

(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2020-501519

(P2020-501519A)

(43) 公表日 令和2年1月23日(2020.1.23)

(51) Int.Cl.	F 1	テーマコード (参考)		
C 12 N 15/52 (2006.01)	C 12 N 15/52	Z N A Z	4 B 0 5 0	
C 12 N 9/00 (2006.01)	C 12 N 9/00		4 B 0 6 4	
C 12 N 15/62 (2006.01)	C 12 N 15/62	Z	4 B 0 6 5	
C 12 N 15/70 (2006.01)	C 12 N 15/70	Z		
C 12 N 15/76 (2006.01)	C 12 N 15/76	Z		

審査請求 未請求 予備審査請求 有 (全 91 頁) 最終頁に続く

(21) 出願番号	特願2019-523736 (P2019-523736)	(71) 出願人	516286741 ギンゴー バイオワークス, インコーポ レイテッド アメリカ合衆国 マサチューセッツ 02 210, ボストン, ドライドック ア ベニュー 27, フロア 8
(86) (22) 出願日	平成29年10月27日 (2017.10.27)	(74) 代理人	100105957 弁理士 恩田 誠
(85) 翻訳文提出日	令和1年6月26日 (2019.6.26)	(74) 代理人	100068755 弁理士 恩田 博宣
(86) 國際出願番号	PCT/US2017/058805	(74) 代理人	100142907 弁理士 本田 淳
(87) 國際公開番号	W02018/081592	(74) 代理人	100152489 弁理士 中村 美樹
(87) 國際公開日	平成30年5月3日 (2018.5.3)		
(31) 優先権主張番号	62/414, 435		
(32) 優先日	平成28年10月28日 (2016.10.28)		
(33) 優先権主張国・地域又は機関	米国(US)		

最終頁に続く

(54) 【発明の名称】 化合物の生産のための組成物および方法

(57) 【要約】

本開示は、目的の化合物の生産に有用なタンパク質、核酸、ベクター、および宿主分子、ならびにその使用方法を提供する。

【特許請求の範囲】

【請求項 1】

未修飾のドメインを備える参照ポリケチドと比較して酵素活性が改変された 1 つまたは複数の修飾されたドメインを備える、遺伝子操作されたポリケチドシンターゼであって、遺伝子操作されたポリケチドシンターゼによる化合物の発現を可能にするのに適した条件下で発現された場合にポリケチドを生産することができる、遺伝子操作されたポリケチドシンターゼ。

【請求項 2】

酵素活性が改変された 2 つ以上の修飾されたドメインを備える、請求項 1 に記載の遺伝子操作されたポリケチドシンターゼ。

10

【請求項 3】

少なくとも 1 つの修飾されたドメインが、低下した酵素活性を有する、請求項 1 または 2 に記載の遺伝子操作されたポリケチドシンターゼ。

【請求項 4】

少なくとも 1 つの修飾されたドメインが、機能的に不活性である、請求項 3 に記載の遺伝子操作されたポリケチドシンターゼ。

【請求項 5】

修飾されたドメインが、 - ケトンプロセシングドメインである、請求項 1 乃至 4 のいずれか 1 項に記載の遺伝子操作されたポリケチドシンターゼ。

【請求項 6】

- ケトンプロセシングドメインが、ケトレダクターゼ、デヒドラターゼ、またはエノイルレダクターゼである、請求項 5 に記載の遺伝子操作されたポリケチドシンターゼ。

20

【請求項 7】

- ケトンプロセシングドメインが、配列番号 1 ~ 9 のいずれか 1 つの保存された領域に対して少なくとも 90 % の配列同一性を有する部分を備える、請求項 6 に記載の遺伝子操作されたポリケチドシンターゼ。

【請求項 8】

- ケトンプロセシングドメインが、ケトレダクターゼであり、ケトレダクターゼが、(a) 保存された Y A A A N 触媒性モチーフ内のチロシンに対応する位置でチロシン以外のアミノ酸を備え、配列番号 1 内の保存された F G ヘリックスを備えていないか、(b) 配列番号 2 内の S 9 - p k s A O R F のアラニン 6 6 3 2 に対応する位置でグルタミン酸残基を備えているか(S 9 における変化)、または(c) 配列番号 3 の W T S 1 2 - p k s B O R F のアミノ酸 3 3 8 6 から 3 5 1 6 に対応するアミノ酸を備えていない、請求項 6 に記載の遺伝子操作されたポリケチドシンターゼ。

30

【請求項 9】

- ケトンプロセシングドメインが、デヒドラターゼであり、デヒドラターゼが、(a) 配列番号 4 の保存された H X X X G X X X X P モチーフ内の S 6 7 9 - p k s B O R F の p k s B における 4 2 8 8 位のグリシンに対応する位置でアスパラギン酸を備えるか、(b) 配列番号 5 内の S 1 2 - p k s B O R F における 3 0 6 6 位から 3 0 7 0 位に対応する位置の保存された L P F X W モチーフにおける置換を備えるか、(c) 配列番号 6 の S 6 7 9 - p k s A O R F の P r o 6 8 4 4 から T r p 6 8 7 4 の間の欠失を備えるか、または(d) 配列番号 7 の A、B、C、および D に対応する位置での置換もしくは欠失を備える、請求項 6 に記載の遺伝子操作されたポリケチドシンターゼ。

40

【請求項 10】

- ケトンプロセシングドメインが、エノイルレダクターゼであり、エノイルレダクターゼが、配列番号 8 内の S 1 2 - p k s B O R F の 1 5 4 6 位に対応する位置でリジンを備えていない、および / または配列番号 8 もしくは 9 内の S 1 2 - p k s B の 1 5 6 8 位に対応する位置でアスパラギン酸を備えていない、請求項 6 に記載の遺伝子操作されたポリケチドシンターゼ。

【請求項 11】

50

(a) 第1のポリケチドシンターゼのドメインの保存された領域を備える、第1のドメイン、および
 (b) 第2のポリケチドシンターゼのドメインの保存された領域を備える、第2のドメイン
 を備える、ポリケチドシンターゼ。

【請求項12】

第1のドメインおよび第2のドメインの少なくとも1つが、-ケトンプロセシングドメインである、請求項11に記載のポリケチドシンターゼ。

【請求項13】

第1のドメインおよび第2のドメインが、共に-ケトンプロセシングドメインである
 10、請求項11または12に記載のポリケチドシンターゼ。

【請求項14】

-ケトンプロセシングドメインが、ケトレダクターゼ、デヒドラターゼ、またはエノイルレダクターゼである、請求項12または13に記載のポリケチドシンターゼ。

【請求項15】

第1のドメインおよび第2のドメインの少なくとも1つが、機能的に不活性なドメイン
 である、請求項11乃至14のいずれか1項に記載のポリケチドシンターゼ。

【請求項16】

第1のドメインおよび第2のドメインの両方が、機能的に不活性なドメインである、請求
 20項15に記載のポリケチドシンターゼ。

【請求項17】

(c) 第3のポリケチドシンターゼのドメインの保存された領域、または第2のポリケ
 チドシンターゼの第2のドメインの保存された領域を備える、請求項11乃至16のいずれか
 1項に記載のポリケチドシンターゼ。

【請求項18】

第3のドメインが、機能的に不活性なドメインである、請求項17に記載のポリケチド
 シンターゼ。

【請求項19】

(d) 第4のポリケチドシンターゼのドメインの保存された領域、第3のポリケチドシ
 ンターゼの第2のドメインの保存された領域、または第2のポリケチドシンターゼの第3
 のドメインの保存された領域を備える、請求項11乃至18のいずれか1項に記載のポリ
 30ケチドシンターゼ。

【請求項20】

第4のドメインが、機能的に不活性である、請求項19に記載のポリケチドシンターゼ
 。

【請求項21】

機能的に不活性なドメインが、配列番号1～9のいずれか1つの保存された領域のアミ
 ノ酸配列を備える、請求項15乃至20のいずれか1項に記載のポリケチドシンターゼ。

【請求項22】

ポリケチドシンターゼの少なくとも1つのドメインが、配列番号10または11の配列
 を有するポリケチドシンターゼと比較して修飾されており、前記修飾が、酵素活性の改
 变をもたらす、キメラポリケチドシンターゼ。

【請求項23】

少なくとも1つのケトレダクターゼドメインが、(a)保存されたY A A A N触媒性モ
 チーフ内のチロシンに対応する位置でチロシン以外のアミノ酸を備え、配列番号1内の保
 存されたF Gヘリックスを備えていないか、(b)配列番号2におけるS 9-p k s A
 40O R Fのアラニン6632に対応する位置でグルタミン酸残基を備えるか、または(c)
)配列番号3のW T S 12-p k s B O R Fのアミノ酸3386から3516に対応
 するアミノ酸を備えていない、キメラポリケチドシンターゼ。

【請求項24】

10

20

30

40

50

少なくとも 1 つのデヒドラターゼドメインが、(a) 配列番号 4 の保存された H X X X G X X X X P モチーフ内の S 6 7 9 - p k s B O R F の p k s B における 4 2 8 8 位のグリシンに対応する位置でアスパラギン酸を備えるか、(b) 配列番号 5 内の S 1 2 - p k s B O R F における 3 0 6 6 位から 3 0 7 0 位に対応する位置の保存された L P F X W モチーフにおける置換を備えるか、(c) 配列番号 6 の S 6 7 9 - p k s A O R F の P r o 6 8 4 4 から T r p 6 8 7 4 の間の位置に対応する欠失を備えるか、または(d) 配列番号 7 の A 、 B 、 C 、および D に対応する位置での置換もしくは欠失を備える、キメラポリケチドシンターゼ。

【請求項 2 5】

少なくとも 1 つのエノイルレダクターゼドメインが、配列番号 8 内の S 1 2 - p k s B O R F の 1 5 4 6 位に対応する位置でリジンを備えていない、および / または配列番号 8 もしくは 9 内の S 1 2 - p k s B の 1 5 6 8 位に対応する位置でアスパラギン酸を備えていない、キメラポリケチドシンターゼ。

10

【請求項 2 6】

(a) 配列番号 7 、 8 、もしくは 9 、(b) 配列番号 1 0 、 1 1 、もしくは 1 2 、(c) 配列番号 1 3 、 1 4 、もしくは 1 5 、(d) 配列番号 1 6 、 1 7 、もしくは 1 8 、(e) 配列番号 1 9 、 2 0 、 2 1 、もしくは 2 2 、(f) 配列番号 2 3 、 2 4 、 2 5 、もしくは 2 6 、(g) 配列番号 2 7 、 2 8 、 2 9 、もしくは 3 0 、または(h) 配列番号 3 1 もしくは 3 2 のアミノ酸配列に対して少なくとも 8 0 % の配列同一性を有するドメインを備えるキメラポリケチドシンターゼ。

20

【請求項 2 7】

請求項 1 乃至 2 6 のいずれか 1 項に記載のポリケチドシンターゼをコードする核酸。

【請求項 2 8】

核酸が L A L をさらにコードし、 L A L をコードする配列が、ポリケチドシンターゼをコードする配列に機能可能に連結されている、請求項 2 7 に記載の核酸。

30

【請求項 2 9】

L A L が異種 L A L である、請求項 2 8 に記載の核酸。

【請求項 3 0】

L A L が、配列番号 3 8 に対して少なくとも 8 0 % の同一性を有する部分を備える、請求項 2 8 または 2 9 に記載の核酸。

30

【請求項 3 1】

L A L が、配列番号 3 8 の配列を有する部分を備える、請求項 3 0 に記載の核酸。

【請求項 3 2】

L A L が配列番号 3 8 の配列を有する、請求項 3 1 に記載の核酸。

【請求項 3 3】

L A L をコードする核酸が、オープンリーディングフレーム内で T T A 阻害コドンを欠いている、請求項 2 8 乃至 3 2 のいずれか 1 項に記載の核酸。

40

【請求項 3 4】

核酸が L A L 結合部位をさらに備え、 L A L 結合部位をコードする配列が、ポリケチドシンターゼをコードする配列に機能可能に連結されている、請求項 2 7 乃至 3 3 のいずれか 1 項に記載の核酸。

【請求項 3 5】

L A L 結合部位が、配列番号 3 9 の配列に対して少なくとも 8 0 % の配列同一性を有する部分を備える、請求項 3 4 に記載の核酸。

【請求項 3 6】

L A L 結合部位が、配列番号 3 9 の配列を有する部分を備える、請求項 3 5 に記載の核酸。

【請求項 3 7】

L A L 結合部位が配列番号 3 9 の配列を有する、請求項 3 6 に記載の核酸。

【請求項 3 8】

50

L A L 結合部位が配列 G G G G G T (配列番号 4 0) を有する、請求項 3 4 に記載の核酸。

【請求項 3 9】

L A L 結合部位への L A L の結合が、ポリケチドシンターゼの発現を促進する、請求項 3 4 乃至 3 8 のいずれか 1 項に記載の核酸。

【請求項 4 0】

非リボソームペプチドシンターゼをさらにコードする、請求項 2 7 乃至 3 9 のいずれか 1 項に記載の核酸。

【請求項 4 1】

第 1 の P 4 5 0 酵素をさらにコードする、請求項 2 7 乃至 4 0 のいずれか 1 項に記載の核酸。 10

【請求項 4 2】

第 2 の P 4 5 0 酵素をさらにコードする、請求項 4 1 に記載の核酸。

【請求項 4 3】

請求項 2 7 乃至 4 2 のいずれか 1 項に記載の核酸を備える発現ベクター。

【請求項 4 4】

人工染色体である、請求項 4 3 に記載の発現ベクター。

【請求項 4 5】

人工染色体が細菌人工染色体である、請求項 4 4 に記載の発現ベクター。

【請求項 4 6】

請求項 4 3 乃至 4 5 のいずれか 1 項に記載の発現ベクターを備える宿主細胞。 20

【請求項 4 7】

ポリケチドが宿主細胞に対して異種である、請求項 1 乃至 2 6 のいずれか 1 項に記載のポリケチドシンターゼを備える宿主細胞。

【請求項 4 8】

L A L を天然に欠く、請求項 4 6 または 4 7 に記載の宿主細胞。

【請求項 4 9】

L A L 結合部位を天然に欠く、請求項 4 6 乃至 4 8 のいずれか 1 項に記載の宿主細胞。

【請求項 5 0】

L A L 結合部位に結合してポリケチドシンターゼの発現を調節することができる L A L を備える、請求項 4 6 乃至 4 9 のいずれか 1 項に記載の宿主細胞。 30

【請求項 5 1】

L A L が異種である、請求項 5 0 に記載の宿主細胞。

【請求項 5 2】

L A L が、配列番号 3 8 の配列に対して少なくとも 8 0 % の同一性を有する部分を備える、請求項 5 0 または 5 1 に記載の宿主細胞。

【請求項 5 3】

宿主細胞が細菌である、請求項 4 6 乃至 5 2 のいずれか 1 項に記載の宿主細胞。

【請求項 5 4】

細菌がアクチノバクテリウムである、請求項 5 3 に記載の宿主細胞。 40

【請求項 5 5】

アクチノバクテリウムが、ストレプトマイセス・アムボファシエンス (S t r e p t o m y c e s a m b o f a c i e n s) 、ストレプトマイセス・ハイグロスコピカス (S t r e p t o m y c e s h y g r o s c o p i c u s) 、またはストレプトマイセス・マラエンシス (S t r e p t o m y c e s m a l a y e n s i s) である、請求項 5 4 に記載の宿主細胞。

【請求項 5 6】

アクチノバクテリウムが、S 1 3 9 1 、S 1 4 9 6 、または S 2 4 4 1 である、請求項 5 5 に記載の宿主細胞。

【請求項 5 7】

50

ポリケチドシンターゼの発現を増強させるように修飾されている、請求項 4 6 乃至 5 6 のいずれか 1 項に記載の宿主細胞。

【請求項 5 8】

(i) 化合物生産タンパク質を発現する内在性遺伝子クラスターの欠失、(i i) 化合物生産タンパク質を発現する異種遺伝子クラスターの挿入、(i i i) 抗生物質での刺激への宿主細胞の曝露、および / または (i v) 同種プロモーターと比較して化合物の発現を少なくとも 2 倍増大させる異種プロモーターの導入によって、化合物生産タンパク質の発現を増強させるように修飾されている、請求項 5 7 に記載の宿主細胞。

【請求項 5 9】

請求項 4 6 乃至 5 8 のいずれか 1 項に記載の宿主細胞を適切な条件下で培養する工程を備える、ポリケチドを生産する方法。 10

【請求項 6 0】

請求項 1 乃至 2 6 のいずれか 1 項に記載のポリケチドシンターゼを発現するように遺伝子操作された宿主細胞を、ポリケチドシンターゼがポリケチドを生産するのに適した条件下で培養する工程を備える、ポリケチドを生産する方法。

【請求項 6 1】

ポリケチドシンターゼの活性を調節する方法であって、

(a) 親ポリケチドシンターゼをコードする親核酸配列を提供する工程、および

(b) 親核酸配列の少なくとも 1 つのコドンを修飾する工程であって、前記コドンが、親ポリケチドシンターゼの少なくとも 1 つのドメインの保存されたモチーフ内の残基を特定し、 20

前記修飾が、少なくとも 1 つのドメインの酵素活性または調節活性の改変をもたらす、工程

を備える、方法。

【請求項 6 2】

化合物を生産する方法であって、

(a) 親ポリケチドシンターゼをコードする親核酸を提供する工程、

(b) 親核酸の少なくとも 1 つのコドンを修飾して、化合物を生産することができる修飾されたポリケチドシンターゼをコードする修飾された核酸を生成させる工程であって、前記コドンが、ポリケチドシンターゼの少なくとも 1 つのドメインの保存されたドメイン内の残基を特定し、前記修飾が、ポリケチドシンターゼの少なくとも 1 つのドメインの酵素活性の改変をもたらす、工程。 30

(c) 修飾された核酸を宿主細胞に導入する工程、および

(d) 修飾されたポリケチドシンターゼによる化合物の発現を可能にするのに適した条件下で宿主細胞を培養する工程

を備え、それによって化合物を生産する、方法。

【請求項 6 3】

化合物を生産する方法であって、

(a) 化合物を生産することができる親ポリケチドシンターゼを提供する工程、

(b) 親ポリケチドシンターゼのアミノ酸配列を決定する工程、

(c) 親ポリケチドシンターゼをコードする親核酸を提供する工程、

(d) 親核酸の少なくとも 1 つのコドンを修飾して、化合物を生産することができる修飾されたポリケチドシンターゼをコードする修飾された核酸配列を生成する工程であって、前記コドンが、ポリケチドシンターゼの少なくとも 1 つのドメインの保存されたドメイン内の残基を特定し、前記修飾が、少なくとも 1 つのドメインの酵素活性の改変をもたらす、工程。 40

(e) 修飾された核酸を宿主細胞に導入する工程、

(f) 修飾されたポリケチドシンターゼによる化合物の発現を可能にするのに適した条件下で宿主細胞を培養する工程、および

(g) 修飾されたポリケチドシンターゼによって生産される化合物を回収する工程

50

を備え、それによって化合物を生産する、方法。

【請求項 6 4】

化合物を生産する方法であって、
 (a) 親ポリケチドシンターゼの構造を決定する工程、
 (b) 親ポリケチドシンターゼをコードする親核酸を生産する工程、
 (c) 核酸を修飾して、修飾されたポリケチドシンターゼをコードする修飾された核酸を生産する工程であって、修飾されたポリケチドシンターゼの少なくとも 1 つのドメインが、親ポリケチドシンターゼと比較して酵素活性が改変される工程、
 (d) 修飾された核酸配列を宿主細胞に導入する工程、および
 (e) 修飾されたポリケチドシンターゼによる化合物の発現を可能にするのに適した条件下で宿主細胞を培養する工程

を備え、それによって化合物を生産する、方法。

【請求項 6 5】

化合物のライブラリーを生産する方法であって、
 (a) 親ポリケチドシンターゼをコードする親核酸配列を提供する工程、
 (b) 親核酸配列の少なくとも 1 つのコドンを修飾して、化合物を生産することができる第 1 の修飾されたポリケチドシンターゼをコードする第 1 の修飾された核酸を生成する工程、
 (c) 親核酸の少なくとも 1 つのコドンを修飾して、化合物を生産することができる第 2 の修飾されたポリケチドシンターゼをコードする第 2 の修飾された核酸を生成する工程であって、第 1 および第 2 の修飾された核酸が異なる、工程、
 (d) 第 1 および第 2 の修飾された核酸配列を 1 つまたは複数の宿主細胞に導入する工程、および
 (e) 第 1 および第 2 の修飾されたポリケチドシンターゼによる化合物の発現を可能にするのに適した条件下で 1 つまたは複数の宿主細胞を培養する工程

を備え、それによって化合物のライブラリーを生産する、方法。

【請求項 6 6】

請求項 5 9 乃至 6 5 のいずれか 1 項に記載の方法によって生産される化合物。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本発明は、化合物の生産のための組成物および方法に関する。

【背景技術】

【0 0 0 2】

ポリケチド天然生成物は、他の特注の酵素と組み合わせて、ポリケチドシンターゼ (P K S : polykettide synthases)、例えば I 型ポリケチドシンターゼによって生合成によって生産される。ポリケチドシンターゼ (P K S) は、その触媒機能がポリケチドを生産するためのモジュールに組織化されている、大きなマルチドメインタンパク質のファミリーである。ポリケチドシンターゼクラスターの基本的な機能ユニットは、例えばマロニル - C o A に由来する 2 炭素からなる伸長ユニットをコードするモジュールである。ポリケチドシンターゼ内に一般に存在するモジュールには、 i) ローディングモジュール、 i i) 伸長モジュール、および i i i) 放出モジュールが含まれる。モジュール内で、ポリケチド鎖の伸長および延長に必要な最小ドメインアーキテクチャーには、ケトシンターゼ (K S : ketosynthase) ドメイン、アシルトランスフェラーゼ (A T : acyl-transferase) ドメイン、および A C P (acyl-carrier protein) (アシル担体タンパク質) ドメインが含まれ、各モジュールの特異的な化学は、 A T ドメインによって、ならびに、 - ケトンプロセシングドメイン、すなわち、ケトレダクターゼ (K R : ketoreductase) ドメイン、デヒドロターゼ (D H : dehydratase) ドメイン、およびエノイルレダクターゼ (E R : enoylreductase) ドメインの存在によってコードされる。ポリ

10

20

30

40

50

ケチドシンターゼの生合成は、2つの鍵となるメカニズム、すなわち、ポリケチドシンターゼ伸長モジュールでのポリケチド鎖の延長、およびモジュール間でのポリケチド中間体の転位によって進む。生産的な鎖の延長は、モジュール内およびモジュール間の両方の、多くの触媒性ドメインの協調した機能に依存する。

【0003】

コンビナトリアル生合成は、ポリケチドシンターゼ（PKS）遺伝子クラスターを遺伝子操作して新規な薬剤候補を生産するために採用されている、一般的な戦略である（非特許文献1）。これまで、これらの戦略は、モジュール内でのPKSドメインの欠失および／もしくはドメインスワッピングの遺伝子操作に依拠するものであるか、または、モジュール全体を別のクラスターからスワッピングしてキメラクラスターを生産することによるものであった。

【先行技術文献】

【非特許文献】

【0004】

【非特許文献1】ワイスマンおよびリードレイ（Weissman and Leadlay）、Nature Reviews Microbiology、2005

【発明の概要】

【発明が解決しようとする課題】

【0005】

このアプローチの課題は、ホールセールドメインの置き換え、挿入、または欠失を介するポリケチドメガシンターゼのタンパク質遺伝子操作が、PKSの「アセンブリライン」アーキテクチャーを混乱させ得、こうして、合成されるポリケチドの量が大幅に低減することである。

【課題を解決するための手段】

【0006】

本開示は、進化によってドメインの活性を「オン」または「オフ」にするメカニズムを模倣し、加速させることによって、化合物生産を著しく損なうことなくポリケチドのコンビナトリアル生合成を容易にするために有用な組成物および方法を提供する（図1）。

【0007】

さらに具体的には、本開示は、酵素活性を制御する-ケトンプロセシングドメイン、すなわち、ケトレダクターゼ（KR）ドメイン、デヒドラターゼ（DH）ドメイン、およびエノイルレダクターゼ（ER）ドメインにおける短鎖タンパク質配列を利用することによる、ドメインレベルでのPKS遺伝子操作のための組成物および方法を提供する（図2Aおよび2B）。推定上の死滅したドメイン配列を生きたドメインにグラフトしてドメイン活性を不活性化し、クラスターによってコードされるポリケチドの化学構造を改変する。ストレプトマイセス（Streptomyces）発現宿主における修飾されたクラスターの異種発現を次いで使用して、新規な化合物を生産する。このアプローチは、1つのクラスターにおいて複数のドメインレベルの遺伝子操作を行って、遺伝子操作された分子のコンビナトリアルライブラーを生成させることに、さらに利用することができる。

【0008】

したがって、一態様では、本開示は、未修飾のドメインを含む参照ポリケチドシンターゼと比較して酵素活性が改変された1つまたは複数の修飾されたドメインを含む、遺伝子操作されたポリケチドシンターゼであって、遺伝子操作されたポリケチドシンターゼによる化合物の発現を可能にするのに適した条件下で発現された場合にポリケチドを生産することができる、遺伝子操作されたポリケチドシンターゼを提供する。

【0009】

一部の実施形態では、遺伝子操作されたポリケチドシンターゼは、酵素活性が改変された2つ以上の修飾されたドメインを含む。

一部の実施形態では、少なくとも1つの修飾されたドメインは、低下した酵素活性を有する（例えば、少なくとも1つの修飾されたドメインが機能的に不活性である）。

10

20

30

40

50

【0010】

一部の実施形態では、修飾されたドメインは、-ケトンプロセシングドメイン（例えば、ケトレダクターゼ、デヒドラターゼ、またはエノイルレダクターゼ）である。

別の様では、本開示は、

(a) 第1のポリケチドシンターゼのドメインの保存された領域を含む、第1のドメイン、および

(b) 第2のポリケチドシンターゼのドメインの保存された領域を含む、第2のドメインを含む、ポリケチドシンターゼを提供する。

【0011】

一部の実施形態では、第1のドメインおよび第2のドメインの少なくとも1つは、-ケトンプロセシングドメイン（例えば、ケトレダクターゼ、デヒドラターゼ、またはエノイルレダクターゼ）である。一部の実施形態では、第1のドメインおよび第2のドメインは、共に-ケトンプロセシングドメインである。

10

【0012】

一部の実施形態では、第1のドメインおよび第2のドメインの少なくとも1つは、機能的に不活性なドメインである。一部の実施形態では、第1のドメインおよび第2のドメインの両方は、機能的に不活性なドメインである。

【0013】

一部の実施形態では、ポリケチドシンターゼは、(c)第3のポリケチドシンターゼのドメインの保存された領域（例えば、機能的に不活性なドメイン）、または第2のポリケチドシンターゼの第2のドメインの保存された領域を含む。

20

【0014】

一部の実施形態では、ポリケチドシンターゼは、(d)第4のポリケチドシンターゼのドメインの保存された領域（例えば、機能的に不活性なドメイン）、第3のポリケチドシンターゼの第2のドメインの保存された領域、または第2のポリケチドシンターゼの第3のドメインの保存された領域を含む。

【0015】

一部の実施形態では、機能的に不活性なドメインは、配列番号1～9のいずれか1つの保存された領域のアミノ酸配列を含む。

30

前述のポリケチドシンターゼのいずれかの一部の実施形態では、-ケトンプロセシングドメインは、配列番号1～9のいずれか1つの保存された領域に対して少なくとも90%の配列同一性を有する部分を含む。

【0016】

配列番号1

D P D G T V L I T G G S G V R A G A L A R H L V T E R G V R H L L L S R T T A
D E E L L N E L G E L G A R V D T A I C D V S D R A R L A Q V L A G V S P E H P
L T A V I H T A G A L D D D V V E S L T A Q R L D T V L R P K A D G A W H L H E
L T R D T D L A A F V M Y S S A A G V M G N P G Q G N F A A A T A F L D A L A E
Q R R A E G L P A L A L A W G S S E E T G G L T G L R A I S A E H G M R L F D S
A S H R R E P L L V A A S M D P V L A A E V P A L L R S L R R P I A R R A A S A
D G V Q W L A G L A P E E R A K A L L K V V C D T A A T V L G H A D A R T I P L
T G A F K D L G V D S L T A V E L R N S L T K A T G L R L P A T L V F D Y P T P
T A L A V R L

40

配列番号2

D P D G T I L I T G G S G V L A G I L A R H L A A E H G A R H L L L S R T A P
D E A L I K E L A E L G A R V E T A A C D V S D R A G L A R V L A G V S P E H P
L T A V I H T A G A L D D G V V E S L T T Q Q L D T V L R P K A D G A W H L H E
L T R D A D L A A F V V Y S S A A A V L G N E G Q G N Y A A A N A F L D A L A E
Q R R T Q G L P A L A L A W G P W E Y T G D L T A Q L T G T D Q D R I R C S G M
R T I T A E D G M R L F D T A S H H G E P L L V P A V L D P T R D G E V P A L L

50

R S L R R P I A R R A A S A D G G V Q W L A A L A P A E R E K A L L K L V C D S
A A M V L G H A D A R S I P A A G A F K D L G V D S L M A V E L R N G L V K A T
G L R L P A T L V F D Y P T P T V L A A R L

配列番号 3

D P D G T V L I T G E R A G A V A R R M A E R G V R H L L L A S G R V P D E L M
D L D T S V E V A V C D V S D R A A L A G V L A G L P S L T G V I Q T A G E D V
L P V L A G A I T P T R D G E I P A S L R L L R R P L V R R R V S A A G D S S L
A A L P P A E R E R A L L K V V R D S A A V V L G H A D G R T V P A T A A F K D
L G L D S L T A V E L R N S L R K A T G L Q L P A T L V F D Y P S P V A L A A R
L G

10

配列番号 4

H P F L G A A L P A P D G D S L T L T G R I T L D A H P W L A D H I I R D T L I
L P G A A F A E C V L R A G R E V G C D L L E E L V I E A P L V L P A T G G V A
V R I A V G E P D D A G R R T F D L Y A R P D A A P G W N R H A G G T L K P G D
A L P A T E A A T E T V A W P P A D A E P V D V D D L Y D R L A A A G Y A Y G P
A F Q S V H A A W R T P D A I W A E V V L D G E P A G F G L H P A L L D G A L Q
L S A L A A T G G D V A Q L P F A W H D V R L P G H G A D R L R V R L

配列番号 5

H P L L G A I V A V P Q S G G V A M T S R L S P R N H P W L A E H T L G G V P T
V P T S V L V E L A V R A G D E V G C G V V E E L T V D A P L L L P E R G G V R
V Q V I V G A T D A N G Q R G L D I F S A P E D T G Q E A W T R H A T G T L A P
G G D I A A D V D L S A W P P A N A Q P V D V T D G Y D L L E R A G Y G Y G P A
F Q G V R A I W R R G E E L F A E V A L E P E L T D T A A R F G L H P A L L D A
A W H P E L R D E V A E T S P D G R R W W S Q P S R W A G L R L H T A G A T V L
R V R L A P V D A D S M S L Q A A D E T G D P V L T V D S L S

20

配列番号 6

H P L L G A G M P I A G T G A V L F G T E V A H P W F D G H E T L P A A A F A E
I A V R A A A E V G S P V V G E L H V E L L P R I P A D G R I R L Q T W V D G P
D P T G V R R F T V H A R P D P T A A W L R V A S G V L T G A E A P V P A F A G
G E P L H I A D G T P A G F L L H P D A T P A A D W F G L V A H G S G A R Q Q H
V Y Q A G E G L C V T D D A G R P I V T A A R V R

30

配列番号 7

H P L L G W G V P V A E A G G R L Y T G R V A R Q D G P V L S V A A F V E M A F
A A A G G R P I R E L S V D A L L Y I P D D G T A E L Q T W V S E H R L T I H A
R Y R D T E P W T R L A T A A L D T T A P T T H T P H P G L I T T A L T L T G
D E A P A I W H D L T L H T S N A T E L H T H I T P G D D G T L T I T A T D T T
G Q P V L T A H T A T

配列番号 8

R L S A L A S L G E P Q I V V R D D T P L V A R L A R E K S P A L T I P G E R A
W V L E P D H S G V L Q E L A L V A A D T D V R P L R P G E V R I E V R A A G L
N F R D V L V A L G T D L G D G V F G A E G A G V V L E T G S D V R D L R P G D
R V F G L L E G G H G S I A I A D R R M L A V I P E G W S F A T A A S V P E V F
V I A Y Y G L V D L A G L R A G E S V L I H A A T G G V G M A A T Q I A R H L G
A Q V Y A T A G V G K Q H I L R D A G L G D D R I A D S R T T D F R E A F R D S
T Q G R G V D V V L N S L K G D F V D A S L D L L A D G G R F L E L G Q T D I R
D A G E I A A E R P G T T Y H S F T R M N A G P D R L R E I I A E L L A L F E Q
G V L R P S P V H T W D I R H A R E A F S W M S G G R H T G K M V L T M P Q R I
D P G G T V L I A G D S E A L A R I A A R H L G V R H L L L D R G V A D A A P D
A V V C D V S D H D A L E R V L A D L S P E H P L T A V I H T G G A A V T D E I
R R L H D L T E S L D L T D F V V F S Q D A P A A V E A F A R S R R A H G L P V

50

R T I A W G I P E A D P V V A D E H L L G R A L A S A E Q A Q I V A R V N T A G
L R A L T A A N A L P T L L R N L I R A E P E E T G Q S A W P H R F E A A G A D
R E E A L L D L I R A N V V D I L S L P T A D R Y A P D R T F R E M G I D S L T
A V G L R N S L A K A T G L P L P T T M V F D Y P T P A V L T A R M R E L

配列番号 9

R L S T L V A L G E P Q I A L R D S T P L V P R L A P E S S T A L T T P A A R A
W V L E P A R S G T L R E L S L V A A D T D A R P L R P G E V R V D V R A A G L
N F R D V L I A L G T Y P G D G V M G G E A A G V V L E V G P E V N D L S V G D
R V F G L V T D G F G P V T I T D R R L L A A M P Q D W S F T T A A S A A M A F
A T A H Y G L V E L A G L K A G E S V L I H A A T G G V G M A A T Q I A H H L G 10
A H I Y A T A S S G K Q H L L R A A G I D D D R I A N S R T T G F R D A F L D S
T G G R G V D V V L N S L S G E F V D S S L D L L A H G G R F I E M S T D I R D
A G R I A A E R P G T T Y Q A F H L V D A D P D R L R E I L T E L L A L F D Q G
I L D P L P V Q A W D I R Q A R E A F S W M S R A R H T G K L V L T I P Q H I D
P D G T V L I T G G S G G L A G V V A R H L V A D K G A R R L L L S C D T L D
A T L A A E L T E S G A R V D T A V C D V S D R A A L A Q V L A G V S P E H P L
T A I V H A G G A A V A D E S R Q L H H L T K N R D L A A F V V F S Q D A P A A
T E A F A G I R Q A E G L P V T T I A W G I P E A E P V V V G Q H L L D R A M A
S A D R A H V A A R V N T A G L R A L A A A N A L P P V L K N L V G A E T D G T
G H Q D W S R R F M V A E A A R Q Q E L L D L I R T T V M E I L S L P T T A R Y 20
F P D R T F R E N G I D S L T A V E L V N S L A K T T G L R L S A T M V F D Y P
T P T A L A G R M R E L

前述のポリケチドシンターゼのいずれかの一部の実施形態では、 - ケトンプロセシングドメインは、 (a) 保存された Y A A A A N 触媒性モチーフ内のチロシンに対応する位置でチロシン以外のアミノ酸を含み、かつ、配列番号 1 内の保存された F G ヘリックスを含まない、 (b) 配列番号 2 内の S 9 - p k s A ORF のアラニン 6 6 3 2 に対応する位置でグルタミン酸残基を含む (S 9 における変化) 、または (c) 配列番号 3 の W T S 1 2 - p k s B ORF のアミノ酸 3 3 8 6 から 3 5 1 6 に対応するアミノ酸を含まない、ケトレダクターゼである。

【 0 0 1 7 】

前述のポリケチドシンターゼのいずれかの一部の実施形態では、 - ケトンプロセシングドメインは、 (a) 配列番号 4 の保存された H X X X G X X X X P モチーフ内の S 6 7 9 - p k s B ORF の p k s B における 4 2 8 8 位のグリシンに対応する位置でアスパラギン酸を含む、 (b) 配列番号 5 内の S 1 2 - p k s B ORF における 3 0 6 6 位から 3 0 7 0 位に対応する位置の保存された L P F X W モチーフにおける置換を含む、 (c) 配列番号 6 の S 6 7 9 - p k s A ORF の P r o 6 8 4 4 から T r p 6 8 7 4 の間の欠失を含む、または (d) 配列番号 7 の A 、 B 、 C 、および D に対応する位置での置換もしくは欠失を含む、デヒドラターゼである。

【 0 0 1 8 】

前述のポリケチドシンターゼのいずれかの一部の実施形態では、 - ケトンプロセシングドメインは、配列番号 8 内の S 1 2 - p k s B ORF の 1 5 4 6 位に対応する位置でリジンを含まない、および / または配列番号 8 もしくは 9 内の S 1 2 - p k s B の 1 5 6 8 位に対応する位置でアスパラギン酸を含まない、エノイルレダクターゼである。

【 0 0 1 9 】

別の態様では、本開示は、ポリケチドシンターゼの少なくとも 1 つのドメインが、配列番号 1 0 または 1 1 の配列を有するポリケチドシンターゼと比較して修飾されており、修飾が、酵素活性の改変をもたらす、キメラポリケチドシンターゼを提供する。

【 0 0 2 0 】

配列番号 1 0

M S R E E F I Q P I H D L L R V N A E R L G D K I A Y A D S R R E L T H A E L R 50

T R T G R I A G H L V D L A V E R G D R V A I L L G N R V E T I
 E S Y L A I A R A G A I A V P L N P D A T G A E V A H F L A D S G A V L V I T D
 S A H L D D V R R A A A A V T V V L V D E G P L P A G T R S F A E L A T A E P P
 T P A R D D L G L D E A A W M L Y T S G T T G T P K G V V S T Q G S G L W S A A
 N C D V P A W E L T E N D V L L W P A P L F H S L A H H L C L L A T T A V G A T
 A R I M S G F V A G E V L H E L E E H A C T V L V G V P T M Y H Y L L G A V G E
 A G P R L P S L K M G L V A G A V S P P A L I E G F E R V F G V P L L D T Y G C
 T E T T G S L T V N R L S G P R M P G S C G Q A V P G I S L R F V D P H T G A E
 V A E G E E G E L W A S G P S L M I G Y H G R P D A T R E V L S D G W Y R T G D
 L A R R S E T G H V T I T G R V K E L I I R G G E N I H P R D I E A V A L E L P 10
 G V R D A A A A G K Q H P V L G E I P A L Y L V P D A D G V D A E A V L A A C R
 E K L S Y F K V P E E I Y R V D A I P R T L S G K V K R A A L T E A P A E L L S
 A A S G N G S L Y R L E W V P A E T P P A G T G G P V A V H V T R R A V A T G P
 A D L P D Q E Q A A T W D A L R G E Q T G P G G P V L I D L D G A D I D D A R L
 S A L A S L G E P Q I V V R D D T P L V A R L A R E K S P A L T I P G E R A W V
 L E P D H S G V L Q E L A L V A A D T D V R P L R P G E V R I E V R A A G L N F
 R D V L V A L G T D L G D G V F G A E G A G V V L E T G S D V R D L R P G D R V
 F G L L E G G H G S I A I A D R R M L A V I P E G W S F A T A A S V P E V F V I
 A Y Y G L V D L A G L R A G E S V L I H A A T G G V G M A A T Q I A R H L G A Q
 V Y A T A G V G K Q H I L R D A G L G D D R I A D S R T T D F R E A F R D S T Q
 G R G V D V V V L N S L K G D F V D A S L D L L A D G G R F L E L G Q T D I R D A
 G E I A A E R P G T T Y H S F T R M N A G P D R L R E I I A E L L A L F E Q G V
 L R P S P V H T W D I R H A R E A F S W M S G G R H T G K M V L T M P Q R I D P
 G G T V L I A G D S E A L A R I A A R H L G V R H L L L D R G V A D A A P D A V
 V C D V S D H D A L E R V L A D L S P E H P L T A V I H T G G A A V T D E I R R
 L H D L T E S L D L T D F V V F S Q D A P A A V E A F A R S R R A H G L P V R T
 I A W G I P E A D P V V A D E H L L G R A L A S A E Q A Q I V A R V N T A G L R
 A L T A A N A L P T L L R N L I R A E P E E T G Q S A W P H R F E A A G A D R E
 E A L L D L I R A N V V D I L S L P T A D R Y A P D R T F R E M G I D S L T A V
 G L R N S L A K A T G L P L P T T M V F D Y P T P A V L T A R M R E L L A G E S
 P A P A R T A A R A V A Q D E P L A I V G M A C R L P G G V S S P D D L W R L V
 A A G T D A I S E F P A D R G W D V D N L Y D P D P D A P G K T Y T V L G G F L
 D G V A G F D A S F F G I S P R E A L A M D P Q Q R L M L E V S W E A F E H A G
 I P P R S V R G S D A G V F M G A F P S G Y D A G L E E F G M T G D A V S V L S
 G R V S Y F F G L E G P A I T V D T A C S S S L V A L H Q A S S A L R Q G E C S
 L A L V G G V T V L A T P Q T F V E F S R Q R G L A L D G R S K A F A D A A D G
 A G W A E G V G V L V V E R L S D A R A K G H Q I W G V I R G S A V N Q D G A S
 N G L S A P N G P S Q Q R V I R Q A L A N A G L A P H E V D V V E A H G T G T T
 L G D P I E A Q A V I A T Y G Q D R E Q P L L L G S L K S N V G H T Q A A A G V
 S G V I K M V M A L Q H D T V P A T L H V D A P S R H V D W T A G A V E L V T E
 N R P W P E T G R V R R A G V S S F G I S G T N A H V I L E S A P E Q P V S P P
 E A V A P V V A S D R V P L V I S A K T P A A L A E M E N R L R A Y L A A A P G
 A D P R A V A S T L A T A R S V F E H R A V L L G E N T I T G T V A G A D P R V
 V F V F P G Q G W Q Q L G M G R A L R E S S P V F A A R M A E C A A A L S E F V
 D W D L F T M L D D P A V I D R I D V L Q P A C W A V M M S L A A V W Q A A G V
 R P D A V I G H S Q G E I A A A C V A G A L S L R D A A R I V A L R S Q L L A R
 E M V G H G V M A A V A L P A D D I P L V D G V W I G A C N G P S S T V I S G T
 P E A V E V V V A A C E E R G A R V R R I T A A V A S H S P L G E K I R T E L L
 G I S A S I P S R T P V V P W L S T A D G I W I E A P L D P A Y W W R N L R E P
 V G F G P A V D L L Q A R G E N V F L E M S A S P V L L P A M N D A V T V A T L 50

R R D D G T P D R M L T A L A E A H A H G V I V D W P R V F G S T T R V L D L P
 T Y A F E H Q R Y W A V S A D R P S D A G H P M V E T V V P L P A S G G V A L T
 G R V S L A T H A W L A D H A V R G T A L L P G T A F V E L V T R A A T E V D C
 P V I D E L V I E A P L P L T Q T G A V Q L S T T V G E A D E S G R R P V T V F
 S Q A D G T D A W T R H V T A T I G R A A S L P D P V A W P P A Q A E P V D V T
 G F Y D E L A A A G Y E Y G P A F Q G L R A A W S D G D T V Y A E V V L A E E Q
 A H E V D R Y A V H P A L L D A A L Q A G M V N T A G T G Q G V R L P F S W N G
 I Q V H S T G A T T L R V A A T P L A D G W S V R A A A D N G R P V A T I G S L
 V T R P V T T D M L G S T T D D L F A V V W T E I T A P E P G D P S D V G V F T
 A L P E A G G D P L T Q T R A L T A Q V L Q T V Q Q W L A G E D R P L V V R T G 10
 T D L A S A A V S G L V R S A Q S E H P G R L I L V E S D D E L T P E Q L A G T
 A G L D E P R I R I D G G H Y E V P R L A R E D A S L T V P E D R A W L L E L P
 G S G T L R D L R V I P T D T A E R P L R W G E V R V G V R A G G L N F R D V V
 V A L G M V T D P R P A G G E A A G V V L E T G P G V E D L S P G D R V F G I L
 D G G F G S V A I A D R R L L A V I P D G W S F T T A A S I P V V F A T A Y Y G
 L V D L A G L R A G E S V L I H A A T G G V G M A A T Q I A R H L G A E I Y G T
 A G I A K Q H V L R D A G L G D D R I A D S R T T G F R E T F R D S T Q G R G V
 D V V L N S L S G D F V D A S L D V L A E G G R F I E M G K T D I R D A E Q I T
 H A T Y R A F D L M D A G P D R V R E I I A E L L G L F E Q G V L R P L P V Q A
 W D I R Q A R D A F T W M S R A R H I G K I V L T I P Q Q L D P D G T V L I S G
 G S G V L A G I L A R H L V A E R G V R H L L L V S R S A P S E A L I S E L T A
 L G A Q V E T V A C D V S D R V A L E Q V L D G V P L T A V F H T A A A L D D G
 V V E S L T P Q R V D T V L R P K A D A A W Y L H E L T R D A D L A A F V M Y S
 S V A G I M G A A G Q G N Y A A A N A F L D A L A A H R R R E G L P A L S L A W
 G L W E D A S G L S A G L T E T D H D R I R R G G L E A I A A E H G M R L F D T
 A T R Q G E P V L L A S P L N L T R Q G E V P A L L R T L H R P V A R R A A T A
 N G R P A D L T P E A L L K L V C G R A A A V L G H V D A D A V P V A V A F R D
 L G V D S L T A V E L R N S L A K A T G L R L P A T L V F D Y P T P T V L A G R
 L G E L L A G G T A P V R A A V V R R A A A S D E P L A I V G M A C R L P G G V
 L S P E D L W R L V E S G G D A I S G F P V D R G W D V E N L F D P D P D A A G
 R T Y A V R G G F L D G A A G F D A S F F G I S P R E A Q A M D P Q Q R L V L E
 V S W E A F E R A G I E P G S V R G S D T G V F M G A Y P G G Y G V G T D L G G
 F G M T S V A V S V L A G R V S Y F F G L E G P A M T V D T A C S S S L V A L H
 Q A G S A L R Q G E C S L A L V G G V T V M P T P Q T F V E F S R Q R G L A A D
 G R C K A F A D A A D G T G F S E G V G V L L V E R L S D A Q A R G H N I L A V
 V R G S A V N Q D G A S N G L T A P N G P S Q Q R V I R Q A L A N A G L A G A E
 V D V V E A H G T G T T L G D P I E A Q A V I A T Y G Q D R D Q P V L L G S L K
 S N L G H T Q A A A G V S G V I K M V M A L R H D T V P A T L H I D E P S R H I
 D W T A G A V E L V T E N Q S W P E T G R A R R A A V S S F G I S G T N A H V I
 L E S A P A Q P V P L V D T P V S A V T A G V V P L P I S A R T V P A L A D L E
 D R L R A Y L T T T P E T D L P A V A S T L A V T R S V F E H R A V L L G E E T
 V T G I A V S D P R V V F V F S G Q G S Q R V G M G E E L A A A F P L F A R L H
 R Q V W D L L D V P D L E V D D T G Y V Q P A L F A L Q V A L F G L L E S W G V
 R P E A V I G H S V G E V A A G Y V A G V W S L E D A C T L V S A R A R L M Q A
 L P A G G A M V A V P V S E E R A R A V L V D G V E I A A V N G P A S V V L S G
 D E S A V L R V A E G L G R W T R L S A S H A F H S V R M E P M L E E F R Q V A
 S E L T Y R E P R I V M A A G E Q V T T P E Y W V R Q V R D T V R F G D Q V A A
 F G D A V F L E I G P D R T L S R L I D G I P T L H G D D E Q H A V V A A L A E
 L H V Q G V P I D W S S I L G A N P A R V L D L P T Y A F Q H E R Y W M V S T G
 R V G G E G H P L L G W G V P V A E A G G R L Y T G R V A R Q D G P V L S V A A 50

