

(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2020-524490

(P2020-524490A)

(43) 公表日 令和2年8月20日(2020.8.20)

(51) Int.Cl.	F I	テーマコード (参考)
C 1 2 N 15/10 (2006.01)	C 1 2 N 15/10 2 0 0 Z	4 B 0 6 3
C 4 0 B 40/06 (2006.01)	C 4 0 B 40/06 Z N A	4 B 0 6 5
C 1 2 Q 1/68 (2018.01)	C 1 2 Q 1/68	
C 1 2 N 15/31 (2006.01)	C 1 2 N 15/31	
C 1 2 N 1/21 (2006.01)	C 1 2 N 1/21	

審査請求 未請求 予備審査請求 未請求 (全 301 頁) 最終頁に続く

(21) 出願番号	特願2019-567283 (P2019-567283)	(71) 出願人	517366688
(86) (22) 出願日	平成30年6月6日 (2018.6.6)		ザイマージェン インコーポレイテッド
(85) 翻訳文提出日	令和2年2月4日 (2020.2.4)		アメリカ合衆国 カリフォルニア 946
(86) 国際出願番号	PCT/US2018/036333		08, エメリービル, ホートン スト
(87) 国際公開番号	W02018/226880		リート 5980, スイート 105
(87) 国際公開日	平成30年12月13日 (2018.12.13)	(74) 代理人	100078282
(31) 優先権主張番号	62/515,870		弁理士 山本 秀策
(32) 優先日	平成29年6月6日 (2017.6.6)	(74) 代理人	100113413
(33) 優先権主張国・地域又は機関	米国 (US)		弁理士 森下 夏樹
		(74) 代理人	100181674
			弁理士 飯田 貴敏
		(74) 代理人	100181641
			弁理士 石川 大輔
		(74) 代理人	230113332
			弁護士 山本 健策

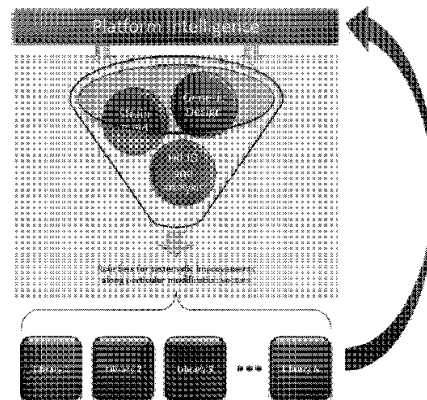
最終頁に続く

(54) 【発明の名称】 *Escherichia Coli* を改良するための HTP ゲノム操作プラットフォーム

(57) 【要約】

本開示は、コンピューターで駆動され、分子生物学、自動化、および高度な機械学習プロトコルを統合する、*Escherichia coli* を改良するための HTP ゲノム操作プラットフォームを提供する。この統合プラットフォームは、一式の HTP 分子ツールセットを利用して、HTP 遺伝子設計ライブラリーを創製する。これらのライブラリーは、とりわけ、科学的洞察および反復パターン認識から得られる。本開示は、伝統的な微生物株改良プログラムに関連した無数の問題を被らない *E. coli* のためのハイスループット (HTP) ゲノム操作プラットフォームを提供する。

FIGURE 3



【特許請求の範囲】

【請求項 1】

所望の表現型を獲得するために E. coli 微生物を進化させるゲノム操作のハイスループット (HTP) 方法であって、

a. 同じゲノム株バックグラウンドを有する最初の複数の E. coli 微生物のゲノムを攪乱することによって、ユニークな遺伝的バリエーションを有する個々の E. coli 株を含む最初の HTP 遺伝子設計 E. coli 株ライブラリーを創製するステップと；

b. 前記最初の HTP 遺伝子設計 E. coli 株ライブラリーの個々の株を、前記所望の表現型についてスクリーニングおよび選択するステップと；

c. それぞれが遺伝的バリエーションのユニークな組合せを含む後続の複数の E. coli 微生物を提供することによって、後続の HTP 遺伝子設計 E. coli 株ライブラリーを創製するステップであって、前記遺伝的バリエーションは、先行するステップにおいてスクリーニングされた少なくとも 2 つの個々の E. coli 株に存在する遺伝的バリエーションから選択される、ステップと；

d. 前記後続の HTP 遺伝子設計 E. coli 株ライブラリーの個々の E. coli 株を、前記所望の表現型についてスクリーニングおよび選択するステップと；

e. E. coli 微生物が前記所望の表現型を獲得するまで、線形または非線形様式でステップ c) ~ d) を 1 回または複数回繰り返すステップであって、各後続の反復により、先行する HTP 遺伝子設計 E. coli 株ライブラリーの少なくとも 2 つの個々の E. coli 株の中から選択される遺伝的バリエーションの組合せであるユニークな遺伝的バリエーションを内包する個々の E. coli 株を含む新しい HTP 遺伝子設計 E. coli 株ライブラリーが創製される、ステップとを含む、方法。

【請求項 2】

前記最初の HTP 遺伝子設計 E. coli 株ライブラリーが、プロモータースワップ微生物株ライブラリー、SNP スワップ微生物株ライブラリー、開始/停止コドン微生物株ライブラリー、最適化配列微生物株ライブラリー、ターミネータースワップ微生物株ライブラリー、タンパク質可溶性タグ微生物株ライブラリー、タンパク質分解タグ微生物株ライブラリー、およびこれらの任意の組合せからなる群から選択される少なくとも 1 つのライブラリーを含む、請求項 1 に記載のゲノム操作の HTP 方法。

【請求項 3】

前記最初の HTP 遺伝子設計 E. coli 株ライブラリーが、プロモータースワップ微生物株ライブラリーを含む、請求項 1 に記載のゲノム操作の HTP 方法。

【請求項 4】

前記最初の HTP 遺伝子設計 E. coli 株ライブラリーが、少なくとも 1 つのバイシストロン設計 (BCD) 調節配列を含有するプロモータースワップ微生物株ライブラリーを含む、請求項 1 または 2 に記載のゲノム操作の HTP 方法。

【請求項 5】

前記最初の HTP 遺伝子設計 E. coli 株ライブラリーが、SNP スワップ微生物株ライブラリーを含む、請求項 1 に記載のゲノム操作の HTP 方法。

【請求項 6】

前記最初の HTP 遺伝子設計 E. coli 株ライブラリーが、

a. キメラ生合成酵素をコードする少なくとも 1 つのポリヌクレオチドであって、前記キメラ生合成酵素は、E. coli において調節経路に参与する酵素を含み、前記酵素に、DNA 結合部位に結合することができる DNA 結合ドメインが翻訳的に融合する、少なくとも 1 つのポリヌクレオチド；および

b. 前記 DNA 結合部位を含む少なくとも 1 つの DNA 足場配列を含む微生物株ライブラリーを含む、請求項 1 または 2 に記載のゲノム操作の HTP 方法。

【請求項 7】

前記後続のH T P遺伝子設計E. coli株ライブラリーが、前記最初のH T P遺伝子設計E. coli株ライブラリーにおける前記遺伝的バリエーションに由来する完全コンビナトリアル株ライブラリーである、請求項1に記載のゲノム操作のH T P方法。

【請求項8】

前記後続のH T P遺伝子設計E. coli株ライブラリーが、前記最初のH T P遺伝子設計E. coli株ライブラリーにおける前記遺伝的バリエーションに由来する完全コンビナトリアル株ライブラリーのサブセットである、請求項1に記載のゲノム操作のH T P方法。

【請求項9】

前記後続のH T P遺伝子設計E. coli株ライブラリーが、先行するH T P遺伝子設計E. coli株ライブラリーにおける前記遺伝的バリエーションに由来する完全コンビナトリアル株ライブラリーである、請求項1に記載のゲノム操作のH T P方法。

10

【請求項10】

前記後続のH T P遺伝子設計E. coli株ライブラリーが、先行するH T P遺伝子設計E. coli株ライブラリーにおける前記遺伝的バリエーションに由来する完全コンビナトリアル株ライブラリーのサブセットである、請求項1に記載のゲノム操作のH T P方法。

【請求項11】

前記ゲノムを攪乱するステップが、ランダム変異誘発、標的配列挿入、標的配列欠失、標的配列の置換、およびこれらの任意の組合せからなる群から選択される少なくとも1つの方法を利用することを含む、請求項1に記載のゲノム操作のH T P方法。

20

【請求項12】

前記最初の複数のE. coli微生物が、工業生産E. coli株に由来するユニークな遺伝的バリエーションを含む、請求項1に記載のゲノム操作のH T P方法。

【請求項13】

前記最初の複数のE. coli微生物が、S₁Gen₁で表される工業生産株微生物およびS_nGen_nで表されるこれらに由来する任意の数の後続の微生物世代を含む、請求項1に記載のゲノム操作のH T P方法。

【請求項14】

SNPスワップE. coli株ライブラリーを生成するための方法であって、

30

a. 参照E. coli株および第2のE. coli株を提供するステップであって、前記第2のE. coli株は、前記参照E. coli株内に存在しない一塩基多型、DNA挿入、およびDNA欠失から選択される複数の同定された遺伝的バリエーションを含む、ステップと；

b. 前記参照E. coli株または前記第2のE. coli株のいずれかのゲノムを攪乱することによって、複数の個々のE. coli株を含む最初のSNPスワップE. coli株ライブラリーを創製するステップであって、前記複数の個々の株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記参照E. coli株と前記第2のE. coli株との間の前記複数の同定された遺伝的バリエーションから選択される単一の遺伝的バリエーションに対応する、ステップと

40

を含む、方法。

【請求項15】

前記参照E. coli株の前記ゲノムが攪乱されて、前記第2のE. coli株内に見出される前記同定された一塩基多型、DNA挿入、またはDNA欠失の1つまたは複数が付加される、請求項14に記載のSNPスワップE. coli株ライブラリーを生成するための方法。

【請求項16】

前記第2のE. coli株の前記ゲノムが攪乱されて、前記参照E. coli株内に見出されない前記同定された一塩基多型、DNA挿入、またはDNA欠失の1つまたは複数

50

が除去される、請求項 14 に記載の SNP スワップ E. coli 株ライブラリーを生成するための方法。

【請求項 17】

ユニークな遺伝的バリエーションを有する結果として生じる前記複数の個々の E. coli 株が、前記参照 E. coli 株と前記第 2 の E. coli 株との間のすべての前記同定された遺伝的バリエーションの完全コンビナトリアルライブラリーと一緒に含む、請求項 14 から 16 までのいずれか一項に記載の SNP スワップ E. coli 株ライブラリーを生成するための方法。

【請求項 18】

ユニークな遺伝的バリエーションを有する結果として生じる前記複数の個々の E. coli 株が、前記参照 E. coli 株と前記第 2 の E. coli 株との間のすべての前記同定された遺伝的バリエーションの完全コンビナトリアルライブラリーのサブセットと一緒に含む、請求項 14 から 16 までのいずれか一項に記載の SNP スワップ E. coli 株ライブラリーを生成するための方法。

【請求項 19】

産生 E. coli 株の表現型性能を再建および改善するための方法であって、

a. 親系統 E. coli 株およびそれに由来する産生 E. coli 株を提供するステップであって、前記産生 E. coli 株は、前記親系統株内に存在しない塩基多型、DNA 挿入、および DNA 欠失から選択される複数の同定された遺伝的バリエーションを含む、ステップと；

b. 前記親系統 E. coli 株または前記産生 E. coli 株のいずれかのゲノムを攪乱して、最初の E. coli 株のライブラリーを創製するステップであって、前記最初のライブラリーにおける各株は、前記親系統 E. coli 株と前記産生 E. coli 株との間の前記複数の同定された遺伝的バリエーションからのユニークな遺伝的バリエーションを含む、ステップと；

c. 参照 E. coli 株に対する表現型性能改善について前記最初のライブラリーの個々の株をスクリーニングおよび選択することによって、表現型性能改善をもたらすユニークな遺伝的バリエーションを同定するステップと；

d. 前記先行するステップにおいてスクリーニングされた少なくとも 2 つの個々の E. coli 株に存在する遺伝的バリエーションからのユニークな遺伝的バリエーションの組合せをそれぞれが含む後続の複数の E. coli 微生物を提供することによって、E. coli 株の後続のライブラリーを創製するステップと；

e. 前記参照 E. coli 株に対する表現型性能改善について前記後続のライブラリーの個々の株をスクリーニングおよび選択することによって、追加の表現型性能改善をもたらす遺伝的バリエーションのユニークな組合せを同定するステップと；

f. E. coli 株が、前記産生 E. coli 株の前記表現型性能と比較して、所望のレベルの表現型性能の改善を呈するまで、線形または非線形様式でステップ d) ~ e) を 1 回または複数回繰り返すステップであって、各後続の反復により、微生物株の新しいライブラリーが創製され、前記新しいライブラリーにおける各株は、先行するライブラリーの少なくとも 2 つの個々の E. coli 株の中から選択される遺伝的バリエーションの組合せである遺伝的バリエーションを含む、ステップとを含む、方法。

【請求項 20】

前記最初の E. coli 株のライブラリーが、前記親系統 E. coli 株と前記産生 E. coli 株との間の前記同定された遺伝的バリエーションのすべてを含む完全コンビナトリアルライブラリーである、請求項 19 に記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

【請求項 21】

前記最初の E. coli 株のライブラリーが、前記親系統 E. coli 株と前記産生 E. coli 株との間の前記同定された遺伝的バリエーションのサブセットを含む完全コン

ビナトリアルライブラリーのサブセットである、請求項 19 に記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

【請求項 22】

前記後続の E. coli 株のライブラリーが、前記最初のライブラリーの完全コンビナトリアルライブラリーである、請求項 19 に記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

【請求項 23】

前記後続の E. coli 株のライブラリーが、前記最初のライブラリーの完全コンビナトリアルライブラリーのサブセットである、請求項 19 に記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

10

【請求項 24】

前記後続の E. coli 株のライブラリーが、先行するライブラリーの完全コンビナトリアルライブラリーである、請求項 19 に記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

【請求項 25】

前記後続の E. coli 株のライブラリーが、先行するライブラリーの完全コンビナトリアルライブラリーのサブセットである、請求項 19 に記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

【請求項 26】

前記親系統 E. coli 株の前記ゲノムが攪乱されて、前記産生 E. coli 株内に見出される前記同定された一塩基多型、DNA 挿入、または DNA 欠失の 1 つまたは複数が付加される、請求項 19 に記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

20

【請求項 27】

前記産生 E. coli 株の前記ゲノムが攪乱されて、前記親系統 E. coli 株内に見出されない前記同定された一塩基多型、DNA 挿入、または DNA 欠失の 1 つまたは複数除去される、請求項 19 に記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

【請求項 28】

前記ゲノムを攪乱するステップが、ランダム変異誘発、標的配列挿入、標的配列欠失、標的配列の置換、およびこれらの組合せからなる群から選択される少なくとも 1 つの方法を利用することを含む、請求項 19 から 25 までのいずれか一項に記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

30

【請求項 29】

後続のライブラリーの E. coli 株の前記表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 10 % の増加を呈するまで、ステップ d) ~ e) を繰り返す、請求項 19 に記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

【請求項 30】

後続のライブラリーの E. coli 株の前記表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 1 倍の増加を呈するまで、ステップ d) ~ e) を繰り返す、請求項 19 に記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

40

【請求項 31】

ステップ f) の表現型性能の改善が、目的の産物の体積生産性、目的の産物の比生産性、目的の産物の収率、目的の産物の力価、およびこれらの組合せからなる群から選択される、請求項 19 に記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

【請求項 32】

ステップ f) の表現型性能の改善が、目的の産物の増大したまたはより効率的な産生で

50

あり、前記目的の産物は、低分子、酵素、ペプチド、アミノ酸、有機酸、合成化合物、燃料、アルコール、一次細胞外代謝産物、二次細胞外代謝産物、細胞内コンポーネント分子、およびこれらの組合せからなる群から選択される、請求項 19 に記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

【請求項 33】

前記同定された遺伝的バリエーションが、プロモータースワップライブラリー由来の人工プロモータースワップ遺伝的バリエーションをさらに含む、請求項 19 に記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

【請求項 34】

前記最初の E. coli 株のライブラリー、または

10

後続の E. coli 株のライブラリー

のいずれかの少なくとも 1 つの微生物株のゲノムを、内在性 E. coli 標的遺伝子に作動可能に連結したプロモーターラダー由来の 1 つまたは複数のプロモーターを含むように操作するステップ

をさらに含む、請求項 19 に記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

【請求項 35】

プロモータースワップ E. coli 株ライブラリーを生成するための方法であって、

a. 基準 E. coli 株に内在する複数の標的遺伝子、およびプロモーターラダーを提供するステップであって、前記プロモーターラダーは、前記基準 E. coli 株において異なる発現プロファイルを呈する複数のプロモーターを含む、ステップと；

20

b. 前記基準 E. coli 株のゲノムを操作することによって、複数の個々の E. coli 株を含む最初のプロモータースワップ E. coli 株ライブラリーを創製するステップであって、前記複数の個々の E. coli 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記基準 E. coli 株に内在する前記標的遺伝子の 1 つに作動可能に連結した前記プロモーターラダー由来の前記プロモーターの 1 つまたは複数を含む、ステップと

を含む、方法。

【請求項 36】

前記複数のプロモーターの少なくとも 1 つが、バイシストロン設計 (BCD) 調節配列を含む、請求項 35 に記載のプロモータースワップ E. coli 株ライブラリーを生成するための方法。

30

【請求項 37】

産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法であって、

a. 基準 E. coli 株に内在する複数の標的遺伝子、およびプロモーターラダーを提供するステップであって、前記プロモーターラダーは、前記基準 E. coli 株において異なる発現プロファイルを呈する複数のプロモーターを含む、ステップと；

b. 前記基準 E. coli 株のゲノムを操作することによって、複数の個々の E. coli 株を含む最初のプロモータースワップ E. coli 株ライブラリーを創製するステップであって、前記複数の個々の E. coli 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記基準 E. coli 株に内在する前記標的遺伝子の 1 つに作動可能に連結した前記プロモーターラダー由来の前記プロモーターの 1 つまたは複数を含む、ステップと；

40

c. 参照 E. coli 株に対する表現型性能改善について前記最初のプロモータースワップ E. coli 株ライブラリーの個々の E. coli 株をスクリーニングおよび選択することによって、表現型性能改善をもたらすユニークな遺伝的バリエーションを同定するステップと；

d. 前記先行するステップにおいてスクリーニングされた少なくとも 2 つの個々の E. coli 株に存在する遺伝的バリエーションからのユニークな遺伝的バリエーションの組

50

合せをそれぞれが含む後続の複数の E. coli 微生物を提供することによって、後続のプロモータースワップ E. coli 株ライブラリーを創製するステップと；

e. 前記参照 E. coli 株に対する表現型性能改善について前記後続のプロモータースワップ E. coli 株ライブラリーの個々の E. coli 株をスクリーニングおよび選択することによって、追加の表現型性能改善をもたらす遺伝的バリエーションのユニークな組合せを同定するステップと；

f. E. coli 株が、前記産生 E. coli 株の前記表現型性能と比較して、所望のレベルの表現型性能の改善を呈するまで、線形または非線形様式でステップ d) ~ e) を 1 回または複数回繰り返すステップであって、各後続の反復により、微生物株の新しいプロモータースワップ E. coli 株ライブラリーが創製され、前記新しいライブラリーにおける各株は、先行するライブラリーの少なくとも 2 つの個々の E. coli 株の中から選択される遺伝的バリエーションの組合せである遺伝的バリエーションを含む、ステップと

を含む、方法。

【請求項 38】

前記後続のプロモータースワップ E. coli 株ライブラリーが、前記最初のプロモータースワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである、請求項 37 に記載の産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法。

【請求項 39】

前記後続のプロモータースワップ E. coli 株ライブラリーが、前記最初のプロモータースワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである、請求項 37 に記載の産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法。

【請求項 40】

前記後続のプロモータースワップ E. coli 株ライブラリーが、先行するプロモータースワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである、請求項 37 に記載の産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法。

【請求項 41】

前記後続のプロモータースワップ E. coli 株ライブラリーが、先行するプロモータースワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである、請求項 37 に記載の産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法。

【請求項 42】

後続のプロモータースワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 10 % の増加を呈するまで、ステップ d) ~ e) を繰り返す、請求項 37 から 41 までのいずれか一項に記載の産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法。

【請求項 43】

後続のプロモータースワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 1 倍の増加を呈するまで、ステップ d) ~ e) を繰り返す、請求項 37 から 41 までのいずれか一項に記載の産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法。

【請求項 44】

ステップ f) の表現型性能の改善が、目的の産物の体積生産性、目的の産物の比生産性、目的の産物の収率、目的の産物の力価、およびこれらの組合せからなる群から選択される、請求項 37 に記載の産生 E. coli 株の表現型性能を改善するためのプロモーター

スワップ方法。

【請求項 45】

ステップ f) の表現型性能の改善が、目的の産物の増大したまたはより効率的な産生であり、前記目的の産物は、低分子、酵素、ペプチド、アミノ酸、有機酸、合成化合物、燃料、アルコール、一次細胞外代謝産物、二次細胞外代謝産物、細胞内コンポーネント分子、およびこれらの組合せからなる群から選択される、請求項 37 に記載の産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法。

【請求項 46】

ターミネータースワップ E. coli 株ライブラリーを生成するための方法であって、

a. 基準 E. coli 株に内在する複数の標的遺伝子、およびターミネーターラダーを提供するステップであって、前記ターミネーターラダーは、前記基準 E. coli 株において異なる発現プロファイルを呈する複数のターミネーターを含む、ステップと；

b. 前記基準 E. coli 株のゲノムを操作することによって、複数の個々の E. coli 株を含む最初のターミネータースワップ E. coli 株ライブラリーを創製するステップであって、前記複数の個々の E. coli 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記基準 E. coli 株に内在する前記標的遺伝子の 1 つに作動可能に連結した前記ターミネーターラダー由来の前記ターミネーターの 1 つまたは複数を含む、ステップとを含む、方法。

【請求項 47】

産生 E. coli 株の表現型性能を改善するためのターミネータースワップ方法であって、

a. 基準 E. coli 株に内在する複数の標的遺伝子、およびターミネーターラダーを提供するステップであって、前記ターミネーターラダーは、前記基準 E. coli 株において異なる発現プロファイルを呈する複数のターミネーターを含む、ステップと；

b. 前記基準 E. coli 株のゲノムを操作することによって、複数の個々の E. coli 株を含む最初のターミネータースワップ E. coli 株ライブラリーを創製するステップであって、前記複数の個々の E. coli 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記基準 E. coli 株に内在する前記標的遺伝子の 1 つに作動可能に連結した前記ターミネーターラダー由来の前記ターミネーターの 1 つまたは複数を含む、ステップと；

c. 参照 E. coli 株に対する表現型性能改善について前記最初のターミネータースワップ E. coli 株ライブラリーの個々の E. coli 株をスクリーニングおよび選択することによって、表現型性能改善をもたらすユニークな遺伝的バリエーションを同定するステップと；

d. 前記先行するステップにおいてスクリーニングされた少なくとも 2 つの個々の E. coli 株に存在する遺伝的バリエーションからのユニークな遺伝的バリエーションの組合せをそれぞれが含む後続の複数の E. coli 微生物を提供することによって、後続のターミネータースワップ E. coli 株ライブラリーを創製するステップと；

e. 前記参照 E. coli 株に対する表現型性能改善について前記後続のターミネータースワップ E. coli 株ライブラリーの個々の E. coli 株をスクリーニングおよび選択することによって、追加の表現型性能改善をもたらす遺伝的バリエーションのユニークな組合せを同定するステップと；

f. E. coli 株が、前記産生 E. coli 株の前記表現型性能と比較して、所望のレベルの表現型性能の改善を呈するまで、線形または非線形様式でステップ d) ~ e) を 1 回または複数回繰り返すステップであって、各後続の反復により、微生物株の新しいターミネータースワップ E. coli 株ライブラリーが創製され、前記新しいライブラリーにおける各株は、先行するライブラリーの少なくとも 2 つの個々の E. coli 株の中から選択される遺伝的バリエーションの組合せである遺伝的バリエーションを含む、ステップと

10

20

30

40

50

を含む、方法。

【請求項 48】

前記後続のターミネーターズワップ E. coli 株ライブラリーが、前記最初のターミネーターズワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである、請求項 47 に記載の産生 E. coli 株の表現型性能を改善するためのターミネーターズワップ方法。

【請求項 49】

前記後続のターミネーターズワップ E. coli 株ライブラリーが、前記最初のターミネーターズワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである、請求項 47 に記載の産生 E. coli 株の表現型性能を改善するためのターミネーターズワップ方法。

10

【請求項 50】

前記後続のターミネーターズワップ E. coli 株ライブラリーが、先行するターミネーターズワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである、請求項 47 に記載の産生 E. coli 株の表現型性能を改善するためのターミネーターズワップ方法。

【請求項 51】

前記後続のターミネーターズワップ E. coli 株ライブラリーが、先行するターミネーターズワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである、請求項 47 に記載の産生 E. coli 株の表現型性能を改善するためのターミネーターズワップ方法。

20

【請求項 52】

後続のターミネーターズワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 10 % の増加を呈するまで、ステップ d) ~ e) を繰り返す、請求項 47 から 51 までのいずれか一項に記載の産生 E. coli 株の表現型性能を改善するためのターミネーターズワップ方法。

【請求項 53】

後続のターミネーターズワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 1 倍の増加を呈するまで、ステップ d) ~ e) を繰り返す、請求項 47 から 51 までのいずれか一項に記載の産生 E. coli 株の表現型性能を改善するためのターミネーターズワップ方法。

30

【請求項 54】

ステップ f) の表現型性能の改善が、目的の産物の体積生産性、目的の産物の比生産性、目的の産物の収率、目的の産物の力価、およびこれらの組合せからなる群から選択される、請求項 47 に記載の産生 E. coli 株の表現型性能を改善するためのターミネーターズワップ方法。

【請求項 55】

ステップ f) の表現型性能の改善が、目的の産物の増大したまたはより効率的な産生であり、前記目的の産物は、低分子、酵素、ペプチド、アミノ酸、有機酸、合成化合物、燃料、アルコール、一次細胞外代謝産物、二次細胞外代謝産物、細胞内コンポーネント分子、およびこれらの組合せからなる群から選択される、請求項 47 に記載の産生 E. coli 株の表現型性能を改善するためのターミネーターズワップ方法。

40

【請求項 56】

E. coli 宿主細胞における生合成経路由来の生合成酵素を共局在化させるためのシステムであって、

a. 酵素反応に関与する 2 種またはそれよりも多くのキメラ酵素タンパク質であって、各キメラ酵素タンパク質は、DNA 結合ドメイン部分とカップリングした酵素部分を含む、2 種またはそれよりも多くのキメラ酵素タンパク質；および

50

b.

i. 少なくとも1つの核酸スパーサーによって分離された2つまたはそれよりも多くの異なるDNA結合部位をそれぞれが含む1つまたは複数のサブユニットを含むDNA足場を含み、

前記キメラ酵素タンパク質は、そのそれぞれが、前記DNA足場内の少なくとも1つのDNA結合部位に結合する、該キメラ酵素タンパク質とカップリングしたDNA結合ドメイン部分によって前記DNA足場に補充される、システム。

【請求項57】

10

前記キメラ酵素タンパク質の前記DNA結合ドメイン部分が、ジンクフィンガーDNA結合ドメインを含み、前記DNA足場の前記DNA結合部位が、対応するジンクフィンガー結合配列を含む、請求項56に記載のシステム。

【請求項58】

前記2種またはそれよりも多くのキメラ酵素タンパク質のそれぞれの前記酵素部分が、そのそれぞれのDNA結合ドメイン部分と、ポリペプチドリinker配列を介してカップリングしている、請求項56に記載のシステム。

【請求項59】

前記2種またはそれよりも多くのキメラ酵素タンパク質のそれぞれの前記酵素部分が、そのそれぞれのDNA結合ドメイン部分と、そのアミノ末端またはそのカルボキシ末端を介してカップリングしている、請求項56に記載のシステム。

20

【請求項60】

前記2種またはそれよりも多くのキメラ酵素タンパク質が、アミノ酸生合成経路の酵素を含む、請求項56に記載のシステム。

【請求項61】

- a. 作動可能に連結したプロモーター；
- b. 第1のリボソーム結合部位(SD1)；
- c. 第1のシストロン配列(Cis1)；
- d. 第2のリボソーム結合部位(SD2)

を順に含むバイシストロン設計調節(BCD)配列であって、

30

標的遺伝子配列(Cis2)に作動可能に連結している、BCD配列。

【請求項62】

SD1およびSD2がそれぞれNNNGGANNNの配列を含む、請求項61に記載のBCD。

【請求項63】

SD1とSD2が異なる、請求項61に記載のBCD。

【請求項64】

Cis1が停止コドンを含み、Cis2が開始コドンを含み、前記Cis1停止コドンと前記Cis2開始コドンが少なくとも1ヌクレオチド重複している、請求項61に記載のBCD。

40

【請求項65】

SD2がCis1内に完全に組み込まれている、請求項61に記載のBCD。

【請求項66】

2種の標的遺伝子タンパク質を宿主生物において発現させるための方法であって、

a. 前記宿主生物に、第1の標的遺伝子タンパク質をコードする第1のポリヌクレオチドを導入するステップであって、前記第1のポリヌクレオチドは、第1の請求項61に記載のバイシストロン設計調節(BCD)配列に作動可能に連結している、ステップと；

b. 前記宿主生物に、第2の標的遺伝子タンパク質をコードする第2のポリヌクレオチドを導入するステップであって、前記第2のポリヌクレオチドは、第2の請求項61に記載のBCDに作動可能に連結している、ステップと

50

を含み、

前記第1および第2のBCDは、それらのそれぞれのCis1配列以外は同一であり、前記標的遺伝子タンパク質は、前記宿主生物においてそれぞれ第1および第2の発現レベルで発現される、方法。

【請求項67】

前記第1の発現レベルが前記第2の発現レベルの1.5倍以内である、請求項66に記載の方法。

【請求項68】

前記第1および第2のポリヌクレオチドが、前記宿主細胞において、前記第1および第2のポリヌクレオチドが同一のBCDによって発現される対照宿主細胞と比較して低レベルの相同組換えを受ける、請求項66に記載の方法。

【請求項69】

タンパク質可溶性タグスワップE. coli株ライブラリーを生成するための方法であって、

a. 基準E. coli株に内在する複数の標的遺伝子、および可溶性タグラーを提供するステップであって、前記可溶性タグラーは、前記基準E. coli株において異なる可溶性プロファイルを呈する複数の可溶性タグを含む、ステップと；

b. 前記基準E. coli株のゲノムを操作することによって、複数の個々のE. coli株を含む最初の可溶性タグスワップE. coli株ライブラリーを創製するステップであって、前記複数の個々のE. coli株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記基準E. coli株に内在する前記標的遺伝子の1つに作動可能に連結した前記可溶性タグラー由来の前記可溶性タグの1つまたは複数を含む、ステップと

を含む、方法。

【請求項70】

産生E. coli株の表現型性能を改善するためのタンパク質可溶性タグスワップ方法であって、

a. 基準E. coli株に内在する複数の標的遺伝子、および可溶性タグラーを提供するステップであって、前記可溶性タグラーは、前記基準E. coli株において異なる発現プロファイルを呈する複数の可溶性タグを含む、ステップと；

b. 前記基準E. coli株のゲノムを操作することによって、複数の個々のE. coli株を含む最初の可溶性タグスワップE. coli株ライブラリーを創製するステップであって、前記複数の個々のE. coli株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記基準E. coli株に内在する前記標的遺伝子の1つに作動可能に連結した前記可溶性タグラー由来の前記可溶性タグの1つまたは複数を含む、ステップと；

c. 参照E. coli株に対する表現型性能改善について前記最初の可溶性タグスワップE. coli株ライブラリーの個々のE. coli株をスクリーニングおよび選択することによって、表現型性能改善をもたらすユニークな遺伝的バリエーションを同定するステップと；

d. 前記先行するステップにおいてスクリーニングされた少なくとも2つの個々のE. coli株に存在する遺伝的バリエーションからのユニークな遺伝的バリエーションの組合せをそれぞれが含む後続の複数のE. coli微生物を提供することによって、後続の可溶性タグスワップE. coli株ライブラリーを創製するステップと；

e. 前記参照E. coli株に対する表現型性能改善について前記後続の可溶性タグスワップE. coli株ライブラリーの個々のE. coli株をスクリーニングおよび選択することによって、追加の表現型性能改善をもたらす遺伝的バリエーションのユニークな組合せを同定するステップと；

f. E. coli株が、前記産生E. coli株の前記表現型性能と比較して、所望のレベルの表現型性能の改善を呈するまで、線形または非線形様式でステップd)～e)を

10

20

30

40

50

1 回または複数回繰り返すステップであって、各後続の反復により、微生物株の新しい可溶性タグスワップ E. coli 株ライブラリーが創製され、前記新しいライブラリー内の各株は、先行するライブラリーの少なくとも 2 つの個々の E. coli 株の中から選択される遺伝的バリエーションの組合せである遺伝的バリエーションを含む、ステップとを含む、方法。

【請求項 7 1】

前記後続の可溶性タグスワップ E. coli 株ライブラリーが、前記最初の可溶性タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである、請求項 7 0 に記載の産生 E. coli 株の表現型性能を改善するための可溶性タグスワップ方法。

10

【請求項 7 2】

前記後続の可溶性タグスワップ E. coli 株ライブラリーが、前記最初の可溶性タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである、請求項 7 0 に記載の産生 E. coli 株の表現型性能を改善するための可溶性タグスワップ方法。

【請求項 7 3】

前記後続の可溶性タグスワップ E. coli 株ライブラリーが、先行する可溶性タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである、請求項 7 0 に記載の産生 E. coli 株の表現型性能を改善するための可溶性タグスワップ方法。

20

【請求項 7 4】

前記後続の可溶性タグスワップ E. coli 株ライブラリーが、先行する可溶性タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである、請求項 7 0 に記載の産生 E. coli 株の表現型性能を改善するための可溶性タグスワップ方法。

【請求項 7 5】

後続の可溶性タグスワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 10 % の増加を呈するまで、ステップ d) ~ e) を繰り返す、請求項 7 0 から 7 4 までのいずれか一項に記載の産生 E. coli 株の表現型性能を改善するための可溶性タグスワップ方法。

30

【請求項 7 6】

後続の可溶性タグスワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 1 倍の増加を呈するまで、ステップ d) ~ e) を繰り返す、請求項 7 0 から 7 4 までのいずれか一項に記載の産生 E. coli 株の表現型性能を改善するための可溶性タグスワップ方法。

【請求項 7 7】

ステップ f) の表現型性能の改善が、目的の産物の体積生産性、目的の産物の比生産性、目的の産物の収率、目的の産物の力価、およびこれらの組合せからなる群から選択される、請求項 7 0 から 7 4 までのいずれか一項に記載の産生 E. coli 株の表現型性能を改善するための可溶性タグスワップ方法。

40

【請求項 7 8】

ステップ f) の表現型性能の改善が、目的の産物の増大したまたはより効率的な産生であり、前記目的の産物は、低分子、酵素、ペプチド、アミノ酸、有機酸、合成化合物、燃料、アルコール、一次細胞外代謝産物、二次的細胞外代謝産物、細胞内コンポーネント分子、およびこれらの組合せからなる群から選択される、請求項 7 0 から 7 4 までのいずれか一項に記載の産生 E. coli 株の表現型性能を改善するための可溶性タグスワップ方法。

【請求項 7 9】

50

タンパク質分解タグスワップ E. coli 株ライブラリーを生成するための方法であって、

a. 基準 E. coli 株に内在する複数の標的遺伝子、および分解タグラダーを提供するステップであって、前記分解タグラダーは、前記基準 E. coli 株において異なる溶解性プロファイルを呈する複数の分解タグを含む、ステップと；

b. 前記基準 E. coli 株のゲノムを操作することによって、複数の個々の E. coli 株を含む最初の分解タグスワップ E. coli 株ライブラリーを創製するステップであって、前記複数の個々の E. coli 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記基準 E. coli 株に内在する前記標的遺伝子の 1 つに作動可能に連結した前記分解タグラダー由来の前記分解タグの 1 つまたは複数を含む、ステップとを含む、方法。

【請求項 80】

産生 E. coli 株の表現型性能を改善するためのタンパク質分解タグスワップ方法であって、

a. 基準 E. coli 株に内在する複数の標的遺伝子、および分解タグラダーを提供するステップであって、前記分解タグラダーは、前記基準 E. coli 株において異なる発現プロファイルを呈する複数の分解タグを含む、ステップと；

b. 前記基準 E. coli 株のゲノムを操作することによって、複数の個々の E. coli 株を含む最初の分解タグスワップ E. coli 株ライブラリーを創製するステップであって、前記複数の個々の E. coli 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記基準 E. coli 株に内在する前記標的遺伝子の 1 つに作動可能に連結した前記分解タグラダー由来の前記分解タグの 1 つまたは複数を含む、ステップと；

c. 参照 E. coli 株に対する表現型性能改善について前記最初の分解タグスワップ E. coli 株ライブラリーの個々の E. coli 株をスクリーニングおよび選択することによって、表現型性能改善をもたらすユニークな遺伝的バリエーションを同定するステップと、

d. 前記先行するステップにおいてスクリーニングされた少なくとも 2 つの個々の E. coli 株に存在する遺伝的バリエーションからのユニークな遺伝的バリエーションの組合せをそれぞれが含む後続の複数の E. coli 微生物を提供することによって、後続の分解タグスワップ E. coli 株ライブラリーを創製するステップと；

e. 前記参照 E. coli 株に対する表現型性能改善について前記後続の分解タグスワップ E. coli 株ライブラリーの個々の E. coli 株をスクリーニングおよび選択することによって、追加の表現型性能改善をもたらす遺伝的バリエーションのユニークな組合せを同定するステップと；

f. E. coli 株が、前記産生 E. coli 株の前記表現型性能と比較して、所望のレベルの表現型性能の改善を呈するまで、線形または非線形様式でステップ d) ~ e) を 1 回または複数回繰り返すステップであって、各後続の反復により、微生物株の新しい分解タグスワップ E. coli 株ライブラリーが創製され、前記新しいライブラリー内の各株は、先行するライブラリーの少なくとも 2 つの個々の E. coli 株の中から選択される遺伝的バリエーションの組合せである遺伝的バリエーションを含む、ステップとを含む、方法。

【請求項 81】

前記後続の分解タグスワップ E. coli 株ライブラリーが、前記最初の分解タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである、請求項 80 に記載の産生 E. coli 株の表現型性能を改善するための分解タグスワップ方法。

【請求項 82】

前記後続の分解タグスワップ E. coli 株ライブラリーが、前記最初の分解タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットであ

10

20

30

40

50

る、請求項 80 に記載の産生 E. coli 株の表現型性能を改善するための分解タグスワップ方法。

【請求項 83】

前記後続の分解タグスワップ E. coli 株ライブラリーが、先行する分解タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである、請求項 80 に記載の産生 E. coli 株の表現型性能を改善するための分解タグスワップ方法。

【請求項 84】

前記後続の分解タグスワップ E. coli 株ライブラリーが、先行する分解タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである、請求項 80 に記載の産生 E. coli 株の表現型性能を改善するための分解タグスワップ方法。

10

【請求項 85】

後続の分解タグスワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 10% の増加を呈するまで、ステップ d) ~ e) を繰り返す、請求項 80 から 84 までのいずれか一項に記載の産生 E. coli 株の表現型性能を改善するための分解タグスワップ方法。

【請求項 86】

後続の分解タグスワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 1 倍の増加を呈するまで、ステップ d) ~ e) を繰り返す、請求項 80 から 84 までのいずれか一項に記載の産生 E. coli 株の表現型性能を改善するための分解タグスワップ方法。

20

【請求項 87】

ステップ f) の表現型性能の改善が、目的の産物の体積生産性、目的の産物の比生産性、目的の産物の収率、目的の産物の力価、およびこれらの組合せからなる群から選択される、請求項 80 から 84 までのいずれか一項に記載の産生 E. coli 株の表現型性能を改善するための分解タグスワップ方法。

【請求項 88】

ステップ f) の表現型性能の改善が、目的の産物の増大したまたはより効率的な産生であり、前記目的の産物は、低分子、酵素、ペプチド、アミノ酸、有機酸、合成化合物、燃料、アルコール、一次細胞外代謝産物、二次的細胞外代謝産物、細胞内コンポーネント分子、およびこれらの組合せからなる群から選択される、請求項 80 から 84 までのいずれか一項に記載の産生 E. coli 株の表現型性能を改善するための分解タグスワップ方法。

30

【請求項 89】

微生物宿主細胞における発現のための、異種遺伝子に作動可能に連結したキメラ合成プロモーターであって、長さが 60 ~ 90 ヌクレオチドであり、ラムダファージ p_R プロモーターの遠位部分、それぞれ長さが 6 ヌクレオチドであるラムダファージ p_L および p_R プロモーターの可変 -35 および -10 領域、ラムダファージ p_L および p_R プロモーターのコア部分、ならびにラムダファージ p_R プロモーターの 5' UTR / リボソーム結合部位 (RBS) 部分からなる、キメラ合成プロモーター。

40

【請求項 90】

前記ラムダファージ p_R プロモーターの前記遠位部分、前記ラムダファージ p_L および p_R プロモーターの前記可変 -35 および -10 領域、前記ラムダファージ p_L および p_R プロモーターの前記コア部分、ならびに前記ラムダファージ p_R プロモーターの前記 5' UTR / リボソーム結合部位 (RBS) 部分の核酸配列が、表 1. 5に見出される核酸配列から選択される、請求項 89 に記載のキメラ合成プロモーター。

【請求項 91】

微生物宿主細胞における発現のための、異種遺伝子に作動可能に連結したキメラ合成プ

50

ロモーターであって、長さが60～90ヌクレオチドであり、ラムダファージ p_R プロモーターの遠位部分、それぞれ長さが6ヌクレオチドであるラムダファージ p_L および p_R プロモーターの可変-35および-10領域、ラムダファージ p_L および p_R プロモーターのコア部分ならびにE. coli *accs* 遺伝子のプロモーターの5' UTR/リボソーム結合部位(RBS)部分からなる、キメラ合成プロモーター。

【請求項92】

前記ラムダファージ p_R プロモーターの前記遠位部分、前記ラムダファージ p_L および p_R プロモーターの前記可変-35および-10領域、前記ラムダファージ p_L および p_R プロモーターの前記コア部分、ならびに前記E. coli *accs* 遺伝子の前記プロモーターの前記5' UTR/リボソーム結合部位(RBS)部分の核酸配列が、表1. 5に見出される核酸配列から選択される、請求項91に記載のキメラ合成プロモーター。

10

【請求項93】

配列番号132～152、159～160、162、165、174～175、188、190、199～201または207から選択される核酸配列からなる、請求項89から90までのいずれかに記載のキメラ合成プロモーター。

【請求項94】

配列番号153～158、161、163～164、166～173、176～187、189、191～198または202～206から選択される核酸配列からなる、請求項91から92までのいずれかに記載のキメラ合成プロモーター。

【請求項95】

前記微生物宿主細胞がE. coliである、請求項89から92までのいずれかに記載のキメラ合成プロモーター。

20

【請求項96】

前記異種遺伝子が、表2に見出される目的のタンパク質産物をコードする、請求項95に記載のキメラ合成プロモーター。

【請求項97】

前記異種遺伝子が、リシン生合成経路の一部である遺伝子である、請求項95に記載のキメラ合成プロモーター。

【請求項98】

前記異種遺伝子が、asd遺伝子、ask遺伝子、hom遺伝子、dapA遺伝子、dapB遺伝子、dapD遺伝子、ddh遺伝子、argD遺伝子、dapE遺伝子、dapF遺伝子、lysA遺伝子、lysE遺伝子、zwf遺伝子、pgi遺伝子、ktk遺伝子、fbp遺伝子、ppc遺伝子、pcK遺伝子、ddx遺伝子、pyc遺伝子またはicd遺伝子から選択される、請求項97に記載のキメラ合成プロモーター。

30

【請求項99】

前記異種遺伝子が、リコピン生合成経路の一部である遺伝子である、請求項95に記載のキメラ合成プロモーター。

【請求項100】

前記異種遺伝子が、dxs遺伝子、ispC遺伝子、ispE遺伝子、ispD遺伝子、ispF遺伝子、ispG遺伝子、ispH遺伝子、idi遺伝子、ispA遺伝子、ispB遺伝子、crtE遺伝子、crtB遺伝子、crtI遺伝子、crtY遺伝子、ymgA遺伝子、dxr遺伝子、elbA遺伝子、gdhA遺伝子、appY遺伝子、elbB遺伝子、またはymgB遺伝子から選択される、請求項99に記載のキメラ合成プロモーター。

40

【請求項101】

前記異種遺伝子が、生物学的製剤をコードするまたは生物学的製剤を生成するための経路内の遺伝子である、請求項95に記載のキメラ合成プロモーター。

【請求項102】

前記生物学的製剤が、ヒューマリン(rhインスリン)、イントロンA(インターフェロンアルファ2b)、ロフェロン(インターフェロンアルファ2a)、ヒューマトロップ

50

(ソマトロピン r h 成長ホルモン)、ニューボジェン (フィルグラスチム)、ベタフェロン (インターフェロンベーター 1 b)、リスプロ (速効型インスリン)、ラピリシン (レテプラゼ)、インフェルゲン (インターフェロンアルファコンー 1)、グルカゴン、ベロムン (タソネルミン)、オンタック (デニロイキンジフチトクス)、ランタス (長時間作用型インスリングラルギン)、キネレット (アナキンラ)、ナトレコール (ネシリチド)、ソマバート (ペグビソマント)、カルシトニン (組換えカルシトニン サケ)、ルセンティス (ラニビズマブ)、プレオタクト (ヒト副甲状腺ホルモン)、クリステクサ (r h 尿酸オキシダーゼ、PEG 化)、ニベスチム (フィルグラスチム、r h G C S F)、ボラキサゼ (グルカルピダーゼ)、またはプレオス (副甲状腺ホルモン) から選択される、請求項 99 に記載のキメラ合成プロモーター。

10

【請求項 103】

配列番号 132 ~ 207 から選択される核酸配列を有するキメラ合成プロモーターに作動可能に連結した異種遺伝子。

【請求項 104】

表 2 に見出される目的のタンパク質産物をコードする、請求項 103 に記載の異種遺伝子。

【請求項 105】

リシン生合成経路の一部である遺伝子である、請求項 103 に記載の異種遺伝子。

【請求項 106】

a s d 遺伝子、a s k 遺伝子、h o m 遺伝子、d a p A 遺伝子、d a p B 遺伝子、d a p D 遺伝子、d d h 遺伝子、a r g D 遺伝子、d a p E 遺伝子、d a p F 遺伝子、l y s A 遺伝子、l y s E 遺伝子、z w f 遺伝子、p g i 遺伝子、k t k 遺伝子、f b p 遺伝子、p p c 遺伝子、p c k 遺伝子、d d x 遺伝子、p y c 遺伝子または i c d 遺伝子から選択される、請求項 105 に記載の異種遺伝子。

20

【請求項 107】

リコピン生合成経路の一部である遺伝子である、請求項 103 に記載の異種遺伝子。

【請求項 108】

d x s 遺伝子、i s p C 遺伝子、i s p E 遺伝子、i s p D 遺伝子、i s p F 遺伝子、i s p G 遺伝子、i s p H 遺伝子、i d i 遺伝子、i s p A 遺伝子、i s p B 遺伝子、c r t E 遺伝子、c r t B 遺伝子、c r t I 遺伝子、c r t Y 遺伝子、y m g A 遺伝子、d x r 遺伝子、e l b A 遺伝子、g d h A 遺伝子、a p p Y 遺伝子、e l b B 遺伝子、または y m g B 遺伝子から選択される、請求項 107 に記載の異種遺伝子。

30

【請求項 109】

生物学的製剤をコードするまたは生物学的製剤を生成するための経路内の遺伝子である、請求項 103 に記載の異種遺伝子。

【請求項 110】

前記生物学的製剤が、ヒューマリン (r h インスリン)、イントロン A (インターフェロンアルファ 2 b)、ロフェロン (インターフェロンアルファ 2 a)、ヒューマトローブ (ソマトロピン r h 成長ホルモン)、ニューボジェン (フィルグラスチム)、ベタフェロン (インターフェロンベーター 1 b)、リスプロ (速効型インスリン)、ラピリシン (レテプラゼ)、インフェルゲン (インターフェロンアルファコンー 1)、グルカゴン、ベロムン (タソネルミン)、オンタック (デニロイキンジフチトクス)、ランタス (長時間作用型インスリングラルギン)、キネレット (アナキンラ)、ナトレコール (ネシリチド)、ソマバート (ペグビソマント)、カルシトニン (組換えカルシトニン サケ)、ルセンティス (ラニビズマブ)、プレオタクト (ヒト副甲状腺ホルモン)、クリステクサ (r h 尿酸オキシダーゼ、PEG 化)、ニベスチム (フィルグラスチム、r h G C S F)、ボラキサゼ (グルカルピダーゼ)、またはプレオス (副甲状腺ホルモン) から選択される、請求項 109 に記載の異種遺伝子。

40

【発明の詳細な説明】

【技術分野】

50

【0001】

関連出願の相互参照

本出願は、2017年6月6日に提出された米国仮出願第62/515,870号の優先権の利益を主張し、これは、すべての目的のために、その全体が参照によって本明細書に組み込まれている。

