(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2013-514080
(P2013-514080A)
(43)公表日　平成25年4月25日(2013.4.25)

(51)Int.Cl.		F I			テーマコード（参考）
C12N 15/09	(2006.01)	C12N	15/00	A	2B030
C12Q 1/04	(2006.01)	C12Q	1/04		4B024
C12Q 1/68	(2006.01)	C12Q	1/68	A	4B063
A01H 1/00	(2006.01)	A01H	1/00	Z	
A01H 5/00	(2006.01)	A01H	5/00	Z	

審査請求 未請求　予備審査請求 未請求　（全 26 頁）

(21)出願番号　特願2012-544420 (P2012-544420)
(86)(22)出願日　平成22年12月17日 (2010.12.17)
(85)翻訳文提出日　平成24年8月2日 (2012.8.2)
(86)国際出願番号　PCT/NL2010/050860
(87)国際公開番号　WO2011/074964
(87)国際公開日　平成23年6月23日 (2011.6.23)
(31)優先権主張番号　61/287,927
(32)優先日　平成21年12月18日 (2009.12.18)
(33)優先権主張国　米国 (US)
(31)優先権主張番号　2004019
(32)優先日　平成21年12月24日 (2009.12.24)
(33)優先権主張国　オランダ (NL)

(71)出願人　509351340
キージーン・エン・フェー
オランダ・６７００・アーエー・ワーゲニンゲン・ピー・オー・ボックス・２１６
(74)代理人　100106002
弁理士　正林　真之
(74)代理人　100120891
弁理士　林　一好
(72)発明者　スツールマン　イェルーン
オランダ国　エヌエル－６７００　アーエー　ワーゲニンゲン　ピー．オー．ボックス

最終頁に続く

(54)【発明の名称】改善されたバルク変異体分析

(57)【要約】
本発明は、生物の中で発現され、かつその生物の特定の表現型（特徴の形質）と因果関係がある核酸配列の特定、および任意の単離、のための新しい戦略に関する。本発明の方法を用いると、当該技術分野で公知の方法とは対照的に、ゲノムに関して情報が入手できないかまたは限られた情報しか入手できない、例えば（作物）植物のような生物において、遺伝子を効率的に濃縮し、特定し、単離し、および／またはクローニングすることが可能になった。
【選択図】なし

【特許請求の範囲】
【請求項１】
生物の特徴に関連する発現された核酸配列の特定、および任意の単離のための方法であって、
ａ）前記特徴の形質Ａを有する前記生物の少なくとも２つのメンバーを準備する工程であって、形質Ａを有する前記メンバーは、形質Ｂを有する前記生物の同質遺伝子的なメンバー（複数可）に由来し、好ましくは独立に由来し、かつ形質Ａおよび形質Ｂは異なる工程と、
ｂ）形質Ａを有する工程ａ）の前記メンバーの各々からｃＤＮＡを得て、形質Ａを有する前記メンバーから得られた個々のｃＤＮＡの、少なくとも一部分、好ましくは各々のヌクレオチド配列を決定する工程と、
ｃ）形質Ａを有する他のメンバーの対応するｃＤＮＡとの比較によって、形質Ａを有するメンバーの各々の個々のｃＤＮＡの中の一塩基多型位置を決定する工程と、
ｄ）形質Ａを有する前記メンバーの少なくとも２つ、好ましくは形質Ａを有する前記メンバーのすべて、が少なくとも１つの一塩基多型を含有する個々のｃＤＮＡを特定する工程であって、好ましくはこれらの一塩基多型は、前記形質Ａを有する少なくとも２つの、好ましくはすべてのメンバーについてのｃＤＮＡの中で異なる位置にある工程と、
ｅ）任意に、工程ｄ）で特定されたｃＤＮＡの各々について、それが、生物の特徴に関連することを検証する工程と
を含むことを特徴とする方法。
【請求項２】
形質Ａを有する前記生物の前記メンバーは前記形質Ｂを有する前記生物のメンバーの変異生成によって得られたものであり、形質Ｂを有する前記メンバーは変異生成処理の前には同質遺伝子的であった、請求項１に記載の方法。
【請求項３】
前記変異生成の方法は点変異を誘導する、請求項２に記載の方法。
【請求項４】
変異生成は、非生物学的変異誘発因子の使用によって実施される、請求項２または請求項３に記載の方法。
【請求項５】
前記生物は、植物、好ましくはトマト、コショウ、ナス、レタス、ニンジン、タマネギ、ニラ、チコリー、ダイコン、パセリ、ホウレンソウ、メロン、キュウリからなる群から選択される作物植物である、請求項１から請求項４のいずれか１項に記載の方法。
【請求項６】
前記生物が植物であり、かつ工程（ａ）に先立って、劣性の形質Ａをコードする対立遺伝子および優性の形質Ｂをコードする第２の対立遺伝子についてヘテロ接合であるＦ１ハイブリッド集団が作成され、かつ前記Ｆ１ハイブリッド集団は突然変異を誘発され、かつ前記変異生成の後に前記Ｆ１集団から、形質Ａを有するメンバーが選択される、請求項１から請求項５のいずれか１項に記載の方法。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、生物の特徴に関連する発現された核酸配列の特定、および任意の単離のための方法に関する。
【背景技術】
【０００２】
表現型によってのみ定められる生命体の形質に関与する遺伝子をクローニングすること（いわゆる順遺伝子クローニング（ｆｏｒｗａｒｄ　ｇｅｎｅ　ｃｌｏｎｉｎｇ））は、遺伝学およびバイオテクノロジーにおいて、積年の課題である。特に（作物）植物の分野では、順遺伝子クローニングのための迅速かつ経済的な方法が大いに求められている。現

在の探索は、順遺伝子（ｆｏｒｗａｒｄ　ｇｅｎｅ）単離のスピードを高め、順遺伝子単離の費用を下げる方法を見出そうとしており、特に、遺伝子クローニングを実施することができる種の範囲を広げようとしている。遺伝子クローニングは、大量の反復するＤＮＡを含有する大きいかまたは特徴づけされていないゲノムを有する種では、とりわけ煩わしい。このような種は、レタス、コショウ、タマネギ、ニラ、マメ、小麦、ライムギ、大麦、ほか多数である。

【０００３】

順遺伝子クローニングの目的は、表現型からのみ知られ、分子情報が入手できないかもしくは非常に限られた分子情報しか入手できないかまたは配列情報が入手できない遺伝子の特定である。順遺伝学（ｆｏｒｗａｒｄ　ｇｅｎｅｔｉｃｓ）における出発点は、例えば、天然に存在する表現型バリアントまたは人工的に誘導された変異体であることができる。

【０００４】

実質的に、順遺伝子クローニングについて２つの群の方法が認識されている：地図作成（マッピング）戦略およびタグ化戦略。これらの戦略は補完的であり、互いがその固有の限界を有する。

【０００５】

地図作成戦略では、表現型は、マーカー支援減数分裂期組換え地図作成によって、染色体の最小のセグメントまで突き止められる。地図に基づくクローニングは多くの時間と労力を要する手順であり、特に豊富なゲノム情報、好ましくは全ゲノムＤＮＡ配列、が入手できる小さい群のモデル種で用いられる。

【０００６】

理論的には、地図に基づくクローニングは、生殖周期を通して繁殖するいずれの生物についても普遍的に適用できる手順である（非特許文献１）。しかしながら、実際は、根本的な生物学的な制約がある。最も重要なことであるが、地図に基づくクローニングは、注目する遺伝子の領域における減数分裂期組換えの良好な頻度に、決定的に依存する。第２に、地図に基づくクローニングは、例えばレタス、コショウ、タマネギ、ほか多数のような植物における特徴づけが不十分な大きいゲノムにおいてはますます困難になる。このようなゲノムでは、反復するＤＮＡがゲノムの大部分を構成し、これがＤＮＡマーカーの明確な地図作成および遺伝子間隔（ｇｅｎｅｔｉｃ　ｉｎｔｅｒｖａｌ）に及ぶ大きい近接するＤＮＡ配列のアセンブリを妨げるからである。

【０００７】

古典的なタグ化戦略では、十分に特徴づけられた生物学的挿入変異誘発因子（トランスポゾンまたはＴ－ＤＮＡ挿入物）は、遺伝子をクローニングするための媒体として使用されてきた。転位因子を遺伝子の中へと挿入することは、その挿入が、例えば遺伝子のコード領域または非コード領域で起こったかどうかによって、機能喪失または機能獲得、発現パターンの変化を導くことができ、または遺伝子機能に影響を与えないことができる。いずれの新たに生じる表現型に関与する遺伝子も、当然その挿入エレメントの近傍に存在し、従って、ゲノムＤＮＡの隣接片とともに挿入配列を読み出すことによりクローニングされる。

【０００８】

挿入変異誘発因子を用いるタグ化は、遺伝子をクローニングするための理論的には好ましいアプローチである。なぜなら、そのようなタグ化は、速く、減数分裂期組換えからは独立であり、そして遺伝子地図または豊富な配列情報などのこれまでのゲノム資源を必要としないからである（非特許文献２を参照）。タグ化により変異遺伝子をクローニングすることは、反復するゲノム配列によってもまたはゲノムサイズによっても妨げられない。

【０００９】

このタグ化アプローチは、自然の遺伝子タグ化（トランスポゾン）系が既知であったか、またはランダムに組み込むＤＮＡ（Ｔ－ＤＮＡ）を用いて大量の個体を形質転換することができた小数のモデル生物では成功であった（例えばトウモロコシを使用する非特許文

献３を参照）。この小さい群の種のほかでは、既知の挿入エレメントが存在しないことに起因するか、または集団サイズに対する物量面の制約に起因するかのいずれかによって、タグ化はまったく使用できないか、またはほとんど使用できない。物量面では、１ゲノムあたりの挿入配列の数は少ない（１～２００）ので、タグ化は、１つまたはいくつかの特定の変異体を見出すために数千もの個体の集団を必要とする。

【先行技術文献】

【非特許文献】

【００１０】

【非特許文献１】Ｐｅｔｅｒｓら、Ｔｒｅｎｄｓ　ｉｎ　Ｐｌａｎｔ　Ｓｃｉｅｎｃｅ、２００３年、第８巻、第１０号、４８４－４９１頁

【非特許文献２】Ｍａｅｓｓら、Ｔｒｅｎｄｓ　Ｐｌａｎｔ　Ｓｃｉ．、１９９９年３月、第４巻、第３号、９０－９６頁

【非特許文献３】Ｓｅｔｔｌｅｓｓら、ＢＭＣ　Ｇｅｎｏｍｉｃｓ、２００７年、第８巻、１１６頁

【発明の概要】

【発明が解決しようとする課題】

【００１１】

本発明の目的のうちの１つは、特定の表現型の顕在化に関わる核酸の効率的な特定を可能にする新しい方法、およびその使用を提供することである。当該方法は、対象の種の豊富な減数分裂期組換え地図作成および事前のゲノムＤＮＡ配列の知識を必要とせず、人工的に交雑し操作することができるいずれの（植物）種でも用いることができる。本発明の他の目的は、本発明の説明、実施形態および特許請求の範囲から明らかになるであろう。