F V E M A F A A A G G R P I R E L S V D A L L Y I P D D G T A E L Q T W V S E H
 R L T I H A R Y R D T E P W T R L A T A A L D T T A P A T T H T P H P G L I T T
 A L T L T G D E A P A I W H D L T L H T S N A T E L H T H I T P G D D G T L T I
 T A T D T T G Q P V L T A H T A T P T T I P V H T P T T P A D D L L T L T W T Q
 I P T P G P G D P T D I A V C T A L P D P D G D P L A Q T R T L T A Q V L Q S I
 Q T T L T G E D R P L V V H T G T G L A S A A V S G L V R S A Q S E H P D R F I
 L V E S D D S L P Q A Q L A A V A G L D E P W L R I T G S C Y E V P R L T K T T
 T A T A T A V S E P V W N P D G T V L I T G G S G A L A G I L A R H L V T E R G
 V R H L L L I S R S T P S T T L T D E L R E L G A H V D V A A C D V S D R D A L
 A R V L D G V D L T A V F H T A G A L D D G V V E S L T P Q R L D T V L T P K A 10
 D G A W H L H E L T R D R D L T A F V M Y S S A A G V M G A A G Q G N Y A A A N
 A F L D A L A E H R H A D G L P A L S L A W G M W D D T D G M T A S L S G T D H
 R R I R R S G Q R A I T A E H G M R L L D K A S G R S E P V L V A T A M N P I P
 D T D L P A L L R S L Y P K T A R K S Q P I Q E L S P E A L L K I V R D S A A L
 M L G H P N T D A I A A T T A F R D L G V D S L I A V E L R N S L A K A T G L R
 L P A T L V F D Y P T P T V L A G R L G E L L A G V T P Q R H A T V R T G T A S
 D E P L A I V G M A C R L P G G V S S P E D L W R L V E S G T D A I T D F P T D
 R G W D T D D L F D P D P D T A G K T Y T V H G G F L D D V A G F D A S F F G I
 S P R E A Q A M D P Q Q R L V L E A A W E A F E R A G I E P G S V R G S D T G V
 F M G A Y P G G Y G I G A D L G G F G A T A G A G S V L S G R L S Y F F G L E G
 P A M T V D T A C S S S L V A L H Q A G S A L R Q G E C S L A L V G G V T V I A
 N P Q I F V E F S R Q R G L A A D G R C K A F A D S A D G T G W S E G V G V L L
 V E R L S D A Q A R G H N I L A V V R G S A V N Q D G A S N G L T A P N G P S Q
 Q R V I R Q A L A N A G L A G A E V D V V E A H G T G T T L G D P I E A Q A V I
 A T Y G Q D R D Q S V L L G S L K S N L G H T Q A A A G V S G V I K M V M A L Q
 N G V V P R T L H A D Q P S R H I D W T A G A V E L V T E N Q P W P E L D R P R
 R A A V S A F G V S G T N A H V I L E S A P D Q P V P L V D T P V S A V T A G V
 V P L P I S A R T V P A L A D L E D Q L R A Y L T T A P E T D L P A V A S T L A
 T T R S V F E H R A V L L G E D T V T G T A I P D P R I V F V F S G Q G S Q R A
 G 30

M G E E L A A A F P L F A R L H R Q V W D L L D V P D L D V D D T G Y V Q P A L
 F A L Q V A L F G L L E S W G V R P R A V I G H S V G E V A A G Y V A G V W S L
 E D A C A L V S A R A R L M Q A L P A G G A M V A V P V S E E R A R A V L V D G
 V E I A A V N G P A S V V L S G D E A A V L R V A E G L G R W T R L S A S H A F
 H S V R M E P M L E E F R Q V V S R L T Y R E P R I V M A A G E Q V T T P E Y W
 V R Q V R E T V R F G D Q V A A F G D A V F L E I G P D R T L S R L I D G I A M
 L D G D D E V R A A V A A L A M M H V Q G V G V D W P A I L G T T T G R V L D L
 P T Y A F Q H E R Y W M A N A D E G H P L L G K V E H P L L G S V M A L P N S D
 G V V L T G R I S L A T H A W L A D H V V R G T V L L P G T G F V E M V A R A A
 A E V G C G V I D E L L I E A P L L L P E H G G V H L S V S V G E A D G A G R R
 P V T V F A Q A D D A E V W V R Q V T A T I S P A G P A V S L P E L E V W P P V
 Q A E P V D V S T F Y E R L A R A D W Q W G P A F Q G L R A A W R D G D T I Y A
 E I V L A D E E A R E A D Q F L V H P A L L D A A L Q T S V L K T P D D L R L P
 F S W N Q I E F H A T G A A I L R V A V T P V A D R W I V H A A D S T G R P V A
 T I G A L V S R P V T A E T L G S N T D D L F A L T W T E I P T P G P G D P A D
 V A V C T A L P E P D S D P L T Q T R T L T A Q V L Q S I Q T S L T G E D R P L
 V V H T G T G L A S A A V S G L V R S A Q S E H P D R F I L V E C D D E T L T P
 D Q L A A T A G L D E P W L R I T G G H Y E V P R L T K T T A A A T T V S E P
 V W D P D G T V L I T G G S G A L A G I L A R H L V T E R S V R H L L L I S R S 50

TPSTTLINE LRELGAHIETAACDVSDRDALARVLDGVDLT
 AVFHTAGALDDGVVESLTPQRLDTVLMPKADAAWHLHELT
 RDRDLAAFVMYSSAAGVMGAAGQGNYAAANANAFLDALAEHR
 RADGLPALS LAWG MWDDADGMTASLSGT DHRRI RSGQRA
 ITAEHGMRLLDKASGRSEPVLVATAMNPAGEGEVPA LLRT
 LHRPVARRAATTNGRPADLTPEALLKVV RDSA AVV LGHAS
 ADTVPAATAFQELGLDSLIAVELRNSLAKATGLR LPATMV
 FDYPTPAALAGRLGELLAGE TTPATAAVV R RATA SDEPLA
 IVGMACRLPGGVSSPEDLWRLVESGFD AITGFPTDRGWDV
 DNLYDPDPDAPGKSTTLHGGFLDDVAGFDASFFG ISPREA 10
 VAMDPQQRLAMEVSWEA FERAGIEPGSVRGSDTGVFMGAY
 PGGYGIGAE LGGFMLTGRAGSVLAGRVSYFFGLEGPAMTV
 DTACSSSLVALHQAA YALRQGECSLALVGGVTVMPTPVMF
 VEFSQQQNLADDGRCKAFAD SADGTGWSEGVGVLLVERLS
 DAQARGHNI LAVV RGS AVNQDGASNGLTAPNGPSQQ RVI
 SALTSAGLTTADVDVVEAHGTGTTLGDPIEAQAVLATY GQ
 DRDQPVLLGSLKSNLGH TQAAAGVSGV I K M V M A L Q N G V V P
 RTLHVEEPSRHVDWTAGA VELVTENQSWPETGRARRA AVS
 SF G FSGTNAHVILE S A PAQPVPPMDTPAPAVTTGVVPLPI
 SAKSLPALADLEDQLRAYLTATPETDLP AVASTLAMTRSV
 FEHRAVLLGEETVTGTAIPDPRIVFVFSQGSQRVGMGEE 20
 LAAA FPLFARLHRQVWDLLDVPDLDVDDTGYVQPALFALQ
 VALFGLLESWGVRPRAVIGH S VGEVAAGYVAGVWSLEDAC
 ALVSARARLMQALPAGGAMVAVPVSEERARVALVDGVEIA
 AVNGPASVVLSGDEAAVLQIAEGLGRWTRL SASHAFHSVR
 M EPMLEEF GQVASELTYQEPRIVMAAGEQVTTPEYWVRQV
 RDTVRFGDQVAAFGDAVFL EIGPDRTLSRLIDGIAML DGD
 D E V R A A V A A L A E L H V Q G V P I D W P A V L G T T G R V L D L P T Y A
 FQHQRYWAASTDRPAGDGHPLLDTVVALPGADGVVLTGRI 30
 SLATHAWLADH A VRGTVLLPGTGF VEMVARAAA EVGCAVV
 DELVIEAPLLL P ASGGVQLS VSVGEADDAGHRPVTVHSQA
 DETEAWVRHVTATISP SGPIVSPPEFEVWPPAQAEPVEVA
 RFYDELAAAGY EYGA AFQGLRAAWRAGETIYAEVVLAEDQ
 TLEAARFTVHPALLDAA LQANILNASGDLRLPFSWGQVQF
 HTTGAATL R VAVTPVADGWTIQATDDAGR PVATVGSVVAR
 PVAGLGATAEDLFALT WNEIPAPGQGGRTVGRFEDLADDG
 PVPELVVFTALPDVDADPLVRTRALTARVLEAIQRWLGE P
 RFADSTLVVVRTGTDLASAAVSGLVRSAQSEHPDRFILVEG
 DSSPVEIGLDEPWLRVDGGRYEVPR LIRL SAE PVQEA AWN
 PDGMVLITGGT GALAGILARHLVAENKARRLL VSRSPD 40
 DALISELTELGAEVGTAVCDVSDRAALARVLAGVPSLTAV
 IHTAGVLDDGVME S LTPQRLDTVLRAKADGAWHLHELT RD
 RDLAAFVMYSSAAGL MGSPGQGNYAAANANAFLDALAVERRA
 EGLPALS LAWG FWEETTGLTANLTGADRDRI RRGGLQTIT
 AERGMRMFD TATQHGEPVLLA API SPVRDGEVPA LLRSLH
 RRGTRRGTTADASAQWL AGLAPEE REGALIKVVRDTAAVV
 LGHADAGTI PVTAAFKDLGLDSLTA VELRNSLAKSTGLRL
 PATMVF D YPTPASLAARLDDLMNPRVSSTALLAE LDRI EG
 MFDSVT FDEKQASLVKDR LSA ALGK WQQISRSADVATVAL
 ANADAGEI LDFIDREFGNPTI 50

配列番号 11

MPDHDKLVEYLRWATAELHTTRAKLQAATEAGTQPLAIVG
 MACRLPGGVSSPEDLWRLVESGTD AISGFPVDRGWDVDGL
 YDPDPDVPGKSYTVEGGFLDAVTGFDAPFFGISPREALAM
 DPQQRLVLEASWEAFERAGIEPGSVRGSDTGVFMGAFPGG
 YGTGADLGGFGMTGGAASVLSGRVSYFFGLEGPAMTVDTV
 CSSSLVALHQAGYALRHGECSLALVGGVTVMSTPQTFVEF
 SRQRGLAADGRCKAFADNADGTGWSEGVGVLVERLSDAQ
 ARGHNI LAVVRGSAVNQDGASNGLTAPNGPSQQRVIRQAL
 ANAGLTGADVDVVEAHGTGTLGDPIEAQAVIATYGRDRD 10
 QPVLLGSLKSNLGHQTAAAGVSGVIKMVMALQNGVVPR
 HIEEPSRHVDWTAGAVQLVTENRPWPELGRARRAAVSSFG
 LSGTNAHVILESAPDQPPAPTTDTPVSAVTAGVVPLP
 ISAKTVPALADLEDRLRTYLTTPD TDLP
 HRAVLLGEDTVTGTAI PDP RVV
 FVFPQGQGWQWQGMGSALL
 TSSTVFAERMAECAA
 AALSEFV
 DWDLLTVLDDPSVV
 DRDV
 VQ PACWAVMISLA
 AAVWQAA
 GIH PDI
 VLGHSQGE
 IAACLA
 GAISLPDAAR
 IVAQRSQ
 LIAHQLTGH
 GAMASISLP
 ADDIP
 TTDKVWIAAHNG
 TSTVIAGDP
 QAVEAVL
 LATCETRGAR
 VRK
 INV
 DYASHTPHVE
 QIRTE
 LLDIT
 TTGIE
 EAHTPAV
 PWL
 STTD
 NTWIDQPLDPT
 YWYRNL
 REPV
 RFGPA
 IDLL
 QTQDNN
 LFIE
 ISASPVLL
 QTMDNA
 ATV
 ATLRR
 DEDTT
 QRL
 LTAF
 AEAH
 VH
 GATIDW
 PTV
 LD
 TTPV
 L
 LD
 PT
 YP
 FQ
 RQ
 RY
 WAT
 SNG
 RSTG
 QGHPL
 LETV
 VAL
 PGTD
 GVAL
 TGR
 ISL
 AT
 HP
 WL
 TD
 HT
 VR
 GT
 VLL
 PGTA
 F
 VEL
 VT
 RA
 AT
 EV
 NC
 QI
 IDE
 LI
 IE
 APL
 PL
 P
 Q
 TDG
 VQL
 S
 VT
 VGE
 A
 DE
 A
 GH
 RP
 VT
 VY
 SQ
 T
 DES
 DD
 WI
 QH
 VT
 AT
 IGP
 GASLP
 ETAA
 WPP
 PA
 HA
 EP
 VNV
 VT
 GL
 YD
 NL
 AA
 AG
 YE
 YG
 PA
 FQ
 GL
 QAA
 WRAG
 DTV
 YAE
 VT
 LA
 EEE
 QA
 QET
 TAR
 FT
 MHP
 ALL
 DA
 ALH
 TIAL
 HDT
 GDL
 HLP
 F
 SWTR
 VQF
 HGT
 GAAT
 LR
 V
 A
 T
 P
 A
 ADG
 W
 NIR
 AT
 DDT
 GRA
 VAT
 IGS
 LV
 TRP
 MAA
 ET
 T
 D
 D
 L
 A
 LT
 TW
 TE
 IP
 A
 P
 EP
 V
 D
 P
 T
 D
 V
 V
 V
 F
 T
 A
 L
 P
 D
 T
 V
 E
 D
 V
 P
 A
 Q
 T
 R
 A
 L
 T
 T
 R
 V
 L
 H
 T
 I
 Q
 E
 WL
 ADD
 D
 R
 T
 L
 I
 V
 R
 T
 G
 D
 L
 A
 S
 A
 A
 V
 S
 G
 L
 V
 R
 S
 A
 Q
 S
 E
 H
 P
 G
 R
 F
 I
 L
 V
 E
 S
 A
 D
 E
 A
 L
 T
 Q
 E
 Q
 L
 A
 A
 T
 A
 G
 L
 D
 E
 P
 R
 L
 R
 I
 T
 G
 G
 R
 Y
 E
 V
 P
 R
 L
 T
 R
 E
 D
 T
 A
 L
 A
 V
 P
 T
 D
 R
 A
 W
 L
 L
 E
 Q
 P
 R
 S
 G
 S
 L
 E
 D
 L
 A
 L
 L
 P
 T
 D
 A
 A
 E
 R
 P
 L
 Q
 A
 G
 E
 V
 R
 I
 G
 V
 R
 A
 A
 G
 M
 N
 F
 R
 D
 V
 V
 V
 A
 L
 G
 M
 V
 T
 D
 T
 R
 L
 A
 G
 G
 E
 A
 A
 G
 V
 V
 L
 E
 V
 G
 T
 DV
 N
 D
 F
 R
 P
 G
 D
 R
 V
 F
 G
 I
 L
 E
 G
 G
 F
 G
 S
 V
 A
 I
 C
 D
 H
 R
 T
 L
 A
 V
 I
 P
 D
 G
 W
 S
 F
 T
 TA
 A
 S
 V
 P
 I
 A
 F
 A
 T
 A
 Y
 Y
 G
 L
 V
 D
 L
 A
 G
 L
 R
 A
 G
 E
 S
 V
 L
 I
 H
 A
 A
 T
 G
 G
 V
 G
 I
 A
 A
 T
 Q
 I
 A
 R
 H
 L
 G
 A
 E
 I
 Y
 G
 T
 A
 S
 V
 G
 K
 Q
 H
 V
 L
 R
 D
 A
 G
 L
 A
 D
 D
 R
 I
 A
 D
 S
 R
 T
 T
 D
 F
 R
 D
 T
 F
 R
 D
 G
 T
 Q
 G
 R
 G
 V
 D
 V
 V
 L
 N
 S
 L
 R
 G
 E
 F
 I
 D
 A
 S
 L
 D
 L
 L
 V
 D
 G
 G
 R
 F
 I
 E
 M
 G
 K
 T
 D
 I
 R
 D
 A
 A
 Q
 I
 P
 D
 A
 T
 Y
 H
 A
 F
 D
 L
 M
 D
 A
 G
 H
 D
 R
 L
 R
 E
 I
 M
 T
 E
 L
 L
 A
 L
 F
 E
 Q
 G
 V
 L
 H
 P
 M
 P
 V
 H
 A
 F
 D
 I
 R
 Q
 A
 R
 E
 A
 F
 S
 W
 M
 S
 R
 A
 R
 H
 I
 G
 K
 L
 V
 L
 T
 I
 P
 Q
 P
 I
 D
 P
 D
 G
 T
 V
 L
 I
 T
 G
 G
 S
 G
 V
 L
 A
 G
 I
 V
 A
 R
 Y
 L
 V
 T
 E
 N
 R
 A
 R
 H
 L
 L
 L
 L
 S
 R
 S
 A
 P
 S
 A
 S
 L
 I
 D
 E
 L
 T
 A
 L
 G
 A
 H
 V
 D
 V
 A
 A
 C
 D
 V
 A
 D
 R
 A
 A
 L
 A
 E
 I
 L
 D
 G
 V
 D
 L
 T
 A
 V
 I
 H
 A
 G
 L
 D
 G
 V
 V
 E
 S
 L
 T
 P
 Q
 R
 L
 D
 T
 V
 L
 T
 P
 K
 A
 D
 G
 A
 W
 H
 L
 H
 E
 L
 T
 R
 D
 R
 D
 L
 A
 A
 F
 I
 V
 Y
 S
 S
 A
 A
 G
 V
 L
 G
 A
 A
 G
 Q
 G
 N
 Y
 A
 A
 A
 N
 A
 F
 L
 D
 A
 L
 A
 V
 R
 R
 L
 E
 G
 L
 P
 G
 L
 S
 L
 A
 W
 G
 L
 W
 E
 D
 A
 S
 G
 L
 T
 A
 D
 L
 T
 D
 A
 R
 D
 R
 I
 R
 R
 S
 G
 Q
 R
 A
 I
 T
 A
 A
 Y
 G
 M
 R
 M
 L
 D
 A
 A
 T
 R
 Q
 S
 E
 A
 I
 L
 L
 A
 A
 P
 I
 S
 P
 I
 Q
 D
 G
 D
 V
 P
 A
 I
 L
 R
 S
 L
 H
 R
 R
 V
 G
 R
 R
 A
 S
 V
 A
 H
 G
 H
 P
 A
 D
 L
 T
 P
 E
 A
 L
 L
 K
 V
 V
 R
 D
 S
 A
 A
 M
 V
 G
 H
 T
 N
 A
 D
 T
 V
 P
 T
 A
 T
 A
 F
 Q
 E
 L
 G
 L
 D
 S
 L
 T
 A
 V
 E
 L
 R
 N
 S
 L
 T
 K
 A
 T
 G
 L
 R
 L
 P
 A
 T
 M
 A
 F
 D
 Y
 P
 T
 P
 D
 A
 L
 A
 A
 R
 L
 G
 E
 L
 L
 A
 G
 E
 A
 A
 P
 K
 A
 A
 A
 V
 R
 R
 A
 T
 A
 S
 D
 E
 10
 20
 30
 40
 50

P L A I V G M A C R L P G G V S S P E D L W R L V E S G T D A I T D F P T D R G
 W D T D T L F D P D P D T P G K T Y T V H G G F L N D V A G F D A P F F G I S P
 R E A V A M D P Q Q R L V L E S S W E A F E R A G I Q P D S I R G S D T G V F M
 G A Y P D G Y G I G A D L A G F G V T A G A G S V L S G R V S Y F F G L E G P A
 M T V D T A C S S S L V A L H Q A A Y A L R Q G E C S L A L V G G V T V M P S P
 R T F I E F S R Q R G L A A D G R S K A F A D A A D G T G F S E G V G V L L V E
 R L S D A Q A K G H N I L A L V R S S A V N Q D G A S N G L T A P N G P S Q Q R
 V I Q S A L A G A G L T S A D V D V V E A H G T G T T L G D P I E A Q A V L A T
 Y G Q D R D Q P V L L G S L K S N L G H T Q A A A G V S G V I K M V M A L Q H N
 T V P A T L H V D A P S R H V D W T A G A V R L A T E N Q P W P E T N R P R R A 10
 G V S S F G V S G T N A H V I L E Q A P A A S P V E P V D T T D V V I P L V V S
 A R S S G S L S D Q A D R L A A L V G S P D A P A L T S L A D A L L T R R T V F
 S Q R A V V V A G S H E Q A A A G L R A L A S G D S H P A L V T G A A G P A R G
 V V L V F P G Q G S Q W A G M G A E L L D T S P V F A A R I A E C A E A L R P W
 V D W S L D E V L R G D A S A D V L G R V D V V Q P A S F A V M V G L A A V W E
 S A G V R P D A V L G H S Q G E I A A A Y V A G A L S L T D A A K I V A V R S R
 L I A A R L A G R G G M A S V A L A P D E A A A K L G R T E L A A V N G P A S V
 V I A G D A E A L D E T L A M L E G E A V R V R R V A V D Y A S H T P H V E E L
 E Q S M A E A L A D V R S R Q P R V G F L S T V T G D W V T E A G A L D G G Y W
 Y R N L R Q P V R F G P A V A S L A E A G Y T V F V E A S A H P V L V Q P V A E 20
 T L D R T D A V V T G T L R R Q D G G L P R L L T S M A E L F V G G V P V N W P
 V L L P A G A V R G W V D L P T Y A F D H Q R Y W L E N R V A T D A A A L G L A
 G A D H P L L G A I V A V P Q S G G V A M T S R L S P R N H P W L A E H T L G G
 V P T V P T S V L V E L A V R A G D E V G C G V V E E L T V D A P L L P E R G
 G V R V Q V I V G A T D A N G Q R G L D I F S A P E D T G Q E A W T R H A T G T
 L A P G G D I A A D V D L S A W P P A N A Q P V D V T D G Y D L L E R A G Y G Y
 G P A F Q G V R A I W R R G E E L F A E V A L E P E L T D T A A R F G L H P A L
 L D A A W H P E L R D E V A E T S P D G R R W W S Q P S R W A G L R L H T A G A
 T V L R V R L A P V D A D S M S L Q A A D E T G D P V L T V D S L S L C A V S A
 D Q L T T A E S S D D A L F R L E W T P L S K A P T A A R S W V P V E T G A D V 30
 A A L D G Q A V V D A V M L E A A G T G D A L E L T C R V L E V V Q A W L T L P
 G W D E S R L V V V T R G A V G A V G D P A G S A V W G L V R A A Q A E N P D R
 I A L L D L D G G R P V E P L L A E E S E P Q L A I R G A E A L V P R L I R A A A
 A T D A P A L F D E S Q T V L I T G G T G S L G G L L A R H L V G R Y G L R R L
 V L V S R R G P D A P G A Y E L A A E L A A H G A E A A L V A C D L T D R D A V
 A R L L T E H H P T A V V H A A G V S D D G V I G T L T S D R L A Y V F G P K A
 T A A R H L D E L T R E L L P D L A A F V T Y S S I S A V F L G A G S G G Y A A
 A N A Y L D G L M A R R H A E G L P G L S L A W G L W D Q E A D G G G M A A G L
 Q D I T R N R M R R R G G V L S F T P A E G M A L F D A A M A T D E A L V V P V
 R L D L P A L R A E A V A E G R S A P V L L R G L V R P G R R L A R T V S G G T 40
 G V L A D L T P E A L L K L V R G R A A A V L G H V D A D A V P V A A A F K D L
 G V D S L T A V E L R N S L A K A T G L R L P A T L V F D Y P T P T V L A G R L
 G E L L A G G T A P V R A A V V R R A A A S D E P L A I V G M A C R L P G G V L
 S P E D L W R L V E S G G D A I S G F P V D R G W D V E N L F D P D P D A A G R
 T Y A V R G G F L D G A A G F D A S F F G I S P R E A Q A M D P Q Q R L V L E V
 S W E A F E R A G I E P G S V R G S D T G V F M G A Y P G G Y G M G T D L G G F
 G M T S V A V S V L A G R V S Y F F G L E G P A M T V D T A C S S S L V A L H Q
 A G S A L R Q G E C S L A L V G G V T V M P T P Q T F V E F S R Q R G L A A D G
 R C K A F A D A A D G T G F S E G V G V L L V E R L S D A Q A R G H N I L A V V
 R G S A V N Q D G A S N G L T A P N G P A Q Q R V I Q S A L A G A G L A S A D V 50

D V V E A H G T G T T L G D P I E A Q A V I A T Y G Q D R D Q P V L L G S L K S
 N L G H T Q A A A G V S G V I K M V M A L Q N G V V P R T L H I D E P S R H I D
 W T A G A V E L V T E N Q S W P E T G R A R R A A V S S F G I S G T N A H V I L
 E S A P A Q P V P L V D T P V S D V T A G V V P L P I S A R T V P A L A D L E D
 Q L R A Y L T T A P E T D L P A V A S T L A M T R S V F E H R A V L L G E E T V
 T G I A V S D P R V V F V F S G Q G S Q R V G M G E E L A A A F P L F A R L H R
 Q V W D L L D V P D L E V D D T G Y V Q P A L F A L Q V A L F G L L E S W G V R
 P R A V I G H S V G E V A A G Y V A G V W S L E D A C T L V S A R A R L M Q A L
 P A G G A M V A V P V S E E R A R A V L V D G V E I A A V N G P A S V V L S G D
 E S A V L R V A E G L G R W T R L S A S H A F H S V R M E P M L E E F R Q V A S
 E L T Y R E P R I V M A A G E Q V T T P E Y W V R Q V R D T V R F G D Q V A A F
 G D A V F L E I G P D R T L S R L I D G I A M L D G D D E V R A A V A A L A M M
 H V Q G V G V D W P A V L G T T G R V L D L P T Y A F Q H E R Y W M V S T G R
 P G G E G H P L L G W G V P V A E A D G R L Y T G R V A R Q D G P V L P V A A F
 V E M A F A A A G G R P I R E L S V D A L L Y I P D D G T A E L Q T W V S E H R
 L T I H A R Y R D T E P W T R L A T A T L D T T E P A T T H T P H P G L I T T A
 L T L T G D E A P A I W H D L T L H T S N A T E L H T H I T P G D D G T L T I T
 A T D A T G Q P V L T A H A A T P T T I P V H T P T T P A D D L L T L T W T Q I
 P T P G P G D G A D I A V C T A L P D P D S D P L A Q T R T L T A Q V L H S I Q
 A S L T G E D R P L V V H T G T G L A S A A V S G L V R S A Q S E H P D R F I L
 V E S D E T L T P D Q L A A V A G L D E P W L R I T D G R Y E V P R L T K T T T
 T A T A T A V S E P V W D P D G T V L I T G G S G A L A G I L A R H L V T E R G
 V R H L L L V S R S T P S T T L I D E L R E L G A H V D V A A C D V S D R A A L
 A R V L D G V D L T A V F H T A G A L D D G V V E S L T P Q R V D A V L R P K A
 D G A W H L H E L T R D R D L T A F V M Y S S A A G V M G A A G Q G N Y A A A N
 A F L D A L A E H R R A D G L P A L S L A W G M W D D A D G M T A S L S G T D H
 R R I R R S G Q R A I T A E H G M R L L D K A S G R S E P V L V A T A M N P I P
 D T D L P A L L R S L Y P K T A R K S Q P I Q E L S P E A L L K I V R D S A A M
 V L G H A N A D T V P T A T A L Q E L G L D S L T A V E L R N S L T K A T G L R
 L

P A T M A F D Y P T P A A L A G R L G E L L A G D T T P A T A A V V R R A T A S
 D E P L A I V G M A C R L P G G V S T P E D L W R L V E S G T D A I T D F P T D
 R G W D T D D L F D P D P D T P G K T Y T V H G G F L D D V A G F D A S F F G I
 S P R E A L A M D S Q Q R L V L E A A W E A F E R A G I E P G S V R G S D T G V
 F M G A Y P D G Y G I G A D L G G F G A T A G A G S V L S G R L S Y F F G L E G
 P A M T V D T A C S S S L V A L H Q A G S A L R Q G E C S L A L V G G V T V I A
 N P Q I F V E F S R Q R G L A A D G R C K A F A D N A D G T G F S E G V G V L L
 V E R L S D A Q A K G H N I L A L V R S S A V N Q D G A S N G L T A P N G P S Q
 Q R V I R Q A L A N A G L T G A E V D V V E A H G T G T T L G D P I E A Q A V L
 A T Y G Q D R D Q P V L L G S L K S N L G H T Q A A A G V S G V I K M V M A L R
 H D T V P A T L H I D E P S R H I D W T A G A V E L V T E N Q P W P V L G R P R
 R A A V S A F G V S G T N A H V I L E S A P D Q P P A P A T D T P A P A A T A G
 V V P L P I S A K T V P A L A D L E D R L R T Y L T T T P E T D L P A V A S T L
 A T T R S L F E H R A V L L G E D T V T G T T I P D P R I V F V F P G Q G W Q W
 Q G M G S A L L T S S T V F A E R M A E C A A A L S E F V D W D L L T V L D D P
 S I V D R V D V V Q P A C W A V M I S L A A V W Q A A G I H P D I V L G H S Q G
 E I A A A C L A G A I S L P D A A R I V A Q R S Q L I A H Q L T G H G A M A S I
 S L P A D D I P T T D K V W I A A H N G T S T V I A G D P Q A L D T V L A T C E
 T H G A R V R K I N V D Y A S H T P H V E Q I R T E L L D I T T D I E A H T P T

10

20

30

40

50

V PWL S T T D N T W I D Q P L D P T Y W Y R N L R E P V R F G P A I D L L Q T
 Q D N N L F I E I S A S P V L L Q T M D N A T T V A T L R R D E D T T Q R L L T
 A F A E A H V H G A T I D W P T V L D T T T P V L D L P T Y P F Q R Q R Y W A
 T S N G R P T S Q G H P L L E T V V A L P G T H G V A L T G R I S L A T H P W L
 T D H T V R G T V L L P G T A F V E L V T H A A T E V N C Q V I D E L I I E A P
 L P L P Q N G G V Q L S V T V G E A D E A G H R P V T V Y S Q T D E S D D W V Q
 H V T A T I A P G V S S S E S A A W P P A Q A E P V N V T G L Y D N L A A A G Y
 E Y G P A F Q G L Q T A W R D G S T V Y A E V T L A E E Q A Q E T A R F T M H P
 A L L D A A L H T I A L H D T A D L Q L P F S W R Q V Q F H G S G A A T L R V A
 V T P A A D G W N I R A T D D T G Q T V A T I G S L V T R P M A A E T T N D L L 10
 A L T W T E I P A P E P V D P A D V V V F T A L P E P G S D P L A Q T R A L T T
 R V L H T I Q E W L A D D D R T L I V R T G T D L A S A A V S G L V R S A Q S E
 H P G R F I L V E S D D E T L T H E Q L A A T A G L D E P R L R I T D G R Y E V
 P R L T R E D T A L A V P E G G A W M L D Q P S R S G T L Q D L R L V P T D A A
 E R P L R P G E V R V G V R A A G L N F R D V A V A L G M V T D T R L I G G E G
 A G V V L E A G P G V E D L R P G D R V F G L L E G G F G P V A V A D R R A L A
 L I P D G W S F T T A A S V P I A F A T A Y Y G L L D L A G L R A G E S V L I H
 A A T G G V G M A A T Q I A R H L G A D V Y A T A S T G K Q H V L R D A G L S D
 D R I A D S R T T G F R E T F R D S T D G R G V D V V L N S L K G D F V D A S L
 D L L V D G G R F I E M G K T D I R D A A Q I P D A T Y R A F D L M D A G P E R
 L R E I I T E L L A L F E Q G V L R P L P V H A F D I R Q A R D A F G W M S R A
 R H I G K L V L T I P Q P I D P D G T V L I T G G S G V L A G I V A R H L V I A
 E G L R N L L L S R S A P S E A L I G E L T A L G A Q V E T A A C D I A D R A
 A L A R V L D G V P L T A V I H T A G A L D D G V V E S L D P Q R L D S V L T P
 K A D G A W H L H E L T R D R D L A A F I M Y S S A A G V L G A A G Q G N Y A A
 A N A F V D A L A V H R R F M G L P A L S L A W G L W D D T S A L T A G L T D S
 D H D R I R R S G A R T I T A E H G M R M F D A A T R Q S E A V L L A A P M G P
 I R G E D V P A L L R G L A T V R Q P R T R A K R D M G P E R L R D R L N G R T
 S V E Q H R I M V E L V L A H A T S V L G H E S P D A I A P D R A F K D L G M D
 S L T A I E L R N H L V A E T G V R L P A T T A F D H P T A D D L A K R L L A E
 V G L T P A P Q R T E A D I R E E V V V R E P A G D D S W T S E P I A I V S M S
 C R A P G G V D S P E S L W R L V E S G T D A I T D F P G D R G W D V A G L Y S
 P D P D T G Y K T Y C V Q G G F L D A A A D F D A A F F G I S P R E A L G M D P
 Q Q R L L L E T S W E A I E R A R I D P R S L R G R N V G V Y V G G A A Q G Y G
 V G A I D Q Q R D N V I T G S S I S L L S G R L S Y A L G L E G P G V T V D T A
 C S S S L V A L H L A C Q A L R Q R E C S M A L V S G V S V I P T P D V F V E F
 S R Q R G L A A D G R C K S F S A S A D G T I W A E G V G V L V L E R L S E A T
 R L G H R V L A V V R G S A V N S D G A S N G L T A P N G V S Q Q R V I R Q A L
 T G A G L T A A D V D V V E A H G T G T K L G D P I E A E A I L A T Y G Q D R S
 T P V C L G S L K S N I G H A M A A S G V L A V I K M V E A M R H G L I P R T L
 H V E E P S P H V D W A S G D V A L L T E N Q P W P D D A K L R R A G V S S F G
 L S G T N A H V V L E Q Y R A P A A P D I T T E H E P L A W T L S A R D P K A
 L R E Q A G R L H A A L T E S P Q W R P L D I G Y S L A T T R S N F A H R A V A
 V G S D R E D L L R A L S K L A D G S A W P A L V T A T A K D R R V A Y L F D G
 Q G S Q R P D M G S G L Y E R F P A F A R A W D R I S A E F G K H L D H S L T D
 V Y L G R G D A A T A D L V D D T L Y A Q A G L F T M E I A L F E L L A E W G V
 R P D F V S G H S I G E T A A A Y A A G V L S L E D V T T L I V A R G R A L R Q
 V P P G A M V A L R A G E D E A R E F L G R T G A A L D L A A V N S P T S V V V
 S G A S E A V A G F R A R W T E S G R E A R T L N V R H A F H S R H V E A V L G
 E F R E V L E S L T F R T P A L P V V S T V T G R L I E P T E L S T S E Y W L R 50

Q V R Q T V R F H D A V R E L S G Q G V G T F V E I G P S G A L A S A G L E C L
 G D E A S F H A V Q R P G S P G D V C L M T A V A E L H A G G T T V D W A T V L
 A G G R A T D L P V Y P F Q H G S Y W L A P V T R A A D G A P S A G V P A P G E
 Y A R P S A P E E P R T M L E L V R L E A A I A L S I T D P G L I A D D S S F L
 D L G F D S I S A L R L S N R L A A V T G L D L P P S L L F D H P T P A E L A A
 R L D E L S A A D L D G A G V Y A L L E E I D E L D D E D L D M T E E E Q T A I
 S E L L T K L S A K W S R

別の態様では、本開示は、少なくとも1つのケトレダクターゼドメインが、(a)保存されたY A A A N触媒性モチーフ内のチロシンに対応する位置でチロシン以外のアミノ酸を含み、かつ、配列番号1内の保存されたF Gヘリックスを含まない、(b)配列番号2におけるS 9 - p k s A ORFのアラニン6 6 3 2に対応する位置でグルタミン酸残基を含む、または(c)配列番号3のW T S 1 2 - p k s B ORFのアミノ酸3 3 8 6から3 5 1 6に対応するアミノ酸を含まない、キメラポリケチドシンターゼを提供する。
 10。

【0021】

別の態様では、本開示は、少なくとも1つのデヒドラターゼドメインが、(a)配列番号4の保存されたH X X X G X X X X Pモチーフ内のS 6 7 9 - p k s B ORFのp k s Bにおける4 2 8 8位のグリシンに対応する位置でアスパラギン酸を含み、(b)配列番号5内のS 1 2 - p k s B ORFにおける3 0 6 6位から3 0 7 0位に対応する位置の保存されたL P F X Wモチーフにおける置換を含み、(c)配列番号6のS 6 7 9 - p k s A ORFのP r o 6 8 4 4からT r p 6 8 7 4の間の位置に対応する欠失を含み、または(d)配列番号7のA、B、C、およびDに対応する位置での置換もしくは欠失を含む、キメラポリケチドシンターゼを提供する。
 20

【0022】

別の態様では、本開示は、少なくとも1つのエノイルレダクターゼドメインが、配列番号8内のS 1 2 - p k s B ORFの1 5 4 6位に対応する位置でリジンを含まない、および/または配列番号8もしくは9内のS 1 2 - p k s Bの1 5 6 8位に対応する位置でアスパラギン酸を含まない、キメラポリケチドシンターゼを提供する。

【0023】

別の態様では、本開示は、(a)配列番号1 2、1 3、もしくは1 4、(b)配列番号1 5、1 6、もしくは1 7、(c)配列番号1 8、1 9、もしくは2 0、(d)配列番号2 1、2 2、もしくは2 3、(e)配列番号2 4、2 5、2 6、もしくは2 7、(f)配列番号2 8、2 9、3 0、もしくは3 1、(g)配列番号3 2、3 3、3 4、もしくは3 5、または(h)配列番号3 6もしくは3 7のアミノ酸配列に対して少なくとも80%の配列同一性を有するドメインを含むキメラポリケチドシンターゼを提供する。
 30

【0024】

配列番号1 2：
 C G C G A C C G G G A C T T G G C C G C G T T C G T C A T G T A C T C C T C C G
 C G G C C G G T G T G A T G G G T G C T G A G G G G C C A G G G C A A C T A C G C
 G G C G G C C A A C G C G T T C C T C G A T G C C C T G G C C G A G C A C C G C
 C G C
 40

配列番号1 3：
 T C A T G T A C T C C T C C G C G G C C G G T G T G A T G G G T G C T G C G G G
 C C A G G G C A A C T T C G C G G C G G C C A A C G C G T T C C T C G A T G C C
 C T G G C C G A G C A C C G C C G C G C T G A C G G C T T G C C C G C A C T C T
 C C C T G G G C A T G G G G T A T G T G G G A C G G A C G G C A G A C G G T A T G A G
 C G G T C A G C G G G C C A T C A C C G C C G A A C A C G G G A T
 40

配列番号1 4：
 G G C G T C G A C C T G A C C G C G G T G T T C C A C A C C G C C G G A G C C C
 T G G A C G A C G G T G T C G T G G A A C T G G T C G C C A C C G C A A T G A A
 50

CCCCGGCGGGGGAGGGTGAAGTCCCCGCGCTGCTGCGTACG
配列番号15:

CGCGACCGGGATCTGGCGGCGTTCGTCATGTACTCCCTCCG
CCGCGGGGCTCATGGCGAGCGAGGGACAGGGCAACTACGC
GGCAGCCAACGCCCTTCCCTGGACGCGCTCGCGGTAGAGCGT
CGT

配列番号16:

TCAATGTACTCCCTCCGCCGCGGGGCTCATGGCGAGCCCCGG
ACAGGGCAACTTCGCGGCAGCCAACCGCCTTCCCTGGACGCG
CTCGCGGTAGAGCGTCGTGCGGAGGGTTTGGCCCGCGCTCT
CGCTGGCGTGGGGTTTCTGGGAGGAAACGACCGGGCTGGG
GGGATTGCAAGACCATCACCGCCGAGCGCGGCAT

10

配列番号17:

GTCGCCGTCCCTGACGGCGGTGATCCACACCGCGGGAGTCC
TCGACGACGGGTGATGGAATTGCTTGGCCGCACCGATGGC
CCCGGTCCGGACGGCGAGGTTCCCGCCCTGCTGCGGTGCG

配列番号18:

GCGGCGGTGTA CGGCCAGAGCGTCCAGGAACCGCGTTGGCC
GCAGCGTAGTTACCTTGTCCCTCAGCGCCCAAGGACGCCGG
CGGCGGACGAGTACACGATGAATGCGGCCAAGTCCCTGTC
GCG

20

配列番号19:

ATCCCCGTACGCCGGCGCTGATGGCACGCTGGCCGCTCAGGC
CGCTCGCGTCCCTCCCACAGTCCCCAGGCCAGGGACAAACC
AGGCAAAACCTCAAGGCGGCAGGTGTACGGCCAGAGCGTCC
AGGAACCGCGTTGGCCGCAGCGAAGTTACCTTGTCCGGCAG
CGCCCAAGGACGCCGGCGGACGAGTACACGA

配列番号20:

TGACCGCAGGATCGCGGGGACGTCCCCGTCCCTGGATCGGG
CTGATCGGCGGCCGAGCAGTTCCACGACACCGTCTGTCGA
GGGCGCCGGCGGTGTGGATCACCGCGGTCAAGGTCTGACGCC

30

配列番号21:

GTCGCCGCCATCAGGCCGTCGAGGTAGGGCGTTCGCGGAGG
GCCGCGTAACCGCCGGAGCCCTCGCCCAAGGAACACCGCGG
AGATGGAGGAGTAGGGTGACGAACGCCGCCAGGTCTGGGGAG
CAA

配列番号22:

ATGCCCTCGGCCGGGTGAAACGACAGCACGCCGCCATGC
CGCCACCGTCTGGCTTCCCTGGTCCCCACAGGCCCAAGGCCAG
GGACAGGCCGGCAGCCCTTCGGCGTGGCCGCCATC
AGGCCGTCGAGGTAGGGCGTTCGCGGCCGCGAAACACCGCGG
AGCCCGCGCCAGGAACACCGCGGAGATGGAGG

40

配列番号23:

GAGCACCAACCCGACCGCGGTCTGTCATGCGGCTGGCGTGT
CCGACGACGGCGTGAATCGGCCTGGTGCCTGGCTCCGGCTCGA
CCTGCCCGCCCTCCGCCGAAGCGGTCTGCCGAGGGGCCGC

配列番号24:

GAGGCCGGGAAGCAGACCAAGTTCCCTGGTGCACCCCGCCC
TGCTGGACGCCCTGGCATCCGGAGGCTGCGCGACGAAGT
GGCCGAGGACGCCCGGACGGCCGGCGCTGGTGGTCTGCAAA

50

CCGTCGCGATGGAACCAAGATCGAGTTCCACGCGACCGGGCG
CGGCGATACTGCGCGTC

配列番号 25 :

GAGGCCGGGAAGCAGACCAAGTTCCCTGGTGCACCCCGCCC
TGCTGACCAACCGCCCTCACCCCTCACCGGGGACGAGGGCACCC
CGCCATCTGGAACCAAGATCGAGTTCCACGCGACCGGGCGCG
GCGATACTGCGCGTC

配列番号 26 :

GTCACGGCACGATCAGCCCCGGCCCTGCCGTCTCGC
TGCCTGGCCTTCGCGGGTGGCGAACCCCTGCACATCGCGGA
CGGCACCCCCGGCCGGCTTCCTCCTGCATCCGGACGGCGACA
CCGGCCGCGACTGGAACCAAGATCGAGTTCCACGCGACCG
GCGCGGGCGATACTGCGC

10

配列番号 27 :

CTCGGTTCGGTGATGGCGTTGCCGAACCTCGGACGGTGTGG
TGCTGACCGGGCAGGATCTCGCGTCAGGACGGTCCGGTTCT
GTCCGTTGCGGCTTCGTTGAAATGGCGTTCGCGGCTGCT
GGTGGTCGCCGATCCGTGAACGTCTGTTGACGGCGCTGC
TGTACATCCCGGACGACGGCACCGCGGAACTGCAGACCTG
GGTCTCTGAACACCGTCTGACCATCCACGACGTTACCGT
GACACCGAACCGTGGACCCGTCTGGCGACCGGCCGCTCTGG
ACACCAACCGCGCCCTGCGACGACCCACACCCCCGCAACCTGG
TCTGATCACCAACGGCGCTGACCCCTGACCGGGTGACGAAGCA
CCGGCGATCTGGAACCAAGATCGAGTTCCACGCGACCGGGCG
CGCGATACTGCGCGTCGGTGAACACCGGTG

20

配列番号 28 :

CAGACGCTGGAGGGCGGCCCGGTTTACGGTGCATCCCGCGC
TGCTGGACGCGGCCCTGGCATCCGGAGGCTGCGCGACGAAGT
GGCCGAGACGAGGCCCGGACGGGCCGGCGCTGGTGGTCGCAA
CCGTCGCGATGGGGTCAGGTTCAAGTTCCATACGACCGGGCG
CGCGACGCTGCGGGTC

30

配列番号 29 :

CAGACGCTGGAGGGCGGCCCGGTTTACGGTGCATCCCGCGC
TGCTGACCAACCGCCCTCACCCCTCACCGGGGACGAGGGCACCC
CGCCATCTGGGGTCAGGTTCAAGTTCCATACGACCGGGCGCG
GCGACGCTGCGGGTC