【0002】

分野

本開示は、*Escherichia coli*についてのハイスループット（HTP）微生物ゲノム操作を対象とする。開示されるHTPゲノム操作プラットフォームは、コンピュータで駆動され、分子生物学、自動化、および高度な機械学習プロトコルを統合する。この統合プラットフォームは、一式のHTP分子ツールセットを利用して、HTP遺伝子設計ライブラリーを生み出す。これらのライブラリーは、とりわけ、科学的洞察および反復パターン認識から得られる。

10

【0003】

配列表に関する陳述

本願に添付された配列表は、紙コピーの代わりにテキストフォーマットで提供されており、参照により本明細書に組み込まれている。配列表を含有するテキストファイルの名称は、ZYMR__012__01WO__SeqList__ST25.txtである。テキストファイルは、約127KBであり、2018年6月6日に作成されたものであり、EFS-Webを介して電子的に提出されている。

20

【背景技術】

【0004】

背景

人間は、目的の産物を生成するのに千年間にわたって微生物細胞の生合成経路の力を利用してきており、その最古の例としては、アルコール、酢、チーズ、およびヨーグルトがある。これらの産物は、今日でも依然として大きな需要があり、また、微生物が産生可能な産物のレパートリーが増え続けている。遺伝子操作技術の出現により、科学者は、様々な生物体中に新規生合成経路を設計およびプログラムして、広い範囲の工業、医療用、および消費者向け製品を生産することが可能になった。実際に、微生物細胞培養が、低分子、抗生物質、ワクチン、殺虫剤、酵素、燃料、および工業用化学物質におよぶ産物を生産するのに現在使用されている。

30

【0005】

最新の工業用微生物によって産生される多数の産物を考慮すると、エンジニアが、所与の微生物が標的産物を産生することができる速度および効率を改善するのに極めて大きいプレッシャーの下にあることは驚くに当たらない。様々な手法が、関与する微生物を「改良する」ことによって生物学に基づく工業プロセスの経済性を改善するのに使用されてきた。例えば、多くの製薬工業および化学工業では、微生物培養の親株が化学物質またはUV放射線への曝露によって連続的に変異され、生産性、収率、および力価などの性能増大について引き続いてスクリーニングされる、微生物株改良プログラムを利用する。この変異誘発プロセスは、株が産物性能の適当な増大を実証するまで広範に繰り返される。次いで後続の「改善された」株は、商業生産で利用される。変異誘発による改善された工業用微生物株の同定は、時間がかかり、非効率である。このプロセスは、まさにその本質から、計画性のないものであり、産物産出について望ましい結果を有する変異の際のつまづきを利用する。伝統的な微生物株改良プログラムが非効率であるだけでなく、このプロセスは、高い程度の有害な変異誘発性ロード（mutagenic load）を有する工業用株に至る場合もある。これらのタイプのプログラムに供された工業用株における変異の蓄積は、かなり大きなものとなり得、性能改善の速度における最終的な停滞に至り得る。

40

【0006】

伝統的な株改良プログラムの結果として生じる停滞に関して既存の最も操作された微生物宿主系の1つである*E. coli*を用いるものよりも良い例はおそらく存在しない。微

50

生物は数十年にわたって上述の伝統的な微生物株改良の方法に供されてきた。E. coli を操作するために膨大な量の労力がつぎ込まれてきたにもかかわらず、なお微生物には膨大な量のまだ利用されていない潜在性がある。これは、E. coli が微生物を産生目的で改良しようと試みる研究者に対してユニークな課題を提示するからである。これらの課題により、E. coli におけるゲノム操作の分野が妨害され、研究者がこの微生物系の完全な潜在性を利用することが妨げられてきた。

【0007】

特に、工業界はE. coli におけるゲノム操作のためのハイスループットなシステムを未だ開発していない。伝統的な株改良の方法はこの生物体系に関して停滞期に達していることが明らかであるが、研究者はこの停滞期を越えるために必要なゲノム操作ツールを未だ有していない。

10

【0008】

したがって、従来の株改良プログラムに固有の上述した欠点を被らないE. coli を産生目的で操作する新しい方法についての当技術分野における大きな必要性が存在する。具体的には、E. coli における有益な変異を発見し、強固にするハイスループットなシステムにより、当該分野が改革され、研究者がこの生物体の完全な潜在性を利用することが可能になる。

【発明の概要】

【課題を解決するための手段】

【0009】

20

開示の概要

本開示は、伝統的な微生物株改良プログラムに関連した無数の問題を被らないE. coli のためのハイスループット (HTP) ゲノム操作プラットフォームを提供する。

【0010】

さらに、本明細書で教示されるHTPプラットフォームは、数10年のランダム変異誘発に基づく株改良プログラムによって有益でない変異を蓄積してきたE. coli 株を再建する (rehabilitate) ことができる。

【0011】

本開示は、E. coli 系におけるHTPプラットフォームの機能性を補強するユニークなゲノム操作ツールセットおよび手順も提供する。

30

【0012】

開示されるHTPゲノム操作プラットフォームは、コンピューターで駆動され、分子生物学、自動化、および高度な機械学習プロトコルを統合する。この統合プラットフォームは、一式のHTP分子ツールセットを利用して、HTP遺伝子設計ライブラリーを生み出す。これらのライブラリーは、とりわけ、科学的洞察および反復パターン認識から得られる。

【0013】

教示されるHTP遺伝子設計ライブラリーは、E. coli において試験するための特定のゲノム変更のライブラリーを提供することによって、ゲノム操作プロセスのドライバーとして機能する。特定のライブラリーまたはライブラリーの組合せを利用して操作された微生物は、結果として生じる成果、例えば、目的の産物の産生についてHTP様式で効率的にスクリーニングされる。HTP遺伝子設計ライブラリーを利用して微生物において試験するための特定のゲノム変更を明らかにし、次いでこれらの変更を内包する宿主微生物ゲノムを引き続いてスクリーニングするこのプロセスは、効率的な反復様式でインプリメントされる。一部の態様では、ゲノム操作キャンペーンの反復サイクルまたは「ラウンド」は、少なくとも1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、またはそれ超の反復／サイクル／ラウンドであり得る。

40

【0014】

よって、一部の態様では、本開示は、E. coli 宿主系において少なくとも1、2、

50

3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100、125、150、175、200、225、250、275、300、325、350、375、400、425、450、475、500、525、550、575、600、625、650、675、700、725、750、775、800、825、850、875、900、925、950、975、1000またはそれ超の「ラウンド」のHTP遺伝子操作（例えば、SNPスワップ、PROスワップ、STOPスワップ、またはこれらの組合せのラウンド）を行う方法を教示する。

10

【0015】

一部の実施形態では、本開示は、各後続のHTP遺伝子操作ラウンドが先のラウンドの遺伝子操作で同定された遺伝的バリエーションに基づく線形手法を教示する。他の実施形態では、本開示は、各後続のHTP遺伝子操作ラウンドが先に行われた分析を含む任意の先のラウンドの遺伝子操作、および別個のHTP遺伝子操作ブランチで同定された遺伝的バリエーションに基づく非線形手法を教示する。

20

【0016】

これらの反復サイクルからのデータは、大規模データ分析およびパターン認識を可能にし、それは統合プラットフォームによって利用され、HTP遺伝子設計ライブラリー実装の後続のラウンドを通知する。したがって、教示されるプラットフォームで利用されるHTP遺伝子設計ライブラリーは、大規模データパターン認識アルゴリズムから利益を得、各反復ラウンドの微生物操作によってより有益となる高度に動的なツールである。このようなシステムはE. coliに対しては開発されたことがなく、当技術分野において切実に必要とされている。

【0017】

一部の実施形態では、本開示の遺伝子設計ライブラリーは、少なくとも1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100、125、150、175、200、225、250、275、300、325、350、375、400、425、450、475、500、525、550、575、600、625、650、675、700、725、750、775、800、825、850、875、900、925、950、975、1000、またはそれ超の個々の遺伝子変化（例えば、少なくともXの数のPROスワップライブラリーにおけるプロモーター：遺伝子組合せの）を含む。

30

40

【0018】

一部の実施形態では、本開示は、所望の表現型を獲得するためにE. coli株を進化させるゲノム操作のハイスループット（HTP）方法であって、a）同じ株バックグラウンドを有する最初の複数のE. coli株のゲノムを攪乱することによって、ユニークな遺伝的バリエーションを有する個々の株を含む最初のHTP遺伝子設計E. coli株ライブラリーを生み出すステップと；b）所望の表現型について最初のHTP遺伝子設計E. coli株ライブラリーの個々の株をスクリーニングおよび選択するステップと；c）それぞれが遺伝的バリエーションのユニークな組合せを含む後続の複数のE. coli微

50

生物を提供することによって、後続のHTP遺伝子設計E. coli株ライブラリーを生み出すステップであって、該遺伝的バリエーションは、先行するステップにおいてスクリーニングされた少なくとも2つの個々のE. coli株に存在する遺伝的バリエーションから選択される、ステップと；d) 所望の表現型について後続のHTP遺伝子設計E. coli株ライブラリーの個々のE. coli株をスクリーニングおよび選択するステップと；e) E. coli株が所望の表現型を獲得するまで、線形または非線形様式でステップc)～d)を1回または複数回繰り返すステップであって、各後続の反復により、先行するHTP遺伝子設計E. coli株ライブラリーの少なくとも2つの個々のE. coli株の中から選択される遺伝的バリエーションの組合せであるユニークな遺伝的バリエーションを内包する個々のE. coli株を含む新しいHTP遺伝子設計E. coli株ライブラリーが創製される、ステップとを含む、方法を教示する。 10

【0019】

一部の実施形態では、本開示は、最初のHTP遺伝子設計E. coli株ライブラリーが、プロモータースワップ微生物株ライブラリー、SNPスワップ微生物株ライブラリー、開始／停止コドン微生物株ライブラリー、最適化配列微生物株ライブラリー、ターミネータースワップ微生物株ライブラリー、タンパク質可溶性タグ微生物株ライブラリー、タンパク質分解タグ微生物株ライブラリーまたはこれらの任意の組合せからなる群から選択される少なくとも1つであることを教示する。

【0020】

一部の実施形態では、本開示は、それぞれが遺伝的バリエーションのユニークな組合せを含む後続の複数のE. coliを作製する方法であって、組み合わせられた遺伝的バリエーションのそれぞれが、最初のHTP遺伝子設計E. coli株ライブラリーまたは先行するステップのHTP遺伝子設計E. coli株ライブラリーに由来する、方法を教示する。 20

【0021】

一部の実施形態では、後続の複数のE. coli株における遺伝的バリエーションの組合せは、最初のHTP遺伝子設計E. coli株ライブラリーまたは先行するステップのHTP遺伝子設計E. coli株ライブラリーにおける遺伝的バリエーションのすべての可能な組合せのサブセットを含むことになる。

【0022】

一部の実施形態では、本開示は、後続のHTP遺伝子設計E. coli株ライブラリーが、最初のHTP遺伝子設計E. coli株ライブラリーまたは先行するステップのHTP遺伝子設計E. coli株ライブラリーにおける遺伝的バリエーションに由来する完全コンビナトリアルE. coli株ライブラリーであることを教示する。 30

【0023】

例えば、先のHTP遺伝子設計E. coli株ライブラリーが遺伝的バリエーションA、B、C、およびDのみを有している場合、前記バリエーションの部分的コンビナトリアルは、それぞれがAB、AC、またはADの遺伝的バリエーションのユニークな組合せのいずれかを含む3種の株を含む後続のHTP遺伝子設計E. coli株ライブラリーを含むことができる（変異が表される順序は重要でない）。先行するステップのHTP遺伝子設計ライブラリーの遺伝的バリエーションに由来する完全コンビナトリアルE. coli株ライブラリーは、それぞれがAB、AC、AD、BC、BD、またはCDの遺伝的バリエーションのユニークな組合せのいずれかを含む6種の微生物を含むはずである。 40

【0024】

一部の実施形態では、本開示の方法は、ランダム変異誘発、標的とした配列挿入、標的とした配列欠失、標的とした配列置き換え、またはこれらの任意の組合せからなる群から選択される少なくとも1つの方法を利用してE. coliのゲノムを攪乱することを教示する。

【0025】

現在開示した方法の一部の実施形態では、最初の複数のE. coliは、工業生産E. 50

c o l i 株に由来するユニークな遺伝的バリエーションを含む。

【0026】

現在開示した方法の一部の実施形態では、最初の複数の E. c o l i は、S₁ G e n₁ で表される工業生産 E. c o l i 株、および S_n G e n_n で表されるこれらに由来する任意の数の後続の微生物世代を含む。

【0027】

一部の実施形態では、本開示は、SNPスワップ E. c o l i 株ライブラリーを生成するための方法であって、a) 参照 E. c o l i 株および第2の E. c o l i 株を提供するステップであって、第2の E. c o l i 株は、参照株内に存在しない一塩基多型、DNA挿入、およびDNA欠失から選択される複数の同定された遺伝的バリエーションを含む、ステップと；b) 参照株または第2の株のいずれかのゲノムを攪乱することによって、複数の個々の E. c o l i 株を含む最初の SNPスワップ E. c o l i 株ライブラリーを生み出すステップであって、前記複数の個々の E. c o l i 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、参照株と第2の株との間の複数の同定された遺伝的バリエーションから選択される単一の遺伝的バリエーションに対応する、ステップとを含む、方法を教示する。

【0028】

SNPスワップライブラリーの一部の実施形態では、参照 E. c o l i 株のゲノムが攪乱されて、第2の E. c o l i 株内に見出される同定された一塩基多型、DNA挿入、またはDNA欠失の1つまたは複数が付加される。

【0029】

SNPスワップライブラリーの一部の実施形態では、第2の E. c o l i 株のゲノムが攪乱されて、参照 E. c o l i 株内に見出されない同定された一塩基多型、DNA挿入、またはDNA欠失の1つまたは複数が除去される。

【0030】

一部の実施形態では、SNPスワップライブラリーの遺伝的バリエーションは、参照 E. c o l i 株と第2の E. c o l i 株との間で同定されたすべての遺伝的バリエーションのサブセットを含む。

【0031】

一部の実施形態では、SNPスワップライブラリーの遺伝的バリエーションは、参照 E. c o l i 株と第2の E. c o l i 株との間の同定された遺伝的バリエーションのすべてを含む。

【0032】

一部の実施形態では、本開示は、工業用 E. c o l i 株の表現型性能を再建および改善するための方法であって、a) 親系統 E. c o l i 株およびそれに由来する工業用 E. c o l i 株を提供するステップであって、工業用株は、親系統株内に存在しない一塩基多型、DNA挿入、およびDNA欠失から選択される複数の同定された遺伝的バリエーションを含む、ステップと；b) 親系統株または工業用株のいずれかのゲノムを攪乱することによって、複数の個々の株を含む最初の SNPスワップ E. c o l i 株ライブラリーを生み出すステップであって、前記複数の個々の株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、親系統株と工業用株との間の複数の同定された遺伝的バリエーションから選択される単一の遺伝的バリエーションに対応する、ステップと；c) 参照 E. c o l i 株に対する表現型性能改善について最初の SNPスワップ E. c o l i 株ライブラリーの個々の株をスクリーニングおよび選択することによって、前記 E. c o l i 株に表現型性能改善をもたらすユニークな遺伝的バリエーションを同定するステップと；d) それぞれが遺伝的バリエーションのユニークな組合せを含む後続の複数の E. c o l i 株を提供することによって、後続の SNPスワップ E. c o l i 株ライブラリーを生み出すステップであって、前記遺伝的バリエーションは、先行するステップにおいてスクリーニングされた少なくとも2つの個々の株に存在する遺伝的バリエーションから選択される、ステップと；e) 参照株に対する表現型性能改

善について後続のSNPスワップE. coli株ライブラリーの個々の株をスクリーニングおよび選択することによって、前記E. coli株に追加の表現型性能改善をもたらす遺伝的バリエーションのユニークな組合せを同定するステップと；f)株が、工業用E. coli株の表現型性能と比較して、所望のレベルの表現型性能の改善を呈するまで、線形または非線形様式でステップd)～e)を1回または複数回繰り返すステップであって、各後続の反復により、先行するSNPスワップE. coli株ライブラリーの少なくとも2つの個々の微生物株の中から選択される遺伝的バリエーションの組合せであるユニークな遺伝的バリエーションを内包する個々の微生物株を含む新しいSNPスワップE. coli株ライブラリーが創製される、ステップとを含む、方法を教示する。

【0033】

一部の実施形態では、本開示は、工業用E. coli株の表現型性能を再建および改善するための方法であって、親系統E. coli株のゲノムが攪乱されて、工業用E. coli株内に見出される同定された一塩基多型、DNA挿入、またはDNA欠失の1つまたは複数が付加される、方法を教示する。

【0034】

一部の実施形態では、本開示は、工業用E. coli株の表現型性能を再建および改善するための方法であって、工業用E. coli株のゲノムが攪乱されて、親系統E. coli株内に見出されない同定された一塩基多型、DNA挿入、またはDNA欠失の1つまたは複数が除去される、方法を教示する。

【0035】

一部の実施形態では、本開示は、プロモータースワップE. coli株ライブラリーを生成するための方法であって、a)基準E. coli株に内在する複数の標的遺伝子、およびプロモーターラダーを提供するステップであって、前記プロモーターラダーは、基準E. coli株において異なる発現プロファイルを呈する複数のプロモーターを含む、ステップと；b)基準E. coli株のゲノムを操作することによって、複数の個々のE. coli株を含む最初のプロモータースワップE. coli株ライブラリーを生み出すステップであって、前記複数の個々の株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、基準E. coli株に内在する標的遺伝子の1つに作動可能に連結したプロモーターラダー由来のプロモーターの1つを含む、ステップとを含む、方法を教示する。

【0036】

一部の実施形態では、本開示は、所望の表現型を獲得するためにE. coli株を進化させるゲノム操作のプロモータースワップ方法であって、a)基準E. coli株に内在する複数の標的遺伝子、およびプロモーターラダーを提供するステップであって、前記プロモーターラダーは、基準E. coli株において異なる発現プロファイルを呈する複数のプロモーターを含む、ステップと；b)基準E. coli株のゲノムを操作することによって、複数の個々のE. coli株を含む最初のプロモータースワップE. coli株ライブラリーを生み出すステップであって、前記複数の個々の株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、基準E. coli株に内在する標的遺伝子の1つに作動可能に連結したプロモーターラダー由来のプロモーターの1つを含む、ステップと；c)所望の表現型について最初のプロモータースワップE. coli株ライブラリーの個々の株をスクリーニングおよび選択するステップと；d)それぞれが遺伝的バリエーションのユニークな組合せを含む後続の複数のE. coli株を提供することによって、後続のプロモータースワップE. coli株ライブラリーを生み出すステップであって、前記遺伝的バリエーションは、先行するステップにおいてスクリーニングされた少なくとも2つの個々の株に存在する遺伝的バリエーションから選択される、ステップと；e)所望の表現型について後続のプロモータースワップE. coli株ライブラリーの個々の株をスクリーニングおよび選択するステップと；f)微生物が所望の表現型を獲得するまで、線形または非線形様式でステップd)～e)を1回または複数回繰り返すステップであって、各後続の反復により、先行するプロモ

10

20

30

40

50

ータースワップ E. coli 株ライブラリーの少なくとも 2 つの個々の株の中から選択される遺伝的バリエーションの組合せであるユニークな遺伝的バリエーションを内包する個々の株を含む新しいプロモータースワップ E. coli 株ライブラリーが創製される、ステップとを含む、方法を教示する。

【0037】

一部の実施形態では、本開示は、ターミネータースワップ E. coli 株ライブラリーを生成するための方法であって、a) 基準 E. coli 株に内在する複数の標的遺伝子、およびターミネーターラダーを提供するステップであって、前記ターミネーターラダーは、基準 E. coli 株において異なる発現プロファイルを呈する複数のターミネーターを含む、ステップと；b) 基準 E. coli 株のゲノムを操作することによって、複数の個々の株を含む最初のターミネータースワップ E. coli 株ライブラリーを生み出すステップであって、前記複数の個々の株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、ターミネーターラダー由来のターミネーターの 1 つまたは複数に作動可能に連結した基準 E. coli 株に内在する標的遺伝子の 1 つを含む、ステップとを含む、方法を教示する。

【0038】

一部の実施形態では、本開示は、所望の表現型を獲得するために E. coli を進化させるゲノム操作のターミネータースワップ方法であって、a) 基準 E. coli 株に内在する複数の標的遺伝子、およびターミネーターラダーを提供するステップであって、前記ターミネーターラダーは、基準 E. coli 株において異なる発現プロファイルを呈する複数のターミネーターを含む、ステップと；b) 基準 E. coli 株のゲノムを操作することによって、複数の個々の E. coli 株を含む最初のターミネータースワップ E. coli 株ライブラリーを生み出すステップであって、前記複数の個々の株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、ターミネーターラダー由来のターミネーターの 1 つまたは複数に作動可能に連結した基準 E. coli 株に内在する標的遺伝子の 1 つを含む、ステップと；c) 所望の表現型について最初のターミネータースワップ E. coli 株ライブラリーの個々の微生物株をスクリーニングおよび選択するステップと；d) それぞれが遺伝的バリエーションのユニークな組合せを含む後続の複数の E. coli 株を提供することによって、後続のターミネータースワップ E. coli 株ライブラリーを生み出すステップであって、前記遺伝的バリエーションは、先行するステップにおいてスクリーニングされた少なくとも 2 つの個々の株に存在する遺伝的バリエーションから選択される、ステップと；e) 所望の表現型について後続のターミネータースワップ E. coli 株ライブラリーの個々の株をスクリーニングおよび選択するステップと；f) 微生物が所望の表現型を獲得するまで、線形または非線形様式でステップ d) ~ e) を 1 回または複数回繰り返すステップであって、各後続の反復により、先行するターミネータースワップ E. coli 株ライブラリーの少なくとも 2 つの個々の株の中から選択される遺伝的バリエーションの組合せであるユニークな遺伝的バリエーションを内包する個々の株を含む新しいターミネータースワップ E. coli 株ライブラリーが創製される、ステップとを含む、方法を教示する。

【0039】

一部の実施形態では、本開示は、(a) (1) 1 つまたは複数のバックグラウンド E. coli 株に対する遺伝子変化を表す入力、および (2) 対応する性能尺度を含む訓練セットで設定された (populated with) 予測モデルにアクセスするステップと；(b) 遺伝子変化を表す予測モデルに試験入力 (test input) を適用するステップであって、試験入力は、これらの遺伝子変化を組み入れる候補 E. coli 株に対応する、ステップと；(c) 予測モデルに少なくとも部分的に基づいて候補 E. coli 株の表現型性能を予測するステップと；(d) これらの予測された性能に少なくとも部分的に基づいて候補 E. coli 株の第 1 のサブセットを選択するステップと；(e) 候補 E. coli 株の第 1 のサブセットの測定された表現型性能を得るステップと；(f) これらの測定された表現型性能に少なくとも部分的に基づいて候補 E. coli 株の第 2 のサブセットの選択物

を得るステップと；（g）（2）選択された候補 E. coli 株の第 2 のサブセットの対応する測定された性能とともに（1）選択された候補 E. coli 株の第 2 のサブセットに対応する入力を予測モデルの訓練セットに加えるステップと；（h）少なくとも 1 つの候補 E. coli 株の測定された表現型性能が性能測定基準を満たすまで、（b）～（g）を繰り返すステップとによって候補 E. coli 株の設計を反復して改善することを教示する。一部の場合では、予測モデルへの試験入力第 1 の適用の間に、試験入力によって表される遺伝子変化が、1 つまたは複数のバックグラウンド E. coli 株に対する遺伝子変化を含み、試験入力の後続の適用の間に、試験入力によって表される遺伝子変化が、候補 E. coli 株の先に選択された第 2 のサブセット内の候補 E. coli 株に対する遺伝子変化を含む。

10

【0040】

一部の実施形態では、第 1 のサブセットの選択は、上位性効果に基づき得る。これは、第 1 のサブセットの第 1 の選択の間に、1 つまたは複数のバックグラウンド E. coli 株に対する遺伝子変化を表す複数のそれぞれの入力の適用に応じて、1 つまたは複数のバックグラウンド E. coli 株の性能尺度間の相違点の程度を決定するステップと、少なくとも 2 つの候補 E. coli 株内に組み入れられる遺伝子変化の適用に応じて、1 つまたは複数のバックグラウンド E. coli 株の性能尺度の相違点の程度に少なくとも部分的に基づいて、少なくとも 2 つの候補 E. coli 株を第 1 のサブセットに含めるために選択するステップとによって達成することができる。

20

【0041】

一部の実施形態では、本発明は、候補 E. coli 株の反復改良において上位性効果を適用する方法であって、少なくとも 1 つの E. coli バックグラウンド株に行われた対応する遺伝子変化に応じて、測定された性能を表すデータを得るステップと；少なくとも 2 つの遺伝子変化の対応する応答性能尺度間の相違点の程度に少なくとも部分的に基づいて、少なくとも 2 つの遺伝子変化の選択物を得るステップであって、相違点の程度が、少なくとも 2 つの遺伝子変化が異なる生物学的経路を通じたこれらの対応する応答性能尺度に影響する程度に関連する、ステップと；選択された遺伝子変化を含む E. coli バックグラウンド株に対する遺伝子変化を設計するステップとを含む、方法を教示する。一部の場合では、少なくとも 2 つの選択された遺伝子変化が設計される E. coli バックグラウンド株は、測定された応答性能を表すデータが得られた少なくとも 1 つの E. coli

30

【0042】

一部の実施形態では、本開示は、単一タイプの遺伝子ライブラリーのみを利用する HTP E. coli 株改良法を教示する。例えば、一部の実施形態では、本開示は、SNP スワップライブラリーのみを利用する HTP E. coli 株改良法を教示する。他の実施形態では、本開示は、PRO スワップライブラリーのみを利用する HTP E. coli 株改良法を教示する。一部の実施形態では、本開示は、STOP スワップライブラリーのみを利用する HTP 株改良法を教示する。一部の実施形態では、本開示は、開始／停止コドン スワップライブラリーのみを利用する HTP E. coli 株改良法を教示する。

【0043】

他の実施形態では、本開示は、2 つまたはそれ超のタイプの遺伝子ライブラリーを利用する HTP E. coli 株改良法を教示する。例えば、一部の実施形態では、本開示は、SNP スワップライブラリーおよび PRO スワップライブラリーを組み合わせる HTP E. coli 株改良法を教示する。一部の実施形態では、本開示は、SNP スワップライブラリーおよび STOP スワップライブラリーを組み合わせる HTP E. coli 株改良法を教示する。一部の実施形態では、本開示は、PRO スワップライブラリーおよび STOP スワップライブラリーを組み合わせる HTP E. coli 株改良法を教示する。

40

【0044】

他の実施形態では、本開示は、複数のタイプの遺伝子ライブラリーを利用する HTP

50

E. coli 株改良法を教示する（例えば、図5を参照のこと）。一部の実施形態では、遺伝子ライブラリーは、組合せ変異（例えば、1つまたは複数の遺伝子に適用されるプロモーター／ターミネーター組合せラダー）を生成するように組み合わせられる。さらに他の実施形態では、本開示のHTP E. coli 株改良法は、1つまたは複数の伝統的な株改良法と組み合わせることができる。

【0045】

一部の実施形態では、本開示のHTP E. coli 株改良法は、改良されたE. coli 宿主細胞をもたらす。すなわち、本開示は、1つまたは複数のE. coli 宿主細胞性質を改善する方法を教示する。一部の実施形態では、改善されるE. coli 宿主細胞性質は、E. coli 宿主細胞によって産生される目的の産物の体積生産性、比生産性、
10 収率、または力価からなる群から選択される。一部の実施形態では、改善されるE. coli 宿主細胞性質は、体積生産性である。一部の実施形態では、改善されるE. coli 宿主細胞性質は、比生産性である。一部の実施形態では、改善されるE. coli 宿主細胞性質は、収率である。

【0046】

一部の実施形態では、本開示のHTP E. coli 株改良法は、HTP株改良法に供されない対照E. coli 宿主細胞に対して1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、150%、200%、250%、300%またはそれよりも大きな少なくとも1つのE. coli 宿主細胞性質の改善（例えば、これらの間の任意の範囲および部分範囲を組み入れての、目的の生体分子の収率または生産性のX%の改善）を呈するE. coli 宿主細胞をもたらす。一部の実施形態では、本開示のHTP E. coli 株改良法は、SNPスワップ、PROスワップ、STOPスワップ、SOLUBILITY TAGスワップ、DEGRADATION TAGスワップ、およびこれらの組合せからなる群から選択される。
20
30

【0047】

よって、一部の実施形態では、本開示のSNPスワップ方法は、SNPスワップ方法に供されない対照E. coli 宿主細胞に対して、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、150%、200%、250%、300%、またはそれ超の少なくとも1つのE. coli 宿主細胞性質の改善（例えば、これらの間の任意の範囲および部分範囲を組み入れての、目的の生体分子の収率または生産性のX%の改善）を呈するE. coli 宿主細胞をもたらす。
40

【0048】

10

20

30

40

50

よって、一部の実施形態では、本開示のPROスワップ方法は、PROスワップ方法に供されない対照E. coli 宿主細胞に対して、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、150%、200%、250%、300%、またはそれ超の少なくとも1つのE. coli 宿主細胞性質の改善（例えば、これらの間の任意の範囲および部分範囲を組み入れての、目的の生体分子の収率または生産性のX%の改善）を呈するE. coli 宿主細胞をもたらす。

10

【0049】

よって、一部の実施形態では、本開示のTERMINATORスワップ方法は、TERMINATORスワップ方法に供されない対照E. coli 宿主細胞に対して、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、150%、200%、250%、300%、またはそれ超の少なくとも1つのE. coli 宿主細胞性質の改善（例えば、これらの間の任意の範囲および部分範囲を組み入れての、目的の生体分子の収率または生産性のX%の改善）を呈するE. coli 宿主細胞をもたらす。

20

30

【0050】

よって、一部の実施形態では、本開示のSOLUBILITY TAGスワップ方法は、SOLUBILITY TAGスワップ方法に供されない対照E. coli 宿主細胞に対して、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、150%、200%、250%、300%、またはそれ超の少なくとも1つのE. coli 宿主細胞性質の改善（例えば、これらの間の任意の範囲および部分範囲を組み入れての、目的の生体分子の収率または生産性のX%の改善）を呈するE. coli 宿主細胞をもたらす。

40

【0051】

よって、一部の実施形態では、本開示のDEGRADATION TAGスワップ方法は、DEGRADATION TAGスワップ方法に供されない対照E. coli 宿主細胞

50

胞に対して、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、150%、200%、250%、300%、またはそれ超の少なくとも1つのE. coli 宿主細胞性質の改善（例えば、これらの間の任意の範囲および部分範囲を組み入れての、目的の生体分子の収率または生産性のX%の改善）を呈するE. coli 宿主細胞をもたらす。

【0052】

一部の実施形態では、本開示は、タンパク質可溶性タグスワップE. coli 株ライブラリーを生成するための方法であって、a. 基準E. coli 株に内在する複数の標的遺伝子、および可溶性タグラダーを提供するステップであって、前記可溶性タグラダーは、基準E. coli 株において異なる可溶性プロファイルを呈する複数の可溶性タグを含む、ステップと；b. 基準E. coli 株のゲノムを操作することによって、複数の個々のE. coli 株を含む最初の可溶性タグスワップE. coli 株ライブラリーを創製するステップであって、前記複数の個々のE. coli 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、基準E. coli 株に内在する標的遺伝子の1つに作動可能に連結した可溶性タグラダー由来の可溶性タグの1つまたは複数を含む、ステップとを含む、方法を教示する。

【0053】

一部の実施形態では、本開示は、産生E. coli 株の表現型性能を改善するためのタンパク質可溶性タグスワップ方法であって、基準E. coli 株に内在する複数の標的遺伝子、および可溶性タグラダーを提供するステップであって、前記可溶性タグラダーは、基準E. coli 株において異なる発現プロファイルを呈する複数の可溶性タグを含む、ステップと；基準E. coli 株のゲノムを操作することによって、複数の個々のE. coli 株を含む最初の可溶性タグスワップE. coli 株ライブラリーを創製するステップであって、前記複数の個々のE. coli 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、基準E. coli 株に内在する標的遺伝子の1つに作動可能に連結した可溶性タグラダー由来の可溶性タグの1つまたは複数を含む、ステップと；参照E. coli 株に対する表現型性能改善について最初の可溶性タグスワップE. coli 株ライブラリーの個々のE. coli 株をスクリーニングおよび選択することによって、表現型性能改善をもたらすユニークな遺伝的バリエーションを同定するステップと；先行するステップにおいてスクリーニングされた少なくとも2つの個々のE. coli 株に存在する遺伝的バリエーションからのユニークな遺伝的バリエーションの組合せをそれぞれが含む後続の複数のE. coli 微生物を提供することによって、後続の可溶性タグスワップE. coli 株ライブラリーを創製するステップと；参照E. coli 株に対する表現型性能改善について後続の可溶性タグスワップE. coli 株ライブラリーの個々のE. coli 株をスクリーニングおよび選択することによって、追加の表現型性能改善をもたらす遺伝的バリエーションのユニークな組合せを同定するステップと；E. coli 株が、産生E. coli 株の表現型性能と比較して、所望のレベルの表現型性能の改善を呈するまで、線形または非線形様式でステップd)～e)を1回または複数回繰り返すステップであって、各後続の反復により、微生物株の新しい可溶性タグスワップE. coli 株ライブラリーが創製され、新しいライブラリー内の各株は、先行するライブラリーの少なくとも2つの個々のE. coli 株の中から

選択される遺伝的バリエーションの組合せである遺伝的バリエーションを含む、ステップとを含む、方法を教示する。

【0054】

一部の実施形態では、後続の可溶性タグスワップ E. coli 株ライブラリーは、最初の可溶性タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである。

【0055】

一部の実施形態では、後続の可溶性タグスワップ E. coli 株ライブラリーは、最初の可溶性タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである。

10

【0056】

一部の実施形態では、後続の可溶性タグスワップ E. coli 株ライブラリーは、先行する可溶性タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである。

【0057】

一部の実施形態では、後続の可溶性タグスワップ E. coli 株ライブラリーは、先行する可溶性タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである。

【0058】

一部の実施形態では、後続の可溶性タグスワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、産生 E. coli 株の表現型性能と比較して、測定された表現型変数における少なくとも 10% の増加を呈するまで、ステップ d) ~ e) を繰り返す。

20

【0059】

一部の実施形態では、後続の可溶性タグスワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、産生 E. coli 株の表現型性能と比較して、測定された表現型変数における少なくとも 1 倍の増加を呈するまで、ステップ d) ~ e) を繰り返す。

【0060】

一部の実施形態では、ステップ f) の表現型性能の改善は、目的の産物の体積生産性、目的の産物の比生産性、目的の産物の収率、目的の産物の力価、およびこれらの組合せからなる群から選択される。

30

【0061】

一部の実施形態では、ステップ f) の表現型性能の改善は、目的の産物の増大したまたはより効率的な産生であり、前記目的の産物は、低分子、酵素、ペプチド、アミノ酸、有機酸、合成化合物、燃料、アルコール、一次細胞外代謝産物、二次的細胞外代謝産物、細胞内コンポーネント分子、およびこれらの組合せからなる群から選択される。

【0062】

一部の実施形態では、本開示は、タンパク質分解タグスワップ E. coli 株ライブラリーを生成するための方法であって、a. 基準 E. coli 株に内在する複数の標的遺伝子、および分解タグラダーを提供するステップであって、前記分解タグラダーは、基準 E. coli 株において異なる可溶性プロファイルを呈する複数の分解タグを含む、ステップと；b. 基準 E. coli 株のゲノムを操作することによって、複数の個々の E. coli 株を含む最初の分解タグスワップ E. coli 株ライブラリーを創製するステップであって、前記複数の個々の E. coli 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、基準 E. coli 株に内在する標的遺伝子の 1 つに作動可能に連結した分解タグラダー由来の分解タグの 1 つまたは複数を含む、ステップとを含む、方法を教示する。

40

【0063】

一部の実施形態では、本開示は、産生 E. coli 株の表現型性能を改善するためのタンパク質分解タグスワップ方法であって、基準 E. coli 株に内在する複数の標的遺伝子、および分解タグラダーを提供するステップであって、前記分解タグラダーは、基準 E

50

． c o l i 株において異なる発現プロファイルを呈する複数の分解タグを含む、ステップと；基準 E． c o l i 株のゲノムを操作することによって、複数の個々の E． c o l i 株を含む最初の分解タグスワップ E． c o l i 株ライブラリーを創製するステップであって、前記複数の個々の E． c o l i 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、基準 E． c o l i 株に内在する標的遺伝子の 1 つに作動可能に連結した分解タグ由来の分解タグの 1 つまたは複数を含む、ステップと；参照 E． c o l i 株に対する表現型性能改善について最初の分解タグスワップ E． c o l i 株ライブラリーの個々の E． c o l i 株をスクリーニングおよび選択することによって、表現型性能改善をもたらすユニークな遺伝的バリエーションを同定するステップと、先行するステップにおいてスクリーニングされた少なくとも 2 つの個々の E． c o l i 株に存在する遺伝的バリエーションからのユニークな遺伝的バリエーションの組合せをそれぞれが含む後続の複数の E． c o l i 微生物を提供することによって、後続の分解タグスワップ E． c o l i 株ライブラリーを創製するステップと；参照 E． c o l i 株に対する表現型性能改善について後続の分解タグスワップ E． c o l i 株ライブラリーの個々の E． c o l i 株をスクリーニングおよび選択することによって、追加の表現型性能改善をもたらす遺伝的バリエーションのユニークな組合せを同定するステップと；E． c o l i 株が、産生 E． c o l i 株の表現型性能と比較して、所望のレベルの表現型性能の改善を呈するまで、線形または非線形様式でステップ d) ～ e) を 1 回または複数回繰り返すステップであって、各後続の反復により、微生物株の新しい分解タグスワップ E． c o l i 株ライブラリーが創製され、新しいライブラリー内の各株は、先行するライブラリーの少なくとも 2 つの個々の E． c o l i 株の中から選択される遺伝的バリエーションの組合せである遺伝的バリエーションを含む、ステップとを含む、方法を教示する。

【0064】

一部の実施形態では、後続の分解タグスワップ E． c o l i 株ライブラリーは、最初の分解タグスワップ E． c o l i 株ライブラリーの完全コンビナトリアルライブラリーである。

【0065】

一部の実施形態では、後続の分解タグスワップ E． c o l i 株ライブラリーは、最初の分解タグスワップ E． c o l i 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである。

【0066】

一部の実施形態では、後続の分解タグスワップ E． c o l i 株ライブラリーは、先行する分解タグスワップ E． c o l i 株ライブラリーの完全コンビナトリアルライブラリーである。

【0067】

一部の実施形態では、後続の分解タグスワップ E． c o l i 株ライブラリーは、先行する分解タグスワップ E． c o l i 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである。

【0068】

一部の実施形態では、後続の分解タグスワップ E． c o l i 株ライブラリーの E． c o l i 株の表現型性能が、産生 E． c o l i 株の表現型性能と比較して、測定された表現型変数における少なくとも 10 % の増加を呈するまで、ステップ d) ～ e) を繰り返す。

【0069】

一部の実施形態では、後続の分解タグスワップ E． c o l i 株ライブラリーの E． c o l i 株の表現型性能が、産生 E． c o l i 株の表現型性能と比較して、測定された表現型変数における少なくとも 1 倍の増加を呈するまで、ステップ d) ～ e) を繰り返す。

【0070】

一部の実施形態では、ステップ f) の表現型性能の改善は、目的の産物の体積生産性、目的の産物の比生産性、目的の産物の収率、目的の産物の力価、およびこれらの組合せか

らなる群から選択される。

【0071】

一部の実施形態では、ステップ f) の表現型性能の改善は、目的の産物の増大したまたはより効率的な産生であり、前記目的の産物は、低分子、酵素、ペプチド、アミノ酸、有機酸、合成化合物、燃料、アルコール、一次細胞外代謝産物、二次的細胞外代謝産物、細胞内コンポーネント分子、およびこれらの組合せからなる群から選択される。

【0072】

一部の実施形態では、本開示は、微生物宿主細胞における発現のための、異種遺伝子に作動可能に連結したキメラ合成プロモーターであって、長さが60～90ヌクレオチドであり、ラムダファージ p_R プロモーターの遠位部分、それぞれ長さが6ヌクレオチドであるラムダファージ p_L および p_R プロモーターの可変-35および-10領域、ラムダファージ p_L および p_R プロモーターのコア部分、ならびにラムダファージ p_R プロモーターの5' UTR/リボソーム結合部位(RBS)部分からなる、キメラ合成プロモーターを教示する。

10

【0073】

一部の実施形態では、ラムダファージ p_R プロモーターの遠位部分、ラムダファージ p_L および p_R プロモーターの可変-35および-10領域、ラムダファージ p_L および p_R プロモーターのコア部分、ならびにラムダファージ p_R プロモーターの5' UTR/リボソーム結合部位(RBS)部分の核酸配列は、表1. 5に見出される核酸配列から選択される。

20

【0074】

一部の実施形態では、本開示は、微生物宿主細胞における発現のための、異種遺伝子に作動可能に連結したキメラ合成プロモーターであって、長さが60～90ヌクレオチドであり、ラムダファージ p_R プロモーターの遠位部分、それぞれ長さが6ヌクレオチドであるラムダファージ p_L および p_R プロモーターの可変-35および-10領域、ラムダファージ p_L および p_R プロモーターのコア部分ならびに E. coli a c s 遺伝子のプロモーターの5' UTR/リボソーム結合部位(RBS)部分からなる、キメラ合成プロモーターを教示する。

【0075】

一部の実施形態では、ラムダファージ p_R プロモーターの遠位部分、ラムダファージ p_L および p_R プロモーターの可変-35および-10領域、ラムダファージ p_L および p_R プロモーターのコア部分、ならびに E. coli a c s 遺伝子のプロモーターの5' UTR/リボソーム結合部位(RBS)部分の核酸配列は、表1. 5に見出される核酸配列から選択される。

30

【0076】

一部の実施形態では、キメラ合成プロモーターは、配列番号132～152、159～160、162、165、174～175、188、190、199～201または207から選択される核酸配列からなる。

【0077】

一部の実施形態では、キメラ合成プロモーターは、配列番号153～158、161、163～164、166～173、176～187、189、191～198または202～206から選択される核酸配列からなる。

40

【0078】

一部の実施形態では、微生物宿主細胞は E. coli である。

【0079】

一部の実施形態では、異種遺伝子は、表2に見出される目的のタンパク質産物をコードする。

【0080】

一部の実施形態では、異種遺伝子は、リシン生合成経路の一部である遺伝子である。

【0081】

50

一部の実施形態では、異種遺伝子は、a s d 遺伝子、a s k 遺伝子、h o m 遺伝子、d a p A 遺伝子、d a p B 遺伝子、d a p D 遺伝子、d d h 遺伝子、a r g D 遺伝子、d a p E 遺伝子、d a p F 遺伝子、l y s A 遺伝子、l y s E 遺伝子、z w f 遺伝子、p g i 遺伝子、k t k 遺伝子、f b p 遺伝子、p p c 遺伝子、p c k 遺伝子、d d x 遺伝子、p y c 遺伝子または i c d 遺伝子から選択される。

【0082】

一部の実施形態では、異種遺伝子は、リコピン生合成経路の一部である遺伝子である。

【0083】

一部の実施形態では、異種遺伝子は、d x s 遺伝子、i s p C 遺伝子、i s p E 遺伝子、i s p D 遺伝子、i s p F 遺伝子、i s p G 遺伝子、i s p H 遺伝子、i d i 遺伝子、i s p A 遺伝子、i s p B 遺伝子、c r t E 遺伝子、c r t B 遺伝子、c r t I 遺伝子、c r t Y 遺伝子、y m g A 遺伝子、d x r 遺伝子、e l b A 遺伝子、g d h A 遺伝子、a p p Y 遺伝子、e l b B 遺伝子、または y m g B 遺伝子から選択される。

10

【0084】

一部の実施形態では、異種遺伝子は、生物学的製剤をコードするまたは生物学的製剤を生成するための経路内の遺伝子である。

【0085】

一部の実施形態では、生物学的製剤は、ヒューマリン（r h インスリン）、イントロン A（インターフェロンアルファ 2 b）、ロフェロン（インターフェロンアルファ 2 a）、ヒューマトロップ（ソマトロピン r h 成長ホルモン）、ニューボジェン（フィルグラスチム）、ベタフェロン（detaferon）（インターフェロンベクター 1 b）、リスプロ（速効型インスリン）、ラピリシン（レテプラザー）、インフェルゲン（infergen）（インターフェロンアルファコン-1）、グルカゴン、ベロムン（タソネルミン）、オンタック（on tak）（デニロイキンジフチトクス）、ランタス（長時間作用型インスリン グラルギン）、キネレット（アナキンラ）、ナトレコール（ネシリチド）、ソマバート（ペグビソマン ト）、カルシトニン（組換えカルシトニン サケ）、ルセンティス（ラニビズマブ）、ブレオタクト（ヒト副甲状腺ホルモン）、クリステクサ（kystexxal）（r h 尿酸オキシダーゼ、P E G 化）、ニベスチム（フィルグラスチム、r h G C S F）、ボラキサゼ（vora xaze）（グルカルピダーゼ）、またはプレオス（preos）（副甲状腺ホルモン）から選択される。

20

30

【0086】

一部の実施形態では、本開示は、配列番号 1 3 2 ~ 2 0 7 から選択される核酸配列を有するキメラ合成プロモーターに作動可能に連結した異種遺伝子を教示する。

【0087】

一部の実施形態では、異種遺伝子は、表 2 に見出される目的のタンパク質産物をコードする。

【0088】

一部の実施形態では、異種遺伝子は、リシン生合成経路の一部である遺伝子である。

【0089】

一部の実施形態では、異種遺伝子は、a s d 遺伝子、a s k 遺伝子、h o m 遺伝子、d a p A 遺伝子、d a p B 遺伝子、d a p D 遺伝子、d d h 遺伝子、a r g D 遺伝子、d a p E 遺伝子、d a p F 遺伝子、l y s A 遺伝子、l y s E 遺伝子、z w f 遺伝子、p g i 遺伝子、k t k 遺伝子、f b p 遺伝子、p p c 遺伝子、p c k 遺伝子、d d x 遺伝子、p y c 遺伝子または i c d 遺伝子から選択される。

40

【0090】

一部の実施形態では、異種遺伝子は、リコピン生合成経路の一部である遺伝子である。

【0091】

一部の実施形態では、異種遺伝子は、d x s 遺伝子、i s p C 遺伝子、i s p E 遺伝子、i s p D 遺伝子、i s p F 遺伝子、i s p G 遺伝子、i s p H 遺伝子、i d i 遺伝子、i s p A 遺伝子、i s p B 遺伝子、c r t E 遺伝子、c r t B 遺伝子、c r t I 遺伝子、

50

c r t Y 遺伝子、y m g A 遺伝子、d x r 遺伝子、e l b A 遺伝子、g d h A 遺伝子、a p p Y 遺伝子、e l b B 遺伝子、または y m g B 遺伝子から選択される。

【0092】

一部の実施形態では、異種遺伝子は、生物学的製剤をコードするまたは生物学的製剤を生成するための経路内の遺伝子である。

【0093】

一部の実施形態では、生物学的製剤は、ヒューマリン（r h インスリン）、イントロン A（インターフェロンアルファ 2 b）、ロフェロン（インターフェロンアルファ 2 a）、ヒューマトロップ（ソマトロピン r h 成長ホルモン）、ニューボジェン（フィルグラスチム）、ベタフェロン（インターフェロンベクター 1 b）、リスプロ（速効型インスリン）、ラピリシン（レテプラザー）、インフェルゲン（インターフェロンアルファコンー 1）、グルカゴン、ベロムン（タソネルミン）、オンタック（デニロイキンジフチトクス）、ランタス（長時間作用型インスリン グラルギン）、キネレット（アナキンラ）、ナトレコール（ネシリチド）、ソマバート（ペグビソマント）、カルシトニン（組換えカルシトニン サケ）、ルセンティス（ラニビズマブ）、プレオタクト（ヒト副甲状腺ホルモン）、クリステクサ（r h 尿酸オキシダーゼ、P E G 化）、ニベスチム（フィルグラスチム、r h G C S F）、ボラキサゼ（グルカルピダーゼ）、またはプレオス（副甲状腺ホルモン）から選択される。

【図面の簡単な説明】

【0094】

【図 1】図 1 は、多様性プール中のバリエーションを増加させるための本開示の D N A 組換えの方法を示す。物理的または酵素的／化学的な手段により、D N A 切片、例えば、関連のある種からのゲノム領域を切断することができる。切断した D N A 領域を融解し、再アニールさせ、その結果、オーバーラップする遺伝子領域が、ポリメラーゼ伸長反応を開始する。産物が 1 つまたは複数の出発配列に由来するエレメントを含むキメラ D N A に再アセンブルされるまで、後続の融解／伸長反応を実施する。

【0095】

【図 2】図 2 は、選択された配列の改変（例えば、スワップするための 100 個の S N P）を用いて、新しい宿主 E. c o l i 株を生成するための本開示の方法の概要を示す。手短に述べると、方法は以下を含む：（1）所望の D N A 挿入断片を、設計し、アセンブリ反応において 1 つまたは複数の合成オリゴを組み合わせることによって生成する、（2）D N A 挿入断片を、形質転換プラスミド中にクローニングする、（3）完成したプラスミドを、所望の産生株中に移入させ、そのプラスミドが宿主株のゲノム中に組み込まれる、および、（4）選択マーカーおよび他の望まれない D N A エレメントを、宿主株からループアウトする。D N A アセンブリの各ステップは、追加の、品質管理（Q C）のステップ、例えば、増幅およびシーケンシングのためにプラスミドを E. c o l i 細菌中にクローニングするステップも含むことができる。

