Bacteria - 3735; Metazoa - 1296; Fungi - 469; Plants - 387; Viruses - 0; Other Eukaryotes - 1660 (source: NCBI BLINK).</p>	<p>MVRTSRAEDADLLTHVPLQAILLADSFATKFR PITLERPKVLLPLVGIPIIDYSLWLESGVEEVF VFCCAHSKVIOHQLAKSQWLKNPDMAVTTIESH NSISAGDAMRVIYERNVHGDFVLVTGDTVSNMS LAEALQEHKDRRKKDSNAVMTMVKKSPSPIT HQSLRGTDLDFVAINPDTKQLLSYEDNANPIDKSL LSDNSSISLRNMQDCYIDICAPEVLSLFTDNFDY QHLRRDFLKGLLVDEITDYKIYTHEIHSSYAARVE NFRSYDTISKDIIHRWTTPLVDPQYFGSSAVKLG RQGYASDIDGLSSNAQIGPFTLIGKGTSVGSNS KVSNSIIGEGGK(GSNVSIEGSHWNNVTIEDGCK LKHALCDGVVMKARAVLEPGVILSKVYIGKDFV VPAYSKVSLYEQPIFQDSDEELEAYDSSNGMAEI SSISASTGMPDGLVSRAAEASEVGDAGAGYIY SVFEGREEEWFHRSVAPISEDKLAELQAADDVLD VPHVENHQHSEELVAPSIDPENLTDNDIDEFEKE VEATFQRVQENIKDDNVVLEVNSLRLSYNMAAA DCAGALFYVMKLAIVETPHSSISELLHNVTLIRIK WAKLLKYYLSSIDEIEIEVILKFEEMCTESTKELSP FTKVLQFLYEDVDLIQEEAALSWESEKMGAEVD RVFLKQAEPIQWLKTAQEEEDSEEEEE</p>
632	sp_039600_nsgq	-0.856331	#N/A	<p>zinc finger (C2H2 type) family protein; FUNCTIONS IN: sequence-specific DNA binding, transcription factor activity; INVOLVED IN: regulation of transcription; LOCATED IN: intracellular; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, C2H2-like (InterPro:IPR015880); BEST Arabidopsis thaliana protein match is: zinc finger (C2H2 type) family protein (TAIR:AT5G16470.1); Has 86 Blast hits to 86 proteins in 31 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 77; Viruses - 0; Other Eukaryotes - 9 (source: NCBI BLINK).</p>	<p>MTGKAKPKKHTAKELAKLDAVMTTNRGGGKAGV ADRTGAVKGGHAKMECPCKVTAPDVKSMKIH EAKHPKIPFDENIKSNLHASTPAAAPSTSVADPNK PKPGIRGSLKK</p>

10

20

30

40

【表 1 4 - 2 3 1】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh sp_107660_klqg (SOAP1)	sp_107670_ptqx (SOAP2-like)		注釈 Peroxidase superfamily protein; FUNCTIONS IN: peroxidase activity, heme binding; INVOLVED IN: oxidation reduction, response to oxidative stress; LOCATED IN: endomembrane system; CONTAINS InterPro DOMAIN/s: Haem peroxidase (InterPro:IPR010255), Plant peroxidase (InterPro:IPR000823), Peroxidases heme-ligand binding site (InterPro:IPR019793), Haem peroxidase, plant/fungal/bacterial (InterPro:IPR02016), Peroxidase, active site (InterPro:IPR019794); BEST Arabidopsis thaliana protein match is: Peroxidase superfamily protein (TAIR:AT5G66390.1); Has 4558 Blast hits to 4529 proteins in 286 species: Archae - 0; Bacteria - 0; Metazoa - 2; Fungi - 246; Plants - 4242; Viruses - 0; Other Eukaryotes - 68 (source: NCBI BLINK).	aa seq
633	sp_039670_occn	-0.253053	-0.271333	#N/A		MGLQRIILQGGCDASILLDSTLANPSEKENLANGQ SLRGFEVIDEAKAELEFNCPGTVSCADILAFARD SVLELGGINYGVPAGRRDGKVSIKDEPTQNLPS INSTVELOQMFAFKGLNLQEMVALSGAHSIGRV TRCSVFSHRLHTFNATHVQDPSMDLKFAGEMKN NCPSTTSGGEDEVKVPDLDGFTFSRLDNRYYRNL ENGRGLLSDDQALMDSPLTAKMVRFFAKYGRV WMPEFVNAMTRMGSIIEVLTGSSQGEIRRNCRFVN L

10

20

30

40

【 1 2 3 1】

【表 1 4 - 2 3 2】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MPPFSRWYVCCCLLSTLCLLFLFTFSVGVGVWYTE RGNLDFSLGNSSALIPDSSEDRHGFGETSEV LSLVDLVSCEDLEGVGSFNITCLLNSNVHLNSSI YAFGINLELPHVLIIVCPIEGCSISFNLFNVKVG PNAAIAGSVISSHSLTIGHNSSVNTTSLGKPP QTSSTPFYEGAGGGHGRGASCLNEQLDAW GGDYVSWNSVSKPWSYSGKSGGRSAQNPFGG NGGGRVKLVKDELYLNGSVTAEGDGGVGGGG GSGGSNITAFKLGKLTSAAGGRGWGGGGGG RISLHCYSVQDVKTVHGGLSIGCPENAGAAATY FDATLLSLRVNDNLSTETETPLDFSTPLWSK VFIENSAKVLVPLWTRVQLRQGISVLSRSRINFG LMEHPISEFELVAEELMSDSQIQVYGALRVNVK MLLMWNSEIQIDGGEAVVTSVLEVRNLAVLREH SVSSNTNALYGGGLLKLTDGDCIKSORLLLSQ FYMITLGRGSLQAPLDDDDASKMTWTKSCENQT CPMDLNPDDCHINSTTSFSLQICRVEDILVDGH VKGSIHFHRTIVVADGMITASELGCRRKGG MGRFLKAGSGAGHGGKGGSGMFMGTISEGGS RYGNANLPCELGSGSEGNESYGHVAGGGMIVI GATQWPLFRLLDLYGSLRADQSHSQKIKNRDGT LIGLGGGGGGTILLFQKLALAEKSSLSAAGK GGPLGGGGGGGGRIHFHW'SNINTGDEYPIAVV NGTYVSGGGGAGDSGGYEGEETITGKCEPKGLY GTFCESEPIGTFKVDGSDQFCAPCPQELPH RAEFYVGGVNEASCFFKICISDKYRMPHCFPL EELMYPVGGPWTFSLLLGCFVILGILLNLRKIM VSGSSVCHVYTFEQHNFEDYQYPYLLSSEVKG PNKAEDTOSHVRYMYFAGNPTTFREPWHLPYSP DEIVEVYEDAFNRFIDEINSAIAYDWWEGSVHSIL LVLAYPCAWSKQWRRRIKIQRLQEVVKSSEYDH SLRCSRARALYKGMKVGVTGTPDLMWAYIDFFLG
634	sp_039820_kppp	#N/A	#N/A	31.8537339	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane_system; EXPRESSED IN: 23 plant_structures; EXPRESSED DURING: 13 growth_stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G11700.2); Has 1807 Blast hits to 1807 proteins in 277 species: Archaea - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	

10

20

30

40

【表 1 4 - 2 3 3】

A	B	C	D	E	F
1	分析に用いたペイトとの相関係数	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	sp_107620_kqnh (SOAP1)			LOW-LEVEL BETA-AMYLASE 1 (LBAT1); FUNCTIONS IN: in 6 functions; INVOLVED IN: sugar mediated signaling pathway, RNA interference, nuclear-transcribed mRNA catabolic process, nonsense-mediated decay, response to sucrose stimulus; LOCATED IN: plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Restriction endonuclease, type I, R subunit/Type III, Res subunit (InterPro:IPR006935), Helicase, superfamily 1/2, ATP-binding domain (InterPro:IPR014021), RNA helicase UPP1, UPP2-interacting domain (InterPro:IPR018999); BEST Arabidopsis thaliana protein match is: DNA-binding protein, putative (TAIR:AT2G03270.1); Has 9640 Blast hits to 8330 proteins in 1330 species: Archaea - 182; Bacteria - 2840; Metazoa - 3001; Fungi - 1369; Plants - 807; Viruses - 148; Other Eukaryotes - 1293 (source: NCBI BLINK).	MDSVNNLCTETASQPTGTGNDYDTTLEPNTQAG GDDDDVNFAYPQFNELSQIFASASPSSSA WPTPSDSIVAPEIKESHQPPESHSSSSGGPTKG RGGGGGGGGVAVSVDLSLAAGMSGLNFEETAG DEDFVDQFGKGVGDFSEHACRYCGVQNPACV VFCNVPSCKRWFCNSRGTSGSHVNHILVRAKH KEVCLHKDPSLGETILECYNCGCRNVLLGFISAK TESVAVLLCREPGLSVNALKDMNWDLSQWGPLI DDRCFLPWLKVPSEQEQLRARQISAQINKIEE LWKSNDASLEDELEKPGVDEPQVWLKYEDAF QYQNVFAPLIKLEADYDKMMKESQSKDNLTYRW DIGLNKRYAYFVFKEDNELRLVPGDELRLRYS GDAAPAWQSVGHVVKLTAQEEVALELRASQGV PVDITHGFSYDFVWVKSTSFDRMQGAMKTFAYDE TSVSGYIYHLLGHEVEMQVVRNTLPRRFGAPGL PELNASQVFAVKSVLQKPIQLIQGPPGTGKTVTS AAIYHMAKGGGGQVLVCAPSNVAVDQLAEKISA TGLKVVFLCAKREAVSSPVEHLTHYQVRLDIT SEKSELYKLOQLKDEQGELESSEDEKYYALKRAT EREIAQSADVICCTCVGAGDPRLSNFRFRQVLID ESTQATEPECLPLVLAQKQVVLVGDHCQLGPVI MCKKAARAGLAQSLFERLLLVGKPIRLQVQYRM HPALSEFPNSFYEGTLQNGVTVNERQSTGIDFP WPVYPRMFFYVQMGQEEISASGTSYLNRTAA NVEKIVTSLRSGVVPVSGIVITPYEGQRAYVNY MSRNGALRQQLYKEIEVASVDSFQGREKDYIILS CVRSNEHQGIGFLNDRRLNVALTRARYGIVILG NPKVLKQPLWNLGTHYKEHECLVEGFLSNLK QSMVQFKPKKIYNDRLFLYGAAGVYVPSDNTV PGSTTGASGDRRSGRGRYMPFGAPNGTHKPA VHPSGFMTRVPIPPYHGAPPVQPYAIPSRGAVH GPVGAVPQVPPQGNFRGFGAGRGSAPIGSHLP
635	sp_039840_utkd	-0.85369	-0.852123		

10

20

30

40

【表 1 4 - 2 3 4】

A	B	C	D	E	F
1	分析に用いたペイトとの相関係数 sp_107620_kqnh (SOAP1)		sp_107670_ptqx (SOAP2-like)		
2				<p>注釈</p> <p>RNA polymerase II large subunit (NRPB1);            FUNCTIONS IN: DNA-directed RNA polymerase activity, DNA binding; INVOLVED IN: transcription, transcription from RNA polymerase II promoter;            LOCATED IN: nucleus, chloroplast, DNA-directed RNA polymerase II, core complex, vacuole;            EXPRESSED IN: 28 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: RNA polymerase, N-terminal (InterPro:IPR008592), RNA polymerase, alpha subunit (InterPro:IPR00722), RNA polymerase II, heptapeptide repeat, eukaryotic (InterPro:IPR00684), RNA polymerase Rpb1, domain 7 (InterPro:IPR007073), RNA polymerase Rpb1, domain 3 (InterPro:IPR007066), RNA polymerase Rpb1, domain 1 (InterPro:IPR007080), RNA polymerase Rpb1, domain 4 (InterPro:IPR007083), RNA polymerase Rpb1, domain 5 (InterPro:IPR007081), RNA polymerase Rpb1, domain 6 (InterPro:IPR007075); BEST Arabidopsis thaliana protein match is: nuclear RNA polymerase C1 (TAIR:AT5G60040.1); Has 181834 Blast hits to 82224 proteins in 9254 species: Archaea - 731; Bacteria - 33255; Metazoa - 56600; Fungi - 34284; Plants - 19037; Viruses - 3576; Other Eukaryotes - 34351 (source: NCBI BLink).</p>	<p>aa seq</p> <p>MDTRFFTSFAEYSSKVFRTQFGILSPDETRMWSYV            QVEFSETTERGKPKPGGLSDMRLGTIDRKLKCD            TCTANMAECPGFHGLELAKPMFHGFKTVLSI            MRCVFNCSKILADEDDHKFKQALKIRNPKNRLK            KILEACKNKFKCEGGDSIDVQDDDEEPAKKS            GGGAQQPTTIEGMKIAEYKMTKKKNDQEQL            PEPAERKQOLSAEKLSVLKRVSDDEDCQLMGFN            PRYARPDWMLQVLPPIPPVPRPSVMMDTSSRS            EDLTHOLAMIRHNDLNRQERNGAPAHISEFA            QLLQFHATYFDNELPGQPRATQSRGPIKSICTR            LKAKEGRIRGNLGMKRVDFSAFTVITPDNINIDE            LGVPWSIALNLTYPETVTPYNLERLKLVDYYPH            PPGKTKGAKYIIRDDGQRDLRLYLKSSDHLLE            GYKVERHLNDGDFVLFNRQPSLHKMSIMGHRIKI            MPYSYTRFLNLSVTSPYNADFDGEMMIMHVQSF            ETRAEVLELMVPKIVSPQANKVPMGIVQDTLL            GCRKVTKRDTFIEKDFMNLIMWWEDEFDGKVPQ            PTLLKPRPIWTGKQVFNLIIPKQINLVRFSAWHQE            TETGFLTPGDTRVLIERGEKLLSGTLCKKTLGSSS            GSLIHVVEEVGPDAAARKFLGHTQWLVNYWLLQ            NAFSIGIDTIADASTMETINVTISQAKNEVKNLIR            AAQEKQLEPEPEGRMTMESFENRVNQVNLKARD            DAGSSAQRLSDETNLIKAMVTAGSKGSFNISQM            TACVQQNVGKRIPIFGVGRTPHFTKDDYGP            ESRGFVENSYLFG. TQEFFHFMGGREGGLDITA            VKTSETGYIQRRLVKAMEDIMVKYDGTVRNSLGD            VIOFLYGEDGMDAVWVIESQKLDLKMKKSEFDAT            YRYEIDAENWEPDYMLKEHVDLKSIREFRNVFD            AEVQKLEDRQCQLATDIATTDNSWPLPLNLKRL            LWNAGKTFKVDMREKSGMHPMEVEAVNKLQER            LRWVPGDDFLSKEAQKNATLFFNILLRSTLSSKRV            LKFEFLNREAFDWWVIGEIESRFLQSLVAPGEMIG</p>
636	sp_040130_pjur #N/A	-0.352337	#N/A		

10

20

30

40

【表 1 4 - 2 3 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1) 分析に用いたペアイトとの相関係数	sp_107660_kfkg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
637	sp_040150_kxyz	#N/A	#N/A	P-loop containing nucleoside triphosphate hydrolases superfamily protein; BEST Arabidopsis thaliana protein match is: P-loop containing nucleoside triphosphate hydrolases superfamily protein (TAIR:AT2G19120.1); Has 5876 Blast hits to 5615 proteins in 1104 species: Archae - 176; Bacteria - 1754; Metazoa - 1255; Fungi - 973; Plants - 640; Viruses - 179; Other Eukaryotes - 899 (source: NCBI BLink).	MGSRGRLFDLNETFVEVDEEIDFFSGVGFQKYL PSLNSCTSDLLVSSVAPTSIVNNNAFSSHASSVS FQPFVRIKESGLDRIVDQQRARDMDCNTSLPEYD ENSDRKEVMKVNNGNIYAQVIEKEEGEWSDGEC SEPHVSRMEDNSSASGNKDLNEQGIPDAANN SSSGTSLKKNRSDSPADSRKIKHDYASAKPDTF PNDITGGTSONMNGNAERNSTVDFQGDSESVSK QKEVKGAEIAHALKANSLSGKHLDEHKEAMLGK KRSRKTMLNIEDAKQAGAVKSSTPKRTQVTPFS SRTTVKGRNALLSAERNGEKQSEPAIRDQNOV ELSCNEESNPMEPIDRKSESNSDVNAGLLARPR RLNSCNDLSTEGNTRHNLWKQPTDPRQFKDPQ TQPRKLASSHGSLVDVQKLNKKNPPKRLQVLT NHQYQDTSVERLLREVTSEKFWHNPCEELQCV PGRFESVEDYIRVFEPLLFEECRAQLYSTWEELT ETSARDAHVMVVKSVERRRERGWYDVVLLPAN ENKWTKEGDVAVLSCPRPFGSVRRNSTSSAE EYEEPEGARTVGTVRRHTPLDVRDPQGAILHFF VGGSYDPNSAADSDHWVVKLQPPSIWFLSVLGS LATTQREYVALHAFRLNVQMQNALKPSAEQFP KYQEQTPIPECFTFNFDVHLHRTFNDFQLSAIQ WAAMHTAAGTNGVTKRQDPWPFTLVQGGPQGT GKTHTVWGLMNVHLVQYHYTALLKLAPESEY KONNERNADTSSAGSIDEVLRSMDENLFRLLPKL CPKPRMLVCAPSNAATDELLARVLDIRGFIDGEMK VYRPDVARVGDVDTQSRAAQAVSVERRTEQLLVK TRDEIFGLWLOLRGREAHLRSROMATLQVQLNSA AAAGRSQGSVGDVDPDYLWARDQNRDALLQTLA AVVEERKVLVEMSRLLILEGKFRSASFNLLEEA RASLETSFANEAEVFTTSSSGRKLFSRLTHGF DMVVIDEAAQAASEYGLPPLALGAARCVLVGDPPQ QLPATVISKAAGTLLYSRSLFERFQAGACPMLL
638	sp_040530_xuja	-0.924213	-0.876287	ECERIFERUM 7 (CER7); CONTAINS InterPro DOMAIN/s: Exoribonuclease, phosphorolytic domain 2 (InterPro:IPR015847), Exoribonuclease, phosphorolytic domain 1 (InterPro:IPR001247), Ribosomal protein S5 domain 2-type fold (InterPro:IPR020558); BEST Arabidopsis thaliana protein match is: ribonuclease PH45A (TAIR:AT3G12990.3).	MEORLANSWNMTVNEKNFIETALLSDLRYDGRR PFDYRRLTIKFGREDGSGSEVQLGQTHVMGFVTA QLVQPKDRPNEGSLAYTEFSPMADSSFETGRP GESAVELGRIDRGLRESRAIDTESLCLVLAGLWV SIRIDLHLDNGGNLVDAAAMAAALLTFRRPECS LGGDEGOELTVHPPEVREPLPLIVHHLPIAVTFAP FTADTVVDPSPHSEEAAMKGRMTVTLNTNGDICA IQKGGGVGMHNVVMQCLRVASVKSADITTKIKE AVESYNTEKAYRKYKRHPSSAGGQTER

10

20

30

40

【 1 2 3 5】

【表 1 4 - 2 3 6】

1	A	B	C	D	E	F
		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiag (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
639	sp_040550_seaa	#N/A	#N/A	-0.361141	unknown protein; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF814 (InterPro:IPR008532); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MVFYKARPEAGDYTFIMGLDKRFENEELIKYGFPP EDIWFHVDKMSSA-HVYVRLHRGQTMIAISEGLL EDCAQLKANSIQGNKVNINIDVYTPWANLKKTA SMDYGVGFHNSKMYRTVVRVKEKRVNRLNR TKVERTPDFKAEREAYAAAEKAEKAKQAKQREKKR REEMDRLEKERVSEMRSYKNLMLVSDKMTSNKQI ASAEKSLQLELEEDFM
640	sp_041120_iasz	-0.36240%	-0.34736%	-0.37237%	uncharacterized protein	MPGYTADSTVVEGDLDDSSPQRTSSAESNFFRELD DAYLQNAQRIWLGVEVLRQIRFDEQIDISDLADGEL LFEVAEIVSRTLPTLSADESHIKAHERKTLASNKS TRRYMPYVNVDSFLRMCKNLGIDGIDLFTPSDVV ERRNTRKVMCLRSLSKSRRLRDLNVPDFDSVT CTLAMPTDMVEGIRKSLKSHSSSSLSVEQEAEK VSRVVRWRMSNSASAVSNGYSDSCDEAESKF VVAESYSSDDAYDAESPDISGSQVVRVYNSL QLSTRDFENEDDLCEYLSPSNAESVGSPCSQ YYSDLEDLSPSSNGLSGRISGKVPPELNIETNHEY QISDVVHPIDWEDHLDVGNKGNVANDGEVD CVSHDNEVDYDTCCHGFYLLIGNDMCNEQLGTSN NPSKVDCDNISELGEPPCEVLTDDFFHVSADAPHS HEDAPSENNATKHDPVCEVDLGTGEHYSHGKT KDTVLYGELLIQSSCVSKPSESPDLLYDSENAQAS NACTKVINIGVNSCCDRVNNENVTSEVAVAGELRF HEEGSPCSVDYDGNITFNCSSLVDYKSNDDYDDHA TFVDPIRTSIPSLFERDVTICTDVHCSPIENDKESL ESKMPDNEQSHDNGMGMEGNASNESDMPQM DLNSSPCVMEAKNITGCSKGGCHDTMEGSEGHNA LVESVDHKPLDNTILVSSFATAVCDPHSEAQHPKI LEDTDHKLQCYKEAEKGESKFKQEGTNKEVP VHKPQRKLLKSMVKGTTLVGVAVFLLHLRKRRE NTSGAKRPAQVIQGGTTKFLQKQMGSGSTSSV YPAEKLHLKL

10

20

30

40

【表 1 4 - 2 3 7】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
					<p>Chaperone DnaJ-domain superfamily protein;            FUNCTIONS IN: heat shock protein binding;            INVOLVED IN: biological process unknown;            LOCATED IN: plasma membrane; EXPRESSED IN:            cultured cell; CONTAINS InterPro DOMAIN/s: Heat            shock protein DnaJ_N-terminal (InterPro:IPR001623);            Has 3526 Blast hits to 2426 proteins in 300 species:            Archaea - 4; Bacteria - 245; Metazoa - 1133; Fungi -            271; Plants - 159; Viruses - 2; Other Eukaryotes -            1712 (source: NCBI BLLink).</p>	<p>MGAPKASSHDVMEPEKNFDERILKKTCKTRKRDNIV            LIEYSDDFGDVILLDVPDSFQKFRKRAKLVINID            DEDGGSSDGVCHTNSDATTSQCFFTEAADP            STTSDVNGEDTPVKLSKSKRTYPGKSIIPRHCG            ISLDESSESSESDDCVLMGSSGKLRHWEK            ASFRRRQDASCAPHDFENEAGTSPVSHVHDYA            EDAEPPVQOSNVRKCSVQQDGDNVFAGYPEKF            SEESKASDQDYDPMIDIGHVNSSSDEKERRGTL            SPESPVHHNTEEPAGDFRDSTGKLAEEILFRRP            DYGDMTKNEVCSVYSTSITKGPSEDVESDHATLJ            DRQGVVEICMLPDHELGGGNQVSSSDQINIID            RCKIKETEYKHAMEEELAAQQQLLIQAEEAR            LNRKKAENMRIQMIENRRQKERVEEIRGTRRRD            EENLNLEKHRSEIKMELRRTEAICSDMASLLRCL            GIHVGQGYPPNVEVQAAKYKRACLKFHPDRMST            TCLRQQVEAEEKFKLISNMKDKFLQNR</p>
641	sp_041560_zjad	#N/A	-0.873708	-0.86334	<p>SUPPRESSOR OF GAMMA RADIATION 1 (SOG1);            FUNCTIONS IN: sequence-specific DNA binding            transcription factor activity; INVOLVED IN:            multicellular organismal development, response to            gamma radiation, DNA damage checkpoint,            regulation of meiosis; LOCATED IN:            cellular component unknown; EXPRESSED IN: 23            plant structures; EXPRESSED DURING: 13 growth            stages; CONTAINS InterPro DOMAIN/s: No apical            meristem (NAM) protein (InterPro:IPR003441); BEST            Arabidopsis thaliana protein match is: NAC domain            containing protein 85 (TAIR:AT5G14490.1); Has 2166            Blast hits to 2160 proteins in 77 species: Archaea - 0;            Bacteria - 0; Metazoa - 22; Fungi - 5; Plants - 2134;            Viruses - 0; Other Eukaryotes - 5 (source: NCBI            BLLink).</p>	<p>RSWLVDGNRFATKIKCASGSCDLGNIKWKSNTPT            RSCPNCHHTIDNNDVGOEWPGLPRGVKFDPTD            QEITOHLLAKSGIGDEAPHPFIDEFIPTVDEDDGIC            YTHPENLFGVKRDRGSSASHFFHRAIKAYNTGTTRK            RKIHGNLGDVVRWHTGRTKPVLLNGEQKGGKK            IMVLYAVKGGKPEKTNWVMHQYHLGIEEDEK            GSHVSKVFYQQQQLVKQSEKTTGYISDVAEDSV            VKVDVPTKSYTPDPPRNMRLHTKNDISSVNC            NSGAQDREMGFADEVQOEITISSGVNDSIKDDCI            DITSQDPVLEHGDNTVEEEKVWDSSEKYLNSQ            QLVLEGLCLDELLOSQSPNRNGSTNTDDVAKKR            VYLSEYARLGPDDLKKDLEDCCQNLVLDPANIDL            TTPDFRLSQLLEFGSQDSYSAWGLGKTTD</p>
642	sp_042010_qucd	-0.868562	#N/A	#N/A		

10

20

30

40

【表 1 4 - 2 3 8】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
643	sp_042030_fpyg	#N/A	-0.963775	Transducin/WD40 repeat-like superfamily protein; CONTAINS InterPro DOMAIN:s: WD40 repeat 2 (InterPro:IPR019782), WD40 repeat, conserved site (InterPro:IPR019775), WD40 repeat (InterPro:IPR01680), G-protein beta WD-40 repeat, region (InterPro:IPR020472), WD40 repeat-like-containing domain (InterPro:IPR011046), WD40-repeat-containing domain (InterPro:IPR017986), WD40/YVTN repeat-like-containing domain (InterPro:IPR015943), WD40 repeat, subgroup (InterPro:IPR019781); BEST Arabidopsis thaliana protein match is: Transducin/WD40 repeat-like superfamily protein (TAIR:AT5G66240.1); Has 2372 Blast hits to 13576 proteins in 557 species: Archae - 62; Bacteria - 5277; Metazoa - 7900; Fungi - 4335; Plants - 2143; Viruses - 0; Other Eukaryotes - 2655 (source: NCBI BLINK).	MAITSLTELDVVFRSMAIGAVFSDFGGQINSIDF HRTSDLLVTASADSVRLYDINAKLLKTTYHKKH GADRICFTHFSAVVCSSKYNLDSGESLRYLSM YDNFCLRYFKGKERIVSLCMSPIINDSFMGSLD HSVRIWDLRVNACQGLRLRGRPAVAYDQOGLV FAVAMEGGAVKLFDSRSYDKGPFDFLYGGDNA EVDGIKFSNDGKSMLLTTNNVYVLDAYAGEKR CGFSLEPSONTTEATFTFDGQYVAVAGSDGSM HAWNINMRSEVARWSSHIGYVPSCLKWAPRRVM LYAASSVLTFWIPNHGKGTGESNANDADVATKLE VLPH
644	sp_042100_dsea	-0.911067	-0.911938	uncharacterized protein	FLFSLSSSHLIFLFFYSSFFHHRKRSHTSAAYVSIASP PRLTSLPIDLVADISPLSNSSVISLPPQACAQR RSPFSYLLTSLWDLRKHKESGRLWIMLESHNLN HPWLLNPLRFKLTCCVN
645	sp_042830_ozyn	-0.961581	-0.967487	uncharacterized protein	MVELTPNLFNRDTKEEEDLQVYDNRSQSPPIRREK ASPTVKFEFEQIAINQSPQIVSDPSLDGYAVHKL GVIQFL
646	sp_042930_kryi	#N/A	-0.968111	Protein of unknown function (DUF1423); CONTAINS InterPro DOMAIN:s: Protein of unknown function DUF1423, plant (InterPro:IPR004082); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT4G14840.1); Has 192 Blast hits to 189 proteins in 31 species: Archae - 0; Bacteria - 0; Metazoa - 2; Fungi - 10; Plants - 173; Viruses - 0; Other Eukaryotes - 7 (source: NCBI BLINK).	MTCENEDASGHGGELMEVDIADASNGCMTEIRE NGLLLWPYLPQESGEGLPYAVVNWPKFGDIWG WRVGRVAASGYLDLYLPLSYORSRRKRDG FASKLVEQYVRAFPGTNTVYAFFSFSWIKFSK QAYLLKDEHGSFRISSEEMAAQSPYMLPGGLG CKAGHFACASLLEDQMSPSATLSDICSEPSFC RECCILCOCKTIDLGLGGYFIRCEATVHDDLICG HAAHINCALRSMAGTYGSGVGLDTEYYCRRCD TRTDLTKHVTRLLKTCESIESVDEIDKMLSVGACI LRGSKKYQSKLLNRIESILHKLFGNDVGDVWVK VEVEILPVTPTPTTGGQLDYANSACTDVGEDDL LYQNGYLKLEEEINQVLOELKSSQEEYFESIAHDKL HACKDNLFSLYQQYVKDKLELLRPSPSKDRDALL RSIAAREEYIKHELKLRDMEKVANGFGKTPKLL KEHFDDLKDK

10

20

30

40

【 1 2 3 8】

【表 1 4 - 2 3 9】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
647	sp_043190_gcix	-0.947319	-0.844454	-0.953366	<p>nine-cis-epoxycarotenoid dioxygenase 5 (NCED5); CONTAINS InterPro DOMAINs: Carotenoid oxygenase (InterPro:IPR004294); BEST Arabidopsis thaliana protein match is: nine-cis-epoxycarotenoid dioxygenase 2 (TAIR:AT4G18350.1); Has 2882 Blast hits to 2845 proteins in 494 species: Archae - 16; Bacteria - 802; Metazoa - 232; Fungi - 202; Plants - 893; Viruses - 0; Other Eukaryotes - 737 (source: NCBI BLINK).</p>	<p>MMNPALTNPCTKSDWAQLAFNSPPKAMICTRK LSKENIGYNYTSSNNNIQCATASPSILHFPNTH DNNVKNPPTKQQQLKKWNIFQKAYGAIDMWET ALVSLERSQPLPKTSDPKVQIAGNFAPVPEMPVK QSLHVTGEIPGCDIGVYVRNGANPLFEPTAGHHL FDGDGMVHAVTIHDGVSAYACRFTESTRLAQER DLGRAVFPKAIAGELHGHSGIARLMLFYSRG_LFGLI DNEHGMGVANAGLVYFNDRLLAMSEDDIPYQVR VTPAGDLKTVGRFDKGGQLTSTMIHPKLDPESK ELFALSVDVLKAPYLKFKFSQDGEKSADVDIPLN EPTMMHDFAITKKFVVPDQQVWFKMKEMITGGS PVVYDKNKTSRFGIMRKDAKADSDITWVTSPTHT CFHLWNAAWEESETEIIVIGSCMTPPDSIFNECD ENLQSLSEIRLNLRTGESTRRPIIDFSQHPHINLE AGMVNRNLLGRKTRFAYLAISEPWPVKVSGFAKV DLITGEVKKFMVYGEKYGGEPFFLPHDINSILDES EDDGYIMGFVHDEECEEKSELLVNAQTQLLEAAV KLPSRVYPYGFHGTFRSEDLSHQA</p>

10

20

30

40

【表 1 4 - 2 4 0】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
648	sp_043260_pwie	-0.337334	-0.337303	#N/A	<p>SUPPRESSOR OF AUXIN RESISTANCE 3 (SAR3);            FUNCTIONS IN: porin activity; INVOLVED IN: mRNA            export from nucleus, defense response signaling            pathway, resistance gene-dependent, response to            auxin stimulus, negative regulation of flower            development, developmental process; LOCATED IN:            nuclear membrane, nucleus; EXPRESSED IN: 23            plant structures; EXPRESSED DURING: 13 growth            stages; CONTAINS InterPro DOMAIN/s: Peptidase            S59, nucleoporin (InterPro:IPR007230), Nuclear            protein 96 (InterPro:IPR021967); BEST Arabidopsis            thaliana protein match is: Nucleoporin autopeptidase            (TAIR:AT1G10390.2); Has 653 Blast hits to 526            proteins in 148 species: Archae - 0; Bacteria - 0;            Metazoa - 269; Fungi - 224; Plants - 93; Viruses - 0;            Other Eukaryotes - 87 (source: NCBI/Link).</p>	<p>MECDTLRDPFSSRTNFEEYSSLSQISSPRKRRKT            NRVKYDGFPPILRSSDYTEPPLKELAELEEIYPG            YCSKIKDFTVGRVGFSGWKFLGETDVRRLDLREI            VKFNRHEWVYEDESSKPDVGEGLNKPAVVTLFL            KFETL SFDRSKFKVVDRLRASTRDQGAEFISYA            PLNGEWFKSVRHFSSFFGLSMDDEDDVMDADAV            NPCEDPEVMNGGVGFGDEDDIGSDGDFMGE            NSAELSHSLPAHLGLDPARMKEMKMLMFSGEDD            EDENYDDLSLEKQGAAYAEKFRLYNSGLRVNYKP            SPPPPRTKTPVPLEYKNSLNDVNAPGTLSSQOH            KGSPLRVVYKPDGFKLNYKQETPVSGYHSRNIVDA            GLFMGRSFRVWGWGPNGLLVHSGVPVSSNDSQG            LDLHKSLDHRILEVEVGSLSLQEVVCDRSALS            DICRRYEVYERQLEVPGMPLPPVIRMSLMHQVIVW            ELIKVLFSERETSGRTVFGVEDHEEDMMLDGKAI            RSEIDPESLPLMRRAEFSYWLQESVYHRRVQEDL            SCLSESDLEQIFLFLTGKQLDAAVELASRGDV            RLACLLEAGGSMVNRSDYARQLELWRITNGMD            FNFKIERIRLFEKLSGNIHSALAGKSIDWKRFGLG            LMWYELPPGTSIPAVFHTYQHLLEIGKAPYVPV            YIDEGPVEETVKVDTGDRYDLAYLMLLHASEDG            KFDGLKSMFSAFASTNDPLDYHMHWHRQGLEAI            GAFSSSDLYLDMAYVSQLCLGKCHWAIYVVLH            MPYCDFFPYLHSTVIREILFQYCDTWTSTQEEQRK            FIEELGIPSOVLHEAMAVFYQYGDIAKALEHYIR            CFCWQKASHFITSVAHSFLSAEEDSEIWRLAMS            LEDHKSEIEDWDLGAGIYISFYLLRSSLCCDEDDI            NKLDLTSKEAACREFFDCLNKSLEVLGSRPLTD            GRVYKMSIEEICSLLVSDSGADSTRETQLSCFD            TIAAAPIFEDLRFSCHLQEAVALLLFTFLSETTT</p>

10

20

30

40

【 1 2 4 0】

【表 1 4 - 2 4 1】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	<p>注釈</p> <p>embryo selective 1075 (emb1075); FUNCTIONS IN: pyridoxal phosphate binding, carboxy-lyase activity, catalytic activity; INVOLVED IN: cellular amino acid metabolic process, embryo development ending in seed dormancy; LOCATED IN: cellular_component unknown; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINs: Pyridoxal phosphate-dependent:transferase, major domain (InterPro:IPR015424), Pyridoxal phosphate-dependent:decarboxylase (InterPro:IPR002129), Pyridoxal-phosphate binding site (InterPro:IPR021115), Pyridoxal phosphate-dependent:transferase, major region, subdomain 1 (InterPro:IPR015421); BEST Arabidopsis thaliana protein match is: glutamate decarboxylase 2 (TAIR:AT1G65960.1); Has 3259 Blast hits to 3251 proteins in 995 species: Archae - 187; Bacteria - 1878; Metazoa - 518; Fungi - 211; Plants - 249; Viruses - 11; Other Eukaryotes - 205 (source: NCBI BLINK).</p>	aa seq
649	sp_043420_onwr	#N/A	#N/A		<p>MAGLAAPQELSDQVDLLNRTDLDPTAVVDPPI PPVAVTSNGENGVKMGREIVMGRNHSSCLDIT EPEANDDYTGDKKEYAMGILARYRQSLLETKYH LGYPYLNDFDYGALSQ_QHFSINILGDPFIESNY GVHSRQFEV(SVLDWFARLWEIEKQYWGTYTNC GTEGNLHGILYGREVFPDGLYASKESHYSIFKAA RMYRMDCKIDTLSSGVEVDCADFRTKLLQNKDK PAINVINIGTVKGAVDLIDLVIQTLEECGFPQNR FYHCDGALFGLMLFPVTRAPQVTFKKPGISVSVS GHKFGVCPMFCGQITRMDHINTLSSNVEYLASR DATIMGSRNGHAFIMWYTLNPKGYRGFGQEVQ KCLRNAHYLKDRLRDVGVSAMLNELSSTVWER PRDEEFIRRWQLACQGEIAHVVMVMPSTYIEKLDI FLDELVEKRLTWFKDGGGNQPPCLASEMGKDNCA CPLH</p>
650	sp_043640_gfmq	#N/A	#N/A	<p>Transcription factor TFIIIE, alpha subunit; FUNCTIONS IN: RNA polymerase II transcription factor activity, transcription initiation factor activity; INVOLVED IN: transcription initiation from RNA polymerase II promoter; LOCATED IN: endomembrane system, transcription factor TFIIIE complex; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINs: Transcription factor TFIIIE, alpha subunit (InterPro:IPR002853), Transcription factor TFE/TFIIIEalpha, HTH domain (InterPro:IPR017919); BEST Arabidopsis thaliana protein match is: Transcription factor TFIIIE, alpha subunit (TAIR:AT4G20340.1); Has 759 Blast hits to 706 proteins in 191 species: Archae - 5; Bacteria - 10; Metazoa - 313; Fungi - 190; Plants - 108; Viruses - 11; Other Eukaryotes - 122 (source: NCBI BLink).</p>	<p>MNSLSIEPFNRLVKLTARAFYDDISTKGDNOQPK TSRSDNRGTAVVILDALTRQWVYREEDLAKDLKL HLKQLRRTLRFEEEEKLVMRDRHRETAKGVRMH NAAVSATANHHYHGGREGDDKIKLHTHSYCCLDYA QIYDVVRYMRHMKKIKLDELDSKNITIQEYICPN CNKRYTALDAARLISEDEGENFLCSCQGGELVAES DKLAGQDYGDADENVRRRRREKLEKMLQKMEE QLKPLVEQLNRVKDLAVPDFGSLQAWVEYKASAA APAAAGDNPDPSRSGGTFPFLGDPKIHVD ISGVEDKESKSGTGSAPMKVLPWMIRQGMKL TKEORGEESKSPDVTSEPAGLSDDKKSTVWK DKSIOQEYFKAYAAALERQRETIDVKGEPDVST EEVTGQNTSLDRQVGLKSKREEQSEDEEDVWEE ATPSAGNTGESYKVDLNVQAEPLPEEEEDNDEE DVEWEDG</p>

10

20

30

40

【表 1 4 - 2 4 2】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 IND1 (iron-sulfur protein required for NADH dehydrogenase)-like (INDL); LOCATED IN: mitochondrion; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: ATPase-like, ParaMinD (InterPro:IPR019591); BEST Arabidopsis thaliana protein match is: ATP binding (TAIR:AT3G24430.1); Has 13193 Blast hits to 13164 proteins in 2587 species: Archaea - 550; Bacteria - 7572; Metazoa - 442; Fungi - 431; Plants - 201; Viruses - 0; Other Eukaryotes - 3997 (source: NCBI BLink).	aa seq MKGAFKINTVFNILRNRYGIRGLSAMAGKAKGLQI SGVKDIIIVASGKGGVGSKSTTAVNLAVALANRCQ LAVGLLDADVYGPSIPTMMKLGKPEATEDLKMI PKENYGVKCMMSMGLLVEKDAPIVWRGPMVMSA LEKMAKGVWEGKLDLIVDMPPGGTDAQLTISQ RLQLSGAIIVSTPDQIALLDARRGAHMFSKVNVPII GIENMSCFMCPHCSPQSFIFGEGGAKRTATEMS LDYLGDIPIEMDIRTCDEGVPIVLSPPDSTSAMK YTDVAQKVMKRIEIEISHKQFSPEISL
651	sp_0444010_lqjo	-0.8738753	#N/A	#N/A	GNARLED (GRL); CONTAINS InterPro DOMAIN/s: Nck-associated protein 1 (InterPro:IPR019137); Has 2387 Blast hits to 741 proteins in 151 species: Archaea - 0; Bacteria - 54; Metazoa - 569; Fungi - 142; Plants - 67; Viruses - 0; Other Eukaryotes - 1555 (source: NCBI BLink).	MSEWRRKISWQRGHVDENLQNMQSSSWTAD MKLKQVPSLLYLFFILDCEWHEANRSDLSPKLVFV DQLCEISPHLPRSSRWNSIIRSLYNYQYGNSSSL PLTLRL
652	sp_044230_medy	-0.866573	-0.933373	-0.875225	SBP (S-ribonuclease binding protein) family protein; FUNCTIONS IN: zinc ion binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, RING-type (InterPro:IPR001841), S-ribonuclease binding protein, SBP1, pollen (InterPro:IPR017066); BEST Arabidopsis thaliana protein match is: SBP (S-ribonuclease binding protein) family protein (TAIR:AT1G60610.3); Has 1212 Blast hits to 1212 proteins in 167 species: Archaea - 0; Bacteria - 2; Metazoa - 524; Fungi - 0; Plants - 416; Viruses - 105; Other Eukaryotes - 165 (source: NCBI BLink).	MFGNNSNHAFPMFLDENRFQLNTNAPNHQLOL FGNX
653	sp_044830_weky	-0.952733	-0.986458	-0.91463		

10

20

30

40

【表 1 4 - 2 4 3】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
654	sp_044870_omf		#N/A	CRINKLY4 related 1 (CCR1); FUNCTIONS IN: protein serine/threonine kinase activity, receptor activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; EXPRESSED IN: 16 plant structures; EXPRESSED DURING: 7 growth stages; CONTAINS InterPro DOMAIN/s: TNFR/CD27/30/40/95 cysteine-rich region (InterPro:IPR001368), Protein kinase, ATP binding site (InterPro:IPR017441), Serine-threonine/tyrosine-protein kinase (InterPro:IPR001245), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271), Regulator of chromosome condensation/beta-lactamase-inhibitor protein II (InterPro:IPR009091), Protein kinase, catalytic domain (InterPro:IPR000719); BEST Arabidopsis thaliana protein match is: CRINKLY4 related 2 (TAIR:AT2G39180.1); Has 92283 Blast hits to 90862 proteins in 3897 species: Archae - 88; Bacteria - 8648; Metazoa - 34704; Fungi - 6843; Plants - 29398; Viruses - 212; Other Eukaryotes - 12390 (source: NCBI BLINK).	MLSSLLFLLLSLPISAYAFSGMGPISAAFGGG DAFFCAIDASGKQGIICWGMSPKTSFNTTIPS MAALSGGEGFMCGILANTSOAYCWDSHGTE LVPPFYRYNNYSSIASGNKHVCAIRGNFYDSSN SGSVDCWEIINGSNSNTFTSKOSALFYDQSISSL VFEKIVSGDFSGVGDDEELICWGPSTNLGIS KISDNFLGLSSRRSSVGVKPSGEVKCWGKND SYDPFSGVRFVALTGGSRHFHFCGIRADTHAVDC WGGFNSSSIPKSGSFFAIATSDVAACGIREEDLV LDCWFVNESSAPTQYNPLQLCSPGLCSFRSCD VGEFTNASALNEPELTNVCIRRDHLHLCSSCGTN CSSGYLLSPPCTGNADRVCTECLCQNSSCWG VCGLAEVELKRRHQROMKIVLILGPSIAGFLILI GWCLFPYLYFFKAKDGPTKQFRSCIGKDDPESD ANGDANPPMAVACCPEAQIFRILSELKDATNGF KEFNELGRGSGYGFYKALLTDGRQVAYKRANAA TIHTNIREFENELEILCKMRHGNVNLGLYCTEMG ERLLVYEMTHGTLHDHLHDGLSPLSWTIFLRIA MOAAKGLVYLHMEANPPVVRDVKTYNILLSD WGARISDFGCLNFNERDELDEKMDVYVNFVGL LEILSGRKAYDVEDYSPNIVEWAVPLJRRGKAAL IDRYVALPRNVEPLKLAIAESAVRESSSDRPTM SAIANSLEQIVKEGLL
655	sp_045040_dcja		#N/A	uncharacterized protein	VPPPLTAWPFPAGRCIPAVPHRRVAPPCL

10

20

30

40

【 1 2 4 3】

【表 1 4 - 2 4 4】

1	A	B	C	D	E	F
1	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa_seq
2					<p>chromatin remodeling 8 (CHR8); FUNCTIONS IN: helicase activity, DNA binding, ATP binding, nucleic acid binding; INVOLVED IN: DNA repair, response to gamma radiation; EXPRESSED IN: 15 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: DEAD-like helicase, N-terminal (InterPro:IPR014001), DNA/RNA helicase, C-terminal (InterPro:IPR01650), Helicase, superfamily 1/2, ATP-binding domain (InterPro:IPR014021), SNF2-related (InterPro:IPR000330); BEST Arabidopsis thaliana protein match is: chromatin remodeling 24 (TAIR:AT5G63950.1); Has 17230 Blast hits to 14463 proteins in 1653 species: Archae - 99; Bacteria - 4180; Metazoa - 3846; Fungi - 4219; Plants - 1677; Viruses - 118; Other Eukaryotes - 3091 (source: NCBI BLink).</p>	<pre> MEEEEDVILNSGVTSWTFDIERFDIKKATIVV ESSOAGRRITTEADFTDITDQDSSSSQVELL KKLRAIEVEIHAVASAVEQTKSGGKNEELDSH GKKEEDDEGIDKAIPRDSALQHAAADRLRSLKK TRAEIRKELSSICQDGKTVGVVEKLLKYLVEES KSKRKLKEVQGPSKESKKAQVYQEDDEFDVAV LNAASGGFVETERDELVRKGLTFPHKLGKFERR IQOPETSSQQNGDIADRASQSVAKAVRSLTEA AGARPTSKLLDPKDLPRLDAPSAFNRLSKSFV SKSLDDGKSNDAKAKRRRPLPGERWRKRIAKE EIEVEEEDVQDDLITGTQDADNEDGDVTPSVLE GGLKIPETIFDKLFDYQKYGQVWLWELHCQRAG GIIGDEMGLGKTQVLAFLGSLHFSKMYKPSIYCP VLLRQWKREARKWYRSHFVHVLHDSAVDGGRK KQSESESESESESESSDODRVPSPKSAKK WDSVIKRVLRSESGLLITTYEQLRILGEKLLDIEW GYAVLDEGHRIKRNPNAEVTVCKQLQTVHRIMTG APIQNKLSLWLFDFVFPKGLGVLVFEAEFV PIQVGGYAMASPLQVSTAYRCVAVLRDLJMPYLL RRMKADVNAHLPKKTEHVLFCSLTLPQRSVYRAF LASSEVEQIFDGNRNSLYGIDVMRKICNHPDILER EHASQNPDYGNLERSGKMKVVAQVLMKMKEQG HRVLLFACTOOMLDLENFLINSEYSYRMDGLT AVKORMALIDEFNNSVDVIFILTTKVGGLGTLNLT GANRVIIFDPDWPSPDMDQARERAWRVGQTKD VTIFRLTRGTIEEKYHRRQIYKHFLLTNKLNPOQ RRFFKARDLKDILFVLTDEEHAGSTETSNIFGOLS EQVNVISKEKEDQGNINACAGATAACADDMLDR ENSSQOELSGNKEKEEAENGGEADEETNFKLSL FDTHGIHSAINHDAIMDAHDEEKIRLEEQAASLVAQ RAAEALRQSRMLRSKESISVPTWTGKSGAAGAP SSVRRFTFGSTVNSQLVRNRSRSEEGAPKNGTTN </pre>
656	sp_045120_u1uo	-0.884377	#N/A	-0.911742	<p>RNA polymerase III RPC4; FUNCTIONS IN: DNA-directed RNA polymerase activity, DNA binding; INVOLVED IN: transcription from RNA polymerase III promoter; LOCATED IN: DNA-directed RNA polymerase III complex; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: RNA polymerase III Rpo4 (InterPro:IPR007811); BEST Arabidopsis thaliana protein match is: RNA polymerase III RPC4 (TAIR:AT4G25180.1); Has 234 Blast hits to 234 proteins in 100 species: Archae - 0; Bacteria - 0; Metazoa - 103; Fungi - 41; Plants - 59; Viruses - 0; Other Eukaryotes - 31 (source: NCBI BLink).</p>	<pre> MDPASFDQKPRVPRKVRFPKAPVRKPKSAAVA KTEVADDAEAAQKKELMRRFQDGLKAPKFEKK TERVQRVAFYGGSSIPGTFQFSANGNNSWRQD KAHNLRENKEYKEPWNYSYYPVTLPLRRPYS GNPDILNDEEFGQLSTFDESTSNPAMELGLMGE APEPRMFLQLPTTIPMTKRNSAAAGQESTSSS KPPSGSNPVEKAAAGLKELOAGYMGKMLVYRS GAVKLGDTLYDVSLSGSDCAFAQDVVAIVAVDK HFSVVGELNKRALTPDVSILDSFADLGL </pre>
657	sp_045280_zpgr	-0.868674	#N/A	-0.880279		

10

20

30

40

【表 1 4 - 2 4 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
658	sp_046480_fmth	#N/A	-0.833727	ENTH/ANTH/VHS superfamily protein; FUNCTIONS IN: phospholipid binding, clathrin binding, binding, phosphatidylinositol binding; INVOLVED IN: N-terminal protein myristoylation, clathrin coat assembly; LOCATED IN: membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINs: Epsin-like, N-terminal (InterPro:IPR013809), ANTH (InterPro:IPR011417), ENTH/VHS (InterPro:IPR008942), Clathrin adaptor, phosphoinositide-binding, GAT-like (InterPro:IPR014712); BEST Arabidopsis thaliana protein match is: ENTH/ANTH/VHS superfamily protein (TAIR:AT1G14910.1); Has 1020 Blast hits to 594 proteins in 190 species: Archae - 0; Bacteria - 14; Metazoa - 316; Fungi - 202; Plants - 440; Viruses - 0; Other Eukaryotes - 48 (source: NCBI Link).	MATLWKRKAYGALKDQTKVGLATINSYKDLDV AIVKATNHVECPKHERHLRKLILLSAIRPRADVA YCLHALARRLGRKTRNWTVALKTLIVVHLLREGD PTFREELLNFQGRVQLQSNFKDDSSPVAWDC SGWVRTYALFLEERLECFRVLKYDIEAERLPKPA EGDEKGYSTRDLGSEELHLPALQQLLYRLIG CRPEGAALGNVIQYALALVKESFKYCAINDGII NLVDFKFFEMPRHEAVKALEIYRRAGQQAASLSEF YEYCKGLELARINFQFPVLRFPQSFLATMDEYIR EAPKMDMPAQOLEPPERLLLTYKPEENSLSSE EKSNVEEPPPMASDDLTSAAEAALATTRALSPAT NMDSGDLLGNDVTPNATAIEEGNALALAIVSSD AASTNSGTVSSSDFPDSGWELALVTAP2D1STH TERQLGGGLDLSLTNSLYDEAAYRANQOPVYGA PAAPNPFVHDPPFGVSHNIAPPTAVQMNMSMTQQ SAPNPFQPYQAPFHACSQPPVGLGAPNPFQDM GFGPYAVTPVNHHPQHAANPFQNTGLL
659	sp_046590_kwap	#N/A	#N/A	CRUCIFERINA (CRA1); FUNCTIONS IN: nutrient reservoir activity; INVOLVED IN: response to abscisic acid stimulus; LOCATED IN: endomembrane system; EXPRESSED IN: guard cell, seed; EXPRESSED DURING: seedling growth, seed development stages; CONTAINS InterPro DOMAINs: 11-S seed storage protein, conserved site (InterPro:IPR022379), Cupin, RmlC-type (InterPro:IPR011051), Cupin 1 (InterPro:IPR006045), RmlC-like jelly roll fold (InterPro:IPR014710), 11-S seed storage protein, plant (InterPro:IPR006044); BEST Arabidopsis thaliana protein match is: cruciferin 2 (TAIR:AT1G03880.1); Has 1047 Blast hits to 1009 proteins in 168 species: Archae - 0; Bacteria - 115; Metazoa - 2; Fungi - 0; Plants - 930; Viruses - 0; Other Eukaryotes - 0 (source: NCBI Link).	MASTTNALLFLACLCLVALFPPTLAQTFLOQPQ QFLRVRGQQQLGHECDIQQLTAAEPTRRVRAEAG VIEWESNTEEFRCCTGVAARYVIEPKGLLLPSY TNAPYVYVMQGGKIQGVILFGCPETFEPSYRFG GSESGKVERDQYQRLRVHEGDVIGSLAGVY QWYTDGDTPIVSTLDDLNSPNQNDLJNFRSFY LAGNPQGGGQEGQKPEVVGKNIENGFDELLA DAFNVDTEIRSIKIGENDERGSIIRVERDLELLSPE WDDMEERMRRLNGLQVLCVTSFKQNDQPSR ADVYTEHGGRINTLNSHKLPLQLHLQLSYERGLV YKNAMMTPHWNIASHIYITRGTGRIQVARENGL VFDGRVQEGQLLVFQNFVYVKKAEQEGLEWVS FKTNDNAMISPLAGKLSAIRGMPPEEVLNMSYGLS RDEVRRILKYGREQLSFLSPRLD

10

20

30

40

【 1 2 4 5】

【表 1 4 - 2 4 6】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
660	sp_047050_qjrz	-0.830244	#N/A	-0.833381	CONTAINS InterPro DOMAINs: WRC (InterPro:IPR014977); Has 219 Blast hits to 219 proteins in 19 species: Archae - 0; Bacteria - 0; Metazoa - 3; Fungi - 0; Plants - 215; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink).	MRIRNRKVLPLLSLLSPVPLSDPHLYSRSPMPVQDVATYSPPEPPSPNDLTTPTSWPSSDLHQHRGLDNNITVGAIRFEDGEOVQISSPAPILHQLSATTSAAGASENNNGRLVFLSNKDRVGRWCDDKAIPLKRRGFGFNITSFAENSSTRITIDEDEDEDEDEDEDEDEDEDDDDTATTATEMIIMKWNIN DDEEKQAVKDDKDAVAPPAPAPAPAPAPAPAPAPAPATTTVKKGNKRGNTILEGSRCSRVNGR GWRCCQPTLVGYSLCEHHLGKGRLSNWSASAKSRSSKKSNSITKFAIGRDEESDHTYDVREKLE RHNISFGNKRVKLVVYKARSLSSLLGQITNAMKNN NIVADNHATMQQEKISNLDYQ
661	sp_047290_czfo	-0.836726	-0.839532	-0.839387	G-box binding factor 1 (GBF1); CONTAINS InterPro DOMAINs: Basic-leucine zipper (bZIP), transcription factor (InterPro:IPR004827), G-box binding, MFMR (InterPro:IPR012900), bZIP, transcription factor, bZIP-1 (InterPro:IPR011616); BEST Arabidopsis thaliana protein match is: basic region/leucine zipper transcription factor 16 (TAIR:AT2G35530.1); Has 3905 Blast hits to 3842 proteins in 356 species: Archae - 8; Bacteria - 123; Metazoa - 1421; Fungi - 442; Plants - 1419; Viruses - 137; Other Eukaryotes - 355 (source: NCBI BLink).	MWAGQHPMPPYGTVPYPALYPPGAVVYAHPNMAMTPNPASVWETKGNKSDYKQNSVYKSKGN SGLGNEKGVETAKATSGSGNEGASQSAESGSE GSSDASDDNNQGEHSATKKGFDLMLEDGANA QHHSVNASAGGSVPKGLGVAPNLNIGMDLWN SSAGGAKLRPSTGVASGVAPAMVGRGLMPDH QWIQDERELKQKQKQSNRESARRSLRKOAE CEVLQKRVESLNENSNLREELKRLSAECEKLTSENKSIKEELTOTCGPDSLQSLLEENNASGEGDS
662	sp_047760_jdxq	-0.871505	-0.883454	-0.882471	SOUL-1; FUNCTIONS IN: binding; INVOLVED IN: N-terminal protein myristoylation, red or far-red light signaling pathway; LOCATED IN: plasma membrane, vacuole; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINs: SOUL haem-binding protein (InterPro:IPR006917); BEST Arabidopsis thaliana protein match is: SOUL heme-binding family protein (TAIR:AT3G10130.1); Has 1375 Blast hits to 1375 proteins in 156 species: Archae - 18; Bacteria - 186; Metazoa - 71; Fungi - 0; Plants - 173; Viruses - 0; Other Eukaryotes - 927 (source: NCBI BLink).	MGMILGITVETPKYEVKIKSTNEYEIRKYPASVIAE VTYDPSQMKGNKDDGGFTILANYIGAIGKQNTT SEKVSMTAPVVTQSSNAAAAAIAIWTAPVITKG REGEKEEVTMQFILPEKYKKAEEAPKPVDERVVI KQEGERKYGVWKFSGVASDGVVWVKVVEGLRIAL DKDGFKVIQEYLLARYNPPFTLFLRLTNEVFLPVE

10

20

30

40

【 1 2 4 6】

【表 1 4 - 2 4 7】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
663	sp_047840_szo0			<p>general regulatory factor 8 (GRF8); FUNCTIONS IN: protein phosphorylated amino acid binding; INVOLVED IN: brassinosteroid mediated signaling pathway; LOCATED IN: in 6 components; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINs: 14-3-3 protein (InterPro:IPR000308); BEST Arabidopsis thaliana protein match is: G-box regulating factor 6 (TAIR:AT5G10450.1); Has 2678 Blast hits to 2669 proteins in 381 species: Archae - 0; Bacteria - 0; Metazoa - 1257; Fungi - 298; Plants - 764; Viruses - 0; Other Eukaryotes - 359 (source: NCBI BLink).</p>	<p>MAIPDNLTRDQYVYMAKLAEOAERYEEMAKFMD KLVVTTSSGGELTVEERNLLSVAYKNVIGSLRA AWRVSSIEQKEEGRKNDHVLVDRYRTKVESE LSDICACILKILDSNLIPTASVSESRVFLKMKGGY HRYLAFFKSGEERKSAEDTMEAYKNAQEIATVD LPPTHPIRLGALNFSVYFYEILNAADKACSMKQ AFEEAIELDTLGEESYKSDSTLIMQLLRDNLTLWT SDQDQLDES</p>
664	sp_048480_img			<p>DHFS-FPGS homolog C (DFC); CONTAINS InterPro DOMAINs: Mur ligase, central (InterPro:IPR013221), Folylpolyglutamate synthetase, conserved site (InterPro:IPR018109), Mur ligase, C-terminal (InterPro:IPR004101), Folylpolyglutamate synthetase (InterPro:IPR001645); BEST Arabidopsis thaliana protein match is: DHFS-FPGS homolog B (TAIR:AT5G05980.2); Has 7721 Blast hits to 7719 proteins in 2542 species: Archae - 43; Bacteria - 4847; Metazoa - 185; Fungi - 368; Plants - 133; Viruses - 0; Other Eukaryotes - 2165 (source: NCBI BLink).</p>	<p>MLHIQSSFLRHGLFRISFCFERRWRLRGVAAR GVSCLPSCLNKHPAPGLRSGMIPTEIVENMTRI EVTEEYSWDLTHRSRYETVMEALSSLLTCRKRGE KPSIGGKYTKLERMTMYLKIIGLEENIAGLNIIHVA GTKGKGSTCAFSEAILRECGFKTGLTSPHLIDVR ERRFRLDGIDIAEEKFLQYFWDQDQLRDNVTKD LPMPLLFQFLTLAFKIFIDEKVDVGFVGLGGEK DSTNVIKPEVAVCGISSLGMIDHLDLSDGLTGKIALH KAGIFKPNVPAFTVPQLPEAMNVISOKAQLRIPL EAVQPLTRGELNGLQLSLAGDHQFYNASLAVSL CKSWLRHTGNSEKVFVNDGCEADMIPDAFLRGL SSARLFGRAQIVHDAIFRSSDTATGDLVYYLDGA HSPESMDACARWFSKMAKGDHKSSTPLSSSKSS CENGLGEAVSKNGFTKQEKGDKSKIKOILLFNC MEVRNPQILLKILGTCCASSGVHFGSALFVPSIST YKVVSGYSITSNIPILRNLSWQFNLRQIWEKTVN GKDFDLDRNSMTECIKNESTDALPFLFLHGDGY RDFPFGDSFAVAVPSLSSAIKLLRDCARENPISR LQVLTGSLHLVGDVLRLLKR</p>

10

20

30

40

【表 1 4 - 2 4 8】

A	B	C	D	E	F
1	分析に用いたペイトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
665	sp_049010_xxhz	-0.331053	#N/A	<p>G2484-1 protein (G2484-1); FUNCTIONS IN: RNA binding; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Agene1 (InterPro:IPR008395), Tudor-like, plant (InterPro:IPR014002); BEST Arabidopsis thaliana protein match is: agene1 domain-containing protein / bromo-adjacent homology (BAH) domain-containing protein (TAIR:AT1G68580.2); Has 6169 Blast hits to 4010 proteins in 631 species: Archaea - 30; Bacteria - 1624; Metazoa - 1931; Fungi - 976; Plants - 400; Viruses - 26; Other Eukaryotes - 1182 (source: NCBI BLINK).</p>	<p>MFVMDVDDNDITRRNICLZAGEGRSSSSALRFPYD  FPKFDFTLQGHRLRFDLSLVETEVLGIGSQEDN  QWIEDFSRVNMGMEFNSTPADSCSISRNRNVWS  EATSSSEVEMLLKSVGQEEEMMPGDSIITESNATH  DQGIPKNVSESSKLIERTDAYVNSDAALPPDEL  LEDPCASGRHIEKQLFDPSAFHAKVESASGISSN  VNGNALCNGGGLTYESNHSDHEERSGEKLEES  QTSVGQTYNFGSGVDSLGFSSITTDNVTNSIGTLS  TSKTQQVMFNAFHESCNDIPKSOVSIAELDDT  TILEGKANASSINFTGSSSHGFVPESGRDCCQEAS  LDFVQESTNILYKGESELKSSSEGVKNTSLTISIQ  TSRHEKVTALDDTAILEGKACVSSINIITGSSSR  GSVPESGRDGREALDFVRESANLVYKGESELKS  SEGVKNTSFTTISQTSKHEKVVGMSEVIEPFD  DCPGRPSYDFASTEQDMDLTMQLPHKQDIEWND  NMEDYSAIRLNHESMSEMPSLIEDNGVPRGSA  DCIRTEVGNSSNISVQLSASDHQHNRVEVEGYH  NEGIVYGEFPGVAPDKVPIQVIGENLYHAQADF  SEIDICKPEVDSEVSTDLGNIDSENVRTLKNDQR  TQSAEVSSEGDDGNNLIHKLQPPQTSGEPEAINVS  PFATTNAAALDDDCPVIGQREPAVAVGLFHSNDD  EEAGRTKNDITNLGAEVDGASKVEDSVADDIYGI  PVSEMAAKSEQPVKIDIDICISGSPSSVQAMNSLY  SDHPQAHSDDGNKPKSLKESADADLINVENHEKA  MKNDEEKESAQIAEIDKTCGSPSTISSSEPSQSE  KDNQEEEDRGYITNEVKDKIQSTSEDSKRKDLLGD  DSSFTFKVNALPDTTTEREPAKNWSPPFDVYEKCK  VVVETTPSGKGTGRTPLKTPQNTQVSEGET  GFGSSKGTSEKGRKRGKGTAKENAKNGNSKGP  TCENASIMMDKYSNMLSTPISNPIVQYGMQOAY  GHTEANNKFPAGAVVPTPSLPDLNASALNNST  PSSTSYRQPFDTLQQVQLFAQIFVYGSLLIQGTAP</p>

10

20

30

40

【 1 2 4 8】

【表 1 4 - 2 4 9】

	A	B	C	D	E	F
1		分析に用いたパイロットとの相関係数				
2	sp_107620_dnoz (SOAP1)	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
666	sp_049520_dnoz	#N/A	#N/A	sp_266333	uncharacterized protein	MSFTNGPEKFFQYDMEELLEENVDELKKNPAM KEVQSLADTKFPDDGPLNEHVTASPLVEFAVSI HSYKSTLFNDGKPCVYGEVKVTNVDTTDHLK RYEESSEIVCLGGSLSLTGPTSSIVPYLTKLNIHL RDRVRDIEIARGQIDLDFRTEDESEALKVGFVQG ECGFAAVLYTLRSAVLANIVARFVNKYGFSNTD VAKQYFSTSNAAQKDCLSAVLFEKPPSQYEY EYEEYERVEHGTGLIGLSRSLVVPAYSSLVIEVD LWDHGTLOFVNGTAEFRARRCGNGDCMNLQG QDCVLIKLRVTWVHEDVCLSPMIQNKRELA VTDKRRQKTRTKFGKSPLEPLRGHFFIIGCCASSVKP KRLCKFPMLVFSVFIGQNNKGSAGLYGKIEVLC ALGRLRIKREKDMFQQQLSPEKNSSTLNFRCRGI PGTRLGIVDFKDAHSPLVIQSLWVWCDFSLEY KSWFDKRICSVIQGEYGYAAIHYTIFSAVQAIVK ASISTNYSDFPVQYGSIVAQYTNVSYSTYEN KYRVSLEKTVSEITDFKLSKYVAVPKDSSLII EVDLYLSSIGESYGHGNDVENGTGHIRASAEFE DDGSGSGSGSGSGESFRTIKGQYNIIEFVVEWT DRLMTIKE

10

20

30

40

【 1 2 4 9】

【表 1 4 - 2 5 0】

A	B	C	D	E	F
1	分析に用いたペイトとの相関係数	sp_107660_kqnh (SOAP1)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2					<p>WIPFSDDEFERIKTADDAASVYVYKRRDSTIRDTVA  NLETVKAEADAASITAEQTCSLLEHKYLSLSAEFS  ALESRYSLELNSLQQRVSELAQVQSDKHQLHLK  CIEKDGDIERLSIEVKELOKSKRWVWLVGVKDAE  ISDKKATIKSVYDKITSLGSRSSDREARLSELAEL  VRCKANCDKLLQEKELIEKHINTWLNHEELTIKVDLSL  LDCHRRSADIADFSMKLADMEKQFNECTTSL  RHKEMVRELESKVESLHRELCSARDASAEQQ  FSAQISTSKLADLYKESSEEWSSKSGELEGVKA  LETHLSQVESDYKDKLEKETAAARKEAKVTVLKL  EKLVKCEAELESSRRANELDIIPLSSFTTKDLMYT  SNTNDIVGDNLAIVHIPAGVSGTALAASLLREG  WSLAKMYTYQEAADALRHEQLGRKQSQAILER  VLLIEEKASVILDEREEHDMRMAEAYALMNGKLQ  QSISEKVNLDKLVILEKADLKKRERECTYAQKEIK  DLQTQVTILKECRDYQLRCGLDSNEYMMDVA  ALSDIDDESDTQKVISEHLVSVKDIKGLVEQNAKL  RSLVFSFSDNIENKELQKQYEEELQKHKEEAA  SKVAAVLERAEEOQRMIESLHSSVAMYRRLYEEE  HRLFSSQPLSTEAPSTAGSKDFMLLLEGSQEAT  KRVEQAGERVRFLEELASPKAEIISLRERDKI  ALLENFAREKLEFRKFEFDHLREEMNGVLA RNIE  FSQLVGYQKLRGAEVSTAEELTRKLNMEIS  VLKREKDLLINSEKRAQDEVDLQSKRVYRLQATL  DTIQSTEQWEAARAEEKRQEEYKIKIEREAWAE  VKKELQEEKNKVTLSLDHTTIKDMIQIENLR  KELASALQVVKDAEVRVAVAEARYADLERKLSKSS  EVEVSEQDGRSIPLSSAAEILGDVGFREEIERLK  AESQASKDHMLQYKSIAEVNEAALKKMEASHDN  FRIEAEQSKSLAEIQSLDRVRELENEYTLKSE  EATSVVQKDEALASALGEINNKKEETS NKYSQVL  AALEVQFTMLKDDLEREHQRFALAAQAANYERQVIL</p>
667	sp_049730_dsjn #N/A	-0.379803	-0.920231	<p>nuclear pore anchor (NUA); FUNCTIONS IN:  molecular_function unknown; INVOLVED IN: ooly(A)-  mRNA export from nucleus, stamen development,  negative regulation of protein sumoylation, negative  regulation of flower development; LOCATED IN:  cytosol, nuclear envelope, nucleolus, chloroplast;  EXPRESSED IN: 26 plant structures; EXPRESSED  DURING: 14 growth stages; CONTAINS InterPro  DOMAIN/s: Tetrairicopeptide, MLP1/MLP2-like  (InterPro:IPR012929); Has 30201 Blast hits to 17322  proteins in 780 species: Archae - 12; Bacteria - 1396;  Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses  - 0; Other Eukaryotes - 2996 (source: NCBI BLink).</p>	

10

20

30

40

【 1 2 5 0】

【表 1 4 - 2 5 1】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
668	sp_050810_qjgh -0.883187	-0.946356	-0.905331	retinoblastoma-related 1 (RBR1); FUNCTIONS IN: transcription factor binding; INVOLVED IN: in 15 processes; LOCATED IN: cytosol; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Retinoblastoma-associated protein, B-box (InterPro:IPR002719), Cyclin-like (InterPro:IPR011028), Cyclin-related (InterPro:IPR013763), Retinoblastoma-associated protein, A-box (InterPro:IPR002720).	MEAKELQANFVHLLSKYKRAYQEFRKSSSEAD KDLVDASSSSSSSTIDYYKFGWLLFLALRVH AFSRFKDLVCTNGLVSVLAIHIIIPARLRKFNVL DSPLFEKKTDRGVNLVSSLQCKYDTSDELKMI EKANELILELKKKPCSESOCKTENLENIDPDDLTF FEDLLEDSSLPTCISMLEKDYETTICNKGDIDERV FVNVDLSLASSLSGGSVNISGKTRTFEAMASP TKTISSPLSPCRSPPTHVNGARGGVNVRMMSTPV STAMTTAKWLRYSISPLSPKSPQLEKYFMSCDR NISNVEVYHRAHIMEAIFPNSALGERNVAGNLPNS NLTDNIWAHQRRLEALKLYRVLEAMCSAEQALL HANNLTSLLTNEFFHRCMLACSAELVLAHTKTVT MLFPVAVLEKTGITAFDLSKVIESFIRHEDSLPRELR RHLSNLEERLESVMWEKGGSMYNSLIJAKPAMS AEINRFGLLAERPMSLDAIAVHINMCSGGLPPLPS SIKHETATGQCQFGSPK-LSTFPSPDMSPFKYAKN VYVSPLRSSKMDALISNSKSYACVGESTHEFFQ SPSKDLTAINDRLNSTRKVRGALNFDTDAGLVSD SLVANSLYLQNGNCTTTTSSSVVPLKTEQPDSS MAKSCGDEERVIILSKMLDEFDYLOSEICLDERKNS DRFESRMSEEDKLAAVGNGLPFTATFRDREYEF KYTRRKGVAINEFDFLARIMWQHPVTITGKRFRLMIS MPNSCYQFFVIPENITGSSVVVADSGKPIFRVGLIP NNVKM
669	sp_051030_zndn #N/A	#N/A	-0.873089	uncharacterized protein	
670	sp_052180_skwk -0.880396	#N/A	#N/A	pseudouridine synthase and archaeosine transglycosylase (PUA) domain-containing protein; FUNCTIONS IN: RNA binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Pseudouridine synthase/archaeosine transglycosylase-like (InterPro:IPR015947), Pseudouridine synthase/archaeosine transglycosylase (InterPro:IPR002478), Translation-associated RNA-binding, predicted (InterPro:IPR016437), Uncharacterised domain 2 (InterPro:IPR004521); BEST Arabidopsis thaliana protein match is: eukaryotic translation initiation factor SU11 family protein (TAIR:AT1G1350.1); Has 885 Blast hits to 883 proteins in 291 species: Archaea - 148; Bacteria - 0; Metazoa - 325; Fungi - 155; Plants - 86; Viruses - 0; Other Eukaryotes - 171 (source: NCBI BLINK).	MFRKFTSEYVSAQNVKASVQRRIRQISAEIYPG LEPVLEDLLPKKSPVVKCQNHNLNVLVYVNVPL FFNIRDGPMPTLRLLHQYFNIMKKFQVDRGAIF VFSGANIMCPGLTSPGGVLDDEEVEIERPVAIMAE GKQHALAIGFTKMSAKNIKAINKIGVDNLHYLND GLWKMETLD

10

20

30

40

【 1 2 5 1】

【表 1 4 - 2 5 2】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
				RNA-binding KH domain-containing protein; FUNCTIONS IN: RNA binding, nucleic acid binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro:IPR018111, K Homology, type 1, subgroup (InterPro:IPR004087), K Homology, type 1 (InterPro:IPR004087), K Homology, type 1 (InterPro:IPR004088); BEST Arabidopsis thaliana protein match is: RNA-binding KH domain-containing protein (TAIR:AT1G14170.3); Has 5625 Blast hits to 2559 proteins in 215 species: Archae - 0; Bacteria - 48; Metazoa - 3662; Fungi - 737; Plants - 967; Viruses - 0; Other Eukaryotes - 211 (source: NCBI BLINK).	MAQQGQYTIPTLQPYTQHQQYIPPPHPLQP PHPHQPLNIPYPHQYHHHPHQYADYASVPLQ QHMLPLPIQYPTYPPIPHYSDRLSSKRRPREVSDS AEVPSFKQAVGNDILFRVVPVSGQIGKVGKGG HRVQKIRDDTKSTIKIADPIAREEYEEVVISSRENEA GGISDAEKALLQIASLILADDTGTSAPVIAAATDP YAAAGISGGHAVANSVKLLIAGSQAGGLIGVSGQ NIEKLRNSSGATIAVLAPNQLPCASAHESDRMV QISGEVPAVMKAVEEIGVELRQNPVKQVVISIPTY NLSIARAAQNYMDPNAAEYVTLDMVISETVMVGG IGRAGSNISIRITSESGAMIKYFGGKGENKHRQIQ IGSAQQVALAKQRVDEYIYTLTQTPAAEQSV
671	sp_052260_lhkm -0.910166	-0.934352	-0.827464	ubiquitin-specific protease 17 (UBP17); FUNCTIONS IN: cysteine-type endopeptidase activity, ubiquitin thiolesterase activity, zinc ion binding; INVOLVED IN: ubiquitin-dependent protein catabolic process; LOCATED IN: mitochondrion; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, MYND-type (InterPro:IPR002893), Peptidase C19, ubiquitin carboxyl-terminal hydrolase 2, conserved site (InterPro:IPR018200), Peptidase C19, ubiquitin carboxyl-terminal hydrolase 2 (InterPro:IPR001394); BEST Arabidopsis thaliana protein match is: ubiquitin- specific protease 16 (TAIR:AT4G24560.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MLEIPVFLSLLIFTVLRLWTLKNEETKEVMVSR EAAMADFDFQGFNSNTYAMKAHECALCSSLT STRCSRCKAVRYCSCQISHWRQSHKYECP QGPSASERDNFNGMAEMMEQHDLSAEDFPRM DEKNTICQLNHQVLDLSPVPRLSVTIEGSDIDTRW SSSGENGTYSYDINAPSMQDKKNT LDNTS MGLASNVVMSAKLNREKSRQNDNRKRLSIS GLSSELKDNSFHRAVSDNIHSSKSKELPFVDNE GLKHSQSKHGKVKSMASKHSGNSLLSSGSS SQLPHNAFKSKKVVHQFNTPRQKSSLSHNSLL GISGNPNYKVIQYDAFLSLYTSNDIELIKPFGLTN CGNSCYANAVLQCLAYTRPLTMYLLQGHKSEC WKDEWCFICEFENLMLKGRGYTP_LSPIRSSNIS HLHGREDDEAHEFLRYAVDTMQSCVELAGFVD PVFEEESTLVGMTFGYHSHKIKCMKCHNSGKHE RMMDLTVEIDGDIRTLQOALRKFTASETLEGNK YYCSRRCKSYEKAKKLLIFQAPNLTMLKRFQA GNFTLKRVRQVPEILNITPYMSDSDKYYPQYSLY AVVHLDVSSADSSGHYCVKYNKHEEWFKIDD SVVSVDLERVLSEKAYMVLARDTPRPPASGMS FHDEKPADYSYGSYVMSKVRHSGSDVSDVPWVIQ HQSSKSGYVKDPSYVEEWHIVHVESIPREDSSSE VSSLFSYSEEGTCTASTKDSGSSGWFI
672	sp_052480_wsyg -0.910274	#N/A	#N/A		

10

20

30

40

【 1 2 5 2】

【表 1 4 - 2 5 3】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1) 分析に用いたバイトとの相関係数	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
673	sp_053060_crdw	#N/A	#N/A	-0.821033	<p>pfkB-like carbohydrate kinase family protein;            FUNCTIONS IN: kinase activity, ribokinase activity;            INVOLVED IN: D-ribose metabolic process;            LOCATED IN: cellular component unknown;            EXPRESSED IN: 22 plant structures; EXPRESSED            DURING: 13 growth stages; CONTAINS InterPro            DOMAIN/s: Carbohydrate/purine kinase            (InterPro:IPR011611), Ribokinase            (InterPro:IPR002139), Carbohydrate/purine kinase,            PfkB, conserved site (InterPro:IPR002173); Has 2896            Blast hits to 2893 proteins in 1061 species: Archae -            1; Bacteria - 2475; Metazoa - 103; Fungi - 59; Plants -            64; Viruses - 0; Other Eukaryotes - 194 (source: NCBI            BLink).</p>	<p>MKKSAMPIIIGMVLDIHATASVPSTAKSTTPGQ            NEFILGGVARNIAECLLKGTKSFMISALGCDLPG            DLLLEQWRASAGLLTEGIORGLDITTSVCNMFDI            NGDLTAGVASVESIEKYVTPQWIQFKSIVHSAP            VMVMDANLSPEALKASCRCMAADFVVPWFEPV            SVTKSERVAAIANYITFASPNHEHELIAMANS.SRK            NKFFPIERDQRKMMASIESLFKELKPAILVLLQGI            KYVLVTLGPDGAFCLSGSGSPGSLRECLSLSTKSF            KKREIFDKVSNRCPHQYLGPTPYKKGSQLFVM            HFPALPATVGRVSGAGDCLVGGMLASLCSGLDV            MQSVAVGVAATAIEVKTNPDKYDLTTVADDA            GAVYSAKVVFQESML</p>
674	sp_053320_mitgc	#N/A	#N/A	-0.853838	uncharacterized protein	<p>SSLLWPPTAPTPSYSGSTLTASLSILALTSLSLTPRC            FSYSTTITINDLCECLHNRSADFWRWGREKP</p>

10

20

30

40

【 1 2 5 3 】

【表 1 4 - 2 5 4】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
675	sp_053470_agdp	-0.3820486	-0.833933	-0.3824337	<p>EMBRYO DEFECTIVE 2754 (EMB2754);</p> <p>FUNCTIONS IN: small GTPase regulator activity; binding; INVOLVED IN: embryo development ending in seed dormancy; LOCATED IN: cellular component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages;</p> <p>CONTAINS InterPro DOMAIN's: Citron-like (InterPro:IPR01180); Vacuolar sorting protein 39/Transforming growth factor beta receptor-associated domain 1 (InterPro:IPR019452); Armadillo-type fold (InterPro:IPR016024); Vacuolar sorting protein 39/Transforming growth factor beta receptor-associated domain 2 (InterPro:IPR019453); Has 871 Blast hits to 660 proteins in 186 species: Archaea - 0; Bacteria - 0; Metazoa - 376; Fungi - 292; Plants - 87; Viruses - 0; Other Eukaryotes - 116 (source: NCBI BLink).</p>	<p>MVHSAYDSFEELFNDICPTKIDAVESYGSKLLIGCS DGSRLRIYAPESSSDASPPSYRSTQLELLEKPEYF QERINQFAKRPITAMEYLSKELLISLSESIVHKL SSFETVAISKAGANAFSWDDSRFLCFKQOR RVYFRHDGGGRGFVVKDYGVDPDVTVKSMSWCG ENIGIFRKEYMILNSSTGASSEVFSSGRISPPV VPLPSGGFLGKDNIGVFDQHGKLGQEGRICW SEAFSQVHIEPYAIGLLPRHVEIRSLQQFPFIQT VGLRNVRHVRFSNGAILVSLKESINGLFPVPLGAQ IVQLTASGNFEEALSCKLLPPEDSSLRAAKESSI HIRYAHYLFENGSEYEEAMEHFAGSQVDITYVLSL YPSITAPKSTALTEPQKVIDFGYDASDLSRCSGGI SDDMEASPRASQSDLEENATIVESKIVSHNTLMALI KFLQKKRYHVIEKAAAEAGTDEAVSDAFGTFKFSY DSNRSKSNKGRGSIPISSSGRERAAILD TALLQA LILTGSSAAMELLRGVNYCDLKISEEFLHRNQP TVLLELYKCNMHRREALKLLQLEESKSEHSRPE LTQKFTPEMIDYLPKLCGTDPMVLVEFAMLVLES CPSQTIELFLSNGIQADLVNSYLKQHFAPNMQATY LEHMLSMNEDGISGNLQNEWQYLVSELDVWYK ELSSQQQWDETAYSPTPKLLSALENSIGYNPE GILLKRLPADALYEERAMLLGKLNQHELALS.YVH KLHVPELALSYCDRVYESQLNQSRSKOSNYLT LLOIYNPRRSTRNFEKQNYLVSTPNAEKRVNS SASIRSKVGRSTRKIAEIEAGEDVRFSPSNTDSGK SDGDADAEANYEGGSNIMLDEALDLLGRRWERIN GAQALKLLPRETKLKNLLSFLGPLLRKSSSEAYRN FAVIKRLRESENLOVKDELYNKRKAVVYKVTGDSM CSLCKKIKGTSVFAYVYNGTTLVHFVCFRDSQSM KAGRGSQLRKISQ</p>
676	sp_054420_ycik	#N/A	#N/A	-0.652133	<p>MAIGO 1 (MAG1); FUNCTIONS IN: hydrolase activity, acting on ester bonds, protein serine/threonine phosphatase activity; INVOLVED IN: endosome to lysosome transport, protein targeting to vacuole; LOCATED IN: microsome, retromer complex, membrane, multivesicular body; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN's: Phosphodiesterase Mj0936 (InterPro:IPR000979); Has 35333 Blast hits to 34131 proteins in 2444 species: Archaea - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLink).</p>	<p>MVLLVLAIGDLHIPHRAPDLPAKFKSMLVFGKIQHII CTGNLCKIEVHDYKLTCLPDLHITRGEVDEDARY PETKTLITGQFKLGLCHGHQVVPWGDLDLAML QRQLDVLIVTGTHTHQFTAYKHGGGVVNPQGSAT GAYSSITYDVNPFVLMIDIDGLRVRVYVYVELIDGE VKYDKIDFKKTTAFQSPH</p>

10

20

30

40

【 1 2 5 4】

【表 1 4 - 2 5 5】

	A	B	C	D	E	F
1		分析に用いたバइटとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
677	sp_054540_thrj	-0.901185	-0.883243	-0.928636	methyl-CpG-binding domain 9 (MBD9); FUNCTIONS IN: methyl-CpG binding, DNA binding; INVOLVED IN: photoperiodism, flowering, secondary shoot formation, regulation of transcription, DNA-dependent; LOCATED IN: nucleus; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Zinc finger, PHD-type, conserved site (InterPro:IPR019786), Zinc finger, RING-type (InterPro:IPR001841), FY-rich, C-terminal (InterPro:IPR003889), Zinc finger, PHD-type (InterPro:IPR001965), FY-rich, N-terminal (InterPro:IPR003888), DNA-binding, integrase-type (InterPro:IPR016177), Zinc finger, FYVE/PHD-type (InterPro:IPR011011), Methyl-CpG DNA binding (InterPro:IPR001739), Zinc finger, PHD-finger (InterPro:IPR019787); BEST Arabidopsis thaliana protein match is: RING/FYVE/PHD-type zinc finger family protein (TAIR:AT1G7250.1); Has 6416 Blast hits to 3988 proteins in 224 species: Archae - 0; Bacteria - 0; Metazoa - 4085; Fungi - 602; Plants - 1260; Viruses - 0; Other Eukaryotes - 469 (source: NCBI BLINK).	MKLNDSFPIVAQSSSTPSSPSSSHRSLAI DLNEVPSSETLPSSSSSSTFILSPTLSLVRAND NPPIEGPPAEIPDDVTPCGCGGSPSELVSDRL VDCGERGFHNCVTGLPVVREEWLGRYCVAE GVGSKRWRLGRKEDRGVRLIDMNASPPSDGDV ESLAAADAAAAAAGSSREVFGFM
678	sp_054640_zodi	-0.855161	#N/A	-0.877034	PLAC8 family protein; CONTAINS InterPro DOMAINs: Protein of unknown function Cys-rich (InterPro:IPR006461); BEST Arabidopsis thaliana protein match is: PLAC8 family protein (TAIR:AT3G18470.1); Has 865 Blast hits to 864 proteins in 99 species: Archae - 0; Bacteria - 0; Metazoa - 128; Fungi - 94; Plants - 626; Viruses - 0; Other Eukaryotes - 17 (source: NCBI BLINK).	MEMIGISVPPPPYIPLSQSEEESEALITGNNSVQ ASASISNGTVQWSSGICACCDDSTSCCIGFCPC YLFKNAEFLGSETLVGPCMAHMLWALASSLS CLYCVLLGLPGCLLSCYACNYRILRTRKYNLPEAP CGDFATHLCHLCCALCQEQYREICERSGDPNYS PKLPVITAPPVQKMEASPG

10

20

30

40

【表 1 4 - 2 5 6】

	A	B	C	D	E	F
1		分析に用いたバइटとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
679	sp_055210_azxq	-0.857227	#N/A	-0.878538	NDR1/HIN1-like 8 (NHL8); FUNCTIONS IN: zinc ion binding; INVOLVED IN: biological_process unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ubiquitin-conjugating enzyme/RWD-like (InterPro:IPR016135), Zinc finger, RING-type, conserved site (InterPro:IPR017907), Zinc finger, C6HC-type (InterPro:IPR002867), Zinc finger, RING-type (InterPro:IPR001841), RWD (InterPro:IPR006575); BEST Arabidopsis thaliana protein match is: RING/U-box superfamily protein (TAIR:AT3G14250.1); Has 3653 Blast hits to 3199 proteins in 223 species: Archaea - 0; Bacteria - 0; Melazoa - 1736; Fungi - 604; Plants - 670; Viruses - 7; Other Eukaryotes - 636 (source: NCBI BLINK).	MREKVRQFEAEVEELSEEQIQINDQVQDELLAM ESYGDWYILDKRQGLRFSQIHINIELPDSYSVST KLKEDSLQTKVKVQCLSPILLTCLLPRSYPSHYPP SFTISVQWLHESEISQLCSMLDLLWNEQPGQEIL YRWVEWLPSSSLQHLGFDEELSIGPYNVVESVES VSDRRSFGSISPDIDIVASMKSYNSERQHENFLN GCHECCICFSEFAGTKFLRLPCDHFVCKMTTY AEVLVKEGTVNKLQCPNKTGGMIPPGLLKRLLG GEEFDRWESLMLQKTLDAMPDVVYCPFCETACL EEKDHHQQSKCFFSFCSLCRERRRHVGLACMTP EMKLSLEDQPIFSLKESQRKQRIVMNELLSV KEILRDAKQCF5FFISRLTGMCDCHHSPAEPVIE PEQNMALQVROILRELQVMVHGN5RPFPCPSCK QFKAKVSLYIQ

10

20

30

40

【 1 2 5 6】



【表 1 4 - 2 5 8】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
681	sp_055650_fznp	-0.851607	-0.877124	#N/A	<p>RNI-like superfamily protein; FUNCTIONS IN: ubiquitin-protein ligase activity; INVOLVED IN: ubiquitin-dependent protein catabolic process; LOCATED IN: endomembrane system; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAIN/s: Leucine-rich repeat, cysteine-containing subtype (InterPro:IPR006553); BEST Arabidopsis thaliana protein match is: F-box family protein (TAIR:AT5G27920.1); Has 15959 Blast hits to 6468 proteins in 357 species: Archae - 0; Bacteria - 920; Metazoa - 6194; Fungi - 1434; Plants - 4975; Viruses - 16; Other Eukaryotes - 2420 (source: NCBI BLINK).</p>	<p>MAKTSSTSSPSSPSTTNVFDLISEELVFHILDYL NDNNPQDKKSFSLVCKSFYSIESKHRRTLKPLRS EHLPSVLNRYPQISTLDLTLCPRTDFTLSVWATTA CVNTRFVDLSRSKFFTGNGLAALVASCRLTAL DLSNATELFDLAAAAGIIEAKNLEKLWLGRCRKLITD IGIGCIAYGCPKLIKIGLKWCLGIGDLGVELIAVKC KDLRTFDLSYLPDVISFVACAFSGVITSKCLPAILK LQYLELLEGGCFGIDDDSLTLKQGSKSLKTLDIS SCQNWSDTGLSSTLSTGIGCLOQLTLAYGSPVTLD LAFSLQRLSKLQSIKLDGCLVTCGLQAINGYCA SLRELSLCKSGVTDEGLSSLVKQKELKLDIT CCRKLTCYSIAHVAYSCTSLSRMESCTLVPRE AFVLGRKCHLLEELDLTDNEIDDEGLKSKSCSNL SVLKGICLNTDVGLSYIGDFCSKLVLEIDL YRSMID ITDSGITAIARGCPLMNTSLGIAAVGCKHLTKLDI CAKLTLECRGCPMLTSLGIAAVGCKHLTKLDI KKCYNINDFGVLAHFQSNLKEINMCSITSIDVG LLSLSSISCSLOSMTLHLKGLSPRGLASALLACGG LTKVKLHSSFRKTLLPQPLFEHLEAFRGCVFQWPREK MFQDELEDSSKSWKLVQVAMNINDS</p>

10

20

30

40

【表 1 4 - 2 5 9】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
682	sp_056150_tflf	-0.2833222	-0.894813	-0.881918	DNA-binding bromodomain-containing protein; CONTAINS InterPro DOMAINs: Bromodomain conserved site (InterPro:IPR018359), Bromodomain (InterPro:IPR001487); BEST Arabidopsis thaliana protein match is: DNA-binding bromodomain-containing protein (TAIR:AT1G76380.2); Has 11096 Blast hits to 8233 proteins in 323 species: Archae - 0; Bacteria - 121; Metazoa - 4363; Fungi - 1457; Plants - 677; Viruses - 7; Other Eukaryotes - 4471 (source: NCBI BLink).	MGKVVVEKKKKKGRPSLLDLKRSLEKQEQEQ QHQQSQKQKQFQNNVKKIPNSKIPNPNSSA GIRRSRRNPNPNSDIDEELPSGVGKKKQQLKAV KLLPNNGGDEDEEVNSDEGRENNRSESGPE GDASPLNLRKRKIDSISHSGSVEETDKSEKQGS AVATTKOASTQECTLDDGSMPLPDKSLLFL DRLOKDTYGVFADPVDPEELPDYHEVISDPMD STYRKKLANGAYATLEQFENDIFLICSNAMQYNA ADTYFRQARSIQELAIRNFDNLRADSDDNPEPE PKVVRGRPPTKNLKTTPGPPLEFRANSDFPEA TLATAGGTTNRSNNEVRKASHSEKFGSTDSFSR SLYGSRVGETYIGWSAERYERADDMSSGASKGI LMKHKKHIVLDETRRNTYNSYSSACGQEPSVLT AFDRERKQLIALDCGGLSCYLARLSMISGYQPKH HNPEDELDTAYWRGPAFNAFCFLTESNRVSMDE EIFWIREY_LPKLTNVGLHTEYGYARSLARFAAK VGPVAVKVASKKIERCLPTAVKFGPGWVGENEA PAERLQQQSPASVGGKTPPPSSLPAPTDTSSA HTQGIADSSGDSV_LKTEGEQLSESQKPCNTSN VNEESGKPLCPPVAGNADATPLSSNVPSTGAGP NLIAPEASAPPSTVWNLINNGSSMRPENPCPI NQTWSVHRSVNHGHNNGSYGLNFAAQMGKMGIA GRPTMFINLOSNPSTVSRTEIQPATDSNGLGE AKLENSCTKIPIDGLQKGTGNTQKVGSTPLDLNVG FQSPGSPVCGGKADSVQTDLALQL
683	sp_056190_lguh	#N/A	-0.8839333	#N/A	uncharacterized protein	MLANEPLHPLSLS
684	sp_056270_qrni	-0.9213358	-0.922287	-0.927139	SULPHUR DEFICIENCY-INDUCED 1 (ATSDH1); FUNCTIONS IN: binding; INVOLVED IN: cellular response to sulfur starvation, regulation of sulfur utilization; LOCATED IN: cellular_component unknown; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAINs: Tetratricopeptide-like helical (InterPro:IPR011990), Tetratricopeptide repeat-containing (InterPro:IPR013026), Tetratricopeptide repeat (InterPro:IPR019734); BEST Arabidopsis thaliana protein match is: Tetratricopeptide repeat (TPP)-like superfamily protein (TAIR:AT1G04770.1); Has 255 Blast hits to 246 proteins in 34 species: Archae - 2; Bacteria - 37; Metazoa - 0; Fungi - 0; Plants - 177; Viruses - 0; Other Eukaryotes - 39 (source: NCBI BLink).	MEKRNRSPEGESLHAYVKVPSGDPYVYRAKH AQAINAGDRVDSALKDMVMVKQLDRTEEAIEAI MSFRGLCSRNAQESLDNVLIDLKCKGKVDQID LLFRKRLIYHGEAFNGKPTKTAHSHGKFKFQVSIK QETSRIKLNGLWAYMQKTNMYMAAEVVYRKAQMI DPDANKACNLGFLIKQARYEARSVLEEVVEAK LPGSEDIRSRNPAEELLDDLESROPPSLLSNIVCL DLEDEFVHGLELDESKWGLPRLPKRPLIFEEISNF KNQLAC

10

20

30

40

【 1 2 5 9】

【表 1 4 - 2 6 0】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
685	sp_057280_dzei	-0.983778	-0.9933354	pleiotropic regulatory locus 1 (PRL1); CONTAINS InterPro:IPR019782, WD40 repeat, conserved site (InterPro:IPR019775), WD40 repeat (InterPro:IPR001680), G-protein beta WD-40 repeat-region (InterPro:IPR020472), WD40 repeat-like-containing domain (InterPro:IPR011046), WD40-repeat-containing domain (InterPro:IPR017986), WD40/YVTN repeat-like-containing domain (InterPro:IPR015943), WD40 repeat, subgroup (InterPro:IPR019781); BEST Arabidopsis thaliana protein match is: Transducin/WD40 repeat-like superfamily protein (TAIR:AT3G16650.1); Has 84704 Blast hits to 34179 proteins in 878 species: Archaea - 64; Bacteria - 9610; Metazoa - 34201; Fungi - 18356; Plants - 10881; Viruses - 0; Other Eukaryotes - 11592 (source: NCBI/Link).	MTQTTPSAPKDDSNRPEATASPINKPENSPEER ENKNLSRRVKDVEISPIVYGTAFWLGKATKES QSHRWTVVYRGGTNEEDLSVVVKRVVFLQHPFSN NPIRVVDSPPFSECGWGEFEIATLFLYFSDAG DRKLDVHFHKLKYLEQNGPQSTKPKVVMETVD EVVFPDFSEQFLARVQNHFAVVVPRLPFAGLELPS SASTEERRGSMKDPHGLGHVFMVNFSEADELLKLA AARQQVQAHIVKLRQMSMVMVEGSSGTADPASAF
686	sp_057580_pspt	-0.963398	-0.992154	GLIOMAS 41 (GAS41); INVOLVED IN: regulation of transcription, DNA-dependent; LOCATED IN: nucleus; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: YEATS (InterPro:IPR005033); BEST Arabidopsis thaliana protein match is: TBP-associated factor 14 (TAIR:AT2G18000.2); Has 807 Blast hits to 807 proteins in 219 species: Archaea - 0; Bacteria - 0; Metazoa - 350; Fungi - 284; Plants - 64; Viruses - 0; Other Eukaryotes - 109 (source: NCBI/Link).	IMKYGVYIYTHRSHTYLTKAIFLPRPEEENOTEKIE SSGRRKRKEASNDGQGGGGGGVGERNREKE KRHHHRHQEHKSKKMSSTTQQVRASSHILKHKQ GSRKASWKDFEGVILNTTKESAIQLKSIRDDI LSGNANFEDVASRISDCSKAKRGGDLGPFGRPQ MQKPFEDATFALKVGEISDVIETDSDSGVHHIKRTG
687	sp_057620_rwai	#N/A	#N/A	peptidylprolyl cis/trans isomerase, NIMA-interacting 1 (PINT1); FUNCTIONS IN: peptidyl-prolyl cis-trans isomerase activity; INVOLVED IN: regulation of cell cycle; LOCATED IN: plasma membrane; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Peptidyl-prolyl cis-trans isomerase, Ppic-type (InterPro:IPR000297); BEST Arabidopsis thaliana protein match is: FKBP-like peptidyl-prolyl cis-trans isomerase family protein (TAIR:AT1G26550.1); Has 6394 Blast hits to 6305 proteins in 1691 species: Archaea - 14; Bacteria - 4880; Metazoa - 261; Fungi - 147; Plants - 126; Viruses - 0; Other Eukaryotes - 966 (source: NCBI/Link).	

10

20

30

40

【 1 2 6 0】

【表 1 4 - 2 6 1】

A	B	C	D	E	F
1	分析に用いたバイトとの相関係数 sp_107620_kqnh	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
688 sp_057830_znmu	-0.8882	#N/A	-0.888188	fringe-related protein; FUNCTIONS IN: transferase activity, transferring glycosyl groups; INVOLVED IN: biological process unknown; LOCATED IN: endomembrane system; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINs: Maleicim/receptor-like protein kinase (InterPro:IPR021720); BEST Arabidopsis thaliana protein match is: Protein kinase superfamily protein (TAIR:AT5G24010.1).	MNSSIIPAFSLSLPLLLFLFLEVSSTPTDHHLI NCGSSAPSTIDADHRKFFGDESSLPHSVRLSASQ TIPLTDLEFKLSDVSPLYNSVFRKPKSKYTFRV KEKGALVRLHFRHFDSSNFDLSDAQFHVLADG YLLNFEFSLTMKSLIIRIYILWVDSDKLELSFKP LKNFNGFVNAIEVISAARDLILDSAQFVDSNGVIK IDGIRSFAFETMFRVNVGGAKVTFFNDSLWRTW TTDDGYLESDDHGSKEAHFGGKIKYQLGGASREV APDNVNSARVSVSSSSGVAKEYLTWGFPIVDG YKYLVRTHFCDIASPLVGLSLYFNVYVNGDMSIENL DLSTITNSLSPYYADVFVNGENGVTVSIGHSKG SSSSGVALLNGIEVFKMNSMNSLSDGMVSAKSI LFRRCRGGSSFLVVAAVGLLVAASVYLRKRLM AASDSVAWSPFLPYDISEVSLKSTYL
689 sp_058230_ygyd	#N/A	-0.821798	-0.822099	Flavodoxin family protein; FUNCTIONS IN: oxidoreductase activity, FMN binding; INVOLVED IN: oxidation reduction; LOCATED IN: cellular component unknown; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 7 growth stages; CONTAINS InterPro DOMAINs: Oxidoreductase FAD/NAD(P)-binding (InterPro:IPR001433), Ferredoxin reductase-type FAD-binding domain (InterPro:IPR017927), Riboflavin synthase-like beta-barrel (InterPro:IPR017938), FAD-binding, type 1 (InterPro:IPR003097), Flavodoxin/nitric oxide synthase (InterPro:IPR008254), Flavoprotein pyridine nucleotide cytochrome reductase (InterPro:IPR001709); BEST Arabidopsis thaliana protein match is: P450 reductase 1 (TAIR:AT4G24520.1); Has 7053 Blast hits to 6598 proteins in 1587 species: Archaea - 9; Bacteria - 3832; Metazoa - 1042; Fungi - 865; Plants - 492; Viruses - 0; Other Eukaryotes - 813 (source: NCBI/BLINK).	MEEKPKLILYATQTGNAMDVAERIGREANRRGC PVSLSIDOFDPGLLPNORAVIFWSTTGGQDTP DSLKVFWRHLLQRHLTRDWLGGVHHAVFGLGD SGYOKYFVAKKLDKRLSDLGGTPIIERGLGDDQ HPSG YEGSLDPWMSLLWSSLYQINPFLPIGVYDV VPDMKLLDQSKVSVRYHVKDGASLJYSTTPDLES RASQIERARLMSPARFPFSDKSRPHCWLMKTRNC LLTRAGSGRDVRHLEFGFHSIEVEYEGDVVEML PSQNPDAVDAFIRRCNLDPESITVQLRDSNPDA DVDADAASGLPVLRLTVESTMVDAVASAPRRYF FEVMSYFATAEHEKERLEYFASAEGRDDLQYQN QKERRTVLEVLEDFPSVQMPFEWLIQLPPLKTR AFSISSSPIAHNEVHLTMSVSVSWVYTPFKRKYG LCSKWLAGLDPOQGVFVYVWFQKGLLAPSPSL PLLIGTGTGCAFRGFVAERAVQSGYGTAPILF FFGCRNKEDFLYSDFWLHSDQGGVLSSEKGG GGFYVAFSRDQPRKYVYVQHKMCEHGQRVWDL SNGASVYIAGSSTKMPDSDVMSTFEEIISKESGYP KEAAARWLRALEKAGRYHYEAVS

10

20

30

40

【表 1 4 - 2 6 2】

	A	B	C	D	E	F
1						
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
690	sp_058260_urcy	-C_850:86	#N/A	#N/A	<p>RADIATION SENSITIVE23D (RAD23D); FUNCTIONS IN: damaged DNA binding, ubiquitin binding, proteasome binding; INVOLVED IN: response to cold, nucleotide-excision repair; LOCATED IN: nucleus; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Heat shock chaperonin-binding (InterPro:IPR006636), Ubiquitin-associated/translation elongation factor EF1B, N-terminal (InterPro:IPR00449), UV excision repair protein Rad23 (InterPro:IPR04806), Ubiquitin-associated/translation elongation factor EF1B, N-terminal eukaryote (InterPro:IPR015940), Ubiquitin (InterPro:IPR00626), XPC-binding domain (InterPro:IPR015360), Ubiquitin supergroup (InterPro:IPR019355), UBA-like (InterPro:IPR009060); BEST Arabidopsis thaliana protein match is: Rad23 UV excision repair protein family (TAIR:AT3G02540.3); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).</p>	<p>MKIFVTLKGTNFEIEADSDSDSVDVVKKSIEKSQG ADVYPASLQMLIFCGKVKDKDATTLGENGVDCSF LVVMISKNKFAAAGASSVSAAAPAPANQAQPS STTPAAVAPASVAPQAQTGTATEPVAAPAAPA ASGDENVYGAASNLVAGSSLETTIQQLDMGG GSWDRDTRVRLRAAYNNPERAVEYLNGIPEQ AEVPPVARTPAAGPTDLQVDDPLVQALNPLAEA GNPLAQAPQAAVVPATGPNASPLDLFPQGLPNVG STAGGASSLDLFRNSQQFQALRAMVQANPQILQ PMLQELGKONPQLMPLIQEHOADEFLRLINEPVEG EGNLLSQFASAVPPTVTSVTTEEREAGRLEGMGF RRATVLOVYFACNKNEELAAANYLLDHMDEFDNE EDEGLS</p>

10

20

30

40

【表 1 4 - 2 6 3】

1	A	B	C	D	E	F
		分析に用いたペアとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)		注釈	aa seq MTGSHKFRGGRGARGGASANKRINSYFFPPFPQ RPSAGRPSLSSRIPPSSSSAATPPATQSEESFS LVPGNPLSFTAIRLAPDLIDEIKRVETQGGAAIK FDSNPNPPGNVIDVGGKEYRFTWSHEMGLC DIYEEROSCEDGNGLLVESGCSWRKLNVRVLD ESTKQVKKRSEYEQKLKSKRAVLDQCNPMS KSQMQALAAAEAPWKSPPKPKRFFPEPQVSG HPKPVKSTTVSSSTATAKGRQSTSPFPQNS VPLTSFVRRTATESQIVAEINTTNOTREKENV SSGNVFPQVSGGMEAVVPRDQAGISDLWN LLVTLKESPKGMSIKALEKAVGDMTNSLRKIEP ILKKIATCEPSGRYILKSGVELESUKRSYSDIGSSP EDNNQKQFVEDNHGGVPAPDSGIAGKSPAQDV EEGVQILDEELNVIEKFDASQQFTDLFGDDEPLD NNEGQAMRTNDNKCESDSDSGSSSHSKEASDVY SRSPAGSGSASSDSDGDDSSSHSKEASDVY DIMSDDKPEPNLLGTEHGLSTAPVPWSMPEVM PLQDGRNGDQGYASEVDVDIENDFPDNDVOLA TAPRNLGIITEKLRVSPDHDQNPCKNADVITEH DVGDDEFQRENHSFAKSGGKSRGSLIQFEHNG DHDNRFKENSQPSMSRHLGSPPKHFAADLINED IHESQNLQKQPNRDDSANAHLGQVQNGYNYVV PSKSAVESLQSGRRPAAARSYNTPEIDKAVKNPSN LGVLELKFQPNSTQEGYSKHEKDYGEAQDIEE DGSNEKVNKGYKESPRGQKHSRPESHRKHHS EWDKNSHSLGTSSRLRTGSSLKDSGKAFFDRSP ADRRLLHRELSDELREPLPEEPAGNKRQFD RLGSRFQSETELTGTSVSRMPELKSSTSKPIVDS GNTSPDFAGAPNISYSSRRRSLRLEPHVEDNYR PHGDLQSQ.LQGSRSQGAEVGSLPIKVVVDHSRSG YTENVANQGTAEAGYGETHKKKSQINSQRDSRQ EIGSRKKGKSGSRKSNASKDANDLTKKKKDASD
691	sp_059170_rqsy	#N/A	-0.374324	#N/A	dentin sialoprotein-related; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Occludin/RNA polymerase II elongation factor_ELL domain (InterPro:IPR010844); Has 17175 Blast hits to 8691 proteins in 788 species: Archaea - 88; Bacteria - 6651; Metazoa - 4182; Fungi - 1520; Plants - 456; Viruses - 109; Other Eukaryotes - 4169 (source: NCBI/BLink).	

10

20

30

40

【表 1 4 - 2 6 4】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MLCRRVYVFDLCSYFLCAVSDLEKVDVGGVGT SNOAALVSVCSDDGVLVWVSRGSGHCRRCRELP PWVGSFVIQALPTNKRYICVASCFTDIDFASDP DYVDSLEGSEVTGDEESLSRKAPKCTIIVDSYSL NIVQTVFHGHLISGP_KFMALVESVDDSQKHSTFI IDSFKSOQIPIKDPNLEEGGAKLHRSSALENIAI LNNGSYDGLVLSVATYGGQLGLVYKTSVFFLV ASATILGVISFEDSHHGGADCILSCVSGGMFFTSE SSDSTCMPESRKIHITGLKFSVWDDRGSVIYEL SYEDDTFKYEVINQIPATTCPINVKLSSKFVQLKQI FLRVESICFHNEEPLWRPRVTIWSLHQVDDDR ENFGKICRLLGEGDFLDWTDGSSYVINSEASKH NSGGIPATEEAPMTTMASALDQGVAGTLHKNDL NCTFIQKESIVSSTMVISLNAHIPYAIYVGYHSGQI ELVHFNLAEEKDSQGPKKQFLGHTGYPVLCLAA HQMVATSTERICSKILLSGSMDCTLRVWDIETCN AIMVMHHVAPVYRQIILPPLQSYHPWWDCLFSLVG EDSCVALVSLLETRIERMEFPGHPSCLKVVWDG VRGYLACLQSNQSVASGTKDTLYWDIKAGSRER VLRSTAAHSMFNFCGMDISINIGHAFSYNSTR SSLPLPSTEGCCSLSPRLNVPFISSEVNLMSL NNSSTRSVKQNTGNSDKLPIRCSCPIGAILSF LSSLLSPDICGDVGGNFPNPLNGSQTSINSERG TISNDSSEKGFQDMKDNOCIDITFRSCLLOFSL SFLHLWGVDFFELDHLTMEVMVKRYENVMLAPG LEGDGFSTIAPAFSSALQLWRSSSEFCAMRSL TILSLAQSLVLSHSFSGASSALSIFYTRNIAEKV PSVKPPSLQLLVSWQDEDEHVRLLAARSLFHCA ASRAIPPP_LCCLTAENTKPVNSVDQVQGSLELN EKFAGTWASFLPEMSTEVOGASLQVGGQVDIRR WLESYELQDWISCVGGTSQDAMASHIIVAAALAI WYPSLIKPSLSMLVHPLVKLVAMMNEKYSCTAA MDIQESAQRITLGMVLFKRFDFDEQKVSLDL DADGLLGSVEGKYGILLRVP MLDTPFGDDNDGASGGONGSQFOMNFSMIKAI CNEVMKAMKQKQANNNGSAHSVAFANSAGISFE CSISCNLVFGDSWIDSGATDHTMTYNAKLFKSRK TLKIPICLPDAGSNQKLGERNPILHIRHLGVFS KTLTAKLELPNYRAAAGQPPATVSPVSSATPGW LPMRVVQAHLH
692	sp_059430_zjkr	-0.9399391	-0.959314		Transducin/WD40 repeat-like superfamily protein; FUNCTIONS IN: nucleotide binding; INVOLVED IN: biological process unknown; LOCATED IN: cytosol; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN(s): WD40 repeat.2 (InterPro:IPR019782), WD40 repeat-like-containing domain (InterPro:IPR011046), WD40 repeat, conserved site (InterPro:IPR019775), WD40-repeat-containing domain (InterPro:IPR017986), WD40/YVTN repeat- like-containing domain (InterPro:IPR015943), WD40 repeat (InterPro:IPR001680), WD40 repeat, subgroup (InterPro:IPR019781); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).
693	sp_059670_uchn	-0.877004	#N/A	#N/A	uncharacterized protein
694	sp_059830_wrtw	-0.977186	-0.976152	-0.999405	uncharacterized protein

10

20

30

40

【 1 2 6 4】

【表 1 4 - 2 6 5】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MEGGDTGEISALGFLRVPFSSFLPDLAAGSGS QILFDVNLGRQIRSFVFEGRVHGIAAGLSEEN ALSSVTFKVIWISGERRVKLFLRQLGSLDKPG VGTDLTFLQLLPRFSSHVVLDGFLDGLSSSEEG FRYLAVGCNDNSVHWDAITSSFIFEICPERCLL YSMKIWGRDFRALLIASGTINEIVVWKLQKNRK QPPTDVGGQDFSVFEDMHP_LDCLHYEVDIISKLA GHQGSIFYLSWSSDGTKLVSDDRSARMWITN AEMDMITDLTKLHGHSAGPVLFGHGARIWDCG ISDSFIVTAGEDCSCRWGLDGKELRVIKEHIGR GIWRCLYDPNLSLVTAGFDSAIKVVHCLHSSFA VSGESIKGVKEREAEVFRVCIPNSSGGQMLMDS KSEYVRCFKFARKDMIVYATNHGYLYHAKLSESG TVQWTELRITEEVP_LICMDLFSRNLSSEFEDWISV GDGRGIVKVIKVVFRDGAQVAVRISWSAEMER QLLGVYWCALGSGGFIFTSNPRGLVKLWKLSDS VSHDAVRNCDVYLVASFMSPPFGARIMCLDASYE EEALVCGDLRGNLILFPLSKRILHDSSILSIEMTPL TYFKGAHGISTVCSISIVRLTSSLVEICSTGGDGI CYLEYNRDQQTVEEVGMRHAKELSLVQSFCSYR SSPDDSANISYAVGFSSTDFLIWNLLSETKVLQVK CGGWRPSPHSYVYVQTPICSGGFAYKDEIIVHR YWIPIGEMKIFPRNLHLQFHGREIHSCLCFIPDDVQ CATHLHPKLSWVATGGEDGTVRLTRYNHGTGK WALSNLLGEHVGGSAVRSLCFVSKIVSAETKIEA TDGTTGGANNSLDKIENPCLLISVGAKRVLTSWLL KKSQRIEGDSVADGLEKCKDCKSHKSSPGHMSMS FKWLSSDMFVKVTRTKKAGSDGNLARPSDEY PGSEAVSSMEIGLKHQHGDAQDNDWRVAVYTSF LVNNAKSRFTVCFVYVACSDATLTLRALVLPYRY WFDVASLAPLSSPVLALQHAVAPLCCLYQKMM TKNIIISGSDGSIASFWDVTEISAEAFMQRVSSLH
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MEGGDTGEISALGFLRVPFSSFLPDLAAGSGS QILFDVNLGRQIRSFVFEGRVHGIAAGLSEEN ALSSVTFKVIWISGERRVKLFLRQLGSLDKPG VGTDLTFLQLLPRFSSHVVLDGFLDGLSSSEEG FRYLAVGCNDNSVHWDAITSSFIFEICPERCLL YSMKIWGRDFRALLIASGTINEIVVWKLQKNRK QPPTDVGGQDFSVFEDMHP_LDCLHYEVDIISKLA GHQGSIFYLSWSSDGTKLVSDDRSARMWITN AEMDMITDLTKLHGHSAGPVLFGHGARIWDCG ISDSFIVTAGEDCSCRWGLDGKELRVIKEHIGR GIWRCLYDPNLSLVTAGFDSAIKVVHCLHSSFA VSGESIKGVKEREAEVFRVCIPNSSGGQMLMDS KSEYVRCFKFARKDMIVYATNHGYLYHAKLSESG TVQWTELRITEEVP_LICMDLFSRNLSSEFEDWISV GDGRGIVKVIKVVFRDGAQVAVRISWSAEMER QLLGVYWCALGSGGFIFTSNPRGLVKLWKLSDS VSHDAVRNCDVYLVASFMSPPFGARIMCLDASYE EEALVCGDLRGNLILFPLSKRILHDSSILSIEMTPL TYFKGAHGISTVCSISIVRLTSSLVEICSTGGDGI CYLEYNRDQQTVEEVGMRHAKELSLVQSFCSYR SSPDDSANISYAVGFSSTDFLIWNLLSETKVLQVK CGGWRPSPHSYVYVQTPICSGGFAYKDEIIVHR YWIPIGEMKIFPRNLHLQFHGREIHSCLCFIPDDVQ CATHLHPKLSWVATGGEDGTVRLTRYNHGTGK WALSNLLGEHVGGSAVRSLCFVSKIVSAETKIEA TDGTTGGANNSLDKIENPCLLISVGAKRVLTSWLL KKSQRIEGDSVADGLEKCKDCKSHKSSPGHMSMS FKWLSSDMFVKVTRTKKAGSDGNLARPSDEY PGSEAVSSMEIGLKHQHGDAQDNDWRVAVYTSF LVNNAKSRFTVCFVYVACSDATLTLRALVLPYRY WFDVASLAPLSSPVLALQHAVAPLCCLYQKMM TKNIIISGSDGSIASFWDVTEISAEAFMQRVSSLH
695	sp_059860_e/ns	-0.381104	-0.933323	Transducin family protein / WD-40 repeat family protein; CONTAINS InterPro DOMAIN:s: WD40 repeat 2 (InterPro:IPR019782), WD40 repeat-like-containing domain (InterPro:IPR011046), WD40 repeat; conserved site (InterPro:IPR019775), WD40-repeat-containing domain (InterPro:IPR017986), WD40/YVTN repeat-like-containing domain (InterPro:IPR015943), WD40 repeat (InterPro:IPR001680), WD40 repeat, subgroup (InterPro:IPR019781); BEST Arabidopsis thaliana protein match is: Transducin/WD40 repeat-like superfamily protein (TAIR:AT2G43770.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archaea - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	

10

20

30

40

【表 1 4 - 2 6 6】

	A	B	C	D	E	F
1		分析に用いたペアとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_ktgg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
696	sp_060250_reut	0.85791	#N/A	#N/A	<p>cysteine-rich RLK (RECEPTOR-like protein kinase) 41 (CRK41); FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: endomembrane system; EXPRESSED IN: 12 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Protein of unknown function DUF26 (InterPro:IPR002902), Serine-threonine/tyrosine-protein kinase (InterPro:IPR001245), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: cysteine-rich RLK (RECEPTOR-like protein kinase) 29 (TAIR:AT4G21410.1); Has 118247 Blast hits to 116820 proteins in 4488 species: Archae - 94; Bacteria - 12773; Metazoa - 43795; Fungi - 10171; Plants - 33437; Viruses - 407; Other Eukaryotes - 17570 (source: NCBI BLINK).</p>	<p>MIVRSFNTGKVKSLKLVCLKVACQISIVYAFKRR          NWXCKTRSSVEDLKWETRYKITLGIARGILYLRE          NSCGAGPSDADKVVQTHATSS</p>
697	sp_060590_znhr	0.876935	#N/A	#N/A	<p>BEST Arabidopsis thaliana protein match is: Ioridin-related (TAIR:AT5G64550.1); Has 3596 Blast hits to 2004 proteins in 150 species: Archae - 0; Bacteria - 54; Metazoa - 1597; Fungi - 38; Plants - 513; Viruses - 9; Other Eukaryotes - 1385 (source: NCBI BLINK).</p>	<p>MGTRNPGAFCSDNLSNGYKNLGDLSQIAGGGS          DHRADTLRLQPIGSIIVPHCPPTSKGMKRWGGR          DGKTGEYVDFSLCLGHRSSSSSDSKGSSATVC          TAISAKETEAESSMDLDFLHLLSSDKSPSPK          KPTVSNKFRASMQNLNLDLELSLSTSPLESEITSF          GGLDTPQHGVPMLMVGATQFVDNESASSHWEV          ENSPSPVLSGGAGGISILKPIPRKVDJASATPDIS          ASTFSPKSSVTCTSGILQHPQRGSSSKLQCFGTG          CGKGARGASGLIAHGGRRCCQRPCKNGAEG          RTALCKAHGGRRRCEFLGCTKSAEGRDIFCIAH          GGGRRCGHEDCSRAARGKSLCIRHGGGKRCQ          RENC TRSAEGLSGLCISHGGRRRCHHLDCNKGA          QGSTMFCKAHGGKRCVTVDGCTKGAEGSTFC          KHGGGKFCYSYSGCPKSVHGGTFCVAHGGGK          RCAVTGCTKSARGRTDFCVRHGGGKRCIFEGGG          KSAQGSTDFCKAHGGKRCQAWGQLGSOYGNK          VDDPQGSFAFGKTLGICAMHGALVQDKRVHGCS          TLGSLQDFCSSDLGNMIGDVSAGIMNVNDMMAM          GPTGHYAAGWPPGCDLMHVAKEKESHSPPTPEGR          VHGGSLLSLLASSSSSYQAKQWP</p>

10

20

30

40

【 1 2 6 6】

【表 1 4 - 2 6 7】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
698	sp_060640_lwad -0.887545	-0.933358	#N/A	XIPOTL 1 (XPL1); CONTAINS InterPro DOMAINs: Methyltransferase type 11 (InterPro:IPR013216); BEST Arabidopsis thaliana protein match is: S- adenosyl-L-methionine-dependent methyltransferases superfamily protein (TAIR:AT1G48600.2); Has 22524 Blast hits to 22132 proteins in 2666 species: Archaea - 571; Bacteria - 16701; Metazoa - 242; Fungi - 923; Plants - 574; Viruses - 10; Other Eukaryotes - 3503 (source: NCBI BLLink).	MAASAMGVLQEREVFKYWIEHSVDLTYEAMML DSQASDLDKVERPEVLSMLPPYEGKSVLELGAGI GRFTGELAEKASQVIALDFIESVKKNESINGHYK NVKFMCADVTSPSLNISPNVDIIFSNWLLMYLSD EEVERLVERMLKWLKPGGYIFFRESCHFQSGDH KPKSNPTHYRPFYTKIFKECHMQDDSGNSYE LSLIGKCKIGAYKSKKNQNSWLVQKVDSEDD KGFQRFLDSSQYKFNILRYERVFPGYVSTGGL ETTKEFVSKLDLKPQGVLDVGGIGGGDFYMA ENYDVEVVGIDLSINMISFALERSIGLKCAVEFEVA DCTKDYPENSDVYYSRDTLHIQDKPALFRSFH KWLKPGGKVLISDYCKSAGTSAEFAAYIFQFGY DLHDVKAYGKMLKDAGFVEVIAENRRTDQFIQVLQ KELDALEQEKDDFDDFSEEDYNDIVDVGWKAKLV RTTEGEQQWGLFIKKM
699	sp_061230_uedu -0.938431	-0.892764	-0.891335	Tetratricopeptide repeat (TPP)-like superfamily protein; CONTAINS InterPro DOMAINs: Pentatricopeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: Tetratricopeptide repeat (TPP)-like superfamily protein (TAIR:AT1G80980.1); Has 37863 Blast hits to 13191 proteins in 288 species: Archaea - 3; Bacteria - 35; Metazoa - 443; Fungi - 342; Plants - 95932; Viruses - 0; Other Eukaryotes - 1108 (source: NCBI BLLink).	MWKFRRGFSLLNFSYDYCLGVRVNSMMLGRTH VPNFVTGYKLFSPAENLKYFFSFHKIQLFCSAT SDCLNLDNSNTELVNSDESGCNDENDELSE DDGIEAGKLGIDEKLDHDLGLVLEILRGSGPSQG ENSCDNPRLLIGAEVTNKLNQCGISLTSSELVVEV LSRIRNDWELAIMFFLWAGKQPDYVHVSROYHSI ISILAKRRKFDTAWSLIDEMRKAGIVNPKTLIMIR RYAATHDVVNAINTFHGKRRNFVLYGVDDFOKLL SALTRYKNVQEAQLFSNRNTYQFNTKSFNILN GWCNIVGSPREANRFWRMLMELGVKRDVDFSSII ISCYSKLNR_LKEVLLKFGQMKEMGILPDRKAYNSI MHALAKEGRAKEALNLOIMEDKGIAPNVVTVNS LKLPLCKAKKRDEAQNVDVMSLRGLKPSIRTFH AFLGIORSGEAAFMLLDKMKELGCCQPNKDSYLIL RKFGRWQQIDIVFTLVLDQMKENGLNDVDVSHVA LHGLFLNGKLEEAHYSEMKEKMDMNPDKVTEE LIQTWLSNKQMELETSVHTSSETLVDSDSHVPG KTKSIRKQIGGGSDATEQPEPRFKVVKERGFSEW E

10

20

30

40

【 1 2 6 7】

【表 1 4 - 2 6 8】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
700	sp_061800_gchx	#N/A	-0.872619	unknown protein; Has 34 Blast hits to 33 proteins in 15 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 32; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLink).	MAKOTPTLFLLEDWLRSSSTSSNVSNSSKNSAN SSAQMIOAWVDLRECHQKGAHFHPRHLHLSKALV INSQASLYVADPQAKLLSILSSPNISLPHDSPPLIF RLLYVYVRKSSKPSLALIDLAMEVIAHFSDQFDA KKSLSLFAEGVLLVGAFCVSPSIAENTKTVCCLKV CRMLEEYOLIGSVGKFIPIVDVLAGIGYALSSSGE VYFVRLNLLGNWVKIEVPHDISHALVILHIEWV VFRLITSSLEKIKIFSCVEMENPKNYAKYAVLIM ATAGALFRASNRPVKHGERLNIVTRFRIAAEAQIEV IASDLISKIATSSNIHIEDAEYSFILRCLLAAASRSGG LISRRGPLLLSLALGLLIEPLRRFYGMIVESQTS DNGLVLPDVKNHLESILFREAGAATGVFCNLAYS ASEKDKSRVEDLWSYCHDIYSGHRLVALLQGR SDOLLNMEKIAESVFLMVAVFAVAVTKHKLTSK FSREMQMDCAVKILISFSCMEYFRHQIQLPEYMEAI RSAATSVOENEIACYSFVESLPSYADLTNPOGLC HLORGNVMSKDEVQTSRIWFYLRVIPTCISHLPT NEFOKVAAATMELVMGHPDNKVARASHLVFSSF VTFEESSQKERMLLKEQLVYVMRRSLEAYPGI TPFDGLSGSVALVRLHPAGSPSYVYCHLSLVEKA NVLSREVTNLKGSSEWKNWQGDSEPCQKIIELL MOLISMVDIOVLPEMMKRLAQLVQLPKDGGQNVS LSVCSQPTATNEACIEVSRSPNSLSLQNVTARL
701	sp_061930_gcjw	-0.862763	-0.866236	HMG (high mobility group) box protein with ARID/BRIGHT DNA-binding domain; FUNCTIONS IN: sequence-specific DNA binding transcription factor activity; INVOLVED IN: regulation of transcription; LOCATED IN: intracellular, nucleus; EXPRESSED IN: 10 plant structures; EXPRESSED DURING: 7 growth stages; CONTAINS InterPro DOMAINs: High mobility group, superfamily (InterPro:IPR009071), High mobility group, HMG1/HMG2 (InterPro:IPR000910), ARID/BRIGHT DNA-binding domain (InterPro:IPR01606); BEST Arabidopsis thaliana protein match is: HMG (high mobility group) box protein with ARID/BRIGHT DNA-binding domain (TAIR:AT1G76110.1); Has 3445 Blast hits to 3110 proteins in 301 species: Archae - 0; Bacteria - 5; Metazoa - 2269; Fungi - 308; Plants - 431; Viruses - 3; Other Eukaryotes - 429 (source: NCBI BLink).	MIDKMGSSGNKHLDAAAAAAAAAAAAAAAAAAYED VYSNRQLFMTTLENLHAMSTKFMIPVIGGRDL LHKLFVEVTSRGGIDKILAEKRWKDVDTATFNFPST ATNASFVLRKYYSLLQHYERYIYFKSQDCPPGV LOSLNATPVAVHGTAKMSPSPPTPSHPANVQINQ INSTEFSQAPSGGYPPIGVIDGKFDSSGYLVTVTIG KDKLKGVLQTPPAKMNQVAQSSGVSTVKTDTST RATSGVPRRRKRKSELKRDPDTPKPNRSGY NFFFAEQHARLKLHPGKDRISRMIGELWNKLL DNEKAVYQAKAKDKERYKTEMEDYKERLRNTQ VISDAIPQORFPDIDVEMAEVGDGKDETERDSQES PENESDSSKSETDEDEDDDEEFCDDADTPPADP NAGTSEASVAAGAGTGTGTGFEFVVKGAEGGG FKESHVNVGDEAIFST

10

20

30

40

【 1 2 6 8】

【表 1 4 - 2 6 9】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
702	sp_062540_czgd	-0.875640	-0.854482	Surfeit locus protein 2 (SURF2); CONTAINS InterPro DOMAIN/s: Surfeit locus 2 (InterPro:IPR008833); BEST Arabidopsis thaliana protein match is: Surfeit locus protein 2 (SURF2) (TAIR:AT5G14440.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archaea - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	MATSEEEKPTNNIKGNSLLGSPFFKQDNGRFK CVETGHELPAHAQDSYSQSKRCRIGLIDFSLAHK KPLLNMFKQDPLNSSLKICKLTDITNKSEEHMK HINGKRFNKLQOMEAGKLSGSDVVEEKGEKQG IKAGLKKEENGAEKKKKQKMKETKKEIVSEM RDSPLGKFRKENGAEKKEKQKMKETKKEIVS DSLEGLEFWMPFAGERWDFDDGGDRWGSDS ESGDOHDEGEKGETSI WGESDVKSKRDESDSETKEESKRSRVREFDKRE RDKDKGRNGEKERRRDREKRDRESRRSERDR SYDDKRRERERERDKGHRDKHRDRDREREKER KDQKDRDGDREDKERKDKAKEREREREER ERERERERERERERERERERERERERERERE EKERERERERERERERERERERERERERERE RRDDEKRRERERERERERERERERERERERER DSQKKGDDEPDRSQEIMERKEQDMEEEOORLL DDEMEKRRRVQEWQELKRRKEEDERKQLGDA NDEEPKAGNWTLEGESDDEEAAATEGSDADM DVRDRFPPTTEAGDPMADIVENENKLENLVPNG GADYPEEDDIDPLDAFNSMVLPEVEKLSSSVLD TENAELNGKEEIVEQNSVEVSKKAPKYLGRIMP EEDSDPYEDQEDGGTNLEDEDDDEFVKRVKKT KAEKLDLHSHKIDYDVRKFNFYIEVKDIARMALED VDSYRTELELVHGVKDVPKPIKTWHQTGLTKL EVIRKMFEPMPQAQALPVMISGRDCGIAKTG SGKTLAFVLMRLHIKQDQPLMSGDDGPIGLMPT RELVQQIHNDIKKFAKVMNLTSPVYGGGVAQ QIGDLKKGAEIVCTPGRMIDILCTSSGKITLRRV TFLVMEADRMDFMGFEQITRIVQNIQIRPDRQTV LFSATFPROVEVLAKVLNKVEIQVGGRSVNS DITQMMEMRPNENRFLLELLEGEWYKGLVLF VQSQDKDALFRDLLKSGYPCLSLHGGKQGGDR ESTLADFKSNVGNLLIATSIAARGLDVKDLVLN YDAPNHVEDYVHRVGRTRAGRPGKHAIHISEDE AKYAPDLVKALELSEQVPVQDGLKLDGFAVAKK EGLEQAHGTGFGGSGFKFDEQDEARKAAKKA QAKYEGFFDKSDSEDEDEDEVRKRSDVTVTPV LGGGQPDITGHNVSQGGTIGSVPGSILSSINSS DNPALAVAVAAAMKLCQNLAKLQQDAMPHEHYEA
703	sp_062800_jhai	-0.876191	-0.932305	P-loop containing nucleoside triphosphate hydrolases superfamily protein; FUNCTIONS IN: helicase activity, nucleic acid binding, ATP-dependent helicase activity, ATP binding; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: DNA/RNA helicase, DEAD/DEAH box type, N-terminal (InterPro:IPR011545), RNA helicase, DEAD-box type, Q motif (InterPro:IPR014014), RNA helicase, ATP-dependent, DEAD-box, conserved site (InterPro:IPR000629), DEAD-like helicase, N-terminal (InterPro:IPR014001), DNA/RNA helicase, C-terminal (InterPro:IPR01650), Helicase, superfamily 1/2, ATP-binding domain (InterPro:IPR014021); BEST Arabidopsis thaliana protein match is: P-loop containing nucleoside triphosphate hydrolases superfamily protein (TAIR:AT3G09620.1); Has 370658 Blast hits to 166919 proteins in 4473 species: Archaea - 2452; Bacteria - 63379; Metazoa - 145404; Fungi - 45975; Plants - 23206; Viruses - 1711; Other Eukaryotes - 88531 (source: NCBI BLINK).	