40

配列番号 30

GTCACGGCACGATCAGCCCCGTCCGGTCCGATCGTCTCGC
CGCCGGCCTTCGCGGGTGGCGAACCCCTGCACATCGCGGA
CGGCACCCCCGGCCGGCTTCCTCCTGCATCCGGACGGCGACA
CCGGCCGCGACTGGGGTCAGGTTCAAGTTCCATACGACCG
GCGCGGGGACGCGTGC

40

配列番号 31 :

CTGGACACCGTCGTGGCGTTGCCGGCGCGGACGGTGTGG
TGCTGACCGGGCAGGATCTCGCGTCAGGACGGTCCGGTTCT
GTCCGTTGCGGCTTCGTTGAAATGGCGTTCGCGGCTGCT
GGTGGTCGCCGATCCGTGAACGTCTGTTGACGGCGCTGC
TGTACATCCCGGACGACGGCACCGCGGAACTGCAGACCTG
GGTCTCTGAACACCGTCTGACCCATCCACGACGCGTACCGT
GACACCGAACCGTGGACCCGTCTGGCGACCGGCCGCTCTGG

50

A C A C C A C C G C G C C T G C G A C G A C C C A C A C C C C G C A C C C T G G
T C T G A T C A C C A C G G G C G C T G A C C C T G A C C G G G T G A C G A A G C A
C C G G C G A T C T G G G G T C A G G T T C A G T T C A T A C G A C C G G C G
C G G C G A C G C T G C G G G T C G C G G T G A C G C C G G T G

配列番号 3 2 :

C G A C C C G T A G C G T C G C C G C G C C G G T A C C G T G G A A C T G C A C
C C G A G C C A T C G C G A C G G T T G C G A C C A C C A G C G C C G G C C G T
C C G G G C T C G T C T C G G C C A C T T C G T C G C G C A G G C T C C G G A T G
C C A G G C C G C G T C G A G C A G G G C G G G A T G C A T G G T G A A G C G G
G C C G T T T C C T G G G C C T G

10

配列番号 3 3 :

C G A C C C G T A G C G T C G C C G C G C C G G T A C C G T G G A A C T G C A C
C C G A G C C A G A T G G C G G G T G C C T C G T C G C C G G T G A G G G T G A
G G G C G G T G G T G A G C A G G G C G G G A T G C A T G G T G A A G C G G G C
C G T T T C C T G G G C C T G

配列番号 3 4 :

C C G T A G C G T C G C C G C G C C G G T A C C G T G G A A C T G C A C C C G A
G T C C A G T C G G C G G C C G G T G T C G C G T C C G G A T G C A G G A G G A
A G C C G G C C G G G T G C C G T C C G C G A T G T G C A G G G G T T C G C C
A C C C G C G A A G G C C G G C A G T G A C G C G C C G G A C C G A T G G T G
G C G G T G A C G T G C T G G A T

20

配列番号 3 5 :

C T G G A A A C C G T C G T G G C A C T G C C C G G C A C C G A C G G G G T G G
C A C T G A C C G G C C G A A T C T C A C G T C A G G A C G G T C C G G T T C T
G T C C G T T G C G G C T T T C G T T G A A A T G G C G T T C G C G G C T G C T
G G T G G T C G C C C G A T C C G T G A A C T G T C T G T T G A C G C G C T G C
T G T A C A T C C C G G A C G A C G G C A C C G C G G A A C T G C A G A C C T G
G G T C T C T G A A C A C C G T C T G A C C A T C C A C G C A C G T T A C C G T
G A C A C C G A A C C G T G G A C C C G T C T G G C G A C C G C C G C T C T G G
A C A C C A C C G C G C C T G C G A C G A C C C A C A C C C C G C A C C C T G G
T C T G A T C A C C A C G G C G C T G A C C C T G A C C G G T G A C G A A G C A
C C G G C G A T C T G G A C T C G G G T G C A G T T C C A C G G T A C C G G C G
C G G C G A C G C T A C G G G T C G C G G T G A C C C C G G C G

30

一部の実施形態では、本発明のポリケチドシンターゼの少なくとも1つのエノイルレダクターゼは、配列番号35～36のいずれか1つに対して少なくとも70%（例えば、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも99%）の配列同一性を有する核酸によってコードされている。

【0025】

配列番号 3 6 :

G A A C A G G G C G A G C A A C T C G G T C A T G A T C T C G C G G A G C C G G
T C G T G G C C G G C A T C C A T C A G G G T G A A G G C A T G G T A G G T G G
C A T C C G G G A T C T G A G C G G C G T C G C G G A T G T C G G T C T G G C C
C A T C T C G A T G A A C C G G C C G C C G T C G A C C A G C A G G T C G A G G
G A G G C G T C G A T G A A C T C

40

配列番号 3 7 :

G A A C A G G G C G A G C A A C T C G G T C A T G A T C T C G C G G A G C C G G
T C G T G G C C G G C A T C C A T C A G G G T G G A A G G C A T G G T A G G T G G
C A T C C G G G A T C T G A G C G G C G T C G C G G A T G T C G G T G C C C A T
C T C G A T G A A C C G G C C G C C G T C G A C C A G C A G G T C G A G G G A G
G C G T C G A T G A A C T C

50

別の態様では、本開示は、前述のポリケチドシンターゼのいずれかをコードする核酸を提供する。

【0026】

本発明の一部の実施形態では、核酸は、配列番号38のアミノ酸配列に対して少なくとも70%（例えば、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも99%）の配列同一性を有する部分を含むLALをさらにコードする。一部の実施形態では、LALは、配列番号38の配列を有する部分を含む。一部の実施形態では、LALは、配列番号38の配列を有する。一部の実施形態では、核酸は、少なくとも1つのオープンリーディングフレーム内でTTA調節性コドンを欠く。

10

【0027】

配列番号38：M P A V E S Y E L D A R D D E L R R L E E A V G Q A G N G R G V
 V V T I T G P I A C G K T E L L D A A A A K S D A I T L
 R A V C S E E E R A L P Y A L I G Q L I D N P A V A S Q L P D P V S M A L P G E
 H L S P E A E N R L R G D L T R T L L A A E R P V L I G I D D M H H A D T A
 S L N C L L H L A R R V G P A R I A M V L T E L R R L T P A H S Q F H A E L L S
 L G H H R E I A L R P L G P K H I A E L A R A G L G P D V D E D V L T G L Y R A
 T G G N L N L G H G L I K D V R E A W A T G G T G I N A G R A Y R L A Y L G S L
 Y R C G P V P L R V A R V A A V L G Q S A N T T L V R W I S G L N A D A V G E A
 T E I L T E G G L L H D L R F P H P A A R S V V L N D L S A R E R R R L H R S A
 L E V L D D V P V E V V A H H Q A G A G F I H G P K A A E I F A K A G Q E L H V
 R G E L D A A S D Y L Q L A H H A S D D A V T R A A L R V E A V A I E R R R N P
 L A S S R H L D E L T V A A R A G L L S L E H A A L M I R W L A L G G R S G E A
 A E V L A A Q R P R A V T D Q D R A H L R A A E V S L A L V S P G A S G V S P G
 A S G P D R R P R P L P P D E L A N L P K A A R L C A I A D N A V I S A L H G R
 P E L A S A E A E N V L K Q A D S A A D G A T A L S A L T A L L Y A E N T D T A
 Q L W A D K L V S E T G A S N E E E G A G Y A G P R A E T A L R R G D L A A A V
 E A G S A I L D H R R G S L L G I T A A L P L S S A V A A A I R L G E T E R A E
 K W L A E P L P E A I R D S L F G L H L L S A R G Q Y C L A T G R H E S A Y T A
 F R T C G E R M R N W G V D V P G L S L W R V D A A E A L L H G R D R D E G R R
 L I D E Q L T H A M G P R S R A L T L R V Q A A Y S P Q A Q R V D L L E E A A D
 L L L S C N D Q Y E R A R V L A D L S E A F S A L R H H S R A R G L L R Q A R H
 L A A Q C G A T P L L R R L G A K P G G P G W L E E S G L P Q R I K S L T D A E
 R R V A S L A A G G Q T N R V I A D Q L F V T A S T V E Q H L T N V F R K L G V
 K G R Q H L P A E L A N A E

20

一部の実施形態では、核酸は、LAL結合部位、例えば、配列番号39の配列（CTA G G G G T T G C）に対して少なくとも80%（例えば、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも99%）の同一性を有するLAL結合部位をさらに含む。一部の実施形態では、LAL結合部位は、配列番号39の配列を含む。一部の実施形態では、LAL結合部位は、配列番号39の配列を有する。一部の実施形態では、LAL結合部位は、配列番号40の配列（G G G G G T）を含む。

30

【0028】

一部の実施形態では、核酸は、LAL結合部位へのLALの結合がオープンリーディングフレームの発現を促進するように位置しているオープンリーディングフレームをさらに含む。一部の実施形態では、オープンリーディングフレームは、化合物生産タンパク質（例えば、ポリケチドシンターゼ）をコードする。

40

【0029】

一部の実施形態では、核酸は、非リボソームペプチドシンターゼをさらにコードする。一部の実施形態では、核酸は、第1のP450酵素をさらにコードする。一部の実施形態では、核酸は、第2のP450酵素をさらにコードする。

50

【0030】

一部の実施形態では、発現ベクターは、前述の核酸のいずれかを含む。一部の実施形態では、発現ベクターは、人工染色体（例えば、細菌人工染色体）である。

別の様では、本開示は、前述のベクターまたはポリケチドシンターゼのいずれかを含む宿主細胞を提供する。一部の実施形態では、ポリケチドシンターゼは、宿主細胞に対して異種である。

【0031】

本発明の一部の実施形態では、宿主細胞（例えば、LALおよび/またはLAL結合部位を天然に欠く宿主細胞）は、組換えLAL（例えば、異種LAL）を発現するように遺伝子操作されている。一部の実施形態では、LALは、構成的に活性である。一部の実施形態では、宿主細胞は、LAL結合部位を核酸内に挿入することによって遺伝子操作されている。一部の実施形態では、LAL結合部位への組換えLALの結合は、核酸（例えば、ポリケチドシンターゼなどの化合物生産タンパク質をコードする核酸）の転写を促進する。一部の実施形態では、LAL結合部位は、LALに対して異種である。一部の実施形態では、LAL結合部位は、LALに対して内在性である。一部の実施形態では、LAL結合部位は、配列GGGGGT（配列番号40）を含む。

10

【0032】

一部の実施形態では、宿主細胞は、異種LAL結合部位へのLALの結合がオープンリーディングフレームの発現を促進するようにオープンリーディングフレームに機能可能に連結している異種LAL結合部位を含む核酸を含む。一部の実施形態では、異種LAL結合部位は、合成LAL結合部位である。一部の実施形態では、異種LAL結合部位は、オープンリーディングフレームに機能可能に連結している内在性LAL結合部位よりも大きな発現を促進する。一部の実施形態では、異種LAL結合部位は、C₁-T₂-A₃-G₄-G₅-G₆-G₇-G₈-T₉-T₁₀-G₁₁-C₁₂（配列番号39）の少なくとも8つの連続的なヌクレオチドを含み、このうち、いずれのヌクレオチドも任意の他のヌクレオチドによって置き換えられていないか、または、G₄、G₅、G₆、G₇、G₈、T₉、およびT₁₀のいずれか3つのヌクレオチド（例えば、G₄、G₇、およびT₉；G₅、G₈、およびT₁₀；またはG₆、G₇、およびG₈）以外の最大6つのヌクレオチドが、任意の他のヌクレオチドによって置き換えられている。

20

【0033】

30

一部の実施形態では、組換えLALは、配列番号38の配列に対して少なくとも70%（例えば、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、または少なくとも99%）の配列同一性を有する部分を含む。一部の実施形態では、組換えLALは、配列番号38の配列を有する部分を含む。一部の実施形態では、組換えLALは、配列番号38のアミノ酸配列を有する。

【0034】

40

一部の実施形態では、宿主細胞は、細菌（例えば、ストレプトマイセス・アムボファシエンス（*Streptomyces ambofaciens*）、ストレプトマイセス・ハイグロスコピカス（*Streptomyces hygroskopicus*）、またはストレプトマイセス・マラエンシス（*Streptomyces malayensis*）などのアクチノバクテリウム）である。一部の実施形態では、アクチノバクテリウムは、S1391、S1496、またはS2441である。

【0035】

50

一部の実施形態では、宿主細胞は、化合物生産タンパク質（例えば、ポリケチドシンターゼ）の発現を増強するように修飾されている。例えば、一部の実施形態では、宿主細胞は、（i）化合物生産タンパク質（例えば、ポリケチドシンターゼ）を発現する内在性遺伝子クラスターの欠失、（ii）化合物生産タンパク質（例えば、ポリケチドシンターゼ）を発現する異種遺伝子クラスターの挿入、（iii）抗生物質刺激への宿主細胞の曝露、および/または（iv）同種プロモーターと比較して化合物の発現を少なくとも2倍増大させる異種プロモーターの導入によって、化合物生産タンパク質（例えば、ポリケチ

ドシンターゼ)の発現を増強させるように修飾されている。ポリケチドの発現を増強させるためのさらなる方法は、培地条件を成長のために最適化することである。これには、発酵が液体培地で行われても固体培地で行われても培地の具体的な化学組成および栄養組成、発酵の時間的経過、ならびに発酵ランの容積／スケールが含まれる。

【0036】

別の態様では、本開示は、前述の宿主細胞のいずれかを適切な条件下で培養する工程を含む、ポリケチドを生産する方法を提供する。

別の態様では、本開示は、前述のポリケチドシンターゼのいずれかを発現するように遺伝子操作された宿主細胞を、ポリケチドシンターゼがポリケチドを生産するのに適した条件下で培養する工程を含む、ポリケチドを生産する方法を提供する。

10

【0037】

別の態様では、本開示は、ポリケチドシンターゼの活性を調節する方法であって、(a)親ポリケチドシンターゼをコードする親核酸配列を提供する工程、および(b)親核酸配列の少なくとも1つのコドンを修飾する工程であって、コドンが、親ポリケチドシンターゼの少なくとも1つのドメインの保存されたモチーフ内の残基を特定し、修飾が、少なくとも1つのドメインの酵素活性または調節活性の改変(例えば、ドメインを不活性にする改変)をもたらす、工程を含む、方法を提供する。

【0038】

別の態様では、本開示は、化合物を生産する方法であって、(a)親ポリケチドシンターゼをコードする親核酸を提供する工程、(b)親核酸の少なくとも1つのコドン(例えば、-ケトンプロセシングドメインをコードする核酸の部分内のコドン)を修飾して、化合物を生産することができる修飾されたポリケチドシンターゼをコードする修飾された核酸を生成させる工程であって、コドンが、ポリケチドシンターゼの少なくとも1つのドメインの保存されたドメイン内の残基を特定し、修飾が、ポリケチドシンターゼの少なくとも1つのドメインの酵素活性の改変をもたらす、工程、(c)修飾された核酸を宿主細胞に導入する工程、および(d)修飾されたポリケチドシンターゼによる化合物の発現を可能にするのに適した条件下で宿主細胞を培養する工程を含み、それによって化合物を生産する、方法を提供する。

20

【0039】

別の態様では、本開示は、化合物を生産する方法であって、(a)化合物を生産することができる親ポリケチドシンターゼを提供する工程、(b)親ポリケチドシンターゼのアミノ酸配列を決定する工程、(c)親ポリケチドシンターゼをコードする親核酸を提供する工程、(d)親核酸の少なくとも1つのコドンを修飾して、化合物を生産することができる修飾されたポリケチドシンターゼをコードする修飾された核酸配列を得る工程であって、コドンが、ポリケチドシンターゼの少なくとも1つのドメインの保存されたドメイン(例えば、-ケトンプロセシングドメイン)内の残基を特定し、修飾が、少なくとも1つのドメインの酵素活性の改変(例えば、活性が低下する)をもたらす、工程、(e)修飾された核酸を宿主細胞に導入する工程、(f)修飾されたポリケチドシンターゼによる化合物の発現を可能にするのに適した条件下で宿主細胞を培養する工程、および(g)修飾されたポリケチドシンターゼによって生産される化合物を回収する工程を含み、それによって化合物を生産する、方法を提供する。

30

【0040】

別の態様では、本開示は、化合物を生産する方法であって、(a)親ポリケチドシンターゼの構造を決定する工程、(b)親ポリケチドシンターゼをコードする親核酸を生産する工程、(c)核酸を修飾して、修飾されたポリケチドシンターゼをコードする修飾された核酸を生産する工程であって、修飾されたポリケチドシンターゼの少なくとも1つのドメイン(例えば、-ケトンプロセシングドメイン)が、親ポリケチドシンターゼと比較して酵素活性が改変される(例えば、酵素活性が低下する)工程、(d)修飾された核酸配列を宿主細胞に導入する工程、および(e)修飾されたポリケチドシンターゼによる化合物の発現を可能にするのに適した条件下で宿主細胞を培養する工程を含み、それによつ

40

50

て化合物を生産する、方法を提供する。

【0041】

別の態様では、本開示は、化合物のライブラリーを生産する方法であって、(a)親ポリケチドシンターゼをコードする親核酸配列を提供する工程、(b)親核酸配列の少なくとも1つのコドンを修飾して、化合物を生産することができる第1の修飾されたポリケチドシンターゼをコードする第1の修飾された核酸を生成する工程、(c)親核酸の少なくとも1つのコドンを修飾して、化合物を生産することができる第2の修飾されたポリケチドシンターゼをコードする第2の修飾された核酸を生成する工程であって、第1および第2の修飾された核酸が異なる、工程、(d)第1および第2の修飾された核酸配列を1つまたは複数の宿主細胞に導入する工程、および(e)第1および第2の修飾されたポリケチドシンターゼによる化合物の発現を可能にするのに適した条件下で1つまたは複数の宿主細胞を培養し、それによって、化合物のライブラリーを生産する工程を含む、方法を提供する。

【0042】

別の態様では、本開示は、前述の方法のいずれかによって生産された化合物を提供する。

定義

用語「ドメインの保存された領域」は、本明細書において使用する場合、活性な同一タイプの全てのドメインにおいて実質的に同一であるポリケチドシンターゼのドメインの部分を指す。

【0043】

本明細書において使用する場合、用語「遺伝子操作されたポリケチドシンターゼ」は、その設計および/または生産が人の手を加えることを伴う、非天然のポリケチドシンターゼを記載するために使用される。例えば、一部の実施形態では、「遺伝子操作された」ポリケチドシンターゼは、ポリケチドシンターゼをコードする非天然ポリヌクレオチドの生産によって調製される。

【0044】

「含有するように遺伝子操作された」および/または「発現するように遺伝子操作された」細胞は、細胞内で天然に生じないタンパク質を含有するおよび/または発現するように修飾されている細胞を指す。細胞は、例えば、核酸を含むベクターの導入によってタンパク質をコードする核酸を導入することによって、タンパク質を含有するように遺伝子操作され得る。

【0045】

用語「機能的に不活性な」は、本明細書において使用する場合、活性を有さない、または検出点を下回る活性を有する、ポリケチドシンターゼのドメインを指す。

用語「低分子を生産する遺伝子クラスター」は、本明細書において使用する場合、1つまたは複数の化合物生産タンパク質をコードする遺伝子のクラスターを指す。

【0046】

用語「異種」は、本明細書において使用する場合、天然に存在しない2つ以上のタンパク質、核酸、化合物、および/または細胞の間の関係を指す。例えば、配列番号38の配列を有するLALは、S18 Streptomyces (S t r e p t o m y c e s) 株において天然に生じ、したがって、その株に対して同種であり、したがって、S12 S t r e p t o m y c e s 株に対して異種である。

【0047】

本明細書において使用される用語「同種の」は、天然に存在する2つ以上のタンパク質、核酸、化合物、および/または細胞の間の関係を指す。例えば、配列番号38の配列を有するLALは、S18 Streptomyces 株において天然に生じ、したがって、その株に対して同種である。

【0048】

用語「修飾されたドメイン」は、本明細書において使用する場合、少なくとも1つのア

10

20

30

40

50

ミノ酸残基が参照配列から改変されているポリケチドシンターゼのドメインを指す。

「ポリケチドシンターゼ」は、ポリケチドを生産することが可能なマルチドメイン酵素のファミリーに属する酵素を指す。ポリケチドシンターゼは、細菌、真菌、植物、または動物で天然に発現され得る。

【0049】

用語「組換えの」は、本明細書において使用する場合、合成方法を使用して生産されるタンパク質を指す。

本明細書において使用する場合、用語「参照ポリケチドシンターゼ」は、修飾されているドメインの配列に対しては除いて、遺伝子操作されたポリケチドシンターゼの配列に対して少なくとも80%の同一性（例えば、少なくとも85%の同一性、少なくとも90%の同一性、少なくとも95%の同一性、少なくとも99%の同一性、または100%の同一性）を有する配列を有するポリケチドシンターゼを指す。

10

【図面の簡単な説明】

【0050】

【図1】ポリケチドシンターゼによるポリケチドの生合成を描写する画像。

【図2A】ポリケチドシンターゼのドメインによるマロニル-ケトンの修飾を描写する画像。

【図2B】ポリケチドシンターゼのドメインによるマロニル-ケトンの修飾を描写する画像。

20

【図3】ポリケチドの生合成を描写する画像。

【図4A】ポリケチドシンターゼ内の不活性なドメインを生じさせる突然変異を描写する配列アラインメント。

【図4B】ポリケチドシンターゼ内の不活性なドメインを生じさせる突然変異を描写する配列アラインメント。

【図5A】ケトレダクターゼドメインの失活を描写する画像。

【図5B】修飾されたポリケチドシンターゼによる化合物の生成を描写する画像。

【図5C】修飾されたポリケチドシンターゼによる化合物の生成を描写する画像。

【図6A】修飾されたポリケチドシンターゼによる化合物の生成を描写する画像。

【図6B】修飾されたポリケチドシンターゼによる化合物の生成を描写する画像。

【図6C】修飾されたポリケチドシンターゼによる化合物の生成を描写する画像。

30

【図7】図7(A)は、修飾されたポリケチドシンターゼによって生成された化合物の画像であり、図7(b)は、修飾されたポリケチドシンターゼによる化合物の生成を描写するグラフであり、図7(C)は、CEP250への化合物の結合を描写する画像であり、図7(D)は、CEP250への化合物の結合を描写するグラフ。

【図8A】粗抽出物における化合物のアッセイに使用するための、target-ID法を描写する画像。

【図8B】CEP250およびCBY1への化合物の結合を描写する画像。

【図9A】ポリケチドシンターゼ内のケトレダクターゼドメインの失活を描写する画像。

【図9B】ポリケチドシンターゼ内のケトレダクターゼドメインの失活による、環が拡大した化合物の生成を描写する画像。

40

【発明を実施するための形態】

【0051】

本発明者らは、失活した-ケトンプロセシングドメイン、すなわち、ケトレダクターゼ(KR)ドメイン、デヒドラターゼ(DH)ドメイン、およびエノイルレダクターゼ(ER)ドメインをもたらす、ポリケチドシンターゼ内の短鎖タンパク質配列を、別のポリケチドシンターゼ内の生きたドメインにグラフトして、ドメインの活性を失活させ、ポリケチドシンターゼによって生産されるポリケチドの化学構造を改変することができるということを発見した。

【0052】

化合物

50

本発明の方法で生産され得る化合物には、限定されないが、エリスロマイシンなどのポリケチドおよびポリケチドマクロライド系抗生物質；ラパマイシンおよびFK506などのハイブリッドポリケチド／非リボソームペプチド；ゲンタマイシン、カナマイシン、ネオマイシン、トプラマイシンなどのアミノグリコシド系抗生物質を含む炭水化物；ベンゾフラノイド；ベンゾピラノイド；フラボノイド；バンコマイシンを含むグリコペプチド；ダブトマイシンを含むリポペプチド；タンニン；リグナン；多環芳香族天然生成物、テルペノイド、ステロイド、ステロール、リネゾリドを含むオキサゾリジノン；ポリミキシン、非リボソームペプチド、カルバペネム、セファロスポリン、およびペニシリンを含むラクタム系抗生物質を含む、アミノ酸、ペプチド、およびペプチド系抗生物質；プリン、ブリジン、ポリピロール、テトラサイクリン、キノロン、およびフルオロキノロン；ならびにスルホンアミドが含まれる。

10

【0053】

タンパク質

ポリケチドシンターゼ

ポリケチドシンターゼ (PKS) は、ポリケチドを生産するマルチドメイン酵素のファミリーである。I型ポリケチドシンターゼは、モジュールに組織化されたいくつかのドメインを含む、大きなモジュラータンパク質である。ポリケチドシンターゼ内に一般に存在するモジュールには、i) ローディングモジュール、ii) 伸長モジュール、および iii) 最終ポリケチドが直鎖状であるかまたは環状であるかに応じて放出モジュールおよび / または環化モジュールが含まれる。モジュール内で一般に見られるドメインは、アシルトランスフェラーゼ、アシル担体タンパク質、ケトシンターゼ、ケトレダクターゼ、デヒドラターゼ、エノイルレダクターゼ、メチルトランスフェラーゼ、スルフヒドロラーゼ、およびチオエステラーゼである。

20

【0054】

ポリケチド鎖およびスターター基は、チオエステル連結を介してケトシンターゼドメイン (ポリケチド鎖) およびアシルトランスフェラーゼドメイン (ローディング基およびマロニル伸長ユニット) 内の活性部位システィンのチオール基に一般に結合している。アシル担体タンパク質 (ACP) への結合は、ACP のセリンヒドロキシルに結合しているホスホパンテイニル基のチオールが介在して、成長しているポリケチド鎖へのチオエステル連結を形成する。成長しているポリケチド鎖は、1つのチオール基から別のチオール基へトランスアシル化によって引き渡され、加水分解または環化による合成の後に放出される。

30

【0055】

ポリケチドの合成は、ローディングモジュールにおけるアシルトランスフェラーゼによって触媒される、PKS のアシル担体タンパク質ドメイン上にローディングされているスターターユニットによって開始される。伸長ユニット、例えばマロニル-COA は、別のアシルトランスフェラーゼドメインによって触媒される現モジュールのアシル担体タンパク質ドメイン上にローディングされる。ポリケチド鎖は次いで、モジュール n のアシル担体タンパク質ドメインから $n+1$ モジュールのケトシンターゼドメインに渡された後、次の伸長モジュールによって延長される。アシル担体タンパク質が結合した伸長ユニットは、ケトシンターゼドメインに結合したポリケチド鎖と反応して CO_2 を排出して、アシル担体タンパク質に結合した伸長したポリケチド鎖を生産する。付加された伸長ユニットの各々は、次いで、-ケトプロセッシングドメイン、すなわち、ケトレダクターゼ (これは、延長基のカルボニルをヒドロキシに還元する)、デヒドラターゼ (これは、 H_2O を排出してアルケンを生産する)、およびエノイルレダクターゼ (これは、アルケンを還元して飽和炭化水素を生産する) によって修飾され得る。ポリケチドの合成が完了したら、放出モジュール内のチオエステラーゼドメインは、完成したポリケチド鎖を、最後の伸長モジュールのアシル担体タンパク質から加水分解する。PKS から放出された化合物は、次いで、他のタンパク質、例えば、非リボソームペプチドシンターゼによってさらに修飾され得る。PKS によるポリケチドの合成の例を図3において描写する。一部のケースで

40

50

は（例えば、化合物 1 をコードするクラスターであるラパマイシンおよび X 1 ）、生合成クラスターは、ポリケチドメガシンターゼおよび非リボソームペプチドシンターゼ（N R P S ）を内在させている。このハイブリッドアーキテクチャーは、ハイブリッド P K S / N R P S と呼ばれる。ラパマイシンおよび化合物 1 のケースでは、N R P S モジュールは、分子の F K B P 1 2 結合領域内にピペコレート部分を挿入する（図 3 ）。

【 0 0 5 6 】

- ケトンプロセシングドメイン

- ケトンプロセシングドメインは、ポリケチドの合成の間に付加された延長基の修飾をもたらす、P K S 内のドメインである。各 - ケトンプロセシングドメインは、延長基の酸化状態を変化させ得る。 - ケトンプロセシングドメインには、ケトレダクター（これは、延長基のカルボニルをヒドロキシに還元する）、デヒドラターゼ（これは、H₂O を排出してアルケンを生産する）、およびエノイルレダクター（これは、アルケンを還元して飽和炭化水素を生産する）が含まれる。

10

【 0 0 5 7 】

非機能的なドメイン

- ケトン修飾ドメインの包括的分析は、ポリケチドの最終的な構造に影響しない非機能的な - ケトン修飾ドメインの存在を示した。これらのドメインは、「死んでいる」可能性が高い（図 3 ）。鋳型として機能的に関連するドメインの結晶構造を使用するホモロジー モデリングと組み合わせた、タンパク質配列アラインメントは、これらの機能的でないドメインが、鍵となる触媒性モチーフおよび基質結合モチーフに、それらを不活性にする突然変異を有することを明らかにした（図 4 A および 4 B ）。それでも、これらの「死んだ」ドメインは、進化を経て遺伝子クラスター内に保持されており、このことは、これらのドメインがその代わりに構造的役割を有すること、すなわち、効率的なアセンブリラインポリケチド合成のためにモジュール内の触媒性ドメインの適切な空間的組織化を維持していることを示唆する。ドメインの活性は、進化によって選択的に「オフ」にされている可能性があり、天然生成物の化学構造、タンパク質の標的の係合、および進化した分子の生理化学的特性を修飾している。

20

【 0 0 5 8 】

ケトレダクター ゼドメインレベルでの遺伝子操作では、3 つの K R 死滅ドメインが分析されている：S 9 の K R 3 、K R 6 - S 3 0 3 、および K R 3 - S 3 9 9 。 S 9 の K R 3 は、保存された触媒性 Y A A A N モチーフの近くに A 1 a から G 1 u への単一の置換を含む。理論に拘束はされないが、ホモロジー モデリング（P D B 2 F R O を使用する）は、この位置のグルタミン酸が近くのアルギニンと塩橋を形成できること、および得られた塩橋が基質キャップ領域（F G ）の可動性をブロックし、ポリケチド基質へのケトレダクター ゼ活性部位のアクセスを防止することを示唆した。S 3 0 3 死滅 K R 6 ドメインおよび S 3 9 9 死滅 K R 6 ドメインは、より顕著な損傷を含む。S 3 0 3 では、触媒性 T y r は P h e によって置き換えられ、F G ヘリックスは欠失していた。S 3 9 9 では、触媒性残基および基質結合残基を包含する、より大きな 1 5 0 残基の欠失が存在する。

30

【 0 0 5 9 】

一部の実施形態では、本発明のポリケチドシンターゼの少なくとも 1 つのケトレダクター ゼドメインは、配列番号 1 2 ~ 2 3 のいずれか 1 つに対して少なくとも 7 0 %（例えば、少なくとも 7 5 %、少なくとも 8 0 %、少なくとも 8 5 %、少なくとも 9 0 %、少なくとも 9 5 %、少なくとも 9 9 %）の配列同一性を有する核酸によってコードされる。

40

【 0 0 6 0 】

デヒドラターゼドメインレベルでの遺伝子操作では、4 つの「死滅した」D H ドメインが分析された：S 6 7 9 - D H 7 、S 1 2 - D H 6 、S 1 2 - D H 7 、S 6 7 9 - D H 4 、および S 1 2 - D H 2 。D H ドメインの必須の活性部位残基は、4 つの鍵となる保存されたモチーフ：H X X X G X X X X P 、G Y X Y G P X F 、D X X X (Q / H) 、および L P F X W にわたって分布している。S 6 7 9 - D H 7 は、ポリケチド基質を脱プロトン化させて脱水反応を開始させる H i s 残基を含有する H X X X G X X X X P モチーフにお

50

いて、G 1 y から A s p への単一の置換を有する。S 1 2 - D H 6 および S 1 2 - D H 7 は、L P F X W モチーフ内に置換を有する。S 6 7 9 - D H 4 は著しい内部欠失を含有し、S 1 2 - D H 2 は、D H 活性部位を含む 4 つの鍵となるモチーフの全てにおいて突然変異を内在させている。ラパマイシンなどのハイブリッド P K S / N R P S クラスターは、大環状化のためにヒドロキシルを要し、したがって、S 1 2 - D H 2 「死滅」 D H ドメインは、環化および生物学的活性について不活性であり続けるに違いない。

【 0 0 6 1 】

一部の実施形態では、本発明のポリケチドシンターゼの少なくとも 1 つのデヒドラターゼドメインは、配列番号 2 4 ~ 3 5 のいずれか 1 つに対して少なくとも 7 0 % (例えれば、少なくとも 7 5 %、少なくとも 8 0 %、少なくとも 8 5 %、少なくとも 9 0 %、少なくとも 9 5 %、少なくとも 9 9 %) の配列同一性を有する核酸によってコードされる。

10

【 0 0 6 2 】

エノイルレダクターゼドメインレベルでの遺伝子操作では、S 1 2 および S 6 1 からの 2 つの E R ドメインを分析した。両死滅 E R ドメインは、各クラスターのローディングモジュール内に位置しており、したがって、スターターユニットの化学と関連しており、マロニル由来のポリケチド鎖の化学とは関連していない。両死滅ドメインにおいて、不变の L y s - A r g 二分子は置換されているかまたは欠失している。

【 0 0 6 3 】

一部の実施形態では、本発明のポリケチドシンターゼの少なくとも 1 つのエノイルレダクターゼドメインは、配列番号 3 6 または 3 7 に対して少なくとも 7 0 % (例えれば、少なくとも 7 5 %、少なくとも 8 0 %、少なくとも 8 5 %、少なくとも 9 0 %、少なくとも 9 5 %、少なくとも 9 9 %) の配列同一性を有する核酸によってコードされる。

20

【 0 0 6 4 】

L A L

L A L は、3 つのドメイン、すなわち、ヌクレオチド結合ドメイン、誘導因子結合ドメイン、および D N A 結合ドメインを含む。L A L を含む調節タンパク質の構造的クラスの決定的な特徴は、A A A + A T P a s e ドメインの存在である。ヌクレオチドの加水分解は、タンパク質における大きな立体構造の変化、および / または多量体化と結び付き、ヌクレオチドの結合および加水分解は、L A L の活性 (例えれば、L A L の活性時間) を制御する「分子タイマー」に相当する。L A L は、誘導因子結合部位への低分子リガンドの結合によって活性化される。ほとんどのケースでは、L A L のアロステリック誘導因子は未知である。関連するタンパク質 M a 1 T のケースでは、アロステリック誘導因子はマルトリオースである。L A L タンパク質の考えられる誘導因子には、化合物 (例えれば、ポリケチド) の生合成を引き起こす環境において見られる低分子が含まれる。L A L の調節は、化合物生産タンパク質 (例えれば、ポリケチドシンターゼ) の生産を制御し、外部環境刺激の存在下での化合物 (例えれば、ポリケチド) 生産の活性化をもたらす。したがって、株内に存在する一方で、(i) L A L が活性化されていない、(i i) 株がコンセンサスと異なる L A L 結合部位を有している、(i i i) 株が L A L 調節因子を欠いている、または (i v) L A L 調節因子が実験室条件下あまり発現しない可能性があるもしくは発現しない、という理由から化合物を生産しない、低分子を生産する遺伝子クラスター (例えれば、P K S 遺伝子クラスター) が存在する。既知の P K S L A L の L A L の D N A 結合領域が高度に保存されているため、既知の L A L は、それらが天然に調節するもの以外の P K S 遺伝子クラスターを活性化するために、区別せずに使用することができる。一部の実施形態では、L A L は融合タンパク質である。

30

【 0 0 6 5 】

一部の実施形態では、L A L は、非 L A L D N A 結合ドメインを含むように修飾され得、これによって、L A L ヌクレオチド結合ドメインおよび非 L A L D N A 結合ドメインを含む融合タンパク質を形成する。ある特定の実施形態では、非 L A L D N A 結合ドメインは、プロモーターのタンパク質結合部位への D N A 結合ドメインの結合が目的の遺伝子 (例えれば、本明細書において記載される、化合物生産タンパク質をコードする遺伝子

40

50

) の発現を促進するように位置しているタンパク質結合部位を含むプロモーターに結合し得る。非 L A L D N A 結合ドメインには、当技術分野において知られているあらゆる D N A 結合ドメインが含まれ得る。一部の場合では、非 L A L D N A 結合ドメインは、転写因子 D N A 結合ドメインである。非 L A L D N A 結合ドメインの例には、限定されないが、塩基性ヘリックス - ループ - ヘリックス (b H L H) ドメイン、ロイシンジッパー ドメイン (例えば、塩基性ロイシンジッパー ドメイン) 、 G C C ボックスドメイン、ヘリックス - ターン - ヘリックスドメイン、ホメオドメイン、 s r f 様ドメイン、ペアードボックスドメイン、ウイングドヘリックスドメイン、ジンクフィンガードメイン、 H M G - ボックスドメイン、 W o r 3 ドメイン、 O B - フォールドドメイン、免疫グロブリンドメイン、 B 3 ドメイン、 T A L エフェクタードメイン、 C a s 9 D N A 結合ドメイン、 G A L 4 D N A 結合ドメイン、および当技術分野において知られているあらゆる他の D N A 結合ドメインが含まれる。一部の場合では、プロモーターは、融合タンパク質がプロモーターに結合することができ、目的の遺伝子の発現を誘発または阻害することができるよう 10 に、目的の遺伝子の上流に位置している。ある特定の場合では、プロモーターは、目的の遺伝子を含有する核酸 (例えば、染色体、プラスミド、フォスマミド、または当技術分野において知られているあらゆる他の核酸構築物) に導入された異種プロモーターである。他の場合では、プロモーターは、目的の遺伝子の上流に位置する既存のプロモーターである。プロモーター内のタンパク質結合部位は、例えば、非 L A L タンパク質結合部位であり得る。ある特定の実施形態では、タンパク質結合部位は非 L A L D N A 結合ドメインに結合し、これによって、同族の D N A 結合ドメイン / タンパク質結合部位対を形成する 20 。

【 0 0 6 6 】

一部の実施形態では、 L A L は、配列番号 4 1 ~ 6 2 のいずれか 1 つに対して少なくとも 7 0 % (例えば、少なくとも 7 5 % 、少なくとも 8 0 % 、少なくとも 8 5 % 、少なくとも 9 0 % 、少なくとも 9 5 % 、少なくとも 9 9 %) の配列同一性を有する核酸によってコードされるか、または、配列番号 6 3 ~ 7 3 のいずれか 1 つに対して少なくとも 7 0 % (例えば、少なくとも 7 5 % 、少なくとも 8 0 % 、少なくとも 8 5 % 、少なくとも 9 0 % 、少なくとも 9 5 % 、少なくとも 9 9 %) の配列同一性を有する配列を有する。 20

【 0 0 6 7 】

配列番号 4 1 :

A T G C C T G C C G T G G A G T G C T A T G A A C T G G A C G C C C G C G A T G
 A C G A G C T C A G A A A A C T G G A G G G A G G T T G T G A C C G G G C G G G C
 C A A C G G C C G G G G T G T G G T G G T C A C C A T C A C C G G A C C G A T C
 G C C T G C G G G C A A G A C C G A A C T G C T C G A C G C A G C C G C C G C G A
 A G G C C G A C G C C A T C A C G T T A C G A G C G G T C T G C T C C G C G G A
 G G A A C A G G G C A C T C C C G T A C G C C C T G A T C G G G C A G C T C A T C
 G A C A A C C C G G C G C T C G C C T C C C A C G C G C T G G A G G C C G G C T
 G C C C G A C C C T C C C G G G C G A G C A C C T G T C G C C G G A G G G C C G A
 G A A C C G G C T G C G C A G C G A C C T C A C C C G T A C C C T G C T G G C G
 C T C G C C G C G A A C G G C C G G T G C T G A T C G G C A T C G A C G A G T
 C A C A C G C G A A C G C T T T G T G T C T G C T C C A C C T G G C C C G A A G
 G G T C G G C T C G G C C C G G A T C G C C A T G G T C C T C A C C G A G T T G
 C G C C G G C T C A C C C C G G C C C A C T C A C A G T T C C A G G C C G A G C
 T G C T C A G C C T G G G G C A C C A C C G C G A G A T C G C G C T G C G C C C
 G C T C A G C C C G A A G C A C A C C G C C G A G G C T G G T C C G C G C C G G T
 C T C G G T C C C G A C G T C G A C G A G G A C G T G C T C A C G G G G T T G T
 A C C G G G C G A C C G G G C G A A C C T G A A C C T C A C C C G C G G A C T
 G A T C A A C G A T G T G C G G G A G G G C T G G G A G A C G G G A G G G A C G
 G G C A T C A G C G C G G G C C G C G C G T A C C G G C T G G C A T A C C T C G
 G T T C C C T C T A C C G C T G C G G C C C G G T C C C G T T G C G G G G T C G C 50

A C G G G T G G C C G C G T G C T G G G G C C A G A G C G C C A A C A C C A C C
 C T G G T G C G C T G G A T C A G C G G G C T C A A C G C G G A C G G C G G T G G
 G C G A G G C A A C C G A G A T C C T C A C C G A A G G C G G C C T G C T G C A
 C G A C C T G C G G T T C C C G C A C C C G G C G G C C C G T T C G G T G G T A
 C T C A A C G A C A T G T C C G C C C A G G A A C G A C G C C G C C T G C A C C
 G G T C C G C T C T G G A A G T G C T G G A C G A C G T G C C C G T G G A A G T
 G G T C G C G C A C C A C C A G G T C G G C G G C C G G T C T C C T G C A C G G C
 C C G A A G G C C G C C G A G A T A T T C G C C A A G G C C G G C C A G G A G C
 T G C A T G T G C G C G G C G A G T T G G A C A C C G C G T C C G A C T A T C T
 G C A A C T G G C C C A C C A G G C C T C C G A C G A C G C C G T C A C C G G G 10
 A T G C G G G C C G A G G C C G T G G C G A T C G A G G C G C C G C C G A A C C
 C G C T G G G C C T C G A G G C C G G C A C C T C G A C G A G G C T G A C C G T C G T
 C G C C C G T G C C G G G C T G C T C T T C C C C G A G C A C A C G G C G C T G
 A T G A T C C G C T G G C T G G C G T C G G C G G G C G G T C C G G C G A G G
 C A G C C G G G C T G C T G G C C T C G C A G C G C C C C C G T G C G G T C A C
 C G A C C A G G A C A G G G C C C A T A T G C G G G C C G C G A G G T A T C G
 C T C G C G C T G G T C A G C C C C G G C A C G T C C G G C C C G G A C C G G C
 G G C C G C G T C C G C T C A C G C C G G A T G A G G C T C G C G A A C C T G C C
 G A A G G C G G C C C G G C T C T G C G C G A T C G C C G A C A A T G C C G T C
 A T G T C G G C C C T G C G C G G T C G T C C C G A G G C T C G C C G G C G G C G 20
 A G G C G G A G A A C G T C C T G C A G C A C G C C G A C T C G G C G G C G G C
 C G G C A C C A C C G C C C T C G C C G G C T G A C C G C C T T G C T G T A C
 G C G G A G A A C A C C G A C A C C G C T C A G C T C T G G G C C G A C A A G C
 T G G T C T C C G A G A C C G G G G C G T C G A A C G A G G A G G A G G G C G G G
 C T A C G C G G G G C C G G C G C G C G A A G C C G C G T T G C G T C G C G G C
 G A C C T G G C C G G C G G G T C G A G G C A G G C A G C A C C G T T C T G G
 A C C A C C G G C G G G C T C T G A C G C T C G G C A T C A C C G C C G G C T
 A C C G C T G A G C A G C G G G T G G C C G C C A T C C G G C T G G G C 30
 G A G A C C G A G C G G G C G G G A G A A G T G G C T C G C C C A G C C G C T G C
 C G C A G G C C A T C C A G G A C G G C C T G T T C G G C C T G C A C C T G C T
 C T C G G C G C G G C C A G T A C A G C C T C G C C A C G G G C C A G C A C
 G A G T C G G C G T A C A C G G C G T T T C G C A C C T G C G G G G A A C G T A
 T G C G G A A C T G G G G C G T T G A C G T G C C G G G T C T G T C C C T G T G
 G C G C G T C G A C G C C G C C G A G G G C G T C G A C G A G C A A C T C A C C C
 C G G G A C G A G G G C C G A C G G G C T C G T C G A C G A G C A A C T C A C C C
 G T G C G A T G G G A C C C C G T T C C C G C G C C T T G A C G C T G C G G G G T
 G C A G G C G G G C G T A C A G C C C G C C G G G C G A A G C G G G G T C G A C C T G
 C T C G A T G A A G C G G G C G A C C T G C T G C T C T C C T G C A A C G A C C
 A G T A C G A G C G G G C A C G G G T G C T C G C C G A C C T G A G C G A G A C
 G T T C A G C G C G T C C G G C A C C A C A G C C G G G C G C G G G G A C T G 40
 C T T C G G G C A G G C C C G G C A C C T G G C C G C C C A G C G C G G C G C G A
 T A C C G C T G C T G C G C C G A C T C G G G G C C A A G C C C G G A G G G C C C
 C G G C T G G C T G G A G G A A T C C G G C C T G C C G C A G C G G A T C A A G
 T C G C T G A C C G A C G C G G A G C G G G C G G G T G G C G T C G C T G G C C G
 C C G G C G G A C A G A C C A A C C G C G T G A T C G C C G A C C A G C T C T T
 C G T C A C G G C C A G C A C G G T G G A G C A G C A C C T C A C G G A C G T C
 T C C A C T G G G T C A A G G C C G C C A G C A C C T G C C G C C G A A C T C G
 T C T A G