【課題を解決するための手段】

【００１２】

本発明は、生物の中で発現され、かつその生物の特定の表現型（特徴の形質）と因果関係がある核酸配列の特定、および任意の単離、のための新しい戦略に関する。本発明の方法を用いると、当該技術分野で公知の方法とは対照的に、ゲノムに関して情報が入手できないかまたは限られた情報しか入手できない、例えば（作物）植物のような生物において、遺伝子を効率的に濃縮し、特定し、単離し、および／またはクローニングすることが可能になった。

【図面の簡単な説明】

【００１３】

【図１】本発明に係る方法（の実施形態）の概略図である。この概略図では、そして本願明細書に開示される本発明に当てはまるが、ホモ接合であるときに観察することができる表現型上の顕在化の根底にあり、かつ本発明に係る方法によって特定される対象の遺伝子である遺伝子に変異（黒丸）を保有する参照対立遺伝子についてヘテロ接合である単一の個体によって描かれる、同質遺伝子的な集団が示されている（図１Ａ、左図）。変異原性処理の後に、この集団は、参照対立遺伝子の表現型上の効果がヘテロ接合性の喪失によって観察可能になるように、遺伝子機能を排除または損なう新規な対立遺伝子（図１Ａ、右図、灰色の丸）を含有するであろう。図１Ａでは、５つの独立の非同義の対立遺伝子が描かれている。図１Ｂは、ヘテロ接合性の喪失に関連づけられる表現型の顕在化についての群として選択されるこれらの５つの植物の対立遺伝子組成を示す。図１Ｂの左図では、選択された表現型上の顕在化に関連しない遺伝子について、これらの５つの植物の遺伝子型が１～５として描かれており、父系のまたは母系の対立遺伝子が接辞「ａ」または「ｂ」を用いて示されている。このような遺伝子が植物および両親の中で関連する多型を含有しないということが理解できる。対照的に、選択された表現型と因果関係がある選択された遺伝子（図１Ｂ、右図）は、すべての５つの同質遺伝子的な植物に共通でありかつ親のうちの一方（この場合、親「ａ」）にはない非同義の、または別の態様で遺伝子的に中立でない対立遺伝子を１つ含有する。加えて、この選択された遺伝子は、これら５つの表現型的に選択された植物の各々に、新規かつ異なる非同義の、または別の態様で遺伝子的に中

立でない多形性を含有する。多形性のこのパターンは、非常に稀である必要があり、従って、選択された表現型によって規定される形質の根底にある核酸配列（遺伝子）を特定するために使用することができる。
【図２】表１。Ａ）ヘテロ接合性の喪失によって赤色の花の色を獲得した３つの独立のペチュニア変異体のトランスクリプトームをシーケンシングした後の合わせた結果（参照トランスクリプトーム）の要約。Ｂ）本発明の工程に従って、ユニジーンにおける多形性パターンに基づく選択基準を適用した後に残るユニジーンの数。参照トランスクリプトームの中のユニジーンの総数の百分率として表される、選択後の濃縮係数（ｅｎｒｉｃｈｍｅｎｔ　ｆａｃｔｏｒ）も与えられている。Ｃ）本発明に係る方法に基づく、残る遺伝子のランク付けされた一覧。３つの遺伝子は、適用したすべての基準に完全に適合した。赤色の花の色の根底にあることが事前に知られていた遺伝子はアントシアニジン３－Ｏ－グルコシルトランスフェラーゼであり、これは、２つの他の遺伝子とランク１を共有する。
【発明を実施するための形態】
【００１４】
定義
以下の説明および実施例では、いくつかの用語が使用される。そのような用語に与えられるべき範囲を含めて、本願明細書および特許請求の範囲の明快で一貫性のある理解を提供するために、以下の定義が与えられる。本願明細書中に特段の記載がない限り、使用されるすべての技術用語および科学用語は、本発明が属する技術分野の当業者が一般に理解する意味と同じ意味を有する。すべての刊行物、特許出願、特許および他の参考文献の開示は、参照によりその全体を本願明細書に援用したものとする。
【００１５】
対立遺伝子：相同染色体上で同じ位置を有することができかつ別の形質に関与する、遺伝子の少なくとも２つの選択的形態のうちの１つ。非限定的な例は花卉の花の色についての遺伝子であり、－　１つの遺伝子が花弁の色をコントロールする可能性もあるが、その遺伝子のいくつかの異なるバージョンまたは対立遺伝子が存在する可能性がある。１つのバージョンは赤色の花弁を生じるかもしれないが、他方で、別のバージョンは白色の花弁を生じるかもしれない。個々の花卉の生じた色は、その花卉が遺伝子についてどちらの２つの対立遺伝子を保有するか、およびその２つがどのように相互作用するかによるであろう。
【００１６】
特徴：生物の表現型の質に関する。特徴は、異なる形質で顕在化することができる。例えば、生物は、花の色を特徴として有する植物であることができ、赤色の花または白色の花がその特徴の形質ＡおよびＢであることができる。本発明の範囲内では、特徴の第１の形質を有する生物のメンバーが、その特徴の第２の形質を有するその生物のメンバーから表現型的に区別することができる限りは、特徴（または形質）はいずれであってもよい。これは、生物の検査によって直接観察することができる差だけに限定されるわけではなく、その生物のさらなる分析の際に、例えば特定の状況に対する（例えば特定の疾患に対する）耐性の分析の際に、またはこのような生物における特定の代謝産物の存在の分析の際に明らかになりうる特徴／形質をも含む。
【００１７】
本発明に関する範囲での「遺伝子の発現」または「発現された核酸」は、適切な調節領域に、特にプロモーターに作動可能に連結されているＤＮＡ領域が、生物活性がある、すなわち生物活性のあるタンパク質またはペプチドへと翻訳されうるか、またはそれ自体で活性であるＲＮＡへと転写されるプロセスを指す。
【００１８】
「遺伝子」は、好適な転写調節領域（例えばプロモーター）に作動可能に連結された、細胞の中でＲＮＡ分子（例えばｍＲＮＡ）へと転写される領域（転写された領域）を含むＤＮＡ配列である。従って、遺伝子は、いくつかの作動可能に連結された配列、例えば、プロモーター、例えば翻訳開始に関わる配列を含む５’非翻訳リーダー配列（５’ＵＴＲ

とも呼ばれ、これは、翻訳開始コドンの上流の転写されるｍＲＮＡ配列に対応する）、（タンパク質）コード領域（ｃＤＮＡまたはゲノムＤＮＡ）ならびに例えば転写終結点およびポリアデニル化部位を含む３’非翻訳配列（３’非翻訳領域、または３’ＵＴＲとも呼ばれる）を含んでもよい。

【００１９】

「同質遺伝子的な」は、遺伝子的に同一であることを意味する。同質遺伝子的な集団内の個々の細胞は、典型的には、単一の祖先の子孫であり、等しい遺伝子構成を有する。本発明の範囲内で、「同質遺伝子的な」は、ｃＤＮＡのレベルで、個々のメンバーが、自然変異、または、本発明の好ましい実施形態では、変異原性処理の結果として生じる可能性があるいずれかの点変異を除いて９５～９９％、さらには１００％同一であると解釈されるものとする。実際、用語「同質遺伝子的な」は周知であり、当業者によって理解される。

【００２０】

変異生成または変異原性処理：この用語は、本願明細書では、核酸、遺伝子、またはゲノムにおける変化の導入を導く処理に関し、例えば点変異および／または１０までの連続するヌクレオチドの挿入もしくは欠失の導入を導く。

【００２１】

核酸：本発明に係る核酸は、ピリミジンおよびプリン塩基、好ましくはそれぞれシトシン、チミン、およびウラシル、ならびにアデニンおよびグアニン、のいずれかのポリマーまたはオリゴマーを包含しうる（Ａｌｂｅｒｔ Ｌ．Ｌｅｈｎｉｎｇｅｒ、「Ｐｒｉｎｃｉｐｌｅｓ ｏｆ Ｂｉｏｃｈｅｍｉｓｔｒｙ」、Ｗｏｒｔｈ Ｐｕｂ．、１９８２年、７９３－８００頁を参照。この文献は、参照により、あらゆる目的のためにその全体を本願明細書に援用したものとする）。核酸は、ＤＮＡ（ｃＤＮＡを含む）、もしくはＲＮＡ、またはこれらの混合物であってもよく、持続的にまたは一過性に一本鎖のまたは二本鎖の形態（ホモ二本鎖、ヘテロ二本鎖、およびハイブリッド状態を含む）で存在してもよい。

【００２２】

シーケンシング：用語「シーケンシング」は、核酸試料、例えばＤＮＡまたはＲＮＡの中のヌクレオチドの順序（塩基配列）を決定することを指す。

【００２３】

形質：生物学では、形質は、同じ生物の（任意の）他の個々のメンバーと比較した、生物の個々のメンバーのいずれかの表現型上の特有の特徴に関する。本発明に関する範囲では、形質は、遺伝で受け継がれることができ、すなわちその生物の遺伝情報によって、その生物の次の世代に伝えることができる。形質の非限定的な例としては、例えば、視覚的に観察することができる形質または（例えば、代謝産物の存在またはレベルについての）分析によって観察することができる形質が挙げられる。

【００２４】

「同じ特徴の形質」または「当該（前記）特徴の形質」：１つの特徴について存在する（または目に見えるようになる）少なくとも２つの形質の群のいずれか１つ。例えば、特徴「花の色」の場合には、表現型の顕在化は、青色、赤色、白色などを含んでもよい。上の例では、青色、赤色および白色は、すべて同じ特徴の異なる形質である。

【００２５】

上記の目的は、驚くべきことに、添付の特許請求の範囲に記載される方法を提供することによって解決される。

【００２６】

より具体的には、生物の特徴に関連する発現された核酸配列の特定、および任意の単離のための方法であって、

ａ）当該特徴の形質Ａを有する当該生物の少なくとも２つのメンバーを準備する工程であって、この形質Ａを有するメンバーは、形質Ｂを有する生物の同質遺伝子的なメンバー（複数可）に由来し、好ましくは独立に由来し、かつ形質Ａおよび形質Ｂは異なる工程と

、
ｂ）この形質Ａを有する工程ａ）のメンバーの各々からｃＤＮＡを得て、形質Ａを有するメンバーから得られた個々のｃＤＮＡの、少なくとも一部分、好ましくは各々のヌクレオチド配列を決定する工程と、
ｃ）形質Ａを有する他のメンバーの対応するｃＤＮＡとの比較によって、形質Ａを有するメンバーの各々の個々のｃＤＮＡの中の一塩基多型位置を決定する工程と、
ｄ）形質Ａを有するメンバーの少なくとも２つ、好ましくは形質Ａを有するメンバーのすべて、が少なくとも１つの一塩基多型を含有する個々のｃＤＮＡを特定する工程であって、好ましくはこれらの一塩基多型は、この形質Ａを有する少なくとも２つの、好ましくはすべてのメンバーについてのｃＤＮＡの中で異なる位置にある工程と、
ｅ）任意に、工程ｄ）で特定されたｃＤＮＡの各々について、それが、生物の特徴に関連することを検証する工程と
を含むことを特徴とする方法が提供される。