【0096】

【図 3】図 3 は、本開示の形質転換プラスミドアセンブリ、および宿主 E. c o l i ゲノム中へのそのプラスミドの組込みを示す。挿入断片 D N A（insert DNA）を、アセンブリ反応において 1 つまたは複数の合成オリゴを組み合わせることによって生成する。所望の配列を含有する D N A 挿入断片に、ゲノムの標的領域に相同な D N A の領域を隣接させる。これらの相同領域は、ゲノムの組込みを促進し、一旦組み込まれると、続くステップでベクター骨格の D N A をループアウトするように設計されたダイレトリピート領域を形成する。アセンブルされたプラスミドは、挿入断片 D N A を含有し、任意選択で、1 つまたは複数の選択マーカーも含有する。

【0097】

【図 4】図 4 は、選択された、宿主 E. c o l i 株に由来する D N A の領域をループアウトするための手順を示す。挿入された D N A および宿主ゲノムのダイレトリピート領域は、組換え事象において「ループアウトする」ことができる。選択マーカーについて対抗

選択された細胞は、ダイレクトリポート領域が隣接するループDNAの欠失を含有する。

【0098】

【図5】図5は、本開示のE. coli株改良プロセスの実施形態を示す。遺伝子改変を含有する宿主株配列（遺伝子設計）を、多様な株バックグラウンドにおいて、株性能の改善（株構築）について試験する。有益な変異を示す株を分析し（ヒットIDおよび分析）、データを、さらに分析するために、ライブラリー（とりわけ、例えば、SNPスワップライブラリー、PROスワップライブラリーおよびそれらの組合せ）に保存する。追加の反復分析のために、本開示の選択規則により、1つまたは複数のライブラリーに由来するエレメントの組合せについて予測される効果に基づいて新たに提案されるE. coli宿主株配列が生成される。

10

【0099】

【図6A】図6A～Bは、本開示の実施形態のうちの1つのDNAアセンブリー、形質転換およびE. coli株スクリーニングのステップを示す。図6Aは、DNA断片を構築するためのステップ、前記DNA断片をベクター中にクローニングするためのステップ、前記ベクターを宿主E. coli株中に形質転換するためのステップ、および選択配列を対抗選択によりループアウトするためのステップを示す。図6Bは、ハイスループット培養するためのステップ、スクリーニングするためのステップ、および選択されたE. coli宿主株を評価するためのステップを示す。この図はまた、培養物タンク中の、選択されたE. coli株の培養、スクリーニングおよび評価を行う任意選択のステップも示す。

20

【図6B】図6A～Bは、本開示の実施形態のうちの1つのDNAアセンブリー、形質転換およびE. coli株スクリーニングのステップを示す。図6Aは、DNA断片を構築するためのステップ、前記DNA断片をベクター中にクローニングするためのステップ、前記ベクターを宿主E. coli株中に形質転換するためのステップ、および選択配列を対抗選択によりループアウトするためのステップを示す。図6Bは、ハイスループット培養するためのステップ、スクリーニングするためのステップ、および選択されたE. coli宿主株を評価するためのステップを示す。この図はまた、培養物タンク中の、選択されたE. coli株の培養、スクリーニングおよび評価を行う任意選択のステップも示す。

【0100】

【図7】図7は、本開示の自動化されたシステムの一実施形態を示す。本開示は、宿主E. coliのクローニング、形質転換、培養、スクリーニングおよび／またはシーケンシングが可能な多様なモジュールを有する自動化されたロボットシステムの使用を教示する。

30

【0101】

【図8】図8は、本開示のE. coli株改良プログラムの実施形態の概要を示す。

【0102】

【図9】図9は、およそ 3.2×10^6 個の塩基対を含むCorynebacterium glutamicumのゲノムを描写する。

【0103】

【図10】図10は、本開示の形質転換の実験結果を示す。0.5kb～5.0kbの範囲のDNA挿入断片を、Corynebacterium glutamicumのゲノムの（相対的な位置1～24として示す）多様な領域への挿入に対して標的化した。明るい色により、組込みの成功を示し、一方、より暗い色により、挿入の失敗を示す。

40

【0104】

【図11】図11は、第2のラウンドのHTP操作PROスワッププログラムの結果を示す。PROスワップの第1のラウンドの間に同定したプロモーター：：遺伝子の上位の組合せを、本開示の方法に従って分析して、宿主の性能に対して相加的またはコンビナトリアルに有益な効果を示す可能性があろう、前記変異の組合せを同定した。したがって、第2のラウンドのPROスワップ変異体は、多様なプロモーター：：遺伝子の変異の対の組

50

合せを含んだ。結果として得られた第2のラウンドの変異体を、選択された生体分子の宿主細胞の収率の差についてスクリーニングした。有益な効果を示すことが予測されていた、変異の組合せの対が、円を用いて強調されている。

【0105】

【図12】図12は、*E. coli*中に形質転換したプラスミドに関するプラスミドアセンブリの成功について試験した実験結果を示す。1 kbおよび2 kbの挿入配列を含有するプラスミドについては、4つのコロニーを選べば、13%の失敗率を実践するのに十分である。より大きな挿入は、一貫性のある結果をなしとげるのに、追加のコロニースクリーニングを必要とする場合がある。

【0106】

【図13】図13は、*Corynebacterium glutamicum*の挿入ベクターによる形質転換の成功について試験した実験結果を示す。2 kbおよび5 kbのDNA挿入断片のサイズは、高い形質転換率を示し、アセンブリの失敗率は低かった。

【0107】

【図14】図14は、*Corynebacterium glutamicum*におけるループアウト選択の結果を示す。形質転換された細菌のスクロース耐性により、sacB選択マーカーがループアウトされたことが示される。DNA挿入断片のサイズが、ループアウト効率に影響を及ぼすようには見えない。

【0108】

【図15】図15は、相関性の尺度を使用して計算した類似性マトリックスである。マトリックスにより、SNPバリエーション間の機能類似性が描写される。低い機能類似性を有するSNPの統合は、より高い機能類似性を有するSNPの統合とは対照的に、株性能を改善する可能性がより高いことが期待される。

【0109】

【図16】図16A～Bは、上位性マッピングの実験結果を示す。低い機能類似性を有するSNPスワップとPROスワップとの組合せにより、改良された株性能が得られる。図16Aは、すべてのSNP/PROスワップの機能類似性によりクラスター化された樹状図を示す。図16Bは、産物の収率により測定した場合の、宿主株の、統合したSNPの性能を示す。より大きなクラスターの距離が、宿主株の統合性能の改善と相関する。

【0110】

【図17】図17A～Bは、多様性プール中の株バリエーション間のSNPの差を示す。図17Aは、この実験の株の間の関係を示す。株Aは、野生型の宿主株である。株Bは、中間の、操作された株である。株Cは、工業生産用の株である。図17Bは、各株において、ユニークなSNPおよび共有されているSNPの数を同定するグラフである。

【0111】

【図18】図18は、本開示の方法に従う第1のラウンドのSNPスワッピングの実験を示す。(1) CからのSNPをすべて、基準株A中に個々および/またはコンビナトリアルにクローニングする(AからCへの「ウェーブアップ」)。(2) CからのSNPをすべて、市販株Cから個々および/またはコンビナトリアルに除去する(CからAへの「ウェーブダウン」)。(3) BからのSNPをすべて、基準株A中に個々および/またはコンビナトリアルにクローニングする(AからBへのウェーブアップ)。(4) BからのSNPをすべて、市販株Bから個々および/またはコンビナトリアルに除去する(BからAへのウェーブダウン)。(5) CにユニークなSNPをすべて、個々および/またはコンビナトリアルに市販株B中にクローニングする(BからCへのウェーブアップ)。(6) CにユニークなSNPをすべて、市販株Cから個々および/またはコンビナトリアルに除去する(CからBへのウェーブダウン)。

【0112】

【図19】図19は、プロモータースワップのプロセスにおいて利用される遺伝子標的の例を示す。4つの下線は、下方制御の標的にすることができる転用遺伝子(diverting gene)であり、経路上の残りの19種の遺伝子は過剰発現の標的とすること

10

20

30

40

50

ができる。

【0113】

【図20】図20は、同定した遺伝子標的についてプロモータースワップのプロセスを実施するために利用されている例示的なプロモーターライブラリーを示す。PROスワップ（すなわち、プロモータースワップ）のプロセスにおいて利用されるプロモーターは、P₁～P₈であり、それらの配列および同一性は、表1に見出すことができる。

【0114】

【図21】図21は、標的化遺伝子がそれ自体のプロモーターを含むか、またはオペロンの一部であるかに応じたプロモータースワッピングのための異なる利用可能な手法を示す。

10

【0115】

【図22】図22は、リシンの収率に対する性能に有意に影響を及ぼす改変を示す、例示的なHTPプロモータースワッピングのデータを示す。X軸は、プロモータースワップ遺伝子設計微生物株ライブラリー内の異なる株を表し、Y軸は、各株についてのリシンの収率の相対的な値を含む。グラフ上の各文字が、PROスワップの標的遺伝子を表す。各データ点が、反復を表す。データは、本明細書に記載の、HTPの適用例に適応させた分子ツール（すなわち、PROスワップ）は、目的の化合物または分子を産生するために、効率的に、微生物株を創製し、その株性能を最適化することが可能であることを実証している。この場合、目的の化合物は、リシンであった；しかし、教示するPROスワップの分子ツールを利用して、任意の目的の化合物の産生を最適化し、かつ／または増加させることができる。当業者であれば、いかにして所望の化合物の産生をコードする標的遺伝子を選び、次いで、教示するPROスワップの手順を利用するかを理解するはずである。本明細書で教示するリシンの収率の増加を例示するのに示したデータと、本出願中に提示する詳細な開示とにより、PROスワップの分子ツールが、HTPゲノム操作の広く適用可能な進歩となり得ることを当業者は容易に認識するはずである。

20

【0116】

【図23】図23は、検討中のインプットデータについての相対的な株性能の分布を示す。ゼロの相対的な性能は、操作された株がプレート中の基準株に等しく良好に機能したことを示す。本明細書に記載するプロセスは、ゼロを有意に上回って機能する可能性がある株を同定するように設計される。

30

【0117】

【図24】図24は、相対的な株性能の平均変化（増加または減少）を示す線形回帰係数値を示し、該株性能は、表示する株中に組み入れている各遺伝子変化と関連がある。

【0118】

【図25】図25は、予測される株の設計の上位100個についての変化の組成（composition of change）を示す。x軸に、潜在的な遺伝子変化（dss変異はSNPスワップであり、pcg変異はPROスワップである）のプールを列挙し、y軸に、順位を示す。黒色のセルにより、候補の設計における特定の変化の存在を示し、一方、白色のセルにより、そうした変化の不在を示す。この特定の例では、上位100個の設計はすべて、pcg3121__pgi、pcg1860__pyc、dss__339およびpcg0007__39__lysaの変化を含有する。加えて、上位の候補の設計は、dss__034、dss__009の変化を含有する。

40

【0119】

【図26】図26は、本開示の実施形態のうちの1つの、DNAアセンブリーおよび形質転換のステップを示す。フローチャートは、DNA断片を構築するためのステップ、前記DNA断片をベクター中にクロニングするためのステップ、前記ベクターを宿主E. coli株中に形質転換するためのステップ、および選択配列を対抗選択によりループアウトするためのステップを示す。

【0120】

【図27】図27は、ハイスループット培養するためのステップ、スクリーニングするた

50

めのステップ、および選択された宿主 *E. coli* 株を評価するためのステップを示す。この図はまた、培養物タンク中の、選択された *E. coli* 株の培養、スクリーニングおよび評価を行う任意選択のステップも示す。

【0121】

【図28】図28は、本開示のプロモーターラダーに従って一連の調節性の発現を示す、例示的なプロモーターの発現プロファイルを示す。プロモーターAの発現は、細菌培養の誘導期にピークに達し、一方、プロモーターBおよびCはそれぞれ、対数期および定常期にピークに達する。

【0122】

【図29】図29は、本開示のプロモーターラダーに従って一連の調節性の発現を示す、例示的なプロモーターの発現プロファイルを示す。プロモーターAの発現は、選択された基質を添加すると、即時にピークに達するが、基質の濃度が低下するにつれて、検出不能なレベルに急速に戻る。プロモーターBの発現は、選択された基質を添加すると、即時にピークに達し、基質の対応する低下と一緒に、ゆっくり低下して、検出不能なレベルに戻る。プロモーターCの発現は、選択された基質を添加すると、ピークに達し、培養全体を通して、基質が消散してしまった後でさえ、高度な発現状態を維持する。

【0123】

【図30】図30は、本開示のプロモーターラダーに従って一連の構成性の発現レベルを示す、例示的なプロモーターの発現プロファイルを示す。プロモーターAは、最も低い発現を示し、それに続いて、プロモーターBおよびCがそれぞれ、発現レベルを増加させる。

【0124】

【図31】図31は、*E. coli* 株改良のための本開示のLIMSシステムの実施形態を図示する。

【0125】

【図32】図32は、本開示のLIMSシステムの実施形態のクラウドコンピューティングの実装を図示する。

【0126】

【図33】図33は、本開示の反復予測株設計ワークフローの実施形態を示す。

【0127】

【図34】図34は、本開示の実施形態に従う、コンピューターシステムの実施形態を図示する。

【0128】

【図35】図35は、本開示の一実施形態に従うDNAアセンブリーに関連するワークフローを示す。このプロセスは、4つの段階：重要部分(part)の生成、プラスミドアセンブリー、プラスミドのQC、および形質転換のためのプラスミドの調製に分けられる。重要部分の生成の間に、オリゴのシーケンシングのベンダーから、Laboratory Information Management System (LIMS)により設計されたオリゴを発注し、宿主生物から標的配列をPCRにより増幅するために使用する。これらのPCRの重要部分は、清浄して、混入物を除去し、断片の分析により、すなわち、観察される断片サイズと理論的な断片サイズとのin silicoでの品質管理比較、およびDNAの定量を行って、成功について評価する。重要部分を、酵母中にアセンブリーベクターと一緒に形質転換し、プラスミド中に相同組換えによりアセンブルする。アセンブルされたプラスミドを、酵母から単離し、*E. coli* 中に形質転換して、続いて、アセンブリーの品質管理および増幅を行う。プラスミドアセンブリーの品質管理の間に、各プラスミドのいくつかの複製物を単離し、ローリングサークル増幅(RCA)を使用して増幅し、正確なアセンブリーについて酵素消化および断片の分析により評価する。QCプロセス中に同定した正確にアセンブルされたプラスミドをヒットピックアップし、恒久的なストックを生成し、プラスミドDNAを抽出し、定量してから、標的宿主生物中に形質転換する。

10

20

30

40

50

【0129】

【図36】図36は、ターミネーターT1～T8の効果を、2つの培地中で2つの時点にわたり特徴付けた実験結果を示す。条件Aおよび条件Cは、BHI培地についての2つの時点を表し、一方、B点およびD点は、HTP試験培地についての2つの時点を表す。

【0130】

【図37】図37は、UV変異誘発などの従来の株改良アプローチの有効性と、本開示のHTP操作の方法論とを比較した実験結果を示す。圧倒的多数のUV変異は、宿主細胞の性能の目立った増加をもたらさなかった。対照的に、本開示のPROスワップの方法論は、宿主細胞の性能の1.2～2倍の増加を示す変異体を高い比率で産生した。

【0131】

【図38】図38は、第1のラウンドのHTP操作SNPスワッププログラムの結果を示す。186個の個別のSNP変異を同定し、基準株上に個々にクローニングした。結果として得られた変異体を、選択された生体分子の宿主細胞の収率の差についてスクリーニングした。

【0132】

【図39】図39は、第2のラウンドのHTP操作SNPスワッププログラムの結果を示す。第1のラウンドのSNPスワッププログラムからの176個の個別のSNP変異を、第1のラウンドのSNPプログラムの間に同定した有益なSNPを含有する第2のラウンドの宿主細胞株中に個々にクローニングした。したがって、結果として得られた変異体は、2つの変異を組み合わせた対の効果を表す。選択された生体分子に関して、宿主細胞の収率（Y軸）および生産性（X軸）における差についてのスクリーニングの結果を示す。

【0133】

【図40】図40は、タンク発酵のバリデーション実験の結果を示す。HTP SNPスワップの第2のラウンドからの変異の上位の対を、発酵タンク中で培養した。選択された生体分子（すなわち、リシン）に関して、宿主細胞の収率および生産性についての結果を示す。認め得るように、本発明者らは、1つのラウンドのゲノム操作において、PROスワップ手順を利用して、特定のPROスワップ変異体（zwf）が、基準株と比較して（すなわち、基準株と、基準株+zwfとを比較して）、選択された生体分子の収率の増加を示すことを決定した。次いで、本発明者らは、別のラウンドのゲノム操作を実施し、この場合、SNPスワップ手順を使用して、前記PROスワップ変異体と組み合わせた場合に、生体分子の収率に影響を及ぼすことができる有益なSNP変異を決定した。PROスワップ手順とSNPスワップ手順との組合せにより、これまでのPROスワップのみの変異体よりも（すなわち、基準株+zwf+SNP121と、これまでに論じた基準株+zwfとを比較して）、さらに高い収率を有する変異体が創製された。この図は、本開示のPROスワップ手順とSNPスワップ手順とを組み合わせることによってなしとげることができる収率の劇的な改善を示す。PROスワップゲノム操作キャンペーンとSNPスワップゲノム操作キャンペーンとを組み合わせる態様では、目的の生体分子／産物の収率および／または生産性を、基準株と比べて、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、15%、20%、25%、30%、40%、45%、50%またはそれ超だけ増加させることができる。

【0134】

【図41】図41は、第1のラウンドのHTP操作PROスワッププログラムの結果を示す。本開示の方法に従って、宿主の性能と関連があると考えられる、選択された遺伝子を、プロモーターラダーと組み合わせて、第1のラウンドのPROスワップライブラリーを創製した。結果として得られた変異体を、選択された生体分子（すなわち、リシン）の宿主細胞の収率の差についてスクリーニングした。

【0135】

【図42】図42は、本開示の実施形態に従って、微生物株を設計するための変異の選択における上位性効果の考慮事項を示すフローチャートである。

【0136】

【図 4 3】図 4 3 は、本開示に従うバイシストロン設計 (BCD) 調節配列を示す。一部の実施形態では、本開示は、PRO スワップにおける異なるプロモーター：：標的遺伝子組合せ間の発現一貫性を改善するために、従来のプロモーターの代わりに BCD を使用できることを教示する。一部の実施形態では、BCD は、プロモーター、第 1 のリボソーム結合部位 (SD1)、第 1 のシストロン配列 (Cis1)、第 2 のリボソーム結合部位 (SD2) を含み、目的の標的遺伝子 (Cis2) に作動可能に連結している。一部の実施形態では、本開示は、Cis1 が任意のペプチドコード配列であり得ることを教示する。BCD の設計および使用に関する追加的な情報は、本明細書の後ろのセクションに提供されている。

【0137】

10

【図 4 4】図 4 4 は、組換え DNA 結合ドメインを介した経路酵素共局在化の図解である。操作された細胞は、DNA 結合ドメインを有する経路酵素 Enz1~3 をコードする。発現されると、これらの酵素は、経路酵素と融合した組換え DNA 結合ドメインによって認識される DNA モチーフを含む足場 DNA または他の標的場所に結合する。融合した DNA 結合ドメインが足場プラスミド上のそれらの同類の DNA モチーフに付着すると、酵素は、空間中に互いと近くに制約され、これにより、経路の生産性が改善され得る。

【0138】

【図 4 5】図 4 5 は、DNA 結合ドメインをコードするヌクレオチド配列の経路酵素への組み入れについての概略図である。GOI は、経路酵素をコードする。E. coli 細胞を、DNA 結合ドメインをコードするヌクレオチド配列 (星印で示されている) を含む GOI の変異体バージョンをコードするプラスミドで形質転換する。該プラスミドは、ルーブイン細胞の選択を可能にする抗生物質耐性マーカー (Ab)、および後続のルーブアウト細胞の対抗選択を可能にする対抗選択マーカー (対抗) もコードする。「ルーブイン」ステップでは、変異体 GOI を含めたプラスミド全体を相同組換え (HR) によってゲノム中に組み入れる。「ルーブアウト」ステップの間に、細胞の一部は HR によって生来の GOI 配列に戻るが、その他の細胞は変異体 GOI をゲノム中に残す HR 事象を受ける。

20

【0139】

【図 4 6】図 4 6 は、本開示の収率モデルに関して、訓練データの測定された性能と比した予測される性能のドットプロットである。基礎をなすモデルは、(4 次多項式カーネルを用いる) カーネルリッジ回帰モデルである。モデルを、1864 個のユニークな遺伝子構築物および関連する表現型の性能に対して訓練する。適合させたモデルは、0.52 の r^2 値を有する。

30

【0140】

【図 4 7】図 4 7 は、本開示の予測アルゴリズムにより生成された候補の設計の遺伝子構成を示す。これらの候補の設計を、HTP の構築および分析に提出した。ここでは、候補の設計は、親株 id と導入された変異 (複数可) との組合せとして定義される。

【0141】

【図 4 8】図 4 8 は、本開示の予測アルゴリズムにより生成され、本開示の HTP 構築法に従って構築した候補の設計の測定された性能と比した予測される性能のドットプロットである。この図は、モデルが、候補株の性能を許容できる精度内で予測することができることを示す。

40

【0142】

【図 4 9】図 4 9 は、親株に対する候補株の収率のパーセント変化を示す箱ひげ図である。y 軸上では、0.01 の値が、1% に相当する。この図は、コンピューターモデルにより設計された株 (薄い灰色) が、それらに対応する親株を上回る測定可能な改善をなしとげること示す。加えて、図は、これらのモデル基準株の改善の大きさが、ヒト専門家が設計した株によりなしとげられた改善に匹敵することも示す。

【0143】

【図 5 0】図 5 0 は、コンピューターモデルにより設計された株 (暗灰色) およびヒト専門家により設計された株 (薄い灰色) についての収率による性能の分布を示す。コンピュ

50

ーターにより設計された株は、より密な分布を示し、獲得中央値がより高かった。

【0144】

【図51】図51は、コンピューターにより生成された候補株（薄い灰色）またはヒト専門家により生成された候補株（暗灰色）の絶対的な収率を示す箱ひげ図である。結果は親株から集計する。

【0145】

【図52】図52は、およそ460万個の塩基対を含む *Escherichia coli* のゲノムの描写である。

【0146】

【図53】図53は、ベクター骨格中のインスレーターおよびターミネーター部分の形質転換およびプラスミド組込みの効率に対する影響を示す。

10

【0147】

【図54】図54は、表1.4からの合成プロモーター-5'UTR配列のコンビナトリアル設計を示す。

【0148】

【図55】図55は、ベクター1骨格のコンポーネントを示すプラスミドマップを示す。

【0149】

【図56】図56は、ベクター2骨格のコンポーネントを示すプラスミドマップを示す。

【0150】

【図57】図57は、ベクター3骨格のコンポーネントを示すプラスミドマップを示す。

20

【0151】

【図58】図58は、ベクター4骨格のコンポーネントを示すプラスミドマップを示す。

【0152】

【図59】図59は、*E. coli* リコピン生合成経路を示す。

【0153】

【図60】図60は、リコピン経路標的 *idi* および *ymgA* におけるターミネーター編集を示す。ターミネーター *TyjbE* により、対照と比べて株性能の低下が実証され、したがって、極めて重要な経路標的の同定に対するこれらのライブラリータイプの有用性が強調される。

【0154】

30

【図61】図61は、多数のリコピン経路標的におけるターミネーター編集を示す。

【0155】

【図62】図62は、リコピン経路標的 *dxs* におけるプロモータースワップ（比較用）、分解タグスワップ、およびターミネータースワップを示す。*ssrA*__*LAA* 分解タグにより、対照と比べて株性能の改善が実証される。この株は単一の経路標的における *PROSWP* と分解タグの組合せであるので、これは予想外である。最初の *PROSWP* によりタンパク質存在量が増大することが予想され、分解タグによりタンパク質存在量が減少することが予想され、したがって、最適な株性能の調整に関するライブラリータイプの組合せの有用性が実証される。

【0156】

40

【図63】図63は、リコピン経路標的 *gdhA* における可溶性タグスワップ、プロモータースワップ、およびターミネータースワップを示す。可溶性タグ *FH8* により、対照と比べて株性能の改善が実証されるが、*GB1* 可溶性タグではそれが実証されず、したがって、各改変タイプのライブラリーを評価する必要性が実証される。

【発明を実施するための形態】

【0157】

詳細な説明

定義

以下の用語は、当業者であればよく理解していると考えられるが、以下の定義を、本開示の主題の説明を容易にするために示す。

50

【0158】

用語「a」または「an」は、そのエンティティの1つまたは複数进行指し、すなわち、複数形の指示対象を指すことができる。したがって、用語「a」または「an」、「1つまたは複数」および「少なくとも1つ」は、本明細書で互換的に使用される。さらに、不定冠詞「a」または「an」による「要素」への言及は、文脈により唯一の要素が存在することが明確に要求されない限り、1つを超える要素が存在する可能性を除外しない。

【0159】

本明細書で使用される場合、用語「細胞生物」、「微生物 (microorganism)」、または「微生物 (microbe)」は、広く解釈されるべきである。これらの用語は、互換的に使用され、これらに限定されないが、2つの原核生物ドメイン、細菌および古細菌、ならびにある特定の真核真菌および原生生物を含む。一部の実施形態では、本開示は、本開示内に存在するリスト／表および図面の「微生物 (microorganism)」または「細胞生物」または「微生物 (microbe)」に言及する。この特徴付けは、表および図面の同定された分類学的な属だけでなく、同定された分類学的な種、ならびに前記表または図面内の任意の生物体の様々な新規のおよび新しく同定または設計された株も参照することができる。同じ特徴付けが、実施例などの本明細書の他の部分におけるこれらの用語の記述にも当てはまる。

【0160】

用語「原核生物」は、当技術分野で認識されており、核または他の細胞小器官を含有しない細胞を指す。原核生物は一般に、2つのドメイン、細菌および古細菌の1つに分類される。古細菌と細菌ドメインの生物体間の決定的な差異は、16SリボソームRNA内のヌクレオチド塩基配列の基本的な差異に基づく。

【0161】

用語「古細菌」は、Mendosicutes門の生物体のカテゴリー分類を指し、典型的には、珍しい環境内で見つかり、リボソームタンパク質の数および細胞壁におけるムラミン酸の欠如を含めたいくつかの判定基準によって原核生物の残りとは区別される。ssrRNA分析に基づいて、古細菌は、2つの系統学的に別個の群：CrenarchaeotaおよびEuryarchaeotaからなる。これらの生理機能に基づいて、古細菌は、3つのタイプ：メタン生成菌（メタンを産生する原核生物）；高度好塩菌（非常に高濃度の塩（NaCl）で生きる原核生物）；および高度（超）好熱菌（thermophilus）（非常に高温で生きる原核生物）へと整理することができる。古細菌を細菌と区別する統一的古細菌フィーチャ（すなわち、細胞壁内にムレインが無いこと、エステル連結膜脂質など）に加えて、これらの原核生物は、これらをその特定の生息環境に適応させるユニークな構造的または生化学的属性を呈する。Crenarchaeotaは、超好熱性硫黄依存性原核生物から主になり、Euryarchaeotaは、メタン生成菌および高度好塩菌を含有する。

【0162】

「細菌」または「真正細菌」は、原核生物体のドメインを指す。細菌は、以下の少なくとも11の別個の群を含む：（1）2つの主要な亜門：（1）高G+C群（Actinomyces、Mycobacteria、Micrococcus、その他）（2）低G+C群（Bacillus、Clostridia、Lactobacillus、Staphylococci、Streptococci、Mycoplasmas）が存在するグラム陽性（グラム+）細菌；（2）Proteobacteria、例えば、紅色光合成+非光合成グラム陰性細菌（最も「一般的な」グラム陰性細菌を含む）；（3）ラン藻類、例えば、酸素発生型光合成生物；（4）スピロヘータおよび関連種；（5）Planctomyces；（6）Bacteroides、Flavobacteri a；（7）Chlamydia；（8）緑色硫黄細菌；（9）緑色非硫黄細菌（嫌気性光合成生物も）；（10）放射線抵抗性マイクロコッカスおよび類縁体；（11）ThermotogaおよびThermosiphon thermophiles。

【0163】

「真核生物」は、細胞が、膜内に囲まれた核および他の小器官を含有する任意の生物体である。真核生物は、分類群 Eukarya または Eukaryota に属する。真核細胞を原核細胞（上述の細菌および古細菌）から分離する決定的なフィーチャは、これらが膜結合型小器官、特に、遺伝物質を含有し、核包膜によって囲まれている核を有することである。

【0164】

用語「遺伝的に改変された宿主細胞」、「組換え宿主細胞」、および「組換え株」は、本明細書で互換的に使用され、本開示のクローニングおよび形質転換方法によって遺伝的に改変された宿主細胞を指す。よって、この用語は、宿主細胞（例えば、細菌、酵母細胞、真菌細胞、CHO、ヒト細胞など）であって、それが由来した天然に存在する生物体と比較して、それが変更された、改変された、または異なる遺伝子型および／または表現型（例えば、遺伝子改変が微生物のコーディング核酸配列に影響するとき）を呈するように遺伝的に変更、改変、または操作された、宿主細胞を含む。一部の実施形態では、この用語は、問題の特定の組換え宿主細胞だけでなく、このような宿主細胞の子孫または潜在的な子孫も指すことが理解される。

10

【0165】

用語「野生型微生物」または「野生型宿主細胞」は、天然に存在する細胞、すなわち、遺伝的に改変されていない細胞を記述する。

【0166】

用語「遺伝子操作された」は、宿主細胞のゲノムの任意のマニピュレーション（例えば、核酸の挿入、欠失、変異、または置き換えによる）を指すことができる。

20

【0167】

用語「対照」または「対照宿主細胞」は、遺伝子改変または実験処置の効果を決定するための適切な比較宿主細胞 (comparator host cell) を指す。一部の実施形態では、対照宿主細胞は、野生型細胞である。他の実施形態では、対照宿主細胞は、処置宿主細胞を差別化する遺伝子改変（複数可）を除いて、遺伝的に改変された宿主細胞と遺伝的に同一である。一部の実施形態では、本開示は、対照宿主細胞としての親株（例えば、株改良プログラムの基礎として使用された S₁ 株）の使用を教示する。他の実施形態では、宿主細胞は、処置宿主細胞において試験されている特異的プロモーターまたは SNP を欠く遺伝的に同一の細胞であり得る。

30

【0168】

本明細書で使用される場合、用語「対立遺伝子（複数可）」は、遺伝子の 1 つまたは複数の代替形態のいずれかを意味し、これらの対立遺伝子すべては、少なくとも 1 つの形質または特性と関係する。二倍体細胞では、所与の遺伝子の 2 つの対立遺伝子は、一对の相同染色体上の対応する遺伝子座を占有する。

【0169】

本明細書で使用される場合、用語「遺伝子座 (locus)」(遺伝子座 (loci) 複数形) は、例えば、遺伝子または遺伝子マーカーが見出される染色体上の具体的な 1 つもしくは複数の場所または部位を意味する。

【0170】

本明細書で使用される場合、用語「遺伝的に連結された」は、2 種またはそれ超の形質であって、これらが交雑によって分離するのが困難であるように、繁殖中に高い割合で共遺伝される、形質を指す。

40

【0171】

本明細書で使用される「組換え」または「組換え事象」は、染色体交差または独立組合せを指す。

【0172】

本明細書で使用される場合、用語「表現型」は、個々の遺伝子構造（すなわち、遺伝子型）と環境との間の相互作用から生じる個々の細胞、細胞培養物、生物体、または生物体の群の観察可能な特性を指す。

50

【0173】

本明細書で使用される場合、用語「キメラの」または「組換えの」は、核酸配列またはタンパク質配列を記述するとき、少なくとも2つの異種ポリヌクレオチドもしくは2つの異種ポリペプチドを連結して単一の巨大分子にする、または少なくとも1つの天然の核酸もしくはタンパク質配列の1つもしくは複数のエレメントを再編成する核酸またはタンパク質配列を指す。例えば、用語「組換えの」は、例えば、化学合成による、または遺伝子操作技法による核酸の単離セグメントのマニピュレーションによる、配列の2つの別の方法で分離されたセグメントの人工の組合せを指すことができる。

【0174】

本明細書で使用される場合、「合成ヌクレオチド配列」または「合成ポリヌクレオチド配列」は、天然に存在することが知られていない、または天然に存在しないヌクレオチド配列である。一般に、このような合成ヌクレオチド配列は、任意の他の天然に存在するヌクレオチド配列と比較したとき、少なくとも1つのヌクレオチド差異を含む。

【0175】

本明細書で使用される場合、用語「核酸」は、リボヌクレオチドもしくはデオキシリボヌクレオチドのいずれか、またはこれらの類似体の、任意の長さのヌクレオチドのポリマー形態を指す。この用語は、分子の一次構造を指し、よって、二本鎖および一本鎖DNA、ならびに二本鎖および一本鎖RNAを含む。これは、修飾核酸、例えば、メチル化および／またはキャッピングされた核酸、修飾塩基、主鎖修飾を含有する核酸なども含む。用語「核酸」および「ヌクレオチド配列」は、互換的に使用される。

【0176】

本明細書で使用される場合、用語「DNA足場」または「核酸足場」は、足場として別の目的のために再利用される人工的に生成されたまたは天然に存在する配列のいずれかである核酸足場を指す。本発明の一実施形態では、核酸足場は、合成デオキシリボ核酸足場である。合成足場のデオキシリボヌクレオチドは、プリンおよびピリミジン塩基、または他の天然、化学的もしくは生化学的に修飾された、非天然、もしくは誘導体化されたデオキシリボヌクレオチド塩基を含み得る。本明細書においてより詳細に記載する通り、本発明の核酸足場は、生物学的経路に関与する2つまたはそれよりも多くのタンパク質、すなわち、生合成酵素を空間的および時間的にアセンブルし、固定化して、機能的な複合体を創製するために利用される。各生物学的経路タンパク質の足場上へのアセンブリーおよび固定化は、足場のタンパク質結合配列、すなわちタンパク質ドッキング部位の1つと、キメラ生合成酵素の対応するDNA結合部分の結合相互作用によって起こる。よって、核酸足場は、1つまたは複数のサブユニットを含み、各サブユニットが、2つまたはそれよりも多くの異なるキメラ生物学的経路タンパク質の結合に適応させるための2つまたはそれよりも多くのタンパク質結合配列を含む。

【0177】

本明細書で使用される場合、「DNA結合配列」または「DNA結合部位」は、本開示の改変された遺伝子によってコードされるキメラ生合成遺伝子（例えば、キメラ生合成酵素）のDNA結合ドメイン部分によって認識され、それが結合する特異的な核酸配列を指す。多くのDNA結合ドメインおよびそれらの同類の結合パートナーDNA認識部位（すなわち、DNA結合部位）が当技術分野で周知である。例えば、多数のジンクフィンガー結合ドメインおよびそれらの対応するDNA結合標的部位が当技術分野で公知であり、本発明における使用に適する。他のDNA結合ドメインとしては、限定することなく、ロイシンジッパー結合ドメインおよびそれらの対応するDNA結合部位、ウィングドヘリックスDNA結合ドメインおよびそれらの対応するDNA結合部位、ウィングドヘリックスターンヘリックスDNA結合ドメインおよびそれらの対応するDNA結合部位、HMGボックスDNA結合ドメインおよびそれらの対応するDNA結合配列、ヘリックス・ループ・ヘリックスDNA結合ドメインおよびそれらの対応するDNA結合配列、ならびにヘリックス・ターン・ヘリックスDNA結合ドメインおよびそれらの対応するDNA結合配列が挙げられる。公知のDNA結合配列を有する他の公知のDNA結合ドメインとしては、

免疫グロブリンDNAドメイン、B3 DNA結合ドメイン、およびTALエフェクターDNA結合ドメインが挙げられる。本発明の核酸足場サブユニットは、上述のDNA結合部位のうちの任意の2つまたはそれよりも多くを含み得る。

【0178】

本明細書で使用される場合、用語「遺伝子」は、生物学的機能と関連したDNAの任意のセグメントを指す。よって、遺伝子は、これらに限定されないが、コード配列および／またはこれらの発現に要求される調節配列を含む。遺伝子は、例えば、他のタンパク質の認識配列を形成する非発現DNAセグメントも含み得る。遺伝子は、目的の源からのクローニング、または公知のもしくは予測された配列情報からの合成を含めた、様々な源から得ることができ、所望のパラメータを有するように設計された配列を含み得る。

10

【0179】

本明細書で使用される場合、用語「相同の」または「相同体」または「オルソログ」(ortholog)または「オルソログ」(orthologue)は、当技術分野で公知であり、共通の祖先またはファミリーメンバーを共有し、配列同一性の程度に基づいて決定される関連配列を指す。用語「相同性」、「相同の」、「実質的に同様の」、「および「実質的に対応する」は、本明細書で互換的に使用される。これらは、核酸断片であって、1つまたは複数のヌクレオチド塩基の変化が、遺伝子発現を媒介し、またはある特定の表現型を生じさせる核酸断片の能力に影響しない、核酸断片を指す。これらの用語は、本開示の核酸断片の修飾、例えば、最初の無修飾断片と比べて得られる核酸断片の機能的性質を実質的に変更しない1つまたは複数のヌクレオチドの欠失または挿入も指す。したがって、当業者が十分に理解するように、本開示は、具体的な例示的な配列より多くを包含することが理解される。これらの用語は、1つの種、亜種、種類、栽培品種、または株内に見出される遺伝子と、別の種、亜種、種類、栽培品種、または株内の対応するまたは等価な遺伝子との間の関係を記述する。本開示の目的に関して、相同配列が比較される。「相同配列」または「相同体」または「オルソログ」は、機能的に関連していると見なされ、考えられ、または知られている。機能的関係は、これらに限定されないが、(a)配列同一性の程度および／または(b)同じもしくは同様の生物学的機能を含めた、いくつかのやり方のいずれか1つにおいて示され得る。好ましくは、(a)および(b)の両方が示される。相同性は、当技術分野で容易に利用可能なソフトウェアプログラム、Current Protocols in Molecular Biology (F.M. Ausubelら編、1987年)補遺30、セクション7.718、表7.71に論じられたものなどを使用して決定することができる。いくつかの整列プログラムは、MacVector (Oxford Molecular Ltd、Oxford、U.K.)、ALIGN Plus (Scientific and Educational Software、Pennsylvania)、およびAlignX (Vector NTI、Invitrogen、Carlsbad、CA)である。別の整列プログラムは、デフォルトパラメータを使用するSequencher (Gene Codes、Ann Arbor、Michigan)である。

20

30

【0180】

本明細書で使用される場合、用語「内在性の」または「内在性遺伝子」は、天然に存在する遺伝子であって、それが宿主細胞ゲノム内に天然に見出される場所における天然に存在する遺伝子を指す。本開示との関連で、異種プロモーターを内在性遺伝子に作動可能に連結することは、既存の遺伝子であって、その遺伝子が天然に存在する場所における、既存の遺伝子の前に異種プロモーター配列を遺伝的に挿入することを意味する。本明細書に記載の内在性遺伝子は、本開示の方法のいずれかによって変異された天然に存在する遺伝子の対立遺伝子を含み得る。

40

【0181】

本明細書で使用される場合、用語「外因性の」は、用語「異種の」と互換的に使用され、その生来の源(native source)以外の何らかの源に由来する物質を指す。例えば、用語「外因性タンパク質」または「外因性遺伝子」は、非生来の源または場所由来であり、生物系に人工的に供給されたタンパク質または遺伝子を指す。

50

【0182】

本明細書で使用される場合、用語「ヌクレオチド変化」は、当技術分野でよく理解されているように、例えば、ヌクレオチド置換、欠失、および／または挿入を指す。例えば、変異は、サイレント置換、付加、または欠失を生じさせるが、コードされるタンパク質の性質もしくはは活性、またはタンパク質が作製される方法を変更しない変更を含有する。

【0183】

本明細書で使用される場合、用語「タンパク質修飾」は、当技術分野でよく理解されているように、例えば、アミノ酸置換、アミノ酸修飾、欠失、および／または挿入を指す。

【0184】

本明細書で使用される場合、用語、核酸またはポリペプチドの「少なくとも一部」または「断片」は、このような配列の最小限のサイズ特性を有する部分、または最大で全長分子の、および全長分子を含めた、全長分子の任意のより大きい断片を意味する。本開示のポリヌクレオチドの断片は、遺伝子調節エレメントの生物活性部分をコードすることができる。遺伝子調節エレメントの生物活性部分は、遺伝子調節エレメントを含む本開示のポリヌクレオチドの1つの部分を単離し、本明細書に記載の活性を評価することによって調製することができる。同様に、ポリペプチドの部分は、4アミノ酸、5アミノ酸、6アミノ酸、7アミノ酸などであり得、最大で全長ポリペプチドまで伸び得る。使用される部分の長さは、特定の用途に依存する。ハイブリダイゼーションプローブとして有用な核酸の部分は、12ヌクレオチドという短さであり得；一部の実施形態では、それは、20ヌクレオチドである。エピトープとして有用なポリペプチドの部分は、4アミノ酸という短さであり得る。完全長ポリペプチドの機能を果たすポリペプチドの部分は一般に、4アミノ酸より長いはずである。

【0185】

バリエーションポリヌクレオチドは、DNAシャッフリングなどの変異誘発および組換え誘導 (recombinogenic) 手順に由来する配列も包含する。このようなDNAシャッフリングの戦略は、当技術分野で公知である。例えば、Stemmer (1994年)、PNAS、91巻：10747～10751頁；Stemmer (1994年)、Nature、370巻：389～391頁；Cramerら (1997年)、Nature Biotech.、15巻：436～438頁；Mooreら (1997年)、J. Mol. Biol.、272巻：336～347頁；Zhangら (1997年)、PNAS、94巻：4504～4509頁；Cramerら (1998年)、Nature、391巻：288～291頁；ならびに米国特許第5,605,793号および同第5,837,458号を参照。

【0186】

本明細書に開示のポリヌクレオチドのPCR増幅に関して、オリゴヌクレオチドプライマーを、目的の任意の生物体から抽出されたcDNAまたはゲノムDNAから対応するDNA配列を増幅するためのPCR反応で使用するのに設計することができる。PCRプライマーを設計しPCRクロニングするための方法は、当技術分野で一般に公知であり、Sambrookら (2001年)、Molecular Cloning: A Laboratory Manual (3版、Cold Spring Harbor Laboratory Press、Plainview、New York) に開示されている。Innisら編 (1990年)、PCR Protocols: A Guide to Methods and Applications (Academic Press、New York)；InnisおよびGelfand編 (1995年)、PCR Strategies (Academic Press、New York)；ならびにInnisおよびGelfand編 (1999年) PCR Methods Manual (Academic Press、New York) も参照。PCRの公知の方法としては、これらに限定されないが、対プライマー、ネステッドプライマー、単一特異的プライマー、縮重プライマー、遺伝子特異的プライマー、ベクター特異的プライマー、部分ミスマッチプライマーなどを使用する方法が挙げられる。

【0187】

本明細書で使用される用語「プライマー」は、増幅標的にアニールし、DNAポリメラーゼを結合させ、それによって、プライマー伸長産物の合成が誘導される条件下で、すなわち、ヌクレオチド、およびDNAポリメラーゼなどの重合のための作用物質の存在下で

、かつ適当な温度およびpHで配置されたとき、DNA合成の開始点として機能を果たすことができるオリゴヌクレオチドを指す。(増幅)プライマーは、好ましくは、増幅の効率を最大にするために一本鎖である。好ましくは、プライマーは、オリゴデオキシリボヌクレオチドである。プライマーは、重合のための作用物質の存在下で伸長産物の合成をプライムするのに十分長くなければならない。プライマーの正確な長さは、プライマーの温度および組成(A/T対G/C含有量)を含めた多くの因子に依存する。一对の双方向性プライマーは、PCR増幅などのDNA増幅の技術分野で一般に使用されるように、1つの順方向プライマーおよび1つの逆方向プライマーからなる。

【0188】

本明細書で使用される場合、「プロモーター」は、コード配列または機能RNAの発現を制御することができるDNA配列を指す。一部の実施形態では、プロモーター配列は、近位およびより遠位の上流エレメントからなり、後者のエレメントは、エンハンサーと呼ばれることが多い。したがって、「エンハンサー」は、プロモーター活性を刺激することができるDNA配列であり、プロモーターの固有のエレメントまたはプロモーターのレベルもしくは組織特異性を増強するために挿入される異種エレメントであり得る。プロモーターは、その全体が生来の遺伝子に由来する場合があります、または天然に見出される異なるプロモーターに由来する異なるエレメントから構成される場合があります、または合成DNAセグメントを含むことさえできる。異なるプロモーターは、異なる組織もしくは細胞型における、または発達の異なる段階における、または異なる環境条件に応じて遺伝子の発現を指示することができることが当業者によって理解されている。ほとんどの場合、調節配列の正確な境界は、完全に画定されていないので、あるバリエーションのDNA断片は、同一のプロモーター活性を有し得ることがさらに認識されている。

【0189】

本明細書で使用される場合、語句「組換えコンストラクト」、「発現コンストラクト」、「キメラコンストラクト」、「コンストラクト」、および「組換えDNAコンストラクト」は、本明細書で互換的に使用される。組換えコンストラクトは、核酸断片の人工の組合せ、例えば、天然と一緒に見出されない調節配列およびコード配列を含む。例えば、キメラコンストラクトは、異なる源に由来する調節配列およびコード配列、または同じ源に由来するが、天然に見出されるものと異なる様式で配列されている調節配列およびコード配列を含み得る。このようなコンストラクトは、それ自体で 사용할ことができ、またはベクターと併せて使用することができる。ベクターが使用される場合、ベクターの選択は、当業者に周知であるように、宿主細胞を形質転換するのに使用される方法に依存する。例えば、プラスミドベクターを使用することができる。当業者は、本開示の単離核酸断片のいずれかを含む宿主細胞を順調に形質転換、選択、および増殖させるためにベクター上に存在しなければならない遺伝子エレメントをよく知っている。当業者は、異なる独立した形質転換事象は、異なるレベルおよびパターンの発現をもたらす(Jonesら(1985年)、EMBO J.、4巻:2411~2418頁;De Almeidaら(1989年)、Mol. Gen. Genetics、218巻:78~86頁)、よって、所望の発現レベルおよびパターンをディスプレイする系(line)を得るために、複数の事象がスクリーニングされなければならないことも認識する。このようなスクリーニングは、とりわけ、DNAのサザン分析、mRNA発現のノーザン分析、タンパク質発現のイムノブロットング分析、または表現型分析によってなしとげることができる。ベクターは、自発的に複製する、または宿主細胞の染色体内に組み込むことができるプラスミド、ウイルス、バクテリオファージ、プロウイルス、ファージミド、トランスポゾン、人工染色体などであり得る。ベクターは、自律複製していない裸のRNAポリヌクレオチド、裸のDNAポリヌクレオチド、同じ鎖内のDNAおよびRNAの両方から構成されるポリヌクレオチド、ポリリシンコンジュゲートDNAまたはRNA、ペプチドコンジュゲートDNAまたはRNA、リボソームコンジュゲートDNAなどでもあり得る。本明細書で使用される場合、用語「発現」は、機能的最終産物、例えば、mRNAまたはタンパク質(前駆体または成熟)の産生を指す。

【0190】

「作動可能に連結した」は、本文脈において、本開示によるプロモーターポリヌクレオチドのさらなるオリゴヌクレオチドまたはポリヌクレオチドとの、前記さらなるポリヌクレオチドの転写をもたらす連続した配置を意味する。

【0191】

本明細書で使用される用語「目的の産物」または「生体分子」は、供給原料から微生物によって産生される任意の産物を指す。一部の場合では、目的の産物は、低分子、酵素、ペプチド、アミノ酸、有機酸、合成化合物、燃料、アルコールなどであり得る。例えば、目的の産物または生体分子は、任意の一次または二次細胞外代謝産物であり得る。一次代謝産物は、とりわけ、エタノール、クエン酸、乳酸、グルタミン酸、グルタメート、リシン、トレオニン、トリプトファン、および他のアミノ酸、ビタミン、多糖などであり得る。二次代謝産物は、とりわけ、ペニシリンのような抗生物質化合物、またはシクロスポリンAのような免疫抑制剤、ジベレリンのような植物ホルモン、ロバスタチンのようなスタチン薬、グリセオフルビンのような殺真菌薬などであり得る。目的の産物、または生体分子は、微生物によって産生される任意の細胞内コンポーネント、例えば、カタラーゼ、アミラーゼ、プロテアーゼ、ペクチナーゼ、グルコースイソメラーゼ、セルラーゼ、ヘミセルラーゼ、リパーゼ、ラクターゼ、ストレプトキナーゼ、および多くのその他を含めた微生物酵素でもあり得る。細胞内コンポーネントとして、組換えタンパク質、例えば、インスリン、B型肝炎ワクチン、インターフェロン、顆粒球コロニー刺激因子、ストレプトキナーゼなども挙げることができる。