配列番号 4 2

A T G C C T G C C G T G G A G T G C T A T G A A C T G G A C G G C C C G C G A T G 50

A C G A G C T C A G A A A A C T G G A G G A G G T T G T G A C C G G G C G G G C
 C A A C G G G C G G G T G T G G T G G T C A C C A T C A C C G G A C C G A T C
 G C C T G C G G C A A G A C C G A A C T G C T C G A C G C A G C C G C C G C G A
 A G G C C G A C G C C A T C A C G C T G C G A G C G G T C T G C T C C G C G G A
 G G A A C A G G C A C T C C C G T A C G C C C T G A T C G G G C A G C T C A T C
 G A C A A C C C G G C G C T C G C C T C C A C G C G C T G G G A G G C C G C T
 G C C C G A C C C T C C C G G G C G A G C A C C T G T C G C C G G A G G C C G A
 G A A C C G G C T G C G C A G C G A C C T C A C C C G T A C C C T G C T G G C G
 C T C G C C G C C G A A C G G C C G G T G C T G A T C G G C A T C G A C G A G T
 C A C A C G C G A A C G C T T T G T G T C T G C T C C A C C T G G C C C G A A G
 G G T C G G C T C G G C C C G G A T C G C C A T G G T C C T C A C C G A G T T G
 C G C C G G C T C A C C C G G C C C A C T C A C A G T T C C A G G C C G A G C
 T G C T C A G C C T G G G G C A C C A C C G C G A G A T C G C G C T G C G C C C
 G C T C A G C C C G A A G C A C A C C G C C G A G C T G G T C C G G C C G G T
 C T C G G T C C C G A C G T C G A C G A G G A C G T G C T C A C G G G G T T G T
 A C C G G G C G A C C G G G C A A C C T G A A C C T C A C C C G C G G A C T
 G A T C A A C G A T G T G C G G G A G G C C T G G G A G A C G G G A G G G A C G
 G G C A T C A G C G C G G G C C G C G C G T A C C G G C T G G C A T A C C T C G
 G T T C C C T C T A C C G C T G C G G C C C G G T C C C G T T G C G G G T C G C
 A C G G G T G G C C G C C G T G C T G G G C C A G A G C G C C A A C A C C A C C
 C T G G T G C G C T G G A T C A G C G G G C T C A A C G C G G A C G G C G G T G G
 G C G A G G C A A C C G A G A T C C C T C A C C G A A G G C G G C C T G C T G C A
 C G A C C T G C G G T T C C C G C A C C C G G C G G C C C G T T C G G T G G T A
 C T C A A C G A C A T G T C C G C C C A G G A A C G A C G C C G C C T G C A C C
 G G T C C G C T C T G G A A G T G C T G G A C G A C G T G C C C G T G G A A G T
 G G T C G C G C A C C A C C A G G T C G G C G C C G G T C T C C T G C A C G G C
 C C G A A G G C C G C C G A G A T A T T C G C C A A G G C C G G C C A G G A G C
 T G C A T G T G C G G C G A G T T G G A C A C C G C G T C C G A C T A T C T
 G C A A C T G G C C C A C C A G G C C T C C G A C G A C G C C G T C A C C G G G
 A T G C G G G C C G A G G C C G T G G C G A T C G A G G C G C C G C C G A A C C
 C G C T G G C C T C G A G G C C G G C A C C T C G A C G A G G C T G A C C G T C G T
 C G C C C G T G C C G G G C T G C T C T T C C C C G A G G C A C A C G G C G C T G
 A T G A T C C G C T G G C T G G G C G T C G G C G G G C G G T C C G G G C G A G G
 C A G C C G G G C T G C T G G C C T C G C A G G G C C C C G T G C G G G T C A C
 C G A C C A G G A C A G G G C C C A T A T G C G G G C C G C C G A G G T A T C G
 C T C G C G C T G G T C A G C C C C G G C A C G T C C G G C C C G G A C C G G C
 G G C C G C G T C C G C T C A C G C C G G A T G A G G C T C G C G A A C C T G C C
 G A A G G C G G C C C G G C T C T G C G G C A T C G C C G A C A A T G C C G T C
 A T G T C G G C C C T G C G C G G T C G T C C C G A G G C T C G C C G G G C C G
 A G G C G G A G A A C G T C C T G C A G C A C G C C G A C T C G G C G G C G G C
 C G G C A C C A C C G C C C T C G C C G G C G T G A C C G C C T T G C T G T A C
 G C G G A G A A C A C C G A C A C C G C T C A G G C T C T G G G C C G A C A A G C
 T G G T C T C C G A G A C C G G G G C G T C G A A C G A G G G A G G A G G G C G G G
 C T A C G C G G G G C C G C G C G C G A A G G C C G G C G T T G C G T C G C G G C
 G A C C T G G C C G C G G G C G G T C G A G G G A G G C A G C A C C G T T C T G G
 A C C A C C G G C G G C T C T G C A G G C T C G G G C A T C A C C G C C G G C T
 A C C G C T G A G C A G C G C G G T G G G C C G C C A T C C G G C T G G G C
 G A G A C C G A G C G G G G C G G A A G T G G C T C G C C C A G C C G C T G C
 C G C A G G C C A T C C A G G A C G G C C T G T T C G G G C C T G C A C C T G C T
 C T C G G C G C G G C C A G T A C A G C C T C G C C A C G G G C C A G C A C

10

20

30

40

50

G A G T C G G C G T A C A C G G C G T T T C G C A C C T G C G G G G A A C G T A
 T G C G G A A C T G G G G C G T T G A C G T G C C G G G T C T G T C C C T G T G
 G C G C G T C G A C G C C G C C G A G G G C G C T G C T G C A C G G C C G C G A C
 C G G G A C G A G G G C C G A C G G C T C G T C G A C G A G C A A C T C A C C C
 G T G C G A T G G G A C C C C G T T C C C G C G C C T T G A C G C T G C G G G T
 G C A G G C G G C G T A C A G C C C G C C G G C G A A G C G G G T C G A C C T G
 C T C G A T G A A G C G G C C G A C C T G C T G C T C T C C T G C A A C G A C C
 A G T A C G A G C G G G C A C G G G T G C T C G C C G A C C T G A G C G A G A C
 G T T C A G C G C G T C C G G C A C C A C A G C C G G G C G C G G G A C T G
 C T T C G G C A G G C C C G G C A C C T G G C C G C C A G C G C G G G C G C G A
 T A C C G C T G C T G C G C C G A C T C G G G G G C C A A G C C C G G A G G C C C
 C G G C T G G C T G G A G G A A T C C G G C C T G C C G C A G C G G A T C A A G
 T C G C T G A C C G A C G C G G A G C G G C G G G T G G C G T C G C T G G C C G
 C C G G C G G A C A G A C C A A C C G C G T G A T C G C C G A C C A G C T C T T
 C G T C A C G G C C A G C A C G G T G G A G C A G C A C C T C A C G G A C G T C
 T C C A C T G G G T C A A G G C C G C A G C A C C T G C C G C C G A A C T C G
 T C T A G

配列番号 4 3

G T G G T T C C T G A A G T G C G A G C A G C C C C C G A C G A A C T G A T C G
 C C C G C G A T G A C G A G C T G A G C C G C C T C C A A C G G G C A C T C A C 20
 C A G G G C G G G G A G C G G A A G G G G C G G C G T C G T C G C C A T C A C C
 G G G C C C A T C G C C A G C G G A A A G A C G G G C G T C G T C G A C G C C G
 G A G C G G C C A A G T C C G G C T T C G T C G C A C T C C G T G C G G T G T G
 C T C C T G G G A A G A G C G G C A C T C T G C C G T A C G G G A T G C T G G G C
 C A G C T C T T C G A C C A T C C C G A A C T G G C C G C C C A G G G C G C C G G
 A C C T T G C C C A C T T C A C G G C T T C G T G C G A G A G C C C T C A G G C
 C G G T A C C G A C A A C C G C C T G C G G G C C G A G T T C A C C C G C A C C
 C T G C T G G C G C T C G C C G C G G A C T G G C C C G T C C T G A T C G G C A
 T C G A C G A C G T G C A C C A C G C C G A C G C G G A A T C A C T G C G C T G
 T C T G C T C C A C C T C G C C C G C C G A T C G G G C C C G G C C G A T C 30
 G C G G T C G T A C T G A C C G A G C T G C G C A G A C C G A C G C C C G C C G
 A C T C C C G C T T C C A G G C G G A A C T G C T G A G C C T G C G C T C C T A
 C C A G G A G A T C G C G C T C A G A C C G C T C A C C G A G G G C G C A G A C C
 G G C G A A C T C G T A C G T C G G C A C C T C G G C G G A G A C C C A C G
 A G G A C G T C T C C G C C G A T A C G T T C C G G G C G A C C G G G C G G G A A
 C C T G C T C C T C G G G C A C G G T T T G A T C A A T G A C A T C C G G G A G
 G C G C G G A C A G C G G G A C G G G C C G G G G T C G T C G C G G G G C G G G
 C G T A C C G G C T C G C G T A C C T C A G C T C G C T C T A C C G C T G C G G
 C C C G A G C G C G C T G C G T G T C G C C C G G G C G T C C G C C G T G C T C
 G G C G C G A G C G C C G A A G C C G T G C T C G T C A G C G G A T G A C C G 40
 G A C T G A A C A A G G A C G C G G T C G A A C A G G T C T A T G A G C A G C T
 G A A C G A G G G A C G G G C T G C T G C A G G G C G A G C G G G T T T C C G C A C
 C C G G C G G C C C G C T C C A T C G T C C T T G A C G A C C T G T C G G C C C
 T G G A A C G C A G A A A C C T G C A C G A G T C G G G C G C T G G A G G C T G C T
 G C G G G A C C A C G G C G T G G C C G G C A A C G T G C T C G C C C G C C A C
 C A G A T C G G C G C C G G C C G G G T G C A C G G C G A G G G A G G G C C G T C G
 A G C T G T T C A C C G G G G C C G C A C G G G A G C A C C A C C T G C G C G G
 T G A A C T G G A C G A C G C G G C G A T A C C T G G A A C T C G C C C A C
 C G T G C C T C C G A C G A C C C C G T C A C G C G C G C C G C A C T A C G C G
 T C G G C G C C G C C G A T C G A G C G C C T C T G C A A T C C G G T A C G 50

G G C A G G C C G G C A T C T G C C C G A G C T G C T C A C C G C G T C G C G C
 G C G G G A C T G C T C T C C A G C G A G C A C G C C G T G T C G C T C G C C G
 A C T G G C T G G C G A T G G G C G G G C G C C C G G G C G A G G C G G C C G A
 G G T C C T C G C G A C G C A G C G T C C C G C G G C C G A C A G C G A G C A G
 C A C C G C G C A C T C C T G C G C A G C G G C G A G T T G T C C C T C G C G C
 T G G T C C A C C C C G G C G C G T G G G A T C C G T T G C G C C G G A C C G A
 T C G G T T C G C C G C G G G C G G G C T C G G C T C G C T T C C C G G A C C C
 G C C C G G C A C C G C G C G G T C G C C G A C C A A G C C G T C A T C G C G G
 C G C T G C G T G G A C G T C T C G A C C G G G C G G A C G C C A A C G C G G A
 G A G C G T T C T C C A G C A C A C C G A C G C C A C G G C G G A C C G G A C C
 10 A C G G C C A T C A T G G G C G T T G C T G G C C C T G C T C T A C G C G G A G A
 A C A C C G A T G C T G T C C A G T T C T G G G T C G A C A A A C T G G C C G G
 T G A C G A G G G C A C C A G G A C A C C G G C C G A C G A G G G C G G T C C A C
 G C G G G G T T C A A C G C C G A G A T C G C G C T G C G C C G C G G C G A C T
 T G A T G A G A G G C G T C G A G T A C G G C G A G G C A G C G G C T C G G C C A
 C C G G C A C C T G C C C A C C T G G G G A A T G G C C G C C G T C T G C C G
 C T G A G C A G C A C C G T G G T T G C C G C G A T C C G G C T C G G C G A C C
 T C G A C A G G G C C G A G C G G T G G C T C G C C G A G C C G C T G C C G C A
 G C A G A C G C C G G A G A G C C T C T T C G G G C T G C A C C T G C T C T G G
 G C C C G C G G G C A G C A C C A C C T C G C G A C C G G G C G G C A C G G G G
 20 C G G C G T A C A C G G C G T T C A G G G A A T G C G G C G A G C G G A T G C G
 G C G G T G G G C C G T C G A C G T G C C G G G C C T G G C C C T G T G G C G G
 G T C G A C G C C G C G A A T C G C T G C T G C T G C T C G G C C G T G A C C
 G T G C C G A A G G A C T G C G G C T C G T C T C C G A G C A G C T G T C C C G
 G C C G A T G C G C C C T C G C G C G C G T G C A G A C G T T A C G G G T A
 C A G G C G G C C T A C A G T C C G C C G C C C A A C G G A T C G A C C T G C
 T C G A A G A G G C C G C G A C C T G C T G G T C A C C T G C A A C G A C C A
 G T A C G A A C T G G C A A A C G T A C T C A G C G A C T T G G C A G A G G C C
 T C C A G C A T G G T C C G G C A G C A C A G C A G G G C G C G G G T C T G C
 30 T C C G C C G G G C A C G G C A C C T C G C C A C C C A G T G C G G G C G C C G T
 G C C G C T C C T G C G G C G G C T C G G C G C G G A A C C C T C G G A C A T C
 G G C G G A G C C T G G G A C G C G A C G G C T G G G A C A G C G G A T C G C G T
 C A C T G A C G G A G T C G G A G C G G C G G G T G G C C G G C G T C G C C G C
 G G T C G G G C G T A C G A A C A G G G A G A T C G C C G A G C A G C T G T T C
 G T C A C G G C C A G C A C G G T G G A A C A G C A C C T C A C G A A C G T G T
 T C C G C A A A C T G G C G G T G A A G G G C C G C A G C A G C T T C C G A A
 G G A A C T G G C C G A C G T C G G C G A G C C G G G A C C G C G A C C G C
 C G G T G C G G G T A G

配列番号 4 4

 A T G G T T C C T G A A G T G C G A G C A G C C C C C G A C G A A C T G A T C G
 C C C G C G A T G A C G A G C T G A G C C G C C T C C A A C G G G C A C T C A C
 C A G G G C G G G G A G C G G A A G G G G C G G C G T C G T C G C C A T C A C C
 G G G C C C A T C G C C A G C G G A A A G A C G G G C G G C G T C G C T C G A C G C C G
 G A G C G G C C A A G T C C G G C T T C G T C G C A C T C C G T G C G G G T G T G
 40 C T C C T G G G A A G A G G C G A C T C T G C C G T A C G G G A T G C T G G G C
 C A G C T C T T C G A C C A T C C C G A A C T G G C C G C C C A G G G C G C C G G
 A C C T T G C C C A C T T C A C G G C T T C G T G C G A G A G G C C C T C A G G C
 C G G T A C C G A C A A C C G C C T G C G G G C C G A G T T C A C C C G C A C C
 C T G C T G G C G C T C G C C G C G G A C T G G C C C G T C C T G A T C G G C A
 T C G A C G A C G T G C A C C A C G C C G A C G C G G A A T C A C T G C G C T G
 50

T C T G C T C C A C C T C G C C C G C A T C G G G C C C G G C C C G C A T C
 G C G G T C G T A C T G A C C G A G G C T G C G C A G A C C G A C G C C C G C C G
 A C T C C C G C T T C C A G G C G G A A C T G C T G A G C C T G C G C T C C T A
 C C A G G A G A T C G C G C T C A G A C C G C T C A C C G A G G G C G C A G A C C
 G G C G A A C T C G T A C G T C G G C A C C T C G G C G C G G A G A C C C A C G
 A G G A C G T C T C C G C C G A T A C G T T C C G G G C G A C C G G C G G G A A
 C C T G C T C C T C G G G C A C G G T T T G A T C A A T G A C A T C C G G G A G
 G C G C G G A C A G C G G G A C G G G C C G G G G T C G T C G C G G G G C G G G
 C G T A C C G G C T C G C G T A C C T C A G C T C G C T C T A C C G C T G C G G
 C C C G A G C G C G T C G C G T G T C G C C C G G G C G T C C G C C G T G C T C
 G G C G C G A G C G C C G A A G C C G T G C T C G T C C A G C G G A T G A C C G
 G A C T G A A C A A G G A C G C G G T C G A A C A G G T C T A T G A G C A G C T
 G A A C G A G G G A C G G G T G C T G C A G G G C G A G C G G T T T C C G C A C
 C C G G C G G C C C G C T C C A T C G T C C T T G A C G A C C T G T C G G C C C
 T G G A A C G C A G A A A C C T G C A C G A G T C G G C G C T G G A G C T G C T
 G C G G G A C C A C G G C G T G G C C G G C A A C G T G C T C G C C C G C C A C
 C A G A T C G G C G C C G G C C G G G T G C A C G G C G A G G G A G G G C G T C G
 A G C T G T T C A C C G G G G C C G C A C G G G A G C A C C A C C T G C G C G G
 T G A A C T G G A C G A C G C G C C G G A T A C C T G G A A C T C G C C C A C
 C G T G C C T C C G A C G A C C C C G T C A C G C G C G C G C A C T A C G C G
 T C G G C G C C G C C G C G A T C G A G G G C C T C T G C A A T C C G G T A C G
 G G C A G G G C G G C A T C T G C C C G A G G C T G C T C A C C G C G T C G C G C
 G C G G G A C T G C T C T C C A G C G A G G C A C G C C G C T G T C G C T C G C C G
 A C T G G C T G G C G A T G G G C G G G C G C C C G G G C G A G G G C G G G C C G A
 G G T C C T C G C G A C G C A G C G T C C C G G G C G A C A G C G A G C A G
 C A C C G C G C A C T C C T G C G C A G C G G G C G A G T T G T C C C T C G C G C
 T G G T C C A C C C C G G G C G G T G G G A T C C G T T G C G C C G G A C C G A
 T C G G T T C G C C G G G C G G G C T C G G C T C G C T T C C C G G A C C C
 G C C C G G G C A C C G C G G G T C G C C G A C C A A G C C G T C A T C G C G G
 C G C T G C G T G G A C G T C T C G A C C C G G G C G G A C G C C A A C G C G G A
 G A G C G T T C T C C A G C A C A C C G A C G C C A C G G C G G A C C C G G A C C
 A C G G C C A T C A T G G G C G T T G C T G G C C C T G C T C T A C G C G G A G A
 A C A C C G A T G C T G T C C A G T T C T G G G T C G A C A A A C T G G C C G G
 T G A C G A G G G G A C C A G G A C A C C G G G C G A C G G G C G G T C C A C
 G C G G G G T T C A A C G C C G A G A T C G C G C T G C G C C G C G G G C G A C T
 T G A T G A G A G G C C G T C G A G T A C G G C G A G G C A G C G G C T C G G C C A
 C C G G C A C C T G C C C A C C T G G G G A A T G G C C G C C G T C T G C C G
 C T G A G C A G C A C C G T G G T T G C C G C G A T C C G G C T C G G C G A C C
 T C G A C A G G G C C G A G C G G G T G G C T C G C C G A G C C G C T G C C G C A
 G C A G A C G C C G G A G A G C C T C T T C G G G C T G C A C C T G C T C T G G
 G C C C G C G G G C A G C A C C A C C T C G C G A C C G G G C G G C A C G G G G
 C G G C G T A C A C G G C G T T C A G G G A A T G C G G C G A G C G G A T G C G
 G C G G T G G G C C G T C G A C G T G C C G G G C C T G G C C C T G T G G C G G
 G T C G A C G C C G C G A A T C G C T G C T G C T G C T C G G C C G T G A C C
 G T G C C G A A G G A C T G C G G C T C G T C T C C G A G C A G C G T G T C C C G
 G C C G A T G C G C C C T C G C G C G C G T G C A G A C G C T G C G G G T A
 C A G G C G G G C C T A C A G T C C G C C G C C C A A C G G A T C G A C C T G C
 T C G A A G A G G C C G C C G A C C T G C T G G T C A C C T G C A A C G A C C A
 G T A C G A A C T G G C A A A C G T A C T C A G C G A C T T G G C A G A G G G C C
 T C C A G C A T G G T C C G G C A G C A C A G C A G G G G C G C G G G G T C T G C 50

T C C G C C G G G C A C G G C A C C T C G C C A C C C A G T G C G G C G C C G T
 G C C G C T C C T G C G G G C G T C G G G C G C G G A A C C C T C G G A C A T C
 G G C G G A G C C T G G G A C G C G A C G C T G G G A C A G C G G A T C G C G T
 C A C T G A C G G A G T C G G A G C G G G G T G G G C G C G C T C G C C G
 G G T C G G G C G T A C G A A C A G G G A G A T C G C C G A G C A G C T G T T C
 G T C A C G G C C A G C A C G G T G G A A C A G C A C C T C A C G A A C G T G T
 T C C G C A A A C T G G C G G T G A A G G G C C G C A G C A G C T T C C G A A
 G G A A C T G G C C G A C G T C G G C G A G C C G G G A C C G C G A C C G C
 C G G T G C G G G T A G

配列番号 45

10

G T G A T A G C G C G C T T A T C T C C C C C A G A C C T G A T C G C C C G C G
 A T G A C G A G T T C G G T T C C C T C C A C C G G G C G T C A C C C G A G C
 G G G G G G C G G G C G G G G C G T C G T C G C C G C G T C A C C G G G C G
 A T C G C C T G C G G C A A G A C C G A A C T C C T C G A C G C C G C C G C G
 C C A A G G C C G G C T T C G T C A C C C T T C G C G C G G T G T G C T C C A T
 G G A G G A G C G G G C C C T G C C G T A C G G C A T G C T C G G C C A G C T C
 C T C G A C C A G C C C G A G C T G G C C G C C C G G A C A C C G G A G C T G G
 T C C G G C T G A C G G C A T C G T G C G A A A A C C T G C C G G C C G A C G T
 C G A C A A C C G C C T G G G G A C C G A A C T C A C C C G C A C G G T G C T G
 A C G C T C G C C G C G G A G C G G C C G T A C T G A T C G G C A T C G A C G

20

A C G T G C A C C A C G C C G A C G C G C C G T C G C T G C G C T G C C T G C T
 C C A C C T C G C G C G C G C A T C A G C C G G G C C C G T G T C G C C A T C
 G T G C T G A C C G A G C T G C T C C G G C C G A C G C C C G C C C A C T C C C
 A A T T C C G G G C G G C A C T G C T G A G T C T G C G C C A C T A C C A G G A
 G A T C G C G C T G C G C C C G C T C A C C G A G G G C G A G A C C A C C G A A
 C T C G T G C G C C G G C A C C T C G G C C A G G A C G C G C A C G A C G A C G
 T G G T G G G C C A G G C G T T C C G G G C G A C C G G C G G C A A C C T G C T
 C C T C G G G C C A C G G C C T G A T C G A C G A C A T C C G G G A G G G C A C G G
 A C A C G G A C C T C A G G G T G C C T G G A A G T G G T C G C G G G G G G G G G

30

C G T A C C G G C T C G C C T A C C T C G G G T C G C T C T A T C G T T G C G G
 C C C G G C C G C G C T G A G C G T C G C C C G A G C T T C C G C C G T G C T C
 G G C G A G A G T G T C G A A C T C A C C C T C G T C C A G C G G A T G A C C G
 G C C T C G A C A C C G A G G G C G G T C G A G C A G G C C C A C G A A C A G C T
 G G T C G A G G G G C G G C T G C T G C G G G A A G G G C G G T T C C C G C A C
 C C C G C G G C C C G C T C C G T C G T A C T C G A C G A C C T C T C C G C C G
 C C G A G C G G C G T G G C C T G C A C G A G C T G G C G C T G G A A C T G C T
 G C G G G A C C G C G G C G T G G C C A G C A A G G T G C T C G C C C G C C A C
 C A G A T G G G T A C C G G C C G G G T G C A C G G C G C C G A G G T C G C C G
 G G C T G T T C A C C G A C G C C G C G C G A G C A C C A C C T G C G C G G

40

C G A G C T C G A C G A G G C C G T C A C C T A C C T G G A G T T C G C C T A C
 C G G G C C T C C G A C G A C C C C G C C G T C C A C G C C G C A C T G C G C G
 T C G A C A C C G C C G C C A T C G A G C G G C T C T G C G A T C C C G C C A G
 A T C C G G G C G C A T G T G C C C G A G C T G C T C A C C G C G T C G C G G
 G A A C G G C T C C T C T C C A G C G A G C A C G C C G T G T C G C T C G C C T
 G C T G G C T G G C G A T G G A C G G G C G G C C G G G C G A G G C C G C C G A
 G G T C C T G G C G G C C A G C G G C T C C G C C G C C C G A G C G A G C A G
 G G C C G G G C G C A C C T G C G C G T C G C G G A C C T G T C C C T C G C G C
 T G A T C T A T C C C G G C G C G G C C G A T C C C G C C G T C C C G G C C G A
 T C C G C C G G C C G A G G A C G G A G G T C G C C T C G T T T T C C G G A G C C
 G T C C G G C A C C G C G C C G T C G C C G A C A A G G C C C T G A G C A A C G

50

CGCTGCGCGCTGGTCCGAACAGGCCGAGGCCAAAGCCGA
 GTACGTGCTCCAGCAGTCCCAGGTCACGACGGACCGGACC
 ACGACCATGATGGCGTTGCTGGCCCTGCTCTACGCCGAGG
 ACACCGATGCCGTCCAGTCCCTGGGTCGACAAAGCTGGCCGG
 TGACGACAACATGCGGACCCCGGCCGACGAGGGCGGTCCAC
 GCGGGGTTCCCGCGCCGAGGCCGCGCTGCGCCGCCGCGACCC
 TGACCGCCGCCGTCGAATGCGGCCGAGGCCGCGCTCGCCCC
 CCGGGTCGTGCCCTCCCTGGGGGATGGCCGCCGCATTGCCG
 CTGAGCAGCACCGTGGCCGCCGCGATCCGACTTGGCGACCC
 TGGACCGGGCGGAGCGGGTGGCTCGCCGAGCCGTTGCCGGA 10
 GGAGACCTCCGACAGCCTCTTCGGACTGCACATGGTCTGG
 GCCCGTGGGCAACACCATCTCGCGGCCGGCGGTACCGGG
 CGCGTACAACCGTTCGGGACTTGCAGGGAGCGGATGCG
 ACGCTGGTCGTCGACGTGCGCGGCCCTGGCCCTGTGGCG
 GTCGACGCCGCCGAAGCGCTTCTGCTGCTCGGCCGCC
 GTGACGAGGGGCTGAGGCTCATCTCCGAGCAGCTGTCCC
 GCCGATGGGTTCCCGGGCGCGCGGGTGATGACGCTGCGGGTG
 CAGGCGGCCTACAGTCCGCCGCCAAGCGGATCGAACTGC
 TCGACGAGGCCGCGATCTGCTCATCATGTGCCGCCGACCA
 GTACGAGCTGGCCCGCGTCCCTCGCCGACATGGGCGAAGCG
 TCGGGCATGCTCCGGGGCACAGCCGTTGCGCGGGGACTGT 20
 TCCGCCGCGCACGGCACCTCGCGACCCAGTGCAGGGAGC
 GCGCTCCCTCCGGGCTCGGTGGGGAGTCCTCGGACGCG
 GACGGCACCCAGGACGTGACGCCGGCGAGCGGGATCACAT
 CGCTGACCGAGGCCGGAGCGGGGGTGCGTGCACGCC
 GGTCTGGGCGCACCAACAAGGAGATCGCCAGCCAGCTGTT
 GTCACCTCCAGCACGGTGGAACAGCACCTCACCAACGTGT
 TCGCAAGCTGGGGTGAAGGGCCGTCAGCAAACGTGAA
 GGAACGTCTCGACGCCGGCTGA

配列番号 4 6

ATGATAGCGCGCTGTCCTCCCCCAGACCTGATCGCCCGCG
 ATGACGAGTTCGGTTCCCTCCACCGGGCGCTCACCCGAGC
 GGGGGGGCGGGCGGGGGCGTCGTCGCCGCCGTCACCGGGCG
 ATCGCCCTGCGGCAAGACCGAACCTCTCGACGCCGCC
 CCAAGGGCGGCTTCGTCACCCCTTCGCGCGGGTGTGCTCCAT
 GGAGGAGCGGGCCCTGCCGTACGGCATGCTCGGCCAGCTC
 CTCGACCAAGCCCCGAGCTGGCCGCCGGACACCGGGAGCTGG
 TCGGGCTGACGGCATCGTGCAGAACCTGCCGGCGACGT
 CGACAACCGCCCTGGGGACCGAACCTCACCCGCACGGTGCTG
 ACGCTCGCCGCCGGAGCGGCCGTACTGATCGGCATCGACG 30
 ACGTGCACCAACGCCGACGCGGCCGTCGCTGCGCTGCC
 CCACCTCGCGGCCGATCACGCCGGGCCGTCGCCATC
 GTGCTGACCGAGCTGCTCCGGCCGACGCCGCCACTCCC
 AATTCCGGGGCGGCACTGCTGAGTCTGCGCCACTACCAAGGA
 GATCGCGCTGCGCCCGCTCACCGAGGGCGAGACCAACCGAA
 CTCGTGCGCCGGCACCTCGGCCAGGGACGGACGCCACGACG
 TGGTGGCCCAGGGCGTTCCGGGCGACCGGGCGGCAACCTGCT 40
 CCTCGGGCACGGCCTGATCGACGACATCCGGGAGGGCACGG
 ACACGGACCTCAGGGTGCCTGGAAAGTGGTGCAGGGGGCGGG
 CGTACCGGGCTCGCCCTACCTCGGGTCGCTCTATCGTTGCCG
 50

C C C G G C C G C G C T G A G C G T C G C C C G A G C T T C C G C C G T G C T C
 G G C G A G A G T G T C G A A C T C A C C C T C G T C C A G C G G A T G A C C G
 G C C T C G A C A C C G A G G C G G T C G A G C A G G C C A C G A A C A G C T
 G G T C G A G G G G C G G G C T G C T G C G G G A A G G G C G G T T C C C G C A C
 C C C G C G G C C C G C T C C G T C G T A C T C G A C G A C C T C T C C G C C G
 C C G A G C G G C G T G G C C T G C A C G A G C T G G C G C T G G A A C T G C T
 G C G G G A C C G C G G C G T G G C C A G C A A G G T G C T C G C C C G C C A C
 C A G A T G G G T A C C G G C C G G G T G C A C G G C G C C G A G G T C G C C G
 G G C T G T T C A C C G A C G C C G C G C G A G C A C C A C C T G C G C G G
 C G A G C T C G A C G A G G C C G T C A C C T A C C T G G A G T T C G C C T A C 10
 C G G G C C T C C G A C G A C C C C G C C G T C C A C G C C G C A C T G C G C G
 T C G A C A C C G C C G C C A T C G A G C G G G C T C T G C G A T C C C G C C A G
 A T C C G G C C G G C A T G T G C C C G A G C T G C T C A C C G C G T C G C G G
 G A A C G G C T C C T C C A G C G A G C A C G C C G T G T C G C T C G C C T
 G C T G G C T G G C G A T G G A C G G G C G G C C G G G C G A G G C C G C C G A
 G G T C C T G G C G G C C A G C G C T C C G C C C C G A G C G A G C A G
 G G C C G G G C G C A C C T G C G C G T C G C G G A C C T G T C C C T C G C G C
 T G A T C T A T C C C G G C G C G G C C G A T C C G C C G T C G T T T C C G G A G C C
 T C C G C C G G C C G A G G A C G G A G G T C G C C T C G T T T C C G G A G C C
 G T C C G G G C A C C G C G C C G T C G C C G A C A A G G C C C T G A G C A A C G 20
 C G C T G C G C G G C T G G T C C G A A C A G G C C G A G G C C A A A G C C G A
 G T A C G T G C T C C A G C A C T C C C G G G T C A C G A C G G A C C G G A C C
 A C G A C C A T G A T G G C G T T G C T G G C C C T G C T C T A C G C C G A G G
 A C A C C G A T G C C G T C C A G T C C T G G G T C G A C A A G C T G G C C G G
 T G A C G A C A A C A T G C G G A C C C C G G C C G A C G A G G G C G G T C C A C
 G C G G G G T T C C G C G C C G A G G C C G G C G T C G C G C C G G C G A C C
 T G A C C G C C G C C G T C G A A T G C G G C G A G G C C G G C G T C G C C C C
 C C G G G T C G T G C C C T C C T G G G G G A T G G C C G C C G C A T T G C C G
 C T G A G C A G C A C C G T G G C C G C C G G C G A T C C G A C T T G G G C G A C C
 T G G A C C G G G C G G A G C G G T G G C T C G C C G A G C C G T T G C C G G A 30
 G G A G A C C T C C G A C A G C C T C T T C G G A C T G C A C A T G G T C T G G
 G C C C G T G G G C A A C A C C A T C T C G C G G C C G G G C G G T A C C G G G
 C G G C G T A C A A C A C G C G T T C C G G G A C T T G C G G G G A G C G G A T G C G
 A C G C T G G T C C G T C G A C G T G C C G G G C C T G G C C C T G T G G C G G G
 G T C G A C G C C G C G A A G C G G C T T C T G C T G C T C G G C C G C G G C C
 G T G A C G A G G G G C T G A G G G C T C A T C T C C G A G C A G C T G T C C C G
 G C C G A T G G G G T C C C G G G C G C G G G T G A T G A C G C T G C G G G T G
 C A G G C G G G C C T A C A G T C C G C C G G C C A A G C G G A T C G A A C T G C
 T C G A C G A G G C C G C G A T C T G C T C A T C A T G T G C C G C G A C C A
 G T A C G A G C T G G C C C G C G T C C T C G C C G A C A T G G G C G A A G C G 40
 T G C G G C A T G C T C C G G C G G C A C A G G C C G T G C G C G G G G A C T G T
 T C C G C C G C G C A C G G C A C C T C G C G A C C C A G T G C G G A G C C G T
 G C C G C T C C T C C G G G C G T C G G T G G G G A G T C C T C G G A C G C G
 G A C G G C A C C C A G G A C G T G A C G C C G G C G A G C G G A T C A C A T
 C G C T G A C C G A G G C G G A G C G G C G G G T G G C G T C G C A C G C C G C
 G G T C G G G C G C A C C A A C A A G G A G A T C G C C A G C C A G C T G T T C
 G T C A C C T C C A G C A C G G T G G A A C A G C A C C T C A C C A A C G T G T
 T C C G C A A G C T G G G G G T G A A G G G C C G T C A G C A A C T G C C C A A
 G G A A C T G T C C G A C G C C G G C T G A

GTGGAGTTTACGACCTGGTCGCCCGCGATGACGAGCTCA
 GAAGGTTGGACCAAGGCCCTCGGCCGCGCCGCCGGCGGACG
 GGGTGTGTCGTGGTCACCGTCACCGGACCGGTCGGCTGCGGC
 AAGACCGAACCTGCTGGACGCGGCCGCGGCCGAGGAGGAAT
 TCATCACGTTGCGTGCGGTCTGCTCGGCCGAGGAGGAGC
 CCTGCCGTACGCCGTGATCGGCCAACCTCCCTGACCATCCC
 GTACTCTCCGCACGCGGCCGACCTGGCCTGCGTGA
 CTCCGGGCCGGACGCTGCCGGCCGACACCGAGAAACCGCC
 CGGCCGACCTCACCCGGGCCCTGCTGGCCCTGGCCTCC
 GAACGACCGGTTCTGATCTGCATCGACGACGTGCA
 10 CCGACACCGCCCTCGCTGAACCTGCCCTGCTGCACCTGGCCCG
 GCGGGTCGCCCTCGGCCGGATCGCCATGATCCTCACCGAG
 TTGCGCCGGCTCACCCGGCTCACCTCCGGTTCGAGGCGG
 AACTGCTCAGCCTGCACCGCCACGAGATCGCGCTGCG
 TCCCCCTCGGCCCGGCCGACACCGCCGAACCTGGCCTGCC
 CGGCTCGGCCGCCGGCTCACCGCCGACGAGCTGGCCCAAGG
 TCCACGAGGCCACCAAGCGGGAACCCCCAACCTGGTCGGAGG
 CCTGGTCAACGACGTGCGAGAGGCCCTGGCGGCCGGTGGC
 ACGGGCATTGCGGGGGCGGGCTACCGGGCTGGCGTACCG
 20 TCAGCTCCGTGTACCGCTGTGGTCCGGTCCCCGTTGCGGAT
 CGCCCAAGGCGGGCGGTGCTGGGTCCCCAGCGCCACCGTC
 ACGCTGGTGCGCCGGATCAGCGGGCTCGACGCCGAGACGG
 TGGACGAGGCAGCCCGGATCCTCACCGAGGGCGGGCTGCT
 CCGGGACCAACCGGTTCCCGCATCCGGCGGCCCGCTGGTC
 GTACTCGACGACATGTCCCGCGCAGGAACGCCGCCCTG
 ACCGGTCCACGCTGGACGCTGGACGGCGTACCCGTCGA
 CGTCTCGCGCACCAAGGCCGGCGCCGGTCTGCTGCAC
 GGCCCAGGGCGAGATGTTGCCCGGGGCCAGCCAGGG
 AGCTGCGGGTACGCGGGCGAGCTGGACGCCGACCGAGTA
 CCTGCAACTGGCCCTACCGGGCTCCGACGACGCCGGCGCC
 30 CGGGCCGCCCTGCAAGGTGGAGACCGGTGGCCGGCGAGCGCC
 GCGCAACCGCTGGCGCCAGCCGGCACCTGGACGAGCT
 GGCGCCGCCGCCGGGCCGGCTGCTGTCGGCCGAGCAC
 GCGCCCTGGTCGTGCACTGGCTGGCCGACGCCGGACGAC
 CGGGCGAGGCCGCCGGAGGTGCTGGCGCTGCA
 40 GGCGCTCACCGACCAAGGCCGGCCCTGCGGGCGGCC
 GAGGTGTCGCTCGCGCTGTTCCACCCGGCTCCCGGTT
 CGGACCCGCGGGCCCTCGCGCCGGAGGAGCTCGCGAGCCT
 GTCCCTGTCGGCCCGACGGTGTGACCGCCGACAACGCG
 GTGCTGGCGCGCTGCGCGGGCTCCGAGTCTGGCCGCC
 CGAGGCGGGAGAACGCTGCTGCGCAACGCCGACGCCGCC
 GTCCGGCCCGACCGCCCTGGCCGCGCTGACGGCCCTGCTC
 TACGCCGAGAACACCGACGCCGCCAGCTCTGGCGGGACA
 AGCTGGCCGCGGGCATCGGGGCGGGGAGGGGGAGGCCGG
 CTACGCGGGGCCGCCGGACCGTGGCCGCCCTGCGTCGCGGC
 GACCTGACCAACCGCGGTCCAGGGCGCCGGCGCGGTCTGG
 ACCGGCGGGCCGCCGTCGCTCGGCATCACCGCCGTGTT
 GCGTGTGAGCGGGCGCGGTCTGCCGCCGCGATCCCGGCTGG
 GAGCTCGAGCGGGCCGAGAAGTGGCTGGCCGAGCCGCTG
 50 CGAAGGCCGTCCACGACGCCCTGTTCTGGCCCTGCA
 CCGACGCCCTGCTGCAACCTGCT

GATGGCGCGGGGCCGCTACAGCCTCGCGGTGGGCCGGCAC
 GAGGCGGCGTACGCCGCGTTCGGGACTGCGGTGAACGGA
 TGCGCCGGTGGGACGCTCGACGTGCCCGGGCTGGCCCTGTG
 GC GG GT TGGACGCGGCCGAGGGCGCTGCTGCCCGGCGATGAC
 CGGGCGGAGGGCCGGCGCTGATCGACGAGCAGCTCACCC
 GCCGATGGGGCCCGGTCACGAGGCCCTGACCCCTGCGGGT
 ACGAGCGGCCCTACGCCCGCCGGCGAAACGGATCGACCTG
 CTCGACGAAGCGGGCGACCTGCTGCTCTCCAG
 CAACGACCAAGTACGAGCGGGCACGGGTGCTGGCCGACCTG
 AGCGAGGGCGTTCAAGCGCGCTCCGGCAGAACGGCCGGCGC 10
 CGGGCATCCTGCGGCAGGGCCCGGCACCTGGCCGCCAGTG
 CGGGCGGTCCCCCTGCTGCGCCGGCTGGGCGTCAAGGCC
 GGCGGTCCGGTCGGCTCGGCCGGCGCAGGGAAATCC
 GCTCCCTGACCGAGGCCGAGCGCCGGTGGCACGCTGGC
 CGCCGCCGGGCAGACCAACCGGGAGATCGCCGACCAAGCTC
 TT CGTCACCGCCAGCACGGTCAGCAGCACCTCACCAACG
 TGTTC CGCAAGCTCGGCGTGAAGGGCCGCCAGCAATTGCC
 GCCGAGCTGGCCGACCTGCGGCCGCCGGCTGA

配列番号 48

ATGGAGTTTACGACCTGGTCGCCCGCGATGACGAGCTCA 20
 GAAGGTTGGACCAAGGCCCTCGGCCGCCGGGAGCG
 GG GT TCGTGGTCACCGTCAACCGGACCGGGTCTGGCTGCGGC
 AAGACCGAACTGCTGGACGCGGCCGCGGGCAGGGAGGAAT
 TCATCACGTTGCGTGC GG TCTGCTCGGCCGAGGAGCGGGC
 CCTGCCGTACGCCGTGATCGGCCAACCTCCCTGACCATCCC
 GTACTCTCCGCACGCGGCCGACCTGGCCCTGCGTGA CGG
 CTCCGGGCCGGACGCTGCCGGCCGACACCGAGAACCGCCCT
 CGGCCGCGACCTCACCCGGGCCCTGCTGGCCCTGGCCTCC
 GAACGACCGGTTCTGATCTGCATCGACGACGTGCA ACCAGG
 CCGACACCGCCCTCGCTGAACCTGCCCTGCTGCA CCTGGCCCG 30
 CGGGGT CGCCTCGGCCCGGATCGCCATGATCCCTCACCGAG
 TT GCGCCGGCTCACCCCGGCTCACCTCCGGTTTCGAGGGCGG
 AACTGCTCAGCCTGCGGCACCGCCACGAGATCGCGCTGCG
 TCCCCCTCGGCCCCGGCCGACACCGCCGAACTGGCCCGGCC
 CGGCTCGGCCGGCGTCAACCGCCGACGAGCTGGCCCAAGG
 TCCACGAGGCCACCAAGCGGGAACCCCCAACCTGGTCGGAGG
 CCTGGTCAACGACGTGCGAGAGGGCTGGGCGGGGGTGGC
 ACGGGCATTGCGGGCGGGCGGGCGTACCGGGCTGGCGTACCC
 TCAGCTCCGTGTACCGCTGTGGTCCGGTCCCCGTTGCGGGAT
 CGCCCAAGGGGGCGGTGCTGGGTCCCAAGCGCCACCGTC 40
 AC GCTGGTGCGCCGGATCAGCGGGCTCGACGCCGAGACGG
 TGGACGAGGCCACCGCGATCCCTCACCGAGGGCGGCCCTGCT
 CGGGGACCAACCGGTTCCCGCATCCGGCGGCCGCTCGGT
 GTACTCGACGACATGTCCTCGCGCAGGAACCGCCGCCCTG
 ACCGGTCCACGCTGGACGCTGGTGGACCGCGTACCCGTCGA
 CGTGCCTCGCGCACCAACCGAGGCCGGCGCCGGTCTGCTGCAC
 GGCCCGCAGGGCGGGCGAGATGTTCGCCCGGGCCAGCCAGG
 AGCTGCGGGTACGCGGGCGAGCTGGACGCGCCGCGACCGAGTA
 CCTGCAACTGGCCCTACCGGGCCTCCGACGACGCCGCCGGCGCC
 CGGGCCGCCCTGCAAGGTGGAGACCGTGGGCCGGCGAGCGCC 50

G C C G C A A C C C G C T G G C C G C C A G C C G G C A C C T G G A C G A G C T
 G G C C G C C G C C G G G G C C G G C C T G C T G T C G G G C C G A G C A C
 G C C G C C C T G G T C G T G C A C T G G C T G G C C G A C G C C G G A C G A C
 C C G G C G A G G C C G C C G A G G T G C T G G C G C T G C A G C G G G C G C T
 G G C C G T C A C C G A C C A C G A C C G G G C C C G C C T G C G G G C G G C C
 G A G G T G T C G C T C G C G C T G T T C C A C C C C G G C G T C C C C G G T T
 C G G A C C C G C G G C C C T C G C G C C G G A G G A G C T C G C G A G C C T
 G T C C C T G T C G G C C C G G C A C G G T G T G A C C G C C G A C A A C G C G
 G T G C T G G C G G C G C T G C G C G G C G T C C C G A G T C G G C C G C C G
 C C G A G G C G G A G A A C G T G C T G C G C A A C G C C G A C G C C G C C G C
 G T C C G G G C C G A C C G C C C T G G G C C G C G C T G A C G G G C C C T G C T C
 T A C G C C G A G A A C A C C G A C G C C G C C C A G C T C T G G G C G G A C A
 A G C T G G C C G C G G G C A T C G G G G C G G G G A G G G G G A G G G C C G G
 C T A C G C G G G G C C G G A C C G C G C G C T G G C C G C C C T G C G T C G C G G C
 G A C C T G A C C A C C G C G G T C C A G G C G G C C G G C G C G G T C C T G G
 A C C G C G G C C G G C C G T C G C T C G G C A T C A C C G C C G T G T T
 G C C G T T G A G C G G C G G C G G T C G C C G C C G C G A T C C G G C T G G G C
 G A G C T C G A G C G G G C C G A G A A G T G G C T G G C C G A G C C G C T G C
 C C G A A G C C G T C C A C G A C A G C C T G T T C G G C C T G C A C C T G C T
 G A T G G C G C G G G C C G C T A C A G C C T C G C G G T G G G C C G G C A C
 G A G G C G G G C G T A C G C C G C G T T C C G G G A C T G C G G G T G A A C G G A
 T G C G C C G G T G G G A C G T C G A C G T G C C C G G G C T G G C C C T G T G
 G C G G G T G G A C G C G G C C G A G G C G C T G C T G C C C G G C G A T G A C
 C G G G C G G A G G G C C G G C G G C T G A T C G A C G A G C A G C T C A C C C
 G G C C G A T G G G G C C C C G G T C A C G A G G C C C T G A C C C T G C G G G T
 A C G A G C G G C C T A C G C C C C G C C G G C G A A A C G G A T C G A C C T G
 C T C G A C G A A G C G G G C G A C C T G C T G C T C T C A G C A A C G A C C
 A G T A C G A G C G G G C A C G G G T G C T G G C C G A C C T G A G C G A G G G C
 G T T C A G C G G C G T C C G G C A G A A C G G C C G G G C G C G G G C A T C
 C T G C G G C A G G C C C G G C A C C T G G C C G C C C A G T G C G G G G C G G
 T C C C C C T G C T G C G C C G G C T G G G C G T C A A G G C C G G C C G G T C
 C G G T C G G C T C G G C C G G C C G C C G A G G G A A T C C G C T C C C T G
 A C C G A G G C C G A G C G C C G G G T G G C C A C G C T G G C C G C C G C C G
 G G C A G A C C A A C C G G G A G A T C G C C G A C C A G C T C T T C G T C A C
 C G C C A G C A C G G T C G A G C A G C A C C T C A C C A A C G T G T T C C G C
 A A G C T C G G C G T G A A G G G C C G C C A G C A A T T G C C G G G C G A G C
 T G G C C G A C C T G C G G C C G C C G G G C T G A

配列番号 4 9

G T G G T C A C C G T C A C C G G C C C A A T C G C C T G C G G C A A G A C A G
 A A C T G C T T G A C G C G G C T G C C G C G A A G G C T G A G G C C A T C A T
 T C T G C G C G C G G T C T G C G C G C C A G A A A G A G C G G G C T A T G C C G
 T A C G C C A T G A T C G G G C A G C T C A T C G A C G A C C C G G G C G T C G
 C G C A T C G G G C G C C G G G G C T G G C T G A T C G G A T A G C C C A G G G
 C G G G C A G C T G T C G C T G A G G G C C G A G A A C C G A C T G C G C A G G
 G A T C T C A C C C G T G C C C T G C T G G C G C T T G C C G T C G A C C G G C
 C T G T G C T G A T C G G C G T C G A C G A T G T G C A T C A C G C C G A C A C
 C G C C T C T T G A A C T G T C T G C T G C A T T T G G C G C C G G G G T C
 C G T C C G G C C G G A T A T C C A T G A T C T T C A C C G A G T T G C G C A
 G C C T C A C C C C T A C T C A G T C A C G G T T C A A G G C G G A G C T G C T
 C A G C C T G C C G T A C C A C C A C G A G A T C G C G C T G C G T C C G T T C 50