【００２７】
任意におよび好ましくは、この方法は、形質Ａを有する少なくとも２つの、好ましくはすべてのメンバーが少なくとも１つの非同義の一塩基多型を有するｄ）で特定されたｃＤＮＡの各々から個々のｃＤＮＡを特定する工程（ｆ）を含むことができる。

【００２８】
次に、好ましい実施形態では、この方法は、形質Ａを有する少なくとも２つの、好ましくはすべてのメンバーは、正確に１つの非同義の一塩基多型を有する上記の工程（ｆ）において特定されたｃＤＮＡの各々から個々のｃＤＮＡを特定する工程（ｇ）を含んでもよい。

【００２９】
また、当該方法は、実施された工程（ｄ）、（ｆ）および／または（ｇ）において特定されたすべての個々のｃＤＮＡを列挙し、すべての個々のｃＤＮＡが適合する実施された工程（ｄ）、（ｆ）および／または（ｇ）の数に従って（最も多い適合が最高のランクである）、すべての個々のｃＤＮＡをランク付けする工程（ｈ）を含んでもよい。

【００３０】
この方法は、好ましくは、この特定されたｃＤＮＡが由来する発現された核酸配列を特定し、任意に，その発現された核酸配列を含む遺伝子をクローニングする工程も含んでもよい。

【００３１】
本発明は、現在の順クローニング戦略に関する上記の課題が、本発明に係る方法によって解決されうるということの実現に基づき、そして本発明は、いずれの生物にも適用でき、かつ１ゲノムあたり数千もの十分に検出可能な変異を生成して、物量面の集団の制約を可能な限り取り除く、非生物学的変異誘発因子を使用し、これを、ある表現型（形質Ａ）を示す（少なくとも２つのメンバーの）変異体コレクションから得られるトランスクリプトーム（ｃＤＮＡ）をシーケンシングすることと組み合わせ、そしてそれらに基づいて、この変異体コレクションのメンバーの各々において（しかし、必ずしも当該遺伝子／ｃＤＮＡ内の同じ位置にある必要はない）、少なくとも１つの一塩基多型、好ましくは少なくとも１つの非同義の一塩基多型を有するいずれかの遺伝子を特定することにより、生物学的変異誘発因子を用いてタグ化することに基づく方法の限界を克服する。加えて、この変異体コレクションは自然変異から得ることができる。

【００３２】
例えば変異誘発因子を用いた処理によって誘導された変異に加えて、さらなる同質遺伝子的な集団において発生する自然突然変異も使用することができる。換言すれば、形質Ｂを有する同質遺伝子的な集団において、（例えば自然突然変異に起因して）形質Ａを有する少なくとも２つのメンバーが見出される場合、その特徴および形質に関与する遺伝子は、本願明細書に開示される方法を用いて特定することができる。

【００３３】

いずれの生物における遺伝子も、点変異のような化学的に誘導された変異を用いてタグ化することができる（そして、これにより異なる表現型を顕在化させうる）ということが実現された。しかしながら、このような小さくかつランダムな変異（タグ）を全ゲノムの中でどのようにして検出するかという大きな問題が存在する。なぜなら、そのような変異は、特異的なＰＣＲ増幅のいずれの形態にも適していないからである。さらに、化学的に誘導された変異は、高頻度でランダムにゲノムに導入される。結果として、もとの種の遺伝子と比較して、多くの遺伝子が、変異を同時に含む可能性があり、観察される表現型に関与するか、または特定の形質を示すための遺伝子の正しい特定をさらに複雑にする。

【００３４】

本発明者らは、解決策は、クローニングされる対象の遺伝子座（しかし、これはまだ知られていない）に独立かつ異なる変異体対立遺伝子のコレクションを作成し、次いでこのコレクションに由来するすべてのｃＤＮＡを複数回リシーケンシングすることから得ることができるということを実現した。このコレクションの中で（すべての）ｃＤＮＡ配列を比較するとき、変異体プールの中の１つのｃＤＮＡが一貫した変異を示すであろう。これは、その変異体表現型の根底にある遺伝子である可能性が極めて高い。

【００３５】

換言すれば、および例えば、特定の形質Ｂ、例えば赤色の花、を顕在化させる同質遺伝子的な集団（の種子）を、ゲノムの中に点変異を誘導する変異誘発因子で処理することができる。この変異誘発因子を用いた処理の結果として、ランダムな変異が、各処理されたメンバーのゲノムに導入される。通常、処理後、処理されたメンバーのほとんどは、いまだ特定の形質Ｂを顕在化させるであろう。それらのメンバーでは、変異誘発因子を用いた処理によって導入された変異は、形質を変えなかった。換言すれば、これらのメンバーでは、その形質の特徴に関連する遺伝子は、当該形質が（例えば、形質Ａへと、例えば、黄色の花であるように）変化するようには変異していない。しかしながら、いくつかのメンバーでは、この変異誘発因子は、それらのメンバーではもともとの形質Ｂはもはや顕在化されないが形質Ａは顕在化されるというように、観察される形質の根底にある遺伝子において変異を誘導した可能性がある。

【００３６】

変異誘発因子のランダムな特徴に起因して、および形質Ａを有する少なくとも２つのメンバー、好ましくはより多くのメンバーが観察される場合、変異誘発因子によって導入され、かつ形質Ｂから形質Ａへの変化に関与する変異は、同じ遺伝子に位置するが、当該遺伝子の中の同じ位置にはないという可能性が非常に高い。例えば、形質Ａを有する第１のメンバーでは、変異はヌクレオチドＲに位置してもよく、形質Ａを有する第２のメンバーでは、変異はヌクレオチドＳに位置してもよく、そして形質Ａを有する第３、第４および第５のメンバーでは、変異は、それぞれ、位置Ｔ、Ｕ、およびＶに位置してもよい（Ｒ、Ｓ、Ｔ、Ｕ、およびＶはその遺伝子内の異なる位置である）。換言すれば、これらのメンバーは、すべて、互いに比較したとき、その関与する遺伝子の異なる位置に変異を示す。同時に、形質Ａを示すメンバーはゲノム全体にわたって他の位置にも変異を保有するが、現在形質Ａを示すメンバーの各々において別の発現された遺伝子が改変されている可能性は、事実上ゼロである。

【００３７】

形質Ａを有するメンバーから得られるｃＤＮＡをここで互いに比較することにより、形質Ａを有するメンバーの各々において少なくとも１つのヌクレオチドの改変を保有し、かつこの改変は、当該メンバーのうちの少なくとも２つにおいて、当該ｃＤＮＡ内の異なる位置に位置するｃＤＮＡが、特定されうる。このｃＤＮＡは同じ特徴の観察される形質に関与する。

【００３８】

本発明に係る方法では、（誘導された）対立遺伝子系列のほかの点では同質遺伝子的な遺伝的バックグラウンドの、（誘導された）対立遺伝子系列からプールが作製される。本発明に係る方法は、クローニングされる対象の遺伝子内に位置するＳＮＰを検出する。

【００３９】
当該方法は、コショウおよびタマネギなどの煩わしいと悪名高い作物を含めた人工的にハイブリダイズし変異化することができるすべての種に適用でき、遺伝子地図の存在とは独立しており、あらゆるサイズのおよび複雑度のゲノムにおいても機能するであろう。
【００４０】
従って、本発明に係る方法を用いると、非常に多くの遺伝子の中から、観察される形質に関わるまたは関与することになる可能性のある候補である遺伝子を特定することが、ここで可能になった。実際、本発明に係る方法を適用するとき、可能性のある候補の群は、例えば１０個、５個、３個、２個、またはさらには１個の候補遺伝子に迅速かつ確実に限定することができる。本発明に係る方法を用いると、特定の形質の顕在化に関与するかまたは関わる遺伝子の迅速な特定がここで可能になった。
【００４１】
より詳細に言えば、本発明に係る方法では、特定の表現型を有する（または特定の形質Ａを有する）生物の少なくとも２つの個々のメンバーが比較される。これらの少なくとも２つのメンバーは、（異なる）形質Ｂを有する同質遺伝子的な生物に由来する。
【００４２】
本発明については、形質Ａまたは形質Ｂを有する生物のメンバーが、（形質Ａの場合には）（例えば変異誘発因子処理によって、もしくは自然突然変異によって）当該生物の同質遺伝子的なメンバーに由来するか、または（形質Ｂの場合には）当該生物の同質遺伝子的なメンバーであること、換言すれば、当該生物が同じ遺伝的バックグラウンドを有する（例えば、同じ近交系に由来する）ことが重要である。形質Ａまたは形質Ｂを有する個々のメンバーは同質遺伝子的であり、すなわち、自然変異に起因して、または、本発明の好ましい実施形態では、変異原性処理に起因して遺伝子材料に導入された変化を除いて、同じ遺伝的バックグラウンドを有する。
【００４３】
観察される表現型が含まれるのは、形質Ａを有するメンバーと形質Ｂを有するメンバーとの間の、遺伝子材料におけるこれらの差異においてである。（形質Ａから形質Ｂへの）特徴の観察される変化の結果として、形質Ａを有するメンバーは、それらが、形質Ｂを有する当該生物のメンバーによって発現された核酸とは異なり、かつ当該生物（形質Ａおよび形質Ｂはこの生物の表現型形態である）の特徴と関連する、少なくとも１つの核酸をここで発現するように改変される必要がある。
【００４４】
好ましくは、形質Ａを有するこの少なくとも２つのメンバーは、独立に、形質Ｂを有する当該生物の同質遺伝子的なメンバーに由来する。
【００４５】
用語「独立に…に由来する」は、形質Ａを有するこの少なくとも２つのメンバーが、例えば変異誘発因子処理の結果として、互いに独立に、当該形質Ａを得たメンバーであるということを示す。当該生物が植物である場合、この少なくとも２つのメンバーは、各々、例えば植物が生育するもとになる種子の変異誘発因子処理の際に、形質Ａを獲得したかもしれない。そのような例では、この形質は、形質Ａを顕在化させるいずれかのさらなるメンバーから独立に第１のメンバーに導入される。換言すれば、好ましくは、形質Ａを有するメンバーは、例えば、すでにこの形質Ａを顕在化させる最後の３世代、４世代、５世代、６世代、７世代、８世代、９世代または１０世代にわたって共通の先祖を共有しない。好ましくは、すべてのメンバーは、独立に、形質Ｂを有する当該生物の同質遺伝子的なメンバーに由来する。
【００４６】
上記のように、２つの異なる形質の存在に顕在化されるように、形質Ａを有するメンバーは、形質Ｂを有する他のメンバーによって発現された核酸とは異なり、かつ当該生物（形質Ａおよび形質Ｂはこの生物の表現型形態である）の特徴と関連する少なくとも１つの核酸を発現する。