【0192】

用語「炭素源」は一般に、細胞増殖のための炭素の源として使用されるのに適した物質を指す。炭素源としては、これらに限定されないが、バイオマス加水分解物、デンプン、スクロース、セルロース、ヘミセルロース、キシロース、およびリグニン、ならびにこれらの基質のモノマーコンポーネントが挙げられる。炭素源は、これらに限定されないが、ポリマー、炭水化物、酸、アルコール、アルデヒド、ケトン、アミノ酸、ペプチドなどを含めた様々な形態での様々な有機化合物を含み得る。これらとしては、例えば、様々な単糖、例えば、グルコース、デキストロース（D-グルコース）など、マルトース、オリゴ糖、多糖、飽和もしくは不飽和脂肪酸、スクシネート、ラクテート、アセテート、エタノールなど、またはこれらの混合物がある。光合成生物は、光合成の産物として炭素源をさらに産生することができる。一部の実施形態では、炭素源は、バイオマス加水分解物およびグルコースから選択され得る。

【0193】

用語「供給原料」は、他の産物が作製され得る微生物または発酵プロセスに供給される原料または原料の混合物として定義される。例えば、バイオマスまたはバイオマスに由来する炭素化合物などの炭素源は、発酵プロセスにおいて目的の産物（例えば、低分子、ペプチド、合成化合物、燃料、アルコールなど）を産生する微生物の供給原料である。しかし、供給原料は、炭素源以外の栄養分を含有し得る。

【0194】

用語「体積生産性」または「生産速度」は、単位時間当たり、媒体の体積当たりに形成される産物の量として定義される。体積生産性は、1時間当たり1リットル当たりのグラム（ g/L/h ）で報告することができる。

【0195】

用語「比生産性」は、産物の形成の速度として定義される。比生産性は、1時間当たり細胞乾燥重量（CDW）1グラム当たりの産物のグラム（ g/g CDW/h ）での比生産性として本明細書でさらに定義される。所与の微生物の OD_{600} に対するCDWの関係を使用して、比生産性を、1時間当たり600nmにおける培養プロセスの光学密度（OD）当たり培養培地1リットル当たりの産物のグラム（ g/L/h/OD ）として表現することもできる。

【0196】

用語「収率」は、原料の単位重量当たりに得られる産物の量として定義され、基質1g

当たりの産物の g (g/g) として表現することができる。収率は、理論的収率のパーセンテージとして表現することができる。「理論的収率」は、産物を作製するのに使用される代謝経路の化学量論比によって決まる所与の量の基質あたりに産生され得る産物の最大量として定義される。

【0197】

用語「力価 (titre)」または「力価 (titer)」は、溶液の強度または溶液中の物質の濃度として定義される。例えば、発酵プロセス中の目的の産物（例えば、低分子、ペプチド、合成化合物、燃料、アルコールなど）の力価は、発酵プロセス 1 リットル当たりの溶液中の目的の産物の g (g/L) として記述される。

【0198】

用語「総力価」は、プロセスにおける初期体積またはプロセスにおける操作体積に対する、プロセスから取り出され、回収される溶液中の目的の産物、適用可能な場合、気相中の目的の産物、および任意の目的の産物を含むがこれらに限定されない、プロセスにおいて産生されるすべての目的の産物の和として定義される。

【0199】

本明細書で使用される場合、用語「HTP 遺伝子設計ライブラリー」または「ライブラリー」は、本開示による遺伝子攪乱のコレクションを指す。一部の実施形態では、本発明のライブラリーは、i) データベースまたは他のコンピューターファイル中の配列情報のコレクション、ii) 上述の一連の遺伝子エレメントをコードする遺伝子コンストラクトのコレクション、または iii) 前記遺伝子エレメントを含む宿主細胞株として顕在化させることができる。一部の実施形態では、本開示のライブラリーは、個々のエレメントのコレクション（例えば、PRO スワップライブラリーのプロモーターのコレクション、STOP スワップライブラリーのターミネーターのコレクション、SOLUBILITY TAG スワップライブラリーのタンパク質可溶性タグのコレクション、または DEGRADATION TAG スワップライブラリーのタンパク質分解タグのコレクション）を指すことができる。他の実施形態では、本開示のライブラリーは、遺伝子エレメントの組合せ、例えば、プロモーター：遺伝子、遺伝子：ターミネーター、またはさらにはプロモーター：遺伝子：ターミネーターの組合せも指すことができる。一部の実施形態では、本開示のライブラリーは、プロモーター、ターミネーター、タンパク質可溶性タグおよび／またはタンパク質分解タグの組合せも指すことができる。一部の実施形態では、本開示のライブラリーは、宿主生物におけるライブラリーの各メンバーを適用する効果と関連したメタデータをさらに含む。例えば、本明細書で使用されるライブラリーは、プロモーター：遺伝子配列組合せのコレクションを、特定の種における 1 つまたは複数の表現型に対するこれらの組合せの結果として生じる効果と一緒に含み、よって、将来のプロモーター スワップにおいて前記組合せを使用することについての将来の予測値を改善する。

【0200】

本明細書で使用される場合、用語「SNP」は、核内低分子多形（複数可）を指す。一部の実施形態では、本開示の SNP は、広く解釈されるべきであり、一塩基多型、配列挿入、欠失、反転、および他の配列置き換えを含む。本明細書で使用される場合、用語「非同義の」または「非同義 SNP」は、宿主細胞タンパク質におけるコード変化をもたらす変異を指す。

【0201】

ゲノム操作の「ハイスルーブット (HTP)」方法は、自動化設備（例えば、液体ハンドラーまたはプレートハンドラーマシン）の少なくとも 1 つのピースを利用して、前記方法の少なくとも 1 つのステップを実行することができる。

【0202】

株改良の伝統的な方法

株改良の伝統的な手法は、2 つのタイプの手法：指向株操作およびランダム変異誘発に広く分類され得る。

【0203】

株改良の指向操作方法は、特定の生物体の少数の遺伝子エレメントの計画された攪乱を伴う。これらの手法は、典型的には、特定の生合成プログラムまたは発生プログラムをモジュレートすることに着目し、前記経路に影響する遺伝因子および代謝因子の先の知識を利用する。その最も単純な実施形態では、指向操作は、特徴付けられた形質（例えば、遺伝子、プロモーター、または測定可能な表現型を生じさせることができる他の遺伝子エレメント）を、1つの生物体から同じまたは異なる種の別の生物体への移入を伴う。

【0204】

株操作のランダム手法は、性能改善を同定するように設計された広範なスクリーニングに加えて、親株のランダム変異誘発を伴う。これらのランダム変異を生じさせる手法には、紫外線への曝露、またはメタンスルホン酸エチルなどの変異誘発化学物質が含まれる。ランダムで概ね予測不可能であるが、株改良のこの伝統的手法は、より指向性の遺伝子マニピュレーションと比較していくつかの利点を有していた。第1に、多くの工業用生物体は、これらの遺伝子レポーターおよび代謝レポーターの観点からあまりよく特徴付けられておらず（かつあまりよく特徴付けられていないままであり）、代替の指向改善手法を不可能でなくとも困難にした。

10

【0205】

第2に、比較的よく特徴付けられたシステムにおいてでさえ、工業的性能改善をもたらす遺伝子型の変化は、予測するのが困難であり、時に、既知および未知の機能の多くの遺伝子における累積的変異を必要とする上位性表現型として現れるだけである。

【0206】

さらに、長年、所与の工業用生物体において指向ゲノム変異を引き起こすのに要求される遺伝的ツールは、利用可能でなく、または非常に遅く、かつ／もしくは使用するのが困難であった。

20

【0207】

しかし、伝統的な株改良プログラムの適用の拡大により、所与の株系統における獲得が次第に低減し、最終的にさらなる株効率の可能性が枯渇する。有益なランダム変異は、相対的に希有な事象であり、大きいスクリーニングプールおよび高い変異率を必要とする。これは必然的に、「改善された」株における多くの中立のかつ／または有害な（もしくは部分的に有害な）変異の不慮の蓄積をもたらし、それは、最終的に将来の効率向上に対して重荷となる。

30

【0208】

伝統的な累積的改善手法の別の制限事項は、任意の株測定基準に対する任意の特定の変異の効果について分かっている情報が皆無かそれに近いことである。これは基本的に、有益な変異を組み合わせ、統合し、または中立のもしくは有害な変異誘発の「荷物 (baggage)」を除去する研究者の能力を制限する。

【0209】

変異誘発系統内の株間で変異をランダムに組み換える他の手法および技術が存在する。例えば、DNAシャッフリング、進化、または分子育種と時に呼ばれる反復的配列組換えのいくつかのフォーマットおよび例は、1994年2月17日に出願された米国特許出願第08/198,431号、1995年2月17日に出願された第PCT/US95/02126号、1995年4月18日に出願された第08/425,684号、1995年10月30日に出願された第08/537,874号、1995年11月30日に出願された第08/564,955号、1996年3月25日に出願された第08/621,859号、1996年3月25日に出願された第08/621,430号、1996年4月18日に出願された第PCT/US96/05480号、1996年5月20日に出願された第08/650,400号、1996年7月3日に出願された第08/675,502号、1996年9月27日に出願された第08/721,824号、および1996年9月27日に出願された第08/722,660号；Stemmer、Science、270巻：1510頁（1995年）；Stemmerら、Gene、164巻：49～53頁（1995年）；Stemmer、Bio/Technology、13巻：549～553頁（1995年）；Stemmer、Proc. Na

40

50

tl. Acad. Sci. U.S.A., 91巻: 10747~10751頁(1994年); Stemmer, Nature, 370巻: 389~391頁(1994年); Cramerら、Nature Medicine、2巻(1号): 1~3頁(1996年); Cramerら、Nature Biotechnology、14巻: 315~319頁(1996年)に記載されている。これらのそれぞれは、すべての目的に関してその全体が参照により本明細書に組み込まれている。

【0210】

これらとしては、プロトプラスト融合、および変異された株全体にわたるゲノム組換えを促進する全ゲノムシャッフリングなどの技法が挙げられる。酵母および糸状菌などの一部の工業用微生物について、天然の接合サイクル(mating cycle)もペアワイズゲノム組換えのために活用することができる。このように、有害な変異を、変異体を親株で「戻し交配する」ことによって除去し、有益な変異を統合することができる。さらに、2つの異なる株系統由来の有益な変異を潜在的に組み合わせることができ、それにより、単一の株系統をそのまま変異させることから入手可能であり得るものよりもさらなる改善可能性がもたらされる。しかし、これらの手法は、多くの制限を受けやすく、それらは、本開示の方法を使用して回避される。

10

【0211】

例えば、上述した伝統的な組換え手法は、遅く、変異をスワップするのに相対的に少数のランダム組換えクロスオーバー事象を利用し、したがって、任意の所与のサイクルまたは期間内に試みることができる組合せの数が限定される。さらに、先行技術における自然組換え事象は、本質的にランダムであるが、これらもゲノムの位置的バイアスを受けやすい。

20

【0212】

最も重要なことに、伝統的な手法はまた、個々の変異の影響についてわずかな情報しか提供せず、組み換えられた変異のランダムな分配に起因して、多くの特定の組合せを生成および評価することができない。

【0213】

伝統的な株改良プログラムに関連した上述の問題の多くを克服するために、本開示は、コンピューターで駆動され、分子生物学、自動化、データ分析、および機械学習プロトコルを統合するユニークなHTPゲノム操作プラットフォームを示す。この統合プラットフォームは、HTP遺伝子設計ライブラリーを構築するのに使用される一式のHTP分子ツールセットを利用する。これらの遺伝子設計ライブラリーは、以下で詳しく述べられる。図8は、本開示のE. coli株改良プログラムの実施形態の概要を示す。

30

【0214】

ここで開示されるHTPプラットフォームおよびそのユニークな微生物遺伝子設計ライブラリーは、微生物株発生および進化のパラダイムを基本的にシフトする。例えば、工業用微生物株を開発する伝統的な変異誘発ベース方法は、最終的に、何年ものランダム変異誘発にわたって蓄積された重い変異誘発性ロードを背負った微生物に至る。

【0215】

この問題を解決する(すなわち、これらの微生物によって蓄積された遺伝的荷物を除去する)能力は、数十年にわたって微生物研究者が成し遂げられなかった。しかし、本明細書に開示のHTPプラットフォームを利用すると、これらの工業用株を「再建する」ことができ、有害である遺伝子変異を同定および除去することができる。調和して、有益と同定される遺伝子変異を保持し、一部の場合では、改善することができる。結果として生じる微生物株は、これらの親株と比較して優れた表現型形質(例えば、目的の化合物の改善された産生)を実証する。

40

【0216】

さらに、本明細書に教示のHTPプラットフォームは、微生物株性能に対して個々の変異が有する効果を同定し、特徴付け、定量することができる。この情報、すなわち、所与の遺伝子変化xが宿主細胞表現型yに対してどのような効果を有するか(例えば、目的の化合物または産物の産生)を生成することができ、次いで、以下で論じられる微生物HT

50

P 遺伝子設計ライブラリーに記憶することができる。すなわち、各遺伝子順列についての配列情報および宿主細胞表現型に対するその効果が1つまたは複数のデータベースに記憶され、後続の分析（例えば、以下で論じられる上位性マッピング（epistasis mapping））に利用可能である。本開示は、遺伝子挿入コンストラクトの形態で、または前記遺伝子順列を含有する1種または複数の宿主細胞生物体の形態で貴重な遺伝子順列を物理的に保存／記憶する方法も教示する（例えば、以下で論じられるライブラリーを参照）。

【0217】

これらのHTP遺伝子設計ライブラリーを洗練されたデータ分析および機械学習プロセスと統合された反復プロセスと結びつけると、宿主細胞を改善するための劇的に異なる方法が出現する。したがって、教示されるプラットフォームは、宿主細胞株を開発する先に論じられた伝統的な方法と基本的に異なる。教示されるHTPプラットフォームは、先の方法に関連した欠点の多くを被らない。これらおよび他の利点は、以下で論じられる、HTP分子ツールセットおよび得られる遺伝子設計ライブラリーを参照して明らかである。

10

【0218】

遺伝子設計 & 微生物操作：一式のHTP分子ツールおよびHTP遺伝子設計ライブラリーを利用する株改良の系統的なコンビナトリアル手法

上述した通り、本開示は、株全体にわたる遺伝子変化の反復性系統的導入および除去によって微生物生物体を操作するための新規HTPプラットフォームおよび遺伝子設計ストラテジーを提供する。プラットフォームは、HTP遺伝子設計ライブラリーの創製を可能にし、所与の宿主株への遺伝子変更の効率的なインプリメンテーションを可能にする一式の分子ツールによって支持される。

20

【0219】

本開示のHTP遺伝子設計ライブラリーは、特定の微生物株バックグラウンド内に導入され得る可能な遺伝子変更の源として機能を果たす。このように、HTP遺伝子設計ライブラリーは、所与の微生物株の初期のまたはさらなる操作に適用され得る遺伝的多様性の宝庫または遺伝子攪乱のコレクションである。宿主株へのインプリメンテーションのために遺伝子設計をプログラムするための技法は、本明細書にそれらのそれぞれのその全体が参照により組み込まれている「Microbial Strain Design System and Methods for Improved Large Scale Production of Engineered Nucleotide Sequences」という表題の、係属中の米国特許出願第15／140,296号および係属中の国際特許出願第PCT／US17／29725号に記載されている。

30

【0220】

このプラットフォームで利用されるHTP分子ツールセットは、とりわけ、（1）プロモータースワップ（PROスワップ）、（2）SNPスワップ、（3）開始／停止コドン交換、（4）STOPスワップ、（5）配列最適化、（6）SOLUBILITY TAGスワップおよび（7）DEGRADATION TAGスワップを含み得る。本開示のHTP方法は、（8）上位性マッピングプロトコル、を含めたHTPツールセットの統合／コンビナトリアル使用を指示するための方法も教示する。上述した通り、この一式の分子ツールは、単独または組合せのいずれかで、HTP遺伝子設計宿主細胞ライブラリーの創製を可能にする。

40

【0221】

実証されるように、教示されるHTP微生物操作プラットフォームとの関連で上述のHTP遺伝子設計ライブラリーを利用すると、有益な「原因となる」変異または遺伝子区画の同定および統合、ならびにまた、受動変異または有害な変異または遺伝子区画の同定および除去が可能になる。この新しい手法は、伝統的なランダム変異誘発または指向遺伝子操作（directed genetic engineering）によって達成することができなかった株性能の急速な改善を可能にする。遺伝的荷重を除去し、または遺伝的荷重を有さない株内に有益な変化を統合すると、さらなる改善を可能にし得る追加のランダム変異誘発の新しい、ロバストな出発点も提供される。

50

【0222】

一部の実施形態では、本開示は、オルソゴナルな有益な変化が、変異誘発株系統の様々な個別分岐にわたって同定されるので、これらを、より良好に働く株へと急速に統合することもできることを教示する。これらの変異はまた、変異誘発系統の一部でない株、例えば、指向遺伝子操作によって獲得される改良を有する株へと統合され得る。

【0223】

一部の実施形態では、本開示は、それが、発現および非発現遺伝子エレメントを含む複数の異なるゲノム領域にわたって変異のゲノムワイドコンビナトリアル効果を分析し、集められた情報（例えば、実験結果）を使用して株を増強させることが予期される変異組合せを予測するという点で公知の株改良手法と異なる。

10

【0224】

一部の実施形態では、本開示は、i) 開示される発明を介した改良に適用できる工業用微生物および他の宿主細胞、ii) 下流分析のための多様性プールの生成、iii) 大きいバリエーションプールをハイスループットスクリーニングおよびシーケンシングするための方法およびハードウェア、iv) ゲノムワイド変異の相乗効果の機械学習コンピューター分析および予測のための方法およびハードウェア、ならびにv) ハイスループット株操作のための方法を教示する。

【0225】

以下の分子ツールおよびライブラリーは、例示的な微生物例の観点から論じられている。当業者は、本開示のHTP分子ツールは、真核細胞およびより高等の生命体を含めた任意の宿主細胞に適合することを認識する。さらに、例示されている実施形態の多くは *Corynebacterium* において行われるが、同じ原理およびプロセスを *Escherichia coli* において展開することができる。

20

【0226】

微生物操作プラットフォームで利用される様々なHTP遺伝子設計ライブラリーの創製を可能にする同定されたHTP分子ツールセットのそれぞれを、次に論じる。

【0227】

1. プロモータースワップ：プロモータースワップ微生物株ライブラリーを得るための分子ツール

一部の実施形態では、本開示は、全宿主株表現型に対する有益な効果（例えば、収率または生産性）を生じさせるのに最適な発現性質を有するプロモーターを選択する方法を教示する。

30

【0228】

例えば、一部の実施形態では、本開示は、一連の発現強度（例えば、以下に論じられるプロモーターラダー）、または優れた調節性質（例えば、選択された遺伝子のより厳格な調節制御）を呈する、宿主細胞内の1つもしくは複数のプロモーターを同定し、かつ／または1つもしくは複数のプロモーターのバリエーションを生成する方法を教示する。これらの同定および／または生成されるプロモーターの特定の組合せを、以下でより詳細に説明されるプロモーターラダーとして一緒にグループ化することができる。

【0229】

40

次いで、問題のプロモーターラダーは、目的の所与の遺伝子と関連させられる。よって、プロモーター $P_1 \sim P_8$ （一連の発現強度を呈するように同定および／または生成された8つのプロモーターを表す）を有し、プロモーターラダーを微生物中の目的の単一遺伝子と関連させる（すなわち、所与の標的遺伝子に作動可能に連結した所与のプロモーターで微生物を遺伝的に操作する）場合、8つのプロモーターの各組合せの効果を、操作された微生物が標的遺伝子と関連した特定のプロモーター（複数可）を除いて他の点では同一の遺伝的バックグラウンドを有することを考慮して、各コンビナトリアル取り組みの結果として生じる操作された株のそれぞれを特徴付けることによって確認することができる。

【0230】

このプロセスを介して操作された結果として生じる微生物は、HTP遺伝子設計ライブ

50

ラリーを形成する。

【0231】

H T P 遺伝子設計ライブラリーは、このプロセスを介して形成される実際の物理的な微生物株コレクションを指すことができ、各メンバー株は、他の点では同一の遺伝的背景グラウンドにおいて、特定の標的遺伝子に作動可能に連結した所与のプロモーターを代表し、前記ライブラリーは、「プロモータースワップ微生物株ライブラリー」と呼ばれる。E. coli の特定の状況では、該ライブラリーを「プロモータースワップ E. coli 株ライブラリー」と呼ぶことができるが、E. coli は微生物の具体的な例であるので、これらの用語を同義に使用することができる。

【0232】

さらに、H T P 遺伝子設計ライブラリーは、遺伝子攪乱のコレクションを指すことができ、この場合、所与のプロモーター x は、所与の遺伝子 y に作動可能に連結しており、前記コレクションは、「プロモータースワップライブラリー」と呼ばれる。

【0233】

さらに、プロモーター $P_1 \sim P_8$ を含む同じプロモーターラダーを利用して微生物を操作することができ、ここで、8つプロモーターのそれぞれは、10の異なる遺伝子標的に作動可能に連結される。この手順の結果は、目的の標的遺伝子に作動可能に連結した特定のプロモーターを除いて、他の点では遺伝的に同一と仮定される80の微生物である。これらの80の微生物は、適切にスクリーニングし、特徴付け、別のH T P 遺伝子設計ライブラリーを生じさせることができる。H T P 遺伝子設計ライブラリー内の微生物株を特徴付けると、リレーショナルデータベース、オブジェクト指向データベース、または大規模分散NoSQLデータベースを含めた任意のデータ記憶コンストラクトに記憶することができる情報およびデータが生じる。このデータ／情報は、例えば、所与の遺伝子標的に作動可能に連結しているときの所与のプロモーター（例えば、 $P_1 \sim P_8$ ）の効果であり得る。このデータ／情報は、プロモーター $P_1 \sim P_8$ のうちの2つまたはそれ超を所与の遺伝子標的に作動可能に連結することから生じるより広いセットのコンビナトリアル効果でもあり得る。

【0234】

8のプロモーターおよび10の標的遺伝子の上述の例は、単に例示的であり、その理由は、この概念は、一連の発現強度および任意の所与の数の標的遺伝子の呈示に基づいて一緒にグループ化された任意の所与の数のプロモーターを用いて適用することができるためである。当業者は、任意の遺伝子標的の前に2つまたはそれ超のプロモーターを作動可能に連結する能力も認識する。よって、一部の実施形態では、本開示は、プロモーターラダー由来の1、2、3、またはそれ超のプロモーターが1つまたは複数の遺伝子に作動可能に連結したプロモータースワップライブラリーを教示する。

【0235】

要約すると、様々なプロモーターを利用して生物体内の様々な遺伝子の発現を駆動することは、目的の形質を最適化する強力なツールである。本発明者らが開発したプロモータースワッピングの分子ツールは、少なくとも1つの条件下で少なくとも1つの遺伝子座の発現を変更することが実証されたプロモーター配列のラダーを使用する。次いでこのラダーは、ハイスループットゲノム操作を使用して生物体内の遺伝子の群に系統的に適用される。この群の遺伝子は、いくつかの方法のいずれか1つに基づいて目的の形質に影響を与える高い見込みを有すると決定される。これらは、目的の形質に対する公知の機能もしくは影響に基づく選択物、または先に決定された有益な遺伝的多様性に基づくアルゴリズム選択物を含み得る。一部の実施形態では、遺伝子の選択物は、所与の宿主内のすべての遺伝子を含み得る。他の実施形態では、遺伝子の選択物は、ランダムに選択された所与の宿主内のすべての遺伝子のサブセットであり得る。

【0236】

次いで、遺伝子に連結したプロモーター配列を含有する生物体の結果として生じるH T P 遺伝子設計微生物株ライブラリーが、ハイスループットスクリーニングモデルにおいて

性能について評価され、性能を増大させるプロモーター遺伝子連鎖が決定され、情報がデータベースに記憶される。遺伝子攪乱（すなわち、所与の遺伝子 y に作動可能に連結した所与のプロモーター x）のコレクションは、「プロモーターズワップライブラリー」を形成し、これは、微生物操作処理で利用される潜在的な遺伝子変更の源として利用することができる。経時的に、より大きいセットの遺伝子攪乱は、より多様な宿主細胞バックグラウンドに対してインプリメントされるので、各ライブラリーは、実験的に確認されたデータのコーパスとしてより強力になり、それを使用して目的の任意のバックグラウンドに対する標的とした変化をより正確かつ予想通りに設計することができる。

【0237】

生物体内の遺伝子の転写レベルは、生物体挙動に影響するための重要な管理点である。転写は、翻訳（タンパク質発現）としっかりと共役し、このタンパク質は、生物体挙動を決定する量で発現される。細胞は、数千の異なるタイプのタンパク質を発現し、これらのタンパク質は、多数の複雑なやり方で相互作用して機能を生み出す。一連のタンパク質の発現レベルを系統的に変更することによって、機能は、複雑性のために、予測することが困難であるように変更され得る。一部の変更は、性能を増大させることができ、したがって、性能を評価するための機構と関連して、この技法は、機能が改善された生物体の生成を可能にする。

【0238】

低分子合成経路との関連で、酵素は、基質で始まり、目的の低分子で終わる、直鎖または分岐鎖中で該酵素の低分子基質および産物を介して相互作用する。これらの相互作用は、順次連結されるので、このシステムは、分散制御を呈し、1つの酵素の発現の増大は、別の酵素が律速となるまで経路フラックスを増大させることができるだけである。

【0239】

代謝制御分析（MCA）は、実験データおよび第1原理から、どの1つまたは複数の酵素が律速であるかを決定するための方法である。しかし、MCAは、それが新しい律速な酵素を決定するのに各発現レベル変化後に広範な実験を必要とするので、限定されている。プロモーターズワッピングは、この状況において有利であり、その理由は、プロモーターラダーを経路内の各酵素に適用することによって、律速の酵素が見出され、同じことを後続のラウンドで行って律速となる新しい酵素を見出すことができるためである。さらに、機能の読み取りは、目的の低分子のより良好な産生であるので、どの酵素が律速であるかを決定する実験は、産生を増大させる操作と同じであり、よって開発時間を短縮する。一部の実施形態では、本開示は、多ユニット酵素の個々のサブユニットをコードする遺伝子へのPROスワップの適用を教示する。さらに他の実施形態では、本開示は、個々の酵素、または生合成経路全体の調節を担う遺伝子にPROスワップ技法を適用する方法を教示する。

【0240】

一部の実施形態では、本開示のプロモーターズワップツールを使用して、選択された遺伝子標的の最適の発現が同定される。一部の実施形態では、プロモーターズワップの目標は、標的遺伝子の発現を増大させて代謝経路または遺伝経路における障壁を低減することであり得る。他の実施形態では、プロモーターズワップの目標は、標的遺伝子の発現が要求されないとき、前記標的遺伝子の発現を低減して宿主細胞内の不要なエネルギー消費を回避することであり得る。

【0241】

転写、輸送、またはシグナル伝達のような他の細胞系という状況で、様々な合理的な方法を使用して、どのタンパク質が発現変化の標的であり、その変化が何であるべきかを実験的に試みて見出すことができる。これらの合理的な方法は、性能を改善するものを見出すのに試験されなければならない攪乱の数を低減するが、これらは、かなりのコストでそれを行う。遺伝子欠失スタディにより、その存在が特定の機能に極めて重要であるタンパク質が同定され、次いで重要な遺伝子を過剰発現させることができる。タンパク質相互作用の複雑性に起因して、これは、性能の増大において効果的でないことが多い。細胞にお

けるタンパク質レベルの関数として転写またはシグナル伝達挙動を第1原理から記述することを試みる異なるタイプのモデルが開発された。これらのモデルは、多くの場合、発現変化が異なる機能または改善された機能をもたらし得る標的を示唆する。これらのモデルの根底をなす仮定は、過度に単純化されており、パラメータは、測定するのが困難であり、したがって、これらが行う予測は、特に非モデル生物体について誤っていることが多い。遺伝子欠失およびモデル化の両方を用いると、ある特定の遺伝子にどのように影響するかを決定するのに要求される実験は、性能を改善する変化を起こすための後続の仕事と異なる。プロモータースワッピングは、これらの課題を回避し、その理由は、特定の攪乱の重要性を強調する構築された株も、すでに改良された株であるためである。

【0242】

10

よって、特定の実施形態では、プロモータースワッピングは、

【0243】

1. 「ラダー」として作用する一連の「x」プロモーターを選択するステップ。理想的にはこれらのプロモーターは、多重ゲノム遺伝子座全体にわたって高度に多様な発現をもたらすことが示されているが、唯一の要件は、これらが遺伝子発現を何らかのやり方で攪乱することである。

【0244】

2. 標的にする一連の「n」遺伝子を選択するステップ。このセットは、ゲノム内のあらゆるオープンリーディングフレーム（ORF）、またはORFのサブセットであり得る。サブセットは、先に実証された有益な攪乱（先のプロモータースワップもしくは先のSNPスワップ）との関連によって、先に生成された攪乱間の上位性相互作用に基づくアルゴリズム選択、標的にする有益なORFに関する仮説に基づく他の選択判定基準によって、またはランダム選択を介して、機能に関連したORFについてのアノテーションを使用して選択することができる。他の実施形態では、「n」標的化遺伝子は、非コードRNAを含めた非タンパク質コード遺伝子を含み得る。

20

【0245】

3. 以下の遺伝子改変を迅速に、かつ一部の実施形態では、並行して実行するためのハイスループット株操作：生来のプロモーターが標的遺伝子nの前に存在し、その配列が公知であるとき、生来のプロモーターをラダー内のxプロモーターのそれぞれで置き換える。生来のプロモーターが存在せず、またはその配列が未知であるとき、ラダー内のxプロモーターのそれぞれを遺伝子nの前に挿入する（例えば、図21を参照）。このように、株の「ライブラリー」（HTP遺伝子設計ライブラリーとも呼ばれる）が構築され、ライブラリーの各メンバーは、他の点では同一の遺伝子状況（genetic context）においてn標的に作動可能に連結したxプロモーターの一事例である。先に記載したように、プロモーターの組合せを挿入し、ライブラリーが構築される組み合わせの可能性（combinatorial possibility）の範囲を拡張することができる。

30

【0246】

4. 1つまたは複数の測定基準に対する株の性能が最適化されている性能を示す状況における株のライブラリーのハイスループットスクリーニング。
を含むマルチステッププロセスである。

40

【0247】

この基本プロセスを拡張して、とりわけ、（1）相互作用プロセスにおいて1つずつ、または単一ステップにおける複数の変化として、単一株バックグラウンド内に複数の有益な攪乱を統合することによって、株性能をさらに改善することができる。複数の攪乱は、確定した変化の特定のセット、または変化の部分的にランダム化されたコンビナトリアルライブラリーのいずれかであり得る。例えば、標的のセットが経路内のあらゆる遺伝子である場合、攪乱のライブラリーを株の先のライブラリーの改善された1つまたは複数のメンバー内に逐次再生すると、任意の所与の反復においてどの遺伝子が律速であるかにかかわらず経路内の各遺伝子の発現レベルを最適化することができ；（2）ライブラリーの個々のおよびコンビナトリアル生成の結果として生じる性能データを、そのデータを使用し

50

て各攪乱の相互作用に基づいて攪乱の最適のセットを予測するアルゴリズムに送り；（3）上記2つの手法の組合せをインプリメントする（図20を参照）。

【0248】

上記に論じた分子ツールまたは技法は、プロモータースワッピングとして特徴付けられるが、プロモーターに限定されず、一連の標的の発現レベルを系統的に変更する他の配列変化を含み得る。一連の遺伝子の発現レベルを変更するための他の方法は、a) リボソーム結合部位（または真核生物内のコザック配列）のラダー；b) 各標的の開始コドンの他の開始コドンのそれぞれでの置き換え（すなわち、以下に論じられる開始／停止コドン交換）；c) 転写物の5'もしくは3'末端への、または任意の他の場所における配列を安定化し、または不安定にする様々なmRNAの結合、d) タンパク質中の任意の場所における配列を安定化し、または不安定にする様々なタンパク質の結合（すなわち、以下に論じられる分解または可溶化タグ交換）を含み得る。

10

【0249】

手法は、工業用微生物を用いて本開示において例示されているが、所望の形質が遺伝子変異体の集団において同定され得る任意の生物体に適用可能である。例えば、これは、CHO細胞、酵母、昆虫細胞、藻類、および植物などの多細胞生物の性能を改善するのに使用することができる。

【0250】

2. SNPスワップ：SNPスワップ微生物株ライブラリーを得るための分子ツール

ある特定の実施形態では、SNPスワッピングは、微生物株の改良のランダム変異誘発手法ではなく、むしろ、株全体にわたる個々の核内低分子多形ヌクレオチド変異（Small Nuclear Polymorphism nucleotide mutation）（すなわち、SNP）の系統的な導入または除去を伴う（それが名称「SNPスワッピング」の由来である）。

20

【0251】

このプロセスを介して操作される結果として生じる微生物は、HTP遺伝子設計ライブラリーを形成する。

【0252】

HTP遺伝子設計ライブラリーは、このプロセスを介して形成される実際の物理的な微生物株コレクションを指すことができ、各メンバー株は、他の点では同一の遺伝的バックグラウンドにおいて、所与のSNPの存在または非存在を代表し、前記ライブラリーは、「SNPスワップ微生物株ライブラリー」と呼ばれる。E. coliの特定の状況では、該ライブラリーを「SNPスワップE. coli株ライブラリー」と呼ぶことができるが、E. coliは微生物の具体的な例であるので、これらの用語を同義に使用することができる。

30

【0253】

さらに、HTP遺伝子設計ライブラリーは、遺伝子攪乱のコレクションを指すことができ、この場合、所与のSNPが存在し、または所与のSNPが存在せず、前記コレクションは、「SNPスワップライブラリー」と呼ばれる。

【0254】

一部の実施形態では、SNPスワッピングでは、同定された有益な性能効果を有する標的SNP「ビルディングブロック」の最適な組合せを用いて宿主生物を再構築する。よって、一部の実施形態では、SNPスワッピングでは、反復プロセスにおいて1つずつ、または単一ステップにおける複数の変化としてのいずれかで、単一株バックグラウンド内に複数の有益な変異を統合することを伴う。複数の変化は、確定した変化の特定のセット、または変異の部分的にランダム化されたコンビナトリアルライブラリーのいずれかであり得る。

40

【0255】

他の実施形態では、SNPスワッピングでは、反復プロセスにおいて1つずつ、または単一ステップにおける複数の変化としてのいずれかで、株から有害と同定された複数の変異の除去も関与する。複数の変化は、規定された変化の特定のセット、または変異の部分

50

的にランダム化されたコンビナトリアルライブラリーのいずれかであり得る。一部の実施形態では、本開示のSNPスワッピング方法は、有益なSNPの付加ならびに有害なおよび／または中立の変異の除去の両方を含む。

【0256】

SNPスワッピングは、変異誘発および目的とする形質改善のための選択に供された株の系統における有益および有害な変異の両方を同定および活用する強力なツールである。SNPスワッピングは、ハイスループットゲノム操作技法を利用して、変異誘発系統における個々の変異の影響を系統的に決定する。ゲノム配列は、既知の性能改善を有する変異誘発系統の1つまたは複数の世代にわたる株について決定される。次いでハイスループットゲノム操作が系統的に使用されて、より早期の系統株における改良された株由来の変異を反復し、かつ／またはより後期の株内の変異をより早期の株配列に戻す。次いでこれらの株の性能が評価され、目的の改善された表現型に対する各個々の変異の寄与を決定することができる。上述した通り、このプロセスから生じる微生物株が分析され／特徴付けられ、宿主株全体にわたる微生物株改良を通知することができるSNPスワップ遺伝子設計ライブラリーの基礎を形成する。

10

【0257】

有害な変異を除去すると、即時に性能を改善することができ、変異誘発荷重に供されていない株バックグラウンド内の有益な変異を統合すると、株性能を迅速かつ大いに改善することができる。SNPスワッピングプロセスを介して生成される様々な微生物株は、HTP遺伝子設計SNPスワッピングライブラリーを形成し、それは、様々な付加／欠失／または統合されたSNPを含むが、他の点では同一の遺伝的バックグラウンドを有する微生物株である。

20

【0258】

先に論じた通り、性能改善のためのランダム変異誘発および後続のスクリーニングは、工業用株改良のために一般に使用される技法であり、大規模製造に現在使用されている多くの株は、長年、時に数十年の期間にわたって反復してこのプロセスを使用して開発されてきた。ゲノム変異の生成のランダム手法、例えば、UV放射線またはメタンスルホン酸エチルなどの化学的変異誘発物質への曝露は、工業用株改良の好適な方法であった。その理由は、1) 工業用生物体は、遺伝的または代謝的にあまりよく特徴付けられておらず、指向改善手法のための標的選択を困難または不可能にする場合があり、2) 相対的に十分に特徴付けられたシステムにおいてでさえ、工業的性能改善をもたらす変化は、予測するのが困難であり、既知の機能を有さない遺伝子の攪乱を必要とする場合があり、3) 所与の工業用生物体において指向ゲノム変異を作製するための遺伝的ツールは、利用可能でなく、または非常に遅く、かつ／もしくは使用するのが困難である場合があるためである。

30

【0259】

しかし、このプロセスの上述の利益にもかかわらず、いくつかの公知の不利点も存在する。有益な変異は、相対的に希有な事象であり、定まったスクリーニング能力を用いてこれらの変異を見出すために、変異率は、十分に高くなければならない。これにより、望まれない中立のおよび部分的に有害な変異が有益な変化とともに株内に組み入れられることが多い。経時的に、この「変異誘発荷重」が蓄積し、全体的なロバスト性、および増殖速度などの重要な形質が欠損した株を生じる。最終的に、「変異誘発荷重」は、ランダム変異誘発による性能のさらなる改善を得ることをますます困難または不可能にする。適当なツールを用いない場合、株系統の個別分岐および並列分岐に見出される有益な変異を統合することが不可能である。

40

【0260】

SNPスワッピングは、変異誘発系統内の株を比較するとき観察される一部またはすべての変異を系統的に反復し、または戻すことによってこれらの制限事項を克服する一手法である。このように、有益な（「原因となる」）変異を同定および統合することもでき、かつ／または有害な変異を同定および除去することもできる。これにより、さらなるランダム変異誘発または標的化遺伝子操作によって達成することができなかった株性能の迅速

50

な改善が可能になる。

【0261】

遺伝的荷重を除去し、または有益な変化を遺伝的荷重が無い株内に統合すると、さらなる改善を可能にすることができる追加のランダム変異誘発の新しいロバストな出発点ももたらされる。

【0262】

さらに、オルソゴナルな有益な変化が、変異誘発株系統の様々な個別分岐にわたって同定されるので、これらを、より良好に働く株へと急速に統合することができる。これらの変異はまた、変異誘発系統の一部でない株、例えば、指向遺伝子操作によって獲得される改良を有する株へと統合され得る。

10

【0263】

変異誘発系統内の株間の変異をランダムに組み換える他の手法および技術が存在する。これらとしては、プロトプラスト融合、および変異された株全体にわたるゲノム組換えを促進する全ゲノムシャッフリングなどの技法が挙げられる。酵母および糸状菌などの一部の工業用微生物について、天然の接合サイクルもペアワイズゲノム組換えのために活用することができる。このように、有害な変異を、変異体を親株で「戻し交配する」ことによって除去し、有益な変異を統合することができる。しかし、これらの手法は、多くの制限を受けやすく、それは、本開示のSNPスワッピング方法を使用して回避される。

【0264】

例えば、これらの手法は、変異をスワップするのに相対的に少数のランダム組換えクロスオーバー事象を利用するので、株性能を最適化するのに組換えおよびスクリーニングの多くのサイクルを要し得る。さらに、天然の組換え事象は、本質的にランダムであるが、これらもゲノムの位置的偏りを受けやすく、一部の変異は、対処するのが困難であり得る。これらの手法はまた、追加のゲノムシーケンシングおよび分析を伴わない個々の変異の影響についてわずかな情報しか提供しない。SNPスワッピングは、これらの基本的な制限事項を克服し、その理由は、それがランダム手法ではなく、むしろ、株全体にわたる個々の変異の系統的な導入または除去であるためである。

20

【0265】

一部の実施形態では、本開示は、多様性プールの生物体の中に存在するSNP配列多様性を同定するための方法を教示する。多様性プールは、分析に利用される所与の数 n の微生物であり得、前記微生物のゲノムは、「多様性プール」を代表する。

30

【0266】

特定の態様では、多様性プールは、特定の時点における「ベースライン」または「参照」遺伝子配列を有する元の親株(S_1) ($S_1 \text{ Gen}_1$)、および次いで前記 S_1 株に由来し／それから開発された、かつ S_1 のベースラインゲノムに関して異なるゲノムを有する任意の数の後続の子孫株(S_{2-n}) ($S_{2-n} \text{ Gen}_{2-n}$)であり得る。

【0267】

例えば、一部の実施形態では、本開示は、多様性プール内の微生物ゲノムをシーケンシングして各株内に存在するSNPを同定することを教示する。一実施形態では、多様性プールの株は、歴史的な微生物産生株である。よって、本開示の多様性プールは、例えば、工業用参照株、および伝統的な株改良プログラムを介して産生された1つまたは複数の変異工業用株を含み得る。

40

【0268】

一部の実施形態では、多様性プール内のSNPは、「参照株」を参照して決定される。一部の実施形態では、参照株は、野生型株である。他の実施形態では、参照株は、任意の変異誘発に供される前の元の工業用株である。参照株は、専門家が定義することができ、元の野生型株または元の工業用株である必要はない。基準株は、前記参照株に由来し、またはそれから開発された後続の株が比較される「基準」、「参照」、または元の遺伝的バックグラウンドと見なされるものを単に代表する。

【0269】

50

多様性プール内のすべてのSNPが同定されると、本開示は、SNPスワッピングの方法およびスクリーニング方法を教示して、個々にかつ／または群内でSNPの効果（例えば、目的の表現型の創製）を描写する（すなわち、定量し、特徴付ける）。

【0270】

一部の実施形態では、本開示のSNPスワッピング方法は、変異株（例えば、 S_{2-n} Gen $_{2-n}$ の中からの株）内で同定された1つまたは複数のSNPを参照株（ S_1 Gen $_1$ ）または野生型株に導入する（「ウェーブアップ」）ステップを含む。

【0271】

他の実施形態では、本開示のSNPスワッピング方法は、変異株（例えば、 S_{2-n} Gen $_{2-n}$ の中からの株）内で同定された1つまたは複数のSNPを除去する（「ウェーブダウン」）ステップを含む。

【0272】

一部の実施形態では、1つまたは複数のSNP変化（導入または除去のいずれか）を含むそれぞれの生成された株は、培養され、本開示の1つまたは複数の判定基準（例えば、目的の化学物質または産物の産生）下で分析される。分析された宿主株のそれぞれからのデータは、宿主株内に存在する特定のSNPまたはSNPの群と関連付けられ、または相関付けられ、将来使用するために記録される。よって、本開示は、任意の数の目的の微生物遺伝的または表現型形質に対する所与のSNPの効果と同定することができる、大きい、高度にアノテートされたHTP遺伝子設計微生物株ライブラリーの創製を可能にする。これらのHTP遺伝子設計ライブラリーに記憶された情報は、HTPゲノム操作プラットフォームの機械学習アルゴリズムに知らせ、プロセスの将来の反復を指示し、それは最終的に、高度に望ましい性質／形質を有する進化した微生物生物体をもたらす。

【0273】

3. 開始／停止コドン交換：開始／停止コドン微生物株ライブラリーを得るための分子ツール

一部の実施形態では、本開示は、開始および停止コドンバリエーションをスワップする方法を教示する。例えば、*S. cerevisiae*および哺乳動物の典型的な停止コドンは、それぞれTAA（UAA）およびTGA（UGA）である。単子葉植物の典型的な停止コドンは、TGA（UGA）であり、一方、昆虫および*E. coli*は一般に、停止コドンとしてTAA（UAA）を使用する（Dalphinら（1996年）、Nucl. Acids Res.、24巻：216～218頁）。他の実施形態では、本開示は、TAG（UAG）停止コドンの使用を教示する。

【0274】

本開示は、開始コドンをスワップすることを同様に教示する。一部の実施形態では、本開示は、ほとんどの生物体（特に真核生物）によって利用されているATG（AUG）開始コドンの使用を教示する。一部の実施形態では、本開示は、原核生物は、ATG（AUG）を最も多く使用し、その後GTT（GUG）およびTTG（UUG）が続くことを教示する。

【0275】

他の実施形態では、本発明は、ATG開始コドンをTTGで置き換えることを教示する。一部の実施形態では、本発明は、ATG開始コドンをGTTで置き換えることを教示する。一部の実施形態では、本発明は、GTT開始コドンをATGで置き換えることを教示する。一部の実施形態では、本発明は、GTT開始コドンをTTGで置き換えることを教示する。一部の実施形態では、本発明は、TTG開始コドンをATGで置き換えることを教示する。一部の実施形態では、本発明は、TTG開始コドンをGTTで置き換えることを教示する。

【0276】

他の実施形態では、本発明は、TAA停止コドンをTAGで置き換えることを教示する。一部の実施形態では、本発明は、TAA停止コドンをTGAで置き換えることを教示する。一部の実施形態では、本発明は、TGA停止コドンをTAAで置き換えることを教示

する。一部の実施形態では、本発明は、TGA停止コドンをTAGで置き換えることを教示する。一部の実施形態では、本発明は、TAG停止コドンをTAAで置き換えることを教示する。一部の実施形態では、本発明は、TAG停止コドンをTGAで置き換えることを教示する。

【0277】

4. Stopスワップ：STOPスワップ微生物株ライブラリーを得るための分子ツール

一部の実施形態では、本開示は、細胞遺伝子転写の最適化によって宿主細胞生産性を改善する方法を教示する。遺伝子転写は、転写開始（RNAp動員および転写複合体形成）、伸長（鎖合成／拡張）、ならびに転写終結（RNAp脱離および終結）を含めたいくつかの別個の生物学的現象の結果である。多くの関心が、遺伝子の転写モジュレーションによる（例えば、プロモーターを変更し、または調節転写因子を誘導することによる）遺伝子発現の制御に注がれているが、比較的少ない取り組みが遺伝子ターミネーター配列のモジュレーションを介した転写のモジュレーションに向けて行われている。

【0278】

転写が遺伝子発現レベルに影響を与える最も明白なやり方は、Pol II開始の速度によるものであり、これは、プロモーターまたはエンハンサー強度およびトランス活性化因子の組合せによってモジュレートすることができる（Kadonaga, J.T., 2004年、「Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors」、Cell., 2004年1月23日；116巻（2号）：247～57頁）。真核生物では、伸長速度も、オルタナティブスプライシングに影響を与えることによって遺伝子発現パターンを決定する場合がある（Cramer P.ら、1997年、「Functional association between promoter structure and transcript alternative splicing.」、Proc Natl Acad Sci U S A., 1997年10月14日；94巻（21号）：11456～60頁）。遺伝子上の終結の失敗は、プロモーターのPol IIへのアクセシビリティを低減することによって下流遺伝子の発現を損なわせ得る（Greger I.H.ら、2000年、「Balancing transcriptional interference and initiation on the GAL7 promoter of *Saccharomyces cerevisiae*.」、Proc Natl Acad Sci U S A., 2000年7月18日；97巻（15号）：8415～20頁）。このプロセスは、転写干渉として公知であり、下等真核生物において特に関連しており、その理由は、これらが密集した遺伝子を有することが多いためである。

【0279】

終結配列も、配列が属する遺伝子の発現に影響し得る。例えば、スタディは、真核生物における非効率な転写終結は、スプライスされていないpre-mRNAを蓄積させることを示す（West, S.およびProudfoot, N.J., 2009年、「Transcriptional Termination Enhances Protein Expression in Human Cells」、Mol Cell., 2009年2月13日；33巻（3～9号）；354～364頁を参照）。他のスタディも、3'末端プロセシングが非効率な終結によって遅延され得ることを示した（West, S.ら、2008年、「Molecular dissection of mammalian RNA polymerase II transcriptional termination.」、Mol Cell., 2008年3月14日；29巻（5号）：600～10頁）。転写終結も、合成の部位から転写物を放出することによってmRNA安定性に影響し得る。さらに、強力な終結配列により、mRNA安定性を増大させ、したがって、タンパク質存在量および全体的な経路活性を増大させることができる。

【0280】

真核生物における転写機構の終結

真核生物における転写終結は、RNAポリメラーゼIIに関連したタンパク質因子によって認識されるターミネーターシグナルを介して働く。一部の実施形態では、切断・ポリアデニル化特異性因子（CPSF）および切断刺激因子（CstF）は、RNAポリメラーゼIIのカルボキシル末端ドメインからポリAシグナルに移動する。一部の実施形態では、CPSF因子およびCstF因子はまた、終結部位に他のタンパク質を補充し、次い

でこれは、転写物を切断し、転写複合体からmRNAを解放する。終結は、mRNA転写物のポリアデニル化も誘発する。バリデートされた真核生物終結因子の実例およびこれらの保存構造は、本文書の後の部分で論じられている。

【0281】

原核生物における転写の終結

原核生物では、Rho非依存性およびRho依存性終結と呼ばれる2つの主要機構が転写終結を媒介する。Rho非依存性終結シグナルは、外因性転写終結因子を要求せず、その理由は、一連のウリジン(U)残基とともにこれらの配列から転写されるRNAにおけるステムループ構造の形成が、転写複合体からのRNA鎖の放出を促進するためである。一方、Rho依存性終結は、mRNA上でRhoと呼ばれる転写終結因子およびシス作用エレメントを必要とする。Rhoの最初の結合部位、Rho利用(rut)部位は、実際のターミネーター配列の上流の、高シチジン/低グアノシン含有量、および合成されているRNAにおける相対的に小さい二次構造によって特徴付けられる伸長された(約70ヌクレオチド、時に80~100ヌクレオチド)一本鎖領域である。ポリメラーゼ休止部位と遭遇すると、終結が起こり、転写物がRhoのヘリカーゼ活性によって放出される。