GGACCGGAGCAATCGGCGGAGCTGGCCCGCGCCGCCCTCG
 GCCCCGGGCCTCGCCGAGGATGTGCTCGTGGGGTTGTATAA
 AACGACCAAGGGCAATCTGAGTCTCAGCCGTGGACTGATC
 AGCGATGTGCGGGAGGCCCTGGCCAACGGAGAGAGCGCC
 TCGAGGGCGGCCGCGCGTTCGGCTGGCGTACCTCGGCTC
 GCTCTACCGCTGTGGCCCGGGTCGCGCTGCGGGTGC
 GTGGCTGCCGTGCTGGGCCCGAGCGGCCACCAACCACGCTGG
 TCGGCCGTCTAACGGGGCTCAGCGCGGGAGACGATAGACCG
 GGCAACCAAGATCCTCACCGAGGGCGGGCTGCTGCTCGAC
 CAGCAGTTCCCGCACCCGGCGCCCGCTCGGTGGTGGTGGT
 ATGACATGTCCGCCAGGAACGACGCGGCCCTGCACACTCT
 CGCCCTGGAACTGCTGGACGAGGGCGCCGGTTGAAGTGCTC
 CGCACCACCAAGGTCGGCGCCGGTCTCATACACGGGCCA
 AGGCTGCGGAGATGTTCGCCAAGGCCGGCAAGGGCTCTGGT
 CGTACGCAACGAGTTGGGCGACGCGGGCAGAATACCTGCAA
 CTGGCTCACCGGGCTCCGACGATGTCTCCACCCGGGCCG
 CCTTACGGGTCGAGGCCGTGGCGATCGAGCGCCGCCGCAA
 TCCGCTGGCCTCCAGTCGGCACATGGACGAGCTGAGCGCC
 GCGGCCCGCGCCGGTCTGCTTTCCCCAAGCATGCGGCCG
 TGGCCGTCTCTGGCTGGCCGACGGCGGGCGATCCGGCGA
 GGCAAGCCGAGGTGCTGGCGTGGAAACGCCCGCTAGCGACC
 ACCGATCAGAACCGGGCCACTTGCAGATTGTCGAGGTGA
 CTCTCGCGCTGTTCTCTCCCGGCGCCTTCGGATCGGACCG
 CGGCCACCTCCGCTGACGCCGGACGAACTCGCCAGCCTG
 CCGAAGGGCGGCCCTGGCAATGCGCGGTGCGCCGACAACGCG
 CCATGACCGCCCTTGCACGGTCACTCAGAACATTGCCACCGC
 TCAGGGGAAACAGTTCTGCAGGCGAGGCTGATTGCGAGCC
 GACGCGATCCCCGCCGCGCTGATCGCCCTGTTGTACGCGG
 AGAACACCGAGTCCGCTCATATCTGGGCCGACAAGCTGGG
 CAGCACGAATGGCGGGGTATCGAACGAGGGCGGAAGCGGGC
 TACGCCGGCCCGTGCGCCGAGATCGCCCTGCGGCCGGCG
 ACCTGGCCACGGCGTTCGAGGGCTGGTAGCACCGTCCCTGGA
 CGACCGGTCGCTGCCGTCGCTCGGCATCACCGCCGCATTG
 CTGTTGAGCAGCAAGACGGCCGCCGCTGTCGCGATGGGCCG
 AACTCGAGCGTGCAGGAGAACGCTGCTCGCCGAGCCGCTTCC
 GAACGGCGTCCAGGGACAGCCCTTTCGGTCTGCACCTGCTC
 TCGGCATACTGGCCAGTACAGGCCCTCGCGATGGGCCGATATG
 AATCGGGCTCTCCGGGCGTTTACACCTGCGGGAGAACGTAT
 CGCGAGCTGGGATGTTGACGTCGCTGGTCTGGCCCTGTGG
 CGTGTGACGCCGCCGAGGGCGCTGCTCAGCCTCGACCGGA
 ACAGGGGCCAGCGGCTCATCGACGAACAACTCACCCGTCC
 GATGGGGCCTCGTTCCCGCGCGTTAACGCTGCGGATCAAG
 CGGGCATACCTCCCGGGACGAAGCGGGATCCCCCTGCTCC
 ATGAGGGCGGCCGAGCGTGCCTGCTCGCCCTGCCGACCCGTA
 CGAGCAAGCGCGGGTGCTCGCCGATCTGGCGACACGCTC
 AGCGCGCTCAGACGCTATAGCCGGCGGGAGTTCTCC
 GGCAAGGCTCGTCACCTGGCCGCCAGTGCAGGTGCTGTCCC
 GCTGCTGCGCAGGCTCGGGGGCGAGGCCCGGCCGGAATCGAC
 GACGCCGGCCTGCGCGACGGAGCACATCGTTGACCGATG
 CGGAGCGGGCGGGTGGCGGCGCTGGCCGCCGGACAGAC

10 20 30 40 50

CAACCGGGAGATCGCCAAACAGCTGTTCTGTCACGGCCAGC
 ACAGTGGAACAGCACCTCACAAAGCGTCTTCCGCAAACCTGG
 GGGTCAAGGGTCGCAAGCAGCTGCCGACCGCGCTGGCCGA
 CGTGGAACAGACCTGA

配列番号 50

ATGTATAGCGGTACCTGCCGTGAAGGATACGAACCTCGTCG
 CACGCGAGGACGAACCTCGGCATTCTACAGAGGTCTCTGG
 ACAAGCGAGCAGCGGCCAGGGCGTCGTGGTCACCGTCACC
 GCCCAATCGCCTGCGGCCAGGGCGTCGTGGTCACCGTCACC
 CTGCCCGAAGGCTGAGGCCATCATTCTGCGCGCGGTCTG 10
 CGCGCCAGAAGAGAGCGGGCTATGCCGTACGCCATGATCGGG
 CAGCTCATCGACGACCCGGCGCTCGCGCATCGGGCGCCGG
 GGCTGGCTGATCGGATAGCCCAGGGCGGGCAGCTGTCGCT
 GAGGGCCGAGAACCGACTGCGCAGGGATCTCACCCGTGCC
 CTGCTGGCGCTTGGCGCCGGTCCGTCGCTGATCGGCG 20
 TCGACGATGTGCATCACGCCGACACCGCCTCTTGAAC TG
 TCTGCTGCATTTGGCGCCGGTCCGTCGCTCCGGCCCGGATA
 TCCATGATCTTCACCGAGTTGCAGCCTCACCCCTACTC
 AGTCACGGTTCAAGGGAGCTGCTCAGCCTGCCGTACCA
 CCACGAGATCGCGCTGCGTCCGTTCCGGACCGGAGCAATCG
 GCGGAGCTGGCCCGCGCCCTTCGGCCCGGGCTCGCCG 30
 AGGATGTGCTCGTGGGTTGTATAAAACGACCAAGGGCAA
 TCTGAGTCTCAGCCGTGGACTGATCAGCGATGTCGGAG
 GCGCTGGCCAACGGAGAGAGCGCCCTTCGAGGGCGCCGCG
 CGTTCCGGCTGGCGTACCTCGGCTCGCTCACCGCTGTTG
 CCCGGTGCCTGGCGGCTGCCCGAGTGGCTGCCGTGCTG
 GGCCCGAGCGCCACCAACGCTGGTGCCTGCGCCGTCTAAGCG
 GGCTCAGCGCGGAGACGATAGACCGGGCAACCAAGATCCT
 CACCGAGGGCGGGCTGCTGCTCGACCAAGCAGTTCCCGCAC
 CCGGCCGCCGCTCGGTGGTCTTGATGACATGTCGCCCG 40
 AGGAACGACGCCCTGCACACTCTCGCCCTGGAACTGCT
 GGACGAGGCGCCGGTTGAAGTGCTCGCGCACCAAGGTC
 GGCGCCGGTCTCATACACGGGCCAACGGCTGCGGGAGATGT
 TCGCCAAGGGCGGCAAGGCTCTGGTCTGCAACGAGTT
 GGGCGACGCGGCAGAAATACCTGCAACTGGCTCACCGGGCC
 TCGGACGATGTCCTCCACCCGGGCCCTGCGGGTCTGAGGG
 CGTGGCGATCGAGCGCCGCCGCAATCCGCTGGCCTCCAG
 TCGGCACATGGACGAGCTGAGCGCCGCCGGCCGCGCCGGT
 CTGCTTCCCCAAGCATGCGGCCGCTGGCCGTCTTCTGGC
 TGGCCGACGGCGGGCGATCCGGCGAGGCAGCCGAGGTGCT
 GCGTGGAAACGCCCGCTAGCGACCAACCGATCAGAACCGG
 GCGCCACTTGCATTTGTCGAGGGTGAACCTCGCGCTGTTCT
 CTCCCCGGCGCCCTTCGGATCGGACCGGGGCCACCTCCGCT
 GACGCCGGACGAACCTGCCAGCCTGCCGAAGGGCGGCCCTGG
 CAATGCGCGGTGCGCCGACAACGCGGCCATGACCGCCCTTGC
 ACGGTCATCCAGAACCTGCCACCGCTCAGGGGAAACAGT
 TCTGCGGCAGGGCTGATTCTGGCAGCCGACGCGATCCCCGCC
 CGCCTGATCGCCCTGTTGTAACGCGGGAGAACACCGAGTCCG
 CTCATATCTGGGCCGACAAGCTGGGCAGCACGAATGGCGG
 GGTATCGAACCGAGGGCGGAAGCGGGCTACGCCGGCCCGTGC 50

G C C G A G A T C G C C C T G C G G C G C G G C G A C C T G G C C A C G G C G T
 T C G A G G C T G G T A G C A C C G T C C T G G A C G A C C G G T C G C T G C C
 G T C G C T C G G C A T C A C C G C C G C A T T G C T G T T G A G C A G C A A G
 A C G G C C G C C G C T G T C C G G C T G G G C G A A C T C G A G C G T G C G G
 A G A A G C T G C T C G C C G A G C C G C T T C C G A A C G G C G T C C A G G A
 C A G C C T T T C G G T C T G C A C C T G C T C T C G G C A T A C G G C C A G
 T A C A G C C T C G C G A T G G G C C G A T A T G A A T C G G C T C T C C G G G
 C G T T T C A C A C C T G C G G A G A A C G T A T G C G C A G C T G G G A T G T
 T G A C G T G C C T G G T C T G G C C C T G T G G C G T G T C G A C G C C G C C
 G A G G C G C T G C T C A G C C T C G A C C G G A A C G A G G G C C A G C G G C 10
 T C A T C G A C G A A C A A C T C A C C C G T C C G A T G G G G C C T C G T T C
 C C G C G C G C T G A C G C T G C G G A T C A A G G C G G C A T A C C T C C C G
 C G G A C G A A G C G G A T C C C C C T G C T C C A T G A G G C G G C G A G C
 T G C T G C T C C C C T G C C C C G A C C C G T A C G A G C A A G C G C G G G T
 G C T C G C C G A T C T G G G C G A C A C G C T C A G C G C G C T C A G A C G C
 T A T A G C C G G G C G C G G G G A G T T C T C C G G C A G G C T C G T C A C C
 T G G C C G C C C A G T G C G G T G C T G T C C C G C T G C T G C G C A G G C T
 C G G G G G C G A G C C C G G C C G G A T C G A C G A C G C C G G C T G C C G 20
 C A G C G G A G C A C A T C G T T G A C C G A T G C G G A G C G G C G G G T G G
 C G G C G C T G G C C G C G G C C G G A C A G C C A A C C G G G A G A T C G C
 C A A A C A G C T G T T C G T C A C G G C C A G C A C A G T G G A A C A G C A C
 C T C A C A A G C G T C T T C C G C A A A C T G G G G G T C A A G G G T C G C A
 A G C A G C T G C C G A C C G C G C T G G C C G A C G T G G A A C A G A C C T G
 A

配列番号 5 1

A T G C C T G C C G T G G A G A G C T A T G A A C T G G A C G C C C G C G A T G
 A C G A G C T C A G A A G A C T G G A G G G A G G G C G G T A G G C C A G G C G G G
 C A A C G G C C G G G G T G T G G T G G G T C A C C A T C A C C G G G C C G A T C
 G C C T G C G G C A A A G A C C G A A C T G C T C G A C G C G G C C G C C G G A 30
 A G A G C G A C G C C A T C A C A T T A C G T G C G G T C T G C T C C G A G G G A
 G G A A C G G G C C C T C C C G T A C G C C C T G A T C G G G C A G C T C A T C
 G A C A A C C C G G C G G T C G C C T C C C A G C T G C C G G A T C C G G T C T
 C C A T G G G C C C T C C C G G G C G A G C A C C T G T C G C C G G A G G G C C G A
 G A A C C G G C T G C G C G G C G A C C T C A C C C G T A C C C T G C T G G C G
 C T C G C C G C C G A A C G G C C G G T G C T G A T C G G C A T C G A C G A C A
 T G C A C C A C G C C G A C A C C G C C T C T T G A A C T G C C T G C T C C A
 C C T G G C C C G G A G G G T C G G C C C G G C G A T C G C C A T G G T C
 C T C A C C G A G C T G C G C C G G C T C A C C C C G G C C C A C T C C C A G T
 T C C A C G C C G A G C T G C T C A G C C T G G G G C A C C A C C G C G A G A T
 C G C G C T G C G C C C G C T C G G C C C G A A G C A C A T C G C C G A G C T G 40
 G C C C G C G C C G G C C T C G G T C C C G A T G T C G A C G A G G G A C G T G C
 T C A C G G G G T T G T A C C G G G C G A C C G G C G G C A A C C T G A A C C T
 C G G C C A C G G A C T G A T C A A G G A T G T G C G G G A G G G C C T G G G C G
 A C G G G C G G G A C G G G C A T C A A C G C G G G C G C G C G T A C C G G C
 T G G C G T A C C T C G G T T C C C T C T A C C G C T G C G G C C C G G T C C C
 G T T G C G G G T C G C A C G G G T G G C C G C C G T G C T G G G C C A G A G C
 G C C A A C A C C A C C C T G G T G C G C T G G A T C A G C G G G C T C A A C G
 C G G A C G C G G T G G G C G A G G G C G A C C G A G A T C C T C A C C G A G G G
 C G G C C T G C T G C A C G A C C T G C G G T T C C C G C A T C C G G C G G C C
 C G T T C G G T C G T A C T C A A C G A C C T G T C C G C C C G G A A C G C C 50

G C C G A C T G C A C C G G T C C G C T C T G G A A G T G C T G G A T G A C G T
 A C C C G T T G A A G T G G T C G C G C A C C A C C A G G G C C G G T G C C G G T
 T T C A T C C A C G G T C C C A A G G C C G C G A G A T C T T C G C C A A G G
 C C G G C C A G G A G C T G C A T G T G C G C G G C G A G C T G G A C G C C G C
 G T C C G A C T A T C T G C A A C T G G C C C A C C A C G C C T C C G A C G A C
 G C C G T C A C C C G G G C C G C G C T G C G G G T C G A G G G C C G T G G C G A
 T C G A G C G C C G C G C A A C C C G C T G G C C T C C A G G C C G C C A C C T
 C G A C G A G G C T G A C C G T C G C C G C C C G T G C C G G T C T G C T C T C C
 C T C G A G G C A C G C C G G C G C T G A T G A T C C C G C T G G C T G G C T C T C G
 G C G G G C G G T C C G G C G A G G G C G G G C C G A G G T G C T G G C C G C G A 10
 G C G C C C G C G T G C G G G T C A C C G A C C A G G A C A G G G C C C A C C T G
 C G G G C C G C C G A G G T A T C G C T G G C G C T G G T C A G C C C G G G C G
 C G T C C G G C G T C A G C C C G G G T G C G T C C G G C C G G A T C G G C G
 G C C G C G T C C G C T C C C G C C G G A T G A G C T C G C G A A C C T G C C G
 A A G G C G G C C C G G C T T T G T G C G A T C G C C G A C A A C G C C G T C A
 T A T C G G C C C T G C A C G G T C G T C C C G A G C T T G C C T C G G C C G A
 G G C G G A G A A C G T C C T G A A G C A G G G C T G A C T C G G C G G G C G A C
 G G C G C C A C C G C C C T C T C C G C G C T G A C G G C C T T G C T G T A C G
 C G G A G A A C A C C G A C A C C G C T C A G C T C T G G G C C G A C A A G C T
 C G T C T C C G A G A C C G G G C G T C G A A C G A G G G A G G G A A G G C G C G 20
 G G C T A C G C G G G G C C G C G C C G A G A C C G C G T T G C G C C G C G
 G C G A C C T G G C C G C G G G T C G A G G G C G G G C A G C G C C A T T C T
 G G A C C A C C G G C G G G G T C G T T G C T C G G C A T C A C C G C C G C G
 C T A C C G C T G A G C A G C G C G G T A G C C G C C G C A T C C G G C T G G
 G C G A G A C C G A G C G G G G C G G A G A A G T G G C T C G C C G A G C C G C T
 G C C G G A G G C C A T T C G G G A C A G C C T G T T C G G G C T G C A C C T G
 C T C T C G G C G C G C G G C C A G T A C T G C C T C G C G A C G G G C C G G C
 A C G A G T C G G C G T A C A C G G C G T T C C G C A C C T G C G G G G A A C G
 G A T G C G G A A C T G G G G C G T C G A C G T G C C G G G T C T G T C C C T G 30
 T G G C G C G T C G A C G C C G C G A G G G C G C T G C T G C A C G G C C G C G
 A C C G G G A C G A G G G C C G A C G G G C T C A T C G A C G A G C A G C T C A C
 C C A T G C G A T G G G A C C C C G T T C C C G C G C T T T G A C G C T G C G G
 G T G C A G G G C G G C G T A C A G C C C G C A G G G C G C A G C G G G T C G A C C
 T G C T C G A A A G A G G C G G C C G A C C T G C T G C T C T C C T G C A A C G A
 C C A G T A C G A G C G G G C G C G G G T G C T C G C C G A T C T G A G C G A G
 G C G T T C A G C G C G C T C A G G C A C C A C A G C C G G G C G C G G G G A C
 T G C T C C G G C A G G C C C G G C A C C T G G C C G C C C A G T G C G G G C G C
 G A C C C C G C T G C T G C G C G G C T C G G G G C C A A G C C C G G A G G C
 C C C G G C T G G C T G G A G G A A T C C G G C C T G C C G C A G C G G A T C A 40
 A G T C G C T G A C C G A C G C G G A G C G G G C G G G T G G C G T C G C T G G C
 C G C C G G G C G G C C A G A C C A A C C G C G T G A T C G C C G A C C A G C T C
 T T C G T C A C G G C C A G C A C G G T G G A G C A G C A C C T C A C G A A C G
 T C T T C C G C A A G C T G G G G C G T C A A G G G C C G C A G C A C C T G C C
 G G C C G A A C T C G C C A A C G C G G A A T A G

配列番号 5 2

A T G C C T G C C G T G G A G A G C T A T G A A C T G G A C G C C C G C G A T G
 A C G A G C T C A G A A G A C T G G A G G G A G G G C G G T A G G C C A G G C G G G
 C A A C G G C C G G G G T G T G G T G G T C A C C A T C A C C G G G C C G A T C
 G C C T G C G G C A A G A C C G A A C T G C T C G A C G C G G G C C G C G C G A 50
 A G A G C G A C G C C A T C A C A C T G C G T G C G G T C T G C T C C G A G G G A

G G A A C G G G C C C T C C C G T A C G C C C T G A T C G G G C A G C T C A T C
 G A C A A C C C G G C G G T C G C C T C C C A G C T G C C G G A T C C G G T C T
 C C A T G G C C C T C C C G G G C G A G C A C C T G T C G C C G G A G G C C G A
 G A A C C G G C T G C G C G G C G A C C T C A C C C G T A C C C T G C T G G C G
 C T C G C C G C C G A A C G G C C G G T G C T G A T C G G C A T C G A C G A C A
 T G C A C C A C G C C G A C A C C G C C T C T T G A A C T G C C T G C T C C A
 C C T G G C C C G G A G G G T C G G C C C G G C C C G G A T C G C C A T G G T C
 C T C A C C G A G C T G C G C C G G C T C A C C C G G C C C A C T C C C A G T
 T C C A C G C C G A G C T G C T C A G C C T G G G G C A C C A C C G C G A G A T
 C G C G C T G C G C C C G C T C G G C C C G A A G C A C A T C G C C G A G C T G 10
 G C C C G C G C C G G C C T C G G T C C C G A T G T C G A C G A G G A C G T G C
 T C A C G G G G T T G T A C C G G G C G A C C G G C G G C A A C C T G A A C C T
 C G G C C A C G G A C T G A T C A A G G A T G T G C G G G A G G G C C T G G G C G
 A C G G G C G G G A C G G G C A T C A A C G C G G G C C G C G C G T A C C G G C
 T G G C G T A C C T C G G T T C C C T C T A C C G C T G C G G C C C G G T C C C
 G T T G C G G G T C G C A C G G G T G G C C G C G T G C T G G G C C A G A G C
 G C C A A C A C C A C C C T G G T G C G C T G G A T C A G C G G G C T C A A C G
 C G G A C G C G G T G G G C G A G G C G A C C G A G A T C C T C A C C G A G G G
 C G G C C T G C T G C A C G A C C T G C G G T T C C C G C A T C C G G C G G C C
 C G T T C G G T C G T A C T C A A C G A C C T G T C C G C C C G G G A A C G C C 20
 G C C G A C T G C A C C G G T C C G C T C T G G A A G T G C T G G A T G A C G T
 A C C C G T T G A A G T G G T C G C G C A C C A C C A G G G C C G G T G C C G G T
 T T C A T C C A C G G T C C C A A G G C C G C G A G A T C T T C G C C A A G G
 C C G G C C A G G A G C T G C A T G T G C G G C G G C G A G C T G G A C G C C G C
 G T C C G A C T A T C T G C A A C T G G C C C A C C A C G C C T C C G A C G A C
 G C C G T C A C C C G G G C C G C G C T G C G G G T C G A G G G C C G T G G C G A
 T C G A G C G C C G C G C A A C C C G C T G G C C T C C A G C C G C C A C C T
 C G A C G A G C T G A C C G T C G C C G C C C G T G C C G G T C T G C T C T C C
 C T C G A G G C A C G C C G C G C T G A T G A T C C G C T G G C T G G C T C T C G
 G C G G G C G G T C C G G G C G A G G G C G G C C G A G G T G C T G G C C G G C A 30
 G C G C C C G C G T G C G G T C A C C G A C C A G G A C A G G G C C C A C C T G
 C G G G C C G C C G A G G T A T C G C T G G C G C T G G T C A G C C C G G G C G
 C G T C C G G C G T C A G C C C G G G T G C G T C C G G C C C G G A T C G G G C G
 G C C G C G T C C G C T C C C G C C G G A T G A G C T C G C G A A C C T G C C G
 A A G G C G G C C C G G C T T T G T G C G A T C G C C G A C A A C G C C G T C A
 T A T C G G G C C C T G C A C G G T C G T C C C G A G C T T G C C T C G G G C C G A
 G G C G G A G A A C G T C C T G A A G C A G G G C T G A C T C G G G C G G G A C
 G G C G C C A C C G C C C T C T C C G C G C T G A C G G G C T T G C T G T A C G
 C G G A G A A C A C C G A C A C C G C T C A G C T C T G G G C C G A C A A G C T
 C G T C T C C G A G A C C G G G G C G T C G A A C G A G G G A G G G A A G G C G C G 40
 G G C T A C G C G G G G C C G C G C C G A G A C C G C G T T G C G C C G C G
 G C G A C C T G G G C C G C G G G G T C G T T G C T C G G C A T C A C C G C C G C G
 G G A C C A C C G G C G G G G G T C G T T G C T C G G C A T C A C C G C C G C G
 C T A C C G C T G A G C A G C G C G G T A G C C G C C G C A T C C G G C T G G
 G C G A G A C C G A G C G G G G C G G A G A A G T G G C T C G C C G A G C C G C T
 G C C G G A G G C C A T T C G G G A C A G C C T G T T C G G G C T G C A C C T G
 C T C T C G G G C G C G C G G C C A G T A C T G C C T C G C G A C G G G C C G G C
 A C G A G T C G G G C G T A C A C G G G C G T T C C G C A C C T G C G G G G A A C G
 G A T G C G G A A C T G G G G G C G T C G A C G T G C C G G G T C T G T C C C T G
 T G G C G C G T C G A C G C C G C C G A G G G C G C T G C T G C A C G G C C G C G 50

A C C G G G A C G A G G G C C G A C G G C T C A T C G A C G A G C A G C T C A C
 C C A T G C G A T G G G A C C C C G T T C C C G C G C T T T G A C G C T G C G G
 G T G C A G G C G G C G T A C A G C C C G C A G G C G C A G C G G G T C G A C C
 T G C T C G A A G A G G C G G C G A C C T G C T G C T C T C C T G C A A C G A
 C C A G T A C G A G G C G G G C G C G G G T G C T C G C C G A T C T G A G C G A G
 G C G T T C A G C G C G C T C A G G C A C C A C A G C C G G G C G C G C G G G A C
 T G C T C C G G C A G G C C C G G C A C C T G G C C G C C C A G T G C G G G C G C
 G A C C C C G C T G C T G G C C G G C T C G G G G C C A A G C C C G G A G G G C
 C C C G G C T G G C T G G A G G A A T C C G G C C T G C C G C A G C G G A T C A
 A G T C G C T G A C C G A C G C G G A G C G G G C G G G T G G C G T C G C T G G C
 C G C C G G C G G C C A G A C C A A C C G C G T G A T C G C C G A C C A G C T C
 T T C G T C A C G G C C A G C A C G G T G G A G G C A G C A C C T C A C G A A C G
 T C T T C C G C A A G C T G G G C G T C A A G G G C C G C A G C A C C T G C C
 G G C C G A A C T C G C C A A C G C G G A A T A G

配列番号 53

G T G A A G C G C A A C G A T C T G G T T G C C C G C G A T G G C G A G C T C A
 G G T G G A T G C A A G A G A T T C T C A G T C A G G C G A G C G G A G G G C C G
 G G G G G C C G T G G T C A C C A T C A C G G G G G C G A T C G C C T G T G G C
 A A G A C G G T G C T G C T G G A C G C C G C G G C A G C C A G T C A A G A C G
 T G A T C C A A C T G C G T G C G G T C T G C T C G G C G G A G G G A G C A G G A
 G C T G C C G T A C G C G A T G G T C G G A C A A C T A C T C G A C A A T C C G
 G T G C T C G C C G C G C G A G T G C C G G C C C T G G G C A A C C T G G C T C T C G C C G A C
 C G G C G G G C G A G C G G G C T G C T G C C G G C A C C G A G A A C A G G A T
 C C G G C G G G A G G C T C A C C C G C A C C C T G C T G G C T C T C G C C G A C
 G A A C G A C C G G G T G C T G A T C G G G G T C G A C G A C A T G C A C C A T G
 C G G A C C C C G C C T C G C T G G A C T G C C T G C T G C A C C T G G C C C G
 G C G G G T C G G C C C G G C C G A T C G C G A T C G T T C T G A C C G A G
 T T G C G C C G G C T C A C C C C G G C T C A C T C G C G C T T C C A G T C C G
 A G C T G C T C A G C C T G C G G T A C C A C C A C G A G A T C G G G T T G C A
 G C C G C T C A C C G C G G A G C A C A C C G C C G A C C T G G G C C C G C G T C
 G G C C T C G G T G C C G A G G G T C G A C G A C G A C G T G C T C A C C G A G C
 T C T A C G A G G G C G A C C G G C G G C A A C C C G A G G A C T G T G C T G C G G
 C C T G A T C A G G G A C G T G C G G C A G G G A C T G G G A G G G C C G G G G T C
 A C C G G T A T C C A C G T C G G C C G G G C G T A C C G G C T G G C C T A T C
 T C A G T T C G C T C T A C C G C T G C G G C C C G G G C G G C G T G C G G A C
 C G C C C G C G C G G C C G G G T G C T G G G C G A C A G C G C C G A C G C C
 T G C C T G A T C C G C C G G G T C A G C G G G C C T C G G T A C G G A G G G C G
 T G G G C C A G G G C G A T C C A G C A G C T C A C C G A G G G G C G G C C T G C T
 G C G T G A C C A G C A G T T C C C G C A C C C G G C G G C C C G C T C G G T C
 G T G C T C G A C G A C A T G T C C G C G C A G G A A C G C C A C G C G A T G T
 A T C G C A G C G C C C G G G A G G G C A G C C G C C G A A G G G T C A G G C C G A
 C C C C G G C A C C C C G G G C G A G C C G C G G G C G G C T A C G G C G T A C
 G C C G G G T G T G G T G A G C A A G C C G G G T G A C T A C C C G G A G G C C G G
 C C G G G C C G G G C T G C G T G G A C G G G T G C C G G G T C C G G C C G A G T A
 C T G C G G C G A C C C G C A C G G C G C C G A C G A C G A C C C G G A C G A G
 C T G G T C G C C G C G C T G G G C G G G C T G C T G C C G A G G C C G G C T C G
 T G G G C G A T G A A G A T C C G G C G C C T G G C G G G T G G G C C G G G C G C C C
 C G G G G C G G G C T G C C G A G C T G C T G A C C T C G C A G C G G G T T G C A C
 G C G G T G A C C A G C G A G G A C C G G G C C A G C C T G C G G G C C G C C G
 A G G T G G G C G C T G C C A C G C T G T G G C C G G G T G C G A C C G G C C C

10

20

30

40

50

GGACCGGCATCCGCTCACGGAGCAGGAGGCAGGCTGAGCCTG
 CCGGAGGGTCCGCGCCTGCTCGCTGCCGCCGACGATGCCG
 TCGGGGCCGCCCTGCGCGGTCGCGGCCGAGTACGCCGCC
 CGAGGCAGGAGAACGTCCTGCGGCACGCCGATCCGGCAGCC
 GGTGGTGACGCCATGCCGATGATCGCCCTGCTGTACA
 CGGAGCACCCCGAGAACGTCGCTGTTCTGGGCCGACAAGCT
 CGACGCCGGCCGCCCGACGAGGAGACCAAGTTATCCC
 CTGCGGGCCGAGACCAGCGGTTGCGGCTCGGTGACCTGGAAA
 CGCGATGGAGCTGGGCCGACGGTGCCTGGACCAAGCGGGCG
 GCTGCCGTCCTGGGTGTCGCCGCGGGCCTGCTCCCTGGGC 10
 GGCGCGGTGACGGCCGCCATCCGGCTCGGCCGACCTCGACC
 GGGCGGGAGAACGTTGGCTCGCCGAGCCGATCCCCGACGCCAT
 CGTACCAAGCCTCTACGGCCCTGCACGTCGCTGGGCCGCGG
 GCCCGGCTCGACCTGGCCGCGGGCCGCTACGAGGCGGGGT
 AACACGGCGTTCCGGCTGTTGGCGAGCGGGATGGCAGGGCTG
 GGATGCCGATGTCCTCCGGGCTGGCGCTGTTGGCGCTGAC
 CGCGCCGAGGCCCTGCTGTCGCCGCGGGCATCCGCCGGACG
 AGGGCCGCAAGCTCATCGACGACCGCTCACCCGTGAGAT
 GGGGCCCGCTTCCCGGGCGCTGACGCTGCGGGCGCAAGCG
 CGTACAGCCTGCCGGTGCACCGGGTGGGCCCTGCTCGACG 20
 AGGCGGCCGGCCTGCTGCTCGCCATGACGGGTACGA
 CGGGCGCGGGTGCCTGCGGACCTGGGGAGACCCCTGCGC
 ACGCTGCGGCACACCGACGCGGCCAGCGGGTGCCTCCGGC
 AGGCCGAGCAGGCGGCCGCGCGGGTGCCTGGTCCCCT
 GCTGCGGCCTCGGGGCCGAACCCGTACGCACTGGCACCC
 CGCGTGGTGAACCCGGCCCTGCCGAGCGGGATCAGGCTGC
 TGACCGATGCCGAGCGGGGTTGCCGCGATGGCCGCCGC
 CGGGCAGACCAACCGGGAGATCGCCGGTCTGGCTCTCGT
 ACGGCCAGCACGGTGGAGCAGCACCTGACCAAGCGTCTTC
 GCAAGCTGGCGTCAAGGGCCGCCGGTTCCCTGCCGACCGA 30
 GCTCGCCCAAGCGTCTGA

配列番号 54

ATGCCCTGCCGTGAAGCGCAACGATCTGGTTGCCCGCGATG
 CGAGGCTCAGGTGGATGCAAGAGAGATTCTCAGTCAGGGCGAG
 CGAGGGCCGGGGGCCGTGGTCACCATCACGGGGCGATC
 GCCTGTGGCAAGACGGTGCCTGGACGCCGCCGAGCCA
 GTCAAGACGTGATCCAACCTGCGTGCCTGCTCGGG
 GGAGCAGGAGCTGCCGTACCGGATGGTCGGACAACTACTC
 GACAATCCGGTGCCTGCCGCGAGTGCCTGGCCCTGGCA
 ACCTGGCTGCCGGCGGGCGAGCGGGCTGCTGCCGGCACCGA 40
 GAAACAGGGATCCGGCGGGAGCTCACCCGCACCCCTGCTGGCT
 CTGCCGACGAACGACCGGTGCTGATCGCGTCTGACGACA
 TGCACCATGCCGACCCCGCCCTCGCTGGACTGCCCTGCTGCA
 CCTGGCCCGGCCGGTCCGGCCCGGCATCGCGATCGTT
 CTGACCGAGTTGCGCCGGCTCACCCCGGCTCACTCGCGCT
 TCCAGTCCGAGCTGCTCACGCCCTGCGGTACCAACGAGAT
 CGGGTTGCAGCCGCTCACCGGGAGCACACCGCCGACCTG
 GCCCAGCTCGGCCCTCGGTGCCGAGGTCGACGACGACGTGC
 TCACCGAGCTCTACGAGGGCAGCCGGCGAGTCT
 GTGCTGCCGCTGATCAGGGACGTGCGGGCAGGACTGGAG 50

G C C G G G G T C A C C G G T A T C C A C G T C G G C C G G G C G T A C C G G C
 T G G C C T A T C T C A G T T C G C T C T A C C G C T G C G G C C C G G C G G C
 G C T G C G G A C C G C C C G C G C G G C C G C G G T G C T G G G C G A C A G C
 G C C G A C G C C T G C C T G A T C C G C C G G G T C A G C G G C C T C G G T A
 C G G A G G C C G T G G G C C A G G C G A T C C A G C A G C T C A C C G A G G G
 C G G C C T G C T G C G T G A C C A G C A G T T C C C G C A C C C G G C G G C C
 C G C T C G G T C G T G C T C G A C G A C A T G T C C G C G C A G G A A C G C C
 A C G C G A T G T A T C G C A G C G C C C G G G A G G C A G C C G C C G A A G G
 T C A G G C C G A C C C C G G C A C C C C G G G C G A G C C G C G G G C G G C T 10
 A C G G C G T A C G C C G G G T G T G G T G A G C A A G C C G G T G A C T A C C
 C G G A G C C G G C C G G G C C G G G C C T G C G T G G A C G G T G C C G G T C C
 G G C C G A G T A C T G C G G C G A C C C G C A C G G C G C C G A C G A C G A C
 C C G G A C G A G C T G G T C G C C G C G C G C T G G G C G G G C T G C T G C C G A
 G C C G G C T C G T G G C G A T G A A G A T C C G G C G C C T G G C G G T G G C
 C G G G C G C C C C G G G G C G G C T G C C G A G G C T G C T G A C C T C G C A G
 C G G T T G C A C G C G G T G A C C A G C G A G G A C C G G G C C A G C C T G C
 G G G C C G C C G A G G T G G C G C T C G C C A C G C T G T G G G C C G G G T G C
 G A C C G G C C C G G A C C G G C A T C C G C T C A C G G A G C A G G A G G G C G
 G C G A G C C T G C C G G A G G G T C C G C G C C T G C T C G C T G C T G C C G C C G
 A C G A T G C C G T C G G G G C C G C C C T G C G C G G T C G C G C C G A G T A 20
 C G C C G C G G C C G A G G C G G A G A A C G T C C T G C G G C A C G C C G A T
 C C G G C A G C C G G T G G T G A C G C C T A C G C C G C C A T G A T C G C C C
 T G C T G T A C A C G G A G C A C C C C G A G A A C G T G C T G T T C T G G G C
 C G A C A A G C T C G A C G C G G C C G C C C C G A C G A G G G A G A C C A G T
 T A T C C C G G G C T G C G G G C C G A G A C C G C G G T G C G G G C T C G G T G
 A C C T G G A A A C G G C G A T G G A G G C T G G G C C G C A C G G T G C T G G A
 C C A G C G G C G G C T G C C G T C C C T G G G T G T C G C C G C G G G C C T G
 C T C C T G G G C G G C G G C G G T G A C G G C C G C A T C C G G C T C G G C G
 A C C T C G A C C G G G C G G A G A A G T G G C T C G C C G A G C C G A T C C C
 C G A C G C C A T C C G T A C C A G C C T C T A C G G G C C T G C A C G T G C T G 30
 G C C G C G C G G G C C G G C T C G A C C T G G C C G C G G G C C G C T A C G
 A G G C G G C G T A C A C G G C G T T C C G G C T G T G T G G C G A G C G G A T
 G G C A G G C T G G G A T G C C G A T G T C T C C G G G C T G G C G C T G T G G
 C G C G T C G A C G C C G C C G A G G C C C T G C T G T C C G C G G G C A T C C
 G C C C G G A C G A G G G C C G C A A G C T C A T C G A C G A C C A G C T C A C
 C C G T G A G A T G G G G G C C C G C T C C C G G G C G C T G A C G C T G C G G
 G C G C A A G C G G G C G T A C A G C C T G C C G G G T G C A C C G G G T G G G C
 T G C T C G A C G A G G C G G C C G G C C T G C T G C T C G C C C T G C C A T G A
 C G G G T A C G A G C G G G C G C G G G T G C T C G C G G A C C T G G G G G A G
 A C C C T G C G C A C G C T G C G G C A C A C C G A C G C G G C C C A G C G G G 40
 T G C T C C G G C A G G C C G A G C A G G C G G C C G C G C G G T G C G G G T C
 G G T C C C G C T G C T G C G G C G G C T C G G G G C C G A A C C C G T A C G C
 A T C G G C A C C C G G C G T G G T G A A C C C G G C C T G C C C G C A G C G G A
 T C A G G C T G C T G A C C G A T G C C G A G C G G C G G G T T G C C G C G A T
 G G C C G C C G C G G G C A G A C C A A C C G G G A G A T C G C C G G T C G G
 C T C T T C G T C A C G G C C A G C A C G G T G G A G C A G C A C C T G A C C A
 G C G T C T T C C G C A A G C T G G G C G T C A A G G G C C G C C G G T T C C T
 G C C G A C C G A G C T C G C C C A A G C C G T C T G A

配列番号 55

G T G G T C A C C G T C A C C G G C C C A A T C G C C T G C G G C A A G A C A G 50

A A C T G C T T G A C G C G G C T G C C G C G A A G G C T G A G G C C A T C A T
 T C T G C G C G C G G T C T G C G C G C C A G A A G A G A G C G G G G C T A T G C C G
 T A C G C C A T G A T C G G G C A G C T C A T C G A C G A C C C G G G G C T C G
 C G C A T C G G G C G C C G G G G C T G G C T G A T C G G A T A G C C C A G G G
 C G G G C A G C T G T C G C T G A G G G C C G A G A A C C G A C T G C G C A G G
 G A T C T C A C C C G T G C C C T G C T G G C G C T T G C C G T G G A C C G G C
 C T G T G C T G A T C G G C G T C G A C G A T G T G C A T C A C G C C G A C A C
 C G C C T C T T G A A C T G T C T G C T G C A T T T G C C C G C C G G G T C
 C G T C C G G C C C G G A T A T C C A T G A T C T T C A C C G A G T T G C G C A
 G C C T C A C C C C T A C T C A G T C A C G G T T C A A G G C G G A G C T G C T 10
 C A G C C T G C C A T A C C A C C A C G A G A T C G C G C T G C G T C C A T T C
 G G A C C G G A G C A A T C G G C G G A G C T G G C T C G C G C C G C C T T C G
 G C C C G G G C C T C G C C G A G G A T G T G C T C G C G G G G T T G T A T A A
 A A C G A C C A G G G G C A A T C T G A G T C T C A G C C G T G G A C T G A T C
 A G C G A T G T G C G G G A G G C C C T G G C C A A C G G A G A G A G C G C T T
 T C G A G G G C G G G C C G C G C G T T C C G G C T G G C G T A C C T C A G C T C
 G C T C T A C C G C T G T G G C C C G G T C G C G C T G C G G G T C G C C C G A
 G T G G C T G C C G T G C T G G G C C C A A G C G C C A C C A C C A C G C T G G
 T G C G C C G G C T A A G C G G G C T C A G C G C G G A G A C G A T A G A C C G
 G G C A A C C A A G A T C C T C A C T G A G G G C G G G C T G C T G C T C G A C 20
 C A G C A G T T C C C G C A C C C G G C C G C C C G C T C G G T G G T G C T C G
 A T G A C A T G T C C G C C C A G G A A C G A C G C A G C C T G C A C A C T C T
 C G C C C T G G A A C T G C T G G A C G A G G G C G C C G G T T G A A G T G C T C
 G C G C A C C A C C A G G T C G G C G C C G G T C T C A T A C A C G G G C C C A
 A G G C T G C G G A G A T G T T C G C C A A G G C C G G C A A G G C T C T G G T
 C G T A C G C A A C G A G T T G G G C G A C G C G G C C G A A T A C C T G C A A
 C T G G C T C A C C G G G C C T C C G A C G A T G T C T C C A C C C G G G C C G 30
 C C T T A C G G G T C G A G G C C G T G G C C A T C G A G C G C C G C G C A A
 T C C G C T G G C C T C C A G T C G G C A C A T G G A C G A A C T G A G C G C C
 G C C G G C C G C G C C G G T C T G C T T T C C C C C A A G C A T G C G G G C G C
 T G G C C G T C T T C T G G C T A G C C G A C G G G C G G G C G A T C C G G C G A 40
 G G C A G C C G A A G T G C T G G C G T C G G A A C G C C C G C T C G C G A C C
 A C C G A T C A G A A C C G G G C C C A C C T G C G A T T T G T C G A G G T G A
 C T C T C G C G C T G T T C T C T C C C G G C C T T C G G A T C G G A C C G
 G C G C C C A C C T C C G C T G A C G C C G G A C G A A C T C G C C A G C C T G
 C C G A A G G C G G C C T G G C A A T G C G C G G T C G C C G A C A A C G C G G
 C C A T G A C C G C C T T G C A C G G C C A T C C A G A A C T T G C C A C C G C
 T C A G G C G G A A A C A G T T C T G C G G C A G G C T G A T T C G G C A G C C
 G A C G C G A T C C C C G C C G G C T G A T C G C C C T G T T G T A C G C G G 50
 A G A A C A C C G A G T C C G C T C A T A T C T G G G C C G A C A A G C T G G G
 C A G C A C G A A T G C C G G G G T A T C G A A C G A G G G C G G A A G C G G G C
 T A C G C C G G C C C G T G C G C C G A G A G T C G C C C T G C G G G C G G C G
 A C C T G G C C A C G G C G T T C G A G G G C T G G T A G C G C C G T C C T G G A
 C G A C C G G T C G C T G C C G T C G C T C G G C A T C A C C G C C G A T T G
 C T G T T G A G C A G C A A G A C G G C C G C C G T G T C C G G C T G G G C G
 A A C T C G A G C G T G C G G G A G A A G C T G C T C G C C G A G C C G C T T C C
 G A A C G G G G T C C A G G A C A G C C C T T T C G G G T C T G C A C C T G C T C
 T C G G C G T A C G G C C A G T A C A G C C T C G C G A T G G G C C G A T A T G
 A A T C A G C T C A C C G G G C G T T T C G C A C C T G C G G G A G A A C G T A T
 G C G C A G C T G G G A T G T T G A C G T G C C T G G G T C T G G C C C T G T G G

C GT G T C G A C G C C G C C G A G G C G C T G C T C A G C C T C G A C C G G A
 A C G A G G G C C A G C G G C T C A T C G A C G A A C A A C T C A C C C G T C C
 G A T G G G G C C T C G T T C C C A C G C G T T A A C G C T G C G G A T C A A G
 G C G G C A T A C C T C C C G C G G A C G A A G C G G A T C C C C T G C T C C
 A T G A G G C G G C C G A G C G C T G C T G C T C C C C T G C C C C G A C C C G T A
 C G A G C A A G C G C G G G T G C T C G C C G A T C T G G G C G A C A C G C T C
 A G C G C G C T C A G A C G C T A T A G C C G G G C G C G G G G A G T T C T C C
 G G C A G G C T C G T C A C C T G G C C A C C C A G T G C G G T G C T G T C C C
 G C T G C T G C G C A G G C T C G G G G G C G A G C C C G G C C G G A T C G A C
 G A C G C C G G C C T G C C G C A G C G G A G C A C A T C G T T G A C C G A T G 10
 C G G A G C G G G C G G G T G G C G G C G T G G C C G G C G G G A C A G A C
 C A A C C G G G A G A T C G C C G A A C A G C T G T T C G T C A C G G C C A G C
 A C A G T G G A A C A G C A C C T C A C A A G C G T C T T C C G C A A G C T G G
 G C G T C A A G G G C C G C A A G C A G C T G C C G A C C G C G C T G G C C G A
 C G T G G A A C A G A C C T G A