【００４７】
このような（例えば点変異によって）タグ化された核酸を検出するために、本発明の方法では、形質Ａを有する工程ａ）のメンバーの各々に由来するｃＤＮＡ、好ましくは全ｃＤＮＡが得られ、これらの個々のｃＤＮＡのヌクレオチド配列、好ましくはこの形質Ａを有するメンバーから得られたｃＤＮＡの各々が決定される。換言すれば、形質Ａを有するすべてのメンバーのこれらのトランスクリプトーム、好ましくは全トランスクリプトームは、独立に形質Ａを獲得したすべての他のメンバーのトランスクリプトーム、好ましくは全トランスクリプトームと比較される。好ましくは、選択された組織の中のすべての発現された遺伝子の完全配列が比較される。

【００４８】
これについて、ｃＤＮＡ、好ましくは全ｃＤＮＡ、は形質Ａを有するこの少なくとも２つのメンバーから得られる。ｃＤＮＡ、好ましくは全ｃＤＮＡの調製は、いずれかの好適な当業者に公知の方法によって行うことができる。例えばＡＢｇｅｎｅ、Ａｍｂｉｏｎ、Ａｐｐｌｉｅｄ Ｂｉｏｓｙｓｔｅｍｓ、ＢｉｏＣｈａｉｎ、Ｂｉｏ－Ｒａｄ、Ｃｌｏｎｔｅｃｈ、ＧＥ Ｈｅａｌｔｈｃａｒｅ、ＧｅｎｅＣｈｏｉｃｅ、Ｉｎｖｉｔｒｏｇｅｎ、Ｎｏｖａｇｅｎ、Ｑｉａｇｅｎ、Ｒｏｃｈｅ Ａｐｐｌｉｅｄ Ｓｃｉｅｎｃｅ、Ｓｔｒａｔａｇｅｎｅなどから、ｃＤＮＡ合成のための多くの市販のキットを購入することができる。このような方法は、例えばＳａｍｂｒｏｏｋら（Ｓａｍｂｒｏｏｋ，Ｊ．、Ｆｒｉｔｓｃｈ，Ｅ．Ｆ．、およびＭａｎｉａｔｉｓ，Ｔ．、「Ｍｏｌｅｃｕｌａｒ Ｃｌｏｎｉｎｇ： Ａ Ｌａｂｏｒａｔｏｒｙ Ｍａｎｕａｌ」、Ｃｏｌｄ Ｓｐｒｉｎｇ Ｈａｒｂｏｒ Ｌａｂｏｒａｔｏｒｙ Ｐｒｅｓｓ、ＮＹ、第１、２、３巻（１９８９））に記載されている。当業者なら、ｃＤＮＡは、例えば個々のｃＤＮＡの提示（ｒｅｐｒｅｓｅｎｔａｔｉｏｎ）の差異を少なくするためのノーマライゼーション（ｎｏｒｍａｌｉｓａｔｉｏｎ）（Ｅｖｒｏｇｅｎなど）によってさらに処理することができるということも理解するであろう。

【００４９】
示したように、本発明に係る方法のためには、形質Ａを有する少なくとも２つのメンバーのｃＤＮＡ、好ましくは全ｃＤＮＡが得られ、シーケンシングされ、そして互いに比較されることが必要とされる。形質Ｂを有するメンバーから得られるｃＤＮＡとの比較は、これまでに説明したように、必要とはされない。より好ましくは、各々が形質Ａを有する当該生物の少なくとも３、４、５、６、７または８のメンバーが、本発明に係る方法で準備され／使用される。

【００５０】
これまでに説明したように、本発明は、独立に表現型上の形質（形質Ａ）を獲得した生物の（好ましくは全体の）トランスクリプトームの配列を比較することにより、特定の特徴（または形質）に関連する遺伝子の差異（ｇｅｎｅｔｉｃ ｄｉｆｆｅｒｅｎｃｅ）を特定することができるということを実現したことに基づく。好ましい実施形態では、形質Ａを有するメンバーは形質Ｂを有するメンバーのランダムな（化学的）変異生成処理によって生成されるので、このような処理された生物の遺伝子材料は、多くのおよびランダムな変異を含むであろうし、そしてそれらの変異のほとんどは観察される形質とは関連づけられないであろう。従って、形質Ａを有する１つのメンバーのｃＤＮＡ、好ましくは全ｃＤＮＡを、形質Ｂを有する１つのメンバーのｃＤＮＡ、好ましくは全ｃＤＮＡと比較することで、観察される表現型と関連する核酸の特定は可能にはならないであろう。しかしながら、各々が好ましくは独立に当該形質Ａを獲得した形質Ａを有する少なくとも２つのメンバーが互いに比較されるとき、その特徴（または形質）に関与する（または関連づけられる）核酸を特定することが可能になるということが、本発明者らによって実現された。

【００５１】
形質Ａを有するメンバーは、（独立に）同じ同質遺伝子的な源に由来する。この同質遺伝子的な源の変異処理の結果として、変化がランダムに遺伝子材料に導入される。第１のメンバーにおいて誘導される変化は、第２のメンバーにおいて誘導される変化とは異なる

であろう。しかしながら、これらの第１のメンバーおよび第２のメンバーの両方が、変異原性処理の結果として、ここである表現型を発現する、すなわち形質Ａを有する場合、両方のメンバーにおいて同じｃＤＮＡが一塩基多型（ＳＮＰ）を呈示するであろうという可能性が極めて高い。この変異処理のランダムなプロセスに起因して、このようなＳＮＰは、そのようなｃＤＮＡ内で同じ位置に位置する可能性は非常に低い。このｃＤＮＡは、ここで、検討対象の生物の特定の形質または特徴に関連づけられる発現された核酸に由来するｃＤＮＡとして特定されることができる。なぜなら、関連づけられないｃＤＮＡが、形質Ａを有するその少なくとも２つのメンバーにおいてそのようなＳＮＰを呈示するという可能性はゼロに近いからである。

【００５２】

換言すれば、本発明は、例えば１つの遺伝子（形質Ａ）における５つの対立遺伝子変異の「変異体プール」に由来するすべてのｃＤＮＡの配列において、互いに比較したときに一貫して配列における差を示している個々のｃＤＮＡはただ１つ存在するであろうということを実現したことに基づく。

【００５３】

形質Ａを有する各メンバーのｃＤＮＡ、好ましくは全ｃＤＮＡ、を比較するために、形質Ａを有するそのメンバーのｃＤＮＡ、好ましくは全ｃＤＮＡはシーケンシングされる必要がある。ｃＤＮＡのシーケンシングは、当業者に公知のいずれの好適な方法によっても行うことができる。しかしながら、特にＭａｒｇｕｌｉｅｓ Ｍ.ら（ｈｔｔｐ：／／ｗｗｗ.４５４.ｃｏｍ、「Ｇｅｎｏｍｅ ｓｅｑｕｅｎｃｉｎｇ ｉｎ ｍｉｃｒｏｆａｂｒｉｃａｔｅｄ ｈｉｇｈ－ｄｅｎｓｉｔｙ ｐｉｃｏｌｉｔｒｅ ｒｅａｃｔｏｒｓ（微細加工した高密度ピコリットル反応装置でのゲノム塩基配列決定）」、Ｎａｔｕｒｅ、２００５年、第４３７巻、３７６－８０頁］によって記載されるもののような方法が非常に好ましく、この方法は、全トランスクリプトーム（すべてのｃＤＮＡ）の迅速かつ効率的なシーケンシングを可能にする。例えば、および好ましい実施形態では、得られるｃＤＮＡ断片のヌクレオチド配列は、例えば国際公開第０３／００４６９０号パンフレット、国際公開第０３／０５４１４２号パンフレット，国際公開第２００４／０６９８４９号パンフレット、国際公開第２００４／０７０００５号パンフレット、国際公開第２００４／０７０００７号パンフレット、および国際公開第２００５／００３３７５号パンフレット、Ｓｅｏら、Ｐｒｏｃ.Ｎａｔｌ.Ａｃａｄ.Ｓｃｉ.ＵＳＡ、２００４年、第１０１巻、５４８８－９３頁に開示されるもの、ならびにＨｅｌｉｃｏｓ、Ｓｏｌｅｘａ、ＵＳ Ｇｅｎｏｍｉｃｓなどの技術（これらは、参照により本願明細書に援用したものとする）のような高スループットシーケンシング方法によって決定される。

【００５４】

本発明に係る方法の次の工程では、ｃＤＮＡにおける一塩基多型頻度を確立するために、形質Ａを有するメンバーのｃＤＮＡ、好ましくは全ｃＤＮＡの得られる配列は、形質Ａを有する他のメンバーのｃＤＮＡ、好ましくは全ｃＤＮＡと比較される。この判定は、当業者に公知のいずれの好適な方法によっても、および例えば後述する実施例に示される方法によって行うことができる。

【００５５】

要するに、例えば、ｃＤＮＡ断片のヌクレオチド配列のアラインメントを使用して、同じ転写された遺伝子に由来する（しかし、形質Ａを有する異なるメンバーに由来する）ヌクレオチド配列を集め、これらのヌクレオチド配列を比較してもよい。ヌクレオチド配列が同じ転写された遺伝子に由来するかどうかは、配列間の相同性に基づいて確立することができる。本発明の目的のために、ヌクレオチド配列が、少なくとも３０、好ましくは少なくとも５０、より好ましくは少なくとも９０、なおより好ましくは少なくとも１００、１５０、２００ヌクレオチドの長さにわたって少なくとも９５、９６、９７、９８、９９、１００％相同的であるとき、それらのヌクレオチド配列は、同じ転写された遺伝子に由来するとみなされる。

【００５６】

次の工程では、そして得られたデータに基づいて、増加したＳＮＰ頻度を有する形質Ａを有するメンバーのｃＤＮＡ、好ましくは全ｃＤＮＡのコレクションの中のｃＤＮＡを特定することができる。本発明に係る方法では、形質Ａを有するメンバーの少なくとも２つが少なくとも１つの一塩基多型を含有し、かつ好ましくはこれらの一塩基多型がｃＤＮＡの中の異なる位置にある、個々のｃＤＮＡが特定される。好ましくは、形質Ａを有する分析されたメンバーの各々は、そのｃＤＮＡの中に、好ましくはｃＤＮＡの中の異なる位置に少なくとも１つの一塩基多型を含有する。例えば、形質Ａを有する４つの植物が分析される場合、これら４つの植物の（同じ、すなわち、これらの植物の各々において同じ遺伝子に由来する）ｃＤＮＡの中に、植物１、２、３、および４についてそれぞれ位置２５、３４、５６および１０１にＳＮＰを含有するｃＤＮＡが特定されてもよい。あるいは、このＳＮＰは、いくつかの植物について同じ位置にあってもよいが、いくつかの他の植物についてはそうではなくてもよい。例えば、これらのＳＮＰは、植物１、２、３、および４についてそれぞれ位置２５、２５、５６および１０１、またはさらには２５、２５、１０１、および１０１にあってもよい。