【0282】

ターミネータースワッピング(STOPスワップ)

一部の実施形態では、本開示は、最適な発現性質を有する終結配列(「ターミネーター」)を選択して、全体的な宿主株生産性に対して有益な効果を生じさせる方法を教示する。

【0283】

例えば、一部の実施形態では、本開示は、一連の発現強度を呈する、宿主細胞内の1つもしくは複数のターミネーターを同定し、かつ/または1つもしくは複数のターミネーターのバリエーションを生成する(例えば、以下に論じられるターミネーターラダー)方法を教示する。これらの同定および/または生成されるターミネーターの特定の組合せを、以下でより詳細に説明されるターミネーターラダーとして一緒にグループ化することができる。

【0284】

次いで、問題のターミネーターラダーは、目的の所与の遺伝子と関連させられる。よって、ターミネーターT₁~T₈(1つまたは複数のプロモーターと組み合わせられるとき、一連の発現強度を呈するように同定および/または生成された8つのターミネーターを表す)を有し、ターミネーターラダーを宿主細胞内の目的の単一遺伝子と関連させる(すなわち、所与の標的遺伝子の3'末端に作動可能に連結した所与のターミネーターで宿主細胞を遺伝的に操作する)場合、ターミネーターの各組合せの効果を、操作された宿主細胞が標的遺伝子と関連した特定のターミネーター(複数可)を除いて他の点では同一の遺伝的背景グラウンドを有することを考慮して、各コンビナトリアル取り組みの結果として生じる操作された株のそれぞれを特徴付けることによって確認することができる。このプロセスを介して操作される結果として生じる宿主細胞は、HTP遺伝子設計ライブラリーを形成する。

【0285】

HTP遺伝子設計ライブラリーは、このプロセスを介して形成される実際の物理的な微生物株コレクションを指すことができ、各メンバー株は、他の点では同一の遺伝的背景グラウンドにおいて、特定の標的遺伝子に作動可能に連結した所与のターミネーターを代表し、前記ライブラリーは、「ターミネータースワップ微生物株ライブラリー」または「STOPスワップ微生物株ライブラリー」と呼ばれる。E. coliの特定の状況では、該ライブラリーを「ターミネータースワップE. coli株ライブラリー」または「STOPスワップE. coli株ライブラリー」と呼ぶことができるが、E. coliは微生物の具体的な例であるので、これらの用語を同義に使用することができる。

【0286】

さらに、HTP遺伝子設計ライブラリーは、遺伝子攪乱のコレクションを指すことがで

き、この場合、所与のターミネーターxは、所与の遺伝子yに作動可能に連結しており、前記コレクションは、「ターミネーターズワップライブラリー」または「STOPスワップライブラリー」と呼ばれる。

【0287】

さらに、ターミネーター $T_1 \sim T_8$ を含む同じターミネーターラダーを利用して微生物を操作することができ、ここで、8つのターミネーターそれぞれは、10の異なる遺伝子標的に作動可能に連結している。この手順の結果は、目的の標的遺伝子に作動可能に連結した特定のターミネーターを除いて、他の点では遺伝的に同一と仮定される80の宿主細胞株である。これらの80の宿主細胞株は、適切にスクリーニングし、特徴付け、別のHTP遺伝子設計ライブラリーを生じさせることができる。HTP遺伝子設計ライブラリー内の微生物株を特徴付けると、リレーショナルデータベース、オブジェクト指向データベース、または大規模分散NoSQLデータベースを限定することなく含めた任意のデータベースに記憶することができる情報およびデータが生じる。このデータ／情報は、例えば、所与の遺伝子標的に作動可能に連結しているときの所与のターミネーター（例えば、 $T_1 \sim T_8$ ）の効果を含み得る。このデータ／情報は、2つまたはそれ超のプロモーター（例えば、 $T_1 \sim T_8$ ）を所与の遺伝子標的に作動可能に連結することから生じるより広いセットのコンビナトリアル効果でもあり得る。

【0288】

8のターミネーターおよび10の標的遺伝子の上述の例は、単に例示的であり、その理由は、この概念は、一連の発現強度および任意の所与の数の標的遺伝子の呈示に基づいて一緒にグループ化された任意の所与の数のターミネーターを用いて適用することができるためである。例えば、本明細書に提供される方法（例えば、STOPスワッピング）において使用することができるターミネーターの別のセットは、核酸配列番号225、226、227、228、229、または230を有する表1.2に見出されるターミネーターのセットである。

【0289】

要約すると、様々なターミネーターを利用して生物体内の様々な遺伝子の発現をモジュレートすることは、目的の形質を最適化する強力なツールである。本発明者らが開発したターミネーターズワッピングの分子ツールは、少なくとも1つの条件下で少なくとも1つの遺伝子座の発現を変更することが実証されたターミネーター配列のラダーを使用する。次いでこのラダーは、ハイスループットゲノム操作を使用して生物体内の遺伝子の群に系統的に適用される。この群の遺伝子は、いくつかの方法のいずれか1つに基づいて目的の形質に影響を与える高い見込みを有すると決定される。これらは、公知の機能に基づく選択、または目的の形質に対する影響、または先に決定された有益な遺伝的多様性に基づくアルゴリズムによる選択を含み得る。

【0290】

次いで、遺伝子に連結したターミネーター配列を含有する生物体の結果として生じるHTP遺伝子設計微生物ライブラリーが、ハイスループットスクリーニングモデルにおいて性能について評価され、性能を増大させるターミネーター遺伝子連鎖が決定され、情報がデータベースに記憶される。遺伝子攪乱（すなわち、所与の遺伝子yに連結した所与のターミネーターx）のコレクションは、「ターミネーターズワップライブラリー」を形成し、これは、微生物操作処理で利用される潜在的な遺伝子変更の源として利用することができる。経時的に、より大きいセットの遺伝子攪乱は、より大きい多様性の微生物バックグラウンドに対してインプリメントされるので、各ライブラリーは、実験的に確認されたデータのコーパスとしてより強力になり、それを使用して目的の任意のバックグラウンドに対する標的とした変化をより正確かつ予想通りに設計することができる。すなわち、一部の実施形態では、本開示は、本発明の遺伝子設計ライブラリーのいずれかと関連したメタデータ内に埋もれた先の実験結果に基づく、宿主細胞内への1つまたは複数の遺伝子変化の導入を教示する。

【0291】

よって、特定の実施形態では、ターミネータースワッピングは、以下を含むマルチステッププロセスである。

【0292】

1. 「ラダー」として作用する一連の「x」ターミネーターを選択するステップ。理想的にはこれらのターミネーターは、多重ゲノム遺伝子座全体にわたって高度に多様な発現をもたらすことが示されているが、唯一の要件は、これらが遺伝子発現を何らかのやり方で攪乱することである。

【0293】

2. 標的にする一連の「n」遺伝子を選択するステップ。このセットは、ゲノム内のあらゆるORF、またはORFのサブセットであり得る。サブセットは、先に実証された有益な攪乱（先のプロモータースワップ、STOPスワップ、SOLUBILITY TAGスワップ、DEGRADATION TAGスワップもしくはSNPスワップ）との関連によって、先に生成された攪乱間の上位性相互作用に基づくアルゴリズム選択、標的にする有益なORFに関する仮説に基づく他の選択判定基準によって、またはランダム選択を介して、機能に関連したORFについてのアノテーションを使用して選択することができる。他の実施形態では、「n」標的化遺伝子は、非コードRNAを含めた非タンパク質コード遺伝子を含み得る。

【0294】

3. 以下の遺伝子改変を迅速に、かつ並行して実行するためのハイスループット株操作：生来のターミネーターが標的遺伝子nの3'末端に存在し、その配列が公知であるとき、生来のターミネーターをラダー内のxターミネーターのそれぞれで置き換える。生来のターミネーターが存在せず、またはその配列が未知であるとき、ラダー内のxターミネーターのそれぞれを遺伝子停止コドンの後に挿入する。

【0295】

このように、株の「ライブラリー」（HTP遺伝子設計ライブラリーとも呼ばれる）が構築され、ライブラリーの各メンバーは、他の点では同一の遺伝子状況においてn標的に連結したxターミネーターの一事例である。先に記載したように、ターミネーターの組合せを挿入し、ライブラリーが構築される組み合わせの可能性の範囲を拡張することができる。

【0296】

4. 1つまたは複数の測定基準に対する株の性能が最適化されている性能を示す状況における株のライブラリーのハイスループットスクリーニング。

【0297】

この基本プロセスを拡張して、とりわけ、（1）相互作用プロセスにおいて1つずつ、または単一ステップにおける複数の変化として、単一株バックグラウンド内に複数の有益な攪乱を統合することによって、株性能のさらなる改善をもたらすことができる。複数の攪乱は、確定した変化の特定のセットまたは変化の部分的にランダム化されたコンビナトリアルライブラリーのいずれかであり得る。例えば、標的のセットが経路内のあらゆる遺伝子である場合、攪乱のライブラリーを株の先のライブラリーの改善された1つまたは複数のメンバー内に逐次再生すると、任意の所与の反復においてどの遺伝子が律速であるかにかかわらず経路内の各遺伝子の発現レベルを最適化することができ；（2）ライブラリーの個々のおよびコンビナトリアル生成の結果として生じる性能データを、そのデータを使用して各攪乱の相互作用に基づいて攪乱の最適のセットを予測するアルゴリズムに送り；（3）上記2つの手法の組合せをインプリメントする。

【0298】

手法は、工業用微生物を用いて本開示において例示されているが、所望の形質が遺伝子変異体の集団において同定され得る任意の生物体に適用可能である。例えば、これは、CHO細胞、酵母、昆虫細胞、藻類、および植物などの多細胞生物の性能を改善するために使用することができる。

【0299】

5. 配列最適化：最適化配列微生物株ライブラリーを得るための分子ツール

一実施形態では、開示の方法は、宿主生物が発現する1つまたは複数のコドン最適化遺伝子を含む。様々な宿主内での発現を改善するためにコドンを最適化するための方法は、当技術分野で公知であり、文献において記載されている（その全体が参照により本明細書に組み込まれている米国特許出願公開第2007/0292918号を参照）。特定の原核生物宿主または真核生物宿主に好適なコドンを含む最適化されたコード配列（Murrayら、（1989年）、Nucl. Acids Res.、17巻：477～508頁も参照）を調製して、例えば、翻訳の速度を増大させ、または非最適配列から産生される転写物と比較して、より長い半減期などの望ましい性質を有する組換えRNA転写物を生成することができる。

10

【0300】

タンパク質発現は、転写、mRNAプロセッシング、ならびに翻訳の安定性および開始に影響するものを含めた多数の因子によって支配される。よって、最適化は、任意の特定の遺伝子のいくつかの配列フィーチャのいずれかに対処することができる。具体例として、希少コドン誘導翻訳休止は、タンパク質発現を低減し得る。希少コドン誘導翻訳休止は、宿主生物において稀に使用される目的のポリヌクレオチド中のコドンの存在が、利用可能なtRNAプールにおけるその不足に起因してタンパク質翻訳に対して負の効果を有し得ることを含む。

【0301】

代替の翻訳開始も、異種タンパク質発現を低減し得る。代替翻訳開始は、リボソーム結合部位（RBS）として機能することができるモチーフを偶然に含有する合成ポリヌクレオチド配列を含み得る。これらの部位は、遺伝子内部部位から短縮タンパク質の翻訳を開始することができる。精製中に除去することが困難であり得る短縮タンパク質を産生する可能性を低減する一方法は、最適化ポリヌクレオチド配列から推定上の内部RBS配列を排除するステップを含む。

20

【0302】

リピート誘導ポリメラーゼのスリッページは、異種タンパク質発現を低減し得る。リピート誘導ポリメラーゼのスリッページは、フレームシフト変異をもたらし得るDNAポリメラーゼのスリッページまたはスタタリングを引き起こすことが示されているヌクレオチド配列リピートを伴う。このようなリピートは、RNAポリメラーゼのスリッページも引き起こし得る。高いG+C含有量偏りを有する生物体では、GまたはCヌクレオチドリピートから構成されるより高い程度のリピートが存在し得る。したがって、RNAポリメラーゼのスリッページを誘導する可能性を低減する一方法は、GまたはCヌクレオチドの長いリピートを変更するステップを含む。

30

【0303】

二次構造の干渉も異種タンパク質発現を低減し得る。二次構造は、RBS配列または開始コドンを隔離することができ、タンパク質発現の低減と相関している。ステムループ構造も、転写休止および減衰に関与することができる。最適化ポリヌクレオチド配列は、ヌクレオチド配列のRBSおよび遺伝子コード領域内に最小限の二次構造を含有して、転写および翻訳の改善を可能にすることができる。

40

【0304】

例えば、最適化プロセスは、宿主が発現する所望のアミノ酸配列を同定することによって始めることができる。アミノ酸配列から、候補ポリヌクレオチドまたはDNA配列を設計することができる。合成DNA配列の設計中に、コドン使用の頻度を、宿主発現生物体のコドン使用と比較することができ、希少な宿主コドンを合成配列から除去することができる。さらに、合成候補DNA配列を、望ましくない酵素制限部位を除去し、任意の所望のシグナル配列、リンカー、または非翻訳領域を付加または除去するために修飾することができる。合成DNA配列は、翻訳プロセスを妨害し得る二次構造、例えば、G/Cリピートおよびステムループ構造の存在について分析することができる。

【0305】

50

6. SOLUBILITY TAGスワップ：SOLUBILITY TAGスワップ 微生物株ライブラリーを得るための分子ツール

一部の実施形態では、本開示は、翻訳後機構の最適化によって宿主細胞生産性を改善する方法を教示する。従来の株改良は、多くの場合、目的のある分子を産生させる経路遺伝子の過剰発現によって達成することができる。典型的には、経路タンパク質存在量を増加させて所与の経路についての速度、力価、または収率の改善を達成する目的で、公知の経路遺伝子を複製させることもでき、強力なプロモーターを挿入してこれらの遺伝子の発現を駆動し、したがって、mRNA転写物レベルを上昇させることもできる。この手法を全ゲノム規模で系統的に適用して、株性能を改善することができるすべての遺伝子を同定することができる。別の頻繁に適用される手法は、所望の経路から炭素を転用し得るタンパク質産物を完全に排除する目的で、潜在的に競合する経路遺伝子を欠失させることであり得る。しかし、当技術分野で公知のこれらの過剰発現および／または欠失による株改良手法にはいくつかの制限があり得る。

【0306】

経路複製または強力なプロモーターの挿入から開始して、mRNA転写物レベルの上昇の予測される効果は、必ずしもタンパク質存在量の増加をもたらすものではない可能性がある。様々なタンパク質産物には、それらの産生において様々な律速ステップがあり得、また、この律速ステップはmRNA転写物レベルではない可能性がある。mRNA転写が律速ステップではないシナリオでは、翻訳後機構が全体的なタンパク質存在量に影響を及ぼし得る可能性がある。例えば、タンパク質可溶性タグの存在を使用して、標的分子の産生に寄与する可能性がある正確にフォールディングされた活性なタンパク質の存在量を増加させることができ、一方、mRNA転写物レベルを単に上昇させることでは、ミスフォールディングされた不活性なタンパク質の増加しか導かれない可能性がある。効果を発揮したタンパク質可溶性タグを使用される可溶性タグの配列に応じて調整可能になるように作製し、標的表現型の的確な最適化を可能にすることもできる。

【0307】

タンパク質可溶性タグスワッピング（SOLUBILITY TAGスワップ）

一部の実施形態では、本開示は、全体的な宿主株生産性に対して有益な効果を生じさせるために、最適なタンパク質可溶性を有するタンパク質可溶性タグ配列（「可溶性タグ」）を選択する方法を教示する。

【0308】

例えば、一部の実施形態では、本開示は、一連の可溶性強度を呈する、宿主細胞の1つもしくは複数のタンパク質可溶性タグを同定し、かつ／または1つもしくは複数のタンパク質可溶性タグのバリエーションを生成する（例えば、以下に論じられるタンパク質可溶性タグ）方法を教示する。これらの同定および／または生成されるタンパク質可溶性タグの特定の組合せを、以下でより詳細に説明されるタンパク質可溶性タグラダーとして一緒にグループ化することができる。

【0309】

次いで、問題のタンパク質可溶性タグラダーは、目的の所与の遺伝子と関連させられる。よって、タンパク質可溶性を増強し、また100アミノ酸よりも小さいことがCostaら、Front Microbiol.、2014年；5巻：63頁から同定されたタンパク質可溶性タグのサブセットを表すタンパク質可溶性タグPST₁～PST₄（表17参照）を有し、タンパク質可溶性タグラダーを宿主細胞内の目的の単一遺伝子と関連させる（すなわち、所与の標的遺伝子に作動可能に連結した所与のタンパク質可溶性タグで宿主細胞を遺伝子操作して、N末端またはC末端のいずれかにおいてタグ付けされた標的タンパク質を生成する場合、タンパク質可溶性タグの各組合せの効果を、操作された宿主細胞が標的遺伝子と関連した特定の可溶性タグ（複数可）を除いて他の点では同一の遺伝的バックグラウンドを有することを考慮して、各コンビナトリアル取り組みの結果として生じる操作された株のそれぞれを特徴付けることによって確認することができる。このプロセスを介して操作される結果として生じる宿主細胞は、HTP遺伝子設計ライブラリーを形成する。

【0310】

H T P 遺伝子設計ライブラリーは、このプロセスを介して形成される実際の物理的な微生物株コレクションを指すことができ、各メンバー株は、他の点では同一の遺伝的バックグラウンドにおいて、特定の標的タンパク質に作動可能に連結した所与のタンパク質可溶性タグを代表し、前記ライブラリーは、「可溶性タグスワップ微生物株ライブラリー」または「S O L U B I L I T Y T A G スワップ微生物株ライブラリー」と呼ばれる。E . c o l i の特定の状況では、ライブラリーを「S O L U B I L I T Y T A G スワップ E . c o l i 株ライブラリー」または「S O L U B I L I T Y T A G スワップ E . c o l i 株ライブラリー」と呼ぶことができるが、E . c o l i は微生物の具体的な例であるので、この用語は同義に使用することができる。

10

【0311】

さらに、H T P 遺伝子設計ライブラリーは、遺伝子攪乱のコレクションを指すことができ、この場合、所与のタンパク質可溶性タグ x は所与の遺伝子 y に作動可能に連結しており、前記コレクションは、「タンパク質可溶性タグスワップライブラリー」または「S O L U B I L I T Y T A G スワップライブラリー」と呼ばれる。

【0312】

さらに、タンパク質可溶性タグ P S T ₁ ~ P S T ₄ を含む同じタンパク質可溶性タグラダーを利用して微生物を操作することができ、ここで、4 つのタンパク質可溶性タグのそれぞれは、10 の異なる遺伝子標的に作動可能に連結される。この手順の結果は、目的の標的遺伝子に作動可能に連結した特定のタンパク質可溶性タグを除いて、他の点では遺伝的に同一と仮定される40の宿主細胞株である。これらの40の宿主細胞株は、適切にスクリーニングし、特徴付け、別のH T P 遺伝子設計ライブラリーを生じさせることができる。H T P 遺伝子設計ライブラリー内の微生物株を特徴付けると、リレーショナルデータベース、オブジェクト指向データベースまたは大規模分散N o S Q L データベースを限定することなく含めた任意のデータベースに記憶することができる情報およびデータが生じる。このデータ／情報は、例えば、所与の遺伝子標的に作動可能に連結しているときの所与のタンパク質可溶性タグ（例えば、P S T ₁ ~ P S T ₄）の効果を含み得る。このデータ／情報は、2つまたはそれよりも多くの可溶性タグ（例えば、P S T ₁ ~ P S T ₄）を所与の遺伝子標的に作動可能に連結することから生じるより広いセットのコンビナトリアル効果でもあり得る。

20

30

【0313】

4 つのタンパク質可溶性タグおよび10の標的遺伝子の上述の例は、単に例示的であり、その理由は、この概念は、一連の可溶性強度および任意の所与の数の標的遺伝子の呈示に基づいて一緒にグループ化された任意の所与の数のタンパク質可溶性タグを用いて適用することができるためである。

【0314】

要約すると、様々なタンパク質可溶性タグを利用して生物体内の様々なタンパク質の可溶性をモジュレートすることは、目的の形質を最適化するための強力なツールである。本発明者らが開発したタンパク質可溶性タグスワッピングの分子ツールは、少なくとも1つの条件下で少なくとも1つのタンパク質の溶解性を変動させる（例えば、増強する）ことが実証されたタンパク質可溶性タグ配列のラダーを使用する。次いでこのラダーは、ハイスループットゲノム操作を使用して生物体内の遺伝子の群に系統的に適用される。この群の遺伝子は、いくつかの方法のいずれか1つに基づいて目的の形質に影響を与える高い見込みを有すると決定される。これらは、目的の形質に対する公知の機能もしくは影響に基づく選択物、または先に決定された有益な遺伝的多様性に基づくアルゴリズム選択物を含み得る。

40

【0315】

次いで、遺伝子に連結したタンパク質可溶性タグ配列を含有する生物体の結果として生じるH T P 遺伝子設計微生物ライブラリーがハイスループットスクリーニングモデルにおいて性能について評価され、性能を増大させるタンパク質可溶性タグー遺伝子連鎖が決定

50

され、情報がデータベースに記憶される。遺伝子攪乱（すなわち、所与の遺伝子 y に連結した所与のタンパク質可溶性タグ x ）のコレクションは、「タンパク質可溶性タグスワップライブラリー」を形成し、これは、微生物操作処理において利用される潜在的な遺伝子変更の源として利用することができる。経時的に、より大きいセットの遺伝子攪乱は、より多様な微生物バックグラウンドに対してインプリメントされるので、各ライブラリーは、実験的に確認されたデータのコーパスとしてより強力になり、それを使用して目的の任意のバックグラウンドに対する標的とした変化をより正確かつ予想通りに設計することができる。すなわち、一部の実施形態では、本開示は、本発明の遺伝子設計ライブラリーのいずれかと関連したメタデータ内に組み込まれた先の実験結果に基づく、宿主細胞内への1つまたは複数の遺伝子変化の導入を教示する。

10

【0316】

したがって、特定の実施形態では、タンパク質可溶性タグスワッピングは、以下を含むマルチステッププロセスである。

【0317】

1. 「ラダー」として作用する一連の「 x 」タンパク質可溶性タグを選択するステップ。理想的にはこれらのタンパク質可溶性タグは、多重ゲノム遺伝子座全体にわたってタンパク質溶解性の増強をもたらすことが示されているが、唯一の要件は、これらが溶解性を何らかのやり方で攪乱することである。

【0318】

2. 標的にする一連の「 n 」遺伝子を選択するステップ。このセットは、ゲノム内のあらゆるORF、またはORFのサブセットであり得る。サブセットは、先に実証された有益な攪乱（先のプロモータースワップ、STOPスワップ、DEGRADATION TAGスワップまたはSNPスワップ）との関連によって、先に生成された攪乱間の上位性相互作用に基づくアルゴリズム選択、標的にする有益なORFに関する仮説に基づく他の選択の判定基準によって、またはランダム選択を介して、機能に関連したORFについてのアノテーションを使用して選択することができる。

20

【0319】

3. 以下の遺伝子改変を迅速に、かつ並行して実行するためのハイスループット株操作：生来のタンパク質可溶性タグが標的遺伝子 n 内に存在し、その配列が公知であるとき、生来のタンパク質可溶性タグをラダー内の x タンパク質可溶性タグのそれぞれで置き換える。生来のタンパク質可溶性タグが存在しない、またはその配列が未知であるとき、ラダー内の x タンパク質可溶性タグのそれぞれを挿入する。

30

【0320】

このように、株の「ライブラリー」（HTP遺伝子設計ライブラリーとも呼ばれる）が構築され、ライブラリーの各メンバーは、他の点では同一の遺伝子状況において n 標的に連結した x タンパク質可溶性タグの一事例である。先に記載したように、タンパク質可溶性タグの組合せを挿入し、ライブラリーが構築される組み合わせの可能性の範囲を拡張することができる。

【0321】

4. 1つまたは複数の測定基準に対する株の性能が最適化されている性能を示す状況における株のライブラリーのハイスループットスクリーニング。

40

【0322】

この基本プロセスを拡張して、とりわけ、（1）相互作用プロセスにおいて1つずつ、または単一ステップにおける複数の変化としてのいずれかで、単一株バックグラウンド内に複数の有益な攪乱を統合することによって、株性能をさらに改善することができる。複数の攪乱は、確定した変化の特定のセット、または変化の部分的にランダム化されたコンビナトリアルライブラリーのいずれかであり得る。例えば、標的のセットが経路内のあらゆる遺伝子である場合、攪乱のライブラリーを株の先のライブラリーの改善された1つまたは複数のメンバー内に逐次再生すると、任意の所与の反復においてどの遺伝子が律速であるかにかかわらず経路内の各遺伝子の発現レベルを最適化することができ；（2）ライ

50

ブラリーの個々のおよびコンビナトリアル生成の結果として生じる性能データを、そのデータを使用して各攪乱の相互作用に基づいて攪乱の最適のセットを予測するアルゴリズムに送り；（３）上記２つの手法の組合せをインプリメントする。

【０３２３】

手法は、工業用微生物を用いて本開示において例示されているが、所望の形質が遺伝子変異体の集団において同定され得る任意の生物体に適用可能である。例えば、これは、CHO細胞、酵母、昆虫細胞、藻類、および植物などの多細胞生物の性能を改善するのに使用することができる。

【０３２４】

７．DEGRADATION TAGスワップ：DEGRADATION TAGスワップ微生物株ライブラリーを得るための分子ツール

10

翻訳後機構の最適化を通じて宿主細胞生産性を改善するための方法に関する上記の実施形態に付け加えて、遺伝子欠失のストラテジーにはまた、本発明のタンパク質分解タグ（ならびにターミネーターおよびタンパク質可溶性タグ）によって対処することができる欠点もあり得る。遺伝子およびその対応するタンパク質産物の大規模な欠失により、一部の場では、細胞に激しい変化が課される。種々の速度でタンパク質を分解の標的にするタンパク質分解タグのライブラリーを通じて、よりの確かつ調整可能な応答を達成することができる。この手法には、細胞生存に必須であり得、完全に欠失すると生存可能でないとされるタンパク質産物のモジュレーションを可能にするという利益もあり得る。これらの分解タグは翻訳後レベルでも機能するので、上記の通りmRNA転写物レベルの変更によりタンパク質レベルの変更がもたらされないシナリオに対処することができ得る。

20

【０３２５】

タンパク質分解タグスワッピング（DEGRADATION TAGスワップ）

一部の実施形態では、本開示は、全体的な宿主株生産性に対して有益な効果を生じさせるために、最適なタンパク質分解性またはタンパク質レベルモジュレーション性を有するタンパク質分解タグ配列（「分解タグ」）を選択する方法を教示する。

【０３２６】

例えば、一部の実施形態では、本開示は、一連の分解強度を呈するまたは標的タンパク質のレベルをモジュレートする、宿主細胞内の１つもしくは複数のタンパク質分解タグを同定し、かつ／または１つもしくは複数のタンパク質分解タグのバリエーションを生成する（例えば、以下に論じられるタンパク質分解タグ）方法を教示する。これらの同定および／または生成されるタンパク質分解タグの特定の組合せを、以下でより詳細に説明されるタンパク質分解タグラダーとして一緒にグループ化することができる。

30

【０３２７】

次いで、問題のタンパク質分解タグラダーは、目的の所与の遺伝子と関連させられる。したがって、表１８において詳述されている通り様々な源から同定されたタンパク質分解タグのサブセットを表すタンパク質分解タグPDT₁～PDT₈。（表１８参照）を有し、タンパク質分解タグラダーを宿主細胞内の目的の単一遺伝子と関連させる（すなわち、所与の標的遺伝子に作動可能に連結した所与のタンパク質分解タグで宿主細胞を遺伝子操作する場合、タンパク質分解タグの各組合せの効果を、操作された宿主細胞が標的遺伝子と関連した特定の分解タグ（複数可）を除いて他の点では同一の遺伝的背景を有することを考慮して、各コンビナトリアル取り組みの結果として生じる操作された株のそれぞれを特徴付けることによって確認することができる。このプロセスを介して操作される結果として生じる宿主細胞は、HTP遺伝子設計ライブラリーを形成する。

40

【０３２８】

HTP遺伝子設計ライブラリーは、このプロセスを介して形成される実際の物理的な微生物株コレクションを指すことができ、各メンバー株は、他の点では同一の遺伝的背景において、特定の標的タンパク質に作動可能に連結した所与のタンパク質分解タグを代表し、前記ライブラリーは、「分解タグスワップ微生物株ライブラリー」または「DEGRADATION TAGスワップ微生物株ライブラリー」と呼ばれる。E. c

50

o l i の特定の状況では、ライブラリーを「D E G R A D A T I O N T A G スワップ E . c o l i 株ライブラリー」または「D E G R A D A T I O N T A G スワップ E . c o l i 株ライブラリー」と呼ぶことができるが、E . c o l i は微生物の具体的な例であるので、この用語は同義に使用することができる。

【0329】

さらに、H T P 遺伝子設計ライブラリーは、遺伝子攪乱のコレクションを指すことができ、この場合、所与のタンパク質分解タグ x は所与の遺伝子 y に作動可能に連結しており、前記コレクションは、「タンパク質分解タグスワップライブラリー」または「D E G R A D A T I O N T A G スワップライブラリー」と呼ばれる。

【0330】

さらに、タンパク質分解タグ P D T₁ ~ P D T₈ を含む同じタンパク質分解タグラダーを利用して微生物を操作することができ、ここで、8つのタンパク質分解タグのそれぞれは、10の異なる遺伝子標的に作動可能に連結される。この手順の結果は、目的の標的遺伝子に作動可能に連結した特定のタンパク質分解タグを除いて、他の点では遺伝的に同一と仮定される80の宿主細胞株である。これらの80の宿主細胞株は、適切にスクリーニングし、特徴付け、別のH T P 遺伝子設計ライブラリーを生じさせることができる。H T P 遺伝子設計ライブラリー内の微生物株を特徴付けると、リレーショナルデータベース、オブジェクト指向データベースまたは大規模分散N o S Q L データベースを限定することなく含めた任意のデータベースに記憶することができる情報およびデータが生じる。このデータ／情報は、例えば、所与の遺伝子標的に作動可能に連結しているときの所与のタンパク質分解タグ（例えば、P D T₁ ~ P D T₈）の効果を含み得る。このデータ／情報は、2つまたはそれよりも多くの分解タグ（例えば、P D T₁ ~ P D T₈）を所与の遺伝子標的に作動可能に連結することから生じるより広いセットのコンビナトリアル効果でもあり得る。

【0331】

8つのタンパク質分解タグおよび10の標的遺伝子の上述の例は、単に例示的であり、その理由は、この概念は、一連の分解強度および任意の所与の数の標的遺伝子の呈示に基づいて一緒にグループ化された任意の所与の数のタンパク質分解タグを用いて適用することができるためである。

【0332】

要約すると、様々なタンパク質分解タグを利用して生物体内の様々なタンパク質の分解をモジュレートすることは、目的の形質を最適化するための強力なツールである。本発明者らが開発したタンパク質分解タグスワッピングの分子ツールは、少なくとも1つの条件下で少なくとも1つのタンパク質の分解を変動させる（例えば、増強する）ことが実証されたタンパク質分解タグ配列のラダーを使用する。次いでこのラダーは、ハイスループットゲノム操作を使用して生物体内の遺伝子の群に系統的に適用される。この群の遺伝子は、いくつかの方法のいずれか1つに基づいて目的の形質に影響を与える高い見込みを有すると決定される。これらは、目的の形質に対する公知の機能もしくは影響に基づく選択物、または先に決定された有益な遺伝的多様性に基づくアルゴリズム選択物を含み得る。

【0333】

次いで、遺伝子に連結したタンパク質分解タグ配列を含有する生物体の結果として生じるH T P 遺伝子設計微生物ライブラリーが、ハイスループットスクリーニングモデルにおいて性能について評価され、性能を増大させるタンパク質分解タグー遺伝子連鎖が決定され、情報がデータベースに記憶される。遺伝子攪乱（すなわち、所与の遺伝子 y に連結した所与のタンパク質分解タグ x）のコレクションは、「タンパク質分解タグスワップライブラリー」を形成し、これは、微生物操作処理において利用される潜在的な遺伝子変更の源として利用することができる。経時的に、より大きいセットの遺伝子攪乱は、より多様な微生物バックグラウンドに対してインプリメントされるので、各ライブラリーは、実験的に確認されたデータのコーパスとしてより強力になり、それを使用して目的の任意のバックグラウンドに対する標的とした変化をより正確かつ予想通りに設計することができる

。すなわち、一部の実施形態では、本開示は、本発明の遺伝子設計ライブラリーのいずれかに関連したメタデータ内に組み込まれた先の実験結果に基づく、宿主細胞内への1つまたは複数の遺伝子変化の導入を教示する。

【0334】

したがって、特定の実施形態では、タンパク質分解タグスワッピングは、以下を含むマルチステッププロセスである。

【0335】

1. 「ラダー」として作用する一連の「x」タンパク質分解タグを選択するステップ。理想的にはこれらのタンパク質分解タグは、多重ゲノム遺伝子座全体にわたってタンパク質分解の増強をもたらすことが示されているが、唯一の要件は、これらが分解を何らかのやり方で攪乱することである。

10

【0336】

2. 標的にする一連の「n」遺伝子を選択するステップ。このセットは、ゲノム内のあらゆるORF、またはORFのサブセットであり得る。サブセットは、先に実証された有益な攪乱（先のプロモータースワップ、STOPスワップ、SOLUBILITY TAGスワップまたはSNPスワップ）との関連によって、先に生成された攪乱間の上位性相互作用に基づくアルゴリズム選択、標的にする有益なORFに関する仮説に基づく他の選択の判定基準によって、またはランダム選択を介して、機能に関連したORFについてのアノテーションを使用して選択することができる。

20

【0337】

3. 以下の遺伝子改変を迅速に、かつ並行して実行するためのハイスループット株操作：生来のタンパク質分解タグが標的遺伝子n内に存在し、その配列が公知であるとき、生来のタンパク質分解タグをラダー内のxタンパク質分解タグのそれぞれで置き換える。生来のタンパク質分解タグが存在しない、またはその配列が未知であるとき、ラダー内のxタンパク質分解タグのそれぞれを挿入する。

【0338】

このように、株の「ライブラリー」（HTP遺伝子設計ライブラリーとも呼ばれる）が構築され、ライブラリーの各メンバーは、他の点では同一の遺伝子状況においてn標的に連結したxタンパク質分解タグの一事例である。先に記載したように、タンパク質分解タグの組合せを挿入し、ライブラリーが構築される組み合わせの可能性の範囲を拡張することができる。

30

【0339】

4. 1つまたは複数の測定基準に対する株の性能が最適化されている性能を示す状況における株のライブラリーのハイスループットスクリーニング。

【0340】

この基本プロセスを拡張して、とりわけ、（1）相互作用プロセスにおいて1つずつ、または単一ステップにおける複数の変化として、単一株バックグラウンド内に複数の有益な攪乱を統合することによって、株性能のさらなる改善をもたらすことができる。複数の攪乱は、確定した変化の特定のセットまたは変化の部分的にランダム化されたコンビナトリアルライブラリーのいずれかであり得る。例えば、標的のセットが経路内のあらゆる遺伝子である場合、攪乱のライブラリーを株の先のライブラリーの改善された1つまたは複数のメンバー内に逐次再生すると、任意の所与の反復においてどの遺伝子が律速であるかにかかわらず経路内の各遺伝子の発現レベルを最適化することができ；（2）ライブラリーの個々のおよびコンビナトリアル生成の結果として生じる性能データを、そのデータを使用して各攪乱の相互作用に基づいて攪乱の最適のセットを予測するアルゴリズムに送り；（3）上記2つの手法の組合せをインプリメントする。

40

【0341】

手法は、工業用微生物を用いて本開示において例示されているが、所望の形質が遺伝子変異体の集団において同定され得る任意の生物体に適用可能である。例えば、これは、CHO細胞、酵母、昆虫細胞、藻類、および植物などの多細胞生物の性能を改善するために

50

使用することができる。

【0342】

8. 上位性マッピング — 有益な遺伝子統合 (genetic consolidation) を可能にする予測的分析用ツール

一部の実施形態では、本開示は、宿主細胞内に有益な遺伝子変更を予測し、組み合わせるための上位性マッピング法 (epistasis mapping method) を教示する。遺伝子変更は、上述のHTP分子ツールセット (例えば、プロモータースワップ、SNPスワップ、開始/停止コドン交換、配列最適化、タンパク質可溶性タグスワップ、タンパク質分解タグスワップおよびSTOPスワップ) のいずれかによって生み出すことができ、これらの遺伝子変更の効果は、得られたHTP遺伝子設計微生物株ライブラリーの特徴付けから分かる。よって、本明細書で使用される場合、用語の上位性マッピングは、宿主性能の増大を生じさせる可能性のある遺伝子変更の組合せ (例えば、有益なSNPまたは有益なプロモーター/標的遺伝子の関連付け) を同定する方法を含む。

10

【0343】

諸実施形態では、本開示の上位性マッピング法は、2つの異なる機能群からの有益な変異の組合せは、同じ機能群からの変異の組合せと比較して、宿主性能を改善する可能性が高いというアイデアに基づく。例えば、Costanzo, The Genetic Landscape of a Cell, Science, 327巻、5964号、2010年1月22日、425~431頁を参照 (その全体が本明細書に参照により組み込まれている)。

【0344】

同じ機能群からの変異は、同じ機構によって働く可能性が高く、よって、全体的な宿主性能に対して負または中立の上位性を呈する可能性が高い。対照的に、異なる機能群からの変異は、独立した機構によって働く可能性が高く、それは、宿主性能の改善、および一部の事例では相乗効果をもたらし得る。例えば、図19を参照すると、lysAおよびzwfは、異なる経路で働いてリシンの産生を達成する遺伝子である。これらの遺伝子の個々の性能の相違点に基づく、これらの遺伝子を使用する遺伝子変化は、相加的な統合効果をもたらすはずである。これは、図16Bおよび実施例6に示した通り、lysAおよびzwfの組合せの統合効果を実際に測定して裏付けられた。

20

【0345】

よって、一部の実施形態では、本開示は、SNP変異を分析して、異なる機能群に属すると予測されるSNPを同定する方法を教示する。一部の実施形態では、SNP機能群類似度は、変異相互作用プロファイルの余弦類似度 (相関係数と類似、図16Aを参照) を計算することによって決定される。本開示は、変異類似度マトリックス (図15、例えば、Corynebacteriumにおいて行われた分析を参照) またはデンドログラム (図16A、例えば、Corynebacteriumにおいて行われた分析を参照) を介してSNPを比較することも例示する。

30

【0346】

よって、上位性マッピング手順は、1つまたは複数の遺伝的バックグラウンドにおいて適用される遺伝子変異の多様性を、1つまたは複数の遺伝的バックグラウンド内に前記変異を効率的かつ有効に統合する目的で、グループ化および/またはランク付けするための方法を提供する。

40

【0347】

諸態様では、統合は、標的生体分子を産生させるのに最適化された新規株を生み出す目的で実施される。教示した上位性マッピング手順によって、変異の機能分類を同定することが可能であり、このような機能分類は、望ましくない上位性効果を最小限にする統合ストラテジーを可能にする。

【0348】

先に説明した通り、工業発酵で使用するための微生物の最適化は、経済、社会、および自然界に関する広い暗示を含む重要で困難な問題である。伝統的に、微生物操作は、ランダム変異誘発の遅く不確かなプロセスによって実施されてきた。このような手法は、細胞

50

の自然進化能力を活用して、人工的に課された選択圧に適応させる。このような手法はまた、有益な変異の珍しさ、基礎をなす適応度地形の荒削りさによって制限されており、より一般に、細胞生物学および分子生物学における現況技術を十分に利用しない。

【0349】

最新の手法は、機構レベルでの細胞機能の新しい理解および特定の表現型目標への標的化遺伝子マニピュレーションを実施する新しい分子生物学ツールを活用する。実際には、このような合理的手法は、生物学の根本的な複雑性によって混乱する。原因となる機構は、特に、それぞれに有益な効果が観察された2つまたはそれ超の変化を組み合わせようと試みる場合、理解が不十分である。時に、遺伝子変化のこのような統合は、正の結果（所望の表現型活性の増大によって測定される）を生じるが、正味の正の結果は、予期されたものより低くなり得、一部の 경우에는、予期されたものより高くなり得る。他の事例では、このような組合せは、正味の中立効果または正味の負の効果を生む。この現象は、上位性と呼ばれ、微生物操作（および一般に遺伝子操作）の基本的な課題の1つである。

10

【0350】

上述した通り、本HTPゲノム操作プラットフォームは、伝統的な微生物操作手法に関連した問題の多くを解決する。本HTPプラットフォームは、自動化技術を使用して、数百または数千の遺伝子変異を一度に実施する。特定の態様では、上述した合理的手法と異なり、開示したHTPプラットフォームは、その全体が本明細書に参照により組み込まれている、Microbial Strain Design System And Methods For Improved Large-Scale Production Of Engineered Nucleotide Sequencesという表題の米国出願第15/140,296号に開示されている通り、数千の変異体の並列構築を可能にして、関連したゲノム空間の大きいサブセットをより有効に探索する。「あらゆること」を試すことによって、本HTPプラットフォームは、本発明者らの限られた生物学的理解によって誘導される困難を回避する。

20

【0351】

しかし同時に、本HTPプラットフォームは、ゲノム空間の組合せによる爆発的サイズ、および遺伝的相互作用の複雑性を考慮して生成されたデータセットを解明する計算手法の有効度によって基本的に制限されるという問題に直面する。所望の結果を生じる組合せの非ランダム選択を最大にするやり方で広大なコンビナトリアル空間のサブセットを探索する技法が必要とされる。

30

【0352】

いくぶん類似のHTP手法が、酵素最適化の場合において有効であると証明された。このニッチな問題において、目的のゲノム配列（1000塩基の程度の）は、物理的構成が幾分複雑なタンパク質鎖をコードする。正確な構成は、その構成要素の原子成分間の集団電磁相互作用によって決定される。短いゲノム配列および物理的に制約された折り畳み問題のこの組合せは、貪欲最適化戦略に具体的に役に立つ。すなわち、あらゆる残基において配列を個々に変異させ、結果として生じる変異体をシャッフルして、配列活性応答モデル化（Sequence Activity Response modeling）に適合する分解能で局所的配列空間を有効にサンプリングすることが可能である。

40

【0353】

しかし、生体分子の完全ゲノム最適化に関して、このような残基中心の手法は、いくつかの重要な理由で不十分である。第1に、生体分子のゲノム最適化に伴う関連した配列空間の指数関数的増大のため。第2に、生体分子合成における調節、発現、および代謝相互作用の複雑性の追加のため。本発明者らは、教示される上位性マッピング手順を介してこれらの問題を解決した。

【0354】

変異のコレクション間の上位性相互作用を、前記変異を1つまたは複数の遺伝的バックグラウンド内により効率的かつ有効に統合する目的で、モデル化するための教示される方法は、当技術分野において革新的であり、高度に必要とされる。

【0355】

50

上位性マッピング手順を記述する場合、用語「より効率的な」および「より有効な」は、特定の表現型の目的に関して統合株の中での望ましくない上位性相互作用の回避を指す。

【0356】

プロセスを上記で一般に詳しく述べてきたので、より具体的なワークフロー例を次に記載する。

【0357】

第1に、M個の変異のライブラリーおよび1つまたは複数の遺伝的バックグラウンド（例えば、親菌株）から始める。ライブラリーの選択も遺伝的バックグラウンドの選択も、ここに記載した方法に特異的でない。しかし特定のインプリメンテーションでは、変異のライブラリーは、もっぱら、または組合せで：SNPスワップライブラリー、プロモータースワップライブラリー、または本明細書に記載の任意の他の変異ライブラリーを含み得る。

【0358】

一インプリメンテーションでは、単一の遺伝的バックグラウンドのみが提供される。この場合、別個の遺伝的バックグラウンド（微生物変異体）のコレクションが、この単一のバックグラウンドから最初に生成される。これは、変異の一次ライブラリー（またはそのあるサブセット）を所与のバックグラウンドに適用して、例えば、特定のSNPのHTP遺伝子設計ライブラリーまたは特定のプロモーターのHTP遺伝子設計ライブラリーを所与の遺伝的バックグラウンドに適用して、それに組み入れられている所与のHTP遺伝子設計ライブラリーからの特定の遺伝子変更を除いて同一の遺伝的バックグラウンドを有する微生物変異体の集団（おそらく100または1,000）を生み出すことによって達成され得る。以下に詳述する通り、この実施形態は、コンビナトリアルライブラリーまたはペアワイズライブラリーをもたらし得る。

【0359】

別のインプリメンテーションでは、別個の公知の遺伝的バックグラウンドのコレクションを単に与えることができる。以下に詳述する通り、この実施形態は、コンビナトリアルライブラリーのサブセットをもたらし得る。

【0360】

特定のインプリメンテーションでは、遺伝的バックグラウンドの数およびこれらのバックグラウンド間の遺伝的多様性（変異の数または配列編集距離などで測定した）は、この方法の有効度を最大にするように決定される。

【0361】

遺伝的バックグラウンドは、天然の、生来の、もしくは野生型株、または変異された、操作された株であり得る。N個の別個のバックグラウンド株は、ベクトルbで表すことができる。一例では、バックグラウンドbは、N個の一次変異 $m_0 = (m_1, m_2, \dots, m_N)$ を野生型バックグラウンド株 b_0 に適用して、N個の変異されたバックグラウンド株 $b = m_0 b_0 = (m_1 b_0, m_2 b_0, \dots, m_N b_0)$ （式中、 $m_i b_0$ は、変異 m_i のバックグラウンド株 b_0 への適用を表す）を形成することによって形成される操作されたバックグラウンドを表すことができる。

【0362】

いずれの場合にも（すなわち、単一の提供された遺伝的バックグラウンドまたは遺伝的バックグラウンドのコレクション）、結果は、N個の遺伝的に別個のバックグラウンドのコレクションである。関連した表現型が各バックグラウンドについて測定される。

【0363】

第2に、M個の変異のコレクション m_1 内の各変異が、N個のバックグラウンド株bのコレクション内の各バックグラウンドに適用されて、 $M \times N$ 個の変異体のコレクションが形成される。N個のバックグラウンドが変異の一次セット m_0 （上述した）を適用することによってそれ自体得られたインプリメンテーションでは、変異体の結果として生じるセットは、時に、コンビナトリアルライブラリーまたはペアワイズライブラリーと呼ばれる

。公知のバックグラウンドのコレクションが明示的に提供された別のインプリメンテーションでは、変異体の結果として生じるセットは、コンビナトリアルライブラリーのサブセットと呼ぶことができる。操作されたバックグラウンドベクトルの生成と同様に、諸実施形態では、入力インターフェース 202（図 31 を参照のこと）は、変異ベクトル m_1 およびバックグラウンドベクトル b 、ならびに外積などの指定されたオペレーションを受け取る。

【0364】

上記操作されたバックグラウンド例を続けて、 $M \times N$ コンビナトリアルライブラリーの形成は、 $m_1 \times m_0$ b_0 、 $b = m_0$ b_0 の N 個のバックグラウンドに適用された m_1 の外積によって形成された行列によって表すことができ、この場合、 m_1 中の各変異は、 b 内の各バックグラウンド株に適用される。結果として生じる $M \times N$ 行列の各 i 番目の行は、バックグラウンドコレクション b 内のすべての株への m_1 内の i 番目の変異の適用を表す。一実施形態では、 $m_1 = m_0$ であり、行列は、出発株 b_0 への同じ変異のペアワイズ適用を表す。その場合、行列は、その対角線回りで対称であり（ $M = N$ ）、対角線は、任意の分析において無視することができ、その理由は、それが同じ変異の 2 回の適用を表すためである。

【0365】

諸実施形態では、 $M \times N$ 行列の形成は、入力インターフェース 202（図 31 を参照のこと）に化合物発現 $m_1 \times m_0$ b_0 を入力することによって達成することができる。発現のコンポーネントベクトルは、直接入力することができ、これらの要素は、1 つもしくは複数の DNA 仕様を介して、またはインタープリター 204 によるインタープリテーション中にベクトルの検索を可能にするライブラリー 206 への呼出しとして明示的に指定される。「Microbial Strain Design System and Methods for Improved Large Scale Production of Engineered Nucleotide Sequences」という表題の米国特許出願第 15/140,296 号に記載されている通り、インタープリター 204、実行エンジン 207、発注エンジン（order placement engine）208、およびファクトリー 210 を介して、LIMS システム 200 は、入力された発現によって指定された微生物株を生成する。

【0366】

第 3 に、図 42 を参照すると、分析装置 214（図 31 を参照のこと）は、 $M \times N$ コンビナトリアルライブラリー行列（4202）内の各変異体について表現型応答を測定する。したがって、応答のコレクションは、 $M \times N$ 応答行列 R と解釈することができる。 R の各要素は、 $r_{ij} = y(m_i, m_j)$ として表すことができ、式中、 y は、変異 m_i によって変異された、操作されたコレクション b 内のバックグラウンド株 b_j の応答（性能）を表す。単純性および実用性のために、本発明者らは、ペアワイズ変異を仮定し、この場合、 $m_1 = m_0$ である。今回の場合のように、変異のセットがペアワイズ変異ライブラリーを表す場合、結果として生じる行列は、遺伝子相互作用行列、またはより特定すれば、変異相互作用行列とも呼ばれる場合がある。