10

20

30

40

【 1 2 6 9】

【表 1 4 - 2 7 0】

	A	B	C	D	E	F
1		分析に用いたペアとの相関係数	sp_107670_ptqx			aa seq
2	共通遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	
704	sp_062990_kkia	-0.8338376	-0.846354	-0.614834		
705	sp_063110_xnoj	-0.8388376	-0.843424	-0.937733	uncharacterized protein	

10

20

30

40

【表 1 4 - 2 7 1】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
706	#N/A	-0.924372	-0.853591	<p>ankyrin repeat family protein; CONTAINS InterPro DOMAIN/s: Ankyrin repeat-containing domain (InterPro:IPR020683); Ankyrin repeat (InterPro:IPR002110); BEST Arabidopsis thaliana protein match is: Ankyrin repeat family protein (TAIR:AT4G03500.1); Has 38664 Blast hits to 17395 proteins in 690 species: Archae - 39; Bacteria - 2692; Metazoa - 19825; Fungi - 3120; Plants - 3302; Viruses - 203; Other Eukaryotes - 9483 (source: NCBI BLINK).</p> <p>IRREGULAR XY LEW IZ (FRKTZ); FUNCTIONS IN: laccase activity; INVOLVED IN: secondary cell wall biogenesis; LOCATED IN: endomembrane system, apoplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Multicopper oxidase, type 3 (InterPro:IPR011707), Laccase (InterPro:IPR017761), Multicopper oxidase, type 2 (InterPro:IPR011706), Cupredoxin (InterPro:IPR008972), Multicopper oxidase, copper-binding site (InterPro:IPR002355), Multicopper oxidase, type 1 (InterPro:IPR001117); BEST Arabidopsis thaliana protein match is: laccase 10 (TAIR:AT5G01190.1); Has 9472 Blast hits to 8326 proteins in 1422 species: Archae - 45; Bacteria - 3703; Metazoa - 487; Fungi - 3359; Plants - 1600; Viruses - 0; Other Eukaryotes - 298 (source: NCBI BLINK).</p>	MEDDTMSLELYAAAVNGDMSIFPNQSGSEEEATS GTSRSNQVDYFLVTLLEERTILHVAARTGHNLS FIAEALKRFPILFQTNKGETALHVSARQGNKEIT KLLVTFYRDVEGAAAGKNGSMLWRVKNSEGD TPLHTAIKRGKIQVALFUSVDNSLAISVNSRETP LHLAAKICSRIGALHVSARQGNKEITKLLVTFYRD AEAAAAGONGSMLWRVKNSEGDTPLHTAIKRG KIQVALFISVDNSLAISVNSRETPHLAAKICSR GASKRDELTPLLRASCCKSNLFFILVLIHNFPQ FAEQDLNGRNLIHQIENYEDGKRALEIPEINAL KDEQDFNGETPLHLAIKNSDFMKVRYLMECSADF TIKSSYGSPWDLIQSDNFANKMKDETDIAMNG MMKELKYASITGNVQFLSRHDIYLLSRSMEN NILHTAMSPESSTQAEIEFKTLQRIPLLCCQTNANGD TPLHFAARARSAHELTISCKNSFKVLDMGAS FLYSTPPWRLRNSGENTPLHVAALRAGNSTCARF LIEYDDEVASFVANNYNETPLHVAARYLQIPGLW GNQISERPLPOFCNRNAERISTQLSELDDLLLEANI SAAFLRDMEGTLPLRAAQCNVNIENIRKILDKCP SCAEIRDLSGCTLLHLIPLKDKHEVISHIHKLSNAR DHEGNTPLHTAINNRQFNKVEGFT
707	#N/A	-0.81623	-0.525327		<p>MSSFMDSPVAVDSTNYGLLRYPKPSLGGSPVL TAMPSLNAOQRNFHSLRSINSKYPKAKVPLSI DHSLLFTMGVGNPFCATCLNGSKLVGNINNVTFV MPTISLLOAHYINISGVFTDDDFGKPLISFDYTTV NYTGSLATMNGTKVYRPLFPNATVQVVLQDTSIA PESHPIHLHGTFNFVYVGTIGNYDQATKDSINFLV DPVERNITDPTGGWTVIRFLADNPQVWFMHCH LEVHTSWGKMAFLIENGDDGNPESIMAPPKDLPK C</p>

10

20

30

40

【表 1 4 - 2 7 2】

A	B	C	D	E	F	
1	分析に用いたバイトとの相関係数	sp_107660_k1qg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	
2	sp_107620_kqnh (SOAP1)	sp_107660_k1qg (SOAP2)	sp_107670_ptqx (SOAP2-like)	casein kinase alpha 1 (CKA1); FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: endomembrane system; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase domain (InterPro:IPR02290), Tyrosine-protein kinase, catalytic domain (InterPro:IPR020635), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Serine/threonine-protein kinase, active site (InterPro:IPR011009), BEST kinase-like domain (InterPro:IPR011009), BEST Arabidopsis thaliana protein match is: casein kinase II, alpha chain 2 (TAIR:AT3G50000.1); Has 85531 Blast hits to 84684 proteins in 2530 species: Archaea - 70; Bacteria - 9546; Metazoa - 32860; Fungi - 11294; Plants - 14699; Viruses - 299; Other Eukaryotes - 16793 (source: NCBI BLINK).	MTWKYQITYTTFSSLSLLFFRLLPILLCALLALR APVALPILRPLPSSSTLNYFSAAINPDYFISAA GKSSMSKSRVYTDVNLVLRPREYWDYFESLTVQW GDQDDYEVVRKIGRGKYSEVFEGINVNSERCVI KILKPVKKKKIKFEIKLQNLCCGGFNIIKLLDIVRDQ HSKTPSLVFEFVNSTDFKVLPTLSYDIRIYYEL LKALDFCHSQGIMHRDVKPHNVMIDHDLRKLRLI DWGLAEFYHPGKEYNVRVASRYFKGPELLVDLQ DYDYSLDMWSLGCMFAGMFERKEPEFFYGHNDH QLVKIAKVLGTDELNAYLNKYLELDPLQDLALYV RHSRKPWSRFVNDQNLHLSVPEAIDFLDKLLRY DHODRLTAKEAMAHPHYFSQVRAAESRTRTQ	
708	sp_064020_rckn	-0.932435	-0.871384	-0.930122	eukaryotic translation initiation factor 4A1 (EIF4A1); FUNCTIONS IN: protein binding, ATP-dependent helicase activity, translation initiation factor activity; INVOLVED IN: response to calcium ion, translational initiation; LOCATED IN: cytosol, nucleolus, cell wall, membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAIN/s: RNA helicase, DEAD-box type, Q motif (InterPro:IPR014014), DNA/RNA helicase, DEAD/DEAH box type, N-terminal (InterPro:IPR011545), RNA helicase, ATP-dependent, DEAD-box, conserved site (InterPro:IPR000629), DEAD-like helicase, N-terminal (InterPro:IPR014001), DNA/RNA helicase, C-terminal (InterPro:IPR001650), Helicase, superfamily 1/2, ATP-binding domain (InterPro:IPR014021); BEST Arabidopsis thaliana protein match is: eif4a-2 (TAIR:AT1G54270.1); Has 49714 Blast hits to 49117 proteins in 3142 species: Archaea - 779; Bacteria - 27442; Metazoa - 6117; Fungi - 4773; Plants - 2624; Viruses - 17; Other Eukaryotes - 7962 (source: NCBI BLINK).	MRLGLFMAGAAPEGSOFDARQYDEKMTVLLNEE GGTEFFTSYDEVHDSFDKMKGLKENLIRGIYAYGF EKPSAIQQRGIPFCFKGLDVIQQAGSGTGKTATF CSGILQDLLEYELCOALVLAFTRELAQQIEKVM RALGLDYLGVKWHACVGGTGVREDQRILAAAGVHV VVGTPGRVDFMLRQSLRANIKMIFVLEADEM LSRFGKQYDFQQQLPPKIQGVFSATMPPEALE ITRKFMMKPYRILVKRDELTLLEGIKQFYVNDKEE WKLDLTDLYETLAIQTQSVIFVYTRRKYVDW_LTDQ MRSRHTVSAATHGMDQNDTRDIIIMREFRSQSSR VLITDLDLARGIDVQQVSLVINYDLPQPENYLHRI GRSGRFRGKGVSNFVITGADDRMLFDIQKFYVNV VIEELPANVADL
709	sp_064530_chim	-0.931495	-0.933685	-0.933487		

10

20

30

40

【 1 2 7 2】

【表 1 4 - 2 7 3】

A	B	C	D	E	F
1		分析に用いたバイトとの相関係数	sp_107670_ptqx		
2	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
710 sp_064900_adyo	#N/A	0.876473	#N/A	<p>RAB GTPase homolog G3F (RAB7B); FUNCTIONS IN: GTP binding; INVOLVED IN: protein transport, small GTPase mediated signal transduction; LOCATED IN: plasma membrane, nucleus; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ras GTPase (InterPro:IPR001806), Small GTP-binding protein (InterPro:IPR020851), Ras GTPase (InterPro:IPR013753), Ras small GTPase, Rab type (InterPro:IPR003579); BEST Arabidopsis thaliana protein match is: RAB GTPase homolog G3E (TAIR:AT1G49300.2); Has 26432 Blast hits to 26399 proteins in 759 species: Archae - 38; Bacteria - 171; Metazoa - 13566; Fungi - 4011; Plants - 2887; Viruses - 20; Other Eukaryotes - 5739 (source: NCBI BLink).</p>	<p>MPSRRRTLLKVIILGDSVGVGKTSMLMNQYVVKKFS  NQYKATIGADFLTKVEQFDDRLFTLQIWDTAGQE  RFQSLGVAFYRGADCCVLVYDVNSMKSFNDLNN  WREEFLIQASPSDPDNFPFVIGNKIDVGGNSR  VVPEKKARAWCASKGNIPYFETSKEGLNVEEA  FQCIANKALKSGEEIEIYLPDTIDVGNQNGQQRST  GCEC</p>

10

20

30

40

【 1 2 7 3】

【表 1 4 - 2 7 4】

A	B	C	D	E	F
1	分析に用いたペイトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
711	sp_084910_gogj	#N/A	-0.884832	2-oxoglutarate dehydrogenase, E1 component; FUNCTIONS IN: oxoglutarate dehydrogenase (succinyl-transferring) activity, cobalt ion binding, zinc ion binding; INVOLVED IN: glycolysis, metabolic process; LOCATED IN: mitochondrion; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: 2-oxoglutarate dehydrogenase, E1 component (InterPro:IPR011603), Dehydrogenase, E1 pyrimidine-binding domain (InterPro:IPR005475); BEST Arabidopsis thaliana protein match is: 2-oxoglutarate dehydrogenase, E1 component (TAIR:AT5G65750.1); Has 10999 Blast hits to 10962 proteins in 1954 species: Archae - 31; Bacteria - 4409; Metazoa - 546; Fungi - 299; Plants - 167; Viruses - 0; Other Eukaryotes - 5547 (source: NCBI BLINK).	MTWFRAGSNVAKLAVRRTVYSGRSYVTRTRVID SQSRFVNSQSRFFHSTVLPKKAAPVPRPVLRL TDSFLDGTSSVYLEELQRAWADPDSVDESWD NFFRNFGQAATSPGISGQTQIESMRLLLLVRAY QVNGHFKAKLDP LSL EQRPVPGDLDL DNYGFT ADLDREFFLVWHRMAGFLSENRPIQTLRNLNRL EQAYCGSIGFEYMHADREKGNWLRDRIETPTPM EYNKERRHVMLDRLWSTQFESFLATKWTAKR FGLEGCELIPGMKEMFDRSADLGVENIVGMSH RGRNLVGNVWRKPLRQIFSEFSGGKIPIDEVGLY TGTGDVYHLGTSYDRPTRGGKRHLHSLVANPS HLEAVDPVWVGKTRAKQYHSNDYDRTKNMAILH GDGSAFAGQGVVYETLHLSALPNYTTGGTHIVN NQVAFITDPTAGRSSQYCTDAKALDAPIFHVNG DDLEAVVHVCLEAAEWRQTFHSDVVDIVCYRR FGHNEIDEPSFTQPKMYKIIRNHPSSLDIYKKVL ESSEMSKEEVQIQKVMITLNEEFDASKEYVPC RRDWLSAYWLGFKSPEQLSRINTGVKPEILKNV GKAITLPETFKPHKAVKRIYEQRAQMIESGEGID WAYGEALAFATLLVEGNHVRVLSGQDVERGTFSH RHSVLHDOESGEKYCPLGHVMINQNDVEMFTVSN SSSEFGVLGFEMGYSMENPSLWVWEAQFGDF SNGAQVIFDQFLSSGESKWLRLQTLVLLPHGY DGGQPEHSSARLERFLQMSDDNYPYAIPEMEPTL PKQIQEGNWQVNVVTPANFYHVLRRQIHREFR KPLIVMAPKNLLRHKECKSNLSEFDDVKGHAGFD KGGTRFKRLIKQNMHSDREEGIRLLVCSGKYV YELDEQRKKEGSDVAICRVEQLCFPFYDLIQRE LKRYPNAEAVWCQEEPMMVGAYSYIAPRLSTAM KSVNRGNADDIKYVGRPPSAATATGFSISVHGKEQ TGLVQKALQPEPIQIFP

10

20

30

40

【 1 2 7 4】

【表 1 4 - 2 7 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiag (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
712	sp_064930_ncsf	-0.301433	-0.855732	<p>CHB3; FUNCTIONS IN: DNA binding, sequence-specific DNA binding transcription factor activity, zinc ion binding; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: SANT, DNA-binding (InterPro:IPR01005), Homeodomain-like (InterPro:IPR009057), Myb, DNA-binding (InterPro:IPR014778), SWIRM (InterPro:IPR007526), SANT, eukarya (InterPro:IPR017884), Zinc finger, ZZ-type (InterPro:IPR000433); BEST Arabidopsis thaliana protein match is: SWITCH/sucrose non-fermenting 3C (TAIR:AT1G21700.1); Has 17747 Blast hits to 11817 proteins in 958 species: Archae - 45; Bacteria - 1957; Metazoa - 7032; Fungi - 2514; Plants - 988; Viruses - 139; Other Eukaryotes - 5072 (source: NCBI BLLink).</p>	<pre> MEERKDFGGSPFAATISGGSSINFSDFATTEFF APRRRNOKRKSNTTSSSAPPKRHAREKANAAA AAASGGFFSLPIIHNGPLTRARQLSDNNAIFSA ONAIKNETSMAAVNDSAASEGEEAKLAKEECE KLCDAVIDAEFEIRNRDSGVHVVAAGWFSWE NIHSIEKQTLCSFFNGKVENRTPDVYKEIRTWIMK RFHANPTTNIKDLTSLGDLARQEIEMFLDH WGLINYHPPQNEPNIDADPNTDADKAEQDLSI EKLYCFESKQSSAQLVPRANVSAAYVSGLFPE SITEELVKQGPVVEYHCNSCADCSBKRYHCQ KQADFDLCSSEYNGKFGSGMCPSPDFILMEPAE ASGGTGGKWTDOETLLLEALELFKENWNEIAE HVATKKAQCILHFLQMPIEDTFLDDDDKDDGAL ENGEPNSTNEQVPKGDSEPLEGDNKGDNVVG ATETGSGKDHSEFSASKTSKDNEEPSETKTDND AAQLVSAQTETSNSEDDLLNGKVVQGRGDDIIKA LKEAFHVAGCLTPEDELSFAEAGNSVMALAAF TQLVEPGLATASSCISLKTISQSSPGSQAARHCF ILEDPPDVDKHEVSKRMATENDPNTSEIAKEES KNDKEIIOIEEKSTPLLDEHEKESNASDDPRLPP GEKDTCADNSQIMEKPEDMDMPQVTEKPEDMDMP QVAEKPEDEMPQVVEKPEKSQLVDSKSGEKSQ PVDKPEKSKLVENPDKKSQVLEKLEKVSQVLE KLVKESQVLEKPEDELSQVTEKPEKPAKPRVTEK PDEEKQVTEKDAEKQVTEKPEKPEKQVTEK PDDEKQVTEKPEKPEKQVTEKPEKPEKQVTEK HDEVTNVEAGMKSVMKPNCSLQNEQSITAPK DQVTSMSEVMPPVSENDSEIFKITEENVOCTD VPKDSDIELSSVSDKVEQEQEAGVPMVRYVASE ADKTVIDKITDPKSTKTARTKDEDENVAKLKRALA TLSAAAANKVLLADQEEEEIKQLAXPAPPPRPLR KLEIKLSYFTEMDNAINRVREQLERSRQKLFNER </pre>

10

20

30

40

【 1 2 7 5】

【表 1 4 - 2 7 6】

A	B	C	D	E	F	
1						
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 RNA-binding (RRM/RBD/RNP motifs) family protein; FUNCTIONS IN: RNA binding, nucleotide binding, nucleic acid binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: RNA recognition motif, RNP-1 (InterPro:IPR000504), Nucleotide-binding, alpha-beta plait (InterPro:IPR012677); BEST Arabidopsis thaliana protein match is: RNA-binding (RRM/RBD/RNP motifs) family protein (TAIR:AT3G52660.2); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLLink).	aa seq	MAEGSEIDERVLDLDDNYYMEETDEDVDVGDIEDE PEEGLDDKAAEQVEELRTDGGSGKQSSSEPPDRS PIGDHEDEEKFASINEEDKEKHAELSLPPY GAEVFIGGLPKVDIEEDLDLCEFFGEIYEVKLIKN KDTGEFKYAFYAFKTKEDAKAIDELKNEKFKG KTVRCSLSDAKHRLFIGNIPKSWSEDDQKQMIHG VGFPAEMIDLKQDGLNPGNRGFAFIEYYNAC ADFARQKMSNSSRKLGNTPVTWADPKSAPDS AAASOVKALYKKNIPESATPEQLKEFQRHGEVT KVVLPAPAKGGQKNGFGFIHGERSSALKAVKD TEKYEIDGQALEVTMAKPTDKKSEAAAASFAAGS NYLPQAAYGAGNPNYGSVAGYGVAGYQQPQM IYGRGPMPSGMQMPMVLPDGRIGYVYLQQPGM QMASVPRFRNRDRSNGSGGPPQGRNGSGNDR TRFRFRPY
714	sp_065040_fsoy #N/A	sp_06527337 #N/A	sp_06535322 #N/A	注釈 HEAT-INTOLERANT 1 (HIT1); CONTAINS InterPro DOMAIN/s: Vps53-like, N-terminal (InterPro:IPR007234); BEST Arabidopsis thaliana protein match is: Membrane trafficking VPS53 family protein (TAIR:AT1G50970.1); Has 898 Blast hits to 860 proteins in 251 species: Archae - 12; Bacteria - 82; Metazoa - 379; Fungi - 169; Plants - 68; Viruses - 0; Other Eukaryotes - 188 (source: NCBI BLLink).	MSVPRMDKSSALEYINQMIFPTEASLSGVEPLM QKIHGEIRRYDASILGAVRQSQSNSTGKAKEDLAA ATCAVEELMYKIREIKNAEQSNTMVEQICRDIKK LDFAKHITTTITLHRLTMLVSAVEQLQVNMASKR QYKEAAAQLEAVNQLCSHFAYRDVPKITELREK FKNIKQLKSHVFSDFSSLGTGKETEETILLQQLS DACLVDALEPSVREELVKIFONRELTSYRQIFEG AELAKLDKERRYAWIKRRLRTNEEIKWIFPSSW HVPYLLCQFQKLTRTQLEELDLNLEKPDVGTLLL ALQRTLEFEELAEKFGSSPAKEVNHADNTE KGNFSAQVSDIKKYEKLLAANQSGGTDETESL KDFAVPGAGFNFRGIFSSCFEHLNVYVELEKTL MEHLEKLVQEEETWDAEEGSQLNLSSSMQVFLII RRSLKRCSTLTKNQLYLNFKVQKILKAYATKLF ARLPKGGTGVAATGTDGQIKTSEKDERMICYIV NTAEYCHKTSGELAENIAKIVSFAERVDMSDV QDEYSAVITKSLITLVQGLETKFDAEMAAMTRVS WATLESVDQSEYANGINTVTSIPVGLTLLSPI YFQFFLDKLASSLGPFPYQNIKCKQISETGAQQ MLLDTQAVKTLLEPVS- AKQTSAAANSYKYSR EMSKAEALLKVLSP- LSVADTYRALLPEGTTFEF QRILEKLGKADQQTLLDDFNKFGSSIKHPSMSP AVAIAPTPLAPPALSSPASSAITASREDVLAARAA LGRGAATTGFKRFLALTEAAKDRKDGPFERKLFNP	
714	sp_065100_ujja #N/A	#N/A	sp_0653971 #N/A			

10

20

30

40

【 1 2 7 6】

【表 1 4 - 2 7 7】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈  Integrase-type DNA-binding superfamily protein; CONTAINS InterPro DOMAINs: DNA-binding, integrase-type (InterPro:IPR016177), Pathogenesis-related transcriptional factor/ERF, DNA-binding (InterPro:IPR001471); BEST Arabidopsis thaliana protein match is: related to AP2 4 (TAIR:AT1G78080.1); Has 5953 Blast hits to 5757 proteins in 257 species: Archae - 0; Bacteria - 6; Metazoa - 0; Fungi - 0; Plants - 5926; Viruses - 6; Other Eukaryotes - 15 (source: NCBI BLINK).	aa seq MEENYQTSSTFTPIIDRDSVSAIILTAGANNNTI LDSIFSYPCQPSNLSLTTNNININLLEPLGSSVYL KQRDLLOQFCLDNQSNKVGYYGSSLSLNQLFQN PSTLLNPMKKKLYRGVQRHWGKWWAEIRLP QRNMRWLGTYETAETAAYAYDRAAYKLRGEYA RNFNPLKDGATKLFQGSARMNNAVKSADVAKI QAICQKLLKERANRIKKKQKEATISEGGIKGGG GEEKSNKSTTHSGDSNLNENDSLMNSYSSSSSF CSTIVNDSGMVSPDISEEGWWSNGNGASPPSSD SFTSYPMMTDDFGFEDCSLEKMPSPDFPELWVEV LAN
715	sp_066580_wrmtd	-0.9836348	-0.9836315	geminivirus rep interacting kinase 2 (GRIK2); CONTAINS InterPro DOMAINs: Calcium/calmodulin-dependent protein kinase kinase (InterPro:IPR020657), Protein kinase, ATP binding site (InterPro:IPR017441), Serine/threonine-protein kinase domain (InterPro:IPR002290), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271), Protein kinase, catalytic domain (InterPro:IPR000719), Calcium/calmodulin-dependent protein kinase-like (InterPro:IPR020636); BEST Arabidopsis thaliana protein match is: geminivirus rep interacting Kinase 1 (TAIR:AT3G45240.2); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MGCCGCGFSRRHKAIARPKFRIDAFSTVLSQTP LLDDGDDADADEDDVDDGLYNGEDTSMQTQGYD ESEVTGGQRAEEIINFKIRINGLIGRQFEVKETRK LVRSEDENGTKMINEYLREFKIGSGSYAKVLYLR SSIDGKHYAMKAFHKSHLLKQRPVAPSETAMTDVF REVLIMKVSHPNVNLIEVINDPDDTHLYMVEYV EGKWDCGEGAGPPGGLGEAIMROYLRDIVSGLMY LHSHNIVHMDIKPQNLVSRNGTYKIADFSVSOVF EGDNVLRSPGTPVITAPCECLGLTYHGKAADT WAVGILYCMITGKFFFIGETLQDTYKIVNDPIDL PDNLSQRLNLYEGLLCKGKDPDQRLTLDMAEHP WVIGEEGPLPEFLCWCCKQNSGRDLHDVDKEVL ARGLEE
716	sp_067140_fgju	-0.9327306	-0.932707	proline-rich family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; Has 103183 Blast hits to 42587 proteins in 2006 species: Archae - 171; Bacteria - 19191; Metazoa - 36983; Fungi - 10058; Plants - 17697; Viruses - 3627; Other Eukaryotes - 15456 (source: NCBI BLINK).	MEFRPPLMKLLILGLLFLFVIESQQSSQQQP QFPQQSQSQNGQNPIDRQPPFPQPQPQPPTP TPPSSFPSPSPSPPTLTPPPSTLSPSPSTIFRP PPSPPLSPSPSPSPSPSPSPSPSPSPSPSPSP PPALPLASPPKLLPVGLSPSPSPSPSPSPSPSP LPPSPSPPLPQSPSPSPSPSPSPSPSPSPSP QNSKLNKGVGLLFLAGIIVLQIGVYVSLAFKRR QLYKMRD
717	sp_067180_yuqe	-0.915319	-0.915315		

10

20

30

40

【表 1 4 - 2 7 8】

A	B	C	D	E	F
1	sp_107620_kqnh	sp_107660_klqg	sp_107670_ptqx		
2	共発現遺伝子	分析に用いたペアとの相関係数	(SOAP2-like)	注釈	aa seq
718	sp_068730_zhiq	-0.8333948	-0.951127	-0.8333948	<p>ACT-like protein tyrosine kinase family protein; FUNCTIONS IN: protein serine/threonine/tyrosine kinase activity, protein kinase activity; INVOLVED IN: protein amino acid phosphorylation, metabolic process; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Protein kinase, catalytic domain (InterPro:IPR00719), Amino acid-binding ACT (InterPro:IPR02912), Tyrosine-protein kinase, catalytic domain (InterPro:IPR020635), Serine-threonine/tyrosine-protein kinase (InterPro:IPR001245), Serine/threonine protein kinase-like, ATMRK (InterPro:IPR015783), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: ACT-like protein tyrosine kinase family protein (TAIR:AT2G17700.1); Has 132714 Blast hits to 130570 proteins in 5062 species: Archaea - 147; Bacteria - 14500; Metazoa - 50865; Fungi - 12053; Plants - 33999; Viruses - 512; Other Eukaryotes - 20638 (source: NCBI BLINK).</p>
719	sp_069220_pcpn	#N/A	-0.8333948	<p>galacturonosyltransferase 6 (GAUT6); FUNCTIONS IN: polygalacturonate 4-alpha-galacturonosyltransferase activity, transferase activity, transferring glycosyl groups; INVOLVED IN: carbohydrate biosynthetic process; LOCATED IN: endomembrane system; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINs: Glycosyl transferase, family 8 (InterPro:IPR002495); BEST Arabidopsis thaliana protein match is: los glycosyltransferase 5 (TAIR:AT2G30575.1); Has 1389 Blast hits to 1379 proteins in 238 species: Archaea - 0; Bacteria - 400; Metazoa - 143; Fungi - 0; Plants - 830; Viruses - 0; Other Eukaryotes - 16 (source: NCBI BLINK).</p>	<p>MELIEGVGESSPPRRFGGGGGVLLDIRDDV YNRLIENGIEEYVSNPDRFRELIDVHFNKLPFSYGL DINIDRAEDVLLHRKLLAKADPEKRPVFWHRCLE NFKSEDESEHLAVNHNHGERVQNNREEGREFEP CSKLEDLNEVQRNMGMMKKSQEISSHVPTHEVI CSTPDKPKLLSLSALLSDVGLNIREAHVFSSTD GYSLDVYVDGWVPEVADADSLRKAAMEAIDKSEG SWSGSSHKFSVKKALAAARGEPGNWEIDRRLLK IGDKVASGCGDLRYGVYLGQDVAVKILRSEHLT EASEDEFAQEVAILEVRHENVVRFIGACTKSPD LCIVTEYMPGGSLYNYLHKIKRALKFSELLKFAID VSKGMEYLHNDIHRDLKSNLLMDSLQVVKVA DFGVARFLHQEGVMVTAETGTYRWMAPVINHQP YDQKADIFSAVWLWELLTAKIPYENMTPLQAAALG VROGLRPEIQEIHPKLCNLMQRCWDTAPSNRP PFSIRVELEDLLENDOETSATVAGNLAGSNIL</p>

10

20

30

40

【 1 2 7 8】

【表 1 4 - 2 7 9】

A	B	C	D	E	F
1	分析に用いたベイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiag (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2				TUDOR-SN protein 1 (Tudor1); FUNCTIONS IN: RNA binding, nuclease activity, nucleic acid binding; INVOLVED IN: response to calcium ion, protein secretion, response to stress; LOCATED IN: in 6 components; EXPRESSED IN: 12 plant structures; EXPRESSED DURING: M germinated pollen stage, seedling growth; CONTAINS InterPro DOMAIN/s: Staphylococcal nuclease (SNase-like) (InterPro:IPR006021), RNA-induced silencing complex, nuclease component Tudor-SN (InterPro:IPR016685), Staphylococcal nuclease (SNase-like), OB-fold (InterPro:IPR016071), Tudor subgroup (InterPro:IPR018351), Tudor domain (InterPro:IPR002999), Maternal tudor protein (InterPro:IPR008191); BEST Arabidopsis thaliana protein match is: TUDOR-SN protein 2 (TAIR:AT5G61780.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MAAADRWLKGIVKAVPSGDTLVIMGIVKGDAPIP PEKAVILKHIMAPRLARRDDGDEPFADWSREYLR KLCIGKEIVFRADSNSTSGKQKQAPPGRFASVLLA DGS DVAKLAVQNGWAKVKTPSAQAKSEATSDLT ELLELEEKAKQOGLGRWSEVPGASEASIRNLPP SAMDPNDFDAVAFASENKGRSLQAIVEQVDRDG STVRVYLLPEFYQVYVAGIQSPSMGRMPPE RTVEAEVSTDGSNGKTASEPRAFTAAQKVASA AAPGVAVPPDFGREGKHFTETRTLHRNVRVILE QLDSFKNLVGSVYDDGGEVKNLALLELVKINGLA KYLEWSAKSLEAKDQKELKLSAELEAKKQRLMMW TNYVPPATNSKAIHQDNFTGKIVIEWVSGDCIIVAD DAVPSGSPSAERRVNISSIKAPKLGPNRWINDNK GFESHREAKFLRQKLIQRQVWVSMESYRKVP LGGATENAAAPSDFRVMDFGSVFLPAQSKGKGG DVTSPVSTAVAGEAGHNIAELLYSRGLATWVSHR DFEERSNYDALLAAEAKARANKGTHGAASP NHIQDLTASGKKAKDFLFLQSRRLSAVVEYV LSGHRFKLLIPKETSIAFSFSGVRCPRGRGEPYS EDAIAFMRRRIMQRDVEIEIETVDRRTGTLGSLWE SKTNNMAAVLLEAGLAKFKQSFQTDRIADAHLLAQ AEESAKCOKKIWENYVEQVEVTNGSVSPESRQ KEVLKVVTEVLGGGKFFVQTVEDKMWVSIQEQ ASLEIGEAPIIGSFNPKKGDLYLAQFSADNSWNR AMVNVPRGGIVQSPKDEFEFYIDYGNQVEVLPY SFLRPLPPSVHSPVGSQAQLCSLAFVKVPSLEEDY GQEAAYLSELTLNSTRREFRAMIEKDTSGGKSK GQGTGTVLLVTLVDVEASSINASMLKVLGLARLE KLIKWDSRERKSAIDNLEEFQEEAKKGRSGMW QYGDVQSDDEDSAPPMPRKAAGKR
720	#N/A	-0.3446	-0.274335	uncharacterized protein	
721	sp_0699770_xshr	-0.317173	#N/A		

10

20

30

40

【表 1 4 - 2 8 0】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	<p>注釈</p> <p>middle-3-acetic acid inducible 12 (AAV12);</p> <p>FUNCTIONS IN: transcription repressor activity, sequence-specific DNA binding, transcription factor activity; INVOLVED IN: xylem and phloem pattern formation, response to auxin stimulus, embryonic pattern specification, root development; LOCATED IN: nucleus; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINS: Aux/IAA-ARF-dimerisation (InterPro:IPR011525), AUX/IAA protein (InterPro:IPR003311); BEST Arabidopsis thaliana protein match is: auxin-induced protein 13 (TAIR:AT2G33310.3); Has 2001 Blast hits to 2000 Metazoa - 0; Fungi - 0; Plants - 2000; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLINK).</p>	aa seq
722	sp_070970_tdhg	#N/A	-0.887767	#N/A		MFLSFVKRLRIMRTSEANGLAPRSEERSAKQRG RPIKSISSSCFIYREKKG

10

20

30

40

【表 1 4 - 2 8 1】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたペイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_ktgg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
723	sp_071150_aeqj	-0.924998	-0.940681	-0.939939	hua enhancer2 (HEN2); CONTAINS InterPro DOMAIN/s: DNA/RNA helicase DEAD/DEAH box type, N-terminal (InterPro:IPR011545), DSH, C-terminal (InterPro:IPR012961), DEAD-like helicase, N-terminal (InterPro:IPR014001), DNA/RNA helicase, C-terminal (InterPro:IPR001650), RNA helicase, ATP-dependent, SK12/DOB1 (InterPro:IPR016438), Helicase, superfamily 1/2, ATP-binding domain (InterPro:IPR014021); BEST Arabidopsis thaliana protein match is: RNA helicase, ATP-dependent, SK12/DOB1 protein (TAIR:AT1G59760.1); Has 9841 Blast hits to 8173 proteins in 1240 species: Archaea - 753; Bacteria - 3176; Metazoa - 1197; Fungi - 1159; Plants - 489; Viruses - 32; Other Eukaryotes - 3035 (source: NCBI/BLINK).	M A E L P S T T T E L P S T T T E A P S T M V E N P P S S N A P V A E Q O T L G K R K Q S E L E E P V P Q O S S L E E Q K Q Q N Q K E S Q E S S E K P G S K R R S S T R T C V H E V A V P K D F I S D K D S S I H G T L A N P S F N G S E A K K Y A F T L D P F O Q V S I A C L E R N E S V L V S A H T S A G K T A V A E Y A I A M S F R D K Q R V I Y T S P L K A L S N Q K Y R E L S O E F T D V G L M T G D V T I S P N A S C L V M T T E I L R G M L Y R G S E V L K E V A W V I F D E I H Y M K D R E R G V V W E E S I V F L P T A I K M V F L S A T M S N A T E F A E W I C N I H K Q P C H V V Y T D F R P T L O H Y V F P V G G S G L Y V D E N E Q F R E D N F L K L O D T R Y K P G H M K A A A D K A S G R I A K A G S G P K G G S D I Y K I V K M I M E R K F O P V I V F S F R R E C E Q N A M S M S K L D F N T D E E K D I V E Q V F K N A I L C L N E E D R G L P A I E L M L P L L Q R G I A V H H S G L L P I I K E L V E L L F Q E G L V K A L F A T E T F A M G L N M P A K T V V F T A V K K F D G S H R Y G S G E Y I Q M S G R A G R R G K D D R G I C I M V D E Q M E M N S L R D M V L G K P A P L V S T F R L S Y S I L N L M C R A E G Q F T A E H V I R N S F H O F Q H E K A L P E I E K R I E K L E O E A A T L D A S G E S E V A E Y H K L R L D L V L L E K K M M S E I T R P E R I L Y Y L L P G R L I K I R E G G T D W G W G V V N V W V K A P A A S G A V O R G G G Y I V D T L L H C S P S T S D S G S R R P K P C P P R P G E K G E M H V V P Q L P L I A A L S K L R L S I P P D L R P V E A R Q S I L L A V O E L E K R F P Q G L P K L N P V K D M G I E D K E I V E L V S Q I E E L E Q K L F S H P M H K S Q D F H Q V K C F Q Q K A E V N Y E I Q L K T M R E S Q L O K F R D E L K N R S R V L K L G H I D G D G V W Q L K G R A A S L I D T G D E L L V T E L M F N G T F N D L D H H Q V A S L A S C F V P C D K S N E Q I Q L R A E L A K F L Q O L Q D S A R R I A E I Q R E C R L D V N V E E Y V E S T A R P Y L M D V I C W S K G A N F S E V I Q M T D I F E G S H I R L A R L D E F I N Q L K A A A S A V G E A N L E N K F G E A S T S L R R G I M F A N S L Y L
724	sp_071240_odew	-0.883795	-0.949557	#N/A	uncharacterized protein	M L N I K H Q I S S S D H Y I S R M L E N G E A N T P T I S S T H L D S Y G I P L L M

10

20

30

40

【表 1 4 - 2 8 2】

A	B	C	D	E	F
1					
2	分析に用いたバイトとの相関係数 sp_107620_kqnh sp_107660_kiqg (SOAP1) (SOAP2)		sp_107670_ptqx (SOAP2-like)	注釈	aa seq
725	sp_071300_xeww -0.839738	-0.97236	-0.826333	MATERNAL EFFECT EMBRYO ARREST 58 (MEE58); FUNCTIONS IN: adenosylhomocysteinase activity, copper ion binding; INVOLVED IN: methylation-dependent chromatin silencing, one-carbon metabolic process, posttranscriptional gene silencing, embryo development ending in seed dormancy; LOCATED IN: plasma membrane, vacuole, membrane; EXPRESSED IN: 36 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAINs: S-adenosyl-L-homocysteine hydrolase, conserved site (InterPro:IPR020082), S-adenosyl-L-homocysteine hydrolase (InterPro:IPR000043), S-adenosyl-L-homocysteine hydrolase, NAD binding (InterPro:IPR015878); BEST Arabidopsis thaliana protein match is: S-adenosyl-L-homocysteine (SAH) hydrolase 2 (TAIR:AT3G23810.1); Has 7964 Blast hits to 7958 proteins in 1477 species: Archae - 229; Bacteria - 2334; Metazoa - 639; Fungi - 145; Plants - 247; Viruses - 0; Other Eukaryotes - 4370 (source: NCBI BLINK).	MATLSVEKTTSGREYKVKDMSQADFGRLIEILA EVEMPGLMSCRTEFGPSQPFKGAKITGSLHMTIQ TAVLIETLALGAEVRCSCNFSTQDHAATAIAR DSAAVFAWKGETLQEYWWCTERALDWAAGG PDLIVDDGGDALLIHEGVKAELEFEKNGTIPDPA STDNVEFQLVGLIKESLKVDPKRYHKMKERLVG VSEETTTGVKRLYQMLNGTLFPAINVANDSVTK SKFDNLYGCRHSLPDGLMRATDVMIAGVGVVC GYDVGKGCALAKAAGARVIVTEIDPIALQAL MEGFQILPLSDVSEADIFVTTTGNKDIIMVDDMR KMKNNAIVCNIGHFDNEIDMAGIENYFGIKRITIKP QTRDFVPEETKTGVILLAEGRMLNLGCATGHPSP VMSCSFTNQVIAQLLELWNEKDSGKYKVEYVLP KHLDEKVAALHLNKLGAKLTKLTKDQADYISVPVE GPYKPAHYRY
726	sp_071360_rloq -0.812739	-0.971848	-0.8313354	CONTAINS InterPro DOMAINs: Nin one binding (NCBI) Zn-ribon like (InterPro:IPR014881), D-site 20S pre-rRNA nuclease (InterPro:IPR017117); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MDSQPPLSTPEAPTVEPPVVCWSNVKQTPSP TPPQPKTPTKLFETGKSAKGIYAIVADANAIOG GERLPSLADKFVSVSEVINEIRDPTSRHLSFLPF SVDTLEPSPESLKKVVSFARATGDLQTLSDVDLK LIALTYLEAQVHGTSHLRDSDPPPIHV_NVRRLE KDLPGWGSNVPNLEWEALDAATGGGPDHNSR ILPLKDLTLNIPVEDQISLQDGSVNGTVNDGAG LTTTRRHRYPKPKREVKLDGKMWADGVDAQOG QHDDNDDGDWRPVAVSRGTHRRFLRRKARRDM YEASLDDDEQQEVNKNTEEEVIMEAKNGLTS DRHVQDEISEDKDDTDTGNSAENISSILEQMLRA EDALKHDEEQDVQIGSHLEVASMDADLEDAHS GEGDNVEDDLLENGSVDTSYAEDDDGSEQSWM LFLSLSQSSVACITGDFAMQNVILQMGLR_LAPGG MQIRELHRWLVKCHACQNVTAEIGRIFCPKCGN GGTLRKAVTVGENGVLQARKQRFSLRGTKFS LPLPQGGRAMSKNPIREDQLPQKFLYPKTKKK SNKDDIFFGDSIFGHHTEKTKTYNRPFRQALAV FSGKRNPNDNHYSFAKH

10

20

30

40

【 1 2 8 2】

【表 1 4 - 2 8 3】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
727	sp_071780_hjrg -0.886388	-0.843572	#N/A	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 12 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G48335.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MLENPSDLSSNPLSAPIKRYAPPNQRSGLSRRKSGDRSERTSNLHLVDGDKNLGTGLGIDHSDTGS SKSVYSEFRPRLTALNGCCRSEASQLLHERWNG AVHCYSDPTVDLSERPVMYTGSAASAWGNFRLP HQLMSAPQMDFLAELRQAINNSNASSMMN
728	sp_071840_nwda #N/A	#N/A	-0.866281	SPT4 homolog 2 (SPT42); FUNCTIONS IN: positive transcription elongation factor activity, zinc ion binding; INVOLVED IN: positive regulation of transcription, N-terminal protein myristoylation; LOCATED IN: nucleus; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Transcription initiation Spt4 (InterPro:IPR009287), Transcription initiation Spt4-like (InterPro:IPR016046); BEST Arabidopsis thaliana protein match is: Transcription initiation Spt4-like protein (TAIR:AT5G08565.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MGSVVAQIPTSFGHELRACLRCLVKTYDQFRES GCENCPFFGLDKDHERVVECTTPNFTGISVMDP SRSWAARWLRIGRFVPGCYTLAVSEALPEEFQAI CEDERVQYVPPKQK
729	sp_072230_dtom -0.952856	-0.846463	-0.866281	plant U-box 9 (PUB9); FUNCTIONS IN: ubiquitin-protein ligase activity; INVOLVED IN: protein ubiquitination; LOCATED IN: nucleus; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: U box domain (InterPro:IPR003613), Armadillo-like helical domain (InterPro:IPR011989), Armadillo-type fold (InterPro:IPR016024); BEST Arabidopsis thaliana protein match is: ARM repeat superfamily protein (TAIR:AT5G18320.1); Has 3234 Blast hits to 3002 proteins in 224 species: Archae - 0; Bacteria - 16; Metazoa - 312; Fungi - 206; Plants - 2395; Viruses - 3; Other Eukaryotes - 302 (source: NCBI BLINK).	MAKTGVSEGDSIRVLELKKELRLLVKTILEEEDH NYNNNDDYSTNSCCSSVIDAIDKAKDLLSNL KEMKMKTTSLRLDDGFSFSCPPHFCCPLSKQL MRDPVIITGQTYDKSFIQKWLKAGHRVCPITQQ VLSHSL TPNVLYREMISQWCESRGIENPDFVGY INEDGANEADRDYFSSLIDKMSKSGEQAQAKEL RLLTKRMPFRALFGESTEAIPLRLKPLTQQRRG SFSSGSNELTHPDLQEDLITLLNISHDSNKKLVA ETPMVLLKLVDAQFGTPETRSNSAAALFTLSAL DSNKVLIGKSGALKPLIELDEGHPMAVKDAASAL FNLCIENKGRAVFDGAVRIFKIKIKINGVHIDELL AILAMLSTHQRAVEELSDLGAVPCLLANLRESSC AFRNKENCIALYITICVNDRTKLRMRDEERAYGTL SQLAQDGTSGRAKRKASGILDRLNRAHNLHTTA

10

20

30

40

【 1 2 8 3】

【表 1 4 - 2 8 4】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
730	sp_072290_ehhq	-0.530707	-0.834724	-0.420317	FUNCTIONS IN: molecular function unknown; INVOLVED IN: ribosome biogenesis; LOCATED IN: nucleus; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: AAP2CN (InterPro:IPR012948); Protein of unknown function DUF663 (InterPro:IPR007034); BEST Arabidopsis thaliana protein match is: P-loop containing nucleoside triphosphate hydrolases superfamily protein (TAIR:AT1G06720.1); Has 2741 Blast hits to 2088 proteins in 291 species: Archae - 2; Bacteria - 131; Metazoa - 833; Fungi - 650; Plants - 171; Viruses - 49; Other Eukaryotes - 905 (source: NCBI BLINK).	MGGFRAQVNAKHAHKTFRFASKSSRNIIHKTSLKDKH QNAKSDRFKGGARAAARIQRNINMIRDQKRTAVLN EKFRSSGSASPLVIVLFGLSANVDLNAVERDLL GLLSECGSEASPTVVSSEYKIRATVLIKAPHGDL SPCMDMAKVADLMAFVTSASSPCEDGSSDYIDS FGLQCLSVFKAIGLFPSTVVLIRDLPSDLKGRNESK KFCNSSLASEFPDPCKFYPADTKDDLHKFLYLLK DQRLKSPHWRNQRAYVASKLMLDHDHDSVSGK CTLQVITGIFGRNLSVNLVHVSAGDYQLSKIDI LRDPYALNPRKEQDTMDADDAPDLQVSSLVPPDD SKQEP LHYENVDPD _AGEQTWPSKEMEDAEEK NLKHKEVKKRRLLPPGTSQYAAWIVDESDDDDD VGSVDESNDGMALDDEEDGFASKSEISFDLDE DQGLDLGESDGETETDSEMGPTREQIEHDIK RLKDAHAEDEEYVDEVDTPLDVPAKRFARYRG LKSFRSSWDPKESLPVEYSRIFAFDNFTTRKH VLAKALEWDNKSKEETVSPGTFARLYIKDVPQAV ASTLYSRAKMIPIIVCGLLQHEKMSVLFHSIKKH DSYHEPIKGEELIFHVGFRQFVARPFISSDNFNS DKHKMEKFLHAGRFVSASVAPIFPPPLCLVAMKK LGDMSGPV_LAAVGSLSVDPDRILLKIKILTGYPQR VSKVKATVRYMFHSPEDVRFVWFKPVVEWTKGGRR GRVKEPVGTHGAMKCLFNGVLQQSDTVCMSLF KRAYPKWPEHKFPLSNF

10

20

30

40

【表 1 4 - 2 8 5】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
731	sp_072340_uoit	#N/A	-0.69111	-0.977553	<p>trehalose-phosphatase/synthase 9 (TPS9);            FUNCTIONS IN: transferase activity, transferring glycosyl groups, trehalose-phosphatase activity;            INVOLVED IN: trehalose biosynthetic process, metabolic process; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: HAD-superfamily hydrolase, subfamily IIB (InterPro:IPR006379), Glycosyl transferase, family 20 (InterPro:IPR001830), Trehalose-phosphatase (InterPro:IPR003337); BEST Arabidopsis thaliana protein match is: trehalose-6-phosphatase synthase S8 (TAIR:AT1G70290.1); Has 4900 Blast hits to 4828 proteins in 1062 species: Archae - 56; Bacteria - 2863; Metazoa - 236; Fungi - 655; Plants - 621; Viruses - 0; Other Eukaryotes - 469 (source: NCBI BLINK).</p>	<p>MGSKSYGNLLGLVSGDLLDIPCTPRTLPRVMTVP            GISDLGRESSEDVDAIVRERTIIVANNMLPLH            ASRDTETGKWRFTVDEDSILLKNGLTPNTEVI            YIGSLKADIDINEQDEVSQKLLREFKCVPTFLPQD            LOKKPYHYCKQYLWPLFHYMLPMCPCDHGDRF            DRALWQAYCANKIFADKAVMEVINPEDDYVWVH            DYHLMVLPFLKRRFRVYKLGFFLHSPFFSSEIYR            TLPVRDELRLGLLNCDLVGFHTFDYARHFLSCCS            RMLGLDYASKRHHGLDYFGRVFIKILPVGIHVG            RIESALNLPCTSVRVKEIQEQKGGKVIIVGIDMDI            FKGISLKLAIQELLQNPEFQGLVLIQIINPARSS            GKDVQEAQKRETYLARRINEMYGSPDYEPYVIDR            PVPFRFEKSAYYAAECCLVNAVDRGMNLVYKYI            VCRQGTPTLDEALGVKSDTPTQTSMLVSEFVGC            SPSLSGAIRVNPWDIDAVAEALNLATMPDAEKH            LRFHEKHRYVSSHDAVYWARFVQDLERACON            AYRRTSKRVIFLDYDGTIVPQTSMNKCPSPSEVFS            HYNKRCWGFGLGLSFRVLSLSPFRKLNVDYIAS            VLNALCNDPKNTVYVSGRGKTKLDEWLSPCQL            LGIAAEHGYFLRPNKATEWESNTAELDWKEIPEP            IMRSYTEATDGSNIETKESALVWHHQDADPDFG            SWQAMELMDHLENVLANEPTVWARGHNIVEVKP            QGVSKGLVVEKVVITTLINGGKRDPFLVCIGDDRS            DEDMFESLLKFFSSPTSSSGLPEVFCCTVGGKPSK            AKYYLDDTSDVLRLLLAGLAAAASTQKPRYNPHAQ            VSFESTF</p>

10

20

30

40

【表 1 4 - 2 8 6】

1	A	B	C	D	E	F
		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
732	sp_072350_yyue	#N/A	#N/A	-0.671384	<p>K+ uptake permease 6 (KUP6); FUNCTIONS IN: potassium ion transmembrane transporter activity; INVOLVED IN: potassium ion transport; LOCATED IN: plasma membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Potassium uptake protein, kup (InterPro:IPR018519); K+ potassium transporter (InterPro:IPR003855); BEST Arabidopsis thaliana protein match is: Potassium transporter family protein (TAIR:AT5G14880.1); Has 3423 Blast hits to 3365 species in 1017 species: Archae - 13; Bacteria - 2376; Metazoa - 1; Fungi - 98; Plants - 811; Viruses - 4; Other Eukaryotes - 120 (source: NCBI BLINK).</p>	<p>MDPESGLSQNQVKKKESWRQVLLLAYQSLGVV YGDLSSTPLVYVYKSAFAEDIQHSSENEEIFGLVLSF VFWTLTIVPLLYVYVIRADDNNEGEGTFALYSLL CRHARVNSLQNCQLADEDLSEYKDKGVVWPAQT SFGSRILKSILEKHRVQLKLLILALIGACMVIQDGV LTPALSVFSAVSGLELSMSHEHKKYVELPVTGVIL IGLFALQHYGTHRVGFLFAPVWVWLFQISTIGLYN IIHWNPVHYKALSPVYMYKFLKKTQTGGWMSLG GILLITGSEAMFADLGHFSQSSIKIAFSVYVPSL ILAYMGQAAYLSRHHVITTTTYQIGFYVSVPEKLRW PVLGAILAAVYVSGSQAITTGTFSIKQCCQALGCFPR VKVHTSSKIHGQIYIPEINWILMILCLAVTIGFRDT KRMGNASGLAVTVMVLTTC.LMSLVVICWRKSV FVALAFVLFQGAFAEALYFASALIKFLEGAWVPIALS LFFMLIMYVWHYGTLLKYEFDVQNKVSNWLLSL GPSLGIVRVVRLGLIHTELVSQIPAFSHFVTLNPA FHQVLFVLCVKSVPVPHVRPEERFLVGHIGPKY RLYRCIVRYGRDFHKDDDFENDLVCSSIAEFIRG ESNMNGLKDESDKDDTEKMTVVGSPSTHVDGI RMHEDVEVEDIPETSEMKEIRSPLETPKKRVRFLL PETPKMEDKTMEELODLMEAREAGIAYLGHSHV RAKQGGWIKKLAINGIDYDFLRRNCRAPSPVLSV PHASTLELVGMICHM</p>
733	sp_072400_ijms	#N/A	-0.886636	-0.652403	<p>lipase class 3 family protein; FUNCTIONS IN: triglyceride lipase activity; INVOLVED IN: lipid metabolic process; LOCATED IN: cellular component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Lipase, class 3 (InterPro:IPR002921); Has 393 Blast hits to 389 proteins in 69 species: Archae - 0; Bacteria - 12; Metazoa - 12; Fungi - 52; Plants - 275; Viruses - 0; Other Eukaryotes - 42 (source: NCBI BLINK).</p>	<p>MDRYLLAEAGDTLFAVSVGTVKQYKDLVANANIFQ GALFDDDDMEIMEETEIVESDIADGRKGYLYSHA KTETPKQKQKHLRPAARHGFMARAKGIPALELY RLAQKKKGLVLCGHSLGGAVKYVC</p>

10

20

30

40

【 1 2 8 6】

【表 1 4 - 2 8 7】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_ktgg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq WVSSHLFDTRRRKPEVLLSARLRYVYAGST DSLSCLETDEFDEKTDVTSSTSSPACLSLSSCDEG QIPSSSLERSCQVNGHADAVSLASGAAGISYPR KNDNEHGSDFVNGWYVLEQGMCGPYIQQ QLFEGLSTGLPDELVPYILNGALLNPVLKFFK QYPDQVTTGFAYLRSITPNVSIITSSHIAPGGDT HKDTGVLNYDTQSHVQQAASHSTDAIVSDSKLS PRTAVVQQSLSWEKCKWFLDGYGRVGTGPHS LSELYYHQYGYLHSSVMIRHTDNKCGPFSLLSL INACGKESSTANTSEKESISSSELISGVSEIIGS QLHDGIMKAAHRYVLEIGNVIADYVYAMKKAQK QLKNEHGREDLIFERNKTKDGENNSSASHAITS CHDDASTATAKETLEEPSFQHMKSIGSVENY QGATEAISRVLDYGMWMAVYEPVANCTL TWRKNRWVSHNSTEEAVALDKNSDFQKVVVEV EASEPGVPADELDCPPGFHSIVQDSLCCDSSLIA QSHFREDKLSFPMQHQDEDEGTIGRVEEELF MLAKTSMDDFLKISIEKEVFKLNKLDKNGTHEVS VENSGQLSKDNDSGYAASEMLOGLVSRRICLDE SHIISQGTGLLHQPAVCPHKDSSDFLGTGFRSL GAPAITAHCCKTDEPLPPGIADLTIDPSAQDLQKF LLDLHLPIQAIRSHEAADLAHGNSSNFIGHAFKR LGMPICSVIDDHRHTDEAQLPAGGGEVKDIVSPTIS EFPPLHLGEGIPKQKFMGFALCRQKLDHDDVLRRE WLSHFDELHCLLQKRDLVQEELIRTEECSKLTH PMQPELDHLVERSKSCYSGSPSEYVWEFTYS RRKKVLKKSQAEPSSPPLRDFKLHQSSKFKFG PISNNVSCGALGIKNVKQKAKGVCQRETGLSS LTSKLGAVKRSIGKDRQSRSSNGRRLOSFAFPS QSTVAVDDPKHGVEKVLDMEAESNDIEELSSSS IHNSNAKDKLANDHSQTVKNSSTVSKLKRKNLVG ESSLSFKAPKLVLDAGAKQAEVKVAERRTKLKKS
734	sp_073500_mwyk	-0.825361	-0.3448353	-0.8950031	SET domain protein 25 (SDG25); CONTAINS InterPro DOMAINs: SET domain (InterPro:IPR001214), GYF (InterPro:IPR003169); BEST Arabidopsis thaliana protein match is: homologue of trithorax (TAIR:AT2G31650.1); Has 5838 Blast hits to 5683 proteins in 501 species: Archae - 3; Bacteria - 461; Metazoa - 2434; Fungi - 507; Plants - 1016; Viruses - 2; Other Eukaryotes - 1415 (source: NCBI B_link).	

10

20

30

40

【表 1 4 - 2 8 8】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
735	sp_073920_ncj	#N/A	0.886616	0.862409	<p>Chaperone DnaJ-domain superfamily protein;            FUNCTIONS IN: heat shock protein binding;            INVOLVED IN: biological_process unknown;            LOCATED IN: cytosol, plasma membrane;            EXPRESSED IN: male gametophyte, guard cell,            cultured cell, pollen tube; EXPRESSED DURING: M            germinated pollen stage; CONTAINS InterPro            DOMAIN/s: Heat shock protein DnaJ, N-terminal            (InterPro:IPR01623); BEST Arabidopsis thaliana            protein match is: Chaperone DnaJ-domain            superfamily protein (TAIR:AT4G12780.1); Has 79658            Blast hits to 47178 proteins in 2717 species: Archaea -            253; Bacteria - 22646; Metazoa - 24220; Fungi -            7704; Plants - 5305; Viruses - 398; Other Eukaryotes -            19032 (source: NCBI BLINK).</p>	<p>MNDFGLLSTDYGFKPQGKAAPMAASKGTNS            SNSAFNFEIGGSKTASFPSSNSKNPTSSSLFDDHD            TLFRSNPKTTSSDDFGILDDVFATGGIKNSAKSS            NNNSVMDSFIFGGGGGGGGGSGPKFSSLPVFD            KPVYDDESSSAKFENIFAPPSSAPAKTKDAFDD            LLGGLGKKEQSTRSVAVGPGSGADKGFSSFD            LLPGGGGSTTSRNPSSDASKKTSVVIDDPFVG            LDSTSPVSSGIFTDPLEEIGNLNNSRRTAES            SSGTAFDDIDPLGGGKSVPSFAEINGEEKEAM            RTRVSGRQTSVSGDASRKSSVRSRDRGSEPKFP            VEDYSEPSVFEMPSPVSAETSRVSNASFPTHSN            ASFRESNSHGHSFPNSENEMGSSDDIWLTVSEV            PLFTHPTAAPPPSRPPRPTHFSASYPASNVHK            KVNDSYFPFGPGPSFHSPKSAPVEARGSFASPL            DELEEFAMGRNPSNEDATHASSEGGSGEDFDS            NSVAAAASAAAMKEAMDRAEIKFRHARGVREREN            AKAKSRESVQVEKEENPTYDAPEREFREHOER            LERERQOREMEEEERRRRTKEEERRRAEKWEE            KEREKRKVEKEROFGROAVEWAAKEARAAIE            ARARADRVAVQRAQAEAREAREAKAAVHRAQSE            APERAAAGAKERAEKAAAEAREKAKAEAREKEA            REREAREKEAREKAAATAKANQSKNENDLESFF            SMGRASSAPPRAPSSDPFEAQFTKPAAEATR            ASVNSSSMRKASSSTNIVDDLSSIFGGPSSGG            FQEVDTDDRRRARDRDRHQRTOERRAAKALAE            KNARDLEVQREQAEKNRISSETLDVEIRRWAAAG            EGNRALLSTMQYVLWPECGWQVPSLTLITGA            AVKKAYRKATLCIHPDKVQQQKATLQQKFAEKV            FDLLKEAWNKFNSEELF</p>

10

20

30

40

【表 1 4 - 2 8 9】

A	B	C	D	E	F
1					
2	sp_107620_kqnh sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_klqg sp_107660_klqg (SOAP2)	sp_107670_ptqx sp_107670_ptqx (SOAP2-like)	注釈	aa seq
736	sp_075180_rzyq -0.485953	#/N/A	#/N/A	DEAD box RNA helicase family protein; FUNCTIONS IN: helicase activity, ATP-dependent helicase activity, ATP binding, nucleic acid binding; LOCATED IN: cellular component unknown; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: DNA/RNA helicase, DEAD/DEAH box type, N-terminal (InterPro:IPR011545), RNA helicase, DEAD-box type, Q motif (InterPro:IPR014014), RNA helicase, ATP-dependent, DEAD-box, conserved site (InterPro:IPR000629), WW/Rsp5/WWP (InterPro:IPR001202), DEAD-like helicase, N-terminal (InterPro:IPR014001), DNA/RNA helicase, C-terminal (InterPro:IPR001650), Helicase, superfamily 1/2, ATP-binding domain (InterPro:IPR014021); BEST Arabidopsis thaliana protein match is: DEAD box RNA helicase 1 (TAIR:AT3G01540.4).	MAATATVSPGRYAPEDPSLPKWPWRLGLVDGKT GYLYFWNPETNVTQYERPGSSSHGCSIPSRPFS GYVSSAQYQQSSNGHKDDNDKYNRSGSGER IQOSARGNTPNEAVSERSTIEHVPVDAVSGLSA EAYFRRRHIEISVSGTGVPVPPFIFEATGFPSELLKE VHQAGFSAPTPIQAQSWPIAMQGDIVAIKGTGS GKTLGYLLPGFMHLKQRRNNPQMGPVTVLSPT RELATQIODEAVKFGRRSRIACTCLYGGAPKGPQ LKELDRGVDIVAVTPGRINDILETRRSLNQVSYL VLDEADRMIDMGFEPQIRKIVKELPTRRQTLMYT ATWPKVEKIAADLLSNAVQVNIQNADELVANKSI TCHVEVLSMEKQRLEQLRAEPEGSKIIIFCST KKMCDLLARNLSRQFGAAAHGDKSQSERDHLV NQFRSRSRPLVATDVAARGLDVKDIRVVVYDF PTGIEDVYHRIGRTGAGATGVAYSFFNDQDSKY ASDLVILEGANQKVPQEVDRMMLRGGGGMG RFRFGSGGGRDGGRRGSDSSYGGRDGGR GRDRYDSYDPGYQARGSRSPSPTGQKAWGG NQVKDQGSWRGSRSRSPKINKGQPGKSFHE MMMEKTRSSPMFPQKTPAADYNGNGQSAWGWKS NDEEEEGAIIPS
737	sp_075680_bzyf #/N/A	-0.395053	-0.853416	ser/arg-rich protein kinase 4 (SRPK4); FUNCTIONS IN: protein kinase activity, kinase activity, INVOLVED IN: response to oxidative stress, protein amino acid phosphorylation; LOCATED IN: cellular component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Serine/threonine-protein kinase domain (InterPro:IPR002290), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Serine/threonine-protein kinase, active site (InterPro:IPR008271), Protein kinase-like domain (InterPro:IPR01009), Protein kinase, catalytic domain (InterPro:IPR000719), Tyrosine-protein kinase, catalytic domain (InterPro:IPR020635); BEST Arabidopsis thaliana protein match is: Protein kinase superfamily protein (TAIR:AT3G44850.1); Has 38460 Blast hits to 29477 proteins in 1034 species: Archae - 2; Bacteria - 1392; Metazoa - 15648; Fungi - 7270; Plants - 6510; Viruses - 16; Other Eukaryotes - 7622 (source: NCBI/BLINK).	MNRRKETVSVTADTEELREKGEPREKEAAVEE EEEEEVSESGDYTSEDEKDDYRFRGGYHAVRI GDSFKNGAYVQSKLWGFHFSTVWLAWDTNHS RYVALIKVQSAQHYTEAAMDEITILKQIADGDPD DKKCVKLLDHFHKGSPNGQHVCMVFEYLGDNL LTLIKSYDRGMPMDKVKIICYHLVGLDYLHKQL SIHTDLKPENILLSVIDSSKDPKRTGAPLVLPS KDKTLYESSTKESKVSNGDLTKNQKKIRKAK RAAQCGAGKEAASDQDPEPESDVGVKRTSG GSSVHKRGSSTRKRLLESVDLKGKLVDFGNGAC WTKQFTNDIQTRQYRCPEVILGSKYSTYADLW SFACICFELATGDVLFDPHSGDNYDRDEDHLALM MELMGMMPRKIALGGYRSRDNFFNRYGDLRHRR LRFWPLNKYVEKVEFNEKANDLAELFIPILDV PDKRPSAGECLSHPWAGGLTGLKASPESQEV DVTITDKKKEKDEREAMEKMGMSIAITIDSKPSK QFLSSPKPDKAAVIGSSR

10

20

30

40

【 1 2 8 9】

【表 1 4 - 2 9 0】

A	B	C	D	E	F
1	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	共発現遺伝子			<p>VACUOLAR MORPHOLOGY 3 (VAM3); FUNCTIONS IN: SNAP receptor activity; INVOLVED IN: in 7 processes; LOCATED IN: late endosome, trans-Golgi network transport vesicle, plant-type vacuole membrane, vacuole, plant-type vacuole; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Target SNARE collect-coil domain (InterPro:IPR000727), t-SNARE (InterPro:IPR010989), Syntaxin/epimorphin, conserved site (InterPro:IPR006012), Syntaxin, N-terminal (InterPro:IPR006011); BEST Arabidopsis thaliana protein match is: syntaxin of plants 23 (TAIR:AT4G17730.2); Has 30201 Blast hits to 17322 proteins in 780 species; Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).</p>	<p>MSFDLEAGRRPLASRRDLINAKQDPTQAVASGI FCINTAVSSFORLVNTLGTPTKDTLQLREKLNKTR LHIGQLVKDTSKDKLQKQASSETDHSNNVNSANKKIAD AKLAKDFOAVLKFQKQARLAAERETAYTPPIQ AVLPSSYTAASEIDVSSDKTPEQRALLVEARRQDVI LLDNEITFNNAVIEEREQIGQIEHQIGVEVNEIFKD LAVLVHDGGMIDDIGTIVENAAQAATVQAKSHLA QASKTQRSNSLTCLLLVIFAIVLVLIIIAA</p>
738	sp_075870_binw #N/A	-0.871473	#N/A	<p>LRR and NB-ARC domains-containing disease resistance protein; FUNCTIONS IN: ATP binding; INVOLVED IN: apoptosis, defense response; CONTAINS InterPro DOMAIN/s: NB-ARC (InterPro:IPR002182), Leucine-rich repeat (InterPro:IPR001611), Disease resistance protein (InterPro:IPR000767); BEST Arabidopsis thaliana protein match is: NB-ARC domain-containing disease resistance protein (TAIR:AT3G14470.1); Has 38623 Blast hits to 22788 proteins in 858 species; Archae - 54; Bacteria - 2838; Metazoa - 8237; Fungi - 656; Plants - 25379; Viruses - 2; Other Eukaryotes - 1455 (source: NCBI BLINK).</p>	<p>MDIVGTLSSAQTLLAALQCSSEQLSIFGYKSQ LDDLQRTVSAINAVLLDAEAKQELSHKAEKHWLEE LKGAVFEADDLDEFVTLAEQQLMEAGSLSKK WPKNTTNNVNDIETEGLYLRNKEHLQGLKDFL HRKVDGVDVTKEARRLMEELQPHSKLKLTVRR YSGVRMPSWTTLLPNLFLALQNGEELGYLPLCL GNLRLHLKLEGLKLEYIGISPSVLGSAEGYFF FASLEYLESLPKLKGWAGVAGAGDDLLFLDDC ENIQQLRLCSLQKLVLLIEVPELTCMPPCFGLK YLQLRKFNRRRIINTSSMFCHEIDICPSPSRSLFD PENRSNRQWTVDDVAVLNSLPIEAFQGLER LAKGQLEAIGEAIVESLEEVKEVFNCSSSLQDL EEREDGTPWRPMSLHLSLGSFVNLPLVNLPIWM QSLAALKHIEDCKTLESMPSWMSKLSLTLR LLECSDSLKRCQKQKPPGEDW/PYIHIKEIDVRE DRC</p>
739	sp_076260_dyus -0.831852	-0.844755	-0.859274		

10

20

30

40

【表 1 4 - 2 9 1】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
740	sp_076300_hzig #N/A	-0.370156	#N/A	RPM1 interacting protein 2 (RING2); CONTAINS InterPro DOMAINs: Zinc finger, RING-type (InterPro:IPR01841), Zinc finger, C3HC4 RING-type (InterPro:IPR018957), Ubiquitin system component Cue (InterPro:IPR003892); BEST Arabidopsis thaliana protein match is: RPM1 interacting protein 3 (TAIR:AT5G51450.1); Has 2005 Blast hits to 1973 proteins in 236 species: Archae - 0; Bacteria - 2; Metazoa - 699; Fungi - 246; Plants - 674; Viruses - 9; Other Eukaryotes - 375 (source: NCBI BLINK).	MGVRYLAISAVCTVSMIGLQCVTEISLDKFKAD GLIGNDTLYSENASRIIEILLRSHATITLAAICVINA FVLFILSIKTIFFGELYAPETRKLVERLVNYYVYKGT FLP_IIPPTIFQSSLWLTWLTVCFLKMFQALARD RLERLNASPSATPWTYFRVYSALLLVAFDSLWI WFCAAVTGPOVSSYTYWLLFFEPLSIAFETLQAV VHGFQIMDVWLHDSAAETSDCQRSKLFDTTAG SLWEWKSSVRNLGFFLDMMLLMALGHVYLIW WLGHTFHLVDAILFLNIFALLSAVVKRIKGFILK IALGALHEALPDATSEEQIAYDDECAICREFMAKA KRLSCNHLFLACLRSWLDQGLNEVYSCPTCRK PLFVGRSESEASSHSRAYSSDEQLARQLSSGVD RVNAEGGGVFPNQNTPEVNAWRFGAGADSS WLQWPSSQGLDVAGPSTAIRSAGLGRVQMMIR HLASVGENYAQTALIEDTSWLSLWPTSSSQATPSN TTFPPAAHRYPATTTTNTAHRRTAPRFANDSIAN LLAMAETVREVLPHVPDEMIFQDLQQTNSAVTV NNLLQM
741	sp_076350_znwf -0.939367	-0.361755	-0.633401	LUC7 N_terminus domain-containing protein; FUNCTIONS IN: RNA binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: LUC7 related (InterPro:IPR004882); BEST Arabidopsis thaliana protein match is: LUC7 related protein (TAIR:AT3G03340.1).	MDAQRALLDELMGAARNLTEEEKYKEIKWDD KEYCAPMIRFCPHDLFVNTRSDLGPRIHDQK LKESEFNSPRHDSYVPRFEAELVQFCEKLVADLD RKVRRGRERLDQEVPPPPPPISAEKAEQLSVLE DKIKNLEQVESLGEAGKVDAAEALMPKVDLSLV EKAILTQQPLNPATMLTQEKKMAALCEICGSFLVA NDAMERTQSHITGKHIGYGLVDFDIAEKAKE KAREEERLAREKEAEERKQKENEENEGKNRRI SSERDCHRDYRDRSFRDRSRENNRNRDE GRGMDRRNDRNRGRDGGGRNRYRGRDRSR SRSPVRHGHR

10

20

30

40

【表 1 4 - 2 9 2】

A	B	C	D	E	F
1					
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2) sp_107670_ptqx (SOAP2-like)	注釈	aa seq
742	sp_076570_mgcf	#N/A	#N/A	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein; FUNCTIONS IN: asparaginase activity, hydrolase activity; INVOLVED IN: glycoprotein catabolic process; LOCATED IN: endomembrane system; EXPRESSED IN: 21 plant structures; EXPRESSED DUJRING: 13 growth stages; CONTAINS InterPro DOMAINs: Peptidase T2, asparaginase 2 (InterPro:IPR000246); BEST Arabidopsis thaliana protein match is: N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein (TAIR:AT3G16150.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archaea - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI/BLINK).	MAKIEMTFVGFPPAGFQSLIIIIITLLVANTVVANEV EENNGKYPIVIVSTWPEVDVAVRAAWSVVDGGFS AVDAVQGSACEQLRCDGTV/GPGGSPDENGE TLDALVMNGATMEIGAVAAVMRYVKDGRRAARLV MQYTEHTLLVGEQASVFAISMLPGPTNLSSEES LHKWNQKDNINCPNFRKNVFPDSDSCGPYHPK ENVGFGTSCPTAYLVETMTRKVPINIHHDITIS MAVIDKTGHVAVGTSTNGATFKIPGRVGDGPIPG SSAYADDEVGACGATGGDMMRFLPCYQVVEES MRRGMPEKFAAEDAISRILKFKFPDFVGVAFADNS SCIGRLCHAESAAEIKFSDVYATELADWGVARQL TKCCFLHQDSEMYLFFVHAMQRKTKTASFVHLV EYTFGRDQLVFNPKSGKNGQNIWGVVAPIF SRAKVRQVTVTERAGHAFDLVSSMSNRELISYD GIVAVGGDLFNEILNGLLSRRFKTEFFPPVPEF MASFGEDPSTSVPIPEINAADVEDVEFSPHEWF RFGLIPAGSTDAVICTTGARDPITSALHIVLGRIG LDIAQVVKWKATMDSDFEFSVHYTASFAGDNTE LRDLRRF
743	sp_077240_nted	-0.3463992	-0.3933598	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G23150.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archaea - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI/BLINK).	MQVLSDVKTFQVLLSEICILEQHHRDYRIAGCSEL LTFRLCOSSRPKVLEKSSGLVAGVALTGHSHWHE KSLH
744	sp_077350_qfhz	-0.353065	-0.3933914	AT-hook motif nuclear-localized protein 1 (AHL1); FUNCTIONS IN: DNA binding; LOCATED IN: mitochondrion, nucleolus, nucleus, cytoplasm; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: LP_04: four leaves visible, 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAINs: Protein of unknown function DUF296 (InterPro:IPR005175), AT hook, DNA-binding motif (InterPro:IPR017956); BEST Arabidopsis thaliana protein match is: AT hook motif DNA-binding family protein (TAIR:AT4G22770.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archaea - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI/BLINK).	MATATPPVNSMAGVGERVPTFTSSDKKKRGRP RKYGPDGAAAAAALSPMPISASIPLSGDFSGWK QIPRRPFQNIKKAKHKVSVQTESLGDRLAYFVYG ANFTPHVITV/NAGEDIAMKVL/SFSQQESRAICVLS ASGISNVTLRQPCNSGGTLTYEGRFEILSLSGS FMPTDNAMEKSRSGMSVSLAGPDGRVLLGGGLA GLLVAASPVQVVLGSLFGHHQEQKPKKQKRVVE PVTAMPPIPNNHQPDPDGGGGYGGVHKNVFT SPPPSPFHTENLPSMNSHILDPMDEADHNNVST PEGSDDDFDDEDDDDSLQI

10

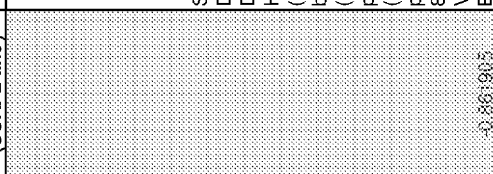
20

30

40

【 1 2 9 2】

【表 1 4 - 2 9 3】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAPf)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
745	sp_077630_mkip	#N/A	#N/A		<p>SPLAYED (SYD); CONTAINS InterPro DOMAIN/s:  DEAD-like helicase, N-terminal (InterPro:IPR014001),  DNA/RNA helicase, C-terminal (InterPro:IPR001650),  Helicase/SANT-associated, DNA binding  (InterPro:IPR014012), Helicase, superfamily 1/2, ATP-  binding domain (InterPro:IPR014021), SNF2-related  (InterPro:IPR000330); BEST Arabidopsis thaliana  protein match is: Homeotic gene regulator  (TAIR:AT3G06010.1); Has 37993 Blast hits to 29119  proteins in 2440 species: Archaea - 313; Bacteria -  8300; Metazoa - 11402; Fungi - 6833; Plants - 2357;  Viruses - 252; Other Eukaryotes - 8536 (source: NCBI  BLink).</p>	MYKELNDVADVPTASDAVDLRSGSTLEAQSRED EOIHGSDDGKVDADFIDISTASGAVDIGIVSTLESLS RVDEQIHGPDGKVDVDAVDAVDAVDAVGLGIASL ESHRSREDDGQYHGLGDKGKVDVAVGAPTASDAVSL GIAATLESRSREDDGQYHGLGDKGKVDVAVGAPTAS DDVRLGIASLTDLSHSNEDEQFHGHVGVKVDAA SCAVAAPSLVLEMEVSRDTNDDTADTNVGGPDS DAAATSSLESKQKELHGPDHDSGELEIGDAPTV CCVLGSEAAAIISITKLINEKEEFLGSEVGYTVNDA ADAPAVSLAVVSTTVTSTWGLEDKGELCSPSDIN AVHDAGNAPVVLSSVSDIAVTPMLASIVTEEPRV ADNSNGDEHVAGPAGNAPSTPNIVEFDTAVCSIQ EKEELEGHNDGGAEILDVGESPTDSTIFGSDIAAV LTSVLTKEDCLGPEAGIVEHDAVDVSSACIVGF DKQAASSKDKKKECHGPDSDYDKLKADDITYNLH VGSDLAETSTLIEEECNGLDRNGELDAVDAPTA FHAVRSGASSESEKKECHGPKDGDVKHDDTGD FKGSDTPTETSTLSSTDKKCKHGCTD

10

20

30

40

【表 1 4 - 2 9 4】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
746 sp_077930_opdp	#N/A	#N/A	-0.9324533	<p>ATP binding microtubule motor family protein;            FUNCTIONS IN: microtubule motor activity, ATP binding; INVOLVED IN: microtubule-based movement; LOCATED IN: chloroplast; CONTAINS InterPro DOMAIN/s: Kinesin, motor region, conserved site (InterPro:IPR019821), Kinesin, motor domain (InterPro:IPR007752); BEST Arabidopsis thaliana protein match is: P-loop containing nucleoside triphosphate hydrolases superfamily protein (TAIR:AT5G60930.1); Has 10625 Blast hits to 10139 proteins in 419 species: Archae - 2; Bacteria - 255; Metazoa - 4769; Fungi - 1373; Plants - 1690; Viruses - 3; Other Eukaryotes - 2533 (source: NCBI Link).</p>	<p>MEWNSPEPRSTVRRNPHRSARPTTTNAAPKT            STSHLSKLNKIPNFDLDDLSMDIQLNTNPKPS            PNPNSLNVFLRVPLPIAPPKSKQIKPSKSKSC            LTINDSOSITSPPIRDNRLKSEVFNFGFSHVFSDDS            SQDEVYEKLMRPVMEEFNIGRSGLLAALGPSPGS            GKTHVFGSAKEPGLVPRALRRIEESDAGKSSLV            SRNYLSVFEIHSEKGGKGLSDLTPDGSDIYMQQ            SAYKGLKELVKDVAQAESLLTHAKSKRATATTN            SNSOSSRSQCIIRCAAEEDTEGESEAVLTFVDLA            GAEREKKTGNKGGRLLESFNINNTSMVFLGLR            SLEHQKNPKPLQKHFKSSLLTRYLRDYLEGRK            RMAILLTVRSGEEDYFDTSSLLRQASPYMEIKFC            DIEETLNLPRNKRHRHQIIGNEQKRRVLLTIADGRN            DVTVISRDSETFQTEAKSREILDGNGGIPVSRVD            ENLLPEVNTDSNIHFQYYFMQGFSAKALWVWLK            QYKERVKTMQKDINILTENLRVGDGAKALVMEEEEL            LKLSQCSQCPNQVIGVTSQSMGGLNSEPSFCEE            EFTDGHGVVSDVOGLACLECKEKEILHSSGKEK            TLHVCELLFGVAPEEPHEPCEVSECKTTNLSCEKA            LHTEGKVDVTLIPKSENHDLSEESQSGSVTDGL            QELGSRVNAKLVHLGDSGSGPKSILSVDESID            TDFILTSPhKANLSSNELKIFSPHNHOSLMEYQTP            HDMSKIDCHESGQRKCKDIQSTDAACCLHRSHV            VDLIDGEGESGRDTPNNIDKENAGNGPQSKSRL            EKKETSHSKPSKAGRPKRRLMPASLLLRDTGSI            GAIEENIHFRRGGKGGQKLVEDEHFKRTKGNACLIM            RLLKNHLPR</p>

10

20

30

40

【 1 2 9 4】

【表 1 4 - 2 9 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
				<p>MATERNAL EFFECT EMBRYO ARREST 5 (MEES5);            FUNCTIONS IN: translation factor activity, nucleic acid binding, GTP binding, translation elongation factor activity, GTPase activity; INVOLVED IN: regulation of embryo sac egg cell differentiation, embryo development ending in seed dormancy; LOCATED IN: nuclear speck, cell wall, plasma membrane; EXPRESSED IN: whole plant, cultured cell; CONTAINS InterPro DOMAIN/s: Protein synthesis factor, GTP-binding (InterPro:IPR000795), Translation elongation factor EFG/EF2, domain IV (InterPro:IPR005517), Small GTP-binding protein (InterPro:IPR005225), Translation elongation factor EFG/EF2, C-terminal (InterPro:IPR00640), Translation elongation factor EFTu/EF1A, domain 2 (InterPro:IPR004161), Ribosomal protein S5 domain 2-type fold (InterPro:IPR020588), Ribosomal protein S5 domain 2-type fold, subgroup (InterPro:IPR014721), Elongation factor G/III/V (InterPro:IPR009022), Translation elongation/initiation factor/Ribosomal, beta-barrel (InterPro:IPR009000); BEST Arabidopsis thaliana protein match is: Ribosomal protein S5/Elongation factor G/III/V family protein (TAIR:AT5G25230.1); Has 61019 Blast hits to 57504 proteins in 4658 species: Archae - 1314; Bacteria - 38464; Metazoa - 2393; Fungi - 1714; Plants - 1333; Viruses - 1; Other Eukaryotes - 15800 (source: NCBI BLINK).</p>	<p>MDDSLYDFGNVIGPEIDSDQDDSDNDIPSSDDA            DKRPSDDEAPSSAAAGPTNGWMTTISHDEEDIEI            VDNQVLAEDKYYPTAEVYVGDVETLVEDED            ELHLEAIKIPV/KALFELGKVDSSSTVSTQFMLG            LMSNPNLVRNVALVGVHGGKTLFMDMLVEQTH            HISTDPKNERHMYTDRVDEGERRISIKSVPM            SLVLEDSNSKFLINIMDAPGHVNFSDMSAALRL            ADGAVLVDAEAGVMVNTERRAIRHATQERIPV            VINKVDRLITELKPPKDAYHKLRHLEIINHITAA            SSSTAGDVQLIDPAAGNVCFASATAGWSFTLLSFA            KLYVKLHGVAFDANIKFASRLWGDMMYYHSDSRFT            KRKPPATGGERSVQFLEPLKYISQVIGEHRS            VEATLAEGLVTLPAAYRLNVRLLRACSTVFG            TATGFTDMLVQHVPCGKEASSRKYVDHIYGPKD            TVVYKSMENCDSGGLMNVNITKLYPKADCSVFA            FGRVYSGTIQTGQAVRVLGEGYSPEDEEDMTVK            EYTKLWVYQARYRIPISKAPPGSWVLIJEGVDASI            MKTATLCNVDFEDAYIFRPLQFNLTLSVVKTATE            PLNPSELPKMWEGRLKISKSYLAITKVEESGEHT            ILGTGELYLDSIMKDLRELYSEVEVKVADPVVSFC            ETVETSSMKCFEETPNKKNKITMIEPLERGLA            EDIENGVSIDWPRKKLGEFFQTKYEWDLAARS            VWAFQPKQGNILLDDTLPSVDKLNLLNAVKDS            IVQGFQWAGREGPLCDEPIRNVKFKIVDARIAPE            PLHRGTGQIPTARRVAYSFAFLMPELPEVY            VEIQTPIDCVSAITVLSRRRGGHVTADVPOPGTPA            YLVKAFPLVIESFGFETDLRYHTQGGAFCSVFD            HWAIVPGDPLDKGIVLRPLEPAPIQLHAREFMVK            TRFRKGMSEDSVSIQKFFEDDAMVMVELAQQAADLH            LQMM            MKQGSRRRMYVGLAMTLMFISIMMPYVCGEETT            LLGNLEKNGNLRKALFRADP</p>
747	sp_077950_eczu	#N/A	#N/A		
748	sp_078240_fxen	#N/A	#N/A	uncharacterized protein	

10

20

30

40

【表 1 4 - 2 9 6】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiag (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
749	sp_078890_ahzz	#N/A	#N/A	<p>prolyl oligopeptidase family protein; FUNCTIONS IN: serine-type peptidase activity; INVOLVED IN: proteolysis; LOCATED IN: membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Peptidase S9, prolyl oligopeptidase, catalytic domain (InterPro:IPR001375), Peptidase S9B, dipeptidylpeptidase IV N-terminal (InterPro:IPR002469); BEST Arabidopsis thaliana protein match is: acylaminoacyl-peptidase-related (TAIR:AT4G14570.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).</p>	<p>MHPVGTAGRKNTLKRSLDMPATVTTDMS MVSOKFAVDSDSGLPLFSVEEIVQYPLPGYVA PTLSFSPDDNLTLYLFSFDHLSLRKVFNFLETG KHELVSFSDVGGDLDESINISEEKLRRERERGL GVTRYEWKMLSSKKIIMVPLPDGVVYLYLSSSS KVDLKLQKTPSSPIIDPHLSPDGNMILAFIRGYELH VFDILYNKEKQLTFGAKGSLTHGLAEYIAQEEMD RKNGYWWSADSKFAFTQVDSSEIPLYRIMHOG KSSVGPDAQEDHAYFPAGASNKVRGLGVSIAG GPITWMDLHCGGNDDEEYLARINWMPGNILIAQV LNRAQTRLVKLFDIRTGQRKVLVMEEHDKWVN LHDCFTPLDKSVSGGGGFLWASEKGTGFRHLYL HDVDGHCLGPITTEGDWVMEQIAGNEATGVLVYF TGTVDGPLENLVSKLVNRSGLQAPVRLTH GNGKHVLDHQMQRFIDIDHSLTSPKKVYVLSLR DGSLLTLYDQPSGVPFRFRKRLKLEPPEIVQIKADD GTVLYASMYKPDRETRFPPPPYKTMISVYGGPSV QLVSDSWINAVDMRAQYLRSKGLVWVKLDNRG SARRGLDFEAAVKHFRGLDAEDQLAGAQLIN KGLAKOGSGLYGVSYGGYLSAMTLARFPEVFT CAVAGAPVTSWDGYSFYTEKYMGLPSENLAGY EFGSVMHHVRNVKGGKLLLVHGMIDENVHFRHTA RLVNAFVSAGKPYELLIPFDERHMPRRLRDRYIM EERIWEFIQRNL</p>
750	sp_079190_rnwe	#N/A	#N/A	<p>histone deacetylase 9 (HDA9); FUNCTIONS IN: histone deacetylase activity; INVOLVED IN: histone deacetylation; LOCATED IN: nucleus; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Histone deacetylase (InterPro:IPR003084), Histone deacetylase superfamily (InterPro:IPR000286); BEST Arabidopsis thaliana protein match is: histone deacetylase 1 (TAIR:AT4G38130.1); Has 8723 Blast hits to 8511 proteins in 1453 species: Archae - 219; Bacteria - 3191; Metazoa - 1512; Fungi - 548; Plants - 472; Viruses - 0; Other Eukaryotes - 2781 (source: NCBI BLINK).</p>	<p>MRSKDKISFYDGDVGVSYVFGPNHPMKPHRLQ TTHLVLSDYDLHKKMEIYRPHKAYVYLAQFHSAD YVDFLHRIPTDKQHFFMSEMAYKYNLGEDCPVFD NLFEFCQIYAGGTIDARRLNKLCDIANWAGGL HHAKKCEASGFCYINDLVLGILELLKYHARVLYDI DVHHGDGVEEAFYFTRVMTVSFFHKYGDLFFPG TGDVYKENGREGKKNYAINVPLKNGIDDA5YTRLF KAISKVVEFYQGVVLOCGADSLSGDRLGCENL SIDGHAEFCRVFKFNPLPLLVTTGGGTYTKENVAR CWTYVETGLDLDLNEIPENYIKYFGPDYSLRI PNGLIENMNSKSYLTTIKMQVLENLRCIQHAPS QMQEYPPDFYIPDFDGDENPDERINQHTQDKLIQ RDDEYEGDHDNDHMHDDV</p>

10

20

30

40

【表 1 4 - 2 9 7】

A	B	C	D	E	F
1	分析に用いたバイトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
751	sp_079390_kohp	#N/A	-0.853219	cell division cycle 5 (CDC5); FUNCTIONS IN: DNA binding, sequence-specific DNA binding transcription factor activity; INVOLVED IN: defense response signaling pathway, resistance gene-independent, defense response to bacterium, defense response to fungus, regulation of transcription; LOCATED IN: chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF3351 (InterPro:IPR021786), SANT, DNA-binding (InterPro:IPR001005), Myb, DNA-binding (InterPro:IPR014778), Homeodomain-like (InterPro:IPR009057), Myb transcription factor (InterPro:IPR015495), Homeodomain-related (InterPro:IPR012287), HTH transcriptional regulator, Myb-type, DNA-binding (InterPro:IPR017930); BEST Arabidopsis thaliana protein match is: myo domain protein 3r-4 (TAIR:AT5G11510.1); Has 13137 Blast hits to 10813 proteins in 768 species; Archae - 14; Bacteria - 356; Metazoa - 3570; Fungi - 914; Plants - 4936; Viruses - 8; Other Eukaryotes - 3339 (source: NCBI BLINK).	MRIMIKGGVWKNTEDEILKAAVMKYGKNQWARI SSLLVRKSAQCKARWYEWLDPISIKKTEWTFREE DEKLLHLAKLMPQWRTIPIVGRTPSQCLERYE KLLDAACTDDNYPGGDDPKRKRPEIDPNPESK PARPDPVMDDEKEMLSEARARLANTRGKKAK RKAREKQLEEARLASLQKRRELKAAGIDPRQR KRRKGGIDYNAEIPFEKRAPAGFYDVAEEDRPVE QPKFPTTIDELEGERRVDREARLRKQDAKNIKAE RQDAPGAIMQANKLNDPETVRKRSKMLLPPQIS DQELEEIARMGYASDLLVGNQELAEQSGATHALL ANYSQTPROGMPMRTPQRTPAKGGDAIMMEA ENLARLFRESQTPLLGGDNPELHPDFSGVTPKK RDVQTPNPMLTPSATPGGAGLTPRIGMTPSRDG SSFGMTPKGTPIRDELFRINEELDMHDSAKLELRR QAEKRNRLTGLSGLPEPKNEYEVVVRPAPEDT EPEEKIEEDMSDRIARERAAEERQOALLKFRS KVLQRELPFPVAVSELIRHSLMRAGEDKSSFPV PTDIOAEIVRKELLSLLEHDMNAKYPUNDKYQKK GSKRSANGVPIIEFEEDVYLKEAASFIKEEAQYLC VAMGHQEETLDFVEAHNTCLNDLMLYFTRNAY GLSSVAGTVEKLAALRNEFENVRKMDMDNKKKA QRLEKIDVLTHGYKMPADKLTQIELTFKQMT GGTELDLQROENLAASYRIKGLGWEVQKQ KEYEKTQKRYGERVAEQERVQVMDQYRLOA TTEVDDSSNFVFKHVEEAPVSLQNDIHVDGDV QQLEIEVKKRTEELAAAFVAKEASFKSIAPPEPSS TTEVDDSSNFVFKHVEEAPVSLQNDIHVDGDV ASNADIKNSLASETPLEDSSKESEAAKEQNVDD NTNVVDGNLENDEAAKEQNVDDNSNVVDGNLE NSDEAAKEQNVDTNVVDGNLENDEAAKEQNA GEAVENNGFVNGEPVIEAGVDGKLGSGQINKAE EGNETMTEDAGNPVESVENGASDNLV
752	sp_079840_xqus	-0.973544	-0.876988	Cytokine-induced anti-apoptosis inhibitor 1, Fe-S biogenesis; FUNCTIONS IN: molecular function unknown; INVOLVED IN: biological process unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF689 (InterPro:IPR007785); BEST Arabidopsis thaliana protein match is: Cytokine-induced anti-apoptosis inhibitor 1, Fe-S biogenesis (TAIR:AT5G18362.1); Has 539 Blast hits to 538 proteins in 227 species; Archae - 0; Bacteria - 0; Metazoa - 183; Fungi - 185; Plants - 75; Viruses - 0; Other Eukaryotes - 96 (source: NCBI BLINK).	MAMLALTDVILPLISVSSSKNFELNILEGPVLDNAL VITLASTLSALPLETSSMDTVIGSKARDFFAVPL LAEMFRILKPKDGRILIHLELSDGLPFDNILLSSLQR KILVAGFLPEEAGQKQPDFAFSGVQSIQVAKAKPS WKIGSSFSIKASKGLPKVQINDMDIDEDILLTE EDLKKPQLPVPVGDCEVGVSTRIKACKNCSCGRAEE EQKVEKGLTMEQLDNPQACGSCGSLGDAFFRC STCPYKGLPAFKLGDYKVSLSNFI_DADI

10

20

30

40

【 1 2 9 7】

【表 1 4 - 2 9 8】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
753	sp_079920_jixy	-0.931248	-0.885115	26S proteasome regulatory complex, non-ATPase subcomplex, Rpn2/Psm1 subunit; FUNCTIONS IN: enzyme regulator activity, binding; INVOLVED IN: protein catabolic process, ubiquitin-dependent: protein catabolic process; LOCATED IN: proteasome regulatory particle, base subcomplex, proteasome complex, nucleus, plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINS: Armadillo-like helical (InterPro:IPR011989), Proteasome/cyclosome, regulatory subunit (InterPro:IPR02015), Armadillo-type fold (InterPro:IPR016024), 26S proteasome regulatory complex, non-ATPase subcomplex, Rpn2/Psm1 subunit (InterPro:IPR016642); BEST Arabidopsis thaliana protein match is: 26S proteasome regulatory complex, non-ATPase subcomplex, Rpn2/Psm1 subunit (TAIR:AT2G32730.1); Has 1171 Blast hits to 1071 proteins in 297 species: Archae - 14; Bacteria - 33; Metazoa - 390; Fungi - 382; Plants - 143; Viruses - 0; Other Eukaryotes - 209 (source: NCBI BLINK).	MFRRSQYGRPLLILSPAQQDSSELVIAASHHTDPG VLTLQLDQDFGGLQVKEGKWNVVKPIQVEYTV KLAASGFVFLKDLRPEEPEVLSLTDAPASASATT TNSDDADAPVQQGTVAAMA VDEEPPQPPPEY TS
754	sp_081000_gwez	-0.930794	-0.880994	protein phosphatase X 2 (PPX2); CONTAINS InterPro DOMAIN/s: Metallophosphoesterase (InterPro:IPR004843), Serine/threonine-specific protein phosphatase/bis(5-nucleosyl)-tetraphosphatase (InterPro:IPR006186); BEST Arabidopsis thaliana protein match is: protein phosphatase X 1 (TAIR:AT4G26720.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MSDLKQIEQLKRCEPLKESEVKALCLKAIELVE ESNVORVDAPVTICGDIHQGFYDMKELKRVGGD CPKTNYLFLGDFVDRGFYSVETFLLLALKVRYYP DRITLIFGNHESRQITQVYGFYDECLRKYGSVNV WRYCTDIFDYLSSLALIENKIFGVHGGLSPTITLD GIQRTIDRKGQEVPHDGMCDLLWSDPEDAVDGGW GMSPRGAGYLFGGSVYTSFNHTNIDYICRAHQ VMEGYKWMFNHQITVWVSAPNYCYRCGNVAAIL ELDENLNKEFRVFOAAPQDARGTPAKKPPADYF L

10

20

30

40

【表 1 4 - 2 9 9】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
755	sp_081150_szix #N/A	sp_083032 #N/A	#N/A	<p>decapping 5 (DCP5); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: negative regulation of translation, deadenylation-independent decapping of nuclear-transcribed mRNA, cytoplasmic mRNA processing body assembly; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: DFDF motif (InterPro:IPR019050), FFD/TFG box motif (InterPro:IPR019053); BEST Arabidopsis thaliana protein match is: decapping 5-like (TAIR:AT5G45330.1); Has 14874 Blast hits to 10327 proteins in 603 species: Archae - 12; Bacteria - 902; Metazoa - 6418; Fungi - 2978; Plants - 1226; Viruses - 261; Other Eukaryotes - 3077 (source: NCBI BLINK).</p>	<p>MAAEASRSATGSADSYLGLSILTSKSEIRYEGVL FNINTDEASIGLRNVRSEFGTEGRRKDGQQIPAND KVVEYVFRGSDIKDLQVKSSPPVQTPAPPNDP AICSHYAAPTASSTVPTASVPAAGLLSDFNNQ MSQPGVPTTYQNSVPMYQPGGNSLSPWNPSTP PTSNVSGVSVPTYWQYGYGSPSGSQMQQOQSL HPVQGLSMPPSPGPPSAQOQLQYPTMNSYLPT GASNISASQSMEEFPPILLQPIVSSTLMSQYNVHP DQSPVLSAESSTNRMFSMYPVPAQAVPVVPOSTKS PLLPVLSNVLDKSTVGP LIPNOAKPISTTSMFPN SFSELSAAGRKSTSPALVTPGQFLQPGPTVSQ TAQTEQKDVWQVQSP LVSPPQPS TTSATTAIV PGSAEQAQPLLPLPSPNSKINEFFSQTRYVNVN RGGRRGGRRGGNEFSRPPVANFTEDFDFEAMNEK FKKDEVWGHGKGNKAKADDGLQDEYDDDDVS EIDSKPVYKDDFFDFTLSCNALDRESRNGRTRFS EQMRLDTEFTGNFRNRFRGGRGPGGGRRSRG GYGRGYGYGYVGRGRGHNNVYRAI</p>
756	sp_081630_zsis #N/A	#N/A	<p>ETHYLENE-DEPENDENT GRAVITROPISM-DEFICIENT AND YELLOW-GREEN-LIKE 3 (EGY3); INVOLVED IN: response to high light intensity, response to hydrogen peroxide, response to heat; LOCATED IN: chloroplast; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 9 growth stages; BEST Arabidopsis thaliana protein match is: Peptidase M50 family protein (TAIR:AT5G35220.1); Has 660 Blast hits to 652 proteins in 203 species: Archae - 107; Bacteria - 240; Metazoa - 53; Fungi - 13; Plants - 117; Viruses - 1; Other Eukaryotes - 129 (source: NCBI BLINK).</p>	<p>MATLFISSSSWVSNKTKKLASFSSNSHFHHPV TLSSYKPSKPRKLRFSQQDEGERESSVWVD EKSKNDEDTONGVEEREVEYDDEKQQLDWDKN DAEFKFMGNPISIAAILEKKRADRKLKELDDE SKSDNPIVGVFNRIAPENILLREKETLEKVEQAFKA LDLNVCLVQPKSEIDLQFESTNLSTPWGYLSAIA LGVTTFTGIALTSGFFVKFGATFDDYLANVPLFG GFLTILGVSEIVTRLTANRYGKLSPSFLVPSNWT GCLGYMNNYELLPSKALFDIPVARTASAYLTS LALAVAFASDGSFNGGDNALVRFQFFENNPLF SFIQYVIGYVADDLGNLVAVEYGVGVVPLAF AGLLGMVVTSLNLLPCGKLEGGRIQAAMFGRNT ATLLSFATS.LLIGIGLSSVICLSWGLFAFFFRG GEENPAKDEITFLGDDRYAWGVWVFLICFLTLFP NGGGTFFSSFFTGPFRRGDL</p>	

10

20

30

40

【表 1 4 - 3 0 0】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh sp_107660_kiqg (SOAP1)	sp_107670_ptqx (SOAP2-like)		注釈	aa_seq WSSDRQLVYTRDLVEEARRRRIIVLSIVVGLSTYIM SLTSTSVWMLPAAAVLIIFRYFSLDFETRRKEA AYNSRSGKITQEKPLDVPKTLVKKQDQWRKVN SPVEDAIDQLTRHIVSEWYTDLWYSXGEVLKEL VOLMNGVLEEVAYCMKHINLIDLKRDVHLICKHL ELFRVCOAKIKKNLASLDFQDRDMELKVLGSEN KLHPLFSIESEHKVLQIMDGLLSLTFKPEDLQC SFFRYVSHELLACAVIRPVNLNANPRFINERIQV VISARKNGEAAATNNSSLLKONGGNSDHFSPFP DPSVGGVELVELKSIQSKSMKLGKEKVDSEDA SKDPLLVDTRSTRSWSSLPFGSAAAAGSHDRS ASGGEGDMILDYFSRKKKQTLAPENFENMMWSK GRDYNKDTQITNHVHPGSSIKRSDAVNSSKVFA KQKERHNIAEAERGVLPGSNVRPQSETMYISAD SNMSTPPLTSDDEDDEQKMETELLSQSSQSSS GDEEDRVTLGDSPTTKVWDGRTNRSVGYSHIRH PLEGLEGGKGNFGQYQKLRQOQSGRKRBSR LGSQKVLWQEVERTNLPFGDGKDVHPLKPNVK DDSSDDYDMEIVGRLOSGTTASSSASSLPGSF TSTVNSPQTAFLDLTFKLRCEVLGANIVTSSKT FAVYSIVIDVERNSWISIKRRYHFEELHRLKEY PEYNLHLPKHFSLSTGLDLVLQICERCQLLGGYLK KLMQLPTVSGSIDWDFLSDISQTYTFSSAFSIIET FSVKPHDSSANMSSLGGPKVNHLPQPRRDAESE NKDQSPKPNVKEVKGKLAGDSDSRIQSNRKGKN EGAVQQNVGKEGDIADDPTFPMWVPPNLSGPI LDLVDVVFLQEGAWIRKAFWVAKQVLQJLGM GDALDDWLVEKIQRLKGSVAVASGKRVQILWP DGIFITKHSRQRPPASGSPQSSDGRKIPDQLS SPKAENFEDIAELQEEEAERAKFVVELIMDNAP AAVVSIVGHKEYEQCAKDLYYFLQSSVCLKQLVM DLQLLVLHAFPELHDVHFKQLDEEKHKFGEFKAT  MKILAVLLPFSQVHLAFRLPAAIAAFAKSVQYRGA SVAPLKNYIMADVALKLHSSSACYLNTSEYS
757	sp_082110_zxgk	-0.833367	-0.8352833	-0.811709	Phox-associated domain;Phox-like;Sorting nexin, C-terminal; CONTAINS InterPro DOMAIN/s: PX-associated, sorting nexin 13 (InterPro:IPR013996), Sorting nexin, C-terminal (InterPro:IPR013937), Phox-like (InterPro:IPR001683), Phox-associated domain (InterPro:IPR003114); BEST Arabidopsis thaliana protein match is: Phox-associated domain;Phox-like;Sorting nexin, C-terminal (TAIR:AT2G15900.1); Has 944 Blast hits to 779 proteins in 177 species: Archae - 0; Bacteria - 4; Metazoa - 546; Fungi - 164; Plants - 152; Viruses - 0; Other Eukaryotes - 78 (source: NCBI BLINK).	
758	sp_082510_cuwe	#N/A	#N/A	-0.88478	uncharacterized protein	

10

20

30

40

【表 1 4 - 3 0 1】

A	B	C	D	E	F
1					
2	分析に用いたペアトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
759	sp_083500_skgc	-0.87782	-0.852477	DWARF 4 (DWF4); CONTAINS InterPro DOMAIN:s: Cytochrome P450 (InterPro:IPR001128), Cytochrome P450, conserved site (InterPro:IPR017972), Cytochrome P450, E-class, group I (InterPro:IPR002401); BEST Arabidopsis thaliana protein match is: Cytochrome P450 superfamily protein (TAIR:AT5G05690.1); Has 31120 Blast hits to 31036 proteins in 1650 species: Archae - 61; Bacteria - 5303; Metazoa - 10591; Fungi - 5722; Plants - 7996; Viruses - 3; Other Eukaryotes - 1444 (source: NCBI BLink).	MSDLEFFLYLIPFVLA-SWTLNLFKRRKPKPNLP PGNSGLPFIGETFGYLKPHATTIGKFMIEQHIERF GKYRANLFEPTIVSADAGLNFRIFQNEGRLLFES SKYPSIGGILGKWSMLV LNGDIHRDMRNISLNFLS QSRLKAHLMEKVERQAVFV LNTWQEDCSFSAQ DEAKKFTFNMAKQILSLDHGEAEETEHLKKEYINF MKGVVSAPIINLPGTPYRRA LKSRRTALDFIEMKME EKLKKKMMNDGSENDEDDL GWV LKHSTLSKE QILDLSLLFAGHETSSVSIALAIFLFGSPSAVEQ LREEHLIARAKQOQSAEKELNWEDYKMKMEFTQC VISETMRLGNVRFHLHRKALKDVRYYKGYDIPCGW KVLPLYAAIHLDFAFYDQPHVFNPRWVQOINGIN RGGVSPSATASNSNSNSPTSGNNFMFPFGGGQ RL
760	sp_085060_arwu	#N/A	-0.86872%	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; Has 428 Blast hits to 428 proteins in 59 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 428; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink).	MASSSTAQNPRHLNHLGGTTSAPPTPSTQ PSLSSSTDAADTLRLLHRLPPSLPTRLSPR APAKSSPHLPTLSLSTSPSSGYDSNLFELGYF QLTHSITPKLAESAESLDFDLPLDEKLSSTP TDWPLGNDDDDDDDDNDGDDHOTGTSFFLDGQ CSATNAELSTLSREL TRELEKLGLEWEKLASA VGFENPREGVGVSNKQSLMWTSSDGDKA RRVYVYGLQYQIRPRKWDLVSDSGSVSVEPQ VGSVLTLDIAQVWVSNGLKVKVGRPAPAFQDI KSNISMTMLTLPLESTVSPLLLVPNYTGEKEANN QOKDKDNQTNFNTNSNDGNDVDNEDORINFSF SFEDYAWRVYHERLLFKDPLDRYRVIH
761	sp_085280_uxwe	#N/A	#N/A	DA1-related protein 2 (DAR2); FUNCTIONS IN: zinc ion binding; LOCATED IN: chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN:s: Zinc finger, LIM-type (InterPro:IPR001781), Protein of unknown function DUF3633 (InterPro:IPR022087); BEST Arabidopsis thaliana protein match is: LIM domain-containing protein (TAIR:AT4G36860.1); Has 1539 Blast hits to 1276 proteins in 109 species: Archae - 0; Bacteria - 10; Metazoa - 1178; Fungi - 26; Plants - 170; Viruses - 0; Other Eukaryotes - 155 (source: NCBI BLink).	MADATQIHLSQPCYIGLSAPPNTEKKSFRMFWL SKIFKGGGGGGSSRSSSAGGRHPQFLEDEN MVWRAPSRSLDRPRASKDQELDRALAFSLSD DLKRPSCKRRTDNENTAKGLQDSLHRPSYTP YMPSPYDYPYDFVFPYPRGYRLCSGCNGEIGY GNYLGMGTFFFHPECFCRSCSYPTTEHEFSL GKDPYHKTCKFKELTHPKCEVCHOFIPTNTAGLIE YRCHPFWSQKYCPSHEHDQIARCCSCERLESW NTRYIALGDGRSLCLECMESAIMDTGDCQPLYHA IRDVYEGMNMRIQQIPMLLVERQALNEAIVEGK NGFHMPETRGLCLSEEQTVTSLRPRRIGGHRLL VGMRTQPKLTRRCEVTAIYGLPRLLLTGAILA HELMHGWLKGYRNLSPVEEGICQVLSYMWL ESEVMFGFNMASTSTASTSSFTSKKGTSDV EKKLGEFFLHQIAHDASPAYGGGFRSANAAYNKY GLRRTLDHRLTGSFPI

10

20

30

40

【 1 3 0 1】

【表 1 4 - 3 0 2】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	分析に用いたベイトとの相関係数 sp_107670_ptqx (SOAP2-like)		注釈	aa seq
2				DEFECTIVE KERNEL_1 (DEK1); FUNCTIONS IN: cysteine-type endopeptidase activity, calcium- dependent cysteine-type endopeptidase activity; INVOLVED IN: cell fate specification, embryo development ending in seed dormancy; LOCATED IN: plasma membrane; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN's: Peptidase C2, calpain, domain III (InterPro:IPR022683), Peptidase C2, calpain, large subunit, domain III (InterPro:IPR022682), Peptidase, cysteine peptidase active site (InterPro:IPR000169), Concanavalin A-like lectin/glucanase (InterPro:IPR008985), Peptidase C2, calpain, catalytic domain (InterPro:IPR001300), Peptidase C2, calpain family (InterPro:IPR022684); Has 3377 Blast hits to 2425 proteins in 256 species: Archae - 0; Bacteria - 87; Metazoa - 1580; Fungi - 264; Plants - 115; Viruses - 0; Other Eukaryotes - 1331 (source: NCBI BLINK).	MESEQRIRIVACVVSGLAFALGGLSSSSFSILVVVVN WRPRLYSWIFARKWPDLFQGPLRGVICGALT LAWFVVSPLLLVGVGSW_LILIGRDIIIGLAVIMP GTALLSFYSIMLWWRTRWOSRAVAFLLFAV ALLCAYELCAVYVTAGSSASERYSPSGFFGVS IALAINMLFICRMIFNGGLDVEYVRRAYKAYS DCVEVGPVTCLEPPDPNELYPHOSSRVSHLGT LYLGSLLLVYSILYGLTAKERMMGAITSAAVIV LDWNMGACLFQFLLKSPIAALFIAGASRIFLIF GVHYWYLGHCISYLVMASVLLGAAVCRHLVTNP SAERRDALQSTVIRLREGFRKEHNTSSSSSEG CGSSMKRSSAEAGHLGHVESVGRNGGCHTG EAVGWNINGTLRTTSSHEGINSEKSMDSGRPSL VLRSSCRSVVQDPDVGPSFVEKSDHHSVVGYS SGLESQGYDSSTSTSVNYLPDLNLALVYQEKLN DPRIAAILKRAKQGDRELTNLLQNKGLDPNFAM MLKEKGLDPTLLALLQRSSLDADRDHCDNTDITV DSNSVGDALPNQISLSEELRLRGEFEKWLQILRSL WHFVAGTPERALACFSSVFILETVGVAFSPRPRV KLVNATHQOQFEFGICVWLLSPVICLIMAYLRSLOA EDTSTSSKPRKYGFFFAWLLSTCVGLLSLSKSS VLLGLSLTFLMVAACLSIAVPLWIRNGYNFVWVOD NPAPRTENHQALRIEGIAFLVCVAVFVGSVISLG VIVSFKPLDELSYNGWADGKIYASPYASPYLGG WAMASGIALLVTVGLVPVISWFATYRFSVFGVCL VLFVSVLVAFCVGSYAEITNSRIDHIPTINGDFLAL LPLLCIPAILSLCTGLHKWKDDDEWRLSCSVYVFTI GLLLLCASIVAVIVOPWTVGVAFLLVLLVLAIG AHLWASNNFYLTRACMFFVCFLAFLLAVAAFLV GRFEAKPFVGSVGYFSFLLAGRALTVLLSPPI VYYSRVLVYVYDAHADCCGKNVSGAFMLLYGIA LATEGWGWASLNIYPPFVGAASVAVTLVVSFGF
762	sp_085410_xpia	-0.912433	-0.449752	-0.893079	

10

20

30

40

【表 1 4 - 3 0 3】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_piqa (SOAP2-like)	注釈	aa seq
2	共発現遺伝子			Protein kinase family protein; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: cellular component unknown; EXPRESSED IN: cultured cell; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: Protein kinase family protein (TAIR:AT5G18190.1); Has 16177 Blast hits to 16113 proteins in 1419 species: Archae - 20; Bacteria - 5310; Metazoa - 4811; Fungi - 1519; Plants - 1815; Viruses - 223; Other Eukaryotes - 2479 (source: NCBI BLINK).	MPELRSARRSKRLGDLOQPAPELNDGGENFLP STQNKTRRRGGRRGANAAAGVAKAAIGARQ TGGAGRGRVVRVIDLDPEIPGDGLPQVAPIGVAG ATHNRVEGAAGKDIAMIEGSGGDKLVAEEDSNL TPVPERVQVGTSPVYKIERKLGKGGFGQVYVGR RMVGGTDRIGPDAIEVAVKFEHRNSKGCNCGPP YEWQVNSLNACYGVPVSHFKGROAEYIILMID MLGPSLWDVWNSLQGSQSPNMVACIAVEAISIL EKHLHKGfVHGDKPENFLGPGTADAEKLLYLI DLGLASRWKDVSSGHVHVEYDORPDVFRGTIRYA SVHAHLGRTGSRRDDLESLAYTLIFLIKGRLPWQ GYOQDNKFLVCKKKMGTSPELMCCFCFAPFKQ FLEAVTIMKFDDEEPYAKULSYFDSJIEPCAPLRPI KIDGALKVGNKRGRMLVNLNLEEDEQPKKKIRLGS PATOWISVYNARRPMKQRYHYNVADARLKHQID KGNEDGLYSCVASATNLWALIMDAGTGFAGAV WELSNVFLHKDWIMEQWKEKWFYSSSIAGASNGS SLVMSKGTSTYQQSQYKVSDFPFKWKINKWKE GFHVTSMTTAGSRWGWVMSRNAGFTDQWVLD FLYPSSEGHRRWESGYRTSMAATPDQAAFLSI PKRRTMDETOETLRTSAPPSTHVKEKWAKNLVIA SICYGRTVC
763	sp_086420_utoj	-0.923259	-0.923259	cytokinin oxidase/dehydrogenase 1 (CKX1); FUNCTIONS IN: cytokinin dehydrogenase activity; INVOLVED IN: N-terminal protein myristoylation, cytokinin catabolic process, meristem development; LOCATED IN: vacuole; EXPRESSED IN: lateral root, shoot apex, hypocotyl, root, flower; CONTAINS InterPro DOMAIN/s: Cytokinin dehydrogenase 1, FAD/cytokinin binding domain (InterPro:IPR015345), FAD-binding, type 2 (InterPro:IPR016166), Oxygen oxidoreductase covalent FAD-binding site (InterPro:IPR006093), FAD-linked oxidase, N- terminal (InterPro:IPR016164), FAD linked oxidase, N- terminal (InterPro:IPR006094); BEST Arabidopsis thaliana protein match is: cytokinin oxidase/dehydrogenase 6 (TAIR:AT3G63440.1); Has 6769 Blast hits to 6763 proteins in 1376 species: Archae - 168; Bacteria - 3882; Metazoa - 142; Fungi - 1302; Plants - 645; Viruses - 0; Other Eukaryotes - 630 (source: NCBI BLINK).	MLLKVFLVLLNLAEHNINHCYLPHTPDFTNLS SLELVSSLKKLNILNGNDFENTRRVANDFGNRYH FLPMAYLLPKTVSDISTTIKHVINMGHNSLTYTA RGHGHSLOGSQSQSHGGIVISMESLEEPKMHVHT GEOPYVDVSGGELWLNILHETLKQGLSPKSWTD YLYLTGGTLNAGISGQAFRYGPQINNVHRLEIV TGKGDILNCSAKENADLFYGLGGLGQFGIITRA RISLEPAPRSETEHLLSKLNYISHTLFSQSEVSYN FLNVRHLSEIKLREKGLWDVPHPWLNLLVPKSKI YSFAEEVFGHLTDSNGPILYFVQNSRWVNHKTS MVTPEDEVFLVAFVLTSAIMPSSSGSDGLENLTK NKRILDFCEMAKLGAKQYLPHFDTDQGWRAHFG PRAWDFARRKLTYPDLSLTPGQRIFRKGRAIS MKAMMEDDQIDGDEIDEGDDAGIEDEEDLVM LRRLLGQKWKAKARVELLSLWKWLNPNPGRVYR KMRK
764	sp_086510_tjgw	#N/A	-0.83072%	uncharacterized protein	
765	sp_087460_pewf	-0.901059	-0.9332963		

10

20

30

40

【 1 3 0 3】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
766 sp_087470_gmaz	-0.880831	-0.926714	-0.883488	Noc2p family; INVOLVED IN: biological process unknown; LOCATED IN: cellular component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Uncharacterised protein family IUPF0120 (InterPro:IPR005343); BEST Arabidopsis thaliana protein match is: Noc2p family (TAIR:AT3G55510.1); Has 4709 Blast hits to 3397 proteins in 331 species: Archae - 4; Bacteria - 179; Metazoa - 1693; Fungi - 681; Plants - 249; Viruses - 70; Other Eukaryotes - 1833 (source: NCBI BLINK).	MRCFRMLNCIAAATDTFIPVSMILLMLLEMKELN CPFSGGVSKAVDLRLTILKEACVSSVIDELGEHLA QRSYSVAFFELSFIPAVRLRNFCKCTKVERCRKD MKELNRQIEANSEFTNKKRATISFLPSEPAISSFL EEEEKQSGYPLSKYVATLRQAKQKNVSVSESS VLYGAEPPIISDKMLAESDDEESDDEDEGDKGEA VFNFSQKQKTRDKPPPEDEKTKQSKKRSRS CKDAIDEDIVEYLVLSSVEYDEPMDSDSPPLEETVK TEIGISKRRNRKRKPTIEEKWFNEE
767 sp_087600_wiun	-0.880852	-0.956705	-0.882487	Pentatricopeptide repeat (PPR) superfamily protein; CONTAINS InterPro DOMAIN/s: Pentatricopeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: Tetraatricopeptide repeat (TPR)-like superfamily protein (TAIR:AT1G68980.1); Has 11641 Blast hits to 5218 proteins in 154 species: Archae - 0; Bacteria - 0; Metazoa - 18; Fungi - 47; Plants - 11427; Viruses - 0; Other Eukaryotes - 149 (source: NCBI BLINK).	MWRRAFSLLRPPQTRPFSTTTTATGETPSL YSFLQPSLSLKPQQTASPPSPPTSSKIITPEHK TTLEITLQOSSLTADTDENWKSFKSLTNTSSYPS KSLTNSLITLSSLDYHNLKRAFTAVVFLVEKINP NLLDFDSVRYLSSMAAANYAAPAFSLVKCMFKN RFFVDFDMCGPVLVEISKNGSFVGLGVFHENC RVAIDEKLDWMKPNLDACNAALEGCCSDLESVEE AERVIELMSVGLVGRPNELSGLLGLYLKGLLEG KIVELGSLMGGFGFTDKMVFLSYLRGYVKAGSM EESVSKTVLCSLKAGEGDGSKFSEETTCYCEIVKGF KHGSIKDLAALVAEAQMSEPEENVGEGSTGFGIV NACTVLGMLDKAHSILDEMKAKGAVVGLGVYLSI LKAYCKEYRTAEAAQMVAEINCSGLHLDVSCYDS LIEASSSQDFQSASFSLFRMDRARISELKGSYL TIMTGLTENNRPPELMAAFLEVEDVEDPVEVGTGD WNSIIHAFCKAGLEDASRTFRMAFLHFGFNDDY TLYSLINGVYVAGKYFNVLMWNEVKRRVSTDDY GGFKLNHKLIDAFLYALYKGGFFDAMQVVDKSR EWKIFIDKWRVYKQAFMETHKHLKLVSKLRRKNFKK MEALIAFKNWAAGLSV

10

20

30

40

【表 1 4 - 3 0 5】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
768	sp_088110_anou	-0.334044	-0.873172	Helicase/SANT-associated, DNA binding protein; FUNCTIONS IN: DNA binding; LOCATED IN: apoplast; EXPRESSED IN: male gametophyte cultured cell, pollen tube; EXPRESSED DURING: M germinated pollen stage; CONTAINS InterPro DOMAIN/s: SANT, DNA-binding (InterPro:IPR010005), HSA (InterPro:IPR006562), Homeodomain-like (InterPro:IPR009057), HAS subgroup (InterPro:IPR013999), Helicase/SANT-associated, DNA binding (InterPro:IPR014012), MYB-like (InterPro:IPR017877); BEST Arabidopsis italiana protein match is: Helicase/SANT-associated, DNA binding protein (TAIR:AT3G24880.1); Has 9946 Blast hits to 7198 proteins in 452 species: Archae - 0; Bacteria - 514; Metazoa - 4085; Fungi - 1476; Plants - 1172; Viruses - 23; Other Eukaryotes - 2576 (source: NCBI BLINK).	MIRGFTSGTFLVWVYVDSMGGVVDGGVIGIKNK PSPRAAIEKAGADLRVEYDVREERRELEFELEK GGNPLDFKFGHAASVSVQSTSHTYQHDPQILTS EAKGSFALTASPHGDSVSESSRGLGVTLCEPNS ADNFDGENEILESERKSKHPTRGTTTPEHSSQL DGSQWKESEDSPIHFHPKGGQAYRRRNRFRTRN DGPRSSDMSARGQNSSLPGRHALAKPVASE YGFLESEGTGHLHTSCSRASEISEGKLTISTK ENFREELQNSSYKADEIPAAATNVEPDVAGKETY VPVDAGGPPCSGTEKEDIVSASRLDESCKFNG DGKPTNDIISNDAWTKGLDSESSCTRTSRV DGNMIEID.YTNLKNVDSNGVPKEKLPGLGKCI NVVDEMLSEKHVTNPLDVSCKEKEDQHSISHGN GFADKQEVDRDRYPSTDKPDPFINREELLALDH AETKPKDKHGHDSVQKVESVCPVRLSSVEPSS HEVNRNDLSAPPPASNTOTSTRENORKQLDKAHE DRVLEEARIIEAKRKRIAEELSGLTPREHRRKSO WDFVLEEMAWLANDFSQERYVWKLTAAGCICHF AAFASRLRLEKQNKSWKQKSVSLTLAKAVMEFW HSAGLRVENVT-HNASSGADKTGQVKDVAAGFE KDDVSEKKLQSTGCKRDLAVYGYAARFLQYNRS SPPFVKVGGSTPDVACDIDLIGVSWQDQLTEEN LFY_VPPGAIGYRKSIECCLSDFERVSNIIQEEV DTSLYDAAATEIGYQENDEYEEDEGENGYDLPFDS GRSSKKRKNMKSVEYGGDFGVRSLLDSRNGT POSALVGRKRPNANNHGVGSIPTKIRFTASRPRVVG PFGAGSAGVTLAPSRADASSGDTNSFQDDIVPR GSEVDSGMGFENQSSFDSPSEISAKPKKKKIKIHP GAGYDPNWWQVDSLSHNECKDHARKRYEYQLE PNGNSGMYAGHTKPKTKFKHPLDNPFDSTLPLMA GSVASPVASQMSNMSSQSKIILIGGRDRSRKAK GLKVPAACTGSGSPWTSFEDOALIVLHDMGPN

10

20

30

40

【 1 3 0 5】

【表 1 4 - 3 0 6】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
769	sp_088380_sicj	-0.883146	#N/A	<p>Pentatricopeptide repeat (PPR-like) superfamily protein; CONTAINS InterPro DOMAIN/s: Pentatricopeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: Tetratricopeptide repeat (TPR)-like superfamily protein (TAIR:AT3G06920.1); Has 55395 Blast hits to 14888 proteins in 319 species: Archae - 4; Bacteria - 71; Metazoa - 778; Fungi - 987; Plants - 51638; Viruses - 0; Other Eukaryotes - 1917 (source: NCBI BLink).</p>	<p>MFLVNPVSDSRISLPSRKSSARMLDERFIRLIKIF KWGPDAEKALEVLKLVKVDHRLVCEVLKIDVDINV KIQFFKWAGKRRNFHEDTTSYMSLIHCLLEAGSA GELWKTQDFVRRSSCHFSPSEFSEVLRVLGHAK MLSKAISIFYHIKRYKCKPTSNYNSLILMLMQGG HYEKVHELYNEMCNDSDSCFPDVTYTGALIQAFG KLGRRDSSAVRILFDEMKENGLVFNNAKIYTTVLAIF KLAADKALDLVKEMKDKGCSLTVYTTTELKGL GKLGVRDEAYDTFLSMLKEGCKPDVVLNNLNLV GKSGLLSEALKVFEDMKSLQCIPTWYTYVIKCL FEANAPAAQASALFEKMKAIQIAPSSYYSIMIDG FCQKNRYEKALMLLEEMDEHGFPCCPAAYCSLIN SLGKARFYDAAKELFQELKETSSTGTSARVYAVMI KQFGKSGC.DEAVDVLNEMKMLKGCSPNVYAYNA LMSALAFAGMINEAQMILLRTMEENGCAFDINSQ NIILNGLARTGGPKQAMEMFEKMKDLSIKPDAVS YNTMLSCLSRAGMFEESVKLMKEMSSNGFKYDQ ITYSSLLEAVGKVDVDDPIPPPTFSQ</p>
770				<p>kinase associated protein phosphatase (KAPP); CONTAINS InterPro DOMAIN/s: Protein phosphatase 2C, manganese/magnesium aspartate binding site (InterPro:IPR000222), SMAD/FHA domain (InterPro:IPR008984), Protein phosphatase 2C-related (InterPro:IPR001932), Kinase associated protein phosphatase (InterPro:IPR014045), protein phosphatase 2C, N-terminal (InterPro:IPR015655), Protein phosphatase 2C (InterPro:IPR015655), Forkhead-associated (FHA) domain (InterPro:IPR000253); BEST Arabidopsis thaliana protein match is: HOPW1-1-interacting 2 (TAIR:AT4G31750.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).</p>	<p>VDEERPLVQEDSDQLLSORNEPLKFFHTSECSHL QTEGLFNLPGSHSLTYKQRLPPAPHPSTLGGSLV LDIGNQAEEDLLVGQTLRRPLLTHFIEEOKCTTKD GLNHDSRLRPNVWFQISIKDVLDDQGSLSLEVI TGPACGLQSSVWSTDTSKLPLTLGRVPPSEVVL KDSEVSGKHAMIKWNSNKLKWLIDMGSINGTL LNTKAVHHPDFGSRWGEPEKELVSGDVTLTGT SQISVQITPLAQSEVPVGVGIASDPMSTRGGKR LPMEDVCYYQWPLFGAEQFGLFGICDGHGGAA AAESVKLIPETVATILDSLRRREKRVFSDASDV LREAIQSQTEACLNHYYEGCTATVLLVWGDGYDSF FAGCAVNGDSACIMNIDGRQIKMTEDEHRTSYSE RQRINKTGNPLKDDGTRLCGLNLGRMLGDKFLK EQDARFSEPEYVQVPIGQASEASVLAASDGLW DVISAKKNDQFVLOAKERYNSNDRKACAOAKIENFL LNEARKLRTKDNTSVVYVDFDLSALRSLCKYEP</p>
770	sp_088430_fcna	#N/A	#N/A		

10

20

30

40

【 1 3 0 6】

【表 1 4 - 3 0 7】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたペアとの相関係数 sp_107670_ptqx (SOAP2-like)	sp_107660_kiqg (SOAP2)	注釈	aa seq MGKKQKQKQVTLPPALPPDVPDEDIVPSDEDV EFVDKNPAYASFVSRDLDTOSITKHVNRVANAKED DLEALYKRLRRKTLNEEKKGIEVDRVDAPIKT LDGQLYKTKVEKAKPKPESDKDEDEDEDKGF KLTAEERFAKLKKKKKDEKQAKPEPVKTEEPKD POSEVLAEVREDLTAEEVAETKQHKIAELGTALL EDPEANIKSLKELLSFCNDDDKAIAKAAALLSLVV FRDIIPGYRIRLPTEKEQEMVYSKAVKKLRYFETT LLSYKAYLQRLISLQPLPEYNRVAVRCLTLDDA HPHFNFRENIPAVKSLASADDVVRKLCSTTIQS LFKNEGKHGGEATIESYLIADLVKIRNQCQLHPES IEVFLALSFDEDLARSKKAEDDEKRKNKKYKKGK KSEEQVQTAVSDVKKRSKELLSKTRDEVNADYK AASLALDEKERRRRLQSRRLTSVFNTRYRILKHTIQ PDATRFKANATASGASLGGHPLLSSCLKGLGMF SHLIDIFMADLMKHLKILASGTLNLSNDSLSLTVS ERIDCCIVAFKVMRKNLDALNVDLQDFIQLNLI EYRTERDQGEILAEALKIMLCCDRDQHDQRAAAF IKRLASSLFCGSAESMAALVAKNLLQKNFKCR YMLENDVGGSGVSGVAKYQYASDPHLSGALS SVLWELNLSKHYPHTVSAMASSISIKNTSSNQV YMPYVSPQQAYNELLEKEQFSIKSDIKLSNKRK RGINPNLLGEDESLEVEDEALRKLSEHYVTVH DIVENEKLRRELGLTTLSQLOVHEEYKVKKIKQEG SKQKRRKQ MERKOGFFSALKEEVWRLSPGKSRARSARSSG SAMTGLLRRKGRHVAIPEVLIARSGSLRPPPEML SPLRGPDGCGAREERDEGGWAHWMKRNPSPIS GSSSSAAACKRSDRLRLGLVGLAPLAPVHVA SEPLPHLSIKDTPLETSSAQYLOQYTAASGGGLK QNSIHNAVAMGKMLASDIETPNKVIKNRSSK SAESGGFVWQMNPDMMWYVEALGGSKVHAG CNGKLVWRHTFPWGAHAAGPVRPLRRLALQGL DPRTAAMFANARCIGEKINGEDCFILKLCADPK TLKARSEGPAPAEIRHVLFYGFYSOKTGLLHIEDSH LTRIQSNGGDVAWYETTINSCLDDYRPRVEGIMIA HSGRSVTLFRFGETAMSHHTKTRMEEAWTIEEV AFNVPLGSLDCFIPPAEIRYSSVSGDSCLELPHGDRL KTMSATAAAAYRAKVASMDRSYETTNTSNMTWQW DNMHV
771	sp_088790_dipa #N/A	#N/A	sp_088722_d #N/A	binding; FUNCTIONS IN: binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant_structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: CCAAT-binding factor (InterPro:IPR005612), Armadillo-type fold (InterPro:IPR016024), Nucleolar complex-associated (InterPro:IPR011501); Has 3184 Blast hits to 2630 proteins in 361 species: Archae - 21; Bacteria - 280; Metazoa - 1055; Fungi - 428; Plants - 179; Viruses - 26; Other Eukaryotes - 1195 (source: NCBI BLink).	
772	sp_089300_mepf #N/A	sp_089303_mepf #N/A	Protein of unknown function (DUF620); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: 21 plant_structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF620 (InterPro:IPR006873); BEST Arabidopsis thaliana protein match is: Protein of unknown function (DUF620) (TAIR:AT1G49840.1); Has 219 Blast hits to 218 proteins in 2 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 219; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLink).		