配列番号 5 6

A T G T A T A G C G G T A C C T G C C G T G A A G G A T A C G A A C T C G T C G
 C A C G C G A G G A C G A A C T C G G T A T T C T A C A G A G G T C T C T G G A
 A C A A G C G A G C A G C G G C C A G G G C G T C G T G G T C A C C G T C A C C
 G G C C C A A T C G C C T G C G G C A A G A C A G A A C T G C T T G A C G C G G 20
 C T G C C G C G A A G G C T G A G G C C A T C A T T C T G C G C G C G G T C T G
 C G C G C C A G A A G A G A G C G G G C T A T G C C G T A C G C C A T G A T C G G G
 C A G C T C A T C G A C G A C C C G G C G C T C G C G C A T C G G G C G C C G G
 G G C T G G C T G A T C G G A T A G C C C A G G G C G G G C A G C T G T C G C T
 G A G G G C C G A G A A C C G A C T G C G C A G G G A T C T C A C C C G T G C C
 C T G C T G G C G C T T G C C G T G G A C C G G C C T G T G C T G A T C G G C G
 T C G A C G A T G T G C A T C A C G C C G A C A C C G C C T C T T G A A C T G 30
 T C T G C T G C A T T T G G C C C G C C G G G T C C G T C C G G C C C G G A T A
 T C C A T G A T C T T C A C C G A G T T G C G C A G C C T C A C C C C T A C T C
 A G T C A C G G T T C A A G G C G G A G G C T G C T C A G C C T G C C A T A C C A
 C C A C G A G A T C G C G C T G C G T C C A T T C G G A C C G G A G C A A T C G 40
 G C G G A G G C T G G C T C G C G C C G C C T T C G G C C C G G G C C T C G C C G
 A G G A T G T G C T C G C G G G G T T G T A T A A A A C G A C C A G G G G C A A
 T C T G A G T C T C A G C C G T G G A C T G A T C A G C G A T G T G C G G G A G
 G C C C T G G C C A A C G G A G A G A G G C G C T T T C G A G G G C G G G C C G C G
 C G T T C C G G C T G G C G T A C C T C A G C T C G C T C T A C C G C T G T G G
 C C C G G T C G C G C T G C G G G G T C G C C C G A G T G G C T G C C G T G C T G
 G G C C C A A G C G C C A C C A C C A C G C T G G T G C G C C G G C T A A G C G
 G G C T C A G C G C G G A G A C G A T A G A C C G G G C A A C C A A G A T C C T
 C A C T G A G G G C G G G C T G C T C G A C C A G C A G T T C C C G C A C 40
 C C G G C C G C C C G C T C G G T G G T G C T C G A T G A C A T G T C C G C C C
 A G G A A C G A C G C A G C C T G C A C A C T C T C G C C C T G G A A C T G C T
 G G A C G A G G C G C C G G T T G A A G T G C T C G C G C A C C A C C A G G T C
 G G C G C C G G T C T C A T A C A C G G G C C A A G G C T G C G G A G A T G T
 T C G C C A A G G C C G G C A A G G C T C T G G T C G T A C G C A A C G A G T T
 G G G C G A C G C G G C C G A A T A C C T G C A A C T G G C T C A C C G G G C C
 T C C G A C G A T G T C T C C A C C C G G G C C G C C C T G C G G G T C G A G G
 C C G T G G C C A T C G A G C G C C G C G C A A T C C G C T G G C C T C C A G
 T C G G G C A C A T G G A C G A A C T G A G G C G C C G C G G G C C G C G C C G G T
 C T G C T T T C C C C A A G C A T G C G G C G C T G G C C G T C T T C T G G C C 50

TAGCCGACGGCGGGCGATCCGGCGAGGCAGCCGAAGTGT
 GCGTCGGAACGCCCCGCTCGCGACCAACGATCAGAACCGG
 GCCCACCTGCGATTGTCGAGGTGACTCTCGCGCTGTTCT
 CTCCC GGCGCCCTTGGATCGGACCGGGCGCCACCTCCGCT
 GACGCCGGACGAACCTCGCCAGCCTGCCGAAGGGCGGCTGG
 CAATGCGCGGTCGCCGACAACGCGGCCATGACCGCCCTTGC
 ACGGCCATCCAGAACCTTGCACCGGCTCAGGGGAAACAGT
 TCTGCGGCAGGGCTGATTGCCAGGGGACGGCGATCCCCGCC
 CGCGT GATCGCCCTGTTGTACGCGGAGAACACCGAGTCCG
 CTCATATCTGGGCCGACAAGCTGGGCAGCACGAATGCCGG 10
 GGTATCGAACGAGGGCGGAAGCGGGCTACGCCGGCCGTG
 GCCGAGATCGCCCTGCGGCGGGGACCTGGCCACGGCGT
 TCGAGGGCTGGTAGCGCCGTCCTGGACGACCGGTCGCTGCC
 GTCGCTCGGCATCACCGCCGCATTGCTGTTGAGCAGCAAG
 ACGGCCGCCGCTGTCGGCTGGCGAACACTCGAGCGTGC
 AGAACAGCTGCTCGCCGAGCCGCTTCCGAACGGCGTCCAGGA
 CAGCCTTTCGGTCTGCACCTGCTCTCGGCGTACGGCCAG
 TACAGCCTCGCGATGGGCCGATATGAATCAGCTCACCGGG
 CGTTTCGCACCTGCGGAGAACGTATGCGCAGCTGGGATGT
 TGACGTGCCTGGTCTGGCCCTGTTGGCGTGTGACGCCGCC 20
 GAGGGCGCTGCTCAGCCTCGACCGGAACGAGGGGCCAGCGGC
 TCATCGACGAAACAACTCACCCGTCCTGGCGATGGGGCCCTCG
 CCACGCGCTGACGCTGCGGATCAAGGCGGCATACCTCCCG
 CGGACGAAGCGGATCCCCCTGCTCCATGAGGGCGGCCGAGC
 TGCTGCTCCCCCTGCCCCGACCCGTACGAGCAAGCGCGGGT
 GCTCGCCGATCTGGCGACACGCTCAGCGCGCTCAGACGC
 TATAGCCGGCGCGGGGAGTTCTCCGGCAGGCTCGTCACC
 TGGCCACCCAGTGCGGTGCTGTCCTGGCGTGCAGCGAGGCT
 CGGGGGCGAGCCCCGGCGGATCGACGACGCCGGCTGCCG 30
 CAGCGGAGCACATCGTTGACCGATGCGGAGCGGGGGTGG
 CGGCGCTGGCCGCGGCCGGACAGACCAACCGGGAGATCGC
 CGAACAGCTGTTGTCACGGCCAGCACAGTGGAAACAGCAC
 CTCACAAAGCGTCTTCCGCAAGCTGGCGTCAAGGGCCGCA
 AGCAGCTGCCGACCGCGCTGGCCGACGTGGAACACAGACCTG
 A

配列番号 57

GTGTATAGCGGTACCTGCCGTGAAGGATAACGAACCTCGTCG
 CCCCGCGAGGACGAACCTCGGCATTCTGCAGAGGTCTCTGGA
 AGAACAGGCAAGCGGCCAGGGCGCCGTGGTCACCGTCACC
 GGCCCGATCGCCTGCGCAAGACAGAACCTGCTTGACGCCGG 40
 CTGCCCGAAGGCTGACGCCATCATTCTGCGCGCGGTCTG
 CGCGCCCGAAGAGCGCGCTATGCCGTACGCCATGATCGGG
 CAGCTCATCGACGACCCGGCGCTCGCGCATCGGGCGCCGG
 AGCTGGCTGATCGGATAGCCCAGGGCGGGCATCTGTCGCT
 GAGGGCGAGAACCGACTGCGCAGGGATCTCACCCGTGCC
 CTGCTGGCGCTTGGCGTCAACCGGCCCTGCTGATCGGCCG
 TCGACGATGTGCATCACGCCGACACCGCCCTCTTGAACCTG
 TCTGCTGCATTAGCCGCCGGTCCGTCCGCCCTACTC
 TCCATGATCTCACCGAGTTGCGCAGCCTCACCCCTACTC
 AGTCACGATTCAAGGGGGAGCTGCTCAGCCTGCCGTACCCA 50

CCACGAGATCGCGCTGCGTCCACCTCGGACCGGAGCAATCG
GCGGAGCTGGCCCACGCCGCTTCGGCCCGGGCCCTCGCCG
AGGATGTGCTCGCGGGTTGTATGGGATGACCAAGGGGCAA
CCTGAGTCTCAGCCGTGGACTGATCAGCGATGTGCGGGAG
GCCCAAGGCCAACGGAGAGAGCGCTTTCGAGGTGGGCCGCG
CGTTCCGGCTGGCGTACCTCAGCTCGCTCTACCGCTGTGG
CCCGATCGCGCTGCGGGTCTGCCGAGTGGCTGCCGTGCTG
GGCCCAAGCGCCACCAACGCTGGTGCGCCGTCTAACCG
GGCTCAGCGGGAGACGATAAGACCGGGCAACCAAGATCCT
CACTGAGGGCGGGCTGCTGCTCGACCAACAGTTCCCGCAC 10
CCGGCCGCCCGCTCGGTGGTGCTCGATGACATGTCCGCC
AGGAACGACGCAGCCCTGCACACTCTCGCCCTGGAACTGCT
GGACGAGGCGCCGGTTGAAGTGCTCGCGCACCAACAGGTC
GGCGCCGGTCTCATACACGGGCCAAGGCTGCGGAGATAT
TCGCCAGGGCTGGCCAGGCTCTGGTTGTACGCAACGAGTT
GGCGACGCGGCCGAATACCTGCAACTGGCTCACCGAGCC
TCCGACGATGTCTCCACCCGGCCGCCCTAACGGTGCAGG
CCGTGGCAATCGAGCGCCGCCGAATCCGCTGGCCTCCAG
TCGTCACATGGACGAGCTGAGCGCCGCCGGCGCCGGT
CTGCTTCCCCAAGCATGCAAGCGCTGGCTGTCTTCTGGC 20
TGGCCGACGGCGGGCGATCCGGCGAGGCAGCCGAGGTGCT
GGCGTCGGAACACCCGCTCGCGACCAACCGATCAGAACCGA
GCACACCTCGGATTTGCCAGGGTGACTCTCGCGCTGTTCT
GTCCCCGGCGCCCTTCGGGTCTGGGACCGGGCGCCACCTCCGCT
GGCGCCGGACGAGCTGCCAGCTTGGCGAAGGGCGGCCCTGG
CAATGCGCGGTGCCGACAACAGCGGGTCATGACAGCGTTGC
ATGCTCATCCAGAACTTGGCACCGCTCAGGGGAAACAGT
TCTGCGGCAGGGCTGATTGGCAGCCGACGCAATCCCCGCC
GCACTGATCGCCCTGTTGTACGAGAGAACACCGAGTCCG
CTCAGATCTGGGCCGACAAGCTGGCGAGCACCAATGCCGG 30
GGTATCGAACGAGGCAGGAAAGCGGGCTACGCCGGCCCGTGC
GCCGAGATCGCCCTGCGGCGGGGACCTGGCCACCGGGGTGC
TCGAGGGCTGGTGGCACCGTCTGGGACGACCGGGCGCTGCC
GTCGCTCGGCATCACCGCCGCATTGCTGTTGAGCAGCAAG
ACGGCAGCCGCTGTCGCCCTGGCGAACACTCGAGCGTGC
AGAACGCTGCTGCTGAGGCCGTTCCGAACGGTGTCCAGGA
CAGCCTTTCGGTCTGCACCTGCTCTCGGCCGACGGCCAG
TACAGCCTCGCGATGGGCCGATATGAATCGGCTCACCGGG
CGTTTCAACCTCGGGAGAACGTATGCGCAGCTGGGTGT
TGACGTGCCCTGGTCTAGCCCTGTTGGCGTGTGACGCCGCC 40
GAGGCAC TGCTCAGCCTCGACCGGAAACGAGGGGCCAGCGGGC
TCATCGACGAACAAACTCGCCCGTCCGATGGGACCTCGTT
CCGCGCATTAAACGCTCGGGATCAAGGCCGACATACCTCCG
CGGACGAAGCGGGATCCCCCTGCTCCATGAGGCCAGCTGAGC
TGCTGCTCTCCCTGCCCGACCCGTACGAGCAAGCGCGGGT
GCTCGCCGATCTGGCGGACACGCTCAGCGCGCTCAGACGC
TATAGCCGGGCGCGGGGAGTTCTCCGGCAGGGCTCGTCACC
TGGCCACCCAGTGCGGTGCTGTCCTCGCTGCTGCCGACT
CGGGGGCGAGCCCCGGCGGATCGACGACGCCGGCCTGCCG
CAGCGGAGCACATGTTGACCGATGCGGGAGCGGGCGGGTGT 50

CGGCCCTGGCCGCGGCCGGACAGACCAACCGGGAGATCGC
 CAAACAGCTATTGTCACGGCCAGCACCGTGGAACAGCAC
 CTACACAAGCGTCTTCCGCAAGCTGGCGTTAAGGGCCGCA
 GGCAAGCTACCGACCGCGCTGGCCGACGTGGAAATAG

配列番号 58

ATGTATAGCGGTACCTGCCGTGAAGGATACGAACCTCGTCG
 CCCGCGAGGACGAACTCGGCATTCTGCAAGAGGTCTCTGGA
 AGAACAGCAGGCAGCGGCCAGGGCGCCGTGGTCACCGTCACC
 GGCCCGATCGCCTGCGGCAAGACAGAACCTGCTTGACGCGG
 CTGCCGCGAAGGCTGACGCCATCATTCTGCGCGCGGTCTG 10
 CGCGCCCGAAGAGAGCGCGCTATGCCGTACGCCATGATCGGG
 CAGCTCATCGACGACCCGGCGCTCGCGCATCGGGCGCCGG
 AGCTGGCTGATCGGATAGCCCAGGGCGGGCATCTGTCGCT
 GAGGGCGAGAACCGACTGCGCAGGGATCTCACCCGTGCC
 CTGCTGGCGCTTGCCGTCGACCGGGCTGTGCTGATCGGCG
 TCGACGATGTGCATCACGCCGACACCGCCTCTTGAAC 20
 TCTGCTGCATCTGGCCCGCCGGTCCGTCGCCGCGGATA
 TCCATGATCTTCACCGAGTTGCGCAGCCTCACCCCTACTC
 AGTCACGATTCAAGGGAGCTGCTCAGCCTGCCGTACCA
 CCACGAGATCGCGCTGCGTCCACTCGGACCGGGAGCAATCG
 CGGGAGCTGGCCCACGCCGCTTCGGCCCGGGCTCGCCG
 AGGATGTGCTCGCGGGTTGTATGGGATGACCAAGGGCAA 30
 CCTGAGTCTCAGCCGTGGACTGATCAGCGATGTTGCGGGAG
 GCGCCAGGCCAACCGGAGAGAGCGCTTTCGAGGTGGGCCGCG
 CGTTCCGGCTGGCGTACCTCAGCTCGCTCTACCGCTGTGG
 CCCGATCGCGCTGCGGGTCGCCCGAGTGGCTGCCGTGCTG
 GGCCCAAGCGCCACCAACCGCTGGTGCAGCCGTC 40
 GGCTCAGCGCGAGACGATAGACCGGGCAACCAAGATCCT
 CACTGAGGGCGGGCTGCTGCTCGACCAACAGTTCCCGCAC
 CCGGCCGCCGCTCGGTGGTGCATGACATGTCGCCGCCC
 AGGAACGACGCAGCCTGCAACACTCTGCCCTGGAACTGCT
 GGACGAGGCGCCGGTTGAAGTGCTCGCGCACCAAGGTC
 GGCGCCGGTCTCATACACGGGCCAACGGCTGCGGGAGATA
 TCGCCAGGGCTGGCCAGGGCTCTGGTTACGCAACGAGTT
 GGGCGACGCGGCCGAATACCTGCAACTGGCTCACCAGGCC
 TCGACGATGTCCTCCACCCGGCCGCCCTGCGGGTCGAGG
 CGTGGCAATCGAGCGCCGCCGCAATCCGCTGGCCTCCAG
 TCGTCACATGGACGAGCTGAGCGCCGCCGGCCGCGCCGGT
 CTGCTTCCCCAAGCATGCAAGCGCTGGCTGTC 50
 TGCGCGACGGCGGGCGATCCGGCGAGGCAGCCGAGGTGCT
 GCGTCGGAACACCCGCTCGCGACCAACCGATCAGAACCGA
 GCACACCTGCGATTGCCGAGGTGACTCTCGCGCTGTTCT
 GTCCCCGGCGCCCTTCGGGTCGGACCGGGGCCACCTCCGCT
 GGCGCCGGACGAGCTGCCAGCTTGCCTGAAGGGCGGCCCTGG
 CAATGCGCGGTGCGCCGACAACCGCGGTGATGACAGCGTTGC
 ATGCTCATCCAGAACCTGCCACCGCTCAGGGGGAAACAGT
 TCTGCGGCAGGGCTGATTGCCAGCCGACGCAATCCCCGCC
 GCACTGATCGCCCTGTTGTCAGCAGAGAACACCGAGTCCG
 CTCAGATCTGGCCGACAAGCTGGGCAGCACCAATGCCGG
 GGTATCGAACCGAGGGCGGAAGCGGGCTACGCCGGCCGTG 50

G C C G A G A T C G C C C T G C G G C G C G G C G A C C T G G C C A C G G C G T
 T C G A G G C T G G T G G C A C C G T C C T G G A C G A C C G G C C G C T G C C
 G T C G C T C G G C A T C A C C G C C G C A T T G C T G T T G A G C A G C A A G
 A C G G C A G C C G C T G T C C G C C T G G G C G A A C T C G A G C G T G C G G
 A G A A G C T G C T C G C T G A G C C G C T T C C G A A C G G T G T C C A G G A
 C A G C C T T T C G G T C T G C A C C T G C T C T C G G C G C A C G G C C A G
 T A C A G C C T C G C G A T G G G C C G A T A T G A A T C G G C T C A C C G G G
 C G T T T C A C A C C T G C G G A G A A C G T A T G C G C A G C T G G G G T G T
 T G A C G T G C C T G G T C T A G C C C T G T G G C G T G T C G A C G C C G C C
 G A G G C A C T G C T C A G C C T G C A C C G G A A C G A G G G C C A G C G G C 10
 T C A T C G A C G A A C A A C T C G C C C G T C C G A T G G G A C C T C G T T C
 C C G C G C A C T G A C G C T G C G G A T C A A G G C G G C A T A C C T C C C G
 C G G A C G A A G C G G A T C C C C C T G C T C C A T G A G G C A G C T G A G C
 T G C T G C T C T C C T G C C C C G A C C C G T A C G A G C A A G C G C G G G T
 G C T C G C C G A T C T G G G C G A C A C G C T C A G C G C G C T C A G A C G C
 T A T A G C C G G G C G C G G G G A G T T C T C C G G C A G G C T C G T C A C C
 T G G C C A C C C A G T G C G G T G C T G T C C C G C T G C T G C G C C G A C T
 C G G G G G C G A G C C C G G C G G A T C G A C G A C G C C G G C T G C C G 20
 C A G C G G A G C A C A T C G T T G A C C G A T G C G G A G C G G C G G G T G T
 C G G C C C T G G C C G C G G C G G A C A G A C C A A C C G G G A G A T C G C
 C A A A C A G C T A T T C G T C A C G G C C A G C A C C G T G G A A C A G C A C
 C T C A C A A G C G T C T T C C G C A A G C T G G G C G T T A A G G G C C G C A
 G G C A G C T A C C G A C C G C G C T G G C C G A C G T G G A A T A G

配列番号 59

G T G T A T A G C G G T A C C T G C C G T G A A G G A T A C G A A C T C G T C G
 C A C G C G A G G A C G A A C T C G G C A T T C T A C A G A G G T C T C T G G A
 A C A A G C G A G C A G C G G C C A G G G C G T C G T G G T C A C C G T C A C C
 G G C C C A A T C G C C T G C G G C A A G A C A G A A C T G C T T G A C G C G G
 C T G C C G C G A A A G G C T G A G G G C C A T C A T T C T G C G C G C G G T C T G
 C G C G C C C G A A A G A G C G G G C T A T G C C G T A C G C C A T G A T C G G G 30
 C A G C T C A T C G A C G A C C C G G C G C T C G C G C A T C G G G C G C C G G
 G G C T G G C T G A T C G G A T A G C C C A G G G C G G G C A G C T G T C G C T
 G A G G G C C G A G A A C C G A C T G C G C G A G G G A T C T C A C C C G T G C C
 C T G C T G G C G C T T G C C G T G C A C C G G C C T G T G C T G A T C G G C G
 T C G A T G A T G T G C A T C A C G C C G A C A C C G C C T C T T G A A C T G
 T C T G C T G C A T T T G G C G C G C G G G T C C G T C C G G C C C G G A T A 40
 T C C A T G A T C T T C A C C G A G T T G C G C A G C C T C A C C C C T A C T C
 A G T C A C G A T T C A A G G C G G A G G C T G C T C A G C C T G C C G T A C C A
 C C A C G A G A T C G C G C T G C G T C C A T T C G G A C C G G A G C A A T C G
 G C G G A G G C T G G C T C G C G C C G C C T T C G G C C C G G G C C T C G C C G
 A G G A T G T G C T C G C G G G G T T G T A T A A A A C G A C C A G G G G C A A 50
 T C T G A G T C T C A G C C G T G G A C T G A T C A G C G A T G T G C G G G A G
 G C C C T G G C C A A C G G A G A G A G G C G C T T T C G A G G G C G G G C C G C G
 C G T T C C G G C T G G C G T A C C C T C A G C T C G C T C T A C C G C T G T G G
 C C C G G T C G C G C T G C G G G G T C G C G C C C G A G T G G C T G C C G T G C T G
 G G C C C A A G C G C C A C C A C C A C G C T G G T G C G C C G G C T A A G C G
 G G C T C A G C G C G G A G A C G A T A G A C C G G G C A A C C A A G A T C C T
 C A C C G A G G G C G G G C T G C T G C T C G A C C A G C A G T T T C C G C A C
 C C G G C C G C C C G C T C G G T G G T G C T C G A T G A C A T G T C C G C C C
 A G G A A C G A C G C G G C C T G C A C A C T C T C G C C C T G G A A C T G C T

GGACGAGGCGCCGGTTGAAGTGCTCGCGCACCAACCAGGTC
 GGCGCCGGTCTCATACACAGGGCCCAAGGGCTGCGGAGATGT
 TCGCCAAGGCCGGCAAGGGCTCTGGTCGTACGCAACGAGTT
 GGGCGACGCGGCCGAATACCTGCAACTGGCTCACCGGGCC
 TCCGACGATGTCTCCACCCGGGCCGCTTACGGGTCGAGG
 CGTGGCGATCGAGCGCCGCCGCAATCCGCTGGCCTCCAG
 TCGGCACATGGACGAGCTGAGCGCCGCCGGCGCCGGGT
 CTGCTTCCCCAAGCATGCGGGCGCTGGCCGTCTTCTGGC
 TGGCCGACGGCGGGCGATCCGGCGAGGCAGCCCAGGTGCT
 GCGTCGGAACGCCGCTCGCGACCAACCGATCAGAACCGG
 GCCCACCTGCGATTTGTGAGGGTGACTCTCGCGCTGTTCT
 CTCCC GGCGCCCTTCGGATCGGACCCGGCGCCACCTCCGCT
 GACGCCGGACGAACCTGCCAGCCTGCCGAAGGGCGGCTGG
 CAATGCGCGTCGCCGACAACGCCGCGCATGACCGCCCTTGC
 ACGGCCATCCAGAACCTGCCACCGCTCAGGGGAAACAGT
 TCTGCGGCAGGGCTGATTGGCAGCCGACGCGATCCCCGCC
 CGGCTGATGCCCTGTTGTACGCCGGAGAACACCGAGTCCG
 CTCATATCTGGGCCGACAAGCTGGCAGCAGCATGAATGCCGG
 GGTATCGAACGAGGGCGAACGGGGCTACGCCGGCGTGC
 GCCGAGATGCCCTGCGGCCGGGACCTGGCCACGGCGT
 TCGAGGGCTGGTAGCACCGTCTGGACGACCGGTCACTGCC
 GTCGCTCGGCATCACCGCCGCATTGCTGTTGAGCAGCAAG
 ACGGCCGCCGCTGTCCGGCTGGCGAACACTCGAGCGTGC
 AGAACGCTGCTCGCCGAGCCGCTTCCGAACGGCGTCCAGGA
 CAGCCTTTCGGTCTGCACCTGCTCTCGCGTACGGCCAG
 TACAGCCTCGCGATGGGCCGATATGAATCGGCTCACCGGG
 CGTTTCGCACCTGCGGAGAACGTATGCGCAGCTGGGATGT
 TGACGTGCCTGGTCTGGCCCTGTTGGCGTGTGACGCCGCC
 GAGGGCGCTGCTCAGCCTCGACCGGAAACGAGGGGCCAGCG
 TCATCGACGAAACAACCTCACCCGTCCTGATGGGACCTCGTT
 CGCGCGGTTAACGCTGCGGATCAAGGGGGCATACCTCCCG
 CGGACGAAGCGGATCCCCCTGCTCCCATGAGGGCGGCCAGC
 TGCTGCTCCCCCTGCCCGACCCGTAACGAGCAAGCGCGGGT
 GCTCGCCGATCTGGGCGACACGCTCAGCGCGCTCAGACGC
 TATAAGCCGGCGCGGGGAGTTCTCCGGCAGGGCTCGTCACC
 TGGCCACCCAGTGCGGTGCTGTCCTCGCTGCGGCCGACT
 CGGGGGCGAGCCCCGGCGGATCGACGACGCCGGCTGCCG
 CAGCGGAGCACATCGTTGACCGATGCGGAGCGGGGGTGG
 CGCGCTGGCGCGGCCGGACAGACCAACCGGGAGATCGC
 CGAACAGCTGTTGTCACGGCCAGCACAGTGGAAACAGCAC
 CTCACAAAGCGTCTTCCGCAAGCTGGGCGTCAAGGGCCGCA
 AGCAGCTGCCGACCGCGCTGGCCGACGTGGAAACAGACCTG
 A

配列番号 60

ATGTATAGCGGTACCTGCCGTGAAGGGATACGAACTCGTCG
 CACGCGAGGACGAACCTCGGCATTCTACAGAGGGTCTCTGG
 ACAAGCGAGCAGCGGCCAGGGCGTCGTGGTCACCGTCACC
 GGCCCAATCGCCTGCGCAAGACAGAACACTGCTTGACGCG
 CTGCCGCGAAGGGCTGAGGCCATCATTCTGCGCGCGGTCTG
 CGCGCCCGAACGAGCGGGCTATGCCGTACGCCATGATCGGG
 10

20

30

40

50

C A G C T C A T C G A C G A C C C G G C G C T C G C G C A T C G G G C G C C G G
 G G C T G G C T G A T C G G A T A G C C C A G G G G C G G G C A G C T G T C G C T
 G A G G G C C G A G A A C C G A C T G C G C A G G G A T C T C A C C C G T G C C
 C T G C T G G C G C T T G C C G T G C A C C G G C C T G T G C T G A T C G G C G
 T C G A T G A T G T G C A T C A C G C C G A C A C C G C C T C T T G A A C T G
 T C T G C T G C A T T T G G C G C G C G G G T C C G T C C G G C C C G G A T A
 T C C A T G A T C T T C A C C G A G T T G C G C A G C C T C A C C C C T A C T C
 A G T C A C G A T T C A A G G C G G A G G C T G C T C A G C C T G C C G T A C C A
 C C A C G A G A T C G C G C T G C G T C C A T T C G G A C C G G A G C A A T C G
 G C G G A G C T G G C T C G C G C C G C C T T C G G C C C G G G C C T G C C G
 10 A G G A T G T G C T C G C G G G G T T G T A T A A A A C G A C C A G G G G C A A
 T C T G A G T C T C A G C C G T G G A C T G A T C A G C G A T G T G C G G G A G
 G C C C T G G C C A A C G G A G A G A G G C G C T T T C G A G G G C G G G C C G C G
 C G T T C C G G C T G G C G T A C C T C A G C T C G C T C T A C C G C T G T G G
 C C C G G T C G C G C T G C G G G T C G C C C G A G T G G C T G C C G T G C T G
 G G C C C A A G C G C C A C C A C C A C G C T G G T G C G C C G G C T A A G C G
 G G C T C A G C G C G G A G A C G A T A G A C C G G G C A A C C A A G A T C C T
 C A C C G A G G G C G G G C T G C T G C T C G A C C A G C A G T T T C C G C A C
 C C G G C C G C C G C T C G G T G G T G C T C G A T G A C A T G T C C G C C C
 A G G A A C G A C G C G G G C C T G C A C A C T C T C G C C C T T G G A A C T G C T
 G G A C G A G G G C G C C G G T T G A A G T G C T C G C G C A C C A C C A G G T C
 G G C G C C G G T C T C A T A C A C G G G C C C A A G G C T G C G G A G A T G T
 T C G C C A A G G C C G G C A A G G C T C T G G T C G T A C G C A A C G A G T T
 G G G C G A C G C G G C C G A A T A C C T G C A A C T G G C T C A C C G G G C C
 T C C G A C G A T G T C T C C A C C C G G G C C G C C T G C G G G T C G A G G
 C C G T G G C G A T C G A G C G C C G C G C A A T C C G C T G G C C T C C A G
 T C G G C A C A T G G A C G A G C T G A G C G C C G C G C G C G C C G G T
 C T G C T T T C C C C C A A G C A T G C G G C G C T G G C C G T C T T C T G G C
 T G G C C G A C G G C G G G C G A T C C G G C G A G G C A G C C C A G G T G C T
 G G C G T C G G A A C G C C C G C T C G C G A C C A C C G A T C A G A A C C G G
 30 G C C C A C C T G C G A T T T G T C G A G G T G A C T C T C G C G C T G T T C T
 C T C C C G G C G C C T T C G G A T C G G A C C G G C G C C C A C C T C C G C T
 G A C G C C G G A C G A A C T C G C C A G C C T G C C G A A G G G C G G G C T G G
 C A A T G C G G C G G T C G C C G A C A A C G C G G C C A T G A C C G C C T T G C
 A C G G C C A T C C A G A A C T T G C C A C C G C T C A G G C G G A A A C A G T
 T C T G C G G C A G G C T G A T T C G G C A G C C G A C G C G A T C C C C G C C
 G C G C T G A T C G C C C T G T T G T A C G C G G A G A A C A C C G A G T C C G
 C T C A T A T C T G G G C C G A C A A G C T G G G C A G C A T G A A T G C C G G
 G G T A T C G A A C G A G G G C G G A A G C G G G C T A C G C C G G C C G T G C
 G C C G A G A T C G C C C T G C G G C G G G C G A C C T G G C C A C G G C G T
 40 T C G A G G G C T G G T A G C A C C G T C C T G G A C G A C C G G G T C A C T G C C
 G T C G C T C G G C A T C A C C G C C G C A T T G C T G T T G A G C A G C A A G
 A C G G C C G C C G C T G T C C G G C T G G G C G A A C T C G A G C G G T G C G G
 A G A A G C T G C T C G C C G A G C C G C T T C C G A A C G G C G T C C A G G A
 C A G C C T T T C G G T C T G C A C C T G C T C T C G G C G T A C G G C C A G
 T A C A G C C T C G C G A T G G G C C G A T A T G A A T C G G C T C A C C G G G
 C G T T T C G C A C C T G C G G A G A A C G T A T G C G C A G C T G G G A T G T
 T G A C G T G C C T G G T C T G G C C C T G T G G G C G T G T C G A C G C C G C C
 G A G G C G C T G C T C A G C C T C G A C C G G A A C G A G G G G C C A G C G G C
 T C A T C G A C G A A C A A C T C A C C C G T C C G A T G G G A C C T C G T T C
 50

CCGGCGCGCTGACGCTGCGGATCAAGGCGGCATACCTCCCG
 CGGACGAAGCGGATCCCCCTGCTCCATGAGGCCGAGC
 TGCTGCTCCCCCTGCCCGACCCGTACGAGCAAGCGCGGGT
 GCTCGCCGATCTGGCGACACGCTCAGCGCGCTCAGACGC
 TATAAGCCGGGCGCGGGGAGTTCTCCGGCAGGCTCGTCACC
 TGGCCACCCAGTGCGGTGCTGTCGGCTGCTGCGCCGACT
 CGGGGGCGAGCCCCGGCGGATCGACGACGCCGGCTGCCG
 CAGCGGAGCACATCGTTGACCGATGCGGAGCGGGGGTGG
 CGCGCTGGCGCGGGACAGACCAACCGGGAGATCGC
 CGAACAGCTGTTGTCACGGCCAGCACAGTGGAACAGCAC
 CTCACAAGCGTCTTCCGCAAGGCTGGCGTCAAGGGCCGCA
 AGCAGCTGCCGACCGCGCTGGCCGACGTGGAACAGACCTG
 A

配列番号 6 1

GTGCGAGCTATTAATGCGTCCGACACCGGTCCCTGAACCTGG
 TCGCCCGCGAAGACGAACCTGGGACGTGTACGAAGTGCCCT
 GAACCGAGCGAACCGGGGCCAAGGTGTCCCTGATCTCCATT
 ACCGGTCCGATCGCCCTGCGGCAAGACCGAACCTGCTTGAGG
 CTGCCGCCCTCGGAAGTTGACGCCATCACTCTGCGCGGGT
 CTGTGCCGCCGAGGAACGGGCGATACCTTATGCCCTGATC
 20 GGGCAGCTTATCGACAACCCCGCGCTCGGCATTCCGGTT
 CGGATCCGGCCGGCTGACCGCCAGGGCGGACGACTGTC
 ATCGAGCGCCGAGAACCGACCTGCGTCGCGACCTCACCGT
 GCCCTGCTGACGCTCGCCACCGAACCGGCTGGTGCTGATCT
 GTGTCGATGACGTGCAGCACGCCGACAACGCCCTCGTTGAG
 CTGCCCTCTGTATCTGGCCCGACGGCTTGTCCCGGGCTCGA
 ATCGCTCTGGTATTCAACCGAGTTGCGAGTCCTCACCTCGT
 CTCAGTTACGGTTCAACCGCGAGCTGCTCAGCTTGC
 30 CCACACTGCGAGATCGCGCTGCGGCCACTCGGCCCGGGGCA
 CGGGCCGAGCTGGCCCGCGCCACCCCTCGGCCCGGCCCT
 CCGACGAAACACTCACGGAGCTGTACCGGGTGACCGGGAGG
 CAACCTGAGTCTCAGCGCGGGCTGATCGACGATGTC
 GACGCCCTGGGCACGAGGGAAACGGGCGTCCAGGTGGGCC
 GGGCGTTCCGGCTGGCCTACCTCGGTTCCACCGCTG
 TGGTCCGCTGGCGTTGCGGGTCGCCCGCGTAGCCGCCGTA
 CTGGGCCCGAGCGCCACCAAGCGTCCCTGGTGC
 40 GTGGGCTCAGCGCGGGAGGCCATGGCCCAGGGGACCGA
 CCTCGCTGACGGCGGCCCTCCCTGCGCGACCAAGCGGTT
 CATCCAGCGGCCCGCTCGGTGGTGCTCGACGACATGT
 CCGAGGAACGACGCAGCGTGCACAGCCCTCGGCCCTGG
 GCTGGACGAGGCACCGGCCGAGATGCTCGCGCACCA
 45 GTCGGCGCCGGTCTCGTGCACGGGCCGAAGGCCG
 CATTCAACGGGGCCGGCCGGGCACTGGCCGTT
 GCTGGCGAGGCAGCCGACTACCTGCAACTGGCGTACCG
 GCCTCCGGCGACGCCGCTACCAAGGCCGCGATAC
 AGTCCGTGGCGGTGAGCGCCGACGCAATCC
 50 CCGTGGTGCAGGACGAGCTGAGCGTGC
 CAGTCGCCATTGGGACGAGCTGAGCGTGC
 GGTCTGCTCTCCCTGCGAGCACGTGT
 GGCTGACCGTCTGGTGGCGGGCCCGGG
 GCTGGCGTCGCAACACCGACGGGTC
 GTCACCGGATCAGGAC

CGGGCCCACCTGC GG GT CGCCGAGTTCTCGCTCGCGCTGC
 TGTACCCGGTACGTCGGCTCGGACCGGGCGCCGCACCC
 GCTCACGTCGGACGAACCTCGCGGCCCTACCGACTGCGACC
 AGACACTGCGCGATCGCCGATAACGCTGTCATGGCTGCCT
 TCGGTGGTCATCCGGAGCTTGCACCGCCGAGGGCAGAAC
 CGTTCTGCAGCAAGCCGACGCGGGCGACGGCGCTGCTCTC
 ACCGGCGCTGATGGCCCTGCTGTACGCGGGAGAGCATTGAGG
 TCGCTGAAGTCTGGGCGGACAAAGCTGGCGGCAGAGGCCGG
 AGCATTGAAACGGGCAGGACGCGGGAGTACGCCGGTATAACGC
 GCGGAAATCGCCCTGCGGCGGGGATCTGACCGCGGCCG
 10 TCGAGACCGGCCGGCATGGTCCTGGGACGGGCCGGCGCTGCC
 GTCGCTCGACATCACCGGCCACGTTGCTGTTGGGCCGGCAGG
 GCGTCCGTCGCCGTCCGGCTGGCGAACCTGACCAACGCCGG
 AGGAGCTGTTGCCCGCGGCCGGAGGGACGCCCTTCCAGGA
 CAGCCTCTTCGGTCTGCATCTGCTCTCGGCGCACGGCCAG
 TACAGCCTCGCGACAGGCCGGCGAGTCGGCATACCGGG
 CCTTTCGTGCCCTGCGGCGAACGTATGCGCGATTGGGGCTT
 CGACGCGCCCGGTGTGGCCCTGTTGGCGCGTCGGCGCCGCC
 GAGGCGCTGCTCGGCCCTGACCGGGAACGAGGGGCCGACGGC
 TCATCGACGAAACAGCTGAGCCGGACGATGGCCCCCGGTC
 20 CCACGCGTTGACGCTGCGGATAAAAGCGGGCGTACATGCCG
 GAGCCGAAGCGGGTCGACCTGCTCTACGAAGCGGGCTGAGC
 TGCTGCTCTCCCTGCCGGGACCCAGTATGAGCGAGCGCGGGT
 GCTCGCCGATCTGGGCGAGGGCGTCAGCGCGCTCGGGAAC
 TACCGGGCAGGGCGCGAGGTGTGCTCCGGCAGGGCTCGGCATC
 TGGCCATGCGAACCGGGCGGGACCCGCTGCTGCGCCGGCT
 CGGAATCAGGCCCGGCCGGCAGGACGACCCCCGACCCGCA
 CC CGGGAGCAGATCGCTGACCAACGCTGAGCGGGCGTGC
 CGTCGCTGGCCGCGACCGGACTGACCAACCGGGAGATCGC
 CGACCGGGCTTTCGTCACCGCCAGCACCGTGGAGCAGCAC
 30 CTCACCAACGTTCTCCGCAAGCTGGCGTCAAGGGCCGCA
 AGCAGCTGCCGGCGAGTTGGACGACATGGAATAG

配列番号 62

ATGCGAGCTATTAATGCGTC CGACACCGGTCCCTGAACCTGG
 TCGCCCGCGAAAGACGAACCTGGGACGTTGACGAAAGTGCCT
 GAACCGAGCGAACCGGGGCCAAGGGTGTCCCTGATCTCCATT
 ACCGGTCCGATCGCCCTGCGGCAAGACCGAACCTGCTTGAGG
 CTGCCGCCCTCGGAAGTTGACGCCATCACTCTGCGCGGGT
 CTGTGCCGCCAGGGAACCGGGGAGTACCTTATGCCCTGATC
 GGGCAGCTTATCGACAAACCCCGCGCTCGGCATTCCGGTT
 40 CGGATCCGGCCGGCCTGACCGCCCAAGGGCGGACGACTGTC
 ATCGAGCGCCGAGAACCGACTGCGTCGCGACCTCACCGT
 GCCCTGCTGACGCTCGCCACCGAACCGGGCTGGTGCCTGATC
 GTGTCGATGACGTTGACGCCAGCACGCGAACACGCCCTCGTT
 CTGCCTTCTGTATCTGGCCCGACGGCTTGTCCCGGGCTCGA
 ATCGCTCTGGTATTCACCGAGTTGCGAGTCCTCACCTCGT
 CTCACTGCGAGATCGCGCTGCGCCACTCGGGCCCGGGCAT
 CGGGCCGAGCTGGCCCGCGCCACCCCTCGGGCCCGGGCTCT
 CCGACGAAACACTCACGGAGCTGTACCGGGTGAACCGGAGG
 50

C A A C C T G A G T C T C A G C C G C G G G C T G A T C G A C G A T G T G C G G
 G A C G C C T G G G C A C G A G G G G A A A C G G G C G T C C A G G T G G G C C
 G G G C G T T C C G G C T G G C C T A C C T C G G T T C C C T C C A C C G C T G
 T G G T C C G C T G G C G T T G C G G G T C G C C C G C G T A G C C G C C G T A
 C T G G G C C C G A G C G C C A C C A G C G T C C T G G T G C G C C G A T C A
 G T G G G C T C A G C G C G G A G G C C A T G G C C C A G G C G A C C G A T A T
 C C T C G C T G A C G G C G G C C T C C T G C G C G A C C A G C G G T T C A C A
 C A T C C A G C G G C C C G C T C G G T G G T G C T C G A C G A C A T G T C C G
 C C G A G G A A C G A C G C A G C G T G C A C A G C C T C G C C C T G G A A C T
 G C T G G A C G A G G C A C C G G C G A G A T G C T C G C G C A C C A C C G G 10
 G T C G G C G C C G G T C T C G T G C A C G G G C G A A G G C C G G G A G A
 C A T T C A C C G G G G C C G G C C G G G C A C T G G C C G T T C G C G G C A T
 G C T G G G C G A G G C A G C C G A C T A C C T G C A A C T G G C G T A C C G G
 G C C T C C G G C G A C G C C G C T A C C A A G G C C G C G A T A C G C G T C G
 A G T C C G T G G C G G T C G A G C G C C G A C G C A A T C C G C T G G T C G T
 C A G T C G C C A T T G G G A C G A G C T G A G C G T C G C G G C C C G G C C
 G G T C T G C T C T C C T G C G A G C A C G T G T C C A G G A C G G G C C C G C T
 G G C T G A C C G T C G G T G G G C G G C C C G G C G A G G C G G C C A G G G T
 G C T G G C G T C G C A A C A C C G A C G G G T C G T C A C C G A T C A G G A C 20
 C G G G C C C A C C T G C G G G T C G C C G A G T T C T C G C T C G C G C T G C
 T G T A C C C C G G T A C G T C C G G C T C G G A C C G G C C C G C A C C C
 G C T C A C G T C G G A C G A A C T C G C G G C C T A C C G A C T G C G A C C
 A G A C A C T G C G C G A T C G C C G A T A A C G C T G T C A T G G C T G C C T
 T G C G T G G T C A T C C G G A G C T T G C C A C C G C C G A G G C A G A A G C
 C G T T C T G C A G C A A G C C G A C G C G G G A C G G G C G T G C T C T C
 A C C G C G C T G A T G G C C C T G C T G T A C G C G G A G A G C A T C G A G G
 T C G C T G A A G T C T G G G C G G A C A A G C T G G C G G C A G A G G C C G G
 A G C A T C G A A C G G G C A G G A C G C G G A G T A C G C C G G T A T A C G C
 G C C G A A A T C G C C C T G C G G C G G C G A T C T G A C C G C G G C C G 30
 T C G A G A C C G C C G G C A T G G T C C T G G A C G G G C C G G C G T G C C
 G T C G C T C G A C A T C A C C G C C A C G T T G C T G T T G G C C G G C A G G
 G C G T C C G T C G C C G T C C G G C T G G G C G A A C T C G A C C A C G C G G
 A G G A G C T G T T C G C C G G C G C C G G A G G A C G C C T T C C A G G A
 C A G C C T C T C G G T C T G C A T C T G C T C T C G G C G C A C G G C C A G
 T A C A G C C T C G C G A C A G G C C G G C C G A G T C G G C A T A C C G G G
 C C T T T C G T G C C T G C G G C G A A C G T A T G C G C G A T T G G G G C T T 40
 C G A C G C G C C C G G T G T G G C C C T G T G G C G C G T C G G C G C C G C C
 G A G G C G C T G C T C G G C C T C G A C C G G A A C G A G G G G C C G A C G G C
 T C A T C G A C G A A C A G C T G A G C C G G A C G A T G G C C C C C G G T C
 C C A C G C G T T G A C G C T G C G G A T A A A A G C G G G C G T A C A T G C C G
 G A G C C G A A G C G G G T C G A C C T G C T C T A C G A A G C G G G C T G A G C
 T G C T G C T C T C C T G C C G G G A C C A G T A T G A G C G A G C G G C G G G T
 G C T C G C C G A T C T G G G C G A G G G G C T C A G C G C G C T C G G G A A C
 T A C C G G C A G G C G C G A G G T G T G C T C C G G C A G G G C T C G G C A T C
 T G G C C A T G C G A A C C G G C G G A C C C G C T G C T G C G C C G G C T
 C G G A A T C A G G C C C G G C G G C A G G A C G A C C C C G A C C C G C A G
 C C G C G G A G C A G A T C G C T G A C C A A C G C T G A G C G G G C G T G C G G
 C G T C G C T G G C C G C G A C C G G A C T G A C C A A C C G G G A G A T C G C
 C G A C C G G C T C T T C G T C A C C G C C A G C A C C C G T G G A G C A G C A C
 C T C A C C A A C G T C T T C C G C A A G C T G G G C G T C A A G G G C C G C A 50

A G C A G C T G C C G G C C G A G T T G G A C G A C A T G G A A T A G

配列番号 6 3

M P A V E C Y E L D A R D D E L R K L E E V V T G R A N G R G V V V T I T G P I
 A C G K T E L L D A A A A K A D A I T L R A V C S A E E Q A L P Y A L I G Q L I
 D N P A L A S H A L E P A C P T L P G E H L S P E A E N R L R S D L T R T L L A
 L A A E R P V L I G I D E S H A N A L C L L H L A R R V G S A R I A M V L T E L
 R R L T P A H S Q F Q A E L L S L G H H R E I A L R P L S P K H T A E L V R A G
 L G P D V D E D V L T G L Y R A T G G N L N L T R G L I N D V R E A W E T G G T
 G I S A G R A Y R L A Y L G S L Y R C G P V P L R V A R V A A V L G Q S A N T T
 L V R W I S G L N A D A V G E A T E I L T E G G L L H D L R F P H P A A R S V V
 L N D M S A Q E R R R L H R S A L E V L D D V P V E V V V A H H Q V G A G L L H G 10
 P K A A E I F A K A G Q E L H V R G E L D T A S D Y L Q L A H Q A S D D A V T G
 M R A E A V A I E R R R N P L A S S R H L D E L T V V A R A G L L F P E H T A L
 M I R W L G V G G R S G E A A G L L A S Q R P R A V T D Q D R A H M R A A E V S
 L A L V S P G T S G P D R R P R P L T P D E L A N L P K A A R L C A I A D N A V
 M S A L R G R P E L A A A E A E N V L Q H A D S A A A G T T A L A A L T A L L Y
 A E N T D T A Q L W A D K L V S E T G A S N E E E A G Y A G P R A E A A L R R G
 D L A A A V E A G S T V L D H R R L S T L G I T A A L P L S S A V A A A I R L G
 E T E R A E K W L A Q P L P Q A I Q D G L F G L H L L S A R G Q Y S L A T G Q H
 E S A Y T A F R T C G E R M R N W G V D V P G L S L W R V D A A E A L L H G R D
 R D E G R R L V D E Q L T R A M G P R S R A L T L R V Q A A Y S P P A K R V D L
 L D E A A D L L L S C N D Q Y E R A R V L A D L S E T F S A L R H H S R A R G L 20
 L R Q A R H L A A Q R G A I P L L R R L G A K P G G P G W L E E S G L P Q R I K
 S L T D A E R R V A S L A A G G Q T N R V I A D Q L F V T A S T V E Q H L T D V
 S T G S R P P A P A A E L V