【0367】

当業者は、一部の実施形態では、上位性効果および予測的株設計に関するオペレーションは、LIMS システム 200 の自動手段を通じて、例えば、分析装置 214（図 31 を参照のこと）によって、または人によるインプリメンテーションによって、または自動化手段およびマニュアル手段の組合せを通じて完全に実施することができることを認識する。オペレーションが完全に自動化されていない場合、LIMS システム 200 のエレメント、例えば、分析装置 214 は、例えば、その独自のオペレーション可能な能力を通じた結果を生成するのではなく、オペレーションを人が遂行した結果を受け取ることができる。本明細書の他の場所で記載されている通り、分析装置 214 などの LIMS システム 200 のコンポーネントは、1 つまたは複数のコンピューターシステムによって完全または部分的にインプリメントすることができる。一部の実施形態では、特に、予測的株設計に関するオペレーションが自動化手段およびマニュアル手段の組合せによって実施される場

合、分析装置 2 1 4 は、コンピューターハードウェア、ソフトウェア、もしくはファームウェア（またはこれらの組合せ）だけでなく、以下の表 5 に列挙されたもの、例えば、「性能を評価する」のカテゴリー下に列挙された装置などの人のオペレーターが運転する装置も含み得る。

【0368】

第 4 に、分析装置 2 1 2（図 3 1 を参照のこと）は、応答行列を正規化する。正規化は、バイアスを除去し、かつ／またはこの方法に特異的な効果の関連した部分を単離する目的で測定された応答値を調整するマニュアルおよび／またはこの実施形態では、自動化プロセスからなる。図 4 2 に関して、第 1 のステップ 4 2 0 2 は、正規化された測定データを得ることを含み得る。一般に、予測的株設計および上位性マッピングを対象とする請求項において、用語「性能尺度」または「測定された性能」などは、未処理であっても、何らかの様式で処理された、例えば、正規化データであっても測定データを反映する測定基準を記述するのに使用され得る。特定のインプリメンテーションでは、正規化は、測定された応答値から先に測定されたバックグラウンド応答を減じることによって実施することができる。そのインプリメンテーションでは、結果として生じる応答要素は、 $r_{ij} = y(m_i, m_j) - y(m_j)$ として形成することができ、式中、 $y(m_j)$ は、一次変異 m_j を親株 b_0 に適用することによって引き起こされた操作されたコレクション b 内の操作されたバックグラウンド株 b_j の応答である。正規化された応答行列の各行は、その対応する変異の応答プロファイルとして扱われることに留意されたい。すなわち、 i 番目の行は、 $j = 1 \sim N$ についてすべてのバックグラウンド株 b_j に適用された対応する変異 m_i の相対的效果を記述する。

【0369】

ペアワイズ変異の例に関して、2 つの変異の結果として生じる株の組み合わせられた性能／応答は、個々に、変異のそれぞれに対する株の性能／応答より大きい、小さい、またはそれに等しい場合がある。この効果は、「上位性」として公知であり、一部の実施形態では、 $e_{ij} = y(m_i, m_j) - (y(m_i) + y(m_j))$ として表すことができる。この数学的表現の変形は、可能であり、例えば、どのように個々の変化が生物学的に相互作用するかに依存し得る。上述したように、同じ機能群からの変異は、同じ機構によって働く可能性が高く、よって、全体的な宿主性能に対して負または中立の上位性を呈する可能性が高い。対照的に、異なる機能群からの変異は、独立した機構によって働く可能性が高く、それは、例えば、重複する変異効果を低減することによって宿主性能の改善をもたらし得る。よって、同じでない応答を生じる変異は、同様の応答を生じる変異より相加的な様式で組み合わせる可能性が高い。これは、次ステップにおける類似度の計算につながる。

【0370】

第 5 に、分析装置 2 1 4 は、応答の中の類似度、ペアワイズ変異例では、応答行列（4 2 0 4）内の i 番目の変異の効果と j 番目（例えば、一次）変異の効果との間の類似度を測定する。 R の i 番目の行は、 N 個のバックグラウンド株上の i 番目の変異 m_i の性能効果を表し、そのそれぞれは、それ自体上述した操作された変異の結果であり得ることを想起されたい。よって、 i 番目の変異の効果と j 番目の変異の効果との間の類似度は、 i 番目の行と j 番目の行との間の類似度 s_{ij} 、それぞれ ρ_i および ρ_j によって表されて類似度行列 S を形成することができ、その一例は、図 1 5 に例示されている。類似度は、多くの公知の技法、例えば、相互相関または絶対余弦類似度、例えば、 $s_{ij} = \cos(\cos(\rho_i, \rho_j))$ を使用して測定され得る。

【0371】

余弦類似度のような測定基準の代替または補足として、応答プロファイルをクラスター化して類似度の程度を決定することができる。クラスターリングは、適当な距離尺度（例えば、Euclidean、Hamming など）と併せて距離に基づくクラスターリングアルゴリズム（例えば、 k -平均、階層的凝縮型など）を使用することによって実施することができる。代替として、クラスターリングは、適当な類似度尺度（例えば、余弦、相関な

ど)とともに類似度に基づくクラスタリングアルゴリズム(例えば、スペクトル、ミンカットなど)を使用して実施することができる。もちろん、任意の数の標準的な機能的なオペレーション(例えば、指数関数)を介して類似度尺度にマッピングすることができ、逆もまた同様である。一インプリメンテーションでは、階層的凝縮型クラスタリングを絶対余弦類似度と併せて使用することができる(図16A、例えば、Corynebacteriumにおいて行った分析を参照)。

【0372】

クラスタリングの一例として、Cを変異 m_i のk個の別個クラスターへのクラスタリングとする。Cを、 c_{ij} が、変異iがクラスターjに属する程度、0から1の値であるクラスターメンバーシップ行列とする。そのとき、変異iとjとの間のクラスターに基づく類似度は、 $C_i \times C_j$ (Cのi番目の行とj番目の行とのドット積)によって与えられる。一般に、クラスターに基づく類似度行列は、 CC^T (すなわち、 $C \times C$ —転置行列)によって与えられる。ハードクラスタリング(変異が正確に1つのクラスターに属する)の場合では、2つの変異間の類似度は、これらが同じクラスターに属する場合、1であり、そうでない場合、0である。

【0373】

Costanzo, The Genetic Landscape of a Cell, Science, 327巻、5964号、2010年1月22日、425~431頁(その全体が本明細書に参照により組み込まれている)に記載されている通り、変異応答プロファイルのこのようなクラスタリングは、細胞の根本的な機能的な機構のおおよそのマッピングに関係する。すなわち、一緒にクラスタ化する変異は、根本的な生物学的プロセスまたは代謝経路に関係付けられる傾向がある。このような変異は、「機能群」と本明細書で呼ばれる。この方法の重要な知見は、2つの変異が同じ生物学的プロセスまたは経路によって働く場合、観察される効果(特に観察される利益)は、重複し得ることである。反対に、2つの変異が遠位機構によって働く場合、有益な効果が重複する可能性は低い。

【0374】

第6に、上位性効果に基づいて、分析装置214は、図42(4206)に示した通り、同じでない応答に至る変異の対を選択し、例えば、これらの余弦類似度測定基準は、類似度閾値未満に入り、またはこれらの応答は、十分に分離したクラスター内に入る(例えば、図15および図16A、例えば、Corynebacteriumにおいて行った分析における)。これらの相違点に基づいて、選択された変異の対は、同様の対より良好にバックグラウンド株内に統合するはずである。

【0375】

十分に異なった応答に至る選択された変異の対に基づいて、LIMSシステム(例えば、インタープリター204、実行エンジン207、発注依頼者(order placer)208、およびファクトリー210のすべて、またはある組合せ)を使用して、これらの選択された変異を有する微生物株を設計することができる(4208)。諸実施形態では、本明細書で以下に、かつ他の場所で記載する通り、上位性効果を、予測モデルに組み入れる、またはそれと併せて使用して株選択を重み付けまたは選別することができる。

【0376】

ある好適な予測モデルを介してライブラリーからの変異のコレクションを特定のバックグラウンド内に統合することによって得られる仮定の株の性能(別名スコア)を推定することが可能であると仮定される。教示される方法で利用される代表的な予測モデルは、「ゲノムワイド遺伝子設計判定基準の効果のコンピューター分析および予測」というより大きいセクションに見出される「予測的株設計」という表題の以下のセクションに提供されている。

【0377】

線形回帰などの予測的株設計技法を使用する場合、分析装置214は、例えば、回帰結果を選別して十分に同じでない変異のみを保存することによって、低い類似度尺度を有する変異にモデルを制限することができる。代替として、予測モデルは、類似度行列を用い

て重み付けることができる。例えば、一部の実施形態は、類似度行列を使用する重み付き最小二乗回帰を使用して、提案された変異の相互依存性を特徴付けることができる。一例として、重み付けは、回帰モデルに「カーネル」トリックを施すことによって実施することができる。（「カーネルトリック」が多くの機械学習モデル化手法に一般的である程度に、この再重み付け戦略は、線形回帰に制限されない）。

【0378】

このような方法は、当業者に公知である。諸実施形態では、カーネルは、要素 $1 - w * s_{ij}$ を有する行列であり、式中、 1 は、恒等行列の要素であり、 w は、 0 から 1 の間の実数値である。 $w = 0$ であるとき、これは、標準回帰モデルに変換される。実際には、 w の値は、ペアワイズコンビナトリアルコンストラクトおよびこれらの付随効果 $y(m_i, m_j)$ に対して評価されるとき、予測モデルの精度 (r^2 値または二乗平均平方根誤差 (RMSE)) に関係する。1つの単純なインプリメンテーションでは、 w は、 $w = 1 - r^2$ と定義される。この場合、モデルが完全に予測的であるとき、 $w = 1 - r^2 = 0$ であり、統合は、予測モデルだけにに基づき、上位性マッピング手順は、役割を果たさない。一方、予測モデルがまったく予測的でない場合、 $w = 1 - r^2 = 1$ であり、統合は、上位性マッピング手順だけにに基づく。各反復中、精度を評価してモデル性能が改善しているかどうかを決定することができる。

【0379】

本明細書に記載の上位性マッピング手順は、どのモデルが分析装置 214 によって使用されているかに依存しないことが明らかであるはずである。このような予測モデルを考慮すると、コンビナトリアル統合 (combinatorial consolidation) を介して変異ライブラリーにアクセス可能なすべての仮定の株をスコア付けおよびランク付けすることが可能である。

【0380】

一部の実施形態では、上位性効果を説明するために、同じでない変異応答プロファイル进行分析装置 214 が使用して、予測モデルからの各仮定の株に関連したスコアおよびランクを強化することができる。この手順は、同じでない応答プロファイルを有する候補株（例えば、多様なクラスターから引き出される株）を支持するための、スコアの再重み付けとして広く考えることができる。1つの単純なインプリメンテーションでは、株は、相違点閾値を満たさない、または同じクラスターから引き出された構成要素の変異（適当な重み付けを有する）の数によって低減されたそのスコアを有し得る。特定のインプリメンテーションでは、仮定の株の性能推定値は、仮定の株に関連した構成要素の変異（やはり適当な重み付けを有する）のすべての対に関連した類似度行列の項の和によって低減される。仮定の株は、これらの強化されたスコアを使用して再ランク付けすることができる。実際には、このような再重み付け算出は、最初のスコア推定値と併せて実施することができる。

【0381】

結果は、混乱させる上位性相互作用をより有効に回避するように強化されたスコアおよびランクを有する仮定の株のコレクションである。仮定の株をこの時構築することができ、またはこれらは、続いて分析または使用するために別の計算方法に送ることができる。

【0382】

当業者は、本明細書に記載の上位性マッピングおよび反復性予測的株設計は、ペアワイズ変異のみを使用することに限定されず、はるかに多くの変異をバックグラウンド株に同時に適用するように拡張することができることを認識する。別の実施形態では、追加の変異を、本明細書に記載の予測方法に従って選択された変異を使用して既に変異された株に順次適用することができる。別の実施形態では、上位性効果は、互いにわずかに異なるいくつかの株バックグラウンドに同じ遺伝子変異を適用し、改変された株バックグラウンドの中の正の応答プロファイルにおける任意の有意差に留意することによってインピュートされる。

【0383】

遺伝子設計に受け入れられる生物体

開示されるH T Pゲノム操作プラットフォームは、工業用微生物細胞培養物（例えば、*Corynebacterium*）で例示されるが、所望の形質を遺伝子変異体の集団内で同定することができる任意の宿主細胞生物体に適用可能である。

【0384】

さらに、イントロダクションに記載されている通り、本開示は、*E. coli*系における宿主細胞特性を改善するためのH T Pゲノム操作プラットフォームを提供し、*E. coli*におけるこのような系の開発をこれまで妨げてきた多くの問題を解決する。

【0385】

よって、本明細書で使用される場合、用語「微生物」は、広く解釈されるべきである。これは、これらに限定されないが、2つの原核生物ドメイン、細菌および古細菌、ならびにある特定の真核真菌および原生生物を含む。しかし、ある特定の態様では、「高等」真核生物、例えば、昆虫、植物、および動物を本明細書に教示の方法において利用することができる。

【0386】

適当な宿主細胞としては、細菌細胞、藻類細胞、植物細胞、真菌細胞、昆虫細胞、および哺乳動物細胞が挙げられるが、これらに限定されない。一例示的な実施形態では、適当な宿主細胞には、*E. coli*（例えば、New England BioLabs in Ipswich, Mass. から入手可能なSHuffle（商標）コンピテント*E. coli*）が含まれる。*E. coli*ゲノムのサイズは4, 646, 332bpである（図52を参照されたい）。

【0387】

*E. coli*種の適当な宿主株は、腸管毒素原性*E. coli*（ETEC）、腸管病原性*E. coli*（EPEC）、腸管侵入性*E. coli*（EIEC）、腸管出血性*E. coli*（EHEC）、尿路病原性*E. coli*（UPEC）、ペロ毒素産生性*E. coli*、*E. coli* O157:H7、*E. coli* O104:H4、*Escherichia coli* O121、*Escherichia coli* O104:H21、*Escherichia coli* K1、および*Escherichia coli* NC101を含む。一部の実施形態では、本開示は、*E. coli* K12、*E. coli* B、および*E. coli* Cのゲノム操作を教示する。

【0388】

一部の実施形態では、本開示は、*E. coli*株NCTC 12757、NCTC 12779、NCTC 12790、NCTC 12796、NCTC 12811、ATCC 11229、ATCC 25922、ATCC 8739、DSM 30083、BC 5849、BC 8265、BC 8267、BC 8268、BC 8270、BC 8271、BC 8272、BC 8273、BC 8276、BC 8277、BC 8278、BC 8279、BC 8312、BC 8317、BC 8319、BC 8320、BC 8321、BC 8322、BC 8326、BC 8327、BC 8331、BC 8335、BC 8338、BC 8341、BC 8344、BC 8345、BC 8346、BC 8347、BC 8348、BC 8863およびBC 8864のゲノム操作を教示する。

【0389】

一部の実施形態では、本開示は、ペロ毒素産生性*E. coli*（VTEC）、例えば、株BC 4734（O26:H11）、BC 4735（O157:H-）、BC 4736、BC 4737（n. d.）、BC 4738（O157:H7）、BC 4945（O26:H-）、BC 4946（O157:H7）、BC 4947（O111:H-）、BC 4948（O157:H）、BC 4949（O5）、BC 5579（O157:H7）、BC 5580（O157:H7）、BC 5582（O3:H）、BC 5643（O2:H5）、BC 5644（O128）、BC 5645（O55:H-）、BC 5646（O69:H-）、BC 5647（O101:H9）、

BC 5648 (O103:H2)、BC 5850 (O22:H8)、BC 5851 (O55:H-)、BC 5852 (O48:H21)、BC 5853 (O26:H11)、BC 5854 (O157:H7)、BC 5855 (O157:H-)、BC 5856 (O26:H-)、BC 5857 (O103:H2)、BC 5858 (O26:H11)、BC 7832、BC 7833 (O raw form:H-)、BC 7834 (ONT:H-)、BC 7835 (O103:H2)、BC 7836 (O57:H-)、BC 7837 (ONT:H-)、BC 7838、BC 7839 (O128:H2)、BC 7840 (O157:H-)、BC 7841 (O23:H-)、BC 7842 (O157:H-)、BC 7843、BC 7844 (O157:H-)、BC 7845 (O103:H2)、BC 7846 (O26:H11)、BC 7847 (O145:H-)、BC 7848 (O157:H-)、BC 7849 (O156:H47)、BC 7850、BC 7851 (O157:H-)、BC 7852 (O157:H-)、BC 7853 (O5:H-)、BC 7854 (O157:H7)、BC 7855 (O157:H7)、BC 7856 (O26:H-)、BC 7857、BC 7858、BC 7859 (ONT:H-)、BC 7860 (O129:H-)、BC 7861、BC 7862 (O103:H2)、BC 7863、BC 7864 (O raw form:H-)、BC 7865、BC 7866 (O26:H-)、BC 7867 (O raw form:H-)、BC 7868、BC 7869 (ONT:H-)、BC 7870 (O113:H-)、BC 7871 (ONT:H-)、BC 7872 (ONT:H-)、BC 7873、BC 7874 (O raw form:H-)、BC 7875 (O157:H-)、BC 7876 (O111:H-)、BC 7877 (O146:H21)、BC 7878 (O145:H-)、BC 7879 (O22:H8)、BC 7880 (O raw form:H-)、BC 7881 (O145:H-)、BC 8275 (O157:H7)、BC 8318 (O55:K-:H-)、BC 8325 (O157:H7)、およびBC 8332 (ONT)、BC 8333を教示する。

【0390】

一部の実施形態では、本開示は、株BC 8246 (O152:K-:H-)、BC 8247 (O124:K(72):H3)、BC 8248 (O124)、BC 8249 (O112)、BC 8250 (O136:K(78):H-)、BC 8251 (O124:H-)、BC 8252 (O144:K-:H-)、BC 8253 (O143:K:H-)、BC 8254 (O143)、BC 8255 (O112)、BC 8256 (O28a.e)、BC 8257 (O124:H-)、BC 8258 (O143)、BC 8259 (O167:K-:H5)、BC 8260 (O128a.c.:H35)、BC 8261 (O164)、BC 8262 (O164:K-:H-)、BC 8263 (O164)、およびBC 8264 (O124)などの腸管侵入性E. coli (EIEC)を教示する。

【0391】

一部の実施形態では、本開示は、腸管毒素原性E. coli (ETEC)、例えば、株BC 5581 (O78:H11)、BC 5583 (O2:K1)、BC 8221 (O118)、BC 8222 (O148:H-)、BC 8223 (O111)、BC 8224 (O110:H-)、BC 8225 (O148)、BC 8226 (O118)、BC 8227 (O25:H42)、BC 8229 (O6)、BC 8231 (O153:H45)、BC 8232 (O9)、BC 8233 (O148)、BC 8234 (O128)、BC 8235 (O118)、BC 8237 (O111)、BC 8238 (O110:H17)、BC 8240 (O148)、BC 8241 (O6H16)、BC 8243 (O153)、BC 8244 (O15:H-)、BC 8245 (O20)、BC 8269 (O125a.c.:H-)、BC 8313 (O6:H6)、BC 8315 (O153:H-)、BC 8329、BC 8334 (O118:H12)、およびBC 8339を教示する。

【0392】

一部の実施形態では、本開示は、腸管毒素原性 *E. coli* (EPEC)、例えば、株 BC 7567 (O86)、BC 7568 (O128)、BC 7571 (O114)、BC 7572 (O119)、BC 7573 (O125)、BC 7574 (O124)、BC 7576 (O127a)、BC 7577 (O126)、BC 7578 (O142)、BC 7579 (O26)、BC 7580 (OK26)、BC 7581 (O142)、BC 7582 (O55)、BC 7583 (O158)、BC 7584 (O-)、BC 7585 (O-)、BC 7586 (O-)、BC 8330、BC 8550 (O26)、BC 8551 (O55)、BC 8552 (O158)、BC 8553 (O26)、BC 8554 (O158)、BC 8555 (O86)、BC 8556 (O128)、BC 8557 (OK26)、BC 8558 (O55)、BC 8560 (O158)、BC 8561 (O158)、BC 8562 (O114)、BC 8563 (O86)、BC 8564 (O128)、BC 8565 (O158)、BC 8566 (O158)、BC 8567 (O158)、BC 8568 (O111)、BC 8569 (O128)、BC 8570 (O114)、BC 8571 (O128)、BC 8572 (O128)、BC 8573 (O158)、BC 8574 (O158)、BC 8575 (O158)、BC 8576 (O158)、BC 8577 (O158)、BC 8578 (O158)、BC 8581 (O158)、BC 8583 (O128)、BC 8584 (O158)、BC 8585 (O128)、BC 8586 (O158)、BC 8588 (O26)、BC 8589 (O86)、BC 8590 (O127)、BC 8591 (O128)、BC 8592 (O114)、BC 8593 (O114)、BC 8594 (O114)、BC 8595 (O125)、BC 8596 (O158)、BC 8597 (O26)、BC 8598 (O26)、BC 8599 (O158)、BC 8605 (O158)、BC 8606 (O158)、BC 8607 (O158)、BC 8608 (O128)、BC 8609 (O55)、BC 8610 (O114)、BC 8615 (O158)、BC 8616 (O128)、BC 8617 (O26)、BC 8618 (O86)、BC 8619、BC 8620、BC 8621、BC 8622、BC 8623、BC 8624 (O158)、および BC 8625 (O158) を教示する。

【0393】

一部の実施形態では、本開示は、*Shigella flexneri*、*Shigella dysenteriae*、*Shigella boydii*、および *Shigella sonnei* を含めた *Shigella* 生物を操作するための方法も教示する。

【0394】

遺伝子設計 & HTP 微生物操作プラットフォームにおいて利用するための遺伝的多様性プールの生成

一部の実施形態では、本開示の方法は、遺伝子設計として特徴付けられる。本明細書で 사용되는場合、用語の遺伝子設計は、新しい優れた宿主細胞を設計および生み出すために特定の遺伝子、遺伝子の部分、プロモーター、停止コドン、5' UTR、3' UTR、または他の DNA 配列の最適のバリエーションを同定および選択することによる宿主生物のゲノムの再構築または変更を指す。

【0395】

一部の実施形態では、本開示の遺伝子設計法における第1のステップは、新しい宿主ゲノムを再構築することができる複数の配列バリエーションを有する最初の遺伝的多様性プール集団を得ることである。

【0396】

一部の実施形態では、本明細書で教示される遺伝子設計方法における後続のステップは、上述の HTP 分子ツールセット (例えば、SNP スワッピング、プロモーター スワッピング、ターミネーター スワッピング、タンパク質可溶性 タグ スワッピング または タンパク質分解 タグ スワッピング) の 1 つまたは複数を使用して HTP 遺伝子設計 ライブラリーを

構築することであり、次いでこのライブラリーは、宿主細胞において試験するための特定のゲノム変更のライブラリーを提供することによってゲノム操作プロセスのドライバーとして機能する。

【0397】

既存の野生型株からの多様性プールの利用

一部の実施形態では、本開示は、所与の野生型集団の微生物の中に存在する配列多様性を同定するための方法を教示する。したがって、多様性プールは、分析に利用される所与の数 n の野生型微生物であり得、前記微生物のゲノムは、「多様性プール」を代表する。

【0398】

一部の実施形態では、多様性プールは、前記野生型微生物の中で天然の遺伝的バリエーション中に存在する既存の多様性の結果であり得る。このバリエーションは、所与の宿主細胞の株バリエーションから生じる場合があり、または完全に異なる種である微生物の結果である場合がある。遺伝的バリエーションは、天然に存在するか否かにかかわらず、株の遺伝子配列の任意の差異を含み得る。一部の実施形態では、遺伝的バリエーションは、とりわけ、SNPスワップ、PROスワップ、開始/停止コドンスワップ、SOLUBILITY TAGスワップ、DEGRADATION TAGスワップまたはSTOPスワップを含み得る。

【0399】

既存の工業用株バリエーションからの多様性プールの利用

本開示の他の実施形態では、多様性プールは、伝統的な株改良プロセスの間に創製された株バリエーション（例えば、ランダム変異を介して生成され、長年にわたって収率改善のために選択された1種または複数の宿主生物株）である。よって、一部の実施形態では、多様性プールまたは宿主生物は、歴史的な産生株のコレクションを含み得る。

【0400】

特定の態様では、多様性プールは、特定の時点で「ベースライン」遺伝子配列を有する元の親微生物株 (S_1) ($S_1 \text{ Gen}_1$)、およびさらに、前記 S_1 株に由来し、/それから開発された、かつ S_1 のベースラインゲノムに関して異なるゲノムを有する任意の数の後続の子孫株 (S_2 、 S_3 、 S_4 、 S_5 など、 S_{2-n} に一般化可能) ($S_{2-n} \text{ Gen}_{2-n}$) であり得る。

【0401】

例えば、一部の実施形態では、本開示は、多様性プール内の微生物ゲノムをシーケンシングして各株に存在するSNPを同定することを教示する。一実施形態では、多様性プールの株は、歴史的な微生物産生株である。よって、本開示の多様性プールは、例えば、工業用基準株、および伝統的な株改良プログラムを介して産生された1種または複数の変異工業用株を含み得る。

【0402】

多様性プール内のすべてのSNPが同定されたら、本開示は、個々にかつ群において、SNPの効果（例えば、目的の表現型の創製）を描写する（すなわち、定量し、特徴付ける）ためのSNPスワッピングの方法およびスクリーニング方法を教示する。よって、上述した通り、教示したプラットフォームにおける最初のステップは、複数の配列バリエーション、例えば、SNPを有する最初の遺伝的多様性プール集団を得ることであり得る。次いで、教示したプラットフォームにおける後続のステップは、上述のHTP分子ツールセット（例えば、SNPスワッピング）の1つまたは複数を使用して、HTP遺伝子設計ライブラリーを構築することであり得、次いでこのライブラリーは、微生物において試験するための特定のゲノム変更のライブラリーを提供することによってゲノム操作プロセスのドライバーとして機能する。

【0403】

一部の実施形態では、本開示のSNPスワッピング方法は、変異株（例えば、 $S_{2-n} \text{ Gen}_{2-n}$ の中からの株）内で同定された1つまたは複数のSNPを基準株 ($S_1 \text{ Gen}_1$) または野生型株に導入するステップを含む。

【0404】

他の実施形態では、本開示のSNPスワッピング方法は、変異株（例えば、 $S_{2-n}G_{en_2-n}$ の中からの株）内で同定された1つまたは複数のSNPを除去するステップを含む。

【0405】

変異誘発を介した多様性プールの創製

一部の実施形態では、細胞の所与の多様性プール集団における目的の変異は、変異誘発化学物質または放射線を含めた、株を変異させる任意の手段によって人工的に生成することができる。用語「変異誘発させること」は、細胞核酸材料において1つまたは複数の遺伝子改変を誘導するための方法を指すのに本明細書で使用される。

10

【0406】

用語「遺伝子改変」は、DNAの任意の変更を指す。代表的な遺伝子改変は、ヌクレオチド挿入、欠失、置換、およびこれらの組合せを含み、単一塩基という小さいものまたは数万塩基という大きいものであり得る。よって、用語「遺伝子改変」は、染色体の領域を含むDNAの位置または配向が変更されるヌクレオチド配列の反転および他の染色体再配列を包含する。染色体再配列は、染色体内再配列または染色体間再配列を含むことができる。

【0407】

一実施形態では、本主張の主題において使用される変異誘発方法は、実質的にランダムであり、その結果、遺伝子改変は、変異誘発される核酸材料内の任意の利用可能なヌクレオチド位置で行うことができる。別の言い方をすれば、一実施形態では、変異誘発させることは、特定のヌクレオチド配列における出現の優先または頻度の増加を示さない。

20

【0408】

本開示の方法は、これらに限定されないが、紫外光、X線照射、ガンマ照射、N-エチル-N-ニトロソ尿素（ENU）、メチルニトロソ尿素（methylnitrosourea）（MNU）、プロカルバジン（PC）、トリエチレンメラミン（TEM）、アクリルアミドモノマー（AA）、クロラムブシル（CHL）、メルファラン（MLP）、シクロホスファミド（CP）、硫酸ジエチル（DES）、エチルメタンサルホネート（EMS）、メチルメタンサルホネート（MMS）、6-メルカプトプリン（6-MP）、マイトマイシンC（MMC）、N-メチル-N'-ニトロ-N-ニトロソグアニジン（MNG）、 ^3H_2O 、およびウレタン（UR）を含めた任意の変異誘発剤を使用することができる（例えば、Rinchik、1991年；Markerら、1997年；およびRussell、1990年を参照）。追加の変異誘発剤は、http://www.iephb.nw.ru/~spirov/hazard/mutagen_lst.htmlに記載されているものを含めて当業者に周知である。

30

【0409】

用語「変異誘発させること」は、細胞機能を変更（例えば、標的化変異によって）またはモジュレートすることによって、変異誘発の速度、品質、または程度を増強する方法も包含する。例えば、細胞を変質させ、またはモジュレートすることによって、DNA修復、変異誘発物質代謝、変異誘発物質感受性、ゲノム安定性、またはこれらの組合せが機能していない、または欠損しているようにすることができる。よって、通常ゲノム安定性を維持する遺伝子機能の破壊を使用して変異誘発を増強することができる。破壊の代表的な標的としては、これらに限定されないが、DNAリガーゼI（Bentleyら、2002年）およびカゼインキナーゼI（米国特許第6,060,296号）が挙げられる。

40

【0410】

一部の実施形態では、本開示の切断酵素をコードする核酸を生成するために、部位特異的変異誘発（例えば、トランスフォーマー部位指向変異誘発キット（Clontech）などの市販キットを使用するプライマー指向変異誘発）を使用して、核酸配列全体にわたって複数の変化をもたらす。

【0411】

1種または複数の変異誘発剤に曝露した際の遺伝子改変の頻度は、処置の用量および／

50

または繰り返しを変更することによってモジュレートすることができ、特定の用途に合わせることができる。

【0412】

よって、一部の実施形態では、本明細書で使用される「変異誘発」は、エラープローンPCR変異誘発、オリゴヌクレオチド指向性変異誘発、部位指向変異誘発、および本明細書に記載の技法のいずれかによる反復的配列組換えを含めて、変異を誘導するためのすべての当技術分野で公知の技法を含む。

【0413】

多様性を生成するための単一遺伝子座変異

一部の実施形態では、本開示は、ゲノムDNAの選択された部分を導入し、欠失させ、または置き換えることによって細胞集団を変異させることを教示する。よって、一部の実施形態では、本開示は、特定の遺伝子座に変異を標的化するための方法を教示する。他の実施形態では、本開示は、標的DNA領域を選択的に編集するための遺伝子編集技術、例えば、ZFN、TALENs、Lambda RedまたはCRISPRの使用を教示する。

10

【0414】

他の実施形態では、本開示は、宿主生物の外で選択されたDNA領域を変異させ、次いで変異配列を宿主生物内に挿入して戻すことを教示する。例えば、一部の実施形態では、本開示は、生来のプロモーターまたは合成プロモーターを変異させて、様々な発現性質を有する一連のプロモーターバリエーション（以下のプロモーターラダーを参照）を生成することを教示する。他の実施形態では、本開示は、ProSARなどの単一遺伝子最適化技法に適合する（本明細書に参照により組み込まれている、Foxら、2007年、「Improving catalytic function by ProSAR-driven enzyme evolution.」、Nature Biotechnology、25巻（3号）、338～343頁）。

20

【0415】

一部の実施形態では、DNAの選択された領域は、自然バリエーションの遺伝子シャッフリング、または合成オリゴでのシャッフリング、プラスミドプラスミド組換え、ウイルスプラスミド組換え、ウイルスウイルス組換えを介してin vitroで生成される。他の実施形態では、ゲノム領域は、エラープローンPCRを介して生成される。

【0416】

一部の実施形態では、選択された遺伝子領域内で変異を生成するステップは、「リアセンブリーPCR」によって実践される。簡単に言えば、オリゴヌクレオチドプライマー（オリゴ）が、目的の核酸配列のセグメントをPCR増幅するために合成され、その結果、オリゴヌクレオチドの配列は、2つのセグメントの接合部と重複する。重複領域は、典型的には、長さが約10～100ヌクレオチドである。セグメントのそれぞれは、一連のこのようなプライマーで増幅される。次いでPCR産物は、アセンブリープロトコルに従って「リアセンブルされる」。簡単に説明すると、アセンブリープロトコルでは、PCR産物は、例えば、ゲル電気泳動またはサイズ排除クロマトグラフィーによって最初に精製されてプライマーから離れる。精製産物は、一緒に混合され、追加のプライマーの非存在下で（「セルフプライミング」）、ポリメラーゼおよびデオキシヌクレオシドトリホスフェート（dNTP）および適切な緩衝塩の存在下で、約1～10サイクルの変性、リアニリング、および伸長に供される。遺伝子に隣接するプライマーを用いた後続のPCRが使用されて、完全にリアセンブルおよびシャッフルされた遺伝子の収量が拡大される。

30

40

【0417】

本開示の一部の実施形態では、上記に論じたものなどの変異DNA領域は、変異体配列について富化され、その結果、複数の変異体スペクトル、すなわち、変異の可能な組合せが、より効率的にサンプリングされる。一部の実施形態では、変異配列は、アセンブリー反応の前に、in vitroで親和性精製材料を増幅する好適なステップを用いて、mutSタンパク質親和性マトリックスを介して同定される（Wagnerら、Nucleic Acids Res.、23巻（19号）：3944～3948頁（1995年）；Suら、Proc. Natl.

50

Acad. Sci.(U.S.A.), 83巻: 5057~5061頁(1986年))。次いでこの増幅された材料は、本願の後の部分に記載されているアセンブリーまたはリアセンブリーPCR反応にかけられる。

【0418】

プロモーターラダー

プロモーターは、遺伝子が転写される速度を調節し、様々なやり方で転写に影響を与えることができる。構成的プロモーターは、例えば、内部または外部の細胞条件にかかわらず、定速でこれらの関連する遺伝子の転写を指示し、一方、調節可能プロモーターは、内部および／または外部の細胞条件、例えば、増殖速度、温度、具体的な環境化学物質に対する応答などに応じて遺伝子が転写される速度を増減する。プロモーターは、これらの正常な細胞コンテキストから単離することができ、操作されて、実質的に任意の遺伝子の発現を調節することができ、細胞増殖、産物収率、および／または目的の他の表現型の有効な改変を可能にする。

【0419】

一部の実施形態では、本開示は、下流の遺伝子設計法で使用するためのプロモーターラダーライブラリーを生成するための方法を教示する。例えば、一部の実施形態では、本開示は、一連の発現強度または優れた調節性質を呈する、1つもしくは複数のプロモーターを同定し、かつ／または宿主細胞内で1つもしくは複数のプロモーターのバリエーションを生成する方法を教示する。これらの同定および／または生成されたプロモーターの特定の組み合わせは、以下でより詳細に説明されるプロモーターラダーと一緒にグループ化することができる。

【0420】

一部の実施形態では、本開示は、プロモーターラダーの使用を教示する。一部の実施形態では、本開示のプロモーターラダーは、連続した範囲の発現プロファイルを呈するプロモーターを含む。例えば、一部の実施形態では、プロモーターラダーは、刺激に応答して、または構成的発現を通じて一連の発現強度を呈する天然の、生来の、または野生型プロモーターを同定することによって創製される(例えば、図20および図28~30を参照)。これらの同定されたプロモーターは、プロモーターラダーと一緒にグループ化することができる。

【0421】

他の実施形態では、本開示は、異なる条件にわたって一連の発現プロファイルを呈するプロモーターラダーの創製を教示する。例えば、一部の実施形態では、本開示は、発現ピークが発酵の異なる段階全体にわたって広がったプロモーターのラダーを生み出すことを教示する(例えば、図28を参照)。他の実施形態では、本開示は、特異的刺激に応答して異なる発現ピークダイナミクスを有するプロモーターのラダーを生み出すことを教示する(例えば、図29を参照)。当業者は、本開示の調節プロモーターラダーは、任意の1つまたは複数の調節プロファイルの代表であり得ることを認識する。

【0422】

一部の実施形態では、本開示のプロモーターラダーは、応答の連続的な範囲にわたって予測可能な様式で遺伝子発現を攪乱するように設計される。一部の実施形態では、プロモーターラダーの連続的な特質は、株改良プログラムに追加の予測力を付与する。例えば、一部の実施形態では、選択された代謝経路のスワッピングプロモーターまたは終結配列は、非常に最適の発現比またはプロファイルを同定する宿主細胞性能曲線を生成することができ、標的とした遺伝子が特定の反応または遺伝的カスケードにとってもはや制限因子でない一方、不適切な環境下で不要な過剰発現または誤発現も回避する株を生成する。一部の実施形態では、プロモーターラダーは、所望のプロファイルを呈する天然の、生来の、または野生型プロモーターを同定することによって創製される。本明細書に提供される方法において使用するための生来のプロモーターの例は、表1.4に見出すことができる。他の実施形態では、プロモーターラダーは、天然に存在するプロモーターを変異させて複数の変異プロモーター配列を得ることによって創製される。これらの変異プロモーターの

それぞれは、標的遺伝子発現に対する効果について試験される。一部の実施形態では、編集されたプロモーターは、様々な条件にわたって発現活性について試験され、その結果、各プロモーターバリエーションの活性が記録され／特徴付けられ／アノテートされ、かつデータベース内に記憶される。結果として生じる編集されたプロモーターバリエーションは、引き続いて組織化されて、これらの発現の強度に基づいて配列されたプロモーターラダーにされる（例えば、頂部付近に高度発現バリエーションおよび底部付近に減弱された発現を有し、したがって、用語「ラダー」をもたらす）。本明細書に提供される方法において使用するための合成プロモーターの例は、表 1. 4 に見出すことができる。

【0423】

一部の実施形態では、本開示は、同定された天然に存在するプロモーターおよび変異されたバリエーションプロモーターの組合せであるプロモーターラダーを教示する。

10

【0424】

一部の実施形態では、本開示は、以下の判定基準：1) 構成的プロモーターのラダーを表した；および2) 短いDNA配列、理想的には100塩基対未満によってコードされ得る、の両方を満たした天然の、生来の、または野生型プロモーターを同定する方法を教示する。一部の実施形態では、本開示の構成的プロモーターは、2つの選択された成長条件（典型的には、工業的な培養中に経験した条件の間で比較した）にわたって一定の遺伝子発現を呈する。本明細書に提供される異なるプロモーターを使用した遺伝子発現の調査の例は、実施例12に見出すことができる。一部の実施形態では、本開示のプロモーターは、約60塩基対のコアプロモーターおよび長さが26から40塩基対の間の5'UTRからなる。

20

【0425】

本明細書に提供されるPROSWP方法において使用するためのプロモーターラダーに含めるための生来のプロモーターは、関連する遺伝子の発現において最小のバリエーションを示す前記生来のプロモーターに基づいて選択することができる。さらに、生来のプロモーターの長さは60～90bpであり得、推定転写開始点の50bp前に位置する配列、および必要に応じて、推定開始コドンは含まずに推定開始コドンまでの配列からなり得る。本明細書に提供される方法において使用するための生来のプロモーターの例は、表 1. 4 に見出すことができる。特に、本明細書に提供される方法において使用するための生来のプロモーターは、表 1. 4 の核酸配列番号71～131から選択することができる。

30

【0426】

一部の実施形態では、上述の同定した天然に存在するプロモーター配列の1つまたは複数、遺伝子編集のために選択される。一部の実施形態では、天然のプロモーターは、上述した変異方法のいずれかを介して編集される。他の実施形態では、本開示のプロモーターは、所望の配列を有する新しいプロモーターバリエーションを合成することによって編集される。

【0427】

本明細書に提供されるPROSWP方法において使用するためのプロモーターラダーに含めるための合成プロモーターは、長さが60～90bpのキメラ配列であり得る。本発明で使用するための合成プロモーターライブラリーは、一連のまたは複数の合成プロモーターを含み得、これらは、構成的に発現し、かつ／または互いと比較して一連の発現強度を表す可能性が高くなるように設計および構築することができる。さらに、合成プロモーターは、E. coli内に存在するいずれかの調節エレメントに結合し、したがって、遺伝子発現を構成的に駆動する可能性が低くなるように設計および構築することができる。

40

【0428】

これらの設計目標を達成するために、キメラ合成プロモーターは、表 1. 5 に見出されるエレメントのすべてまたは組合せを含み得る。特に、図 5 4 に示されている通り、合成プロモーターは、転写開始点に対して、遠位領域、-35領域、コア領域、-10領域および5'UTR／リボソーム結合部位(RBS)領域を含み得る、またはそれからなり得る。図 5 4 に示されている通り、遠位領域は-35領域のすぐ上流に位置し得、コア領域

50

は-35と-10領域の間に位置し得る。遠位領域およびコア領域はどちらも調節エレメントとの結合のために重要であり得る (Coxら、Mol Syst Biol., 2007年; 3巻: 145頁を参照されたい)。ラムダファージ p_R プロモーターは、発現を構成的に駆動することが予想されるので、このプロモーター由来の遠位およびコア領域を設計ストラテジーに使用することができる。ラムダファージ p_L プロモーターのコア領域も同じ理由で、ならびに追加の多様さをライブラリーに加えるために、含めることができる。

【0429】

-35および-10領域は、原核生物においてRNAポリメラーゼの結合に特に重要であることが分かっており、したがって、発現の度合いをモジュレートするために極めて重要であり得るので、これらを含めることができる。一実施形態では、ラムダファージ p_R プロモーターおよび p_L プロモーター由来の-35および-10領域を使用する。 p_R および p_L 由来の-35および-10領域は、強力な発現を駆動することが予測されるので、これらを使用することができる。さらに、-35および-10領域が p_R および p_L に対する小さなバリエーションを示し、また、 p_R および p_L と比較してプロモーター強度が低下すると予測することができるような、多くの生来の E. coli プロモーターに見出される-35および-10領域を使用することができる。-35および-10領域を構成する可変性の 6 bp 配列は、表 1. 5に見出される-35および-10配列から選択することができる。

【0430】

上記のエレメントに加えて、キメラ合成プロモーターは、原核生物においてリボソームの結合に特に重要であり得、したがって、タンパク質発現の度合いをモジュレートするために極めて重要であり得るリボソーム結合部位 (RBS) を含む 5' 非翻訳領域 (5' UTR) を含み得る。一実施形態では、生来の E. coli 遺伝子 *acs* 由来の 5' UTR/RBS を使用して、追加の多様さをライブラリーに加えることができる。別の実施形態では、ラムダファージ p_R プロモーター由来の 5' UTR/RBS を使用することができる。

【0431】

本明細書に提供される方法において使用するための合成プロモーターの例は、表 1. 4に見出すことができる。特に、本明細書に提供される方法において使用するための合成プロモーターは、表 1. 4の核酸配列番号 132~207から選択することができる。

【0432】

2015年12月7日に出願された米国特許出願第 62/264, 232号の開示全体は、すべての目的に関してその全体が参照により本明細書に組み込まれている。

【0433】

本開示のプロモーターの非網羅的なリストを、以下の表 1 および/または表 1. 4に提供する。プロモーター配列のそれぞれは、異種プロモーターまたは異種プロモーターポリヌクレオチドと呼ぶことができる。

【表 1】

表1. 本開示の選択されたプロモーター配列

配列番号	プロモーター省略名	プロモーター名
1	P1	Pcg0007_lib_39
2	P2	Pcg0007
3	P3	Pcg1860
4	P4	Pcg0755
5	P5	Pcg0007_265
6	P6	Pcg3381
7	P7	Pcg0007_119
8	P8	Pcg3121

10

20

【表 1 . 4 - 1】

表 1.4 本開示の追加のプロモーター配列。

プロモーター名	配列 番号	タイプ*
b0904_プロモーター	71	生来の
b2405_プロモーター	72	生来の
b0096_プロモーター	73	生来の
b0576_プロモーター	74	生来の
b2017_プロモーター	75	生来の
b1278_プロモーター	76	生来の
b4255_プロモーター	77	生来の
b0786_プロモーター	78	生来の
b0605_プロモーター	79	生来の
b1824_プロモーター	80	生来の
b1061_プロモーター	81	生来の
b0313_プロモーター	82	生来の
b0814_プロモーター	83	生来の
b4133_プロモーター	84	生来の
b4268_プロモーター	85	生来の
b0345_プロモーター	86	生来の
b2096_プロモーター	87	生来の
b1277_プロモーター	88	生来の
b1646_プロモーター	89	生来の
b4177_プロモーター	90	生来の
b0369_プロモーター	91	生来の
b1920_プロモーター	92	生来の
b3742_プロモーター	93	生来の
b3929_プロモーター	94	生来の
b3743_プロモーター	95	生来の
b1613_プロモーター	96	生来の

10

20

30

【表 1. 4 - 2】

b1749_プロモーター	97	生来の
b2478_プロモーター	98	生来の
b0031_プロモーター	99	生来の
b2414_プロモーター	100	生来の
b1183_プロモーター	101	生来の
b0159_プロモーター	102	生来の
b2837_プロモーター	103	生来の
b3237_プロモーター	104	生来の
b3778_プロモーター	105	生来の
b2349_プロモーター	106	生来の
b1434_プロモーター	107	生来の
b3617_プロモーター	108	生来の
b0237_プロモーター	109	生来の
b4063_プロモーター	110	生来の
b0564_プロモーター	111	生来の
b0019_プロモーター	112	生来の
b2375_プロモーター	113	生来の
b1187_プロモーター	114	生来の
b2388_プロモーター	115	生来の
b1051_プロモーター	116	生来の
b4241_プロモーター	117	生来の
b4054_プロモーター	118	生来の
b2425_プロモーター	119	生来の
b0995_プロモーター	120	生来の
b1399_プロモーター	121	生来の
b3298_プロモーター	122	生来の
b2114_プロモーター	123	生来の
b2779_プロモーター	124	生来の
b1114_プロモーター	125	生来の
b3730_プロモーター	126	生来の
b3025_プロモーター	127	生来の
b0850_プロモーター	128	生来の

10

20

30

【表 1 . 4 - 3】

b2365_プロモーター	129	生来の
b4117_プロモーター	130	生来の
b2213_プロモーター	131	生来の
pMB029_プロモーター	132	合成
pMB023_プロモーター	133	合成
pMB025_プロモーター	134	合成
pMB019_プロモーター	135	合成
pMB008_プロモーター	136	合成
pMB020_プロモーター	137	合成
pMB022_プロモーター	138	合成
pMB089_プロモーター	139	合成
pMB001_プロモーター	140	合成
pMB051_プロモーター	141	合成
pMB070_プロモーター	142	合成
pMB074_プロモーター	143	合成
pMB046_プロモーター	144	合成
pMB071_プロモーター	145	合成
pMB013_プロモーター	146	合成
pMB080_プロモーター	147	合成
pMB038_プロモーター	148	合成
pMB060_プロモーター	149	合成
pMB064_プロモーター	150	合成
pMB058_プロモーター	151	合成
pMB085_プロモーター	152	合成
pMB081_プロモーター	153	合成
pMB091_プロモーター	154	合成
pMB027_プロモーター	155	合成
pMB048_プロモーター	156	合成
pMB055_プロモーター	157	合成
pMB006_プロモーター	158	合成
pMB012_プロモーター	159	合成
pMB014_プロモーター	160	合成

10

20

30

【表 1 . 4 - 4】

pMB028_プロモーター	161	合成
pMB059_プロモーター	162	合成
pMB061_プロモーター	163	合成
pMB043_プロモーター	164	合成
pMB066_プロモーター	165	合成
pMB079_プロモーター	166	合成
pMB032_プロモーター	167	合成
pMB068_プロモーター	168	合成
pMB082_プロモーター	169	合成
pMB030_プロモーター	170	合成
pMB067_プロモーター	171	合成
pMB050_プロモーター	172	合成
pMB069_プロモーター	173	合成
pMB017_プロモーター	174	合成
pMB039_プロモーター	175	合成
pMB011_プロモーター	176	合成
pMB072_プロモーター	177	合成
pMB016_プロモーター	178	合成
pMB077_プロモーター	179	合成
pMB047_プロモーター	180	合成
pMB052_プロモーター	181	合成
pMB090_プロモーター	182	合成
pMB035_プロモーター	183	合成
pMB073_プロモーター	184	合成
pMB004_プロモーター	185	合成
pMB054_プロモーター	186	合成
pMB024_プロモーター	187	合成
pMB007_プロモーター	188	合成
pMB005_プロモーター	189	合成
pMB003_プロモーター	190	合成
pMB088_プロモーター	191	合成
pMB065_プロモーター	192	合成

10

20

30

【表 1. 4 - 5】

pMB037_プロモーター	193	合成
pMB009_プロモーター	194	合成
pMB041_プロモーター	195	合成
pMB036_プロモーター	196	合成
pMB049_プロモーター	197	合成
pMB044_プロモーター	198	合成
pMB042_プロモーター	199	合成
pMB086_プロモーター	200	合成
pMB053_プロモーター	201	合成
pMB057_プロモーター	202	合成
pMB018_プロモーター	203	合成
pMB002_プロモーター	204	合成
pMB015_プロモーター	205	合成
pMB087_プロモーター	206	合成
pMB063_プロモーター	207	合成