10

20

30

40

【 1 3 0 7】

【表 1 4 - 3 0 8】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたベクトルとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa_sqg MEGRVFSGSDSGVYVNRGSSGGLIKRKRKDG SGYGSKGRKLPESKEKRRSLNASEVSSDEE LLEFYPRRGVSGMNVGGASFGYESNDFEDVS FGRNGLMREAGPKRSKLDVYDFDEYDLDSDAEM MRMRHYDQGGEGEERRFRAMPGMNNGSGLD FESGSPKDIYTRRNSFHGGSSKKINGGDLYS GKRRFEMSSELVSMGKELGGSDEAIRLQGN GVLLRVYKHKEISEQFSKYYVHVGNNAVLR QEFASKNIVAFSPYSEKNVSKPIAVDRDKTQ RKRRKPPESSEEDSANDSDSPKLTQDCVQV HRSMKRGRSGGSTESPOSAQPIKKEGKVRGS GTEKQLLREKIRNMLLTAGWTIDYRPRRGRDYS AVYVPTGTAYWIIKAYDALQKLEEEGNSIKV GDDPSPFAPISENVLKLTQTRKRMKELKRRK REEAGGRKKEEFTSGSFEEDSDTDSGGSGRQ DDKLSFMKRSKSLKGSKESCRDRSNSKIM QERGESSSMSNSKFIHARKSKIGRCTLVRNS GKGLSSESDGFIPYTGKRNLLSWLIDSGTVNLSE KVHYMNRCTRVMLEGMWITREGHCGCCSKLTV SKFEIHAGSKLRQPYQNIYSSGASLMQCQVDA WNRKDSLNRNGYNKININGDDPNDDTCAICCDG GNLICDSCPSTFHQSCLGIEKLPDGDWHPHC SCKFCGLADMSVTGENNKRODPVLTCSFCEGKY HCSCEVGMGTGLINLNSPPFCGKTCQELSEHFQ KLLGVKHEMEAGLFWSLIHRDTLSDTSVRRGYP QRVEWNSKLAVALNMDECFMPIADRRSGINMI QSVLYNSGSIARLDYSGFYTAILEKGDDEIAAAGI RIRSTOFAEMPFIGTRHIYRQGMCRRLFAIEAA LPSLKYKELVPAISELMDTWTTFIFGSSLEELD QDMKLGVLVFPQDMLQKLLNQDSINRDLV TGADSFETDQEAASVMLGLRKKSDIDVPPGNRSV APEPEENEIMKGPVTEVCSODTYCTDDNSLHSA MLSLFRFLASELYMVCSSALQFKIGLFGALAS SASLKLRPKLTC
773	sp_089350_jiky	-0.313375	-0.335324	Acyl-CoA N-acyltransferase with RING/FYVE/PHD-type zinc finger protein; FUNCTIONS IN: DNA binding, zinc ion binding; INVOLVED IN: regulation of transcription, DNA-dependent; LOCATED IN: nucleus; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN(s): Zinc finger, PHD-type, conserved site (InterPro:IPR019786); Zinc finger, PHD-type (InterPro:IPR001965); Zinc finger, FYVE/PHD-type (InterPro:IPR011011); Acyl-CoA N-acyltransferase (InterPro:IPR016181); Zinc finger, PHD-finger (InterPro:IPR019787); BEST Arabidopsis thaliana protein match is: Acyl-CoA N-acyltransferase with RING/FYVE/PHD-type zinc finger protein (TAIR:AT1G05380.2); Has 1402 Blast hits to 1245 proteins in 121 species: Archae - 0; Bacteria - 0; Metazoa - 811; Fungi - 61; Plants - 450; Viruses - 0; Other Eukaryotes - 80 (source: NCBI/Link).	
774	sp_089480_dnr	#N/A	#N/A	uncharacterized protein	

10

20

30

40

【表 1 4 - 3 0 9】

	A	B	C	D	E	F
1						
2	共森 現遺 佐子	分析に用いたペイトとの相関係数 sp_107620_kqph (SOAP1) sp_107660_king sp_107670_ptqx (SOAP2-like)			注釈 ATATADIFSTIS THALIANA CATION/H+ EXCHANGER 18 (ATCHX18); FUNCTIONS IN: monovalent cation:hydrogen antiporter activity, sodium:hydrogen antiporter activity; INVOLVED IN: cation transport; LOCATED IN: integral to membrane; EXPRESSED IN: 10 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: Cation/H+ exchanger (InterPro:IPR006153); BEST Arabidopsis thaliana protein match is: cation/H+ exchanger 17 (TAIR:AT4G23700.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	
775	sp_090650_yxdc	#N/A	-0.876727	-0.891388	alpha/beta-Hydrolases superfamily protein; FUNCTIONS IN: lipase activity; INVOLVED IN: lipid metabolic process; LOCATED IN: cellular component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: AB-hydrolase-associated lipase region (InterPro:IPR006993); BEST Arabidopsis thaliana protein match is: alpha/beta-Hydrolases superfamily protein (TAIR:AT1G18460.1); Has 1707 Blast hits to 1690 proteins in 186 species: Archae - 0; Bacteria - 0; Metazoa - 1166; Fungi - 276; Plants - 162; Viruses - 0; Other Eukaryotes - 103 (source: NCBI BLINK).	MGRAEIVKMLDTAKDERTRKEALKALSAPSVSD MLSEKRRITGNKAMSIAGVICLIPFAMGTGSSVFL RATISKVYSGQAFVGMCLITVPCRIYLFEDVLSI ALTDOFSDPERGKSELGRSQ
776	sp_090980_fmms	-0.887258	#N/A	#N/A	serine-rich protein-related; BEST Arabidopsis thaliana protein match is: serine-rich protein-related (TAIR:AT5G11090.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MMQRIVDNAIAITKESVKTVYESLNVRLINSISA ILLAILPGKASILEGLHGWELEPFRGPRLPWMM GSGVSNQLIHLSVSDASSVEYSSGDEDDG YSSPASPQSSRRSVWASSVNEYDNHRTLIGSQ HAQCPSPVRSRKLPSLREHFVHRATDRRRGVI EDIHLEATEIVIEAVFDVAKHTAHLFLLSPEAFRLS GLIFPQGNVHVHHEAENVPTAVLGDHEPAPT ERKTFPHPLNTDARTCDQVITEFGYPYEAHVVTE DGYILLERIPRRDSRKAVYLQHGVLDSMGMWVS NGVVGSPFAAFDQGDYDFLGNFRGLVSRHVD RNISSREYWRYSINEHGTDIPAMLEKIHVKTSE LKLRYPDYDEDNSDEVPYVCAICHSLGGAGILM YVITRLEQKPHRLSRLLSPAGFHEDSNLAFTV VEHIFLGLAPLVKIVPAFYIPTFRFRLNKLARDF HNYPAVGLVQTLMSYVVGSDSNWVIGVGLPH YNMNDMPGVSFFVALHMAQMKRTGKFRMFDYG NPDANFEVYGSPEPVDLGEFCFDLIDIPVDLVAGR KDKVIRPAMVRYKYNLMTAEAGVEASNEYDYAHL DFTFSHRELLAYVLTRLQLVKSTPKQRQLDSVVE RSKEDLDVQSD
777	sp_091100_firh	-0.882148	#N/A	#N/A	uncharacterized protein	MAATSSSRHQHSFSPSGRFFCASSISSSSSFTSS SPTFTSQSTDLFHRSSPTIRVSIYGCSSPQTITSS SVRFSIDNRPTSPNRSFAAVNRNRNPFPAASGKK KTCMCSPTTHPGSFRCSYHKLSLEAKKQGNMVSYP PNLNMRRSAMTNSLVRIGTVEGGDSVKRALA ALIRPCAHQQRRRSDFQRPFRSRLSIMSKAGDSD P
778	sp_091370_hwci	#N/A	#N/A	-0.830365		IMDQETLSCSSAYRYGDCDGNIEPQVNSSVQPYYL KKLDEEQSSLRRTPDFNDREPLAMEQTLSAKK KFPACINEGYGGEGVKLEDLRKADDEVVVSDFSP ENDVQ

【 1 3 0 9】

10

20

30

40

【表 1 4 - 3 1 0】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたバइटとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
779	sp_093490_xtok	-0.88373	-0.88387	uncharacterized protein	MTMQRGEWCFVAVAGQGEDCGRAGLQDGNIRILLKVDSIMQGRSYPYHDSACTCKGMPGFHKP PWLIFITMPPHYIHSKIEVQDLFLGLCLDSRT NNHSPMSQLQTRFOARQLQRGNPFSSTAAGHA DGNPQLPSL
780	sp_093850_hswl	-0.84243	#N/A	bromo-adjacent homology (BAH) domain-containing protein; FUNCTIONS IN: DNA binding, nucleic acid binding; INVOLVED IN: biological process unknown; LOCATED IN: cellular component unknown; EXPRESSED IN: 9 plant structures; EXPRESSED DURING: 4 anthesis, C globular stage, petal differentiation and expansion stage, E expanded cotyledon stage, D bilateral stage; CONTAINS InterPro DOMAINs: RNA recognition motif, RNP-1 (InterPro:IPR00504), Bromo adjacent homology (BAH) domain (InterPro:IPR01025); BEST Arabidopsis thaliana protein match is: nucleic acid binding (TAIR:AT3G15605.4); Has 602 Blast hits to 478 proteins in 106 species: Archae - 0; Bacteria - 17; Metazoa - 295; Fungi - 38; Plants - 91; Viruses - 4; Other Eukaryotes - 157 (source: NCBI BLINK).	MSKAEDNERESIDFSWGTKRGRGGQNKGVQF YESFTYDGEVEFLYDSVYLSKEGEAKPYVGKLIK WEQPNARKVKLLWFFYPREISNYLKGENVLENE LF_LAKGDGKGLANVNP_LESIGGKCNVVCISKDER NRQP_SVEELKMAHYVYRTFDVQHHTVTDVLGE KIAMTEVNLLLNKSDFLS.DLPWVDSANAVAGLE NPLISHTYTLVEDGTMECDATGGELLEKVNKGD ILLDKSKYHPGGHIDAAGLDNTGLVGRERKLVN KDFGGVDKVVYDDSLGSLLSNDDKSGKGSDA AVPDDLGLASVPATEVALASGSKRRILFSPESKEN DLDEKYVLPVTDAAEKISDHDVKNHRAVGA VAVASNKGSQSVVVGTDNSNIGGRLTKFEFGISNG HPKLLKPEKSDRRLREMSVKQPKDKQESAATD IAANKASQPGNDVETDDKGNLSRKVASTSNGL PKKRSSTIDDKSNDKLRKTFVKEPKDKEEAY GCYKEVTERPPDRSKWFTASPWEGRMKDAHDQ GRLVLLHNLDPYTTSTEVENIWHGFGESCTGKV VQRTATSSRCGQAYVIFKTRVAKV/LKLDNE CFMLPNGRPLVGSRDVVFPGKQSQFPHLVP KPMQRESREAVSTSHSAQSNITLEYEMALEWR LLQERSELAWKMLYKQQGQEKIKLKATMKSE

10

20

30

40

【表 1 4 - 3 1 1】

1	A	B	C	D	E	F
1	共発現遺伝子	分析に用いたペアトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2					<p>RNA helicase. ATP-dependent. SK12/DOB1 protein; FUNCTIONS IN: hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides, helicase activity, ATP binding, ATP-dependent; helicase activity, nucleic acid binding; INVOLVED IN: N-terminal protein myristoylation; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: DNA/RNA helicase, DEAD/DEAH box type, N-terminal (InterPro:IPR011545), DSH, C-terminal (InterPro:IPR012961), DEAD-like helicase, N-terminal (InterPro:IPR014001), DNA/RNA helicase, C-terminal (InterPro:IPR001650), RNA helicase, ATP-dependent, SK12/DOB1 (InterPro:IPR016438), Helicase, superfamily 1/2, ATP-binding domain (InterPro:IPR014021); BEST Arabidopsis thaliana protein match is: RNA helicase, ATP-dependent. SK12/DOB1 protein (TAIR:AT2G06990.1); Has 9909 Blast hits to 8438 proteins in 1290 species: Archae - 828; Bacteria - 3072; Metazoa - 1170; Fungi - 1231; Plants - 476; Viruses - 32; Other Eukaryotes - 3100 (source: NCBI BLINK).</p>	<p>MGSVKRKSVEEENF-SMESNPPPPQKVMRENSNS GTVACFHDVSYSDSYPPRSSTNSSSSTSTPAKE FPFSLDPFOQAIAKLENGESVMVSAHTSAGKTV VASYAIAMSLRNKQRVYVTSPIKALSNQKYREFKE EFSVDGLMTGDVTDIPNASCLVMTTEWRSMQY RGSEITREVAWAIFDEVHYMRDKERGVVWEEESI VMAPKSRFVLSATVYNAKEFADWVAKVHQQP CHIVYTDYRPTLQHYFPPSGGKGLYLVDKAKF REDSFOKALNALAPGGGKRENGKWKQGWV GKLGEEEDFKLVKMLQRYQYDPVILFVSKRECE FLAMQAKMDLNGDDEKYNIETIFWSAMDMLSD DDKKLPQVNTLLPLLKRGIGVHSGLLPILKEVIEI LFQEGIKCLFATETFSIGLNMPAKTVVFTNVRKF DGDKFRWISSGGEYIQMGRAGRRGIDERICILM VDEKLEPSTAKMMLKGNADSLNSAFHLSYNMILN QLRFCEGDPENLLRNSFYQFOADRAIPDLERHA KQLKEEKDSIIIEEETLKNYSLEQYKSLKDDVR VIVYSPKYC:LPFLQGRILVCIETKTDEDPPFSI EDNVTWGVINFORIKSTSEDESCKKPEDASYTL DVLAKCMVSRREGGAKTKVPLKDPGVPVVS PLSQVNLSCIRVIMSKDLLPLEARENLTKKVSEVI SRCSREGVPLDPEADMVKVQSSYRKAARRIEAL ESLFDKHEVAKSPLIKEKQLTLHKKQELAAKIKS RQMRASALAFKDELKARKVLRLLGYTNDV VELKGIASEISADELTLTEILINGVFKDIKVEDM VALLSCLVWQEKLDQAQKPREELLLFAQLQDT ARTVAVQLECKVQIDVENFVSSFRPDIMEAVYA WAKGSKFYEMEITQVFEGLIFAIRRLEEVLQQLI LAAKSIGETELESKFEDAVSKIKRDIVFAASLYL</p>
781	sp_094100_cqxz	-0.313431	-0.357143	#N/A		
782	sp_094760_mpcz	#N/A	-0.655327	#N/A	<p>putative recombination initiation defects 3 (PRD3); Has 279 Blast hits to 273 proteins in 87 species: Archae - 2; Bacteria - 106; Metazoa - 105; Fungi - 5; Plants - 32; Viruses - 0; Other Eukaryotes - 29 (source: NCBI BLINK).</p>	

10

20

30

40

【 1 3 1 1】

【表 1 4 - 3 1 2】

1	A 共発現遺伝子	B 分析に用いたバइटとの相関係数 sp_107620_kqnh (SOAP1)	C sp_107660_kiqg (SOAP2)	D sp_107670_ptqx (SOAP2-like)	E 注釈	F aa seq
783	sp_095490_okqa	c_334223	#N/A	#N/A	<p>Protein prenilyltransferase superfamily protein;            FUNCTIONS IN: binding; INVOLVED IN:            biological_process unknown; LOCATED IN:            cellular_component unknown; EXPRESSED IN: 21            plant structures; EXPRESSED DURING: 10 growth            stages; CONTAINS InterPro DOMAIN/s:            Tetrairicopeptide TPR-1 (InterPro:IPR001440),            Tetrairicopeptide-like helical (InterPro:IPR011990),            Tetrairicopeptide repeat-containing            (InterPro:IPR013026), Tetrairicopeptide repeat            (InterPro:IPR019734); BEST Arabidopsis thaliana            protein match is: Protein prenilyltransferase            superfamily protein (TAIR:AT5G37130.1); Has 1807            Blast hits to 1807 proteins in 277 species: Archaee - 0;            Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385;            Viruses - 0; Other Eukaryotes - 339 (source: NCBI            BLINK).</p>	<p>MAEIVQHLLLEYELRLLRCTLQPPTNRRFLQP            HDTPTSHPPFHLIDNIVSLIESGRYVDSLSSAGVR            TVFQFSNSFESLDPSQFYAGLEKAVESFLVDDSA            EGHASFKFFLVVIAIAALAFVOCNFTGVPVELPS            VPLPWVKEEKWEAWARIQLINSDGSLAAKF            SNLQYLFAKMLLMKTRDLSFGKTYCEAGLGS            TWWFARVIFLQQRILDECCFSLNLLQVFMSES            NRFNVEKVTSYWGASLAEDEALTVSMLQLEA            GLVEHTFGRVDHARQYFEAAAFSVGLELSVTGA            LGVRTVHQYDVPVAGSVLVARKSSFPSSMSPLP            NASHVRDASHDPNHNQNHQPYSDTSDVLMAPK            FMDSSGNDLNLIEHGCAIDRTLTAIQQAVVLAHC            LLIEKRARFDEQQSWDMAPYIEAIDSQLSLYFTIK            CSONILRIKWESTRSFTKERALLMMNMKNVENLYE            PSPQVLERVHSCFSVLIAPAPILRKEYGELLSSGL            IAEALGIFEDLELWDLNIFCYRLLLEKKAANAELIKV            RLLETSPDRLLWCSLGGVMNDSCYEKALEVSN            DRCARAKRSLARSAYNRGDYKSKVLWDSAMA            LNALYPDGWFAFGAAALKARDVSKALDGFNRVAV            QLDPDNGEAWNIACIHKIRKKSQEAFAIAFKALK            SKHNSWQMWENYGVAAADVGNVFEALKAACQM            VLIDITGNKRYDGDLLERIVQEMEERFSSYSPITED            DKDCNSQTPFSGSELAESRDTKLLMELLGKVLQO            VVRSVSGGDSWGMVYARWHKMRGDLTMCSEAL            LKQVRSYQGSSEVWKNARFRKYANASLELCKY            MVISSSESLELHSDMLHRLNTIKQAISFTDTEE            FKALQVCLDDYQMLFKSSLVPTS</p>

10

20

30

40

【 1 3 1 2】

【表 1 4 - 3 1 3】

I	A	B	C	D	E	F
1	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2						<p>IMKRSRDDVFTGQVYKRRPQFVNSYRGSSEGGPQ  MVGGGVGAPTQRLTNDALAYLKAVKDMFQ  DKKEYDEFLEVMKDKAQRIDTTGVIARVKDLF  KGRNLILGFNTLPKGYEITLPPEDDNPAPKPV  EFEAISVFNKIKLRFHGDDEVYKSELILNMYRK  ENKTISEVYQEVAAALFRDHRDILLEFTHFLPDTSA  ATPGONGPSSGKNSILRDRNSGMPSTRLLHVDK  DHEFNLDGSDADNDKLGKAERDQRRRLDKERD  RKDGINRVESITDHGAENYGPRAVSGSYDDKS  GLKSMYSQELAFCEKVKKEGLDNLDDYQEFKCH  IYCREITRQELQALVRDILANYQDLVDGDFELAR  CERPDGFLSNVMSKSLWPEFLCSDGQSKSVKT  EDRERDRMVRDDRRDRRESRERDRTRDGVY  GNKEVANLRLSSNSNKDKYIGKPIQELDSLNCER  CTPSYRLLPKNYPIPI_LASHRTELGAGVLDHWVS  VTSGSEDSYFKHMRKNQYEESELFRCEDDRFELD  MLLESVNTIKRVEDLLDKINSMMIKTENPIRIEDH  FTPINLRICIERLYGDHG_LDMVDYLKMAILSLPVIL  TRLKQKQEEWARCRADFNKVAEYKSNYHKSL  DHRSFYFKQDQTKSLSTKALLAEVKEISEQKCKE  DDVLLAVAGNRRPIPHMEFEYPTDELHEDLYQL  MKYSCEGVECTAEQLEKVMKIWTAFLEPVLGVP  RAQGAEDSEDAIRAKNEASKTDGTPATGESDRSP  SGGAATRKHESSPEEQSSGNHAKDEGYPDAS  SSGLHQVRSQSNVIDEFGKQTTSNERLTASN  PSLPSGVEQFHGRNIENPSGHNTTQLRVGNAA  NSGIELRSNGGNIPSLAVVDSTRVAASNGVITAA  TKIIRYCEESNGHAKYEREELSPNGDFEEDN  FAVFGDASTEPANKPNDGASSAPHRSGPRGARR  ENDVDADDEGEESAHRSSDESENASENGVDYSH  SESGEGEDCSPENEEDGEHDHENDKVESEGEA  EGMADAQDGGISLPLSERYLQTVKPLAKYVPAS</p>
784	sp_095600_pxrk	-0.873303	#N/A	#N/A	<p>SIN3-like 4 (SNL4); INVOLVED IN: regulation of transcription; DNA-dependent; LOCATED IN: nucleus; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Histone acetylase interacting (InterPro:IPR013194), Paired amphipathic helix (InterPro:IPR003822); BEST Arabidopsis thaliana protein match is: SIN3-like 3 (TAIR:AT1G24190.2); Has 2571 Blast hits to 1108 proteins in 232 species: Archae - 4; Bacteria - 56; Metazoa - 1028; Fungi - 808; Plants - 415; Viruses - 14; Other Eukaryotes - 246 (source: NCBI BLINK).</p>	

10

20

30

40

【表 1 4 - 3 1 4】

A	B	C	D	E	F	
1						
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	
				<p>polypyrimidine tract-binding protein 2 (PTB2);            FUNCTIONS IN: RNA binding, nucleotide binding, nucleic acid binding; INVOLVED IN: pollen germination, regulation of RNA splicing, regulation of translation; LOCATED IN: nucleus, cytoplasmic mRNA processing body, cytoplasm; EXPRESSED IN: 31 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS Inter-Pro DOMAIN/s: RNA recognition motif, RNP-1 (InterPro:IPR000504), Nucleotide-binding, alpha-beta plat (InterPro:IPR012677); BEST Arabidopsis thaliana protein match is: polypyrimidine tract-binding protein 1 (TAIR:AT3G01150.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).</p>		
785	sp_096400_wtuh	#N/A	#N/A		<p>MASVSSQPFQFRYTPQPPSMVLHLRNLRPWECTKE            ELIELRNPFGKVVNTKCNVGNANKIQITFKFADLSL            SIAMISYAYASSVPAQIRG</p>	
786	sp_096760_dcxs	#N/A	#N/A	<p>DDT domain superfamily; CONTAINS InterPro DOMAIN/s: DDT domain (InterPro:IPR004022); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G25580.1); Has 20293 Blast hits to 13585 proteins in 823 species: Archae - 41; Bacteria - 1919; Metazoa - 8043; Fungi - 2539; Plants - 1031; Viruses - 220; Other Eukaryotes - 6500 (source: NCBI BLINK).</p>	<p>MQSYDLLPNPTLISPNIHLLDLPMEHVVSPIHR            SLGPEFTNGDAPKPESENNYPETGPEAGNKEGSN            PIEEESKELSGNDVGVSKSSKNGKKNENSSSN            NNVIVRERPTRACTARPAKYVDPVIERPRPP            KRERVEKVAEEVSEEEEDDLPEQCSKVWTSL            VNPPTQEEELMPRWELRSMWELASLNLFHIFRAV            LNINAEFTVEELESALLSPNSTLSDIHIPLKAIPPV            TRMACGRDTWYTVLCKRKLKDWVHWVAEGECP            VYTDHGAETEAYKNLDPKRVVILKALCDIRVEQ            EDIRSYDSSLKHGAQLSAFRKERIGDSSHGTSF            WYEDDEFVGHRLYREIRKVEKVEFKGKTGKS            NVLPCTTYEWETVATNLDFLDVSEKLFSTKNT            EASLKKLRNDMLPEIEKIHKRKEKMLKKQORQA            LLLDNFVMDGLAGGRSLRDRKPVTYTFDDFRSI            NEAIKTKRKQPSPEPGVKREVMKPEPSTNGRW            NGSSNSLLHVSFSPSPVDYDEMVDVHNSEPLD            RGNRRRQRORYSEKEFEVAVSDNDEEYFYSD            DDVGEAVYDEAYMQQRKQKQKRTSSSEGEGD            EEYQWSDENGVDVEEEEDDSVSNSESDSQCP            RGRFKPEGRTRSRFGTKIRSVDDIDIDISGLRRS            NFRATRNIDYRTYELSDSETETSTKGGKANLHAL            DNAGYSESQSDSEGNEEEEEQEKGVGSESS            APMAVENYDNEPEPEPLAKSESPKQNDPERRIE            ATRGYLDLNLNELAPGSGFDDGGALLKDDDDDKQ            GE</p>	

10

20

30

40

【 1 3 1 4】

【表 1 4 - 3 1 5】

A	B	C	D	E	F
1					
2	分析に用いたベクトルとの相関係数 sp_107620_kqnh sp_107660_kiqg sp_107670_ptqx (SOAP1) (SOAP2) (SOAP2-like)			注釈	aa seq METFSRKRTRAEVSTSPFSSSSVIGPTTFRVKKRTRLR SSSSSTTAATPATRSRTRSNISVSPATMDESSS RGGRRGRNHHGADKNDKGEKEHEVEKEREN ELGGLDYFDDDDDDNDNDSEGGVGLLHQN LTSASSALQGLLRKLGAGLDDLPSAMGSA SHQSGRLKKILSGLRADGEDGKQVEALTQCELL SIGTEDSLTFSVDSFVPLVGLLNYESNPDMILL AAPAITHLVDVLPSSCAAVVHYGAVNCFCARLLTI EYMDLAEQSLQALKISQEHPTACLRAALMAVL SYLDFSTGQVRVALATAINMCKKLSDDADFV MEAVPHLTNLQYHDAKVEHASVGLTRIVEAFA SSADKLDLCHNGLVTQVASLISTSNSSGGQASL STPTYTGLVRLLATCAGSTVGAKTLLLVGSGIL KDILAGSGLYGNMISVPALSRPPEQIFEVNLANE LLPLPQGTISLPASSNLFVRGPLEFKKSPMSSSG KODDNGNVQDVSAREKLNDOPELLOQFGIDL PVLQIHGSSVNGPVRHKCLSVIGKLMFFSNADMI SSUNNTNISFLSGVLAWKDPQVLVPAQIAEL MEKLPETFSKMFVREGVVHVDLSIASPNSVSA HTATPEKDTESLTPGSAARARLRSGNHQQDA EELKNSVSVGGSPPTSVEVPTASSNLRTTVSGC AKSFKDKYFSDPGTVDVIGISDDLHLKLNCAKL NAGVVEQKSKVKGSKTAGAVYAEFVTGKEESL VGVSEMLGELSKGDGVSSTFEFVGGVVAALLNY FSCGYFNKERILEVNLPKLROQALKRYKLFISIALP NEGNVPMITLVQKLNALSLERFPVVL.SHSAR SSGGARSLLSGLSALSOPFKLRCRAAGEKSLR DYSNTVVIDPLASLAAVEFDLWPRVORTESGOK ASTSVNGSEGGPLPTGAGASPAASTPATSTRR QSTRARSGVNIQDTSRKDASQEKNNSSKKGKGGK AALKPGEIEKGPQTRNAARRAYSDKDAQKSP TEESSEDELEMSPYDIDDALVIEDDDISDDEDD MLKYAGKYGLVLYLHRTGLAKKCAELRMGAK VRAFFGFVLGDL
787	sp_097120_jway	-0.304084	-0.247532	-0.31531	KAKTUS (KAK); FUNCTIONS IN: ubiquitin-protein ligase activity; INVOLVED IN: irichome branching, DNA endoreduplication; LOCATED IN: plasma membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN(s): Armadillo-like helical (InterPro:IPR011989), Armadillo (InterPro:IPR000225), Armadillo-type fold (InterPro:IPR016024), HECT (InterPro:IPR000569); BEST Arabidopsis thaliana protein match is: ubiquitin- protein ligase 4 (TAIR:AT5G02880.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI/Link).
788	sp_097150_jzay	-0.8552339	#N/A	#N/A	uncharacterized protein

10

20

30

40

【 1 3 1 5】

【表 1 4 - 3 1 6】

	A	B	C	D	E	F
1		分析に用いたペアとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	<p>注釈</p> <p>RNA-binding (TRF/WTED/TRNF motifs); family protein;            FUNCTIONS IN: RNA binding, nucleotide binding,            nucleic acid binding; INVOLVED IN: leaf senescence,            cell death, ethylene biosynthetic process, defense            response; LOCATED IN: nucleolus, nucleus;            EXPRESSED DURING: 24 plant stages; EXPRESSED            DURING: 13 growth stages; CONTAINS InterPro            DOMAIN/s: RNA recognition motif, RNP-1            (InterPro:IPR000504), Nucleotide-binding, alpha-beta            motif (InterPro:IPR012677); BEST Arabidopsis            thaliana protein match is: RNA-binding            (RRM/RBD/RNP motifs) family protein            (TAIR:AT2G41060.2); Has 35333 Blast hits to 34131            proteins in 2444 species: Archae - 798; Bacteria -            22429; Metazoa - 974; Fungi - 991; Plants - 531;            Viruses - 0; Other Eukaryotes - 9610 (source: NCBI            BLINK).</p>	aa seq
789	sp_097220_oxuk	-0.281724	#N/A	-0.31953		MDASKKRKADENGNSFLSDGTLTSPSDLSPPAQ AQTLTPDDARKLIEPFTIDQLRTLQDAVVRHPEV LVSTRALADSDVSRRLKLVRLGLPTTTDSLKSFL TSYGELDEAVVISDKTTGKSKGYGVTFKHVDGA LLALRQPSKKIDGRMTVTHLASAGRGNSEDOA TRKIYVGNVVPYEISAERLLDYFAAFGEVEEGLGF DKOGGKARGFAFVYKTEEGARMSLMEPMKID GHQVMCKLAEDGKKGKMGPNPVMPEGMGD RMSGSMNPQTGSYGNYPGPGYNSGFGMGMAA QGGGGPYGLHSQYPCSGEYASYRMPFPQSS MGFPPYQEGSGYGPFPSSGGGGGGYGGPPSSG GGGYMGQNPQFPSSQGGPPQGGPPGGMYQRMPP SY

10

20

30

40

【 1 3 1 6】

【表 1 4 - 3 1 7】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたペアとの相関係数 sp_107620_kqnh sp_107660_kiqg (SOAP1)	sp_107670_ptqx (SOAP2-like)	注釈		aa seq WALFRLEFYKFPDRLLERSERVVFUCCPSTDAL DEDEYTEYLGIVTQLQDYCPDASFMVFNFRREG DRRSHADIILSDYDMVTMVEYPRQYKCPPLLEM VHFHFRSSESWLSLEGQQNVLHLCERGGWPV LAFMLAGLLYRKNYSGEOKTLEMVNAQASKELL QLLSPLNPQSPSQRYLQYISRSLGSDWPPSDT PLILDCLIRVLPFLDEGKGRPVRVYVYGDPLTK SSRSTKQLFSSLSKSKQVPRYSETECLLVKLDIR CRVQGDVLECVHLCDLIREIMFRIMFHSSFR SNILSLSRDEVDKLDWAKVQISKDFKAEVLFSDV GAMPINTELESDDDDNETDSASPEEFEEAEIIFS NAVDAQGKTFDFGPFHDSYNEVINGEQKED REPPSFVDCGPDTRAHRRDKELVEIETVKDIGLD GWSKTLHRHVESDLNAIKDILMDANQKEDVKW AVDDIVLHEGGHKEVHYDLNEVKDIALDAYKE ADTMIVVSDTLIKQNIKTGFYDGNMVGELDRKQT VEDWEMGKSDSSNSPRKLNKNGDVTKQKQERV LQFSPIKRPPTFNQKASDSALGKHRGKQQAAG DKSGKQKPTHRWIPSNKGSYTNMHHVAYPTTR VNSAPPSLGGTKNQKSGSVGKLDKRAASTDF GFGATSIDRRPHKFDVNPVESAKFSWVETPDD GLNQEVSQKCFPPSPSSLVNPNSSQSLAESVE LEEEIKAPLTLPPQSPVHSSSDMPTVEVLPQOR PPSPNSINSCGNFKISPMPLQSPSLFSSDDKQS PPPPPPPPGRRGPPPPPPPPPPGRRGPPPPPP GGRRGPPPPPPGRRGAPPPPPPPPPGRRGAPPP PPPPGRRGPPPPPPPPPPGRRGPPPPPPPPGAP GFPAPRPPGAGPPPPPPPPGAKAAPGAPNAPPS LGRGRGLPRSAPGVATKTTLPYHWIKVPRLIQ GSLWEELQTOAGSQVAEFDMSIEIKLFSNQVSK PAGAGGKSGEKKKAVGSKSEVWHLIDTRRVNNT EIMLTQVKMPLPDMMAAVLAMDESILDADQVENL
790	sp_098670_qwid	-0.8553864	#N/A	#N/A	actin binding; FUNCTIONS IN: actin binding; INVOLVED IN: cellular component organization, actin cytoskeleton organization; EXPRESSED IN: shoot apex, stem; CONTAINS InterPro DOMAIN/s: Actin- binding FH2/DRF autoregulatory (InterPro:IPR003104), C2 calcium/lipid-binding domain, CalB (InterPro:IPR008973), Tensin phosphatase, C2 domain (InterPro:IPR014020), Actin- binding FH2 (InterPro:IPR015425); BEST Arabidopsis thaliana protein match is: Actin-binding FH2 protein (TAIR:AT5G07770.1).	
791	sp_098770_mxac	-0.8413651	-0.8152305	-0.859234	RAN GTPase 3 (RAN3); FUNCTIONS IN: protein binding, GTP binding, GTPase activity; INVOLVED IN: protein import into nucleus; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Ran GTPase (InterPro:IPR002041), Small GTP-binding protein (InterPro:IPR005225), Ras (InterPro:IPR013753); BEST Arabidopsis thaliana protein match is: RAS- related nuclear protein-1 (TAIR:AT5G20010.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archaea - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI Link).	MALPNQQTVDYPSFKLVLDGGGTGKTKTTEVKRHR LSGFEFEKKYEPTIGVEVHPLDFHTNCGQIRFYCW DTAGQEKFGGLRDGYIHHGOCALIMFDVITARLTY KNWPTWHRDLRCRCENIPVILCGNKVYDVKNRQV KAKQVTFHRKKNLQYYEISAKSNYNFEKPLYLA RKLAGDQNLHFVESPALAPPEVQIDLAEQQRHE QELAVAAAGPLPDDDDDDAFE

10

20

30

40

【 1 3 1 7】

【表 1 4 - 3 1 8】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2				Major facilitator superfamily protein; FUNCTIONS IN: transporter activity; INVOLVED IN: oligopeptide transport, pollen tube growth; LOCATED IN: membrane; EXPRESSED IN: petal, leaf whorl, sepal, male gametophyte, flower; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage, 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Oligopeptide transporter (InterPro:IPR00109), Major facilitator superfamily, general substrate transporter (InterPro:IPR016196); BEST Arabidopsis thaliana protein match is: nitrate transporter 1.7 (TAIR:AT1G69870.1); Has 3642 Blast hits to 3525 proteins in 525 species: Archae - 0; Bacteria - 558; Metazoa - 480; Fungi - 335; Plants - 2158; Viruses - 0; Other Eukaryotes - 111 (source: NCBI BLINK).	MGLTVSVGVPKLPPKCEPPAECLPSESQWLAI LYAGLTLVAIGGCVPCNIGFADQDFDVTTEKG RASARYFNWVWYFSFTFALLVALIGVWVWQVHIS WAIACVVATVCFASSTIIFLIGSPTFINKKPKQGSYI DMARVVAAFGRNVQNHQTYDPPCCISNSNDVI APPKLRTNRLLYLDKASLITQPEELNNSGLAINK WLSLVSQQVEQLKFLGILPVWLTVGFCFVSDIQ HNSYGILQAMQMNRRFTRKGFEPAAWIGFASM ASLSLWAYERICLPLAKKITKNDVRLSINTRIKI GIVMSVLSSTAADVVECKRRLAINKNSYSPLHV GYLAPQQVICGLTEAFAGVTVMEFYTTQMPESM RSLSGSIFFLNLAMASYISAIINIITKIQHTHQGS WLGNDLNIHRLDYLFAIVACLGVNYYFFITFFA SKFVFDINNIETEAASSSTEVHPRVE
792	sp_099050_hcsn	0.823335	#N/A	NADPH-dependent thioredoxin reductase A (NTRA); FUNCTIONS IN: thioredoxin-disulfide reductase activity; INVOLVED IN: in 6 processes; LOCATED IN: cytosol, mitochondrial matrix; EXPRESSED IN: male gametophyte, guard cell, cultured cell, pollen tube, leaf; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage, seedling growth, seed development; stages; CONTAINS InterPro DOMAIN/s: Pyridine nucleotide-disulphide oxidoreductase, class-II, active site (InterPro:IPR008255), FAD-dependent pyridine nucleotide-disulphide oxidoreductase (InterPro:IPR013027), Pyridine nucleotide-disulphide oxidoreductase class-II (InterPro:IPR000103), Pyridine nucleotide-disulphide oxidoreductase, NAD-binding region (InterPro:IPR001327), Thioredoxin reductase (InterPro:IPR005982); BEST Arabidopsis thaliana protein match is: NADPH-dependent thioredoxin reductase B (TAIR:AT4G35460.1); Has 20135 Blast hits to 20133 proteins in 2901 species: Archae - 588; Bacteria - 13597; Metazoa - 140; Fungi - 273; Plants - 163; Viruses - 0; Other Eukaryotes - 5374 (source: NCBI BLINK).	MSSSSPPNAQMSRCLKLRTLKRLNLFPLPLSL SSSSPAAATTTTSSAANSLSTVAAPTMEDF KQALRTLCLIGSGPAAHTAAIYASRAELKPIFLFEG WMANGIAPGGQLTTTDDVENFPFPEGIFGLDIT DKFRQOSLRFGTTEILTETVTSVDFSTTPFKLSAE SKTVLADSVIATGAVAKKLEFPGSGEEGGFWNK GISACAVCDGAAPIFRNKPLAVIGGGDSAMEEAN FLT KYGSKVYIHRGFRFRASKIMQKRAIENPKIEV LWNSVTVEAYGDAEKENRYLGLLKVQDVVSGEV KDLQVSLFFAIGHEPATKFLGGQLELDADGYVV TTPGTTKTSVKGVFAAGDVQDKKYRQAVTAAGT GCMALDAEHYLOEIGSQEGKSD
793	sp_099410_ujgo	-0.873537	#N/A		

10

20

30

40

【 1 3 1 8】

【表 1 4 - 3 1 9】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたペアとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
794	sp_100510_fgkr	#N/A	-0.873323	#N/A	co-factor for nitrate, reductase and xanthine dehydrogenase 7 (CNX7); CONTAINS InterPro DOMAIN/s: ThiamineS (InterPro:IPR003749), Molybdopterin converting factor, subunit 1 (InterPro:IPR010034), Molybdopterin synthase/thiamin biosynthesis sulphur carrier, beta-grasp (InterPro:IPR016155), Beta-grasp fold, ferredoxin-type (InterPro:IPR012675); Has 1000 Blast hits to 1000 proteins in 381 species: Archae - 21; Bacteria - 814; Metazoa - 52; Fungi - 2; Plants - 54; Viruses - 0; Other Eukaryotes - 57 (source: NCBI BLINK).	MEQKNQIESQGNNTHERNEKSSIQLKVLFFAFAR DLTGLADMPLEVTSGSTARDCLDKIVTQFFGLEEI RGCIMVLAALNEEYASESAIVQHRDELAIIPISGG
795	sp_100590_qqjs	-0.835385	-0.9466726	-0.9152887	Ribosomal protein L10 family protein; FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation, ribosome biogenesis; LOCATED IN: cytosolic ribosome, ribosome, chloroplast, chloroplast stroma; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Ribosomal protein L10 (InterPro:IPR001790); BEST Arabidopsis thaliana protein match is: Ribosomal protein L10 family protein (TAIR:AT3G12370.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MGATISTTTTSLSHHTLTLIRTQSHHHQLRHFAK YTPYKPLPKVHQKLRRAIKTKLSHLQDSYLSPF PCHGFSANKFKPLKPLPDPSSAAKLVFKNSLV SQAFAGTQFEALGSCMKGMNSLLFYKKEADAVGP VMGGVKSAVKEGKLEFNDFPGAVVNGQLYGHLD LEVFQNTPSKMEAHGMVLSLFGPSSTLMAVLE GFCADDSANEAAATDAPADGGIDAKDVSESPAQ
796	sp_102600_rozo	#N/A	-0.863854	#N/A	uncharacterized protein	MGTSDILPEKADTHGCLYCCSRDIHHPADICSEFR SGGEPHFDLCSKLEKFNKYLEKGLRKRKDKED LSLRFDPDLGLEVRGNNTCTDMARNASVALRLEN GEKGSNFEVKIGRFEVVLGFLQHMNFEAKN KDDPDAEYMTFFAEVIRRRYFFVDFCQCLGPSF LLPDEAELKGCYCKCIHHPPGPGYLIQRELD
797	sp_103080_jhua	-0.885894	-0.889073	-0.858349	unknown protein; Has 19 Blast hits to 19 proteins in 10 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 19; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).	MEVGSKRENEKQKTEKPPYVAVKDDTKPVLQD PLWSSDPTQAEQAVLRLPPFPYKSSNDQLN

10

20

30

40

【 1 3 1 9】

【表 1 4 - 3 2 0】

A	B	C	D	E	F	
1						
	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	
				FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: membrane; EXPRESSED IN: 14 plant structures; EXPRESSED DURING: 4 anthesis, C globular stage, F mature embryo stage, petal differentiation and expansion stage, E expanded cotyledon stage; CONTAINS InterPro DOMAIN/s: Aminotransferase-like, plant mobile domain (InterPro:IPR019557), Protein of unknown function DUF716 (InterPro:IPR006904); BEST Arabidopsis thaliana protein match is: Aminotransferase-like, plant mobile domain family protein (TAIR:AT1G51538.1); Has 16736 Blast hits to 9656 proteins in 576 species: Archae - 4; Bacteria - 1182; Metazoa - 7098; Fungi - 2631; Plants - 1178; Viruses - 174; Other Eukaryotes - 4469 (source: NCBI BLINK).		MGSFKGHVVPGLLLVGLWHMWCVYRYVRS PNTFRIRVWNPVPGFDGRIKYLELYVYVIGSFIDL CIELYVSPHLKYFVNGVNLNPSHMNMFHSGMLIM FFIFLGLVLLSVKTSILPLPDGALTLIGATAFSSEY LLFYHSTNHKGLGEGYYHYLLVLLIGCVLTSVVG TLVPSNFPLDLSNSIATLQGLWFYQTAFTLYGPM MYPHCQLKGDIVCHSADSIQIGELLANFQLFS MNVVGVFAVISVYFSFATSRFGHSSQERLQEADG DL MPPELVEVAFRSVAERFCIGKILRSTIADDSKVIN GVRSRHFEAALLNKTVAAHRKGNKMNWLQLDCP PFPTFOFGMAGAVYIKGVAVTKYKRSVSDNDE WPSKYKFFIENDGLEMSFTDKRRFATVRLDD PASAPPISALGPDALLEPMTGEFENKLSLKKIAI KPLLLDQSFISGIGNWVADEVLYOARIHPSQNAS SLSKDESDRLLKSITEVLEKAVEVGDADSSQYPSG WIFHSREKPKGAFVDGKKIDFTLISGVQVATEAP AKPRKRASKKDTDDHNDGDDTNGDVRREEETE DTKSKQTKTKSKSGKQPARRKSTKSNDDNDDDD DVDGGGGGGEDDGGDDKTTGTTRGGKVIK RKLIGTNSODANKSKKTRGGGKDDTAEPKSRK SNDSADSDDDGNENGDQKAGAKKPKQKTRKKA K
798	sp_103380_mzes -0.846754	#N/A	#N/A	MUTM homolog-1 (MMH-1); CONTAINS InterPro DOMAIN/s: Ribosomal protein S13-like, H2TH (InterPro:IPR010979), DNA glycosylase/AP lyase, H2TH DNA-binding (InterPro:IPR015886), DNA glycosylase/AP lyase, catalytic domain (InterPro:IPR012319), DNA glycosylase/AP lyase (InterPro:IPR000191); Has 11102 Blast hits to 10156 proteins in 2240 species: Archae - 10; Bacteria - 5065; Metazoa - 1812; Fungi - 570; Plants - 321; Viruses - 42; Other Eukaryotes - 3282 (source: NCBI BLINK).		
799	sp_103510_diku #N/A	-0.837155	-0.850333	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein; FUNCTIONS IN: oxidoreductase activity; INVOLVED IN: oxidation reduction; LOCATED IN: nucleus; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Taurine catabolism dioxygenase TauD/TdA (InterPro:IPR003819); Has 1029 Blast hits to 1021 proteins in 229 species: Archae - 0; Bacteria - 729; Metazoa - 46; Fungi - 17; Plants - 110; Viruses - 0; Other Eukaryotes - 127 (source: NCBI BLINK).		
800	sp_103740_qrtz #N/A	-0.827031	#N/A			

10

20

30

40

【 1 3 2 0】

【表 1 4 - 3 2 1】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
801	sp_104730_tqk	#N/A	#N/A	-0.83333339	<p>glutamate dehydrogenase 3 (GDH3); FUNCTIONS IN: oxidoreductase activity, acting on the CH-NH2 group of donors, NAD or NADP as acceptor, oxidoreductase activity, binding, catalytic activity; INVOLVED IN: regulation of nitrogen compound metabolic process; LOCATED IN: mitochondrion; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAIN/s: Glutamate/phenylalanine/leucine/valine dehydrogenase (InterPro:IPR006095), Glutamate/phenylalanine/leucine/valine dehydrogenase, C-terminal (InterPro:IPR006096), Glutamate dehydrogenase (InterPro:IPR014362), NAD(P)-binding domain (InterPro:IPR016040), Glutamate/phenylalanine/leucine/valine dehydrogenase, dimerisation domain (InterPro:IPR006097); BEST Arabidopsis thaliana protein match is: glutamate dehydrogenase 1 (TAIR:AT5G18170.1); Has 8937 Blast hits to 8931 proteins in 2306 species: Archae - 305; Bacteria - 5351; Metazoa - 356; Fungi - 208; Plants - 399; Viruses - 0; Other Eukaryotes - 2318 (source: NCBI BLINK).</p> <p>FAINB2; CONTAINS InterPro DOMAIN/s: Pyruvate/Phosphoenolpyruvate kinase, catalytic core (InterPro:IPR015813), Ketopantoate hydroxyme.thyltransferase (InterPro:IPR003700); BEST Arabidopsis thaliana protein match is: ketopantoate hydroxymethyltransferase 1 (TAIR:AT2G46110.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).</p>	<p>MNLAATNRNFKLASRILGLDAKLEKSLIIPFREIK VECTIPKDDGTLASIFGRVQHDNSRRGPMKGGIR YHFEVDPDEVNALAQLMTWKTAVANIPYGGAKG GIGCNPSDLSISELERLTRVFTQRIHDLIGIHTDVP APDMGTNSQTMAWILDEYSKFHGHSPAVVTGKP IDLGGSLGRDAATGRGVLFAFEALLNESGKTISS QRFVIOGFGNVGSWAARLISELGGKVVAVSDISS AIKKNGLDIDSLKHVQENRNGYKGFHGSDAIDS DSILVEDCDVLIIPALGGVINRENANEIKAKYVEA ANHPTDEADEILKKNGVILPDIYANISGGVTVSY FEWVQIQGFMWEEERVNNELKTYMTKGLKNV KEMGKTHNCDLRMGAFLLGLNRVARATVLRGW EA</p>
802	sp_104780_rkym	#N/A	#N/A	-0.87777776		<p>MSTHVISSILLHISLPSLRSSFNPSGAEFGTVKV VETALALQEAGSFAFVLECYVAPVAAAATFVLGIP TIGGAGPFCSSGQADSSVRIKKEGMDAIKLEGA SPSRITAAKAVARIATVTPQAISVLGGFSRPQGR NISTAVKVAELVPPKAAAMCSEKYNQTKVGAER GFRVTNYMTLVPVKVFEFQGGGGQTHTIAAEVNGA DVPVAAA</p>

10

20

30

40

【 1 3 2 1】

【表 1 4 - 3 2 2】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
803	sp_105380_osku	#N/A	-0.653333	#N/A	phosphoenolpyruvate carboxylase kinase 1 (PCK1); FUNCTIONS IN: protein serine/threonine kinase activity, kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: cellular component unknown; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase domain (InterPro:IPR002290), Calcium-dependent protein kinase (InterPro:IPR020642), Calcium/calmodulin-dependent protein kinase-like (InterPro:IPR020636), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR01009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: phosphoenolpyruvate carboxylase kinase 2 (TAIR:AT3G04530.1); Has 121330 Blast hits to 119354 proteins in 3292 species: Archae - 175; Bacteria - 14755; Metazoa - 43495; Fungi - 13131; Plants - 28233; Viruses - 522; Other Eukaryotes - 21019 (source: NCBI BLINK).	MYDSFKNEYQLCEEIGRGRFGIVYRCFSPSSLAS FAVKSIDKRHLDDDDLRQCLDKPEKILQILSPHP NILQHNIFDSDSRLIVTDLQKESTLYDRIVSSGPF TEPDACSITQLAEALHCHHRNYCHRIDKPDNILF DTRNRVKLDFGSAEFFCGGGEREKMSGVVGT PYYVAPEVLAGREYNEKVDVWSAGVILYIMLGGV PPFYGDVSDIFAAVLRGNLRFPTKIFRSVSPEA KDLMRKMLCKDVSRRRFSAEQVLRHPWPVVTSGGL SNTMSDMS
804	sp_105590_iumr	-0.362633	-0.375151	-0.353367	pre-mRNA-processing protein 40C (PRP40C); CONTAINS InterPro DOMAIN/s: FF domain (InterPro:IPR002713), WW/Rsp5WWP (InterPro:IPR001202); BEST Arabidopsis thaliana protein match is: pre-mRNA-processing protein 40A (TAIR:AT1G44910.2); Has 7513 Blast hits to 5748 proteins in 436 species: Archae - 33; Bacteria - 336; Metazoa - 3654; Fungi - 967; Plants - 481; Viruses - 7; Other Eukaryotes - 2035 (source: NCBI BLink).	MFSSRSRQLNRLVIRIDSKVFSCLCTYLLOFGD MLKERGAVPFSKWDKELPKIVFDPFRKEIPTYAE RRSLEFHYVTRAEAEERKEKRAAQAAVEGFKQ LLEAKEDIDHTTYHTFRKRWGHQDRFEALERK DREVFNLNERSLKRVAEEKIRAEERTVYVSNFKS MLRERGDVTASSRWSRVKDLRDDSRVKAVKH EEREALFNAYISELKAADDEVEREAKAKREEQEK LKERERLKRKEREEOQMERVRSKVRKEAVA SYLALLVETIKSHQASWTESKPKLVKDPQGFASN PDLDESLERLFREHYKMLYERACACDFRALLVEK ITPEAATRETEDGKTVLNSWSTAKFKLSDPPRYT KMPRPKDREALWKRHAEELQRKQKGAVEQKGDQ YNESKNRNRANDSGRPLSVSRSGQDWR

10

20

30

40

【 1 3 2 2】

【表 1 4 - 3 2 3】

A	B	C	D	E	F
1	分析に用いたバイトとの相関係数	sp_107660_kiqg	sp_107670_ptqx	注釈	aa seq
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	(SOAP2-like)		
805	sp_105830_yqri	-0.927577	-0.934248	-0.930305	<p>unknown protein; Has 30 Blast hits to 30 proteins in 10 species: Archaea - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 30; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).</p> <p>De protein kinase like z (DBPKLZ); FUNCTIONS IN: kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: nucleolus, nucleus, cytoplasm; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase domain (InterPro:IPR002290), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: D6 protein kinase like 1 (TAIR:AT4G26610.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archaea - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).</p>
806	sp_105910_hknn	-0.855386	#N/A	-0.91186	<p>MKSNQFACRVVSPSGESSNCQVSGFESKCSIK YPKVDSRPMIQEAHWVTLQSVYRGGQRHRSVT EASECHSIALFRNRKVERTINPFDVERRKVSSTY RCSTGSDVNDVRIARKAQSNYGLLQFTHFRLL NHLGCGDVGKYLSEFVGTHTFAMKVMKKAAL ESRKKLRAQTEREILQSLDHPFLPTLYTHFETEN YSCLVMEFCGGGLHSLRQRQP-HKCFTEHAAR</p>

10

20

30

40

【 1 3 2 3】

【表 1 4 - 3 2 4】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
807	#N/A	0.826238	#N/A	uncharacterized protein	MALMREQLRARLASSKTSVPEKNEGSMKAKFDR MTLMRAQLRARLASSKTSVPKQKEGLMAQKQFDR MTLMRAQLRARLASGKMSFPKAPGQSKSPKAP RQRDVEMQDNVMEQDQGVHHTEDRFVFSQVSG SPVLPGYKRMRTDEEDESSTGPTTAFLLSDSRER QRRVRSVMNLCRDADDEDYLDSDMDPTQEVNEM HOAWAEFSIRATSMRRRYRYFRELODKVDELEA SAKKHPEVATVNAKITRLEGMYNLADSOAKADV KRIASLKNRNGKKIAALAIKVAKLKEELERAKE SASASRQTHVDEWQKSEDEGVTYLDVAQRSMDI GEKTHDRMREALKRCFFPEVDYINIVWDECEAIRL AEVDALAQEIYDLEHFRLY
808	sp_106560_tkfx	#N/A	#N/A	autophagy 9 (APG9); CONTAINS InterPro DOMAIN/s: Autophagy-related protein 9 (InterPro:IPR007241); Has 489 Blast hits to 477 proteins in 190 species: Archaea - 0; Bacteria - 0; Metazoa - 214; Fungi - 166; Plants - 51; Viruses - 0; Other Eukaryotes - 58 (source: NCBI/BLINK).	MSGWQGRNVLIFKWKPKGDPSPISIGLLNDNP PEIELSGYGRIPSPGSESPSELLNGESLKVEPIAD LDLFFERLYSYCDKGLWCIWIKWVILLSLAFTIC FSGFLLYVDWDLRNKAKGIDAVESGKTKPCDL AIEALHAHLPLTLPKAILGLYGIFSIYWIFCLRF FAQLKEILRIRHYNCLHVTDDDELKTMWPSPILE KVVRVQNSQKLCVYKDLAHEIVMLRMKENYLI GMINKGILAFPIKWIPIGAGPTVKSGPNRGRQRLI LTKLLEWTLNWSILOSMFDRKFCIRGEFVSNPAT LRKRLVVGLAMLLSPFLVIFMLVY_LFLRAEQF YNHPTTASSRRWSNLSKWMFFREFNEVEHFFRH RINSSVHASDYLRLQFPSPISVYAKFISFISGGFAA VLJIIIFLDESLEGHIFGRNLFWYAAVFGTITAIR AAVTDELLVLDPDGTMSLVQVHTHMPKRWRG KENTEFVRVEFETLFQYTMWMLLEEIASIFLTPYL LIFVVPKQVNDILOFITEFTVYIDGVGDVCSFSVFD FQKHGNGYASPYNSLPTRRSSQGMKESLSLF QTSYPSWQPDAGGMQLTLTRFRNEKLRGPIP GYAYSPRRWQQSPFRFRDHSRNSYFRRES YNGGAPGCHLGSLLWLGITDPKTHPYLLDWFVYS GARHIGNNTNDGFSVSLDIDNEQTNDFWMPSL NNEDNNQGEESWPNFLDRSKNHPEASTSS PVFLEESVLRHNDSSSTMLPSARSHWNNRRG GGTT'SADP'QQAASFELEPPEFFHRRGGPHQASFE PPIFRHSGGTQQASFELEPTEFQHRGGTQQTSL EPTFQHRGGRRGLQTSFLEPTEFQHRGGTQQT SFLEPTEFQHRGGPQQTSTFLEPPIFRHSFEPFNH HNADNYCESHSRSDIDDEHDEQPLDWGNYNQL SRTTDPDEFGGFEKLFHFDVYVYKPESEPPMNLDH DKYHVS

10

20

30

40

【 1 3 2 4】

【表 1 4 - 3 2 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたペイトとの相関係数 sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
809	sp_106720_dzxp #N/A	-0.874456	-0.916383	Protein kinase family protein; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase domain (InterPro:IPR032290), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: Protein kinase family protein (TAIR:AT3G13670.1); Has 17667 Blast hits to 17599 proteins in 1508 species: Archaea - 20; Bacteria - 6581; Metazoa - 4779; Fungi - 1455; Plants - 2098; Viruses - 217; Other Eukaryotes - 2517 (source: NCBI BLINK).	MPELRSGARRRGRASNTNPTNNNNQEEIGGOKAT VGRTRRRGGAVAGGRGNKGGSSRRGRNRNQ QPAAAVGGGEGEAGEADEKVVGVQKESFRKT ASLKAEAAAEEVGVKRKEIGEEMDEYDSSGNA AGRSGDKGLNDDDDSTPLPEKQVQVGGSPAYKV DRKLGKGGFGQVYVRRSNGGNTSERNGPGAYE VALKFEHFSSKGGNYGPPYEWQVYNALGGSHG VPRVHYKGRGGDYVMVMDILGPSLWVDVWNNN SHSMSIEWACIAEIAISILEKMHSRGYVHGDKVP ENFLGPPGTPPEKKLFLVDLGLATKWFRDSTTGL HVEYDQRPDVFRGTVRYASVHAHLGRTGSRRD DLESAYTLVFLRGRLPWQQYGGENKGLVCK KKMATSPETLCCFCQPFRRHFVDVYVNLKFDDE PNYAKYSLFDGIVGNPDIRPINTDGAOKLIFQV GHKFRGLTMEEDDDEQPKKIRMGMPATQWISV YNARRPMKGSFSPPPXVADIRLSQHDYKLNEDGL FISSVASCSSLWALMDAGTGFSAQVYELSPHFL HKEWIMEQWEKNFYISSIAGANGSSLLVMSKG TCYLQOQYKVSADAFKFKWINKWREGFYVTSMA TAGSRWAVMSRGAQYSDQAVELDFLYPSEGV HRRWDGGYRITSTAATFDSQALVLSIPRRKPNDE TCQETLRTSAFFSTHVKEKWKAKNLYIASICYGRTV S
810	sp_107050_d1ss #N/A	-0.845243	#N/A	SAND family protein; LOCATED IN: chloroplast; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Vacuolar fusion protein MON1 (InterPro:IPR004353); Has 753 Blast hits to 594 proteins in 226 species: Archaea - 6; Bacteria - 38; Metazoa - 289; Fungi - 213; Plants - 47; Viruses - 3; Other Eukaryotes - 157 (source: NCBI BLINK).	MADSTSSSDSAYDSDAADHNPSDTPINNHS AIDHLLTSLTLPLDEDLAPEKEREIETDOTPMP EIGVTEVEEAAADVLASEVVEEEREESNSVSATG DVGLEWREEVEAPASPGSSGYAGERGSSNAS SASGIEEIGDQVAERDGGVNGVADSQP7AWIP GKRHYNEDDTSILWRRRKHFFILSHSGKPIYSR YGDHKLAFSATLQAIISFVENGDRVRLVLRAG KHQVFLVKGPIYLVGISCTEPEYESLRGOLELLY GOMLILTKSVNRCEKNAKFMPTLLGGTDAVF SSUHSFWSNPATFLHAYTCLLAYPTROAAGAIL QDVADSGVLFALMCKHKVWSLVGAEKASLHPDD MLLSNFVMSSESFRTESEFSLPKYNSMAFL YAYVQYLDVNTYLMILLTRSDAFYHLKECRASIE AVFLKSNVLEVQRSMLDGGMHIEDLFPVDPSSR SGSASSHLGQHSPEFRPINTYTSLGGPGLWHPFI YRSYLDQYYSSEESAPINTPQQQKRLRYAYQKIY ASMHDKGFGPNKQFRRDDNYVLLCWVTQDFE LYAAFDPLADKAMAIAIKTCNRVQCQWVKDVENEIFL LGASFFSW

10

20

30

40

【 1 3 2 5】

【表 1 4 - 3 2 6】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 Acid phosphatase/vanadium-dependent haloperoxidase-related protein; CONTAINS InterPro DOMAIN/s: Acid phosphatase/vanadium-dependent haloperoxidase related (InterPro:IPR003832); BEST Arabidopsis thaliana protein match is: Acid phosphatase/vanadium-dependent haloperoxidase-related protein (TAIR:AT1G67600.1); Has 11 15 Blast hits to 1115 proteins in 408 species: Archae - 0; Bacteria - 729; Metazoa - 0; Fungi - 0; Plants - 228; Viruses - 0; Other Eukaryotes - 158 (source: NCBI BLINK).	aa seq
811	sp_107230_joya -0.851343	#N/A	#N/A		MTGRNIDVLLASSNSSTYSQDLLMFPNFFLVSS LLAISLAQFLKIFTSWFKKWKDSKRMILSSGGMP SSHAGVSALAVAIGVQEGTSGSAFALAVLACV VMYDASGVRLHAGHQAQLLNQINVCLEPPEHPLS NCKPLRDSLGHHTPLQVAAGAILGCIVAYLMOFYR
812	sp_107870_nair #N/A	-0.964394	#N/A	uncharacterized protein	MGRKTRSRKRGKVLINSEEEGNSVVDSSSRSE QDDSDDEDFKADGYQDLDLDDDDDEVQLDVGSG VKFKRGRSKLLYLDRLDKIPVQWGIFFRLKAWDM KEILEAKRAYRESGKGDYGLVGYIPYMDNEILR LKQRKGNSEIRIPMKRQMAKERAKKNSVKLER FELDMDTVNEEAITPPPPRPESLSERCQQVTSS PPPQHKVTPPLPSPKVPPPPQPKVTSPLRPPSLPK VTSPLPPSPPKVTSPLPPSPPKVTSPLPPSP PKVTSPLPQSQANSTSPPLPPPKVNSLPPPS QPKATSP_LPHSPPKVTSPLPHSPPKVTSPLPP PSPQFTLPRSPQLCYDPIGGANLEAVNLMES PLPGSLDREKGTIVCGLSAVPLGHDIIDCDNKMD GLASSALLFORSPRRGRKRVALADRGRGWFF TSEIKKFSYNLVEFVNHNGGLLSGRGNLKGRLYY NYYVEIKMKMCPVS

10

20

30

40

【 1 3 2 6】

【表 1 4 - 3 2 7】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
813	sp_108470_hhjh	-0.433333	-0.473333	#N/A	Transducin/WD40 repeat-like superfamily protein; FUNCTIONS IN: nucleotide binding; INVOLVED IN: biological_process unknown; LOCATED IN: CUL4 RING ubiquitin ligase complex; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: WD40 repeat 2 (InterPro:IPR019782), BING4, C-terminal domain (InterPro:IPR011046), WD40-repeat-containing domain (InterPro:IPR017986), WD40/YVTN repeat-like-containing domain (InterPro:IPR015943), WD40 repeat (InterPro:IPR01680), WD40 repeat, subgroup (InterPro:IPR019781); BEST Arabidopsis thaliana protein match is: MOS4-associated complex 3B (TAIR:AT2G33340.2); Has 9286 Blast hits to 5748 proteins in 422 species: Archae - 16; Bacteria - 3151; Metazoa - 2206; Fungi - 2085; Plants - 594; Viruses - 0; Other Eukaryotes - 1234 (source: NCBI B_link).	MVRRAEKIVPAEAEQDEEEVDVNVKYYLRGESAN LEVLDQKKLKGELAVKEELYKRSAKAAANVEKWL LPSEGGFLESEGIETWRIKQDQDAGKEVDILSSKN QYDMVLPDLGPKYKLDFTASGRYMAAGRKGLG VLDQKMSLVKFEQVRETVRDVVFLHNEFFAAA QKKYPIYNTDGVVELHCLKEHGAVSKLQFLRNFH LLASINFKGQLRYQDVTMGKMIANYKSGLRITNV MOVNPNYGVIGLGHANGSTTMMWKPTSPNPVYVK MLCHHGPLSALAFHNSHGLMATAAGMERKIKLWD LRTFKVLOTMPGHAESLEFSQKGLLAASTGSFY QVFRDGGSSQNYRYSMSHSMAGYKIEKVSFRP YEDVLGIGHSMGWSSILIPGSGEPNFDSSWANP FETTKQRREKEVRSLEKLPESIMMDPKKIGTLK PPRKEKPTRKEREEEMEAAVEAAKGGKLNKT KGRNKDGGKRAKKQEVVQNAKKQFIEQNIKENE ELSKKKRKRTEEVVLELPKSLQRFKAKKAD
814	sp_109590_uamn	-0.375316	-0.395331	-0.383332	RING/FYVE/PHD zinc finger superfamily protein; FUNCTIONS IN: zinc ion binding; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, C3HC4 RING-type (InterPro:IPR018957), Zinc finger, RING-CH-type (InterPro:IPR011016); BEST Arabidopsis thaliana protein match is: RING/FYVE/PHD zinc finger superfamily protein (TAIR:AT5G63780.1); Has 319 Blast hits to 315 proteins in 19 species: Archae - 0; Bacteria - 0; Metazoa - 1; Fungi - 0; Plants - 317; Viruses - 0; Other Eukaryotes - 1 (source: NCBI BLink).	MMGKEIQKETVNOQVCSGSSSPTKGDSSRASTA SQLSIDRDLGLPSCRVCQCQVESDRRGGDAALGFL GIDPLLVESHNVKELPEPGTKKALNDVEGSASVD TSKKLEPGYVQFSPTGTEVFCNTDLEMGACYNR DTLJELGCAKNDLALVHYACALKWFINHGSTVC EICGCIASNIRIADFVKVLSLAKDYEARLDRRTANG VFNPAHVHCNSDVPDPAVAAIRQRLEIYL WFNPNYGLRHNNNTSENINAPSVLVVTPEPANIV FQDVTQNAATKWAVEGTGILLATGLLTVTLAWL IAPRVGKRTAKNGLHLLGGVICALAVVFFRCVIL TRIKYGPARYWAILFVFWFLVFGIWSRTHAAHT T

10

20

30

40

【 1 3 2 7】

【表 1 4 - 3 2 8】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
815	sp_109630_kwcn	-0.85:2405	-0.24:7186	-0.94:2345	polyprenyltransferase 1 (PPT1); FUNCTIONS IN: 4-hydroxybenzoate nonaprenyltransferase activity; INVOLVED IN: ubiquinone biosynthetic process, embryo development; LOCATED IN: integral to membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAINs: 4-hydroxybenzoate polyprenyl transferase (InterPro:IPR006370), UbiA prenyltransferase (InterPro:IPR000537); Has 8781 Blast hits to 8779 proteins in 1934 species: Archaea - 174; Bacteria - 4966; Metazoa - 236; Fungi - 294; Plants - 110; Viruses - 0; Other Eukaryotes - 3001 (source: NCBI/BLINK).	RKIANPPDSDQSFITTPSMAALCLRLSRSTRKFLP FSLSYHSPHLSPSSTHPNSNPNFYHTRSPNHYYT PNFNISFNWDHFTLCQKTFSSAYSTVPQGSKLA NSGKEDDDDEQLKSKQRGEASWVDLYLPRRSAR PYAKLARLDKPIGTWLLAWPCMWSITLAAEPGNL PDFKMLALFGCGAFLLRGAGCTVNDLLDQDIDVK VERTKLRPIASGLLTPFOGIFSLGVQLLLGLGILLQ LNDYSRILGASSLLLVFSYPLMKRLLTFWPAQYLG LTFNWGALLGWSAVRGSLLDPYVVPLVYSGVCW TLVYDTYAHQDKEDDLKVGKSTALRFGDLTKE WLTGFGFVYSISGLALSGYNAHLGWPPYACLTA SGQMAWQIYAADLSNRADCNKRFVSNKWFGLI FCGILFGKLA
816	sp_110690_skox	#N/A	-0.84:9663	#N/A	sumo conjugation enzyme 1 (SCE1); CONTAINS InterPro DOMAINs: Ubiquitin-conjugating enzyme/RWD-like (InterPro:IPR016135), Ubiquitin-conjugating enzyme, E2 (InterPro:IPR000608); BEST Arabidopsis thaliana protein match is: ubiquitin-conjugating enzyme 2 (TAIR:AT2G02760.1); Has 9584 Blast hits to 9566 proteins in 388 species: Archaea - 0; Bacteria - 0; Metazoa - 4206; Fungi - 2037; Plants - 1834; Viruses - 20; Other Eukaryotes - 1487 (source: NCBI/BLINK).	MSGGIARGLAEERKQWRKNHPHGFVAKPDTQ ADGSMNLLVWOCIPGKAGTDWESGYPLTMVF SEEPKPKCKPPPPFFHPNVPYPSGTVCLSLN EEDCWRRPAITVKQVLVGIQDILLDQPNAADPAQT EGYHLFIQDITTEYKRVRVQQAKQYPAISI
817	sp_110830_qnkm	-0.33:0.701	-0.93:1246	-0.985:115	AWPVM19-like family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; CONTAINS InterPro DOMAINs: AWPVM19-like (InterPro:IPR008390); BEST Arabidopsis thaliana protein match is: AWPVM19-like family protein (TAIR:AT5G46530.1); Has 177 Blast hits to 177 proteins in 19 species: Archaea - 0; Bacteria - 0; Metazoa - 0; Fungi - 177; Viruses - 0; Other Eukaryotes - 0 (source: NCBI/BLINK).	MGTVGRNIAGPLFLNLMYLVIGFASWCINNYIN FSTHPGLGGNGATSFLLVLSILAGVIGVSKLAG ANHIRAWRNDLSAAAGSSALVAWALTALAMGLA CKEINIGGWRGWRLLRVLEAFIILLTFEFLFYVLAIH AGLFSRRYGPYREHYDYGVMGVSGETTVDPKGT TRV

10

20

30

40

【 1 3 2 8】

【表 1 4 - 3 2 9】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈  Protein kinase superfamily protein; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: plasma membrane; EXPRESSED IN: cultured cell; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase domain (InterPro:IPR002290), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), BEST Arabidopsis thaliana protein match is: Protein kinase superfamily protein (TAIR:AT4G10730.1); Has 125358 Blast hits to 123605 proteins in 4454 species: Archaea - 114; Bacteria - 14309; Metazoa - 46798; Fungi - 12057; Plants - 31507; Viruses - 632; Other Eukaryotes - 19941 (source: NCBI BLink).	aa seq  MGRSGHAIRSYTVNSNDYKLLLEEIGFGASATVFR AIYIPYNEWAIKLDLDRCNNSNLDRIREAQMMS LIDHPNVVRAFCVDFVDRNLWVMPFMAEGSCL HLMKTAYPDGFEEPVICSLKESLKALEYLHGHG IHRDVKAGNILLDTNGLVKLADFGVSACMFDSDG RQRSRNTFVGTCPWMAPEVLQFGSGYNFKADI WSFGITALELAHGHPFSKYPPMKVLLMTIQNAP PGLDYDRDKKFSKFKEMVAMCLVKDQAKRPTA EKLKHSFFKHAKPPEFSYKVKLLADLPLWDRVK ALQLRDAALALKRMPSAEQEAISQNEYKRGVS AWNFIEDLKAQAALYQEEEDLAEIKEADESVKS ATSSKADAPYGVSYGNSNPPNFINPKQGNVDE VAVTPTGASCKIRGTGNEFDGFCSHGKNIKDGV DLCSGAPACEMDAAGVKNVQVVKGRQSQSGP LMPGIVLQGTSTERSVSRPENENQSAAEKSN RGSRRVPSFGPLVLPTRASANSLSAPIKSSGGY RDSLDDKSKANLVQIKGRFVTSLSLVVKDMM TSTSTRRSSQSGPLPKSASVGDWISSEKKTITS QPSGDINSGNVSALLPHLQNLFOQTAIQQDLIT SLLNSLOSTEGDGSQNGKPPAVPTSGSPENGIV ETVASERERLLLNKISDLOARMAISSEELSAEKH KHLLLOQQLNSVSSSQEEDDEVVN
818 sp_110880_aplz	#N/A	-0.98387	#N/A	RESISTANCE TO P. SYRINGAE PV MACULICOLA 1 (RPM1); FUNCTIONS IN: protein binding, nucleotide binding; INVOLVED IN: defense response, plant-type hypersensitive response; LOCATED IN: plasma membrane; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: NB-APC (InterPro:IPR002182), Leucine-rich repeat (InterPro:IPR001611), Disease resistance protein (InterPro:IPR000767); BEST Arabidopsis thaliana protein match is: Disease resistance protein (CC-NBS-LRR class) family (TAIR:AT5G35450.1); Has 20241 Blast hits to 16785 proteins in 626 species: Archaea - 16; Bacteria - 1308; Metazoa - 4084; Fungi - 259; Plants - 14046; Viruses - 4; Other Eukaryotes - 524 (source: NCBI BLink).	MISKVKELGFCQIFSNSEITEEFVQPHARRLSIHM DHSFHDADHQLMQSITKRKSSIRSLLYAKMYLQP QVSTNLFSKFNLLKLVDFCNALIDYLPKEVGELL NLHYLSRNTQIKRIPNSIGKLOYLQTLDLKGTVP YELPVDLNTLHKLRRHLLTYCYKYDSSGFSLTKL TGVKLCEKTEKFGELQKLAFLVDVGHGTQNSW KELRNLQRLRRLGITELKFKEDGKDLCTAIEEMKFL ETFSVYVTFEIDLDFQSASELILKRLYLINGPLKQ HNTGWLLEHSLVKIRLRWSRLEDDPLHILETLPN LVELQMLEAYDGDNLHFGNTGFKKLLHLLDLQ NLKSLSIWKGALPLLQMLAIGOSMKLQVPLGIKHL SRLETLNFFDMPDFARSLQHGKGFHSIVKHPVN VLFQKSDGIGWETCFL
819 sp_111520_konm	-0.999352	-0.982138	-0.952316		

10

20

30

40

【 1 3 2 9】

【表 1 4 - 3 3 0】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 CLIP-associated protein (CLASP); FUNCTIONS IN: binding; INVOLVED IN: mitosis, protein stabilization, cell growth; LOCATED IN: spindle microtubule, plasma membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: HEAT (InterPro:IPR011989), Armadillo-like helical (InterPro:IPR021133), Armadillo-type fold (InterPro:IPR016024); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLink).	aa seq
820	sp_111640_tmsu -0.3221165	-0.371331	-0.362332	mediator subunit 8 (MED8); Has 21670 Blast hits to 11104 proteins in 489 species: Archae - 4; Bacteria - 472; Metazoa - 8306; Fungi - 2813; Plants - 2364; Viruses - 51; Other Eukaryotes - 7660 (source: NCBI BLink). FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion, endoplasmic reticulum; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DJF106, transmembrane (InterPro:IPR002809), Uncharacterised conserved protein UCPO10045, transmembrane eukaryotic (InterPro:IPR008568); Has 409 Blast hits to 409 proteins in 203 species: Archae - 0; Bacteria - 0; Metazoa - 152; Fungi - 136; Plants - 53; Viruses - 0; Other Eukaryotes - 68 (source: NCBI BLink).	MEEALELARA KDKERMAGVEHLHOLLEASRKP LTSSEVTSLSVDTCIDLLKDNFRV/SQGGLLALASA AVLSGDHFHFKHFNALVPATVERLGDGAKQPVPRDAA RRLLVTLMEVSSPTIIVERAGSYAWEHKSWRVRE EFARTVTSAINLFASTELPLQRAILPILHMLNDPN HSVREAAISCEIEEMVMQISFKMSYKDTSFHHIW  MEAMMGTOGGQTCAAPPQQPPEKVVVEKLSA VQQQLNLDVSKTRAI SLFKAISRILEDLDFVLSRSN ALPKWQDVVGGQFMSVNLLEFNIVEDIKVSKAFV VHPRNVAENATILPVMSSKLLPEMEVDDNSKR DQLLLGMQNLPIQIDKLRIDMIAAACEAEK VIGETRAYGLDTRQRMSTVPTLDKAFSAKIQEK ENLLRAAVNHGEGRLPVDORHVTSP LPSHLAD VLTVDGGQTYDGGSGMYSKNTPVISGNHNG GLPLOGAAGGQLGRSVPSPSTATVSTAENT ASPMQYANSPRSGTNMINAPSSQQQLTQQHQ LQQQRKMMQLPTQQQQQLLAQQCFRQSAMQ GLAQLGQSQMFSQPLASQQFQSRQLPTGHI GQSQLSQGNLNRHIGQFSGAANTALFNASQNS PNASVMSMPATMTSGSMLPRMQFALPFGNPFQ RTHFSQMLSDQMFMNMGANQASMMSLPQQQQV PQGAFGMSQNSQNLQSNMVPLOQATQNLQNF QQQQRQGNQQ
821	sp_112110_iddz #N/A	-0.371325	#N/A		
822	sp_113170_cfim -0.341639%	-0.343367	-0.376952		

10

20

30

40

【 1 3 3 0】

【表 1 4 - 3 3 1】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	共発現遺伝子				<p>MKRHTFTAGEGSKGCEACGMDFVAFESTGTFDFTVDIS          SOLMQRYSSSSAPQHRHLMATAAAVKSILLESLSL          PLSPLCYFAATINALSQPNQDAASASALATFLSIV          LSLVPEDYINPEKAAEAAILRVVLGSEEGEMWV          GSASSARCWVKCLGILLGFCDELDWDSVELPFQT          LLKWSIDKRPKVKRCALVYLENALKSFQSETVKR          KASILVMSLLESYMSIAVKLGIQKVDSESSERQC          ETEQLDVVHALSWRVALPYICDKSIKKTIKLIJELL          DSRSSFTFRPVLNAIEALLEKVEGGIIVQVAEDIID          ALSSFISSETNPNDISIISANLLKISLDKLSHSEDID          KWNKSLVGVNAIAGLLTREDEVASKSSNILKNLI          VDQLSDDIESQPAGEESKGITSMCSTFMEMLD          SLAGIPNEHTLAVLSVFLKLGRRSSCSHMKNIVLK          LAEIFTHADGNKHDTVHLQNSLGCVAVAMGAEN          VLRLVPISFDEENSTYANMWMPLIKKYTCGASL          GYFTEHIVPIAOSLOKASRIKVKKSVTGODLOALA          HGLWGLLPAFCRHFPTDVSQSFQFAKMLLVQLK          KDAASMREEIANLSQELVKNKCLPKFDDAAIVEPIK          HSSGFSVDDYLYEKRNLLSYKKAIAIKNIALSLY          SNELLALMTVFLNSKPAKRSSLKKAISCLASISD          SSVTKFISSLERFPFVKAVSVSEDSDPSSNALLS          DVPSDSTTIGENSQWSLMLDLASSIVERADEPLV          GLFLKTKHVLAQGDGKCIETEAYTLCVLEEHP          WFCSLKMDLVDLIMGLKAPADISFSSFNRLSCLH          LLLVHSLKSSSDVENTKPFILINDIILALKDSREEA          RKIAYDVLKLSCLVNSSDSDSDGPPYKUSIMIM          GYLSGSPSPHITSAAISALSMFVKEPELCLVKVPHV          VSSVASLLOTKAVETIKAVLGFVKLVVSSLEAKDL          QNFPALIDGLVLPWSSSRHFRFSKVTVILEIVIRK          CGFPTIKSATPERYQSFIKRVKSGNRHGNNTTSKEA          DSADAKTQLSNVPINGSKRYSYEETNNNRNGSE          LQRNRRNIQIDFQSPGGTRRNVAAGFDKFKPNKF          LQNFRRNIDSLDIPDVSSTKILMGLDDPADIWRASS          VSRFLWRDFVTNGIGKHLCLRFLPQLSGIDRVYDI          SNKQENSVDGSSSVELEHLKKEHRAFVALARG          LTSSEYGECLSDAISATSDNYPEEGVHNTLVSR          DRIGRRASYWSSSGQSNPEVPERLTKLSDLCII          TEMNVHPFOAFFQSNPIYSAKAVFRFMGHRILP          EGDEEDIVHEGQECAEKFWVYTSPEFPMAQE          NRVONFKLEPILCFGGYLOIELLGRVQRQEMDG          LFYICVAHVQVMGRSLPFPVGVGVELDIPNGRFTLT          YNPFHAQVSPHNGNEESDEISVTPEMQRHVRG          WEQILNMLRGLTGVVEVYDSDDEQHDSEDEMPFEE          FAL</p>
823	sp_113760_euny	-0.831363	-0.828046	-0.985625	<p>ARM repeat superfamily protein; FUNCTIONS IN:          binding; INVOLVED IN: biological process unknown;          LOCATED IN: cellular component unknown;          EXPRESSED IN: 21 plant structures; EXPRESSED          DURING: 13 growth stages; CONTAINS InterPro          DOMAIN/s: Armacillo-type fold (InterPro:IPR016024),          Domain of unknown function, NUC173          (InterPro:IPR012978); BEST Arabidopsis thaliana          protein match is: ARM repeat superfamily protein          (TAIR:AT2G34357.1); Has 378 Blast hits to 366          proteins in 171 species: Archae - 0; Bacteria - 6;          Metazoa - 125; Fungi - 137; Plants - 63; Viruses - 0;          Other Eukaryotes - 47 (source: NCBI BLINK).</p>
824	sp_114360_gqqp	#N/A	#N/A	-0.832297	<p>F-box family protein; CONTAINS InterPro DOMAIN/s:          F-box domain, cyclin-like (InterPro:IPR001810), F-box          domain, Skp2-like (InterPro:IPR022364); Has 80 Blast          hits to 80 proteins in 29 species: Archae - 0; Bacteria -          0; Metazoa - 0; Fungi - 11; Plants - 57; Viruses - 0;          Other Eukaryotes - 12 (source: NCBI BLINK).</p>

10

20

30

40

【 1 3 3 1】

【表 1 4 - 3 3 2】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
825	sp_114370_keez -0.833763	-0.837437	#N/A	phloem protein 2-A13 (PP2-A13); FUNCTIONS IN: carbohydrate binding; INVOLVED IN: N-terminal protein myristoylation, response to wounding; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: F-box domain, cyclin-like (InterPro:IPR001810), F-box domain, Skp2-like (InterPro:IPR022364); BEST Arabidopsis thaliana protein match is: phloem protein 2-A12 (TAIR:AT1G2710.1); Has 444 Blast hits to 438 proteins in 25 species; Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 444; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).	MGVGMSSDGGNHGGTPEYRRKLEDIPESCIA MVLGHLDPEICRLARLNRAFRGASSADFIWETK LPTNYSFVNVKISELNHNKDDMMKNYGGSKDIST GLNKDFARLSRHLNLFDSGNKEIWLDKKTGGVC VSISKALSTIGIDRRYWNHIPTESRRFQTVAYL LQIWWFEVCGELEFFKFPAGTYSLFFRIQLGKSSK RFGRRSCNVEHIHGWNLRPVRFQLTSDGQHAS SECYLDSPGSWILYHGGDFVWKETDKLTKIFSA RQIDCTHIKGLSFDSDVMILPKSIASKAR
826	sp_114620_aksi #N/A	-0.873053	#N/A	VQ motif-containing protein; CONTAINS InterPro DOMAIN/s: VQ (InterPro:IPR008889); BEST Arabidopsis thaliana protein match is: nucleotide binding;nucleic acid binding (TAIR:AT1G21320.1); Has 294 Blast hits to 294 proteins in 63 species; Archae - 0; Bacteria - 0; Metazoa - 108; Fungi - 21; Plants - 144; Viruses - 0; Other Eukaryotes - 21 (source: NCBI BLINK).	MDHPDSPGGRRPRKELHGPRPAPLKVREKESH KIRKPPVAPPAYPTQNLQQLPPVLPFRQPVIYTV SPKVIHTHPSEFMTLVQLRGTSDATSSSYATSSN VTSSSSTVDFGSSSTTTTRSHHVTGGISPAARFA SIEKTKMHSEAKKFKVGPSEVFRREVMIDIEEVVNS GGGGGGGGGGGIEGVERTMVFVPEIACVEA
827	sp_115131z4387 #N/A	-0.875285	#N/A	uncharacterized protein	#N/A
828	sp_116390_ztnwk -0.8333537	#N/A	-0.833714	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G11290.1); Has 881 Blast hits to 512 proteins in 30 species; Archae - 0; Bacteria - 2; Metazoa - 0; Fungi - 38; Plants - 833; Viruses - 0; Other Eukaryotes - 8 (source: NCBI BLINK).	MNEQKVRKATWDDSTRFCEYCADEVHAGN RPNTHFRVGVNDVVKFKQLRTGKRYDQKQLK NKWEKLSSEFSTWKNLHKETGLGWNHDNNTID ASEDWENKEQVDANAIFREGGIKILTEQEIIM FSKIVASDATSWNPYSEEEVETEVEVEDDSIV GNFDVEEQVDITLDSNTTQTPLQPESENIK SNVTIPTGVKVNCKGKSKVSTAKLMQNELKR IVGAMESFSQSAATVTSRRRDLPLGCSIQECLTL GVTPGVEGSEALYMLGTRLFMKKRECFREMFVALP GEDIRSAWLDQELLREKENKT

10

20

30

40

【 1 3 3 2】

【表 1 4 - 3 3 3】

	A	B	C	D	E	F
1						
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
829	sp_116910_xylr	#N/A	-0.872326	#N/A	Oxoglutarate/iron-dependent oxygenase; FUNCTIONS IN: oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, 2-oxoglutarate as one donor, and incorporation of one atom each of oxygen into both donors, oxidoreductase activity, oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, iron ion binding, L-ascorbic acid binding, INVOLVED IN: oxidation reduction, peptidyl-proline hydroxylation to 4-hydroxy-L-proline; LOCATED IN: endomembrane system; EXPRESSED IN: petal, sepal, male gametophyte, flower, pollen tube; EXPRESSED DURING: L mature pollen stage, M germinatec pollen stage, 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: Prolyl 4-hydroxylase, alpha subunit (InterPro:IPR006620), Oxoglutarate/iron-dependent oxygenase (InterPro:IPR005123), Metridin-like ShK toxin (InterPro:IPR03582); BEST Arabidopsis thaliana protein match is: Oxoglutarate/iron-dependent oxygenase (TAIR:AT3G28480.1); Has 2477 Blast hits to 2439 proteins in 330 species: Archae - 0; Bacteria - 383; Metazoa - 1060; Fungi - 71; Plants - 412; Viruses - 15; Other Eukaryotes - 536 (source: NCBI BLINK).	MKFDICRSLHSSVSVYLLLVVPSMILLCSW SEVLVRCSLVSLFNLLVSLVFLVYRGRFTEEC DHLIPLAHGOKEGTNSVISDTQGVARSVLVLDL
830	sp_117090_zisk	-0.933347	-0.931657	-0.810314	CONTAINS InterPro DOMAIN/s: P60-like (InterPro:IPR011687), Tumour suppressor protein Gtiscr2 (InterPro:IPR011211); Has 709 Blast hits to 643 proteins in 201 species: Archae - 0; Bacteria - 32; Metazoa - 224; Fungi - 154; Plants - 45; Viruses - 0; Other Eukaryotes - 254 (source: NCBI BLINK).	MGKINRAKVKRAWRNISTKEFDDYNEKATRDAN SGGSLASAPSDSLFVVDKSTDVPRRRIKIEKKREK VLFRESVLQRNPFVKAVPSSILKLLKDKTKVDLA QDASKDDDKKGPESGKFDIWDQGGKVNIGKKK ANFSLIPAVEVEPPGGSFNPENESHKDSLALAVA EEMKKIYKKELEPEPIPLVPGPEVTEDDRYFLEV DHEDDDVDEMEENQTEDQINEDENPKSKEKR ITTVVKNKRIIRKEQOKAEAEAKKIKDLKSKEISIP DILQEIKADEEEKKNRHLRMRVAKERLKSQPPR LGKHKFEFAPSQVLLTEETIGSLRKLKACCTLAR DRFKSLEKRGILVPSAKRSRK

10

20

30

40

【 1 3 3 3】

【表 1 4 - 3 3 4】

A	B	C	D	E	F	
1	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	
831	sp_117540_guqq	-0.865657	-0.997968	-0.977733	<p>arogenate dehydratase 2 (ADT2); FUNCTIONS IN: arogenate dehydratase activity, prephenate dehydratase activity; INVOLVED IN: L-phenylalanine biosynthetic process, metabolic process; LOCATED IN: chloroplast; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Prephenate dehydratase (InterPro:IPR01086), Amino acid-binding ACT (InterPro:IPR002912), Prephenate dehydratase, conserved site (InterPro:IPR018528); BEST Arabidopsis thaliana protein match is: arogenate dehydratase 1 (TAIR:AT1G1790.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLink).</p> <p>S-adenosyl-L-methionine-dependent methyltransferases superfamily protein; FUNCTIONS IN: methyltransferase activity, catalytic activity; INVOLVED IN: metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Methyltransferase type 11 (InterPro:IPR013216); BEST Arabidopsis thaliana protein match is: S-adenosyl-L-methionine-dependent methyltransferases superfamily protein (TAIR:AT3G60910.1); Has 2390 Blast hits to 2390 proteins in 771 species: Archae - 92; Bacteria - 1226; Metazoa - 381; Fungi - 88; Plants - 206; Viruses - 0; Other Eukaryotes - 397 (source: NCBI BLink).</p> <p>Protein of unknown function (DUF803); INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAINs: Protein of unknown function DUF803 (InterPro:IPR008521); BEST Arabidopsis thaliana protein match is: Protein of unknown function (DUF803) (TAIR:AT3G23870.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1398; Metazoa - 17338; Fungi - 9422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLink).</p>	<p>MLDVASSNSLTSSLIQFRVQFTQPSHHHYATLN        FPKTHFKRCRLNIDSSLHPTSNYIPOAILSQKLPE        DSPSAADATNLPFRPLTSTVASNSDGSSLRVAY        QVSGVAYSELAETAYPNCEAVPCDQFDTAFEA        LFERWLVDRAVLPINSLGSSHRNYDLLLRHRLHI        VGEVKLPIRHCLLANPGVKIDDLKRVLSHPQALS        QCEHTLAKFGVTRAEVDDTAGAAKFVAFNKLQD        AGAVASAAARYGLNLAIEDIQDSDSNVTRFLML        AREPIIFGSDRPFKTSIVFSHTEGPGALFKALAVF        ALRQINLTKIESRPLRNSNNGSSKCFDYLFYLDFF        ASMADPNACNALRHGLEGFATLLRVLGSPADTS        RR</p> <p>MMVPSMSEIKAMVADMMLDLPFADDSFDVYIEKGT        MLL</p> <p>MMMSDNNKGLVAVSSSIFIGSSFIKKGKLMKA        GTTGTRAGSGHSYLYEPWWWAGMITMIVGEIA        NFAAYAFAPALVTPGLAL-SIFSSVLAHFILKEKLIH        FGMVLCVVGSTTIVLHAPHERDIESVKQVWT        LATEPGLVYAAIVVMVVLVIFYVYPRNGRTHMI        VYIGCSLMSGLTYMVCVKAVSIALKLTFFEGNNGFK        FFGAWFFTVVCCILQINLYLNKALDFTNTAVVSP        VYVYVFTSLTILASIMIMFKDWDSONGSOIATELC        GFVTILSGTLLHKTKDMGSTPAAAAVPVTELYR        QYSSNCIRDSSSKHSDR-HGSPDRNCRE</p>
832	sp_117640_yejk	#N/A	#N/A			
833	sp_117670_ndhn	-0.830367	-0.914029	-0.932377		

【 1 3 3 4】

10

20

30

40

【表 1 4 - 3 3 5】

I	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数 sp_107620_kqmh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
834	sp_118770_kxzo	#N/A	#N/A	-0.35421	<p>U3 ribonucleoprotein (Utp) family protein; FUNCTIONS IN: rRNA processing; LOCATED IN: small- INVOLVED IN: processome; EXPRESSED IN: 23 plant subunit; processome; EXPRESSED DURING: 13 growth stages; structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Small-subunit processome, Utp14 (InterPro: PR008709); BEST Arabidopsis thaliana protein match is: U3 ribonucleoprotein (Utp) family protein (TAIR:AT5G08600.2); Has 7468 Blast hits to 4514 proteins in 366 species: Archae - 14; Bacteria - 554; Metazoa - 2650; Fungi - 816; Plants - 392; Viruses - 146; Other Eukaryotes - 2896 (source: NCBI BLINK).</p>	<p>MADNKRSSGGAKQLORREKTPKFKSSKSNPK SSGKRGGGGGGGERRRPR_LPSAFRKLQDNF KNPNNGDDVDEVIDSDGELNYVKDFYVE EEVAEESSKNRRFDHVENYEQLPKFKDEDIL SDDDEDDLDODKDESDVDEKGGDDGRHARM TDITGHTFAFAGKKRKNVVEAYPESEFNPS DVL_SGNGKISIQDILLPLRGTSSYSKLRKRIEHE KGETVQAPLSKPDRELERLGRVAVDHWKDKITK WEHLVKNREAPTLYFDQGNMGLPTVGTMAS EFKPRDFEKAIASIMDNKEVAEAYQSDGAKLLE LNKISVEEVKDRQDR_LAKMRSLLFRHEMKAKQVK KIKSKTYHRLKKKOKKLLKAASDPQTDPEAAREYA EKQERKRAEARMTLFRHKNQGWAKRVLRLRGLS KQDEGTQAAIAEQQLHLHLLTRKMNSVKENSSS DSSCEDDEDMDSSDEEGASKLLAKAKEKTIKL MEQEDEVDPDGLSLPFMVRGKLEKRKE/MNGEA KRALDDYELSLNQTAEKGLKNGTSPSGRRYFVGT KKVEEANKHRPVDVNYGSDNEDTIEEKDDIHE CAKGNHMEKVDVLDLREDLETGDDDSAFKSVI DVAEDSGPRTSRDMKNDGCKGGSTHDANVSK PEVAKPSPSDSGEGEDDSDTDSGGHMIDGMLS SGARQTYELPTNEELIKRAFAFGDDVEEFKIKE DVLNEENPEPEKPLLLPGWGWQWTDVOKKRGLP SWMLKDHANAKOKREEA_LKRRKDARLKHVISEK IDKKVSLFJFDKYHLPLLLKFTWWKSLDAGQV VN</p>

10

20

30

40

【 1 3 3 5】

【表 1 4 - 3 3 6】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)		
2	共発現遺伝子			注釈	aa seq
835	sp_119020_netp	-0.947759	-0.955392	INVOLVED IN: protein localization; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Vacuolar protein sorting-associated protein (InterPro:IPR009543); BEST Arabidopsis thaliana protein match is: calcium-dependent lipid-binding family protein (TAIR:AT1G48090.1); Has 2219 Blast hits to 1322 proteins in 202 species: Archae - 0; Bacteria - 2; Metazoa - 1077; Fungi - 361; Plants - 231; Viruses - 0; Other Eukaryotes - 548 (source: NCBI BLINK).	MFEQVPRQLLLGYLGKIKDFHQKQLKTLDEIVL KDVELILEAFDYQLPVDLKRCHVGRRLRIGMSWK KLSWESWDPIVLSLEEDVLCAGPRDDHEWSID AVERRELAAKAKAAAELEKLSRRVCDNQAQ SLSHVTGRILDSIQVSMRNHVWYCDIQSOSMLF GLFFSSLEMKQAADGPLSSKARGQGVNKIIEIRG LGHCSTLEGLGSLGVNGVGETELWRRVTVDV DKYDYLLMPFDLSQVTKSGKLEFNAPQYSVYVSE LTKLVMTINPAQLQHILMFSDDYICTRLREKYGRY RPWDFPLSKKLGWQKSWWHYAKESVLYDYVRK RRKETSWRHLGKQKIVLRYKYNLYKVKLDLFLOR EQLVDDQDTLLKLEQMEKDSIEDILRFRFSYAEREL EDSQRSVSSDAIEGSGTPEAENLPNDEVPTGGARG WLNWLSRGMLAGGTTDDSSQFSGVVSDEVYKLD ICKATKQPMFVSVTVAANGVKFLFAVKFHYSKMS ILINSTKLDQEIASELSCFATQIEFNWVEESTKILAK VESAEVIDSSKMAILQMKQKYSTCQNVLESKKPF ASTKVDISVKNQMEEFKVKYMIQPIVEIYCDLKFLL RVEVFYSVMESHKSLQERVYLLSLNGIEDANCRLO SKAAE MAEEERECLTMDKDTNLLSKVAANFLYLTQFEA FRSTILSLRHRNPNLALSLLQTVVNGGNFKNAIY STNCSPPALLTWLSSLELFOFENCTSIWSSSTSA DSLRLRVFELLYIQMVSSRVSESVNRKAEMEEEE AEEELRELKSMILLDDGIMELGLRRLKEDLIVSLG EERGSVEEEGVVIREEELGCLRVVLLKNADYFVS LCENIQGQVGVLEKDDNLELAITLRREGRTV LSEEDQRVLSIQKYAQIAHLEEMRQCVIMDDVD EAILHRYLHLVYGVSKEEYRMVFQDILLQKFGSGK EGLGDTSCCTTRDKLLYYREALLSNCTHLVQMIQP CW
836	sp_120190_wezk	-0.862343	-0.857336	unknown protein; Has 157 Blast hits to 144 proteins in 62 species: Archae - 0; Bacteria - 0; Metazoa - 101; Fungi - 0; Plants - 35; Viruses - 0; Other Eukaryotes - 21 (source: NCBI BLINK).	

10

20

30

40

【 1 3 3 6】

【表 1 4 - 3 3 7】

A	B	C	D	E	F
1	分析に用いたペイトとの相関係数	sp_107660_kicg	sp_107670_piqx	注釈	aa seq
2	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_piqx (SOAP2-like)	注釈	aa seq
837	sp_121190_ouox #N/A	-0.3332	#N/A	<p>PHD finger family protein; FUNCTIONS IN: zinc ion binding; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, PHD-type, conserved site (InterPro:IPR019786), Zinc finger, PHD-type (InterPro:IPR001965), Zinc finger, FYVE/PHD-type (InterPro:IPR011011), Zinc finger, PHD-finger (InterPro:IPR019787); BEST Arabidopsis thaliana protein match is: homologue of trithorax (TAIR:AT2G31650.1).</p>	<p>MTDNRFCGGRKRWIMGRGREGGGGGTEERYCPVY  RSPPPASSPAPVSSPENLEKVNAGVLDLYSQAFK  VLSERSPPDSDEVLAKTVATLPAGLWLLSKQSD  SKKKHHPHSETKSSRKSGSTEKSKSSRLWTGLE  DYFRELTPDIERLYRSPFNSLINSSTSNSSLK  VPVLGNGVRTSSDNDNVSKIENRHANAATAAAY  KDEEODEVLDANVPLDGTALQDKKEEFDKLDHV  KVENGSRDLSLSEKKGPFPEYSLOSCESSMEWVL  GSRNRVLLTTRPTTKRKLGVAGLERLMAAQ  PCEGSSSLCHVCSMGETGDLNQLVVCNMGVY  AVHQKCYGQGEVDGGSWLCWCKKHYNARAG  PSSGNDIPCVLCPKSGGALKPFQONARSGDGSV  EFSHLFCMMWPEAYIDDTLMEPVLCRGMKDV  RGLVLCNLGKMKLGCVCVRCSDGSCRTSFHPLCA  REAKNLEIWGRFGCDDIELRAFCAKHSGNLDLD  DSVALESHQNHHTYGSDSLSSIESALLPGNKIK  KLKIGCKNGDKVTFQMRISDAKSNLGVRRERQDT  GIKDADSNTKLKSQGVVEEYEDVSTQLSDIGLGG  TGDDVKTSRVLDLGLVLLKLIHRGKVSLLKQIAAEI  GIOEASLASRLADNHLSPDLYCKLLKLLWLGKRSYL  DTAEKDWRAKLSCLVPSKAEKTKDDNTDAVAVR  DLDSIDVPVKSVPVRRRTVGNIRILKEDNAVSLCI  DKSLGNLTKEGTDPEDTAEHDNLKDSNLDG  MEEMVETNGFGHPTAEPSTSDTLKPNLSLFD  AERNGAEEGTSSGKSSLLDLDQETPSCSSAYRV  GDCDGNIEPINGSVQPNQKKLDEEQSLLSRRT  PDFNGCIDREPLAMEQTLVKKKFPACIDEVDGG  EGMKLEDLRAKEDVWVDFSPENDVOREIIFQO  QLSHNAVALKNHSDQLVCKVEKLPNEMDSVVRKR  RWDSVHVNQYIELKEARKQGRKERHKEAQAAY  LAAATAAAAASRYSSLLRKDAVDESSQEQENLQK  VEIGRVGRGSSLSLVPPRAKVSNPVSLDVPKASDGG</p>

10

20

30

40

【 1 3 3 7】

【表 1 4 - 3 3 8】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
	分析に用いたバイトとの相関係数				
838	sp_121220_xnqe -0.336244	#N/A	-0.358383	VPS9A; FUNCTIONS IN: Rho guanyl-nucleotide exchange factor activity; INVOLVED IN: cell plate assembly, transport, cell wall biogenesis, post-embryonic root development, embryo development; LOCATED IN: cytosol, nucleus, membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Vacuolar sorting protein 9, subgroup (InterPro:IPR013995), Vacuolar sorting protein 9 (InterPro:IPR003123); BEST Arabidopsis thaliana protein match is: Vacuolar sorting protein 9 (VPS9) domain (TAIR:AT5G09320.1); Has 1001 Blast hits to 992 proteins in 208 species: Archae - 0; Bacteria - 4; Metazoa - 518; Fungi - 219; Plants - 81; Viruses - 0; Other Eukaryotes - 179 (source: NCBI BLINK).	MENSEMFSSSTVQHTWHDFLDRMRQPSAAEFV KAIKSFIVSFNPPDPEKDSAAVQKFFANMEVA FRAHPLVWTGSEEELESAGELEKYYMTKLFTR VFA5FPDDVKTDFQLSEKIALVQQFIRPENLDMP TFQNESSWLLAQKELQKINMYKAPREKLACLINC CKVITNLLHNAAVTANENTPGADELPLVLYVYTIKA NPPQLHLNLQYIQRYYRRESRLVSETAYYFTNILSA VSFISNIDAQISIMDENEFERNMESAOALLSGLTT DFQITPNAVDNNTSSSKPPMEEFQQTLPSNEND STKQPNPNPTOPEPNPMTKNMSISDLEEKGAT TLKEDRSINNVFOEFPYFVSVGDLTLRDVGDLL NNYKQIVKYVCLAKGLGVTFPSLSTSSVSQVDDQ SLVEKTTGSEPKRETEIETDEQLCRKQDSSKT DTGISEKSPKESDEVLSLKEDDSEGNQVDIA ATTTKESQDDVQ
839	sp_121450_ixur #N/A	#N/A	-0.330535	Cytochrome b-c1 complex, subunit 8 protein; CONTAINS InterPro DOMAIN/s: Cytochrome b-c1 complex, subunit 8 (InterPro:IPR020101); BEST Arabidopsis thaliana protein match is: Cytochrome b-c1 complex, subunit 8 protein (TAIR:AT3G10860.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MGKIPVKMKMVQYALSPYQOKVMPGLWKDLPG KLSHKVTEYISTITLLGLPLGVYVYVQWVYLEKE KLSHRF
840	sp_121460_jegf #N/A	-0.362181	-0.366362	RNA-binding (RRM/RBD/RNP motifs) family protein with retrovirus zinc finger-like domain; FUNCTIONS IN: RNA binding, nucleotide binding, zinc ion binding, nucleic acid binding; INVOLVED IN: biological process unknown; LOCATED IN: cellular component unknown; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: RNA recognition motif, RNP-1 (InterPro:IPR000504), Nucleotide-binding, alpha-beta plait (InterPro:IPR012677), Zinc finger, CCHC-type (InterPro:IPR001878); BEST Arabidopsis thaliana protein match is: RNA-binding (RRM/RBD/RNP motifs) family protein with retrovirus zinc finger-like domain (TAIR:AT5G04280.1); Has 996832 Blast hits to 989721 proteins in 35749 species: Archae - 21265; Bacteria - 589617; Metazoa - 208621; Fungi - 26263; Plants - 58603; Viruses - 68007; Other Eukaryotes - 24436 (source: NCBI BLINK).	MAARDGYRIFVGG-LGWNTSQRHLEDAFVRYGKII DSLVMVDRETRPRGFGFITFADRRAMEDAIPS MHGRELDGRVIVNKAEPFMGAEDPYDAYGGD RRSGGRESYRGGGERPERSDECFCGGRAGHW ARDCPNSGGGRYSSQSEFGRSGLGGDRDYDRH YDRAGGGRYDSRDRLESRDVYVSRDRYS'TERY PPAGDRFGSDFRMDRYPQNGYVDRDRDYERDGG IRGGDRYAAGGPTRYDKGSYRDRVHPYDAPRRT SRR

10

20

30

40

【 1 3 3 8】

【表 1 4 - 3 3 9】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
842 sp_121550_jwro	-0.461317	#N/A	-0.373958	<p>casein kinase I-like 10 (ck110); FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: cellular component unknown; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271), BEST Arabidopsis thaliana protein match is: casein kinase I (TAIR:AT4G14340.1); Has 56550 Blast hits to 56234 proteins in 2232 species: Archaea - 39; Bacteria - 9117; Metazoa - 21673; Fungi - 6267; Plants - 8904; Viruses - 331; Other Eukaryotes - 10319 (source: NCBI BLINK).</p> <p>Transcription factor TFIIC, tau55-related protein; CONTAINS InterPro DOMAIN/s: Transcription factor TFIIC, tau55-related (InterPro:IPR019481); BEST Arabidopsis thaliana protein match is: Transcription factor TFIIC, tau55-related protein (TAIR:AT1G80745.1); Has 31 Blast hits to 31 proteins in 10 species: Archaea - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 31; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).</p>	<p>MDHVIGGFKLGRKIGSGSFGELYLGINVTSEE VAYKMEAKTKHPQLHYESKLYMLLOGGTGVPH LKWYGVGEYVNVAVIDLLGPSLEDLIFYCTRKFS LKTVMADQMINRVEYMHSRGFLHRDVKPDNIFL MGLGRKANQVYVDFGLAKKYRDLQTHKHIPYRE NKNLTGTARYASVNTHLGVEQRRRDLSDLESGLYV LMYFLRGLSPWQGLKAGTKKOKYDKISEKMLT PIEVLKKNYSEFTSYFYHCRSLRFDKDPDYSYL KRLFTKALNPGPSAERAERTSGQDIRDRFSGAV EGFARRNFGTGGQDQSRHRSADDPSSKDVQ ADDRGRISRNSTSKRAIASSSRPSSSGEPSSDS RSSRLMSTSGRLSTQRVQPGFESKSTLARVS AARFGRDDALRSEDFLSIGAGRRK</p>
842 sp_121670_qnej	-0.936341	-0.913723	-0.938865	<p>casein kinase I-like 10 (ck110); FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: cellular component unknown; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271), BEST Arabidopsis thaliana protein match is: casein kinase I (TAIR:AT4G14340.1); Has 56550 Blast hits to 56234 proteins in 2232 species: Archaea - 39; Bacteria - 9117; Metazoa - 21673; Fungi - 6267; Plants - 8904; Viruses - 331; Other Eukaryotes - 10319 (source: NCBI BLINK).</p> <p>Transcription factor TFIIC, tau55-related protein; CONTAINS InterPro DOMAIN/s: Transcription factor TFIIC, tau55-related (InterPro:IPR019481); BEST Arabidopsis thaliana protein match is: Transcription factor TFIIC, tau55-related protein (TAIR:AT1G80745.1); Has 31 Blast hits to 31 proteins in 10 species: Archaea - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 31; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).</p>	<p>MELDVTIDHNQDEYVLLNLDVAVCGQVNLINLNE PYVLTGLDTANPVLTIQDKLKLIGEYEETIGTCLIF SDKVLYEETSGSSERNLFGACIVDPDQGPSKQIK PVTSLLHKILKFRLLLETDMPQKSPDLPROEVN</p>
843 sp_121690_zyjp	-0.332423	-0.952613	-0.387452	<p>ankyrin repeat family protein; FUNCTIONS IN: methyltransferase activity; INVOLVED IN: metabolic process; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ankyrin repeat-containing domain (InterPro:IPR020683), Arginine N-methyltransferase 2 (InterPro:IPR017409), Methyltransferase type 11 (InterPro:IPR013216), Ankyrin repeat (InterPro:IPR021110); Has 14669 Blast hits to 7820 proteins in 738 species: Archaea - 65; Bacteria - 1760; Metazoa - 7798; Fungi - 1170; Plants - 439; Viruses - 38; Other Eukaryotes - 3399 (source: NCBI BLINK).</p>	<p>MEEKGEKLEAARIGDVAQVTSIDAGYDVTYFD ADGFTPLMYAAKLGHEHVEISKLEAGAPWNAALSP SSLSAGDFAMQNGHQDAFDLLNTGVQAEILIGT IARKEKENSDDGGYLYDRVSESEKLMDDADSK AVNMAWKEKPLMESHAKAVCSGGGHVNVGFG MGLVDMIAIQYNPVSHITIEAHPYVYDRMLRNG WGEKENVKILFGRWQDVLQPLGSDYDGIFFDTYG EYVEMREFHQHLPKLLKPGGIYFFNGLCGDMA FFHVYCNVLSLELENLGYTTQLIPLPVKDCLAEE VWEGVKQYWLQDLYLPLVPCQYTDDESE</p>

10

20

30

40

【 1 3 3 9】

【表 1 4 - 3 4 0】

I	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 alpha/beta-Hydrolases superfamily protein; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF676, hydrolase-like (InterPro:IPR007751); BEST Arabidopsis thaliana protein match is: alpha/beta-Hydrolases superfamily protein (TAIR:AT1G10040.1).	aa seq
844	sp_121800_mdna	-0.922332	#N/A	#N/A		MLLQTRVDTPTVFLSAADWKFGVEQFKMLPD KVHVHRGKETGVLVREGTRFTSVRAMWPLEFPS MTEFFPDEQLLSLRIDTPWYADMVNYLVTRTIP SYIV
845	sp_121920_kswo	#N/A	-0.351311	#N/A	uncharacterized protein	
846	sp_122360_dtaq	#N/A	-0.393274	-0.862445	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 9 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF2365 (InterPro:IPR019314).	MEEGKKEETLSIENSIMKEKETMEASKNFNQDD LFKALEVVERDLSAIAQSYTALFASLRALSSEVTS TSVDHIQCFSDAAGKLOES
847	sp_122450_neft	-0.922304	-0.325738	-0.922313	LATERAL ROOT STIMULATOR 1 (LRS1); FUNCTIONS IN: nucleotide binding; INVOLVED IN: lateral root formation; LOCATED IN: CUL4 RING ubiquitin ligase complex; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: WD40 repeat 2 (InterPro:IPR019782), Protein of unknown function DUF3337 (InterPro:IPR021772), WD40 repeat, conserved site (InterPro:IPR019775), WD40 repeat (InterPro:IPR01680), G-protein beta WD-40 repeat, region (InterPro:IPR020472), WD40 repeat-like-containing domain (InterPro:IPR011046), WD40-repeat-containing domain (InterPro:IPR017986), WD40/YVTN repeat-like-containing domain (InterPro:IPR015943), WD40 repeat, subgroup (InterPro:IPR019781), BEST Arabidopsis thaliana protein match is: Coatomer, alpha subunit (TAIR:AT1G62020.1); Has 59803 Blast hits to 26869 proteins in 729 species: Archaea - 50; Bacteria - 7042; Metazoa - 23770; Fungi - 14096; Plants - 7202; Viruses - 3; Other Eukaryotes - 7640 (source: NCBI BLINK).	MHRVGSAGNTSGSNRPRKEKRLTYVLNASDDTK HSAGVNC.LAV.LKSNDMCNSEY.LFTGSRDGIKLR WAIEEAAATCSATFESHVDVWVNDVAVLGDHTLV SCSSDLSLTKWNCMSNGVCTKLRHHSDYVTCI AAAGKSNIVASEAALAPVSKNSDSTEEEAASNGI NGGNSLPITSLRTINSSNSLHSTPESHGYVPTPA KGHKESYALAMNDSGTILVSGGTEKVVVRWDF RSGSKTKLRGHTDNVALLDSTGRYCLSGSS DSMIRLWDLGQQRVHSHYAVHTDSVWALASTTS FSHVYSGGRDQSLY.LD.LSTRESILLCTEDYPIIQ LALDDSEIWAASVTDSSVHRWPAEGHNPQKMFQ RGGFLAGNLSFSRARASLEGSTPIPVYREPTLK VPGIPAIVQHEILNRRRVLTKDAGSVKLVWEITR GTVIDYGEVSFEKKEKELFEMVSIWPAWFTVDTR LGSLSIHLDTPOCFSAEMYSADLADIAGKPEDDKV NLARETLKGLLAHW.LAKRKQRLGSSQSANGEVQ SVKDISTRSMTHSRIEVDGSSSESDSMIYPPFEFSA VYPPSIITEGSSQGWARKKVTGMDGTEDEKDLP WVWMDCVLNRL.PPRENTKCSFYLHPYEGSTIQ ILTQGLSAPRILRLKVINVYVEKMW.LDKSMEGG ATGDGTFTAGLAGGFLQHSITTVGDGSLRSLRIP WQRPKPAIEILCNQV.L.TAEMSLATVRYVWVKPP EDLVLYHRVMQGR

10

20

30

40

【 1 3 4 0】



【表 1 4 - 3 4 2】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
852	sp_126050_udqo	0.885523	0.94573028	0.940221	<p>Small nuclear ribonucleoprotein family protein;                      FUNCTIONS IN: molecular function unknown;                      LOCATED IN: nucleolus, nucleus, small nucleolar                      ribonucleoprotein complex; EXPRESSED IN: 22 plant                      structures; EXPRESSED DURING: 13 growth stages;                      CONTAINS InterPro DOMAINs: Like-Sm                      ribonucleoprotein (LSM) domain                      (InterPro:IPR001163), Like-Sm ribonucleoprotein                      (LSM) domain, eukaryotic/archaea-type                      (InterPro:IPR006649), Like-Sm ribonucleoprotein                      (LSM)-related domain (InterPro:IPR010920); BEST                      Arabidopsis thaliana protein match is: probable small                      nuclear ribonucleoprotein G (TAIR:AT2G23930.1);                      Has 1287 Blast hits to 1287 proteins in 269 species:                      Archae - 120; Bacteria - 0; Metazoa - 486; Fungi -                      285; Plants - 213; Viruses - 0; Other Eukaryotes - 183                      (source: NCBI/BLINK).</p>	MVVGTLRGFDQFMNLLVDNITVEVNGDNTSDIGM VVIRGNSWVTVEALEPVSKNMMQ
853	sp_127020_qwyy	#N/A	#N/A	0.6523333	<p>ATB2; FUNCTIONS IN: oxidoreductase activity;                      INVOLVED IN: response to calcium ion;                      EXPRESSED IN: 23 plant structures; EXPRESSED                      DURING: 14 growth stages; CONTAINS InterPro                      DOMAINs: Aldo/keto reductase                      (InterPro:IPR001395), Aldo/keto reductase subgroup                      (InterPro:IPR020471); BEST Arabidopsis thaliana                      protein match is: NAD(P)-linked oxidoreductase                      superfamily protein (TAIR:AT1G60730.1); Has 30719                      Blast hits to 30695 proteins in 2595 species: Archae -                      650; Bacteria - 20319; Metazoa - 1822; Fungi - 2308;                      Plants - 1286; Viruses - 0; Other Eukaryotes - 4334                      (source: NCBI/BLINK).</p>	MAGVRRIKLGSQGLEVSAGLGCMGMSAFYGP PKAAEEDMPLIQYAISGGITFLDTSDFYGPHTINELL LGKASTGGWREKVELATKFGVKFTDGGKKEILGD PAYVRAACEGSLKRLDIDCIDLYQHRRIDTKLPIEV TMGELKLVVEEGKIKYGVSEASATIRRAHAVHP ITAVQMEWSLWTRDVEEDIVPTCRELGIGIVAYS PLGKGFSSGPKMTENFTKEDFRQHLPRFKPENI EHNKHIFERVSEIATRKGCTPGQLALAWSITKEM MCVPSLEPLRLRTRSTRLEPCL

10

20

30

40

【 1 3 4 2】

【表 1 4 - 3 4 3】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MTLSDMQLDQVVTGSPFDSTFTSQDFRNRVSSQ HWSVLRGCIHRLIMFCKQLVQLPEAFDVKTAEL SFCRKISLSNLIKLLRNLARSVPYVEVNSELQSI SVFADMLPFLFRPDLDFISSNIASENYEGVNLIL QEFHLVQLLSCGRSVFQNVRACLMASILDLLDS SVWRYNKSANNKPPPLVYFPCVLYFVKLIGDIK NQGDEKQVEKDLGADLSFDIADPKARCVCVHRS ENILVKSYTAEELLEIFIPKSIHWVDNLLHLLYLLH TEGVKLOPKLEKSCSGAKSSICDMENGTCHDD EALFGDLFSEGGRSVGSNDGYEQAPSNPLTGF YNTIEAAADLSVFLKSCIFSPWCAPYYEDAACK LNFRHIDSLMLGCQICYSEGFDGTTVYQEQRR IEHQEICFEMLHGLKRHAFSDLLEEHVLQILKY ENGSYYNDQTLVLAQTLVCRVYGSAGSNLRKK VCQQFVNFIVEKEKDVRSNCSILSELLSLPNVFF MEIILLTFHLSSEKVALAELIFSNIKTASSLOVS SSTQLSCWALVSRLLVFRHMLNMRACPTSL RDFRSKLNAAARLVGSHLHLNDFVTSWSSVWD NVMSLFIKEEPLLNLLNHLVDFESVPVRLFSSDA ANVNSLSWNDIYTVFSPILELWKGKKAATIEDLV LERYMFLCWDIPMAGLSPNHMSSAWANCNLD PSNVEQFILFSHLLGHCDLWKSVDVVDVIVKFL OHLCSVMPADIHEFGWDMRNGQWLSLVLSF LHTFGFSYCVKNSVPGANFVWTFENAFDAEFVSL AEGSISMILKTKVPIVILKSLSSLLAKYRTYQRAF LFTISENQSHAQIFSPILLKHTFEDECMDKDELVK NGLVHAGLESVDDLLPKIDNVNKGSGSHLLGVISS ESMLHGFPPTHPTSSGALVSSILSRGICLLDLIFK IRGTRGDISVETDVLHEILDSVMTKYDRVLKNLS GKCRDVCRCGTGAEGPDYSLLFLKIQIEGFLSDIN VRGNVHVKVEQLVVKAVDMDSLRRDRLGIDVF RFYLGAEVDPQHVLDLYGSGQRGDLVLLDLSLDSVC
854	sp_127340_gjsu	-0.954355	-0.375501	-0.953734	BIG (BIG); FUNCTIONS IN: ubiquitin-protein ligase activity, binding, zinc ion binding; INVOLVED IN: response to auxin stimulus, indeterminate inflorescence morphogenesis, photomorphogenesis, root development, auxin polar transport; LOCATED IN: membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN's: WD40 repeat-like-containing domain (InterPro:IPR011046), Zinc finger, N-recognition, metazoa (InterPro:IPR013993), Armadillo-type fold (InterPro:IPR016024), Zinc finger, ZZ-type (InterPro:IPR000433), Zinc finger, N-recognition (InterPro:IPR003126); Has 1182 Blast hits to 997 proteins in 123 species: Archaea - 0; Bacteria - 4; Metazoa - 768; Fungi - 44; Plants - 189; Viruses - 0; Other Eukaryotes - 177 (source: NCBI BLINK).	