【００５７】

本発明に係る方法では、形質Ａを有するメンバーの中のシーケンシングされるすべてのｃＤＮＡの中で、少なくとも１つの特定のｃＤＮＡが、形質Ａを有する少なくとも２つのメンバー、しかし好ましくは形質Ａを有するメンバーの各々、において少なくとも１つの遺伝子上の（好ましくは非同義の）一塩基多型（ＳＮＰ）を保有するであろう。

【００５８】

換言すれば、この特定のｃＤＮＡは、形質Ａを有する分析されたメンバーの各々において（必ずしも対応するｃＤＮＡの同じ位置にではない）少なくとも１つのＳＮＰを含むであろう。このようなｃＤＮＡは、観察される形質Ａに関わるかまたは関与する遺伝子に由来するｃＤＮＡである可能性が高い候補であり、この情報は、その後、発現された核酸配列を特定し、当業者に公知の方法によって対応する遺伝子をクローニングするために使用することができる。

【００５９】

任意におよび好ましくは、本発明に係る方法の次の工程では、形質Ａを有する少なくとも２つの、好ましくはすべての、メンバーが少なくとも１つの非同義の一塩基多型を有する、これまでの工程で特定されたｃＤＮＡの各々からの個々のｃＤＮＡが、特定される。

【００６０】

当業者なら理解するとおり、観察される表現型（形質Ａ）に関わるか、またはそれに関連するかまたはその原因となるために、これまでの工程で特定されたｃＤＮＡにおいて見出される少なくとも１つのＳＮＰは、コードされるタンパク質の中のアミノ酸の変化に関連づけされるに違いないという可能性が非常に高い。それゆえ、少なくとも１つの非同義のもの（コドンの変化を導き、コードされるアミノ酸の変化を導く）の存在を特定することが、本発明に係る方法をさらに改善し、観察される形質Ａに関わる遺伝子のさらにより信頼できる特定を可能にする。

【００６１】

実際、当業者は、本発明に関する用語「非同義（の）」を、塩基のトリプレットにおける少なくとも１つのヌクレオチドの変化に起因して、コドンの変化が発生し、この変化したコドンがもとのコドンと比較して異なるアミノ酸をコードするということを示すものとして理解する。

【００６２】

形質Ａを有する少なくとも２つの、好ましくはすべての、メンバーが少なくとも１つの非同義の一塩基多型を有するｃＤＮＡを特定するというこれまでの工程が実施される場合、任意に、しかし好ましくは、次の工程において、少なくとも２つの、しかし好ましくは形質Ａを有するすべてのメンバーが１以下の非同義の一塩基多型しか有しない、これまでの工程のｃＤＮＡの各々から個々のｃＤＮＡが特定される。

【００６３】

実際に、そのようなｃＤＮＡが、観察される表現型（形質Ａ）に関与するかまたは観察される表現型（形質Ａ）と因果関係がある遺伝子に由来する可能性が高い候補であり、従って、本発明に係る方法がさらに改善されることが、本発明者らによって見出された。

【００６４】

本発明に係る方法のさらなる改善は、これまでの工程のいずれかで特定されたｃＤＮＡの各々から個々のｃＤＮＡを特定することを含むことができ、この特定において、少なくとも２つの、好ましくは形質Ａを有するすべてのメンバーは進化上保存されたアミノ酸位置でアミノ酸変化を引き起こす非同義の一塩基多型を有し、好ましくはこれらの新しいアミノ酸は、もとの変異していないアミノ酸とは化学的に異なる（すなわち芳香族　対　非芳香族、極性　対　非極性、特定のｐＨにおいて陰性、陽性または中性、正または負の疎水性親水性指標（親水性または疎水性など）。

【００６５】

本発明に係る方法で実施されるこれまでに開示された工程によっては、実施された工程（ｄ）、（ｆ）、（ｇ）、および／または（ｈ）で特定されたすべての個々のｃＤＮＡは、それらｃＤＮＡが適合する（実施された）工程ｄ～ｆの数に従ってランク付けすることができる。次の工程で、すべての工程（ｄ）、（ｆ）、（ｇ）、および／または（ｈ）に適合して、従って最高ランクにある発現された核酸配列を、観察される表現型に関わる、観察される表現型に関与するまたは観察される表現型と因果関係がある可能性が最も高い候補として特定することができる。

【００６６】

上記の事柄に基づいて、発現された核酸配列を特定することができ、その遺伝子をクローニングすることができる。

【００６７】

好ましい実施形態では、当該形質Ａを有する生物のメンバーは形質Ｂを有する当該生物のメンバーの変異生成によって得られたものであり、この形質Ｂを有するメンバーはこの変異生成処理の前には同質遺伝子的であった、本発明に係る方法が提供される。

【００６８】

本発明に係る方法の１つの実施形態では、異なる特徴の形質Ａを有する少なくとも２つのメンバーは、自然変異の結果であることができ、すなわち以前に形質Ｂを顕在化させていた生物における核酸の自然のまたは自発的な変化に起因してもよい。このような遺伝情報における変異は、意図的なものではなく、その生物が（例えば、形質Ｂから形質Ａへの）表現型における観察される変化に関与する遺伝情報を保有するということを明らかにする。

【００６９】

しかしながら、好ましくは、本発明に係る方法は、このような意図的なものではなく制御できない遺伝情報のばらつきには依存しないが、代わりに、計画的な変異生成の結果である特定の形質を有する変異体に依存する。当業者は、例えば既知の変異誘発因子の使用によって任意の生物の遺伝情報を変異させることができる方法を理解する。このような変異誘発因子の使用によれば、変異は、生物のメンバーのゲノムにおいてランダムに発生することができる。換言すれば、遺伝子のような核酸は、（化学的に、生物学的にまたは放射線によって）誘導された変異（例えば、例えばメタンスルホン酸エチルによる点変異）を用いて、（変異していない核酸との比較によって）タグ化することができる。

【００７０】

「照射による変異生成」の例としては、Ｘ線、γ線、ＵＶ光、または電離粒子が挙げられる。「生物学的変異生成」の例としては、例えば、例えばｒｅｃＡのようなリコンビナーゼを使用する国際公開第０１５０８４７号パンフレットに記載されているもののような方法が挙げられる。本発明に係る方法で使用することができる「化学的変異誘発因子」の例としては、メタンスルホン酸エチル（ＥＭＳ）、硫酸ジエチル（ＤＥＳ）、Ｎ－ニトロソ－Ｎ－エチル尿素（ＥＮＵ）、ジエポキシブタン、２－アミノプリン、５－ブロモウラシル、臭化エチジウム、亜硝酸、ニトロソグアニジン、ヒドロキシルアミン、アジ化ナト

リウム、またはホルムアルデヒドなどの特定の化学物質の使用が挙げられる。

【００７１】

好ましくは、変異生成の方法は、ゲノムにおける点変異、または１０までの連続するヌクレオチドの挿入、置換もしくは欠失を誘導する。

【００７２】

当業者なら、この実施形態に関する範囲内で、形質Ａを有するメンバーおよび形質Ｂを有するメンバーは、上記変異生成処理によって導入された変異を除いて、（これらはともに同じ遺伝的バックグラウンドに由来するので）同質遺伝子的であるということを理解するであろう。

【００７３】

しかしながら、挿入変異を用いる旧来の地図作成およびタグ化戦略とは対照的に、このような小さくかつランダムな変異は、検出するのが困難である。なぜなら、そのような変異は、特異的なＰＣＲ増幅のいずれの形態にも適していないからである。しかしながら、上記の戦略を使用する妥当な理由がいくつかある。当該変異生成は、実質的にいずれの注目する種にも適用することができ、誘導された変異体の範囲は、タグ化アプローチを用いる場合よりも広く、変異生成は、通常、より効率的であり、かつ第二部位突然変異がより容易に得られる。

【００７４】

実に様々な好適な変異生成方法を用いることができるが、変異誘発因子が、トランスポゾン挿入物またはＴ－ＤＮＡ挿入物からなる群から選択される生物学的変異誘発因子ではないことが好ましい。好ましくは、使用される変異生成方法は、遺伝子材料に点変異を導入する。

【００７５】

別の実施形態では、当該生物が植物であり、好ましくはトマト、コショウ、ナス、レタス、ニンジン、タマネギ、ニラ、チコリー、ダイコン、パセリ、ホウレンソウ、メロン、およびキュウリからなる群から選択される作物植物である、本発明に係る方法が提供される。

【００７６】

本発明に供することができる生物は、細菌、原核生物および真核生物を含めたいずれの生物であることができる。しかしながら、好ましくはこの生物は、真核生物、特に植物、より特定すれば作物植物である。好ましくは、この植物は、トマト、コショウ、ナス、レタス、ニンジンを含めた重要な作物植物からなる群に属する植物であるが、当該方法は、実質的にすべての他の植物に適用することができる。

【００７７】

特に、本発明は、当該技術分野で利用可能な地図作成およびタグ化技法が不可能であるかまたは煩わしいと判明している作物植物由来の、遺伝子のような核酸の特定およびクローニングを可能にする。

【００７８】

さらなる実施形態では、当該生物が植物であり、かつ工程ａ）に先立って、劣性の形質Ａをコードする対立遺伝子および優性の形質Ｂをコードする第２の対立遺伝子についてヘテロ接合であるＦ１ハイブリッド集団が作成され、かつこのＦ１ハイブリッド集団が変異生成に供され、かつ変異生成の後にこのＦ１集団から、形質Ａを有するメンバーが選択される、本発明に係る方法が提供される。

【００７９】

好ましくは、この対立遺伝子系列は、例えば、１つの遺伝子座で新しい対立遺伝子を選択するための標準的な遺伝子アプローチによって構築することができる。このアプローチは、１つの参照対立遺伝子（これは、この遺伝子座のこれまでに認識された劣性の表現型、すなわち形質Ａを産生する）およびこの遺伝子座の形質Ｂを産生する１つの優性の対立遺伝子についてヘテロ接合であるＦ１集団を作成することからなる。Ｆ１の変異生成により、対立遺伝子Ｂにおいて有害な変異が起こるとき形質Ａを産生する新しい対立遺伝子が