10

*Escherichia coli 由来の生来のプロモーター

20

【表 1. 5 - 1】

表1.5.コンビナトリアル合成プロモーター-5' UTRライブラリーに使用される配列部分

部分名	部分配列	起源
遠位	ACCGTGCGTG (配列番号 208)	λ ファージ、 P_R プロモーター
-35	TTTACA	バリエーション
	TTGACT	λ ファージ、 P_L プロモーター
	TTGACA	λ ファージ、 P_R プロモーター
	TAGGCT	バリエーション
	TAGACT	バリエーション
コア	ATTTTACCTCTGGCGGT (配列番号 209)	λ ファージ、 P_R プロモーター
	TAAATACCACTGGCGGT (配列番号 210)	λ ファージ、 P_L プロモーター
-10	GATAAT	λ ファージ、 P_L プロモーター
	GATACT	λ ファージ、 P_R プロモーター
	TAGAGT	バリエーション
	TATAAT	バリエーション
	TATTTT	バリエーション

30

40

【表 1. 5 - 2】

5'- UTR/RBS	GGTTGCATGTACTAAGGAGGTTGT (配列番号 211)	λ ファージ、 P_R プロモーター
	TAACATCCTACAAGGAGAACAAAAGC (配列番号 212)	<i>E. coli</i> , <i>acs</i> 遺伝子のプロモーター

【0434】

一部の実施形態では、本開示のプロモーターは、表 1 および／または表 1. 4 由来のプロモーターと少なくとも 100%、99%、98%、97%、96%、95%、94%、93%、92%、91%、90%、89%、88%、87%、86%、85%、84%、83%、82%、81%、80%、79%、78%、77%、76%、または75%の配列同一性を呈する。

【0435】

バイシストロン性調節エレメント設計

効率的かつスケラブルなHTP遺伝子設計への関門の1つは、新規の組合せで確実に再使用することができる標準部分がないことである。E. coli 内での多くの例により、表面上は単純な遺伝子機能が異なる状況でいかに違うように挙動するかが強調される。例えば、一部の実施形態では、1つのコード配列の翻訳を開始する原核生物リボソーム結合部位(RBS)エレメントは別のコード配列では全く機能しない可能性がある(Salis, H.M.ら、「Automated design of synthetic ribosome binding sites to control protein expression」、Nat. Biotechnol.、27巻、946～950頁(2009年)を参照されたい)。転写および翻訳などの中心細胞プロセスの制御をコードする遺伝子エレメントが確実に再使用されない可能性がある場合、このような基本エレメントにコードされる高次の対象物がより大きな規模のシステムにおいて信頼できるものになる見込みはほとんどない。一部の実施形態では、本開示の方法では、バイシストロン設計調節配列を使用することによって上述の課題が克服される。

【0436】

一部の実施形態では、本開示のバイシストロン設計により、様々なコード遺伝子に対する所与のプロモーターの発現強度の状況依存的な変動性を著しく低下させることができる(Mutalikら、「Precise and reliable gene expression via standard transcription and translation initiation elements」、Nat. Biotechnol.、10巻(4号)、354～368頁(2013年))。一部の実施形態では、本開示は、バイシストロン設計(BCD)が、プロモーターにより2つのコード配列の発現が駆動されるヌクレオチド配列であり、同じヌクレオチド塩基において第1のコード配列(シストロン1)が終止し、第2のコード配列を開始する(シストロン2/標的遺伝子)ことを教示する。この戦略により、プロモーターと第2のコード配列の間の予測不可能な相互作用に起因する第2のコード配列の発現強度の変動性を回避する手段がもたらされる。

【0437】

一部の実施形態では、本開示のプロモーターは、バイシストロン設計に従う化合物調節配列である。すなわち、一部の実施形態では、本開示のプロモーターラダー内のプロモーターは、i) 作動可能に連結したプロモーター、ii) 第1のリボソーム結合部位(SD1)、iii) 第1のシストロン配列(Cis1)、iv) 第2のリボソーム結合部位(SD2)、v) 標的遺伝子コード配列(Cis2)を含むより大きな調節配列であり、SD1がCis1に作動可能に連結しており、Cis1がSD2と重複しており、次いでSD2がCis2に作動可能に連結している(図43を参照されたい)。一部の実施形態では、本開示は、エレメントi)～iv)の組合せを「バイシストロン設計」または「バイシストロン設計調節配列」(BCD)と呼ぶ。

【0438】

10

20

30

40

50

一部の実施形態では、本開示のBCDを任意の標的遺伝子に作動可能に連結することができる。よって、一部の実施形態では、本開示のBCDを従来のプロモーターの代わりに使用することができる。一部の実施形態では、本開示は、PROスワップツールボックスにBCDを使用することにより、発現された転写物が翻訳される一貫性が増大することを教示する。いかなる一理論によっても束縛されることを望むことなく、本発明者らは、標的遺伝子に作動可能に連結したSD1およびCis1リーダー配列が存在することにより、活性なりボソーム複合体が補充され、次いで、それが一貫してSD2リボソーム結合部位を介して標的遺伝子の翻訳を再開することができると考える。

【0439】

プロモーターおよびバイシストロン設計エレメントのコレクションは、HTPゲノム操作のために使用できることが報告されている(Mutalikら、「Precise and reliable gene expression via standard transcription and translation initiation elements」、Nat. Biotechnol.、10巻(4号)、354~368頁(2013年)を参照されたい)。しかし、これらの報告された配列はすべて、48ntの調節バイシストロン設計配列の最初の35ntに同一のDNA配列を含有する(図43のMutalikの現況技術配列を参照されたい)。

【0440】

一部の実施形態では、本開示は、MutalikらのBCDを使用して、単一の生物体において多数の標的遺伝子を有効に操作できないことを教示する。すなわち、一部の実施形態では、本開示は、MutalikのBCDの宿主細胞のゲノムへの多重組込みに逆らう教示をする。いかなる一理論によっても束縛されることを望むことなく、本発明者らは、MutalikらのBCDを繰り返し使用することにより、ゲノム全体を通して高度に相同な配列が存在することによって誘発される望ましくない相同組換え(HR)の率が上昇すると考える。

【0441】

一部の実施形態では、本開示は、同一でないヌクレオチド配列を有する新規のBCDを記述することによってこの問題を解決する。これらの新規のBCDをE.coliにおいてHTPゲノム操作に使用して、単一のゲノム内の多数の遺伝子の予測可能な発現の変化を、これらの遺伝子のコード配列とは関係なく、望ましくない相同組換えを誘導することなく、もたらすことができる。

【0442】

一部の実施形態では、本開示は、宿主生物において2種の標的遺伝子タンパク質を比較的同様のレベルで発現させる方法を教示する。よって、一部の実施形態では、本開示は、2種またはそれよりも多くの標的遺伝子タンパク質を互いに0.2、0.4、0.6、0.8、1、1.2、1.4、1.6、1.8、2、2.2、2.4、2.6、2.8、または3倍以内で発現させることを教示する。

【0443】

一部の実施形態では、本開示は、宿主生物において2種の標的遺伝子タンパク質を同様のレベルで発現させると同時に、同一の調節配列の使用によって誘発される望ましくない相同組換え(HR)事象のリスクを低下させる方法を教示する。したがって、一部の実施形態では、本開示は、発現レベルが維持されるやり方でBCD配列を変動させると同時に、HRのリスクを低下させる方法を教示する。すなわち、一部の実施形態では、本開示は、タンパク質をコードする2種またはそれよりも多くの遺伝子を、Cis1配列を除いて同一のBCDによって発現させることを教示する。

【0444】

BCDプロモーター

一部の実施形態では、本開示のBCDは、プロモーター配列を含む。一部の実施形態では、BCDに含まれるプロモーターは、宿主細胞において発現することができる任意のプロモーターであり得る。よって、一部の実施形態では、プロモーターは、本明細書に開示される任意のプロモーターであり得る。一部の実施形態では、プロモーターは、E.coliにおいて機能することが公知の任意のプロモーターであり得る。他の実施形態では、

プロモーターは、表 1 および／または表 1. 4 に開示されている任意のプロモーターであり得る。

【0445】

第 1 および第 2 のリボソーム結合部位 (SD1 および SD2)

一部の実施形態では、本開示の BCD は、第 1 および第 2 のリボソーム結合部位を含み、これらは、それぞれ SD1 および SD2 と呼ばれる。一部の実施形態では、SD1 および SD2 の配列は同じであり得る。他の実施形態では、SD1 および SD2 の配列は異なり得る。

【0446】

一部の実施形態では、SD 配列は、HTP ゲノム操作を受ける宿主において機能的な任意の公知のリボソーム結合部位であり得る。他の実施形態では、本開示は、NNNGGANNN の SD 配列を教示し、ここで、N は、任意のヌクレオチドを指す。他の実施形態では、本開示は、表 1. 1 に開示されている配列から選択される SD 配列を教示する。

【表 1. 1 - 1】

表1.1- SD1およびSD2への使用に適したリボソーム結合部位の限定されない一覧。

SD #	配列	SD #	配列	SD #	配列
1	TGCGGAGGG	50	AGTGGGAACC	99	GGAGGAGAC
2	CACGGAGGC	51	GCCGGAGTG	100	TAGGGAACG
3	TGGGGAGGG	52	GCGGGACAG	101	TGAGGACGT
4	AGGGGAGGC	53	CCGGGATAA	102	ATCGGAGGT
5	GGGGGAGGG	54	CCAGGAACG	103	CCGGGAGGG
6	TATGGAGGT	55	GACGGAGCA	104	ACGGGACGG
7	GACGGAGCG	56	AGCGGATTG	105	GTGGGAGAG
8	AGGGGAGTC	57	ATGGGATAT	106	CGAGGATAG
9	ACAGGAGGC	58	CTGGGAAGT	107	TGGGGAGCC
10	GGGGGAGCG	59	TGGGGAGCC	108	CGTGGAGTG
11	ACAGGAGTC	60	GTGGGACAT	109	TGGGGACTG
12	ACAGGAGAC	61	GCCGGAGCG	110	GGTGGAAGC
13	TGGGGAGGC	62	AATGGAGGC	111	GATGGACGG
14	TGGGGAGAT	63	GGAGGATTG	112	AATGGAGGT
15	GGGGGAGAA	64	GATGGACTC	113	GGAGGATCC
16	AGGGGAGGC	65	GAGGGATGG	114	ACAGGATCT
17	AGAGGAGTC	66	GAAGGACAG	115	CCCGGACAG

10

20

30

40

【表 1. 1 - 2】

SD #	配列	SD #	配列	SD #	配列
18	AGGGGATAT	67	TAGGGAAGG	116	ACGGGAAAC
19	CGTGGAGTG	68	TGGGGAACC	117	AGTGGACCG
20	GGGGGAAGG	69	TTAGGAGTC	118	CATGGATCA
21	AGGGGAATC	70	GGAGGAGGA	119	CTGGGATGT
22	CCCGGAGGT	71	CCGGGATCT	120	TTAGGATGG
23	AGGGGAGGG	72	GGGGGATGA	121	GTAGGATTC
24	TTTGGAGTC	73	CTGGGAGTG	122	TTTGGAGTT
25	AGGGGACAC	74	CAAGGAACC	123	CGCGGACTC
26	CGGGGAGAC	75	GCTGGAGGC	124	CTCGGACAG
27	GGGGGAGGG	76	ATGGGACCT	125	CAAGGAGTC
28	AAGGAGATC	77	GGAGGAGGG	126	CTTGGACAG
29	TAAGGAGGT	78	CTGGGATGC	127	GGCGGATCG
30	GGGGGAGTC	79	ACAGGATAC	128	GGGGGACAG
31	GGAGGATCG	80	GAGGGAAGG	129	ATTGGATGG
32	CTGGGATCG	81	AGTGGATCT	130	CCTGGATAT
33	ATCGGACCG	82	CCGGGAGTT	131	CTCGGATAC
34	GGGGGAGTG	83	GGTGGAGGC	132	GGGGGAGCC

10

20

30

40

【表 1. 1 - 3】

SD #	配列	SD #	配列	SD #	配列
35	TAGGGAGCA	84	CCAGGAAGA	133	TGGGGATTG
36	GGTGGAGGG	85	GGGGGATTT	134	CCAGGATCA
37	GCCGGAGGT	86	CGCGGAGTA	135	CGGGGACGG
38	TGAGGAAAG	87	CCAGGAATC	136	TCAGGACAA
39	CTTGGAGGG	88	ATTGGAGTG	137	GTAGGATGG
40	TGAGGAGAT	89	CTAGGAAGT	138	TCCGGACTA
41	CGTGGAGGG	90	GAAGGATAG	139	GTCGGATCA
42	GCGGGAGGG	91	AAAGGACAC	140	TGCGGAGTT
43	GGGGGATAG	92	CCGGGACGT	141	GGTGGACGG
44	GGGGGAGCG	93	ATGGGAGTG	142	CTTGGACGA
45	CCGGGAGCA	94	AGCGGACAG	143	CTGGGACTT
46	GCGGGAGTA	95	GGCGGATCT	144	TATGGAGTA
47	GTAGGACCG	96	ATAGGAGGG	145	ACAGGAGGC
48	CCGGGAAGG	97	TATGGAGGG		
49	GACGGAGTC	98	GGTGGACTC		

【0447】

本明細書に記載の上位性マッピング手順は、どのモデルが分析装置 214 によって使用されているかに依存しないことが明らかであるはずである。このような予測モデルを考慮すると、コンビナトリアル統合を介して変異ライブラリーにアクセス可能なすべての仮定の株をスコア付けおよびランク付けすることが可能である。

【0448】

一部の実施形態では、本開示は、BCDにおける個々のSD配列を変動させることが、BCDの全般的な発現に影響を及ぼすことを教示する。一部のSD配列は、BCDの全般的な発現潜在性を増大させるまたは低減するために役立ち得る。しかし、各BCDは、異なる標的遺伝子Cis2配列にわたって組み合わせるとき、一貫した発現結果を呈することが予測される。

10

20

30

40

50

50

一部の実施形態では、C i s 1 配列は、C i s 2 配列と、これらの中のすべての範囲および部分範囲を含め、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、100ヌクレオチドまたはそれよりも長く重複している。他の実施形態では、本開示のB C Dは、C i s 1 配列がC i s 2 標的遺伝子コード配列と1ヌクレオチド重複し、したがって、最後の数ヌク

レオチドが−1フレームシフトによって停止および開始コドンの両方をコードするように設計される（図43を参照されたい）。一部の実施形態では、C i s 1およびC i s 2配列は、C i s 1およびC i s 2の配列が組み合わさったキメラタンパク質の形成を防止するために、異なるオープンリーディングフレーム上にコードされていなければならない。

【0455】

一部の実施形態では、本開示は、C i s 1配列の開始コドンが任意の機能的開始コドンであり得ることを教示する。一部の実施形態では、本開示は、原核生物では最も一般的な開始コドンにA T G（A U G）が使用され、G T G（G U G）およびT T G（U U G）がそれに続くことを教示する。

【0456】

一部の実施形態では、本開示は、C i s 1配列がいかなる中途停止コドンも有さないことを教示する。他の実施形態では、本開示は、C i s 1配列内の希少コドンにより、C i s 2の翻訳効率が低下し得ることを教示する。したがって、一部の実施形態では、C i s 1は、最大の発現を達成するために、いかなる希少コドンも有さないペプチドをコードする。他の実施形態では、C i s 1は、C i s 2の発現をモジュレートするために、1つまたは複数の希少コドンを含むペプチドをコードする。

【0457】

他の実施形態では、本開示は、C i s 1配列内の多数のコドンリピートにより、C i s 2の翻訳効率が低下し得ることを教示する。したがって、一部の実施形態では、C i s 1は、最大の発現を達成するために、いかなるコドンリピートも有さないペプチドをコードする。他の実施形態では、C i s 1は、C i s 2の発現をモジュレートするために、1つまたは複数のコドンリピートを含むペプチドをコードする。

【0458】

第2のシストロン配列（C i s 2—標的遺伝子）

一部の実施形態では、本開示は、本開示のB C DがC i s 2標的遺伝子配列に、P R Oスワップライブラリーのプロモーターが標的配列に作動可能に連結されるのとはほぼ同じやり方で作動可能に連結されることを教示する。すなわち、一部の実施形態では、本開示のP R Oスワップライブラリーおよび方法において本開示のB C Dを従来のプロモーターの代わりにすることができる。一部の実施形態では、C i s 2配列は、任意の目的の配列であり得る。

【0459】

本開示は、一部の実施形態では、ポリペプチドをコードする標的遺伝子が、B C Dにより、プロモーターによるよりも有効に調節されることを教示する。すなわち、一部の実施形態では、B C Dは、非コードR N Aの発現を、プロモーターによって可能であるよりも大きくモジュレートしない。

【0460】

ターミネーターラダー

一部の実施形態では、本開示は、R N Aコードエレメントの末端の3'位で1つまたは複数の転写終結配列を提供することによって、遺伝子操作された宿主株を改良する方法を教示する。一部の実施形態では、本開示は、終結配列を付加すると、遺伝子操作された宿主内の選択された遺伝子のR N A転写の効率が改善されることを教示する。他の実施形態では、本開示は、終結配列を付加すると、遺伝子操作された宿主内の選択された遺伝子のR N A転写の効率が低減されることを教示する。よって、一部の実施形態では、本開示のターミネーターラダーは、一連の転写効率を呈する一連のターミネーター配列（例えば、1つの弱いターミネーター、1つの平均のターミネーター、および1つの強いプロモーター）を含む。

【0461】

転写終結配列は、転写で、オープンリーディングフレームをコードするヌクレオチド配列の下流に配置されたとき、オープンリーディングフレームの転写を終了させる任意のヌクレオチド配列であり得る。このような配列は、当技術分野で公知であり、原核生物、真

10

20

30

40

50

核生物、またはファージ起源のものであり得る。ターミネーター配列の例としては、これらに限定されないが、P T H-ターミネーター、p E T-T 7ターミネーター、T 3-T ϕターミネーター、p B R 3 2 2-P 4ターミネーター、水疱性口内炎 (vesicular stomatitis) ウイルスターミネーター、r r n B-T 1ターミネーター、r r n Cターミネーター、T T a d c 転写ターミネーター、ならびに酵母認識終結配列、例えば、M a t α (α-因子) 転写ターミネーター、生来のα-因子転写終結配列、A D R 1 転写終結配列、A D H 2 転写終結配列、およびG A P D 転写終結配列が挙げられる。転写ターミネーター配列の非網羅的なリストは、<http://partsregistry.org/Terminators/Catalog>で利用可能である。

【0462】

10

一部の実施形態では、転写終結配列は、ポリメラーゼ特異的であっても非特異的であってもよいが、本実施形態において使用するために選択される転写ターミネーターは、選択されたプロモーターとの「機能的組合せ」を形成するべきであり、ターミネーター配列は、プロモーターで開始するRNAポリメラーゼのタイプによって転写を終結することができるべきであることを意味する。例えば、一部の実施形態では、本開示は、真核生物RNA p o l I I プロモーターおよび真核生物RNA p o l I I ターミネーター、T 7 プロモーターおよびT 7 ターミネーター、T 3 プロモーターおよびT 3 ターミネーター、酵母認識プロモーターおよび酵母認識終結配列などは一般に、機能的組合せを形成するはずであることを教示する。使用される転写終結配列の実体 (identity) は、転写が所与のプロモーターから終結される効率に基づいて選択することもできる。例えば、異種転写ターミネーター配列は、RNAコードエレメントの転写的に下流に提供されて、所与のプロモーターから少なくとも60%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも91%、少なくとも92%、少なくとも93%、少なくとも94%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、または少なくとも99%の終結効率を達成することができる。

20

【0463】

一部の実施形態では、操作された発現コンストラクトからのRNA転写の効率は、RNAコードエレメントの末端の3'位で2つまたはそれ超のヘアピンを含む二次構造を形成する核酸配列を提供することによって改善することができる。特定の理論によって束縛されることを望むことなく、二次構造は、転写伸長複合体を不安定化させ、ポリメラーゼをDNA鋳型から解離した状態にし、それによって、非機能的配列の非生産的転写を最小限にし、所望のRNAの転写を増大させる。したがって、2つまたはそれ超の隣接するヘアピンを含む二次構造を形成する終結配列を提供することができる。一般に、ヘアピンは、それ自体で折り返してアームが一本鎖ループによって接続されている対を成したステム領域を形成することができるパリンドロームヌクレオチド配列によって形成することができる。一部の実施形態では、終結配列は、2、3、4、5、6、7、8、9、10、またはそれ超の隣接するヘアピンを含む。一部の実施形態では、隣接するヘアピンは、0、1、2、3、4、5、6、7、8、9、10、11、12、13、14、または15の対を成していないヌクレオチドによって分離されている。一部の実施形態では、ヘアピンステムは、長さが4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、またはそれ超の塩基対を含む。ある特定の実施形態では、ヘアピンステムは、長さが12~30塩基対である。ある特定の実施形態では、終結配列は、約9~25塩基対を含むステム領域を有する2つまたはそれ超の中型のヘアピンを含む。一部の実施形態では、ヘアピンは、1、2、3、4、5、6、7、8、9、または10ヌクレオチドのループ形成領域を含む。一部の実施形態では、ループ形成領域は、4~8ヌクレオチドを含む。特定の理論によって束縛されることを望むことなく、二次構造の安定性は、終結効率と相関し得る。ヘアピン安定性は、その長さ、それが含有するミスマッチまたはバルジの数、および対を成した領域の塩基組成によって決定される。グアニンとシトシンとの間の対合は、3

30

40

50

つの水素結合を有し、2つしか有さないアデニン-チミン対合と比較してより安定である。ヘアピン形成パリンδροームヌクレオチド配列のG/C含有量は、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%またはそれ超であり得る。一部の実施形態では、ヘアピン形成パリンδροームヌクレオチド配列のG/C含有量は、少なくとも80%である。一部の実施形態では、終結配列は、原核生物、真核生物、またはファージ起源の1つまたは複数の転写ターミネーター配列に由来する。一部の実施形態では、一連の4、5、6、7、8、9、10またはそれ超のアデニン(A)をコードするヌクレオチド配列が、終結配列の3'に提供される。

【0464】

一部の実施形態では、本開示は、一連のタンデム終結配列の使用を教示する。一部の実施形態では、一連の2、3、4、5、6、7、またはそれ超の第1の転写ターミネーター配列を、dsRNAコードエレメントの最終ヌクレオチドの3'に直接、またはdsRNAコードエレメントの最終ヌクレオチドの3'に、少なくとも1~5、5~10、10~15、15~20、20~25、25~30、30~35、35~40、40~45、45~50、50~100、100~150、150~200、200~300、300~400、400~500、500~1,000、もしくはそれ超のヌクレオチドの距離で配置することができる。タンデム転写ターミネーター配列間のヌクレオチドの数は、様々であり得、例えば、転写ターミネーター配列は、0、1、2、3、4、5、6、7、8、9、10、10~15、15~20、20~25、25~30、30~35、35~40、40~45、45~50、またはそれ超のヌクレオチドによって分離され得る。一部の実施形態では、転写ターミネーター配列は、構造予測アルゴリズムによって決定されるこれらの予測された二次構造に基づいて選択することができる。構造予測プログラムは、当技術分野で周知であり、例えば、CLC Main Workbenchを含む。

【0465】

当業者は、本開示の方法は、任意の終結配列に適合することを認識する。一部の実施形態では、本開示は、Pfeifer-Sancarら、2013年「Comprehensive analysis of the *Corynebacterium glutamicum* transcriptome using an improved RNAseq technique」、Pfeifer-Sancarら、BMC Genomics、2013年、14巻：888頁）に開示されたアノテートされた*Corynebacterium glutamicum*ターミネーターの使用を教示する。他の実施形態では、本開示は、<http://partsregistry.org/Terminators/Catalog>で利用可能であるiGEMレジストリにおいて見出される転写ターミネーター配列の使用を教示する。本開示の転写ターミネーター配列の非網羅的なリストを以下の表1.2に提供する。

10

20

30

【表 1. 2 - 1】

表1.2.本開示の終結配列の非網羅的なリスト。

<i>E. coli</i>			
名称	説明	方向	長さ
BBa_B0010	<i>E. coli</i> rrnB 由来の T1	順方向	80
BBa_B0012	大腸菌ファージT7由来のTE	順方向	41
BBa_B0013	大腸菌ファージT7由来のTE (+/-)	順方向	47
BBa_B0015	二重ターミネーター (B0010-B0012)	順方向	129
BBa_B0017	二重ターミネーター (B0010-B0010)	順方向	168
BBa_B0053	ターミネーター (His)	順方向	72
BBa_B0055	-- 説明なし --		78
BBa_B1002	ターミネーター (人工、 小さい、%T _~ =85%)	順方向	34
BBa_B1003	ターミネーター (人工、小さい、%T _~ =80)	順方向	34
BBa_B1004	ターミネーター (人工、小さい、%T _~ =55)	順方向	34
BBa_B1005	ターミネーター (人工、 小さい、%T _~ =25%)	順方向	34
BBa_B1006	ターミネーター (人工、大きい、%T _~ >90)	順方向	39

10

20

30

【表 1. 2 - 2】

BBa_B1010	ターミネーター (人工、大きい、%T<10)	順方向	40
BBa_I11013	バイオブリック部分BBa_B0015の修飾		129
BBa_I51003	-- 説明なし --		110
BBa_J61048	[rnpB-T1] ターミネーター	順方向	113
BBa_K1392970	ターミネーター+Tetr プロモーター +T4エンドリシン		623
BBa_K1486001	アラビノースプロモーター + CpxR	順方向	1924
BBa_K1486005	アラビノースプロモーター + sfGFP-CpxR [Cterm]	順方向	2668
BBa_K1486009	CxpR & Split IFP1.4 [Nterm + Nterm]	順方向	3726
BBa_K780000	Bacillus subtilis のターミネーター		54
BBa_K864501	T22, P22 後期ターミネーター	順方向	42
BBa_K864600	T0 (21 imm) 転写ターミネーター	順方向	52
BBa_K864601	ラムダt1転写ターミネーター	順方向	
BBa_B0011	LuxICDABEG (+/-)	双方向性	46
BBa_B0014	二重ターミネーター (B0012-B0011)	双方向性	95
BBa_B0021	LuxICDABEG (+/-)、逆向き	双方向性	46

10

20

30

40

【表 1. 2 - 3】

BBa_B0024	二重ターミネーター (B0012-B0011) 、逆向き	双方向性	95
BBa_B0050	ターミネーター (pBR322, +/-)	双方向性	33
BBa_B0051	ターミネーター (yciA/tonA, +/-)	双方向性	35
BBa_B1001	ターミネーター (人工、小さい、%T $\sim$ 90)	双方向性	34
BBa_B1007	ターミネーター (人工、大きい、%T $\sim$ 80)	双方向性	40
BBa_B1008	ターミネーター (人工、大きい、%T $\sim$ 70)	双方向性	40
BBa_B1009	ターミネーター (人工、大きい、%T $\sim$ 40%)	双方向性	40
BBa_K187025	pABにおけるターミネーター、 BioBytesプラスミド		60
BBa_K259006	GFP-ターミネーター	双方向性	823
BBa_B0020	ターミネーター (逆方向 B0010)	逆方向	82
BBa_B0022	大腸菌ファージT7由来のTE、逆向き	逆方向	41
BBa_B0023	大腸菌ファージT7由来のTE、逆向き	逆方向	47
BBa_B0025	二重ターミネーター (B0015)、逆向き	逆方向	129
BBa_B0052	ターミネーター (rrnC)	順方向	41
BBa_B0060	ターミネーター (逆方向 B0050)	双方向性	33
BBa_B0061	ターミネーター (逆方向 B0051)	双方向性	35
BBa_B0063	ターミネーター (逆方向 B0053)	逆方向	72

10

20

30

40

【表 1. 2 - 4】

Spy	ターミネーター (配列番号 225)				90
pheA	ターミネーター (配列番号 226)				51
osmE	ターミネーター (配列番号 227)				42
rpoH	ターミネーター (配列番号 228)				41
vibE	ターミネーター (配列番号 229)				71
Thrl_ABC	ターミネーター (配列番号 230)				57
Corynebacterium					
ターミネーター	ターミネーター開始	ターミネーター終了	鎖	転写物終了	DNA 配列
cg0001 T1	1628	1647	+	ループ	配列番号 9
cg0007 T2	7504	7529	+	ステム 1	配列番号 10
cg0371 T3	322229	322252	+	ステム 1	配列番号 11
cg0480 T4	421697	421720	-	ステム 1	配列番号 12
cg0494 T5	436587	436608	+	ループ	配列番号 13
cg0564 T6	499895	499917	+	ステム 1	配列番号 14

10

20

30

40

【表 1. 2 - 5】

cg0610 T7	541016	541039	+	ステム 2	配列番号 15
cg0695 T8	613847	613868	-	ループ	配列番号 16

10

【0466】

タンパク質可溶性タグラダー

一部の実施形態では、本開示は、標的遺伝子由来の標的タンパク質と作動可能に連結した1つまたは複数のタンパク質可溶性タグ配列を提供することによって遺伝子操作された宿主株を改良する方法を教示する。可溶性タグは、標的タンパク質のN末端またはC末端のいずれかにおいて標的タンパク質と作動可能に連結した融合パートナーであり得る。一部の実施形態では、本開示は、可溶性タグ配列の付加により、遺伝子操作された宿主において選択された遺伝子から翻訳されたタンパク質の溶解性が改善されることを教示する。他の実施形態では、可溶性タグを使用して標的タンパク質の精製を補助することもできる。

20

【0467】

本開示のタンパク質可溶性タグラダーにおいて使用するための有効なタグは、独立した、十分にフォールディングされたドメイン、高度に可溶性のドメインを形成する当技術分野で公知の任意の可溶性タグであり得る。これらのドメインは、相加効果を通じてそれらの標的タンパク質の溶解性に寄与し得る、またはN末端タグとして使用されるとき、リボソームから出現した後迅速にフォールディングされ、ミスフォールディングを引き起こす可能性がある出現してくる標的タンパク質のアミノ酸の鎖と他の細胞コンポーネントとの相互作用を立体的に遮断し得る。さらに、可溶性タグラダーに含めるための可溶性タグは、小さな密接にフォールディングされたドメインであるまたは高度に可溶性であることが公知のタンパク質のリーダー配列であるなどの共通した性質を有し得る。タンパク質可溶性タグ配列は、例えば、その内容全体が参照によって本明細書に組み込まれるCostaら、Front Microbiol.、2014年；5巻：63頁に見出されるタグのいずれかなどの、当技術分野で公知の任意のこのようなタグであり得る。一実施形態では、可溶性タグ配列は、表17に見出されるタグを含む。

30

【0468】

一実施形態では、タンパク質可溶性タグは、融合パートナーである。融合パートナーコード遺伝子は、本明細書に提供されるベクター（例えば、シャトルベクター）のいずれかに存在し得、したがって、ベクター内の標的タンパク質に対する遺伝子の組み込みにより、融合パートナーコード遺伝子が標的遺伝子に作動可能に連結される。本発明で使用するための可溶性タグを含むE. coli 発現ベクターは、必要に応じて、可溶性タグ融合パートナーコード遺伝子と標的タンパク質コード遺伝子の間に、タグの除去を可能にすることができるプロテアーゼ認識配列を含み得る。本明細書に提供される可溶性スワップ方法において使用するための融合パートナーの選択は以下に依存する。

40

【0469】

(i) 融合の目的：溶解性を改善するためかまたは親和性精製のためか？ 異なる目的を与える様々な融合タグが利用可能であり、迅速な「一段階」タンパク質産生を得るために、例えば二重ヘキサヒスチジン (hexahistidine) (His6) -MBPタグのような可溶性タグおよび親和性タグの両方を含有するシステムを設計することができる。一部のタンパク質タグはまた、例えばMBPまたはグルタチオン-S-トランスフェラーゼ (GST; EspositoおよびChatterjee, Curr Opin Biotechnol.、2006年、8月；17巻

50

(4号) : 353～8頁、Epub、2006年6月15日)のように、親和性および溶解性役割の両方でも機能し得る。

【0470】

(ii) アミノ酸組成およびサイズ：標的タンパク質には、それらの適用に応じて、より大きなまたはより小さなタグが必要になり得る。より大きなタグは、アミノ酸含有量の主要な多様性を示し得、宿主細胞において小さなタグによって課される代謝的荷重とは異なる代謝的荷重を課し得る。

【0471】

(iii) 要求される産生レベル：構造研究には、強力な翻訳開始シグナルを有するより大きな融合タグを用いて急速に達成することができるより高いタンパク質産生レベルが要求され得るが、生理的相互作用の研究では、より低い産生レベルおよび小さなタグが要求され得る。

【0472】

(iv) タグの場所：融合パートナーは、標的タンパク質のN末端またはC末端に位置するとき、異なる効果を促進し得る。N末端タグは、多くの場合、C末端タグよりも有利であり得、これは、(1) N末端タグにより、融合タンパク質がタグ上の効率的な翻訳開始部位を利用する、効率的な翻訳開始のための信頼できる状況がもたらされること；(2) エンドプロテアーゼの大部分はそれらの認識部位のC末端またはその付近を切断するので、N末端タグは標的タンパク質の生来のN末端配列に追加の残基を全くまたはほとんど残さずに除去できることが理由である。

【表17】

表 17. 本開示のタンパク質溶解性タグ配列の非包括的一覧。

溶解性タグ名 (タンパク質 溶解性タグ番 号)	説明	生物	核酸配列番 号	アミノ酸配 列番号	サ イ ズ (aa)
GB1 (PST1)	Protein G の IgG ドメイン B1	Streptococcus sp.	231	235	56
FH8 (PST2)	Fasciola hepatica 8-kDa 抗原	F. hepatica	232	236	69
ユビキチン (PST3)			233	237	
SUMO (PST4)	低分子ユビキ チン修飾	Homo sapiens	234	238	約100

【0473】

タンパク質分解タグラダー

一部の実施形態では、本開示は、標的遺伝子由来の標的タンパク質に作動可能に連結した1つまたは複数のタンパク質分解タグ配列を提供することによって遺伝子操作された宿主株を改良する方法を教示する。本明細書に提供される方法を使用した分解タグ配列の付加により、標的タンパク質に分解の印を付けることができる。標的タンパク質に分解の印を付けることにより、細胞内の標的タンパク質存在量を低減またはモジュレートすることができる。細胞内の標的タンパク質のレベルまたは存在量を低減またはモジュレートすることにより、標的タンパク質への分解タグ配列の付加が結果として生じる株の全般的な表現型に最終的に影響を及ぼし得る。

【0474】

本開示のタンパク質分解タグラダーに使用するための有効なタグは、宿主生物（例えば、E. coli）における公知の分解経路の一部である、当技術分野で公知の任意の分解

タグであり得る。例えば、E. coliにおける公知の分解経路は、clpXP/clpAP系、HflB系、ftsH系およびlon系を含み得る。したがって、本明細書に提供される分解タグスワップ方法において使用するための分解タグは、これらのE. coliタンパク質分解系のいずれかにおいて機能することが分かっている任意のタグを含み得る。一部の場合では、分解タグを、結果として生じる変異体タグにその活性が調整される能力を付与するように変異させることができる。例えば、ssrAクラスのタグを、変異ssrA分解タグにより、タグ付けされたタンパク質にClpXP分解経路を介して異なる程度の効率で分解の印が付けられるように変異させることができる。一例では、ssrAタグは、AANDENYALAAコンセンサス配列の最後の3残基に単一のアミノ酸変異を含有することができ、したがって、C末端に変異ssrAタグを含む標的タンパク質を、いずれのアミノ酸のssrAタグコンセンサス配列の最後の3残基が変異しているかに応じて、ある特定の細胞内尾部特異的プロテアーゼ（例えば、Tspプロテアーゼ）によって種々のレベルの効率で分解することができる（その内容全体が参照によって本明細書に組み込まれているKeiler K C、Sauer R T. Sequence determinants of C-terminal substrate recognition by the Tsp protease. J Biol Chem.、1996年；271巻：2589～2593頁を参照されたい）。よって、本開示の分解タグスワップ方法を使用すると、Tspコンセンサス配列に軽微な変更を伴うC末端ペプチドタグを有するバリエーションを構築することにより、種々の安定性の標的タンパク質を有する宿主細胞の株を得ることが可能である。本明細書の方法において使用するための変異体ssrAタグの例は、アミノ酸配列番号248、249または250を含み得る。

10

20

【0475】

本明細書に提供される方法において使用するための変異ssrAタグの別の例は、その内容全体が参照によって本明細書に組み込まれる、McGinnessら、「Engineering Controllable Protein Degradation」、Mol. Cell、22巻（5号）、2006年6月に見出されるDASタグであり得る。DASタグでは、ssrAタグの2つの残基が置き換えられ、結果として、SspB認識を減弱させることなく、弱まったClpX結合を呈する変異ssrAタグが生じている。したがって、DASタグを有する標的タンパク質は、SspBが存在するとき、ClpXPだけで効率的に分解することができ、それにより、SspBレベルを制御することによって細胞内分解を調節することが可能になる。

【0476】

本開示のタンパク質分解タグ配列の非包括的一覧は、表18に見出すことができる。

30

【表 18】

表 18. 本開示のタンパク質分解タグ配列の非包括的一覧。

分解タグ名(タンパク質分解タグ番号)	説明	生物	核酸配列番号	アミノ酸配列番号	源
ssrA_LAA (PDT1)	生来	<i>E. coli</i>	239	247	Andersen et al., Appl Environ Microbiol. 1998 Jun; 64(6): 2240-2246.
ssrA_LVA (PDT2)	変異体	<i>E. coli</i>	240	248	Andersen et al., Appl Environ Microbiol. 1998 Jun; 64(6): 2240-2246.
ssrA_AAV (PDT3)	変異体	<i>E. coli</i>	241	249	Andersen et al., Appl Environ Microbiol. 1998 Jun; 64(6): 2240-2246.
ssrA_ASV (PDT4)	変異体	<i>E. coli</i>	242	250	Andersen et al., Appl Environ Microbiol. 1998 Jun; 64(6): 2240-2246.
ftsH-clI89-97 (PDT5)	生来	<i>E. coli</i>	243	251	Kobiler O , Koby S , Teff D , Court D , Oppenheim AB. PNAS. 23 Oct 2002, 99(23):14964-14969.
clI08 (PDT6)	生来	<i>E. coli</i>	244	252	Herman et al., Genes Dev. 1998 May 1; 12(9): 1348-1355.
sul20 (PDT7)	生来	<i>E. coli</i>	245	253	Wohlever et al., Protein Eng Des Sel. 2013 Apr; 26(4): 299-305.
β 20 (PDT8)	生来	<i>E. coli</i>	246	254	Wohlever et al., Protein Eng Des Sel. 2013 Apr; 26(4): 299-305.

【0477】

分解タグは、標的タンパク質のN末端またはC末端のいずれかにおいて標的タンパク質と作動可能に連結した融合パートナーであり得る。よって、融合パートナーコード遺伝子は、本明細書に提供されるベクター（例えば、シャトルベクター）のいずれかに、ベクターへの標的タンパク質の遺伝子の組込みにより、融合パートナーコード遺伝子が標的遺伝子に作動可能に連結され、したがって、コンストラクトの翻訳により、所望の通り標的タンパク質のN末端またはC末端のいずれかに分解タグが存在する融合タンパク質が生成するように存在し得る。一実施形態では、分解タグ（またはその変異体）の標的タンパク質のN末端またはC末端への位置付けは、使用されるタグに依存し得る。例えば、clpXP/clpAP系、HflB系、ftsH系またはlon系のsul20タグに関連する分解タグ（またはその変異体）は、標的タンパク質のC末端に作動可能に連結することができ、lon系の β 20分解タグ（またはその変異体）は、標的タンパク質のN末端または内部に作動可能に連結することができる。一実施形態では、分解タグは、その全体が参照によって本明細書に組み込まれているSekar K、Gentile AM、Bostick JW、Tyo KEJ（2016年）N-Terminal-Based Targeted, Inducible Protein Degradation in

Escherichia coli. PLoS ONE、11巻(2号): e0149746 頁に見出される、*E. coli* に対するN-タグロンタグ(Ntag)である。Ntagは、目的の標的タンパク質のN末端に位置付けることができ、目的の標的タンパク質に*E. coli* 宿主細胞におけるclpXP/clpAP系を介した分解の印を付けるために役立ち得る。別の実施形態では、分解タグは、その全体が参照によって本明細書に組み込まれているButzら、*Biochemistry*、2011年、50巻(40号)、8594~8602 頁に記載されている通り標的タンパク質のN末端に位置させることができるRepAタグである。N末端RepAタグは、目的の標的タンパク質に*E. coli* 宿主細胞におけるclpXP/clpAP系を介した分解の印を付けるために役立ち得る。

【0478】

仮説主導型多様性プールおよびヒルクライミング

本開示は、本開示のHTPゲノム操作方法は、宿主細胞性能の有意な獲得を達成するのに先の遺伝的知識を必要としないことを教示する。実際に、本開示は、ランダム変異誘発、および既存の宿主細胞バリエーションの中での遺伝的多様性の同定(例えば、野生型宿主細胞と工業用バリエーションとの間の比較など)を含めたいくつかの機能にとらわれない(functionally agnostic)手法を介して多様性プール(例えば、図1)を生成する方法を教示する。

【0479】

しかし、一部の実施形態では、本開示は、下流のHTP操作に使用される遺伝的多様性変異を設計する仮説主導方法も教示する。すなわち、一部の実施形態では、本開示は、選択された変異の指向設計を教示する。一部の実施形態では、指向変異が、本開示の操作ライブラリー(例えば、SNPスワップ、PROスワップ、STOPスワップ、SOLUBILITY TAGスワップ、またはDEGRADATION TAGスワップ)中に組み入れられる。

【0480】

一部の実施形態では、本開示は、遺伝子アノテーション、仮定された(もしくは確認された)遺伝子機能、またはゲノム内の場所に基づく指向変異の創製を教示する。本開示の多様性プールは、文献において宿主細胞の性能の増大に関連した特異的代謝経路または遺伝経路に関与すると仮定された遺伝子内の変異を含み得る。他の実施形態では、本開示の多様性プールは、宿主性能の改善に関連したオペロン内に存在する遺伝子への変異も含み得る。さらに他の実施形態では、本開示の多様性プールは、アルゴリズムが予測した機能または他の遺伝子アノテーションに基づく遺伝子への変異も含み得る。

【0481】

一部の実施形態では、本開示は、仮説主導型変異の標的に優先順位をつけるための「シェル」ベースの手法を教示する。標的優先順位付けのシェルメタファーは、少数の主要遺伝子のみが宿主細胞の性能の特定の態様(例えば、単一生体分子の産生)のほとんどを担うという仮説に基づく。これらの主要遺伝子は、シェルのコアに位置しており、その後、第2の層における二次効果遺伝子、第3のシェルにおける三次効果などが続く。例えば、一実施形態では、シェルのコアは、選択された代謝経路内の極めて重要な生合成酵素(例えば、クエン酸の産生)をコードする遺伝子を含み得る。第2のシェルに位置した遺伝子は、産物転換またはフィードバックシグナル伝達を担う生合成経路内の他の酵素をコードする遺伝子を含み得る。この例示的なメタファー下の第3の段の遺伝子は、生合成経路の発現のモジュレーションまたは宿主細胞内の全般的な炭素フラックスの調節を担う調節遺伝子を含みやすい。

【0482】

本開示は、あらゆる同定された変異からの性能獲得を最適化するための「ヒルクライム」法も教示する。一部の実施形態では、本開示は、HTP多様性ライブラリー内のランダム、天然の、または仮説主導型変異が、宿主細胞性能に関連した遺伝子の同定をもたらすことができることを教示する。例えば、本方法は、遺伝子コード配列に、またはその付近に位置した1つまたは複数の有益なSNPを同定することができる。この遺伝子は、宿主

10

20

30

40

50

細胞性能に関連している場合があり、その同定は、生物体のコンビナトリアル遺伝子変異空間における性能「ヒル」の発見に類似し得る。

【0483】

一部の実施形態では、本開示は、SNP変異において具現化された同定されたヒル周囲のコンビナトリアル空間を探索する方法を教示する。すなわち、一部の実施形態では、本開示は、その遺伝子ノード（すなわち、ヒルクライミング）から得られる性能獲得を最適化するための同定された遺伝子および関連した調節配列の攪乱を教示する。よって、本開示の方法によれば、遺伝子は、ランダム変異誘発から供給された多様性ライブラリー内で最初に同定することができるが、同じ遺伝子内の別の配列の指向変異によって株改良プログラムにおいて使用するために後に改善され得る。

10

【0484】

ヒルクライミングの概念は、単一遺伝子配列を圍繞するコンビナトリアル空間の探索を超えて拡張することもできる。一部の実施形態では、特定の遺伝子内の変異は、宿主細胞性能に対する特定の代謝経路または遺伝経路の重要性を明らかにすることができる。例えば、一部の実施形態では、単一RNA分解遺伝子内の変異が有意な宿主性能獲得をもたらしたという発見を、宿主生物から追加の性能獲得を引き出すための手段として関連したRNA分解遺伝子を変異させるための基礎として使用することができる。当業者は、上記のバリエーションが指向遺伝子設計へのシェルおよびヒルクライム手法を記述することを認識する。

【0485】

20

生合成経路足場

一部の実施形態では、本開示は、一部のバイオ産業プロセスの生産性は宿主細胞内の基質、中間体、および生合成酵素のランダムな拡散によって限定されることを教示する。一部の実施形態では、本開示は、宿主細胞培養の生産性を経路内に生合成酵素を共局在化させることによって増大させることができることを教示する。よって、一部の実施形態では、本開示は、生合成酵素のDNAまたはタンパク質足場などの足場への係留を教示する。

【0486】

一部の実施形態では、共局在化は、DNA結合ドメインを経路内の生合成酵素と組換え融合し、次いでそれをDNA足場領域に結合させ、したがって、経路酵素を細胞内で互いの近くに制約することによって達成される。他の実施形態では、共局在化は、タンパク質結合ドメインを経路内の生合成酵素と組換え融合し、次いでそれをタンパク質足場領域に結合させ、したがって、経路酵素を細胞内で互いの近くに制約することによって達成される。一部の実施形態では、共局在化により、産生の速度が増大し、細胞内の経路中間体の濃度が低下する（図44を参照されたい）。

30

【0487】

一部の実施形態では、本開示は、*Escherichia coli*のゲノムを操作するためのハイスループット方法であって、DNA結合ドメインまたはタンパク質結合ドメインをコードするヌクレオチド配列を生合成経路内の酵素をコードする遺伝子に挿入し、DNA足場プラスミドまたは足場タンパク質を細胞に導入する、方法を教示する。本発明の一態様によれば、生合成遺伝子に係留されたDNAまたはタンパク質結合ドメインにより、組換え経路酵素と一緒に足場プラスミドまたはペプチド上に局在化し、したがって、標的産物の生産性の改善が導かれると考えられる。

40

【0488】

一部の実施形態では、本発明は、ハイスループット方法によって操作されたゲノムを有する*E. coli*細胞における低分子の拡散律速生産性の問題を解決する。現在、生合成酵素を局在化させるためのDNA足場の唯一の報告された例は、組換え経路酵素がプラスミド上にコードされるロースループットプロセスである（Leeら、「Improved Production of L-Threonine in *Escherichia coli* by Use of a DNA Scaffold System」*App. And Environ. Microbiol.*、79巻（3号）、774～782頁（2013年））。一部の実施形態では、本発明は、DNA結合ドメインを染色体性にコードされる経

50

路酵素にハイスループットな様式で組み入れるための手段を提供する。

【0489】

一部の実施形態では、本開示は、キメラ生合成酵素ならびに足場DNAおよびタンパク質を教示する。この技術の様々な態様を以下により詳細に考察する。

【0490】

DNA結合キメラタンパク質

一部の実施形態では、本開示は、DNA結合ドメインに係留された選択された生合成酵素を含むキメラタンパク質を教示する。これらの実施形態によれば、キメラ生合成酵素がそれらのDNA結合ドメインによってDNA足場に補充され、それにより、様々な生合成活性が宿主細胞のある領域に濃縮されることが予想される。

10

【0491】

一部の実施形態では、生合成酵素とDNA結合ドメインは共有結合により連結している。一部の実施形態では、生合成酵素はDNA結合ドメインに翻訳的に (translationally) 融合する。よって、一部の実施形態では、DNA結合ドメインと生合成経路タンパク質のアミノ末端、カルボキシ末端、または内部の部位にカップリングすることによってキメラ生合成酵素を形成する。DNA結合ドメインの付加により生合成酵素の活性が実質的に低下しないことを確実にする必要性が当業者には認識されよう。

【0492】

本開示の一部の実施形態では、生合成酵素をそのDNA結合ドメインに短いポリペプチドリンカー配列を介してカップリングさせる。適当なリンカーは、長さが約6から約40アミノ酸の間のペプチドを含む。好ましいリンカー配列は、グリシンリッチ (例えば、G3~5)、セリンリッチ (例えば、GSG、GSGS (配列番号18)、GSGSG (配列番号19)、GSNG (配列番号20)、またはアラニンリッチ (例えば、TSAAG (配列番号21)) リンカー配列を含む。他の例示的なリンカー配列は、例えば、AAGGM (配列番号22) ; AAGGMPPAAGGM (配列番号23) ; AAGGM (配列番号24) ; およびPPAAGGMM (配列番号25) など、グリシン、アラニン、プロリンおよびメチオニン残基の組合せを有する。リンカーは、概して柔軟なキメラ生物学的経路タンパク質をもたらす事実上あらゆる配列を有し得る。

20

【0493】

一部の実施形態では、本開示の方法は、生合成酵素とシスで機能することができるあらゆるDNA結合ドメインと適合する。一部の実施形態では、DNA結合ドメインは、宿主生物に対して外因性であることが好ましい。他の実施形態では、本開示は、設計された足場DNAの外側への過剰な結合を回避するために十分に選択的であるDNA結合ドメインの選択を教示する。