10

20

30

40

【 1 3 4 3】

【表 1 4 - 3 4 4】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたペアとの相関係数 sp_107620_kqnh sp_107660_klqg (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
855	sp_127930_wgty	-0.870671	-0.865252	-0.876705	<p>vernalization5VIN3-like (VEL-1); CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF1423, plant (InterPro:IPR004082), Fibronectin, type III-like fold (InterPro:IPR008957), Fibronectin, type III (InterPro:IPR003961); BEST Arabidopsis thaliana protein match is: Fibronectin type III domain-containing protein (TAIR:AT5G57380.1); Has 270 Blast hits to 250 proteins in 39 species: Archae - 0; Bacteria - 2; Metazoa - 8; Fungi - 14; Plants - 239; Viruses - 0; Other Eukaryotes - 7 (source: NCBI BLink).</p>	MEERREIVYELAKWISHGAAEMLQSWSRDILQIL CSEMGKERKYTLTKLIVHEHLLKVVSEKNSWNH DNEVYRGOQSSPITQKVKPKRKPENPSHLVLS ATNISSCNSDSLVTATYCKNSACKASLNKEDVF CKFRCTCCICQKYDDNKDPSLWLTCSTEPPFVGD SCNMSCHLECAIKHDRSGIARNGFVSQLDGSFYC ISCGKMNDLMSCWRKQIJAKDTRRVDILCYRISL SQKLLAGTKKYLKACDIVDDAIKLLAAEVGALTGV PVKMGGRGIVNRLLSSGQEVORLCSALELFEFSTY HQTPHLGNVAVSRTIAPSIIRVESINATAVTVL GPDNSSLSEKNVGYTLWHRKYGDRKKEOPTC TLFPQPQTRFIVSGLSPATEYIFKAIISFDNTELSSS EVRLTNIISTLERSOSPITTCSTLSNPSSVEDET NNNAPCSNLNENKAENYFNCSDNDQNPSCASDE ENNATGKDSDDLSDTKVVEIKHSHSDGIGSKDTIK TNTPCVRPETSLPITPSRQESSKEGTVKTGPYKS RNKDQDSRLDKRKAERENGECVGDGNSNSGGS NIEYIKMIRWLECEGHIDKSRQKFLTWYSIRAT POETRIVKAFVDLIDDPASLSEQLQDTFEEIISSK RSSAVPPGFCLKLWH

10

20

30

40

【 1 3 4 4】

【表 1 4 - 3 4 5】

A	B	C	D	E	F
1					
2	分析に用いたペイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
856	sp_127950_nmud -0.3027339	#N/A	#N/A	<p>GFY domain-containing protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 25 plant_structures; EXPRESSED DURING: 14 growth_stages; CONTAINS InterPro DOMAIN/s: GFY (InterPro:IPR003169); BEST Arabidopsis thaliana protein match is: GFY domain-containing protein (TAIR:AT1G24300.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archaea - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).</p>	<p>MSSDSRRPILTPFNQIAKDKITGPDNIPFESQW LLPKGETKTGMSLGENRPSPTYGYGDRSDVTK SPNGEEMHDNOKKRDVFRPTFLDAETGRRDR WRDEERTNFSARKDRWREGEKEIGDGRKTR WVDNTPKNFGEARRAPSKWADPSKENNYDQ RRESKWNRWGPDDKEADKWTADAGKHGDMPO DKGSQLANHGKDEREGDNHRPWRPSSLQNRV KAETLHHHPQMTSKEGPTFGHGRGRGNATFS VGRGRANFAGSISNYSVPHSGFNDRGDNHG EFPSPFSYNRTKLLDIYRLTEIRSSNKILDGLKQVS SLTQEELPLALCAPTEESAIMKGDIDKGDIISS GAPQNSKEGSTGRSPNEYAQSRRTKHGRDDV SLPPDEDNDKVDHAKGTYKISELSPREKQMINV YGSNVIMDGNENPQYAEKYRVAARDDESSLKK GDEVSDNRDPGKIGRYSQSGTWRSPSLEEHA STHSRDWFNSPAGGGIRSKSSDISWQAQPKDNS SVLPDFNAHNLKQVGEDSAAARKHAVTVVDR EQETRKLHTSPDMLLYKDPQGGQGFSGS DIIGWFEAGYFGIDLVRPVPNDVAFASLGDV MPHLRAKAPPPPGFNATKQSEVDPVSSRSSSSG FGKVLPGFSGVDLLGNEHRHTHGSDAENKFL ESLMAGEITTSHLERFASSEGLORFYQGNSSGM SPSVAESGDALNLAQRMLNERQKSLTQSPYVW ASREGAPVDFVFNSTIPQANFLSTYVDIPIRQS LPCNVEMMSIQGLSDRSSGGVNNGLGGWPKF SVQGGIDPFKENVDLHSGQNFQPMAYGMQQHK QQPQNHPSMTNLLAQTLDNPAIQSDKLLSAGLTH DPOALNLLQQYLMQLHSQOAVPGPQLSILDKFL LLKQQQKQEEQQQLRQQQQQLLSQVLSQEQQ LVGRFGDFAFGQLQTSASPIGNAPIEQSAPQLSP SLFQANSQVPSVSSMPDNRNYSIANMPPKFSQGL GHQVPADGSSLHLPQHMLGLTACKNWPSPVSMSTHA</p>

10

20

30

40

【 1 3 4 5】

【表 1 4 - 3 4 6】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
857	sp_128300_onxz	-0.388354%	#N/A	#N/A	HEAT SHOCK PROTEIN 81.4 (Hsp81.4); FUNCTIONS IN: unfolded protein binding, ATP binding; INVOLVED IN: protein folding, response to stress; LOCATED IN: cytosol, apoplast, cell wall, nucleus, plasma membrane; EXPRESSED IN: male gametophyte, cultured cell, pollen tube; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage; CONTAINS InterPro DOMAINs: Chaperone protein hspG (InterPro:IPR001404), Heat shock protein Hsp90, conserved site (InterPro:IPR019805), Heat shock protein Hsp90, C-terminal (InterPro:IPR020576), Heat shock protein Hsp90, N-terminal (InterPro:IPR020575), Ribosomal protein S5 domain 2-type fold (InterPro:IPR020568), ATPase-like, ATP-binding domain (InterPro:IPR003594); BEST Arabidopsis thaliana protein match is: heat shock protein 81-3 (TAIR:AT5G56010.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MVEAETFAFAEINQLLSLINTFYSNKEIFLRELIS NSSDALDKIRFESLTDKSKLDAQPELFHIVPDKT NNSLSIDSGIMTKADLVNINLGTIARSGTKEFME ALSAGADYSMIGQFGVGFYSAYLVAEKVITTKH NDDEQYVWESQAGGSFTVTRDPGESIGRGTKIT LFLKEDQLEYLEERRLKDLIKHSSEFISYPSISLWTE KTIEKISDDEDEEEKDEEGKVEDVDEEKEKEE KIKKKIKIKEYSNEWTVVVKQKPIWMRKPEEITKEE YAAFYKSLTNDWEEHLAVKHFSVGGQLEFKAVLF VPKRPFDLFDTKKPNNIKLYVRRVFMINDCEDL IPEYLSFYKGVDSLDLPNISREMLQQNKILKIVIR KNLVKCKVELFEEIAENKEDYKFFYEAFSKNLKLG IHEDSTNRTKFAELLRYHSTKSGEEMTSKLDYVT RMKESQSEIYYITGESKKAVENTSPFELEKLRKKG EVLVYMDAIDEYAVGQKKEFEGKLYSATKEGLK LEESEDEKKAEEQKKEFEGKVKIKDVLDGKVE KVVSRRVDSFCCCLVTGEYGSANMERIMKAQ ALRDSSMAGYMSKKTMEINPENINDELKPKRA DADKNDKSVKDLVLLLFETALLTSGFSLDEPNTF GNRIHRMLKLGSLIDDDDEGEADVDMPALEEVDTE GSKMEEVD

10

20

30

40

【 1 3 4 6】

【表 1 4 - 3 4 7】

A	B	C	D	E	F
1	sp_107620_kqnh	分析に用いたバイトとの相関係数	sp_107670_ptqx		
2	sp_107660_ktqg	(SOAP2)	(SOAP2-like)	注釈	aa sec
858	sp_128340_jksw	-0.925707	-0.88943	<p>global transcription factor group A2 (GTA2);            FUNCTIONS IN: transcription elongation regulator            activity, structural constituent of ribosome, sequence-            specific DNA binding transcription factor activity;            INVOLVED IN: translation, regulation of transcription            from RNA polymerase II promoter, positive regulation            of RNA elongation from RNA polymerase II promoter;            LOCATED IN: ribosome, intracellular; EXPRESSED            IN: guard cell; CONTAINS InterPro DOMAIN/s:            Translation protein SH3-like (InterPro:IPR008391),            Transcription elongation factor Spt5            (InterPro:IPR017071), Transcription antitermination            protein, NusG, N-terminal (InterPro:IPR006645),            KOW (InterPro:IPR005824), Ribosomal protein            L24/L26, conserved site (InterPro:IPR005825),            Transcription elongation factor Spt5, NGN domain            (InterPro:IPR005100); BEST Arabidopsis thaliana            protein match is: Transcription elongation factor Spt5            (TAIR:AT2G34210.1); Has 14630 Blast hits to 9620            proteins in 607 species: Archae - 121; Bacteria - 647;            Metazoa - 6069; Fungi - 2592; Plants - 1061; Viruses -            307; Other Eukaryotes - 3833 (source: NCBI BLINK).</p>	<p>MGRDDYEDDDIEPEDGDEYGEDLYEEPEEEE            DEDDDYGGRRGGGGKGGKSVKQRQRSDFIDD            VAEEDDDDEEVEYVARGRRHKSRGGGFIDDI            AAEVDDDDDEEEDGEDDFIVDTTDDVVEDERRM            HRRPLLSREEEQEEDVEDLERRIQERYARSSQSE            YDEETTDVEQQALLPSVDPKLMWVKCAIGRER            EVAVGLMQKCIDKGSSELQIRSVWALDHLKNFVYV            EADKEAHVREAVKGMRIYPSKILLVPKEMTDVL            SVETKAIDLSRDTWVRMKTGTGKDLAKVLDDVD            NVRQRVTKLIPRIDLQVLANKLEGRDVAKKKAF            VPPFRFMTEEARDLHIRVERRRDSMTGDYENI            NGMMFKDGLIKSVSISISTQNIQPTFDELEKFR            HPGEGDDIGS.LTLFATRKKGFMKGDVAIVVK            GDNLKLGWVEKVEDGNVHIRPNEKGLPKTLAIN            EKDLCKFFEPGNHVKVISGSHGEGATGMVVKVDF            HVLIIISDTTKEDIRVFADHVESSEVTTGVTKIGG            FELHDLVLDNATFGVIIRVESEAFQVKGVPDRP            EVLVRLREIRSKIDKKLFAQDRSKNPVAVKDVVK            VLEGPCPKQGGPVEHYRGLVLFIDRRHLEHAGFI            CAKSQSLVVGSGRGNREKMGDPLSYLRTPQRI            PQSPGRPPRGQSDFGGRGGRGGGRGHDSLIG            TTKIRLGPYKGRVVKDKGTTVRVELESQMKE            VTVERSISDNVTVTPYRDTSTRYGMGSETPMH            PSRTPMHPVMTMRDARATPIHDGMRTPMRDR            AWPYAPMSPARENWEDANPASWSSPQYRP            GSPSPRTYDAPTGVGWASTPFGGNYNDPPTPR            GNAPSPYLPTSGTQPTTSSVSYLPTGPGGQP            TTPFGGGGLDMMSPMIGGENEGLWLLPDILNVVVK            SGDDSMVGVVQEVLPDGSRCRVALGTSNGEVIT            ALPNEIDLIPPRKADKIKIMGGAQRGFTGKLGIVD            GADGIVKVEDSLDVKILDMGILAKLA</p>

10

20

30

40

【 1 3 4 7】

【表 1 4 - 3 4 8】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	<p>Pre-mRNA-splicing factor 3; CONTAINS InterPro DOMAIN/s: Pre-mRNA-splicing factor 3 (InterPro:IPR013881); BEST Arabidopsis thaliana protein match is: Pre-mRNA-splicing factor 3 (TAIR:AT3G55930.1); Has 22923 Blast hits to 12760 proteins in 707 species: Archae - 18; Bacteria - 1085; Metazoa - 11525; Fungi - 3053; Plants - 1999; Viruses - 93; Other Eukaryotes - 5150 (source: NCBI/BLINK).</p> <p>homeobox 51 (HB51); FUNCTIONS IN: sequence-specific DNA binding, DNA binding, sequence-specific DNA binding transcription factor activity; INVOLVED IN: in 6 processes; LOCATED IN: nucleus; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: Homeobox, conserved site (InterPro:IPR017970), Homeobox (InterPro:IPR001356), Homeodomain-like (InterPro:IPR009057), Helix-turn-helix motif, lambda-like repressor (InterPro:IPR000047), Homeodomain-related (InterPro:IPR012287); BEST Arabidopsis thaliana protein match is: homeobox protein 22 (TAIR:AT2G36610.1); Has 11020 Blast hits to 10992 proteins in 578 species: Archae - 2; Bacteria - 0; Metazoa - 8622; Fungi - 156; Plants - 2028; Viruses - 0; Other Eukaryotes - 212 (source: NCBI/BLINK).</p>	<p>MEKATEKSSRKSRREERTEHKKRRHHRHRSRHR  HRYEDEDDEHRRSRREKSHDDEHHRHHRHKK  SREYSVDEQDEEDARDSDRRREGSRDQKDKD  RKRKRREVSEERIDDKRAKVSDDVKKERRRFD  DDNGRIVKDERKEKRRFEDDGGNGRVKIEEAKTK  EDGGRGDSGEIVDNNNGRRDRRRRFFEDGDVK  KESVAEPTGRSGSFPNDGAEKSLNVSSRFSSET  FFANNHTATKYSSISTTHENKGVNITGSHAVTGG  SSTDGTPSTAGKTGNISLDALAKAKKALQWQKEL  AAKMKIPLLNRGTSSSNAYNAQVGSKEEIKPPS  AAGIQQTPSTSTFPFAASAKTASMDPAKYEAV  KRAQELAAKMGFRQDTEFAPLNLNFPGLPLDIT  ALQKPAKAPVLRDLDAHGREVDEEDGNVVISKPS  LSTLVNINKQKDAFOIHKPELEVDPKNPHFD  PRMGIDKTKLLRPKRMNLFVEEGKWSKDAEIK  LKSQFGEAOAKELKAKOAGLARAKAEPNINPLI  EVSERVITKKEKESPEVEWWDVPLLRSGDYAD  IADGSSIEDRVKMEKITTYLEHPRIEPPAEPAPQ  POPPLKLTKEOKKLTORRLAKEKDRQEMIFROG  LIEPPKVKYKMSNIMRV/LASEATDDPTKMEMEIVR  TAAAEERQOAHVDRIARPLTPAERREKTKERKLF  DPNLTETIVSYKINDFSHPQNRKVDINAGENRL  TGCVAISEGFSVAVVVEGGSKSIRYKGLMLRRIN  WLAKVLNEDEEEDDEEKPKNCSLWQGSVA  KPSFNRFIHDRCRTEAAARKVFSDAGVGHYWEL  AANFVDDQI</p>
859	sp_128920_jkap	-0.987017	-0.975912	-0.970757		
860	sp_128970_nft	-0.972049	-0.980073	#N/A		<p>MKRRLLTKQLESLESSEFEEERKLDPRKMKLAR  DLALQPFQVAVWFQNRFRARWAKKQLQHLQYDAL  KLDLFELVSKCHKLQEQEVLRLKSIMVMEETKQ  VYSSYTEASCEETVESTSAAAANQYTGRRNSGNS  NGADDIVLN</p>

10

20

30

40

【 1 3 4 8】

【表 1 4 - 3 4 9】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
861	sp_129010_fpak	-0.926337	-0.634236	Ubiquitin-like superfamily protein; CONTAINS InterPro DOMAIN/s: Ubiquitin subgroup (InterPro:IPR019956), Ubiquitin (InterPro:IPR00626), Ubiquitin supergroup (InterPro:IPR019955); BEST Arabidopsis thaliana protein match is: Ubiquitin-like superfamily protein (TAIR:AT5G42220.1); Has 11489 Blast hits to 5746 proteins in 742 species: Archaea - 0; Bacteria - 401; Metazoa - 4954; Fungi - 1390; Plants - 2430; Viruses - 186; Other Eukaryotes - 2128 (source: NCBI BLINK).	MAEQQSTEGASTSSVPVADSNMELNIKTLDSQ MYNFHVDKNMSVSSFKGMVADKVGVPYGOORLI FRGKVLKDEHLLSEYHVENGDTLHLVERQPNQP QTSSGTSTASGNIGTQGMGDNAGPRRIGQI SHSVLGLNVDGQEGGVPLSRVIGAVLNSL GVNGQTTNAAFGAPNTAPANSSASVAQANDTEQ GRGSGGNQAQFGQFLSQPFHYPFLAGISVPPP SLNMPISDSLNTLSEFMSRMEGVLSONGYQSP PPSGTADQPGTELPNFRGMATPEALSIVLRHAE RLLSGPAVALSHIAGRMEQEQAGSTDVSVRSQIQ SESVQVGLVMQHLGALFLELGRITLMLQMRSP AEALVAGHAVYISPSGPNPLMVQFPFHQTSPLF NAPAVSATNTGTSPVPGIGSAPRNINIIHAGTPL APLASAVGARTAGGEGTQGEHGGVTAASDSGG RVISVRNVAASVFSHTATIAVSPVTTASQPVGV VSTSQPPDAASLSSVISHVNSQIFRSLLENMRGDN QNPSGQQDNPVIPNSSAGSGGGREAHESHTSSL PROOAEGLDKTDRRENIQSDSSOSNEEESPSCMS GDPAAKSIGPKVEVPLGLGGGLQPKRRGRSQM QPVAGEAGPSSSFDQSQOARISGQQLQSLASR NTAGRSSLINPGSGQTGGIEQSVAGTQVTOG SDGOFDITGMSQLLOSPALDGLLSGVAQOAGV GSPNVLNMLQQFTQNPAMRNTVNLQAGQVDR EEFRNMFAGMGTDGTDGTFDSSMICOMMPVVS QVLGGGSIQNEQLRNMPQLMPQSDDRSIEDSS LGVVSOEYVRRIALHNSSEDIFRSVQVQNAVENTE NTETATREFIDDLCSNEDLIDEFVEMLRCDLRRRL RDNFV
862	sp_129111z4911	#N/A	#N/A	uncharacterized protein	#N/A
863	sp_129120_exgt	-0.863244	#N/A	ubiquitin conjugating enzyme 8 (UBC8); CONTAINS InterPro DOMAIN/s: Ubiquitin-conjugating enzyme/RWD-like (InterPro:IPR016135), Ubiquitin-conjugating enzyme, E2 (InterPro:IPR000608); BEST Arabidopsis thaliana protein match is: ubiquitin-conjugating enzyme 28 (TAIR:AT1G64230.2).	MASKRILKELKDLQKDPPTSCSAGPVAEDMFHW QATIMGFADSPYAGGVFLVTHFFPDYPRKPKV AFRTKVFHPNINSNGSICLDILKEQWSPALTSKVL LSICSLTDPNPDPLVPEIAHMYKTDFRSKYEATA RSWTQKYAMG

10

20

30

40

【 1 3 4 9】

【表 1 4 - 3 5 0】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
864	sp_129360_krgp	-0.438725	-0.094599	-0.873879	Sec14p-like phosphatidylinositol transfer family protein; FUNCTIONS IN: phosphatidylinositol transporter activity, transporter activity; INVOLVED IN: transport; LOCATED IN: plasma membrane; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Cellular retinaldehyde-binding/triple function, C-terminal (InterPro:IPR001251), Cellular retinaldehyde-binding/triple function, N-terminal (InterPro:IPR008273), Cellular retinaldehyde binding/alpha-tocopherol transport (InterPro:IPR001071), Phosphatidylinositol transfer protein-like, N-terminal (InterPro:IPR011074); BEST Arabidopsis thaliana protein match is: Sec14p-like phosphatidylinositol transfer family protein (TAIR:AT2G21520.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLink).	MQRQCILNKGPKWDFEIAKLIKSGATKLAISSEE RVIADNKMVKNNNDVSTAESGSEVMHLFSSRAPG AAALAKWKLMIPVDEEI

10

20

30

40

【 1 3 5 0】

【表 1 4 - 3 5 1】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
865	sp_129740_zedu	-0.872884	-0.951729	-0.893886	THO5; CONTAINS InterPro DOMAIN/s: THO complex, subunit 5 (InterPro:IPR019163); BEST Arabidopsis thaliana protein match is: THO complex, subunit 5 (TAIR:AT1G45233.2); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	MDVKMEDAAAAAATATATSATEDIIVLS KTAEKVEKSPYDMLKESKISVEEIVAKMLSLKKDD KAKPELRELVTOMLHFVSLROANRNILIEEDRVK AETERVKAPVDSTSLQHLNLMYEKSHYLKAIKAC KDFKSKYPDIQLVTEAEFFSQAPEEFKGIASNET AHNVMKRLNYELFQRKELCKFREKLDQKKKSL ETIANPKKFLSSLPFSLKALKKASLPAGNQLGLH TKKOKQLHAAELPPPLYIYISQFMACKAEAFGENI DLEITGSLKDAQVFRRTQYKDTGTSSMQESSRI DDVVDEEDDQRRRRKPKKVASKENVDTAGY GAHPLRIILHVYDNEVCEPKSARLVTLKFEYLFKL NVVCGVEGSDQGSENNILCNLFFDDKGLLELPQ QSAKLLDDGAEFDERRTMRPKYWAQHLGIDF LPEVSPSSTNEAGKFETGKDAVVSGLSLYRHQ NRVQTVQVRVLRRTKAQLALSEQLDLSTALKWP TLICQSPWAAHEHLCFLHRWSHYGSSLNQDGS LSRVDVEODHDADFANMAGRASSKKEEGIMTED GELPSLTPATVVTDEKSSAYLNQGGVVKHKRM ALISKSIAPVKRTARNQSGFRGYEEDTDLMMDSSES EVDDTTNNDKIECAVDEGNSWVDYGIKEFNLVL TRKRDAGEGNVLDLVKIRISKEYPLRPPFLKLSLL SQYSYSGGNSSPIDHSVWYNELRAMEAEVNVHILK MLPPEENENILSHQVHCLAMIFDLYLNEASQPSQ GRKCKDSDIGSVSGGLKARFRFRDRRRMIPWK DMELTPGYPF

10

20

30

40

【 1 3 5 1】

【表 1 4 - 3 5 2】

1	A	B	C	D	E	F
		分析に用いたベイトとの相関係数				
2	sp_107620_kqmh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)		注釈	aa seq
866	sp_130120_mkrp	-0.555315	-0.842325	-0.346389	secret agent (SEC); CONTAINS InterPro DOMAIN/s: Teiratricopeptide TPR-1 (InterPro:IPR01440), Teiratricopeptide-like helical (InterPro:IPR011990), Teiratricopeptide repeat-containing (InterPro:IPR013026), Teiratricopeptide repeat (InterPro:IPR019734); BEST Arabidopsis thaliana protein match is: Teiratricopeptide repeat (TPR)-like superfamily protein (TAIR:AT3G11540.1); Has 114710 Blast hits to 35863 proteins in 2480 species: Archaea - 3197; Bacteria - 54774; Metazoa - 16661; Fungi - 4580; Plants - 3873; Viruses - 0; Other Eukaryotes - 31625 (source: NCBI BLINK).	MISLRTESRQOCHHHKNSDIQLSKSSFDLKSDVK FNHFKGTESREDVKEDVLLGAAHQLYKAGDYKK ALEHSNTVYERNPLRTDNLILLGAIHYQLHDFDM CIARNEEAIQIDPHFAECFCGNMANAWKEKGNVDL AIRYYLAIELRPNFADAWSNLASAYMRNRLIEA ACCORQALANPLLVDASHNSLGNLMAQGLVQE AYSCYLEAIRIQPTFAVAVSNLAGLFMESGDLNR ALQYYKEAVKLPSPFPDAYLNLGNVYKALGMSQ EAVTCYQHALQARPNYAMAFGNLASYEQQGL DLAILHYKEAINCDSGFLEAYNNGNALKDVGVR DEAIOCYNQGLSLOPHHPQALTNLGNIYMEWMM VAVAATYYKATLSVTSGLSAPFNILAVYKQQGN YADAISCYNEVLRIDPLAADGLVNRGNTYKELGR VSEALQDYINAINIRPTMAEAHANLASAYKDSGHV EAAIKSYKQALELRFDPFEATCNLLHTLQCVCSW ENKDKMFSEVEGIKRQVKMSVLPVSPQPHAIAY PLDSMLALEISRKYAQHCSSVASYALSPFKHPQ PITIKSEGGFQRLRVGVYSSDFGNHPLSHLMGSV FGMHNRENVEVFCYALSPNDGTEWRQRTQSEA EHFVDVSAMSSDMIKLNQDNIQILVNLNGYTKG AFNEIFAMQAPAIQVSYMGFPGTMGASYDYLVLT DEFVSPFRSAHIYSEKLVHLPKCYFVNDYKQKNR DLLEPNCHKPDSYGLPEDKFLFACFNQLYKMD PEFTTWCNLIKRVNSALWLLKFFPAAGESRLRA YAAQQVGGDQIIFDVAAKNEHIRRSALADLFLD TPLONGHTTGDILWAGLPMITLPLEKMATRVAG SLCLATGLGEEEMVHSIEDYEERAVALAINPSKLE SLRNQLKAVRLTCLPLFDTARWVKNLERAFYKMW NLVYCTDKHPQPKVKTENDVDVDFLDR

10

20

30

40

【 1 3 5 2】

【表 1 4 - 3 5 3】

	A	B	C	D	E	F
1						
2	共発現遺伝子	sp_107620_kqnh (SOAP1) 分析に用いたバイトとの相関係数	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
867	sp_130680_kgmh	-0.888780	#N/A	-0.877736	P-loop containing nucleoside triphosphate hydrolases superfamily protein; FUNCTIONS IN: helicase activity, ATP binding, nucleic acid binding, ATP-dependent; helicase activity; LOCATED IN: cellular component unknown; EXPRESSED IN: 19 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAINs: RNA helicase, DEAD-box type, Q motif (InterPro:IPR014014), DNA/RNA helicase, DEAD/DEAH box type, N-terminal (InterPro:IPR011545), DEAD-like helicase, N-terminal (InterPro:IPR014001), DNA/RNA helicase, C-terminal (InterPro:IPR001650), Helicase, superfamily 1/2, ATP-binding domain (InterPro:IPR014021); BEST Arabidopsis thaliana protein match is: DEAF(D/H)-box RNA helicase family protein (TAIR:AT5G60990.1); Has 35248 Blast hits to 34671 proteins in 2939 species: Archaea - 524; Bacteria - 16250; Metazoa - 5646; Fungi - 4238; Plants - 2336; Viruses - 4; Other Eukaryotes - 6248 (source: NCBI BLINK).	MANSSNTTQKEEIEIDDEENQSFEDLSLDPRLLR ALIKKGIKKPTPIQRVAIPILLEGKDVVARAKTGS KIFAYLLPMLQKLLDSSGNKRLAPSSFILVPTRE LCHQVYSEVTSIELCKAQLKAVQLIRGMTAAQMK QTMGPPDILVSTPACVLTCLKAGILQPKAIDSL SMLVLEADALLNYGHGDDLRLSLTAHVPKRCQC LLSATPSADVEKELKLLHNPYLTLPEVGGNKD ELIPNVNQSWYSCSKSDHDKFIYIALLKYEEVAK KMLIFNTIDMGFRFLKLFLEKFAFRSAVLNAELPQ NSRLHMIIEFNAGLFDYLIATIDIGQTKDQEDSVK KPEGNSRKRKPKLDAEFGVVRGIDFRKNVLTVI NFEVQTAGSYVHRVGRGTGRAFTTGASVSLVSP DEMAFADEIKATLSDKESGKAYFINFSKGLDE AVKSLRYRAEDILRSVTSFCVRESRAQDLFMEILN SEKLSHFENPRDLDLKHDKVLKSKKEPPAHLR NIPDYLDPATQEASNVKLTRLAMGNRPKPSQR PSQKRKSKRKRDRPLKTLSGEGSNQRNKRKRSN OK
868	sp_131650_jnii	-0.854076	-0.845831	#N/A	CONTAINS InterPro DOMAINs: Membrane protein, Tact1/CMV receptor (InterPro:IPR008010); Has 447 Blast hits to 428 proteins in 176 species: Archaea - 0; Bacteria - 0; Metazoa - 190; Fungi - 133; Plants - 49; Viruses - 0; Other Eukaryotes - 75 (source: NCBI BLINK).	MASKSRGKQLSFEILSSTYAVNSDSIDRSSNDP LIQNNKSNRRRRKRKRKKSSTPNSPAPHEFYTFD NGGIDLEERESPMSMNSHSAASSFVVCVREE SLGTVCCQVSGELQRQNVSSGGGGDLVPPVEI PEAEVKSFPNLINGKLDKSELDWKRFAEED PMYPPVENSPIKFFMEEMNGNSIRSTTRGTE KERERFVDTIFRLPWRCELLIDVGYFVCLNSFLSL LTIIPRALMTLWRF_LTKRQPRTPSAEELSDFGC LLVAVGVTIGTRDLSLIYHMIIRGGQIVKLYAVN VLEIFDKLGSFGGDFVQTLFNSAEGLACSSPES SGYWLERYTIDLILAAAAASIVHSFILLAQAILSTCI VAHNALLALLVSNNAEIKSNVFKRLSKDNVHGI VYSDSIERFHISAFVIFVLAQIQIEAEGPWFNGFL SNALLVYFSEMVDIKHSHIAKFNIDKPIVYSEFLE DLCKQTMNIETDGAKKTLTFVPVAPACVIRVITP VYAAARLPCSPLPWRYFYFWMICLTMLSYVMVLSLK VMIGMGLQKHAAWYIKRCQKRKLHSD

10

20

30

40

【 1 3 5 3】

【表 1 4 - 3 5 4】

A	B	C	D	E	F
1	分析に用いたバイトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
869	#N/A	0.852455	0.830155	POZ/BTB containin G-protein 1 (POB1); CONTAINS InterPro DOMAIN/s: BTB/POZ (InterPro:IPR013069), BTB/Kelch-associated (InterPro:IPR011705), BTB/POZ fold (InterPro:IPR011333), Kelch related (InterPro:IPR013089), BTB/POZ-like (InterPro:IPR000210); BEST Arabidopsis thaliana protein match is: BTB/POZ/Kelch-associated protein (TAIR:AT2G46260.1); Has 3471 Blast hits to 3456 proteins in 95 species: Archae - 0; Bacteria - 0; Metazoa - 3227; Fungi - 0; Plants - 131; Viruses - 11; Other Eukaryotes - 102 (source: NCBI BLink).	MKNMSRDLDPRTVMDSEYDRSAKEGNFAFAN DSNFSDRVLRIEIMAGPSYSKSDADGCCSLFDW ARNRKRREDLKKDQIDNDTIEGEDVTVGPEEQ VLHOSDGDVLAACENDEEAVAMIEELPSGDEG TNSNDSLWSMECSITVSVKTLHISPIAASKPFF YKLFNGMRESEQRHVTLRINASEEDALNELLNF MYNNTLSANTAPALLDYLMAADKFEVASCMRYC SRQRLNLPMPKESALLYLDPSSVLMAEAVQPLT DAAKQYLAERYKIDTKFQEEVYMGPLIAGIEAILAS DDLQVASEDAVDFVVKWRSRQYKLEERREILA SHLARYIRFPYMTCKRKLKVLSCPFDHEVSSKL VLEALFFKAEAPHRQRKRLAAEESASTNRRFLERA YKYPVYKVEFELPRQQCVVYLDLKRKECANLFP SGRYYSQAFLHGGGGFFLSAHCNMDQQSSFHC FGLFLGMQEKGSVTFADVFEFAARSKPTTEEFQS KYKGNVTFMGGKAVGYRNLFGIPWTFMFAEDSL YFINGILHLRAELTIRH
870	#N/A	#N/A	0.655372	SUPPRESSOR OF ACTIN 1B (ATSAC1B); CONTAINS InterPro DOMAIN/s: Synaptojanin, N-terminal (InterPro:IPR002013); BEST Arabidopsis thaliana protein match is: Phosphoinositide phosphatase family protein (TAIR:AT3G51460.1); Has 1740 Blast hits to 1632 proteins in 227 species: Archae - 0; Bacteria - 0; Metazoa - 569; Fungi - 574; Plants - 290; Viruses - 0; Other Eukaryotes - 307 (source: NCBI BLink).	MRLWEPDEYVEPTDGTGGFSLSVSRVDGSMK LLDELPOCSAERYPKIOTIYGMGALKLLAGMYLF VITERESVGYLGHPIFKVSSMKILPCDOSLXSSP VEQKRMETEFSSLLAAVKTGILFFSYDQVWVTL MORLHDLGNEEKTIPLWRQADPRYCWNNYMM MLINDKLDPYLLPILOGTFHFQAGVGNDDTDVTL ARRCTRRTGTRMWRRGADPDGYVANFVETEIQI MQLNGYASVQVQVRSMPFLWEQVNLKYPKPF EIVNPEAPRVAERHFLDLRKRKYGAVLAVDLVNK NGGEGQLCKMYGSAMOKIVNEELRYLHDFDHHI CGMHFDRLSILYDQISDFLNKNRYLLLNKGEKI EEOQGVRTNCDICLDRNTVQSMIGRKMLELQL RRIGVGAEEITSSHPKFDENFKILWANHGDNVSI QYSGTALKGDVFRYGRKTFEGVFNDRGNALAR YYLNNFCODGTRQDAIDMWHGHLIASVSRDMTTQ POKGGLEAIASVPVGLSLVLAGIVFAFMSLKQVG LDWRHLLFSLVWASLSIAIAIFVRANGRVFCNRRP LHKPRC
871	#N/A	0.30287	#N/A	uncharacterized protein	MSYRRRGLNGRGAIRPPGYRSNGLTNTHSTG EDTPDQLPNYGHVAEETLNHVRNSRQYQAAASM ATTNTMEATNTRQLPNTVHQVLRQVPHQVPH TIPLHNADVGNHETITNITDTSNPEKDEHVLHP DGLWFP HRAVVKVNSATSYSYKFGPYCNWKM TPPQVQQRWVNAFKHEFSWDFSVAKLVKKGW KSKVSRRLTGIVSKEKKLGYQFSPSEVYETHAD SGGCFVNAKAEAVWEKLGYPSPSEVYETHA DSGGCFVNAKAEAVWVYSIME

【 1 3 5 4】

10

20

30

40

【表 1 4 - 3 5 5】

A	B	C	D	E	F
1	分析に用いたバイトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
872	sp_132510_stum	#N/A	#N/A	<p>decapping 5 (DCP5); FUNCTIONS IN: molecular function unknown; INVOLVED IN: negative regulation of translation, deacylation-independent decapping of nuclear-transcribed mRNA, cytoplasmic mRNA processing body assembly; LOCATED IN: cellular_component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: DFDF motif (InterPro:IPR019050), FFD/TFG box motif (InterPro:IPR019053); BEST Arabidopsis thaliana protein match is: decapping 5-like (TAIR:AT5G45330.1); Has 14874 Blast hits to 10327 proteins in 603 species: Archae - 12; Bacteria - 902; Metazoa - 6418; Fungi - 2978; Plants - 1226; Viruses - 261; Other Eukaryotes - 3077 (source: NCBI BLINK).</p>	<p>MVDFGFDGNLSISESVIDKNLKSWRPKLRSLRLLHPP SYIGSVSLTSSKEIRYEGILFSINT</p>
873	sp_133880_nmnmw	#N/A	#N/A	<p>Heat shock protein 70 (Hsp 70) family protein; FUNCTIONS IN: ATP binding; LOCATED IN: cell wall, plasma membrane; EXPRESSED IN: male gametophyte, guard cell, cultured cell, pollen tube; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage; CONTAINS InterPro DOMAIN/s: Heat shock protein Hsp70 (InterPro:IPR001023), Heat shock protein 70 (InterPro:IPR013126); BEST Arabidopsis thaliana protein match is: heat shock protein 91 (TAIR:AT1G79930.1); Has 26529 Blast hits to 25703 proteins in 4372 species: Archae - 153; Bacteria - 11749; Metazoa - 3974; Fungi - 1605; Plants - 1160; Viruses - 126; Other Eukaryotes - 7762 (source: NCBI BLINK).</p>	<p>MSVVGDFGNECVAVARQRFQIGDVLNDESNR ETPSVCFGEKGRFIGTAGAASAMNPKTTITQIK RLIGRQFSDPELQRDIKSMPPKVTTEGPDGFLIEV RYLGETRTFTPTQLLGMFLSNLKGIAEKLNAAV VDCGIGIPTYFTDLQRRRAVIDAATIAGLHPLRLIHE TTATALAYGIYKTDLPENDPINIVFVDIGHSSLQVC IASFKKAQLKVLAHAFDRSLGRRDFDEALFQHFHA AKFKEEYKIDVYQNAACLRLRSSCEKLVLSA NPEAPLNIECLMEEKDVRGFIKREEFEQISLPIER VKKPLEKALADAGLSVENIHAVELVGSASRIPSVL KILTEFFGKDRRTTWNASECVAKGCALQCALSP TFKVKFEQVNESFPFISLWSKGVADGONGAA EGQQTITVPIKGSPIPCVKAVTMLRSRGTFTVDYQ YADVSELQAPAKISSYTIYIGPFOSSNGDRAKLKV VRLSLHGVISIESATLLEEEVEVYVVKETAKMDT DDAPTETPSGEADVNMQDVKGGAEIPSAENGTP ESAEKSTSEGGDDKPVQMDTDVKADALKKVKKIT VPVSEILVHGLSQDLOKAVEKEFELALQDRVM EETKDRKNAVEAYVDMRNKLNKDYHDFVDPSE KEKFIARLQEVEDWLYEDGEDGEIKSYVYVAKLEE KKGGDPIEERYKDSMERGVINQLTHGQSFRE AAAASNDPKFHDIAEKVKVINACSEALWLGK KQQDALPKYATPALRSADVVKKSEEVDRLCRPI MTKPKPTPKPATPEPQOQOQOQOQOQOQOQOQ DTDSDGANDSNETPDAEPMETEKPENAA</p>

10

20

30

40

【 1 3 5 5】

【表 1 4 - 3 5 6】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
874	sp_135390_aamj	-0.933334	-0.933335	SHORT ROOT (SHR); CONTAINS InterPro DOMAINs: Transcription factor GRAS (InterPro:IPR005202); BEST Arabidopsis thaliana protein match is: GRAS family transcription factor (TAIR:AT3G49950.1); Has 2470 Blast hits to 2429 proteins in 304 species: Archaea - 0; Bacteria - 4; Metazoa - 9; Fungi - 15; Plants - 2427; Viruses - 0; Other Eukaryotes - 15 (source: NCBI BLINK).	MPYSIENPLSLRSDHSLVPRKSTSFPLQLPGKWL MNQTIKKPLKNVHRESSICPPSSSTDNKSCVG WVSSLLCECVKAIQGNPKSQQLLWLVNEVAS AYGDCDQLAYFLQALFAKANNGLQFYSTIKF VEEKNYCFDYYKMLKFKQEVSPWTTFGHVASN GAILLEAGENLHIDLNTLCTQWPMILLESLAT RNDETPHLRLTLVPTSLIITKEISKMEKFFARL MGVPPKFFHVISGLDCLGQIKIEDLCIQDGEAVVV NCIQALQRVHVDRRVIDMIRSIKPRVVTIVEEE VDLTTITSSDFGKCFEECLKFYKLYFEMLEESFPP ISNERLKLKSRNRVYNALACENGEIGEYWRP DKASQWSKRFNLQAFTPFHYNDELIGDVRALLR RYNTRWDLALPQQQHDHQVGIHLKWKDEPVPVWA SAWKPMLSASPPNPS
875	sp_135490_lciw	-0.9226473	-0.8834254	C2H2-like zinc finger protein; FUNCTIONS IN: zinc ion binding, nucleic acid binding; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system, intracellular; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Zinc finger, C2H2-like (InterPro:IPR015880), Zinc finger, C2H2-type (InterPro:IPR007087); BEST Arabidopsis thaliana protein match is: zinc finger (C2H2 type) family protein (TAIR:AT5G40710.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archaea - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MVTIRSKSIGFTCFLLFVMLTNSLILGISAAPLH HQDQGLGENAESRDLLKEHEHIHVCHSRERSR AAWEIVKYLIMHFVEREYERISRNGRLHPDNDLF RDQENHKIKEDVNEWQCGYCKKSLSEEFIDRH FDSRRHYLLNFTGSKCLADLCLGALHCDHVLGSK SFPKACNFAAARNRHLCESLADACFPVYSGGFS ASFLHELFLQQFCYAHTCTGGIKPFSKGGRKETS VLYFAASVLTMLLPFYVVLVCLHQSETRKDTQEL KRISRLGRKAKPM
876	sp_135520_mahn	-0.330649	#N/A	uncharacterized protein	MFEORRRASRGKKAIVAVAAVATATATSAKKY

10

20

30

40

【 1 3 5 6】

【表 1 4 - 3 5 7】

A	B	C	D	E	F
1	分析に用いたペアトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
					<p>MATEEKLKSEEQTLLETVSATISAVSTAKYYVD          DWILAHVAVLFPIDTSSLSGVISEKYSQDLNQA          ETTYTSERRINQWVEFYKGYSFPTLSRFUJYDVA          SNWLPFPFLTARKLVYDIFFVEGLVLEIVQALIPSL          GHNVSVRSNAERLVELLQLENEGVLQARELGL          AYHESKSSNSLVKLTVSRVAQLVASVPDKARLGL          APKSSHLFFKSVTIQLLDGAEKESWELYDKEV          LFHGTVDVGTQFVAEMFARICRRGSTTELLVNEV          ILRIRRIQRFLLSHKDTITENSFKNSDFQFWFL          MIEILRDAYVEKMEIKLHLQAAECEHDETEAYWIL          WILFHAFQHQITIRAMFVENFLLSKVFPFRCLR          WILQFAIFGCPFDDRLKKGDNHNVLLEAAQRVV          QMWSRREFVQSAPLEQQVYVAVAGLLENMS          REDLDSTKDMRSILEGYSRCLESPDELVRRMAS          NIALVFSKVDPKNPLYLDSDCKEGNVDWFEFGS          ASSNSALNSKDRTSDNLTNTKTDKGVHMKEQN          HGDTHERESKSNPGKKKPFHGKLVDPDEIVDPA          SLNDELSSGEEVDDNASEDSEASSDSSLQPYDM          TDDDLTKKRISQLVDVVGALRKPDDADGVESAL          DVAEKLVRASPDRLRHASDLVRTLVQVRCADTV          LFEFEESAEKPKQKALVALLVMCPFESLNAINKLL          FSPHLVDSQRMILDMVTDAEELANAKTIVRKNQ          PKLSISSASEPQVWFLPSSKRPFGAGSWEIPIDT          LPFSALNLTNRFERELPSKPGVIKGTTRRWGHPK          INLQANEINWSQNKFFLYAAAFMLPAMQGFDDKKS          HGVDDLGRDFLVGLKLYMLGVCIRCTALHPEASA          LAPSLLDMISSREICHHEKFAFVRRSVLFAVCSLK          ALHIPFSIASAIMEGNEEVSRLGLEWVVRTWALRAIE          SDTDTTECOAMAMACQLQHAEMSLQASRAFNATAE          SSSLQKKSIGLPSNMMSKGVIRIPGNSNSELG</p>
877	sp_135530_zhgh	-0.8530,33	#N/A	<p>EMBRYO DEFECTIVE 2423 (EMB2423); CONTAINS InterPro DOMAINs: Telomere length regulation protein (InterPro:IPR019337); Has 246 Blast hits to 237 proteins in 108 species: Archae - 0; Bacteria - 0; Metazoa - 102; Fungi - 75; Plants - 52; Viruses - 0; Other Eukaryotes - 17 (source: NCBI BLINK).</p>	
878	sp_135710_jiya	#N/A	-0.86637,37	<p>unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological process unknown; LOCATED IN: endomembrane system; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G57080.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).</p>	<p>MNGKFGGKLPDTGTPSMARSTVVVLSLLGGAS          VVHNIYKPDLTLPSPPEEYSEDKQKNEKQ</p>

10

20

30

40

【 1 3 5 7】

【表 1 4 - 3 5 8】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
879	sp_136540_cxet	#N/A	-0.314974	ENDOPLASMIC RETICULUM MORPHOLOGY 2 (ERM2); FUNCTIONS IN: transporter activity, zinc ion binding; INVOLVED IN: transport, ER body organization; LOCATED IN: COPII vesicle coat; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Sec23/Sec24, helical domain (InterPro:IPR006900), Sec23/Sec24 beta-sandwich (InterPro:IPR012990), Sec23/Sec24, trunk domain (InterPro:IPR008896), Zinc finger, Sec23/Sec24-type (InterPro:IPR007123); BEST Arabidopsis thaliana protein match is: clone eighty-four (TAIR:AT3G44340.1); Has 86818 Blast hits to 46504 proteins in 1607 species: Archaea - 60; Bacteria - 12149; Metazoa - 40940; Fungi - 13838; Plants - 8596; Viruses - 2013; Other Eukaryotes - 9220 (source: NCBI BLINK).	MRPFLSSGVPVAGRPDGDGVRPPPPSAPGSAMP FASPRMAGPPVPGFRPSPVVGFTDQIRP SHAPPTSNINQFSAQYSGPPTNPPIRTPVVG PPNINMASPMPPNIPPIQIRPQVAPPTPLTFR PNSQVPPQYQAQAPPQFAQGGPPQYQAQGGPPQ PYAQQGPPQYQAQGGPPQYQAQGGPPQYQAQGGPPQ PYAQQGPPQYQAQGGPPQYQAQGGPPQYQAQGGPPQ QGFVEDFSSLSLGA VPGSMDPGLDSKALRPLE GDVEPTVNDMTMNCSPKYMRLTTSAINPSQSL ASRWHLPLGAVSPLAEASEGEEVPIINFGSAGII RCRRCTYINPYITFTDGGKRWFCNMCALLNDV PGDYVAPLDATGRRIDADQRPRLRGSIEFVAPTE YMWRAVMPPLYFFLIDVSVSAVRSGMVEVVAQTI KSCLEIPGYRPTQIGFITDSTHFYNMKSTLTQP QIMVVSDDVDFVPLPDDLNLNLSERNVAVESFL DSLPSMFQDVTNVEFAFGPALKAAAYMVMSQLGG KLLIFONTLPSIGLRLRGLRDEPRVYGTDKKASL RVPEDPFYKQMAAECTYQIGINVYAFSDKYTDIA SLGTLSKYTGQVYYPSPFRSTIHADKLRHELAR DLTRETAEAVWRIRCGKGVKFNYSHGNFMLRS LLKFLYPSLRVDFELKASSPDAADLNKVEKRLPL TAESLDARGFILDDGLRFVWIGRMFPPELAASL NIDFASDLQVNGSNEMSRRLMGLLKKLRESL PAYYQFCHLVRQGEQPKGEGFYVLSNLIEDQTAGT NGYVDWIAQIHRQYQONA
880	sp_137640_jigi	#N/A	-0.849215	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 17 plant structures; EXPRESSED DURING: 10 growth stages; Has 35333 Blast hits to 34131 proteins in 2444 species: Archaea - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	MASFITLALASAIIVSMALVPSLSSALVSSSPDVNVP PPPKAISDLKEAVVKGFGFTDDFKITGDMRDT LVGHSVAYEFDVEIDNKVLPVKLLEDVKRWVEYVD LPIFRVEKVPKPGEEENGLIERKMMMEENMGVPLAP FOLAGPMEIWDQDAKHMRLSLPHDVEDAGELRKYVI LAEGAVTVKGAKSVALRQPLELPLNQIQNGF ASGSLSLAQHLHVSGSKGAPLLSLRVLVVGPTSLA SPFSSSSSSNRLLKRLAPLGLVELVSSPSTNKPV EIATIDIPSEVPTLLTPEOFTLLWPMSSLNGSNS NLLGLEKLLSSVLGPKANKKGSFRLLKADVSAQT YUKVGFKEVKQLQEGDGINLEGFPEWRTPKEKV KMHFEALAKVEGDRVIPEKVTQVHPVILEDTVAP SVLLGNVTMSKMPHFHQPSNPLTL

10

20

30

40

【 1 3 5 8】

【表 1 4 - 3 5 9】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MEGENRDGGSTFRVNFSDGGAAKLAALLTDTLK EYMGDDVDDLTKVYVIVLLGNGRLKEEVKEDLDV FLGDKSDSFVTWLDHLDLWLDLNLVYVGSQSPN EVAKAKPELVKTSKDHSDHPVSESDRGKSSKT FRSRHNRDWKGLVTDNSQPPLIRSSVTNNLQNE EKTYPNALHGRQSVSPERPVHRRKRYRPRGPRRP KREISOATIAAPRLLQFAVREAVATFRSSNSYSE PARKRLRSVSTTTGNSTEEVYHRRIQSVAVRLQNI PMATAINVAEASKDVRRTGNVDFDLRSRGGG SSDKVPVKELAVQDDHDHNDYDHIHEKSRHRYV EKFENNGHHNVNASIVASSTGLASDSASDNEGY DDVNVGARGVDVSETGTSKAKADDSLWVHYSV QNTDGVTRQSRKQDELPAASAANTKHKIVNISVN VNTWKPAAHTQKLNKTAGSSLPGAHSTGRPLED SNGNPAADTKLNKTAGSSLPGAHSTGRPLED ADSRITIFVNNVHFAATKDTLSRHFNFKFGVEYLVIL TDAATGHPKGSAYVEFMKEAADNA.LLDGTSF MSRMVVKLVKSSVQCEANPVMWMPRIITRGSPP AAPRFACAPFKVFPFGTQARAPVVKPGMRSMQ WKRDTQQTSVESVAVSPGNLNAGRGLTYVRTDA K
881	sp_137760_ximr	-0.81934%	-0.83697%	RNA binding (RRM/RBD/RNP motifs) family protein; FUNCTIONS IN: RNA binding, nucleotide binding, nucleic acid binding; INVOLVED IN: mRNA processing; LOCATED IN: cellular_component unknown; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: RNA recognition motif, RNP-1 (InterPro:IPR012677), binding, alpha-beta platt (InterPro:IPR02483); BEST Splicing factor PWI (InterPro:IPR002483); BEST Arabidopsis thaliana protein match is: RNA binding (RRM/RBD/RNP motifs) family protein (TAIR:AT2G24350.1); Has 3110 Blast hits to 3077 proteins in 726 species: Archaea - 48; Bacteria - 800; Metazoa - 1243; Fungi - 266; Plants - 328; Viruses - 0; Other Eukaryotes - 425 (source: NCBI BLINK).	
882	sp_138010_ajjs	#N/A	#N/A	NADH-ubiquinone oxidoreductase 24 kDa subunit, putative: FUNCTIONS IN: NADH dehydrogenase (ubiquinone) activity, zinc ion binding; INVOLVED IN: response to oxidative stress, mitochondrial electron transport, NADH to ubiquinone; LOCATED IN: mitochondrion, mitochondrial respiratory chain complex I, respiratory chain complex I; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Thioredoxin fold (InterPro:IPR012335), NADH ubiquinone oxidoreductase, 24kDa subunit (InterPro:IPR002023), Thioredoxin-like fold (InterPro:IPR012336); Has 5564 Blast hits to 5564 proteins in 1535 species: Archaea - 26; Bacteria - 3396; Metazoa - 195; Fungi - 114; Plants - 49; Viruses - 0; Other Eukaryotes - 1784 (source: NCBI BLINK).	MFARRAAQRLLLEIRQVFRQSPCASRFSFSSALNPH LDTPDNPNQWKFSETNQQVKEILSFYPTNY KOSAVIPLDLAQOQHGGWLPVSAMNEVARIEV APIRYEVATFYVSMFNRSKVGKYHLLVCGTTPCM IRGSREIEEALLNHLGVKRNEVTKDGLFSVGEME CMGCCVNPAMITADYSNGSEGYTNYVIEDVTPK RVVEIVEMLKRGEKPPPGTQNPINRPFKCAPEGGH KTLLEGPKPPFCRDLDAC

10

20

30

40

【 1 3 5 9】

【表 1 4 - 3 6 0】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
883	sp_138210_mzrw -0.884134	-0.922452	-0.863813	Transducin/WD40 repeat-like superfamily protein; FUNCTIONS IN: nucleotide binding; INVOLVED IN: biological_process unknown; LOCATED IN: CUL4 RING ubiquitin ligase complex; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: WD40 repeat-2 (InterPro:IPR019782), WD40 repeat-like-containing domain (InterPro:IPR011046), WD40 repeat-containing site (InterPro:IPR019775), WD40-repeat-containing domain (InterPro:IPR017986), WD40/YVTN repeat-like-containing domain (InterPro:IPR015943), WD40 repeat (InterPro:IPR001680), WD40 repeat, subgroup (InterPro:IPR019781); BEST Arabidopsis thaliana protein match is: TBP-associated factor 5 (TAIR:AT5G25150.1); Has 6115 Blast hits to 4690 proteins in 400 species: Archae - 26; Bacteria - 2036; Metazoa - 1593; Fungi - 1149; Plants - 369; Viruses - 2; Other Eukaryotes - 940 (source: NCBI/BLINK).	MSLSQNAIKNKVKKKSKRNINDESEEPSLLEDR EEVLENEGLDGDAAALS.LKRRKRRNDAAEKRLIVE EELMQLGNFLFGSLHSPVEFGNVVEDENENG GLDVEDSALFVMDRSGNKDVSVYGDEGDQVD ENGDQRKFAWLDDDEEDTVNISKVRLRKLK EEDETEISGAYSVSRRAHAKLNFGTGWANLD SKVRTYDSESESEVENGAVASGYENVDDILRTN EELVKFGAKLLPRILEYSRLVDANAEDP/SNGPIN AVQFHRNHOLLAAAGLDKLRFFQIDGKRKNMM DSFIEDGPIRNASFLPDGSSQVIVSGRRKFFYSVDI SKGNVDKIGLTKREEKSLVFAVSPDSRTIAFLG NEGVLVLSAKTKEJGTLKMMGTARS/AFDIDDGQ QLLSSGGDG-HVYLWDLRTRTCLHKGVDEGSING TALCTAFVGNLFAAGSSGIVNVNKEEFLLGGKR KPIKTLHLRTEASFMKFNHDAQILCAICSTMQKSS LKLVPISFTVFSNWP/PPASKTLHYPRCLDFSPAG GFMAAGNAAGNVLLYKLHHYQQA
884	sp_138410_xiut -0.91282	#N/A	-0.892369	DNase I-like superfamily protein; CONTAINS InterPro DOMAIN/s: Endonuclease/exonuclease/phosphatase (InterPro:IPR005135); BEST Arabidopsis thaliana protein match is: DNase I-like superfamily protein (TAIR:AT3G58580.1); Has 1372 Blast hits to 1328 proteins in 220 species: Archae - 0; Bacteria - 20; Metazoa - 540; Fungi - 247; Plants - 315; Viruses - 0; Other Eukaryotes - 250 (source: NCBI/BLINK).	MLSVRVHLPSDIPIVGCCELSPVYLLRKPDKSVIN DDVPEAPIDGGQFLRYKWKYRIGSDKKVPVCSVH PTESATLQCLGCVKGIPIVTKSYHCSPKCFSDA WQHHRVLDHRAASAYNENGGEEDELFGFRNSS GSGVLTGLPPSASSPNLTNGSTPLYPAAVTR SGGETWFEVGRSKTTPSTDDIGHVLKFECCVID AETKQTVGHPMTLTSRVPAPPTPRRMSVSG ADVMGHLDPDGRISAGTFTVVSYNILADAYATS DTYSYCPSWALSWPYRRQNLLREIVTYRADIVCL QEVQDFHEFEFFSPELDKHGHALYKRKTSEVT HTVDGCATFRDRDFSHVKKYVEVEFNKAAQSLT EAVVPAQKLSALSRVLDNNVALIAVLEAKFSHQ GADTTPGRQLLVCVANTHVNVHQLKDKVLLWQV HTLLKLEKIAASADIPMLVCGDFNSTPASAPHAL LAMGKVDPLHLDLQVPLLLRPHHTKLAHLQLPLVS AYSSFARIGIGFALQQRRRMDPSTNEPLFNTCT RDFIGTHDYFYTADSLTVESLLELLELLEDSLRKDT ALPSPWSSDHIALLAEFRCKSRIRR

10

20

30

40

【 1 3 6 0】

【表 1 4 - 3 6 1】

A	B	C	D	E	F
1					
2	分析に用いたペイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
885 sp_138600_dhjc	#N/A	-0.925315	-0.883398	Pseudouridine synthase family protein; FUNCTIONS IN: pseudouridine synthase activity, RNA binding; INVOLVED IN: pseudouridine synthesis, RNA modification; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Pseudouridine synthase, catalytic domain (InterPro:IPR020103); Has 545 Blast hits to 513 proteins in 206 species: Archaea - 168; Bacteria - 1; Metazoa - 125; Fungi - 4; Plants - 38; Viruses - 0; Other Eukaryotes - 209 (source: NCBI BLINK).	MGASNADEELRLDDVLRSPFAAVNDLLTLRLC PQCILRFLFGKYGYPIYTHYSRISTVLAEVVQHQPS DDPLLKESANEDGICSVCFGLQFCVDEKQTLVN KGCASDFAALIYELARKDGHQADSLSEVSIPIIV ENENAVRLSMKQKQYSSDWFQEMLQSEFISVKD ALKLSITEPLETLGKVGSSSFRVRLTYTSSEMP MKTDNGKENQGSKKRIGRRSEITDTEVHLKQ CSDSLQEKISFRLLKISACSLEVLCSSTPYIGGR YLKYSRNVSQSRWIIDDERMGEASVEEIVGTSILP LCKGDNYKFAAGREDIDVRLMGSGRPFLEIQN PFMLPSELSLKEMETKINSLEGKLVKVKNLKVVYD EOGWTLMRGEAEKQKQYALVWISROLEDKD LVTISSLDKLLQRTFVRVLRHRRSPLEREKIHSM SIEKMSGSTQYFLLHLCTOAGTYKEFVHGDGGR TYPSSLGSILGCRVLELQLDVTDVKMDCFLTE
886 sp_139080_gwju	-0.833398	-0.853313	-0.884508	cytochrome P450, family 72, subfamily A, polypeptide 7 (CYP72A7); FUNCTIONS IN: electron carrier activity, monooxygenase activity, iron ion binding, oxygen binding, heme binding; INVOLVED IN: oxidation reduction; LOCATED IN: endoplasmic reticulum; EXPRESSED IN: 16 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: Cytochrome P450 (InterPro:IPR001128), Cytochrome P450, E-class, group 1 (InterPro:IPR002401), Cytochrome P450, conserved site (InterPro:IPR017972); BEST Arabidopsis thaliana protein match is: cytochrome P450, family 72, subfamily A, polypeptide 13 (TAIR:AT3G14660.1); Has 33011 Blast hits to 32901 proteins in 1695 species: Archaea - 71; Bacteria - 5150; Metazoa - 11421; Fungi - 6601; Plants - 8312; Viruses - 3; Other Eukaryotes - 1453 (source: NCBI BLINK).	MGQTSQDDEFPCLPYGEVEGQSLKILLXLINAG KNYFLWSGSPVPMINISKLEIREAFTKTNEFRKTK VNFIFHKLVPGLVSLKGEKWKVHRRMMNPAFRM EKLK

10

20

30

40

【 1 3 6 1】

【表 1 4 - 3 6 2】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたペアとの相関係数 sp_107620_kqnh sp_107660_kiqg (SOAP1)	sp_107670_ptqx (SOAP2-like)		注釈	aa seq
887	sp_139830_cwdw	-0.333315	-0.89437	-0.303373	<p>DUO POLLEN 3 (DUO3); CONTAINS InterPro DOMAIN/s: Homeodomain-like (InterPro:IPR009057); Has 1836 Blast hits to 1575 proteins in 289 species: Archae - 4; Bacteria - 248; Metazoa - 589; Fungi - 287; Plants - 112; Viruses - 46; Other Eukaryotes - 550 (source: NCBI BLink).</p> <p>Late embryogenesis abundant protein; group 2; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: response to calcium ion, response to desiccation, embryo development, ending in seed dormancy; LOCATED IN: plasma membrane, membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Water stress and hypersensitive response domain (InterPro:IPR013990), Late embryogenesis abundant protein, group 2 (InterPro:IPR004864); BEST Arabidopsis thaliana protein match is: Late embryogenesis abundant protein (TAIR:AT1G01470.1); Has 294 Blast hits to 286 proteins in 78 species: Archae - 4; Bacteria - 52; Metazoa - 0; Fungi - 0; Plants - 235; Viruses - 0; Other Eukaryotes - 3 (source: NCBI BLink).</p>	<p>MKTSPLTDEEKARIQEGKAFKLDWVYSWVRTMV PYRDPNLLPRQWRIAIGTQKSKYKADSAKKEKHRV YESNRRRSKMAVVDYFOTGSEKENESGEDCVE NDNEAYVHEAFLADWRPPTGTVQNGIDNQRN QASOSGKVHAQKVSAA TEDSHLVHCKSVHPHITS SFPPPISETFGESVFRSQRAGKPNLVKLAPLPL HYKLPFAVRVISQASLRSSHLGSSGVEVGGVNCVS ADLVKHMQDKSMISNPNSLSVTNELLKNGYSLTK ERGANLDFQMHP LLFRITIEGNLPCYPVNSPMRI PTFSFFPPTLLQVDPDAGLKSSSSKKETAAPPSSSL DFHPLLQRAKEFHANSLDTGCVLNTDLELSRDQF VEDIYDAGATAAAQTPSADANDLDLNIHPSSA SKFRKVSQQGAEINPVSVDGSRVYESVKETSSD FSEFAELQPTADLVAKRHTAEPNPEKSVAVSHK LVSSNDDVGIQPLLEVVMEQEEELSDSDEEWEDIE FECCEEMADSDGKESDVAQVVDADAFHKEQVD DAALVTATHAQLLKSLSAQEQTQTKRGRTPRKG AGKGGKHTKGSWSLNSHHPKPLSRKRTISPNRD CSKEISSSSKDVSCSTEPTKGTGISEKDSVSSPSRRT RKHASKSSSSPLEVNSNSCCVN</p> <p>MASDDNPELVERAFKKGKDSKKEEKEEGFIEK VKDFIHDGIEKIEGAGFGKPTADVSGIIVPHISLH KLDIVDVLVKPNPPIPLDINCYCIDSEGRKLSIG LIPDAGTIHAGSETVKIPATLIFGDKETKDKIPFG SIPYRLKVDLIDVPIGRITIPLEKTGKIPYKPD DIEKIKFRFEETAAIHLKIENKNDPDLGLKDL YEVWLGHVNIAGDLEKSEILPQNGISNLHLVPTF KPKDCGSALWDMRGGKGTGYSMKGNINVDTPFG EMKLPISKEGGTTRLTTKREAGGSDDD</p>
888	sp_139820_mofj	#N/A	#N/A	-0.275333		

10

20

30

40

【 1 3 6 2】

【表 1 4 - 3 6 3】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
889	sp_140140_mgwa	-0.386332	-0.382675	#N/A	APS reductase 1 (APR1); FUNCTIONS IN: acenyl/sulfate reductase activity; INVOLVED IN: response to oxidative stress, sulfate reduction, sulfate assimilation; LOCATED IN: chloroplast, chloroplast stroma, plastid; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Rossmann-like alpha/beta/alpha sandwich fold (InterPro:IPR014729), Thioredoxin-independent 5&apos;-adenylisulphate reductase (InterPro:IPR004508), Thioredoxin fold (InterPro:IPR012335), Phosphocadenosine phosphosulphate reductase (InterPro:IPR002500), Thioredoxin-like (InterPro:IPR017936), Thioredoxin domain (InterPro:IPR013766), Thioredoxin-like fold (InterPro:IPR012336); BEST Arabidopsis thaliana protein match is: APS reductase 3 (TAIR:AT4G21990.1); Has 5999 Blast hits to 5419 proteins in 1658 species: Archaea - 169; Bacteria - 2693; Metazoa - 1288; Fungi - 563; Plants - 600; Viruses - 0; Other Eukaryotes - 686 (source: NCBI BLink).	MALAVTSSLSSEKAPSIGVFDQRQYQNLATISL TTFRAGSAVKPLRAPAESRRSDSVPLAATTFAA PKTEEGYKVEEVLVEEENEFEDLAKELANASPLEI MDRALEKFGDDIAIFSGAEDVALIEYAKLTGRPF RVFSLDTGRNLPETYRFFDEVEKHYGIRIEMFP DSVEVQGLVRNKGFLSFYEDHGQECRCRVRKVRP LRRALKGLRAWITGQRKQDQSPGTRSEVPVQVD PVFEGLDGGVGLVKWPNVANVEGNWVNFLLR TMNVNINSLHAQGYVSGICEPTTRAV_LPGQHERE GRWWWEDAKAKECGLHKGKLNKEESENQNGG SAGANGEVADIFESENLVTLSRAGIENLMKLEDR RDPWVVLVYAPWQCFQAMEGSYLELADKLAGS GVKVGKFRADGDKAFKSELQLQGSFPTLLFPK HSSQPIKYPSENRRDADSLAFVQALR
890	sp_141200_kpwp	-0.386325	-0.384328	-0.384639	DnaJ domain ;Myb-like DNA-binding domain; FUNCTIONS IN: heat shock protein binding, DNA binding; INVOLVED IN: protein folding; LOCATED IN: cellular component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Molecular chaperone, heat shock protein, Hsp40, DnaJ (InterPro:IPR015609), Heat shock protein DnaJ, N-terminal (InterPro:IPR001623), MYB-like (InterPro:IPR017877), Heat shock protein DnaJ, conserved site (InterPro:IPR018253), SANT, DNA-binding (InterPro:IPR001005), Homeodomain-like (InterPro:IPR009057), Myb, DNA-binding (InterPro:IPR014778); BEST Arabidopsis thaliana protein match is: DnaJ domain ;Myb-like DNA-binding domain (TAIR:AT5G06110.2). Has 60375 Blast hits to 43868 proteins in 3410 species: Archaea - 228; Bacteria - 14410; Metazoa - 19184; Fungi - 5541; Plants - 3499; Viruses - 192; Other Eukaryotes - 17321 (source: NCBI BLink).	MMAQRSHVLLISYQELKDGKPLVYSSNCLPLK YSKFEFAGHSFHSAAALRSGYCEEDSDKDDDKS VSNDKETFTPSFESYSGKSKKSGGDKGQQQD HYALLGLAHLRYLATEEQIRKSYRDVALKYHPDK QAAALLLEKTEASKQAKKDEIENHFKAIQEAYELLI DPQKRMVDSTDEFDEIPTDCAPQDFKVFPGP AFMRNSRWSVQAPAPSLGDDNSSFEEVEKFN FWYYSKSWREFPHEDEFDLEQAESRDHKRWME RQNAKLEKARKEEYTRVRSLVDNAYKRDPRIAK RKEFEKAEKRRKESFLAKLLEEEEAARVAEEE RLKKEEDKAAEAQVQKKAKEKELKVKRER TRLRSLASEVAGSADFISGDDVETLMSLDMER LKNLCNKGKAKEEQVKLLKDS-SGKVSSEENG TPAPNPTPKPTGAVKSNVAVSSSVLEKKEK EKPVVKEIDLIRKGIQKFKPKSRRWVEISEYI GTGRSVEELKATKTVLLOKPDSSKAFDSFLEKR KPAVISPLSTRDEVAVNGVNGVNGVNGVNGVNGV VADATSSTTTSDSPNVSSSTDIPSSSDPDMVWS SVQERLLOALKTFPKETAQRWVERVAAPVPGKT VNNCKKLLASLKETFRSKTTA
891	sp_143030_jcaf	-0.38541	-0.396332	#N/A	uncharacterized protein	MSLPEVTITLGRVCLVWFLSRLEEKGDITNLFLT

10

20

30

40

【 1 3 6 3】

【表 1 4 - 3 6 4】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
892	sp_143320_njor	-0.912332	-0.924573	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G61450.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MDDTGTILFQISTLKMMLDKVNDIEITNIOITREIE SEVKEIEETALDARESELMKLYLSQYELSGFIAY TVNSRKSVAALTEEINHQRKADLLERLNKKREI FSTACIEFQKNIESEECNRVSRMLLEKESLEDEIQ KLDRIDTKFRIPMTAVSDEILEDLHSSNSALAMEIQ ERQAEHERLLKEIQFKSVLLSTISDVDDPW
893	sp_143380_yymq	-0.9332509	-0.9332594	hydroxyproline-rich glycoprotein family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cytosol, nucleus, plasma membrane; EXPRESSED IN: 14 plant structures; EXPRESSED DURING: 9 growth stages; BEST Arabidopsis thaliana protein match is: hydroxyproline-rich glycoprotein family protein (TAIR:AT5G57070.1); Has 16953 Blast hits to 10626 proteins in 654 species: Archae - 27; Bacteria - 1228; Metazoa - 7030; Fungi - 2381; Plants - 3856; Viruses - 536; Other Eukaryotes - 1895 (source: NCBI BLINK).	MEQKDEDPFWMQNTTALRRADRRHRRNQL ASSLFLNTGLVLLAAAVFLAIVFVLSLRSI FNPSLVKSSWNSLNILVFAIVCGF_LGRKDDQ STPVSPMKQEAVKVSSSPRQWYEQDSMR RLRRNISSYPDLRQEGQWLSENWRWRSYDDTH VYNHQFQRSESRSRSMGVLDDEIDVKNIGV DTFVGGGIEVAPPSTRTPSPPTPPPAASHNVHV RSKQVADDEDEQYVTKNIEVMVTKSESIPIPSVHQS IPPSPPLSPLSPISPTSYFWPSPSPLPPIPPPPPPQ RAHSSMSYKHKHETKHKHENNNPKRESRSKVP IQAPPPPPPPQSRPIRQGFSEDTDRKKGIV VRTLSRLRRKRKQARRSYENLTEIQPKTSIQR SKSTPEFPASPPIPPPPPPPPPPSPMFQNFITKK GKGRKRVHSHFTTTSSPRHVPSPARPQQHGY QSHRPSPKTTTTPTATRTVPLPPAKTDSFSSMA GNNGNSESLPIIPPPPPPPPKMSAWRFVYTO GDYVKIESNNNSPCSFSGFSADEDESSPNNQKS SNLFCPSFDVDTKADDFAKFRARLRKLDKMNISIKR NHGTGPSPLGSDSGQP

10

20

30

40

【 1 3 6 4】

【表 1 4 - 3 6 5】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたペアとの相関係数 sp_107620_kqnh sp_107660_kiqg (SOAP1)	sp_107670_ptqx (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MIGLGSINSVLLSAGVGGSGKEEGGLGYSVYVAV EKSOEEGLSLFPSDVQDKDEKQSTRLGSTLYFEL HGVPQRSGESSSDVKSSSSTMIHMPDLHLRKE DILLKQCEQYNIPIEQRFSLTRIRYAHAFRSL RSGRLYSRCLLAYLVQSDANDELVSFFANEP EYTNELKIVRSEEVISGTVRTLAMLALGAOLAAY SSSHDRARMLSGSSINFAGGNRLILNLVLOKAILS LKNSNDPSSLAFIEALLOFYLLHVISTSSGSIIRG SGMVPFLPLEDSDPHMLVYAVYVTKLQKLM DYSNAAVTLKDSGGVELLQORLHIEVQVIGTA GSDGTVMV/EGESSSESDHYGQKRLIRALLKA LGSATYAPTNTSSOSSHDSTLPSLALIFSNVYG KFGGDYCSAVTMSEILHKDPTCFPI/HEMGLPD AFLLSSVAVGLPSSKALTCVPSGLGAIC/LNTKGL EAVKSSALRFLLEIFTSKYYVYVAMNEAIVPLANG VEELLRHVSSLFNTGVDVIEIINKISSLGECSGG SSEKVNESNMESMDSGEKGGESSSTVTPS DSAGEDITDEHFINLSVFHVMVLIHRTMENSETCR LFVEKSGIDALLKLLRPSIAGSSEGMALHSTM VFKGFTQHHSSPLARAFCSFLRDLHKKALTGFSL VAGSFLDSDSKLDHGVFSSFLVEFLFLAAAKE TRWISALLTEFGGDKVDLEDIGRVHRETLWQIA LLEDKAVEQEGAGSVSASESQQVSSSAGPEEQ RVTSFRQLLDPLFRRRMSGW/SVESQFFDLINLYR DLGRAGGIQRSTSDSSNLR/SGSSQQSRESAGSD TDIALRKEEDKQKIYSSCCDMVRS/LALHIMHLQ ELGKSM/LQPSRRRDDAINSPSSKAVVSTFASICL GHMNFEGHMS/PSGSEASVSGKRYGKVVDFID GILMDRHDSCNPVYLLNCLYGGQVVGQSVLTTFEAT SQLLFTV/PASPMETDDGNLKHKEEKESEHLWLY GPLASYGK/LMDHLVTSSYLLSPFTKHL/LLVQPLANS NVFPFQDAETTFVKV/LQSMV/LKVVLVYVWTHPQFV
894	sp_143570_pccor	-0.36:1043	-0.32:6604	-0.30:3356	ubiquitin-pro:ein ligase 1 (UPL1); CONTAINS InterPro DOMAIN/s: Ubiquitin-associated/translation elongation factor EF1B; N-terminal (InterPro:IPR000449), Ubiquitin-associated/translation elongation factor EF1B; N-terminal, eukaryote (InterPro:IPR015940), E3 ubiquitin ligase domain of unknown function DUF913 (InterPro:IPR010314), Ubiquitin interacting motif (InterPro:IPR003903), E3 ubiquitin ligase, domain of unknown function DUF908 (InterPro:IPR010309), HECT (InterPro:IPR000569), Armadillo-type fold (InterPro:IPR016024), UBA-like (InterPro:IPR009060); BEST Arabidopsis thaliana protein match is: ubiquitin-pro:ein ligase 2 (TAIR:AT1G0320.1).	

10

20

30

40

【 1 3 6 5】

【表 1 4 - 3 6 6】

I	A	B	C	D	E	F
1						
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)		注釈	aa seq
					<p>NAD(P)-binding Rossmann-fold superfamily protein; FUNCTIONS IN: oxidoreductase activity, binding, catalytic activity; INVOLVED IN: oxidation reduction, metabolic process; LOCATED IN: endoplasmic reticulum, plasma membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Short-chain dehydrogenase/reductase, conserved site (InterPro:IPR020904), NAD(P)-binding domain (InterPro:IPR016040), Glucose/ribitol dehydrogenase (InterPro:IPR002347), Short-chain dehydrogenase/reductase SDR (InterPro:IPR002198); BEST Arabidopsis thaliana protein match is: NAD(P)-binding Rossmann-fold superfamily protein (TAIR:AT3G03350.1); Has 112445 Blast hits to 112247 proteins in 3561 species: Archae - 906; Bacteria - 75731; Metazoa - 6658; Fungi - 6086; Plants - 2548; Viruses - 2; Other Eukaryotes - 20514 (source: NCBI BLINK).</p>	<p>M L P L P V L F L I L V I Y V K F A I V D A D F T L R S K K L K R K E F E D K V V W I T G A S R G I G E V L A K F A S I G A K L I L S A R N E A E L E R V K N Q L S G K H A P H D V K I L P _ D L A S G E E S L R E A V E K A E A F F G G V D Y M I H N A A Y E R P K S A A L D V P E E S L K A T F E V N V F G T I S L T K L L A P Y M L K R G R G H F V M S S A A G K T P A P Q O A M Y C A S K F A L N G Y F H C L R S E L F Q K G I K V T V C P G P I E T S T N A E A G S S G Q A K S I E K R V S A E R C A Q L T I I A A S H G L K E V M I S Y Q P V L G W M Y L V Q Y M P T I G F W L M D K I G E K R V E T A A Q K G N T Y S L R L L F G K S K R E</p>
895	sp_143960_yemn	-0.924727	#N/A	#N/A	<p>translocase inner membrane subunit 17-2 (TIM17-2); FUNCTIONS IN: protein transporter activity, P-P-bond-hydrolysis-driven protein transmembrane transporter activity; INVOLVED IN: protein transport, protein import into mitochondrial inner membrane; LOCATED IN: mitochondrial outer membrane, mitochondrion, mitochondrial inner membrane, mitochondrial inner membrane presequence translocase complex; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Mitochondrial inner membrane translocase complex, subunit Tim17/22 (InterPro:IPR003397); BEST Arabidopsis thaliana protein match is: translocase inner membrane subunit 17-1 (TAIR:AT1G20350.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).</p>	<p>M G T P T S R E P C P D R I L D L G G A F A M G A V G G S V F H Y V R G S I N A P K G E R I L R A L E T V R M S A P R T G G G F A V W G G L F S T F D C T L V Y I R Q K E D P W N S I F A G A A T G G F L Q V R Q G I R A S G R S A A F G G I L L A L I E G A G I M L N R V M S P P A Q I M Y E D P S M G A G A P Q O O M G G Y G F G F P G O A A N V G G V L E S F D T P T P M P N F D Y K</p>
896	sp_144160_wuxy	-0.877376	#N/A	-0.830677		

10

20

30

40

【表 1 4 - 3 6 7】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
897	sp_144200_quwk #N/A	#N/A	-3.853022	<p>SIZ1; CONTAINS InterPro DOMAIN/s: DNA-binding SAP (InterPro:IPR003034), Zinc finger, MIZ-type (InterPro:IPR004181), Zinc finger, PHD-type, conserved site (InterPro:IPR019786), Zinc finger, PHD-type (InterPro:IPR001965), Zinc finger, FYVE/PHD-type (InterPro:IPR011011), Zinc finger, PHD-finger (InterPro:IPR019787); BEST Arabidopsis thaliana protein match is: RING/U-box superfamily protein (TAIR:AT5G41580.1); Has 1186 Blast hits to 1185 proteins in 207 species: Archae - 0; Bacteria - 6; Metazoa - 654; Fungi - 241; Plants - 117; Viruses - 0; Other Eukaryotes - 168 (source: NCBI BLINK).</p>	<p>MCESSGNSNARKQNCSSGSSRTQPREEAESN ELERKLRCPGSDVPGGGPLVQCLDPDCNVWQ HSSCVIPEKLEKISMPIHFLCEICRVKRADPFWK TMAHPLYPCKLGSALSIDGTNPHMILEKNFKLTK ADRELLQKAEHEIQVWCILFNDRVYRMOQWPLY AELLVNGKPVRTVSRATGQLLGANGRDDCAKIM AYIVEGVNKISLSGTDARFCFGVKLVRRTLEQ MSGSRMRTAGRRACAHMGCDFLOTFLVGLNER SRKWQCPICLNKYCLEVDVDPFFNLILKMQRC GEEIHEIDMKPDGSRVRYKNAREVGDLAQWHFP DGLIQVSSISQTAAREHVSERNTSKLTQDNVANW HAYVQGSDEVGNIDQNIITMSSTTTGDKDDGDP SINQDVEQFGVSDMNEPEITSGENCNLPFAFAD VIVLSDSEDDVNLVHIPETSHVCHTKNEECLISS PGENLNSYPDVASDEGPGIFRGSAGDFLIPQSPT PSFLPEGSGDFFGTSGPLLDEHNSLTSITYPRP DNCELTAASAVDPAAHVIPSYDPLNLSFDHQFAS ANGTIPLEALPCTLISGNELPILNSSPGEDWISLK TGANGNESSSNVEAHQESADENGLDLQINQFR SKEDSLLQRTREGESGSRSDIAASRKRSDGPFPTFP RQQRQRRTTCVQTSKDLASKOS</p>
898	sp_144590_rgox #N/A	#N/A	-3.891672	<p>Plant Tudor-like RNA-binding protein; FUNCTIONS IN: RNA binding; INVOLVED IN: biological_process unknown; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 10 growth stages; CONTAINS InterPro DOMAIN/s: ENT (InterPro:IPR005491), Tudor-like, plant (InterPro:IPR014002), Agenet (InterPro:IPR008395); BEST Arabidopsis thaliana protein match is: Plant Tudor-like protein (TAIR:AT2G25590.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).</p>	<p>MMRFKKGTKVEVLTQKDVPSGSRCAEVGGN GHNSYKVEGDVELSGNGTVERVSRKAIRPCPP SYDVLQDWSVEDVVEVLHNFWSKMAIVLKVGL NYLLVRLVGVSTVELKANKFDVRRARLQWQDGEWA VIGKGNKGEDSRFRIPSSRRKGNQSKKIDAIWKR PTKGYQISTNDNALESHVYVPSRRLTKRQSHCH FKIEAYAGPQKTRVEKRDNNMRHQLNLFVPRQTQ LPDKVDDVASSRMEMLHEKGMHPSFTNKVTGYSE LWEREKDGVGCSRSISLESDDDDNSVSSSVAS CSVYGNYSHDFHRVFATGPGYEDNDNSQNSDAESF SOLRPEENSFLPAKEELTEELHRLHLEHAYRCTIG ALHASGPLSWEQELLNLRHLSHISNDEHLKEIR NLVSSSTCIPSS</p>

10

20

30

40

【 1 3 6 7】

【表 1 4 - 3 6 8】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたペアとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
899	sp_145200_eggu	-0.853665	-0.853944	#N/A	uncharacterized protein	RHEQKTRAGGEERKERREEGKKEEGTAAQPG RRDSEAEKGTAGRGDGGCRVAVYAEIIMNSS VFADITLKFVHGGRFRKRVGVGELVYMGKGRTV AVDPDELFCYFDLVLAKKCGYKAGLYYRVQGG ISLEDGLRQFNDEEVAEIGELVWKQRTIELVYVHC DDEIVHPDTPDKHSQPLTHPSPFAGNAKLTPKR APPKSLMSPTKSSPRRQFQTTASNTNDNVTLSS PLKKTQLASTLSKKDYFSTKYVASSSLVEPSLYK QIENPDVNTAAQNTNLSQNTNSTYDSSIPSDY EWEDPRPESPKYNELISESDDDHDPLYPDS GVLFDHSEDDLLNYEGEEGLNLEFEEGEHEEE NGRSRFRGGQDDVYGFAGRGSRRGGGGCA ATGRGKGRGRGKGNQVPGVGVYMGADGT PFSNVS
900	sp_145430_gltw	#N/A	-0.882057	#N/A	isopentenyltransferase 1 (IPT1); FUNCTIONS IN: AMP dimethylallyltransferase activity; ATP/ADP dimethylallyltransferase activity; INVOLVED IN: cytokinin biosynthetic process, secondary growth; LOCATED IN: chloroplast; EXPRESSED IN: 6 plant structures; EXPRESSED DURING: 4 anthesis, C globular stage, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: tRNA isopentenyltransferase (InterPro:IPR002627); BEST Arabidopsis thaliana protein match is: ATP/ADP isopentenyltransferases (TAIR:AT3G19160.1); Has 7944 Blast hits to 7761 proteins in 2606 species: Archaea - 0; Bacteria - 5388; Metazoa - 167; Fungi - 152; Plants - 293; Viruses - 0; Other Eukaryotes - 1944 (source: NCBI/BLINK).	MAPTASGKSRLSVEVASRYPSEIINSKIQVYKGL DITTNKITLEEQLGVPHLLGNIDSVHPELSAAEF RSRASSIIEISGRGNLPVVGGSNSFYALLANE FNKDLNVFSESGAQVCSSELRNYNCCLIWIDVSLPV LDEYIKRVDDMDSGGLNELAKFHQYDMLHLEP NVGLTKSIGVPEFKEYVSRVYGNEMFSSKGDQKQ KDLYDKAVNLMKENTCKLAKACMEKIKFLKQRG WDIKRLDATKSFMALLOGSKNWSWTWETQVLSV VGLLQDEVEVSMVKVEKAPLESFADIGGLDPOIQ EIKEAVELPLTHPELYEDIGIKPKKGVILYGEFGTG KTLAKREAGAYNRERTGNLTIERGKGGQETVL VKKEKSPKLNPTLELRKLLCELGDHLFTGTCYFSP KNHMTKDRPRKVPVSA
901	sp_145440_nqwd	-0.8552435	#N/A	#N/A	uncharacterized protein	MNSERTAITKLSFTTEKNKNTDEKEEEETSITGT DP

10

20

30

40

【 1 3 6 8】

【表 1 4 - 3 6 9】

1	A	B	C	D	E	F
		分析に用いたペアトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)		注釈	aa seq
						SFKTRDHRDLSPAVQYSSTAVTLRRDSPLPEN ISEKFAGKHFSPGASSMAKLETSKSRVTKSL AYFDLSLTPSSSFLSLPSFTCVAVATLSDSOLV YAGTDSGALLLTLNPDLDPSSTAAQIDADYK VKFLRSPIRSALKSHIVDQIGRVLVISDGFLLV DLQLQPVKLILGKLVNVAARRVCSSTFGSSN WFHEEGVNSSSVGGRLMKKLSVGRSINGVVK EFDLRGDNNDGACLFVVTMGKVVFIENLISIK HDRVVDGORGFSVVLKEMPCIDGKVTMWLDD SVILGTDNAYSLMSCNTGQVSLIFSLPDMASPC LKLLSKDKVLLLDNAGITLDSNGQPVGGSVFG CAPHISIAELSTYVVVAGDGKIMELYHKRSGTCVQ TFSGGNFGSSKCYLADEEYVNGRFVVAASSK VFCYRKIAPEEQIKDLRKKFRFRAISLVEEVE GDLREMLSFVHACVGFLLFELHFKAVDNFL SETMQPSEIFPFVIRDPNRWSSLIPRNRVWGLHP PPVHIEDVINDGMLAERAPFLKAGVDVAVNNEF LLNSPNRDDILDSAYRNIIRYLVQMRNKNINQOVK EGVDTLMLYRILKWTYMEKLASSENSCVIEE LESLEESEGLHRLTFLAYANKGMSSNALAIWRVL ARNHPGTHODTAETYHLSDSCSPVSSRETAAT EASKILEESADVESLOHLGWISEVNEELAVRVL SCKRISPLPDDVVAIDPRKIEILORYLOWLIED QDSTDTHTFLYALSLAKTALLESLETDSQHANH KSMREPIYVSGRILPSENPVRRLOIFLHSSDSY DAEEVLEIEGSELWLEKALYRKGGQETLVLOILA LKLEDEAAEQYCVLEGRPDAYMQLDMMYDPO NAFDSSITQRSLSPDMPQLQASETVLRMLRARLH HYHOGKIVHSLSRAVDVDRARLARLEERARNVLN DDSLDCSCHARLGTGLFAMYPDDSVICYKCFRRL GESTSVARRNFKDALIKPGWLVLR
902	sp_146110_fhiw	-0.92358	-0.96714	-0.939046	<p>Vacuolar sorting protein 39; FUNCTIONS IN: small GTPase regulator activity; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant_structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Citron-like (InterPro:IPR001180), Vacuolar sorting protein 39/Transforming growth factor beta receptor-associated domain 1 (InterPro:IPR019452), Vacuolar sorting protein 39/Transforming growth factor beta receptor-associated domain 2 (InterPro:IPR019453); Has 401 Blast hits to 372 proteins in 133 species: Archaea - 0; Bacteria - 2; Metazoa - 168; Fungi - 105; Plants - 74; Viruses - 0; Other Eukaryotes - 52 (source: NCBI/BLINK).</p>	

10

20

30

40

【 1 3 6 9】

【表 1 4 - 3 7 0】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	sp_107620_kqnh	sp_107660_klqg	sp_107670_ptqx	sp_107670_ptqx	注釈	aa seq
	共発現遺伝子	(SOAP2)	(SOAP2)	(SOAP2-like)	translocase inner membrane subunit 17-2 (TIM17-2); FUNCTIONS IN: protein transporter activity, P-P-bond- hydrolysis-driven protein transmembrane transporter activity; INVOLVED IN: protein transport, protein import into mitochondrial inner membrane; LOCATED IN: mitochondrial outer membrane, mitochondrial inner mitochondrial inner membrane, mitochondrial inner membrane presequence translocase complex; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Mitochondrial inner membrane translocase complex, subunit Tim17/22 (InterPro:IPR003397); BEST Arabidopsis thaliana protein match is: translocase inner membrane subunit 17-1 (TAIR:AT1G20350.1); Has 35333 Blast hits to 34131 proteins in 2444 species; Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	
903	sp_146150_zkdp	-0.321300	#N/A	#N/A		MGTPETSRPEPCDRILLDDLGGAFAMGAISSVVF HYVGRSINAPKGERILRALETVRMSAPRTGGGFA VWGGLFSTFDCTMVIYRQKEDPWNISIFAGAATG GFLQVRRQGLRASGRSAVFGGILLALIEGAGIMLN RVVSPPAQIMYEDPMSGAGAPQMGQYGFQFPG QARAPANVGGVLESFDTPTMPNFDYK MGVCSKMKRGRVYRDKDPRMFQKMKLQQLQEI SAILKQREETEVEYERELMLFAFRETEWKRDRKK LREEVKGLRRSLEDREQRIRMEVEPEGITDKSEQ DNHTTHTSTTTTTTTTTSSSFLLENMREERATRDE AVEKWKTLYLAIKHELDHLIQRTHQGTLYWRADD EDLMEDLQKELKAKDETIQVLKQQLASMEDEYK RKREVDILRQSLRIMSINSNKGPTSSAKVKKLST NSIRFTTKLAFVN
904	sp_146630_dzru	-0.306451	#N/A	-0.383797	unknown protein; INVOLVED IN: N-terminal protein myristoylation; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G23930.1); Has 4658 Blast hits to 3455 proteins in 431 species; Archae - 39; Bacteria - 387; Metazoa - 1842; Fungi - 436; Plants - 133; Viruses - 38; Other Eukaryotes - 1783 (source: NCBI BLINK).	#N/A
905	sp_146991z5722	#N/A	-0.651386	#N/A	uncharacterized protein	#N/A

10

20

30

40

【 1 3 7 0】

【表 1 4 - 3 7 1】

	A	B	C	D	E	F
1						
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
906	sp_147060_zmpi	-0.874351	#N/A	-0.855187	Protein kinase superfamily protein; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro:IPRO17441), Protein kinase, binding site (InterPro:IPRO17441), Protein kinase, catalytic domain (InterPro:IPRO00719), Serine/threonine-protein kinase domain (InterPro:IPRO02290), Serine/threonine-protein kinase-like domain (InterPro:IPRO17442), Protein kinase-like domain (InterPro:IPRO11009); BEST Arabidopsis thaliana protein match is: Protein kinase superfamily protein (TAIR:AT1G79640.1); Has 122716 Blast hits to 120970 proteins in 3289 species: Archae - 117; Bacteria - 14305; Metazoa - 44836; Fungi - 12007; Plants - 31268; Viruses - 500; Other Eukaryotes - 19683 (source: NCBI BLINK).	QLTSPFRLLPPPSGVLASSSALRRSDCLLLYIS CIITSTRCLCHCRFLWETFNLSFSAAIIFGVGR QYIFILLLSLWCWFKKALTLGVLLHYLXDAKDYKLY EEVGEVSATVYRALCIPFNEVVAIKVLDLERCN NDLIDGIRREVHTMSLDHSNVLRAHCSFTAGHSL WVWMPYMAGGSCLHIMKSTYEGFEPEVIATLLC ETLKALLYLHAGHGHRRDKAGNILLDTNGTVKLA DFGYSACMFDTGDRQRSRNTFVGTPCWMAPEV MQQLHGYYDFKADIWVSGITALELAHGHPFSSKYP PMKVLLMTLQAPPGLDYERDKRFRSFKDMVA ACLVKDPKRRPTSEKLFKHPFKHAKSYEFLARTI LDGLPPLGERFRVLYKVEADLLVQNNALNGDKE QLSQOEYIRGISAWNFNLEDLKSQAALIQDDSS STAENPDLGLNRMIDLDDQSFPTDKPSSEGTNH SNTSNSHEDGFDDLDLSEGLSASFPIQLQALKG YEDIGEDMKSIGYKKNASLPDAKVHFQKLTENEQL GRTDGDLVFRSSLPQFVGTQKKVLSGPLLYE NASSWVKQIGDGRDNMQPRHPSDRNYSGPS YRQRDTNNNTLVEDSSEGSVWQRKGRFQVMS DATPKGSTCLSSVPGSTAGLTPPYLSASAILPTLQ YILQQNNLRDEITKLIKHVELDNLNLCESADAMSNG ISQTLPIVKEGELLTQVMQLQQRIGSLVDELQIR QKLNKQLEKQLNAFVGRPEKTKD
907	sp_147390_mnam	-0.936176	#N/A	-0.855819	uncharacterized protein	DEL
908	sp_147820_pwmt	-0.870473	#N/A	#N/A	Ribosomal protein L4/L1 family; FUNCTIONS IN: structural constituent of ribosome, rRNA binding; INVOLVED IN: translation; LOCATED IN: ribosome, intracellular, chloroplast, large ribosomal subunit; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Ribosomal protein L4 (InterPro:IPRO15498), Ribosomal protein L4/L1e, bacterial-type (InterPro:IPRO13005), Ribosomal protein L4/L1e (InterPro:IPRO02136); BEST Arabidopsis thaliana protein match is: ribosomal protein L4 (TAIR:AT1G07320.4); Has 8236 Blast hits to 8236 proteins in 2694 species: Archae - 70; Bacteria - 5573; Metazoa - 121; Fungi - 129; Plants - 100; Viruses - 0; Other Eukaryotes - 2243 (source: NCBI BLINK).	MAKRILASLHHLLRPHALHLPGLVYSSCR:JSTTT VTSNSNDEIPSDLMPKKTAFIPDRNQGVYQDV VIPVTNFHEENKGLMSLAGDVFDPVPIRKDVHRV RWQLAKRQQGTHSTKTISEVSGTGRKPHPOKGG TGRARQGSRRSPHMRGGSTMHGPKPRSHAFKL NKVYRRLGLKIALSARAAEGLKLVIFDDLEVPHTKT KNIVHWVNGLDDSKLLLVDDGVINEKLIKLATQNI HYVNLPSIGLNVYSIQHDTLVMSRAAVNRIVER MNTPIKFR

10

20

30

40

【 1 3 7 1】

【表 1 4 - 3 7 2】

	A	B	C	D	E	F	
1							
2	共通遺伝子	分析に用いたペイトとの相関係数 sp_107620_kqnh (SOAP1) sp_107660_kiqg (SOAP2)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	
909	sp_147880_qyhp	-0.986383	-0.972512	-0.978486	Leucine-rich repeat protein kinase family protein; FUNCTIONS IN: protein kinase activity, kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: endomembrane system; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Serine-threonine/tyrosine-protein kinase (InterPro:IPR001245), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: Leucine-rich repeat protein kinase family protein (TAIR:AT2G37050.3); Has 172111 Blast hits to 130582 proteins in 4641 species: Archae - 151; Bacteria - 14946; Metazoa - 47791; Fungi - 11011; Plants - 76783; Viruses - 518; Other Eukaryotes - 20911 (source: NCBI BLINK).	#N/A	#N/A
910	sp_148251z5778	-0.88267	#N/A	#N/A	uncharacterized protein	#N/A	
911	sp_148610_dwxx	#N/A	#N/A	-0.824233	ORMDL family protein; FUNCTIONS IN: molecular function unknown; INVOLVED IN: protein folding; LOCATED IN: integral to membrane, endoplasmic reticulum; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINs: ORMDL (InterPro:IPR007203); BEST Arabidopsis thaliana protein match is: ORMDL family protein (TAIR:AT1G01230.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 947; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MEKMMYVRAEASKEVKNKNTEWVTPGVVWTTYIF ILFFSWLLVLSLLGCSAGMAWTIVHLSHFFVYTH CFHWKKGTFPSDDQGGYNGLTWWEIENGQQQL TRNRKFLTIPVVLVLIASHTTDYQNPMLMLNTLA VFVLVAKFPRMHRVRFINGEN VYKESRKYVPGRTYLTNPEKIEENIEKIDALESTQ DGGGNSSEDLISEVIQGGPKRKRKSRVPLYGKG VTKSLDKKMGKIRILDSVRVLAEYENRVGASTC TPRCVYDCISTRSR	
912	sp_148750_yecd	#N/A	-0.915728	#N/A	uncharacterized protein		

10

20

30

40

【 1 3 7 2】

【表 1 4 - 3 7 3】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
913	sp_148970_rkyg -0.687244	-0.607765	-0.645044	SU(VAR)3-9 homolog 4 (SUVH4); FUNCTIONS IN: double-stranded methylated DNA binding, methyl-CpNpG binding, methyl-CpG binding, histone methyltransferase activity (H3-K9 specific), methyl-CpNpN binding; INVOLVED IN: maintenance of DNA methylation, histone methylation, peptidyl-lysine methylation, histone H3-K9 methylation; LOCATED IN: nucleus; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: SET domain (InterPro:IPR001214), SRA-YDG (InterPro:IPR003105), Pre-SET domain (InterPro:IPR007728), Post-SET domain (InterPro:IPR003616); BEST Arabidopsis thaliana protein match is: SU(VAR)3-9 homolog 6 (TAIR:AT2G22740.1); Has 1807 Blast hits to 1807 Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MEVVVEVGDKVAWAEGRRTRDRTRGRISYNETI LAKIEEVSPPKRYQYQINGGKRTKTVKVEQKESTE MEMVRGLLGVDDGLTGLGFKLEVEEDGNGV VEKSDHLRVKETLRFNNYYLHLIQDEENRCKLIE AEKQKQAKGVEDKRIH-KRPDLKAITKMMNTG KVLVPSKRFGAIPGVGVGHQFFSRAEMAVGFH SHWLNIGDYMGTSTVSTGEYSEYTFPLAVAVLSG QYEDDLDNSDDVVTGGGNNLLGDKKQKQDQ KMLRGNLALKNSMEQDLFVVRVGHSSKSYVG KVVYDGLYKVVNYWAEKGISGFTVYKFRKRLR GQPELTTNQVQFIRIPIDSVSEVRGLVCADISE GOENIPATNLVDDPPVPTGKYRKSQIUVDNV KIPAASSGDCVGGCTNPNICACARLNGSDFFPV HRDGRRLIEAKAVFECGPKCCGGSSCINRTSQ KGLRYRFEVFRTPKKGWGVRSWVYIPAGAPICE YTGFLKFSDEGDSVLENNIFIDIDLQTMKGLDG RQMLRLDVSIPISDGIHEKPEEKVDSWVVEYCIDA GVTGNLARFINHSCEPNLFVQCVLSAHHDIKLAR VMLFAADNIPPLQELTYDYGVALDSVIGFDGKVK EMACYCGAEGCRKRLF
914	sp_149110_oyql #N/A	#N/A	-0.633679	SALT OVERLY SENSITIVE 3 (SOS3); CONTAINS InterPro DOMAIN/s: Recoverin (InterPro:IPR001125), EF-HAND 2 (InterPro:IPR018249), Calcineurin B protein (InterPro:IPR015757), EF-hand-like domain (InterPro:IPR011992), Calcium-binding EF-hand (InterPro:IPR002048); BEST Arabidopsis thaliana protein match is: calcineurin B-like protein 8 (TAIR:AT1G64480.1).	MGC5IKRSKANYWDFASLASETFMSLDELVEAV YVLVYKVCVDDGLHKEEFROALLRMKGPDL FVDRVFDLFLKRVGVLDFGFEVQSLSFHPKSP AADKILFTNLVDLQHTGCIEREVEKEMVLAIKE SDLEPLDEVEIEVENAFNEADSNQDKIDLEEW KEFVGKPNLSLLKMMITPYLMKERNRFDKSAVLA QTHCAFHLVDVWDTGYEPEDEVVMILAFLESD LLLSDDIQTIQDKTFEEADFKGDGKIDLEEWV ARNP5ILKNMTPYIKDIPTGTFPFRHRNDESE KMY
915	sp_149150_laof #N/A	-0.854253	#N/A	GRF zinc finger / Zinc knuckle protein; FUNCTIONS IN: zinc ion binding, nucleic acid binding; INVOLVED IN: biological_process unknown; CONTAINS InterPro DOMAIN/s: Zinc finger, CCHC-type (InterPro:IPR001878), Zinc finger, GRF-type (InterPro:IPR010686); BEST Arabidopsis thaliana protein match is: GRF zinc finger / Zinc knuckle protein (TAIR:AT5G13920.1); Has 109 Blast hits to 86 proteins in 31 species: Archaea - 2; Bacteria - 6; Metazoa - 13; Fungi - 4; Plants - 79; Viruses - 0; Other Eukaryotes - 5 (source: NCBI BLINK).	MELTPRSAGSSSSNSRASLIEIDVKGNHNITAVLR TVKKGSRNIGKRFYGCPLWPRGDCGFQWEC SGCAELQSKLSMKKELTVEEMVEKMKAKKVG HEEIVRELKAELSHTRIELMRSARVEKYAAMAVLL SWVIVVILMLIN
916	sp_149930_uprms #N/A	-0.905206	#N/A	uncharacterized protein	MMVAESVPAADAAVAADDVPAVTVLLRFCCSCC KGCSCWRIKSLLLSFFYFLFHGFLYFLSYLG

10

20

30

40

【 1 3 7 3】

【表 1 4 - 3 7 4】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
917	sp_150040_honf	#N/A	-0.851359	#N/A	alpha/beta-Hydrolases superfamily protein; LOCATED IN: endomembrane system; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; BEST Arabidopsis thaliana protein match is: alpha/beta-Hydrolases superfamily protein (TAIR:AT1G23330.1).	MKRLWSSLQLVELFVSSWHLLYGFYIFSTAVA YDLNQLCDYFKPNSINQIEAAAPPQDSSVDDL PPIVLVHGIFGQGLGGLSYFAGAEKKDERVL VPDLGSLTSIYDRARELFYLLKGGQVDYGEESHK ACGHSQFRIYEQGHPEWDEDEHPIHFVHSAG AOVVRVLOQMLADKAFEGHDTSPNWLITSLS GAFNGTTRTYDGMQPEDGVTLKPVCLLQCLRL GVIYDWFDPWLKSYNFGFDHNL.SWRKIGVW GLVDCLLGNAGPFASGDWILPDLTIQGSMLNSH LQTFPDTYFSAKTRTROLLGFTAPSGIMGIHPL LFIKQMCWRLPADPPPYKGYRDEDDWWDN DGLNTTSMYPRFPVEHPHLLVNDKDCRPLQ PGIWIYKIVEGDHILFVNRERAGVQFDLIYDSIFE RCRKHYFRKFPQTL.PNETPQTVTS
918	sp_150990_rzcx	-0.924826	#N/A	-0.898037	PETER PAN-like protein (PPAN); CONTAINS InterPro DOMAIN/s: Brix domain (InterPro:IPR007109); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI/BLINK).	MGRFRFRKRVVYKPNVKNQQLPTVDQITGKRP QSFVSRTKLPVLRQLEMDLRLKMLPFTALKLK EKKRNLSKDFLNVAGPMGVTHLILSKTNVGPYL RVAKTPOGPTLTKIHEYSLAVDVARSQLRPRCP QDLKNSPLLVL.SGFGTGEELTLITRMFQHMFP SIDVNTKALCQRIVLVHNKESKRIDFRHSYIRL QPVGVSRRLRKFVQNHQVPLRNLQDVSDFYTR AGYGSSEADDEAAATVLSVLDLGRVNRASATKSA VKLQEIQRMTLELIKIEDGMCQGMLF.SNDGKC GGKRTQEDENDEEEEEESQEEEEEDDEKEEG DEEE
919	sp_151700_phuf	-0.865175	-0.976623	-0.883659	Pentatricopeptide repeat (PPR) superfamily protein; CONTAINS InterPro DOMAIN/s: Pentatricopeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: Pentatricopeptide repeat (PPR) superfamily protein (TAIR:AT5G55840.1); Has 45179 Blast hits to 13990 proteins in 287 species: Archae - 4; Bacteria - 20; Metazoa - 378; Fungi - 562; Plants - 43035; Viruses - 0; Other Eukaryotes - 1180 (source: NCBI/BLINK).	MNRLSSFAAPRILL.SNYSTKSALTPSSTSFNPP1 VPNSKPSLNP.LYSFLPETQNPNLNLICLALKKT NFQITHLANQLGDLIPHLGTQEISRVLLRCQESQ SALIFFNWVKNDLGVKPNCHNYGIIVHILVWSRNF VHGKMLSELIRFQSESEKVIDIFQGLVSCSEDC NWDVPVFDMLIKAYMKDILIKEGYRVFKMKVMK GYVPHVAFNCLL.SGLSKANCVDNCWEL.YKEM WRIGINPNTCTFNIFTNVLCCKDQSDVKVNEFLK MEEEGFVPLDVTYNTLIGNYCRSRRLLDDAF.YLKI MYRRGVPLDVTYTSMLNGLCIKGKLEAKHQVF HRMVDRGLTPDIVSYNSLHGYCIEGRMKDSRLL LHDMIAKISPDFTFQVLFVGFVKAGSLRSLN MIVELKRFVGTSSDDVYKLV.LV.SGENHPAAVKL LDPMLEDGHESDVEIKLIKFCDSLELLEGALLH KSKMLERRVKPC.LAIYQNLISLCLRLRRCLEA.VL MQEMDLFDLKPNNICRALIFAYCTENNFDK.AES LLLSLSIKYQLYD.TDSYVNLKIRSEQSSITELMEL QDKMLKLG.FAPNGLTFKYVVLQGLQNTVMKDKGS FQAC.SQLQ

10

20

30

40

【 1 3 7 4】

【表 1 4 - 3 7 5】

A	B	C	D	E	F
1	分析に用いたバイトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
920	sp_152040_toyr	-0.940164	-0.88137	<p>LOCATED IN: cytosol, plasma membrane;  EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Beta tubulin, autoregulation binding site (InterPro:IPR013838), Misato Segment II, myosin-like (InterPro:IPR019605), Tubulin/Fisz, N-terminal (InterPro:IPR019746); Has 345 Blast hits to 341 proteins in 161 species: Archae - 0; Bacteria - 0; Metazoa - 131; Fungi - 140; Plants - 55; Viruses - 0; Other Eukaryotes - 19 (source: NCBI/BLINK).</p>	<p>MREVVYIQVGGFAMFIGSHFWNFODELLGLASD  HMDPVPQNPQVMDVLYRSGETQEGVLTYP  RLLSVNYQGSLSGSMSSHGTLNQTPFLPSPVAT  WKGSIATORSESLRNRLFQSLYKEEQKESVGG  DSKHADSPDNIGDQEIEMESLENGTQFWTDFSKV  HYHPHLSYEINGLWMDVVEGFDNYGAGKDVFAES  SRGFEEMTERLRFVEEEDHVGIOCVDDSSGGF  SSVAADFLOQIADEYNTVPVLLYTARDPGSFMSS  KSRKRNRISRNLDVAFKLSPFCKLIVPVGLPFL  GGSKASTYLRYNDEKPYHGSVYAAALHALSLPF  RIDSLSGPTATSDCTLGALDFNSIVQMLSGOSRON  MVAIIDTTMPAPTIGEQLLSKQLTLPFISDDMED  LHSVETLVIHGALKSDEORASVFEVODASAAVE  QASSRRPFGCHLSATTCLPLIPLFPFISFQTVGRC  GQLLSNPSSSSASKGSLDWHSPMAARLRSASV  LPFLTNRLLENLRFKGIQHGAMGELLKSWGFGK  EELEEMGETLSNMVRTLAPYSDLTSDSD</p>
921	sp_152600_cocp	#N/A	#N/A	<p>LOB domain-containing protein 21 (LBD21);  CONTAINS InterPro DOMAIN/s: Lateral organ boundaries, LOB (InterPro:IPR004883); BEST Arabidopsis thaliana protein match is: LOB domain-containing protein 25 (TAIR:AT3G27650.1); Has 1016 Blast hits to 1011 proteins in 25 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 1016; Viruses - 0; Other Eukaryotes - 0 (source: NCBI/BLINK).</p>	<p>MKCYNNDLNRSTSSCAACKLLKRRRCIPNCIFAPH  FRADEPKKFAKVHKVFGASNVSKILNEVPEEHRE  DTVNSLSYEAEARLDPVYGCIGAVVLLQRRMFQ  LQHDLALARARLAHYKASSLHTNTNNVVVANINT  DDNGLFGLDDHNTVSIITNMINLNGEELISSDQLSQ  SDIGVYNTTDFGQVDPDLLPL</p>
922	sp_153960_hwei	-0.987496	-0.957951	<p>thylakoid lumen 15.0 kDa protein; FUNCTIONS IN: molecular function unknown; INVOLVED IN: biological process unknown; LOCATED IN: thylakoid, thylakoid lumen, chloroplast thylakoid lumen, chloroplast; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 13 growth stages; Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI/BLINK).</p>	<p>MVSHVLOQGEVSKDFFEGEDASIEAAVNAISYFLR  EPVLSLNCSEVM</p>

10

20

30

40

【 1 3 7 5】

【表 1 4 - 3 7 6】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAPf)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
923	sp_153970_gmxy	-0.99671	-0.974129	-0.957377	eukaryotic elongation factor 5A-1 (ELF5A-1); FUNCTIONS IN: ribosome binding, RNA binding, translation elongation factor activity, translation initiation factor activity; INVOLVED IN: translational initiation, xylem development; LOCATED IN: cellular_component unknown; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: Nucleic acid- binding, OB-fold (InterPro:IPR012340), Translation elongation factor, IF5A, hypusine site (InterPro:IPR019789), Translation protein SH3-like, subgroup (InterPro:IPR014722), Translation elongation factor, IF5A (InterPro:IPR001884), Translation elongation factor, IF5A C-terminal (InterPro:IPR020189), Translation protein SH3-like (InterPro:IPR008991), Nucleic acid-binding, OB-fold- like (InterPro:IPR016027), KOW (InterPro:IPR005824); BEST Arabidopsis thaliana protein match is: eukaryotic elongation factor 5A-3 (TAIR:AT1G69410.1); Has 1356 Blast hits to 1355 proteins in 415 species: Archae - 264; Bacteria - 0; Metazoa - 366; Fungi - 246; Plants - 261; Viruses - 0; Other Eukaryotes - 219 (source: NCBI BLINK).	MSDEEHQFESKADAGASKTYPQQAGTIRKNGYI VIKNRPGKWEVSTSKTGKHHAKCNFVGIIDFN GKKLEDIVPSSHNCVPHVNRDYLQIDIAEDGFV SLLTESGGTKDDLLKPLTDETLKQLKEGFADGKD LIVSYQNAMGEEQICGVKDIGGGSK

10

20

30

40

【 1 3 7 6】

【表 1 4 - 3 7 7】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MSVAVGVGVNLRFSKVGSDSFTSNKQTCNRFGLAVWPOLCSFGKRSSRSYAVRRSEAKICEKFKFNRLRSELRPLVLRDPNARFVIQSALSQVPEQPLGLYDPKYDKDSOGVGEVAELSGENSRKTVDAVEMLVRMAHRGACGCETNTGDGAGILVGLPHDFFTVTKDLGFELPALGKYAVGMFFLPTSESRREEKSNVFTKVAESLGHVTLGWRSVPTDINSGLGKSAVQTEPVVEQVFLTPTPRSKADLEOQMYILRRVSMVAIRAALNLOHGGVKDFYICLSLSSRTIVYKGLKPSQLKEYYWDLGNERTSYMALVHSRFSNTTFFSWDRACPMRVLGHNGEINTLRGNVNMWRAREGLLCKELGSKNEMKLLPIVDASSSDSGAFDGVLELLVRAGRTIPEAVMMMIPEAWQNDKNMDDPKRKGFYEYLSALMEPWGGPALITFTDGRYLGATLDRNGLRPGRFYTHSGRRVIMASEVGVVDIAPEDVSKKGRLLNP'GMMLLVDFDKHIVVDEALKKQYSLARPYGELKROKLVLDIVDSVAEAERVAAPASGAMPVSNQDEDMQNGIHGLLAPLKAFGYSVEALEMLLIPMAKDATEALGSMGNDTTPLAVMSNREKLTFOYFKQLFAQVTNPIDPIREKIVTSMEGMIGPEGDLTETMEEQCHRLSLKGPLLSIEEMEA/VKMDYRGRWSKV/DITYPKTLGRKGLEOTLDRICAEARDAIKEGYTVLSDRAFSPSRVAVSLLAVGAVHOHLVKTLERTRVGLIIEAEPREVHHFCTLVGFVDAICPYLAVETIWRLQVDGKIPPRSTGEFRSKEELVKRYFKASNYGMMKVLAKMGISTLASYGKAQIFEAVGLSSEVIERCTGTASRVEGATFEMLASDALKHEMAFPRRALPVGSAESLALPNPGDYHWKDKGELHNDPLAMAKLQEAAKSNSVGYAYKEYSKRINEINKSSNLRLGLKFKVEKVEKVPLEVEVPEVSEIVKRFCTGAMSYGSSISLEAHSALAIAMNKLKQVGSNTGEGG ENP'SRLEPLSDGSRNPKRSIAIKQVASGRFGYSS
924	sp_154380_osza	-0.887376	-0.843224	-0.872245	NADH-dependent glutamate synthase 1 (GLT1); CONTAINS InterPro DOMAIN's: FAD-dependent pyridine nucleotide-disulphide oxidoreductase (InterPro:IPR013027), Glutamine amidotransferase, class-II (InterPro:IPR000583), Aldolase-type TIM barrel (InterPro:IPR013785), Glutamate synthase, alpha subunit, C-terminal (InterPro:IPR002489), Glutamate synthase, NADH/NADPH, small subunit 1 (InterPro:IPR006005), Adrenodoxin reductase (InterPro:IPR000759), Fumarate reductase, C-terminal (InterPro:IPR012285), Glutamate synthase, central-N (InterPro:IPR006982), NAD(P)-binding domain (InterPro:IPR016040), Glutamate synthase, eukaryotic (InterPro:IPR012220), Alpha-helical ferredoxin (InterPro:IPR009051), Glutamate synthase, central-C (InterPro:IPR002932), Glutamine amidotransferase, type II (InterPro:IPR017932); BEST Arabidopsis thaliana protein match is: glutamate synthase 1 (TAIR:AT5G04140.1).

10

20

30

40

【 1 3 7 7】

【表 1 4 - 3 7 8】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1) 分析に用いたバイトとの相関係数 sp_107670_ptqx (SOAP2-like)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq WATTFSSLRPSSAGSGTGNLDAINRVLSDLCIKGI PIKEGASSALRRHIEEAAHDLNGEAFSRFMDQLY DRISTLLESGEVAENLALGALRAINEIDLPLGESASR VSRFSNYRTVFEFKREREILVLSRYLGLHARAG GAMTADVEHQIRNALDWLGRFEREYRLLAAVLI LKEMAENASTVFNVHVQEFVDIWWVALRDPDLSL VREHRAVEALRACLRVIKRETRWRVQWYYCMFE ATQEGLGKNAPVHSHIGSLAVGELLRNTGEFM MSRYKEVAEIVKYLEHRDRRLVRLSITLPRAYF LDRFVTVNLYSTCMDHILAVLKIAPERATGFIALG EMAGALDGEILHYLPTIMSHLRDAIAPRRGRPSVE ALACVGSIAKAMPDMEPHIRGLLDAMFSAGLSR QLVEALDQIAISIPSLSTIQDRLLLECISLVSRSHH AQVPAAGIARGNVANTPLLVDVSGPVLVQLAL QTLAHFNFKGHDLEFGRESVTSYLEDADAVTRK DAALCCCRGLVANSISGTSSTSSQFSARSNRTGRR RRLIEIVEKLLAAVADADVSVRLSIFQSLHEDGG FDEFLAQAADSLSAIFAALENDEDFQVREYSAVAGR LSEKNPAYVLPALRRYLIQLLYLEQSVDSKCRREE SAKSLGCLIFRSCERLLPYAPIHAKLVAKLREGST LSANSGIITGLVITVGDRLRVGGFAMRQYIPELMP LIVDALLDGAATVKREAVATLGGVWQSTGYVITP YNDYPQLLGLLKLKLLNGDLVWSTREVLKVLGIM GALDPYVHKRNQPNLPGSHAEVHTTTTETGQHI RSMDDLPMDLWPSFTTSSSEDYYSYTVAINSLMRV LFPDPSLSSYHQVWVGLSMFIFKSMGLGCVPYLQK VLPDLFQINRVCDNLLKDFITWFKFTLVSIVRQHI RKYLGDLALLVAELWSSFLPVAHVSHVRLGPLYLH LVEQLVALNDEFRTYLPALPCCIQVLSDAEKCNC DYTYVPDILHTLEVFVGGTLDHEHMLLHPALIRLFK VDASVEIRRASIKTLTRLIPRVQVTHISSLVHHLK LVLDGKNDLRLKDAVDALCCLAHALGEDFSIFIPS MESKGSYAKEWVYTWKQIRETLNSENSEYLNST QIPFESAVNSVWEQLNAIVKGDGCMTPSSEKRRVVG AIYAGISLPVTEINRHLNDIAVSNMVERFFSDKN LAVTKKVVHTLAHKRAHGVIAVAVFGQYINRKYD VELVALLFSANVALEVLGSDVDEKIESKNEWP HVTIWTASGVAAKQOASTLQPLVAQGGKASRIEIP PVVIQGVVQVFF
925	sp_154470_uati -0.837395	#N/A -0.842854	-0.834322	target of rapamycin (TOR); FUNCTIONS IN: protein binding, 1-phosphatidylinositol-3-kinase activity; INVOLVED IN: embryo development ending in seed dormancy; LOCATED IN: chloroplast; EXPRESSED IN: 28 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Phosphatidylinositol 3-/4-kinase, catalytic (InterPro:IPR000403), Armadillo-like helical (InterPro:IPR011989), FKBP12-rapamycin-associated protein, FKBP12-rapamycin-binding (InterPro:IPR009076), Protein kinase-like domain (InterPro:IPR011009), PIK-related kinase, FAT (InterPro:IPR003151), PIK-related kinase (InterPro:IPR014009), Armadillo-type fold (InterPro:IPR016024), PIK-related kinase, FATC (InterPro:IPR003152), Phosphatidylinositol 3/4-kinase, conserved site (InterPro:IPR018936); BEST Arabidopsis thaliana protein match is: Ataxia telangiectasia-mutated and RAD3-related (TAIR:AT5G40820.1); Has 4645 Blast hits to 4114 proteins in 299 species: Archaea - 0; Bacteria - 46; Metazoa - 1929; Fungi - 1123; Plants - 507; Viruses - 4; Other Eukaryotes - 1036 (source: NCBI BLink).	
926	sp_154920_okfa #N/A	#N/A	-0.857419	RNA ligase (RNL); CONTAINS InterPro DOMAIN/s: tRNA ligase, phosphodiesterase, fungi (InterPro:IPR015965); Has 209 Blast hits to 196 proteins in 23 species: Archaea - 0; Bacteria - 0; Metazoa - 0; Fungi - 3; Plants - 200; Viruses - 0; Other Eukaryotes - 6 (source: NCBI BLink).	