明らかになるであろう。形質Ａを有するこのような変異体Ｆ１植物はある頻度で現れるであろうが、一方で、Ｆ１集団の大部分は形質Ｂを示すであろう。１つの例では、（形質Ｂをコードする）遺伝子座の優性の対立遺伝子を保有する赤色の花をつける栽培品種の種子（または花粉、またはいずれかの他の組織）を変異化し、この材料を、（形質Ａをコードする）当該遺伝子座の劣性の参照対立遺伝子を保有する突然変異を誘発されていない白色の花をつける栽培品種への交雑（ｓｅｘｕａｌ　ｃｒｏｓｓ）のために使用することができる。得られたＦ１子孫では、形質Ａを発現するように優性のＢ対立遺伝子において新しく誘導された、独立の変異を保有するいくつかの白色の個体を除いて、大多数の個体は赤色であろう。第２の例では、Ｆ１材料（例えば種子、またはいずれかの好適な組織）を直接変異化させて、突然変異を誘発された材料から成長する得られた個々の植物において、形質Ａの発現について採点してもよい。例１および例２の優位点は、当該遺伝子座においてＡおよびＢ対立遺伝子についてヘテロ接合のＦ１材料の使用によって、新しく生じる形質Ａ対立遺伝子が、相対的な確からしさをもって、クローニングされる対象の遺伝子座に割り当てられるようになるということである。Ｆ１アプローチを用いなければ、例えば初代の突然変異を誘発された材料の自家受精によって誘導されるＦ２集団において、白色の花をつける変異体植物が、クローニングされる対象の遺伝子とは異なる遺伝子における変異の結果として生じるということが理論的には可能である。従って、これらの新しく生じる変異は非対立性である可能性があり、本発明に関しては混乱させるものとなるであろう。

【００８０】

本発明に係る方法によって特定されることになる核酸は、その後単離し、（遺伝子を）クローニングし、または宿主細胞に導入し、または例えば、植物育種プログラムで使用することができる。

【００８１】

要約すれば、本発明は、生物において発現され、かつその生物の特定の表現型（特徴の形質）に関連する核酸配列の特定、および任意の単離のための新しい戦略に関する。本発明の方法を用いると、当該技術分野で公知の方法とは対照的に、ゲノムに関して情報が入手できないかまたは限られた情報しか入手できない、例えば（作物）植物のような生物において、遺伝子を効率的に特定し、単離し、またはクローニングすることが可能になった。加えて、本発明では、例えばＴ－ＤＮＡまたはトランスポゾン系を使用することによるよりもはるかに高い変異頻度が生成され、これにより当該方法で必要とされる生物、例えば植物の量を減少させるので、点変異を誘導する化学物質を、変異化された集団を作成するための汎用性がある、広く適用できる方法として使用することを利用することが可能である。さらに、当該方法は注目する遺伝子における変化を直接検出するので、当該方法はゲノムにおけるマーカーの使用または存在に頼らない。

【実施例】

【００８２】

ペチュニアにおける花の色素沈着についての遺伝子のＤｅ　ｎｏｖｏクローニング

１つの花の色の遺伝子における変異体の生成

本発明に係る上記の方法を実証するために、本発明者らは、ペチュニアのこれまでに知られている花の色の遺伝子（ＲＴ）のｄｅ　ｎｏｖｏ特定を意図した。赤紫色の花をつけるＦ１ハイブリッドのペチュニアを、近交系Ｗ５（ｒｔ：：ｄＴｐｈ３；ＳｔｕｕｒｍａｎおよびＫｕｈｌｅｍｅｉｅｒ、Ｐｌａｎｔ　Ｊ．、２００５年、第４１巻、第６号、９４５－５５頁）とＭ１（ＲＴ，Ｓｎｏｗｄｅｎ　ＫＣおよびＮａｐｏｌｉ　ＣＡ、「Ｐｓｌ：ａ　ｎｏｖｅｌ　Ｓｐｍ－ｌｉｋｅ　ｔｒａｎｓｐｏｓａｂｌｅ　ｅｌｅｍｅｎｔ　ｆｒｏｍ　Ｐｅｔｕｎｉａ　ｈｙｂｒｉｄａ」、Ｐｌａｎｔ　Ｊ．、１９９８年、第１４巻、４３－５４頁）との間の交雑から生成し、これにより、ＵＤＰラムノース：アントシアニン－３－グルコシドラムノシルトランスフェラーゼをコードする遺伝子ＲＴに劣性の対立遺伝子を保有する赤紫色の花をつけるヘテロ接合の遺伝子型を生じた。この実施例で使用した遺伝的バックグラウンドでは、ＲＴのヌル変異体は、シアニジン－３－グルコ

シドの蓄積の結果として赤色の花を生成する（Ｋｒｏｏｎら、Ｐｌａｎｔ　Ｊ．、１９９４年、第５巻、第１号、６９－８０頁；Ｂｒｕｇｌｉｅｒａら、Ｐｌａｎｔ　Ｊ．、１９９４年、第５巻、第１号、８１－９２頁）。メタンスルホン酸エチル（ＥＭＳ）を用いたＦ１ハイブリッド種子の変異原性処理の後に、得られた植物のいくつかは、ＲＴの野生型対立遺伝子において新規な変異を保有し、ヘテロ接合性の喪失および赤色の花冠の色の発現を生じるであろう。

【００８３】

３６００のＦ１植物の集団を、ＥＭＳ処理した種子から生育させ、すべての個々の植物に対して花の色を視覚的に記録した。７つの赤色の花をつける変異体を観察した。赤色は劣性であるので、初代の変異化された集団におけるこの色の発現は、赤色の花をつける変異体が、ＲＴ遺伝子の野生型コピーにＥＭＳによって誘導された点変異を保有したことを示す。これらの新しく単離した変異は、必然的にヘテロ接合である。

【００８４】

新しく特定された変異体対立遺伝子から赤色の根底にある遺伝子をｄｅ　ｎｏｖｏ特定するために、アントシアニン色素沈着が目に見えるようになる発生上の時点における花冠の舷部の完全かつノーマライズしたトランスクリプトームを、３つの独立の変異体においてシーケンシングした。これらの３つの変異体の各々の全ＲＮＡを、ＳＭＡＲＴ　ｃＤＮＡ合成キット（Ｃｌｏｎｔｅｃｈ　Ｃｏｍｐａｎｙ、米国、カリフォルニア州　９４０４３、Ｍｏｕｎｔａｉｎ　Ｖｉｅｗ、Ｔｅｒｒａ　Ｂｅｌｌａ　Ａｖｅｎｕｅ　１２９０）を使用してｄｓ－ｃＤＮＡへと変換した。二本鎖の（ｄｓ－）ｃＤＮＡを、Ｔｒｉｍｍｅｒキット（Ｅｖｒｏｇｅｎ　Ｃｏｍｐａｎｙ、ロシア、モスクワ、１１７９９７、１６／１０、Ｍｉｋｌｕｋｈｏ－Ｍａｋｌａｙａ　ｓｔｒ）を使用してノーマライズした。各個々の変異体植物に由来するノーマライズしたｃＤＮＡを、ＧＳ－ＦＬＸ　Ｔｉｔａｎｉｕｍシーケンサー（４５４　Ｌｉｆｅ　Ｓｃｉｅｎｃｅｓ）でのシーケンシングのために、別々にプロセシングした。プロセシングは、ｃＤＮＡを約７００ｎｔへと断片化すること、および１０ｂｐの多重識別子（ＭＩＤ）を用いたタグ化を含んでいた。その後、すべての変異体に由来するすべてのプロセシングしたｃＤＮＡ試料をプールし、２回のランのＧＳ－ＦＬＸ　Ｔｉｔａｎｉｕｍシーケンシングでシーケンシングした。平均３００ヌクレオチドの長さの合計１’９７７’４２２のリード（ｒｅａｄ）を得た。

【００８５】

すべてのプールした配列データをトリミングしてアダプター配列およびＭＩＤ配列を取り除き、その後、アセンブリプログラムＮｅｗｂｌｅｒ（Ｒｏｃｈｅ　４５４　Ｓｏｆｔｗａｒｅ　Ｒｅｌｅａｓｅ：２．３－ＰｒｅＲｅｌｅａｓｅ－９／１４／２００９）を使用し、デフォルト実行アセンブリ設定を使用して、重複のない組のコンティグへとアセンブリングした。結果の概略を表１Ａに提示する。花冠で発現されたユニジーンの完全補体（ｆｕｌｌ　ｃｏｍｐｌｅｍｅｎｔ）を表す合計３９３０６のユニークなコンティグ（平均サイズ　９５５ヌクレオチド、メジアンサイズ　１０２５ヌクレオチド）を得た。１塩基あたりのカバレッジは２４×であり、これは、平均のヌクレオチドが２４回シーケンシングされたことを意味する。この補体は、本発明に係る方法でこの後使用するための共通の参照トランスクリプトーム配列を表すので、この補体を、以降、「参照」と呼ぶ。同様に、コンティグを、以降、「ユニジーン（ｕｎｉｇｅｎｅ）」と呼ぶ。

【００８６】

本発明に係る方法の前提は、複数のユニジーンの中で、１つの特定のユニジーンが、トランスクリプトーム配列を生成するために使用する変異体のほとんどまたは各々において遺伝子的に非同義の一塩基多型（ＳＮＰ）を保有するであろうということである。この実施例で提示する具体例では、赤色の花の色を生成することに関与する遺伝子は、３回、このｃＤＮＡプールを構成した各個々の変異体植物において１回変異されるべきである。ランダムなＥＭＳ変異の頻度は、典型的には、２００キロベースあたり約１であるので、３つの異なるヌクレオチド変異がいずれかの単独のユニジーンで一貫して発生するであろうという可能性は、そのユニジーンが変異体植物を選択する際に基づいた表現型と因果関係

がないならば、ゼロに近い。この変異基準に合致せず、かつ中の塩基変化も非同義的である（これは、その遺伝子のタンパク質産物においてアミノ酸変化を引き起こすということを意味する）特定の遺伝子は、赤色の花の色を引き起こす遺伝子である可能性が非常に高い。この後の記載では、塩基変化を、一ヌクレオチド変異（ＳＮＭ）として示し、非同義の一ヌクレオチド変異をｎＳＮＭとして示す。同義の塩基変化、つまりアミノ酸変化を引き起こさない塩基変化は、それが、例えばイントロン－エクソンスプライス部位に存在する場合には、遺伝的に非中立でありうるときがあるということを注記することができる。しかしながら、対立遺伝子の少なくとも一対または好ましくは一群が比較のために使用されるので、そしてスプライシングエラーは大きく変化したｃＤＮＡを導くであろうから、本発明の方法は、そのような比較的稀なケースに対しても強力である、ということに留意されたい。

【００８７】

本発明者らは、花冠の４５４配列データセットを、すべての個々の４５４配列のリードを上記の参照の花冠のトランスクリプトームに対してアラインメントすることにより、すべての個々の４５４の配列のリードにおけるＳＮＭの発生について分析した。リードが参照と少なくとも９８％同一であるときだけ、４５４のリードのアラインメントをＳＮＭ分析のために使用した。アラインメントは、当業者にとっては周知である種々のバイオインフォーマティクス方法によって行うことができる。ＳＮＭは、本実施例では、以下のように定義する：

（１）４５４のリードの完全セットにおいて少なくとも３回発生する（３×カバレッジ）
（２）３つの個々の変異体植物の１つだけにおいて発生する
（３）個々の変異体植物由来のすべてのリードの１０％＜Ｎ＜９０％に発生する。