配列番号 6 4

M V P E V R A A P D E L I A R D D E L S R L Q R A L T R A G S G R G G V V A I T
 G P I A S G K T A L L D A G A A K S G F V A L R A V C S W E E R T L P Y G M L G
 Q L F D H P E L A A Q A P D L A H F T A S C E S P Q A G T D N R L R A E F T R T
 L L A L A A D W P V L I G I D D V H H A D A E S L R C L L H L A R R I G P A R I 30
 A V V L T E L R R P T P A D S R F Q A E L L S L R S Y Q E I A L R P L T E A Q T
 G E L V R R H L G A E T H E D V S A D T F R A T G G N L L L G H G L I N D I R E
 A R T A G R P G V V A G R A Y R L A Y L S S L Y R C G P S A L R V A R A S A V L
 G A S A E A V L V Q R M T G L N K D A V E Q V Y E Q L N E G R L L Q G E R F P H
 P A A R S I V L D D L S A L E R R N L H E S A L E L L R D H G V A G N V L A R H
 Q I G A G R V H G E E A V E L F T G A A R E H H L R G E L D D A A G Y L E L A H
 R A S D D P V T R A A L R V G A A A I E R L C N P V R A G R H L P E L L T A S R
 A G L L S S E H A V S L A D W L A M G G R P G E A A E V L A T Q R P A A D S E Q
 H R A L L R S G E L S L A L V H P G A W D P L R R T D R F A A G G L G S L P G P
 A R H R A V A D Q A V I A A L R G R L D R A D A N A E S V L Q H T D A T A D R T
 T A I M A L L A L L Y A E N T D A V Q F W V D K L A G D E G T R T P A D E A V H
 A G F N A E I A L R R G D L M R A V E Y G E A A L G H R H L P T W G M A A A L P 40
 L S S T V V A A I R L G D L D R A E R W L A E P L P Q Q T P E S L F G L H L L W
 A R G Q H H L A T G R H G A A Y T A F R E C G E R M R R W A V D V P G L A L W R
 V D A A E S L L L L G R D R A E G L R L V S E Q L S R P M R P R A R V Q T L R V
 Q A A Y S P P P Q R I D L L E E A A D L L V T C N D Q Y E L A N V L S D L A E A
 S S M V R Q H S R A R G L L R R A R H L A T Q C G A V P L L R R L G A E P S D I
 G G A W D A T L G Q R I A S L T E S E R R V A A L A A V G R T N R E I A E Q L F
 V T A S T V E Q H L T N V F R K L A V K G R Q Q L P K E L A D V G E P A D R D R
 R C G 50

配列番号 6 5

M I A R L S P P D L I A R D D E F G S L H R A L T R A G G G R G V V A A V T G P
 I A C G K T E L L D A A A A K A G F V T L R A V C S M E E R A L P Y G M L G Q L
 L D Q P E L A A R T P E L V R L T A S C E N L P A D V D N R L G T E L T R T V L
 T L A A E R P V L I G I D D V H H A D A P S L R C L L H L A R R I S R A R V A I
 V L T E L L R P T P A H S Q F R A A L L S L R H Y Q E I A L R P L T E A Q T T E
 L V R R H L G Q D A H D D V V A Q A F R A T G G N L L L G H G L I D D I R E A R
 T R T S G C L E V V A G R A Y R L A Y L G S L Y R C G P A A L S V A R A S A V L
 G E S V E L T L V Q R M T G L D T E A V E Q A H E Q L V E G R L L R E G R F P H
 P A A R S V V L D D L S A A E R R G L H E L A L E L L R D R G V A S K V L A R H 10
 Q M G T G R V H G A E V A G L F T D A A R E H H L R G E L D E A V T Y L E F A Y
 R A S D D P A V H A A L R V D T A A I E R L C D P A R S G R H V P E L L T A S R
 E R L L S S E H A V S L A C W L A M D G R P G E A A E V L A A Q R S A A P S E Q
 G R A H L R V A D L S L A L I Y P G A A D P P R P A D P P A E D E V A S F S G A
 V R H R A V A D K A L S N A L R G W S E Q A E A K A E Y V L Q H S R V T T D R T
 T T M M A L L A L L Y A E D T D A V Q S W V D K L A G D D N M R T P A D E A V H
 A G F R A E A A L R R G D L T A A V E C G E A A L A P R V V P S W G M A A A L P
 L S S T V A A A I R L G D L D R A E R W L A E P L P E E T S D S L F G L H M V W
 A R G Q H H L A A G R Y R A A Y N A F R D C G E R M R R W S V D V P G L A L W R
 V D A A E A L L L L G R G R D E G L R L I S E Q L S R P M G S R A R V M T L R V
 Q A A Y S P P A K R I E L L D E A A D L L I M C R D Q Y E L A R V L A D M G E A 20
 C G M L R R H S R A R G L F R R A R H L A T Q C G A V P L L R R L G G E S S D A
 D G T Q D V T P A Q R I T S L T E A E R R V A S H A A V G R T N K E I A S Q L F
 V T S S T V E Q H L T N V F R K L G V K G R Q Q L P K E L S D A G

配列番号 6 6

M E F Y D L V A R D D E L R R L D Q A L G R A A G G R G V V V T V T G P V G C G
 K T E L L D A A A A E E E F I T L R A V C S A E E R A L P Y A V I G Q L L D H P
 V L S A R A P D L A C V T A P G R T L P A D T E N R L R R D I L T R A L L A L S
 E R P V L I C I D D V H Q A D T A S L N C L L H L A R R V A S A R I A M I L T E 30
 L R R L T P A H S R F E A E L L S L R H R H E I A L R P L G P A D T A E L A R A
 R L G A G V T A D E L A Q V H E A T S G N P N L V G G L V N D V R E A W A A G G
 T G I A A G R A Y R L A Y L S S V Y R C G P V P L R I A Q A A A V L G P S A T V
 T L V R R I S G L D A E T V D E A T A I L T E G G L L R D H R F P H P A A R S V
 V L D D M S A Q E R R R L H R S T L D V L D G V P V D V L A H H Q A G A G L L H
 G P Q A A E M F A R A S Q E L R V R G E L D A A T E Y L Q L A Y R A S D D A G A
 R A A L Q V E T V A G E R R R N P L A A S R H L D E L A A A A R A G L L S A E H
 A A L V V H W L A D A G R P G E A A E V L A L Q R A L A V T D H D R A R L R A A
 E V S L A L F H P G V P G S D P R P L A P E E L A S L S L S A R H G V T A D N A
 V L A A L R G R P E S A A A E A E N V L R N A D A A A S G P T A L A A L T A L L
 Y A E N T D A A Q L W A D K L A A G I G A G E G E A G Y A G P R T V A A L R R G
 D L T T A V Q A A G A V L D R G R P S S L G I T A V L P L S G A V A A A I R L G 40
 E L E R A E K W L A E P L P E A V H D S L F G L H L L M A R G R Y S L A V G R H
 E A A Y A A F R D C G E R M R R W D V D V P G L A L W R V D A A E A L L P G D D
 R A E G R R L I D E Q L T R P M G P R S R A L T L R V R A A Y A P P A K R I D L
 L D E A A D L L L S S N D Q Y E R A R V L A D L S E A F S A L R Q N G R A R G I
 L R Q A R H L A A Q C G A V P L L R R L G V K A G R S G R L G R P P Q G I R S L
 T E A E R R V A T L A A A G Q T N R E I A D Q L F V T A S T V E Q H L T N V F R
 K L G V K G R Q Q L P A E L A D L R P P G

配列番号 6 7

MY S G T C R E G Y E L V A R E D E L G I L Q R S L E Q A S S G Q G V V V T V T 50

G P I A C G K T E L L D A A A A K A E A I I L R A V C A P E E R A M P Y A M I G
 Q L I D D P A L A H R A P G L A D R I A Q G G Q L S L R A E N R L R R D L T R A
 L L A L A V D R P V L I G V D D V H H A D T A S L N C L L H L A R R V R P A R I
 S M I F T E L R S L T P T Q S R F K A E L L S L P Y H H E I A L R P F G P E Q S
 A E L A R A A F G P G L A E D V L V G L Y K T T R G N L S L S R G L I S D V R E
 A L A N G E S A F E A G R A F R L A Y L G S L Y R C G P V A L R V A R V A A V L
 G P S A T T T L V R R L S G L S A E T I D R A T K I L T E G G L L L D Q Q F P H
 P A A R S V V L D D M S A Q E R R G L H T L A L E L L D E A P V E V L A H H Q V
 G A G L I H G P K A A E M F A K A G K A L V V R N E L G D A A E Y L Q L A H R A
 S D D V S T R A A L R V E A V A I E R R R N P L A S S R H M D E L S A A G R A G 10
 L L S P K H A A L A V F W L A D G G R S G E A A E V L A S E R P L A T T D Q N R
 A H L R F V E V T L A L F S P G A F G S D R R P P P L T P D E L A S L P K A A W
 Q C A V A D N A A M T A L H G H P E L A T A Q A E T V L R Q A D S A A D A I P A
 A L I A L L Y A E N T E S A H I W A D K L G S T N G G V S N E A E A G Y A G P C
 A E I A L R R G D L A T A F E A G S T V L D D R S L P S L G I T A A L L L S S K
 T A A A V R L G E L E R A E K L L A E P L P N G V Q D S L F G L H L L S A Y G Q
 Y S L A M G R Y E S A L R A F H T C G E R M R S W D V D V P G L A L W R V D A A
 E A L L S L D R N E G Q R L I D E Q L T R P M G P R S R A L T L R I K A A Y L P
 R T K R I P L L H E A A E L L L P C P D P Y E Q A R V L A D L G D T L S A L R R
 Y S R A R G V L R Q A R H L A A Q C G A V P L L R R L G G E P G R I D D A G L P 20
 Q R S T S L T D A E R R V A A L A A A G Q T N R E I A K Q L F V T A S T V E Q H
 L T S V F R K L G V K G R K Q L P T A L A D V E Q T

配列番号 6 8

M P A V E S Y E L D A R D D E L R R L E E A V G Q A G N G R G V V V T I T G P I
 A C G K T E L L D A A A A K S D A I T L R A V C S E E E R A L P Y A L I G Q L I
 D N P A V A S Q L P D P V S M A L P G E H L S P E A E N R L R G D L T R T L L A
 L A A E R P V L I G I D D M H H A D T A S L N C L L H L A R R V G P A R I A M V
 L T E L R R L T P A H S Q F H A E L L S L G H H R E I A L R P L G P K H I A E L
 A R A G L G P D V D E D V L T G L Y R A T G G N L N L G H G L I K D V R E A W A
 T G G T G I N A G R A Y R L A Y L G S L Y R C G P V P L R V A R V A A V L G Q S
 A N T T L V R W I S G L N A D A V G E A T E I L T E G G L L H D L R F P H P A A 30
 R S V V L N D L S A R E R R R L H R S A L E V L D D V P V E V V A H H Q A G A G
 F I H G P K A A E I F A K A G Q E L H V R G E L D A A S D Y L Q L A H H A S D D
 A V T R A A L R V E A V A I E R R R N P L A S S R H L D E L T V A A R A G L L S
 L E H A A L M I R W L A L G G R S G E A A E V L A A Q R P R A V T D Q D R A H L
 R A A E V S L A L V S P G A S G V S P G A S G P D R R P R P L P P D E L A N L P
 K A A R L C A I A D N A V I S A L H G R P E L A S A E A E N V L K Q A D S A A D
 G A T A L S A L T A L L Y A E N T D T A Q L W A D K L V S E T G A S N E E E G A
 G Y A G P R A E T A L R R G D L A A A V E A G S A I L D H R R G S L L G I T A A
 L P L S S A V A A A I R L G E T E R A E K W L A E P L P E A I R D S L F G L H L
 L S A R G Q Y C L A T G R H E S A Y T A F R T C G E R M R N W G V D V P G L S L 40
 W R V D A A E A L L H G R D R D E G R R L I D E Q L T H A M G P R S R A L T L R
 V Q A A Y S P Q A Q R V D L L E E A A D L L L S C N D Q Y E R A R V L A D L S E
 A F S A L R H H S R A R G L L R Q A R H L A A Q C G A T P L L R R L G A K P G G
 P G W L E E S G L P Q R I K S L T D A E R R V A S L A A G G Q T N R V I A D Q L
 F V T A S T V E Q H L T N V F R K L G V K G R Q H L P A E L A N A E

配列番号 6 9

M P A V K R N D L V A R D G E L R W M Q E I L S Q A S E G R G A V V T I T G A I
 A C G K T V L L D A A A A S Q D V I Q L R A V C S A E E Q E L P Y A M V G Q L L
 D N P V L A A R V P A L G N L A A A G E R L L P G T E N R I R R E L T R T L L A 50

LADERPVVLIGVDDMHADPASLDCLLHLARRVGPARIAIV
 LTEELRRRTPAHSRFQSELLSLRYHHEIGLQPLTAEHTADL
 ARVGLGAEVDDDVLTELYEATGGNPSSLCCGLIRDVRQDWE
 AGVTGIGHVGRAYRLAYLSSLYRCGPAALRTARAABAVLGDS
 ADAACLIRRVSGLGTEAVGQAIQQLTEGGLLRDQQFPHPAA
 RSVVVLDDMSAQERHAMYRSAREAAAEGQADPGTPGEPRAA
 TAYAGCGEQAGDYPEPAGRACVDGAGPAEYCGDPHGADD
 PDELVAALGGLLPSRLVAMKIRRLAVAGRPGAAAELLTSQ
 RLHAVTSEDRAASLRAAEVALATLWPGATGPDRHPLTEQEA
 ASLPEGPRLAAADDAVGAALRGRAEYAAAEEAENVLRHAD 10
 PAAGGDAYAAMIAALLYTEHPENVLFWADKLDAGRDPDEETS
 YPGLRAETAVRLGDLTAMELGRTVLDQRRLPGLGVAAGL
 LLGGAVTAAIRLGDLDRAEKWLAEPIPDAIRTSLYGLHVL
 AARGRLDLAAGRYEAAYTAFRLCGERMAGWDADVSGLALW
 RVDAAEALLSAGIRPDEGRKLIDDQLTREMGARSRALTLR
 AQAAAYS LPVHRVGLLDEAAGLLLACHDGYERARVLA
 DLGETLRTLRTDAAQRVLRQAEQAAARCGSVP
 LRRLGAEPVRI
 GTRRGE PGLPQRIRLLTDAERRVAAMAAAAGQT
 NREIAGR 20
 LFVTTASTVEQH
 LTSVFRKLGVKGRRLPTELAQAV

配列番号 7 0

MYSGTCREGYELVAREDELGILQRSLEQASSGQGVVVT
 GPIACGKTELLDAAA
 KAEAIILRAVCAPEERAM
 PYAMIGQLIDDP
 PALAHRAPGLADRIA
 QGGQLS
 LRAENRLRDLTRA
 LLALAVDRPVLIGV
 DDVHHADTASLNCLL
 HLA
 RRV
 PARI
 SMIFTEL
 RS
 LTPTQS
 RFKAE
 LLSPYH
 HEIALR
 PFGPEQS
 AELARA
 AFGPGL
 AEDV
 LAGLYK
 TTRGN
 LSLSRGL
 ISDV
 RE
 ALANG
 ESAFEAGRA
 FRLAYL
 SLYRCGP
 VALR
 VARVA
 AVL
 GPSATT
 TLV
 RRLS
 GLSAET
 TIDR
 ATKIL
 TEG
 GL
 LDQQ
 FPH
 PAARS
 VVL
 DDMS
 AQR
 RSL
 HTL
 ALE
 LDE
 AP
 VEV
 LAHH
 QVGAGL
 IHGP
 KAA
 EMFA
 KAG
 KALV
 VRN
 ELD
 GDA
 AEY
 LQL
 AHR
 SDDV
 STRA
 ALR
 VEA
 VAI
 IERR
 RNPL
 ASSR
 HM
 DEL
 SAAG
 RAG
 LLSP
 KHA
 ALA
 VF
 WLAD
 GGR
 SGE
 AAE
 EVL
 LASER
 PLATT
 DQ
 NRA
 HLR
 FVE
 VT
 LAL
 FSP
 GAF
 GSD
 RPP
 PLTP
 DEL
 ASLP
 KAA
 W
 QCA
 AV
 DNA
 A
 MT
 ALH
 GH
 PEL
 ATA
 QA
 ET
 VL
 RQ
 A
 DSA
 ADA
 IPA
 ALI
 ALL
 YA
 EN
 TE
 SA
 HI
 WAD
 KLG
 STN
 AGV
 SNE
 A
 AGY
 AG
 PC
 AE
 IAL
 RRG
 DL
 AT
 A
 FE
 AG
 SAV
 L
 D
 DR
 S
 L
 P
 SL
 G
 IT
 A
 ALL
 L
 S
 SK
 TAA
 A
 V
 RL
 G
 E
 L
 ERA
 EK
 L
 LAE
 P
 L
 P
 NGV
 Q
 D
 S
 L
 F
 GL
 H
 L
 L
 S
 A
 Y
 G
 Q
 Y
 S
 L
 A
 M
 G
 R
 Y
 E
 S
 A
 H
 R
 A
 F
 R
 T
 C
 G
 E
 R
 M
 R
 S
 W
 D
 V
 D
 V
 P
 G
 L
 A
 L
 W
 R
 V
 D
 A
 A
 E
 A
 L
 L
 P
 R
 T
 K
 R
 I
 P
 L
 L
 H
 E
 A
 A
 E
 L
 L
 L
 P
 C
 P
 D
 P
 Y
 E
 Q
 A
 R
 V
 L
 A
 D
 L
 G
 D
 T
 L
 S
 A
 L
 R
 R
 Y
 S
 R
 A
 R
 G
 V
 L
 R
 Q
 A
 R
 H
 L
 A
 T
 Q
 C
 G
 A
 V
 P
 L
 L
 R
 R
 L
 G
 G
 E
 P
 G
 R
 I
 D
 D
 A
 G
 L
 P
 Q
 R
 S
 T
 S
 L
 T
 D
 A
 E
 R
 R
 V
 A
 A
 L
 A
 A
 A
 G
 Q
 T
 N
 R
 E
 I
 A
 E
 Q
 L
 F
 V
 T
 A
 S
 T
 V
 E
 Q
 H
 L
 T
 S
 V
 F
 R
 K
 L
 G
 V
 K
 G
 R
 K
 Q
 L
 P
 T
 A
 L
 A
 D
 V
 E
 Q
 T

配列番号 7 1

MYSGTCREGYELVAREDELGILQRSLEEAGSGQGAVV
 V
 GPIACGKTELLDAAA
 KADAIILRAVCAPEERAM
 PYAMIGQLIDDP
 PALAHRAPELADRIA
 QGGHLS
 LRAENRLRDLTRA
 LLALAVDRPVLIGV
 DDVHHADTASLNCLL
 HLA
 RRV
 PARI
 SMIFTEL
 RS
 LTPTQS
 RFKAE
 LLSPYH
 HEIALR
 P
 LGPEQS
 AELA
 HAA
 FGPGL
 AEDV
 LAGLY
 GMTR
 RGN
 LSLSRGL
 ISDV
 RE

50

A Q A N G E S A F E V G R A F R L A Y L S S L Y R C G P I A L R V A R V A A V L
 G P S A T T T L V R R L S G L S A E T I D R A T K I L T E G G L L L D H Q F P H
 P A A R S V V L D D M S A Q E R R S L H T L A L E L L D E A P V E V L A H H Q V
 G A G L I H G P K A A E I F A R A G Q A L V V R N E L G D A A E Y L Q L A H R A
 S D D V S T R A A L R V E A V A I E R R R N P L A S S R H M D E L S A A G R A G
 L L S P K H A A L A V F W L A D G G R S G E A A E V L A S E H P L A T T D Q N R
 A H L R F A E V T L A L F C P G A F G S D R R P P P L A P D E L A S L P K A A W
 Q C A V A D N A V M T A L H A H P E L A T A Q A E T V L R Q A D S A A D A I P A
 A L I A L L Y A E N T E S A Q I W A D K L G S T N A G V S N E A E A G Y A G P C
 A E I A L R R G D L A T A F E A G G T V L D D R P L P S L G I T A A L L L S S K
 T A A A V R L G E L E R A E K L L A E P L P N G V Q D S L F G L H L L S A H G Q
 Y S L A M G R Y E S A H R A F H T C G E R M R S W G V D V P G L A L W R V D A A
 E A L L S L D R N E G Q R L I D E Q L A R P M G P R S R A L T L R I K A A Y L P
 R T K R I P L L H E A A E L L L S C P D P Y E Q A R V L A D L G D T L S A L R R
 Y S R A R G V L R Q A R H L A T Q C G A V P L L R R L G G E P G R I D D A G L P
 Q R S T S L T D A E R R V S A L A A A G Q T N R E I A K Q L F V T A S T V E Q H
 L T S V F R K L G V K G R R Q L P T A L A D V E

配列番号 7 2

M Y S G T C R E G Y E L V A R E D E L G I L Q R S L E Q A S S G Q G V V V T V T
 G P I A C G K T E L L D A A A A K A E A I I L R A V C A P E E R A M P Y A M I G
 Q L I D D P A L A H R A P G L A D R I A Q G G Q L S L R A E N R L R R D L T R A
 L L A L A V H R P V L I G V D D V H H A D T A S L N C L L H L A R R V R P A R I
 S M I F T E L R S L T P T Q S R F K A E L L S L P Y H H E I A L R P F G P E Q S
 A E L A R A A F G P G L A E D V I A G L Y K T T R G N L S L S R G L I S D V R E
 A L A N G E S A F E A G R A F R L A Y L S S L Y R C G P V A L R V A R V A A V L
 G P S A T T T L V R R L S G L S A E T I D R A T K I L T E G G L L L D Q Q F P H
 P A A R S V V L D D M S A Q E R R G L H T L A L E L L D E A P V E V L A H H Q V
 G A G L I H G P K A A E M F A K A G K A L V V R N E L G D A A E Y L Q L A H R A
 S D D V S T R A A L R V E A V A I E R R R N P L A S S R H M D E L S A A G R A G
 L L S P K H A A L A V F W L A D G G R S G E A A Q V L A S E R P L A T T D Q N R
 A H L R F V E V T L A L F S P G A F G S D R R P P P L T P D E L A S L P K A A W
 Q C A V A D N A A M T A L H G H P E L A T A Q A E T V L R Q A D S A A D A I P A
 A L I A L L Y A E N T E S A H I W A D K L G S M N A G V S N E A E A G Y A G P C
 A E I A L R R G D L A T A F E A G S T V L D D R S L P S L G I T A A L L L S S K
 T A A A V R L G E L E R A E K L L A E P L P N G V Q D S L F G L H L L S A Y G Q
 Y S L A M G R Y E S A H R A F R T C G E R M R S W D V D V P G L A L W R V D A A
 E A L L S L D R N E G Q R L I D E Q L T R P M G P R S R A L T L R I K A A Y L P
 R T K R I P L L H E A A E L L L P C P D P Y E Q A R V L A D L G D T L S A L R R
 Y S R A R G V L R Q A R H L A T Q C G A V P L L R R L G G E P G R I D D A G L P
 Q R S T S L T D A E R R V A A L A A A G Q T N R E I A E Q L F V T A S T V E Q H
 L T S V F R K L G V K G R K Q L P T A L A D V E Q T

配列番号 7 3

M R A I N A S D T G P E L V A R E D E L G R V R S A L N R A N G G Q G V L I S I
 T G P I A C G K T E L L E A A A S E V D A I T L R A V C A A E E R A I P Y A L I
 G Q L I D N P A L G I P V P D P A G L T A Q G G G R L S S S A E N R L R R D L T R
 A L L T L A T D R L V L I C V D D V Q H A D N A S L S C L L Y L A R R L V P A R
 I A L V F T E L R V L T S S Q L R F N A E L L S L R N H C E I A L R P L G P G H
 A A E L A R A T L G P G L S D E T L T E L Y R V T G G N L S L S R G L I D D V R
 D A W A R G E T G V Q V G R A F R L A Y L G S L H R C G P L A L R V A R V A A V
 L G P S A T S V L V R R I S G L S A E A M A Q A T D I L A D G G L L R D Q R F T

10

20

30

40

50

H P A A R S V V L D D M S A E E R R S V H S L A L L D E A P A E M L A H H R
 V G A G L V H G P K A A E T F T G A G R A L A V R G M L G E A A D Y L Q L A Y R
 A S G D A A T K A A I R V E S V A V E R R R N P L V V S R H W D E L S V A A R A
 G L L S C E H V S R T A R W L T V G G R P G E A A R V L A S Q H R R V V T D Q D
 R A H L R V A E F S L A L L Y P G T S G S D R R P H P L T S D E L A A L P T A T
 R H C A I A D N A V M A A L R G H P E L A T A E A E A V L Q Q A D A A D G A A L
 T A L M A L L Y A E S I E V A E V W A D K L A A E A G A S N G Q D A E Y A G I R
 A E I A L R R G D L T A A V E T A G M V L D G R P L P S L D I T A T L L L A G R
 A S V A V R L G E L D H A E E L F A A P P E D A F Q D S L F G L H L L S A H G Q
 Y S L A T G R P E S A Y R A F R A C G E R M R D W G F D A P G V A L W R V G A A
 E A L L G L D R N E G R R L I D E Q L S R T M A P R S H A L T L R I K A A Y M P
 E P K R V D L L Y E A A E L L L S C R D Q Y E R A R V L A D L G E A L S A L G N
 Y R Q A R G V L R Q A R H L A M R T G A D P L L R R L G I R P G R Q D D P D P Q
 P R S R S L T N A E R R A A S L A A T G L T N R E I A D R L F V T A S T V E Q H
 L T N V F R K L G V K G R K Q L P A E L D D M E

10

L A L 結合部位

一部の実施形態では、遺伝子クラスター（例えば、PKS 遺伝子クラスター）は、1つまたは複数の L A L 結合部位を含む1つまたは複数のプロモーターを含む。L A L 結合部位は、ポリヌクレオチドコンセンサス L A L 結合部位配列（例えば、本明細書において記載される）を含み得る。一部の場合では、L A L 結合部位は、コア A G G G G G G (配列番号 74) モチーフを含む。ある特定の場合では、L A L 結合部位は、配列番号 39 に対して少なくとも 80%（例えば、80%、85%、90%、95%、97%、98%、99%、または 100%）の相同性を有する配列を含む。L A L 結合部位は、コンセンサスまたは最適化された L A L 結合部位の配列にマッチするように修復されている突然変異部位を含み得る。一部の実施形態では、L A L 結合部位は、合成 L A L 結合部位である。一部の実施形態では、合成 L A L 結合部位は、(a) 少なくとも 8 つのヌクレオチドを含む複数個の合成核酸を提供する工程、(b) 少なくとも 8 つのヌクレオチドを含む複数個のヌクレオチドの 1 つまたは複数を、1 つまたは複数の L A L と接触させる工程、(c) 工程 (a) の核酸と工程 (b) の L A L との間の結合親和性を判定する工程であって、合成核酸と L A L との間の親和性が X よりも大きければ、合成核酸を合成 L A L 結合部位であると同定する、工程によって同定され得る。同定された合成 L A L 結合部位は、次いで、化合物生産クラスター（例えば、PKS クラスター）の宿主細胞内に導入され得る。

20

【0068】

一部の実施形態では、L A L 結合部位と異種 L A L との対、または、異種 L A L 結合部位と、天然の対と比較して発現が増大した L A L との対は、(a) 1 つまたは複数の L A L 結合部位を提供する工程、(b) L A L 結合部位の 1 つまたは複数と、1 つまたは複数の L A L とを接触させる工程、(c) L A L 結合部位と L A L との間の結合親和性を判定する工程であって、L A L 結合部位と L A L との間の親和性が、L A L 結合部位と、その同種 L A L および / またはその同種 L A L 結合部位にある L A L との間の親和性よりも大きければ、発現が増大した対が同定される、工程によって同定され得る。一部の実施形態では、L A L 結合部位と L A L との間の結合親和性は、L A L および L A L 結合部位の両方を含む細胞によるタンパク質または化合物の発現を判定することによって判定される。

30

【0069】

構成的に活性な L A L

一部の実施形態では、組換え L A L は、構成的に活性な L A L である。例えば、L A L のアミノ酸配列は、改変された L A L がその同族結合部位に係合し、化合物生産タンパク質（例えば、ポリケチドシンターゼ）の転写を活性化するために、誘導因子化合物の存在を要しないように、修飾されている。宿主細胞への構成的に活性な L A L の導入は、化合物生産タンパク質（例えば、ポリケチドシンターゼ）の発現を増大させ、ひいては、対応する化合物（例えば、ポリケチド）の生産を増大させる可能性が高い。

40

50

【0070】

一方向性 L A L の遺伝子操作

F k P h D 遺伝子クラスターは、L A L 結合部位であると推定される保存された（単一または複数の、互いに反転されたおよび／または直接的に反復している）G G G G G T（配列番号40）モチーフを内在させている複数の二方向性プロモーター・オペレーターによって駆動されるマルチシストロン性アーキテクチャーで整理される。二方向性のL A L プロモーターは、対立するプロモーターの1つを戦略的に欠失させるがタンデムL A L 結合部位は維持することによって、一方向性のL A L プロモーター（U n i L A L）に変換され得る（M a l Tについて実証されたように、ネイティブプロモーターにおけるL A L の結合が協調的であるケースにおいて）。機能的に、これは、アンチセンス鎖上に存在する保存されたG G G G G T（配列番号40）モチーフの全ての3'配列（-35プロモーター配列および-10プロモーター配列を含有する可能性が高い）を除去するが、センス鎖上の全配列は無傷のまま残すことによって達成される。この欠失の結果、転写は、一方向のみで活性化される。このフィード・フォワードサーキットアーキテクチャーの利点は、ストレプトマイセス属の栄養成長条件および発酵成長条件の複雑なライフサイクルの間のL A L の発現を調和させるおよび／または最大化することである。

10

【0071】

宿主細胞

一部の実施形態では、宿主細胞は、アクチノバクテリウムなどの細菌である。例えば、一部の実施形態では、宿主細胞は、ストレプトマイセス（*Streptomyces*）株である。一部の実施形態では、宿主細胞は、ストレプトマイセス・アヌラトス（*Streptomyces anulatus*）、ストレプトマイセス・アンチバイオティカス（*Streptomyces antibioticus*）、ストレプトマイセス・コエリカラー（*Streptomyces coelicolor*）、ストレプトマイセス・ペウセティウス（*Streptomyces peucetius*）、ストレプトマイセス（*Streptomyces*）種ATCC 700974、ストレプトマイセス・カヌス（*Streptomyces canus*）、ストレプトマイセス・ノドサス（*Streptomyces nodosus*）、ストレプトマイセス（*Streptomyces*）（複数種）、ストレプトアロテイクス・ヒンダスタヌス（*Streptoaaloticus hindustanus*）、ストレプトマイセス・ハイグロスコピカス（*Streptomyces hygroscopicus*）、ストレプトマイセス・アベルミティリス（*Streptomyces avermitilis*）、ストレプトマイセス・ビリドクロモゲネス（*Streptomyces viridochromogenes*）、ストレプトマイセス・ベルティシラス（*Streptomyces verticillatus*）、ストレプトマイセス・チャルトルエンシス（*Streptomyces chartrensis*）、ストレプトマイセス（*Streptomyces*）（複数種）、サッカロトリックス・ムタビリス（*Saccharothrix mutabilis*）、ストレプトマイセス・ハルステディイ（*Streptomyces halstedii*）、ストレプトマイセス・クラブリゲルス（*Streptomyces clavuligerus*）、ストレプトマイセス・ベネズエラエ（*Streptomyces venezuelae*）、ストレプトマイセス・ロゼオクロモゲネス（*Streptomyces roseochromogenes*）、アミコラトブシス・オリエンタリス（*Amycolatopsis orientalis*）、ストレプトマイセス・クラブリゲルス（*Streptomyces clavuligerus*）、ストレプトマイセス・リシリエンシス（*Streptomyces rishiriensis*）、ストレプトマイセス・ラベンデュラエ（*Streptomyces lavendulae*）、ストレプトマイセス・ロゼオスポルス（*Streptomyces roseosporus*）、ノノムラエア（*Nonomuraea*）種、ストレプトマイセス・ペウセティウス（*Streptomyces peucetius*）、サッカロポリスボラ・エリスラエア（*Saccharopolyspora erythraea*）、ストレプトマイ

20

30

40

50

セス・フィリピネンシス (*Streptomyces filipinensis*)、ストレプトマイセス・ハイグロスコピカス (*Streptomyces hygroskopicus*)、ミクロモノスpora・ブルプレア (*Microomonospora purpurea*)、ストレプトマイセス・ハイグロスコピカス (*Streptomyces hygroskopicus*)、ストレプトマイセス・ナルボネンシス (*Streptomyces narbonensis*)、ストレプトマイセス・カナマイセティカス (*Streptomyces kanamyceticus*)、ストレプトマイセス・コリヌス (*Streptomyces collinus*)、ストレプトマイセス・ラサリエンシス (*Streptomyces lasaliensis*)、ストレプトマイセス・リンカネンシス (*Streptomyces lincolnensis*)、ダクトスボランギウム・アウランティアカム (*Dactylosporangium aurantiacum*)、ストレプトマイセス・トキシトリシニ (*Streptomyces toxicinii*)、ストレプトマイセス・ハイグロスコピカス (*Streptomyces hygroskopicus*)、ストレプトマイセス・プリカタス (*Streptomyces plicatus*)、ストレプトマイセス・ラベンデュラエ (*Streptomyces lavendulae*)、ストレプトマイセス・ガーナエンシス (*Streptomyces ghanaensis*)、ストレプトマイセス・シンナモネンシス (*Streptomyces cinnamomensis*)、ストレプトマイセス・アウレオファシエンス (*Streptomyces aureofaciens*)、ストレプトマイセス・ナタレンシス (*Streptomyces natalensis*)、ストレプトマイセス・チャタノオゲンシス (*Streptomyces chattanoogensis*) L 10、ストレプトマイセス・ライディカス (*Streptomyces lydicus*) A 02、ストレプトマイセス・フライアエ (*Streptomyces fradiae*)、ストレプトマイセス・アムボファシエンス (*Streptomyces ambofaciens*)、ストレプトマイセス・テンダエ (*Streptomyces tendae*)、ストレプトマイセス・ノウルセイ (*Streptomyces noursei*)、ストレプトマイセス・アベルミティリス (*Streptomyces avermitilis*)、ストレプトマイセス・リモーサス (*Streptomyces rimosus*)、ストレプトマイセス・ウェッドモレンシス (*Streptomyces wedmorensis*)、ストレプトマイセス・カカオイ (*Streptomyces cacaoi*)、ストレプトマイセス・プリスチナエスピラリス (*Streptomyces pristinaespinalis*)、ストレプトマイセス・プリスチナエスピラリス (*Streptomyces pristinaespinalis*)、アクチノプラネス (*Actinoplanes*) 種 ATCC 33076、ストレプトマイセス・ハイグロスコピカス (*Streptomyces hygroskopicus*)、ルシェバリエリア・エロコロニゲネス (*Lechevalieria aerocolonogenes*)、アミコラトプシス・メディテラネイ (*Amycolatopsis mediterranei*)、アミコラトプシス・ルリダ (*Amycolatopsis lurida*)、ストレプトマイセス・アルブス (*Streptomyces albus*)、ストレプトマイセス・グリセオルス (*Streptomyces griseolus*)、ストレプトマイセス・スペクタビリス (*Streptomyces spectabilis*)、サッカロポリスピローラ・スピノサ (*Saccharopolyspora spinosa*)、ストレプトマイセス・アムボファシエンス (*Streptomyces ambofaciens*)、ストレプトマイセス・スタウロスボレウス (*Streptomyces staurosporeus*)、ストレプトマイセス・グリゼウス (*Streptomyces griseus*)、ストレプトマイセス (*Streptomyces*) (複数種)、ストレプトマイセス・アクロモゲネス (*Streptomyces acromogenes*)、ストレプトマイセス・ツクバエンシス (*Streptomyces tsukubaensis*)、アクチノプラネス・テイコマイセチクス (*Actinoplanes teichomyceticus*)、

ストレプトマイセス・グロセッセンス (*Streptomyces glaucescens*)、ストレプトマイセス・リモーサス (*Streptomyces rimosus*)、ストレプトマイセス・キャトルヤ (*Streptomyces cattleya*)、ストレプトマイセス・アズレウス (*Streptomyces azureus*)、ストレプトアロテイカス・ヒンダスタヌス (*Streptomyces aloluteicus hindustanus*)、ストレプトマイセス・チャートレウシス (*Streptomyces chartreusis*)、ストレプトマイセス・フラディアエ (*Streptomyces fradiae*)、ストレプトマイセス・コエリカラ (*Streptomyces coelicolor*)、ストレプトマイセス・ハイグロスコピカス (*Streptomyces hygroskopicus*)、ストレプトマイセス (*Streptomyces*) 種 11861、ストレプトマイセス・バージニアエ (*Streptomyces virginiae*)、アミコラトプシス・ジャポニカム (*Amycolatopsis japonicum*)、アミコラトプシス・バルヒマイシニ (*Amycolatopsis balhimycinii*)、ストレプトマイセス・アルブス (*Streptomyces albus*) J 1074、ストレプトマイセス・コエリカラ (*Streptomyces coelicolor*) M 1146、ストレプトマイセス・リビダンス (*Streptomyces lividans*)、ストレプトマイセス・インカーネイト (*Streptomyces incarnatus*)、ストレプトマイセス・バイオラセオルバー (*Streptomyces violaceoruber*)、またはストレプトマイセス・グリセオフュスカス (*Streptomyces griseofuscus*) である。一部の実施形態では、宿主細胞は、エシェリキア・コリ (*Escherichia coli*) などのエシェリキア (*Escherichia*) 株である。一部の実施形態では、宿主細胞は、バチルス・サブティリス (*Bacillus subtilis*) などのバチルス (*Bacillus*) 株である。一部の実施形態では、宿主細胞は、シュードモナス・プチダ (*Pseudomonas putida*) などのシュードモナス (*Pseudomonas*) 株である。一部の実施形態では、宿主細胞は、ミクソコッカス・キサンタス (*Myxococcus xanthus*) などのミクソコッカス (*Myxococcus*) 株である。

【0072】

方法

本発明のタンパク質、核酸、ベクター、および宿主細胞は、化合物（例えば、ポリケチド）の生産のために使用することができる。タンパク質への異種ドメインの導入によって、タンパク質によって生産されるポリケチドの化学構造の改変が可能になる。

【0073】

異種ドメインの導入

- ケトンプロセシングドメインの活性は、他のポリケチドシンターゼのドメインの配列を導入することによって改変することができる。複数の異種配列を、発現するポリケチドの量を劇的に低減させることなく特異的なドメインの活性を改変するそれらの能力について試験することができる。ポリケチドシンターゼの新規な変異体を、厳しい品質管理に供することができる（目的の領域のサンガーシーケンシング、クラスターの完全性を確認するためのPCRに基づく「タイリング」、およびBAC全体をシーケンシングするためのイルミナシーケンシング）。BACを次いで、2つの最適化された *Streptomyces* 生産株にコンジュゲートすることができ、固相抽出された (SPE: solid-phase extracted) 試料を、精製された FKP12 タンパク質と共にトップダウン質量分析に供して、生産された化合物を同定することができる。

【0074】

キメラポリケチドシンターゼから化合物を生成するためのワークフローの代表例には、1つのポリケチドシンターゼのドメイン、例えばケトレダクターゼドメインの短鎖ペプチド配列を、別のポリケチドシンターゼ上に、ホモロジーに基づくクローニングを使用してグラフトすることが含まれる。例えば、1つのケトレダクターゼの触媒性 Tyr を Phe

10

20

30

40

50

で置き換えることができ、活性部位 F G ループもまた欠失させてドメインを不活性化させることができる。得られたクローンは次いで、S t r e p t o m y c e s 発現宿主内にコンジュゲートし、発酵させることができる。化合物は次いで、分画されていない S P E 試料の比較 L C - T O F 分析を使用して同定され得る。トップダウン質量分析もまた、精製されたネイティブ F K B P 1 2 および修飾されたポリケチドシンターゼから得られる化合物を、未修飾のポリケチドシンターゼから得られる化合物と共に注入することによって行うことができる。この分析は、ドメインの活性の変化と一致する 2 つの化合物間の質量の差、例えば、不活性化したケトレダクターゼドメインについての 2 つの間の差を示し得る。

【 0 0 7 5 】

10

複数の構造的变化を有する化合物は、K R 単一変異体、D H 単一変異体、または E R 単一変異体の組み合わせを使用して生成され得る。

遺伝子操作されたポリケチドシンターゼのライプラリーの生産

コンビナトリアルドメインレベルでの遺伝子操作は、単一タンパク質骨格上の複数のドメインレベルでの変異体を組み合わせ、こうして、薬剤開発のための多様な P K S / N R P S 分子のライプラリースケールでの構築を可能にすることによって行うことができる。

【 0 0 7 6 】

20

あるいは、マルチプレックス並行遺伝子操作（例えば、部位特異的突然変異生成による）を使用して、薬剤開発のための遺伝子操作された P K S / N R P S 分子のライプラリーを生産することができる。例えば、親ポリケチドシンターゼをコードするポリヌクレオチドの部位特異的突然変異生成を使用して、並行して、複数個の遺伝子操作されたポリケチドシンターゼをコードする複数個のポリヌクレオチドを生成することができる。一部の実施形態では、複数個の遺伝子操作されたポリケチドシンターゼの各々は、親ポリケチドシンターゼと比較して少なくとも 1 つのコドン修飾（例えば、親ポリケチドシンターゼの少なくとも 1 つのドメインの保存されたモチーフにおける残基を特定するコドン）を含む。

【 0 0 7 7 】

単一分子の長鎖読み取りシーケンシングによる遺伝子操作された P K S ライプラリーの特徴付け

30

本発明の一部の実施形態では、単一分子の長鎖読み取りシーケンシング技術（例えば、N a n o p o r e シーケンシングまたは S M R T シーケンシング）を使用して、本明細書において記載される方法のいずれかによって生産される遺伝子操作されたポリケチドシンターゼまたは非リボソームペプチドシンターゼのライプラリーを特徴付けることができる。特に、単一分子の長鎖読み取りシーケンシング（例えば、N a n o p o r e シーケンシングまたは S M R T シーケンシング）を使用して、遺伝子操作されたポリケチドシンターゼまたは非リボソームペプチドシンターゼのコンビナトリアルライプラリーまたはマルチプレックスライプラリー（例えば、並行遺伝子操作によって生成するマルチプレックスライプラリー）を特徴付ける（例えば、デコンボリューションする）ことができる。単一分子長鎖読み取りシーケンシングによって、コンビナトリアルライプラリーに組み込まれている 1 つまたは複数のモジュールの同定が可能になる。これによってさらに、得られた複数個の遺伝子操作されたポリケチドシンターゼまたは非リボソームペプチドシンターゼの化学を予測することができる。予測された酵素化学は、したがって、遺伝子操作されたポリケチドシンターゼまたは非リボソームペプチドシンターゼによって生産された化合物に結び付けることができる。得られた化合物は、当業者に知られている化学的分析方法（例えば、質量分析または高速液体クロマトグラフィー）によって同定することができる。さらに、予測される酵素化学は、得られた化合物の機能（例えば、標的タンパク質への結合、または細胞に基づく表現型などの表現型の誘発）に結び付けることができる。したがって、遺伝子によってコードされた分子の長鎖読み取りシーケンシングは、遺伝子型 - 表現型の連結を可能にし得る。

40

【 0 0 7 8 】

50

単一分子長鎖読み取りシーケンシング技術は、単一分子のバイオポリマー（例えば、D

N A または R N A などのポリヌクレオチド) のシーケンシングを可能にする、および 2 キロベースを超える(例えば、5 キロベースを超える、10 キロベースを超える、20 キロベースを超える、50 キロベースを超える、または 100 キロベースを超える) 長さの読み取りを可能にする、あらゆるシーケンシング技術を含むと見なされ得る。単一分子長鎖読み取りシーケンシング技術によって、D N A または R N A の複数の単一分子を並行してシーケンシングすることが可能になり得る。単一分子長鎖読み取りシーケンシング技術には、シーケンシング対象の D N A または R N A の各分子の個々の区画化に依拠するシーケンシング技術が含まれ得る。

【0079】

Nanopore シーケンシングは、本明細書において記載される方法のいずれかによって調製される遺伝子操作されたポリケチドシンターゼまたは非リボソームペプチドシナーゼのライブラリーを特徴付けるために使用することができる典型的な単一分子長鎖読み取りシーケンシング技術である。Nanopore シーケンシングによって、単一分子のバイオポリマー(例えば、D N A または R N A などのポリヌクレオチド)の長鎖読み取りシーケンシングが可能になる。Nanopore シーケンシングは、耐電性ポリマー膜内にセットされたタンパク質ナノポアに依拠する。イオン電流が、この膜に電圧をかけることによってナノポアを通過する。分析物(例えば、D N A または R N A などのバイオポリマー)がポアまたはその開口部の近くを通過すれば、この事象は、特徴的な電流崩壊を生じさせる。ナノポアの表面全体にわたる電流密度の程度は、ナノポアを占めるD N A または R N A の組成(例えば、具体的な塩基)に依存する。したがって、電流の測定は、問題となっている分子の配列の同定を可能にする。

10

20

30

30

【0080】

単一分子リアルタイム(S M R T : s i n g l e m o l e c u l e r e a l - t i m e) シーケンシング(P a c B i o) は、本明細書において記載される方法のいずれかによって調製される遺伝子操作されたポリケチドシンターゼまたは非リボソームペプチドシナーゼのライブラリーを特徴付けるために使用することができる、典型的な単一分子長鎖読み取りシーケンシング技術である。S M R T は、並行化された単一分子D N A シーケンシング方法である。S M R T は、ゼロモードウェーブガイド(Z M W : z e r o - m o d e w a v e g u i d e) を利用する。単一D N A ポリメラーゼ酵素は、Z M W の底に、鋳型である単一分子のD N A と共に固定される。Z M W は、D N A ポリメラーゼによって組み込まれる単一ヌクレオチドのD N A のみを観察するために十分小さい、明るく照らされた観察容積を生じさせる構造である。4つのD N A 塩基の各々に、4つの異なる蛍光色素の1つを付着させる。ヌクレオチドがD N A ポリメラーゼによって組み込まれると、蛍光タグは切り落とされ、蛍光がもはや観察され得ない場所であるZ M W 観察区域の外側に拡散する。検出器は、ヌクレオチドの組み込みの蛍光シグナルを検出し、ベースコードが、色素の対応する蛍光に従って作成される。