10

20

30

40

【 1 3 7 8】

【表 1 4 - 3 7 9】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2				<p>SERRATE (SE); FUNCTIONS IN: DNA binding, sequence-specific DNA binding transcription factor activity; INVOLVED IN: in 7 processes; LOCATED IN: cytosol, nuclear speck, nucleolus, nucleus, chloroplast; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Zinc finger, C2H2-like (InterPro:IPR015880), Arsenite-resistance protein 2 (InterPro:IPR007042), Protein of unknown function DJF3546 (InterPro:IPR021933), Zinc finger, C2H2-type (InterPro:IPR007087); Has 5797 Blast hits to 4596 proteins in 474 species: Archaea - 2; Bacteria - 583; Metazoa - 2822; Fungi - 981; Plants - 712; Viruses - 137; Other Eukaryotes - 560 (source: NCBI BLink).</p>	<p>MAEVLNMPVEPIDNRRRNKDKDEPPTDSSPPP PPPPNRRDRDRSRRRDDDFDRPPNRYD RNRRGGGGGGRRDRDYNINNNNNKRNRL SPSPHPPPPYRDNRNNYSPPRRSPYPPFK RSRRDEGGYDGRRRSPRGGYGGYGGDRRFGYD HPGGYDREMGGRPYADDPHGRFRSFGYPG GPSDWSGRGGYGDSSNPNQAQREGLMSYKQ FIQELEDVLPAAEERRRYQYKSEYISTQRTYFE SHKDEDWLNKHYHPANLLTVIERRNELARKVAD FLDLHSGSLDMGSPSANASTMKNKSGQTSEPNSE DETDVNGKRRRHRGAQAKAHEADISAAPKAHLVSS EPRRIQIDIDLACALVTKLDESKRIEDNILSKDDTD KMSRRSHSGSIGPVIIRGLTSVKGLEIELDLTL ATLVWRVHGVYDYYMIETAEPKGLRHVRAEGKN SDVSTNAAEWEKKLDLSDHWQDRLRSLDPLEVMT AKEKIDAAAVALDFVVKIRDEKYGWKGCGAK SCTKLFAHAEYVHKHLKLPHELSDATTKVREEI YFQNYMNDPNAPGGQPVMQQFVPKDKLQRRRP GADNRLKDRGRNRRDRRNGGGERFERPENPQA DFPNKDGPKGNSDEPIDNFQGGGMHVAPFA TDIAPPVLMPPVPGAGPLGPFVPAPEVAMRML REQGGPFPFEGPGRNGMGAPPIILPPSFRQDPR RLRSYQDLDAPEDEVTVIDYRSL</p>
927	sp_155680_xsno	-0.873036	-0.940877	-0.95187	<p>MASASATAMVEDPNFODDQLANLTTEDIARASRI LDNEIRILKEDVQRTTLELDSFKKIKENQEKIKLN KQLFYLGNIVEILEMNPDEAEEDGANIDLDSQR KGGKWLKTSTRTQITLFPVVLVDPDKLKPGLV GVNKDLYLIDLTPSEYDSRVKAMEVDEKPTEDY SDIGGLEKIQIQLVAVLPMTHKERFQTIGIRPPK GVLLYGPPGTGKTLMARACAAQTATFLKLAGP QLYOMIFGDGAKLVRDAFQAKAKAPCIFIDEIDA IGTKRFDSEVSGDREVRTMELLLNQLDGFSSD ERIKVIAATNRADILDIPALMRSGRLDRKIEFPHT EEARARILQHSRKMVHPD/VNFEELARSDDFN GAQLKAVCVAEAGMIALRRDRATEVNHEDFNEGIIQ VQAKKASLNYA</p>
928	sp_156090_qzmp	#N/A	#N/A	-0.800587	<p>regulatory particle triple-A ATPase 5A (RPT5A); CONTAINS InterPro DOMAINs: ATPase, AAA+ type, core (InterPro:IPR003593), ATPase, AAA-type, core (InterPro:IPR003959), ATPase, AAA-type, conserved site (InterPro:IPR003960), 26S proteasome subunit P45 (InterPro:IPR005937); BEST Arabidopsis thaliana protein match is: 26S proteasome AAA-ATPase subunit RPT5B (TAIR:AT1G09100.1); Has 33265 Blast hits to 30927 proteins in 3145 species: Archaea - 1437; Bacteria - 12044; Metazoa - 4947; Fungi - 3672; Plants - 3322; Viruses - 56; Other Eukaryotes - 7787 (source: NCBI BLink).</p>

10

20

30

40

【表 1 4 - 3 8 0】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_king (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
929	sp_156270_kwyn	-0.334847	-0.324382	Protein kinase superfamily protein with ociticosapeptide/Phox/Bem1p domain; FUNCTIONS IN: protein serine/threonine/tyrosine kinase activity, protein kinase activity; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: cellular component unknown; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Ociticosapeptide/Phox/Bem1p (InterPro:IPR000270), Protein kinase, ATP binding site (InterPro:IPR017441), Protein kinase, catalytic domain (InterPro:IPR000719), Serine-threonine/tyrosine-protein kinase (InterPro:IPR001245), Protein kinase-like domain (InterPro:IPR011099), Serine/threonine-protein kinase, active site (InterPro:IPR008271); BEST Arabidopsis thaliana protein match is: Protein kinase superfamily protein with ociticosapeptide/Phox/Bem1p domain (TAIR:AT1G16270.2); Has 119780 Blast hits to 118332 proteins in 4796 species: Archae - 124; Bacteria - 12838; Metazoa - 46034; Fungi - 10270; Plants - 32306; Viruses - 451; Other Eukaryotes - 17757 (source: NCBI/BLINK).	MIDWIKVAGDEVCLEEEESLLFEQNLISGALLIG MDSNSITSNINNNMGNLGNDEMVSPPRVKFL CSFGSILPRPQDGKLRYYGGGETRIVSVPRDIGY EELMGRMRELYESTAILKYQQPEEDLDALSVVN DDDVTMMIEEYDKLGGDGFTRLRFIFSHPEQ DVSMMYMDGGGDERDERRRYVDALNSLCEP YTRQQQQQLINDSPIIMPVDELHIAQLFNSLN LEGSVHNQRNFDMGMAQYNLRQLAIPITMGSGQ HLGPTAQRYTEEAQAPWSPAFYSPRCAGPHDT RTFTYPSPPSSARYRMPFQEMPKCVDRMPD DATTLLQVNSQIPDPQQYSDNVAVVPTGGM PPMDRAGFPGNLNGHSVLDLDRPSVCEQCKMTFQ RNPTSSDYHWKAEQGQCHVEQSTVANGFPVHGN GCVECLPNREPVMINTDPKLOHSIPGVGRMSEH YIRAGAEVGEVFNHQAAGSQVHYVLEDRGV HAANLPYSYGTDPHQVNSHPSAQTTWVFNVPTP VHVSVPYDPSNLQLRNGIPSPVPRANVEASPRF FVGVDSGSPWGESRSRLGVEGLATPESPYGHF SKLSAGSPLQEVHPTFPQDPSRNTPNMGNIPRP TELTMINENMLSSGARDHTSVQREDAGNAGVPS DERSNHHELNETIPLRKAELLDVDFGICTEQNNN AEKHSITVTSLNAMKLNLDTKPAEEIKCGIELGETNT CLPTEKADLSLPELVASAKKAKLESVERVKARA KEDGCNTARVANIDETASDPDVGNSGEGSDENI DTSKIEHTKADEALAQGLQIYNADLEIRMLGS GTYGAVFVHGKWKGSDDVAIKRIKASCFAFPSPER ERLIADFWKEAFLLSSLHHPNVVSYFYVIRVDPD GT LATVTEFMINGSKQYLRKKDRITDRRKLIIA MDAAFGMEYLHGKNIVHFDLKCENLLVNMMDPH RPVCKIGDLGSKVKQHTLVSGGVRGTLPWMMAP ELLSGKSNMYTEKIDVYVYVSGVIMWELLTGEFPYS
930	sp_156450_izyj	-0.373736	#N/A	Transcription factor IIIC, subunit 5; CONTAINS InterPro DOMAIN/s: Transcription factor IIIC, subunit 5 (InterPro:IPR019136); BEST Arabidopsis thaliana protein match is: Transcription factor IIIC, subunit 5 (TAIR:AT3G49410.1); Has 680 Blast hits to 642 proteins in 201 species: Archae - 0; Bacteria - 19; Metazoa - 202; Fungi - 159; Plants - 80; Viruses - 10; Other Eukaryotes - 210 (source: NCBI/BLINK).	MGVIEDGTAKGCLPSEE-LFVHYPAYPSPMSRA VETLGGIDAIAKARSSKLNKLELFRPEPCSHPA FGYIHSNNFLIKISGNQNNNCSSSQSKDCSS GKYENLNVLSAEIVARVPEAYTFTGMVDYQHVLP VHVDVARRKGRCDCEVQPFKGGDLGDAEDEL IMMLVPSLFSKDVPEENVLRRPSTITLISKKKEGVV CHRWEMEIEPCLDAIDFNKIDILT

10

20

30

40

【表 1 4 - 3 8 1】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	分析に用いたペアとの相関係数			注釈	<p>global transcription factor group B1 (GTB1);  FUNCTIONS IN: transcription elongation regulator  activity, hydrolase activity, acting on ester bonds, RNA  binding; INVOLVED IN: chromatin assembly or  disassembly, transcription initiation; LOCATED IN:  cellular component unknown; EXPRESSED IN: 25  plant structures; EXPRESSED DURING: 13 growth  stages; CONTAINS InterPro DOMAIN/s: Resolvase,  RNase H-like fold (InterPro:IPR006641), Nucleic acid-  binding, OB-fold (InterPro:IPR012340), Nucleic acid-  binding, OB-fold-like (InterPro:IPR016027),  Ribosomal protein S1, RNA-binding domain  (InterPro:IPR003029), Transcription elongation factor  Spt6 (InterPro:IPR017072); BEST Arabidopsis  thaliana protein match is: Transcription elongation  factor Spt6 (TAIR:AT1G63210.1); Has 59515 Blast  hits to 32300 proteins in 2671 species: Archaea - 146;  Bacteria - 13496; Metazoa - 17051; Fungi - 6893;  Plants - 5011; Viruses - 737; Other Eukaryotes -  16181 (source: NCBI BLINK).</p>
931	sp_156470_anam	-0.310793	-0.337635	-0.3341236	<pre> MIGKAVVSVDEEELELVQDDEREQEEQEEEGEELDDV GGGGGVDDEDEDEDEEGQDEFEFDDGDFIVEGD EDEEEPEEEDDEERHKKKRRKRESFELDDDDY ELLLDNNVKVAKADSKKFRLLKARRDTDAGH SGLFDDDEFLLSGGKGARTAQEQAKHTLFGDDDG GPVEDMPEEDEQPEADVDV.LADEDDMADFVSE EDEMDADGMPVRRMDKKKKHQAAGVLPSSL QDAQDIFGDVEELDMRRRANLEKDRYDENGLR KERNIGEEFEPTILAEARYMTEKDDRIDIDIPERM QISEESTGPRSTDAEAVDEAKWIKHLDGQVVP GLVNIASGEEMEYHIRNFLNFHYQKLDLPYAMY RKEDIFSLLDKDPETGEDASHDKPALKWHKVLWA VQDLDDKLLQKRSALDGYNKRFEESRRYI DETRFLNQQLFESIIGALRDSEREREVDVDDAK FNLHFPFGAEGVDEGQYKRPKRKSYSLONKAG LWEVASKFGHSSEQFGAISISLAVRVDEPDTTET PEEVASNFTCAMFDFSPQAVLKGARHMAAVEISC EPYVKHFRNNYFDNVVSTCTPTDGNVVIDAFH QFAGKWLNRNPLNKFEDAQWLLIKKAEKELLL QVTIKWPEEHLKLLIGLELLYTSLAVSKLAKPW NEQRRLLQDAIYNFILPALEKEARLLLSRAKNW LLMDYKCLWDKSVAPYQRKESDLSSEDEEAAP RVMACCGYGGKPPPTFVMLDSFGEVVDVLAAG SISLRSQNVTDQQKHKHQQLLKFMTHEHQPOV VVLGAVSLSCTRLKDDIYEIVFKMVEENPRDVGH EMDGLSVFYGDESLPRLYENSISSDQIPGHKGV VRRGVALGRYLQNPFLAVNANLCGPGREILSWKIS SSESFLEDEKYGMVEQIMVDVTVQVIGIDLNLISA HHDWL_FAP_LQFISGLGPRKAATLQKSLVRAGAV TRKDLLTAHGLGRKVFISAAGFLRRRSGGLAVSTS QFVDVLDTRIHPESYALAQEMAKDIYREIVGDD NLDDDDAIDMAIENLRDRPSALKSFNVDAYAKDT </pre>

10

20

30

40

【表 1 4 - 3 8 2】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
932	sp_157110_zdeg #N/A	-0.87876	-0.855241	protein arginine methyltransferase 10 (PRMT10); FUNCTIONS IN: histone-arginine N-methyltransferase activity, methyltransferase activity, protein-arginine omega-N asymmetric methyltransferase activity, [myelin basic protein]-arginine N-methyltransferase activity, protein-arginine omega-N monomethyltransferase activity; INVOLVED IN: vegetative to reproductive phase transition of meristem, peptidyl-arginine methylation, to asymmetrical-dimethyl arginine; LOCATED IN: cytoplasm; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAINs: Skb1 methyltransferase (InterPro:IPF007857); BEST Arabidopsis thaliana protein match is: protein arginine methyltransferase 6 (TAIR:AT3G20020.1); Has 3346 Blast hits to 3335 proteins in 890 species: Archaea - 50; Bacteria - 1167; Metazoa - 1181; Fungi - 251; Plants - 354; Viruses - 1; Other Eukaryotes - 342 (source: NCBI BLINK).	MGSPFANGGGERGLGGAANGSVDKGVYANYF CTYSYLHQKEMLSDRVRMDAYNSIQNTHF QGKTVLDVGTGSGILAIWSAQAGARKYAVEAT KMSEHARLLVKAAGLEDVVEIEGSMEDITLPEK VDVISEWMGYFLRESMFDVSVICARDRWLKP TG VMYPSHARIWLAPIKSHLIDQKRDFHFDGGMEEW HGFVDDTKNSYGVDMISILTESFEEQKYYLOTS LWNRLPEQVIGTAAIKEMDCLTATVDDILVIKTO FFSILEEDTRFSGFGWFDVHFRGRSESPAKC EIELSTAPSVENGTWGGQVFLNLSPIRSNKGDN MSISFVMNRSKENHRLLEAEFEGCEFKLASSQP LQ PFRKFFIE
933	sp_157330_kft #N/A	-0.85538	#N/A	RESURRECTION1 (RST1); FUNCTIONS IN: binding; INVOLVED IN: cuticle hydrocarbon biosynthetic process, embryo development ending in seed dormancy; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Armaclillo-type fold (InterPro:IPR016024); Has 92 Blast hits to 88 proteins in 41 species: Archaea - 0; Bacteria - 0; Metazoa - 50; Fungi - 0; Plants - 29; Viruses - 0; Other Eukaryotes - 13 (source: NCBI BLINK).	MEYAPLLVRRVYVQPSLQRYAVIQIFDKLRSTPS ILSISDFEREAITCQLSSSLAVVDQSVRELRL VKDSKLDPSFALLELOSSLEGPPRFVNVFKGI GFLVRFGCTRRFGSQNFVDAPENHPFIKVTOR SEAGSELYQQVLFVHGKOCGIAEVCDFLRPLL TFSVLRACSSDSSLSSFLSLVYSLMSLCCSLLHE ANPFLLEMLIGLCRCFNLRGAEDLAKVAIFAEFLVD AYSWLRFHIVKMGSMVTEVQLIGVKKLLDALLTCS DFEKRNIESKPVLELSRRLLSFQIELKQYVYPQFS SPMLSLCVLQAQLELEDEKLYVLDLALYFLRWRL QDENVAGAACNLSDSLIFFFPIISLMSPPSKTVN OSASOFLMLEKLLVSHSTTSRDELSVQRGSSFV SKPETHIRILOHLWLODLPSCSFLLNINAYSGNG NYEGSGKSWLSQLGKYLVSFAERHKSCSEISS SQEVMSTEISLLGALAGSLVWHQSLGSSYAIETLA ALNSMEPKLGVPLLAVLYYGNLRLRSNIKYGDELL LKLLTMLSSLASNSAMLP LIYQITLPLMLHKDNPV LYATAIRLLCKAWEINDRVFGLSRFLLPEYFVAS RFDRTIATSLAASLRDLCRKSPERGVQLLSVAACI ECSVSTIRALGFQSLGLCEADVDFYATMNVIA KYGGEYSTDPMIAYR
934	sp_158380_ikcs -0.87883	#N/A	#N/A	uncharacterized protein	MFFLHPAAFTLLHLLSARGQGGRKAQLRDSMK GFCAVQSLSPYHPHYGARR

10

20

30

40

【 1 3 8 2】

【表 1 4 - 3 8 3】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたペアとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiag (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
935	sp_158850_ykm	#N/A	#N/A	-0.836342	MAP kinase 6 (MPK6); FUNCTIONS IN: MAP kinase activity, kinase activity; INVOLVED IN: in 20 processes; LOCATED IN: trans-Golgi network, preprophase band, phragmoplast; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPR017441), JNK MAP kinase (InterPro:IPR008351), Serine/threonine-protein kinase domain (InterPro:IPR002290), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR01009), Serine/threonine-protein kinase, active site (InterPro:IPR008271), MAP kinase, conserved site (InterPro:IPR003527), Protein kinase, catalytic domain (InterPro:IPR000719); BEST Arabidopsis thaliana protein match is: mitogen-activated protein kinase 3 (TAIR:AT3G45640.1); Has 124968 Blast hits to 123484 proteins in 4565 species: Archae - 94; Bacteria - 13025; Metazoa - 47222; Fungi - 12634; Plants - 30512; Viruses - 574; Other Eukaryotes - 20907 (source: NCBI BLINK).	MDALAPQSDTDMMDAGAPPENLPPAADHS QTPPLQQQVSIPTLSHGKFKIYQNFNIFEV TAKYKPIPIGKAGYIVCSSLNSDNTNEHVAIKKI ANAFDNKYDAKRTLREIKLLRHMDHENVWAIIRDII PPPRREAFNDVYIAYLEIMDTDLHQIRSNQGLSEE HCQYFLYQLRGLKYHSANVLRDLKFSNLLUNA NCDLKICDFGLARVSETDFMTEYVYVTRWYRAP ELLNSSDYTAANDVSVSGCFIMELMDRKLFPFG RDHVHQLRLLMELIGTGPSEHELGLFNENAKFYIR QLPAYRRGSLAEKFSHVHPSAIDLVEKMLTFDPR QRITVEDALAHFYLNLSLHDISDEPICVSPFCDFDE QHALSEEQMRELIYREALAFNPEYHQE
936	sp_158850_ueym	-0.632416	-0.322812	-0.673484	nodulin MN21 /EamA-like transporter family protein; LOCATED IN: membrane; EXPRESSED IN: 13 plant structures; EXPRESSED DURING: 7 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF6, transmembrane (InterPro:IPR000620); BEST Arabidopsis thaliana protein match is: nodulin MN21 /EamA-like transporter family protein (TAIR:AT1G44800.1); Has 3616 Blast hits to 3605 proteins in 598 species: Archae - 32; Bacteria - 1860; Metazoa - 4; Fungi - 2; Plants - 1222; Viruses - 0; Other Eukaryotes - 496 (source: NCBI BLINK).	MKDFAGTPTGITGYLIAMGLQVWVSLALMDA YALAKRVLHNSLVSLFVGDWQLLHLQELRFSCT SMISSETIGWTLTYFTEASFNFHGMNPPYVITYRHH VAGVVMFPFAYFLERKAPKLTVALAEIFVLSLL GVGLALNMVYFASLKYTSPTFVASMINTIASLTFVIA VALKMEVLDIWNFRGAACKVGLTSLSGGVLAMTL YKGSVKSINWNPLLHIQGSTTTNENWLMGSILTTA SCLSWSIWYIMQAFILKRYPALLSLTTWMSFYGA ACSAVFTACVQRKPAAWAIGFNIDLLSTVYAGVV CSGLVFLQWCTEEKGPVFTMFPNPLSTILVALL AFVVLGKLYVGSILGAIVIFGMVLLMWKKEEA QTNLAGIFRNYKDSRLKPEAFITTTSEGREIETEP
937	sp_158841z6514	-0.874437	#N/A	#N/A	uncharacterized protein	#N/A

10

20

30

40

【 1 3 8 3】

【表 1 4 - 3 8 4】

I	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAPT)	sp_107660_king (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
938	sp_159350_fmnp	#N/A	#N/A	-0.913332	<p>C-terminal domain phosphatase-like 1 (CPL1);  FUNCTIONS IN: double-stranded RNA binding,  phosphatase activity, nucleotide phosphatase activity;  INVOLVED IN: abscisic acid mediated signaling  pathway, negative regulation of transcription,  response to abiotic stimulus; LOCATED IN:  intracellular; CONTAINS InterPro DOMAIN/s: Double-  stranded RNA-binding (InterPro:IPR001159), NLI  interacting factor (InterPro:IPR004274); BEST  Arabidopsis thaliana protein match is: carboxyl-  terminal domain (ctd) phosphatase-like 2  (TAIR:AT5G01270.2); Has 330 Blast hits to 317  proteins in 105 species: Archae - 0; Bacteria - 31;  Metazoa - 57; Fungi - 81; Plants - 112; Viruses - 1;  Other Eukaryotes - 48 (source: NCBI BLink).</p>	<p>MMKTVYEGDNLGVEIYFENNANKIEMMKKG  MLMRISHYSEASERCPPLAVLHTTKSTGVSFK  MMEKSLYFOQHNDQJFALHSSCLRGNKTAWS  LGEQEIHLVAMRSRMDGVTPCFWGFVMPGL  YESCLGMLNLRGLGIVFDLDELIVANTLRSFEDRI  EALORKMTVEADPORMMAGMAEVKRYQEDKAIL  KQYAEQDQVVDNGKVKIQAEVIPA.LDNHQTVV  RPLRLQDKNIVLTRINPQIRDTSVLRLRPARED  LRSYLTARGRKFEVYVCTMAERDYALEMWRLL  DPSNLIIGRELLDRIVCVKSGSRKSLFNVFQGGI  CHPKMALVIDRLLKVVWDEKQDQPRVHVWPAFAPY  YAPQAEANNAIVLCVARNVACNVRGGFFKEFD  EGLLRMSDVYFEDDPKDFPSPDYVSNYLVSSED  DGGSNANKEPICDFGMADAEVERRLKEAVLSS  SSAPSLPCVTSLATVNLDRHLASSLFFSVAASSM  TIPQAPQASIAFFHANLFSOAGPLARTLASIGPK  DLGLHSSPAREEGEVEPESELDPDTRRLLJLQHG  QDMREGLPNEPPFPGRPPVQAPVAGPQSGGPGP  VVPVGPVAVAGSASISVPGPGVPMMSGVPA  PAPVVPVPRVQSRGSWFQVEDHMNPSPLGRS  ATKEFMPSPDAVHEKQKPPPPFPKRVENPVWS  DRSPEKQRLPREASRRDERLRSNYSVPSHQSF  RGDEISLRSVSSNKGFVEPEKGGSSLSNPSVA  LHDIAMRGAKVEFKLGLVATSELKFFTEAYFVG  EKIGEGTTRREAQYRAEAALMLNADRYLTHI  KSDASTPOSTSRGSPKDMGFASDANSQGDG  TSKKEETTPSSELTRLDLSILEGSKDSMGVSV  LKELCMIEGLVVEFKGQSPSTNPNVHGDEIHAEV  EINGQVLKGGTGLTWDKAKMOAAELALASLKSM  LGQITKRPSSPRLLQGMASKRKLKPEYARVLEHMP  SSRYPRNASPVP</p>
939	sp_159420_duwo	#N/A	#N/A	#N/A	<p>Arabidopsis protein of unknown function (DUF241);  FUNCTIONS IN: molecular_function unknown;  INVOLVED IN: biological_process unknown;  LOCATED IN: chloroplast; EXPRESSED IN: root;  CONTAINS InterPro DOMAIN/s: Protein of unknown  function DUF241, plant (InterPro:IPR004320); BEST  Arabidopsis thaliana protein match is: Arabidopsis  protein of unknown function (DUF241  (TAIR:AT2G17070.1); Has 750 Blast hits to 734  proteins in 87 species: Archae - 2; Bacteria - 37;  Metazoa - 79; Fungi - 28; Plants - 534; Viruses - 2;  Other Eukaryotes - 68 (source: NCBI BLink).</p>	<p>MASYHTRFSFPNSHPVADELDQAQLRRLRSSQ  AASSTSLTSKGLNDLYKCVVEEFLQLPQNKTV  SQSQGENVVEQVLDGSLRLLDIVSTSRD/LALSK  ERLQDQSVLRRRCSSGEIDITNEVSEYKTRKASK  KIIKCKLDINAVDKTNESIAIESMLKDVQTLTDIF  KSLLSYISGQKSSWSFKLRQGSKDKAEATSISD  FEAVDATLSQKNKANVDMSQLVKLESEIQEIDEV  LELFRHLIKTRATLLNV_LSN</p>

10

20

30

40

【 1 3 8 4】

【表 1 4 - 3 8 5】

A	B	C	D	E	F
1					
2	sp_107620_kqmh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
940	sp_160190_qfxc	#N/A	#N/A	uncharacterized protein	MFLGESSGSGFRVGGGGRKGGARAGARGSAHG GARGSPHITQSTQDTFVDDDMTQGEYDSEFVPE TQQDEEVVEPQVFWMTGELWFDHSSLARHITN VIK
941	sp_160250_wrxk	#N/A	#N/A	Peptidase S24/S26A/S26B/S26C family protein; FUNCTIONS IN: serine-type peptidase activity, peptidase activity; INVOLVED IN: proteolysis; LOCATED IN: integral to membrane, membrane; EXPRESSED IN: male gametophyte, pollen tube; EXPRESSED DURING: L mature pollen stage; CONTAINS InterPro DOMAINs: Peptidase S24/S26A/S26B/S26C, beta-ribbon domain (InterPro:IPR011056), Peptidase S24/S26A/S26B/S26C (InterPro:IPR015927), Peptidase S26A, signal peptidase I, conserved site (InterPro:IPR019758), Peptidase S26, conserved region (InterPro:IPR019533), Peptidase S26A, signal peptidase I (InterPro:IPR000223), Peptidase S26A, signal peptidase I, serine active site (InterPro:IPR019756); BEST Arabidopsis thaliana protein match is: thylakoid processing peptide (TAIR:AT2G30440.1); Has 9578 Blast hits to 9341 proteins in 2371 species: Archae - 0; Bacteria - 7160; Metazoa - 214; Fungi - 105; Plants - 244; Viruses - 0; Other Eukaryotes - 1855 (source: NCBI BLINK).	MAIRYTVAFSATIAGNIASSATYRVTNRCVHEFV SKPRPTVNTTPAISNPNLEPSKNYNHYPYCKNSQ PDARRSKSPFPPRQRRTAFSTFAELITGNSSSS SPVYVGLISLVNAAVFGGSISSGSMGTMGVTLRIG ASIPFLQVSKWLPQNEPALMRSTSSDSDVDRGGTS NILEESKIKKVFESRNSAAVDVSKVIAQKVIDRK SWLSKMFNFSDDDAKAVFTAVTVNLLYKSSLAEP RSIPSSMCPTLDVGDRLAEKVSVFREPEVSDI VIFKAPPIQEIIGFATGDDVFKRIVAKAGDCVEVKN GOLLVNGIVREEIFEILEPLAYEMEPVIVPEGVYFV MGDNRNNSFDSHNWGPLPVKNIIVGRSVFRYWP PTKLTDTITVDQHAKEKFNVAVS

10

20

30

40

【 1 3 8 5】

【表 1 4 - 3 8 6】

I	A	B	C	D	E	F
2	sp_107620_kqnh (SOAPT)	sp_107660_king (SOAP2)	sp_107670_ptqx (SOAP2-like)		注釈	aa seq
942	sp_160320_mtxu	#N/A	-0.87471	-0.838423	TBP-associated factor 5 (TAF5); FUNCTIONS IN: transcription regulator activity, nucleotide binding; INVOLVED IN: regulation of transcription; LOCATED IN: nucleus; EXPRESSED IN: guard cell, root, inflorescence, cultured cell, leaf; CONTAINS InterPro DOMAIN/s: WD40 repeat 2 (InterPro:IPR019782), WD40 repeat, conserved site (InterPro:IPR019775), WD40 repeat (InterPro:IPR001680), G-protein beta WD-40 repeat, region (InterPro:IPR020472), WD40 repeat-like-containing domain (InterPro:IPR011046), WD40-repeat-containing domain (InterPro:IPR017986), WD40/YVTN repeat-like-containing domain (InterPro:IPR015943), WD40 repeat, subgroup (InterPro:IPR019781), TFIID subunit, WD40-associated region (InterPro:IPR007592); BEST Arabidopsis thaliana protein match is: Transducin/WD40 repeat-like superfamily protein (TAIR:AT3G49660.1); Has 114463 Blast hits to 42274 proteins in 991 species: Archae - 68; Bacteria - 11258; Metazoa - 46869; Fungi - 25620; Plants - 15010; Viruses - 3; Other Eukaryotes - 15635 (source: NCBI BLINK).	MEEEEEKAVLAVLKKKFRQTEVAFQEEQQQLQ QOQKHPOQQQSKNSNSHSDPDVAKLILFTS ESELPAQYDGYGSKLRSAHNSLDLYKHLLR VLYPVFHCFMDLVARFGHLOEARFFIFREDHET MHERDLQKLEGLVSPSHLKEMEFAHTRQSKVNI KICQYSHELLQLYLKIQSMILGVVNEHIFQVCPG QFGISDDADAVTLVGGQDAAANLNQKEIQWGM LEDSIEDRLKGGGLLSDSKVDGENKEGENEE NKKRSADGGKGAAPYKSKKDKVGGAGGKAAR PEVITPLAPRVKPELTPAMPTEVEHSILEDLRNR VQLSSVALPVSFYTFINTHNLNCASISHDGSLV AGGFSDSSLKVWDMKLGQQMDNTALQGENDT AGNEHMHANGGKIPYTLFGQHSVPVYSASFSP LGDFILSSADSTVRLWSTKLNANLVGYKGHNY VWDVQFSPYGHYFASCSHRTARISMDRIQPL RIMAGHLSDVDCVFWHVNCNYIATGSSDKTVRL WDIQSGECVRFIFGHRSMVLSLAMPDGRYMAS GDEGAINMWDLASGRCITPLVGHVSCVWSLAF SGEGSLASGSADCTVKLWDVNASTKTPRSEEK TGSSNRLRSRLTPTKSSPVVYSLMFSRRNLLFAA GVLSKRS
943	sp_160380_zosc	#N/A	#N/A	-0.838414	unknown protein; Has 158 Blast hits to 158 proteins in 77 species: Archae - 0; Bacteria - 0; Metazoa - 104; Fungi - 0; Plants - 33; Viruses - 0; Other Eukaryotes - 21 (source: NCBI BLINK). Protein of unknown function; DUF599; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF599 (InterPro:IPR006747); BEST Arabidopsis thaliana protein match is: Protein of unknown function, DUF599 (TAIR:AT5G24600.2); Has 303 Blast hits to 303 proteins in 73 species: Archae - 0; Bacteria - 107; Metazoa - 0; Fungi - 0; Plants - 182; Viruses - 0; Other Eukaryotes - 14 (source: NCBI BLINK).	MARRNHTPLGCIACDELSELGAGTEFWLINPNS ILSALDTHSIALSNKSLILFGWSDDFDNIKVIHP ELSPIDAEYITAEIWLVFDDIRVLAVGISSGYLLIYS LDGGLIHKIYPTIRILKLRVGRGTRKRDISQDISNEE VCIVMNGVIARFDGSDLODMLRFRWRETQTNFW DQKRKRD.LDDFEDSFERYVQLWSVKNYGAC ADAAITGVWPPPLMDFEVLVY
944	sp_160490_zngu	#N/A	-0.873385	#N/A		YSQSPVEFAAIQDKNFANFSQSYLLNPPRTDLE FQKIKISETSSDPFYTLNLRNLSHASFYTLIAD ASSLSIRSFAGKVSAGKVSARKTTRYAHISF LLNVPESSEDEGEFIEFVSOQLNRGCLFWSLGL RAFVFAPVLLWIFGAIAMFACSCSMALLFYLDL ASSPSPSYLNGVNVKTRNSPDIESRRVH

10

20

30

40

【表 1 4 - 3 8 7】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2					
945	sp_160690_zgoh	#N/A	#N/A	villin-like 1 (VLN1); CONTAINS InterPro DOMAIN/s: Gelsolin (InterPro:IPR007122), Villin headpiece (InterPro:IPR003128), Gelsolin domain (InterPro:IPR007123); BEST Arabidopsis thaliana protein match is: villin 4 (TAIR:AT4G30160.1); Has 3095 Blast hits to 1746 proteins in 186 species: Archaea - 0; Bacteria - 6; Metazoa - 2441; Fungi - 102; Plants - 219; Viruses - 0; Other Eukaryotes - 327 (source: NCBI/BLINK).	GLEIWCVENLSLVPKSSHGKLFSGSVYLINLTV LLKDGMLHDVHWYVWLGKDANEVESLWSDKALE LDLALGSCTVQYQEKQGGESAKFLSYFKPCIIPC EGVYTSGPERSNRGTYHITLLACKGEHAISVKEV PFSRSLNHDDVFDLDTESKVYLFSGHNSIQERA KALEVIKYEKGGKCCQVATIEDGKFKVGDSDA GEFWSIFGGYAPIPRDSSFFKQHDVFPVKLFLI TIQKLCETSSTGTFREMLDREKCYLDCDAEYV VWMGRCTISERKTSITATEDFLRSQGRSTRSLV MVLTEGSETPKFSNFVGPQTDNPKLYEEGKG KVAEIFQQGYNIKELPEEYQCQLINCRGKLEWV GVNGDHLIMPSEGDIKLFSGDCHVYKTYQSSQ RDESILYAWLGTASIKEDRDVAISHIINMAKSTKG TPVLAQIFEGKEPDQFFWVTQRLVVLKGGTFTRYI NFIQKGFADNTYSAPSSALFRVQSSLENMQAI QVNMVAGSLNSSCYILOTEAAVFSWIGSLTSR DHLLDRMLELLNPAWQPKSFRREGGEPDEFWS ALGGKVDYPREKEIRKHIEDPHLFCVNCSEDDFK VKEIFSTQDLDLTTEDVFDLDCYTEIYVWLGSSQN VQPNLQTIQGLKYLEGDLAQGLSLETLPLVMYTE RHEPEFFTCFFDWDVSKSHMDGNSFEKRLALLR GEKLETPRNSRRTYSSEYQCSGSRKSVSST GMPRSAPSLSGSQLASSTAIAKLLSEQPVNHS STVDFHKLERNPADGADLSLVAENDDDDANL LKYPYERLTLTSDPEVSDIDLTKREAYLSKAEFLE KFGMTTRAFYALPKWRQNKLIKSLNLF
946	sp_161180_kmky	#N/A	#N/A	uncharacterized protein	IMVTVEARFSLPYLSLSSPQVTSFHYDPNLSL ALMHSDSTFSLFPFSFPCISNLPPTQTLIPPPS SSACFVPLNANPNKVQDGVSPDDKRWVFTA GPHNGSKVVLRFHLLGIDGKFSVATQVDCSQN GLCFDRKLGVIDVSHGMKWLSSGVNYLALYSA SGGKVVFGVKLVGNRGLRLVKCAVIDCRLRI SSYLSFGFLLGEMRGVRIPIRAMVKGVGSGR RLRVAKGKESNVVEGRDVKVPIGIRIPFAESNE LNKNTGNSNVSKLLPNGHACSVLDGVSSSNAD GAYKRSLLFRQDSGEWGMQFVLFNDCVHEGIK SGVMOSNKRKATSIQAISSPEKFLLDNGLDLHL VSKSGSSRMRLTCTIQVQQLVAFPDASDSQIV WISDGLYSHVMMIPGSDTKEKESGVSNDEREII QSSVTAQIFTSIQHVVPVLAANRILLGQDLSLAY GIS
947	sp_161260_nwdi	#N/A	#N/A	uncharacterized protein	MNDHEAH

10

20

30

40

【 1 3 8 7】

【表 1 4 - 3 8 8】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq ASLTLKLPKPRKGEKLTSPSGTLAATIDSLGRIL LDDTQALVWRLWKGYRDANCLFMEMLVKQDSA STSYHEYKSDYCLCLAIHAPRKGIHEVWQMRIG PRLLTIOCAKGNACPNKVQSHHMLVERLNIIPD QTQENNSTLDPNVIPGKWIKVTPIKKSAPRTR KPVYKLDGGIHLTPQNCVSIQVSTPKHQTPA ANQIQIHSNMYYDQLLATENFKDMLREASDHEALL LENTMSAMSSSKASEFSSPLFEYGSNLLPLGTGLS PKKEEREFF
948	sp_161340_pmdg -0.850334	#N/A	-0.912897	unknown protein; Has 158 Blast hits to 158 proteins in 77 species: Archae - 0; Bacteria - 0; Metazoa - 104; Fungi - 0; Plants - 33; Viruses - 0; Other Eukaryotes - 21 (source: NCBI BLINK).  Protein kinase superfamily protein; FUNCTIONS IN: protein serine/threonine kinase activity, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: protein amino acid phosphorylation; LOCATED IN: cellular component unknown; EXPRESSED IN: sperm cell, male gametophyte, pollen tube; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage; CONTAINS InterPro DOMAIN/s: Protein kinase, ATP binding site (InterPro:IPRO17441), Protein kinase, catalytic domain (InterPro:IPRO00719), Serine/threonine-protein kinase domain (InterPro:IPRO02290), Serine/threonine-protein kinase-like domain (InterPro:IPRO17442), Protein kinase-like domain (InterPro:IPRO11009), Serine/threonine-protein kinase, active site (InterPro:IPRO08271); BEST Arabidopsis thaliana protein match is: Protein kinase superfamily protein (TAIR:AT5G45430.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MERYKLIKEVGDGTFGCVWRAINQSGEVVAIKK MKKKYSWEECINLRVKSRLKMSHPNIVLKKEV IREHDLHFVFEYMECNLYQLMKDRGRPFSENDV RNWCFQVQALAYMHQRGYFHRDLKPENLLYSK DVKIADFLAREVCVSGPPYTEYVSTRWYRAPEV LLQSPITYTSAVDMWAMGAIMAELFTLRPLFPGLS EADYKICSVIGTPTENSWGHGLLSRAMNYQF PQLPGVHLATLIPASDDAISLRSLCSWDPIQRP TAAEVLQHPFFKNCFYPPSLRTRATVNRTPPSL AARGFYDQKRSARKSSGGTLPAKATNLKPHVA FGASVQRKLEMDGKDVNKNDSIKSSAVYKQPKY RPFARNSPSSIYPGRTSRGVSSETAEKLSNMSLG SGRQTMRSQSVPPPMPKAGGWHGQSNMFMGRRS QEIQTRTRTPRKVAG
949	sp_162000_iday #N/A	#N/A	-0.8627431		

10

20

30

40

【 1 3 8 8】

【表 1 4 - 3 8 9】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
950	sp_162050_cedy	-0.943451	-0.836336	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein; FUNCTIONS IN: oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, 2-oxoglutarate as one donor, and incorporation of one atom each of oxygen into both donors, oxidoreductase activity, iron ion binding; INVOLVED IN: oxidation reduction; LOCATED IN: cellular component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Isopenicillin N synthase (InterPro:IPR005123); BEST dependent oxygenase (InterPro:IPR005123); BEST Arabidopsis thaliana protein match is: 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein (TAIR:AT3G19010.1); Has 8949 Blast hits to 8896 proteins in 1015 species: Archae - 0; Bacteria - 1163; Metazoa - 122; Fungi - 1047; Plants - 5005; Viruses - 0; Other Eukaryotes - 1612 (source: NCBI BLink).	MGDYNPDFIQTEYRPNPTSTIEAKEIPVIDISSAS DEVIAEIKACEKWFQFFVYVNHGVPLKREIVEL AARFKFGLSMEEKKKVRRDEENPLGYDSEHTK NVPDWKEIFDFTVQYPTLIPASHEIGDDEIKELRN QWPDYPEFFRVACEEYVQEMVKLSRKLLELIASS LGVPADRLSGYFKDHSFVRINRYPPCPYDIAL GVGPHKDSGALTVLAQDDIGGLEVKRKMIDGAWI LVKPIPDAYINVGDIQVWNSNDKYESVEHVRVKN PKKERLSIPFFNPS-HYVMVKPLNELIDKQNPARY REYNYGKFTTRRLSDFKKLNVENIQVLFHFRIOE
951	sp_162340_oprm	-0.871027	#N/A	plant U-box 38 (PUB38); FUNCTIONS IN: ubiquitin-protein ligase activity; INVOLVED IN: protein ubiquitination; LOCATED IN: ubiquitin ligase complex; EXPRESSED IN: stem; CONTAINS InterPro DOMAIN/s: U box domain (InterPro:IPR003613), Armadillo-like helical (InterPro:IPR011989), Armadillo (InterPro:IPR000225), Armadillo-type fold (InterPro:IPR016024); BEST Arabidopsis thaliana protein match is: RING/U-box superfamily protein with ARM repeat domain (TAIR:AT5G62560.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).	IMGGNGKFLKLNFSFYRNSHSRSSKQSQSPPPP PEYICPINSLFTDPVYVGGSHTVDRLEAEVCEV HNFKPDFDGTFRPDFSTKIPNLALRPTAIPWCSK NGVALPQEKENAVRIVKELLEKEEREENGLIRP SERFLDSVKDKADPDLSHAVTELTFRANRFDTS TSSEESVIAAASPLTLPPTTRPASCWSPFHTSS SDTLVSDAEMNPNFSDDEEFISFKSCDPFEQ EQGAISLRKATRNDDEEARVSVCTPRLLLAGSLF SSYEKQVINAIAALVNLSEKSNKIKVRSQVPTVI QLLRGRLESAEQEHAAGLIFSLSEKKNRRTAIGVLG ALPPLLNCLLRSDESTRSDAALALYNLSDHSN RVKLVQFNVPPLAVLRQEKESIAGRLLVLY QLAASTEGKAAMLDGNVAVHVMMLFRKGGKGF ESESTRENCVAALYALSHGSMRFRKALAKEARAV EVLKEVEENGSERAKEKARRILVILKGRIESEDG LGEQIDWEALEIPDGGLSRTNRVAVGVPGFQKGT TEF

10

20

30

40

【 1 3 8 9】

【表 1 4 - 3 9 0】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1) 分析に用いたバイトとの相関係数 sp_107660_kicg (SOAP2) sp_107670_ptqx (SOAP2-like)	sp_107660_kicg (SOAP2) -0.97415	sp_107670_ptqx (SOAP2-like) -0.953339	注釈 PIN domain-like family protein; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF652 (InterPro:IPR006984), Nucleotide binding protein, PINc (InterPro:IPR006596); BEST Arabidopsis thaliana protein match is: PIN domain-like family protein (TAIR:AT1G26530.1); Has 697 Blast hits to 697 proteins in 237 species: Archaea - 32; Bacteria - 0; Metazoa - 241; Fungi - 212; Plants - 105; Viruses - 0; Other Eukaryotes - 107 (source: NCBI BLLink).	aa seq
952	sp_162460_epug	-0.941360	-0.953339	RING/U-box superfamily protein; CONTAINS InterPro DOMAIN/s: Zinc finger, RING-type (InterPro:IPR01844), Zinc finger, C3HC4 RING-type (InterPro:IPR018957), Zinc finger, PHD-type (InterPro:IPR019665); BEST Arabidopsis thaliana protein match is: RING/U-box superfamily protein (TAIR:AT2G47700.1); Has 1560 Blast hits to 1269 proteins in 195 species: Archaea - 0; Bacteria - 91; Metazoa - 627; Fungi - 277; Plants - 288; Viruses - 4; Other Eukaryotes - 273 (source: NCBI BLLink).	MGRSKKAPKFAVRKLVKLTQKHKQEVLPNK KKDLAANLPRNPNVSSALFFKYNTALGPYRV IVDTNFNFISQNKLDLEKAMMDCLYAKCTPCITD CVMAELEKLGQYRVALRIAKDPFRDLPLCTHRG TYADDICIVRVNQHCKYIVATCDRDLKRRIRKIPG VPIMYITOHKYSIERLPEATMGGAPRY
953	sp_162750_qchk #N/A	-0.902544	#N/A	EMBRYO DEFECTIVE 2766 (EMB2766); FUNCTIONS IN: structural constituent of nuclear pore; INVOLVED IN: biological process unknown; LOCATED IN: chloroplast, nuclear pore; CONTAINS InterPro DOMAIN/s: Nucleoparin, Nsp 1-like, C-terminal (InterPro:IPR007758); Has 235342 Blast hits to 95277 proteins in 3388 species: Archaea - 819; Bacteria - 58190; Metazoa - 65863; Fungi - 46438; Plants - 10106; Viruses - 2707; Other Eukaryotes - 51219 (source: NCBI BLLink).	MGLDETLIDDDGGGKSLSSNVSCSICLEPVT DNGRSWAKLQCGHEFHLDICGSAFNIKAMQC PNCRIEKQWL YATGCRSFTDIMDDWPHDEDL YDLNYSFMFGVQWCP L PFRSLPLLEEGEPFSTA YHDFLQQAVFAEHTAVSSSTHQCPYVAYGPI HSSSNSVGSVDGNSYNSNRWGGFSGPNEMP NSYAFTGMDPHYHWEQHSLSFSASGTRLGG DQPSAPLTPRSSRVNSDLRSRGSFVHPVNHHS SGGRVGSSTSMIPYPGSAARAPDRVQALQA YFCPPSSPPMVAWIPMSSTRSSSGHRGMSQG AAPMPSSDDGGFFVYPPGSSSTRSFPEFENP TSRPHGWERDHLASFPLNQGD RDRSSWLPHQV VGGSSSFRQRHGSERTPQNRS
954	sp_162840_gkat #N/A	-0.806751	-0.925154		MAQFTGFNFNTSPSTSPFANFSPSTAASSSTF FGGAPSSPAFGAPTSSSAFAGSSSSSSSPSPA FGFGAASSATPFSLGSAAASAPFSGFGSSSTTS VPGFGGSSSTPSSSAFGLSPSSAAPAFSFLP ASASTTTTTSAVDPPFSSAPNLFSSSTSAFSTNT TLFGANNNTTTTPTFGSSSPATPP L FASASSA ATPSSLFASASSAAATPSSLFASASSASPSLFL ASASSLAAASSGFSFLKSSSTQTATTTASPPSRF SFTPPASVAPASAAATSLFNFMNAAASSTAGAS ATKPLFGANATPTTSPVAASLFLSTPTPTTTTTTS TPSATVSSPFAATTPAASSQPAATSTAPAFAFSGV SUNSTVTTPVASASPASTSFGSGFGSSGSSLT PTASFSSFTVPAVKPTTIPASSSQSQPQQTAAITA TTGGAAATTTTTTTTTTAPATGFLASTVSAAISS STTTAVQTSSAVAGSSSGTTSVITPVTSAPTP KLPSEITGKVEEIKIEWNTELDQDRTKFKQKQAN AIADWDRILQNRDVLRLLETEVAKVWETQSNLE RQLLEIETHQCEVDKLSLEIEEAEERIKDFERNLSL LDEEAASTRDAMYDQAESVERELEHMEHQIKSII QTLNLSAQSGELDTTEGMSPLD VVVRILNQLSL MWIDEKAEFFRIONLASEGSAADRELVAAPRL WS

10

20

30

40

【 1 3 9 0】

【表 1 4 - 3 9 1】

A	B	C	D	E	F	
1						
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	
				cleavage stimulating factor 64 (CSTF64); CONTAINS InterPro DOMAINs: RNA recognition motif, RNP-1 (InterPro:IPR000504), Nucleotide-binding, alpha-beta plait (InterPro:IPR012677); BEST Arabidopsis thaliana protein match is: RNA-binding (RRM/RBD/RNP motifs) family protein (TAIR:AT4G00830.3); Has 21430 Blast hits to 18533 proteins in 865 species: Archae - 10; Bacteria - 1688; Metazoa - 9678; Fungi - 3064; Plants - 4656; Viruses - 0; Other Eukaryotes - 2334 (source: NCBI/BLINK).		MASTQQHRCVFNIPYDATEEQIICEEYGPV VSFRLVIDRDTGKPKGYGFCYKDEETSLSARRN LQYEINGRQLRVDFEAENDKGAADRNRREQGRGG PGLATNVDAAQKQVSPATLVDASFHQPIGLPIAM TASSVMAGALGGSTGKVNQNGFQNSLIGND PLTHLANMSRSQLNQVIHEIKMLSMQNKDLARQ LLVSPQLKALFQAQIILGMVTPQMLQMPNLRQ QTPNPAVPTPRSPMQHTLOGHRPPVQSLTGLPS FTPLSSVSNPLAHNQFSAPLPLHPQLRPPQPNL NQMVAPSQSTVPAHSGLSSHLSLRPQLSGLHM PPVSSSFRPVQAPLQHPGHAGIKTTQNVNLQ PTFGSHSDPQYKVAASSVRRPASVEAIKKDTPVRAK TARDTTLASRTNTLNKLASVASVDHKDSQDQM VRFQSKVMKLNNERGTFFSSAGLATPSATAYASS SGAAGQSSSLEAQNIKQAPQFPDVPDLPALQQ VLNLTPEQLSSLPQQQLQVLIQLQQVMLR
955	sp_163080_qqgu	-0.971752	-0.958712	Homeodomain-like superfamily protein; CONTAINS InterPro DOMAINs: Homeodomain-like (InterPro:IPR009057), Myb, DNA-binding (InterPro:IPR014778), HTH transcriptional regulator, Myb-type, DNA-binding (InterPro:IPR017930), Myb-like DNA-binding comain, SHAKKYF class (InterPro:IPR006447), Homeodomain-related (InterPro:IPR012287); BEST Arabidopsis thaliana protein match is: Homeodomain-like superfamily protein (TAIR:AT3G46640.2); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI/BLINK).	MEDQWEMGLPIPEDLTPISQSLISPGIALAFGITP DNTRAQFDLNLHDNCSTYSPSEENLESLSLRLFS GRSDFDCSNMRRGDSFVNDVNSGSKDDGDE NGLNRAVKRPLVWTPQLHKRFVDVIEHLGIDKA VPKTIMEMMINVEGLTRENVAASHLOKRYLYLKRIM EGLSSEGSQSNRPRDTPMFDCCDGFSGMTPMHPH PPPPQQQQCNLPLMPMPMVMVPPPPGAAYGFG HLGLPLPSSSPSGYCGLESGLAYNMFWRQRG	
956	sp_163340_lswk	-0.883523	-0.892876			

10

20

30

40

【表 1 4 - 3 9 2】

1	A	B	C	D	E	F
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh sp_107660_kiqg (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
957	sp_163450_kfrn	-0.306263	-0.373232	-0.35055	uridine kinase/uracil phosphoribosyltransferase 1 (UK/UPRT1); FUNCTIONS IN: uridine kinase activity, uracil phosphoribosyltransferase activity, kinase activity, ATP binding; INVOLVED IN: biosynthetic process, nucleoside metabolic process, metabolic process; LOCATED IN: cytoplasm; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Phosphoribulokinase/uridine kinase (InterPro:IPR006089), Phosphoribosyltransferase (InterPro:IPR000836), Uridine kinase (InterPro:IPR000764); BEST Arabidopsis thaliana protein match is: uridine kinase-like 2 (TAIR:AT3G27190.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MPEATTADYYMEKASGPHFSGRLRSPSTTSSPR ASASGSPFSASFVDSLDLPNPFVIGVSGGTAS GKTTVCDMIQQLHHRVVLVNQDSFYRGLTDEE LECVHEYNFDPDAFDETEQLLECVGKLGSGQPV QVPYDFKTHORRTDTRQVNASDVIIEGILVPH DSRVRLMSMKIFVTDADVRLARRIRRDIVERG RDINSVLEQYAKFKVPAFDVFLPSKKYADVIIPR GGDNHVAIDLIVQHISTKLGQHDLCVKYPNVYVIQ STFQIRGMHTLIRDQEIHKHDFVYSDRLRLVVE HGLGLHLPFTEKQVWPTTGSVYTGVDVDFCKKCGV SIIRSGESMENALRACCKGKIGIKILHRDGDNGK QLIYEKLPKDISERHVLLDFVLATGNSACQAIELL IQKGVPEAHIFLNLISAPEGVHCVCVKRYPYSLKIVT SEIDAALNEEYRVPVGMGEFGDRYFGTDD
958	sp_163460_kccu	-0.954412	-0.446955	-0.672334	ECERIFERUM 7 (CER7); CONTAINS InterPro DOMAIN/s: Exoribonuclease, phosphorolytic domain 2 (InterPro:IPR015847), Exoribonuclease, phosphorolytic domain 1 (InterPro:IPR001247), Ribosomal protein S5 domain 2-type fold (InterPro:IPR020568); BEST Arabidopsis thaliana protein match is: ribonuclease PH45A (TAIR:AT3G12990.3).	MSFVTAQLVQPYRDRNEGSLSIYTEFSPMADPS FEAGRPGESAIELGRIDRGLRESRAVDTESLCLV AGKLVWSRVYDLHIDNGGNLDAANVAALALM TFRRPECTLGGVYDQGEVIVHPPEVQEPRALIVHH LPIAVTFAPMTEEDNILVVDPTYEAEAIMKGRFIATI NASEDVCAIKKAGEVGIPTVMQCLQIAAVKAA DITSKIESAVKLYNTERENRSVKRLSSVALLNISSP SSTMTDCKSNSICERTVYVLOQMRKSPSFRERHISE SMCMETEAPSCRKGEDASRLMKPGSCWDFPFK GVDSEKHLKASLASCCKGSKMAIKENEKIVESKLLNLE VNSSELKRNLSASSDKATDVLPKSCEIKTLKDA VKPKSMRRRVISASNTR

10

20

30

40

【表 1 4 - 3 9 3】

I	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
959	sp_164020_cmpg	#N/A	-0.3317963	-0.937377	Magnesium transporter CorA-like family protein; FUNCTIONS IN: metal ion transmembrane transporter activity; INVOLVED IN: metal ion transport, transmembrane transport; LOCATED IN: membrane; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Mg2+ transporter protein, CorA-like (InterPro:IPR02523); BEST Arabidopsis thaliana protein match is: Magnesium transporter CorA-like family protein (TAIR:AT1G29830.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17398; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MSGRRDINSRRNAPTKEVPKSYNNKDVIPG TELWV DGLCGFEYVYVGRKRKEQKQKTSKFPYS RKLESDSHKTKTYAYGEADARQDQNGIYNPDQSC HTELGSPVSTLDGSMNDNQYKAEQYHGSYVW PIWTRLFEIVQDVQVDVWESKQFFMETEDD VVAELAAPVWERPVVWVCHVDASHQNIIS WLSNGQWLHPAVSIALRDELSRERMKHLLEYE PVRVAGLLFEILGQSAGDPINEEDDIPVLRSW CAQNFLVTAMHVKGSTSNVNLGVMEVQELIA GGYNAPKTVHEVIAHLACRLARWDDRRFRKSF GAADEVLFVNRNINHEMHLFGAILNQEVRKLSR QVIRKWSLHAREEITFELLHQLRGNAAARSLLEMI RKSTREMLDEQAVRGRFLTQDVMQSNIRAWL QDRSLRVTHNLGVFGGGLVLSIITGLFGINVDGI PGAENTPYAFGLFTGILLIGAILIGVGLYLGLKQP VSEEQVEVRKMELOELVKMFQHEAETHAQRVQ ENRQSLPPTARVQFPLDADYVIQ
960	sp_164030_dxyq	-0.866753	-0.927521	-0.854394	YUCCA 8 (YUC8); FUNCTIONS IN: NADP or NADPH binding, oxidoreductase activity, monooxygenase activity, FAD binding, flavin-containing monooxygenase activity; INVOLVED IN: brassinosteroid mediated signaling pathway, auxin biosynthetic process; EXPRESSED IN: 11 plant structures; EXPRESSED DURING: 4 anthesis; CONTAINS InterPro DOMAINs: Pyridine nucleotide-disulphide oxidoreductase, class-II (InterPro:IPR001103), FAD-dependent pyridine nucleotide-disulphide oxidoreductase (InterPro:IPR013027), Flavir-containing monooxygenase-like (InterPro:IPR020946); BEST Arabidopsis thaliana protein match is: Flavir-binding monooxygenase family protein (TAIR:AT5G43890.1); Has 11557 Blast hits to 11539 proteins in 1628 species: Archae - 52; Bacteria - 6582; Metazoa - 787; Fungi - 1560; Plants - 663; Viruses - 0; Other Eukaryotes - 1913 (source: NCBI BLINK).	MDQMFLNLPVVENGGDDHRRRCVWVNGPVVVGAG PSGLAVGACLRDQGVFVLERADCIASLWQRR TYDRLLHLPKQFCGLPKMPPEYFPEYPTKRFQ IDYLESYATHFDIDFRNECVQYAKYDETFGLWR VKTVGQGNKPEVEYVCRWLWATGENAECVV PDIDGLANFNKVIHSSEYKSGESYRGGKVLVYG CGNSGMEVSLDLSHNAQPTVCRSSVHILPREI MGRSTFELAMLMLKWIPLWLVKILLFFAWFLLG DMQKYGLRPNIGPLELKNKQKQKTPVLDIGALAKI RSGDIKVVSGIKSFNANSVELIDGETVEVDSVILAT GYRSNVPSWLOEGEFFFNKGFPKTPFPSAWKG KSGLYAVGFTRKLAGASFDAMKIAQDIDGQVWK HETKQMKLTPRRHRCISTF

10

20

30

40

【 1 3 9 3】

【表 1 4 - 3 9 4】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kirg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
961	sp_164130_mmma -0.883828	#N/A	#N/A	cell division cycle 48C (CDC48C); FUNCTIONS IN: nucleoside-triphosphatase activity, ATPase activity, nucleotide binding, ATP binding; INVOLVED IN: embryo development, ending in seed dormancy; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 11 growth stages; CONTAINS InterPro DOMAIN/s: ATPase, AAA-type, core (InterPro:IPR003959), ATPase, AAA+ type, core (InterPro:IPR003593), ATPase, AAA-type, conserved site (InterPro:IPR003960); BEST Arabidopsis thaliana protein match is: ATPase, AAA-type, CDC48 protein (TAIR:AT5G03340.1); Has 65227 Blast hits to 39412 proteins in 3299 species: Archae - 1912; Bacteria - 27659; Metazoa - 8516; Fungi - 6329; Plants - 5583; Viruses - 58; Other Eukaryotes - 15170 (source: NCBI BLink).	MKRRSSHVAFSSYWNHSMLLRRLGSCNRYES VEDVVKDLCRSRYEYKRLSAAILTKNVQEAALSEL NWIKNKDSSGRDRVNVVRSQSTSSSSSSFFESN SMSAPTGOVSNPKFDFMKSTLRETYCHGGFES DKTEESMKKKEITNVELLEKEEENDAMEDIVMH ERIDGGTGRKDLLEHGRDKVQGPMPKDFGG IMHDVLEELKYNVLIPLRYPELLNRLSQLHAGPPT GILLHGGPPGGCKTMLAQAIANEANVPLYQIAAPEI VSGVGGSEENIRDLFKARRTSPSIIIDEIDAIE KRENLOKMEGRIVTQLKCMDNIYGLALVADPK SSNCESRPSLSPAVEISVPERRQSHVLIAGTN RPNSLDPALRRPGRFSWEIELGVPDENARFEILS VLTHDLGVOEGLDKMARDTQGFVGDADLEEV RKASLHAMQTMINKKAGDNKERALEHDDNW WKNLHFHMGVEVETLYPTMADFVAVKRVQPSLK REGFSTVNPVKWEDVGGGLTLEEFQSGIINPK YPEEYGAFFREDLPAGFLLYGPPGGCKTLIAKAVA NEAGASFIYIKGPELMDKYVGESELAVRKIFSRAR TCSPCILFFDEVDALTANRGESSGVTDRVLKQLLI ELDGAEEERRGYIIGATNRPELIDRALLRPGREE LLYPLPSQOERGLILRALARKRPIDITNDVLIIEAER TACDNFSGADVAELMRRASRVSHQEMQLSKQK GSGAILWINTSHFEKALYGMSPSVNKKQREFYD SLSRFSFKPT
962	sp_164330_upic -0.94782	-0.923816	-0.97903	ENTH/VHS family protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF618 (InterPro:IPR006903), RNA polymerase II, large subunit, CTD (InterPro:IPR006569), ENTH/VHS (InterPro:IPR008942); BEST Arabidopsis thaliana protein match is: ENTH/VHS family protein (TAIR:AT5G65180.1); Has 1003 Blast hits to 995 proteins in 197 species: Archae - 0; Bacteria - 21; Metazoa - 515; Fungi - 219; Plants - 159; Viruses - 1; Other Eukaryotes - 88 (source: NCBI BLink).	MGSTFNPOLVEKLTKLNNSSQSIETLSHWCIFH MNAKLVETWDRQYHSSPREQLAFLYLANDIL QNSRRKGAEEVAFWRYLPEALRDVIRGDDDFG KNAARLNINWDERKVFVGSRGRHILKEEFVARQPE NIVKSGKQLSIKFRQSIGIMFDKIISDYETIYGGQL DEDALLSRTNTVSCLEKLIKDNSSGSSGNFSG SGFAEEFQGHAAALRECIQLTLEASFRASLVSH LRDALQEQLKLDQVRRQLQAAQLQSDQDVANVS QOQFVSPNIVAMVEERVKDDHGLSLAPPGFSTG DROANPVVYTRQVPVSDKSAAAEEDSRKSA AVAAKLTASTASAEMLSHVLTSLVSGVAFNSTD NDPHHSDKRPKLENEASSYASAPPLPPONQQP PSPILTFFPHPELSQQNQONANNSTFTNQQSAG GEQPLPSSPPPLPPLPMPPLYPMPQPFIQTAG SMNVAYSYGATLQFPPLSLGSPVSPMSPYAAP PVNNPYQPPFGSESGFVSPSSVPMAPISRQQT L
963	sp_164840_tepj -0.921455	-0.983035	-0.988522	uncharacterized protein	MAVVEVEVSPASALREKAVL

10

20

30

40

【 1 3 9 4】

【表 1 4 - 3 9 5】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
964	sp_166290_yrd	-0.813645	#N/A	Mitochondrial substrate carrier family protein; FUNCTIONS IN: binding; INVOLVED IN: transport, mitochondrial transport, transmembrane transport; LOCATED IN: mitochondrial inner membrane, membrane; EXPRESSED IN: 20 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINs: Mitochondrial carrier protein (InterPro:IPR002067), Mitochondrial substrate carrier (InterPro:IPR001993), Mitochondrial substrate/solute carrier (InterPro:IPR018108); BEST Arabidopsis thaliana protein match is: Mitochondrial substrate carrier family protein (TAIR:AT2G30160.1); Has 27351 Blast hits to 14015 proteins in 457 species: Archae - 0; Bacteria - 6; Metazoa - 11451; Fungi - 7735; Plants - 5292; Viruses - 0; Other Eukaryotes - 2867 (source: NCBI BLINK).	MATTTDATPKFHNIHELRLLSLEEQQQDLTMD PSPSHTHDGLLFWQFVAVAGCAEHMAMFP VDTKTRMOALGSCPPIKTAGVRTALRSILQSDGAA GLYRGIAMGLGAGPAHAYVYFMYEFCFKSFSG GNPNHVAHAMSGVCATVMSDAVLTMDMVKQ RLQSGGGGAYRGVWDCVKLVKKEGFGAFY ASYRTTVLMNAPFVAVHFATYEAAKKGLTEISP TADDERLVHATAGAAAAGALAAVITPLDVVKTKQ LQCQGYCGDRFTSGSIRDVVIKTIKKDGYRGLM RGWAPRMFMHAPAAACWSTYEAAKTFFQELND QKNSSSLT
965	sp_166920_sydp	#N/A	#N/A	ubiquitin-specific protease 25 (UBP25); FUNCTIONS IN: ubiquitin-specific protease activity, ubiquitin thiolesterase activity; INVOLVED IN: ubiquitin-dependent protein catabolic process; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINs: Peptidase C19, ubiquitin carboxyl-terminal hydrolase 2, conserved site (InterPro:IPR018200), Peptidase C19, ubiquitin carboxyl-terminal hydrolase 2 (InterPro:IPR001394); BEST Arabidopsis thaliana protein match is: ubiquitin-specific protease 23 (TAIR:AT5G57990.1); Has 8659 Blast hits to 7314 proteins in 254 species: Archae - 0; Bacteria - 9; Metazoa - 4447; Fungi - 1449; Plants - 1041; Viruses - 9; Other Eukaryotes - 1704 (source: NCBI BLINK).	MSWQPSLVHQKRKNGGGLGKLNLSNCSY_LNSV LOCLTYTPPIANFCLKNQHSSFCDSSEFPDRKR DCPFGIEKRIARSLSDALDTPSKIVSCLRIFAEH FKCGQEDAHEFLRYVIDACHNTCLRLLKQQQ RPRKITNGSGGESFGKTVMDIFGGSLSQSKVC LSCGTESNKVDEIMDICLEISNCSLKDAMKRFQ AEILDGNNKYKERCCKLVAACKQMSILQAPNVL VVQLKRFEGFGKIDRVIAFEELVLSFFMCKAS QDPHPYSLFATVHSGFSPESGHYYAYIKDAMG RWYCCNDSYVTLSTQOQEVLSKAYILFFTRTGOR PVSVDTAAVSNVYKPKHEMNGSNMAKSPRSSIIP KAINTKSNAPSFVKSIPLASKNNNDASAVLNKFG GITSSLGKRISGSDNGITEVQKRETVGSNGKLGH SVCKOMNGNIIHSLDNNNGNCSNGEHIWVANSQ VLVNGDAHLKNGTGLPLNASLHENHNGLSKATA GKMLDNGVHNGQNNCEIDVSGPRGKLETRETA RKQLDNGAINNGQNHKCTNVSGKRNSEDRDSCI LLGEDAQSLARVEEFKLLQKEAALVLTQTCGWS DKVYVEYMRLLKRVGVNACSSSTDKNELKMLIADA RTTIFSIQIPGSLGSLIEHIRFSFSQK

10

20

30

40

【表 1 4 - 3 9 6】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1) 分析に用いたペイトとの相関係数	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 peroxisomal adenine nucleotide carrier 1 (PNC1); CONTAINS InterPro DOMAIN's: Mitochondrial substrate carrier (InterPro:IPR001993), Mitochondrial substrate/solute carrier (InterPro:IPR018108); BEST Arabidopsis :thaliana protein match is: peroxisomal adenine nucleotide carrier 2 (TAIR:AT5G27520.1); Has 7303 Blast hits to 6606 proteins in 315 species: Archae - 0; Bacteria - 0; Metazoa - 2896; Fungi - 2326; Plants - 1372; Viruses - 0; Other Eukaryotes - 709 (source: NCBI BLINK).	aa seq MALDFESLTEATSGAVGALVSTTVLYPLDQKTK YQAEARSDHQQRVRYRNLSDLVWEAIRNGQVLSLY QGGLTKNLSVQIQFIYFYGYFSLKRLYLKKSQNFK SMGTKANLIVAAAAGALTVIITQPLDTASSKMQTS DFGKAKGLWKTLETKGKWSGAFDGLGISLLLTAN PSIQYTVFDQLKNRLLKTNNTNSDLSQPALSAST AFLGALSKTIATCLTYPAIRCKVVIQAAAEEREKAL PKSEKARKTISRVFYSIWKNEGILGFFKGLHAQIL KTVLSSALLMIKEIKITKSTVWVMMLALRKYLLTTS TRLKSA
966	sp_167360_ikor -0.8863	#N/A	#N/A	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: plasma membrane; EXPRESSED IN: 8 plant structures; EXPRESSED DURING: LP.04 :our leaves visible, 4 :anthesis, petal differentiation and expansion stage; BEST Arabidopsis :thaliana protein match is: unknown protein (TAIR:AT1G20310.1); Has 66 Blast hits to 66 proteins in 15 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 2; Plants - 64; Viruses - 0; Other Eukaryotes - 0 (source: NCBI BLINK).	MDKEKENPKSLKLPKAAASAVSQNFHSPG RDHHHNNHHAHHNNHQYSSKLTTHAHKGFSG PINYSYLPDEAREKPKKDGSEFEAHEPTSPKYSCM GOIKYKHTKNLKRKIANNNNSLIFDNDLKKKTT KSSAISKFFNISKKSMMGYPSEEHGKEGRATLD NNNSKKTNPVPSLGMNKNKFASSGRESLASFDTW AQIAPIDHDDRYSDDDQERRRDDDDSDDEEERE VIIPYSAPNLIGRSRGGGDDDDDDDDNVMRTRK EVLNWKRRMTMAPPPPLALNSKVRS
967	sp_167710_zhhd -0.84697	#N/A	#N/A	Galactosyltransferase family protein; FUNCTIONS IN: transferase activity, transferring hexosyl groups, transferase activity, transferring glycosyl groups; INVOLVED IN: protein amino acid glycosylation; LOCATED IN: membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN's: Glycosyl transferase, family 31 (InterPro:IPR002659); BEST Arabidopsis :thaliana protein match is: Galactosyltransferase family protein (TAIR:AT4G32120.1); Has 600 Blast hits to 599 proteins in 74 species: Archae - 0; Bacteria - 0; Metazoa - 168; Fungi - 0; Plants - 429; Viruses - 0; Other Eukaryotes - 3 (source: NCBI BLINK).	MEGSLPLTTAAKGERRGVGRSSASSYTSATTK PSLLLAFSLVACLYVAGRLWQDAENRLLVNL QKNSDQRPKVLTVEDKLMVLGCKDLKRIAEAE MDLTLARSGYLRNQLRHESDVKLLAVIGVYT GFGSLRNTFRGSWMPFGDNLKLEEKGVIR FVYGRSPNFGDSLDFRIDEENRQTKDFLIIEGHE EAQEELPKAKAFFSTAVQNWDAEFYKVDNDID LDLDLGLLLESRGQGSAYIGCMKSGEVAEEG KQWYEPDWWKFGDQKSYFRHASSSLLLSKLS EYIFVNSVLKTYAHDHDTVSGSWMMGLQATYIDD NRLCCSTNRHDKVCSVA
968	sp_168480_jwzx -0.91334	-0.8865	-0.837883		

10

20

30

40

【 1 3 9 6】

【表 1 4 - 3 9 7】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOA1)	sp_107660_kriq (SOA2)	sp_107670_ptqx (SOA2-like)	注釈	aa seq
				cytochrome c oxidase 17 (COX17); FUNCTIONS IN: copper chaperone activity; INVOLVED IN: response to bacterium, response to copper ion; LOCATED IN: mitochondrial intermembrane space; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Cytochrome c oxidase copper chaperone (InterPro:IPR007745); BEST Arabidopsis thaliana protein match is: Cytochrome C oxidase copper chaperone (COX17) (TAIR:AT1G53030.1); Has 294 Blast hits to 294 proteins in 132 species: Archae - 0; Bacteria - 0; Metazoa - 138; Fungi - 51; Plants - 55; Viruses - 0; Other Eukaryotes - 50 (source: NCBI BLink).	MRGVTASVGVPAEQKNPETVTKATDOEMVTKA SEQDQPKKKICCACPDTKKLRDECIVQNGEDAC GKWIEAHKQC.LRAEGFNV
969	sp_169340_gtah	#N/A	#N/A	uncharacterized protein	PRQSDATREAEFRAMESIEMKRIKINPPRRVVSS FSATKEAFHKVINDSSMAASTPQSIQQARNHAIQ AQQEGCTTFRIFDPSFGNFIPIVITRAELSG
970	sp_169550_esgj	#N/A	#N/A	sorting nexin 2B (SNX2b); FUNCTIONS IN: phospholipid binding; INVOLVED IN: vesicle-mediated transport, intracellular signaling pathway; LOCATED IN: membrane; EXPRESSED IN: cultured cell; CONTAINS InterPro DOMAIN/s: Vps5 C-terminal (InterPro:IPR015404), Phox-like (InterPro:IPR001683); BEST Arabidopsis thaliana protein match is: sorting nexin 2A (TAIR:AT5G58440.1); Has 2399 Blast hits to 2386 proteins in 279 species: Archae - 13; Bacteria - 100; Metazoa - 1366; Fungi - 555; Plants - 146; Viruses - 0; Other Eukaryotes - 219 (source: NCBI BLink).	MMGSENEEGAHPAPSDEMETLILDDPLSSSSSF ANYKSALISLSDSHNHP LSPSIVAKSVGSDPLLGP LHEEDQIRPKVEYDFSHFEDSDDYSDVDFDPIITG GIGGGNGVGGSPKVNRSRFSRSESLSGEYLNI TVLNPQKEHEVSSNSVPGGNTVYTYLITKTNM RGFGGDFSVRRRFRKDVVT LSDLSESYRGGFIA PRPDKSVESQVMHKQEFVEQRRGALEKYLRLK AAHPVIRTSDLRVFLQVQGLPLPTSTDVASRM LDGAVNLKQLFGDSRSVAPHEAVQPAKGGRD LLRLFKELKQSVTNDWGNRSPAAVEEDKEFELEK KELLDFENRLTDVSKQAESLYKAAQQDMGDTMG ELGJLSPKLTKEFENEAILGQTRATEMKNIATAT VRGSRMYRELTAKTVKNLDTLHEYLGVMLAAHN AYSORANALLTVQTLSEVSSLSQSRAEKLEAASS RVFGSDKTRNHKLEELKETHMTKMTEDSKNCAVREY ERIKENRTELELEREKHTDFLKMVLFVNNQV NSAEKMANLWSKVAEDTRGYVKGSS
971	sp_169750_peix	#N/A	#N/A		

10

20

30

40

【 1 3 9 7】

【表 1 4 - 3 9 8】

	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)		auxin response factor 2 (ARF2); FUNCTIONS IN: protein binding, sequence-specific DNA binding transcription factor activity; INVOLVED IN: 7 processes; LOCATED IN: nucleus, chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Aux/IAA-ARF-cimerisation (InterPro:IPR011525), Transcriptional factor B3 (InterPro:IPR003340), AUX/IAA protein (InterPro:IPR003311), Auxin response factor (InterPro:IPR010525); BEST Arabidopsis thaliana protein match is: auxin response factor 1 (TAIR:AT1G59750.2); Has 35333 Blast hits to 34131 proteins in 2444 species: Archaea - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLink).	aa seq
973	sp_169770_doaq	-0.86823	-0.915353	#N/A	root FNR 2 (RFNR2); FUNCTIONS IN: oxidoreductase activity, copper ion binding; INVOLVED IN: oxidation reduction; LOCATED IN: chloroplast, thylakoid membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Oxidoreductase FAD/NAD(P)-binding (InterPro:IPR001433), Ferredoxin reductase-type FAD-binding domain (InterPro:IPR017927), Riboflavin synthase-like beta-barrel (InterPro:IPR017938), Ferredoxin Reductase (InterPro:IPR015701), Flavoprotein pyridine nucleotide cytochrome reductase (InterPro:IPR01709), Ferredoxin--NADP reductase (InterPro:IPR012146); BEST Arabidopsis thaliana protein match is: root FNR 1 (TAIR:AT4G05390.1); Has 6645 Blast hits to 6645 proteins in 1680 species: Archaea - 12; Bacteria - 3566; Metazoa - 779; Fungi - 733; Plants - 567; Viruses - 0; Other Eukaryotes - 988 (source: NCBI BLink).	MAHSAIAKVSVAVPVSNESAVKSGTFRKTYVSFG SKSWVPGYTLDLKTNLRLRKQSVVCMVSQDEAA ITKVAVSPLELEDAKEPPMHVYKAKGPTGTIVSV ERVVGPAPGETCHVIDHAGNVPYWEGOSYGV PPGENPKPGNPHNVRLYSIASTRYGDYFDGKT ATLCVRRALYYDPETGKEDPAKMGVCSNFLCDS KPGDKVQIAGPSGKVMLLPEDPNATHIMATGTG VAPYRGLRRMFEMEDVPAYKFKGLAWLFLGVAN TDSLLYDVEFSKYLKDPDNFRYDRTALSREHQNK KGGKMYVDKIEEY/SDEIFKLLDNGAHYFCGLK GMMPGIQDTLKRVAEERGESWDEKLSQLKKNK QWHEVEY
974	sp_169880_deaj	#N/A	-0.852407	-0.812916		

10

20

30

40

【 1 3 9 8】

【表 1 4 - 3 9 9】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたペアとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 Tetratricopeptide repeat (TPR)-like superfamily protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; EXPRESSED IN: fruit; CONTAINS InterPro DOMAIN/s: Pentatricopeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: Tetratricopeptide repeat (TPR)-like superfamily protein (TAIR:AT4G36680.1); Has 20922 Blast hits to 6391 proteins in 198 species: Archae - 3; Bacteria - 20; Metazoa - 73; Fungi - 116; Plants - 20269; Viruses - 0; Other Eukaryotes - 441 (source: NCBI BLINK).	aa seq LVANSRIEKFEEPALIKLKKERDPDKLFLKFKDN SHNRLVIENKYAFEDTVSRLAGAKRKYDYIEHLEH QKSLPQGRREGFVVRVIMLYGRAGWMQAVDTF YGMHLFGCRRTVKSLWALKVLSQTHNLEAINSF LKEVPWEYGIQLDVISNIVINAFCCGMGILDKAYLV MVMEKAGIKPDVSYTTLISAFYKYGRWEVGN GLWNLWVYKRCYPNVATFNARIQFLVNKGRW QANDLLRMENIELQPDDEVYLVKIKGFFRAGLFD MAKKVYSALHGRGYEPNLKIYQTMVHYLCKAGD FNCAYTMCKECLKKNWIMNVDITICTLLTGLKSTG KVEKANIFSLVQKRHRPPFPARQLDTKAMLSRR LQIKQSEGEFGFSGDVYFGGGGAGTGGGRAGM ASFRQOOQQ
974	sp_170130_kxis	#N/A	0.877316	0.863482		
975	sp_170260_iqun	#N/A	0.874113	#N/A	uncharacterized protein	
976	sp_170360_uufl	#N/A	#N/A	0.838284	CONSTITUTIVE EXPRESSER OF PR GENES 30 (CPR30); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: negative regulation of defense response; LOCATED IN: nucleus, cytoplasm; EXPRESSED IN: 8 plant structures; CONTAINS InterPro DOMAIN/s: F-box domain, cyclin-like (InterPro:IPR001810), F-box domain, Skp2-like (InterPro:IPR022364), F-box associated domain, type 1 (InterPro:IPR006527), F-box associated interaction domain (InterPro:IPR017451); BEST Arabidopsis thaliana protein match is: F-box associated ubiquitination effector family protein (TAIR:AT4G22390.1); Has 1743 Blast hits to 1730 proteins in 49 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 1741; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLINK).	RAHPAKIAGKTLIRFKVLSTFWHSTVSDPKFAKYY LGLTSRHHFLFSPETKFLLLPYDEFNNLKOLEKI DIGGRRDVRVFGCCNGLVCLYNLNLVLSKRNYG SRYRRCGFLSICNPVTHONREITNPGKSELALGY WFVYASSIDDYNIVALFSSFRVRYNEQYVHFQVF SMRTGIMWKRYCHLDDYDYVYKMFSCKTTVMNDI MYWHPGCLGHEWEDKRICAFDVLNVEEWKEYPW LDFLNEYFGADFFVRNGCLSLYCHKSHMVDVVW TLKHYYDRNSWEFLFSVNFNGDWRFRFYTSIPTKIL VRSLLWLFKIFDVSNDRPEQGDARSLSLDYLDIRAE SFVETLTSFPGTINFSKDDNEY

10

20

30

40

【表 1 4 - 4 0 0】

A	B	C	D	E	F
1	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)		
2	共発現遺伝子			注釈	aa seq WWWGNWVWVPPFPTGTSIPFPFWVQVSTVYLC PPQQQEESEPAEAEARLEEKARKWMLNSK RYGDKRKFGEVETQKEDMPPEHVRIIRDHGD SSKKFRHDKRVYLGALKFVPHAVYKLLNMPMP WEQVRDVKLYHITGAIYVNEIPWVVEPIYLAQ WGTMMIMMRREKRDRRHRFRFPFDEEPP LDYADNLDVDPLEPIQLELDEEEDSACTWFD HKPLVTKLINGPSYRRWHLSPIMATLYRFAGQ LLSDMVDNRYFLDFMESFFTAKALNMCIPGGPK FEPLYRDMKEGDWNEFNDFNKLIIRSPLRTEY RIAFPHLYNRRPKVRLCYHTPMVWYKSEDPD LPAFYDFLIHPTSSHKDRDKKAYEEDDDDFEL PEVVEPLLGDQLYDTTAAAGISLLCAPPFNMR SGRTRRAEDIPLVSEWYKEHCPSYVVKVRSY QKLLKCFVLNHLHRPKAQKKHLFRSLAATKF FQSTELDWVEGLQVCRGGYNMLNLIHRKLN YLHLDYFNLFVKVTLTKERKSRFNGAFHLCR EILRLTKLVNANVQFRLGNVDAFQADGLQYIFS HVGQLTGMRYRYRMRQIPMKDKLKHLYYRFN TGPVKGPGCGFWAPMWVWVWVLFLLRIVPLLE RWLGNLLARQEGRHSKGVAKTVTKQRVESHF DLELRAAVMHDVLDAMPEGIKQNKARTILQHLSE AWRCWKNIPWYVPLVPIENMILRYVKAKAD WWTNVAHYNRIRRGATVDTKTCRKNLGRLTR LWLKAEQERQNYLKDGPVYVTPPEEAVAIYTTVH WLESRFKSPIFPFPLSYKHDTKLLALALERKESY SVAVRLNQQPREELGLEQAYDNPHEALSRIKRH LLTQRAKFEVGEIEMDLYSYLIPVYIEIPLKITA YLDQYLWYEGDKRHLFPNWVKPADTEPPPLVY KWCCQGINN.LQGIWDTSDGQCVWMLQTKFEKFE KIDTLMNRLRLILDHNIADYVTAKNNFASFVYQY YGLVLDLILLGLTRGSEIAGQPMPNIEFMFTWDT
977	sp_170680_hsig #N/A	0.303532	#N/A	ABNORMAL SUSPENSOR 2 (SUS2); INVOLVED IN: nuclear mRNA splicing, via spliceosome; LOCATED IN: chloroplast, membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN's: Mov34/MPN/PAD-1 (InterPro:IPR000555), Pre-mRNA-processing-splicing factor 8, U5-snRNA-binding (InterPro:IPR019581), Pre-mRNA-processing-splicing factor 8 (InterPro:IPR012591), PROCN (InterPro:IPR012592), Pre-mRNA-processing-splicing factor 8, U6-snRNA- binding (InterPro:IPR019580), PRP8 domain IV core (InterPro:IPR021983), PPO, C-terminal (InterPro:IPR012984), RNA recognition motif, spliceosomal Prp8 (InterPro:IPR019582); BEST Arabidopsis thaliana protein match is: Pre-mRNA- processing-splicing factor (TAIR:AT4G38780.1); Has 828 Blast hits to 729 proteins in 312 species: Archaea - 0; Bacteria - 2; Metazoa - 333; Fungi - 220; Plants - 69; Viruses - 2; Other Eukaryotes - 202 (source: NCBI BLink).	

10

20

30

40

【表 1 4 - 4 0 1】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_piqx (SOAP2-like)	注釈	aa seq
978	sp_170710_zfhk	-0.826153	#N/A	Y-family DNA polymerase H (POLH); CONTAINS InterPro DOMAIN/s: DNA polymerase eta (InterPro:IPR017061), DNA-repair protein, UmuC-like (InterPro:IPR001126), DNA-repair protein, UmuC-like, N-terminal (InterPro:IPR017963); BEST Arabidopsis thaliana protein match is: DNA-directed DNA polymerases (TAIR:AT5G44750.1); Has 8088 Blast hits to 8015 proteins in 2292 species: Archae - 210; Bacteria - 5635; Metazoa - 500; Fungi - 401; Plants - 163; Viruses - 3; Other Eukaryotes - 1176 (source: NCBI BLINK).	MPYAKPASSESRIAHLDMDCFYVQVEQRKQPQ LRGLPTAVQYNEWKGGGLIAYSEARKYGVKR SMRFGDEAKSVCPQLVQVPVARGKADLSIYRDA GSEVISLAKKGRPCERASIDEDDDTEVKERVKWLQ RSDADHNDKLLSCGALIVAEIRKHVYLKETEFTCS AGIAHNKMLAKLASGLINKPAQOQIVPSASVKGFL GPWPIKMKQLGGKLGSLQIDLGINTVGDLLQF SEOKLOERYGNTGTWLNWIARGVNGEEVKGRI LPKSHGSGKTFPGPRALRSIDTVSHWLKQCEEL SERLLADLQNKRIHTLTLHVSAYKSSDDTDSQK SFPKSCPLRYGTTKIQEDVLSIFHAGLREFMGT GAKSDKNQPGWGVGTGLSVSASKISVPOGTC KKFFPSQCAVCTSSSETDNNELHERMOPVPRD WNLDVSAISIGREDNAVMDHDSGLHONEDKRSQ EDQGEATSFSLDKATTSCS-LDQTTQTCSLLGQ ITDGRTLQLGREGSQSQNDQDEPOKFLDKPEA CLAVSPAKIWDREYKKSPPKALINQGTASILRFL KRPEQSCSP_LKPEPNGTLKDTSVPTCSGGSSLLQ SCGENQKFRERWSYDVHEIDPTIMEELPADIQEEL RSWMRPLKQHOSAKRSSITSYFLPSKDSS
979	sp_170780_demq	-0.9228563	-0.881375	sodium:hydrogen antiporter 1 (NHD1); FUNCTIONS IN: potassium:hydrogen antiporter activity, sodium:hydrogen antiporter activity, INVOLVED IN: sodium ion transport; LOCATED IN: integral to membrane, chloroplast; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Na+/H+ antiporter NhaD-like (InterPro:IPR018390), Divalent ion symporter (InterPro:IPR004680); BEST Arabidopsis thaliana protein match is: Na+/H+ antiporter 2 (TAIR:AT1G49810.1); Has 2659 Blast hits to 2653 proteins in 794 species: Archae - 96; Bacteria - 1787; Metazoa - 39; Fungi - 48; Plants - 76; Viruses - 0; Other Eukaryotes - 613 (source: NCBI BLINK).	MVLLFVIGYTGIFEEESLAFNKSGVGLLMAVSLWA VRWIGAPSNEVAMSELEHATAEYSQIVFFLLGAM TVVEVDAHQGFKLVTEITRKPRTLLWVVGFTV FFLSAILDNLSTIVMVSLLRKLVPPEYRKLGA VYIAANAGGAWTPIGDVTMTMLWIHQJSTLPTM KALIPSVSLAVPLALMSLTSEVNGKGLSSVVM ASEQMAPRGLVFAVGIGALVFPVFKSLTGLPP YMGMLLGLGLVLLTDAIHGSESERQKLVPOA LSRIDTQGALFFLLSVASLEAAIGILRELANLYLD AHIPNIEIASAIGVFSVSAIDNIVPLVAATMGMYDVT SFPKSEFWQLVAYCAGTGGSMLIIGSAAAGVAF MGMEKVDFFWYFKYKVSQGFAGYAAAGIATYAL HNLDISLPLTTLTAHIPFLSG

10

20

30

40

【 1 4 0 1】

【表 1 4 - 4 0 2】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
980	sp_170840_tgxn	-0.952884	-0.988794	-0.961889	SUPPRESSOR OF RPS4-RLD 1 (SRFR1); FUNCTIONS IN: protein complex scaffold; INVOLVED IN: defense response to bacterium; incompatible interaction, defense response to bacterium, negative regulation of transcription, negative regulation of defense response; LOCATED IN: nucleus, perinuclear region of cytoplasm; cytoplasm; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Tetratricopeptide TPR-1 (InterPro:IPR01440), Tetratricopeptide-like helical (InterPro:IPR011990), Tetratricopeptide repeat-containing (InterPro:IPR013026), Tetratricopeptide repeat (InterPro:IPR019734), BEST Arabidopsis thaliana protein match is: Tetratricopeptide repeat (TPR)-like superfamily protein (TAIR:AT3G04240.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI/BLINK).	MASERVEIAKLCSSRNWSKAIRIFDLSLLARSSTVQ DLGNRAFCYSQLELHKHVIKDCDRALEIDPCLLQA YILK
981	sp_170870_apfp	-0.884811	-0.948425	-0.91188	unknown protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: N-terminal protein myristoylation; LOCATED IN: plasma membrane; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF632 (InterPro:IPR006867); BEST Arabidopsis thaliana protein match is: Protein of unknown function (DUF630 and DUF632) (TAIR:AT5G25590.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI/BLINK).	MGCSQKIENEEAVSRCKDRKVFMKSAVTARNA FAAAHCYSAMSLKNAGAALSDFAQGEVYNPLSL SSAPVPTSGVGTMGAIPTSQAPPPPPVQDL PPPPPPIGMPQPLQRAVSMPEFAVKAPGAKLV AEDTIEEEEEEEEEEOEEEEVGRLLRRDN KGGSVGRRMETTPSISNLGTANTNIVGPPSPPP PDERAIRTPPPPLPEVHKQTWDFYFFNTEDIPGG SLRDDDDIRGEYDRTYVYHQRMRSSASASITTA LPPSIEMVEEMWPEKPPPMGGGGVMKKKQV VGGSSGGEGRRKTVNVLVQIFSDLDDCFKA SESAAHYSKMLEATRLHYHSNFADNRGHIDHSA RYMRVITWNRSFRGLPNAEEFKDDFDSEENETH ATVLDKLAWEKKLYDEVKAGEIMKFYQKRVAV LNKQKRGTNPESLEKTKAAVSHLHTRIVDMQOS MDSTVSEISRLRDEQLPKLVAVDAMATMWEA MQVHHEINQLKIASHLRFIDISQSTKETIDHHERTI QLWAIVQEWNTHFEKMWNHQKEYIKALNSWLK NLPIESLLKEKVSPPRPQNPPIQNLLAWHDHL EKLPDEHAKSAIHNFAAVLQTVQQQDEIRKRR VFEESKRDLEKQRHYNKWLSEYMQKKIPNDJGE PGSVEDEFVHKDEIMVEKEFVVENAMKKLEEDS EDYERHCTQVREKSLASLQTRLPFLFRTLSEFAL SASDMYSNLKGLQSFQSHINKHKSVEDA

10

20

30

40

【 1 4 0 2】

【表 1 4 - 4 0 3】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2					Galactose oxidase/kelch repeat superfamily protein; CONTAINS InterPro DOMAINs: Galactose oxidase/kelch, beta-propeller (InterPro:IPR011043), Kelch repeat type 1 (InterPro:IPR006652), Kelch related (InterPro:IPR013089), Kelch-type beta propeller (InterPro:IPR015915); BEST Arabidopsis thaliana protein match is: Galactose oxidase/kelch repeat superfamily protein (TAIR:AT1G14330.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	MLEGISFSSRDLPNACGGERNWWMFDFSLMETS NYKFKSEYGDNDKLGNSIKLLKINKAGDMEIDFLN LSLSLVDFPDSIDNGREMSDTSLSLHQIGRDNIS DCVARCSDYSGSIALVNRNFRSLIEGKELYKLR RQMGIVHEWHYFSCSLEWEVDFPIRLRWKHLPI MPSNDCFRHSDKESLAVGTELLFGKGITSHLYK YSLLTNTWSTGMEMNTPRCLFGSASLGEIAVAG GCDFWGNVFSLAELYNSTDMWTTLPMSMKAR KKCSAVFMDGKFYVIGGIGVDNSNPLTCGEVFDL ERLTWTEIPDMFPLRIVEPGAPGASAMSEAPPLV TVWNNELYTADYARKEIRKYNKSSNSWATVGSLLP ERVVSMHGWGLAFRCGDKLIVGGPRGLEGG YIEVNSWAPSESPQQWTVLGRKPGGFFVYNCNTIM GC
982	sp_171060_mnxd	#N/A	#N/A	-0.6533035	Uncharacterised protein family (UPF0041); INVOLVED IN: biological_process unknown; LOCATED IN: mitochondrion; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Uncharacterised protein family UPF0041 (InterPro:IPR005336); BEST Arabidopsis thaliana protein match is: Uncharacterised protein family (UPF0041) (TAIR:AT4G14695.1); Has 886 Blast hits to 886 proteins in 209 species: Archae - 0; Bacteria - 0; Metazoa - 387; Fungi - 234; Plants - 170; Viruses - 0; Other Eukaryotes - 95 (source: NCBI BLINK).	MAASKFKAFWNHPAGLKTTHFWAPTKWGLSIA NIADFTKPPENISYPQQFVVAASGLIWIWTRYGMVIT PRNWNLCSVNFAMSVTGFCCQLSRKLRHHDVSTEK EPAAAAAAVVKE
983	sp_172480_kejg	#N/A	-0.87348335	#N/A		

10

20

30

40

【表 1 4 - 4 0 4】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたベクトルとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
984 sp_172680_nxoz	-0.336918	#N/A	#N/A	P-loop containing nucleoside triphosphate hydrolases superfamily protein; FUNCTIONS IN: helicase activity, ATP binding, ATP-dependent helicase activity, nucleic acid binding; LOCATED IN: nucleolus; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: RNA helicase, DEAD-box type, Q motif (InterPro:IPR014014), DNA/RNA helicase, DEAD/DEAH box type, N-terminal (InterPro:IPR011545), RNA helicase, ATP-dependent, DEAD-box, conserved site (InterPro:IPR000629), DEAD-like helicase, N-terminal (InterPro:IPR014001), DNA/RNA helicase, C-terminal (InterPro:IPR001650), Helicase, superfamily 1/2, ATP-binding domain (InterPro:IPR014021); BEST Arabidopsis thaliana protein match is: DEA(D/H)-box RNA helicase family protein (TAIR:AT5G65900.1); Has 44310 Blast hits to 43105 proteins in 3082 species: Archae - 753; Bacteria - 22440; Metazoa - 6214; Fungi - 4682; Plants - 2500; Viruses - 12; Other Eukaryotes - 7709 (source: NCBI BLINK).	IQARSPPLLEKGDVILGAARTGSGKTLAFLVPSVE LLHQARFTSRNGVGVVICPTRELAIQTHAVAKDL LKNHSQTLGLVIGGARRRGEAEFLAKGVNLLVAT PGRLLDLHONTKGFVKNLKCLVIDEADRILEANF EEMRQILKILPKERQALFSAQTQTKKVEDLARLS LSKPIYVDDEGRRKVTNEGLOQGYCVFSAQRLL ILLYSFLRNLSKVMVFFSNCNSVFKFSELLRYI QIDCFDIHGKQKQTKRTRTTFFDFCKAEKILLCTD VAARGLDIPSDWIVQYDPPDPKEYIHRVGRRTA RGEAGKGNALLFIPEELQFLGYLKSEKVPVKEY EFDTKLANVQSHLEKLVLTNNYYLNKSAKEAYRA YILSYNSHSMKDVFNVHKLDLOAVATSFQFSSPP KVDIKLDSNASKFRKSRKIEGHRFFNESPNPYGKRG GDSSRQFVRY
985 sp_172700_hx'a	-0.8822861	-0.938114	-0.9463454	Phosphoenolpyruvate carboxylase family protein; FUNCTIONS IN: isocitrate lyase activity, catalytic activity; INVOLVED IN: metabolic process; LOCATED IN: chloroplast, chloroplast stroma; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Pyruvate/Phosphoenolpyruvate kinase, catalytic core (InterPro:IPR015813), Isocitrate lyase/phosphorylimutase, conserved site (InterPro:IPR018523), Isocitrate lyase/phosphorylimutase (InterPro:IPR000918); BEST Arabidopsis thaliana protein match is: Phosphoenolpyruvate carboxylase family protein (TAIR:AT1G77060.1); Has 8771 Blast hits to 8771 proteins in 1489 species: Archae - 119; Bacteria - 4900; Metazoa - 32; Fungi - 52; Plants - 172; Viruses - 0; Other Eukaryotes - 3026 (source: NCBI BLINK).	MGSLEASAPLGRRTTVHRLIAEEGSVLMPIQDAF SAAICAKGFKACFVSGFGVSAALLGLPFDGFLT TTEVVETVRITAAAPDLVWVDGDTGGGGPLN VGRFIKELIAGAKGVLEQVWPKCGHMRGK SVVPAEEHALKIAAREAGSDFFLVARSARAP HGLQEAIRRANLYREAGADATFVEAPANLEELKE VVRDVKGLRIANMIEGGKTPLHTPTFEKEMGFHLI AFHLSITYATTKALVDIMKVLKDKGTTDRDLEKIVT FSEFNDMSLESWYETESKFNFTPKSLES

10

20

30

40

【表 1 4 - 4 0 5】

1	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
986	sp_172750_ooff	-0.364873	-0.348356	#N/A	seuss (SEU); FUNCTIONS IN: protein binding, transcription cofactor activity, DNA binding, protein heterodimerization activity; INVOLVED IN: multicellular organismal development, gynoeccium development, embryo development, ovule development, regulation of flower development; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; BEST Arabidopsis thaliana protein match is: SEUSS-like 3 (TAIR:AT4G25515.1); Has 30201 Blast hits to 17322 proteins in 780 species; Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MVPSPPTPIGGTQSVNPSLLFRSNGMLGGQG GMPFSQGGFSPVSRNQFNINWLLGNVANVSS LLNQSYGNGGANPGISGAGNGPRGGLDAGAES DPLSGMNGMGFNPSYGAANVIVNIPSSSGQS QGQFPNSSNOMLSDQAGSQQHEPQNFHQ QQQFSGNQOQQYFSALGGVGSVKLEPQASG DQASAQQQHQMGSLRGLNQVKLEPHQLQSMR MGTYKMEPQHSEQSMFMHQOQQOQHQLM QQMSMPTNTQAAAAQNLQOQRYIQOQQ QHQQOQHQLMKGIPQFSQMHQQFSPQNLPL RSPVKPYEPGMCARLTHYMYQQQHRPEDNII DFWRKFAEYFAPNAKKKWCYSMYGSGRQTG VFQDQVWHCEICKRKPGRGFEATVE/LPRLFKIK YESGTLLELLYVDMPREYQNSSGGQVLDYAKAIQ ESVFEQLRVRDQQLRIVFSPDLKICSWFCARR HEELIPRLLIPQVSQLGTAQKYQAATQNSSLN MSVSDLQSNCFIVASARQLAKTLEVLVNDLG YTKRYVRCLQISEVNSMKDLIDYSRSTGAPME SLAKFPRTGPPPSMQGQSQOPEEQOQQOQQ PQQQHHQQQPSQPTMGPNNSNDQGSVQAAG MQIASSNGANSVNNNSLNVASTSTQSTIAGL LHQNMSNRQGNPMNNANSPYGGGSSVQMP SPGSSSTMQQQTQPNPFPQSPPTSSNPPQT TSHGAMAATTTNHINSANSPVMSNIHQPAISGDG DTIDTQSSVQKLIQEMMMSSQLNGGGMMGGIGS LGTDAKNVNGILPMSGNSALNGLVNGPMMNSN SSMGGGSLGMMGAMGQASVMVNGIRAAAMGN NSVMNNGRVMASMARDOQSYNHQOQD.LNQIM LSLGLAVNGFNSLQYDWKPS
987	sp_173030_cchg	-0.364853	#N/A	#N/A	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein; FUNCTIONS IN: methyltransferase activity; INVOLVED IN: metabolic process; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Uncharacterised protein family, methyltransferase, Williams-Beuren syndrome (InterPro:IPRO13216); Methyltransferase type 11 (InterPro:IPRO13216); Has 984 Blast hits to 969 proteins in 433 species; Archae - 16; Bacteria - 445; Metazoa - 145; Fungi - 145; Plants - 54; Viruses - 0; Other Eukaryotes - 179 (source: NCBI BLINK).	MAARPELQAPPEIFYDQSEARKYTSRSSRIIDIQ

10

20

30

40

【 1 4 0 5】

【表 1 4 - 4 0 6】

I	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
988	sp_173590_qkol	-0.868167	-0.837514	-0.83307	glycine-rich protein; CONTAINS InterPro DOMAIN/s: Plant specific eukaryotic initiation factor 4B (InterPro:IPR010433); BEST Arabidopsis thaliana protein match is: eukaryotic translation initiation factor 4B1 (TAIR:AT3G26400.1); Has 13357 Blast hits to 7743 proteins in 694 species: Archae - 46; Bacteria - 1331; Metazoa - 4294; Fungi - 1200; Plants - 649; Viruses - 189; Other Eukaryotes - 5648 (source: NCBI BLink).	MAATVSSAWAKPGAWALDSEEHDELKQPIMA QPVKPEPPAADFPFLAAAAAKPKKKNKAQTLSL AEFSTYSAPKPKAFVAGQKLTHTDEILTLPGPPE RTAEELERPLGGFRNYYGGGASDDSRVRYSD RGSSRRDRMGPSRADDEADDWSTKTKKFSAFGGG APERSDRREKVGFNNSRADEVDNWNASNKSFV AAPVDGARRRERMGFESYNSGSGDPSGWSWSP REERGSGVEADTWGRRESVGGNGNGRPRLL NLQPRTLPVENNDNGOSVIPTEVKVASPVLIPT PTEKVKGANPFGEARPREEVLAEKGDWKEIDE KLEAVKKEAGNGEKSEGGFGRHYGLGRSESGD EKTYSWRKPEADAVSSPFGSSADTTENSANT ADSSAESTEKVSEEN
989	sp_173630_lgmz	-0.817337	#N/A	-0.862403	polypyrimidine tract-binding protein 2 (PTB2); FUNCTIONS IN: RNA binding, nucleotide binding, nucleic acid binding; INVOLVED IN: pollen germination, regulation of RNA splicing, regulation of translation; LOCATED IN: nucleus, cytoplasmic mRNA processing body, cytoplasm; EXPRESSED IN: 31 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAIN/s: RNA recognition motif, RNP-1 (InterPro:IPR000504), Nucleotide-binding, alpha-beta plait (InterPro:IPR012677); BEST Arabidopsis thaliana protein match is: polypyrimidine tract-binding protein 1 (TAIR:AT3G01150.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).	MASVSSQPQFRYTOPPSKVLHLRMLPWECTEEE LIELGKPFQKVVNTKCNVGNKNAQAFIEADLNQ AIAMISYASSSEPAQIRGKTYLYQVSNRQEIVNN KTTADTAGNVLVLTIEGDEARQVSEIVLHLVFSAF GFVHKITTFEKTAGFQALIQFSDTETASSAKNALD SRSDYTNQYLPVAPSAIDATGQLSVGDGKKME PESNVLASINMQYAVTLDLVNTVFSAFGPVQKI AMFDKNGGLQALIQYDQVQVAVVAKEALEGHCIY DGGFCKLHLSYRHRVDSLKVVNDRSRDYTVPIH GGMIQSSILGGQPPGGFMQGGNAPPYNGGQ YPPGPPPSGGWAPPQCHMHMQMMPMNYQYNNY QFGHGATPPGPPGTMPPGPMHMQNPPGMPY SQ

10

20

30

40

【表 1 4 - 4 0 7】

	A	B	C	D	E	F
1		分析に用いたペアトとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)		注釈	aa seq
990 sp_174500_wkrw	-0.924113	#N/A	#N/A	#N/A	<p>DnaJ heat shock N-terminal domain-containing protein; FUNCTIONS IN: unfolded protein binding, heat shock protein binding, nucleic acid binding; INVOLVED IN: protein folding; LOCATED IN: cellular component unknown; EXPRESSED IN: male gametophyte; EXPRESSED DURING: M germinated pollen stage; CONTAINS InterPro DOMAIN/s: Molecular chaperone, heat shock protein, Hsp40, DnaJ (InterPro:IPR015609), RNA recognition motif, RNP-1 (InterPro:IPR00504), Heat shock protein DnaJ, N-terminal (InterPro:IPR01623), Heat shock protein DnaJ (InterPro:IPR03095), Heat shock protein DnaJ, conserved site (InterPro:IPR018253); BEST Arabidopsis thaliana protein match is: ARG1-like 2 (TAIR:AT1G59980.1); Has 16747 Blast hits to 16541 proteins in 2695 species: Archae - 126; Bacteria - 6883; Metazoa - 3195; Fungi - 1541; Plants - 1234; Viruses - 13; Other Eukaryotes - 3755 (source: NCBI/Link).</p>	<p>MFLSPFPTERSQIPYSMLGLVPGEEASAKLTEQEI SSAYRKKARESHIPDKNINDPDALAKFKQLQLSYE ILKDPKTRKEFFDLDLLRVRREKVKYRDNKKFKDAK</p>
991 sp_174830_eocrm	-0.954114	-0.935132	-0.950776		<p>YELLOW STRIPE like 3 (YSL3); CONTAINS InterPro DOMAIN/s: Oligopeptide transporter OPT superfamily (InterPro:IPR004813); BEST Arabidopsis thaliana protein match is: YELLOW STRIPE like 2 (TAIR:AT5G24380.1).</p>	<p>MGSDGEVREMEREDYEEQNRVQDEIKRVPWPWQ KQITIRGIASILIGVIYSVIVMKLNLTTGLVPLNINVS AALLAFVLLRSWTSLLHKLGIISTTPTFRQENTVQ TCVACYSISVGGGFGSLLGLNKNITYEQYVD TEGNYPNYSKEPRIGWMTGFLVSSFYGLLALVP LRKIMIDYKLSYPSGTATAVLINGFHFKGDKIAK KQVRGFVKYSLSFVWVGFQWFTYGGDQCGFS QFPTFGLKAWKNSFYDFSMTYIGAGMICHLVN LSLLGAVL.SWGVMMWPLIGDLKWKWFLSSLPE SMKSLNGYKVFISISLIGDGLYNFLKILFFTRSIH MRLKGNLNSCPNDOKRSLDDLQRNEFFMRERI PIWLACSGYTFFSIISIIPKMFPELKWYYVAAVYL APSLFCNAYGAGLTDMMMAVNYGKVALFVLSA LAGKDSGVVAGLVGCGLKSIVSISDLMHDFKTKG HLTLTSPRSMLLSQAIGTAVGCVVAPLTFELFYKA FDVGNPDGEYKAPYALYRNMAILGVQGFSAIPQ HCLQLCYAFFGFVAVANILVRDFSPKKGKWWPLP MAMAVPFLVYGAYFAIDMVCVGLSVFVWHRLNSQ KASLMVPAVASGLICGDGLWILPSSILALAKLKP MCMTFQSS</p>

10

20

30

40

【表 1 4 - 4 0 8】

	A	B	C	D	E	F
1		分析に用いたペアとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_ktqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
992	sp_174890_kiax	-0.356557	-0.962383	-0.357437	NRPB7; FUNCTIONS IN: DNA-directed RNA polymerase activity, RNA binding; INVOLVED IN: transcription; LOCATED IN: DNA-directed RNA polymerase II, core complex; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Nucleic acid-binding, OB-fold (InterPro:IPR012340), Ribosomal protein S1, RNA-binding domain (InterPro:IPR003029), RNA polymerase Rpb7, N-terminal (InterPro:IPR005576); BEST Arabidopsis thaliana protein match is: RNA polymerase Rpb7-like, N-terminal domain (TAIR:AT4G14660.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).	MFFHIIILERNMQLHPRHFGPRLRDKLVSKLMDKV EGTCSGRHGFWAIMGIEHVKGGLRDTGTFVTF PVKYKVVFRPKGEILEAVVTMVKMGFFAEAG PVQVFSVNHLPDDMEFQSDTPNYTTSDNSVKI QKDSEVRLKIGTRVDATEIFEICIGTIKDDFLGWSDP GAAS MESSKQSGNSHHFFEDFDEEGENDEWEAVQA HVKKKRPSTDHVLGARRYDRIMAAFKPICLALTS EDTMVVEWIRLLDIRLDNIIAQRSONGDSLTSQ PSVQNKRFNSVTKPTSGSKSLGNCSSKRKEPSP PAEQKSHTESSNSTPYVKDPISRKKAHSTCEV SLDQPTTNVAIVDQCVTCIPTMDLFGDFKQVGS AYGGEADLIFENAMSVGGIELLHKDGSIGNFTOM VGTNOESECEGGHY
993	sp_175130_zkyx	-0.8833711	#N/A	#N/A	uncharacterized protein	MLFFTQGYHNLVDKHKHVNLSRVVKVNCNSAVH LRKPKYLRYPRTKLPDFGVDTFKPRNEVKD HSEMVDQIDWSSDDQVTRYDFSDDETEAEDGH DVWEADEIAISSLFGRIHQKPGKLNRRERPLP PLPHQIRPLKLPITTKKHVRISSPAALSRLSLTKQY KDPVFLKLAIREIRSLPAEENSVLNLKNGWGRYLR KGSLSLTIKELGHMGLPDRALLTFCWAQKQTHLF PDDRILATVVEILARHDLKLPDFLEKFKVNLSSRG VLEAIARGFKGGSNLALFKLSAARFKRLLDSSI YVKLLLELGNPKDEEVLPLLEELGERDSLNSL QDCTAIMKVCYRGLNDFVESLDFWFOQNRQNP SVVMYTTVYVSRSLDNKYREALALVWEMERLNV LDMPAYRYVWIKLVALNDLPRAVRYFSKIEAGFS PTYDIYRDMISYLLSERLAKCKEVCKEVELAGYK LDANLKSYYLQIFER
994	sp_175420_ayyz	-0.9386593	-0.957722	-0.9178615	EMBRYO DEFECTIVE 975 (EMB975); FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: chloroplast; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Pentatricopeptide repeat (InterPro:IPR002885); BEST Arabidopsis thaliana protein match is: LATERAL ORGAN JUNCTION (TAIR:AT2G39230.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLink).	

10

20

30

40

【表 1 4 - 4 0 9】

A	B	C	D	E	F
1	sp_107660_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2				nuclear transport factor 2B (NTF2B); FUNCTIONS IN: protein transporter activity, Ran GTPase binding; INVOLVED IN: nucleocytoplasmic transport, protein import into nucleus; LOCATED IN: nuclear envelope, intracellular, nucleus, plasma membrane; EXPRESSED IN: 26 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Nuclear transport factor 2 (InterPro:IPR002075), Nuclear transport factor 2, Eukaryote (InterPro:IPR018222); BEST Arabidopsis thaliana protein match is: nuclear transport factor 2A (TAIR:AT1G27310.1); Has 10 18 Blast hits to 10 18 proteins in 227 species: Archae - 0; Bacteria - 0; Metazoa - 491; Fungi - 232; Plants - 204; Viruses - 0; Other Eukaryotes - 91 (source: NCBI/Link).	
995 sp_175560_xezi	-0.948140	-0.948140	-0.919582	RNA-binding (RRM/RBD/RNP motifs) family protein; FUNCTIONS IN: RNA binding, nucleotide binding, nucleic acid binding; INVOLVED IN: biological process unknown; LOCATED IN: cellular component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: RNA recognition motif, RNP-1 (InterPro:IPR000504), Nucleotide-binding, alpha-beta plait (InterPro:IPR012677); BEST Arabidopsis thaliana protein match is: RNA-binding (RRM/RBD/RNP motifs) family protein (TAIR:AT2G33410.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI/Link).	MDFDAVSKAFVEHYSTFDANRFLANLYQDSS MLTFEGKIQGQAITAKLNSLPFQQCQHHISTV DCQPSGTGGMLVYVSGTLQLAGEQHALKFSQM FFLMPTQQSFYQNDIFRLNYA
996 sp_176670_gxpo	-0.967217	-0.9668	-0.9384375	RING/U-box superfamily protein; FUNCTIONS IN: zinc ion binding; EXPRESSED IN: leaf; EXPRESSED DURING: LP_04 'our leaves visible; CONTAINS InterPro DOMAIN/s: Zinc finger, RING-type conserved site (InterPro:IPR017907), Zinc finger, RING-type (InterPro:IPR001841), Zinc finger, C3HC4 RING-type (InterPro:IPR018957); Has 1875 Blast hits to 1873 proteins in 229 species: Archae - 0; Bacteria - 0; Metazoa - 654; Fungi - 156; Plants - 687; Viruses - 149; Other Eukaryotes - 229 (source: NCBI/Link).	MDSQDGKLVFGISWETTEEKLKDYFGQYGDVY QVNIMRDKATGRPRGFGVLFSDPSVLDRLVDY KHTIDRPFVEAKRAMSHEEQSSGRTGNNGG GRSGGGGNFKTKIFVGGGLPPTLTDEGFRQYF ENYGOVTDVIMYDQNTERTPRGFGFISFSEDVAV DRVLONNFDLNGKQVEVKRALPKDANPGMGS RGNCGGYNSSGGQINAYGNRMDSYMSQINA GGGFPAGYSSGGYGAAYGYGFPASGGMGYGG YAGYGGANPGYGGPASAAYGNPNVNAQRSS WNAQAASGYGAAGYGSAGPWAASGAGSGGPT GQSPSGASGYGNQGYGYGSYGGSDGYGSVGN FRAGAPASNAAGGGGEMOQGGYGDANGHGGYG AGGYGAQANGPQMGYGGYGGGARQAQ
997 sp_176860_cqhg	-0.861025	-0.87147	#N/A		MEKHSSMRPPSPNCNNKSKTNSISSIKGK TCPICLSIGHRRPAVITACFHAYCLRCLLRWSNY KFNCPCLNSHFDSY/LSPSSPSSSFHNRLLPLS LRRSPDRPQRRAAAVQORLLRRRRRDEHAVN RQSRPVPRRRSGFLVDSAPSEVVTTERILOWRASI YKHGISAVPFSGRNLOKLNMSVSRESILHRIEPMI QRELAILEDPTIIVHVAATSLYLSISIEKKTTEALS TSSADHKDAFLAPLRFPLFHELTDAFWHEHLCFAE SCFTYDVTYDEYVEYNNKPK

10

20

30

40

【 1 4 0 9】

【表 1 4 - 4 1 0】

I	A	B	C	D	E	F
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈 Calcium-binding EF hand family protein; FUNCTIONS IN: calcium ion binding; LOCATED IN: plasma membrane; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: EF-HAND 2 (InterPro:IPR018249), EPS15 homology (EH) (InterPro:IPR000261), EF-hand-like domain (InterPro:IPR011992), Calcium-binding EF-hand (InterPro:IPR002048); BEST Arabidopsis thaliana protein match is: Calcium-binding EF hand family protein (TAIR:AT1G21630.1); Has 12017 Blast hits to 4944 proteins in 443 species: Archae - 20; Bacteria - 3590; Metazoa - 3650; Fungi - 1736; Plants - 412; Viruses - 16; Other Eukaryotes - 2593 (source: NCBI BLINK).	aa seq
998	sp_176960_help	-0.852032	-0.843565	-0.937133	aconitase 3 (ACO3); FUNCTIONS IN: aconitate hydratase activity, copper ion binding, ATP binding; INVOLVED IN: in 6 processes; LOCATED IN: cytosol, mitochondrion, cell wall, plasma membrane, chloroplast; EXPRESSED IN: 27 plant structures; EXPRESSED DURING: 16 growth stages; CONTAINS InterPro DOMAIN/s: Aconitase/3-isopropylmalate dehydratase large subunit, alpha/beta/alpha (InterPro:IPR01030), Aconitase family, 4Fe-4S cluster binding site (InterPro:IPR018136), Aconitase A/isopropylmalate dehydratase small subunit, swivel (InterPro:IPR000573), Aconitase/3-isopropylmalate dehydratase large subunit, alpha/beta/alpha, subdomain 2 (InterPro:IPR015932), Aconitase/Iron regulatory protein 2/2-methylisocitrate dehydratase (InterPro:IPR015934), Aconitase-like core (InterPro:IPR015937), Aconitase/3-isopropylmalate dehydratase, swivel (InterPro:IPR015928), Aconitase/iron regulatory protein 2 (InterPro:IPR006249), Aconitase/3-isopropylmalate dehydratase large subunit, alpha/beta/alpha, subdomain 1/3 (InterPro:IPR015931); BEST Arabidopsis thaliana protein match is: aconitase 1 (TAIR:AT4G35830.1); Has 20552 Blast hits to 20369 proteins in 2575 species: Archae - 558; Bacteria - 10777; Metazoa - 503; Fungi - 660; Plants - 232; Viruses - 0; Other Eukaryotes - 7822 (source: NCBI BLINK).	
999	sp_177160_edou	-0.887025	-0.886375	-0.923145		

【 1 4 1 0】

10

20

30

40

【表 1 4 - 4 1 1】

A	B	C	D	E	F
1					
2	分析に用いたペアとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
1000  sp_177190_opod	-0.926332	-0.919481	-0.973171	NHL domain-containing protein; LOCATED IN: endomembrane system; EXPRESSED IN: 7 plant structures; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage, 4 anthesis, petal differentiation and expansion stage; CONTAINS InterPro DOMAIN/s: NHL repeat (InterPro:IPR001258). Six-bladed beta-propeller, TolB-like (InterPro:IPR011042); BEST Arabidopsis thaliana protein match is: NHL domain-containing protein (TAIR:AT1G70280.2); Has 2211 Blast hits to 1172 proteins in 193 species: Archae - 21; Bacteria - 928; Metazoa - 143; Fungi - 0; Plants - 261; Viruses - 0; Other Eukaryotes - 858 (source: NCBI BLINK).	MVEKMSKEINTAAGTEPELVSLLHQQQVIDENEA LIMESFSSNNSCSFFCFGGQSPNFRSSGLGG STSSPMSWERLQQTAEIENKWWAGPLOGLKKM REWSEIAGPKWKTFFRRNRGGKKSATGSGKFKO YDPLSYALNFDGPGQNGHLDDEDDGYFPDFSSRF ASIPKSSRSSIDLGKNGIPSLs
1001  sp_177210_cfrw	-0.885386	-0.916781	-0.927344	trehalose phosphate synthase (TPS10); FUNCTIONS IN: transferase activity, transferring glycosyl groups, trehalose-phosphatase activity; INVOLVED IN: trehalose biosynthetic process, metabolic process; LOCATED IN: cellular_component unknown; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: HAD-superfamily hydrolase, subfamily IIB (InterPro:IPR006379). Glycosyl transferase, family 20 (InterPro:IPR001830). Trehalose-phosphatase (InterPro:IPR003337); BEST Arabidopsis thaliana protein match is: trehalose-phosphatase/synthase 9 (TAIR:AT1G23870.1); Has 4634 Blast hits to 4564 proteins in 1018 species: Archae - 52; Bacteria - 2697; Metazoa - 182; Fungi - 655; Plants - 601; Viruses - 0; Other Eukaryotes - 467 (source: NCBI BLINK).	MSSKSYTSFLDLAAGGYLDLPQTPRNLPKVMTV PGVIPDADSTESKDDILDASTRYDRKIVVSNFL PINAQDKPKSGKWSFSLDDDSLYLHLKDGSPES EVYVGS�KVDVVKOEVEVAQILLLENFRCPVTF LPADLQKFKYHGFGKQYLWPLFHYMLPMSPDHG SRFDRGLWGAYVSANKKFKADKVEILNPEDDLV WYHDYHLMLLPTFLRRFRYRVLKGFELHSPFPSS EYRTLVRDEILKAFNLNSDLVGFHTFDYARHFLS CCSRLGLDYESKRGYLGIDYFGRTVFVKILPVGL QLRLEVAMNHFSAVKVKEIQKQFEGKILQV DDMDIFGILSLKLLAAEHLNQHPELQGLVLIQV NPARSTGKDVQEAQKSETYEITKRINMKFGFPGYE FVILIDFPVFLYEKTAYYAMAECCEIVNAVVDGLNL TPYKYVCRGGSPKIDEAMGVSPGETRSSMLVVS EFVCGSPSLSGAIRVNPWNIEAVADAMKMAISIS ELEKQLRHEKHRYVSSHVDVAYVRSRFSVQDLE AKGDHYNKRCWGIQGLSFRILASPSFRKLSVE YVWSAYKATKRAIFLDYDGTVSDMSIVKTPSSE LISVLQNLCSDTKNTFVSGRGRDLSLEWLSPC DKLGLIAAEHGFLRWKQNCDEWISTVTTDDDWK KIAEPVMQLYTESTDGSSTIESKSAIVWHHQDAD VKPOGVSQKGLVTERVLSLNTDNGKAPDFLLCIGD PDFGSCQAKELLDHLENVLSNEPVVVTRGHNIVE DLSDMDIFESILSKASCSTSTGVSPEIFACTVQV KPSKAKYLLNDTKHVMRLLQGLTFASNGKPEPCI PPCQVSFESFA

10

20

30

40

【表 1 4 - 4 1 2】

A	B	C	D	E	F
1					
共発現遺伝子	分析に用いたペイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
1002 sp_178000_xisc	#N/A	-0.81982%	-0.92437%	GPI transamidase subunit; PIG-U; INVOLVED IN: GPI anchor biosynthetic process; LOCATED IN: plasma membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: GPI transamidase subunit; PIG-U (InterPro:IPR009600); BEST Arabidopsis thaliana protein match is: GPI transamidase subunit PIG-U (TAIR:AT1G12730.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	MGEDSTKSKTKVEYTRFWWEWAIAIIFRLVLIY FFPDILSOLASRPEVSTPVTSLRRLAEGYWLKQK ASIPYAGSMYHGSPLLSVLGLPTVNRINDGQPS HILCSLLFVIADLIALLRATGQKLOSAYGCSLKA LGVEKILGSEFYASGDISALFYLNPLTILTCMG SSTSPIENFLVILSYGACKRLIPLAAVGVWVLA SLYPVAVLLPVILLGYPDAPPKKFLQKISRETA KDDSSDLSLQHVKPKITGSALPSPFSWRPVLFLWA SAWYSVLLCGISVSRONGGLQEMFKRTYGFILTV EDFSPNIGVLYWYFFAEVDFFRNFFFLIVHMNLFM ILPLAIRMKHRPCFLAFVYVAICSMILKTYPSVADS ALYLGFLGFLINQLADMHFSEVFLFCGYGVNLLSP VMHNLWVWRGTGNANFYATAIYAALQIILLVYVDS ASAMLNHDRRLRKLTTA
1003 sp_178120_ypai	-0.922775%	-0.85193%	-0.91832%	FUC95A; FUNCTIONS IN: 1,2-alpha-L-fucosidase activity; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Six-hairpin glycosidase-like (InterPro:IPR008928), Uncharacterised conserved protein UCP007663 (InterPro:IPR016518); Has 1349 Blast hits to 1316 proteins in 300 species: Archae - 2; Bacteria - 1123; Metazoa - 0; Fungi - 116; Plants - 52; Viruses - 0; Other Eukaryotes - 56 (source: NCBI BLINK).	MDPQLQLPSLVLPCLFSLFSLVILGRFGAMEEHN WVMGRKPEWEKDIWPKFSTSLVMGVEESRPLKIIF KGPAKNWTDALPIGNRFRGAMVHGGLSSEINLN DDTLWTGVPDYDTPNAPQAIANVRKLVDSGNY SEATAFADKELSGDSSNAYQLLGDILLEFDDHDS AYIEQTYHRELDLDTATVTKYSVGDVVEFTRQLF ASSPDQVYVAKISSKQGSLSFTASLDSKLDHHSY TNEKNQIIMEGSCSGYPKGIQFSLVNLQIDGEG KIHNLQRDLRVEHADSAVILLVASSDFDGPFTKP SDSKKDPDPTSEALSRSININILYSTRHLDYQ KLFHRVSLQLSKNSNIDEKISTAQRVQSFQDDED PSLVELLFYGRYLLISSSRPQTQPANLQGIWSK DVPYPPWSCAPHNLINLQMINYWLSSCNLQELQK PLFDYLSFLAINGKKTANVNYKASGWVAHQVTDI WAKTSPDSGQAVWALWPMGGAWLSTHLWEHY TYTMDKNFLNKAYPILEGAASFLDDWIEDKDG YLESNPSTSEHMFMTSLNGKFPASVYSSTMDVAI TKEVFIISAAEELGRTRDLDVFNVLKAKALPP YRIAKDGLMEWAQDFDDPDPHHRVSHLFGIF PGHSITPEEPPDLCKAANYSLYKRGKEGPGWST VWKAALWAHLHSSEHAQYMKVHFLDLVDPERE GSYEGGVYVSNLFTAHPFOIDANLGFSAIAEML VGSTKDYLLPALPRDKWSNGCVKGLRARGGY TVNICWTDGDLHEVGLWNSGNSVNRKRLNRYRGT VMADILPGTYVTFDSQLKCVNTQSL

10

20

30

40

【 1 4 1 2 】



【表 1 4 - 4 1 4】

I	A	B	C	D	E	F
1		分析に用いたバイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
1007	sp_180150_fjqs	#N/A	-0.335736	#N/A	<p>P-loop containing nucleoside triphosphate hydrolases superfamily protein; FUNCTIONS IN: DNA helicase activity, nucleoside-triphosphatase activity, nucleotide binding, ATP binding; LOCATED IN: nucleolus, chloroplast; EXPRESSED IN: 21 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: TIP49, C-terminal (InterPro:IPR010339), ATPase, AAA+ type, core (InterPro:IPR003593); BEST Arabidopsis thaliana protein match is: P-loop containing nucleoside triphosphate hydrolases superfamily protein (TAIR:AT3G49830.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).</p>	<p>MAELKLSKSKDLTRIERIGAHSHIRGLGLDSTLEP RLTSEGLVGLPARKAAGVILQMIKDKGIARGV LIAGQPGTGKTAIAMGLAKSLGHETPFAMLSGSEI FSLEMSKTEALMOAFRAIGVRIKEEAEVIEGEV EYOMDRPAVAGAAKTKGLTKTTDMETVYDLG TKMIEALSKEKVKQSGDVIADKASGKITKLGSRFS RSREYDAVGPHTKFCQCPDQELOKRKEVHVCVT LHEIDVINSRTQGFALFTGDTGEIRAEVREQIDT KVAEWREEGKAEIVPGVLFIDEVHMLDIECFSLN RALENEMAPILVAVATNRGHTFRIGTNYKSPHGPI DLLDRLLIATQPYSEEEIRKLEIRCEEEDEVMSE EAKALLTKIAVETSLRYSINLINTSALACLKRKAKS VEMEDVSRVYQLFLDKRSTQYLMEYQNCQYMFN DVPFGVEVNEDENTDMTS</p>
1008	sp_180170_rpek	#N/A	-0.362141	-0.333135	<p>Nucleotide-sugar transporter family protein; FUNCTIONS IN: nucleotide-sugar transmembrane transporter activity, CMP-sialic acid transmembrane transporter activity; INVOLVED IN: carbohydrate transport, nucleotide-sugar transport; LOCATED IN: endomembrane system, integral to membrane, Golgi membrane; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Nucleotide-sugar transporter (InterPro:IPR007271), UDP/CMP-sugar transporter (InterPro:IPR021189), UDP-galactose transporter (InterPro:IPR004689); BEST Arabidopsis thaliana protein match is: Nucleotide-sugar transporter family protein (TAIR:AT4G35335.1); Has 1068 Blast hits to 1053 proteins in 185 species: Archae - 0; Bacteria - 3; Metazoa - 567; Fungi - 123; Plants - 182; Viruses - 0; Other Eukaryotes - 193 (source: NCBI BLINK).</p>	<p>MQWYFVAALLTVLTSSQGLTTLSOSNGGYKYDY ATVPFLAEVFKLVVSSILLWREVGTPSPVMTTD WKTVRFPIPSVIYLHNNVGFATLTYDTSTYQIM GNLKVITGLFRFLFKKLSNLQWLAILLAVGTTT SQVKCGESSCDLSAPIQGYMFGILSACLAL AGVYTFELMKKNDLSLWQNVQLYTFGTIFNLAR LVLDFTSGFEKGPWWQRLFDGYSMATWLVVL NLGSSGLLVSWIMKYADNIKVKYVSTSMAMLLTMI LSVYLFDFKPTLQLFLGVICMMSLHMVYFSPANAL VESPANMAKPEHESLKEVNIERLNS</p>

10

20

30

40

【表 1 4 - 4 1 5】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2				cold shock domain protein 1 (CSDP1); CONTAINS InterPro DOMAINs: Nucleic acid-binding, OB-fold (InterPro:IPR012340), Cold-shock conserved site (InterPro:IPR019844), Zinc finger, CCHC retroviral-type (InterPro:IPR013084), Cold shock protein (InterPro:IPR011129), Nucleic acid-binding, OB-fold-like (InterPro:IPR018027), Zinc finger, CCHC-type (InterPro:IPR01878), Cold-shock protein, DNA-binding (InterPro:IPR002059); BEST Arabidopsis thaliana protein match is: cold shock domain protein 3 (TAIR:AT2G17870.1); Has 93964 Blast hits to 39618 proteins in 2725 species: Archae - 55; Bacteria - 19157; Metazoa - 4501; Fungi - 1919; Plants - 2885; Viruses - 60267; Other Eukaryotes - 5180 (source: NCBI BLink).	MAQESRCRGVQVRFNDTRFGFIKPEDDGEDLFVHSHDIKSDGYRTLTAGATVEFTFSVGSNDKPKAVDVTGPDGAPILPPAAARRSSNGGGGRQRNNYSDGGKCYNCGESGHIARDCEGGSDNLGGGSGGGNCYNCGKSGHFARACQSGSAGGGSNVGRGGRYADGDGACFCNCGDIGHMARDCCGQDKVNSGGRYGRRAGGGGACYNCGDIGHLARDCAQERASSGGRFFGRAGGSSGACYTCGEPGHLARDCSQDRPSSGRRSSGGVSGGRNCYQCCEPFGHIARDCSNAVA
1009	sp_180270_gdyp	#N/A	#N/A	atypical CYS HIS rich thioredoxin 4 (ACHT4); FUNCTIONS IN: oxidoreductase activity, acting on sulfur group of donors, disulfide as acceptor; INVOLVED IN: cell redox homeostasis; LOCATED IN: chloroplast membrane, chloroplast; chloroplast stroma; EXPRESSED DURING: 13 growth stages; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Thioredoxin fold (InterPro:IPR012335), Thioredoxin, core (InterPro:IPR015467), Thioredoxin-like (InterPro:IPR017936), Thioredoxin domain (InterPro:IPR013786), Thioredoxin-like fold (InterPro:IPR012336); BEST Arabidopsis thaliana protein match is: atypical CYS HIS rich thioredoxin 3 (TAIR:AT2G33270.1); Has 2537 Blast hits to 2529 proteins in 580 species: Archae - 19; Bacteria - 606; Metazoa - 627; Fungi - 356; Plants - 522; Viruses - 3; Other Eukaryotes - 404 (source: NCBI BLink).	MADVLRKSNVSPKTSHRNHNHKKRFFFIKCSFPPANNGGRLVGISSRSSFSGVGTAVDQNRNIWSQKSRNFQSIGAKMGLSIGKAHKAWWKEGLQNMREITSAEDLVDLSLLNAGDNLVVDVDFPFGGCGCRALHPKLCQLAEMNPDVQFLQVNYEEHKSVMCYSLNVHVLPPFRFYRGAEGRCVCSFCTNATIKFKDALAKHGTRSSLAPTKGLDEKELVALAANKELSFYTSKPVEAEPVPEHQEVLVNVADDKAAAKESLSSNPKTLGSSSTLRSQAENEDKTLATARR
1010	sp_181010_dffa	#N/A	#N/A		

10

20

30

40

【 1 4 1 5】

【表 1 4 - 4 1 6】

	A	B	C	D	E	F
1						
2	共通遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh sp_107660_kiqg sp_107670_ptqx (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
1011	sp_182620_smwtd	-0.877683	-0.855523	-0.929336	sequence-specific DNA binding transcription factors; CONTAINS InterPro DOMAIN/s: SANT, DNA-binding (InterPro:IPR001005); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT1G76870.1); Has 317 Blast hits to 297 proteins in 56 species: Archae - 2; Bacteria - 28; Metazoa - 67; Fungi - 2; Plants - 135; Viruses - 9; Other Eukaryotes - 74 (source: NCBI BLINK).	MDNLSLANMPPGGGSAAYVGSMDMQGSIRVNIHH QNPRLQGLQGLMTQPSSMHEVFLKMGQGHK RDOPNSMADYKNGERKTSPTDEDEHSSNGEG MDGSDPSKGGKVPQWHRVYKWNSSMWKLLITA VSYLSEDAVFEGCSQGRKYLNLHKGGKWKXIS KVMASERFFVSPQOCEDKFNDLNKRYKKLNEIL GKGTSQCQVVEKPALDMMDBHLSDKTKDEVKIL SSKHLFYEELCSYHNRLHLPHDPALQQSLQLA LRSRDDTDLDFRKHHLGAEIEDDDGSESDHED YGENCSKAYPLSGDPSKRMKQGGQLENSDFR POSTDFSKSFQQLQNFADMHPEFFPENQKAA LLQKORLKAQSFQLEEQKQINSEALELEKQRFK WQRFCKRKDRELEKLEKIENDRMKLDNERLALLEK RMEMSFTSNH
1012	sp_183351z8694	#N/A	-0.852085	#N/A	uncharacterized protein KOW domain-containing protein; FUNCTIONS IN: structural constituent of ribosome; INVOLVED IN: translation; LOCATED IN: ribosome, intracellular, large ribosomal subunit; CONTAINS InterPro DOMAIN/s: Translation protein SH3-like (InterPro:IPR008991), Ribosomal protein L24/L26, conserved site (InterPro:IPR005825), KOW (InterPro:IPR005824), Ribosomal protein L24 (InterPro:IPR003236); BEST Arabidopsis thaliana protein match is: Translation protein SH3-like family protein (TAIR:AT5G54600.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	#N/A
1013	sp_183530_yzwm	-0.851386	-0.854419	-0.869342		MGWKAAQKLRNWKILRIGDVMIRGKDKGETG TVKRVVRSNRVVEGKLVVFFLLSL

10

20

30

40

【表 1 4 - 4 1 7】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MIGKVKVAVTAFKKEKEREKEREKEREKERVV STSSPNVSLQSSVTADDGPPVLQKRDACV HLEKGDLEKFLKIESAESTGCECDRDGVNDRK ACKGKGGKKGASDSTRANKATWICLCCGHF SCGAVGLPTSP.LSHALHRSRQRHPLVQLNPL LRWCFSCNTLLPYDKLVDFGEHKDFLDIVKLLK QSSKVRSDVEDVWFGSGSVVSEVKEEKAINL VNGKDFSVRGLVNLGNTCFNFSVMQNLLAMDV LRECLLNLEGSYGPLTSAKKFFVESPESGLRN VINPKSLFGVCYKAKASQFRGYQQDSDHELLRCLL DGLSSEDLIAIKSAGSNGKSKPYSSTFYDAIFGG QTSSSVCCVECGHSSVYEPFLDLSLVSPTKNPL FKKARPVHRAKTKLPKRGRGRIRPKLNKINEDM GPAPSNCTSSSTSISSSSGLKGLKGDMDGTAYSG CNTSTSCESSSLLFAEAPSCAPSAFIPINLVSR NAMMTQDNNSVEPDDASTEMLQSDAQDADGMS WLTVEPISQSDGSGGADCNWMDYLGPILSLD DPNSAGKVGFSVTQDFGSGQDGNHDIISIDGTSKP IVEANLQQDSTMEGDNLDGVTVPDTSKSGGEVS EQPQYVNLKAESTADDWGEERAVQVQENEILL TYKENGPDVGEEMKYDAEISSAIVDGEDSLGFD GLGDLFNEPEAAEAGMNTFVSDGLGTGFLAGN STESDPDEVDNTSPVSVESCLAHIKPELLSGE HAWHCENCSKLVKEGRKRMKQKQKTALIQRN KSSNVCSSTSEILHMSHSFPDSLAGLNGSSDDE NLKKHDTTRTNEQLQDLPQEEIEIRCVPEAAHLSA SGPAVYSPISDGGKSSDTHDNGPVEKPSGLSEEH ESGEEEEVVDLKCVMKMPDATKRILSKIPILTI HLKRFSDQARGLSKLNHGVGFKEYIDLGPYMD CRSTERDNCIYHLIGWEHSGSMRGGHYVAYVR GGDRRKRGRSEETIRGRSVYWHYHSDAYVRETTL EEVVLGCEAYILFYEKAKLKL
1014	sp_184260_ofeo	#N/A	-0.932533	ubiquitin-specific protease 2 (UBP2); FUNCTIONS IN: ubiquitin-specific protease activity; INVOLVED IN: ubiquitin-dependent protein catabolic process; LOCATED IN: proteasome complex, chloroplast; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: Zinc finger, UB2-type (InterPro:IPR001607), Peptidase C19, ubiquitin carboxyl-terminal hydrolase 2, conserved site (InterPro:IPR018200), Peptidase C19, ubiquitin carboxyl-terminal hydrolase 2 (InterPro:IPR001394); BEST Arabidopsis thaliana protein match is: ubiquitin- specific protease 1 (TAIR:AT2G32780.1); Has 10400 Blast hits to 6714 proteins in 274 species: Archae - 0; Bacteria - 50; Metazoa - 5279; Fungi - 1927; Plants - 1459; Viruses - 5; Other Eukaryotes - 1680 (source: NCBI BLink).	