これらの３つの基準は、ＳＮＭがシーケンシングエラーではなく（１）、独立であり（２）、かつヘテロ接合の条件で発生する（３）ということを保証する。当業者は、これらの基準は、プールの中の変異体植物の数、およびＳＮＭ組み入れ（ｃａｌｌｉｎｇ）におけるタイプ１およびタイプ２のエラーのレベルに応じて、より厳格にするかまたは緩和することができるということを理解するであろう。加えて、この実施例ではＧ→ＡまたはＣ→Ｔの移行だけを可能なＳＮＭとして採用した。なぜなら、これらは、使用した突然変異誘発物質（ＥＭＳ）によって誘導される初代変異であるからである。また、当業者は、マッピングされたリードの９８％同一性という閾値をより厳格にするかまたは緩和することができること、およびそのような閾値は、見出されるＳＮＭの数に影響を及ぼす可能性があるということも理解するであろう。

【００８８】

次の工程として、本発明者らは、当該プールの中の３つの変異体の各々について１つで、かつ上で明記したＳＮＭ組み入れ基準（１、２、３）によって特定された少なくとも３つのＳＮＭが存在するユニジーンを検索した。シーケンシングカバレッジが、すべてのＳＮＭがＳＮＭ組み入れについての３つの基準を満たすのに十分であれば、赤色の花をつける表現型と因果関係があるいずれかのユニジーンは、理論上、このルールに適合する必要がある。２６のユニジーンが、各々異なる変異体植物に由来する少なくとも３つのＳＮＭをコード領域に含有するということが判明した。これらのうちで、２０個は、ちょうど３つのＳＮＭを含有していたが、６つは４～７つのＳＮＭを含有していた。

【００８９】

すべての上記のＳＮＭの形式上の遺伝子パターンに適合したこれら２６個のユニジーンを、その後、ｎＳＮＭの発生について調べた。まず、ＤＮＡトリプレットをアミノ酸へと翻訳するための標準的な真核生物の遺伝コードに従って、最長の予測されるオープンリーディングフレーム（ＯＲＦ）を見出すことにより、これら２６個のユニジーンのＤＮＡ配列をアミノ酸配列へと変換した。次いで、相同タンパク質を検出し、これにより予測されたＯＲＦが正しい生物学的に意味のある翻訳リーディングフレーム（ｔｒａｎｓｌａｔｉｏｎａｌ　ｒｅａｄｉｎｇ　ｆｒａｍｅ）の中にあることを確認するために、標準的なＢＬＡＳＴｐ手順を使用してこれらのＯＲＦをタンパク質配列のＵｎｉＰｒｏｔデータベー

スと比較した。次いで、これら２６個の参照ユニジーンの各々のタンパク質配列を、ＳＮＭを含有する変異体植物の、上記タンパク質配列の対照物タンパク質配列と比較した。ＳＮＭを含むコドンでアミノ酸が発生していた場合には、そのＳＮＭを、ｎＳＮＭとして標識した。当然、ｎＳＮＭだけが、コード領域内で表現型の変化を引き起こすことができる。それゆえ、各変異体植物が少なくとも１つのｎＳＮＭを含有しつつ少なくとも３つのｎＳＮＭを含有しないこれら２６個のユニジーンはいずれも、赤色の花をつける表現型を引き起こす可能性のある候補としては、捨てた。この操作の後、９つのユニジーンが候補の一覧に残った（表１Ｂ）。

【００９０】

この後、残りの９つのユニジーンのセットを、形式上の遺伝基準に関するランク付け手順にかけた。ちょうど３つの変異体植物をこの分析で使用したという事実によって導かれるとおり、完全なパターンはちょうど３であるので、３つを超えるｎＳＮＭの存在についてはペナルティーを与えた。ｎＳＮＭが多いほど、ランクは低い。この操作の後、３つのユニジーンがランク１を共有した。９つのユニジーンおよびそれらのランク付けの全一覧を、ＵｎｉＰｒｏｔデータベースへの最良のＢＬＡＳＴマッチを含めて、表１Ｃに提示する。赤色の花をつける表現型を引き起こすことが事前に知られており、かつＵＤＰラムノース：アントシアニン－３－グルコシド　ラムノシルトランスフェラーゼをコードするＲＴ遺伝子がランク１を有することが分かる。

【００９１】

これらのデータから、本発明に係る方法は、３９３０６のユニジーンのバックグラウンドの範囲内で小さい組の候補の中からの、変異体表現型を引き起こす単独の遺伝子の特定を可能にすると結論される。当業者は、本発明に係る方法はペチュニア以外のいずれの種においても実施することができるということを理解するであろう。当業者は、当該方法は、１プールあたりの異なる数の変異体個体、参照トランスクリプトームへのリードのマッピングの際の異なる同一性％、１ＳＮＭあたりのカバレッジの異なる厳格さ、または検索に適用される他の測定可能な設定を用いて実施することができるということも理解するであろう。候補をランク付けするためのさらなる客観的な基準を単独でまたは組み合わせて使用することができることも明らかであり、このような客観的な基準としては、

・生のシーケンシングデータから得られる、可能性のあるＳＮＭの塩基組み入れクオリティスコア、

・プールの組成および使用した変異体植物のヘテロ接合のまたはホモ接合の状態から得られる形式上の遺伝予測に従う、プールの中の個々の対立遺伝子にわたるＳＮＭの完全分布に対するマッチ度、

・野生型アミノ酸と比較した、ｎＳＮＭに起因するアミノ酸置換の化学的類似性スコア

が挙げられる。

【００９２】

物質および方法

１）ＧＳ－ＦＬＸ　ｔｉｔａｎｉｕｍシーケンシングのためのｃＤＮＡの調製

分析に好適なｃＤＮＡ試料を調製するために、以下の工程に従った。

Ｉ．製造業者の取扱説明書に従ってＱｉａｇｅｎ　ＲＮｅａｓｙ　Ｐｌａｎｔ　Ｍｉｎｉキット（Ｑｉａｇｅｎ　Ｃｏｍｐａｎｙ、カリフォルニア州　９１３５５、Ｖａｌｅｎｃｉａ、Ｔｕｒｎｂｅｒｒｙ　Ｌａｎｅ　２７２２０）を使用する、色素沈着の開始時の各変異体植物の花弁組織の全ＲＮＡ単離

ＩＩ．製造業者の取扱説明書に従った、ＳＭＡＲＴ　ｃＤＮＡ　ｓｙｎｔｈｅｓｉｓ　ｆｏｒ　Ｌｉｂｒａｒｙ　Ｃｏｎｓｔｒｕｃｔｉｏｎプロトコル（Ｃｌｏｎｔｅｃｈ　Ｃｏｍｐａｎｙ、米国、カリフォルニア州　９４０４３、Ｍｏｕｎｔａｉｎ　Ｖｉｅｗ、Ｔｅｒｒａ　Ｂｅｌｌａ　Ａｖｅｎｕｅ　１２９０）による、ＳＭＡＲＴ二本鎖（ｄｓ－）ｃＤＮＡ合成

ＩＩＩ．　製造業者の取扱説明書に従ってＴＲＩＭＭＥＲ　ｃＤＮＡ　Ｎｏｒｍａｌｉｚａｔｉｏｎキット（Ｅｖｒｏｇｅｎ　Ｃｏｍｐａｎｙ、ロシア、モスクワ、１１７９９

７、１６／１０、Ｍｉｋｌｕｋｈｏ－Ｍａｋｌａｙａ　ｓｔｒ）を使用するｃＤＮＡノーマライゼーション。

ＩＶ．以下の反応設定：

【表１】

2 x BAL31 反応バッファー	25.0	µl
1 U/µl BAL-31	0.25	µl
cDNA (1 µg)	x	µl
再蒸留水	y	µl
合計	50	µl

を使用する、ＢＡＬ３１　エキソヌクレアーゼ（Ｎｅｗ　Ｅｎｇｌａｎｄ　Ｂｉｏｌａｂｓ、米国、マサチューセッツ州　０１９３８－２７２３、Ｉｐｓｗｉｃｈ、Ｃｏｕｎｔｙ　Ｒｏａｄ　２４０）によるｄｓ－ｃＤＮＡからのポリＡ尾部の除去。

ＢＡＬ３１エキソヌクレアーゼによる反応を、アガロースゲル電気泳動および５０＋５０＝１００塩基対の平均分子サイズシフトを観察することから判断して、ｃＤＮＡの５’および３’末端から約５０ヌクレオチドが取り除かれるまで、製造業者の取扱説明書に従って進行させた。次いで、処理したｃＤＮＡ試料を、製造業者の取扱説明書（Ｑｉａｇｅｎ　Ｃｏｍｐａｎｙ、カリフォルニア州　９１３５５、Ｖａｌｅｎｃｉａ、Ｔｕｒｎｂｅｒｒｙ　Ｌａｎｅ　２７２２０）に従ってＱｉａｇｅｎ　ＰＣＲ精製キットを使用して精製した。

Ｖ．ＢＡＬ３１処理したｃＤＮＡ試料を１５の小分け試料へと分割し、各小分け試料を、以下の一覧からの１つの制限酵素（それらのＤＮＡ認識配列をカッコ内に示す）を用いた消化によって断片化した：ＡｃｃＩ（ＧＴＭＫＡＣ）、ＡｌｕＩ（ＡＧＣＴ）、ＡｐｏＩ（ＲＡＡＴＴＹ）、ＡｖａＩ（ＣＹＣＧＲＧ）、ＡｖａＩＩ（ＧＧＷＣＣ）、ＢｓａＡＩ（ＹＡＣＧＴＲ）、ＢｓｉＥＩ（ＣＧＲＹＣＧ）、ＢｓｔＹＩ（ＲＧＡＴＣＹ）、ＨａｅＩＩ（ＲＧＣＧＣＹ）、Ｈｐｙ９９Ｉ（ＣＧＷＣＧ）、ＭｓｅＩ（ＴＴＡＡ）、ＮｓｐＩ（ＲＣＡＴＧＹ）、ＲｓａＩ（ＧＴＡＣ）、ＳｆｃＩ（ＣＴＲＹＡＧ）、ＳｔｙＩ（ＣＧＷＷＧＧ）。これらの制限酵素は、それらが、コンティグ構築を可能にするためのｃＤＮＡの良好なカバレッジおよび断片間の良好なオーバーラップの可能性を確保しつつ、ＧＳ－ＦＬＸ　Ｔｉｔａｎｉｕｍシーケンシングに好適なサイズ範囲の断片を生成することができることから選んだ。これらの反応を完結するまで進行させ、その後、酵素製造業者の取扱説明書に従って不活化した。すべての酵素をＮｅｗ　Ｅｎｇｌａｎｄ　Ｂｉｏｌａｂｓ（米国、マサチューセッツ州　０１９３８－２７２３、Ｉｐｓｗｉｃｈ、Ｃｏｕｎｔｙ　Ｒｏａｄ　２４０）から得た。完全消化の後、１つの変異体植物あたりすべての１５種の制限反応物をプールし、製造業者の取扱説明書（Ｑｉａｇｅｎ　Ｃｏｍｐａｎｙ、カリフォルニア州　９１３５５、Ｖａｌｅｎｃｉａ、Ｔｕｒｎｂｅｒｒｙ　Ｌａｎｅ　２７２２０）に従ってＱｉａｇｅｎ　ＰＣＲ精製キットを用いて精製した。