【0081】

(実施例)

実施例 1 . ケトレダクターゼドメインの不活化

S 3 0 3 - K R 6 の短鎖ペプチド配列を、ホモロジーに基づくクローニングを使用して X 1 - K R 6 にグラフトした。触媒性 T y r を P h e で置き換え(赤で示す)、活性部位 F G ループ(青で示す)も欠失させる(図 5 A)。得られたクローンを S t r e p t o m y c e s 発現宿主内にコンジュゲートし、発酵させた。化合物 1 および C 1 6 ケト化合物 1 の分画されていない S P E 試料の比較 L C - T O F 分析は、新規な化合物が 6 0 8 . 3 5 という所望の M + H 質量を有していることを示した(図 5 B)。本発明者らは次いで、精製されたネイティブ F K B P 1 2 および化合物 1 または C 1 6 ケト化合物 1 を共注入することによってトップダウン質量分析を行った。この分析もまた、ヒドロキシルの C 1 6 でのケトンへの変換と一致する、2つの間の質量差を示した(図 5 C)。C 1 6 ケト化合物 1 を大規模に再発酵させ、精製して均質にし、構造を N M R 分光法によって確認した。

40

50

【0082】

実施例2. デヒドラターゼドメインおよびエノイルレダクターゼドメインの不活化

実施例1のプロトコールを使用して、化合物1を生産するPKS内のDHドメインおよびERドメインを、図6Aに示すように成功裏に失活させた。

【0083】

実施例3. 複数のドメインの同時の不活化

実施例1のプロトコールを使用して、図6Bで示すように、2つのドメインを同時に失活させた。

【0084】

さらに、モジュール6のKRドメインを不活化することによって生成する化合物であるC16ケト化合物1であるE-06 (KR6^{*})、およびモジュール4のDHドメインを不活化することによって生成するヒドロキシ化合物1類似体であるE36 (DH4^{*})の発現プロファイルを比較した。確認されたKRおよびDHの修飾を単一の構築物に組み合わせると、得られたコンビナトリアル化合物E-74 (KR6^{*}-DH4^{*})は、トップダウンアッセイによって検出すると、625.36の予想化合物質量を高収率で生産した(図6C)。

10

【0085】

実施例4. 定常領域の遺伝子操作

ラパマイシン/FKB506の「定常領域」は、FKB P12を結合するマクロライド環の保存された部分である。化合物1を生産するPKS内のDH8を、L P F X Wモチーフを突然変異させることによって不活化させて、定常領域のピラン環にヒドロキシルを有する化合物2を生成した(図7A)。611.38という予想質量(図7B)がトップダウンアッセイによって観察され、このことによって、化合物2がFKB P12結合親和性を保持していることが確認された。FKB P12:化合物2:CEP250複合体の構造(図7C)を結晶化によって明らかにした。構造によって、(1)FKB P12:CEP250界面が、ピラン環へのヒドロキシルの付加を融通し得ること、(2)新たに取り付けられた-OH基の立体化学、および予想した通り、(3)CEP250の結合もまた保持されていることが確認された。TR-FRETによって測定される、FKB P12に依存する、化合物2へのCEP250の結合は、化合物2と比較して増大した(図7D)。

20

【0086】

上記のデータは、生物学的機能(すなわち、標的タンパク質の結合)を保持するPKS天然生成物の化学的に新規な誘導体を生成するためのドメインレベルでの遺伝子操作の利用性を確立する。

30

【0087】

実施例5. コンビナトリアルドメインの遺伝子操作

粗抽出物における化合物のタンパク質標的の同定を可能にする、FKB P12親和性の強化およびトリプシンペプチドのLC-MS/MSシーケンシングに基づく最適化されたTarget-IDアッセイを開発した(図8A)。Target-ID分析によって、化合物1がCEP250およびCBY1 Ain 293T溶解物の両方を結合する一方で、化合物1のコンビナトリアル化合物誘導体である化合物3は、CBY1を選択的に結合し、CEP250は結合しないことが確認された。質量分析に基づくTarget-IDの結果は、TR-FRETデータで確認された。TR-FRETアッセイによって、CBY1が化合物1に結合することが確認された(図8B)。このデータによって、化合物3がCBY1に特異的であり、CEP250をもはや係合できないことも確認された。さらに、化合物2(図7A)はCEP250に特異的であり、CBY1には特異的でない。化合物3の構造をNMRによって確認し、これは、上記のドメインレベルでの遺伝子操作アプローチを介するKR6ドメイン、DH4ドメイン、およびER5ドメインの成功裏の不活化を示した。化合物2はまた、定常領域において第3のカルボニルを欠いており、このことは、化合物1の合成における最後のテラリング工程であるCypBが、化合物2を生産基質として利用できなかったことを示唆する。

40

50

【 0 0 8 8 】

上記のデータは、ドメインレベルでの遺伝子操作が、「リプログラミングされた」または改变された標的結合を有する化合物を生じさせること、ひいては、ドメインの遺伝子操作が新規な潜在的生物学的機能を有する分子を生成するために利用され得ることを実証する。

【 0 0 8 9 】**実施例 6 . ケトレダクターゼの失活によって生じる環の拡大**

化合物 1 を生産する P K S 内のモジュール 3 ~ 6 内の各 K R ドメインを、体系的に失活させた。6 つの配列を、K R ドメインを不活性化するそれらの能力について試験した(図 9 A)。予想外に、化合物の質量が 4 4 増えた化合物 4 が観察され、これを精製し、構造を N M R によって決定した。構造は、K R 5 の不活性化によってケトを取り付けることよりも、化合物 4 の環のサイズは、P K S 鎖の伸長を介するマロニル組み込みのさらなるラウンドに対応する 2 つの炭素によって拡大することを示した(図 9 B)。化合物 4 をもたらした、化合物 1 を生産する P K S におけるドメインレベルでの化合物変異体は、保存された触媒性 Y A A A N モチーフの近くの K R 5 における单一の A 1 a から G 1 u への置換であった。この突然変異は、ケトレダクターゼ活性部位のアクセスを防ぎ得、そうすることによって、次のサイクルのポリケチド延長におけるモジュール内ドメイン - ドメインハンドオフとモジュール間移動との間の動態バランスを改変することができる。このモデルによって、モジュールの繰り返しが、K R 5 突然変異の結果、有利であり、これにより、さらなるマロニル組み込み事象および環のサイズの拡大が生じることが予測される。

10

20

30

30

40

【 0 0 9 0 】**他の実施形態**

本開示をその詳細な説明と共に記載してきたが、先の記載は、本開示を説明するためのものであり、添付の特許請求の範囲によって定義される本開示の範囲を限定するものではないことを理解されたい。他の様様、利点、および改変は、以下の特許請求の範囲の範囲内である。

【 0 0 9 1 】

当業者であれば、本明細書において記載される発明に従って、具体的な実施形態の多くの同等物を認識しているか、または通常の実験のみを使用して確認することができる。本発明の範囲は、上記の詳細な説明に限定されるものではなく、むしろ、添付の特許請求の範囲で説明されるものである。

【 0 0 9 2 】

特許請求の範囲において、「a」、「a n」、および「t h e」などの冠詞は、逆の事実が示されない限りまたは文脈から明らかでない限り、1 つまたは 2 つ以上を意味し得る。グループの 1 つまたは複数のメンバー間の「または」を含む請求項または記載は、逆の事実が示されない限りまたは文脈から明らかでない限り、グループメンバーの 1 つ、2 つ以上、または全てが、所与の生成物もしくはプロセスにおいて存在する、所与の生成物もしくはプロセスにおいて採用される、または所与の生成物もしくはプロセスに関連するならば、十分であると見なされる。本発明は、グループの厳密に 1 つのメンバーが所与の生成物もしくはプロセスにおいて存在する、所与の生成物もしくはプロセスにおいて採用される、または所与の生成物もしくはプロセスに関連する、実施形態を含む。本発明は、グループメンバーの 2 つ以上または全てが所与の生成物もしくはプロセスにおいて存在する、所与の生成物もしくはプロセスにおいて採用される、または所与の生成物もしくはプロセスに関連する、実施形態を含む。

【 0 0 9 3 】

また用語「備える」は、非限定であることが意図され、さらなる要素または工程を含むことを許容するが必須ではないことにも留意されたい。用語「備える」が本明細書において用いられる場合、用語「からなる」も含まれ、開示される。

【 0 0 9 4 】

領域が示される場合、終点が含まれる。さらに、別段の指示がない限り、または文脈お

50

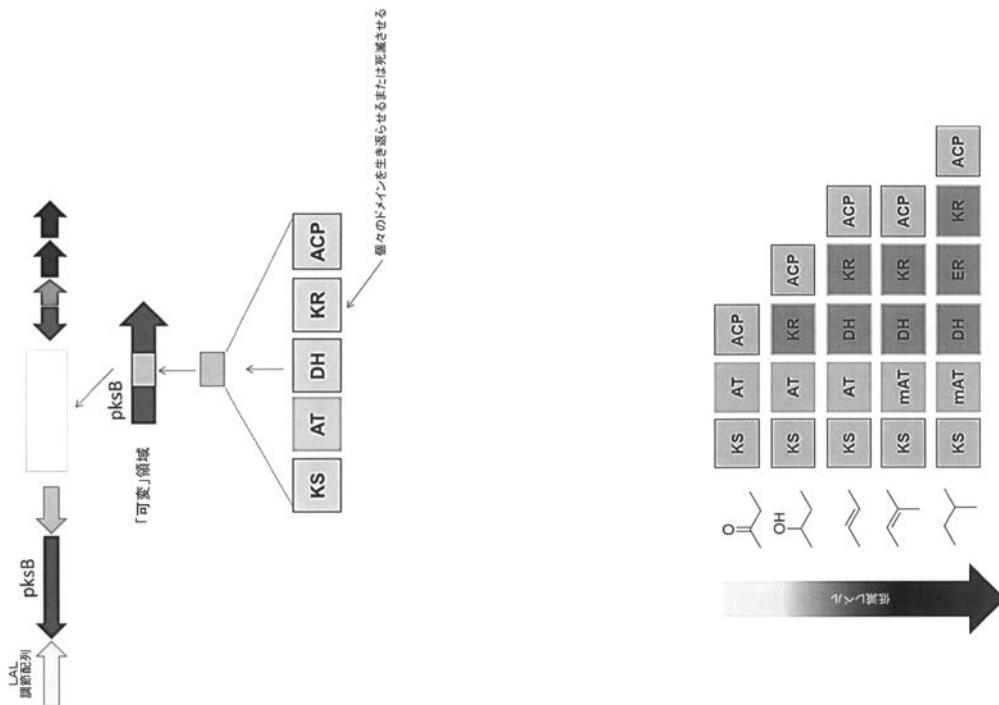
より当業者の知識から明らかでない限り、範囲として表される値は、文脈から別段のことが明らかに示されない限りその範囲の下限の単位の10分の1までの、異なる本発明の実施形態における言及された範囲内の任意の具体的な値または部分範囲であることを理解されたい。

【0095】

さらに、先行技術の範囲内にある本発明の任意の特定の実施形態は、請求項の任意の1つまたは複数からも明確に除外され得ることを理解されたい。このような実施形態は当業者に既知であると見なされるため、これらの実施形態は、除外されることが本明細書において明確に説明されていなくても、除外され得る。本発明の組成物の任意の特定の実施形態（例えば、任意のポリヌクレオチドまたはそれによってコードされるタンパク質、任意の生産方法、任意の使用方法）は、先行技術の存在に関連していてもいなくても、いかなる理由によっても、任意の1つまたは複数の請求項から除外され得る。

10

【図1】



【図 2B】

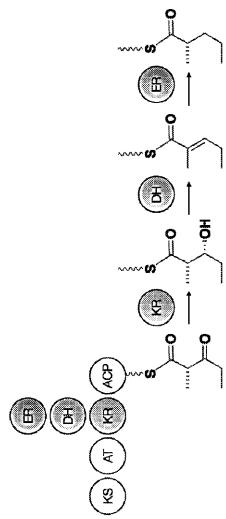
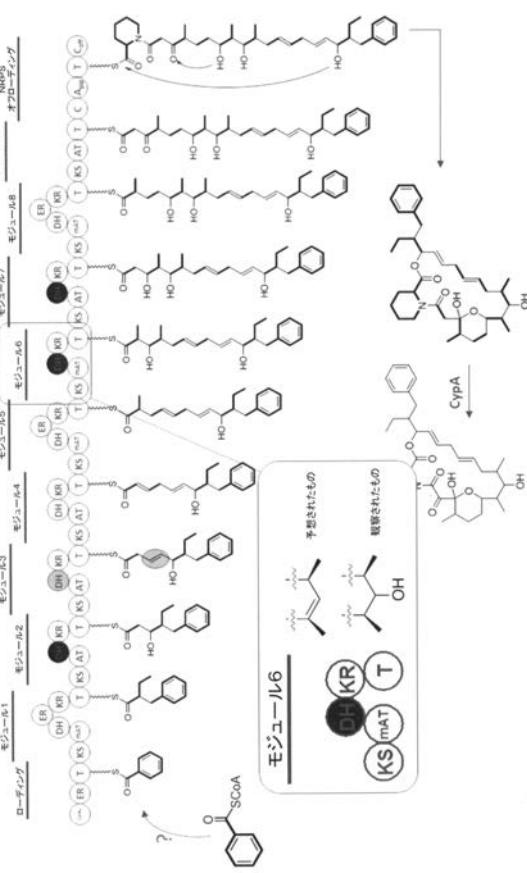


FIG. 2B

【図 3】



【図 4A】

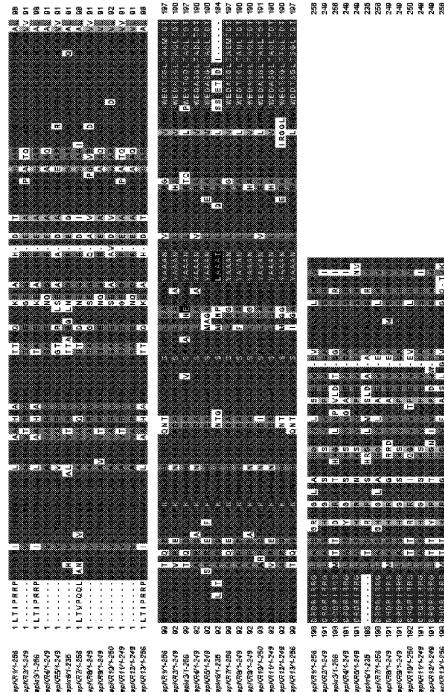


FIG. 4A

【図 4B】

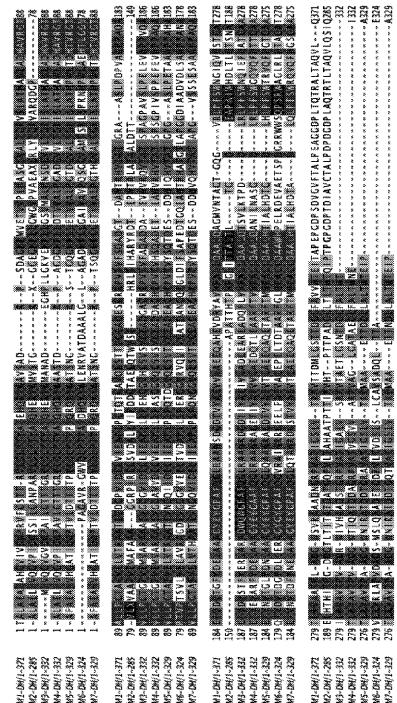
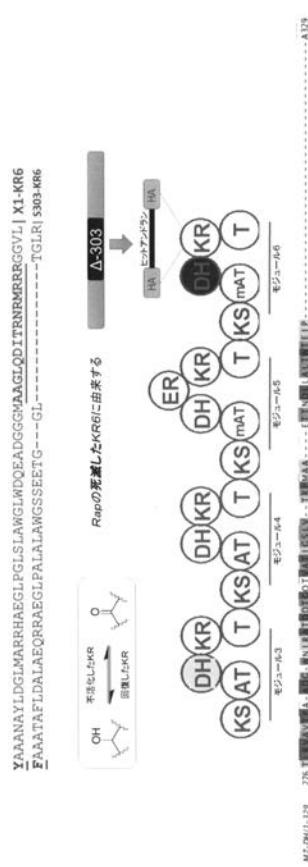
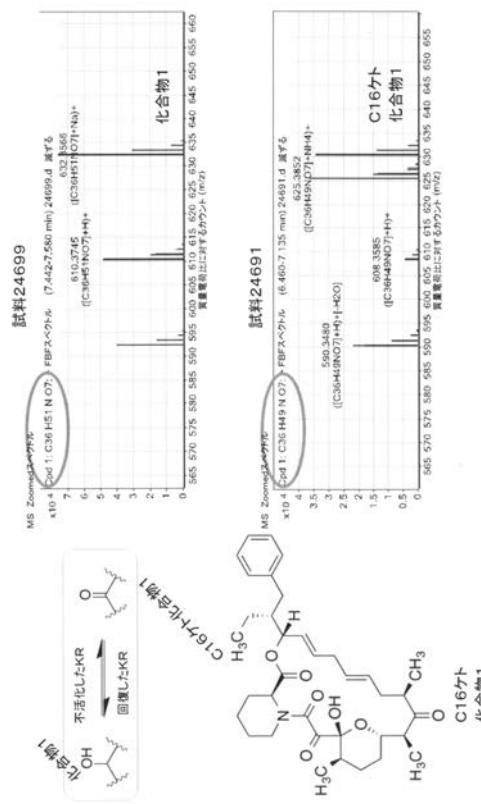


FIG. 4B

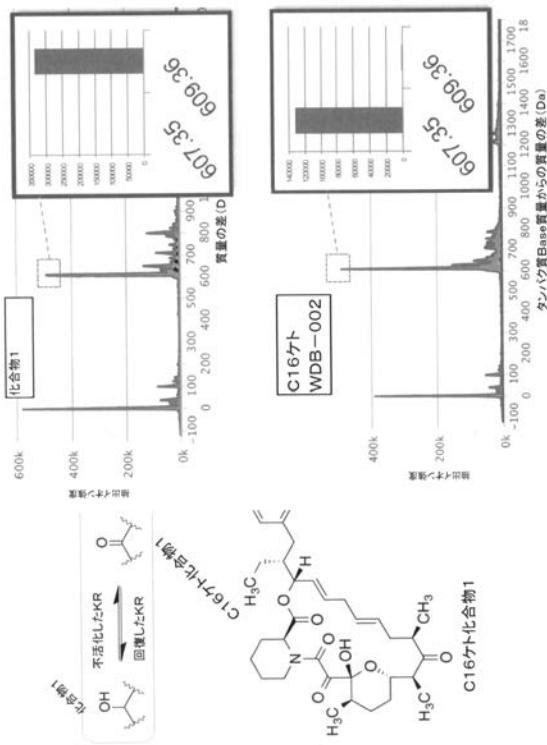
【 図 5 A 】



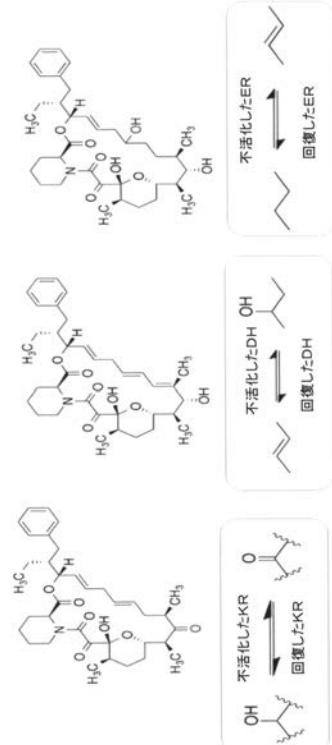
【図5B】



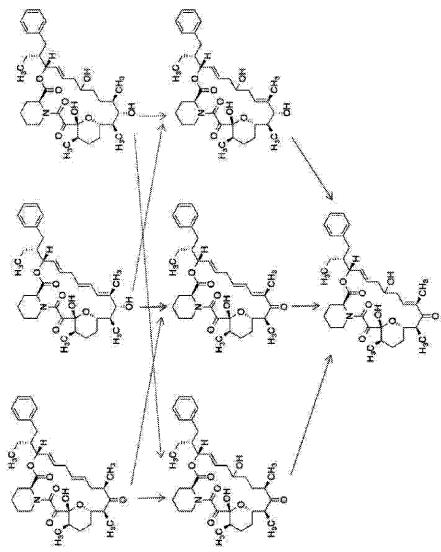
【図5C】



【図 6 A】



【図 6 B】



【 図 6 C 】

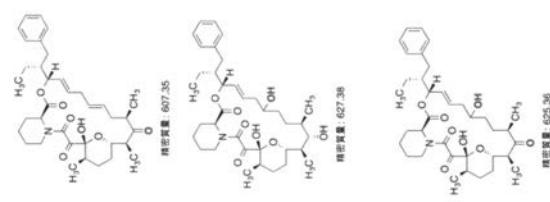
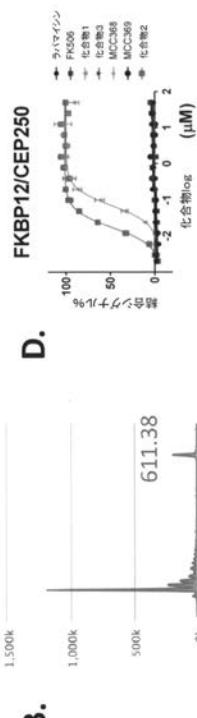
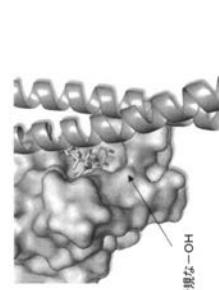
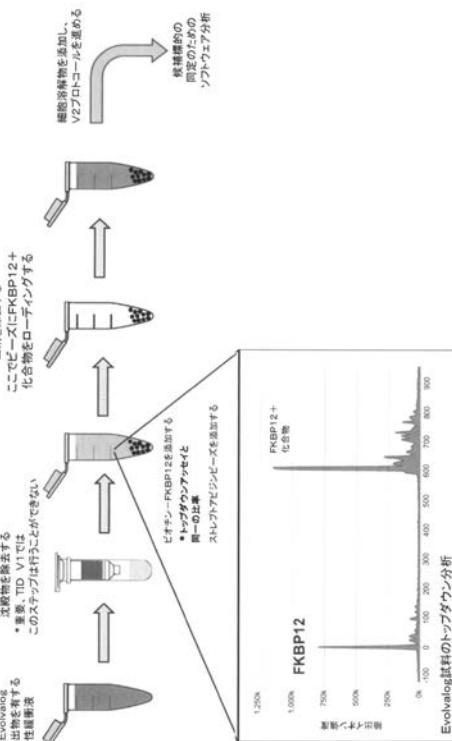


FIG. 6B

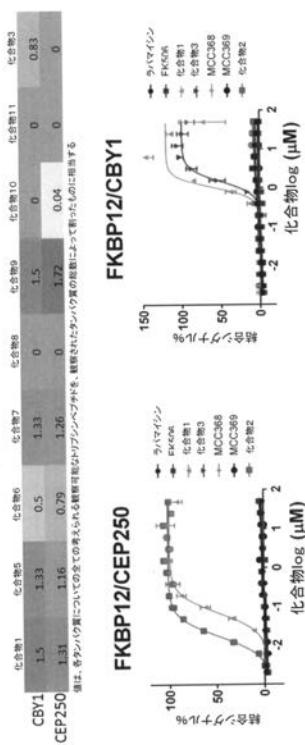
【 図 7 】



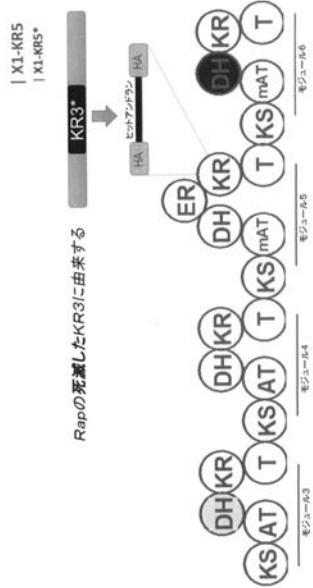
【図8A】



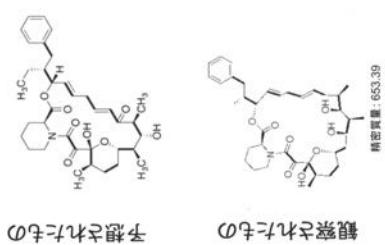
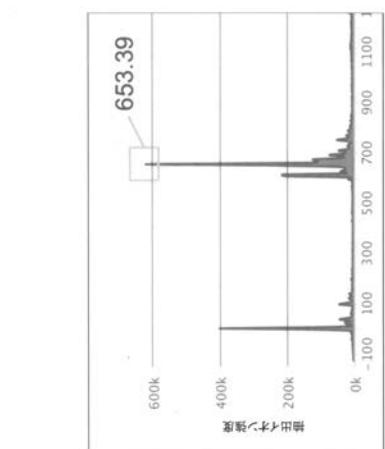
【図 8 B】



【図 9 A】



【図 9 B】



分子式: C26H34N2O10 構造式: C26H34N2O10

【手続補正書】

【提出日】平成30年9月4日(2018.9.4)

【手続補正1】

【補正対象書類名】図面

【補正対象項目名】図4 A

【補正方法】変更

【補正の内容】

【図4A】

配列番号:75 rapKR1/1-256	1LT1P RRP DPDGT I I TGGSGVLAG I ARLHAAEHQABHLLLSRTTDDQALIKELAELGAH . VDTAT CDS DRA GLARVLAGVSPEHPLTAVIHTAGAL 98
配列番号:76 rapKR2/1-249	1LT1P RRP DPDGT I I TGGSGVLAG I ARLHAAEHQABHLLLSRTTDDQALIKELAELGAH . VETAA CDS DPA ALTQVLAGVSPEHPLTAVIHTAGAL 91
配列番号:77 rapKR3/1-256	1LT1P RRP DPDGT I I TGGSGVLAG I ARLHAAEHQABHLLLSRTTDDQALIKELAELGAH . VETAA CDS DPA ALTQVLAGVSPEHPLTAVIHTAGAL 98
配列番号:78 rapKR4/1-249	1LT1P RRP DPDGT I I TGGSGVLAG I ARLHAAEHQABHLLLSRTTDDQALIKELAELGAH . VETAA CDS DPA ALTQVLAGVSPEHPLTAVIHTAGAL 91
配列番号:79 rapKR5/1-249	1LT1P RRP DPDGT I I TGGSGVLAG I ARLHAAEHQABHLLLSRTTDDQALIKELAELGAH . VETAA CDS DPA ALTQVLAGVSPEHPLTAVIHTAGAL 91
配列番号:80 rapKR6/1-235	1LT1P RRP DPDGT I I TGGSGVLAG I ARLHAAEHQABHLLLSRTTDDQALIKELAELGAH . VETAA CDS DPA ALTQVLAGVSPEHPLTAVIHTAGAL 91
配列番号:81 rapKR7/1-256	1LT1P RRP DPDGT I I TGGSGVLAG I ARLHAAEHQABHLLLSRTTDDQALIKELAELGAH . VETAA CDS DPA ALTQVLAGVSPEHPLTAVIHTAGAL 98
配列番号:82 rapKR8/1-249	1LT1P RRP DPDGT I I TGGSGVLAG I ARLHAAEHQABHLLLSRTTDDQALIKELAELGAH . VETAA CDS DPA ALTQVLAGVSPEHPLTAVIHTAGAL 91
配列番号:83 rapKR9/1-249	1LT1P RRP DPDGT I I TGGSGVLAG I ARLHAAEHQABHLLLSRTTDDQALIKELAELGAH . VETAA CDS DPA ALTQVLAGVSPEHPLTAVIHTAGAL 91
配列番号:84 rapKR10/1-250	1LT1P RRP DPDGT I I TGGSGVLAG I ARLHAAEHQABHLLLSRTTDDQALIKELAELGAH . VETAA CDS DPA ALTQVLAGVSPEHPLTAVIHTAGAL 92
配列番号:85 rapKR11/1-249	1LT1P RRP DPDGT I I TGGSGVLAG I ARLHAAEHQABHLLLSRTTDDQALIKELAELGAH . VETAA CDS DPA ALTQVLAGVSPEHPLTAVIHTAGAL 91
配列番号:86 rapKR12/1-249	1LT1P RRP DPDGT I I TGGSGVLAG I ARLHAAEHQABHLLLSRTTDDQALIKELAELGAH . VETAA CDS DPA ALTQVLAGVSPEHPLTAVIHTAGAL 91
配列番号:87 rapKR13/1-256	1LT1P RRP DPDGT I I TGGSGVLAG I ARLHAAEHQABHLLLSRTTDDQALIKELAELGAH . VETAA CDS DPA ALTQVLAGVSPEHPLTAVIHTAGAL 98
配列番号:75 rapKR1/1-256	99 DDGVESLT TDQ DTVLRPKADGAVHHLHETLQNTDLAFAVHLLSAGQANYAAANAFV DVALEAEDRGEGLPALAVAGLWEDTSGLTAKMTDT 197
配列番号:76 rapKR2/1-249	92 DDGVESLT TDQ DTVLRPKADGAVNHLHETLQNTDLAFAVHLLSAGQANYAAANAFV DVALEAEDRGEGLPALAVAGLWEDASGLTAQLTDT 190
配列番号:77 rapKR3/1-256	99 DDGVESLT TDQ DTVLRPKADGAVNHLHETLQNTDLAFAVHLLSAGQANYAAANAFV DVALEAEDRRTQGGLPALALAWG PWEYTGDLTAQLTGT 197
配列番号:78 rapKR4/1-249	92 DDGVESLT TDQ DTVLRPKADGAVNHLHETLQNTDLAFAVHLLSAGQANYAAANAFV DVALEAEDRRAEGPLALAWG LWE DASGLTAQLTDT 190
配列番号:79 rapKR5/1-249	92 DDGVESLT TDQ DTVLRPKADGAVNHLHETLQNTDLAFAVHLLSAGQANYAAANAFV DVALEAEDRRAEGPLALAWG LWE DASGLTAQLTDT 190
配列番号:80 rapKR6/1-235	92 DDGVESLT TDQ DTVLRPKADGAVNHLHETLQNTDLAFAVHLLSAGQANYAAANAFV DVALEAEDRRAEGPLALAWG LWE DASGLTAQLTDT 190
配列番号:81 rapKR7/1-256	99 DDGVESLT TDQ DTVLRPKADGAVNHLHETLQNTDLAFAVHLLSAGQANYAAANAFV DVALEAEDRRAEGPLALAWG LWE DASGLTAQLTDT 197
配列番号:82 rapKR8/1-249	92 DDGVESLT TDQ DTVLRPKADGAVNHLHETLQNTDLAFAVHLLSAGQANYAAANAFV DVALEAEDRRAEGPLALAWG LWE DASGLTAQLTDT 190
配列番号:83 rapKR9/1-249	92 DDGVESLT TDQ DTVLRPKADGAVNHLHETLQNTDLAFAVHLLSAGQANYAAANAFV DVALEAEDRRAEGPLALAWG LWE DASGLTAQLTDT 190
配列番号:84 rapKR10/1-250	93 DDGVESLT TDQ DTVLRPKADGAVNHLHETLQNTDLAFAVHLLSAGQANYAAANAFV DVALEAEDRRAEGPLALAWG LWE DASGLTAQLTDT 191
配列番号:85 rapKR11/1-249	92 DDGVESLT TDQ DTVLRPKADGAVNHLHETLQNTDLAFAVHLLSAGQANYAAANAFV DVALEAEDRRAEGPLALAWG LWE DASGLTAQLTDT 190
配列番号:86 rapKR12/1-249	92 DDGVESLT TDQ DTVLRPKADGAVNHLHETLQNTDLAFAVHLLSAGQANYAAANAFV DVALEAEDRRAEGPLALAWG LWE DASGLTAQLTDT 190
配列番号:87 rapKR13/1-256	99 DDGVESLT TDQ DTVLRPKADGAVNHLHETLQNTDLAFAVHLLSAGQANYAAANAFV DVALEAEDRRAEGPLALAWG LWE DASGLTAQLTDT 197
配列番号:75 rapKR1/1-256	100 DRDRIRRGOLRAISAGRHGG I LDAA SRHGEPLVLLASMEPVR - DVEVPALRL L HRFV I AR 250
配列番号:76 rapKR2/1-249	191 DRDRIRRGOLRAISAGRHGG I LDAA SRHGEPLVLLASMEPVR - DAEVPALRLSRSLH RPI AR 240
配列番号:77 rapKR3/1-256	198 DRDRIRRGOLRAISAGRHGG I LDAA SRHGEPLVLLASMEPVR - DAEVPALRLSRSLH RPI AR 250
配列番号:78 rapKR4/1-249	191 DRDRIRRGOLRAISAGRHGG I LDAA SRHGEPLVLLASMEPVR - DAEVPALRLSRSLH RPI AR 240
配列番号:79 rapKR5/1-249	191 DRDRIRRGOLRAISAGRHGG I LDAA SRHGEPLVLLASMEPVR - DAEVPALRLSRSLH RPI AR 240
配列番号:80 rapKR6/1-235	185 - - - - - CLRAT SAGRDRFEDSASHRGEPLLIMASLDPA - AAEVPALRLSRSLRPPV I AR 235
配列番号:81 rapKR7/1-256	108 DRDRIRRGOLRAISAGRHGG I LDAA SRHGEPLVLLASMEPVR - EAEVPALRLSRSLH RPPV I AR 250
配列番号:82 rapKR8/1-249	191 DHDRIRRGOLRAISAGRHGG I LDAA SRHGEPLVLLASMEPVR - EAEVPALRLSRSLH RPPV I AR 240
配列番号:83 rapKR9/1-249	191 DRDRIRRGOLRAISAGRHGG I LDAA SRHGEPLVLLASMEPVR - DAEVPALRLSRSLH RPPV I AR 240
配列番号:84 rapKR10/1-250	192 DHDRIRRGOLRAISAGRHGG I LDAA SRHGEPLVLLASMEPVR - EAEVPALRLSRSLH RPPV I AR 250
配列番号:85 rapKR11/1-249	191 DRDRIRRGOLRAISAGRHGG I LDAA SRHGEPLVLLASMEPVR - DAEVPALRLSRSLH RPPV I AR 240
配列番号:86 rapKR12/1-249	191 DRDRIRRGOLRAISAGRHGG I LDAA SRHGEPLVLLASMEPVR - DAEVPALRLSRSLH RPPV I AR 240
配列番号:87 rapKR13/1-256	198 DHDRIRRGOLRAISAGRHGG I LDAA SRHGEPLVLLASMEPVR - DAEVPALRLSRSLH RPPV I AR 250

【手続補正2】

【補正対象書類名】図面

【補正対象項目名】図4B

【補正方法】変更

【補正の内容】

【 図 4 B 】

【手続補正3】

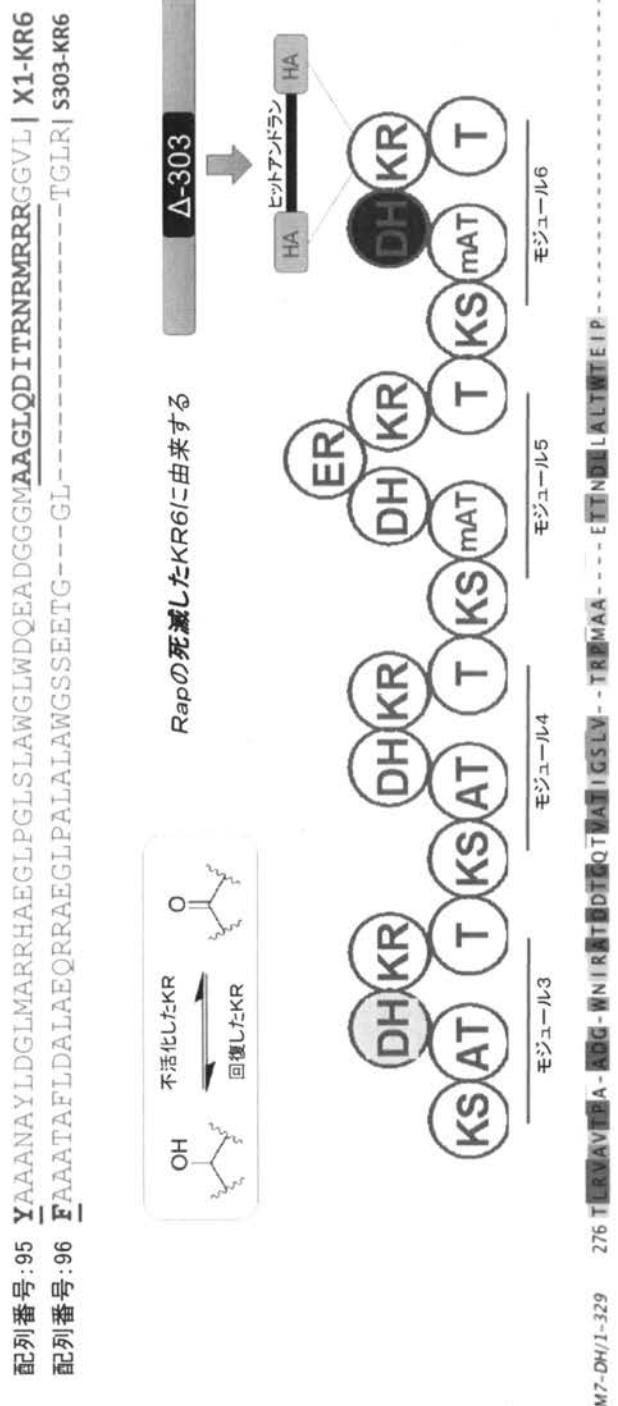
【補正対象書類名】図面

【補正対象項目名】図5A

【補正方法】変更

【補正の内容】

【図 5 A】



【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US 17/58805
A. CLASSIFICATION OF SUBJECT MATTER IPC(8) - C12P 19/62, C12N 9/10, C12N 15/74 (2018.01) CPC - C12P 19/62, C12N 9/10, C12N 15/12, C12N 15/74		
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) See Search History Document Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched See Search History Document Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) See Search History Document		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2003/0153053 A1 (Reid) 14 August 2003 (14.08.2003) para [0012], [0014], [0024], [0048], [0052]	1-4, 11-13
A	UniProtKB accession number Q54297, 01 November 1996, [retrieved 30 July 2018, https://www.uniprot.org/uniprot/Q54297.txt] whole doc.	22, 23, 26
A	UniProtKB accession number A0A061A618, 03 September 2014 [retrieved 30 July 2018, https://www.uniprot.org/uniprot/A0A061A618.txt] whole doc.	22, 23, 26
A	US 2011/0117606 A1 (Jergensen et al.) 19 May 2011 (19.05.2011) para [0034], Table 1, SEQ ID NO: 34	22, 23, 26
T	WO 2018/081592 A2 (Warp Drive Bio, Inc.) 03 May 2018 (03.05.2018) whole doc.	1, 11
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input 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 30 July 2018	Date of mailing of the international search report 27 AUG 2018	
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774	

INTERNATIONAL SEARCH REPORT		International application No.
PCT/US 17/58805		
Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)		
<p>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:</p> <p>a. <input type="checkbox"/> forming part of the international application as filed: <input checked="" type="checkbox"/> in the form of an Annex C/ST.25 text file. <input type="checkbox"/> on paper or in the form of an image file.</p> <p>b. <input type="checkbox"/> furnished together with the international application under PCT Rule 13<i>ter</i>.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.</p> <p>c. <input checked="" type="checkbox"/> furnished subsequent to the international filing date for the purposes of international search only: <input checked="" type="checkbox"/> in the form of an Annex C/ST.25 text file (Rule 13<i>ter</i>.1(a)). <input type="checkbox"/> on paper or in the form of an image file (Rule 13<i>ter</i>.1(b) and Administrative Instructions, Section 713).</p>		
<p>2. <input checked="" type="checkbox"/> 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.</p>		
<p>3. Additional comments:</p>		

INTERNATIONAL SEARCH REPORT		International application No. PCT/US 17/58805
Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)		
This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:		
1.	<input type="checkbox"/>	Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:
2.	<input type="checkbox"/>	Claims Nos.: because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3.	<input checked="" type="checkbox"/>	Claims Nos.: 5-10, 14-21, 27-60, 66 because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).
Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)		
This International Searching Authority found multiple inventions in this international application, as follows:		
- Please see extra sheet for Box No. III Observations where unity of invention is lacking -		
1.	<input type="checkbox"/>	As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.	<input type="checkbox"/>	As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3.	<input type="checkbox"/>	As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4.	<input checked="" type="checkbox"/>	No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-4, 11-13, 22, 23 and 26 limited to SEQ ID NOS: 1, 10, 11, 12
Remark on Protest		<input type="checkbox"/> The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee. <input type="checkbox"/> The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation. <input type="checkbox"/> No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 17/58805

Continuation of:
Box NO. III. Observations where unity of invention is lacking

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I+: Claims 1-4, 11-13 and 22-26, directed to an engineered/chimeric polyketide synthase (PKS) comprising one or more modified domains. The PKS will be searched to the extent that the modified domain(s) encompass ketoreductase domain(s) comprising an amino acid other than tyrosine at the position corresponding to the tyrosine in the conserved YAAAN catalytic motif and does not comprise the conserved α .FG helix in SEQ ID NO: 1 (and the first corresponding claimed ketoreductase domain amino acid sequences SEQ ID NOs: 1 and 12). It is believed that claims 1-4, 11-13, 22, 23 and 26 encompass this first named invention (because SEQ ID NOs: 10 and 11 of claim 22 include ketoreductase domains), and thus these claims will be searched without fee to the extent that the PKS encompasses modified ketoreductase domain(s). Additional modified PKSs will be searched upon the payment of additional fees. Applicants must specify the claims that encompass any additionally elected modified PKSs. Applicants must further indicate, if applicable, the claims which encompass the first named invention, if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched. An exemplary election would be PKS with modified dehydratase domain(s) comprising an aspartic acid at the position corresponding to the glycine at position 4288 in pksB of S679-pksB ORF in the conserved HXXXGXXXXP motif of SEQ ID NO: 4 (claims 1-4, 11-13, 22, 24, 26).

[Note, applicant specification defines different beta-ketone processing domain types (ketoreductase, dehydratase, or enoylreductase) and corresponding sequences at p. 21, ln 5-26]

Group II, claim 61-65, directed to a method of modulating the activity of a PKS, and a method of producing a compound or a library of compounds.

The inventions listed as Groups I+ and II do not relate to a single special technical feature under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Special technical features

Group I+ includes the special technical feature of a modified, chimeric PKS, that is not required by Group II.

Group II includes the special technical feature of modulating the activity of a PKS, comprising providing a parent nucleic acid sequence encoding a parent PKS, and modifying at least one codon, that is not required by Group I+.

The inventions of Group I+ each include the special technical feature of a unique type of beta-ketone processing domain. Each domain type has a different mechanism of action, is encoded by a unique peptide, and is considered a distinct technical feature.

Shared technical features

The inventions of Groups I+ and II share the common technical feature of an engineered polyketide synthase that comprises one or more modified domains having altered enzymatic activity relative to a reference polyketide comprising unmodified domains, wherein the engineered polyketide synthase is capable of producing a polyketide when expressed under conditions suitable to allow expression of a compound by the engineered polyketide synthase.

The inventions of Group I+ further share the common technical feature of a chimeric PKS, wherein the one or more domains comprise a beta-ketone processing domain selected from a ketoreductase, a dehydratase, or an enoylreductase; and a PKS comprising a first domain comprising a conserved region of a domain of a first polyketide synthase, and a second domain comprising a conserved region of a domain of a second polyketide synthase.

However, these shared technical features were previously disclosed by US 2003/0153053 A1 (Reid). Reid teaches an engineered chimeric polyketide synthase that comprises one or more modified ketoreductase or dehydratase domains having altered enzymatic activity relative to a reference polyketide comprising unmodified domains, wherein the engineered polyketide synthase is capable of producing a polyketide when expressed under conditions suitable to allow expression of a compound by the engineered polyketide synthase (para [0014] "a method to alter a module of a modular PKS such that said module will introduce a cis double bond into a polyketide produced by said PKS, said method comprising, either (A) replacing an entire module for the position at which the cis double bond is desired with a module having a type 2 KR and dehydratase (DH) domains, (B) exchanging a portion of a module between an AT and an ACP of said module for a DH plus a type 2 KR domain of another module, (C) in a module already producing a trans double bond, replacing a type 1 KR domain with a type 2 KR domain, (D) in a module containing a type 1 KR domain, changing the KR to a type 2 KR domain by point mutation or replacing the KR with a type 2 KR; and (E) inserting a DH into a module containing a type 2 KR"; [0051] "the expression plasmid...colonies from each transformation were selected to screen for polyketide production", said PKS comprising a first domain comprising a conserved region of a domain of a first polyketide synthase, and a second domain comprising a conserved region of a domain of a second polyketide synthase (para [0014] "changing the KR to a type 2 KR domain by point mutation or replacing the KR with a type 2 KR; and (E) inserting a DH into a module containing a type 2 KR"; [0024] "At the N-termini, a Rossmann fold region corresponds to the SDR Rossmann fold. An absolutely conserved Lys in the ketoreductase family corresponds to a very highly conserved Asn in the SDR superfamily generally, including the tropinone reductases").

As the technical features were known in the art at the time of the invention, they cannot be considered special technical features that would otherwise unify the groups.

Therefore, Groups I+ and II inventions lack unity under PCT Rule 13 because they do not share the same or corresponding special technical feature.

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 N 1/21 (2006.01)	C 1 2 N 1/21	
C 1 2 P 13/00 (2006.01)	C 1 2 P 13/00	

(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,R0,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,RW,SA,SC,SD,SE,SG,SK,SL,SM,ST,SV,SY,TH,TJ,TM,TN,TR,TT

(72)発明者 グレイ、ダニエル シー .

アメリカ合衆国 0 2 1 5 5 マサチューセッツ州 メドフォード ハーバード ストリート 2
4 6 アパートメント 3

(72)発明者 ボウマン、ブライアン アール .

アメリカ合衆国 1 0 8 0 4 ニューヨーク州 ニュー ロシェル シートン ドライブ 1 9 9

(72)発明者 バーダイン、グレゴリー エル .

アメリカ合衆国 0 2 1 1 0 マサチューセッツ州 ボストン ロウズ ワーフ 2 0 ユニット
4 1 0

(72)発明者 ゾーバ、マシュー エドワード

アメリカ合衆国 0 2 4 7 2 - 3 3 1 7 マサチューセッツ州 ウォータータウン ウェルコット
ロード 2 4

F ターム(参考) 4B050 CC04 CC05 DD02 LL05

4B064 AE01 CA19 CA21 CC24 DA01 DA20

4B065 AA50X AA50Y AB01 AC14 BA02 CA16 CA44 CA60