10

20

30

40

【表 1 4 - 4 1 8】

1	A 共発現遺伝子	B sp_107620_kqnh (SOAP1)	C sp_107660_kiqg (SOAP2)	D sp_107670_ptqx (SOAP2-like)	E 注釈	F aa seq
1015	sp_185550_myud	-0.323326	-0.333647	-0.348193	<p>DDB1-CUL4 associated factor 1 (DCAF1);          FUNCTIONS IN: nucleotide binding; INVOLVED IN: in          6 processes; LOCATED IN: nucleus, CUL4 RING          ubiquitin ligase complex; EXPRESSED IN: 29 plant          structures; EXPRESSED DURING: 14 growth stages;          CONTAINS InterPro DOMAIN's: WD40 repeat-like-          containing domain (InterPro:IPR011046), Lish          dimerisation motif, subgroup (InterPro:IPR013720),          WD40-repeat-containing domain          (InterPro:IPR017986), WD40 repeat          (InterPro:IPR001680), WD40/YVTN repeat-like-          containing domain (InterPro:IPR015943), Lish          dimerisation motif (InterPro:IPR006594); Has 5754          Blast hits to 3482 proteins in 410 species: Archaea -          46; Bacteria - 1277; Metazoa - 1225; Fungi - 769;          Plants - 278; Viruses - 87; Other Eukaryotes - 2072          (source: NCBI/BLINK).</p>	<p>MMIEATTSVEEFPSSSSQFPSPFPFPPEUTED          SAAGEEVEDDEDEDDVSVSKAQKFMIEITSS          PDNPNLKALHALASLETTQESRYMIEADHSSSN          QTRASHSIGRLGNIVRENDEFELISSKYLSESRY          STSIQAASARVLLSCSVTWMPHYVFEDAVTENIK          NWVLDDSDSCDGGGLWKQGFGEKCKCSDFD          MRKTYSTGLLALGLTSGSQVVEDVLTSGLSAKLM          RCLRNRYLGETSISOKDTHVLDKSKASLGGGI          GRDESRRARLPSLEAASLDDSRIDEDGGSADLV          GERDHRFSNWQVHGEEQWINAEPDGLAEEV          DIHADIDGEGCWGRDVLDRRLKIGERNPGPR          SLHDEDYNESTRDDSARRRVRNWRMRSGKTRV          HESTIEPEQPI MSPSGSRGQLRSSRERSTTR          NLDGKRGTDNKKSSGRTTAEKAVDREDIDEIY          ECRIGTKDISDLIKTAVRSAAEAARAVDAPIEAIKV          AGEFAAELKSAAMDEFAATNSEDAFLAASRA          ASTVIDAGHAANTDEYSRNSONASVDLPKDGAA          EPDQNDIEFFIPDSDTLAQOMEKYICQCLVLLGE          YVEVLGPV/LHEKGVDSVLA/LQRYSEKKESEIV          SSVLPEVLKJLALAAHRKFAAVFVDRGGIQLLAV          PRVLHTFFG-LSSCLFTIGSLQGMERYCALPSDVI          RRLVELGLQLLECPDQARKNAALFFAAAFVFR          ILDVFDAQDGLQKSLGLLRDAAASVRSVGNSTLQ          ASAPGALNRDSNADVLTSSSEKQIAYHTCVLRQ          YFRAQLLLLVDTVRPNKSNRSIARHIPSIRAGYKP          LDISNEAMDVAVLLIOLKDRKLPALVBARWPAVE          RFLACNGHVTMLELCOAPPVERYLHDLLOYALGI          LHVTPYGRKLVLTATLSNDRVGI/VILDAASGP          GYIDPDIQALNVLINLVCPPPSISNKSPLLAQNQ          QSISVQTSNWSVETETDRNTERSVSDRSHPVPGQ          NEARERNGEFPSTAERGLSTNIMQSPSPNIAASGL          VGDRRISLGAGAGCAGLAQAQLEQGYRQAREAVR</p>

10

20

30

40

【表 1 4 - 4 1 9】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	sp_107660_kicg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2				relative of early flowering 6 (REF6); FUNCTIONS IN: sequence-specific DNA binding, transcription factor activity, zinc ion binding, nucleic acid binding; INVOLVED IN: in 6 processes; LOCATED IN: nucleus; EXPRESSED IN: 8 plant structures; CONTAINS InterPro DOMAIN/s: Transcription factor jumonji/asparyl beta-hydroxylase (InterPro:IPR003347), Zinc finger, C2H2-like (InterPro:IPR015880), Transcription factor jumonji (InterPro:IPR013129), Zinc finger, C2H2-type (InterPro:IPR007087), Transcription factor jumonji, JmjN (InterPro:IPR003349), Zinc finger, C2H2-type/integrase, DNA-binding (InterPro:IPR013087); BEST Arabidopsis thaliana protein match is: Zinc finger (C2H2 type) family protein / transcription factor jumonji (jmi) family protein (TAIR:AT5G04240.1); Has 105567 Blast hits to 35250 proteins in 1010 species: Archae - 0; Bacteria - 23; Metazoa - 99094; Fungi - 2247; Plants - 673; Viruses - 7; Other Eukaryotes - 3523 (source: NCBI BLINK).	MGSSSVITERSPEVFVWVKLCPVAFETPTIQ EFQDPIYIFKIEKEASKFGICKIVPPVPPQSKSVI SNLNQSLNSKNPNPQPTFTTRQQQVGFQPR KORPIQSVWQSGECYTLQDFEAKAKAFENYL KNRAKAVSPLVETLYWKANGDRPFSVEYAND MPGSGFMPMKERKRTGEIVAGAAANNVGETAWN MRGVARENGSLLRFNKEEIPGVTSMPVYVAMIF SWFVHVEDHDLHSLNYLHLGAGKTWYGVPKD AASAFEEVIRVYGFEEINPLVTFQAQLGKTTVMS PEVLINAGVPCRLVQNAgefVTFPRAYHSGFS HGFNCGEASNIATPEWLFRAKDAAPRASINYP MVSHFQLLDLALICSRASIGNSAEPRSSRLKDK KKGEGEMLVKQMFVQDVQINNELLYLQGGSEV VLLPHNSSEIFVWNLRVGSKYKVKPGLPFSLSYS SEEAIKASDDIMLARDDRKQKAFSSVYKTKSGGCM STTPHEIQHSETKEKGGGAGDGFSDRGLFSCV CGIWTACVAIVQPTESAQYLMSADONSFNDWI ACSGVSSHGMDPIDGEANISDPNSFCGSMKHP PDGAYDIPVHSTDYHAQSIGNTSKLSNTAEIVEV GSHTEKRESSALLAMTYGNSDSDDEDVQV NSPVEDNLSSGDGSGWGFHQRDQDASPVFEQ GYDSGAERGSQISSRSECEDEDESSQSRDFEYEH CGHRRVNGDDNEYDHNCSAKFTFEEDILTSEQN YSPADEHDTAKISCAIDPVGKPNLSFAHRCDEDS SRMHVFCLEHAVEVEKQLRPIGGVHILLCHPDY PNVVEAKAAEELEMDYGVKWDIAFSMATKED ERHMLQSEESTPKNGDWAVKLGILNLFYSAILS RSFLYNKQMPYNSIYNFAGCTSPSKSSPEEAKV RGKGFGRQKLLVMAGKWCQKVMWSNQVHPLL LHRDPDEEERNFNACMKSDKVGKRSKSETSHKQ TTYTNRKVGKRRSMPFESTKVKMKFATEFAD TDFEDSVDDIKPEVETRTRSVRKSKTQCRSIRKVK
1016 sp_185680_1szf	-0.951354	-0.814884	-0.670111	vacuolar protein sorting-associated protein 20.2 (VPS20.2); INVOLVED IN: vesicle-mediated transport, N-terminal protein myristoylation; LOCATED IN: ESCRT III complex, plasma membrane; CONTAINS InterPro DOMAIN/s: Snf7 (InterPro:IPR005024); BEST Arabidopsis thaliana protein match is: SNF7 family protein (TAIR:AT5G63880.1); Has 1828 Blast hits to 1807 proteins in 295 species: Archae - 35; Bacteria - 99; Metazoa - 731; Fungi - 367; Plants - 251; Viruses - 5; Other Eukaryotes - 340 (source: NCBI BLINK).	MGNFVKKPKITEVDRAILSLKTRRKLGGYQQQ LEAVIEAEKQAARDLIREKRDORALLAKKKLADI ELSSKQKAVFESLKAGNNTVMKIQSELNLDVQK LMDTTEEARAYQE
1017 sp_186130_dyxo	-0.684379	#N/A	#N/A	uncharacterized protein	#N/A
1018 sp_186271z9033	#N/A	-0.91033	-0.826227		#N/A

10

20

30

40

【 1 4 1 9】

【表 1 4 - 4 2 0】

A	B	C	D	E	F
1					
共発現遺伝子	分析に用いたペアとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
1019 sp_186300_kxyq	#N/A	-0.333226	-0.333104	ubiquitin-specific protease 8 (UBP8); FUNCTIONS IN: ubiquitin-specific protease activity, ubiquitin thiolesterase activity; INVOLVED IN: ubiquitin-dependent protein catabolic process; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINs: Peptidase C19, ubiquitin carboxyl-terminal hydrolase 2, conserved site (InterPro:IPR018200), Peptidase C19, ubiquitin carboxyl-terminal hydrolase 2 (InterPro:IPR001394); BEST Arabidopsis thaliana protein match is: ubiquitin-specific protease 5 (TAIR:AT2G40930.1); Has 10748 Blast hits to 7392 proteins in 262 species: Archae - 0; Bacteria - 6; Metazoa - 5718; Fungi - 1753; Plants - 1297; Viruses - 10; Other Eukaryotes - 1964 (source: NCBI BLINK).	MEPPATESSAAPQSTANNDEQRVYLVFPRWW KEAQDVSTSDSSGERGVLYSASSSSSYAGPL KLINIFGSDLLFNLRREEEISNKNNGEVGVRDY ALVSGEMWLQAFKWHSDSKSGMKRVGFSFAGD VDIADVYPLQLRLSVLREANSILGVKISKDNVAEL FRRACKIFINTEAELLHIWDFSGRTSHFFVNDSS PPKDIQRQSDQELLELIQIYGLSDSPKSRDGKRD EFSEMANAAGSSSSASTTMNGSTGNSYRFRPTN SLTFSSPAESGSLGLTQLNGLNCTCFMSALQ CLAHTPKLVDFHGEYKNEINHNDP_LGMDGEIAL AFGDILIRKLWAPGATPAPRIFKSKLRFAPQFS GFNQHDSQELLAFLLDGLHEDLNVRKLLKPYAEAK DGDGRPDEDVADVEYWLNLHAFRNDISIIVDIOCGQ YRSTLVCPVKVKVITDFPFMYLSLPMPTTMR MTLTMNTDSSMPTFTVYSPKYGKGEDLVKA LTTACSLQIDETLLVAEYHRIIRIFLESDSDLSLJ RDEDRLVAYRFPKATQKATVFMHGEVEEHVH GKLTSGWKAFGLVARICNNANGSDICNLTKLL NPFLMPNEDDLQDDASKSTATENLVEITNLIAI ETSDDAEPTNEEGVNLAADSEFEFFLTDEKGLR NSKEMADPLTSEDLPKRLNIWVSWPEKKEIFDT VLLGLTLEIFKSGFLAKRPQEVSLYRCLAEFLKEE PLGPEDMWYCPGCKQHRQATKLDLWRLPEILV IHLKRFYSRYTKNKLVAVDFVDDLLNSMYIAN KNTPQSFYRYLVYAINSHYGGGLGGGHYTFVHHG GGRWYDFDDSNVPIPIEERIKTSAAYVLYFRRVA DS
1020 sp_187080_ueem	-0.857108	#N/A	#N/A	Protein of unknown function, DUF593; FUNCTIONS IN: molecular function unknown; INVOLVED IN: biological process unknown; LOCATED IN: cellular component unknown; EXPRESSED IN: 18 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAINs: Protein of unknown function DUF593 (InterPro:IPR007656); BEST Arabidopsis thaliana protein match is: Protein of unknown function, DUF593 (TAIR:AT4G30830.1); Has 35333 Blast hits to 34131 proteins in 2444 species: Archae - 798; Bacteria - 22429; Metazoa - 974; Fungi - 991; Plants - 531; Viruses - 0; Other Eukaryotes - 9610 (source: NCBI BLINK).	MSLMATRLSLATHGSNFLKETLQAOQELLOKLY NELDMERESSATAADEALSMILRLQGEKAAYGM EASQYKRIAEKISHVQETLEVFGEIYQKEMQIAS LEFOVQAYRFKLLSMGCEDLGAGELHFPDNMLS RKNVSEESSTRSLRRLSSLPNPLNGVLSRRTS NEGTTTRSTSSPIPWQESKSDP5VRKEHGFDFH SYWDEIRKLDLDRVKELDSLKVSNSPSTEDTNDV PLTKKSAIMPHPSPENFPAKTFVYESLEKVASVQ AGESSETTAQDYVEPICSYSDSSNIKKESQNV TLGSETRNGKLDLWVGEIWKSDVKEPELLSPESD NSSKTNDSKDKIDWEKKPLLPSPGSSNESPKLKD RKTDKPLLFKHPTLGGIADFELOQFNRRRLKQLED QRFVTRGHHEMINSEKAEIEELRLVREIRDKIDLWQ ADITRWGIKKPLPADSRSPGCLKEVILCFWPS

10

20

30

40

【 1 4 2 0】

【表 1 4 - 4 2 1】

	A	B	C	D	E	F
1		分析に用いたペイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-lite)	注釈	aa seq
1021	sp_187500_fols	-0.853313	-0.887372	#N/A	<p>RNA-binding (RRM/RBD/RNP motifs) family protein;            FUNCTIONS IN: RNA binding, nucleotide binding,            nucleic acid binding, EXPRESSED IN: 22 plant            structures; EXPRESSED DURING: 12 growth stages;            CONTAINS InterPro DOMAIN/s: RNA recognition            motif, RNP-1 (InterPro:IPRO00504), Nucleotide-            binding, alpha-beta plat (InterPro:IPRO12677); BEST            Arabidopsis thaliana protein match is: RNA-binding            (RRM/RBD/RNP motifs) family protein            (TAIR:AT4G00830.2); Has 31429 Blast hits to 23398            proteins in 2989 species: Archae - 812; Bacteria -            21040; Metazoa - 874; Fungi - 1027; Plants - 329;            Viruses - 0; Other Eukaryotes - 7347 (source: NCBI            BLink).</p>	<p>MPRTKKSASTPAKSGRLSRTAAKTEAKSDDDGI            NEVVRDEINDNVKVELDNDVDSSADLNVEPQS            QANGSFILTKEDVEKDEYKDERLLEDNEPEY            EPEEYGGYEKEDVDQDDFOEDGDEGMEVEE            SDLAEIEEIEEEDMAEEDIEEVEELGVEEVE            HADADAREHREHGHHDHGHDDHAELEFDDAE            EREHHELKERRKRKEFEVFGGLDKDATEEHL            KEVFSVGEVTEVRLMNPQTKNKGFAFLRFS            TIEQAKRAVSELKPFVNGKQGVTPSQSDSTLFL            LGNICKTWKEKLEKLLYGIENIEDTLVEDSN            NEGKNRGFALEFFSSRSEAMDFAKRLQKRDVLL            GVERPAKVSFADSFIDPGEIMAQVKTVIDGLPA            SWDEDRVILLKNFGEIEKVELARMPASARRKDF            GFVFSHSDAAVTCAKSINNAELGEGDHAKVPR            ARLSRPLQRGKGNPRVDHNGGRRVAGRRPRA            PWNRPRLPRLPPRMRAIGSRPPIVGRGLKRPV            GVRDRRPPVPARARPLPPPPRSPYERRAPGPP            PKSSLKRDYIRREELPPRSRAVAEYGSRLPPER            RPSYDEYAAARSPYVDVHRSSASRSATRRPY            EDNYPPRYERPPSYREGRRDYDYSIGSKRKY            SAMDDVPSRYTDPGHGMRHSRARVDYDVPSSG            SQYADTYGDRMGRSTAGYSGSGSRSSYSQAQDS            HGLYNSRQGMYSYNGSYGNNDGVYSSYSSEY            ASRGSVDVGGSSYSSLYSGRGMSSGGYMGGS            SGSY</p>
1022	sp_188550_huor	-0.831494	#N/A	#N/A	<p>TIMING OF CAB EXPRESSION 1 (TOC1);            CONTAINS InterPro DOMAIN/s: CheY-like            (InterPro:IPRO11006), Signal transduction response            regulator, receiver domain (InterPro:IPRO01789), CCT            domain (InterPro:IPRO10402); BEST Arabidopsis            thaliana protein match is: pseudo-response regulator            3 (TAIR:AT5G60100.3); Has 1807 Blast hits to 1807            proteins in 277 species: Archae - 0; Bacteria - 0;            Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0;            Other Eukaryotes - 339 (source: NCBI BLink).</p>	<p>MEIREVSPNGKSTQDFDRSKVRILLCDNDSKS            SQEVFSLKCSYQVTSVKSARQVVDALNAERS            EIDIILSEVDLPTKKKFKLKYIMRDEQLRIPVIM            MSTODEVSLVKCLRGAADFLVKPLRTNELLNL            WTHMWRRLMLGLAEKNIVSYDFDLAASDHSE            NSTTLFSDDDTDKSRRSVNPESISFTMQEEECN            KVTSDAVPANSFDCRSTLPALGDRRTGVPFVSK            KGEFPLKKEPFPKREIKIGECSAFFTYVKTGA            VESNCQTVVDVEGNAAKPSRPDDKLDTWNGHVI            NGNMTYENGDWENNHSHEGDFPSSTSVPDSTS            MERSCPTVTVFEMQNKSNDEGLSEVHMQPKN            GKFHFDVSTLAAHAAAYPFVMSGMVQNMFMPS            SAHANQKDLQNHGSSTMPHYNHVSHSSSQVP            GMAPFTYYP LGVC.LPQGQTPANHOVPSFSGNSFS            SEAKLSKWNRRREALKFRKQKRCRFDKIRIYV            NRKKLAEFRVRVGGQFVRKRVNGVTVDLNGDPSA            DLYEEDEPNSSPEDDTF</p>

10

20

30

40

【表 1 4 - 4 2 2】

A	B	C	D	E	F	
1						
2	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	
1023	sp_189180_odscl -0.882749	#N/A	#N/A	uncharacterized protein	MVVDLMTPTRQNEQEVNDTDVAGMGOKTPPL MPPTLHTSSGYRPSQSAGTSSGLNMGVFRP GEWIEKTPICLSPKMKMYFEVPPGEPDEPWNPKI DLLRGEISILTDDCKGGGCMGWRAKDLATPR	
1024	sp_189400_twmw -0.885039	-0.879935	-0.837472	RNA-binding (RRM/RBD/RNP motifs) family protein; FUNCTIONS IN: RNA binding, nucleotide binding, nucleic acid binding; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: RNA recognition motif, RNP-1 (InterPro:IPR00504); Nucleotide- binding, alpha-beta plat (InterPro:IPR012677); BEST Arabidopsis thaliana protein match is: RNA-binding (RRM/RBD/RNP motifs) family protein (TAIR:AT5G47620.2); Has 54138 Blast hits to 30457 proteins in 1475 species: Archae - 38; Bacteria - 13880; Metazoa - 18730; Fungi - 4978; Plants - 7803; Viruses - 149; Other Eukaryotes - 8580 (source: NCBI BLINK).	PIF / Ping-Pong family of plant transposases; LOCATED IN: chloroplast; EXPRESSED IN: 16 plant structures; EXPRESSED DURING: 6 growth stages; CONTAINS InterPro DOMAIN/s: Putative harbinger transposase-derived nuclease (InterPro:IPR006912); BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT5G12010.1); Has 644 Blast hits to 642 proteins in 76 species: Archae - 0; Bacteria - 0; Metazoa - 379; Fungi - 49; Plants - 189; Viruses - 3; Other Eukaryotes - 24 (source: NCBI/BLINK).	MLFDQGLIVDFVFLVLLIKETEMDQDFLVMLSN LLHLQNHLDPSSSVFSSDATLSSSSSTSSVSSLL TSSSAAPLLFFTIASVLSFVAASRPSSSSSPTSSS RSPSRSSSSPSDHLSAFRALSTELIWSMDPPV RDAQWRSSYGLSYVPVFTTVVDKLPYIAQCNSL PPDYAVAMVLSRLSHGLSAKTLAKRHSLDPYLVS KITNMVTRLLSTKLYPEFIKIPAGRRRLHETTCAFE ELTSLPNLCGAIDSTPVKLRRLPTDVSDDPSYHC RYGYPALQVAVADHKKIFWDCVCKAPGSYDDA THFRDSVLYNKLMGSDVVDKVVNVVRGHPVVP YVWGDFPLLSKLLTFPFGSTGRGTPAQNTFNAA LMKARLASVEAVGLLKGWRIRLODLNVGLHHAP QTVVACCVLHNLQCIAREPEPELFLKPEENGSLP RVLESESFYYYYGENLRQLTADDLQQLSSR
1025	sp_189670_pjdr -0.833013	-0.870968	-0.813293			

10

20

30

40

【表 1 4 - 4 2 3】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2				LOW EXPRESSION OF OSMOTICALLY RESPONSIVE GENES 1 (LOST); FUNCTIONS IN: translation factor activity, nucleic acid binding, copper ion binding, translation elongation factor activity; INVOLVED IN: response to cold; LOCATED IN: cytosol, nucleolus, chloroplast, plasma membrane, membrane; EXPRESSED IN: 29 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Protein synthesis factor, GTP-binding (InterPro:IPR000795), Translation elongation factor EFG/EF2, domain IV (InterPro:IPR005517), Translation elongation factor EFG/EF2, C-terminal (InterPro:IPR000640), Translation elongation factor EFTu/EF1A, domain 2 (InterPro:IPR004161), Ribosomal protein S5 domain 2-type fold (InterPro:IPR020568), Ribosomal protein S5 domain 2-type fold, subgroup (InterPro:IPR014721), Elongation factor G/III/V (InterPro:IPR009022), Translation elongation/initiation factor/Ribosomal, beta-barrel (InterPro:IPR009000); BEST Arabidopsis thaliana protein match is: Ribosomal protein S5/Elongation factor G/III/V family protein (TAIR:AT3G12915.1); Has 64297 Blast hits to 57476 proteins in 4618 species: Archaea - 1324; Bacteria - 41101; Metazoa - 3172; Fungi - 1934; Plants - 1269; Viruses - 1; Other Eukaryotes - 15496 (source: NCBI Link).	MMVKTADLRRIMDCKHNRMSVIAHVDHGKST LTDLSVAAGIIAQETAGDVRMTDTRADAEERGIT IKSTGISLYEMTDEALAAFKGERAGNDYLNILIDS PGHVDFSSVTAALRITDGLVYVDCIEGVCVQQT ETVLRGALGERIRPVTYVNMKDRCFLEIQVNGEE AYTTFQKVIENANVMATYEDPLLDGQCYYPDKG TVAFSAGLHGWAFTLSNFAMVYASKFGVDESKM MERLWGENFFDPATKKWTSKNTGTATCKRGFV QFCYEPIKVIWATCMNDQDKLWPMMLAKLVGQLK SEEKELMGKALMKRMVMTWLPASSALLEMMIYH LPSPTAQRVRYVENLYEGPMDDDYYATAIRNCDP EGPLMLYVSKMIPAPDKGRFFAFGRVFAKVAT GMKVRIMGPNYVPEGEKDLYKKNVQRTVIMWGGK RQETVEDVPCGNTVAMVGLDQFITKNAATLTNEKE SDAHPIRAMKFSVSPVVRVAVOCKVASDLPKLVE GLKRLAKSDPMVQCSITESGEHIIAGAGELHLEIC LKDLQDDFMGGAETKSDPVVVSFRETVLEKSSRV VMSKSPNKHNRLYMEARPMEDGLAEADDDGRIG PHDDPKIRSKLAEFEFGWDKDLAKKIWCFCPETT GFNMVDMCKGVQYVLEIKDSVVVAGFOWASKE GALAEENMRGICFEVDVLIHADAIHRGGGQVIP TARRVIYASQLTAKPRLEPPVYLVEIQAPENALGG IYSLVNLKRGHVFEEMQRPGTPLYNIKAYLPVVE SFGFSSTLRAATSGQAAPQCVCYFDHWEIMMQDP LDPTSQAAALYNIQIRKRLKGLKEQITPLSDFEDKL
1026 sp_189760_jspe	#N/A	0.871845	#N/A		

10

20

30

40

【表 1 4 - 4 2 4】

A	B	C	D	E	F
1	sp_107620_kqnh (SOAP1)	分析に用いたベイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MGRKGGFSSVEEIVVEGETSNEKLGDKTKG KKGKKGKNGPRSEGGSDGPSTPASSAKPME RRKRRKLLDKVPHSSATESKEIKPAQETPWKFE DFGTASSNGQPVLHMEVTELASADASKREAA AERLAVELQEIQRAYKGGNEDSVGGGLQLEAA KDDGLDNCAPPVRYAVRRLIRGVSSSRECARQG FALGLTIIVSTNSKIRVDSVLKLVDSLEVTSSMKG QEARDCLLGRLFAYGAIARSGRLKEECASDINS YIKFTSLMISVASKKRYLQETVSVLVNLMQVIL PIEAVGNHIVAPGLREWFAAATEGNPDALLA LKIRQKATDSVFDLLPRFSPFRFTVDHLTS IANCLKESTFCQPRVHGWPFLVLSLLPDTVLODA VILEGSLTSSHDKHLAFDILLMLPRLPAFHPIV LSHKFYQGLMDLSTKDSWLYKVAEHFLKQLD WVGSDDIKRVVVALQKHSGGFRFCITRKTIK DLMAGFTTESGCMFLMQKLTSMFYNEEHASDEP SDKNLATDNMSEIGSVEDKESADSNLDLMSWI VESLPTLLKYTLTLDPEAKFRVQKQKVMFLAVQGL FSASLGEVTSFELQEKFEWPKTPISSALCKMCI DQLQLLSNAQKAEGPHALPNSLETNDLGCYFM RELGLTIPSVSLVRLSNEDEKTKKLOALEST LSREERNRGLSADANKLHALRYLLIQLLQLLQRP GEFTEASELICCCKAFSIHDPSESEDEAEEDG ESPELMDVLDLTLVSLPQSSAPLRSVAEQVFKY FCNDVTDGGLRMLRVIKKDLKPSRRKADASED GSDADDDDLLEEAESDEADAVETAAMTDDT EATVTADASDEEFPDASDDSDSDDEDMDDEAMF RMDSYLAKIFQERKINQAGGDTAQAQLIFKLRVL SLLLEYLHENPGSSRVLTVYSNLVQAFVNPSTTE GSEQLAQRWIGLQKFKAKADYKPKSEDYQLVTL EGLEKLSRLASRPFRRKKTSGSVQKQKKSFSW
2	sp_107620_kqnh (SOAP1)	分析に用いたベイトとの相関係数 sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MGRKGGFSSVEEIVVEGETSNEKLGDKTKG KKGKKGKNGPRSEGGSDGPSTPASSAKPME RRKRRKLLDKVPHSSATESKEIKPAQETPWKFE DFGTASSNGQPVLHMEVTELASADASKREAA AERLAVELQEIQRAYKGGNEDSVGGGLQLEAA KDDGLDNCAPPVRYAVRRLIRGVSSSRECARQG FALGLTIIVSTNSKIRVDSVLKLVDSLEVTSSMKG QEARDCLLGRLFAYGAIARSGRLKEECASDINS YIKFTSLMISVASKKRYLQETVSVLVNLMQVIL PIEAVGNHIVAPGLREWFAAATEGNPDALLA LKIRQKATDSVFDLLPRFSPFRFTVDHLTS IANCLKESTFCQPRVHGWPFLVLSLLPDTVLODA VILEGSLTSSHDKHLAFDILLMLPRLPAFHPIV LSHKFYQGLMDLSTKDSWLYKVAEHFLKQLD WVGSDDIKRVVVALQKHSGGFRFCITRKTIK DLMAGFTTESGCMFLMQKLTSMFYNEEHASDEP SDKNLATDNMSEIGSVEDKESADSNLDLMSWI VESLPTLLKYTLTLDPEAKFRVQKQKVMFLAVQGL FSASLGEVTSFELQEKFEWPKTPISSALCKMCI DQLQLLSNAQKAEGPHALPNSLETNDLGCYFM RELGLTIPSVSLVRLSNEDEKTKKLOALEST LSREERNRGLSADANKLHALRYLLIQLLQLLQRP GEFTEASELICCCKAFSIHDPSESEDEAEEDG ESPELMDVLDLTLVSLPQSSAPLRSVAEQVFKY FCNDVTDGGLRMLRVIKKDLKPSRRKADASED GSDADDDDLLEEAESDEADAVETAAMTDDT EATVTADASDEEFPDASDDSDSDDEDMDDEAMF RMDSYLAKIFQERKINQAGGDTAQAQLIFKLRVL SLLLEYLHENPGSSRVLTVYSNLVQAFVNPSTTE GSEQLAQRWIGLQKFKAKADYKPKSEDYQLVTL EGLEKLSRLASRPFRRKKTSGSVQKQKKSFSW
1027	sp_190700_syla	-0.381127	-0.914224	DNA polymerase V family; FUNCTIONS IN: DNA-directed DNA polymerase activity, DNA binding; INVOLVED IN: DNA replication, transcription; LOCATED IN: cellular component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAIN/s: DNA polymerase V (InterPro:IPR007015); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	

10

20

30

40

【 1 4 2 4】

【表 1 4 - 4 2 5】

A	B	C	D	E	F
1					
共発現遺伝子	分析に用いたバイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2				<p>RNA helicase family protein; FUNCTIONS IN: helicase activity, nucleic acid binding, ATP-dependent helicase activity, ATP binding, INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant_structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Helicase-associated domain (InterPro:IPR007502), DNA/RNA helicase, DEAD/DEAH box type, N-terminal (InterPro:IPR011545), Domain of unknown function DUF1605 (InterPro:IPR011709), DNA/RNA helicase, ATP-dependent, DEAH-box type, conserved site (InterPro:IPR002464), DEAD-like helicase, N-terminal (InterPro:IPR014001), DNA/RNA helicase, C-terminal (InterPro:IPR001650), Helicase, superfamily 1/2, ATP-binding domain (InterPro:IPR014021); BEST Arabidopsis thaliana protein match is: DEA(D/H)-box RNA helicase family protein (TAIR:AT5G04895.1); Has 15667 Blast hits to 10761 proteins in 1709 species: Archae - 0; Bacteria - 5024; Metazoa - 4470; Fungi - 1848; Plants - 1329; Viruses - 777; Other Eukaryotes - 2219 (source: NCBI BLLink).</p>	<p>MSMRCFPEPTTSVPTTFSLLKHKFTSSFPPEPKPP LLHLSRPLFLSSMSRPNYQGGFRGGGGGGGR NRGGNRGGGGGGGGGGGGGGGGGGGGGGGG DPVWRAQLAQAAEYEVLDENEWNNKMEGF YKGGQDELIKHLFRREDQQLFDMAQQLLGFH AYNKGKTLVASKVPLPNYRADLDERHGSTQKEIR MSTDIERRVGSLLDTSKGTSTENSIGESSOSSK NSSTISSYKVPPILESHVAKESIAELKHQEEKL KASASVQAMQSFRAKLPANVRSKFLAVAKNQ VLVWGETGCGKTTQLPQFILEEISLLRGDCSII CTQPRRISAVARISSENGENLGETVGYQIRLE TKRSARTRLLFCCTGVLLRRLVQDPNLTGVSHLL VDEIHERGMEDFLILKDLPRPDLRLMSATI NAELFSQYFGNAPSHIPLGLTPVAEFFLEDVLEK TRYSIKQESDSFGSSRRRVKQDSKKDPVTLEF EDVDIDSHYRSTINTRKSLAWSSQLDLSLVQ STIEHICRHEGAILVLTGWDEIKLLDSIKANH FLGDTKFLVPLHGSMPYVQREIFDKPPASTR KIVLATNIAESSITDDVYVIDCGKAKETSYDALN KLACLLPSWISKASAHQRRGRAGRVQPGVYRL YPKLVYDAMLQYQLPELFRTPLOELCLNKSLLQ TVGSLAKALQPPDSLSVQNAIELLKTIGALDERE DLTPLGRHLCTLPLEPNIGKMLLMGVSFVQCLNPA LTTAAALAYRNPVLPINRKEEADAANKSFAGDSC SDHIALLKAFEGWKEAKSSRTEKSFQWENFLSP VTMQMDDMRKQFVDLLSDVIGFVYDVKSKGPNAYN QYSHDMEMVCAILCAGLYPNVVOCKRRRKRRTAF YTKEVGQVDIHPASVYVAGVHLFPLPYWYSEKVK TTSIYIRDSTNISDYALLFGGSLPKSGEGEIML GGYLHFSASKSVLEMLMKLRLELDKLLSRKIEEPR LDITAEKGSVVAAYWELLHS</p>
1028 sp_192520_rhen	#N/A	-0.336977	-0.312352	unknown protein; Has 1784 Blast hits to 634 proteins in 116 species: Archae - 0; Bacteria - 0; Metazoa - 1013; Fungi - 200; Plants - 288; Viruses - 0; Other Eukaryotes - 283 (source: NCBI BLLink).	<p>IMVKGIDELTEDEKALRRRKRFPAPLPKPSSSCS SLPRLAHPGGPVTTKAVAKLAKFLERKLPQFGL DTLKLPELIELAVYNAKAVETVNASGTSKHGKIIRHVT SFDKDNEDYSSKDWPTPEYKIKLKEKKDKKKKKK KKKKNKKKVVDDQ</p>
1029 sp_193100_dzum	-0.372353	-0.383565	-0.323834	uncharacterized protein	<p>MIQPLYLRGDEIHLVLYIGEDGLHILQLLHRSF GELDDDEETKHHFIDLKLSKGRFRL</p>
1030 sp_193160_ktrmf	-0.377466	-0.391833	-0.38367		

10

20

30

40

【 1 4 2 5】

【表 1 4 - 4 2 6】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたベイトとの相関係数 sp_107660_kiqg (SOAP2) sp_107670_ptqx (SOAP2-like)		注釈	aa seq
1031	sp_193390_qijv	#N/A	#N/A	SNF7 family protein; FUNCTIONS IN: molecular function unknown; INVOLVED IN: protein transport; LOCATED IN: cellular_component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 12 growth stages; CONTAINS InterPro DOMAIN/s: Snf7 (InterPro:IPR005024); Has 30201 Blast hits to 17322 proteins in 780 species: Archaea - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MAELKVFIREQVPDWDVDDVKATARFKAFSG QFSDWESQFFWRDLIKIAREFDLLFLSPSQMK NCWFNRGGLSPLCLNDSLEMYKAGNVLRVDF GDPTSGRLSQLKLIHFSIVFRPSTNDILNDRLI LRVMLEKSADVIKALSETHWTSJCITMRKFESL FREHNEAYAVLCYLSGYGKARYSVKKEELIEG VKVLSAAAPILPSTDSAILQLVWTLLEELQKLL VIDEQYKRLRISALAAKSGDKKALRNARQLKA SESREKLTSLFNRVEEVLRITDAESAKKVSFAIQ LGALALKENRTDVEKVMCLQELDEFRDSQEIHH NALESTSYTAMEDEDEIEELKRLDQIRDENLWV SALKSKDDAIRNAGASVTPDSLCAAIVDLKLESET QETTQGPGRNRTSLELEAA
1032	sp_193890_zsia	#N/A	#N/A	uncharacterized protein	MFNMTVQTQNT
1033	sp_194030_anpy	#N/A	#N/A	uncharacterized protein	MCNLPFVYSLFIPTQHLFLPLQKDDIAKAAATEM VRTKPIVLQGLPTVALSKSNALQWCLKVALEEN TEEASEENEENDQAALEAASEENDQAALENTAAP KK
1034	sp_194100_fsiz	#N/A	#N/A	uncharacterized protein	MKNVAKDVTWCELOQNPVNHRYFERKLRPKPFG RGHACLGVTHRVSPNPRTAGRRRMMASHPSPG VDGLNKPPVTNCRSDRW
1035	sp_194830_isew	#N/A	#N/A	uncharacterized protein	MAPLLODLRAVIREQVIRRLGAKESVAEKDPSPG VAEKDPSPGVAEKDPSPASSHHRQISDNAYINVD GEPMAVDKSNKVLSELDVIVLEIDNVQHRVE QPLGGEQNVFNYSQSLPVEQPVFEFLHVPAVEA VONEPMEAEQVQQGP
1036	sp_195680_enhj	#N/A	#N/A	uncharacterized protein	MFNCLERMVYNITAFHERIWNKNTTSMLKVVICIP YSRLGCLHFHFESSFVRIVYVYGMGP
1037	sp_195880_ggoo	#N/A	#N/A	uncharacterized protein	WPTRHRLDGGCSFSPYSASPLSAGSYGCSGAA AGCCCCSLSTRSATGRRRSAGSGPVSVE

10

20

30

40

【表 1 4 - 4 2 7】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたペアとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
1038	sp_196220_mdwg	-0.88143	#N/A	-0.845357	<p>DEAD/DEAH box RNA helicase family protein ;            FUNCTIONS IN: helicase activity, ATP-dependent            helicase activity, nucleic acid binding, ATP binding;            INVOLVED IN: response to cadmium ion; LOCATED            IN: nucleolus; EXPRESSED IN: guard cell, root,            cultured cell; CONTAINS InterPro DOMAIN/s:            DNA/RNA helicase, DEAD/DEAH box type, N-terminal            (InterPro:IPR011545), RNA helicase, DEAD-box type,            Q motif (InterPro:IPR014014), DEAD-like helicase, N-            terminal (InterPro:IPR014001), DNA/RNA helicase, C-            terminal (InterPro:IPR001650), Helicase, superfamily            1/2, ATP-binding domain (InterPro:IPR014021); BEST            Arabidopsis thaliana protein match is: DEAD/DEAH            box RNA helicase family protein            (TAIR:AT5G11200.1); Has 30201 Blast hits to 17322            proteins in 780 species; Archae - 12; Bacteria - 1396;            Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses            - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).</p>	<p>MGETRENDAYEEELLDYDEEEKVPDPVTKTN            GEAAKKGYSYSSGFRDFLLKPELLRSIVDCGFE            HPSEVQHECIPQAILGMDVICQAKSGMGKTVFV            LSTLQIQIDPVAGQVAALVLCHTRELAYQICHEFE            RFSTYLPQVKVAVFYGGVNIKHKDLLKNECPHIV            VGTGPRILALARDKDLNKNVRRHFLDECCKMLES            LDMRRDVQEIFKMTPHDKQVMFSAATLSKAIKRPV            CKKFMQDPMEIYYDDEAKLTLHGLVQHYIKLSEL            EKNRKNLDDLALDFNQVIVFKSVSRAAELNKLKLL            VECNFFSICHSGMSQEEERLARVYKGFKDGKQRIL            VATDLVGRGIDIERVINIVINDYMPDSADTYLHRVG            RAGRFGTKGLAITFVSSASDSDVNLNQVQERFEVD            IKELPEQIDTSTYMPS</p>

10

20

30

40

【表 1 4 - 4 2 8】

A	B	C	D	E	F
1.	sp_107620_kqnh (SOAP1)	分析に用いたバイトとの相関係数	sp_107670_ptqx (SOAP2-like)	注釈	aa seq MSTRGMSVAFHFRKSLTDNRKYMIGDGEIGKGYGRR VYKGLDLENGDFVAIKQVLSLENIQAQEDLNIHQEID LLKLNHNKIVKYLGSLKTRSHLHILEYVENGSLA NIIKPNKFGPFPEQLVAHYIAQVWEGLVY_LHEQGV IHRDIKGANILTTKEGLVKLADFGVATKLTADYVN THSVWGTPYWMAPVEMVMSGVCAASDIWVSGCT VIELLTCVPPYYDLQMPALFRVQDEHPIPIQDGL SADMADFLRFCOKDARLRPDAKTLLSHPWILNY RRTLQPSLPASDAIRNTOEGLMDITELSNGENG VAESHSDVEARRETLTADTSTADDDNOGLSE EYSGFCDDLSADQVPTLAIHEKSVAKSTSMISARI QGAGSEQISEQGHITDPDEVLTNGEAAAPKSKRK DGVSGKYDGGKGSVHKMNSLFSFGTRNQDENL CKVAKKGISTGNELSKFSFDPGDASLEDLFLPLD KNTEADLTSASTSTSHLNPSSITLESKKNDLA TKLRAAIAQKQOMENEMGOANGNLLRVMIRALK EDVIDBSSLVFDKMPGENLFLQAVVEFSVLIAS LRPEEDEDVIVNACLIAFFNQVSEKQKVVFKQ HGLLPLMELLEVRPRPVICSVLQVINGIKONTDFQ ENACHVGLIPVIMGFVAPDRPREVRMEAAVFLQQ LCOSSSLTLQMFACRGPVLYGFLEADYAKYRE MVHLVADGMWQVKLQKSTPRNDFCRIAAKNGI LLRLINTLHLSNEATRLASISGSLTSDGVSQRP RSGOLDPSHPVFSPENEATHSGSDPVDLVRHSA VEHSSSTVDADKPPQSSHAIVESSASSRSTDAADK FVNVKAEATRAEIDLROHRVAGSSNRSTDRPPK VEGGNGLPYAVSSQDQVVRPLLLEKEPPSR HFGQLDLROLAGMERHESILPLLHASNDDKGN GELDFLMAEFAEVSGRGRHNGNSRSETPRISNKTA VKKLSVASNDGAASTGSLASQTASGVLSSGVL NARPGASTSSGLLHVMVSAMINADYSREYLEKVA ELLLEFAQADTTVKSYMCSQSLTLRFLHMFNRVE
2	sp_107660_kiqg (SOAP2)			mitogen-activated protein kinase kinase 7 (MAPKK7); FUNCTIONS IN: protein serine/threonine kinase activity, binding, protein kinase activity, kinase activity, ATP binding; INVOLVED IN: plasma membrane organization, pollen development; LOCATED IN: cytosol, plasma membrane; EXPRESSED IN: 6 plant structures; EXPRESSED DURING: ovule developmental stages; CONTAINS InterPro DOMAIN(s): Protein kinase, ATP binding site (InterPro:IPR017441), Serine/threonine-protein kinase domain (InterPro:IPR002290), Armadillo-like helical (InterPro:IPR011989), Armadillo (InterPro:IPR000225), Serine/threonine-protein kinase-like domain (InterPro:IPR017442), Protein kinase-like domain (InterPro:IPR011009), Serine/threonine-protein kinase, active site (InterPro:IPR008271), Protein kinase catalytic domain (InterPro:IPR000719), Armadillo-type fold (InterPro:IPR016024); BEST Arabidopsis thaliana protein match is: mitogen-activated protein kinase kinase 6 (TAIR:AT3G07980.1); Has 137235 Blast hits to 134768 proteins in 5401 species: Archaea - 190; Bacteria - 15422; Metazoa - 52001; Fungi - 13364; Plants - 33616; Viruses - 670; Other Eukaryotes - 21972 (source: NCBI BLINK).	
1039 sp_196330_cuqo	#N/A	-0.352633	#N/A	RHOMBOID-like 2 (RBL2); CONTAINS InterPro DOMAIN(s): Peptidase S54, rhomboid (InterPro:IPR002610); BEST Arabidopsis thaliana protein match is: RHOMBOID-like protein 6 (TAIR:AT1G12750.3); Has 6856 Blast hits to 6855 proteins in 1947 species: Archaea - 152; Bacteria - 4411; Metazoa - 521; Fungi - 179; Plants - 374; Viruses - 0; Other Eukaryotes - 1219 (source: NCBI BLINK).	MGSRDLKKNQDYNSSYRNYNSYNTQPNHPVA PSGYHIESSARQWTSWLYPMVVAVVAVMVFVVMF INNCPKNFNGFCVAFRLGRFSSFOPLRENPLLG PSANTLMNLGALQWDKVVHQHGGWRLLTCWLW AGVHLLVNM LSLVFIGRLEQOFGFVGRVGLVVIS GLGGVMSALFLOKSISVSGSAGALFGLGAMLAE LLTNWTYANKAAALFTLLIMAINLAVLILPHVDNF AHIGGSGFLGLFVLLRPPQYQVGLERHQYPEHA RVKSKFMPYQMVLLFFIGAAALITGLTIALVMLFOG KNGNDHCRWCRYLSCVPTSKWQCCGN
1040 sp_196400_wumz	#N/A	-0.2856272	#N/A	uncharacterized protein	#N/A
1041 sp_197361z10483	#N/A	-0.2831523	#N/A		

10

20

30

40

【 1 4 2 8】

【表 1 4 - 4 2 9】

A	B	C	D	E	F
1					
	分析に用いたペアとの相関係数				
2	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
1042 sp_197390_sacn	-0.971351	-0.951784	-0.920002	small ubiquitin-like modifier 2 (SUMO2); CONTAINS InterPro DOMAINs: Ubiquitin (InterPro:IPR000626), Ubiquitin supergroup (InterPro:IPR019955); BEST Arabidopsis thaliana protein match is: small ubiquitin-like modifier 1 (TAIR:AT4G26840.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).	MSGTQEEDKKPGGDQAAHINLIKVKGQDGNVEVFF RIKRSTQLKLMNAYCDRQSVELNSIAFLFDGRR LRAEQTPDELEMEDGDEIDAMLHQTGGAMI
1043 sp_197640_frpi	-0.934363	-0.940742	-0.934354	CW7; CONTAINS InterPro DOMAINs: Protein of unknown function DUF2045 (InterPro:IPR019141); Has 145 Blast hits to 144 proteins in 52 species: Archae - 0; Bacteria - 0; Metazoa - 91; Fungi - 0; Plants - 50; Viruses - 0; Other Eukaryotes - 4 (source: NCBI BLink).	MKGFGRGEVEVAVSGVEDQSEHSSGSPGSG SKGAFGLATIVRRAASVASVAAKQAYAAASINAD DDMLPLNCSLSVSLPWEHIAHDLLLVKGGPPVSMI
1044 sp_198000_xzyn	#N/A	-0.974941	-0.959317	profilin 5 (PRF5); CONTAINS InterPro DOMAINs: Profilin/allergen (InterPro:IPR02097), Profilin, plant (InterPro:IPR005455); BEST Arabidopsis thaliana protein match is: profilin 4 (TAIR:AT4G29340.1); Has 915 Blast hits to 914 proteins in 244 species: Archae - 0; Bacteria - 2; Metazoa - 138; Fungi - 121; Plants - 583; Viruses - 0; Other Eukaryotes - 71 (source: NCBI BLink).	MSWQTYVDDHLMCEIEGTTNHLAAAAIIGTDGSS VWAQASAFFQFKPDEIAAVVWDFDEPGTLAPTGL HLGGVKYVMIQGEPEGAVIRGKKGSGGICVKKTA QALVIGLYDEPVTPTGCCNMIVERLGDYLDIQGL
1045 sp_198640_nawc	#N/A	-0.925733	-0.8594179	PAR1 protein; FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: endomembrane system; EXPRESSED IN: 13 plant sirucures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAINs: PAR1 (InterPro:IPR009489); BEST Arabidopsis thaliana protein match is: PAR1 protein (TAIR:AT3G54040.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archae - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLink).	MASHFNMSALVTLISLLSSLIQGLTGEVTCEQLN KEACAFVSSIGKRCVLEKYVRRSGLLEYKCTAS SIEADKLDLVETNHCIEACGLYRETFGSSDSL ESRFTQKLCSECYQRCRNVDLYFNLAAGEGV FLPKLCETQSENPRRGLSDIKSSGIVAPGVPVSGN DDSKLFGITISIGSPLPSMELAAPAPAPAMPPL

10

20

30

40

【表 1 4 - 4 3 0】

A	B	C	D	E	F	
1						
2	分析に用いたベクトルとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq	
1046 sp_198840_rzcc	-0.832312	#N/A	-0.834634	ARM repeat superfamily protein; FUNCTIONS IN: protein transporter activity, binding; INVOLVED IN: intracellular protein transport, protein import into nucleus, docking; LOCATED IN: nucleus, nuclear pore, cytoplasm; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Importin-beta, N-terminal (InterPro:IPR001494), Armadillo-like helical (InterPro:IPR01989), Armadillo-type fold (InterPro:IPR016024), BEST Arabidopsis thaliana protein match is: ARM repeat superfamily protein (TAIR:AT3G59020.2); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).	MTNAIDLHDQOQWLNCLNATLDTNQOQVRSFAEA SLNQATLQHGFGSALAKVAANRELPMGLRQLAA VLLKQFIKHKHDEDEDNFEHPVWACDEKEVIRR LLLLCLDDSHRKTCTAVSMASAVASIAHYDWPENWPP ELLPLMLKINDQTNMNGVLGALRCIALLLSGDIDD KAVPSLPIFLPCLLSIVSAPQIYEKSMRTKALSIV FYCTSMILGTMISGAYKAETSYLMSPLMQPWMEQ FSAVLSHPQPEPEDWVSRMEVYLKCLNQFQIN FPSATEKFTSILGPIWQTFVTVKVKYKSSVEGA QDSFDGRYSDGGAERLESFVQLFEFELLTVGSS RFVKVADNLQELVYTFGFMQITEHOMHAWSL DANQYSDDEDDMMYNCRVSGVLLLEIEISSCGED GLYTLVGAARRRFFIESQQKADGFFEFWRIREA VIYAMASLDQFQEAEEVCGPDGDLKFKVEQMIM EDYGAGVHEQPYLHARVFSVPKFFSSMLSPGVV EHFLFASLRAIGMDVPPPVKVGSCRVLSQLLSEA NKEALPQIQLQFSSLDLNLQASEES.HLVLETL QAAMSGHEASLSVEPVSPILNITWVSHVSDPFIS IDAIEVLEAIKNAPGCIQPLVSRVLPYIGPILSQPKQ QADGLVAGSLDLVTMLLNKAPGPVVKALYDVCF NPIQIVLQSDDHSEMONATECLAAFVCGGKSELL AWGGDSGFTMRSLLDALS	MEGPIESKKETASVSASQVQSEDHEIKARVDVAV WQQMNGYSVKTMNSIVNKSSSTGKTMTQTDK ASINWMSLLGLGSKTKTPTVASTKGNSSPFAQN GTSDEARLKLAAALAAARDAATAASASAMNRGKIE VTEVRDFAGEEIEVKKLIDPSSKAAERAGPSSA VDSVLEQIKKPKLNLDDKTKDWGFEKGENKDY EEELDAYKSSNQYLDKV/SF.LSPADYREFEFERER DVRLAQQAKRKPD
1047 sp_199140_gfod	#N/A	#N/A	-0.882336	FUNCTIONS IN: molecular_function unknown; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; EXPRESSED IN: 22 plant structures; EXPRESSED DURING: 13 growth stages; CONTAINS InterPro DOMAINs: Craniofacial development protein 1/Bucantaur (InterPro:IPR011421); Has 333 Blast hits to 324 proteins in 149 species: Archae - 0; Bacteria - 18; Metazoa - 117; Fungi - 96; Plants - 49; Viruses - 0; Other Eukaryotes - 53 (source: NCBI BLINK).	MESGDEARHRRLLRRFLIQRRNRARNRNINPTHV MVTDSRPRHLPVIVPQHPMPMMNAVQVRGMD NGDEAKHRRLLRRFLIQRRNRARNRNINPTHVFRVT DSRPRHLPVIVPQHPVASMNAVERAHRSLTSR GETSAPNNVIGSGTSPRRRRSENRESTSSLR RRLNSSLPPDGSMTPTTETDIGNTRSPQPNLRV GSNISRNTLEGSASHLHRRKYTRPWTFDITTG VETPLDNATGSNTTVARHLLDNASENDRVTRVET ARKPRNRNEMETDKETTITANREN	
1048 sp_199170_awnz	#N/A	-0.9039406	#N/A	uncharacterized protein	#N/A	
1049 sp_199381z10789	-0.9130365	#N/A	#N/A	uncharacterized protein	#N/A	

10

20

30

40

【 1 4 3 0】

【表 1 4 - 4 3 1】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
1050 sp_199410_irgg	-0.856776	#N/A	#N/A	ribonuclease 2 (RNS2); FUNCTIONS IN: ribonuclease T2 activity, endoribonuclease activity, RNA binding; INVOLVED IN: aging; LOCATED IN: intracellular, vacuole; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAINs: Ribonuclease T2 (InterPro:IPR001588), Ribonuclease T2, active site protein match is: Ribonuclease T2 family protein (TAIR:AT1G14220.1); Has 2318 Blast hits to 2317 proteins in 471 species: Archae - 0; Bacteria - 380; Metazoa - 295; Fungi - 231; Plants - 1299; Viruses - 5; Other Eukaryotes - 108 (source: NCBI BLINK).	MTMALSISIFALLLITFTVAVASTADVSSLLSSN QREFDYQLALQWPGYCYRTHKCCOKNGCCR NSGAPQHFTHGLWPDYNDGTTWSPSCCTKSNFD EKEISTLRKDLDEWYWPVSLSCGSPSSCHGKGSF WGHVQWVPLPTQYELDYICFPNMFVTLVLCFLV PGSHDLRISGYYKKGHTCSPVVKDEYNYFRGTLT VYFKYNITEALSDAGYVPSNSEKYPVPLGGIISAENA FHSPVVCSDAVKELNLCFYKDFKPRDCATRS RMDNKSITSCPKFVSLPTTYDSSRLGIVGALESFVI
1051 sp_199570_huwe	-0.866353	#N/A	#N/A	STELAR K+ outward rectifier (SKOR); FUNCTIONS IN: outward rectifier potassium channel activity, cyclic nucleotide binding; INVOLVED IN: potassium ion transport, response to abscisic acid stimulus; LOCATED IN: membrane; EXPRESSED IN: 8 plant structures; EXPRESSED DURING: L mature pollen stage, M germinated pollen stage, 4 anthesis; CONTAINS InterPro DOMAINs: Cyclic nucleotide-binding (InterPro:IPR000595), Potassium channel, voltage-dependent, EAG/ELK/ERG (InterPro:IPR003938), Protein of unknown function DUF3354 (InterPro:IPR021789), Ankyrin repeat-containing domain (InterPro:IPR020689), Ion transport (InterPro:IPR05821), Cyclic nucleotide-binding-like (InterPro:IPR018490), RmlC-like jelly roll fold (InterPro:IPR014710), Ankyrin repeat (InterPro:IPR002110); BEST Arabidopsis thaliana protein match is: gated outwardly-rectifying K+ channel (TAIR:AT5G37500.1); Has 82938 Blast hits to 34972 proteins in 1815 species: Archae - 147; Bacteria - 10317; Metazoa - 39460; Fungi - 7131; Viruses - 3765; Other Eukaryotes - 20980 (source: NCBI BLINK).	EQGNVVDQLYL VCHGKLVTKVHEEFFLPGEVILE QGNVVDQLYL VCHGKLEEVISIGDQSGEQLISRLE PDSTFGQNSIFCNIPQSCVTRVVDLRLRLRIDKQS LSNIDYFFDGGKIFDNLKNGNDGKFNLTVESDI ASHITQOESEALAKVNNAAHNGDL YOLKSLKAG ADLNKTDYNERSPHIAASKGHENIVLFIQAGAL VNQSDNFGNTPLLEAAKNGHDQIAAVLYREGARI DIKVDGTYLCTVLSKESKDVDFLKRLLNNGVDP NSRDYDGRTPLVHACSHGLYLMAKLVEAGALV LKDRWDNTPLEARLFGHNNHLIKLEAAKSAQLS QFAGQIEQSDRKLHKKCTVFPVHPWDTQEKRIIP GIVLWIPRTIDELVKTASEKLNLPATATCVLSEDDGG KILDIEMIDNGGQKLYLITETD
1052 sp_200130_aejf	-0.863333	#N/A	#N/A	uncharacterized protein	MIDAIEKSGLR

10

20

30

40

【 1 4 3 1】

【表 1 4 - 4 3 2】

	A	B	C	D	E	F
1						
2	共発現遺伝子	分析に用いたペアとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
1053	sp_200840_pngi	-0.85788	-0.319373	-0.858004	Protein of unknown function (DUF3537); CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF3537 (InterPro:IPR021924); BEST Arabidopsis thaliana protein match is: Protein of unknown function (DUF3537) (TAIR:AT3G20300.1); Has 140 Blast hits to 140 proteins in 15 species: Archae - 0; Bacteria - 0; Metazoa - 0; Fungi - 0; Plants - 138; Viruses - 0; Other Eukaryotes - 2 (source: NCBI BLink).	MAYSLPSPHYSPLEEQWYNPDSAQTTETLILPLKH SLRRLKTLRFLGFHDSRFLSLSSFSLLFGIAA PSLIIVFFYCGSGSASACRKYHIKGFLETLFF QATAAAYSLLCFSLNFRKYGLRVFLVDNYQQNV DRFQFEYIPKIDNFRSLAVVLPVCVILKTAREIVR VIYVHNDLSVQSMVMVFLVLSWYTYTATVYLTGS ALFSLVCNLQVHFHEKYKNLLDRNLDLVLYIEEHS RLTFNLKISHRRFRIFLILEFLIVTASQIVTLLETTGN NGTINFNAADFAISSIVQVGIICLHAAAKISHRA QSLASFASRWHALLTCSRNETLGAASENGNLNV TGSAVAAMNYPESDMDSDVDYLPPINEEMNSNV SYLKRQAFVVMYMQSNPQGGVTFGWTVNRAFMD TIFCVELTVVTVFLVGLKTLTFTTNT
1054	sp_201140_jppd	-0.893288	-0.3436245	-0.8586535	AT-HSFB2B; FUNCTIONS IN: transcription repressor activity, sequence-specific DNA binding transcription factor activity; INVOLVED IN: response to chitin; LOCATED IN: nucleus; EXPRESSED IN: 24 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAIN/s: Winged helix-turn-helix transcription repressor DNA-binding (InterPro:IPR011991), Heat shock factor (HSF)-type, DNA-binding (InterPro:IPR00232); BEST Arabidopsis thaliana protein match is: heat shock transcription factor B2A (TAIR:AT5G62020.1); Has 32476 Blast hits to 10764 proteins in 847 species: Archae - 36; Bacteria - 10454; Metazoa - 9035; Fungi - 1651; Plants - 6330; Viruses - 666; Other Eukaryotes - 4304 (source: NCBI BLink).	MIPVP_LDQNGDSTGCGGTSGNASDLQQRLLPT PELTKTYLLVDDPVVDDLISWINDDGTSEIWRPA EFARDLLPKYFKHNNFSSVROLNTYGFRRKVPD RWEFANDCKRGEKALLRDIQRKRKISAPAINSP TAAACPTAIPA VVTVAAPPVTVRAVSPANSGD EQVHSSNSSPAPGVTVQVAVASCASCTTELEENE RLRSENFLLSQELNHLRGLCNLILAMMSNYATGQ PETSGLQEGRP.LDLPTKPVTVAEADPCGASSG GATASGGSGKEEKMEEEEEMEGCPCLRFLGVSIG AKFVRREDEEENRPAVSERSRSREGSDPIQTQQE QQCSGE

10

20

30

40

【表 1 4 - 4 3 3】

A	B	C	D	E	F
1					
2	sp_107620_kqnh (SOAP1)	分析に用いたペアとの相関係数 sp_107660_ktgg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
1055	sp_201480_aomj #N/A	-0.998173	-0.928974	cytochrome P450, family 90, subfamily D, polypeptide 1 (CYP90D1); FUNCTIONS IN: oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, NADH or NADPH as one donor, and incorporation of one atom of oxygen, oxygen binding; INVOLVED IN: stamen development, petal development, leaf development, brassinosteroid biosynthetic process; LOCATED IN: endomembrane system; EXPRESSED IN: 15 plant structures; EXPRESSED DURING: 8 growth stages; CONTAINS InterPro DOMAINs: Cytochrome P450 (InterPro:IPR001128), Cytochrome P450, E-class, group 1 (InterPro:IPR002401), Cytochrome P450, conserved site (InterPro:IPR017972); BEST Arabidopsis thaliana protein match is: Cytochrome P450 superfamily protein (TAIR:AT4G36380.1); Has 27722 Blast hits to 27661 proteins in 1592 species: Archaea - 80; Bacteria - 4717; Metazoa - 10119; Fungi - 4209; Plants - 7293; Viruses - 6; Other Eukaryotes - 1298 (source: NCBI BLINK).	MDLCSWTLISSLVIAISLSYIIFHL YRNTLKNFP KKLTWTPLGTLGWPLLGGETMEFISCAYSDHPES FMDRRRRKYGNVFKSHIFGKQTIHSTDAEVRNFL QSDAKLFVPCYPKSLTELMGKSSILLINGSQRRR HGLIASFFKSSHLKAQVTRDMQCVCYVQQAQMDTWV DNIHSPILIQNAKHAIFQVLVKALISLDPGKEMHF LKKQFQEVSG.LMSL.PVNLPGLTQLYRSLQAKKE MVKLMKKIIVTKRNKMIQAGDPNDVLDVLLSDQS GOLTDLLMSDMMIMMIPGEDSVPLLITLAIKYL DPLALQKFTFEENVKYNKIKARLGEPMYWNIDYL SLPFTQMVRVLFREILSKN
1056	sp_201820_wssd #N/A	#N/A	-0.989529	unknown protein; BEST Arabidopsis thaliana protein match is: unknown protein (TAIR:AT3G03460.1); Has 1807 Blast hits to 1807 proteins in 277 species: Archaea - 0; Bacteria - 0; Metazoa - 736; Fungi - 347; Plants - 385; Viruses - 0; Other Eukaryotes - 339 (source: NCBI BLINK).	MDEAKQLMAGRQHQLLMQQQQQQQQQQQQQQQQ YLLHFQQQQQQQQQQQQQQQQHQHQQQQQQQQQ HHQQQQQQQQQQQQQQQQPQNQLHQQQQQQ QQQAAMARFFSNIDAHLRPLGGANRPLGGPQN QNPNSNSIHPHQQQNPFSSSQSQHQQLLQ QQLQQQGMNRYLSQAQQQQAQVPTLAGNQAEL QMA'YQDAWRVCHPDKRPFASLEDACERLLPYH VVADYEAEEEDRILDSDTTGQMFSSRQQWDHNI AAKVAEFTSTFEKQALAFNIITRKRALGEFRSEER LMIEQALLQEEKRAIMEVRAEIESRAGRE/AHEAK MRMAAQAEQARAESQAHAELMARAPIRVNAL GSQGNDA5MNHDMTDOEQGVIPDEMINGWGN ACRVEREP5DDFLNDEEAENGDTGMGEEWRES GFDLNSR
1057	sp_202750_xnjg -0.938814	-0.99457	-0.935649	uncharacterized protein	MGLDQE
1058	sp_202791z11248 -0.877333	-0.98692	#N/A	uncharacterized protein	#N/A
1059	sp_202860_owok #N/A	-0.877879	#N/A	uncharacterized protein	MARKAAA5WECKWVASEKRLVDLAAVYVSKDE QLEGKDDQIADASKELERVKGELASTVEEMQGL RILYDALQEEFKDCEALAVWRTRAQMMF5CLKG EVSLLWPCQKEVDYLANGGTLLEELSPVVDADA LAMEADTAGTNGADADVAPLVEDYSSVDREGEA SMEQGE5DVPVAVPEV5DLADATVNVDMPPISD QV5T5PLPDEXDN

10

20

30

40

【表 1 4 - 4 3 4】

A	B	C	D	E	F	
1						
2	共発現遺伝子 sp_107620_kqnh (SOAP1) #/N/A	分析に用いたペアとの相関係数 sp_107660_klqg (SOAP2) 0.352233	sp_107670_ptqx (SOAP2-like) #/N/A	注釈 uncharacterized protein	aa seq #/N/A	
1060	sp_203331z11568			squalene epoxidase 3 (SQE3); FUNCTIONS IN: squalene monooxygenase activity; INVOLVED IN: response to jasmonic acid stimulus, response to wounding, sterol biosynthetic process; LOCATED IN: endomembrane system, integral to membrane; EXPRESSED IN: 23 plant structures; COMPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Squalene epoxidase (InterPro:IPR013698), Fumarate reductase/succinate dehydrogenase flavoprotein, N-terminal (InterPro:IPR003953); BEST Arabidopsis thaliana protein match is: squalene epoxidase 2 (TAIR:AT2G22830.1); Has 30201 Blast hits to 17322 proteins in 780 species: Archae - 12; Bacteria - 1396; Metazoa - 17338; Fungi - 3422; Plants - 5037; Viruses - 0; Other Eukaryotes - 2996 (source: NCBI BLINK).		MFFDQYIFGGVIASLLAVAFWLRLRSTKREDKY SQEDDVSQVTTTSPVGECRSDVTDSDVIIVGAGV AGAAIAHLGKDGRRVHVERDLTEPDRIVGELL QFGGKLDLIDLGLEDVCVENIDAQRVFGYALFKDG KNITKSYPLEKFKHDDVSGRGRFHNGRFORMREK ASSLPNVLEQGAQVTSVLEENGIKGVRYKTKAG KELCAYAPLTIIVCDGCFSNLRRSLCTPKIDQPSCF VGLVLENCLELPIYENHGHVVLGDPSPILFYPSISTE VRCLVDVPGKEKVPVANGELAKYLLTVVAPQIPP QLYNFAVAIDKGNKSNMKNKMPAALPKPGAL LMGDANMRHPLTGGGMTVALADIVLLRDLRPL KDLNDAPNLCRYLESFYTLRKPVASTINTLAGALY KVFRAFSPDQARKEVMOACFDVYLSLGGIFSNGPS LLSGLNPKPLSLVHFFAVAFVGFVGRLLLPFSPK RLWIGARLISGASGIIFPIIAKAEVGRKMFPAIVPAY YRYPVVKVS
1061	sp_203380_qtyw			uncharacterized protein	#/N/A	
1062	sp_203681z11614			UBP1-associated protein 2A (UBA2A); CONTAINS InterPro DOMAIN/s: RNA recognition motif, RNP-1 (InterPro:IPR000504), Nucleotide-binding, alpha-beta plait (InterPro:IPR012677); BEST Arabidopsis thaliana protein match is: RNA-binding (RRM/RBD/RNP motifs) family protein (TAIR:AT2G41060.2).	MAKRRKLSAKSAETATSAQQKLOKQKTONNEEQ EQQPEAQKQPEQEQPEKLVAVEEEEEQPPMEE EVEVEEVEEVEEVEEVEEVEEVEEVEEVEE GDEIETEETEEPKPIEEDHTTTQSSDDEPIPLLE PFTKEQLMILCEAAINHPDVASRIKRVADIEDTVH RKIFVHGLGWDTTADTLRSFASQYGEIEDCKAVT DKISGKSGYGFILFKSRSGARKALEMPQKRIGG RMTACQLAAVGPVPTSSMGTAAVVQFPQPMSE YTORKIYVSNVSDIDPHKLSYFVKFGEIEEGPL GLDKHTGPKGFCFLVYFVSVESAKKALEEPHKH FEGHVLHCQQAIDGPKPKQHMHGHPQGHFG GRVDNPNLAGAIGGPPATGAGLMGSGISGGVYN QTGAAPGPNPAGQALTALLASQAGAGLGNLL GTLGGMAGAPGASHGYGAPNVGVPNSVLLGGY GQGG-LQGAYPNQHMGMGGGAGGRSSQHGHHM GVAPYGGH	
1063	sp_203690_pchy					

10

20

30

40

【表 1 4 - 4 3 5】

	A	B	C	D	E	F
1		分析に用いたベイトとの相関係数				
2	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_klqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
1064	sp_203710_mgoq	-0.872488	-0.883554	-0.95183	N2,N2-dimethylguanosine tRNA methyltransferase; FUNCTIONS IN: RNA binding, tRNA (guanine-N2-)- methyltransferase activity, INVOLVED IN: tRNA processing; LOCATED IN: cellular component unknown; EXPRESSED IN: 23 plant structures; EXPRESSED DURING: 14 growth stages; CONTAINS InterPro DOMAINs: N2,N2- dimethylguanosine tRNA methyltransferase (InterPro:IPR002905); BEST Arabidopsis thaliana protein match is: N2,N2-dimethylguanosine tRNA methyltransferase (TAIR:AT3G02320.1); Has 1017 Blast hits to 963 proteins in 359 species: Archaea - 255; Bacteria - 70; Metazoa - 198; Fungi - 150; Plants - 103; Viruses - 0; Other Eukaryotes - 241 (source: NCBI BLINK).	MGSSKDGSYDEEQPKGNPNTSVDFNDYAVIKEG EAELIMHRNNQVFFNKTOVNNPDIJSIAVLRTEINK RKEEHEAFLSRKAKGAAKTSEEKSDTPASEDYVT MESKADSHTSNGIHDITNATDVKPEDASCGTAAE PAQPSNTPGGKFCRELKPPRVLEALSASGLRAL RYAREVEGLGQIVALDNDKASVEACQQRNIKFNNGS VASSKVESNLADARVYMLTHPKFEFVVDLDPYG SPSVFLDSAVQSVADGGMLMCTATDMAVLGGG NAEVCYSKYNSIPLKSKYCHEMALRILLASIESHA NRYKRYIVPVLVFRMDFYIRVVFVRIYSSASAMKET PLKLSYVYQCYGCDSEFLQPLQGRAVSKNNMKNKY YFGYGPVVPKECADCCKRFRNMGGPIWSAPIHDQ DWVTSMLSDVKAMIKDRYPAYNRIIAVLTIVSEEL PDVPLFLCLHNLGGTLKCTSPSAVIFRSAVLNAGY RISGTHASPLGLKTDAPMSVWDMIRCWVKNHP VKPQPADQSGTIVLSKEPELQANFTRAVASLSKA QAKKVARFLPNPERYWGPKVFRGGRIQTSKHVSL MGPEANGSVNHEEGEESDAKROKTDHHTSDS

10

20

30

40

【 1 4 3 5】

【表 1 4 - 4 3 6】

	A	B	C	D	E	F
1		分析に用いたベクトルの相関係数				
2	sp_107620_kqmh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈		aa seq WVGFVRSKRWKRRFTTEIGKLFGLPTFSFIMSEGREE VDDVEHSSRNLYVEAVLNMIVQRLLHKEWQCPSPQ KLSVCVVSSYKAQVKEIERLGNRYEKLNNFELK VKSVEEYHEGGEADVJISTVRSNSAGSVGNSLD RQVANALNSSRHCLWILGNKKNLVEADSVWED IVSDAKERCLFNAEEDEDLAKVIIDVKSKLDELE DILLKDSNIFKAARWVVFSENFVKSFAKLPSAF SKLLVNVFVLANGWRPKSKINMVCKSSVGI KQFKVEGRYICTNDIIRVPGYATQILKVDIHLH DVPALVNLKDIYGALSDDFLICCKESCLOGNLE PVRYASFDVARYKDKNGGSTGSLDAVNGGRSFI EKAKVKSLLMKFYFSAVYLSHLLSASDGTVL NVFVSDQEREIILCEKSSFILGRSGTGKTTVL MKLFQEQQHASEGNVHDALMRGSPSNEPG RIGTSLCQLFVTVNPKLCFAVQKHNNLRSAELFL LYRRFRFHGGNSGEAYISSDVNEMDEYEFKDV PDSLEEPSHVYPLVITFNMLLMDRTGLSYED RFPDIRKYSLETTASSRLVAQAVRRKEVYDKFRT SWMPHNLQLTRNLDPSPRVFTEINSCIKGSLHSR ESDDIKLSASYLQLSKRRVSNFSEDEREMIKIF EDYKMKVYRGEYDLADLVNDLHHRFKCKEYEG EFMDFVYDEVDQDLTMKQLGLFKYICSNVEEGFV FAGDQTAQTARGIDFRFEDIRCLFYTEFLHGEERK KDKGLISATSHLSQNFTRTHAGVNVLAQSVIDLHYH FFPNSIDPLNPETSLCCGELPILLDCA TREDAIKIF QGGKSGHDNSISFGAEQVMVRRDDYARDKLANIKI GRQALVITFDCKGLEFEDVLYNFFGSSPLKEQ WRRIEYMNEKLCLLDPETFPSTNAKHDLVLCYEL KQLYVAITRTRORLWICEDRDGFLSPMADYWGK LNLVQVEILDSDSFOAMQVAVASLEDWKKWRGMK MLEVRNYKAAATCCFERARDFYWEKFAASDLKE TANNWRGLRDLKSLLEMLREAAEFESIKNIKKAIE
1065 sp_204240_dptm	-0.852118	#N/A	#N/A	P-loop containing nucleoside triphosphate hydrolases superfamily protein; BEST Arabidopsis thaliana protein match is: P-loop containing nucleoside triphosphate hydrolases superfamily protein (TAIR:AT1G65780.1); Has 6257 Blast hits to 5371 proteins in 1056 species: Archaea - 197; Bacteria - 1787; Metazoa - 1336; Fungi - 1200; Plants - 672; Viruses - 106; Other Eukaryotes - 959 (source: NCBI BLINK).		
1066 sp_204260_ksmi	-0.974276	-0.940711	-0.988603	uncharacterized protein		

10

20

30

40

【 1 4 3 6】

【表 1 4 - 4 3 7】

	A	B	C	D	E	F
1	共発現遺伝子	sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
1067	sp_204310_synth	sp_107620_kqnh	sp_107660_kiqg	sp_107670_ptqx	<p>Peptidase M1 family protein; FUNCTIONS IN: metalloprotease activity, zinc ion binding; INVOLVED IN: proteolysis; LOCATED IN: apoplast, chloroplast, plasma membrane; EXPRESSED IN: 25 plant structures; EXPRESSED DURING: 15 growth stages; CONTAINS InterPro DOMAIN/s: Peptidase M1, membrane alanine aminopeptidase, N-terminal (InterPro:IPR014782), Peptidase M1, membrane alanine aminopeptidase (InterPro:IPR01930), Peptidase M1, alanyl aminopeptidase (InterPro:IPR012779); BEST Arabidopsis thaliana protein match is: aminopeptidase M1 (TAIR:AT4G33090.1); Has 7901 Blast hits to 7867 proteins in 1782 species: Archaea - 112; Bacteria - 3789; Metazoa - 1713; Fungi - 432; Plants - 227; Viruses - 0; Other Eukaryotes - 1628 (source: NCBI BLINK).</p>	<p>MARLVPCGKGPALSKTNFLGLVSSPLQALS RVN CVQRPVKNL TRFRKFLSLETSLRRNCHFSYPALY RSQEWRRLLCSVATPESPKQAE DLNMEAPKEIF LKDYKKPAYFFDTVDLKLFLLEDKTHVFSKISVFP ADEASASSPLILDGSDKLVSKIDGKDLNEGFEFH VDSRHLLTSPTEKFTLEIVTEIYQNNITSLEGL YKSSGNFCTQCEAEGFRKITFFQDRPDIMAKYTV RLEADTKYPVLLSNGNLVGGQDIEGDKHFAIWE DPYKKPCYLFALVAGPLESRDDTFVTRSGRNVTL RIWTPADDLSKTAHAMYSLKASKMKWDEDVFGLE YDLDLFNIVAVPDFNMGAMENKSLNIFNSKLVLAT PETASDADYAAILGVIGHHEYFHNWVTGNRVTCRD WFQLSLKEGLTVFRDQEFSSDMGSRPVKRIADV SKLRSYQYPQDSGPMAPVPRP-HSYIKMDNFYTE VVRMYKTLGSGQFRKGM DLYFKRHDGQAVTC EDFYAAMPDANSADFANFLWYSQAGTPIVKVT SSYDAEKRTFALKFSQEVVPTPGQSTKEPMFIPV AIGLLDSHGKDMPLSSVYHDGKLESVAIDGKPVH TTVLRVTKKEEFVFSDDVAEQVPVSLLRGFSAPI RLES DLSDSDLSFLLAHSDDEFNRWEAGQLARK LMLS LVADFOQNKPLHLNANFVDGLRSILCDNSL DKEFIKAITLPEGEIMDMMAVADPDVAVHVRT FIKKEALQLKTDLLR TVLENKSSSEQYEFNHSNM AFRALKNISLAYLTSLDDPELTELAVNEYKAATNM TDQFAALAAITQNPCKTHDEVLADFYNKWQHFEF LVNWKVFAALQAVSDVPVGNVENIKLLDHPAFDLR NPNKYSLIGFCGSPVNFHAKDGGSGYQFLGEL VVQLDKINPQVASRMVSAFSAFRWVKRYDETFRNLA KAKLEKIMSANGLSENVFEIASKSLAA</p>

10

20

30

40

【表 1 4 - 4 3 8】

I	A	B	C	D	E	F
1		分析に用いたベイトとの相関係数 sp_107620_kqnh (SOAP1)	sp_107660_kiqg (SOAP2)	sp_107670_ptqx (SOAP2-like)	注釈	aa seq
2	共発現遺伝子				RNA-binding (RRM/RBD/RNP motifs) family protein; FUNCTIONS IN: RNA binding, nucleotide binding, nucleic acid binding; INVOLVED IN: biological_process unknown; LOCATED IN: cellular_component unknown; CONTAINS InterPro DOMAIN/s: RNA recognition motif RNP-1 (InterPro:IPR000504), Nucleotide-binding, alpha-beta plait (InterPro:IPR012677); BEST Arabidopsis thaliana protein match is: U2 snRNP auxiliary factor, large subunit, splicing factor (TAIR:AT4G36690.1); Has 62164 Blast hits to 35633 proteins in 1636 species: Archaea - 185; Bacteria - 4712; Metazoa - 29770; Fungi - 5811; Plants - 3284; Viruses - 214; Other Eukaryotes - 18188 (source: NCBI BLINK).	MSSSTRKRSKSAICERTLDITSEESSARTPLPST DEIMLKRKSKKEGVEEAAQDAETSRAHDETVR NPEHRTRDTNGASMKNSRDSQFKAPKRVKEG YRKEHTHTSDSRDYOREASPKVKDKSSRMEHT HIKRDSSQRESSPRKVEKSYRTERTHIKSDSR DYOREASPRKVEKSSRKEHTHVKSDPRDYOREA CPRKADKSSRTEHTHVKSDPQDSQREGSPRK KVEKSSKTEHADIKSDPRNSOLEASPRKVEKSSR TERTHSKSENKERRDVECRSKPKHDNDASNATE SKTDROGGMFSKIEERSRDARYTNEKNKHRDS VAQDRYVERNGGASERENKRRDRSGDDEKHRM GDVAVKRHDGSDSRNSSEISERKENMELSHSRHDE VGHVLDHGDSTHSEDRDRSRHSDVDRSKMSS TRSKRRRSQERTKDKARRSTSSSPKACKTVL VGHVLDHGDSTHSEDRDRSRHSDVDRSKMSS DSSSSHYRRQGGSSGLGGYSPFRKRKSEAAVKT PSPIHSPEKRSAGWDLPHFAADKLNLTSTVPSGT QLVSQTKPSNMIELAAGFSVASSMKGKPTSAIFSN SLSAFKQATIDSVQLTQATRPMMRLYDNIPIGSAT EKDIMECFNSFLSSGSSNYIRGTRPCISOMIHKKEK GOALVEFLTPEDASAALSFDGRSLSGSLKIRRPK DRFEVPSGSSERLKAGAVSNRGDSDSVREVSITAI SDFVEDSSHKIFVGGIAEGLSSEMIVEISAFGLT KAYHFEVNEDLKEPCAFLEYADQSVTLNACASLN GMRLGGRVLTVACAVPNTQSVESRSRISPFYGIPE HVKPLLEMPQTQVLSNVL DQKTMDSISDLEEL VEDVRLCCSRFGSVSNVVKGGKTHASISEAKE VTYDVNTIQPEQIHOHDMNPVEVAADHYDRDG SNVEPPTDVKEFDAIGVDSGDKHELVPATDFEI GAAGTSIFEESEEAADKISQNADEAAVIDDSAD DMARVNDSEENRQEVSIQDSGVSDLNPNNGSM GIEKSGDGTVYHAEDVPELDKNNNDITQIEVPLPT NLGSEEGIKDEDSYDRGNIFEVGCVLVEFRRIPE

10

20

30

40

【 1 4 3 8】

SOAP3及びSOAP4発現のダウンレギュレーションによってメディカゲン酸由来サポニンの蓄積が低下し、グリコシル化中間体(アウグスチン酸、バヨゲニン及びヘデラゲニン)が蓄積した(図25、図26A~26L、図27、図28A~28C)。

【 1 4 3 9】

50

蓄積した経路中間体の構造から、SOAP3 (配列番号52) はC-2ヒドロキシラーゼ、SOAP4 (配列番号54) はC-23オキシダーゼであることが示された(図20A)。CYP2 (CYP72A656; Sp\_\_148230\_dgra) (配列番号72 [遺伝子] 及び配列番号73 [ポリペプチド]) も同定され、これらはSOAP4 に対して高い相同性を示し(アミノ酸レベルではほぼ92%)、サポニン生合成経路において同じ活性を呈するものと考えられる。また、以下がサイレンシングされた植物では、サポニン含有量の低下が認められた: SOAP6 (UGT74BB2; Sp\_\_170930\_hjgq; 配列番号55 [遺伝子] 及び配列番号56 [ポリペプチド])、SOAP7 (UGT79K1; Sp\_\_020820\_yeau; 配列番号57 [遺伝子] 及び配列番号58 [ポリペプチド])、SOAP8 (UGT79L2; Sp\_\_113700\_suxh; 配列番号59 [遺伝子] 及び配列番号60 [ポリペプチド]) 及びSOAP9 (UGT73BS1; Sp\_\_170320\_dmqi; 配列番号61 [遺伝子] 及び配列番号62 [ポリペプチド]) グリコシルトランスフェラーゼ(図29)。

#### 【1440】

同じ植物では、特定の糖、すなわち、デオキシヘキソース(SOAP6 (配列番号56) 及びSOAP7 (配列番号58))、ヘキソース(SOAP8 (配列番号60)) 及びペントース(SOAP9 (配列番号62)) が欠如した経路中間体が蓄積した(図30A~30C)。さらに、5つの共発現アシルトランスフェラーゼ(AT)のうちの1つ(すなわち、SOAP10; Sp\_\_125800\_kzws; 配列番号63 [遺伝子]、配列番号64 [ポリペプチド])のVIGSの結果、アセチル化サポニンに対するデアセチルサポニンの比が増加した(20から80%) (図31A~31C)。次いで、大腸菌において発現した組換えSOAP10のアシルトランスフェラーゼ活性は、2つのデオキシヘキソースがC-28位に結合したヨッソシドIV及びその他のヨッソシド(例えば、ヨッソシドVI及びヨッソシドXII)をアセチル化することから、*in vitro*で確認した(図14A~14C)。

#### 【1441】

実施例18: SOAP遺伝子の検証

#### 【1442】

目的: 9つの遺伝子全て[すなわち、-アミリン合成酵素(bAS; SOAP1; 配列番号45); シトクロムP450(CYP450s; SOAP2、3、4; それぞれ配列番号46、51、53); UDP-グリコシルトランスフェラーゼ(UGTs; SOAP6、7、8、9; それぞれ配列番号55、57、59、61) 及びアシルトランスフェラーゼ(AT; SOAP10; 配列番号63)]の機能を検証すること。

#### 【1443】

方法: 上記の材料及び方法を参照されたい。

#### 【1444】

結果: 対照とする9つの遺伝子(SOAP1~4及び6~10)は、ベンサミアナタバコの葉において一過性発現した。最初の4種の生合成遺伝子(SOAP1~SOAP4)を発現させたところ、メディカゲン酸が産生した(図20A)。奇妙なことに、9つの遺伝子全てが共発現しても、ハウレンソウの葉で以前に検出されたいずれのサポニンも産生されなかった。それにもかかわらず、ハウレンソウの葉には通常存在しないメディカゲン酸及びグリコシル化誘導体[MA+hex; MA+2xhex、MA+3xhex]が検出された(図32A~32D)。

#### 【1445】

要約: メディカゲン酸の再構成が実現された一方で、9つの遺伝子SOAP1~4及び6~10が発現しても、いずれのサポニンも産生されなかった。

#### 【1446】

実施例19: サポニン生合成経路の構成要素としてのセルロース合成酵素様Gハウレンソウホモログの解析

#### 【1447】

10

20

30

40

50

目的：サポニンの生合成経路の再構成に必要な、欠損している1つまたは複数の遺伝子を同定すること。

【1448】

方法：上記の材料及び方法を参照されたい。

【1449】

結論：この時点で、サポニン生合成経路を再構築するための、少なくとも1つの酵素が存在していないことが分かった。欠損している1つまたは複数の酵素を同定する試みとして、共発現解析によって得られた遺伝子のリストを改訂した。糖代謝に関連すると思われる共発現セットにおける全てのベイトで共発現された唯一の遺伝子は、セルロース合成酵素様Gハウレンソウホモログ(SoCSLG) (「T. Richmond, Higher plant cellulose synthases, Genome Biology. 1, 3001.1-3001.6 (2000).」)であった。この機能予測に基づくと、サポニン生合成へのその寄与の可能性は低いものと考えられたが、その経路内の機能について検討した。

10

【1450】

注目すべきことに、SoCSLGをサイレンシングしたところ、ハウレンソウの葉においてメディカゲン酸の高い蓄積が認められ(図33A~33F、図34A及び34B)、トリテルペノイド系サポニン生合成における役割が示唆された。次に、ベンサミアタバコの葉において、SoCSLG(配列番号65または93)をSOAP1~4(それぞれ配列番号45、配列番号46、配列番号51、配列番号53)と共に一過性発現した。5つの酵素全ての活性を合わせると、ハウレンソウトリテルペノイド系サポニン生合成経路における最初のグリコシル化中間体であるメディカゲン酸3-O-グルクロニド(MA-3-GlcA)が産生した(図20A~20C及び図32A~33F、図35A~35C)。

20

【1451】

さらに、CSLG(SOAP5、Sp\_\_076690\_\_ejcm(配列番号65または93[遺伝子]、配列番号66[ポリペプチド])とbAS(SOAP1;配列番号45)及びCYP450(SOAP2[配列番号46])、SOAP3[配列番号51]及びSOAP4[配列番号53])との組み合わせ発現により、SOAP5[配列番号66]は、メディカゲン酸の他に、オレアノール酸、アウグスチン酸、ヘデラゲニン、ギブソゲニン及びギブソゲン酸などの他のトリテルペノイド系アグリコンをグルクロン酸抱合できることが示されたが、これらはハウレンソウでは有意な量が検出されることはなかった(図36A~36F)。SOAP5(SoCSLG)は、ハウレンソウサポニングルクロン酸トランスフェラーゼとして同定された。逆に、SOAP5(配列番号65または93)の発現を阻害したところ、いくつかの実施形態において、これらに限定されるものではないが、メディカゲン酸、オレアノール酸、アウグスチン酸、ヘデラゲニン、ギブソゲニン、及びギブソゲン酸などのトリテルペノイド系アグリコンのグルクロン酸抱合が低下または阻害される結果となった。

30

【1452】

SOAP5がハウレンソウサポニン経路において確かに欠損している構成要素であるか否かを検討するため、ベンサミアタバコの葉において、10のSOAP遺伝子全て(表16を参照)の組み合わせを一過性発現した。代謝物の分析によると、最も豊富なハウレンソウサポニンヨソシドVの存在が示され、自然の経路において必要な全てのステップ及び酵素が同定されたことが示された(図33A~33F)。

40

【1453】

要約：ハウレンソウサポニンヨソシドV(化合物11)の自然の経路において必要な全てのステップの酵素が同定され、異種植物系で機能的に活性であることが示された。

【1454】

実施例20：トリテルペノイド生合成経路の特徴

【1455】

目的：ハウレンソウサポニン代謝経路のさらなる詳細を解明すること。

50

## 【 1 4 5 6 】

方法：上記の材料及び方法を参照されたい。

## 【 1 4 5 7 】

結果：S O A P 5（配列番号 6 6）をハウレンソウサポニングルクロン酸トランスフェラーゼと同定することによって、経路をさらに解明することができた。S O A P 6（配列番号 5 6）は、フコースをアグリコンの C - 2 8 位のカルボキシル基に結合させる M A - 3 - G l c A に直接作用し、ヨッソシド I（化合物 7）を産生することが分かった。実際に、これによって、S O A P 6 は、トリテルペノイド代謝に關与する、特性が解明された初のフコシルトランスフェラーゼとなっている（「T. Louveau et al., Analysis of two new arabinosyltransferases belonging to the carbohydrate-active enzyme (CAZY) glycosyl transferase family 1 provides insights into disease resistance and sugar donor specificity. The Plant Cell. 30, 3038-3057 (2018).」）（図 3 7 A ~ 3 7 F、図 3 8 A ~ 3 8 E）。

10

## 【 1 4 5 8 】

S O A P 7（配列番号 5 7）の発現により、他の 6 種の S O A P 1 ~ 6 遺伝子（表 1 6 参照）と共に発現したときのみ、ヨッソシド I 生成物が産生した。L C - M S / M S 分析に基づくと、S O A P 7（配列番号 5 8）は、ラムノースをヨッソシド I のフトースに転移させるグリコシルトランスフェラーゼである（図 3 7 A ~ 3 7 F、図 3 8 A ~ 3 8 E）。ヨッソシド I は、C - 2 8 位に結合したフコース及びラムノースを有するアグリコンのみを基質として受け入れるグルコシルトランスフェラーゼである S O A P 8（配列番号 6 0）によって、ヨッソシド I にさらに変換される（図 2 0 A、図 3 7 A ~ 3 7 F、図 3 8 A ~ 3 8 E）。S O A P 9（配列番号 6 1 [ 遺伝子 ]、配列番号 6 2 [ ポリペプチド ]）の発現により、S O A P 9 は、M A に C - 3 で結合したグルクロン酸にペントースを結合させるキシロシルトランスフェラーゼであることが示された。S O A P 9 の場合、その活性にとって、C - 2 8 のトリテルペノイド骨格を修飾する糖部分は重要ではない（図 3 9 A ~ 3 9 C）。in planta 試験及び in vitro 試験により、S O A P 1 0 は、アセチル基をフコース部分に転移させる、ハウレンソウにおけるサポニン合成の最終ステップに關与していることが示された（図 3 1 A ~ 3 1 F）。比較に用いたアセチルトランスフェラーゼのヌクレオチド配列及びアミノ酸配列を、下記の表 1 5 に示す。

20

30

## 【 1 4 5 9 】

表 1 5 : アセチルトランスフェラーゼ酵素のヌクレオチド配列及びアミノ酸配列

## 【 1 4 6 0 】

【表 15 - 1】

名称	SEQ ID NO:	AA SEQ / NT SEQ
sp_074 630_yg ho_Spo 04549_ AT1 遺伝子	109	<p>ATGGCCAAATCTGAGCAAGAAACAATGGCCAAATCTGAGCAAGAAACATCCAT  CAAACCTGTTTCAGAATGCTTTGTAACCAAAAATATGAGATTAAATCCGCTAA  GCAACCTTACCACCTTAGGTCCCATGGATCTAGTTATGTTAACTATCGATCCTAT  ACAAAAAGGTCTTGTCTTTACAATAAAGAATCCCCACTCTTTTTGTCATCCGA  ATCCCATGATAATATTGAAATTATTCGAACCAAAGTTGTGTCACGTTTATTAGA  AAAGCTTAAACACTCACTTTCTATAGCTCTAGTCCACTTCTACCCGTTAGCAGG  TCGTTTCACTACACAAAAACAACCCGAGCATAACACGAGCTTGGTCTTTATTGA  TTGCAACAAAGGTCCCAGGAGCGGTTTCATCCACGCTACTTCCCCTGACTTTAC  TATCTCCGATATACTTTACCCGTTGATGTTTCCATCGTTCATTCTTTGATC  TCGGTGAAAAGCATGTAAACTACGATTGTCATACTAAGGCGTTGCTATCGATCC  AGGTAACAGAACTTTTAGATGGGGTGTATTGGGGTTAGCATGAGTCATAGTG  TGGTTGATGGTACCTCTTTTATTCAATTTGTCAATACCTTGTCTGAAATTTTTAA  ATCTGATGATTTTACCCTATTTACGTCGCCCCAATACTTAATTATAGGCCCTGT  GATATTCCGATCCTTAAATTTCCGTTTCTTGATGTGGAGGGGTTATATGTCGTG  CGTATAACCCTGGGCCGTTAAGGGAAAGAATCTCCACTTTTCACTAAATTCGA  TGCTGAGACTCAAGGCCATGGCTAACCAAGAATGTGGTACCCAAAATGTTTTA  TCATCTTTCCAAGCTTTGACTGCGGTTGTATGGAGGTCCATCACCCGAGTTCGG  AACTTACCAAAGGATGAGCAAACACGCTGTTTTATGGCTATGGGTTCTCGAAT  AGGCTCAACCCGCTCTTTCCGATGACTATTTTGGGAATTTTATGATTAGTACC  AAATTTGCTTGCAAGGCAGAGGAATTATTGGGTAACAGTTTAGGTTGGGTAGC  AATGAATTTACGTAAAATCATTATGTCCACTGACGAGAAATCGATACTTGCTAC  GTACAAAGCATTGGCTGATTCCCCAATAGTGATTCCGCGTGAAACGATCCCCG  GTCCTCATGGGATGACCAGAGTAATAATTGGAGGATCTTCAAGGTTGATATG  TATGGCCCTGAATTTGGATTGGGTCGAGCTTTGGCCGCTCGCATGGGTTATGGG  AATAAGGATGATGGGAAAATAACAGCAAATCCTGGGTGTGAAGGAGGTGGAA  GTGTTGATTTGGAAATTTGCCTTAGGCCTCATATTATGGCCTCTCTGAAGTTGA  TCAAGAGTTTATGGGTTTTGTGTCCTAG</p>
sp_123 780_pg iy_Spo 21561_ AT2 遺伝子	110	<p>ATGACTCCAAATCTGCAAATAGTAACCAACGGAGGCAAACCGGAAAATGATG  AAGCAGAACCCGATCACCTACCGGACAATACTTCAACAGCAAAGTGTTGTCT  GTTTGTGTCTTTGCCATTTAGAAATTGATGTTCTATAGATGACTCGTGTGTAA  TTCCACAACCCGATGATGTTCTTCCATCCCATGAACCCAGATTTTTCATATCAT  GATATCTGACAATAAAGATGTAAAACAATGGAAAAGAGTGGAAGTGAACCTTC  AAGATCATGTTGTGTCGTCCTAGCGTCCCAGATGGCTTATCGGTTGAATCATACG  ACAAGTACTTTGATGAATATCTGACAAAATAACAGTGGATCCATTACCACAG  GATAGGCCTTTATGGGAACTTCATGTTATAAAAATACCCAACAAGCAAAGCAGC  GGGTCATTTTCATCTGGAAGCTTACCATGCACCTGGTACCGGCTACACTCTAAT  GGGAGTACTTCTGTCCGGCGTGAACAGAGCAGATGATCCTTCCCTTCCGTTAAC  TTCCCTTCAACACGATCAAGCTCACTAGTTACAAACAACAAGATGAATATTAT  CAGCTGGGTGCCAAGAACTTTTTAGCAATCTACAACGGTGTTTATAATTTTGG  ATGGAGTTTTCTAAAAGCACTTGCAAGGCAGATGATAAGACACCTATCAGAT  CCGGAAATGAAGGTCTGGGTTTCCACCCAATGAAGATCTCGACAATAGAACTA  TCCCTAGACCAAATCAAATTTATCAAACAAAACCTCGGCGCAACGGTAAATGA  CATTCTTGACGGCATAATTTTCCCTCGGGTTTCGAAAATACATGCAAGCAACTGA  TACAGAATCTGGAACTCAGAATCAACGGCATTGGTGCTGTTAACACTAGGA  ACATTGGAGGTTATATGACCGCTGAGCAAATGAAGAAAGCACAAATGAAAAT  ATGGGGGAACCAATTTGCATTTTTGCATATAGCAATACCTCAATTAATCAATGA  CAAATGCTCGAACCCCTTGACTATGTCTATGAAGCACGAAAACAGATCTCTA  GGTTCAAAGCTCACCATCAGTCTATCTAACAGCTCAGTGCCTAGAGCTGCTAG  GAAATGCAAAGGACCTGAGGCAGCAGCTGAATTTATCCAGAGTACAACGAATA  AAGCAAGCATATTA</p>
sp_149	111	ATGTTAGAGCTAGCAGAAGACGAGGTGAAGAATCTTCAAGGTATGGGCAAT

10

20

30

40

【表 15 - 2】

<p>180_nw my_Sp o15788 _AT3 遺伝子</p>		<p>AGTTTTTGCATCTTTAAGCTATTGTTATTACATAGGCAAGCTAATTAATCCAAA AGGTTATACAAGATTAGTAGCAATAATCCCAATTATTACTCTCTTTTTAGCACT TCCTTAAAATCTCACATCTTTTCATCTTGGTGGTATGACTTGTTCTTTATTGCTT GGCTTGCTAATTTCAAACCTCTTGCTTTTTGCTTTTGATAAGGGCCCACTTTGCGC TAATTCCTCAATCTCATTTCGCCAAATTTCTTGCACTTTCTTGCTTACCCATCAA ATCCAACACCCACCTCATAAAAAGTCATTAAAATCACACCCATCTATTTATAAT TACATCATTAAAGGGATACTTTTATGTCTAATAATTAATAATCTATGATTATGGT GATTACATTCATCCAAAAATCATATGGCTAATCTTTTTCTTCCACTCCTATTTTA CCATAGAGTTAGTCTTTGCATTCCCTAGCAACATCGACTAATATTTTGTAGGGC TCGAACTGGAGCCACAGTTCAATGAACCCCTAATATCAACCTCATTGCAAGACT TTTGGGGTAAGAGATGGAATATCATGGTGACAAGGATACTTAGGCCTACGGTG TACCTTCCCACACTAGAGTACTCCACTAAGGTCGTTGGACGCACGTGGGCCAC ACTTCCGGCGGTGATGCCACGTTCTTTGTGTCAGCCATTATGCACGAGCTCAT CTTCTACTACTTGGGGCGCAACTGGCCACATTCGAGGTGACGTGGTCTTTCT CCTGCATGGATTATGTCTTTGTGTTGAGATTGTCGCTAAGAAGTTAGTTGGTGG GAAATGGAGGATCCCACGGTGGATTTCGGGCCCTGCCACGGTGTGTTTGTGGT GGGTAAGTGGGTTTGGCTGTTCTTGGCCCGTGTGTAAGGCTGGGTTGGATAC TAGACCGTTTCAAGAGTTTGGCCCGTTCGCAAGTTTGTAAAGGATTGTAAGGC AGCTCTCACATTTTG</p>	<p>10</p>
<p>sp_198 340_fo cw_Spo 13090_ _AT4 遺伝子</p>	<p>112</p>	<p>ATGGCTCCTCCTTCTTCTTCAACTACGGGTTCTGGTAATGGTTCTAGTTTTG CAGTCAATATAATGGCGTCTTCTACATTTCCCAACAACCTTCCACCACAA ATTCACATTCTATCCCTCTCACTTTCTTTGACATTCCTTGGCTTCAATATCCTCC GCTCCAACCTCTCTTCTTCTTCAACTTCCATCTACACCCCAATCTTCTTCTTCT CTTCTTCTTCTTCTTCTTCTTCTTTCGACCACAACCTGTACTTGGAGTTAGCTCC ACCATCCTCCCTAGGCTCAAACACTCCCTCGCTTCTGCCTTGCAATATTACTTTC CCTTTCTGGAAAACCTCACCCTACTACCCATACTATCCCGAATAACCTAGTTT TCTCGACAGACTCATCAGATTCTGTTGAGTTGACTGTTTCTCTGTGTGATGCTGA TTTTAATGGTCTATGCAGCTTTCTACCCAGGCTACTCATCTTCCAACAACATTG GTTCCCTCCTTGCCAAATATTGAATCCTTCCAACCTCACTACTACCTCCACCTT TATTAGCTATTCAGATCATTCTTCCACCTCTTCTCCTGGTTCTCTATTGG CTTTGCTTCTCATCCTGTGCTTTCTGATCAGAGGACCTTCAGTAACTTCCCTTAC TCTTGGGCCTCTTCTCCAAGTTTGATAATCTAAACATTTCACTTGGCCCTTCT TCCCTGTCTCTGACAGGCTGTGATTCTCGACCCTGATAGACTTGGAGCCCTTCT GTTGGAGCAGTGGTTGGGATTGGAGTCCAAACCAACCATGTCAACAAAGATGA AGCTACGTCTCCTCCTGCTTATGTCGGTGGCTCGCTCCGGTCCACATTGTCAT GGGCCATCTGATATTGCTAATGCTACACAATGGTTACAAACCCAGTGTGAGA AGCTCAACAGATCATATCCTGTTCTTGTACCCCTACGTCGTCCTTGTGCCTT TATATGGACCTGTTTTCTGAGAGCCCGAGTCCAGAACAGTGTGTTACTAAAGC CAAAGCAAAAGGCACCATGTAATTTGGATTTATTGCTGGTGGTATTACCCGTTT ACCTATCGGGTACCTGCTAAGTATCTTGGCAACTGTGTGGGTTGGACGGGC AGCAGCGCAGAGGGAGGACTACTGAAGGAAGGTGAGGGATGTTGGCAGCT GCTGATGCAATTGGGCTAACCATTAAGGTTGGATAAAGATGTTTTAGGAGG AGCTGAGAAATGGATATATGAATGGCAGACATTAAATGGAATCCGAAGATCATA TTCATGTGGTTGGGTCGCCAAGGTGAACCTTTATGAGACGGATTTTTGGTGGG GGAAACCGAAGAAGATAGAGGAAATTTCAACTGATGTTACCAGAGCCATCTCT CTTACACAGAGCAGGGACATGAAAAGGGGAATTGAAATTGGCCTCACTTTACC AAACTCCATTATGGATGACTTCTCCTCTATCTTCACTCAAGGCCTCCTGTTTTT CAAATTAG</p>	<p>20 30</p>
<p>sp_074 630_yg ho_Spo 04549_ _AT1 ホリヘチ</p>	<p>113</p>	<p>MAKSEQETMAKSEQETSIKLVSECFVKPKYEIKSAKQPYHLGPMDLVMLTIDPIQK GLVFTIKNSPLFLSSESHDNIEIIRTKVVSRLLEKLKHSLSIALVHFYPLAGRFTTQKQ PEHNTSLVFIDCNKGGPARGFIHATSLDFTISDILSPVDVSIHVSFFDLGEKHNVDCH TKALLSIQVTELLDGVFIGFSMSHSVVDGTSFIHFVNTLSEIFKSDDFTTISRAPILNY RPCDIPLKFPFLDVEGFICRAYNPGPLRERIFHFSLNSMLRLKAMANQECGTQNVL SSFQALTAVVWRSITRVRNLPKDEQTTFCMAMGSRTRLNPPLSDDYFGNFMISTKF ACKAEELLGNSLWVAMNLRKIIMSTDEKSILATYKALADSPVIPRETIPGPHGMT RVIIGSSRFDMYGPEFGLGRALAARMGYGNKDDGKITANPGCEGGGSVDLEICLR PHIMASLEVDQEFMGFVS</p>	<p>40</p>
<p>sp_123 780_pg iy_Spo 21561_</p>	<p>114</p>	<p>MTPNLQIVTNGGKPEDEAEVPSPTGQYFNKVLVSVLAILEIDVPIDDSVIPQLR DVFLPMNPRFSSIMISDNKDVKQWKRVEVNLQDHVVVPSVPDGLSVESYDKYFDE YLTKITVDPLPQDRPLWELHVIKYPTSKAAGHFIWKLHHALGDGYTLMGVLLSGV NRADDPSLPLTFPSTRSSSLVTNNKMNIISWVPRTFSAIYNGVYNFGWSFLKSTCKA</p>	

【表 15 - 3】

AT2 ホ <sup>リ</sup> ハ <sup>フ</sup> チ <sup>ド</sup>		DDKTPIRSGNEGLGFHPMKISTIELSLDQIKFIKTKLGATVNDILAGIIFLGVRKYM QATDTESGNESESTALVLFNTRNIGGYMTAEQMKAQMKIWNQFAFLHIAIPQLIND KCSNPLDYVYEARQKISRFSKSSPSVYLTAQCLELLGNAKDLRQQLNLSRVQRIKQAY
sp_149 180_nw my_Sp o15788 _AT3 ホ <sup>リ</sup> ハ <sup>フ</sup> チ <sup>ド</sup>	115	MLELAEDEVKNFFKVVWAIVFASLSYCYIYGKLNPKGYTRLVAIPIITLFLALPLNL TSFHLGGMTCFFIAWLANFKLLFAFDKGPLCANSSISFAKFLALSCLPIKIQHPPHK KSLKSHPSIYNYIIGILLCLIKIYDYGDIHPKIIWLIFFFHSYFTIELVFAFLATSTNI LLGLELEPQFNEPLISTSLQDFWGWKRWIMVTRILRPTVYLPYLEYSTKVVGRTWA TLPVAVMSTFFVSAIMHELIFYYLGRNWPTFEVTWFFLLHGLCLCVEIVAKKLVGGK WRIPRWISGPATVLFVVGTFWFLPPLKAGLDTRPFQEFAAVAKFVRSLSKAALTF
sp_198 340_fo cw_Spo 13090_ AT4 ホ <sup>リ</sup> ハ <sup>フ</sup> チ <sup>ド</sup>	116	MAPPSSSTTGSGNGSSFAVNIMASFYISPPQSTTNSHSIPLTFFDIPWLQYPPLQPL FFFQLPSTPQSSSSSSSSSSSFDHNLYLEFSSTILPRLKHSLSALQYYFPFSGKLT THTIPNNLVFSTDSSDSVELTVSLCDADFNGLCDFLPRSTHLFQQLVPSLPNIESSNLT TFPAPLLAIQITFFPTSSPGFSIGFASHPVLSQRTFSNFLYSWASFSKFDNLNLSLAPS FPVSDRSVILDPDRLEPLLEQWLGLESKPTMSTKMKLRPPPAYVRGSLRSTFVMG PSDIANATQWLQTQCEKLNRSYPVLLSPYVVTCAFIWTFCFLRARVQNSAVTKAKA KGTMYFGFIAGGITRLPYRVPKYLGNVGVFGRAAAQREELLKEGEGMLAAADAI GLTIKKLDKVDLGGAEKWYEWQTLMESEDHIHVVGSPKVNLYETDFWVGKPKK IEEISTDVTRAI SLTQSRDMKRGIEIGLTL PNSIMDDFSSIFTQGLLVFQN

10

## 【 1 4 6 3 】

本発明者らの知る限り、SOAP10（配列番号64）は、アシルトランスフェラーゼのうちの、ベンジルアルコールアセチル - 、アントシアニン - O - ヒドロキシ - シンナモイル - 、アントラニル酸 - N - ヒドロキシ - シンナモイル / ベンゾイル - 、デアセチルピンドリンアセチルトランスフェラーゼ（BAHD）スーパーファミリーの中で、トリテルペノイド系サポニン生合成への関与が報告された初のメンバーである。

20

## 【 1 4 6 4 】

以下の表16は、ハウレンソウにおいて化合物11（ヨッソシドV）を産生する生合成トリテルペノイド系サポニン生合成経路を構成する、10種の酵素のアミノ酸配列及びヌクレオチド配列、並びにこれらの酵素をコードする遺伝子を示す。

## 【 1 4 6 5 】

表16：ハウレンソウにおけるトリテルペノイド生合成経路酵素のヌクレオチド配列及びアミノ酸配列

30

## 【 1 4 6 6 】

【表 16 - 1】

名称	酵素活性	SEQ ID NO:	AA SEQ / NT SEQ
SOAP1	サポニンβ- アミリン 合成酵素	45	atgtggaggttgaaggttggagaaggggctaataccatactatataagcactaataactttgttggcgtcaaaacttgggagttgatcctaactatggcaccctgaggatattcaagaggtcgaagatgctcgcgcgattttacaataatcggttaaagtgaagccttgaacgatctctatggcgtttcagttcttaagagagaaaaactcaagcaaccatacctcaagtgaaggtgggtgacggggaggagatcacatatgagaccgctcgacgacattaaagagagcggtaaatatttcacagccttgacgtctgaacatggccattggcggctgaaattgctggccctcagttcttctctctcttggattttgcttafacattacaggagatcctaactctgtttcggaccagaacatcgtagagaaattcttcgcagcatctatacaccagaacgaagatggaggtggggattacatattgaaggacatagcaccatgttctgta ccgactgaattacatatttacgaatgcttggataggacctgatgaaggtgatgacaacgcgtgccctagagcgcgtaaatggattctcgaccatggtagcgttacacatcccttcttgggtaaaa cttggtatctatactgggctgtcgttggcttggagtaaccaatgccacctgagttctggatccttctacttttctccctatgcattccagcaaaaatgtggtgctactgtcgaatgggtgtatataccaatg tcatactgtatgggaagagattcgtaggtccaatcacacctctcattaaacaacttagggaagaa ctctacaacgaacccttgaacaaatagttggaagaaaatgcgacattgtgtgcaccggaggat ctctactatctcatcattgattcaagacttgatgtgggacgctctttaccttttacggaacctctcc tgaccgttggcctttcaacaagttgatacgaagaaagcattagaggttacaatggaacacata cattatgaagatgagaacagtcgttacataacaattggatgtgctgagaaggtttatgtatgttagc

10

20

【 1 4 6 7 】

【表 16 - 2】

			ctgtgggtggaagaccctaaaggggatcattacaagaaacatcttgaagagtacaagattaca tttgattgctgaagatggattgaaatgcagagtttggagtcacaatgggattgtgggtttca gtacaggcattatagcttctaactttagctcgcgaaattggacctgctttaaagaaaggccattf cttcattaaggagtcacagggtgaaggacaatccatccggcgacttcaaaagctatgcatcgccata tctcaaaaggatcgtggactttctccgaccaagatcatggttggcaagctccgattgactgccg aaggccttaagtgtgtctaatctatcaacaatgccccggaaattgttgagaaaagatggacc ctgaacgccttatgattctgcaatgcttctctacagagtafaaaaggaggctagctgccc tgggaaccagcaggggctcaagaatggttggaggtcctaaaccaacagaattcttgaagaca ttgtattgaacatgagtagtaggttacggctcagcaattcaagcttaataatgtcaagaag ttataccaggacacagaaaaagagattgaaaattttagtaaaagcagtcgaagtacctgaa aacaccaatcttagtggagatggtatgaaaattggggattgtttcatataggaacatggt ttgactaggaaggctagcagcagggtgggaagacatactataattgtctgctgttaggaagggt gttgattttgcttactacaaaaaggagatggtggtggggtaaaagtataattctgtcccaat aaggaatttggcaatagagggaagccaatttggctcagactggttgggctttagtggtgctc cttcactggacaggcggagagggatccaactcctctcagcagcaaaagctttgattaat tcacaactcgaatggcgatttccctcaacaggaaataacaggagtcttcatgaagaattgcatg ttacattatccgatgacagaagcattatccactgtggcaattgcagaatcacagaagcgtgtt cattaccttataacttactga	10
SOAP2	シトクロム P450	46	atggaactcttctttagtgggctagctcttctctcctatctctagcctcttcttcttctata accaccatagaaccgggtacaagctacccccgggcaagatgggggtgccgggtgggg cagtcatttgaatttttcaaacgggtggaagggtaccggaaaagtcatattttagagactg aacaagtacacccaagccaagtgtcaagactccatcgtaggagaaaaggttgcggtttatgt ggcgcggcggtacaagttcttactcaaacgagaacaagtagtacaagcttggtgcccta gctctgtgataagatcttcttcttctacccaaactcctccaagaaggtaagaagatgcg gaaactcctcctacttctcaagcccaggcttatacaggtacatacccatcatgatagcatt gccatccggcacatggagtcgggtgggaggaagacaagtagaagcttcccttggct aagaattacacttctggctggctgcccactctttaaagcgtcagggaccggctcatgagcc aagttctcgaaccattcaacgacatagccgcagggatcatctcagccaatcgacctccccg gaacaccctcaaccaggatcaagtcgttaacgtcgtaaaggaaagattgaggccatcat aaagcagaggaaacttacttagcagatggcaaggcttccctacacaagattctgtctcat gttgtgacttactgaagatggcaagttatgagtaaatggatattgctgataagattctgggac ttctattgtggacatgactgctagcttcttcttcttggtttaagttcttctgagcttctc acatataaggtgtctacaagaagcaaatggagatgcaaatcaaaaaagcaggagaact tctaaattgggaggacatacaaaaaatgaaatactcatggaatgagcttgaagttatgctttg gctcctcactcaaggtggttcagggaagccccttctgatttcatgataacggattccaatecc caagggtggaagttataggatgcaaatcaacacatagaaaccggatgctccggagc ccaagacgtcgacctcaggttcgacggtacgggaccagcaccatacacatagctccctt cggaggaggaccgagaatgtcccgggcaaggatgcaaggctagagatattaggttcat gcacaacgttgaaggtttaaagggaataatgcttctgatgagaaggttattgcaatcc catgcctatcccagaacatggccttctgccccttctcctcactcgaactgtagctgcttaa	20
SOAP2 -Like		47	atggagttcttcttctgtgtggtctagcttcttctctcctatctctagcctcttcttcttctata ccaccatacaaccgggttaccgctaccgcccgggagatgggggtggcgggtggtgggcg actegtttgaatttttcaaacgggtggaacggttaccggaaaatttcatctttagactcaac aaatacaccccaagccaagtgtcaagacttctcctaaagaaaaggttgtgtttatgagaa gggttag	30
SOAP1	サポニンβ- アミリン 合成酵素	48	MELFFMCLVLFSLSLASFLLFYNNHRRTRGYKLPPGKMGW PVVGESFEFFQTGWKGYPEKFIFDRLNKYTPSQVFKTSIVGEK VAVLCGAAGNKFLYSNENKLVQAWWPSSVDKIFPSSTQTSS KEEAKKMRKLLPNFLKPEALHRYIPIMDSIAIRHMESGWEGK DKVEVFPLAKNYTFWLACRLFLSVEDPAHVAKFSEPFNDIAA GIISMPIDLPGTFNRIKSSNVVRKELRAIKQRKLDLADGKA SPTQDILSHMLLTCTEDGKFMSEMDIADKILGLLIGGHTAS ASCTFVVKFLAELPHIYEGVYKEQMEIANSKKAGELLNWEDI QKMKYSWNVACEVMRLAPPLQGGFREALSDFMYNGFQIPK GWKLYWSANSTHMNPECFPEPKTFDPSRFDGTGPAPYTYVP FGGGPRMCPGKEYARLEILVFMHNVVVRKFKWEKMLPDEKVI	40

【表 16 - 3】

			VNPMPIPEHGLPVRLLFPHPTVAA	
SOAP2	シトクロム P450	49	MWRLKVGEGANDPYLYSTNNFVGRQTWEFDPNYGTPEDIQ EVEDARRDFYNNRFKVKPCNDLLWRFQFLREKNFKQTIQV KVGDGEEITYETASTTLKRAVNIFTALQSEHGHWP AEIAGPQ FFLPPLVFCLYITGDLNSVFGPEHRREILRSIYYHQ NEDGGWG LHIEGHSTMFCTALNYICLRMLGIGPDEGDDNACPRARKWIL DHGSVTHIPSWGKTWLSILGLFDWSGSNMPPEFWILPTFLP MHPAKMW CYCRMVYMPMSYLYGKRFVGPITPLIKQLREEL YNEPFEQISWKKMRHLCAPELDLYPHPLIQDLMWDALYLFT EPLLTRWPFNKLIRKKALEVTMEHIHYEDENSR YITIGCVEKV LCMLACWVEDPKGDHYK KHLARVQDYIWAEDGLKMQSFG SQQWDCGFSVQALLASNLSLDEIGPALKKGGHFFIKESQVKDN PSGDFKAMHRHISKGSWTFSDQDHGWQVSDCTA EGLKCCLI LSTMPPEIVGEKMDPERLYDSVNVLLSLQSNKGG LAAWEP GAQEWLEVLNPTEFFEDIVIEHEYVECTASAIQALIMFKKLYP GHRKKEIENFV VNAV KYLENTQYPSGGWYGNW GICFIYGT WFALGGLAAGGKTYNCAA VRKGVDFLLTTQKEDGGWGE SYISCPNKEFVPIEGKSNLVQTGWALMGLLHAGQAERDPTPL HRAAKLLINSQLENGDFPQQEITGVFMKNCMLHYPMYRSIYP LWAI AEYRKRVS LPSINST	10
SOAP2 -Like		50	MEFFFLCGLV FYFSISLASFFLFYNNHHTTRVYPLPAGEMGWP VVGDSFEFFQTGWNGYPENFIFDRLNKYTPSQVFKTFILREK VVFYLRRL	20
SOAP3	シトクロム P450 (C-2 ヒドロキシラーゼ)	51	atgata gaaatcgggtatattgaaaatgggtaattgttagtgattgtagatggggtatggaagatt gtgaattgggttggttacacaaaaaggcttgagaagtttcaagaaaacaaggtttagatgga aattcatacagattttgtgggtgatctcaaatatgtctaaaatgcgtaaaagactagacaaa aacctattcctttactcatgactctttcctcgtatcttgccttccacaatcaccatttcaataaacg gggaaagctcttttcatggatggggcctataaccagttgtgaatgtgcagaacaagagcaagtaa agggtgtgtcactaggaataaaggattcagaaggccaaattaaccactgtgtcattgctgtt ccctggactgtgagcgtgaagggtataaatgggtcaagcacaggaagctcatcaaccggct tttcatatggaaaagcttaagcttatgcatccagcattggcggcagtggtttgatatggtgaaca gtgggagaagatagatctaaaacagggttctctgaagtggatgtggcctgtttgtccagct gactgcagatgctatctctgctgcttttggcagtagctatgatgaaggagaagatattgtg ttggttctgaacaaactgaaatcaccctacgccttctgcaatcagttatcctgtgagatgat gtgccaacaaagaccacaggatgaaaacagtaactctgaaatacaaaaattaitaacgg ggataatcgttaagagaaaggcaatggaggccggcgaagctgccaagatgatttggg ggatattgtggagtccaactacaagataactgaaaatgttctcagtaataagaaaaactaagca tgacttccaggaattgattgatgagtgcaaacgttctacttagcagggcaagagtcgacctgg tgtgctagcaggacaatgattctgttggaaagcacacagagtgccaagcagcagcagaga agaagtagtgcaacgtttgtaaaaacgaacctgatttgaaggcttaaacattgaagatagtg acaatgatactgaatgaggtgtgagggtgtaccctccagtggtacaatcaccgtaagaattca accacgactacagcttgaaatctgacagccctcgtggtgctatggttacgatgacgatate gtattcaagagatcctaaaatattgggggtgatgatgcaaaagatttaaccacagattttcag aaggggttgcaaaaggctacaaaggggaatattgacttctccgttgggtggggccgcgaattt gcatcggacagaacttgcacttattgaagtaaaaatggcagtgccatggtttacaacgctttct ttgagctateaccgtcttatactcatgctcctaccactatcctcactcttaacccaacaagggtct catctcattatacataagctcagggaactaa	30 40
SOAP3	シトクロム P450 (C-23 オキシダーゼ)	52	MIEIGYIVKWVICLVIVRWVWKIVNWVWFTPKRLEKFLRKQ GLDGNSYRFLLDLKDMSKMRKEARQKPIPFTHDFFHRILPF HNHHFNKYGESFFSWMGPIPVVNVAEQEQVKGVFTRIKEFQ KAKLNPLVALLVPGLVSAEGDKWVKHRKLNPAFHMEKLL LMHPAFGASVLDMVNKWEKIVSKTGSSEVDVWPFVSSLTAD AISRAAFGSSYDEGRKIFELVLEQTEITLRLQLSVYIPGWMYV PTKTNRRMKTVNSEIQNLLTGIIVKRKKAMEAGEAAKDDLL GILLESNYKDTENVLSNKKLSMTFQELIDECKLFYLAGQES TSVLLAWTMILLGKHTEWQARAREEVVATFGKNEPDFEGLN	



【表 16 - 5】

			gcagcaaaagtggaaaagacttggcttaaggaggcagtggtgaaggaggcagttcagataagaat atcgacgagatcataactggcttgcgtccctctaa	
SOAP6	グリコシル トランスフェラーゼ (UDP- グリコシル トランスフェラーゼ; フコシル トランスフェラーゼ)	56	MTGKGRTEVIMMPFHHQGHLPMLQFAKRFAWKGAGSIR ITLATTSTAQNMTNSKNNNNNDYDFLTVESIYDDTDDSQL KFMGRMGKFKSEASLQLGRLITTKSIDNNKCMVYDAYLPW ALDVGKDHNIQAAFFVQACAYMASFYPMFLEEFSGDDQHP VVAAAKAESVPSLSVELPSREEMERYAPKCAQSPSSDDKPNT VKKSLHPVYRMVVSSITLHLADFVLINSFDHLEHQLDVLAH EAVGCFITHCGWNSIIEATNFGVPMLGMPQFMDQFLDAHFM EKVWGVGIRAKADEKNFVTCDEIKCGVNEIMYGDKANMIK ENAAKWKDLAKEAVGEGGSSDKNIDEIINWLASS	10
SOAP7	グリコシル トランスフェラーゼ (UDP- グリコシル トランスフェラーゼ)	57	atgggtaaaacagtagcagcttcagacagttacacatagtaatgatcccatggttcttatgga cacattctccctattttgagcttcaaaaacttctgaaaaaggccataaaatcaccttgtagtt ccaacaagaactcaactgattagaaccaagatccgtcatctcttaatacagttacatgacatt cactgtccacacattgaaccttacctccgggactgagacatgttcagacgtcccatgaact tcagaccaccttctgttgcctggacagggcccggcctgaggtggagtcctcatatagacc attgacgacccgaagccggatctgttctacgataacgcttactgggtgctgagatagccaca aagctggggatgaagctctgtttaccagattgcatgtcttaagatcacccgcaatgaagaaa ccccgagtgcgagtgcgagtgcgagtgccaagctcttacttaccacaagtggtgct gacgctaaggtgttaggcgacgcaagagccaattatggagaaggatcacctattaccagag agtgaagaaggctctaaagtctctgatgctatgccttacgacatgccgggagattgaagggg aatctctgatctctggctgcacaataacaagccagctcttctaacaggtccgctcctaccga ggttgaattctcccccttggacaattctgggctgagtggttagccaagtttgggctaagtc cgtggtttatgttcttggagccagtagtccctgacaaggctcaactacaggagatggccct tgccctgagatactggtcttccctcttctgatgtccgttaagccaccacagagtgccaccata gaggaggcgttgcagaagggttctcagagagagcaaggaacgcgggtggttcatggtgg atgggtgcaacagctacaatactagctcaccatcagtggttcttattgtcattgtgggtac gggtcaatgtggagggtgttgtagtgataaccagctagcttattaccacagcttctgaccag ttaatgatggctcaaatgttgcagaaaaagctcaaggtgggtgtgatggtgacagagaagaag atgatgggtgggttccaggaagaactgtgccaagcagtcagctctgcatgategctattccg agtttgacgttactcaagaacaacctgtaactcagagacaagtgtaaccaacgggtttat ggctaattacctgaaagtttggaccaggattgaaacgcttcttactgcaaataa	20
SOAP7	グリコシル トランスフェラーゼ (UDP- グリコシル トランスフェラーゼ)	58	MGKTVAAASEQLHIVMIPWFA YGHILPYFELSNKLAEKGHKIT LVVFNKVKLDLEPKIRHPSLISLHAFTVPHIEPLPPGTETCSDV PIELQHHLAVAMDRARPEVESIIS AIDDPKPDLLFYDNAYWV PEIATKLGKMSVFIYQIACALSITRIKQTPSASASASAKLFTL PKWVLPKVLGDARANYGEGITYYQRVKKALSSCDIALRT CREIEGESSDILAAQYNKPVFLTGPVLPEVEFLPPLDNSWAE WLAKFGPKSVLCCFGSQYVPDKAQLQEMALALEDTGLPFL MSVKPPTFCATIEEALPEGFSERVKERGVVHGGVWVQQLQILA HPSVGCFCICHCGYGSMEWGLLSDNQLVLLPQLPDQLMMAQ MLAEKLVGVMVDREEDDGWVSRKNLCQAVKSVMDPHSE FAALLKNNHANFRDKLLTNGFMANYLEVFDQLKRFLTAN	30
SOAP8	グリコシル トランスフェラーゼ (UDP- グリコシル トランスフェラーゼ)	59	atgggtggagagaagagttgcggaatgtagtgcctccatggcttgccttggacattttatcccat accttcaccttcaaaaacttctgaaaaaggccacaaaatcaccttggcttcccaacaaag ctaggcttcagttggagtcacttaaccttcatctctcataacttccattcaattactgcccacc cctcgaactctcccttatggcactgaacaactcggatctcctcagaccacatggtgaact ctcgatttccatggaccgactcggcccgggtggagcttctctateaacccataagcccgacc tcgctctctacgacatggccatgggtagccgagattgctgctaaggtcgggattaaagtcagttt catacaacgttgatgtctattgctgtagctcatgtagacctagcctcctcttccaaaaggaaag gcagcacatgtaacctgcaattgctctgctcctaaagtgagcttfaatcagcaggttcatcaa caccatatttggggaaggataactgactgaacggctgtaactcctctctgctcggatgc aatagccatccgcacgtgcagggagattgaaggggatattgtgacctgtgctgccacatca acaagcctgctctgaccagccacgcttgcctgatctgactcgaactctctcgttggaga ctcgtggccgagtgctagctaggttcgagccagggtcagtgatcttctgcttggtagtc agcatgcttagacgcaccccaactgcaagagttggccctggggttggaaatgacaggactacc	40

【表 16 - 6】

			cttcttgatggctgtaaaacccctgtggggtgtacctccttgaggaggtgcttcagaagggtt taatgatcggggttagcgggcgaggggtggtcacggtgggtgggtgcagcagcagcagataat ggcgcacccatcgttaggggtgctttgtgacctttgtgggtcttcgctgatgtggaggggttagt gagtgaagtcagttggtattactcccacaactggcagaccaaactctgatccaagtaatggc agatgagctcaaggtgggtgtgaaggtggagagagaagaacgggtggatgacgaagcga agtctatgtgaagctatcaagagtgatggatgaagatagtgatataagtcatgtagttaggaaa aatcatgctaaatagaagtagttgattagccctggctttattagtgctacattgacaactcacc aaggattacaagccctgttccttag	
SOAP8	グリコシル トランスフェラーゼ (UDP- グリコシル トランスフェラーゼ)	60	MGGEKELRIVMFPWLAFGHFIPYLHLSNKLAEKGHKITLLLP NKARLQLESNLHPSLITFHSITVPPLETLPGTETTADISLDQ HGELSISMDRTRHEVESFLSTHKPDLVLYDMAHWVPEIAAKV GIKSVSYNVVCAlAVSHVRPSLPLPKGTA AHVPLPLSSVPKW SLNQHSSTPYFGEGITLLERSVISLSSADAI AIRT CREIEGVYC DRVAATFNKPVLVTSHALPDLELELSPLETRWAEWLARFEP GSVIFCCLGSQHVL DAPQLQELALGLEMTGLPFLMAVKPPV GCTSLEEVLP EGFNDRVSGRGV VHGGWVQQQIMAHPSLG CFVTLCGSSMW EGLVSESQLVLLPQLADQTLYAKLMADDEL KVGVKVEREENGWMTKRSLCEAIKSVMEDESDISHVVRKN HAKYRSMLISPGFISGYIDNFIKDLQALVP	10
SOAP9	グリコシル トランスフェラーゼ (UDP- グリコシル トランスフェラーゼ; キシロシル トランスフェラーゼ)	61	atggagcttcaaacctaccacaacccctaccttaaacgcaaccaacccctacgaggctattc attccattaatcacagttccaagccacattccaattctgtgacattgctaactctctcatcacgg ggagtacatgtgactatcctcaccaccacacacctccctccgctcaacaatccatacatgat tggggctcacaatcgacctccacatcgtcgaactcccgttcagggaagtcggctaccggaagg agtggaaaattacagtgatgccacccctgagcaagcaagccagctttccaggcctttatgatgct tcagaagcctatggaggtgccattcgggtgtaagcccgaactcattccgataggtattat cattggtctactgatctgcacgtgagcttgctattccacggctcattctcatgcatgattttg cattgtgtgctgctgaggtgtgccaagttgccctcatgagaaggtgaaactgacactgacct cttttctctcctgacctcctgataccatccacatgaccggttgcaagcttccgaatggattcaga cccgaacatgttctactgttctcaatgagagaatggacgagggcgataggagtgctacgtgtt attgtaaacagctgctacagtgaggagagacttatgctgacttctaccgcagcaactgggtcga cgtgctgtgtgategtcctatccggtaactgcgacaaggtgtaaaaagaaggagatga cagcaaaaaacattcattgtttgaatggctgataaaatgggagaaggagaagtatatactgag ttttggcactctgctgtttcagccctgctcaaatctcagagctggctactgcaactgaaatgctg gtcaccggttatctggglagtaaggaatggtgagaatgttacctgatggattgaagaaagaat tacagagcaggacaaaagggtgttaataaaagactggcgccacaagtgaataacttgagca cccagctgtaggcggatttctgactcattgtggatggaactcaactgtagaaagttagcagcagg tgtccaatggtcacatggcgcctgtgtgccgagcaattctcaatgaaagttagttaggtt ttgaaggtgggggtcagagtcgggtctgagaagtggagtaggggtatagaccgaatactgata tgattgagaaggataaaatagaaaggc gattaaagagctgatgagtaaagaaccgagggcg aggaaaaggagcagaaggtaaggagttgagtaaggctccaagaatcggttgaagaagg ggctgctcgtacaatttaagtagttgattgaaaaattacaacgtttaaaggcgaatgaaatc agtgtccacaactcagaataa	20 30
SOAP9	グリコシル トランスフェラーゼ (UDP- グリコシル トランスフェラーゼ; キシロシル トランスフェラーゼ)	62	MELSNPTTPTLNATQPLRGYFIPLITVPSHISNLVDIAKLFSSR GVHVTILTTHTSLRFKQSIHDWGFKIDLHIVDFPFREVGLPE GVENYSDATPEQASQLFQAFMMLQKPMEDAIRAAKPDFIVS DRYYHWSTDLARELAIPRLIFHVRCYFALCAA E V VAKFAPHE KVESD TDLFFLPDLPDIHMTRLQLPEWIQTRNMFTVLNERM DEADRECYGVIVNSCYELERA YADFYRSNLGRRAWCIGPYP VHCDKVGKKKGDDSKKHSCFEWLDKMGEGEVIYVSFGTLS CFSPAQISELATALEMSGHPFIWVVRNGEKLLPDGFEERITEQ DKGVLIKDWAPQVKILEHPAVGGFLTHCGWNSTVESLAAGV PMVTWPLGAEQFFNEKLISGVLKVGVEVGSEKWSRGIVPNT DMIEKDKIERAIKELMSKEPEAEERRQKVKELSKAPRNAVEE GGSSRNLSDLIEKLQRLKANEISVSTNSE	40
SOAP 10		63	atgggggaagtcaacatgaagaagtgaatgaataatataatagaaccataaaaccatc atcactactccacaaaactcctccaaaacatcacactttctcactcagcaagctgccct	