ＶＩ　３～５μｇの、各変異体植物から個々に得た精製したｃＤＮＡを使用して、製造業者の取扱説明書に従ってＧＳ－ＦＬＸ　Ｔｉｔａｎｉｕｍ　Ｇｅｎｅｒａｌ　Ｌｉｂｒａｒｙ　Ｐｒｅｐａｒａｔｉｏｎ　Ｋｉｔ（Ｒｏｃｈｅ　Ｄｉａｇｎｏｓｔｉｃｓ　Ｃｏｒｐｏｒａｔｉｏｎ、Ｒｏｃｈｅ　Ａｐｐｌｉｅｄ　Ｓｃｉｅｎｃｅ、米国、インディアナ州　４６２５０－０４１４、インディアナポリス、Ｈａｇｕｅ　Ｒｏａｄ　９１１５、私書箱５０４１４）を使用して、シーケンシングのためのＧＳ－ＦＬＸ　Ｔｉｔａｎｉｕｍライブラリーを生成した。これらの個々のライブラリーを、異なる多重識別子（ＭＩＤ）を用いて調製した。次いで、本研究で使用した３つの変異体植物のＭＩＤで区別可能なライブラリーをプールし、２回のランのＧＳ－ＦＬＸ　Ｔｉｔａｎｉｕｍを使用してシーケンシングした。

【００９３】

バイオインフォーマティクス解析

生のＧＳ－ＦＬＸ　Ｔｉｔａｎｉｕｍリードの事前処理

ＧＳ－ＦＬＸ　Ｔｉｔａｎｉｕｍ　Ｇｅｎｅｒａｌ　Ｌｉｂｒａｒｙ　Ｐｒｅｐａｒａｔｉｏｎ　Ｋｉｔ（Ｒｏｃｈｅ　Ｄｉａｇｎｏｓｔｉｃｓ　Ｃｏｒｐｏｒａｔｉｏｎ、Ｒｏｃｈｅ　Ａｐｐｌｉｅｄ　Ｓｃｉｅｎｃｅ、米国、インディアナ州　４６２５０－０４１４、インディアナポリス、Ｈａｇｕｅ　Ｒｏａｄ　９１１５、私書箱５０４１４）で与えたＭＩＤおよびアダプター配列を取り除くために、リードを事前処理した。

【００９４】

参照トランスクリプトーム配列のアセンブリ

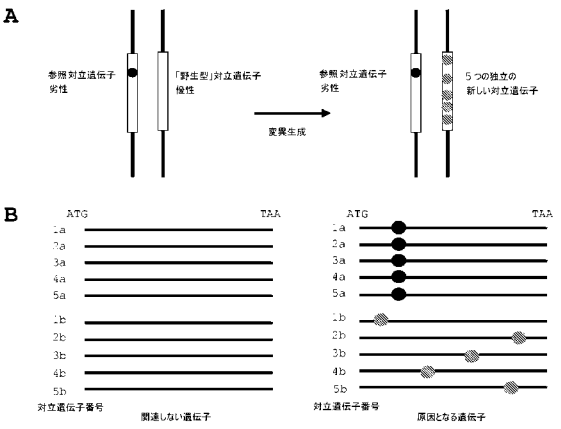
事前処理した配列をインプットとして用い、４５４のデフォルト実行アセンブリ設定を使用して、４５４　Ｌｉｆｅ　Ｓｃｉｅｎｃｅｓ（４５４　Ｌｉｆｅ　Ｓｃｉｅｎｃｅｓ、米国、コネティカット州　０６４０５、Ｂｒａｎｆｏｒｄ、Ｃｏｍｍｅｒｃｉａｌ　Ｓｔｒｅｅｔ　１５）のＮｅｗｂｌｅｒアセンブリソフトウェアを用いてアセンブリを行った。

【００９５】

解析

事前処理したリードを、９８％の最小同一性を使用して、アセンブリングした参照コンティグ上にマッピングした。ユニジーンの中の異なる位置に少なくとも３つのＳＮＭを有することにより、および異なる変異体植物個体との関連性により、候補のユニジーンを選択した。３×の最小カバレッジからＳＮＭを組み入れた。このＳＮＭの特性を、Ｇ→ＡまたはＣ→Ｔであるように強制した。１つのＳＮＭ位置あたり、唯一の、かつヘテロ接合の、変異体植物を許容した。同じヌクレオチドが個々の植物あたり９０％以上のケースで見られた場合に、ヌクレオチド位置をホモ接合と考えた。

【図１】



【図２】

表１

A

GS-FLX Titanium シーケンシング	参照トランスクリプトームの特徴
GS-FLX Titanium リードの数	1977422
GS-FLX Titanium 塩基の数	592944352
平均リード長 (nt)	299.86
コンティグの総数	39305
平均コンティグ長 (nt)	955
メジアンコンティグ長 (nt)	1025
総コンティグ構成度 (nt)	24841945
コンティグ中の1塩基あたりの平均カバレッジ	23.87

B

標準 SNM 基準 - 3 つのヘテロ接合の個体 , ≥3 異なる位置
参照コンティグに対する 98% リード同一性 , ≥3X 1対立遺伝子あたりのカバレッジ

ステップ	基準	ユニジーンの数	参照のも
ステップ d	標準	9	0.023
ステップ e	≥2 の個体において nSNMs	9	0.023
ステップ f	>3 の異なる nSNMs についてはペナルティー	3	0.008

C

ステップ g	特定された cDNA 相同性	受入番号	nSNM 数
ランク 1	アントシアニジン 3-O-グルコシルトランスフェラーゼ	>sp|Q43716	3
ランク 1	小胞結合膜タンパク質	>gi|18415696	3
ランク 1	相同性なし	-	3
ランク 2	RNA ポリメラーゼ関連	>gi|22331241	4
ランク 3	RNA 結合タンパク質	>gi|18651027(	4
ランク 4	未知のタンパク質	>gi|15242587	5
ランク 5	グルタチオン S-トランスフェラーゼ	>gi|15218517	5
ランク 6	リボソームタンパク質 S13	>gi|18411224	6
ランク 7	母性胚発育停止因子(maternal embryo arrest)S9	>gi|15235514	7

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No
PCT/NL2010/050860

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/68
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SARIN SUMEET ET AL: "Caenorhabditis elegans mutant allele identification by whole-genome sequencing", NATURE METHODS, NATURE PUBLISHING GROUP, GB LNKD- DOI:10.1038/NMETH.1249, vol. 5, no. 10, 1 October 2008 (2008-10-01), pages 865-867, XP002572952, ISSN: 1548-7091 [retrieved on 2008-08-01] * abstract figure 2	1-6
A	WO 2007/037678 A2 (KEYGENE NV [NL]; VAN EIJK MICHAEL JOSEPHUS THER [NL]; VAN TUNEN ADRIAN) 5 April 2007 (2007-04-05) example III claims 1-14	1-6

-/--

[X] Further documents are listed in the continuation of Box C. [X] 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 document 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	Date of mailing of the international search report
17 March 2011	29/03/2011

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Schmitt-Humbert, C

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/NL2010/050860

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BRUGLIERA F ET AL: "Isolation and characterization of a cDNA clone corresponding to the Rt locus of Petunia hybrida", PLANT JOURNAL, BLACKWELL SCIENTIFIC PUBLICATIONS, OXFORD, GB, vol. 5, no. 1, 1 January 1994 (1994-01-01) , pages 81-92, XP002572953, ISSN: 0960-7412 * abstract page 83, last paragraph	1-6
A	PETERS JANNY L ET AL: "Forward genetics and map-based cloning approaches", TRENDS IN PLANT SCIENCE, ELSEVIER SCIENCE, OXFORD, GB LNKD-DOI:10.1016/J.TPLANTS.2003.09.002, vol. 8, no. 10, 1 October 2003 (2003-10-01), pages 484-491, XP002535337, ISSN: 1360-1385 the whole document	1-6
X,P	WO 2010/056115 A1 (KEYGENE NV [NL]; STUURMAN JEROEN [NL]) 20 May 2010 (2010-05-20) the whole document	1-6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/NL2010/050860

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. [X] Claims Nos.: 1-6(partially)
because they relate to subject matter not required to be searched by this Authority, namely:

see FURTHER INFORMATION sheet PCT/ISA/210

2. [ ] 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. [ ] Claims Nos.:
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:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. [ ] 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. [ ] 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.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

[ ] 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.

[ ] No protest accompanied the payment of additional search fees.

International Application No. PCT/NL2010/050860

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box II.1

Claims Nos.: 1-6(partially)

Rule 39.1(iv) PCT - Method for treatment of the human or animal body by surgery Independent claim 1 relates to a method for identification of an expressed nucleic acid sequence that is associated with a character of an organism comprising the step of obtaining cDNA from members of said organism. Since the term "organism" includes human/animal beings, the scope of such claim embraces screening methods which are carried out on humans/animals and encompass a step of surgery practised on humans/animals. Independent claim 1, dependent claims 2-6 and the non-numbered claim relate thus to subject-matter considered by this Authority to be covered by the provisions of Rule 39.1(iv). The search of claims 1-6 has been carried out without taking the step of obtaining cDNA from members of an organism.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/NL2010/050860

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2007037678 A2	05-04-2007	AT 453728 T	15-01-2010
		AU 2006295556 A1	05-04-2007
		CA 2623539 A1	05-04-2007
		CN 101313078 A	26-11-2008
		DK 1929039 T3	08-03-2010
		EP 1929039 A2	11-06-2008
		ES 2338459 T3	07-05-2010
		JP 2009509527 T	12-03-2009
		US 2009170713 A1	02-07-2009
WO 2010056115 A1	20-05-2010	NONE	

Form PCT/ISA/210 (patent family annex) (April 2005)

フロントページの続き

(81)指定国 AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CZ,DE,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IS,JP,KE,KG,KM,KN,KP,KR,KZ,LA,LC,LK,LR,LS,LT,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PE,PG,PH,PL,PT,RO,RS,RU,SC,SD,SE,SG,SK,SL,SM,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,ZA,ZM,ZW

(72)発明者 ヤルサン フェイルス
オランダ国 エヌエル－６７００ アーエー ワーゲニンゲン ピー．オー．ボックス ２１６

Ｆターム(参考) 2B030 AA02 AB03 AD12 CA08
4B024 AA08 AA20 CA04 DA01 GA25 GA27 HA14
4B063 QA18 QQ04 QQ09 QQ43 QQ53 QR32 QR36 QR73 QR78 QS34
QX01