【表 16 - 7】

	<p>アシル トランスフェラーゼ (BAHD アシル トランスフェラーゼ)</p>	<p>ttgtactactatccttacttftaactacactaactactactactacccaacatcacaaaatcgag          ttgacataacaagfacctaaaactcacttagcaaaacacttgacaatfccacctattgcagg          tcgatgtgggacgactctacaattgttgaaccaccaaggataaccattcattgaaacaaaagt          gactccaatatcttgatgcatgaactcgccctgagaaaatgaagtgcttatacaagttctccctca          tgcagagttcaagatgtagctgaccagctcggattaaaccattggcgtttcaagtcaatggtt          ccggtgtgggtggggtgatcattggctcctatgtgctccacaagctccttgatggaatctctcttga          acttcttataaaattggcaaccattgctaagatgagcgaagtaaggacgacgactagtacaac          ctgacttgaagccactattaaggcgttccctccgctacagcaactccaatgcttctctgtaatca          acaactccaaggcggctgaaaaccaataataatccagtcaaaagttcttgacaaaagact          tcgtattgacattgttcttaagaagatgattcatggctaaagatggaattggtctctaaacca          ccaaatggagaccgtgacagggttatttgggaacaaacctatcaacattgctgaattctggagt          gaagtgacatacatcgttataactgtaaacatccgcccaagtagtgaccgacccacc          aagaggatccatggtaactgtcgaagatgcaaggcagcagcccaacccagcagcagca          atgggctcaagaccctgtaagaatccattcatcttctcaacaaccagaaaataact          cctctcctcctcctcctcctactactactgctacaacaatccattcatcttctcaacaacaa          cccagaaaataactcctcctcctactactacaacaatccattcatcttctcaacaaccag          aaaataactactactactacagcagaggtattttgactaaacggaaagtgacaatccagtt          cacagaatcgagaaggaaactacctctaccagttggcgcaagattgggtggatgagcgtga          ctctgggttcgaaaagcccgttgggtaattcccaacgatgggagacccectaaagtcaggaat          atgatttcttactgattatagcattcccgaacaggcgttgaaggaaatgacagcattgattacgt          tggagagaaaacaaatgcaatgtttaaagtcacaaccagaattccttcttctactcctaattag</p>	10
<p>SOAP 10</p>	<p>アシル トランスフェラーゼ (BAHD アシル トランスフェラーゼ)</p>	<p>64          MGEVNHEEVEIEIISIETIKPSSLLPPKTPPKTTITLSHLDQAAPL          YYYPLLLYYTNTTTTTPTSQIRVDITSTLKTSLSKTLDKFHPIA          GRCVDDSTICCNHQGIPFIETKVDSNILDVMNSPEKMKLLIKF          LPHAEFQDVTRPVSDLNHLAFQVNVFRCGGVIIGSYVLHKL          DGISLGTFFKNWSTIANDERVKDDDLVQPDFEATIKAFPPRT          ATPMLPRNQQLPKAAEKPNNNPVKVLVTKSFVFDIVSLKKM          MFMAKSELVPKPTKFETVTGFIWEQTLSTLRNSGVEVEHTSL          IIPVNIRPRMSPPLPRGSMGNLLKNAKAQANTSSSNGLQDLV          KEIHSSLQTTQKINTPPPPPPPPPTTTATTIHSSLQTTQKINTP          PPTTTTIHSSLQTTQKINTTTTTAEVILTKRKVDNPVTQNRE          GNYLFTSWCKIGLDEADFGFGKPVWVIPNDGRPPKVRNMIFL          TDYRHPETGVEGIAAWITLEEKQMQLKSNPEFLAFATPN</p>	20
<p>SOAP5</p>	<p>セルロース 合成酵素様G (グルクロン酸 トランスフェラーゼ)</p>	<p>65          atggcaactctcacattcgcaatgccaatlaaccagaccattgtaaccgtctccacatcttct          ccattccgtagccattctctcctcttactaccgttcaactctcttcaactccgacatctccata          ctgttactcttactaccaccgccgaactcttctaaccttctatgggttttactcaggcttcc          ggtggcgtcccgaatgagggagctcgggtacgaatccatcaaacccgaacaactaccgg          gttggatgtctcattgctactgctgaccgacaaaaggaccagttctggaggtgatgaactccg          tgatatacctatggcttggattatccggttgatagactggcggttactgtcggatgacggtgt          tctccgttctgaaggaggcattaagaaggcttatgattgctaaagcttggattccttttgaat          aagtataatgtaagacaaggctgctcaggttcttctcgecctctgctgatgggaaaggctg          attggaattctgatttatggctgatcaattggaactccagaccataatgaagctttgagaactat          gtggagaaaagaaaggagataacaccaaagctgactgcagttcatgacgacctcttgcgttga          gattatacatgacaacaacagacggagaaagtgatgtgaagatgccctctggtttatgtagc          cagggaaaagagacctgtcgtcctcatcggttcaagctggagccctaatgctctctctgagt          atccagttatgagcaatgcaacttactattggtgttgattgtgatgactgcatgatccaac          ttctgctcgtcaatctatgcttccatcttgacacaacatggctcctctctgcatatgtgcaatac          cctcaaatcttataatgttagcaaaaatgacatctatgatggccaagccagatcagctcatatga          cgaatggaaaggcatggatgactcagaggccggcttgaatggaactgggtattatttgaag          gcgaaaagcattattggaaaagcctaataacgaagatgaatacctcaacagtcaccagaaaag          gccttggctcctccacaaaatfaattgctgactaagagagaactccaagcaaaaatctgccata          aaggaaatgacagaagatgagttgaccaagaggctagaatggctactgacatatgaagc          aaacacactatgggagcagtgaggtaggatattctgatgagtgcttggagagatcattactgg          atatatgttacattgcagaggatggaatctgtgatcttaccacaaaagaccatgcttctgggat          gcacaacgattgatagaggatgctacggttcaactaataaaatggacctcctcattactggaat          tgcctctcgaagcttagccctctaaccttggccatgccaagatgcaatcctgcaagcattggt</p>	30
40	50		



【表 16 - 9】

			<p>GGAATTGACAGAAGATGAGTTGTACCAAGAGGCTAGAAA  TTTGGCTACTTGCACATATGAAGCAAACACACTATGGGGC  AGTGAGGTAGGATATTCGTATGAGTGCTTGTGGAGAGTA  CATTCACTGGATATATGTTACATTGCAGAGGATGGAAATC  TGTGTATCTTTACCCAAAAAGACCATGCTTCTTGGGATGC  ACAACGATTGATATGAAGGATGCTACGGTTCAACTAATAA  AATGGACCTCCTCATTACTTGGGAATTGCCCTGTCGAAGTCT  AGCCCTCTAACTTTGGCCATGTCCAGTATGTCAATCCTGCA  AAGCATGTGTTACGCGTACATCACATTTACAGGCCTTTTGG  CAGCTCCATTGGTTATATATGGTGTGTCCTTCCAATAAGC  CTATTGAAGGGCTTTCCTATTTTCCCTAAGGTATCGGATCC  ATGGATTTTGCCATTTGTGTTGATATTTGTATCCTCCCATC  TTCAACATCTATATGAGGTCCTGGAAAGTGACAAATCAGC  AACACAATGGTGGGAATGAGGTGAGAATTTGGATGATGAA  ATCAGTGACAGCCTGTTTGTGTTGGGTTGACGGAAGCGATA  ATGAAGAAGATTGGAGTACAACTGCAACATTCAGATTA  ACAAATAAGGTAGTTGAGAAGGAAAAGATGGATAAATAC  GAGAAGGAGAGGTTTGTATTTCTCAGGAGCAGCTATGCTTA  TGGTTCCTCTTAATATTTTGGTGGTACTAAATATGGTGTCA  TTCATTGGTGGACTCATGAGGGTCATAATCAACAACAGTT  ATGATCAAATGTTTGCACAACCTTTTCTCTCTTTTTTGTG  CTACTTCTTAGCTACCCTGTTGTTAAGGGATGGTTA</p>	10
				20

## 【1475】

要約：本試験により、数々の驚くべき詳細が明らかとなった。例えば、キラヤ (*Quillaja saponaria*) 由来の強力なワクチンアジュバントである QS - 21 (「D.J. Marciani, Is fucose the answer to the immunomodulatory paradox of Quillaja saponins? *Int Immunopharmacol.* 29, 908-913 (2015).」) の生合成経路においてキラヤ酸を修飾するフコシルトランスフェラーゼについての相同性に基づく検索において、SOAP6 は、テンプレートの役割を果たす可能性がある。また、アシルトランスフェラーゼの BAH Dスーパーファミリーのメンバーである SOAP10 は、トリテルペノイド系サポニン生合成に  
関与している。

30

## 【1476】

実施例 21：酵母におけるトリテルペノイド中間体の産生

## 【1477】

目的：酵母 *サッカロミセス・セレビスエ* の場合に、さらなる異種系において、トリテルペノイド中間体を産生すること。

## 【1478】

方法：上記の材料及び方法を参照されたい。

40

## 【1479】

結果：MA - 3 - GlcA の産生におけるさらなる植物酵素の関与を除外するため、本発明者らは、*サッカロミセス・セレビスエ* において SOAP5 を発現させた (図 33D)。酵母はアミリン型トリテルペノイドも UDP - グルコン酸も産生しないため、SOAP5 活性に必須の基質を得るためには、これらの前駆体を産生する酵素の発現が必要であった (「P. Arendt et al., An endoplasmic reticulum-engineered yeast platform for overproduction of triterpenoids. *Metabolic Engineering.* 40, 165-175 (2017)」; 「T. Oka, Y. Jigami, Reconstruction of de novo pathway for synthesis of UDP-glucuronic acid and UDP-xylose from intrinsic UDP-glucose in *Saccharomyces cerevisiae*. *FEBS Journal.* 273, 2645-2657 (2006)」)。

50

## 【1480】

酵母においてSOAP 1、2、3、及び4を発現させたところ、メディカゲン酸が産生した(図40A~40B)。しかしながら、SOAP 5の添加は、メディカゲン酸3-O-グルクロニドの産生には不十分であった(図33A~33F)。SOAP 5に活性化糖を提供するため、UDP-グルコースからUDP-グルクロン酸への変換を触媒する、ホウレンソウ(SoUGD1、Sp\_\_189830\_\_psca)由来のUDP-グルコース6-デヒドロゲナーゼ1を発現させた(配列番号74[遺伝子]及び配列番号75[ポリペプチド])。

## 【1481】

ATGGGTGAAGATTTGCTGCATTTGGGGCTGGTTATGTAGGA 10  
GGCCCTACTATGGCTGTTATAGCACTCAAGTGCCCAAGA  
TTGAAGTTGTAGTGGTTGATATATCTGTGTCTCGGATCAC  
TGCATGGAACAGCGAGCAGCTTCCAATCTATGAGCCGGGT  
CTAGATGATGTGGTTAAGGAATGCCGTGGAAGGAACCTTT  
TCTTTCAGCACTGATGTAGAAAAGCATGTTGCTGAGGCTGA  
TATTGTTTTTGTCTCTGTGAATACCCCTACCAAAAACACA  
GGTCTTGGAGCAGGCAAGCTGCTGATTTGACCTACTGGG  
AGAGTGCTGCCCGTATGATTTGCTGATGTTTCAAAGTCTGA  
CAAAAATCGTTTGTGAGAAATCAACTGTGCCAGTGA AAACT  
GCTGAGGCAATCGAAAAGATTTCTGACGCAACAACAGCAAGG 20  
GAATCAACTACCAAGATCCTTTCAAATCCGGAGTTCTTTGC  
TGAAGGTACTGCTATTCAGGACCTTTTTTCAACCCTGACAGG  
GTTCTCATCGGTGGCCGGGA AACCCAGCAGGCCCTCAAAGG  
CAGTCCAAGCATTTGAAGGATGTGTATGCTCAATGGGTTCC  
TGATGAACGGATCTTAACCAACCAATCTTTGGTCTGCTGAG  
CTCTCAAAGCTTTGCTGCCAACGCCTTCTTAGCACAGAGGA  
TTTTCATCTGTCAATGCAATGTCAGCTCTTTGTGAGGCTAC  
TGGAGCAGATGTTACCCAAGTCCGCATATGCTGTTTGGTAAG  
GACAGTAGGATTTGGGCAAAAAGTTTTTGAACGCTAGTGTTG  
GTTTTGGAGGGTCTTTGCTTCCAGAAAAGACATTTCTGA ACTT 30  
GGTTTACATTTGTGAGTGC AACGGTCTCCCTGAGGTGGCC  
GAGTATTTGGA AACAGGTAATCAAAGGTGAATGATTAATCAGA  
AGAAATCGTTTTTGTGAATAGGGTTGTGGCCCTCCATGTTCAA  
CACTGTATCGAACAAGAAGATTTGCTGTTCTTTGGATTTGCA  
TTCAA AAAAGGATACAGGGGATACTAGGGAGACACCCGCCA  
TAGATGTGTGCAAGGGTTTGTGAGGAGACAAGGCCAAGGTT  
GAGCATCTACGATCCACAAGTCACTGAGGATCAGATTCAG  
CGAGATCTCACCATGAACAAGTTTTGACTGGGACCACCCAA  
TTCACCTCCAGCCCAACAAGTCCCAACA ACTGTTAAGCAAGT  
GAGTGTTGTTTGGGACGCTTATGAGGCCACCAAGAATGCT 40  
CATGCTGTGTGTATTTCTGACTGAGTGGGATGAATTTAAGA  
AACTTGATTA CA AAAAGGATTTTTTGAACAACATGCAGAAAGCC  
AGCTTTTTGTGTTTGTATGGAAGGAACATTTGTGAATGCAGAT  
GAGCTGAGGCAGATTTGGGTTCAATTTGTGTACTCAATTTGGTA  
AACCTTTGGATTCATGGCTCAAGGACATGCCTGCTGTGGC  
TTAA (配列番号74)。

## 【1482】

MVKICCI GAGYVGGPTMAVIALKCPKIEVVVVDISVSRI  
TAWNSEQLP IYEPGLDDVVKECRGRNLFSTDVEKHVAEA  
DIVFVSVNTPTKTTGLGAGKAAADLTYWESAA RMIADVSKS 50

D K I V V E K S T V P V K T A E A I E K I L T H N S K G I N Y Q I L S N P E F L  
 A E G T A I Q D L F H P D R V L I G G R E T P A G L K A V Q A L K D V Y A Q W V  
 P D E R I L T T N L W S A E L S K L A A N A F L A Q R I S S V N A M S A L C E A  
 T G A D V T Q V A Y A V G K D S R I G Q K F L N A S V G F G G S C F Q K D I L N  
 L V Y I C E C N G L P E V A E Y W K Q V I K V N D Y Q K N R F V N R V V A S M F  
 N T V S N K K I A V L G F A F K K D T G D T R E T P A I D V C K G L L G D K A R  
 L S I Y D P Q V T E D Q I Q R D L T M N K F D W D H P I H L Q P T S P T T V K Q  
 V S V V W D A Y E A T K D A H A V C I L T E W D E F K K L D Y K R I F D N M Q K  
 P A F V F D G R N I V N A D E L R Q I G F I V Y S I G K P L D S W L K D M P A V  
 A ( 配列番号 75 )。

10

## 【 1483 】

その結果、他の5つの遺伝子全てと共にSOAP5を発現する酵母細胞のみが、MA-3-GlcAを産生することができた(図33A~33F、図40A~40B)。ベンサミアナタバコのアッセイと同様に、SOAP5は、酵母細胞において、トリテルペノイド中間体(オレアノール酸、ギブソゲン酸及びバヨゲニン)をグルクロン酸抱合することができた(図41A~41F)。

## 【 1484 】

要約：元の酵素とは異種の系において、トリテルペノイド中間体を産生しかつグルクロン酸抱合することが可能であることが示された。

20

## 【 1485 】

実施例22：さらなる植物種におけるセルロース合成酵素様酵素の検討

## 【 1486 】

目的：非ハウレンソウ植物種における機能性セルロース合成酵素様G(CSLG)酵素を同定し、解析すること。

## 【 1487 】

方法：上記の材料及び方法を参照されたい。

## 【 1488 】

結果：C-3位のヒドロキシルを介してサポニンアグリコンに結合したグルクロン酸は、植物、特に、ナデシコ亜綱(Caryophyllidae)、原始的なバラ亜目(Rosidae)及びキク亜目(Asteridae)において見られる多くのサポニンについて共通の特徴である(「M. Henry, Saponins and Phylogeny: Example of the "Gypsogenin group" Saponins. Phytochem Rev. 4, 89-94 (2005)」)。CSLG酵素によるグルクロン酸のトリテルペノイド骨格への抱合がハウレンソウ及びその科に限定されるものであるか否かを調べるため、CSLGタンパク質の系統発生解析を、コケ、裸子植物、及び顕花植物を含む、70を超える植物種において行った。ハウレンソウSOAP5は、ナデシコ目(Caryophyllales)、アオイ目(Malvales)、セリ目(Apiales)、及びマメ目(Fabales)由来のCSLGタンパク質に関連していることが明らかとなった(図42A~42E)。

30

## 【 1489 】

選択した種及び公表データのサポニンプロファイリングにより、ダイズ(Glycine max)、アルファルファ(Medicago sativa)、ミヤコグサ(Lotus japonicus)及びカンゾウ(ウラルカンゾウ(Glycyrrhiza uralensis))並びにビートルート(Beta vulgaris)及びキノア(Chenopodium quinoa)のようなナデシコ目(Caryophyllales)の種において、グルクロニド-オレアナン型サポニンの存在が示された(図43A~43C)。

40

## 【 1490 】

次いで、ハウレンソウSOAP5のうちの最も近いホモログを、前述の種の全てからクローニングし、ベンサミアナタバコにおいて、SOAP1~4と共に、一過性に発現させた。試験した全てのCSLG酵素について、MA-3-GlcAの産生が認められた(図42A~42E)。以下の表17は、これらのSOAP5ホモログの核酸配列及びアミノ酸配列を示す。

## 【 1491 】

50

表 17 : S O A P 5 セルロース合成酵素様 G ホモログ ( グルクロン酸トランスフェラーゼ活性を有する )

【 1 4 9 2 】

【表 17 - 1】

名称	SEQ ID NO:	核酸／アミノ酸配列
赤カブ由来の BvCSLG ポリペプチド (XM_010673823.2_)	94	MSSLHICKVQTTRAILSRFHILFHSALALFYRFTSFSTTKSGILPWTLTAEV VLGFVWALTQAFRWRPVLRDVAGWDSIKEEQLPGVDFICTADPIKEPVLEVM NTVLSAMALDYPAEKLGVYLSDDGGSPLTREAIKEASKFAKVWLPFCISKYGIKT RCPQAFSSFCDGERLDWNQDFKADELVLKSKYEAFKNYVEKASEDESKCTMA HDRSPCVEIIHDNKQNGEGEVKMPLLVYVSREKRPNRPHRFKAGALNALLRVS GVLSNGPYLLVLDCEMYCNDPTSARQSMCFHLDPKLAPSLAFVQYQIFYNYSK NDIYDGGQARSAYKTKWQGMGIRGPVLTGTGYLKRKALYQPHNEDEFLIN QPEKAFGSSTKFIASVSSNSKQNMALKEMTRDDLLEAKNLATCAYESNTEWG NKIGYSYECLLESTFTGYLLHCKGWISVYLYPKRPCFLGCTTIDMKDAMVQLM KWTSGLLGVGISKFSPLTYAFSRMSILQSMCYGYFTFSALFGVSLIYGIVLPVCL LKGVPVFPKVSDPWIGVFVVVFASLLQHLVEVLSSDDSIKTWWNEIRWIKSV TASLFGTMDAIMKKIGIQKASFRLTNKVVDKEKLEKYEKGFDFQGAAVFVMP LILVVLNMVSFVGGRLRAIINKNCDEMFQQLFLSFFLLVLSYPVLEGIVTKVRKG RD
赤カブ由来の BvCSLG 遺伝子 (XM_010673823.2)	95	CATCACAGCCACATGGGAAACAAAAAGTCTTAACCTCGGCCATCTCGTGTG CCTCGTGCACATTCAACATTGATTTTAATTGTGTGCTTTCTCCAAAATGTACC ATACTATATCATCTTGTGAGAATCCAAACTCCAATAACCACCATTGAA GTCATCTACACACAAACACACACTCTTTCTCCCTAAAAATGTCTTCTC CACATTTGCAAAGTCCAAACAACAAGAGCAATACTTAGCCGTTCCACATA CTCTTCCACTCCTTAGCCATCCTTGCTTTATTCTACTACCGTTTTACATCGTTC TCTACCACCAATCAGGCATACTTCCATGGACCTTACTAACCACAGCAGAG GTGGTCTAGGCTTTGTATGGGCGTTAACACAGGCCTTTCGATGGCGGCCCTG TGTTGCGAGATGTAGCTGGATGGGATTCCATCAAGGAGGAACAACCTGCCAG GGGTGGACGTGTTTCATATGCACAGCTGATCCAATAAAGGAGCCGGTGTAG AGGTGATGAACACGGTGCCTTCGGCGATGGCATTGGATTACCCGGCAGAGA AGTTGGGTGTTTATCTTTTCGGATGATGGAGGTTCTCCCTTGACTAGGGAGGC TATTAAGGAGGCTTCTAAGTTTGCTAAGGTTTGGCTTCCTTTTTGTAGTAAGT ATGGTATCAAGACTAGGTGCCTCAGGCTTCTTCTCTCTTTTTGTGATGGG GAAAGACTTGATTGGAATCAGGACTTTAAGGCTGATGAATTGGTGCTCAAG TCAAAAATATGAAGCTTTTAAGAATTATGTGGAGAAAGCAAGTGAAGATGAA AGCAAATGCACCATGGCACATGATCGTTCCCCTTGCCTGAGATTATACATG ACAACAAGCAAAATGGAGAAGGCGAAGTGAATAATGCCCTTTTGGTCTACG TATCCAGGGAAAAGAGACCAAATCGTCCTCATCGTTTCAAAGCCGGAGCTC TTAATGCTCTTCTCAGAGTATCAGGTGATTGAGCAACGGGCCTTACTTATT GGTGTGGACTGTGATATGTAAGTCAATGATCCAATCTGCTCGTCAATCT ATGTGCTTTCATCTTGACCAAAAATTGGCTCCTTCACTTGCATTTGTGCAATA CCCACAAATTTTCTACAACACCAGTAAAAATGATATCTATGATGGCCAAGC TAGATCCGCGTACAAGACAAAATGGCAAGGAATGGATGGTATTAGAGGAC CAGTCTTGACAGGAACAGGGTATTACTTGAAGAGGAAAGCATTGTATGGAC AACCTCATAACGAAGATGAATTTCTCATTAAATCAACCAGAGAAGGCCTTCG GCTCCTCCACAAAATTCATTGCGTCAGTTAGTTCAAACCTCCAAGCAAAAATAT GGCCTTGAAGGAAATGACAAGAGACGACTTGTTAGAAGAGGCTAAAAATTT GGCTACTTGTGCATATGAATCAAACACTGAATGGGGTAACAAGATTGGATA TTCGTATGAGTGTGTTGGAGAGTACATTTACCGGATATCTCTTACATTGC AAAGGATGGATTCCGTGTATCTTTACCAAAAAGACCCTGCTTCTTAGGAT GCACGACGATTGACATGAAAGATGCCATGGTTCAACTAATGAAATGGACCT CTGGATTACTAGGAGTTGGCATATCAAAGTTAGCCCTCTAAGCTTATGCCTT TTCGAGGATGTCTATATTACAAAGCATGTGCTACGGTACTTACATTTTCA GCCCTTTTCGGAGTTTCGTTCTTAATATATGGCATCGTCCTTCCAGTATGCCT ATTGAAGGTGTTCTGTTTTCCTAAGGTATCGGATCCATGGATTGGAGTT TTCGTGGTAGTATTTGCATCCTCCTCCTTCAACATTTATACGAGTTTCTCTC AAGTGACGATTCCATTAACATGGTGGAAACGAGATCAGAATTTGGATCAT CAAATCGGTAACAGCTTCTTATTTGGAACAATGGATGCAATAATGAAAAA GATCGGCATACAAAAGGCTAGTTTCCGATTAACATAACAAGGTTGTGGACAA GGAAAAGCTCGAAAAATATGAGAAGGGCAAGTTTGATTTCCAAGGAGCAG

10

20

30

40

【表 17 - 2】

		<p>CTGTGTCATGGTTCCTCTTATCATTTTTAGTGGTACTAAATATGGTGTCAATTT  GTTGGCGGATTAAGAAGGGCAATAATCAACAAGAATTGTGATGAAATGTTT  GGGCAACTTTTCCTCTCATTCTTTCTCTTAGTTCTTAGCTACCCCGTTTTAGA  AGGGATAGTAACAAAAGTAAGAAAAGGACGTGATTGAGATGAATTTGCATT  GTTTGGTAAAAGATCCAAACTTAGAGAAAAGAGATTGCGTAGGAGATCAAAG  GAAACAATGTGAGAGATTTACAGGCTTCATGAGGCTTAAGACCTCATTAAT  TTTTGTGACAATTTACAAATTCTGTCTCTATATTTTTGGTCAAGACGTATCATT  TGAAAATTTCCATGGTTAGGTAGTTAGATTTCAATGTTCCACGTTTGTAAT  AAGAAGATAAAAATAAGGAAATTTGTGATTTTAGCTTTACATTTCTATGAGAT  AGTCCTTTGTTGTGATGAAAGTTGTGTTCTTAGTGAAAATAATAAACGTGA  CATCAAAATTTTGAGTATAT</p>	10
<p>キノア 由来の CqCSLG タンパク質 _(XM_021 866098.1)_</p>	96	<p>MAATHICKVQTKRVIINRIHILFHSAILALFYRFSSFSNPHISLFPWVLL  TIADLVFTFIWAMTQAFRWRPVLHDVSGYESINPRDLPKIDIFICTADPTK  EPVLEVMNSVISSMALDYPPEKMAVYLSDDGGSPLTREAIKKAVEFAKV  WIPFCNMYGIKTRCPDAFFSALGNDRLHRDQDFNAHESLLKSKYEAFK  KYVEKESGDINKCTVVHDREPCIIEIHDQDGEAEVKMPLVVYVAREK  RPGHPHRFKAGALNALLRVSGLLSNAPYLLVLDCDMYCHDPTSARQSM  CFHLDPNMSPSLAFVQYPQIFYNTSKNDIYDQARSATTKWQGMGDL  RGPVLNGTGYLKKKAIYGRPHNEDEYLINEPEKAFGSSTKFIASLKENS  NQDLVLKEFTNDLLQEARNLATCTYEANSLWGVEVGFSDCLLESSYT  GYLLHCKGWRSVYLYPKRPCFLGCTTIDMKDAIVQLIKWTSGLLGVAM  SKFSPLTYAMSRMSILQSMCYAYITCSGLLAVPLFIYGVVLPFLLKGV  VFPKVSDPWMLGFVVFVSSHVQHLFEVLASDHSVQWVWNEVRIWIM  KAITACLFGSTEAIMKKIGIQKTTFRLTNKVVEKEKLDKYEKGFDFSGA  AMLMLVPLIILTLNLVSVFVGGVLRVINHNNDYDDMFGQLFLSFYLLLSYP  TFEGIVTKVTDKLRKKE</p>	20
<p>キノア 由来の CqCSLG 遺伝子_(X M_021866 098.1)_</p>	97	<p>AAAAGTGAAACTGGTGACTTAGTCTTGTGTCACCCGGGTTCCCTCGAGCACAT  ATCTTCAAGATTTTGGTTTTTGTACGGAGTATTTATATACACAAAAATTA  GGAACCAAATAGGAAGACTCATATCATTTCAAAATGGCGGCAACACACATT  TGCAAAGTCCAAACAAAAGAGTCATTATCAACCGTATTCATATCTCTTTC  ACTCTTTAGCCATTCTGTCTCTCTACTACCGTTTCTCGTCTTTCTCCAACC  CTCATATCTCCCTTTTCCATGGGTATTATTGACTATCGCCGACCTCGTTTTC  ACCTTCATTTGGGCCATGACTCAGGCCTTCCGTTGGCGCCCGTCTTGACAG  ATGTGTCTGGCTATGAGTCCATCAATCCACGCGATCTTCCAAAGATCGATAT  TTTTATATGCACCGCTGATCCCACCAAGGAGCCTGTGTTGGAAGTGATGAAC  TCGGTGATATCATCCATGGCGCTCGATTATCCGCTGAAAAAATGGCGGTGT  ATTTGTGCGATGATGGTGGTTCTCCTTTGACTAGAGAGGCTATTAAGAAGGC  TGTTGAATTTGCTAAGGTTTGGATTCCTTTTTGTAATATGTATGGTATTAAGA  CTAGGTGTCTGATGCTTTCTTCTCCGCTTTGGGTAATGATGAAAGACTTCA  TCGTGATCAAGACTTTAACGCTCATGAATCACTCCTCAAGTCGAAATACGA  AGCTTTTAAGAAATATGTGGAGAAAGAAAGCGGTGATTAATAAATGCAC  CGTTGTGCATGATCGTGAACCTTGCAATTGAGATTATACATGACAGTAAACA  GGATGGAGAAGCTGAAGTGAAAATGCCCTTGTAGTTTATGTAGCCAGGGA  AAAGAGACCAGGTCATCCTCATCGTTTTCAAAGCTGGAGCCCTTAACGCTCTT  CTCCGAGTATCAGGACTATTGAGCAATGCGCCTTACTTATTGGTGTAGACT  GTGATATGTAATGTCATGATCCAACCTCTGCTCGTCAATCTATGTGCTTCCAT  CTTGACCCGAACATGTCTCCCTCTCTTGCCTTTGTTCATACCCTCAAATTTT  CTACAACACTAGTAAAAATGATATCTATGATGGTCAAGCCAGATCAGCTCA  TACGACGAAATGGCAAGGCATGGATGGACTCAGAGGACCGGTCTTGAATGG  AACTGGGTATTATCTGAAGAAGAAGGCGATATATGGAAGGCCCCATAATGA  AGATGAATACCTCATCAATGAACCAGAAAAGGCTTTTGGTTCTTCCACAAA  ATTCATTGCTTCACTTAAAGAAAACCTCGAACCAGGATCTTGTCTTGAAGGAA  TTCACAAACGATTTGTTACAAGAGGCTAGAAAATTTGGCTACTTGCATTTATG  AAGCAAACCTCGTATGGGGTGTGAGGTAGGGTTTTTCGTATGATGCTGTG  GGAGAGTTCATACACTGGATATCTCTTACATTGTAAGGATGGAGACTGTG  GTATCTTTATCCCAAAGACCGTCTTCTTGGGATGCACGACAATTGACATG  AAGGATGCTATTGTTCAATTAATAAAAATGGACTTCCGGATTACTTGGAGTTG  CCATGTCAAAGTTTAGCCCTCTTACTTATGCCATGTCCAGAATGTCTATATT  GCAAAGCATGTGTTACGCGTACATCACGTGTTTCAGGTCTTCTAGCAGTTCCA  CTCTTATATATGGTGTGTTCTACCATTCTGCTACTTAAGGGCGTTCTCTGT</p>	30 40

【表 17 - 3】

		<p>TTTTCCTAAGGTATCGGATCCATGGATGTTGGGTTTCGTGTTTGTATTTGTAT  CCTCCCATGTTCAACATCTATTTCGAAGTGCTAGCAAGTGATCATTCAAGTGCA  ACAGTGGTGGAAATGAGGTGAGAATCTGGATCATGAAAGCGATAACAGCCTG  CTTGTTGGATCAACTGAAGCAATAATGAAGAAGATTGGGATACAGAAAAC  AACATTCAGATTAACAAATAAGGTTGTGGAGAAAGAGAAGTTGGATAAATA  CGAGAAGGGAAAGTTCGATTTCTCAGGAGCAGCAATGCTAATGGTTCCTCT  CATCATTTTGACTATACTAAATTTGGTGTCTGTTCCGTTGGGGACTTGTAAGG  GTGATCAACCACAACAACATGATGATATGTTCCGGGCAACTTTTCCTGTCAT  TTATCTCCTACTTCTTAGCTACCCTACTTTTCGAAGGGATTGTTACAAAAGTT  ACAGACAACTTAGAAAGAAAAGAATAAGGAGTGATTGAGTAACTGCCTAG  TACAGTTTTCACTTCACTTCTAGATTAGTCCTGTTTTTGTATGTTTATT  AAGATCAGCAACACTTGTAGACGGTTGCAATAATGAGTTCAATACCGTTTG  TCTGTCCCTCTGCAAGAACAAGTATATAAATACTTTTCATTAGCCGGTTG  CATTGTTGGATTCATATAGATGAATATTTCAAATATTTTCATCTTTTGAAGTT  ACACACTAATGATTTATATCTGGAATTTTGAAA</p>	10
<p>アルファルファ 由来の MsCSLG タンパク質_ (MSAD_2 99835)</p>	98	<p>MATFTFHKETVQPLLPLRRAYIIFHFTCVLFLFYRISNLFISYPWFLMTIAEILSF  LWFFNQAFRWRLVNRVSMTEKLPPEEKLPLDIFVCTIDPEKEPTVDVMNTVIS  AIAMDYPSNKL SIYLSDDGGSPITLFGIKEAFEFKAVWVPFCKKYDVKSRCPKFF  FTALGENERLHRPREFEEVDQIKKRLNRVDPNSSQVENSKEHMTKAKYEKM  QKNIEKFGSNLKNLCMVTD RPSRIEINDQKEMPLVYVVSREKRPSVPHRFKGG  ALNLLRVSGLISNGPYVLVVD CDMNCNDASSAKQSMCFFLDPETSKDVAFVQ  FPQMFHNL SKKDIYDSQTRT AFTTKWKGM DGLRGPGLT GSGNYISRSALLFGSP  NQKGDYLLDALYNFGKSNMYVESL KALRGQQT KKQNSRDVILQEACEVASCS  YERNTNWGNVGFYSYAIKLESTVTGYLLHCRGW RSTYLYPKRPF LGCAPTDM  KEGLIQPIKWSSELLLLAISKYSPFTYGLSRLPTIHCLTFCYL VSTTQFATAYILYG  FVPQICFLKGIPVYPKVTD PWFIVFTVL YLSSQIHHYIEVISTGGSSMIWWNEQRS  GIVKSIGCVFAI IETAKKKFGLNKAKFTLSDKAIDKDKLKKYEQGFNFDGAAL  LMAPVIVLLTINIVCF FGGWLRLNVRDFDEMFQQLFLIYILALSHPIVEGIISMK  RKSG</p>	20
<p>アルファルファ 由来の MsCSLG 遺伝子_ (MSAD_2 99835)</p>	99	<p>ATGGCAACCTTCACATTTACAAAAGAAACAGTTCAACCATTGTTACCTCTAA  GAAGAGCTTACATAATCTTCCACTTCACATGTGCTTGTCTTCTTTACTAC  CGTATCAGCAATTTGTTTATTTTCATATCCATGGTTTCTAATGACAATAGCTG  AGATTATTCTATCATTTCTATGGTTTTTCAACCAAGCATTCCGTTGGAGGCTG  GTGAATCGTTCAGTTATGACCGAGAAATTACCGCCGGAGGAGAAGTTGCCG  GGACTCGACATATTTGTGTGTACCATTGATCCTGAAAAAGAACCAACGGTT  GATGTTATGAACACTGTTATTTCTGCTATTGCAATGGATTACCCTTCTAATA  AACTTTCTATTTATCTTTCTGATGATGGAGGTTCTCCTATTACTCTTTTTGGG  ATCAAAGAGGCTTTTGAATTTGCTAAAGTTTGGGTTCCTTTTTGTAAAAAAT  ATGATGTTAAGTCAAGGTGTCCTAAGTTTTTCTTCACTGCTTTGGGTGAGAA  TGAACGACTTCATCGACCTCGTGAATTTGAAGAAGTGAGGGACCAGATTA  GAAGAGATTAATAGAGTGGATCCTAACTCACAAGTTGAAAACCTCAA  GGAACATATGCCACCAAGCCAAATACGAGAAAATGCAGAAAAAATATTG  AGAAATTCGGAAGCAACCTAAAGAATCTTTGTATGGTGACCGATAGACCTT  CTCGGATCGAGATCATTAAATGACCAAAAAGAAATGCCACTAGTTGTTTATG  TATCTCGTGAAAAAAGACCATCTGTTCTCACAGATTCAAAGGAGGAGCTC  TCAATACATTGCTTAGGGTGTGAGGGCTAATCAGCAATGGACCTTATGTACT  TGTCGTAGATTGTGATATGAATTGTAATGATGCATCATCAGCCAAACAATCC  ATGTGCTTTTTTCTGATCCTGAAACCTCTAAAGATGTTGCTTTTGTCAATT  CCCTCAAATGTTTCAACCTTAGCAAGAAAGACATATATGATAGTCAGAC  TAGGACTGCTTTTACGACAAAGTGGAAGGGAATGGATGGATTAAGAGGTCC  AGGTCTAACTGGCAGTGGAATTAATAAGTAGAAGTGCATTACTCTTTGG  AAGTCAAACCAAAAAGGGGACTATCTACTTGTGCTCTATAACAACCTTTGG  CAAGTCTAACATGTATGTAGAATCACTAAAAGCGTTACGTGGTCAACAAAC  TAAGAAGCAGAATATTTCAAGAGATGTAATTTTACAAGAAGCATGTGAAGT  GGCTTCTTGTTCCTATGAGAGAAACACAAATGGGGTAATGAGGTGGGATT  CTCGTATGCTATAAAACTTGAGAGTACCGTFACTGGCTATCTCCTCCATTGT  AGAGGATGGAGATCAACTTATCTTTACCCTAAAAGACCATGTTTCTTAGGAT  GTGCTCCAACCTGACATGAAAGAGGGATTGATTCAACCGATAAAGTGGTCAT  CTGAACCTTTGTTGCTTGAATCTCTAAATATAGCCATTCACTTATGGCCTT  TCAAGATTGCCACTATTCATGTTTAACTTTTTGTTACTTGGTAAGCACAAAC  CCAATTTGCAACAGCCTACATCTTATATGGATTCCGTTCTCAGATTTGCTTCT</p>	30 40

【表 17 - 4】

		TGAAGGGAATACCTGTATATCCAAAGGTTACAGATCCTTGGTTTATAGTGT TACAGTATTGTATCTATCCAGTCAAATTCATCATTATATTGAGGTAATTTCA ACTGGTGGCTCCTCGATGATTTGGTGGGAATGAACAAAGAAGGGGATTGTA AAATCAATTGGGTGCGTTTTTCGCAATTATAGAAACAGCGAAAAAAGTTT GGGTTGAACAAGGCAAATTCACCTTTATCGGACAAAGCAATTGACAAAGAT AAGCTAAAGAAATATGAGCAGGGTAAGTTTAAATTTTGTGGTGCAGCATTG CTCATGGCACCAGTGATTGTGTTACTCACAATAAATATTGTTTGCTCTTTGG TGGTTTATGGAGACTACTCAATGTGAGGGATTTTGATGAAATGTTTGGTCAA CTTTTCCTCATTATCTATATACTTGTCTAAGTCATCCTATTGTGGAGGGGAT TATATCTATGAAGCGGAAGAGTGGGTAG
ダイズ 由来の GmCSLG タンパク質_ (NM_0013 65113.1_so ja)	100	MATFHTETVQSGLALSRLHILFHSVALLFLYYYRISHILLEPSFVWIFMTI AELIFGELWLFKQAFWRPVSRAVMPEKLPDGLPALDIFVCTVDPEK EPTVQVMDTVISAIAMDYPSNKLAVYLSDDGGCPVTLYGIREASRFAKE WVPFCRKYGINSRCPKAFFSPMGEDERELLLRNHEFLAEQEQLKAKYN IMQKNIDEFGRDPKNRSIVFDRPARIEIINEQSEIPLVVYVSRERRPNVPH YKGGALNTLLRVSGLFSNGPYVLVVDCEMYCNDPSSAKQAMCFLLDPE TSKDIAFVQFPQMFHNL SMKDIYDSQHRHAFTTMWQGM DGLRGPGLS GSGNYLSRSALIFPSPYEKDG YEHNAQNKFGNSTMYIESLKAIQGQQT KTSISRNVILQEAQAVASCSYEIDTNWGNEVGFSYVILLESTVTGYLLHC RGWRSTYLYPKRPCFLGCAPTDFMEGMLQLVKWSSEFLLGISKYSPFT YGISRIPILHNFTFCYFTSTCQYIVALIVYGIIPQVCFKGTVPFVKVTEPW FVVFAILYVSSQSQHLIEVLYGGGSLGTWWDEQRIWIVKSIVGGIFGSILA IKKRFGLNKAKFILSNKVVAKKFEKYEQGFEFEDAALFMSPLVGLLIV NILCFFGGLWRLFVNDFEKMSGQLFLGLYLAALSYPIFEGIITMKSQVQ
ダイズ 由来の GmCSLG 遺伝子_ (NM_0013 65113.1_so ja)	101	TATATGCATGTTGACCGGTAACATGGCGACCTTCCACACAGAAACC GTGCAATCAGGGTTGGCCTTGAGCAGACTCCACATCCTATTCCACTC GGTGGCACTCTTGTCTTCTATTACTACCGCATAAGCCACATCTTACT GGAACCAAGCTTTGTATGGATTTTCATGACCATAGCGGAGCTTATCT TCGGCGAGCTCTGGCTCTTCAAACAGGCGTTCGGGTGGCGGCCCGTG TCGAGGGCCGTCATGCCGAGAAGCTGCCGAGCGACGGCAAGCTTC CGGCGCTCGACATCTTCGTCTGCACGGTTGACCCCGAAAAGGAGCCG ACGGTGCAGGTGATGGACACCGTCATCTCCGCCATTGCCATGGACTA CCCCTCCAACAAGCTCGCCGTGTACCTTTCGACGATGGCGGGTGTCT CGGTGACTCTGTATGGGATCAGAGAGGCTTCTCGGTTTCGCAAAGGAG TGGGTTCCGTTCTGCAGAAAGTATGGGATCAATTCACGGTGCCCCAA GGCCTTCTTCTCTCCCATGGGGGAGGATGAACGTGAAGTCTTCTTCT TCGCAACCATGAATTTCTTGGCAGAGCAAGAACA ACTCAAGGCTAAA TACAATATAATGCAAAAAAATATTGACGAATTTGGAAGAGACCCTA AAAATCGTTCCATTGTGTTTGATAGACCAGCTCGCATTGAGATTATA AATGAGCAATCCGAAATACCACTGGTTGTTTATGTGTCTCGTGAAAG AAGGCCAAATGTTCCCTACATACATAAAGGGGGAGCCCTCAACACA TTGCTCAGAGTCTCAGGGCTATTCAAGTAAACGGGCCCTATGTAATTGT AGTTGATTGTGATATGTATTGCAATGATCCATCATCAGCTAAACAAG CCATGTGCTTTTTTCTTGATCCTGAAACCTCCAAAGATATTGCTTTTG TCCAATTCCCTCAAATGTTTCACAACCTTAGCATGAAAGACATCTAC GATAGTCAACATAGGCATGCTTTTACAACAATGTGGCAAGGAATGG ATGGACTAAGAGGTCCAGGTCTTTCTGGTAGTGGCAATTACTTAAGT AGAAGTGCATTAATCTTTCCAAGCCATATGAAAAAGACGGCTATGA ACATAATGCCCAAAACAAATTTGGCAACTCTACCATGTACATTGAAT CATTAAAGGCCATTCAAGGACAACAACTTATAAAACGAGCATTTC AGAAATGTGATTTTACAGGAAGCACAAGCAGTGGCCTCTTGTTCCTA TGAAATAGACACAAATTTGGGGTAATGAGGTAGGATTCTCATATGTTA TATTACTGGAGAGTACAGTTACTGGCTATCTTCTTCACTGTAGAGGA TGGAGATCAACTTACCTTTACCCCAAAAGACCTTGTTTCTTGGGATGT GCCCCACTGACTTCATGGAAGGCATGCTTCAGTTGGTGAATGGAG TTCTGAACTTTTCTTGCTAGGAATATCCAATAACAGCCCTTCACTTA TGGGATTTCAAGAATTCCTATTCTGCACAACCTTACCTTTTGCTACTT

10

20

30

40



【表 17 - 6】

		<p>TCTCCCTTGTACATTGAATCATTGAAGGCCATCCGTGGACAACAAAC  TACCAAAAAGAATATCTCAAGAGACGAAAGTTTACTAGAAGCTAAA  GTGGTGGCCTCTGCTTCTACGAGACAAACACAGAATGGGGCTCAGA  GGTTGGATTCTCATATGGCATCTTACTGGAGAGTACTATTACTGGTTA  CCTTTTGCAGTGCAGAGGATGGAAATCAGCTTATCTTTACCCAAAAA  CACCATGTTTCTTAGGGTGTGCCCCACTGACATTAAGAAGGCATG  CTTCAGTTGGTGAAGTGGTTGTCTGAGCTTIGCTTGTITGCTGICICT  AAGTACAGCCCTTTACATATGGGTTTTCAAGATTGCCATTATGCCT  ACCTTCACTTATTGTTTCTGGCAGCTTCATCCCTATATGCTATTGTCT  TCATCCTTATGGCATTGTACCTCAAGTGTGCTTCTTGAAGGAATCC  CTGTGTTTCCAAAGGCCACAGACCCTTGGTTTGCAGTGTTCAGTAT  TGTATGTAGCCACCCAGATTCAACATTTGATTGAAGTCCTTTCTGGCA  ATGGCTCGGTCTCGATGTGGTGGGATGAACAAAGAATTTGGATTCTG  AAGTCAGTTACTAGCGTATTTGCAATGATAGAGGGAATCAAGAAAT  GGTTAGGATTGAACAAGAAAAAATCAACCTGTCAAACAAAGCGGT  TGACAAGGAGAAGGTCAAGAAATATGAGCAAGGTAGGTTTGATTTC  CAAGGAGCAGCTCTGTACATGTCTCCAATGGTTGTGTTGCTCCTAGT  GAACATTGTTTGCTTCTTTGGCGGTTTATGGAGACTGTTTAAGGAGA  AAGATTTTGCAGATATGTTTGGTCAACTTTTCTACTCAGCTATGTGA  TGGCTCTCAGTTATCCCATCTTGAGGGGATAGTAACTATGAAAATG  AAGAGTGGGTAG</p>
--	--	---

10

20

## 【 1 4 9 8 】

ハウレンソウ S O A P 5 の場合と同様にビートルート C S L G ( B v S O A P 5 ) をサイレンシングしたところ、サポニンのレベルが低下し、オレアノール酸の蓄積が上昇した ( 図 4 4 A ~ 4 4 C ) 。同様に、アルファルファ毛状根中において C L S G の発現を抑制したところ、グルクロン酸が結合したサポニンの蓄積が低下した ( 図 4 5 A 及び 4 5 B )

## 【 1 4 9 9 】

要約：これらの結果は、サポニン生合成経路内の酵素の異種機能性を示すものである。

## 【 1 5 0 0 】

30

実施例 2 3 : カンゾウ由来のセルロース合成酵素様 G を用いたグリチルリチンの異種産生

## 【 1 5 0 1 】

目的：C S L G グルクロノシルトランスフェラーゼ活性を見出すことにより、数々のグルクロニド型トリテルペノイド系サポニンの生合成経路における大きな知見の不足が埋められる。現在までに部分的にしか解読されていないこのような経路の 1 つでは、カンゾウで産生され、C - 3 位に 2 つのグルクロン酸部分で修飾されたグリチルレチン酸からなるトリテルペノイド系サポニンである、グリチルリチンを産生する。グリチルリチン ( すなわち G L ) 及びその単一のグルクロン酸抱合誘導体であるグリチルレチン酸モノグルクロニド ( すなわち G A M G ) は、伝統的な中国及び日本の医学で長く使用され、現在では食品、化粧品及び医薬品産業で使用されている重要な天然物である。これらの分子は血糖指数がゼロの低カロリー甘味料で、G A M G はスクロースよりも 9 4 1 倍甘く、G L よりも 5 倍甘い ( 「K. Mizutani, T. Kuramoto, Y. Tamura, N. Ohtake, S. Doi, M. Nakaura, O. Tanaka, Sweetness of glycyrrhetic acid 3-O-beta-D-monoglucuronide and the related glycosides. Biosci Biotechnol Biochem. 58, 554-5 (1994)」 ) 。 G L 及び G A M G の商業生産は、現在では入手が限られている野生のウラルカンゾウと、非効率的な工業的処理に依拠している。最初のグルクロン酸のアグリコンへの結合を触媒する酵素が同定されていないため、異種系における G L 及び G A M G の持続可能な産生は、その重要性にも関わらず、現在のところ不可能である ( 「Y. Nomura et al., Functional specialization of UDP-glycosyltransferase 73P12 in licorice to produce a sweet triterpe

40

50

noid saponin, glycyrrhizin. Plant J. doi: 10.1111/tpj.14409 (2019)」。そこで、異種系においてグリチルリチンを産生することが目標となった。

## 【1502】

方法：上記の材料及び方法を参照されたい。

## 【1503】

結果：サポニンのうちのいくつか、特にマメ科の植物によって産生されるものは、それらの特有の特性のため、重要性が高い。C - 3位の2つのグルクロン酸部分によって修飾されたグリチルレチン酸からなるトリテルペノイドであるグリチルリチンはウラルカンゾウによって産生され、ノンカロリー甘味料及び薬剤として多くのアジア諸国で広く使用されている。グリチルリチン産生に關与する5つの酵素のうち4つは、既に特性が解明されている (b A S、C Y P 8 8 D 6 及び C Y P 7 2 A 1 5 4、U G T 7 3 P 1 2) が、最初のグルクロン酸をアグリコンに結合する役割を果たす酵素は不明であった (「Y. Nomura et al., Functional specialization of UDP-glycosyltransferase 73P12 in licorice to produce a sweet triterpenoid saponin, glycyrrhizin. Plant J. doi: 10.1111/tpj.14409 (2019).」)。ここで疑問となるのは、カンゾウ由来のセルロース合成酵素様 G (G u C S L) が、この反応を行うことができるか否かということであった。

10

## 【1504】

ウラルカンゾウ由来のセルロース合成酵素を同定し、異種系における発現に用いた。発現された遺伝子及びコードされたポリペプチドのヌクレオチド配列及びアミノ酸配列を、下記の表 18 に示す。

20

## 【1505】

表 18 : ウラルカンゾウにおけるトリテルペノイドグリチルリチン生合成経路の構成要素

## 【1506】

## 【表 18 - 1】

名称	酵素活性	SEQ ID NO:	ヌクレオチド配列/アミノ酸配列
<i>GuCYP88D6_GLYUR00056</i> IS00023451.1	$\beta$ -アミリン 11- オキシダーゼ	76	ATGGAAGTACATFGGGTTTGCATGTCCGCTGCCACTTTGTT GGTATGCTACATTTTTGGAAGCAAGTTTGTGAGGAATTTGA ATGGGTGGTATTATGATGTAAACTAAGAAGGAAAGAACA CCCACTACCCCCAGGTGACATGGGATGGCCTCTTATCGGCG ATCTATTGTCCTTCATCAAAGATTTCTCATCGGGTCACCCTG ATTCATTCATCAACAACCTTGTCTCAAATATGGACGAAGT GGTATCTACAAGACTCACTTGTGTTGGGAATCCAAGCATCAT TGTTTGYGAGCCTCAGATGTGTAGGCGAGTTCTCACTGATG ATGTGAACTTTAAGCTTGGTTATCCAAAATCTATCAAAGAG TTGGCACGATGTAGACCCATGATTGATGTCT CTAATGCGGAACATAGGCTTTTTTCGACGCCTCATTACTTCC CCAATCGTGGGTCACAAGGCGCTAGCAATGTACCTAGAGC GTCTTGAGGAAATTGTGATCAATTCGTTGGAAGAATTGTCC AGCATGAAGCACCCCGTTGAGCTCTTGAAGAGATGAAGA AGGTTTCCTTTAAAGCCATTGTCCACGTYTTTCATGGGCTCTT CCAATCAGGACATCATTAAAAAAATTGGAAGTTCGTTTACT GATTTGTACAATGGCATGTTCTCTATCCCCATTAACGTACCT

30

40

## 【1507】

【表 18 - 2】

			<p>GGTTTTACATTCCACAAAGCACTCGAGGCACGTAAGAAGCT  AGCCAAAATAGTTCAACCCGTTGTGGATGAAAGGCGGTTG  ATGATAGAAAATGGTCCACAAGAAGGGAGCCA  AAGAAAAGATCTTATTGATATTCTTTTGGAAAGTCAAAGATG  AGAATGGACGAAAATTGGAGGACGAGGATATTAGCGATTT  ATTAATAGGGCTTTTGTGCTGGCCATGAAAGTACAGCAA  CCAGTTTAATGTGGTCAATTACATATCTTACACAGCATCCC  CATATCTTGAAAAAGGCTAAGGAAGAGCAGGAAGAAATAA  CGAGGACAAGATTTTCCCTCGCAGAAACAATTAAGTCTTAAG  GAAATTAAGCAAATGGTTTATCTTTCTCAGGTAATTGATGA  AACTTTACGATGTGCCAATATTGCCTTTGCAACTTTTCGAG  AGGCAACTGCTGATGTGAACATCAATGGTTATATCATACCA  AAGGGATGGAGAGTGCTAATTTGGGCAAGAGCC  ATTCATATGGATTCTGAATATTACCCAAATCCAGAAGAATT  TAATCCATCGAGATGGGATGATTACAATGCCAAAGCAGGA  ACCTTCCTCCCTTTTGGAGCAGGAAGTAGACTTTGTCCTGG  AGCCGACTTGGCGAAACTTGAAATTTCCATATTTCTTCATT  ATTCCTCCTTAATTACAGGTTGGAGAGAATAAATCCAGAA  TGTCACGTTACCAGCTTACCAGTATCTAARCCCACAGACAA  TTGCTCGCTAAGGTGATAAAGGTCTCATGTGCTTAG</p>	10
<p>GuCYP88D6_  GLYUR00056  1S00023451.1</p>	<p>β-アミリン  11-  オキシダーゼ</p>	77	<p>MEVHWVCMsaatLLVCYIFGSKFVRNLNGWYYDVKLRRKE  HPLPPGDMGWPLIGDLLSFIKDFSSGHPDSFINNLVLYGRSGI  YKTHLFGNPSIIVXEPQMCRRVLTDDVNFGLGYPKSIKELARC  RPMIDVSNAEHRLFRRLITSPVGHKALAMYLERLEEIVINSLE  ELSSMKHPVELLKEMKKVSFKAIHVFMGSSNQDIKKIGSSF  TDLYNGMFSIPINVPGFTHKALEARKKLAKIVQPVDERRLM  IENGPQEGSQRKDLIDLILEVKDENGRKLEDEDISDLLIGLLFA  GHESTATSLMWSITYLTQHPHILKKAKEEQEETRFRSSQKQL  SLKEIKQMVYLSQVIDETLRCANIAFATFREATADVNINGYIIP  KGWRVLIWARAIHMDSEYYPNPEEFNPSRWDDYNAKAGTFL  PFGAGSRLCPGADLAKLEISIFLHYFLLNYRLERINPECHVTSLP  VSXPTDNCLAKVIKVS</p>	20
<p>GuCYP72A15  4_GLYUR000  890S0001907  1.1</p>	<p>11-  オキソ-β-  アミリン30-  オキシダーゼ</p>	78	<p>ATGGATGCATCTTCCACACCAGGGGCTATCTGGGTTGTTCT  GACAGTGATACTAGCTGCGATTCCCATATGGGCATGCCATA  TGGTGAACACGCTGTGGCTGAGGCCAAAGAGGTTGGAAAG  GCATCTCAGAGCTCAAGGTCTTCATGGTGACCCTTACAAGC  TCTCACTTGACAACCTCCAAGCAAACCTATATGCTCAAGTTG  CAACAAGAAGCACAATCAAATCCATTGGTCTCTCCAAAG  ATGATGCTGCACCACGAATCTTCTCCCTTGCCCATCAAAC  GTACACAAATATGGAAAGAACTCCTTTGCATGGGAAGGGA  CAGCACCAAAGGTGATCATCACAGACCCAGAGCAAATTA  GGAAGTCTTTAACAAGATTCAGGACTTCCCAACCAAAAT  TAAATCCCATCGCCAAGTATATTAGCATCGGTCTAATACAG  TATGAGGGTGACAAATGGGCCAAACATCGAAAGATTATCA  ATCCGGCATTCCACTTAGAAAAATTGAAAGGTATGCTGCCA  GCATTTTCTCATAGCTGCCATGAAATGATTAGCAAATGGAA  GGGGTTATTGTCATCAGATGGAACATGTGAGGTTGATGTTT  GGCCCTTCCTTCAAATCTCACTTGTGATGTAATTTCTAGG  ACGGCATTTCGGAAGCAGCTATGCAGAAGGAGCAAAAATAT  TTGAACTTTGAAAAGGCAGGGATATGCTTTGATGACAGCA  CGATACGCACGCATTCCATTATGGTGGCTTCTACCATCAAC  TACCAAAGGAGGATGAAGGAAATGAAAGAGGCATACGT  GATTCATTGAAGGTATCATTAGAAAACGAGAAAAAGCAT  TGAAGAGTGGCAAAAGCACCGATGACGACTTATTAGGCAT  ACTTTTGCAATCAAATCACATTGAAAATAAAGGAGATGAA  AACAGTAAGAGTGCTGGAATGACCACCCAAGAAGTAATGG  AGGAATGCAAACCTTTTTTACCTGGCAGGGCAAGAGNTTGA  AAATAAAGGAGATGAAAACAGTAAGAGTGCTGGAATGACC  ACCCAAGAAGTAATGGAGGAATGCAAACCTTTTTTACCTGGC  AGGGCAAGAGACCACCGCAGCTTTGCTGGCCTGGACAATG</p>	30 40

【表 18 - 3】

			GTGTTATTAGGCAAGCATCCTGAATGGCAAGCACGCGCAA GGCAGGAAGTTTTGCAAGTAACCATGATTTTATATGAGGTA CTCAGGCTGTACCCACCTGGGATTTACCTCACCCGAGCTCT TCGAAAGGATTTGAAACTTGAAAACCTTTTGCTACCTGCTG GAGTACAGGTTTCCGTACCAATACTTTTGATTACCATGAT GAAGGTATATGGGGCAATGATGCAAAGGAGTTCAATCCTG AAAGGTTTGCTGAAGGAATTGCAAAGGCAACAAAAGGCCA AGTTTGCTATTTCCCTTTTGGATGGGGTCTAGAATATGTGT TGGGCAAAACTTTGCCTTATTAGAAGCCAAGATTGTATTGT CATTGCTGCTGCAGAATTTCTCATTTGAGCTATCTCCGACTT ATGCACATGTTCTACCACGGTGCTTACTTTGCAGCCAAAA CATGGGGCACCCATCATTCTGCATAAACTGTAA	10
GuCYP72A15 4_GLYUR000 890S0001907 1.1	11- オキソ-β- アミリン30- オキシダーゼ	79	MDASSTPGAIWVVLTVILAAIPIWACHMVNTLWLRPKRLERH LRAQGLHGDOPYKLSLDNSKQTYMLKLQQAQSKSIGLSKDDA APRIFSLAHQTVHKYGKNSFAWEGTAPKVIITDPEQIKEVENKI QDFPKPLNPIAKYISIGLIQYEGDKWAKHRKIINPAFHLEKLG GMLPAFHSHEMISKWKGLLSSDGTCEVDVWPFLQNLTCDV ISRTAFGSSYAEGAKIFELLKRQYALMTARYARIPLWLLPS TKRRMKEIERGIRDSLEGIIRKREKALKSGKSTDDDLLGILLQ SNHIENKGDENSKSAGMTTQEVMEECKLFYLAGQEXENKGD ENSKSAGMTTQEVMEECKLFYLAGQETTAALLAWTMVLLGK HPEWQARARQEVLVQVTMILYEVLRLYPPGIYLTRALRKDLKL GNLLLPAGVQVSPILLIHHDEGIWGNDAKEFNPERFAEGIAK ATKQVQCYFPFGWGPRIQVGNFALLEAKIVLSLLLQNFSEFEL SPTYAHVPTTVLTLQPKHGAPIILHKL	20
GuCSL_GLY UR003152S00 037491.1  (SOAP5 ホモログ; セルロース 合成酵素様G)	セルロース 合成酵素様G (グルクロン酸 トランスフェラーゼ)	80	ATGGCAAGCTTCACCCTTCACACAGAAACCGTTCAGTCATG GCTACTCCTCAGCAGACTTCACATACTGCTGCACCTCGCAG TTGTACTGCTCCTCTTATACTACCGCATCACAGTTTCCCCT TCCATGCTCCSACTCTACCGTGGACTCTGATGACCGTAGGT GAGGCTATTATGGCAGTGCTGTGGTTCCTCAACCAGGCCTT CCGGTGGCGGCCGGTGAGCCGCTCGGTGATGACGGAGAAG CTGCCCAGCGACGCGAAGCTGCCGGGGCTTGACATATTCGT GTGCACGCTTGACCCCGAGAAGGAGCCACCGTGGAGGTG ATGAACACTCTGGTCTCTGCCCTTGCCATGGACTACCCCC TGACAAGCTCTCCGTTTACCTCTCCGACGATGGCGCCGCC CGGTCACTCTTTACGGCGTGAGAGAGGCTTCTGAGTTCGCG AGGGTGTGGGTCCCTTTCTGCAAAAAGTATGGGATCAAGTC AAGGTGTCCAAGGTTTTCTTCTCTCCAGTGCTGAGGATG AACACCTTCTTCGCACCGACGAGTTCAGGTGAGAGCGAGA CCTCATCAAGGCTAAATACGAGAAAATGCAGAAAAATATT GAGAAATTTGGTTCGGATGCCAAAAATTGTCGTATGGTGAC TGACAGACCTCCTCGGATCGAGATATTGATTGACCAACCAG ACATGCCACGTGTTGTTTACGTGTCTCGGGAAAGAAGGCCA TCACTCCCTCACAAGTTCAAAGGAGGAGCCCTCAATACATT GCTCAGAGTCTCAGGTCTAATCAGCAATGGGCCTTATGTAC TTGTAGTGGACTGTGATATGTATTGCAATGACCCATCCTCA GCCAAACAAGCCATGTGTTTCTTTCTTGATCCTGAAACCTC TAAATYTATTGCATTTGTCCAATTCCTCAAATGTTTCACAA CCTTGCAAAAAAGACATCTATGACAATCAATCTAGGACT GCTTTTAAGACAATGTGGCAAGGGATGGATGGACTAAGAG GTCCTGGTCTTTCTGGCAGCGGTAATTAATTGAATAGAAGT GCATTACTATTTGGAAGTCCAAATCAAAAAGATGACTATCT GGATGATGCCAAAACACTTRGGCAAGTCTACCATGTACA TAGAATCACTAAAGGCCATTCGTGGACAAAAAACTATGAA AAAGAATATTTCAAGAGATGAAATTTTACGAGAAGCTCAA GTATTAGCCTCTTGTTCCTATGAGACAAACACAGAATGGGG AGCAGAGGTAGGATTCTCATATGGCATCTTACTGGAGAGTT	30 40



【表 18 - 5】

			<p>CAGCTTCTACGAACTTGAACCGGTTTACGCGGATCATTACA  GGAACGAGCTTGGGAGAAGAGCTTGGCATTGGGTCCGGT  TTCATTATGCAGTAGGGACACGGAGGAAAAATCGCGGAGG  GGAAGGGAAGCTGCAATTGATGAGAACGAGTGCTTGAAGT  GGCTCAATCAAAGGAACCCAATTCGGTTGTTTATGTTTGT  TTCGGTAGCATGATGGTTTTTCAGTGACGCTCAGCTAAAAGA  GATTGCGATGGGTCTTGAGGCTTCAGGGAAGCCATTATAT  GGGTGGTGAAGAAAGGAGGGGCTAAAAGTGAAGGTGAGA  AATTGGAGTGGCTTCCAGAAGGGTTTGAGGAGAGAATGGG  GGAAAGTAATAAGGGACTAATCATAAGGGGTTGGGCACCA  CAGGTGATGATTTTGACCATGGAGCGGTTGGAGGGTTTGT  GACACATTGTGGGTGGAATCAACGCTGGAAGGAGTGTGT  GCAGGGGTGCCAATGGTGACTTGGCCCATGTATGGGGAAC  AATTTTACAACGCCAAGTTTCTGACGGACATAGTAAAATT  GGGGTGGGTGTTGGGGTTCAAACGTGGATTGGGATGGGAG  GAGGAGAGCCTGTGAAGAAGGAAGTGATAGAGCAGGCAG  TGAGAAGGATAATGGTGGGGCAGGAAGCAGAGGAAATGA  GAAACAGAGCCAAGGAACTGAGCCAGATGGCAAAGCGTGC  TGTGGAGGAAGGAGGATCGTCTCACAAAGATTTAACTCTT  TAATTGAGGATTTGAGGTCGCGTGCCCATTA</p>	10
<p>GuUGAT_KT 759000.1</p>	<p>グリチルレチン酸  グルクロノシル  トランスフェラーゼ</p>	<p>83</p>	<p>MTMGNENRELHIIFFPFLANGHIIPCVDLARVFAARGIRATIVT  THLNVPIYSRTIGKANINIRTIKFPSTEDSGLPEGCENTESALAP  DKFIKFMKATLLLRDPLEHVLQEEQPHCLVADMFFPWATDSA  AKFGIPRIVFHGLGYFPLCVLACTRQYKPKQDKVSSYTEPFVVP  NLPGEITLTKMQLPQLPQHDKVFTQLLEESNESELKSFGVIVNS  FYELEPVYADHYRNELGRRAWHLGPVSLCSRDTEEKSRRGRE  AAIDENECLKWLQSKEPNSVVYVCFGSMVMFSDAQLKEIAM  GLEASGKPFIVVVKKGGAKSEGEKLEWLPEGFERMGESNK  GLIIRGWAPQVMILDHGA VGGFVTHCGWNSTLEGVCAGVPM  VTWPMYGEQFYNAKFLTDIVKIGVGVGVQTWIGMGGGEPVK  KEVIEQAVRRIMVQGAEEMRNRKELSQMAKRAVEEGSS  HNDFNSLIEDLRSRAH</p>	20
<p>GuUGT73P12 _SCAFFOLD 00629 (LC314779)</p>	<p>グリチルレチン  酸3-O-  モノグルクロニド  グルクロノシル  トランスフェラーゼ</p>	<p>84</p>	<p>ATGGACTCCTTTGGGGTTGAAGGTGATCACCAAGCCGACAC  CACAGTGCTGAAGGCGGTTTTTCTTCCCTTCATCTCAAAAA  GTCATCTCATCCGTGAGGTGGACAAAGCAAGGATCTTCGCC  ATGCACGGCGTGATGTCACCATCATCACCACGCCGCCA  ACGCTGCCACTTTCCAAACCTCCATTGACCGCGACTCCAGC  CGCGGCCGCTCCATCAGAACGCACATCGTTCCGTTCCCCCA  AGTCCCCGGTCTACCACAGGGACTCGAGAGACTCGACGCC  GACACTCCTCAACACTTGCTCTCCAAGATCTACCATGGACT  ATCCATTCTGCAAGAGCAGTTCCAACAAGTTCCTGAAA  TGAGGCCAGATTTTCATAGTCACTGACATGTACTACCTTGG  AGCGTGCATGCCGCCGAGTTGGGGATTCCGAGGTTGGT  TTGTAACGGTGGAAAGCTACTTCGCTCAGTCAGCTGTTAACT  CCGTTGAGCTATTTTACCACAAGCCAAGGTTGATTCAAAT  ACCGAGACTTTTCTGCTTCTGCGTTACCCATGAGGTTGA  GATGACACGTTTGAACACTACCGGATTGGCTTAGAGGAGCA  CCGAATGAGTACACCTATTTGATGAAGATGATCAAGGATTC  AGAGAGGAAGAGTTATGGGTGATTGTTCAATAGCTTTTATG  AGCTTGAAGGGACTTATGAGGAACATTACAAGAAAGCCAT  GGGAACCAAGAGTTGGAGTGTGGGGCCAGTTTCTTTGTGG  GTGAACCAAGATGCTTCTGATAAGGCTTGTAGGGGGGATG  TTAAAGAAGGAAAAGGAGATGGGGTGGTGCTTACTTGGCT  GGATTCTAAAACAGAGGACTCTGTTTTGTATGTGAGTTTTG  GGAGCATGAACAAGTTCCTAAAACACTCAGCTTGTGAGATA  GCTCATGCCCTCGAAGATTCTGGCCATGATTTCAATTTGGGT  CGTTGGCAAAATTGAAGAAGGTGAAGGTGGTGCTGATTTTT  TGAGGGAATTTGAGAAGAAAGTGAAGAAAAAACAGAG</p>	30
				40

【表 18 - 6】

			GTTATCTGATATGGGGTTGGGCACCACAGCTTCTGATTCTG GAGCATCCTGCGGTTGGAGCAGTGGTGACTCATTGTGGGGTG GAACACCGTTATGGAAAAGTGTGAATGCAAGTTTGCCATTGG CAACTGGCCATTGTTTGGCGGAGCAGTCTTCAATGAGAAG CTAGTGGTTGATGTGGTGAAGATTGGTGTGCCAGTTGGGGT TAAGGAATGGAGAAATTGGAATGAGTTTGGGGATGAGGTT GTGAAGAGGGAGGACATAGGAAAGGCCATTGCTTTTTTGA TGGGTGGTGGGGATGAATCCTTGGAAATGAGGAAGAGGGT CAAGGTGCTCAGTGGTGTACAAAAGAAAGCTATTGAGGTT GGTGGGTCTTCTCACACCAAGTTGAAAGAAGCTCATAGAAG AGCTCAAGTCAATCAAGCTACAAAAGGTCAACAACAATT AATGGAGGCAGTGGCTTAA	10
GuUGT73P12 _SCAFFOLD 00629 (LC314779)	グリチルレチン 酸3-0- モノグルクロニド グルクロノシル トランスフェラーゼ	85	MDSFGVEGDHQADTTVLKAVFLPFISKSHLIREVDKARI FAMHGVDVTIITTPANAATFQTSIDRDSRRGRSIRTHIVPF PQVPGLPQGLERLDADTPQHLLSKIYHGLSILQEQQQLF REMRPDFIVTDMYYPWSVDAEAELGIPRLVCNGGSYFA QSAVNSVELFSPQAKVDSNTETFLPLPHEVEMTRLQL PDWLRGAPNEYTYLMKMIKDSERKSYGSLFNSFYELEG TYEEHYKKAMGTKSWSVGPVSLWVNQDASDKACRGD VKEGKGDGVVLTWLDKTEDSVLYVSFGSMNKFPTQ LVEIAHALEDSGHDFIWWVGKIEEGEGGADFLREFEKKV KEKNRGYLIWGWAPQLLILEHPAVGAVVTHCGWNTVM ESVNASLPLATWPLFAEQFFNEKLVDVVKIGVPVGVK EWRNWNEFGDEVVKREDIGKAI AFLMGGGDESLEMRK RVKVLSGATKKAIQVGGSSHTKLKELIEELKSIKLQKVN NKLMEAVA	20
中国カンゾウ 由来の GuCSLG タンパク質  _(Glyur00315 2s00037491.1 _Glycyrrhiza ウラルカンゾウ)  (SOAP5 ホモログ; セルロース 合成酵素様G)	セルロース 合成酵素様G (グルクロン酸 トランスフェラーゼ)	102	MASFTLHTETVQSWLLLSRLHILLHLAVVLLLLYYRITR FPFHAPTLPWTLMTVGEAIMAVLWFFNQAFRWRPVSRS VMTEKLPDAKLPGLDIFVCTLDPEKEPTVEVMNTLVSA LAMDYPPDKLSVYLSDDGAAPVTLYGVREASEFARVW VPFCKKYGIKSRCPKVFSPSAEDEHLLRTDEFRRSERDLI KAKYEKMQKNIEKFGSDAKNCRMVTDPRPRIEILIDQPD MPRVVYVSRERRPSLPHKFKGGALNTLLRVSGLISNGPY VLVVDCEMYCNDPSSAKQAMCFDLPETSKSIAFVQFP QMFHNLGKKDIYDNQSRFAFKTMWQGMGDLRGPGLS GSGNYLNRSALLFGSPNQKDDYLDDAQNYLKGSTMYIE SLKAIRGQKTMKKNISRDEILREAQVLASCSYETNTEWG AEVGFYSYGILLESSITGYLFHCRGWKSAYLYPKTPCFLG CAPTDIKEGMLQLVKWLSEYCLLGFSKYSPTFYGFSRMP IMPTLVYCFLLTTTLYSIVFILYGIVPQVCFLKIPVFPKV TDPWFAVFATLYISTQIQHLIEVLSGDGVSAMWWDEQG IWILKSVTSVFALIEAAKKGLGLNKKKFMLSNKAIDKEK LKKYEQGRFDFQGAALFMSPMVLLIVNVVSFIGGIWR LFNAKDIEDMFGQLFLVSYVMALSYPFEGIITMKS KSG	30
中国カンゾウ 由来の GuCSLG タンパク質  _(Glyur00315 2s00037491.1 _Glycyrrhiza ウラルカンゾウ)	セルロース 合成酵素様G (グルクロン酸 トランスフェラーゼ)	103	ATGGCAAGCTTCACCCTTCACACAGAAACCGTTTCAGT CATGGCTACTCCTCAGCAGACTTCACATACTGCTGCA CCTCGCAGTTGTACTGCTCCTCTTATACTACCGCATCA CACGTTTTCCCTTCCATGCTCCGACTCTACCGTGGACT CTGATGACCGTAGGTGAGGCTATTATGGCAGTGCTGT GGTTCTTCAACCAGGCCTTCCGGTGGCGGCGCGGTGAG CCGCTCGGTGATGACGGAGAAGCTGCCAGCGACGC GAAGCTGCCGGGGCTTGACATATTCGTGTGCACGCTT GACCCCGAGAAGGAGCCCACCGTGGAGGTGATGAAC ACTCTGGTCTCTGCCCTTGCCATGGACTACCCCTTGA CAAGCTCTCCGTTTACCTCTCCGACGATGGCGCGCC	40

【表 18 - 7】

(SOAP5 ホモログ; セルロース 合成酵素様G)		<pre> CCGGTCACTCTTTACGGCGTGAGAGAGGCTTCTGAGT TCGCGAGGGTGTGGGTCCCTTTCTGCAAAAAGTATGG GATCAAGTCAAGGTGTCCCAAGGTTTTCTTCTCTCCCA GTGCTGAGGATGAACACCTTCTTCGCACCGACGAGTT CAGGTCAGAGCGAGACCTCATCAAGGCTAAATACGA GAAAATGCAGAAAAATATTGAGAAATTTGTTTCGGAT GCCAAAAATTGTCGTATGGTGACTGACAGACCTCCTC GGATCGAGATATTGATTGACCAACCAGACATGCCACG TGTGTTTACGTGTCTCGGAAAGAAGGCCATCACTC CCTACAAGTTCAAAGGAGGAGCCCTCAATACATTGC TCAGAGTCTCAGGTCTAATCAGCAATGGGCCTTATGT ACTTGTAGTGGACTGTGATATGTATTGCAATGACCCA TCCTCAGCCAAACAAGCCATGTGTTTCTTTCTTGATCC TGAAACCTCTAAATCTATTGCATTTGTCCAATCCCTC AAATGTTTACAACCTTGGCAAAAAAGACATCTATGA CAATCAATCTAGGACTGCTTTTAAGACAATGTGGCAA GGGATGGATGGACTAAGAGGTCCTGGTCTTTCTGGCA GCGGTAATTACTTGAATAGAAGTGCATTACTATTTGG AAGTCCAATCAAAAAGATGACTATCTGGATGATGCC CAAACTACTTAGGCAAGTCTACCATGTACATAGAAT CACTAAAGGCCATTTCGTGGACAAAAACTATGAAAA AGAATATTTCAAGAGATGAAATTTTACGAGAAGCTCA AGTATTAGCCTCTTGTTCTATGAGACAAACACAGAA TGGGGAGCAGAGGTAGGATTCTCATATGGCATCTTAC TGGAGAGTTCAATCACTGGCTATCTTTCCACTGCAG AGGATGGAAATCAGCATATCTTTACCCAAAGACACCA TGTTTCTTAGGGTGTGCCCCAACTGACATCAAGGAAG GAATGCTCCAATTGGTGAAGTGGTTGTCTGAATACTG CTTGCTAGGATTCTCTAAATACAGCCCTTTCACCTTATG GCTTTTCAAGAATGCCATTATGCCTACCTTAGTCTAT TGCTTCTTGACAACAACAACCCTTTATTCCATTGTCTT CATCCTTTATGGCATTGTCCCCAAGTTTGCTTCTTAA AAGGAATACCCGTGTTTCCAAAGGTCACAGACCCTTG GTTTGCAGTGTGTTGCAACACTGTATATCCACCCAG ATTCAACATTTGATAGAGGTCCTTTCTGGTGATGGCTC TGTGGCAATGTGGTGGGATGAACAGGGAATCTGGATT CTGAAGTCAGTCACTAGCGTGTTCGCAATCATAGAGG CAGCTAAGAAAGGGTTAGGATTGAACAAGAAGAAAT TCATGTTGTCAAACAAGCAATTGACAAGGAGAAGCT CAAGAAGTATGAGCAAGGTAGGTTTGATTTCCAAGGT GCAGCTCTGTTTCATGTCCCCAATGGTTGTGTTGTCAT AGTGAACGTTGTTTCCTTTCATTGGTGGCATATGGAGA CTATTCAATGCAAAGGATATTGAAGATATGTTTGGTC AGCTTTTCCTAGTTAGTTATGTAATGGCCCTTAGTTAT CCCATTTTGAAGGGATAATAACCATGAAAAGCAAGA GTGGATAG </pre>	<p>10</p> <p>20</p> <p>30</p> <p>40</p>
-------------------------------------	--	--	---

【1513】

ベンサミアナタバコにおいてハウレンソウ b A S ( S o S O A P 1 ; 配列番号 4 5 )、GuCYP88D6 ( 配列番号 7 6 )、GuCYP72A154 ( 配列番号 7 8 )、及びGuCSL ( 配列番号 8 0 ) を発現させたところ、グリチルレチン酸 3 - O - モノグルクロニド ( G A - 3 - G l c A ) が産生した ( 図 4 6 A ~ 4 6 E )。一方、グリチルレチン酸の連続する 2 段階のグルクロン酸抱合を触媒し、グリチルリチンを産生するものと考えられている既報の GuUGAT ( 配列番号 8 2 ) は、今回行ったアッセイにおいて、G l c A をアグリコンに結合することはできなかった ( 「G. Xu, W. Cai, W. Gao, C Liu, A novel glucuronosyltransferase has an unprecedented ability to catalyse continuou

s two-step glucuronosylation of glycyrrhetic acid to yield glycyrrhizin. *New Phytol.* 212, 123-35 (2016).」)。

【1514】

G u U G T 7 3 P 1 2 (配列番号84)は、2番目のG l c Aを結合させることが以前示されていたが、これをG u C S L及び他の酵素と共に一過性発現させると、ベンサミアナタバコにおける最終産物としてグリチルリチンが産生された(図47A~47F)。これらのアッセイによると、G u C S L (配列番号81)は、2番目のグルクロン酸部分をG A - 3 - G l c Aに付加することができるが、効率が非常に低いことも示された。

【1515】

さらに、異種系(ベンサミアナタバコ)において、キラヤ(*Q. saponaria*)由来のQ S - 21の生合成における最初のグリコシル化ステップは、グルクロン酸をC - 3位のキラヤ酸に結合させるC S L Gによって行われ得ることが示されるであろう。

10

【1516】

要約:セルロース合成酵素様G酵素によって行われる新たな形態の生合成反応が見いだされたことにより、異種系(植物または真菌)において、重要な特化した代謝物の経路と、それらの持続可能な産生とを操作する上で有益な戦略がもたらされる。

【1517】

実施例24:セルロース合成酵素様Gの構造解析

【1518】

目的:セルロース合成酵素様G(C S L G)の構造を解析し、古典的なセルロース合成酵素A(C E S A)と比較すること。

20

【1519】

方法:上記の材料及び方法を参照されたい。

【1520】

結果:系統発生解析によると、C S L Gは、セルロース生合成を媒介する古典的なセルロース合成酵素A(C E S A)と進化的に異なることが示された(約30%のアミノ酸配列類似性)。しかしながら、これら2つのタンパク質群には、いくつかの共通の構造的特徴がある。ハウレンソウS O A P 5 C S L Gタンパク質をテンプレートベースで三次構造をモデル化して得られた予測3Dモデルによると、それは、アミノ末端の脂質二重層(T M D)と、カルボキシ末端の別の4つのT M Dとを横断する、2つのドメインからなることが示された(「M. Kallberg et al., Template-based protein structure modeling using the RaptorX web server. *Nature Protocols.* 7, 1511-1522 (2012).」)。脂質二重層から突出した445アミノ酸長のループが、膜貫通ドメイン間に位置している(図48、図49A)。ハウレンソウS O A P 5酵素、シロイヌナズナ属由来のセルロース合成酵素様G酵素、並びにシロイヌナズナ属由来のセルロース合成酵素Aサブユニット及びハウレンソウ中のそれらのオルソログを、以下の表19に示す。また、A、B、及びEファミリー由来の他のセルロース合成酵素様遺伝子を示す。現在の理解では、A、B、及びEファミリー由来のセルロース合成酵素様遺伝子は、それぞれ全く異なる活性を有する。本明細書に示した研究までは、数少ないC S Lが知られていただけで、C S L Gファミリー由来のC S Lは皆無であったが、今回の研究において、C S Lの特性が解明された。解析したものは、-グルカン及びヘミセルロースの糖代謝、生合成に関与している。いずれのC S Lも(それが属するファミリーであるにも関わらず)、トリテルペノイドに作用できることは、全く示されていなかった。したがって、実施例を通して示した結果は、セルロース合成酵素様酵素について、驚くべき予想外の活性をもたらすものであった。

30

40

【1521】

表19:シロイヌナズナ属C S L G酵素及びC E S A酵素のアミノ酸配列及びヌクレオチド配列

【1522】

【表 19 - 1】

名称	酵素活性	SEQ ID NO:	AA SEQ
シロイヌナズナ属 CESA1	セルロース 合成酵素A (サブユニット1)	67	MEASAGLVAGSYRRNELVRIRHESDGGTKPLKNMN GQICQICGDDVGLAETGDVVFVACNECAFPVCRPCYE YERKDGTOCCPQCKTRFRRHRGSPRVEGDEDEDDV DDIENEFNYAQGANRHRHQRHGEEFSSSRHESQPIP LLTHGHTVSGEIRTPDTQSVRTTSGPLGPSDRNAISSP YIDPRQVPVVRIVDPSKDLNSYGLGNVDWKERVEGW KLKQEKMLQMTGKYHEGKGGEEIEGTGSNGEELQM ADDTRLPMSRVVPISSRLTPYRVVILRLILCFFLQY RTTHPVKNAYPLWLTSVICEIWFVAFSWLLDQFPKWY PINRETYLDRLAIRYDRDGEPSQLVPVDVFVSTVDPL KEPPLVTANTVLSILSVDYPVDKVACYVSDDGSAML TFESLSETAEFAKKWVPFCKKFNIEPRAPEFYFAQKID YLKDKIQPSFVKERRAMKREYEEFKVRINALVAKAQ KIPEEGWTMQDGPWPGNNTRDHPGMIQVFLGHSG GLDTDGNELPRLIYVSREKRPGFQHHKKAGAMNALI RVSAVLTNGAYLLNVDCDHVFNNKAIKEAMCFMM DPAIGKCCYVQFPQRFQDGLDHRVANRNIVFFDIN MKGLDGIQGPVYVGTGCCFNRQALYGYDPVLTEED LEPNIIKSCCGSRKKGKSSKKYNYEKRRGINRSDSN APLNFMEDIDEGFEGYDDERSILMSQRSVEKRFQSP VFIAATFMEQGGIPPTTNPATLLKEAIHVISCYEDKT EWGKEIGWIYGSVTEDILTGFKMHARGWISYCNPPR PAFKGSAPINLSDRLNQVLRWALGSIEILLSRHCPWIY GYHGRLRLLERIA YINTIVYPITSIPLIAYCILPAFLIT DRFIPEISNYASIWFILLFISIAVTGILELRWSGVSIED WWRNEQFWVIGGTAHLFAVFQGLLKVLAGIDTNTFT VTSKATDEDGDFAELEYFKWTALLIPPTTVLLVNLIGI VAGVSYAVNSGYQSWGPLFGKLFALWVIAHLYPFL KGLLGRQNRTPPTIVVWSVLLASIFSLWVRINPFVDA NPNANNFNGKGGVF
シロイヌナズナ属 CESA3	セルロース 合成酵素A (サブユニット3)	68	MESEGETAGKPMKNIVPQTCQICSDNVGKTVDGDRF VACDICSFPVCRPCYEYERKDGNOQSCPQCKTRYKRL KGSPAIPGDKDEDGLADEGTVEFNYPQKEKISERML GWHLTRGKGEEMGEPQYDKEVSHNHLPRLTSRQDT SGEFSAAASPERLSVSSTIAGGKRLPYSSDVNQSPNRI VDPVGLGNVAWKERVDGWKMKQEKNTGPVSTQAA SERGGVDIDASTDILADEALLNDEARQPLSRKVSIPSS RINPYRMVIMLRLVILCLFLHYRITNPVNAFALWL SVICEIWFALSWILDQFPKWFPVNRETYLDRLALRYD REGEPSQLAAVDIFVSTVDPLKEPPLVTANTVLSILAV DYPVDKVSCYVFDDGAAMLSFESLAETSEFARKWVP FCKKYSIEPRAPEWYFAAKIDYLDKQVQTSFVKDRR AMKREYEEFKIRINALVSKALCPEEGWVMQDGT WPGNNTGDHPGMIQVFLGQNGGLDAEGNELPRLVY VSREKRPGFQHHKKAGAMNALRVSAVLTNGPFILN

10

20

30

40

【 1 5 2 3 】

【表 19 - 2】

			LDCDHYINNSKALREAMCFLMDPNLKGQVCYVQFP QRFDGIDKNDRYANRNTVFFDINLRGLDGIQGPVYV GTGCVFNRTALYGYEPPKVKHKKPSLLSKLGGSR KKNSKAKKESDKKSGRHTDSTVPVFNLDIEEGVE GAGFDDEKALLMSQMSLEKRFQSAVFASTLMEN GGVPPSATPENLLKEAIHVISCYEDKSDWGMIEIGWI YGSVTEDILTGFKMHARGWRSIYCMPLPAFKGSAPI NLSDRLNQVLRWALGSVEILFSRHCPIWYGYNGRLK FLERFAYVNTTIYPITSIPLLMYCTLLAVCLFTNQFIIP QISNIASIWFLSLFLSIFATGILEMR WSGVGIDEWWRN EQFWVIGGVS AHLFAVFQGILKVLGIDTNTFTVTSKA SDEDGDFAEYLKFWTLLIPPTLLIVNLVGVVAGV SYAINSGYQSWGPLFGKLFPAFWVIVHLYPFLKGLM GRQNRTPTIVVVWSVLLASIFSLWVRIDPFTSRVTGP DILECGINC	10
シロイヌナズナ属 CSLG1	セルロース 合成酵素様G	69	METHRKNSVVGNILHTCHPCRRTIPYRIYAIFHTCGII ALMYHHVHSLVTANNTLITCLLLSDIVLAFMWATT TSLRLNPVHRTECEPEKYAAKPEDFPKLDVFICTADPY KEPPMMVVNTALSVMAYEYPSDKISVYVSDDGGSSL TFFALIEAAKFSKQWLPFCCKNNVQDRSPEVYFSSSES HSRSDAENLKMMEYEDMKSRVEHVESGKVVETAFI TCDQFRGVFDLWTDKFSRHDHPTIIQVLQNSSETDMD NTRKYIMPNIYVSREKSKVSPHHFKAGALNTLLRVS GVMTNSPIILTLDCDMYSNDPATLVRALCYLTDPEIK SGLGYVQFPQKFLGISKNDIYACENKRLFIINMVGFD GLMGPTHVGTGCFNRRAFYGGPYMLILPEINELKPY RIADKSIKAQDVLSLAHNVAGCIYEYNTNWGSKIGFR YGSLVEDYTGFMHCEGWRSVFCNPKKAAFYGD PKCLVDLVGQQIRWAVGLFEMSFYSKYPITYGKSLD LLMGLGYCNSPFKPFWSIPLTVYGLLPQALISGVS FPKASDPWFVLYIILFFGAYAQDLSDFLLEGGTYRK WWNDQRMLMIKGLSSFFFGFIEFILKTLNLSTPKFNV TSKANDDDEQRKRYEQEIFDFGTSSSMFLPLTTVAIV NLLAFVWGLYGILFCGGELYELMLVSFAVNCPLPI YGAMVLRKDDGKLSKRTCFLAGNLHVGSYCVKLLR PQVTSPLRLIHNNTSGWFKRKKHNMNESV	20 30
ハウレンソウ CESA1	セルロース 合成酵素A 触媒 サブユニット1 [UDP産生]	70	MEATGGMVAGSYKRNELVRIRHDSTDSGSKSLKNL DGQICQICGDTVGVTSNGGVFVACNECAFVCRPCY EYERKDGNOCCPCKTRYKRQKGLRVEGDDEED VDDLNEFN YERGTSKARHQWQGEDVDLSSSRHG SQIPLLTNGQVVS GEIPSATPDNQS VRSTSGPIGPEKR GNHSLPYIDPCLPVPVRIVDPSKDLNSYGLGSVDWKE RVESWKLKQEKNMHTHTGNRYSEGKGGDVEGSGSNG EELQLADDVRQPM SRIVPISSHLTPYRAVIIFRLIILVF FLQFRITHPVEDAYPLWLTSVCEIWFAMSWILDQFP KWYPINRETYLDRLAFRHDREGEPSQLAPIDVFVSTV DPLKEPPIITANTVLSILAVDYPVDK VSCYVSDDGSA MLTFEGLSETAEFARKWVPFCKKFSIEPRAPEFYFQQ KIDYLDKDIQPSFVKERRAMKREYEEFKVRINALVA KAQKVPEEGWTMQDGTAWPGNNPRDHPGMIQVFL GHSGLDMDGNELPRLVYVSREKRPGFQHKKAGA MNALIRVSAVLTNGAYILNVDCDHYFNNSKCLKEA MCFMMDPALGKVCYVQFPQRFDGIDLHRYANRN IVFFDINMKGQDGIQGPVYVGTGCCFNQALYGYDP VLTEEDFEPNFIKNCFGSRKKGKSGNKKYMDKKRG PKRSESSIPIFNMEDIEEGVEGYEDEKSLMSQKRLEK	40

【表 19 - 3】

			RFGQSPVFAATFMEMGGIPPTTNPATLLKEAIHVISC GYEDKSEWGKEIGWIYGSVTEIDLTGFKMHARGWM SIYCMPPRPAFKGSAPLNLSDRLNQVLRWALGSIEIM LSRHCPIWYGKGRRLRFLERLAYINTVVYPLTSIPLIA YCILPAICLLTNKFIPTLSNFASILFIMLFMSIAATGILE LRWSGVSIEDWWRNEQFWVIGGTS AHLFAVFQGLL KVLGIDTNFTVTSKAADEDGDFAEYIFKWTALLIP PTTVLIVNLVGVVAGVSYAINSGYQSWGPLFGKLFSS FWVIAHLYPFLKGLLGRQNRTPVIVVWSVLLASIFSL LWVRINPFTTDAEKAAAGNQCING	10
ホウレンソウ CESA3	セルロース 合成酵素A 触媒 サブユニット3 [UDP産生]	71	MMEDSQSGVKPTKQANEQVCQICSDNIGTTVDGEPF VACDVCSFPVCRACYEYERKDGTSQCPQCKTRYKR QKGS PAIHGEKVEDSDVEDVSDVNEPLGSSILKEKP QERMLGWHMNHGQSGELGPPTYDKEAPISHIPRLAT GRTVSGDLSAASPRFSMPSPGASTGANIRVSREFAS PGFGNVAWKERIDGWKMKQEKSTGPPSVSHAPSEG RFANDIDASTEIAMDDPLLNDETRQPLSRKVPISSRI NPYRMVIVLRLAVLGIFLHYRVTNPVFNAYALWLIS VICEIWF AFSWILDQFPKWLPINRETYLDRALR YDR EGEPSQLAAVDIFVSTVDPLKEPPLVTANTVLSILAVD YPVDKVSCYVSDDGAAMLTFEALSETSEFARKWVPF TKKYNIEPRAPEWYFSQKIDYLDKDVQTTFFVKDRRA MKREYEEFKIRINGLVAKATKVPEEGWVMQDGPW PGNNTRDHPGMIQVFLGQSGGLDTDGNELPRLVYVS REKRPGFTHHKKAGAMNSLVRVSAVLTNGPFMLNL DCDHYINNSKALREAMCFMMDPNLKYCCYVQFPQ RFDGIDRNDRYANRNTVFFDINLRGLDGIQGPVYVG TGCVFNRALYGYEPPKPKPKKGGILSSCFGGSRKK SSKKDSKKKSKHADPTVPIFNLEDIEEGVEGTGFDDE KSLMSQISLEKRFKSEVFASTLMENGGVPQSATP DTLLKEAIHVISCGYEDKTDWGAEIGWIYGSVTEIDIL TGFKMHARGWRSIYCMPLKAAFKGSAPINLSDRLNQ VLRWALGSVEILFSRHCPLWYGYGGRLKWLERFAYI NTTIYPLTSIPLLAYCTLPVAVCLLTGKFIIPQISNLASV WFLSLFLSIFATGILEMRWSGVGIDEWWRNEQFWVI GGISAHLFAVFQGLLKVLGIDTNFTVTSKASDEDGD FTELYLFKWTLLIPPTLIVNLVAVVAGISYAINSGY QSWGPLFGKLFFAFWVIVHLYPFLKGLMGRQNRTP IVVWSILLASIFSLWVRVDPFTTRVTGPDVHICGIN C	20 30
ATCSLA1 _At4g1659 0		86	ATGTCTCTATTTCTGAAGCCCTTCCTCTTCCTATAC GACACCACTCTTAGTCTTCTCTTACTTCTGTTCAT GGATGGAGTCTTGAGGATACAGCAGCAGCCCAA AGAGGCGTGAAGCAGACAAAATGCTGCAGAAAC TGAATGGATCCAACCTCAATACTTGTGGACCAAAA CAAGGAGTGTGTACTACTTCCCCTTTCAAGGGT TTGGTGGTTATGTGTTTGGTTCTATCCATTATAGTG TTCTTCGAGAGTTTTTACATGAACCTTGTGATACTC TTCGTCAAGTTATTTAAACGTAACCCCATAAAGT GTACAAATGGGAGGCCATGCAAGAAGATGTTGAG GTTGGACCCGATAACTACCAATGGTTCTTATCCA AATACCAATGTACAATGAAAAAGAGGAAGGTGTG GACGTAGAGATTGCAAAATGGCAAAGCCAAGGCA TAAACATAAGGTGTGAAAGGAGAGATAACAGGAA CGGCTACAAAGCCGGAGCTATGAAAGAAGCTCTT ACGCAGAGCTACGTCAAGCAATGCGACTTCGTAGC	40



【表 19 - 5】

			<p>TTATTTCTCTCTCAAGGAAGCTTCTAAGTTCGCCAA  GATTTGGGTACCGTTCTGCAAAAAGTACAACACTA  GAGTTAGAGCTCCTTCTAGATATTTTCTGAAACCT  ATAAGCGTCGCAACAGAGGATTATGAATTC AATAG  AGACTGGGAAAAGACGAAGAGGGAGTACGAGAA  GTTGAGGCGGAAAAGTGGAAGATGCCACCGGAGAT  TCTCATATGTTGGATGTAGAAGATGATTTTGAAGC  ATTCTCAAACACAAAACCAAATGATCATTCAACTC  TAGTTAAGGTGGTATGGGAGAACAAGGGAGGTGT  AGGAGACGAGAAAAGAGATCCCTCATATCATATAC  ATATCAAGAGAGAAAAGACCAAATTATGTTCATA  ATCAAAAATGTGGAGCCATGAACTTTCTGGCAAGA  GTGTCAGGGTTGATGACAAACGCACCATACATCTT  GAACGTGGATTGCGACATGTATGCCAATGATGCAG  ATGTAGTCCGACAAGCAATGTGTATACTTCTGCAA  GAATCATTAAATATGAAACATTGTGCTTTTGTTC A  ATTCCGTCAAGAATTCTATGATTCAAGCACCGAGC  TAATAGTCGTCCTACAATCACATTTGGGACGAGGA  ATCGCGGGAATCCAAGGACCGATATATATAGGATC  AGGATGCGTCCACACGAGAAGAGTTATGTATGGTT  TATCTCCAGACGATTTTCGAAGTTGATGGAAGTCTT  TCTTCAGTTGCTACAAGGGAGTTTTTGGTTAAGGA  TAGTTTAGCGAGAAGATTTGGTAATTCTAAAGAGA  TGATGAAATCAGTGGTTGATGCAATACAAAGAAAT  CCAAATCCACAAAATATACTTACAAACTCCATAGA  AGCGGCTCGAGAAGTGGGACATTGTCAGTACGAG  TACCAAACCAGCTGGGGAAACACCATCGGCTGGTT  ATATGATTCAGTGGCGGAAGATTTAAACACGAGTA  TCGGAATACATTTCGAGAGGTTGGACTAGCTCATA C  ATTTCTCCGGATACACCTGCATTTCTTGGATCTATG  CCGGCAGGAGTACCCGAGGCGTACTCCAGCAGC  GTCGATGGGCGACAGGATGGATCGAAATCCTTTTC  AACAAAGCAAAGTCCGTTGCGAGGATTGTTTAGCAA  GAAAATAAGATTCCGACAACGATTAGCTTATCTTT  GCATTATCACCTGTCTAAGGTCAATCCCTGAGCTT  ATTTATTGTCTCCTTCCTGCTTATTGCCTACTCCAC  AACTCTACCTTATTCCCAAGGGACTTTATTTAGGC  ATAACTGTCACACTTGTGGGATACATTGTCTCTAT  ACTCTATGGGAATTTATGAGCCTTGGTTATTCGTA  CAATCGTGGCTAGTCTCCAATCAGTTTGGAGAAT  AGTAGCCACTAGTAGTTGGTTATTTAGCATCTTTG  ATATCACACTCAAGCTTCTTGGCATCTCGGAAACG  GTGTTCATAATCACTAAAAAGACTGTGGCTGGGAC  CAAGTCAGCATTAGGGTCTGGACCCTCTCAAGGAG  AAGACGTTGGTCCAAACTCAGACTTGTTTAAATTT  GAATTTGATGGCTCACTTTGTTTCTTGCCTGGCACA  TTTATTGTGTTGGTGAATATAGCCGCTCTAGCTGTT  TTTTCTGTGGGTCTACAACGGTCGAGTTACAGCCA  TGAAGGAGGTGGTTCGGGTCTGGCAGAGGCTTGC  GATGTGTTTTGGTAATGATGTTGTTCCCTCCATTT C  TAATGGGTTTGTTTAAGAAAGGAAAATATGGAACC  CCATTGTCTACTCTCTATAGCTGGCTTTTTAGCA  GTTTTATTTGTTGTTTTCTCTGTTTGA</p>	<p>10</p> <p>20</p> <p>30</p> <p>40</p>
ATCSLB1 _AT2G326 10		89	<p>MADSSFSPLPLCERISYTNVFLRAVYLTVLGLFFSLLL  HRIRHTSEYDNVWLVAFFCESCFVLCLLITCLKWSP  ADTKPFPDRLDERVHDLPSVDMFVPTADPVREPPIM</p>	

【表 19 - 6】

		<p>VVDTVLSLLAVNYPANKLACYVSDDGCSPLTYFSLK  EASKFAKIWVPFCKKYNTRVRAPSRYFLKPISVATED  YEFNRDWEKTKREYEKLRKVEDATGDSHMLDVED  DFEAFSNTKPNHSTLVKVVWENKGGVGDKEIPIHII  YISREKRPNYVHNQKCGAMNFLARVSLMTNAPYIL  NVDCDMYANDADVVRQAMCILLQESLNMKHCASFV  QFRQEFYDSSTELIVVLQSHLGRGIAGIQGPIYIGSGC  VHTRRVMYGLSPDDFEVDGSLSSVATREFLVKDSL  RRFGNSKEMMKSVVD AIQRNPNPQNILNLSIEAAREV  GHCQYEYQTSWGN TIGWLYDSVAEDLNTSIGIHSRG  WTSSYISPDTPAFLGSM PAGVPEALLQORRWATGWI  EILFNKQSPLRGLF SKKIRFRQLAYLCIITCLRSIPELI  YCLLPAYCLLHNSTLFPKGLYLGITVTLVGIHCLYTL  WEFMSLGYSVQSWLVSQSVWRIVATSSWLFSIFDITL  KLLGISETVFIITKKT VAGTKSALGSGPSQGEDVGPNS  DLFKFEFDGSLCFLPGTFIVLVNIAALAVFSVGLQRSS  YSHEGGGSGLAEACGCVLVMMLFLPFLMGLFKKGK  YGTPLSTLSIAGFLAVLFV VFSV</p>	10
<p>ATCSLE1  _AT1G558  50</p>	<p>90</p>	<p>ATGGTAAACAAAGACGACCGGATTAGACCGGTTTC  ATGAAGCCGACGGTGAACCGCTTTTTGAGACTAGG  AGAAGAACCGGTAGAGTGATTGCGTACCGGTTTTT  CTCAGCCTCGGTTTTCGTGTGTATCTGTTTGATTG  GTTCTACAGAATTGGTGAGATTGGTGATAACCGTA  CCGTTTTAGATCGATTAATCTGGTTTGTTATGTTA  TTGTGGAGATTGGTTCGGTTTATATTGGGTAGTCA  CACAACTTCCCGGTGGAATCCGGTTTGCGGATTT  CCCTTCTCCGATAGACTCTCTCGGAGATACGGAAG  CGACCTCCGAGGCTCGACGCTTTCGTTTGACCGG  CGGATCCGGTGATTGAGCCGCCGTTGTTGGTGGTA  AACACAGTCTTATCTGTGACGGCTCTTACTACCC  ACCGGAGAACTCGCCGTTTATCTCTCAGATGACG  GTGGTTCTGAGCTGACGTTCTATGCTCTCACGGAG  GCAGCTGAGTTTGCTAAACTTGGGTTCCCTTCTG  CAAGAAGTTCAACGTTGAGCCAACATCTCCCGCTG  CTTACTTGTCTTCCAAGGCAAACTGTCTTGATTCTG  CGGCTGAGGAGGTGGCTAAGCTGTATAGAGAAAT  GGCGGCGAGGATTGAAACGGCGGCGAGACTGGGA  CGAATACCGGAGGAGGCGCGGGTGAAGTACGGTG  ACGGGTTTTACAGTGGGATGCTGACGCTACTCGA  AGAAACCATGGAACCATCTTCAAGTTTTGGTAGA  TGGAAGAGAAGGGAATACAATAGCAATACCAACG  TTGGTGTATTTATCAAGAGAAAAGAGACCTCAACA  TCATCATAACTTCAAGGCTGGAGCAATGAACGCAT  TGCTGAGGGTTTCTTCGAAAATTACTTGTGGGAAA  ATCATACTAAACTTGGACTGTGATATGTACGCAAA  CAACTCAAAGTCAACACGCGACGCGCTCTGCATCC  TCCTCGATGAGAAAGAGGGAAAAGAGATTGCTTTC  GTGCAGTTTCCGCAGTGTGTTTGACAATGTTACAAG  AAATGATTTGTATGGAAGCATGATGCGAGTAGGA  ATTGATGTGGAATTTCTTGGATTGGATGAAATGG  TGGTCCGTTATACATTGGAAGTGGATGCTTTCACA  GAAGAGATGTGATCTGTGGAAGAAAGTATGGAGA  GGAAGAAGAAGAAGAAGTCTGAGAGAATTAC  GAAAATTTAGAGCCTGAGATGATTAAGGCTCTCGC  GAGCTGCACTTATGAGGAAAACACTCAATGGGGA  AAGGAGATGGGTGTGAAATATGGTTGCCCGGTAG</p>	20
			30
			40



【表 19 - 8】

		<p>GGTTGTTAACACCGCTTTATCGGTGATGGCTTACG  AGTATCCGTCAGATAAGATCTCGGTGTATGTATCG  GACGATGGAGGATCGTCGTTGACTTTCTTTGCTCTT  ATTGAAGCTGCTAAGTTCTCTAAGCAGTGGTTGCC  CTTTTGCAAGAAGAATAATGTTCAAGATCGGTCTC  CTGAAGTTTATTTCTCTTCAGAGTCACATTCTCGAA  GTGATGAAGCTGAAAACCTTAAGATGATGTACGA  AGACATGAAGAGTAGAGTAGAACATGTGGTGGAG  AGTGAAAAGTTGAAACTGCGTTTATCACATGCGA  CCAATTTCTGTTGGGTATTCTGATTTGTGGACCGACA  AATTCAGTCGTCATGACCATCCCACAATTATTCAG  GTGTTGCAAAATAGCGAGACAGATATGGACAATA  CCAGAAAATATATAATGCCAAACCTAATCTATGTT  TCAAGAGAGAAGAGTAAAGTTTCACCACATCATT  CAAAGCTGGTGCTCTTAATACTTTGCTACGAGTAT  CAGGGGTGATGACAAATTCACCGATCATTCTAACA  CTAGACTGTGATATGTATTCTGAACGACCCGGCAAC  ACTGGTTCGTGCTTTGTGCTATTTAACAGATCCTGA  AATCAAATCCGGTTTAGGATATGTGCAGTTTCCTC  AGAAATTTCTAGGAATAAGCAAAAATGATATATAT  GCTTGTGAAAACAAACGCCTCTTCATTATTAATAT  GGTTGGGTTTGATGGTCTAATGGGTCCAACCTCATG  TGGGAACCTGGTTGTTTCTTTAATCGACGAGCTTTCT  ATGGACCTCCATATATGTTGATTTTACCGGAGATA  AATGAACTAAAGCCTTATCGGATTGCGGATAAGTC  TATCAAAGCCCAAGATGTTTTGTTCATTAGCACACA  ATGTAGCAGGATGTATCTATGAGTACAATACCAAT  TGGGGATCCAAGATTGGATTGAGATATGGGTCATT  AGTAGAAGACTACTACACAGGGTTTATGCTCCATT  GTGAAGGATGGAGATCAGTATTTTGCAACCCAAAA  AAAGCTGCATTTTATGGAGATTCCCCAAAGTGCCT  AGTTGATCTTGTGGGTCAACAAATCCGTTGGGCAG  TTGGGCTTTTCGAAATGTCCTTTTCAAAGTATAGCC  CAATTACCTATGGAATCAAGTCACTGGACCTTTTA  ATGGGTTTAGGTTATTGCAACTCTCCGTTTAAGCC  ATTTTGGTCAATTCCTCTGACCGTCTATGGACTTTT  ACCACAGCTTGCACTCATTCTGGAGTTAGTGTCTT  CCCCAAGGCATCTGATCCGTGGTTTTGGCTTTACAT  CATTTTATTCTTTGGGGCTTATGCCAAGATCTATC  AGACTTTTTATTGGAAGGAGGAACTTATCGGAAAT  GGTGAACGATCAAAGAATGTTGATGATAAAAGG  ACTCTCTTCATTCTTCTTTGGTTTTATAGAGTTCATT  CTCAAAACCCTAAACCTCTCCACACCTAAGTTCAA  CGTCACCAGTAAAGCCAATGATGATGACGAACAG  AGGAAGCGGTACGAGCAAGAAATCTTTGATTTTCGG  AACCTCTTCGTCCATGTTCTTGCCCTTGACCACGGT  TGCCATAGTGAATCTGCTTGCTTTTGTCTGGGGGCT  TTATGGTATTCTCTTCTGCGGAGGAGAACTCTACCT  TGAGCTGATGCTGGTGGCTTCGCGGTGGTGAATT  GCTTACCGATCTACGGGGCTATGGTGTGAGGAAA  GATGATGGAAAATTATCAAAAAGAACTTGTCTT  AGCTGGGAACCTCCACGTTGGTTCTTATTGTGTCA  AGTTACTTCGTCCTCAAGTAACTTCACCCCTTAGGT  TAATTCACAACAATAATACGTCTGGCTGGTTCAAG  CGGAAGAAACACAATATGAATGAATCTGTGTAA</p>
--	--	--

10

20

30

40

トリテルペノイド系アグリコンの酵素活性部位への通路を形成するものと考えられる3つのチャンネルと、グルクロン酸抱合物の出口通路であり得る1つのチャンネルも検出された(図49A~49C)。セルロース合成酵素A(CESA)タンパク質において、膜貫通ドメイン(TMD)は、多糖鎖を膜を介して細胞外空間へと通過させることを可能とする細孔を形成する。CESAにおける膜外ドメインは、酵素活性部位と、このようなタンパク質全てに保存され、基質及び/または受容体結合への関与が予測される4つのモチーフ、すなわち、DD、DCD、ED及びQVLRWとを保持している(「L. Sethaphong et al., Tertiary model of a plant cellulose synthase, Proc. Natl. Acad. Sci. U.S.A. 18, 7512-7 (2013).」)(図48 - 赤色のボックス)。ハウレンソウ由来のCESA及びSOAP5のアミノ酸配列のアライメントにより、ED(SOAP5中:ES)及びQVLRW(SOAP5中:QLIKW)モチーフは、SOAP5には保存されていないことが示された(図48)。

10

#### 【1531】

CESA型アミノ酸モチーフの存在がSOAP5活性に影響を及ぼし得るか否かを確認するため、SOAP5の以下のような3つの突然変異型を作製した:(i)セリン442をアスパラギン酸で置換したもの(S442D)、(ii)リジン483をアルギニンで置換したもの(K483R)及び(iii)両者を置換したもの(S442D/K483R)。突然変異タンパク質をSOAP1~4と共にベンサミアナタバコにおいて発現させたところ、SOAP5中のこれらのモチーフの修飾により、酵素性能が変化することが示された。野生型SOAP5はメディカゲン酸プール全体をMA-3-GlcAに変換し(MA-3-GlcA:MA比:0:1)、SOAP5-S442DまたはSOAP5-K483Rは依然として効果的にMAをグルクロン酸抱合し得る(それぞれMA-3-GlcA:MA比0.03:1及び0.09:1)のに対し、SOAP5-S442D/K483Rは、MAを部分的にプロセシングできたのみであった(MA-3-GlcA:MA比0.90:1)(図50A及び50B)。本発明者らの所見によれば、中性アミノ酸が負に帯電したアミノ酸に変化することによって、アニオン性基質(UDP-GlcA)の結合に影響を及ぼすのに加え、酵素活性部位において立体障害を引き起こし、そのため酵素効率を同時に低下させることが示される。酵素が発達する際の活性部位におけるアミノ酸の変化(S442D及びK483R)が、UDP-Glcの代わりにUDP-GlcAを受け入れるために不可欠となり、その結果、CESAタンパク質の新機能獲得をもたらした可能性がある。

20

30

#### 【1532】

要約:CESAの重要なアミノ酸残基の違いにより、CSLG酵素活性に対して酵素の機能分化をもたらされるものと考えられる。その新たなトリテルペノイド関連機能に不可欠なCSLGのもう1つの特徴は、小胞体(ER)への局在化であった。

#### 【1533】

実施例25:トリテルペノイド系サポニン生合成経路の細胞内局在化

#### 【1534】

目的:トリテルペノイド系サポニン生合成経路の細胞内局在化について調査すること。

#### 【1535】

方法:上記の材料及び方法を参照されたい。

40

#### 【1536】

結果:トリテルペノイド系サポニン生合成経路の細胞内局在化について検討した。ER膜系は、ステロイド及びトリテルペノイドなどの特化した代謝物の産生に関連する環境である。スクアレン合成酵素、スクアレンエポキシダーゼ、オキシドスクアレンシクラーゼ及び既知のCYP450のほとんどを含むトリテルペノイド生合成に関与する多くの酵素は膜結合型であり、ER区画内で作用する(「C.A. Hasemann et al., Structure and function of cytochromes P450: a comparative analysis of three crystal structures. Structure. 1, 41-62 (1995)」; 「K. B. Linscott, T. D. Niehaus, X. Zhuang, S. A. Bell, J. Chappell, Mapping a kingdom-specific functional domain of squalene synt

50

hase. BBA - Molecular and Cell Biology of Lipids. 1861, 1049-1057 (2016)」; 「M. Christen et al., Structural insights on cholesterol endosynthesis: Binding of squalene and 2,3-oxidosqualene to supernatant protein factor. Journal of Structural Biology. 190, 261-270 (2015)」; 「R. Thoma et al., Insight into steroid scaffold formation from the structure of human oxidosqualene cyclase, Nature. 7013, 118-122 (2004).」)。

#### 【1537】

蛍光レポーターに融合したSOAP5 (SOAP5:mRFP)を細胞区画マーカー ( GFPと融合したERマーカー)と共に一過性発現させたところ、ERネットワークへの局在化が示された(図51A~51F)。以前の報告では、発芽ダイズ種子由来のミクロソーム分画は、ソヤサポゲノールBをグルクロン酸抱合することが可能なタンパク質を含むことが示されており、SOAP5の膜貫通局在化を支持するものである(「Y. Kurosawa, H. Takahara, M Shiraiwa, UDP-glucuronic acid:soyasapogenol glucuronosyltransferase involved in saponin biosynthesis in germinating soybean seeds. Planta. 215, 620-9 (2002)」)。次に、SOAP1、2、3、4及び5の蛍光標識型を様々な組み合わせで発現させ、試験した全てのタンパク質のER共局在化を確認した(図35A~35F、図52A~52F)。全ての融合タンパク質を同時に発現させた葉において多量のMA-3-GlcAが検出され、確認した蛍光シグナルが、機能的、したがって適切に折りたたまれて細胞内に局在化したタンパク質によって生じたことが示された(図53)。蛍光共鳴エネルギー移動(FRET)を用い、SOAP1~5タンパク質は互いに近接して位置しており、このため相互作用する可能性があることも認められた(図54A~54B)。試験したタンパク質が正確に共局在化していることから、中間体の拡散を低下させることによって行った反応の高い効率性を説明することができる。

10

20

#### 【1538】

要約：これらの試験により、少なくとも酵素SOAP1~5が、ER中で互いに近接して局在化していることが示された。

#### 【1539】

実施例からの全般的な結論

#### 【1540】

トマト、ジャガイモ、ナス、及びホウレンソウにおけるステロイド系アルカロイド、ステロイド系サポニン、及びトリテルペノイド系サポニン生合成についてのこれらの試験の結果、トリテルペノイド系サポニン複合体の生合成経路を完結する遺伝子を見出し、特性評価を行い、複合型ステロイド系アルカロイド及びステロイド系サポニンの生合成経路を解明した。2つの異なる目に属する7種類の植物種において、トリテルペノイド系アグリコンのグルクロノシルトランスフェラーゼとして、セルロース合成酵素様Gタンパク質の前例のない機能が見出され、これによって、特化した代謝物のグリコシル化が、1つのファミリー炭水化物活性酵素データベース(1 family Carbohydrate-Active enzymes Database)(CAZY)に属するGTによって限定的に行われるわけではないことが証明される。この知見はおそらく、他の特化した代謝物(SM)に関連したCSL酵素についてさらなる機能が見出されるきっかけとなるものと考えられる。また、特有のSOAP6フコシルトランスフェラーゼ活性が見いだされ、トリテルペノイド系サポニンの糖部分をアセチル化可能な初のBAHDアシルトランスフェラーゼ(SOAP10)の特性評価が示された。同様に、GAME15は、特定の実施形態において、ステロイド系アルカロイド及びステロイド系サポニンをグリコシル化し得る。いくつかの実施形態において、GAME15のグリコシル化は、グルクロン酸の転移(グルクロノシルトランスフェラーゼ活性)を含む。さらに、ステロイド系アルカロイド及びステロイド系サポニンの発現が低下することから、栄養剤、化粧品、及び医薬品として有用な、植物コレステロール、フィトコレステロール、コレスタノール、フィトコレスタノール、及びフィトステロールの含有量を増加させる新たな手法がもたらされる。

30

40

#### 【1541】

50

トリテルペノイドグルクロノシルトランスフェラーゼとしてのCSLGタンパク質活性が予想外に見いだされたことは、SMのグリコシル化が、ファミリー1型UGTによって限定的に行われるわけではないことを示している。したがって、この報告は、他の種類のテルペノイド及びSMの修飾において、並びに他の種類のステロイド系アルカロイド及び/またはステロイド系サポニンにおいて機能するCSL酵素を見出すきっかけとなるものと考えられる。さらに、数少ない主要な植物のプロセスの1つ、すなわち、細胞壁を構築するセルロース合成機構における酵素活性が、無関係な、防御に特化した一連の代謝物を産生するために、いかにして「乗っ取られるか」といった、基本的な例が提供される。SOAP5がほとんどの他の経路タンパク質とともに、細胞質基質(UGTが局在化する)ではなく、ハウレンソウER膜に共同在化したことは、その新たな機能の発達的一端であり、効率的なトリテルペノイド産生に関するこれらの酵素間の物理的接近性の重要性を示唆している。この物理的接近性はまた、トマトにおける遺伝子クラスターとしてのGAME遺伝子に関して、遺伝子レベルでも反映された。これまで、全ての植物フコシルトランスフェラーゼは細胞壁多糖類及び糖タンパク質のフコシル化に関連していたが、グルクロン酸トランスフェラーゼ以外の、代謝物に対して作用するこれまで未開示であったフコシルトランスフェラーゼ(すなわちSOAP6)が同定された。BAHD型アシルトランスフェラーゼはSMの修飾に関連することが多い一方で、ハウレンソウトリテルペノイド系サポニンのC-28フコース部分をアセチル化するSOAP10について本明細書で示されたような、トリテルペノイドのアシル化については報告がなかった。グルクロン酸抱合と同様に、トリテルペノイド系サポニンにおけるアセチル化フコースドメインは、治療薬としての有効性、例えば、キラヤ由来の強力なサポニンワクチンアジュバントである、よく知られたQS-21のアジュバント活性にとって重要である。

#### 【1542】

本研究によれば、多くのグルクロニド型トリテルペノイド系サポニンの生合成経路における不足を埋めることにより、食品及び医薬品産業向けの高価値代謝物(例えば、グリチルリチン、ソヤサポニン、QS-21アジュバント)の持続可能な産生のために、植物または酵母または他の真核細胞系を異種操作する上で、無類の戦略がもたらされる。さらに、本研究によれば、植物トリテルペノイド系サポニン生合成経路内の内在性遺伝子を操作して、有益なトリテルペノイド系サポニン、例えば、これに限定するものではないが、中国カンゾウ中のグリチルリチンを増加させる、または、苦い若しくは望ましくないトリテルペノイド系サポニンを低下させる、例えば、これに限定するものではないが、キノア中のこのようなサポニンを低下させることが支援される。同様に、望ましくないステロイド系アルカロイド及び/またはステロイド系サポニンを低下させることによって、毒素及び苦味化合物が低下される。

#### 【1543】

ステロイド系アルカロイド、ステロイド系サポニン、及びトリテルペノイド系サポニンを産生するための遺伝子組換え細胞の特定の特徴及びこれらの使用について本明細書で示し、説明してきたが、ここで、多くの変形、置換、変更、及び同等物が、当業者には思い浮かぶことであろう。具体的な実施形態の前述の説明は、本発明の一般的な性質をよく表しているため、他者は、現在の知識を応用することにより、過度の試験を行うことなく、かつ一般的な概念を逸脱することなく、このような具体的な実施形態を様々な用途のために容易に変形し、かつ/または適応することができる。したがって、これらの適応及び変形は、開示された実施形態の等価物の意味や範囲内で理解されるべきであり、そのように意図されている。本明細書で用いられた語法または用語は、説明を目的とし、限定を目的とするものではないことを理解されたい。様々な開示された機能を行うための手段、材料、及びステップは、本発明から逸脱することなく、様々な代替の形態をとることができる。したがって、添付の特許請求の範囲は、本明細書の開示の真の精神の範囲内にあるような全てのこのような変形及び変更を網羅することを意図していることを理解されたい。

10

20

30

40

【 図 1 】

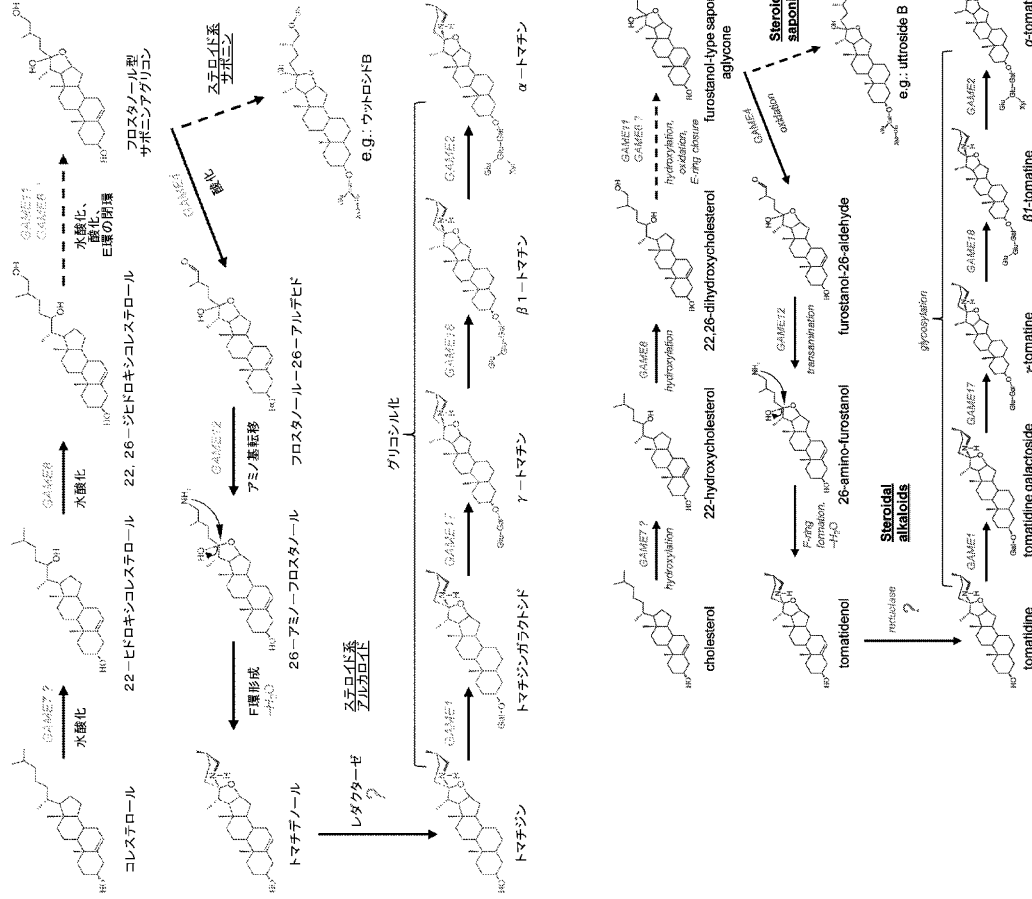
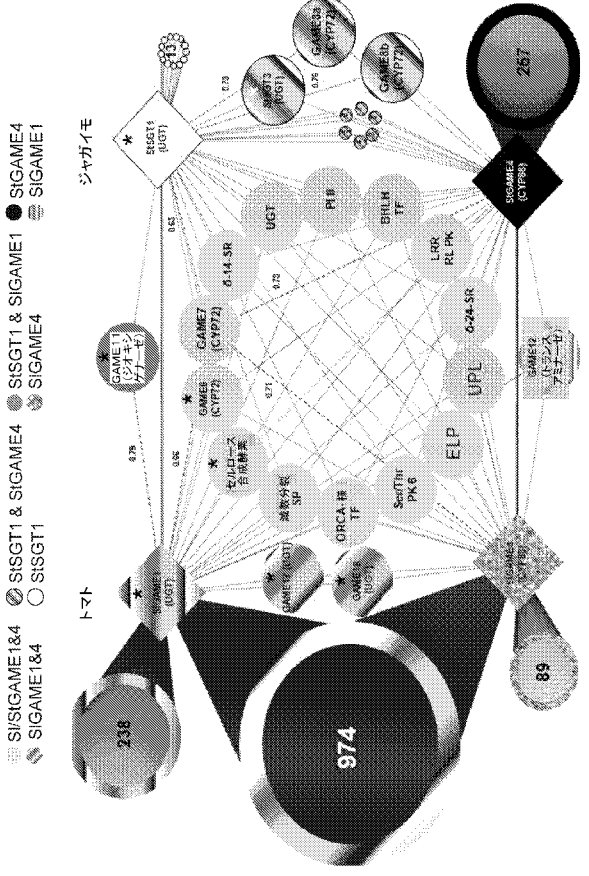
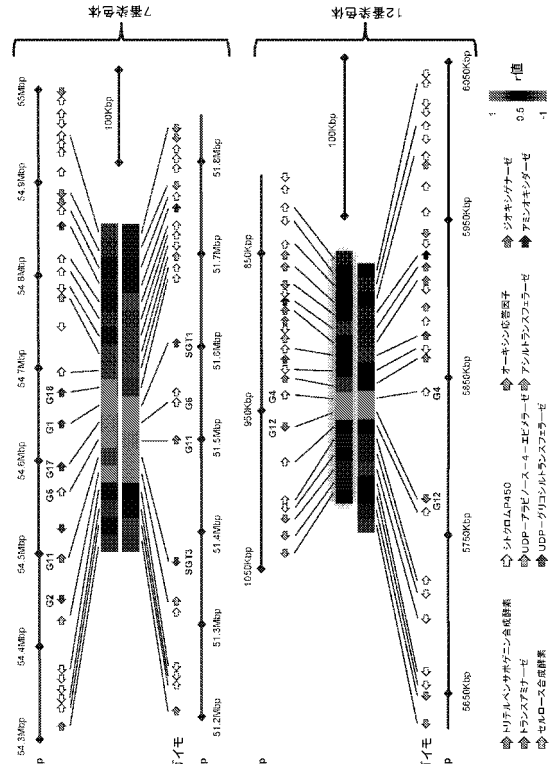


FIGURE 1

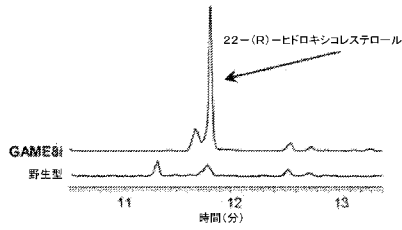
【 図 2 】



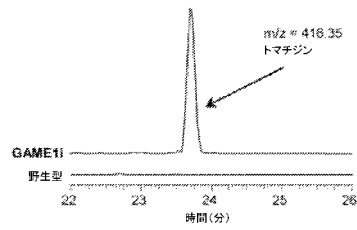
【 図 3 】



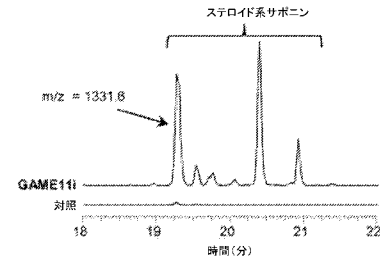
【 図 4 A 】



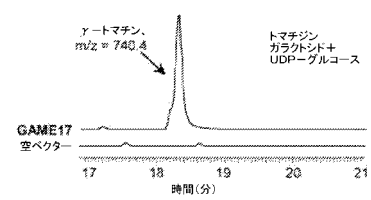
【 図 4 D 】



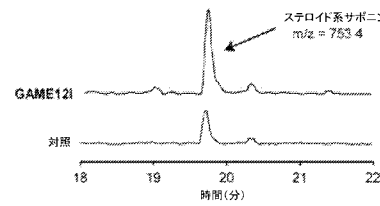
【 図 4 B 】



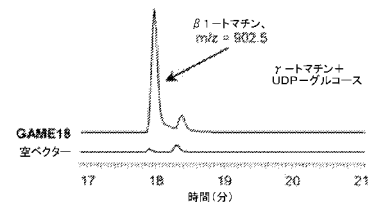
【 図 4 E 】



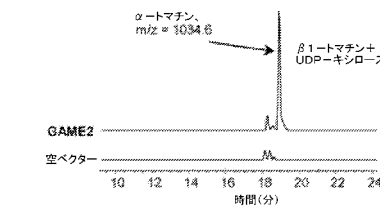
【 図 4 C 】



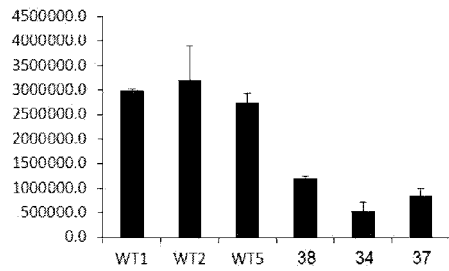
【 図 4 F 】



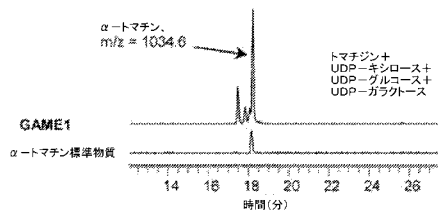
【 図 4 G 】



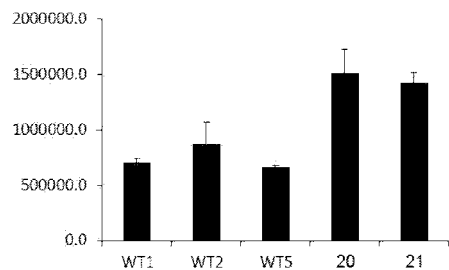
【 図 5 B 】



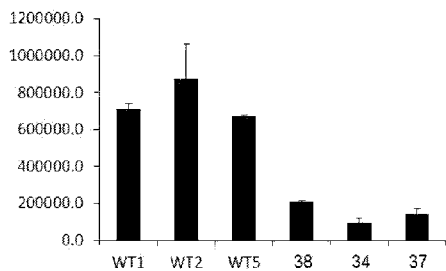
【 図 4 H 】



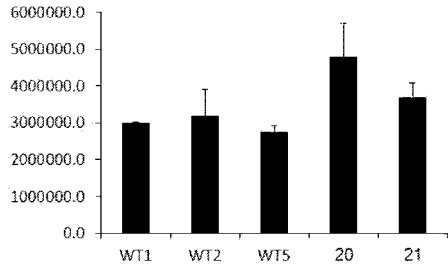
【 図 5 C 】



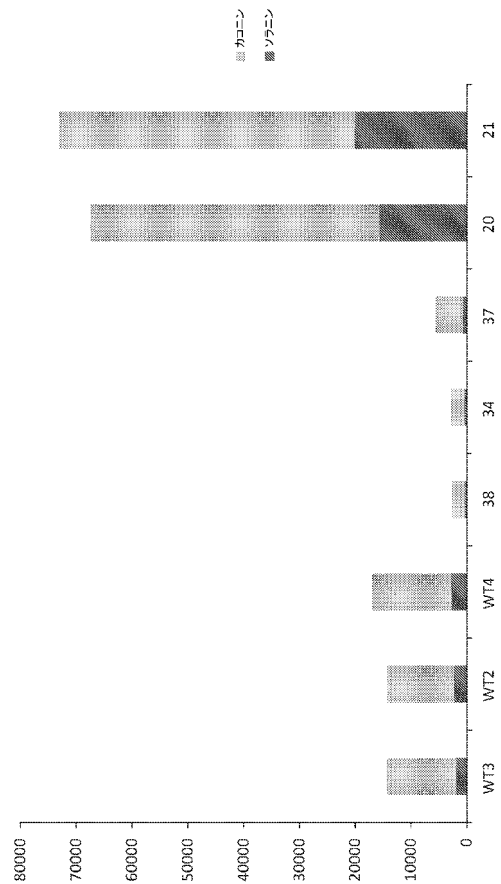
【 図 5 A 】



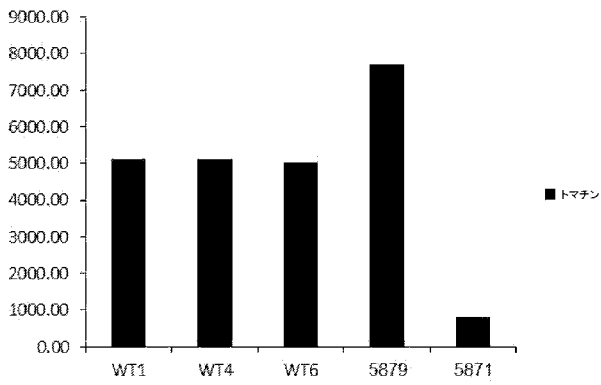
【 図 5 D 】



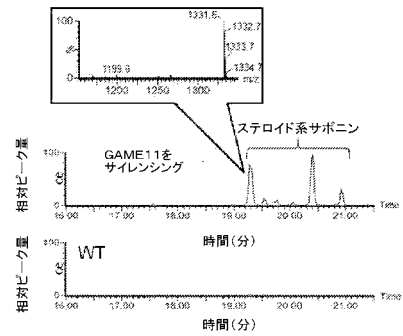
【 図 6 】



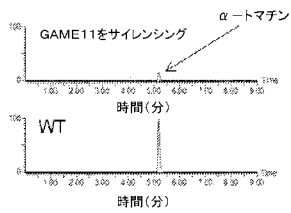
【 図 7 】



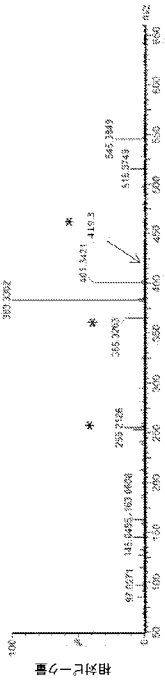
【 図 8 B 】



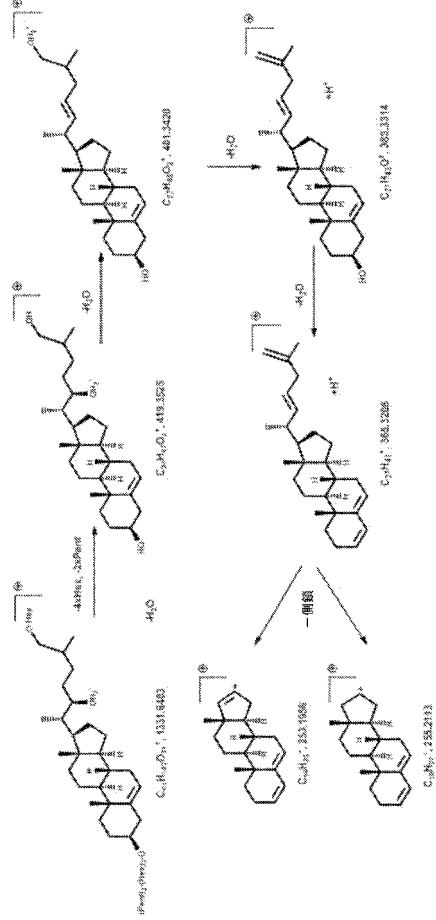
【 図 8 A 】



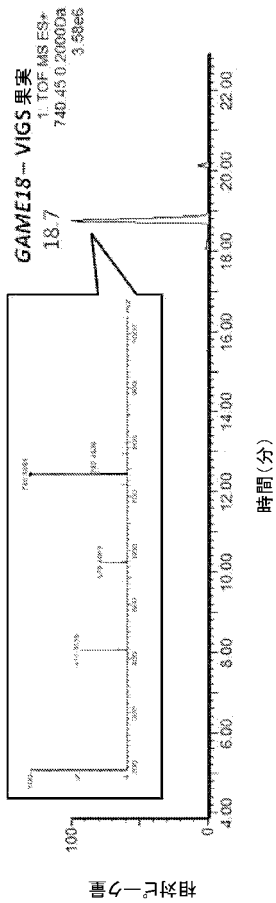
【 図 8 C 】



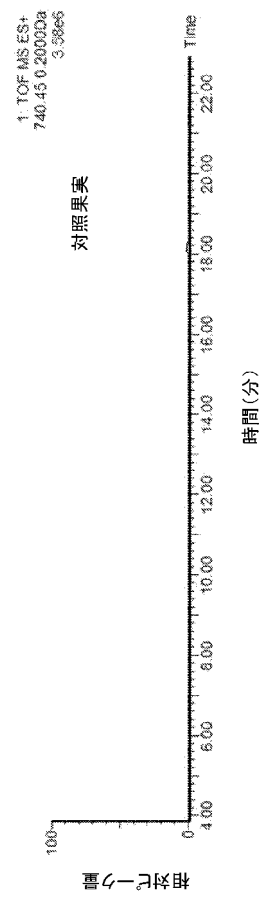
【 図 8 D 】



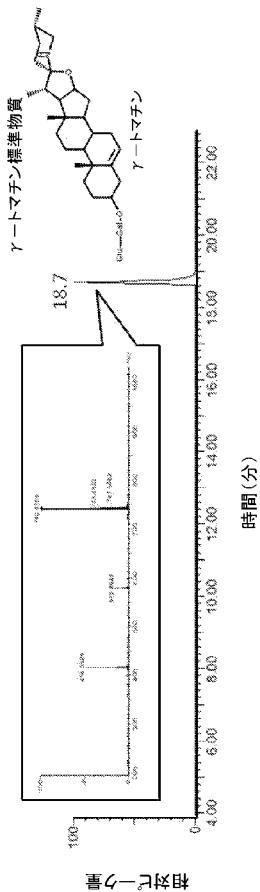
【 図 9 A 】



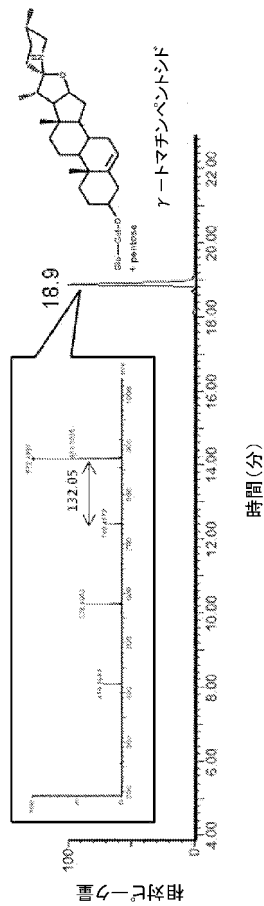
【 図 9 B 】



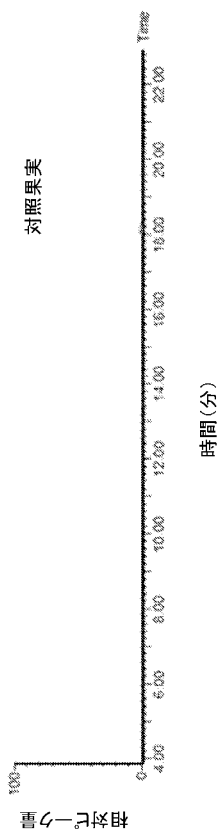
【 図 9 C 】



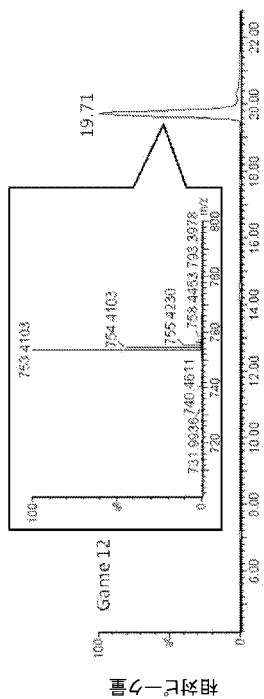
【 図 9 D 】



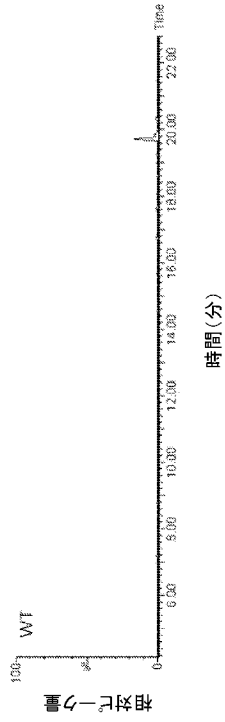
【 図 9 E 】



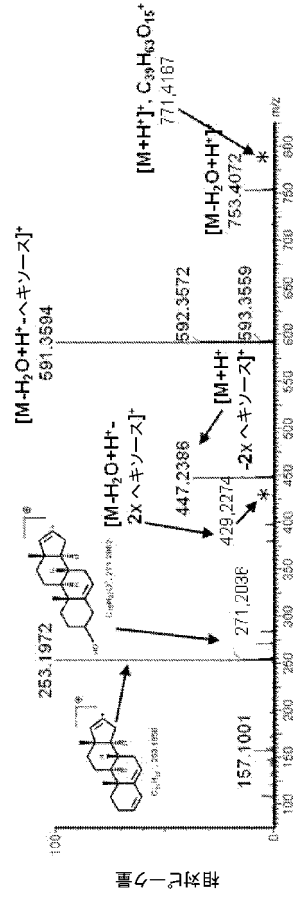
【 図 10 A 】



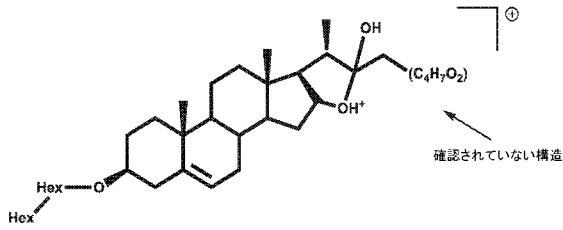
【 図 1 0 B 】



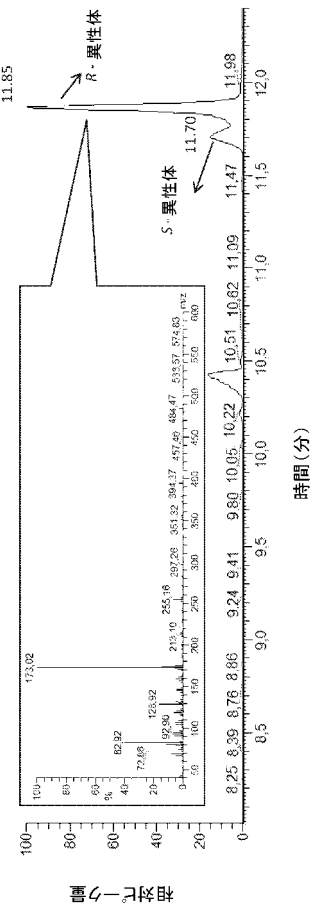
【 図 1 0 C 】



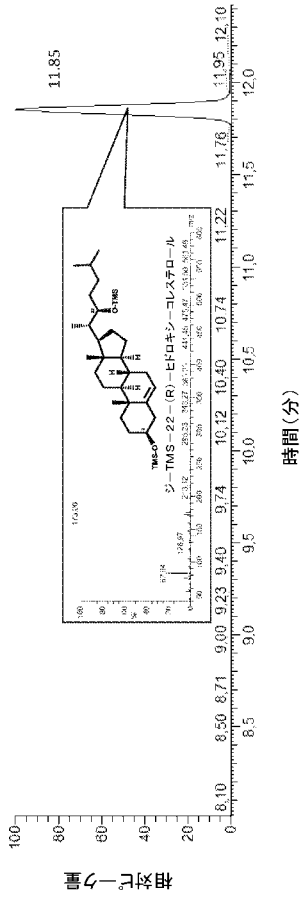
【 図 1 0 D 】



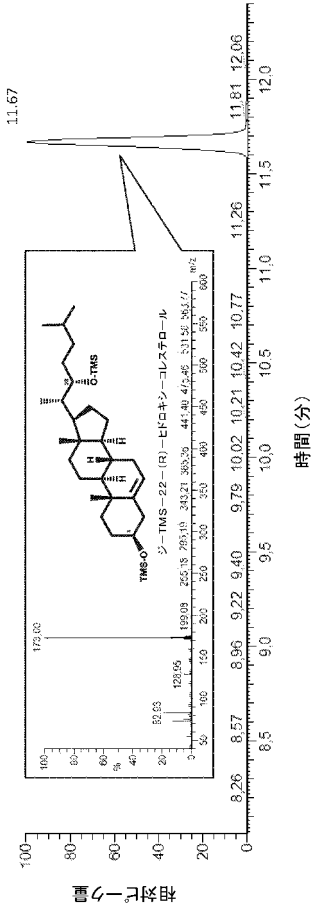
【 図 1 1 A 】



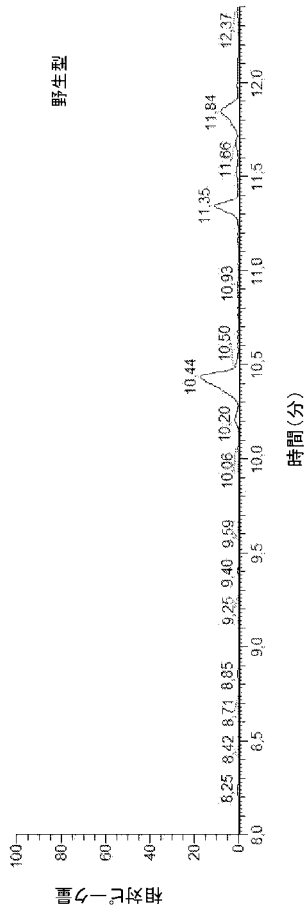
【図 1 1 B】



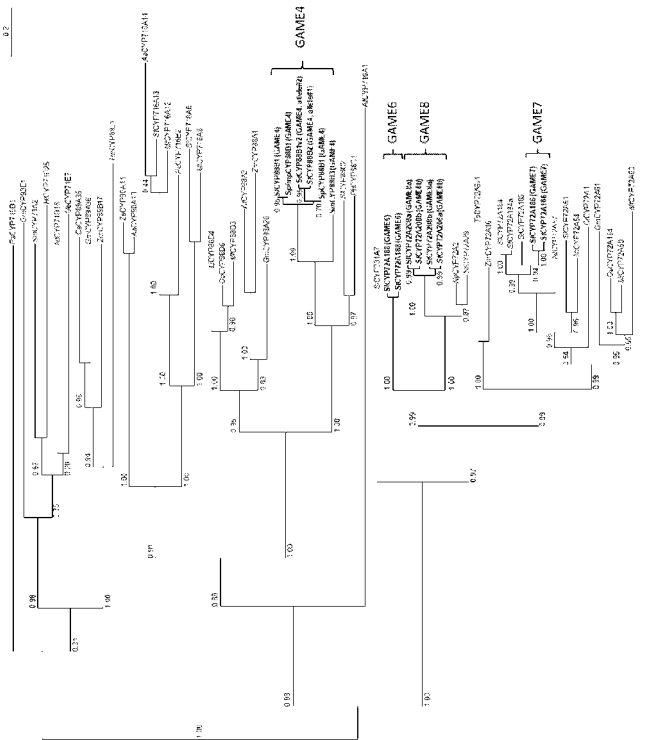
【図 1 1 C】



【図 1 1 D】

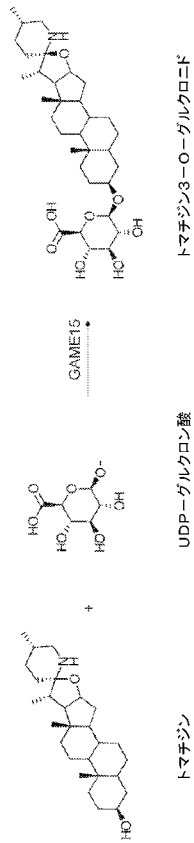


【図 1 2】

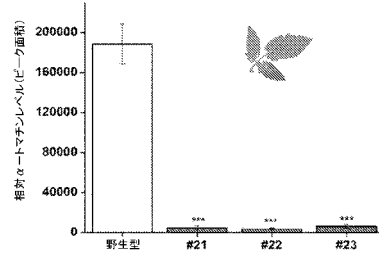




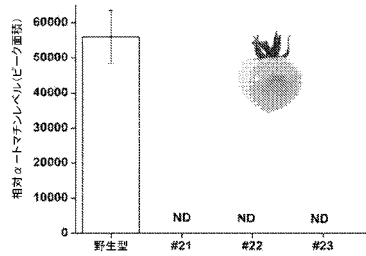
【 図 1 4 D 】



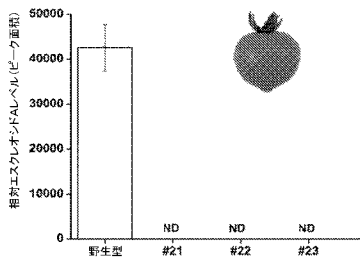
【 図 1 5 A 】



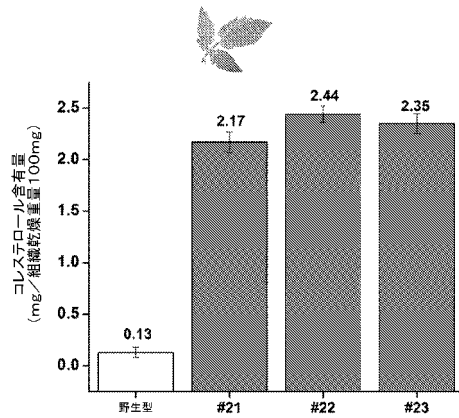
【 図 1 5 B 】



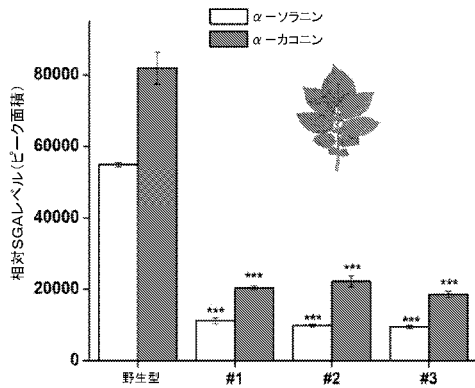
【 図 1 5 C 】



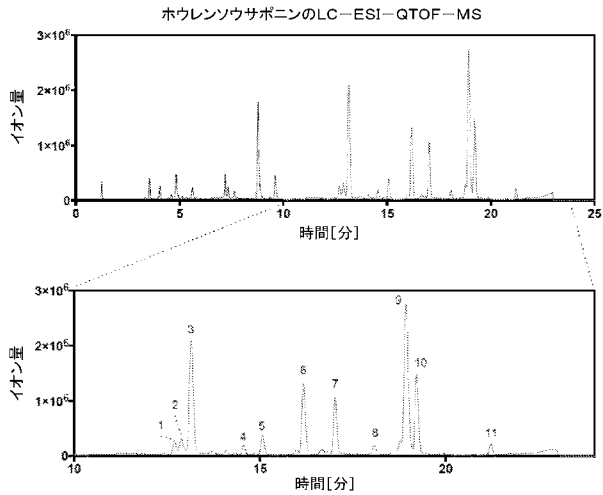
【 図 1 7 】



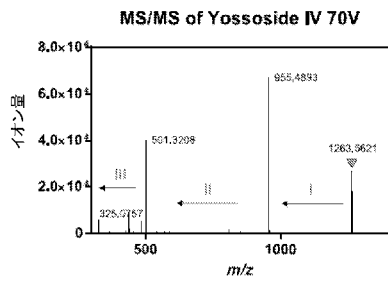
【 図 1 6 】



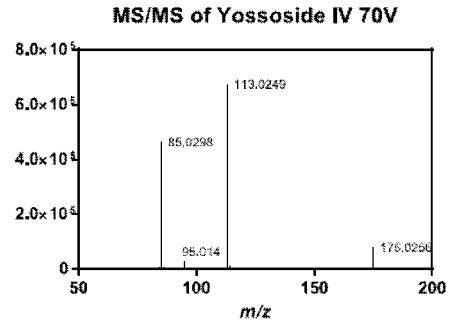
【 図 1 8 】



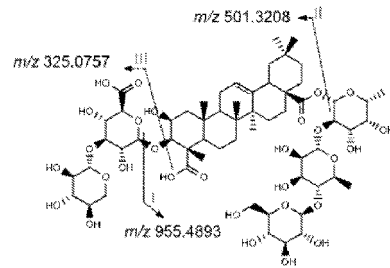
【 図 1 9 A 】



【 図 1 9 B 】



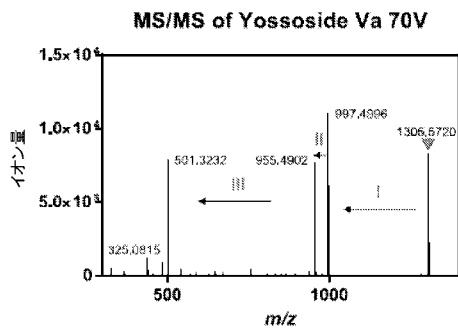
【 図 1 9 C 】



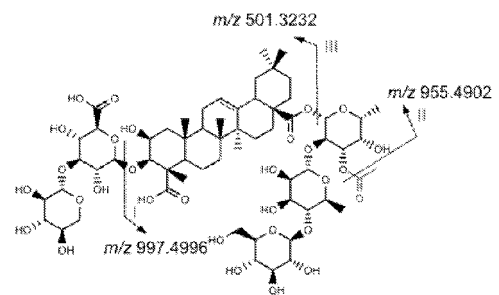
【 図 1 9 D 】

m/z (実測値)	予測式	イオンの帰属	m/z (理論値)	$\Delta m/z$ (ppm)
175.0256	C6H7O6	[GlcA-H <sub>2</sub> O]	175.0246	7.4
113.0249	C5H5O3	[GlcA-2H <sub>2</sub> O-CO <sub>2</sub> ]	113.0239	8.8
95.0140	C5H3O2	[GlcA-3H <sub>2</sub> O-CO <sub>2</sub> ]	95.0133	7.4
85.0298	C4H5O2	[GlcA-2H <sub>2</sub> O-CO <sub>2</sub> -CO]	85.0290	9.4

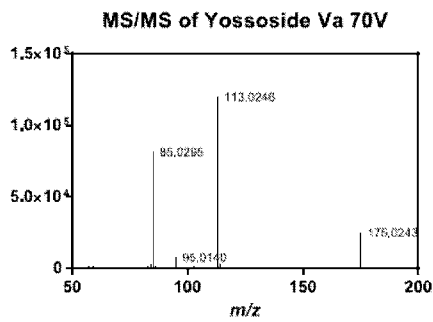
【 図 1 9 E 】



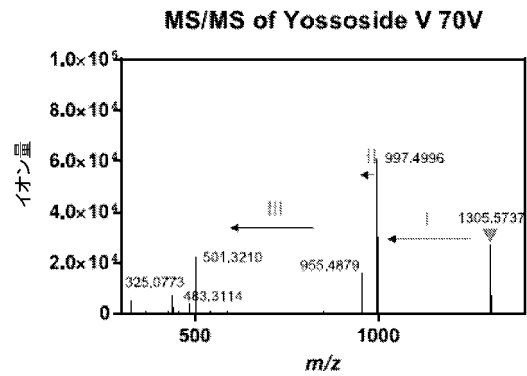
【 図 1 9 G 】



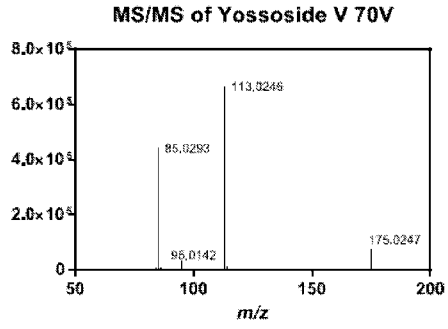
【 図 1 9 F 】



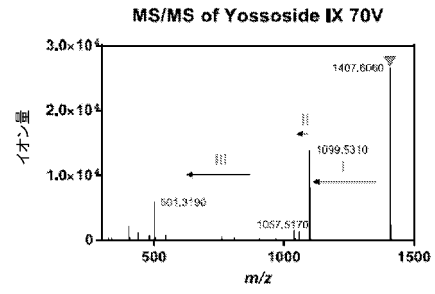
【 図 1 9 H 】



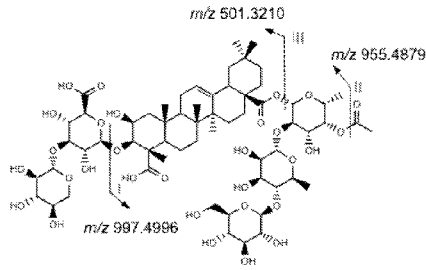
【 図 1 9 I 】



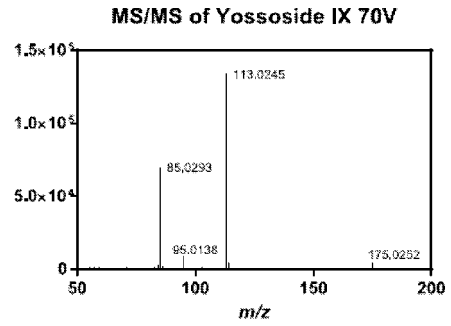
【 図 1 9 K 】



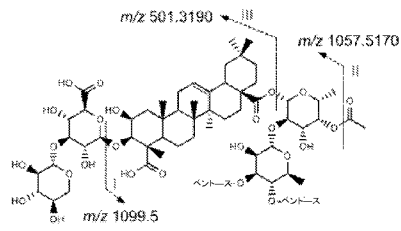
【 図 1 9 J 】



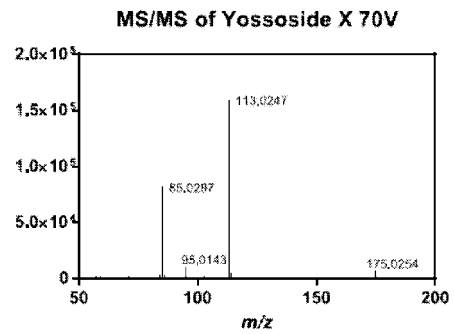
【 図 1 9 L 】



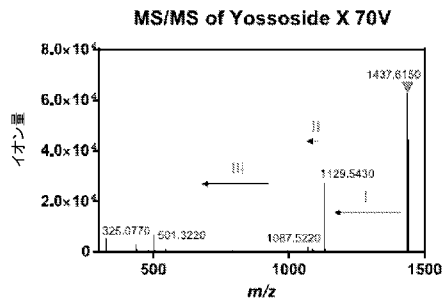
【 図 1 9 M 】



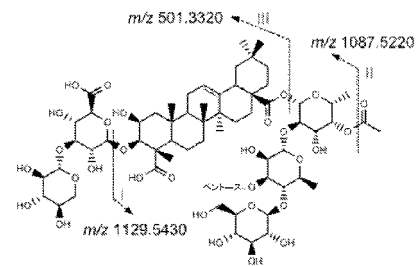
【 図 1 9 O 】



【 図 1 9 N 】

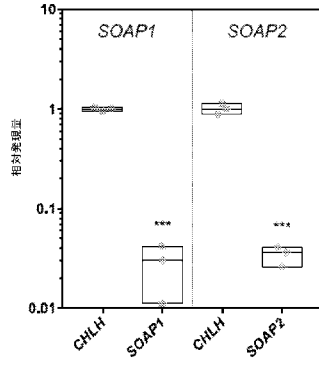


【 図 1 9 P 】

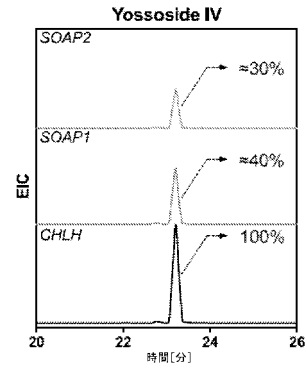




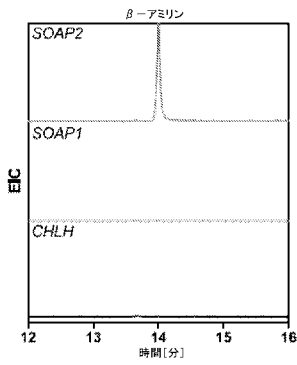
【 図 2 3 B 】



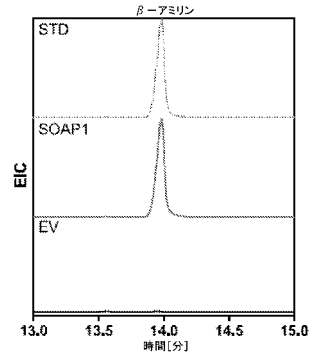
【 図 2 3 D 】



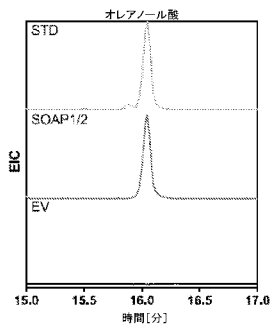
【 図 2 3 C 】



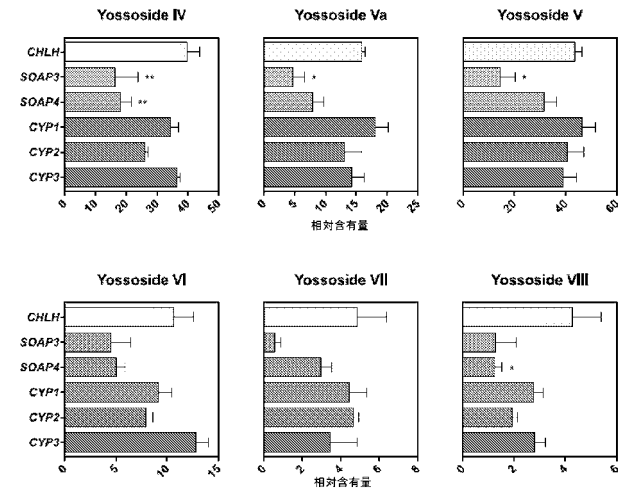
【 図 2 4 A 】



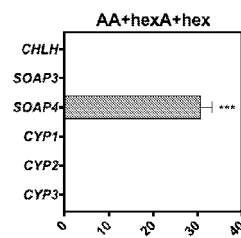
【 図 2 4 B 】



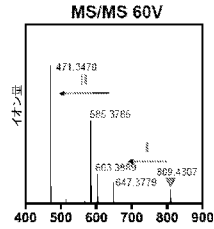
【 図 2 5 】



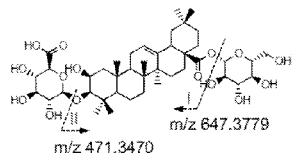
【 図 2 6 A 】



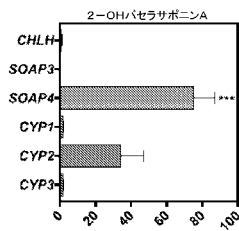
【 図 2 6 B 】



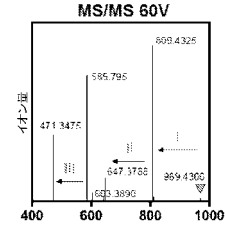
【 図 2 6 C 】



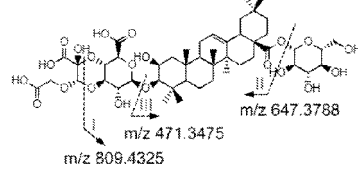
【 図 2 6 D 】



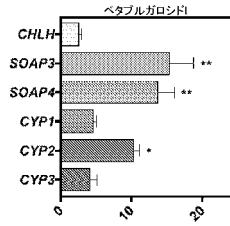
【 図 2 6 E 】



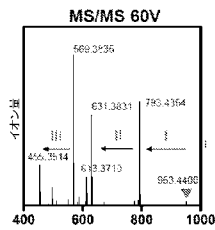
【 図 2 6 F 】



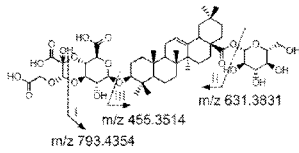
【 図 2 6 G 】



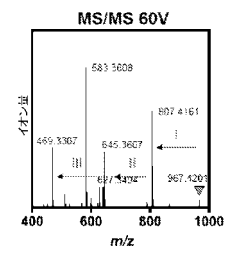
【 図 2 6 H 】



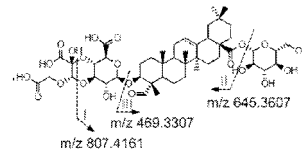
【 図 2 6 I 】



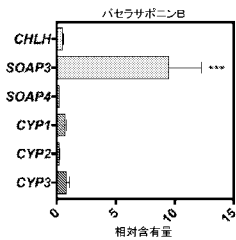
【 図 2 6 K 】



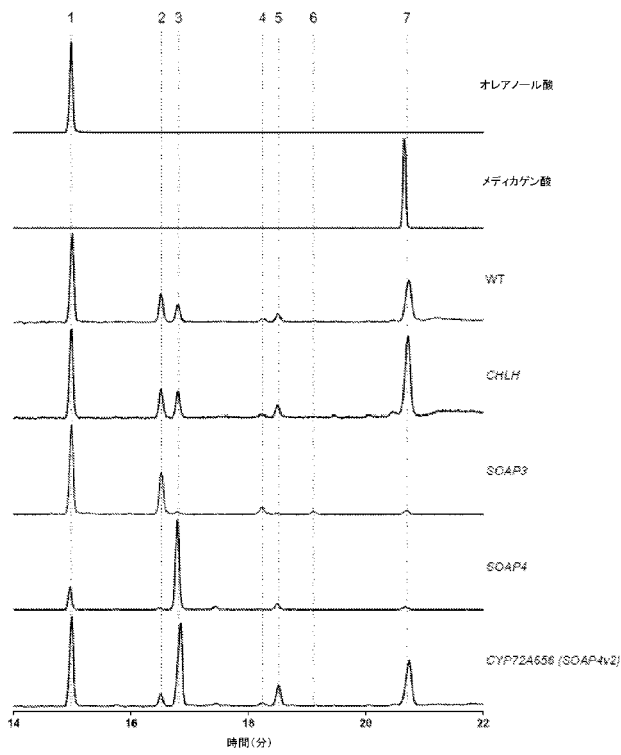
【 図 2 6 L 】



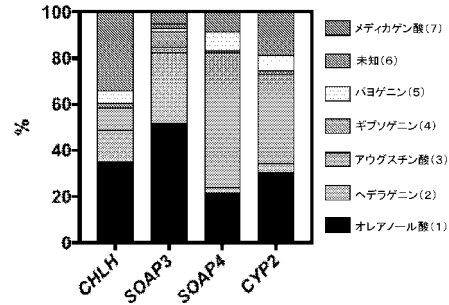
【 図 2 6 J 】



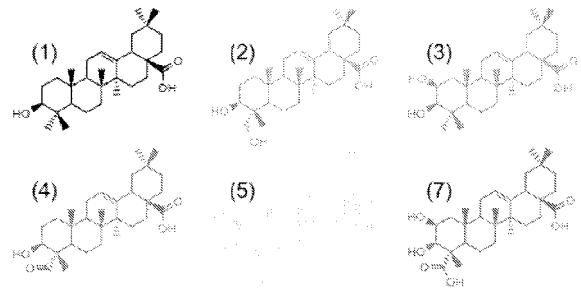
【 図 2 7 】



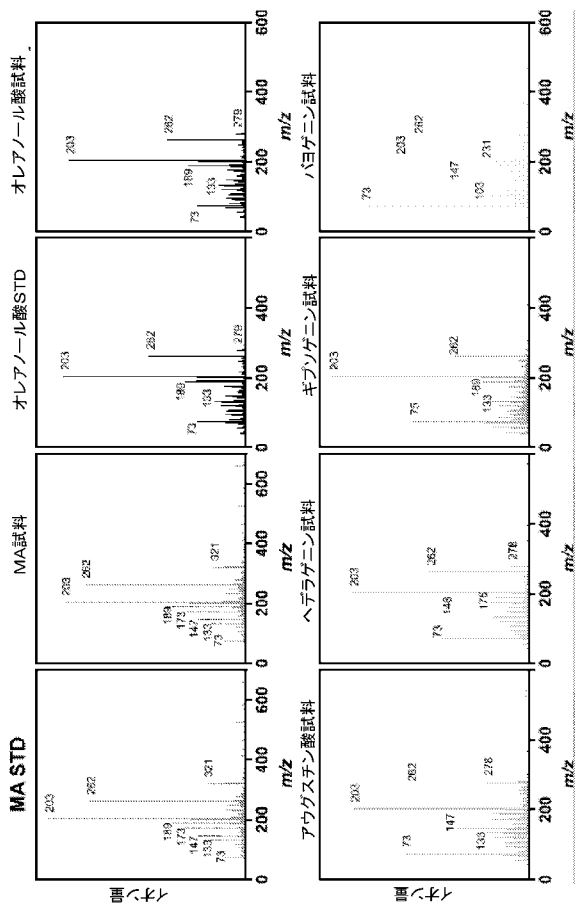
【 図 2 8 A 】



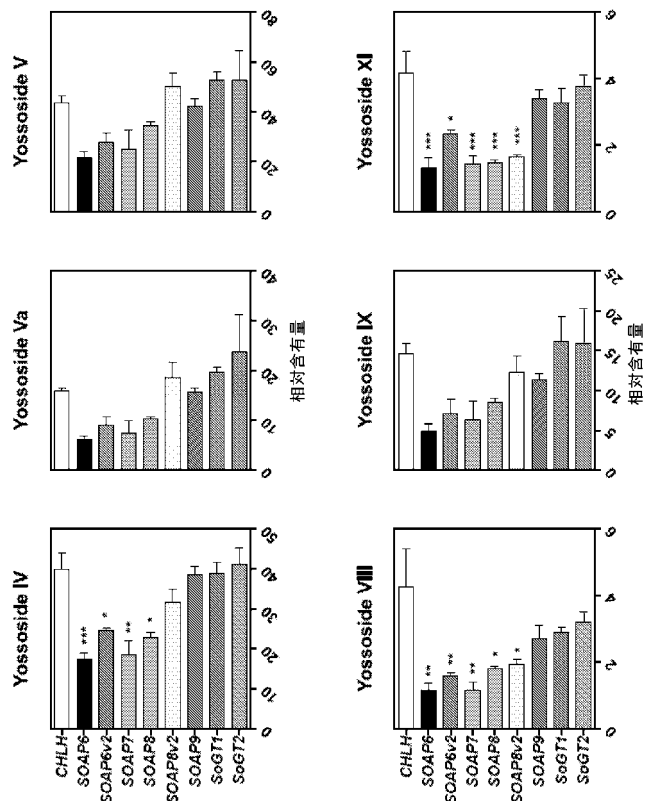
【 図 2 8 B 】



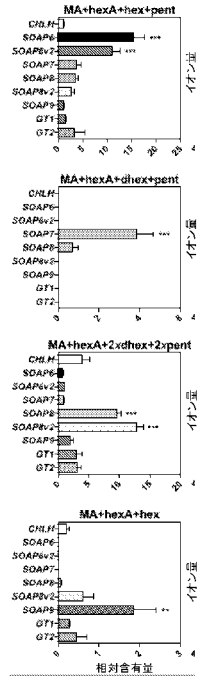
【 図 2 8 C 】



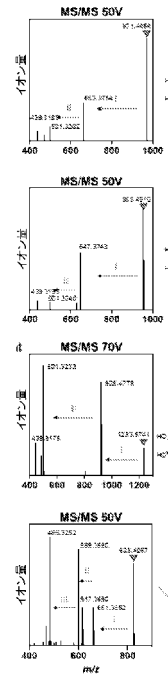
【 図 2 9 】



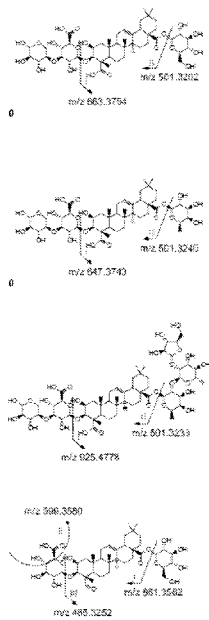
【図30A】



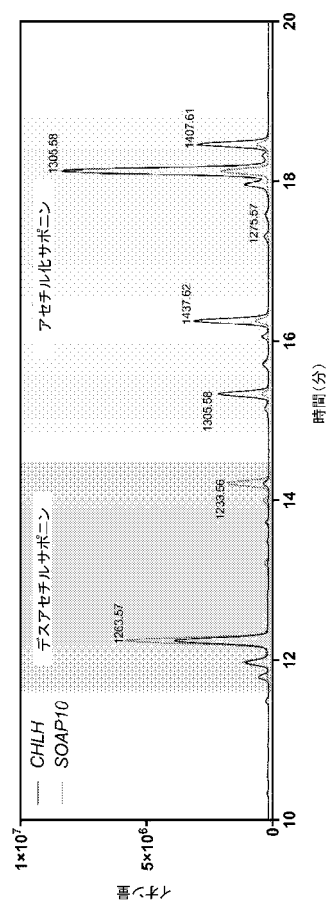
【図30B】



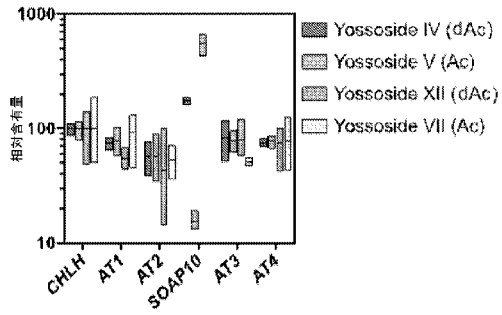
【図30C】



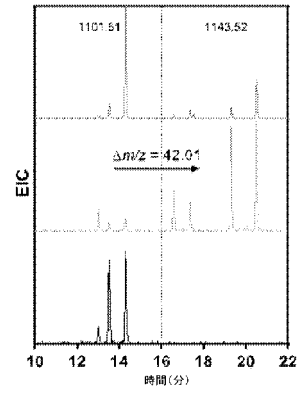
【図31A】



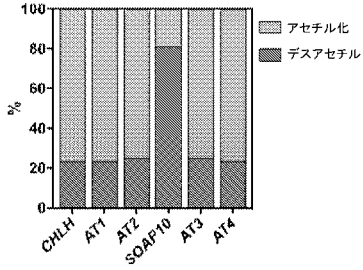
【 図 3 1 B 】



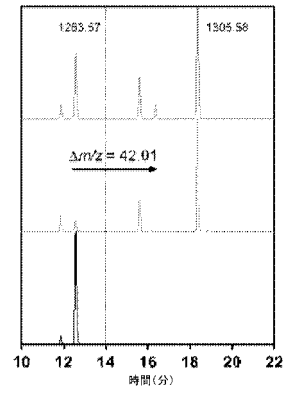
【 図 3 1 D 】



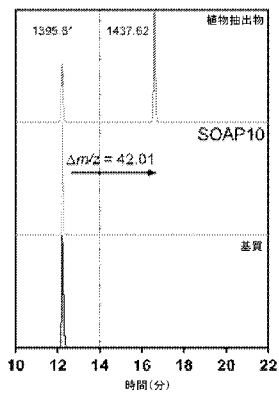
【 図 3 1 C 】



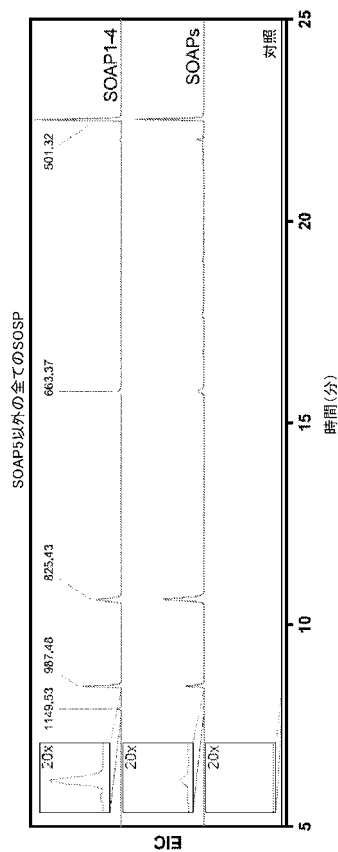
【 図 3 1 E 】



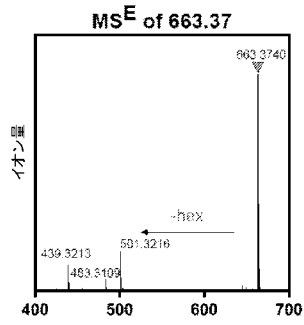
【 図 3 1 F 】



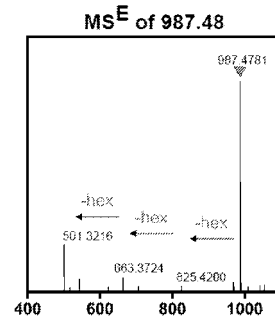
【 図 3 2 A 】



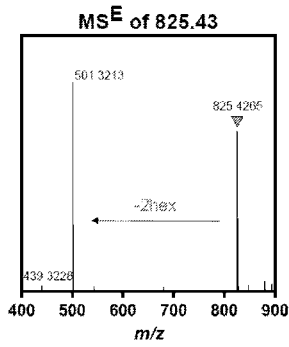
【 図 3 2 B 】



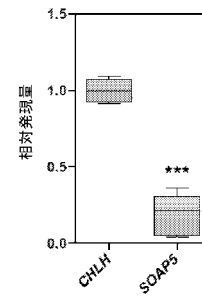
【 図 3 2 D 】



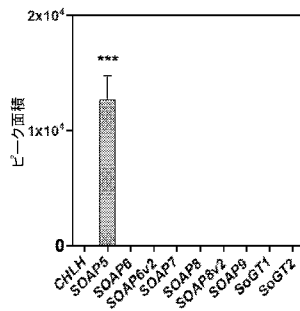
【 図 3 2 C 】



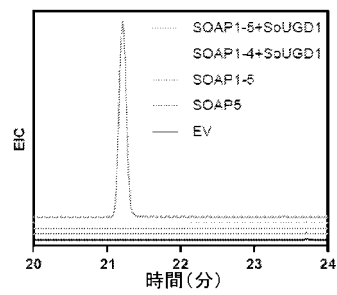
【 図 3 3 A 】



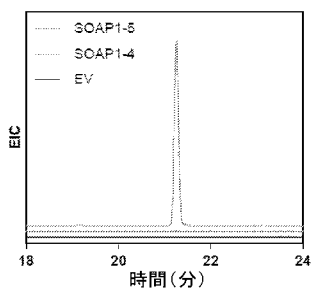
【 図 3 3 B 】



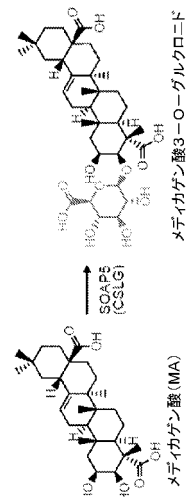
【 図 3 3 D 】



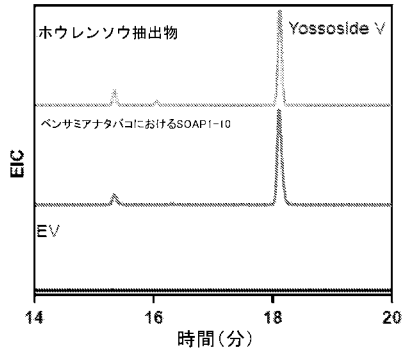
【 図 3 3 C 】



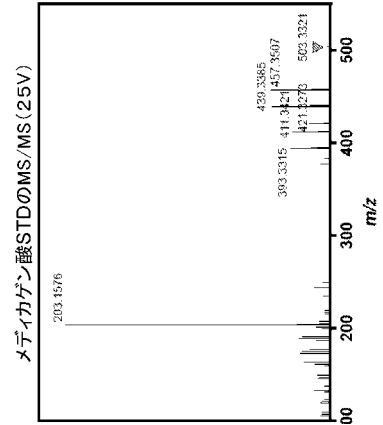
【 図 3 3 E 】



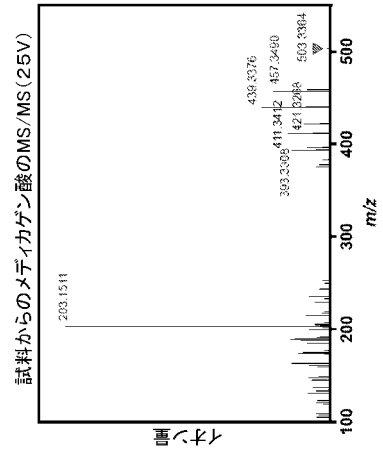
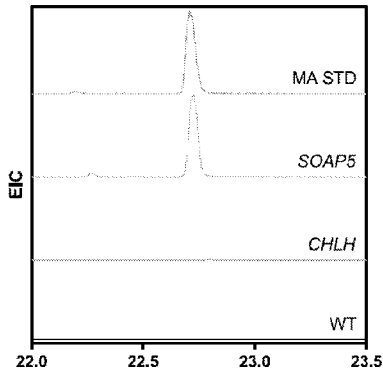
【 図 3 3 F 】



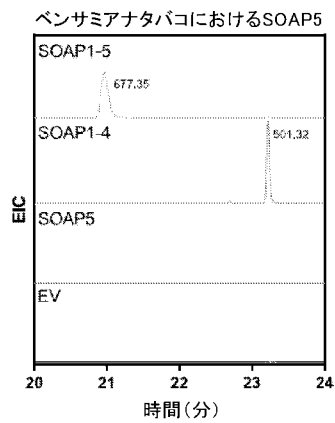
【 図 3 4 B 】



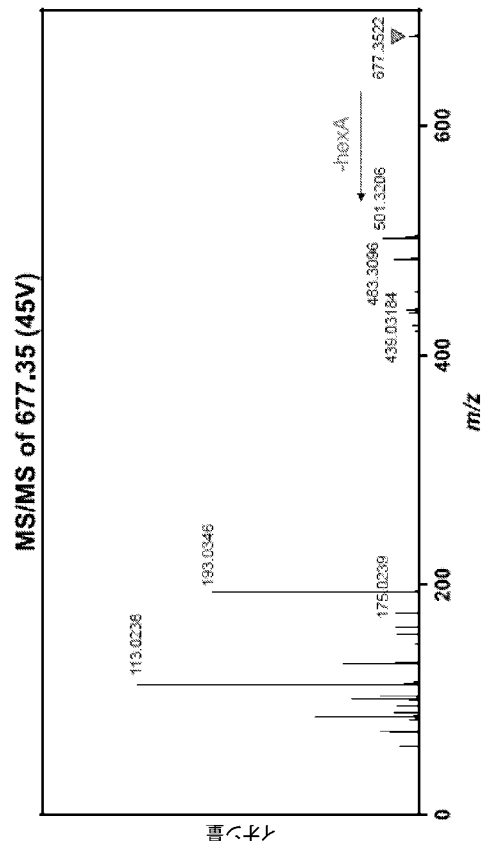
【 図 3 4 A 】



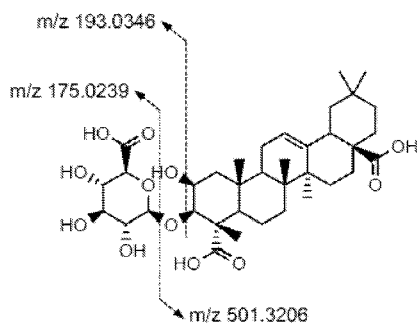
【 図 3 5 A 】



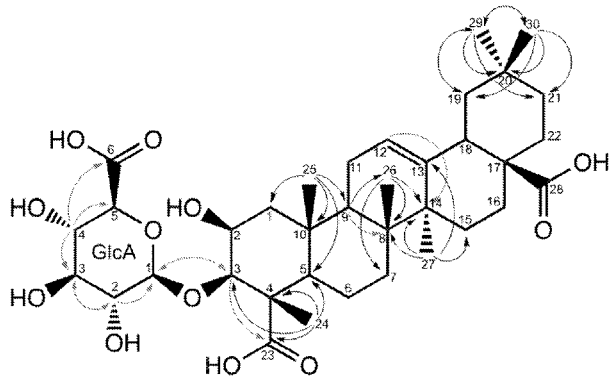
【 図 3 5 C 】



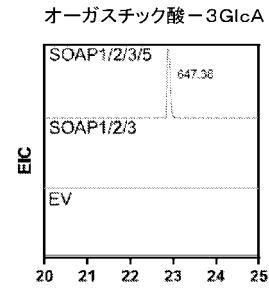
【 図 3 5 B 】



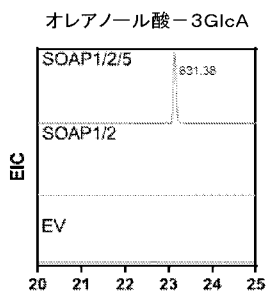
【 図 3 5 D 】



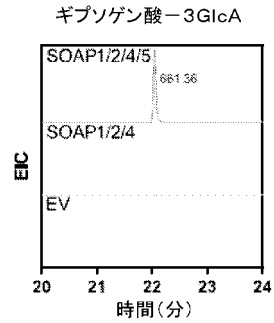
【 図 3 6 B 】



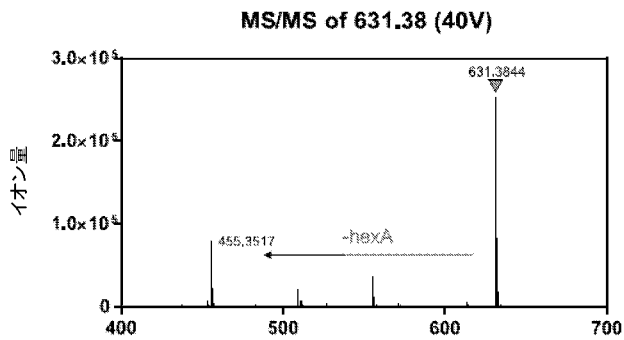
【 図 3 6 A 】



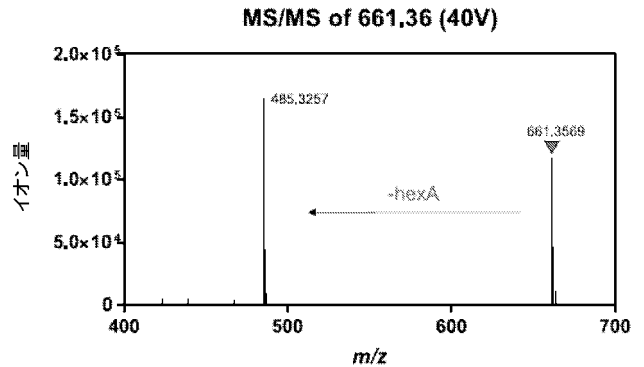
【 図 3 6 C 】



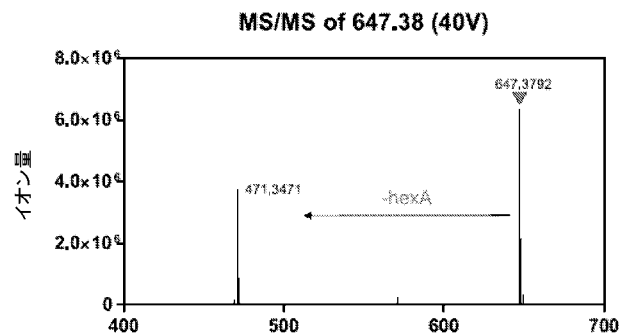
【 図 3 6 D 】



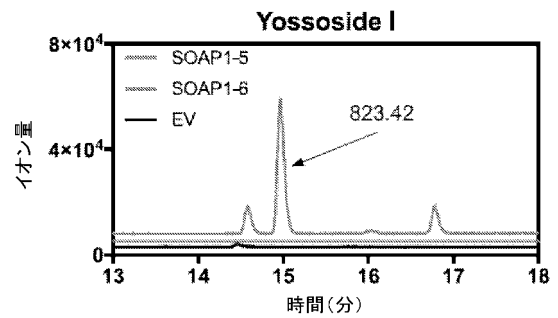
【 図 3 6 F 】



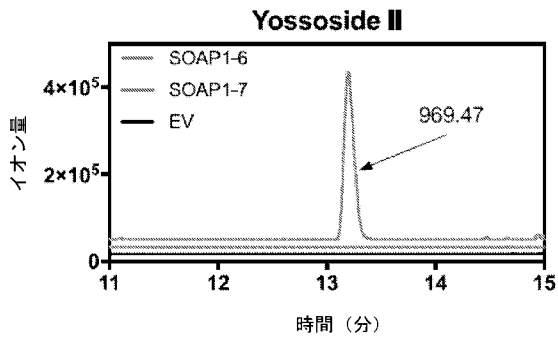
【 図 3 6 E 】



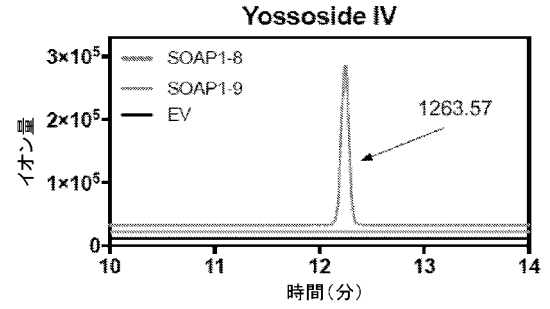
【 図 3 7 A 】



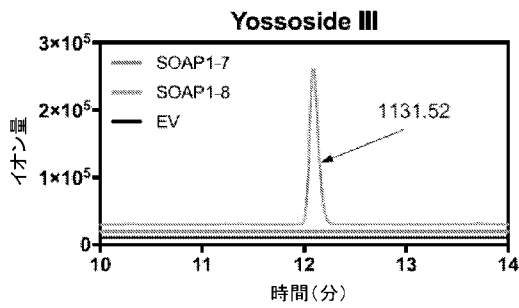
【 図 3 7 B 】



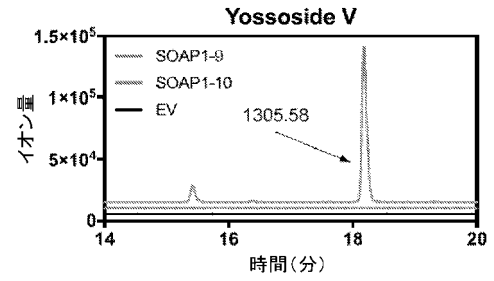
【 図 3 7 D 】



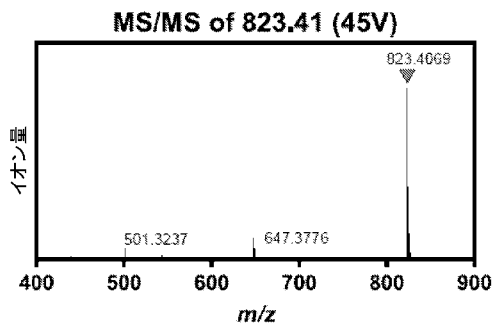
【 図 3 7 C 】



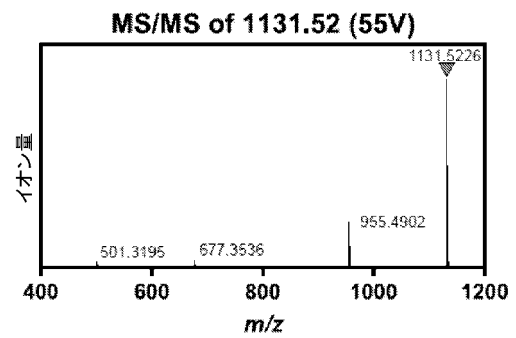
【 図 3 7 E 】



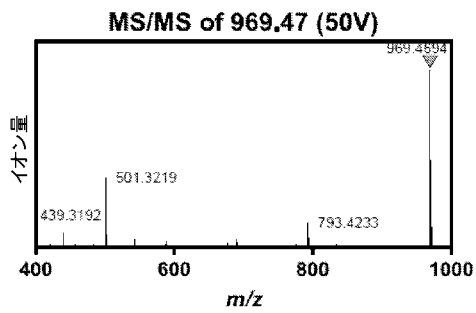
【 図 3 8 A 】



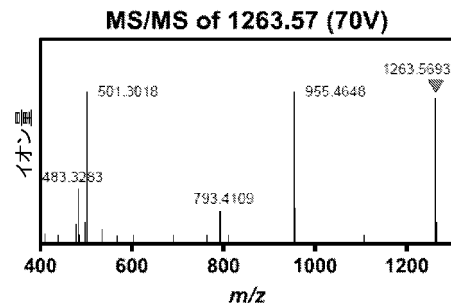
【 図 3 8 C 】



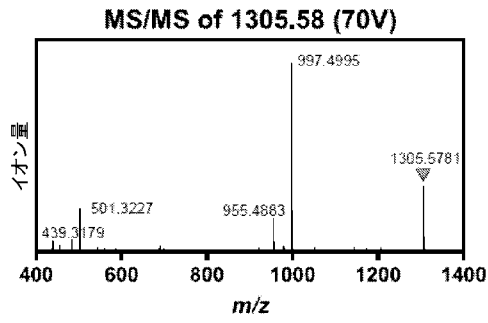
【 図 3 8 B 】



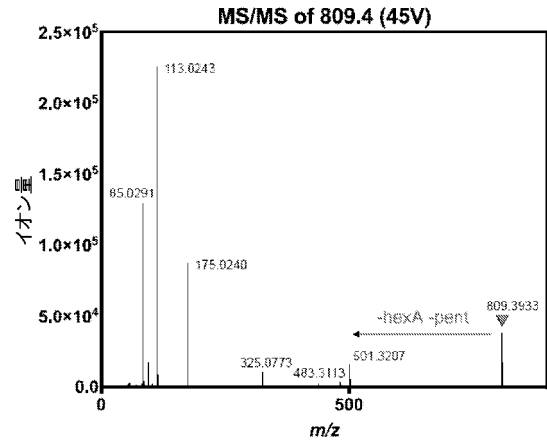
【 図 3 8 D 】



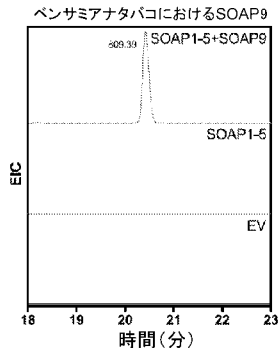
【 図 3 8 E 】



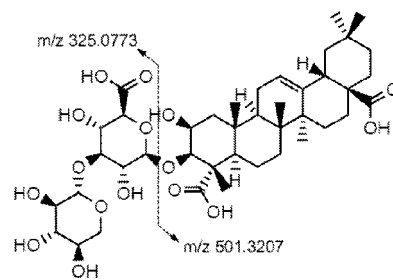
【 図 3 9 B 】



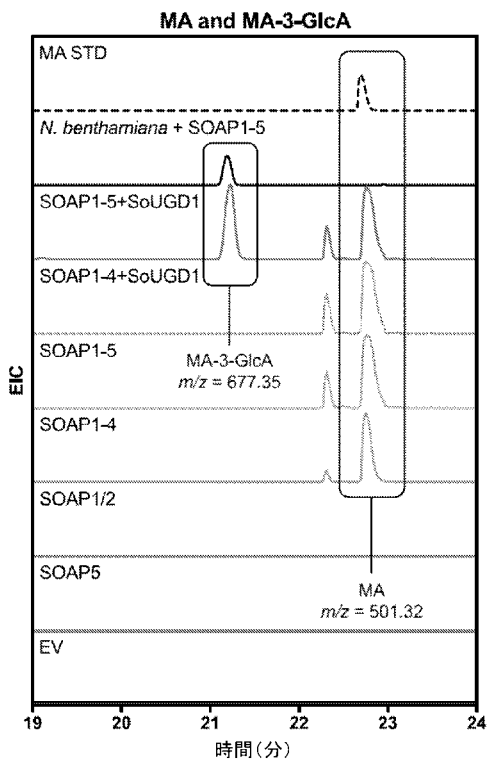
【 図 3 9 A 】



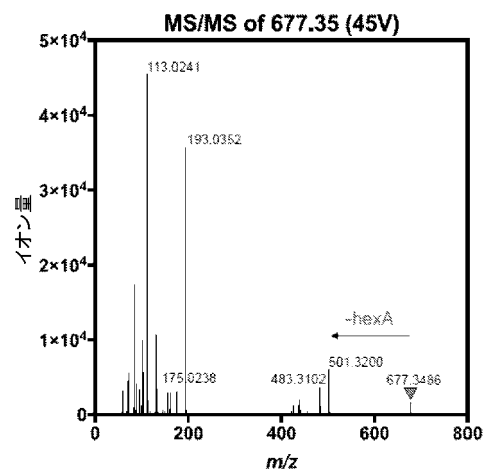
【 図 3 9 C 】



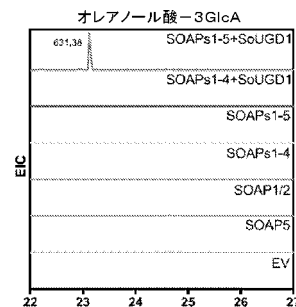
【 図 4 0 A 】



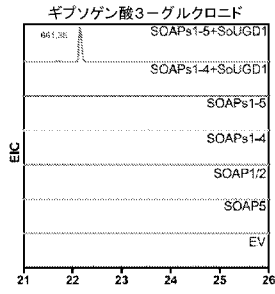
【 図 4 0 B 】



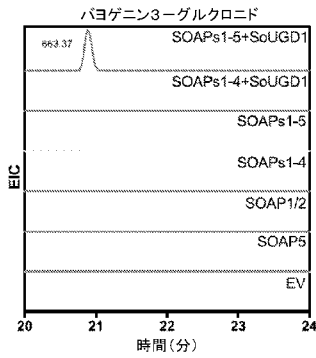
【 図 4 1 A 】



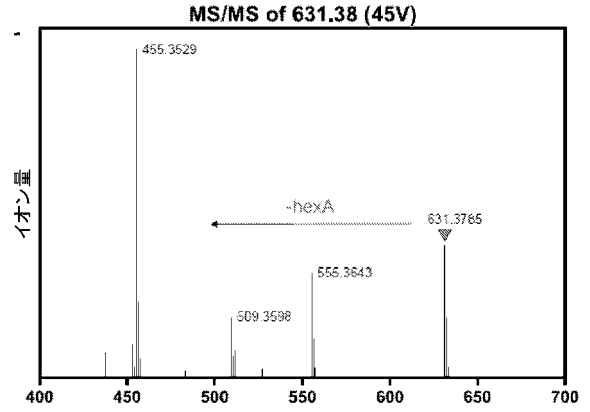
【 図 4 1 B 】



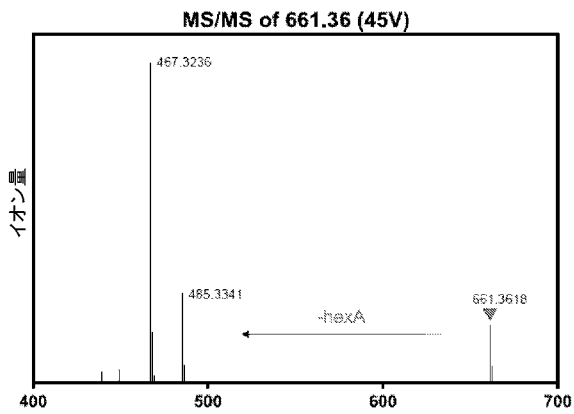
【 図 4 1 C 】



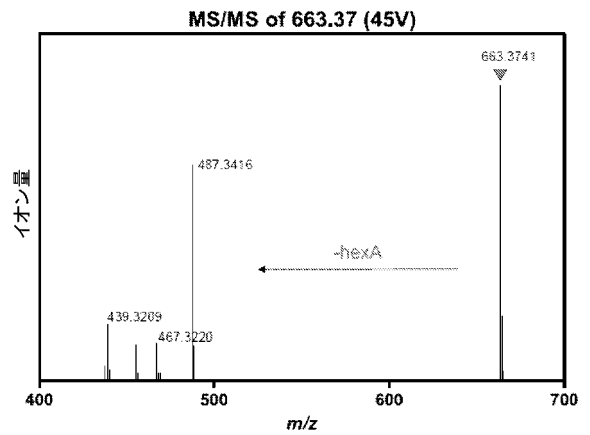
【 図 4 1 D 】



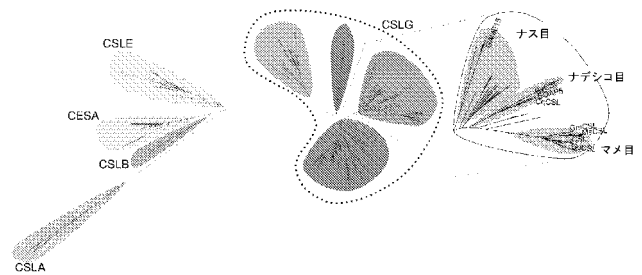
【 図 4 1 E 】



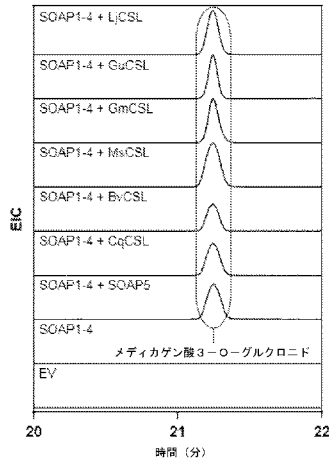
【 図 4 1 F 】



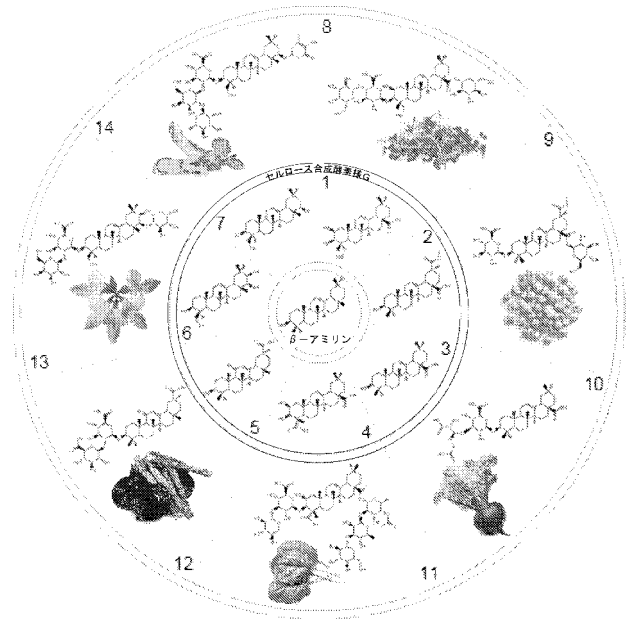
【 図 4 2 A 】



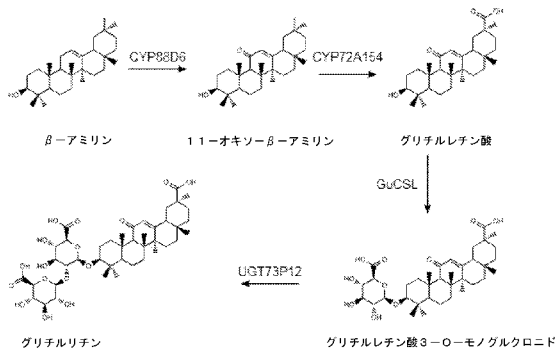
【 図 4 2 B 】



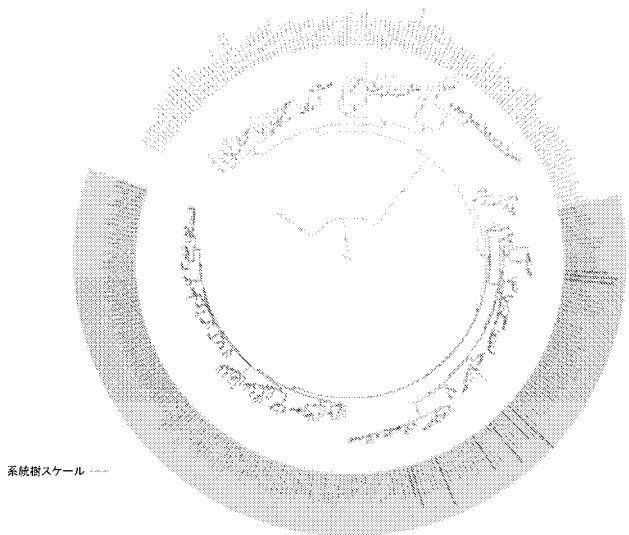
【 図 4 2 C 】



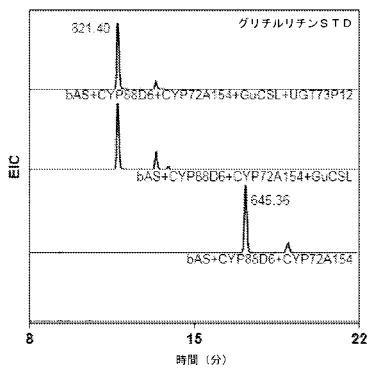
【 図 4 2 D 】



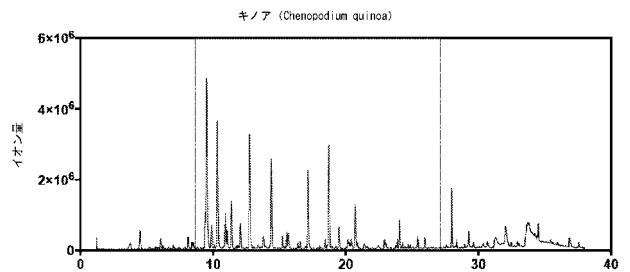
【 図 4 2 F 】



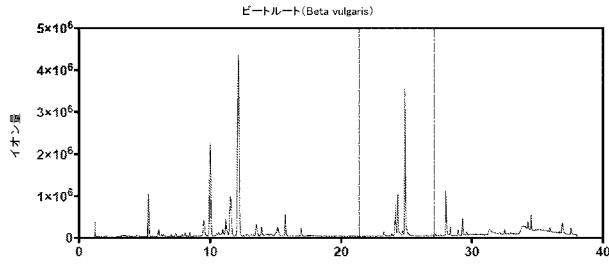
【 図 4 2 E 】



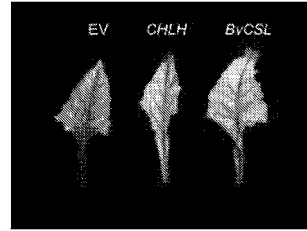
【 図 4 3 A 】



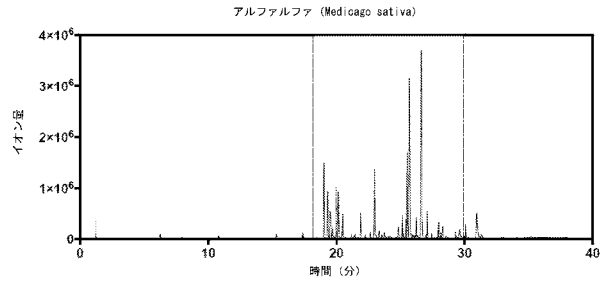
【 図 4 3 B 】



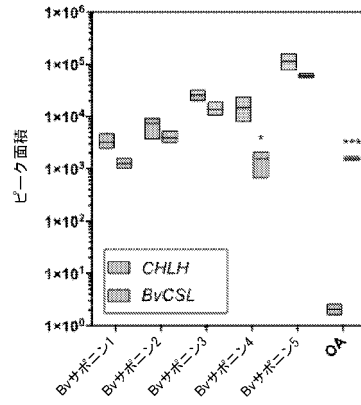
【 図 4 4 A 】



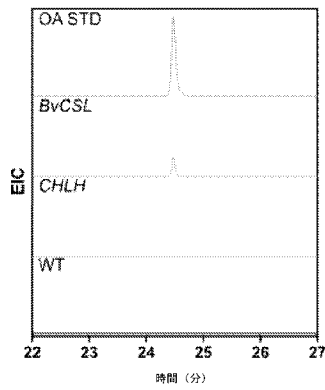
【 図 4 3 C 】



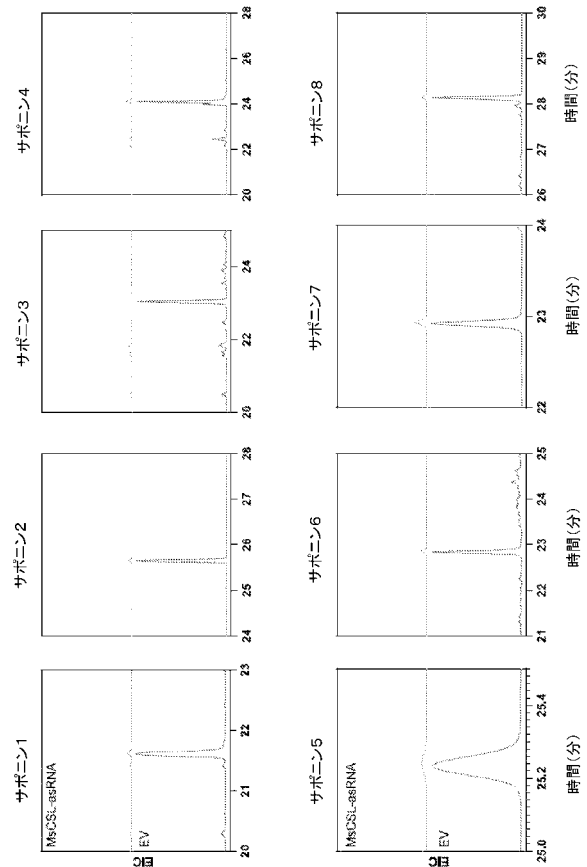
【 図 4 4 B 】



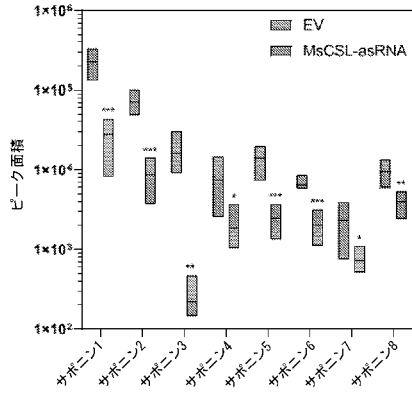
【 図 4 4 C 】



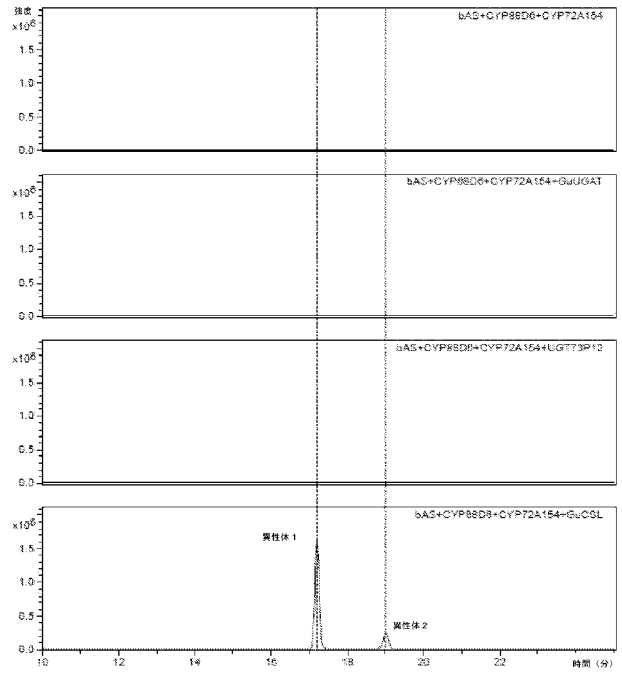
【 図 4 5 A 】



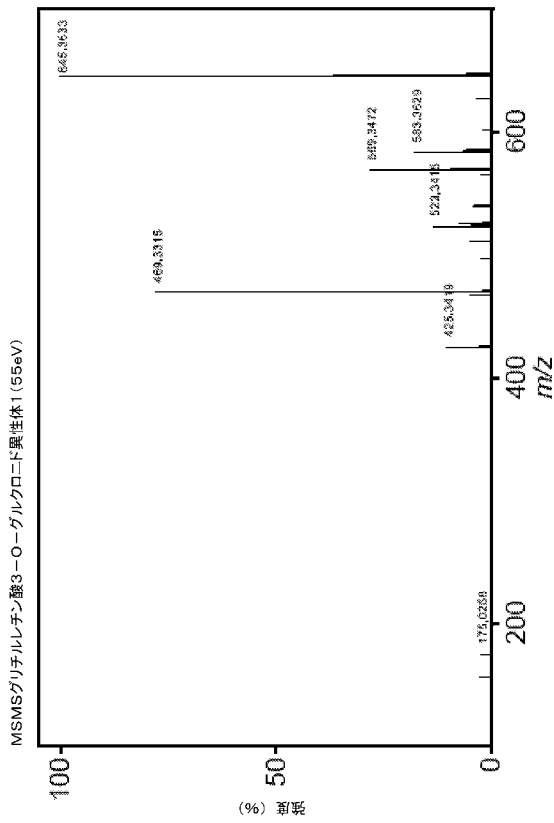
【 図 4 5 B 】



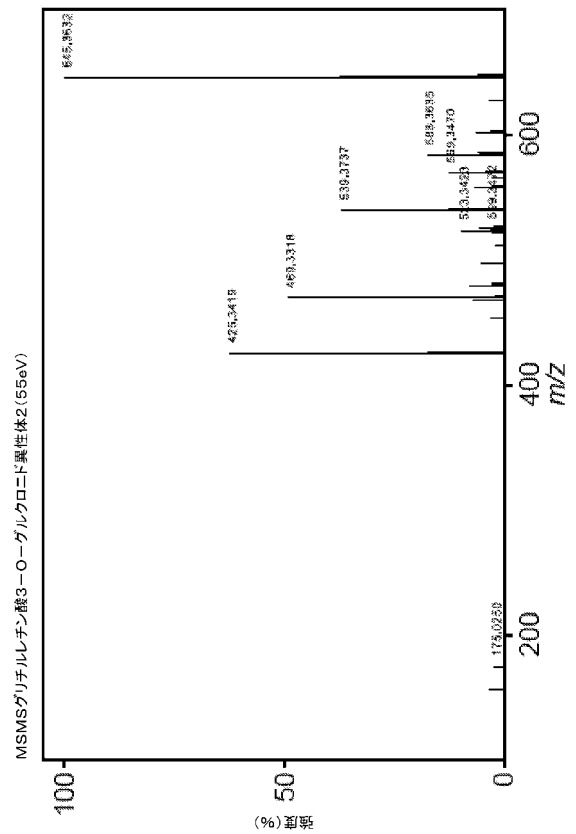
【 図 4 6 A 】



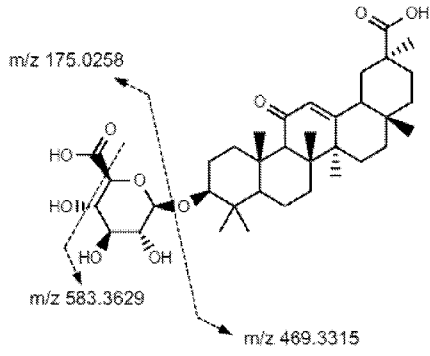
【 図 4 6 B 】



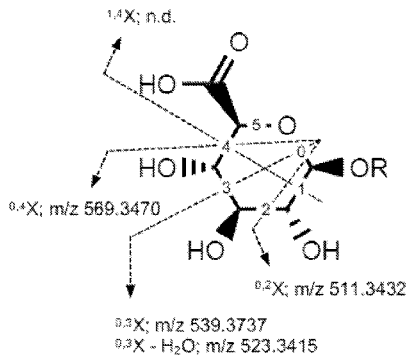
【 図 4 6 C 】



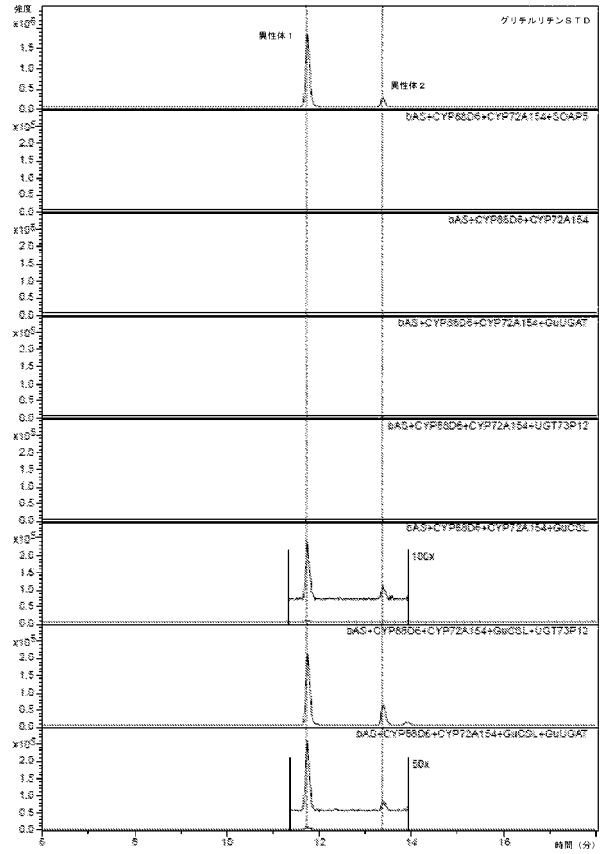
【 図 4 6 D 】



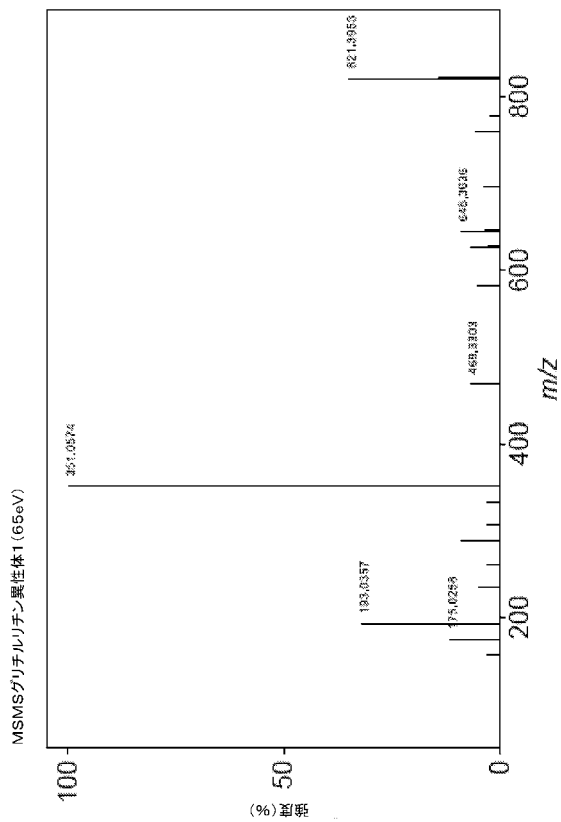
【 図 4 6 E 】



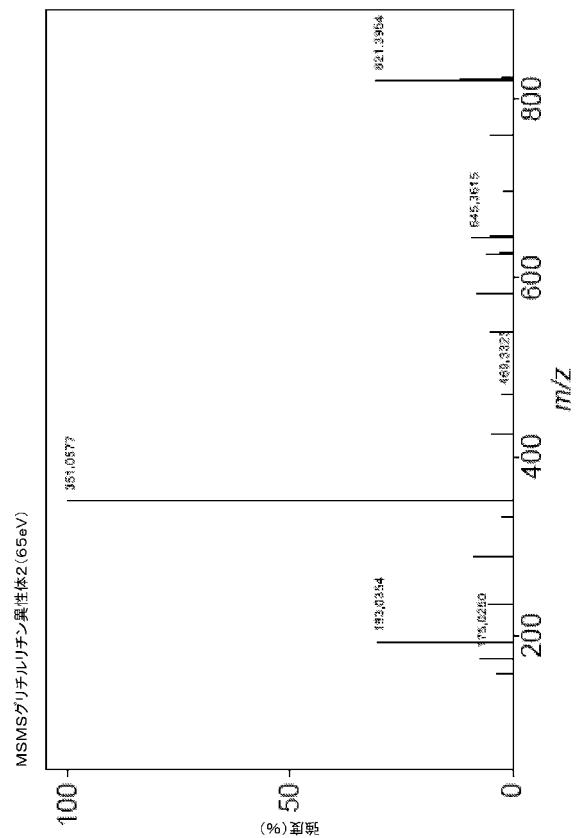
【 図 4 7 A 】



【 図 4 7 B 】

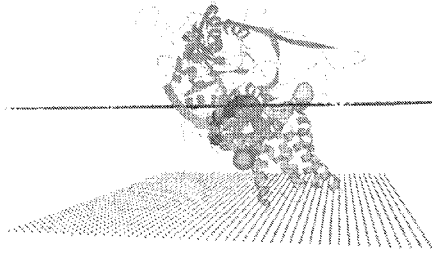


【 図 4 7 C 】

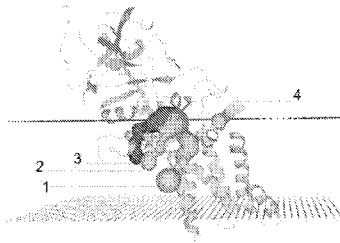




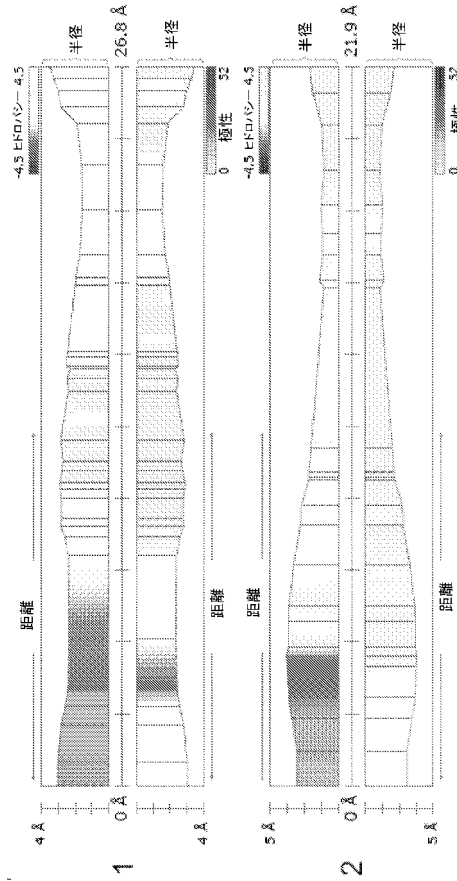
【図 49 A】



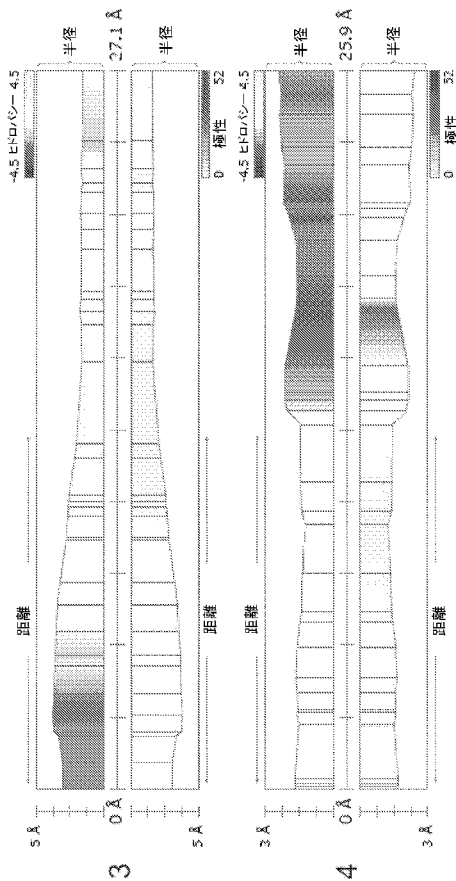
【図 49 B】



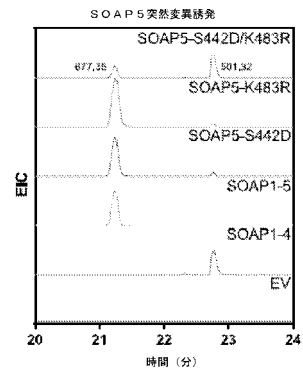
【図 49 C - 1】



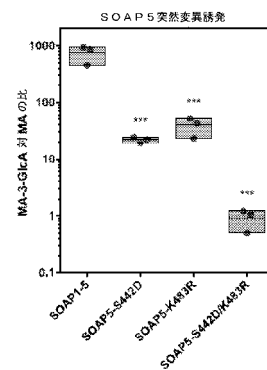
【図 49 C - 2】



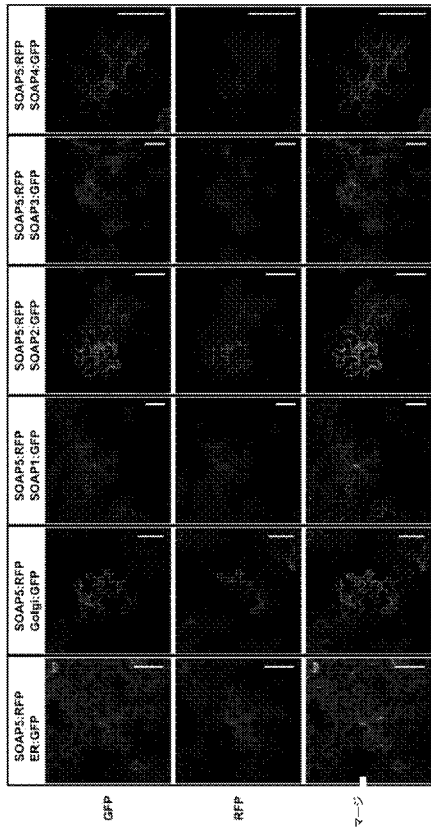
【図 50 A】



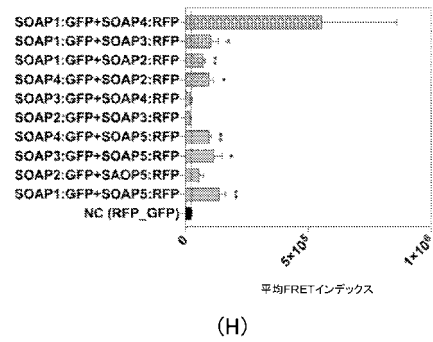
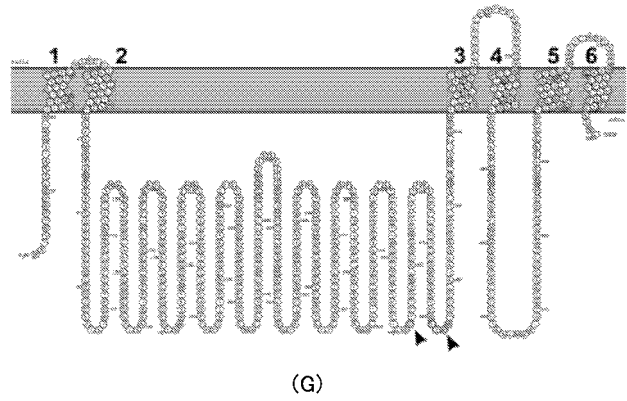
【図 50 B】



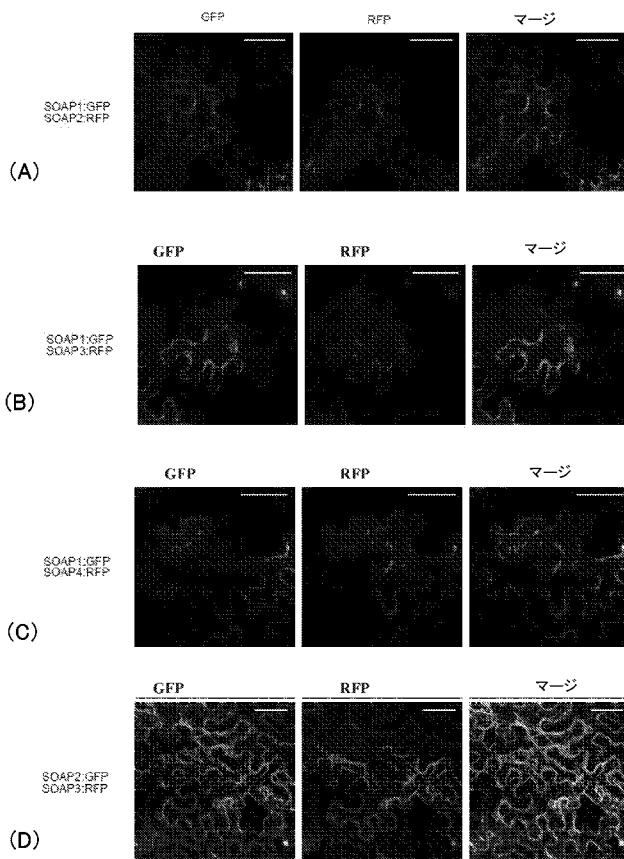
【 図 5 1 - 1 】



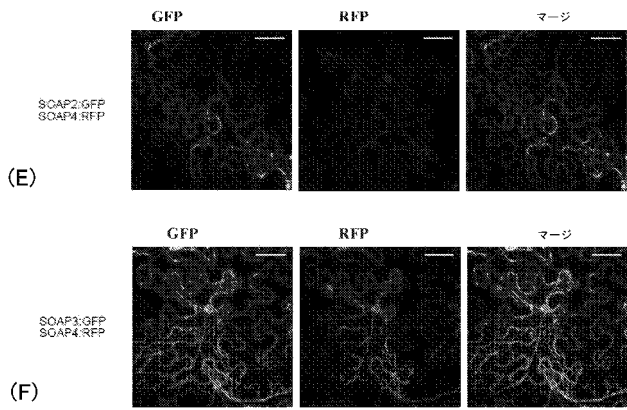
【 図 5 1 - 2 】



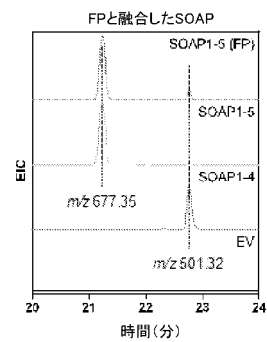
【 図 5 2 - 1 】



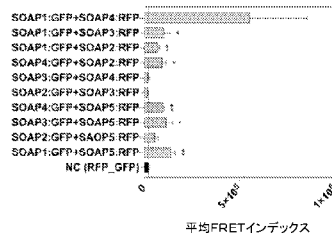
【 図 5 2 - 2 】



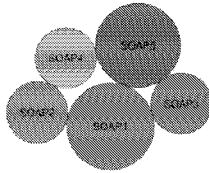
【 図 5 3 】



【図 5 4 A】



【図 5 4 B】



## 【配列表】

2022500023000001.app

## 【手続補正書】

【提出日】令和3年5月7日(2021.5.7)

## 【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

植物またはその少なくとも一部の少なくとも1つの細胞における、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を低減させる方法であって、

前記少なくとも1つの細胞の遺伝子を組み換える遺伝子組換えステップを含み、

前記遺伝子組換えステップは、

(a) セルロース合成酵素様 G (CSLG) 酵素をコードする少なくとも1つの遺伝子を標的とする少なくとも1つのサイレンシング分子を用いて前記少なくとも1つの植物細胞を形質転換するステップであって、前記CSLG酵素をコードする前記少なくとも1つの遺伝子の核酸配列は、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、または、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、もしくは105のいずれか1つに記載の核酸配列に

対して少なくとも55%の同一性及び少なくとも55%のカバレッジを有するそのホモログである、該ステップ、または、

(b) 前記CSLG酵素をコードする前記少なくとも1つの遺伝子の核酸配列を突然変異させるステップであって、前記CSLG酵素をコードする前記少なくとも1つの遺伝子の核酸配列は、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、または、配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、もしくは105のいずれか1つに記載の核酸配列に対して少なくとも55%の同一性及び少なくとも55%のカバレッジを有するそのホモログであり、前記突然変異が、(1)前記核酸配列内に1以上の点突然変異を導入すること、(2)前記核酸配列内に欠失を導入すること、(3)前記核酸配列内に挿入を導入すること、または、(4)それらの任意の組み合わせを含み、前記導入が、コード配列または非コード配列を突然変異させることを含む、該ステップと、を含み、

(c) 遺伝子が組み換えられた遺伝子組換え細胞における前記CSLG酵素をコードする遺伝子の発現は、対応する非組換え植物の細胞と比較して減少し、

前記遺伝子組換え細胞を含む植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物と比較して減少する、方法。

#### 【請求項2】

請求項1に記載の方法であって、

(a) 前記CSLG酵素をコードする前記少なくとも1つの遺伝子の前記核酸配列が配列番号30に記載されている核酸配列である場合、サイレンシング分子は、配列番号42に記載の核酸配列またはその相補的配列であるか、または、配列番号42に記載の核酸配列またはその相補的配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログである、

(b) 前記CSLG酵素をコードする前記少なくとも1つの遺伝子の前記核酸配列が配列番号34に記載されている核酸配列である場合、サイレンシング分子は、配列番号43に記載の核酸配列またはその相補的配列であるか、または、配列番号43に記載の核酸配列またはその相補的配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログである、

(c) 前記CSLG酵素をコードする前記少なくとも1つの遺伝子の前記核酸配列が配列番号38に記載されている核酸配列である場合、サイレンシング分子は、配列番号44に記載の核酸配列またはその相補的配列であるか、または、配列番号44に記載の核酸配列またはその相補的配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログである、

(d) 前記CSLG酵素をコードする前記少なくとも1つの遺伝子の前記核酸配列が配列番号65または配列番号93に記載されている核酸配列である場合、サイレンシング分子は、配列番号106に記載の核酸配列またはその相補的配列であるか、または、配列番号106に記載の核酸配列またはその相補的配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログである、

(e) 前記CSLG酵素をコードする前記少なくとも1つの遺伝子の前記核酸配列が配列番号95に記載されている核酸配列である場合、サイレンシング分子は、配列番号107に記載の核酸配列またはその相補的配列であるか、または、配列番号107に記載の核酸配列またはその相補的配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログである、方法。

#### 【請求項3】

請求項1または2に記載の方法であって、

前記ステロイド系アルカロイド、前記ステロイド系サポニン、または前記トリテルペノ

イド系サポニンは、毒素または苦味化合物を含むか、ホルモン模倣特性を有するか、またはそれらの組み合わせである、方法。

**【請求項 4】**

セルロース合成酵素様 G (CSLG) 酵素をコードする少なくとも 1 つの異種遺伝子の発現が、対応する非組換え細胞と比較して増加した遺伝子組換え細胞であって、

前記 CSLG 酵素のアミノ酸配列は、配列番号 31、33、35、37、39、41、66、81、94、96、98、100、102、または 104 のいずれか 1 つに記載のアミノ酸配列であるか、または、配列番号 31、33、35、37、39、41、66、81、94、96、98、100、102、または 104 のいずれか 1 つに記載のアミノ酸配列に対して少なくとも 55% の相同性及び少なくとも 55% のカバレッジを有するそのホモログであり、

少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、

少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、

少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体

の含有量が、対応する非組換え細胞と比較して増加した、遺伝子組換え細胞。

**【請求項 5】**

セルロース合成酵素様 G (CSLG) 酵素をコードする少なくとも 1 つの CSLG 遺伝子の発現が、対応する非組換え植物における前記 CSLG 遺伝子の発現と比較して改変された少なくとも 1 つの細胞を含む遺伝子組換え植物であって、

前記 CSLG 酵素のアミノ酸配列は、配列番号 31、33、35、37、39、41、66、81、94、96、98、100、102、または 104 のいずれか 1 つに記載のアミノ酸配列であるか、または、配列番号 31、33、35、37、39、41、66、81、94、96、98、100、102、または 104 のいずれか 1 つに記載のアミノ酸配列に対して少なくとも 55% の相同性及び少なくとも 55% のカバレッジを有するそのホモログであり、

少なくとも 1 つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、

少なくとも 1 つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、

少なくとも 1 つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体

の含有量が、対応する非組換え植物と比較して改変された、遺伝子組換え植物。

**【請求項 6】**

遺伝子組換え細胞内にステロイド系アルカロイド、ステロイド系サポニン、またはトリテルペノイド系サポニンを産生する方法であって、

(a) 前記細胞内に少なくとも 1 つの異種遺伝子を導入する導入ステップであって、前記少なくとも 1 つの異種遺伝子がセルロース合成酵素様 G (CSLG) 酵素をコードしており、前記コードされた CSLG 酵素のアミノ酸配列は、配列番号 31、33、35、37、39、41、66、81、94、96、98、100、102、または 104 のいずれか 1 つに記載のアミノ酸配列であるか、または、配列番号 31、33、35、37、39、41、66、81、94、96、98、100、102、または 104 のいずれか 1 つに記載のアミノ酸配列に対して少なくとも 55% の相同性及び少なくとも 55% のカバレッジを有するそのホモログであり、前記少なくとも 1 つの異種遺伝子が任意選択でベクターに含まれている、該ステップと、

(b) 前記細胞内で前記少なくとも 1 つの異種遺伝子を発現させる発現ステップと、を含み、

前記細胞は、少なくとも 1 つのステロイド系アルカロイド、少なくとも 1 つのステロイ

ド系サポニン、または少なくとも1つのトリテルペノイド系サポニンの含有量が、対応する非組換え細胞と比較して増加する、方法。

**【請求項7】**

植物またはその少なくとも一部の少なくとも1つの細胞における、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量を増加させる方法であって、

前記少なくとも1つの細胞の遺伝子を組み換える遺伝子組換えステップを含み、

前記遺伝子組換えステップは、

(a) セルロース合成酵素様 G (CSLG) 酵素をコードする少なくとも1つの遺伝子の核酸配列を突然変異させるステップであって、前記突然変異が、(1) 前記核酸配列内に1以上の点突然変異を導入すること、(2) 前記核酸配列内に欠失を導入すること、(3) 前記核酸配列内に挿入を導入すること、または、(4) それらの任意の組み合わせを含み、前記導入が、コード配列または非コード配列を突然変異させることを含む、該ステップと、を含み、

(b) 前記CSLG酵素をコードする前記核酸配列を発現させるステップと、を含み、  
遺伝子が組み換えられた遺伝子組換え細胞を含む植物は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の含有量が、対応する非組換え植物にと比較して増加する、方法。

**【請求項8】**

前記少なくとも1つの異種遺伝子をコードする核酸配列は、

配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列であるか、または、  
配列番号30、32、34、36、38、40、65、80、93、95、97、99、101、103、または105のいずれか1つに記載の核酸配列に対して少なくとも55%の相同性及び少なくとも55%のカバレッジを有するそのホモログである、

請求項4に記載の遺伝子組換え細胞、請求項5に記載の遺伝子組換え植物、請求項6に記載の方法、または、請求項7に記載の方法。

**【請求項9】**

前記ステロイド系アルカロイド、前記ステロイド系サポニン、前記トリテルペノイド系サポニン、または、それらの誘導体、その代謝物、またはその生合成中間体は、甘味料、発泡剤、乳化剤、防腐剤、抗発癌剤、コレステロール低下剤、抗炎症剤、抗酸化剤、生物学的アジュバント、抗微生物剤、殺虫剤、摂食阻害剤、抗真菌剤、またはそれらの任意の組み合わせを含む、

請求項4に記載の遺伝子組換え細胞、請求項5に記載の遺伝子組換え植物、請求項6に記載の方法、または、請求項7に記載の方法。

**【請求項10】**

前記細胞は、植物細胞、酵母細胞、または藻類細胞を含む、

請求項4、8及び9のいずれかに記載の遺伝子組換え細胞、または、請求項3に記載の方法。

**【請求項11】**

前記遺伝子組換え細胞は、酵素をコードする少なくとも1つの追加の異種遺伝子をさらに発現し、

前記酵素は、サポニン - アミリン合成酵素、シトクロムP450、グリコシルトランスフェラーゼ、アシルトランスフェラーゼ、UDP - グルコース6 - デヒドロゲナーゼ1、及びそれらの任意の組み合わせからなる群から選択され、

前記少なくとも1つの異種遺伝子が任意選択でベクターに含まれている、

請求項 4、及び 8～10 のいずれかに記載の遺伝子組換え細胞、または、請求項 6 に記載の方法。

【請求項 12】

前記少なくとも 1 つの追加の異種遺伝子によりコードされる前記酵素、及び、そのアミノ酸配列は、

(a) サポニン - アミリン合成酵素、及び、配列番号 48 に記載のアミノ酸配列、または配列番号 48 に記載のアミノ酸配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログ、

(b) シトクロム P 450、及び、配列番号 49、52、または 54 のいずれか 1 つに記載のアミノ酸配列、または、配列番号 49、52、または 54 のいずれか 1 つに記載のアミノ酸配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログ、

(c) グリコシルトランスフェラーゼ、及び、配列番号 56、58、60、または 62 のいずれか 1 つに記載のアミノ酸配列、または、配列番号 56、58、60、または 62 のいずれか 1 つに記載のアミノ酸配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログ、

(d) アセチルトランスフェラーゼ、及び、配列番号 64 に記載のアミノ酸配列、または配列番号 64 に記載のアミノ酸配列に対して少なくとも 80% の同一性及び 80% のカバレッジを有するそのホモログ、

(e) UDP - グルコース 6 - デヒドロゲナーゼ 1、及び、配列番号 75 に記載のアミノ酸配列、または配列番号 75 に記載のアミノ酸配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログ、または、

(f) 上記の (a)、(b)、(c)、(d) 及び (e) の任意の組み合わせであり、前記少なくとも 1 つの追加の異種遺伝子によりコードされる前記酵素、及び、該異種遺伝子をコードする核酸配列は、

(g) サポニン - アミリン合成酵素、及び、配列番号 45 に記載の核酸配列、または、配列番号 45 に記載の核酸配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログ、

(h) シトクロム P 450、及び、配列番号 46、51、または 53 のいずれか 1 つに記載の核酸配列、または、配列番号 46、51、または 53 のいずれか 1 つに記載の核酸配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログ、

(i) グリコシルトランスフェラーゼ、及び、配列番号 55、57、59、または 61 のいずれか 1 つに記載の核酸配列、または、配列番号 55、57、59、または 61 のいずれか 1 つに記載の核酸配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログ、

(j) アセチルトランスフェラーゼ、及び、配列番号 63 に記載の核酸配列であるか、または、配列番号 63 に記載の核酸配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログ、

(k) UDP - グルコース 6 - デヒドロゲナーゼ 1、及び、配列番号 74 に記載の核酸配列であるか、または、配列番号 74 に記載の核酸配列に対して少なくとも 80% の同一性及び少なくとも 80% のカバレッジを有するそのホモログ、または、

(l) 上記の (a)、(b)、(c)、(d) 及び (e) の任意の組み合わせである、請求項 4、及び 8～10 のいずれかに記載の遺伝子組換え細胞、または、請求項 6、8、及び 10～11 のいずれかに記載の方法。

【請求項 13】

(a) 前記 CSLG 遺伝子は、異種 CSLG 遺伝子を含み、前記発現の改変は、前記異種 CSLG 遺伝子のデノボ発現を含む、または、

(b) 前記 CSLG 遺伝子は、内在性 CSLG 遺伝子を含み、前記内在性 CSLG 遺伝子は、突然変異を含み、前記突然変異は、少なくとも 1 以上の点突然変異、挿入、欠失、

またはそれらの任意の組み合わせを含み、

前記CSLG酵素は、安定性の向上、活性の増加、またはその両方を含み、前記含有量の改変は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が、対応する非組換え植物と比較して増加すること、または、

前記CSLG酵素は、安定性の低下、活性の低下、またはその両方を含み、前記含有量の改変は、少なくとも1つのステロイド系アルカロイド、その誘導体、その代謝物、またはその生合成中間体、少なくとも1つのステロイド系サポニン、その誘導体、その代謝物、またはその生合成中間体、または、少なくとも1つのトリテルペノイド系サポニン、その誘導体、その代謝物、またはその生合成中間体の量が、対応する非組換え植物と比較して減少することである、

請求項5、8及び9のいずれかに記載の遺伝子組換え植物。

【請求項14】

前記少なくとも1つのCSLG遺伝子の前記発現の改変は、前記少なくとも1つのCSLG遺伝子を標的とするサイレンシング分子を導入することにより生じ、かつ、

(a) 前記CSLG遺伝子をコードする核酸配列が配列番号30に記載されている核酸配列である場合、サイレンシング分子は、配列番号42に記載の核酸配列またはその相補的配列であるか、または、配列番号42に記載の核酸配列またはその相補的配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログである、

(b) 前記CSLG遺伝子をコードする核酸配列が配列番号34に記載されている核酸配列である場合、サイレンシング分子は、配列番号43に記載の核酸配列またはその相補的配列であるか、または、配列番号43に記載の核酸配列またはその相補的配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログである、

(c) 前記CSLG遺伝子をコードする核酸配列が配列番号38に記載されている核酸配列である場合、サイレンシング分子は、配列番号44に記載の核酸配列またはその相補的配列であるか、または、配列番号44に記載の核酸配列またはその相補的配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログである、

(d) 前記CSLG遺伝子をコードする核酸配列が配列番号65または配列番号93に記載されている核酸配列である場合、サイレンシング分子は、配列番号106に記載の核酸配列またはその相補的配列であるか、または、配列番号106に記載の核酸配列またはその相補的配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログである、

(e) 前記CSLG遺伝子をコードする核酸配列が配列番号95に記載されている核酸配列である場合、サイレンシング分子は、配列番号107に記載の核酸配列またはその相補的配列であるか、または、配列番号107に記載の核酸配列またはその相補的配列に対して少なくとも80%の同一性及び少なくとも80%のカバレッジを有するそのホモログである、

請求項5、8、9及び13のいずれかに記載の遺伝子組換え植物。

【請求項15】

(a) 前記ステロイド系アルカロイドは、エスクレオシド、デヒドロエスクレオシド、トマチン、トマチン、デヒドロトマチン、カコニン、ソラニン、ソラニン、ソラマルギン、またはそれらの任意の組み合わせを含む、

(b) 前記ステロイド系サポニンは、ウットロシドB、トマトシド、またはそれらの任意の組み合わせを含む、

(c) 前記トリテルペノイド系サポニンは、メディカゲン酸3-O-グルクロニド(M

A - 3 - G l c A ) ( 化合物 6 )、化合物 7、化合物 8、化合物 9、化合物 1 0、化合物 1 1、グリチルリチン ( 化合物 1 4 )、グリチルレチン酸 3 - O - モノグルクロニド ( 化合物 1 5 )、パヨゲニン ( 化合物 2 5 )、パヨゲニン - ヘキサ A - ヘキサ - ヘキサ ( 化合物 3 1 )、セルジャン酸 ( 化合物 2 6 )、セルジャン酸 - ヘキサ A - ヘキサ ( 化合物 3 2 )、ソヤサボゲノール A ( 化合物 2 9 )、ソヤサボゲノール B ( 化合物 3 0 )、ソヤサボゲノール A - ヘキサ A - ヘキサ - ペント ( 化合物 3 4 )、ソヤサボニン V I ( 化合物 3 5 )、ベタブルガロシド I V ( 化合物 3 3 )、ヘデラゲニン - 3 G l c A、ギブソゲニン - 3 G l c A、ギブソゲン酸 - 3 G l c A、Q S - 2 1 アジュバント、またはそれらの任意の組み合わせを含む、

( d ) 前記トリテルペノイド系サボニンの前記生合成中間体は、化合物 1、化合物 2、メディカゲン酸 ( 化合物 5 )、オレアノール酸 ( 化合物 3 )、オレアノール酸 - 3 G l c A、アウグスチン酸 ( 化合物 4 )、アウグスチン酸 - 3 G l c A、グリチルレチン酸 ( 化合物 1 3 )、またはそれらの任意の組み合わせを含む、または、

( e ) 上記の ( a )、( b )、( c )、及び ( d ) の任意の組み合わせである、

請求項 4 に記載の遺伝子組換え細胞、請求項 5 に記載の遺伝子組換え植物、請求項 6 に記載の方法、請求項 7 に記載の方法、または、請求項 1 に記載の方法。

## 【 国際調査報告 】

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/IL2019/051000

A. CLASSIFICATION OF SUBJECT MATTER INV. C12N9/10 C12N15/82 A01H5/02 A01H5/10 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) C12N A01H		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, Sequence Search, BIOSIS, EMBASE, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 2 982 752 A1 (RIKEN [JP]) 10 February 2016 (2016-02-10) page 3 - page 13; claims 7-12,14,15; examples 6-15	1-62
X	----- US 2011/219476 A1 (ONO EIICHIRO [JP] ET AL) 8 September 2011 (2011-09-08)  the whole document	1,7, 9-17,19, 30,32-37
A	----- WO 2014/195944 A1 (YEDA RES & DEV [IL]) 11 December 2014 (2014-12-11) the whole document  ----- -/--	1
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents :		
"A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed		"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family
Date of the actual completion of the international search  2 December 2019		Date of mailing of the international search report  12/12/2019
Name and mailing address of the ISA/ European Patent Office, P.B. 5618 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040 Fax: (+31-70) 340-3016		Authorized officer  Oderwald, Harald

## INTERNATIONAL SEARCH REPORT

International application No PCT/IL2019/051000
---

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>Marie-Claire Hérold ET AL:  "UDP-Glucuronosyltransferase activity is correlated to saponin production in Gypsophila paniculata root in vitro cultures",  Biotechnology Letters,  March 2001 (2001-03), pages 335-337,  XP055647834,  Dordrecht  DOI: 10.1023/A:1005689104107  Retrieved from the Internet:  URL:https://link.springer.com/content/pdf/10.1023/A:1005689104107.pdf  the whole document  -----</p>	1

**INTERNATIONAL SEARCH REPORT**

International application No.

PCT/IL2019/051000

**Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)**

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
- a.  forming part of the international application as filed:
- in the form of an Annex C/ST.25 text file.
- on paper or in the form of an image file.
- b.  furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
- c.  furnished subsequent to the international filing date for the purposes of international search only:
- in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).
- on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).
2.  In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No

PCT/IL2019/051000

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 2982752	A1	10-02-2016	EP 2982752 A1 10-02-2016
			JP 6344774 B2 20-06-2018
			JP WO2014163174 A1 16-02-2017
			KR 20150139597 A 11-12-2015
			US 2016046912 A1 18-02-2016
			US 2019119655 A1 25-04-2019
			WO 2014163174 A1 09-10-2014
-----			
US 2011219476	A1	08-09-2011	AU 2008361302 A1 11-03-2010
			BR PI0823018 A2 26-02-2019
			CA 2735965 A1 11-03-2010
			EC SP11010938 A 30-06-2011
			JP WO2010026666 A1 26-01-2012
			KR 20110056518 A 30-05-2011
			US 2011219476 A1 08-09-2011
			WO 2010026666 A1 11-03-2010
-----			
WO 2014195944	A1	11-12-2014	CN 105592692 A 18-05-2016
			EP 3003013 A1 13-04-2016
			US 2016122775 A1 05-05-2016
			WO 2014195944 A1 11-12-2014
-----			

## フロントページの続き

(51) Int.Cl.	F I		テーマコード(参考)
C 1 2 P 15/00 (2006.01)	C 1 2 P	15/00	
C 1 2 N 15/113 (2010.01)	C 1 2 N	15/113	1 3 0 Z
A 0 1 H 5/00 (2018.01)	A 0 1 H	5/00	A
C 1 2 N 9/99 (2006.01)	C 1 2 N	9/99	
C 1 2 N 9/10 (2006.01)	C 1 2 N	9/10	
C 1 2 N 15/52 (2006.01)	C 1 2 N	15/52	
C 1 2 N 15/29 (2006.01)	C 1 2 N	15/29	

(81) 指定国・地域 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, TJ, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT

(特許庁注：以下のものは登録商標)

1 . T R I T O N

(72) 発明者 アハロニ、アサフ

イスラエル国 6 9 0 2 5 1 0 ・テルアビブ・ラマ アビブ・カルニ イェフダ 1 0

(72) 発明者 ソナワネ、プラシャント

イスラエル国 7 6 1 0 0 0 2 ・レホボト・ピー． オー． ボックス 9 5 ・イエダ リサーチ  
アンド ディベロップメント カンパニー リミテッド アット ザ ワイズマン インスティテ  
ュート オブ サイエンス内

(72) 発明者 イトキン、マクシム

イスラエル国 7 6 1 0 0 0 2 ・レホボト・ピー． オー． ボックス 9 5 ・イエダ リサーチ  
アンド ディベロップメント カンパニー リミテッド アット ザ ワイズマン インスティテ  
ュート オブ サイエンス内

(72) 発明者 ヨズヴィアク、アダム

イスラエル国 7 6 1 0 0 0 2 ・レホボト・ピー． オー． ボックス 9 5 ・イエダ リサーチ  
アンド ディベロップメント カンパニー リミテッド アット ザ ワイズマン インスティテ  
ュート オブ サイエンス内

F ターム(参考) 2B030 AA02 AB03 AB04 AD08 CA14 CA19

4B050 CC03 DD13 LL01 LL02 LL10

4B064 AB05 AE01 AH08 BH07 CA11 CA19 CA21 CC15 CC24 CE10

DA01 DA10 DA12 DA20

4B065 AA72X AA83X AA88X AA88Y AB01 AC14 BA02 BD14 CA16 CA18

CA29 CA41 CA44 CA47