30

【0494】

様々なDNA結合ドメインが、それらの対応するDNA内のヌクレオチド認識部位 (すなわち、DNA結合部位) と一緒に当技術分野で公知であり、本発明のシステムおよび方法における使用に適当である。例えば、本発明の一実施形態では、キメラ生物学的経路タンパク質のDNA結合部分は、ロイシンジッパーDNA結合ドメインを含み、ここで、足場は、対応するロイシンジッパーDNA結合配列を含む。本発明の別の実施形態では、キメラ生物学的経路タンパク質のDNA結合部分は、ヘリックス・ループ・ヘリックスDNA結合ドメインを含み、ここで、足場は、対応するヘリックス・ループ・ヘリックスDNA結合配列を含む。別の実施形態では、キメラ生物学的経路タンパク質のDNA結合部分は、ウィングドヘリックスDNA結合ドメインを含み、ここで、足場は、対応するウィングドヘリックスDNA結合配列を含む。別の実施形態では、キメラ生物学的経路タンパク質のDNA結合部分は、ウィングドヘリックス・ターン・ヘリックスDNA結合ドメインを含み、ここで、足場は、対応するウィングドヘリックス・ターン・ヘリックスDNA結合配列を含む。別の実施形態では、キメラ生物学的経路タンパク質のDNA結合部分は、ヘリックス・ターン・ヘリックスDNA結合を含み、ここで、足場は、対応するヘリックス・ターン・ヘリックスDNA結合配列を含む。別の実施形態では、キメラ生物学的経路

40

50

タンパク質のDNA結合部分は、HMG-ボックスDNA結合ドメインを含み、ここで、足場は、対応するHMG-ボックスDNA結合配列を含む。別の実施形態では、キメラ生物学的経路タンパク質のDNA結合部分は、特注設計されたTALE DNA結合ドメインを含み、ここで、足場は、対応する設計されたTALE DNA結合配列を含む。本発明の別の実施形態では、キメラ生物学的経路タンパク質のDNA結合部分は、ジンクフィンガーDNA結合ドメインを含み、ここで、足場は、対応するジンクフィンガーDNA結合配列を含む。

【0495】

例示的なジンクフィンガーDNA結合ドメイン配列および対応するDNA結合部位を表1. 3に提供する。当技術分野で公知の他のジンクフィンガーDNA結合ドメインおよびそれらの対応する標的DNA結合配列も本発明における使用に相当である（例えば、それぞれの全体が参照により本明細書に組み込まれている、Greisman H AおよびPabo C O、「A General Strategy for Selecting High-Affinity Zinc Finger Proteins for Diverse DNA Target Sites」、Science、275巻：657～661頁（1997年）、Rebar E JおよびPabo C O、「Zinc Finger Phage: Affinity Selection of Fingers with New DNA-Binding Specificities」、Science、263巻：671～673頁（1994年）；Maederら、「Rapid "Open-Source" Engineering of Customized Zinc-Finger Nucleases for Highly Efficient Gene Modification」、Mol. Cell.、31巻：294～301頁（2008年）、Sanderら、「Selection-Free Zinc-Finger-Nuclease Engineering by Context-Dependent Assembly (CoDA)」、Nat. Methods、8巻：67～69頁（2011年）、Rebarに対する米国特許第5, 5789, 538号、Greismanに対する米国特許第6, 410, 248号、Rebarに対する米国特許第7, 605, 140号、Barbasに対する米国特許第6, 140, 081号、Barbasに対する米国特許第7, 067, 617号、Michaelsに対する米国特許第6, 205, 404号、およびJoungに対する米国特許出願公開第20070178454号を参照されたい）。

【0496】

ジンクフィンガードメインのDNA結合特異性を最適化するための方法および合成DNA結合部位を操作する方法も当技術分野で公知であり、本発明において、新しいジンクフィンガー結合パートナー生成するために利用することができる（例えば、そのそれぞれの全体が参照によって本明細書に組み込まれる、Bulykら、「Exploring the DNA-binding Specificities of Zinc Fingers with DNA Microarrays」、Proc. Nat'l Acad. Sci. U.S.A.、98巻（13号）：7158～63頁（2001年）およびHurtら、「Highly Specific Zinc Finger Proteins Obtained by Directed Domain Shuffling and Cell-based Selection」、Proc. Nat'l Acad. Sci. U.S.A.、100巻（21号）：12271～6頁（2003年）、Rebarに対する米国特許第5, 5789, 538号、Greismanに対する米国特許第6, 410, 248号、Rebarに対する米国特許第7, 605, 140号、Barbasに対する米国特許第6, 140, 081号、Barbasに対する米国特許第7, 067, 617号、Michaelsに対する米国特許第6, 205, 404号、およびJoungに対する米国特許出願公開第20070178454号を参照されたい）。

。

【表 1. 3 - 1】

表1.3 - DNA結合ドメインの限定されない一覧。

ジンク フィンガー	DNA 結合ドメイン配列	DNA 結合配列 (5'→3')
Zif268	PGEKPYACPVESECDRRFSRSDDELTRHIRIHTGQKPFQC RICMRNFSRSDHLTTHIRIHTGKPFACDICGRKFARS DERKRHTKIHT (配列番号 26)	GCGTGGGCG GCG GGG GCG
PBSII	PGEKPYACPECGKSFSQRANLRAHQRTHTGEKPYKC PECGKSFSRSDHLTTHQRTHTGEKPYKCPECGKSFSR SDVLVRHQRTHT (配列番号 27)	GTGTGGAAA

【表 1. 3 - 2】

ジンク フィンガー	DNA 結合ドメイン配列	DNA 結合配列 (5'→3')
ZFa	PGERPFQCRICMRNFSDSPTLRRHTRTHTGEKPFQCRI CMRNFSVRHNLTRHLRTHHTGEKPFQCRICMRNFSDRT SLARHLKTH (配列番号 28)	GTCGATGCC
ZFb	PGERPFQCRICMRNFSSKDLHRHTRTHTGEKPFQCR ICMRNFSLSQTLKRHLRTHHTGEKPFQCRICMRNFSRL DMLARHLKTH (配列番号 29)	GCGGCTGGG
ZFc	PGERPFQCRICMRNFSSPSKLIRHTRTHTGEKPFQCRI MRNFSDGSNLARHLRTHHTGEKPFQCRICMRNFSRVD NLPRHLKTH (配列番号 30)	GAGGACGGC
Tyr123	EKPYKCPECCKSFSRSDSNLTRHQRTHTGEKPYKCPEC GKSFSSTSNLARHQRTHTGEKPFKCPECCKSFSRSDA LTRHQRTHT (配列番号 31)	GTGGATGAC
Tyr456	EKPYKCPECCKSFSQSSNLARHQRTHTGEKPYKCPEC GKSFSRSDHLTKHQRTHTGEKPFKCPECCKSFSQSSN LARHQRTHT (配列番号 32)	GAAGGGGAA
Blues	ASDDRPYACPVESECDRRFSRRDVLNMNHIRHTGQKPF QCRICMRNFSRSDHLTTHIRTHHTGEKPFACDICGRKFA NRDTLTRHSKIHLRQNDLE (配列番号 33)	GTTTGGATG
Jazz	ASDDRPYACPVESECDRRFSRSDVLRHNRTHHTGEKPFQ CRICMRNFSRSDVLRHNRTHHTGEKPFACDICGRKFA SRDVLRRHNRHILRQNDLE (配列番号 34)	GCTGCTGCG

10

20

30

40

【表 1. 3 - 3】

ジンク フィンガー	DNA 結合ドメイン配列	DNA 結合配列 (5'→3')
Bagly	EFMTGDRPYACPVESCDRRFSRSDDELTRHIRIHTGQKP FQCRICMRNFSRDVLRHRNRTHTGEKPFACDICGRK FASRDVLRHRNRIHLRQGRSHVCAECGKAFVESSKLK RHQLVHTGEKPFQLE (配列番号 35)	CGGGCTGCTGC (配列番号 36)
Gli1	KREPESVYETDCRWDGCSQEFDSQEQLVHHINSEHH GERKEFVCHWGGCSRELRPFKAQYMLVVHMRRTHTG EKPHKCTFEGCRKSYSRLNLKTHLRSHHTGEKPYMCE HEGCSKAFSNASDRAKHQNRTHSNEKPYVCKLPGCT KRYTDPSSLRKHVKTVHGPDAAHVTKRHRGD (配列番号 37)	GACCACCCAAG ACGA (配列番号 38)
HIVC	PFQCRICMRNFSRLTDLDRHTRTHTGEKPFQCRICMR NFSLSQTLRRHLRTHHTGEKPFQCRICMRNFSRLSNLG RHLKTHHTGEK (配列番号 39)	GATGCTGCA
B3	AQAALPKEKPYACPECCKSFSDPGNLVRHQRTHTG EKPYKCPECCKSFSRSDKLVRHQRTHTGEKPYKCPEC GKSFSQSSHLVRHQRTHTGKKTSGQAG (配列番号 40)	GACGGGGG
N1	AQAALPKEKPYACPECCKSFSQSSSLVRHQRTHTGE KPYKCPECCKSFSQSSNLVRHQRTHTGEKPYKCPEC KSFSRSDKLVRHQRTHTGKKTSGQAG (配列番号 41)	GTAGAAGGG

10

20

30

【表 1. 3 - 4】

ジンク フィンガー	DNA 結合ドメイン配列	DNA 結合配列 (5'→3')
Sp-1	PGKKKQHIHCHIQGCGKVYGKTSHLRAHLRWHTGERP FMCTWSYCGKRFRTRSDDELQRHQRTHITGEKKFACPEC PKRFRMRSDHLSKHIKTHQNKKG (配列番号 42) PGKKKQHACPECCKSFSKSSHLRAHQRTHTGERPYK CPECCKSFSRSDELQRHQRTHTITGEKPYKCPECCKSFS RSDHLSKHQRTHQNKKG (配列番号 43)	GGGGCGGGG

10

【0497】

核酸足場配列

一部の実施形態では、本開示は、キメラ生合成酵素内に含有されるDNA結合ドメインに対応するDNA結合配列の1つまたは複数を含むDNA足場を教示する。一部の実施形態では、DNA足場は、染色体外プラスミドまたは他のベクターである。他の実施形態では、DNA足場は、宿主細胞のゲノム内にコードされる。

20

【0498】

適当な核酸ベクターとしては、限定することなく、プラスミド、バキュロウイルスベクター、バクテリオファージベクター、ファージミド、コスミド、フォスミド、細菌人工染色体、ウイルスベクター（例えば、ワクシニアウイルス、ポリオウイルス、アデノウイルス、アデノ随伴ウイルス、SV40、単純ヘルペスウイルスなどに基づくウイルスベクター）、人工染色体、酵母プラスミド、酵母人工染色体、および他のベクターが挙げられる。本発明の一部の実施形態では、原核生物宿主細胞における使用に適当なベクターが好ましい。したがって、Escherichia coliなどの原核生物において使用するための例示的なベクターとして、これらに限定されないが、pACYC184、pBel 30 oBac11、pBR332、pBAD33、pBBR1MCSおよびその誘導体、pSC101、SuperCos（コスミド）、pWE15（コスミド）、pTrc99A、pBAD24、ColE1複製起点およびその誘導体を含むベクター、pUC、pBluescript、pGEM、Ori_P1smd27（配列番号213）、ベクター骨格1（配列番号214）、ベクター骨格2（配列番号215）、ベクター骨格3（配列番号216）、ベクター骨格4（配列番号217）、ならびにpTZベクターが挙げられる。

30

【0499】

一部の実施形態では、本開示は、核酸足場サブユニットは、カップリングし合って2、3、4、5、6、7、8、9、10、15、20、25、またはそれよりも多くの異なる生合成経路タンパク質の結合および固定化を促進する2、3、4、5、6、7、8、9、10、15、20、25、またはそれよりも多くの異なるDNA結合部位を含み得ることを教示する。一部の実施形態では、本開示は、DNA足場が対応するキメラ生合成タンパク質それぞれに対して単一のDNA結合部位を有することを教示する。

40

【0500】

他の実施形態では、核酸足場は、同じDNA結合部位の2つまたはそれよりも多くのコピーを含み得る。このアーキテクチャにより、生物学的タンパク質化学量論比の最適化を達成することが可能になる。本発明の本実施形態によれば、同じDNA結合部位をカップリングし合わせて、特定の化学反応のための酵素中心を創製することができる。よって、一部の実施形態では、DNA足場は、多数のDNA結合部位の群を含み、各群が特定のキ

50

メラ生合成遺伝子／酵素に対応する。

【0501】

本発明の一部の実施形態では、合成生物学的経路をアセンブルする方法は、少なくとも第1のキメラ生合成遺伝子（例えば、酵素）および第2のキメラ生合成酵素を核酸足場に固定化するステップを伴う。第1のキメラ生合成酵素により、第2のキメラ生物学的経路タンパク質の基質である第1の産物が産生される。第2のキメラ生合成酵素を足場コンストラクトに、第1のキメラ生合成酵素に隣接してまたはすぐ近くに位置するように固定化する。このように、第1の産物の有効濃度は高く、第2のキメラ生合成酵素は第1の産物に対して効率的に作用し得る。例として、合成核酸足場には、足場コンストラクトの3'→5'または5'→3'の順に、a) 第1のキメラ生合成酵素、およびb) 第2のキメラ生合成酵素が固定化されて、「足場サブユニット」が形成されている。足場サブユニットを合成核酸足場内で2回またはそれよりも多く繰り返すことができる。

10

【0502】

本発明のこの態様およびすべての態様によれば、各キメラ生合成酵素の2つまたはそれよりも多くのコピー（例えば、2、3、4、5、6、7、8、9、10、またはそれよりも多くの分子）を足場サブユニットに固定化することができる。例えば、一部の実施形態では、足場サブユニットには、a) 第1のキメラ生合成酵素の1つの分子（コピー）およびb) 第2のキメラ生合成酵素の1つの分子が固定化されている。他の実施形態では、足場サブユニットには、a) 第1のキメラ生合成酵素の1つの分子およびb) 第2のキメラ生合成酵素の2つまたはそれよりも多くの分子（例えば、2、3、4、5、6、またはそれよりも多くの分子）が固定化されている。よって、生物学的経路内の任意の所与のタンパク質の経路内の任意の他のタンパク質に対する比は変動し得る。単に例として、第1のキメラ生物学的経路タンパク質と第2のキメラ生物学的経路タンパク質の比は、約0.1:10から約10:0.1まで、例えば、約0.1:10から約0.5:10まで、約0.5:10から約1.0:10まで、約1.0:10から約2:10まで、約2:10から約5:10まで、約5:10から約7:10まで、約7:10から約10:10まで、約10:7から約10:5まで、約10:5から約10:2まで、約10:2から約10:1まで、約10:1から約10:0.5まで、または約10:0.5から約10:1まで変動し得る。

20

【0503】

一部の実施形態では、少なくとも3つのキメラ生合成酵素を合成核酸足場に固定化して、足場サブユニットを構成する。本発明の本実施形態によれば、第1のキメラ生合成酵素により、第2のキメラ生合成酵素の基質である第1の産物が産生され、第2のキメラ生物学的経路タンパク質により、第3のキメラ生合成酵素の基質である第2の産物が産生される。これらの実施形態では、足場サブユニットには、足場の3'→5'または5'→3'の順に、a) 第1のキメラ生合成酵素、b) 第2のキメラ生合成酵素、およびc) 第3の生合成酵素が固定化されている。上記の通り足場単位を核酸コンストラクト内で2回またはそれよりも多く繰り返すことができる。

30

【0504】

本発明の別の実施形態では、少なくとも4つのキメラ生合成酵素を核酸足場に固定化する。本発明の別の実施形態では、少なくとも5つのキメラ生合成酵素を核酸足場に固定化する。これらの実施例から、第6の、第7の、第8の、第9の、第10のなどのキメラ生合成酵素を核酸足場に固定化できること、キメラタンパク質はそれらが経路において機能する順序で空間的に固定化されること、および各タンパク質を足場に1、2、3、4、5、6、7、8、9、10、またはそれよりも多くのコピー（または分子）で固定化できることが明らかになるであろう。

40

【0505】

一部の実施形態では、本開示は、足場核酸内の各DNA結合部位の間隔を教示する。本発明のこの態様によれば、2つまたはそれよりも多くのDNA結合部位は、足場サブユニット内で互いに隣接して、互いにタンデムにカップリングして、または少なくとも1つの

50

スパーサーヌクレオチドによって隔てられて、位置する。2つまたはそれよりも多くのDNA結合部位は、互いに2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、またはそれよりも多くのスパーサーヌクレオチドで隔てることができる。異なるDNA結合部位間の間隔は、1足場単位内で変動し得る（すなわち、第1および第2のDNA結合部位の間隔は、第2および第3のDNA結合部位の間隔とは異なり得る）。足場サブユニット内の異なるDNA結合部位間の最適な間隔は、生合成酵素要件および再構築される生物学的経路に応じて変動し、最適な生物学的経路生産性が達成されるように最適化すべきである。

【0506】

ペプチド足場

一部の実施形態では、本開示の足場方法は、細胞内のタンパク質／構造足場にも適用することができる。一部の実施形態では、本開示は、その全体が参照によって本明細書に組み込まれている米国特許出願公開第20110008829号に開示されている方法の適用を教示する。

【0507】

タンパク質結合キメラタンパク質

一部の実施形態では、本開示は、動員ペプチドに結合することができる1つまたは複数のタンパク質結合ドメイン（複数可）に係留された選択された生合成酵素を含むキメラタンパク質を教示する。これらの実施形態によれば、キメラ生合成酵素は、足場ペプチド内に含有される動員ペプチドと相互作用することによって足場ペプチドに補充されることが予想される。

【0508】

一部の実施形態では、生合成酵素とタンパク質結合ドメインは共有結合により連結している。一部の実施形態では、生合成酵素はタンパク質結合ドメインに翻訳的に融合する。したがって、一部の実施形態では、タンパク質結合ドメインを生合成経路タンパク質のアミノ末端、カルボキシ末端、または内部の部位にカップリングすることによってキメラ生合成酵素を形成する。タンパク質結合ドメインの付加により生合成酵素の活性が実質的に低下しないことを確実にする必要性が当業者には認識されよう。

【0509】

本開示の一部の実施形態では、本開示の先の部分に記載の通り、生合成酵素をそのタンパク質結合ドメインに短いポリペプチドリンカー配列を介してカップリングさせる。

【0510】

様々なタンパク質結合ドメイン（PBD）が、それらの対応する動員ペプチド配列と一緒に当技術分野で公知であり、本発明のシステムおよび方法における使用に適當である。適當なPBDの限定されない例示的考察が続く。

【0511】

SH3

【0512】

適當なPBDは、SH3ドメインを含む。SH3ドメインとしては、クラスI SH3ドメイン；クラスII SH3ドメイン；および非慣習的SH3ドメインが挙げられる。SH3ドメインのアミノ酸配列は当技術分野で公知である。例えば、GenBank受託番号NP_058431（Homo sapiens Crkタンパク質）で提供されるアミノ酸配列のアミノ酸136～189；GenBank受託番号AAH31149で提供されるアミノ酸配列のアミノ酸136～189（Mus musculus Crkタンパク質）；およびGenBank受託番号P27986で提供されるアミノ酸配列のアミノ酸4～77（Homo sapiens ホスファチジルイノシトール3-キナーゼのp85サブユニット）を参照されたい。

【0513】

一部の実施形態では、SH3ドメインは、クラスI SH3ドメインであり、これらの

10

20

30

40

50

間のすべての範囲および部分範囲を含め、アミノ酸配列：EGYQYRA L Y D Y K K E R E E D I D L H L G D I L T V N K G S L V A L G F S D G Q E A R P E E I G W L N G Y N E T T G E R G D F P G T Y V E Y I（配列番号44）に対して少なくとも約75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、または100%のアミノ酸配列類似性を有するアミノ酸配列を含む。

【0514】

一部の実施形態では、SH3ドメインは、クラスII SH3ドメインであり、アミノ酸配列：YVRALFDFNGNDEEDLPFKKGDILRIRDKPEEQWWNAEDSEGKRGMI PVPYVEK（配列番号45）に対して少なくとも約75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、または100%のアミノ酸配列類似性を有するアミノ酸配列を含む。1つの限定されない例として、SH3ドメインは、これらの間のすべての範囲および部分範囲を含め、アミノ酸配列：MAEYVRALFDFNGNDEEDLPFKKGDILRIRDKPEEQWWNAEDSEGKRGMI PVPYVEKY（配列番号46）を含む。

【0515】

SH3ドメインは、左巻きポリプロリントタイプIIヘリックスを形成するプロリンリッチペプチドに結合し、ここで、このようなペプチドは、最小のコンセンサス配列Pro-X-Proを含む。一部の実施形態では、各Proには脂肪族残基が先行する。一部の実施形態では、動員ペプチドは、SH3ドメインリガンドである。SH3ドメインは、左巻きポリプロリントタイプIIヘリックスを形成するプロリンリッチペプチドに結合し、ここで、このようなペプチドは、最小のコンセンサス配列Pro-X-X-Proを含む。一部の実施形態では、各Proには脂肪族残基が先行する。SH3ドメインリガンドを含むペプチドのアミノ酸配列の例示的な限定されない例としては、RPLPVAP（配列番号47；クラスI SH3ドメインが結合する）；PPPALPPKRRRPPG（配列番号48）；およびPPPALPPKKR（配列番号49；クラスII SH3ドメインが結合する）が挙げられる。

【0516】

PDZ

適当なPBDは、PDZドメインを含む。PDZドメインのアミノ酸配列は当技術分野で公知である。例えば、Gen Bank受託番号AAC52113（Homo sapiens シナプス後密度タンパク質95）で提供されるアミノ酸配列のアミノ酸108～191、アミノ酸201～287、およびアミノ酸354～434ならびにGen Bank受託番号NP_033254（Mus musculus シントロフィン）で提供されるアミノ酸配列のアミノ酸80～161を参照されたい。

【0517】

一部の実施形態では、適当なPDZドメインは、これらの間のすべての範囲および部分範囲を含め、アミノ酸配列：EITLERGNSGLGFSSIAGGTDNPHIGDDPSIFIT KII PGGA A A Q D G R L R V N D S I L F V N E V D V R E V T H S A A V E A L K E A G S I V R L Y V（配列番号50）に対して少なくとも約75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、または100%のアミノ酸配列類似性を有するアミノ酸配列を含む。

【0518】

一部の実施形態では、適当なPDZドメインは、これらの間のすべての範囲および部分範囲を含め、アミノ酸配列：VMEIKLIKGP KGLGFSSIAGGVGNQH I P G

DN S I Y V T K I I E G G A A H K D G R L Q I G D K I L A V N S V G L E D V
M H E D A V A A L K N T Y D V V Y L K V A (配列番号51)に対して少なくとも約7
5%、76%、77%、78%、79%、80%、81%、82%、83%、84%、8
5%、86%、87%、88%、89%、90%、91%、92%、93%、94%、9
5%、96%、97%、98%、99%、または100%のアミノ酸配列類似性を有する
アミノ酸配列を含む。

【0519】

一部の実施形態では、適当なPDZドメインは、これらの間のすべての範囲および部分
範囲を含め、アミノ酸配列：R I V I H R G S T G L G F N I V G G E D G E G I F I S
F I L A G G P A D L S G E L R K G D Q I L S V N G V D L R N A S H E Q A A I A L
K N A G Q T V T I I A Q (配列番号52)に対して少なくとも約75%、76%、77
%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87
%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97
%、98%、99%、または100%のアミノ酸配列類似性を有するアミノ酸配列を含む
。

10

【0520】

一部の実施形態では、適当なPDZドメインは、これらの間のすべての範囲および部分
範囲を含め、アミノ酸配列：R R V T V R K A D A G G L G I S I K G G R E N K M P I
L I S K I F K G L A A D Q T E A L F V G D A I L S V N G E D L S S A T H D E A V
Q A L K K T G K E V V L E V K (配列番号53)に対して少なくとも約75%、76%
、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%
、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%
、97%、98%、99%、または100%のアミノ酸配列類似性を有するアミノ酸配列
を含む。例えば、PDZドメインは、アミノ酸配列M L Q R R R V T V R K A D A G G L
G I S I K G G R E N K M P I L I S K I F K G L A A D Q T E A L F V G D A I L S V
N G E D L S S A T H D E A V Q A L K K T G K E V V L E V K Y M K E V S P Y F K
G S (配列番号54)を含み得る。

20

【0521】

一部の実施形態では、動員ペプチドは、PDZドメインリガンドである。PDZドメイ
ンは、標的タンパク質のC末端の4～5残基に結合する。一部の実施形態では、コンセン
サスPDZドメインリガンドは、カルボキシル末端に疎水性残基、例えば、V a lまたは
I l eを含む。PDZドメインリガンドを含むペプチドのアミノ酸配列の例示的な限定さ
れない例としては、I E S D V (配列番号55)；V K E S L V (配列番号56)；G V
K E S L V (配列番号57)；G V K Q S L L (配列番号58)；G V K E S G A (配列
番号59)；Y V K E S L V (配列番号60)；およびV E T D V (配列番号61)が挙
げられる。

30

【0522】

G B D

適当なPBDは、当技術分野においてC R I B (C d c 4 2 / R a c 相互作用的結合)
モチーフとも呼ばれるG T Pアーゼ結合ドメイン (G B D)を含む。一部の実施形態では
、G B Dは、C d c 4 2 p様および／またはR h o様低分子G T Pアーゼに結合する。G
B Dのアミノ酸配列は当技術分野で公知である。例えば、GenBank受託番号NP_—
001103835 (R a t t u s n o r v e g i c u s ウィスコット・アルドリッチ
症候群様タンパク質 (W A S P))で提供されるアミノ酸配列のアミノ酸198～240
；GenBank受託番号Q13177 (H o m o s a p i e n s P A K-2)で提
供されるアミノ酸配列のアミノ酸69～112；およびGenBank受託番号P354
65 (R a t t u s n o r v e g i c u s P A K-1)で提供されるアミノ酸配列の
アミノ酸70～105を参照されたい。Garrardら (2003年) EMBO J.、22巻：1
125頁の図3Aに示されているアミノ酸配列P A K (75～111)、A C K (504
～549)、およびW A S P (232～274)も参照されたい。BishopおよびHall (2

40

50

000年) Biochem. J., 348巻: 241頁の図1Aに示されているアミノ酸配列ACK(505~531)、WASP(236~258)、PAK1(70~94)、PAK2(71~91)、PAK~4(6~30)も参照されたい。

【0523】

一部の実施形態では、適当なGBDは、アミノ酸配列: ADI GTPSNFQHIG
HVGWDPNTGF DLNNLDPELK NLFDMCGISE (配列番号62)
)、およびこれらの間のすべての範囲または部分範囲に対して少なくとも約75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、または100%のアミノ酸配列類似性を有するアミノ酸配列を含む。

10

【0524】

一部の実施形態では、適当なGBDは、アミノ酸配列: KERPEISLPSEFETH
TIVGVGFDVTFGEFTGMPEQWAR (配列番号63)、およびこれらの間の
すべての範囲または部分範囲に対して少なくとも約75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、または100%のアミノ酸配列類似性を有するアミノ酸配列を含む。

【0525】

一部の実施形態では、適当なGBDは、アミノ酸配列: MTKADIGTPSNFQH
IGHVGWDPNTGFDLNNLDPELKNLFDLMCGISEAQKLDRETS
SKVIYDFIEK TGGVEAVKNELELRQAP (配列番号64)、およびこれ
らの間のすべての範囲および部分範囲に対して少なくとも約75%、76%、77%、
78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、
88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、
98%、99%、または100%のアミノ酸配列類似性を有するアミノ酸配列を含む。

20

【0526】

一部の実施形態では、動員ペプチドは、GBDリガンドである。例示的な、限定されないGBDリガンドは、アミノ酸配列LVGALMHVMQKRSRAIHSSDEGED
QAGDEDED (配列番号65)を含む。

30

【0527】

ロイシンジッパーペプチド

適当なPBDは、ロイシンジッパーペプチドを含む。一部の実施形態では、ロイシンジ
ッパーペプチドは、コイルドコイルドメインを介して相互作用するペプチドである。ロイ
シンジッパードメインのアミノ酸配列は、当技術分野で公知である。ロイシンジッパー
ペプチドとしては、EE12RR345Lロイシンジッパーペプチド; RR12EE354
Lロイシンジッパーペプチドなどが挙げられる。

【0528】

ロイシンジッパーペプチドのアミノ酸配列の例は、アミノ酸配列: LEIEAAFLER
RENTALETRVAELRQRVQRLR NRVSYRTRYGPLGGGK (配列番号66)のEE12RR345Lロイシンジッパーである。

40

【0529】

一部の実施形態では、ロイシンジッパーペプチドは、アミノ酸配列: LEIEAA
FLERENTALETRVAELRQRVQRLRNRVSYRTRYGPLGGGK (配列番号67)、およびこれらの間のすべての範囲および部分範囲に対して少なくとも
約75%、76%、77%、78%、79%、80%、81%、82%、83%、84%
、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%
、95%、96%、97%、98%、99%、または100%のアミノ酸配列類似性を有
するアミノ酸配列を含む。このようなロイシンジッパーペプチドは、PBDとしてまたは
動員ペプチドとしての機能を果し得る。

50

【0530】

ロイシンジッパーペプチドのアミノ酸配列の別の限定されない例は、アミノ酸配列：L E I R A A F L R Q R N T A L R T E V A E L E Q E V Q R L E N E V S Q Y E T R Y G P L G G G K（配列番号68）のRR12EE345Lロイシンジッパーペプチドである。

【0531】

上記の記載は、足場ポリペプチド上に位置する1つまたは複数の動員ペプチドを標的にする（それに結合する）ように設計されたタンパク質結合ドメインを含むキメラ生合成タンパク質の産生について記載している。この配置の他の適合するバリエーションが当業者には認識されよう。例えば、一部の実施形態では、本開示は、足場ポリペプチド上に位置するタンパク質結合ドメインによって標的にされる動員ペプチドを含むキメラ生合成タンパク質の産生を教示する。他の実施形態では、タンパク質結合ドメインおよび動員ペプチドをそれぞれ2つまたはそれよりも多くのキメラ生合成タンパク質に組み入れ、したがって、キメラタンパク質が複合体（例えば、二量体またはヘテロ二量体）を形成する。複合体の形成の1つの実例は、2つまたはそれよりも多くのキメラ生合成タンパク質がロイシンジッパードメインを介して複合体を形成することが可能になるように、キメラ生合成タンパク質上に位置付けられた適合するロイシンジッパードメインを使用することである。

【0532】

足場ポリペプチド

一部の実施形態では、本開示は、生合成経路酵素を機能的な複合体に組織化する足場ポリペプチドを教示する。一部の実施形態では、本開示の足場ポリペプチドは、2つまたはそれよりも多くの動員ペプチドを含む。すなわち、一部の実施形態では、本開示の足場ポリペプチドは、2つまたはそれよりも多くのキメラ生合成タンパク質を補充することができる。

【0533】

一部の実施形態では、本開示の足場ポリペプチドは、宿主細胞に導入された外因性ペプチドである（例えば、足場ポリペプチドをコードするDNA配列の形質転換、またはペプチドの直接導入によって）。他の実施形態では、足場ポリペプチドは、宿主細胞内に天然に存在する構造である。すなわち、一部の実施形態では、足場ポリペプチドは、細胞小器官または膜（例えば、小胞体またはゴルジ装置）である。よって、一部の実施形態では、本開示の足場ポリペプチドは、1つよりも多くのペプチド配列で構成される宿主細胞構造を含む。

【0534】

一部の実施形態では、足場ポリペプチド内の動員ペプチド配列を、酵素が補充される標的生合成経路が最適化されるように組織化する。一部の実施形態では、足場ポリペプチドの組織化は、上でDNA足場について記載したものと同様である。したがって、一部の実施形態では、足場ポリペプチドは、様々なキメラ生合成タンパク質の順序および比を調節するための動員ペプチドのグループ化を含有する。

【0535】

細胞培養および発酵

本開示の細胞は、任意の所望の生合成反応または選択に適切に改変された慣例的な栄養培地中で培養することができる。一部の実施形態では、本開示は、プロモーターを活性化するための誘導培地中での培養を教示する。一部の実施形態では、本開示は、形質転換体（例えば、抗生物質）の選択剤、または阻害条件（例えば、高エタノール条件）下で増殖させるのに適した生物体の選択を含めた選択剤を含む培地を教示する。一部の実施形態では、本開示は、細胞増殖のために最適化された培地中の増殖細胞培養を教示する。他の実施形態では、本開示は、産物収率のために最適化された培地中の増殖細胞培養を教示する。一部の実施形態では、本開示は、細胞増殖を誘導することができ、最終産物産生に必要な前駆体（例えば、エタノール産生のための高レベルの糖）も含有する培地中の増殖培養を教示する。

10

20

30

40

50

【0536】

例えば、温度、pHなどの培養条件は、発現のために選択された宿主細胞とともに使用するのに適したものであり、当業者に明らかである。述べたように、多くの参考文献が、細菌、植物、動物（哺乳動物を含む）、および古細菌起源の細胞を含めた多くの細胞の培養および産生に利用可能である。例えば、Sambrook、Ausubel（すべて上記）、ならびにBerger、Guide to Molecular Cloning Techniques、Methods in Enzymology、152巻、Academic Press, Inc., San Diego, CA;ならびにFreshney（1994年）、Culture of Animal Cells、a Manual of Basic Technique、3版、Wiley-Liss、New York、およびそれに引用された参考文献；DoyleおよびGriffiths（1997年）、Mammalian Cell Culture:Essential Techniques、John Wiley and Sons、NY；Humason（1979年）、Animal Tissue Techniques、4版、W.H. Freeman and Company；ならびにRicciardelleら、（1989年）、In Vitro Cell Dev. Biol., 25巻：1016～1024頁を参照。これらのすべては、参照により本明細書に組み込まれている。植物細胞培養および再生について、Payneら（1992年）、Plant Cell and Tissue Culture in Liquid Systems、John Wiley & Sons, Inc., New York, N.Y.; GamborgおよびPhillips（編）（1995年）、Plant Cell, Tissue and Organ Culture; Fundamental Methods Springer Lab Manual、Springer-Verlag（Berlin Heidelberg N.Y.）；Jones編（1984年）、Plant Gene Transfer and Expression Protocols、Humana Press、Totowa, N.J.、ならびにPlant Molecular Biology（1993年）、R. R. D. Croy編、Bios Scientific Publishers、Oxford、U.K. ISBN 0 12 198370 6。これらのすべては、参照により本明細書に組み込まれている。一般に細胞培養培地は、AtlasおよびParks（編）、The Handbook of Microbiological Media（1993年）、CRC Press、Boca Raton, Fla.に示されており、これは、参照により本明細書に組み込まれている。細胞培養の追加の情報は、Sigma-Aldrich, Inc（St Louis, Mo.）からのLife Science Research Cell Culture Catalogue（「Sigma-LSRCCC」）、および例えば、やはりSigma-Aldrich, Inc（St Louis, Mo.）からのThe Plant Culture Catalogue and supplement（「Sigma-PCCS」）などの入手可能な市販の文献に見出され、これらのすべては、参照により本明細書に組み込まれている。

10

20

【0537】

使用される培養培地は、適当な様式で、それぞれの株の要求を満たさなければならない。様々な微生物の培養培地の記載は、American Society for Bacteriologyの「Manual of Methods for General Bacteriology」（Washington D.C., USA、1981年）に存在する。

30

【0538】

本開示は、目的の産物を発酵調製するためのプロセスであって、a) 適当な培地中で本開示による微生物を培養し、発酵プロセスをもたらすステップと；b) a) の発酵プロセス中、かつ／または微生物の細胞内で目的の産物を濃縮するステップとを含む、プロセスをさらに提供する。

【0539】

一部の実施形態では、本開示は、生成される微生物を、所望の有機－化学化合物を産生させる目的で、例えば、WO 05/021772に記載されているように連続的に、あるいはバッチプロセス（バッチ培養）またはフェドバッチもしくは繰り返しフェドバッチプロセスで不連続に培養することができることを教示する。公知の培養方法についての全般的な特質の概要は、Chmiel（Bioprozesstechnik. 1: Einfuehrung in die Bioverfahrenstechnik（Gustav Fischer Verlag, Stuttgart, 1991年））による教科書、またはStorhas（Bioreaktoren and periphere Einrichtungen（Vieweg Verlag, Braunschweig/Wiesbaden, 1994年））による教科書において入手可能である。

40

【0540】

一部の実施形態では、本開示の細胞は、バッチまたは連続発酵条件下で増殖される。

【0541】

50

古典的なバッチ発酵は、クローズドシステムであり、培地の組成は、発酵の開始時に設定され、発酵中に人為的な変更には供されない。バッチシステムの変形は、フェドバッチ発酵であり、これは、本開示においても使用を見出す。この変形では、基質は、発酵が進行するにつれて徐々に添加される。フェドバッチシステムは、異化産物抑制が細胞の代謝を阻害する可能性がある、および培地中の基質の量が限定されていることが望ましい場合、有用である。バッチ発酵およびフェドバッチ発酵は、一般的であり、当技術分野で周知である。

【0542】

連続発酵は、目的の所望の生体分子産物を処理および回収するために、規定された発酵培地がバイオリアクターに連続的に添加され、等量の馴化培地が同時に除去されるシステムである。一部の実施形態では、連続発酵は一般に、一定の高密度で培養物を維持し、この場合、細胞は、主に、対数期増殖にある。一部の実施形態では、連続発酵は一般に、静止または後期対数／静止期増殖で培養物を維持する。連続発酵システムは、定常状態増殖条件を維持しようと努める。

10

【0543】

連続発酵プロセスのための栄養分および成長因子をモジュレートするための方法、ならびに産物形成の速度を最大化するための技法は、工業微生物学の技術分野で周知である。

【0544】

例えば、本開示の培養のための炭素源の限定されないリストとしては、糖および炭水化物、例えば、グルコース、スクロース、ラクトース、フルクトース、マルトース、糖蜜、サトウダイコンまたはサトウキビ加工からのスクロース含有溶液、デンプン、デンプン加水分解物、およびセルロースなど；油脂、例えば、ダイズ油、ヒマワリ油、落花生油、およびココナツ脂肪など；脂肪酸、例えば、パルミチン酸、ステアリン酸、およびリノール酸など；アルコール、例えば、グリセロール、メタノール、およびエタノールなど；ならびに有機酸、例えば、酢酸または乳酸などが挙げられる。

20

【0545】

本開示の培養のための窒素源の限定されないリストとしては、有機窒素含有化合物、例えば、ペプトン、酵母エキス、肉エキス、麦芽エキス、トウモロコシ浸出液、ダイズ粉、および尿素；または無機化合物、例えば、硫酸アンモニウム、アンモニウムクロリド、リン酸アンモニウム、炭酸アンモニウム、および硝酸アンモニウムが挙げられる。窒素源は、個々に、または混合物として使用され得る。

30

【0546】

本開示の培養のための可能なリン源の限定されないリストとしては、リン酸、リン酸二水素カリウムもしくはリン酸水素二カリウム、または対応するナトリウム含有塩が挙げられる。

【0547】

培養培地は、例えば、金属、例えば、ナトリウム、カリウム、マグネシウム、カルシウム、および鉄などの塩化物または硫酸塩、例えば、硫酸マグネシウムまたは硫酸鉄などの形態で塩をさらに含み得、これらは、増殖に必要である。

【0548】

最後に、アミノ酸などの必須の成長因子、例えば、ホモセリンおよびビタミン、例えば、チアミン、ビオチン、またはパントテン酸を、上述した物質に加えて使用することができる。

40

【0549】

一部の実施形態では、培養物のpHは、これらに限定されないが、水酸化ナトリウム、水酸化カリウム、アンモニア、もしくはアンモニア水を含めた任意の酸もしくは塩基または緩衝塩；あるいは適当な様式でのリン酸または硫酸などの酸性化合物によって制御することができる。一部の実施形態では、pHは一般に、6.0～8.5、好ましくは6.5～8の値に調整される。

【0550】

50

一部の実施形態では、本開示の培養物は、例えば、脂肪酸ポリグリコールエステルなどの消泡剤を含み得る。一部の実施形態では、本開示の培養物は、例えば、抗生物質などの適当な選択物質を添加することによって、培養物のプラスミドを安定化するように改変される。

【0551】

一部の実施形態では、培養は、好気性条件下で実行される。これらの条件を維持するために、酸素または酸素含有気体混合物、例えば、空気などが、培養物に導入される。過酸化水素を富化させた液体を使用することが同様に可能である。発酵は、適切な場合、高压で、例えば、0.03~0.2 MPaの高压で実行される。培養物の温度は、通常20℃~45℃、好ましくは25℃~40℃、特に好ましくは30℃~37℃である。バッチプロセスまたはフェドバッチプロセスでは、培養は、好ましくは、回収されるのに十分な目的の所望の産物（例えば、有機-化学化合物）の量が形成されるまで継続される。この目的は通常、10時間~160時間以内に達成することができる。連続プロセスでは、より長い培養時間が可能である。微生物の活動により、発酵培地中で、かつ／または前記微生物の細胞内で目的の産物が濃縮（蓄積）される。

【0552】

一部の実施形態では、培養は、嫌気性条件下で実行される。

【0553】

スクリーニング

一部の実施形態では、本開示は、ハイスループット初期スクリーニングを教示する。他の実施形態では、本開示は、性能データのロバストなタンクベースの検証も教示する（図6Bを参照）。

【0554】

一部の実施形態では、ハイスループットスクリーニングプロセスが、バイオリアクター内の株の性能を予測するのに設計される。先に記載した通り、培養条件は、生物体に適しており、バイオリアクター条件を反映しているように選択される。個々のコロニーが精選され、96ウェルプレート中に移され、適当な量の時間にわたってインキュベートされる。細胞は、追加のシード培養のために新しい96ウェルプレートに、または産生培養に引き続いて移される。培養物は、様々な長さの時間にわたってインキュベートされ、この場合、複数の測定を行うことができる。これらは、産物、バイオマス、またはバイオリアクター内の株の性能を予測する他の特性の測定を含み得る。ハイスループット培養結果は、バイオリアクター性能を予測するのに使用される。

【0555】

一部の実施形態では、タンクベースの性能検証が使用されて、ハイスループットスクリーニングによって単離された株の性能が確認される。候補株は、生産性または収率などの関連した株性能特性について、ベンチスケール発酵反応器（例えば、本開示の表5に開示した反応器）を使用してスクリーニングされる。

【0556】

産物回収および定量

目的の産物の産生についてスクリーニングするための方法は、当業者に公知であり、本明細書全体にわたって論じられている。このような方法は、本開示の株をスクリーニングするとき使用することができる。

【0557】

一部の実施形態では、本開示は、非分泌型細胞内産物を産生するように設計された株を改善する方法を教示する。例えば、本開示は、細胞内酵素、油、医薬品、または他の貴重な低分子もしくはペプチドを産生する細胞培養物のロバスト性、収率、効率、または全体的な望ましさを改善する方法を教示する。非分泌型細胞内産物の回収または単離は、本明細書に記載のものを含めて当技術分野で周知である溶解技法および回収技法によって達成することができる。

【0558】

例えば、一部の実施形態では、本開示の細胞は、遠心分離、濾過、沈殿、または他の方法によって回収することができる。次いで回収された細胞は、当業者に周知である凍結融解サイクリング、超音波処理、機械的破壊、もしくは細胞溶解剤の使用、または他の方法を含めた任意の好都合な方法によって破壊される。

【0559】

結果として生じる目的の産物、例えば、ポリペプチドは、当技術分野で公知のいくつかの方法のいずれかによって回収／単離および任意選択で精製することができる。例えば、産物ポリペプチドは、これらに限定されないが、遠心分離、濾過、抽出、噴霧乾燥、蒸発、クロマトグラフィー（例えば、イオン交換、親和性、疎水性相互作用、クロマトフォーカシング、およびサイズ排除）、または沈降を含めた慣例的な手順によって栄養培地から単離することができる。最後に、高速液体クロマトグラフィー（HPLC）を、最終精製ステップで使用するすることができる。（例えば、ともに参照により本明細書に組み込まれている、Parryら、2001年、Biochem. J.、353巻：117頁、およびHongら、2007年、Appl. Microbiol. Biotechnol.、73巻：1331頁に記載の細胞内タンパク質の精製を参照）。

【0560】

上述した参考文献に加えて、様々な精製方法は、例えば、Sandana（1997年）、Bio separation of Proteins、Academic Press, Inc.; Bollagら（1996年）、Protein Methods、2版、Wiley-Liss、NY; Walker（1996年）、The Protein Protocols Handbook、Humana Press、NJ; HarrisおよびAngal（1990年）、Protein Purification Applications: A Practical Approach、IRL Press at Oxford、Oxford、England; HarrisおよびAngal Protein Purification Methods: A Practical Approach、IRL Press at Oxford、Oxford、England; Scopes（1993年）、Protein Purification: Principles and Practice、3版、Springer Verlag、NY; JansonおよびRyden（1998年）、Protein Purification: Principles, High Resolution Methods and Applications、2版、Wiley-VCH、NY; ならびにWalker（1998年）、Protein Protocols on CD-ROM、Humana Press、NJに示されたものを含めて、当技術分野で周知である。これらの文献のすべては、参照により本明細書に組み込まれている。

【0561】

一部の実施形態では、本開示は、分泌産物を産生させるように設計された株を改良する方法を教示する。例えば、本開示は、貴重な低分子またはペプチドを産生する細胞培養物のロバスト性、収率、効率、または全体的な望ましさを改善する方法を教示する。

【0562】

一部の実施形態では、免疫学的方法は、本開示の細胞によって産生される分泌型または非分泌型産物を検出および／または精製するのに使用することができる。一例示的手法では、従来の方法を使用して産物分子に対して（例えば、インスリンポリペプチドまたはその免疫原性断片に対して）産生される抗体は、ビーズ上に固定化され、エンドグルカナーゼが結合している条件下で細胞培養培地と混合され、沈降される。一部の実施形態では、本開示は、酵素結合免疫吸着測定法（ELISA）の使用を教示する。

【0563】

他の関連した実施形態では、米国特許第5,591,645号、米国特許第4,855,240号、米国特許第4,435,504号、米国特許第4,980,298号、およびSe-Hwan Paekら、「Development of rapid One-Step Immunochromatographic assay, Methods」、22巻、53～60頁、2000年）に開示された免疫クロマトグラフィーが使用される。これらの文献のそれぞれは、本明細書に参照により組み込まれている。一般的な免疫クロマトグラフィーは、2つの抗体を使用することによって検体を検出する。第1の抗体は、試験溶液中に、または試験溶液が滴下される、多孔質膜から作製されたおよそ矩形状をした試験片の末端の部分に存在する。この抗体は、ラテックス粒子または金コロイド粒子で標識される（この抗体は、以下で標識抗体と呼ばれる）。滴下された試験溶液が検出されるべき検体を含むとき、標識抗体は、検体を認識して検体と結合す

る。検体と標識抗体との複合体は、アブソーバーに向けて毛管現象によって流れ、このアブソーバーは、濾紙から作製され、標識抗体を含んだ末端と反対の末端に結合している。流れている間に、検体と標識抗体との複合体は、多孔質膜の中間に存在している第2の抗体（これは、以下でタッピング抗体（tapping antibody）と呼ばれる）によって認識および捕捉され、この結果として、複合体は、可視シグナルとして多孔質膜上の検出部分に現れ、検出される。

【0564】

一部の実施形態では、本開示のスクリーニング方法は、測光検出技法（吸収、蛍光）に基づく。例えば、一部の実施形態では、検出は、抗体に結合したGFPなどのフルオロフォア検出剤の存在に基づき得る。他の実施形態では、測光検出は、細胞培養からの所望の産物の蓄積に基づき得る。一部の実施形態では、産物は、培養物または前記培養物からの抽出物のUVを介して検出可能であり得る。

【0565】

当業者は、本開示の方法が、目的の任意の望ましい生体分子産物を産生する宿主細胞に適合することを認識する。以下の表2は、本開示の範囲内に含まれる産物カテゴリー、生体分子、および宿主細胞の限定されないリストを提示する。これらの例は、例示的な目的で提供されており、決して本開示の技術の適用性を限定するように意味していない。

【表2-1】

表2. 一本開示の目的の宿主細胞および産物の限定されないリスト。

産物カテゴリー	産物	宿主カテゴリー	宿主
アミノ酸	リシン	細菌	<i>Escherichia coli</i>
アミノ酸	メチオニン	細菌	<i>Escherichia coli</i>
アミノ酸	MSG	細菌	<i>Escherichia coli</i>
アミノ酸	トレオニン	細菌	<i>Escherichia coli</i>
アミノ酸	トレオニン	細菌	<i>Escherichia coli</i>
アミノ酸	トリプトファン	細菌	<i>Escherichia coli</i>
香味&芳香	アガーウッド	細菌	<i>Escherichia coli</i>
香味&芳香	アンブロックス	細菌	<i>Escherichia coli</i>
香味&芳香	ヌートカトン	細菌	<i>Escherichia coli</i>
香味&芳香	パチュリ油	細菌	<i>Escherichia coli</i>
香味&芳香	サフラン	細菌	<i>Escherichia coli</i>

【表 2 - 2】

産物カテゴリー	産物	宿主カテゴリー	宿主
香味&芳香	ビャクダン油	細菌	<i>Escherichia coli</i>
香味&芳香	バレンセン	細菌	<i>Escherichia coli</i>
香味&芳香	バニリン	細菌	<i>Escherichia coli</i>
食品	CoQ10/ユビキノール	細菌	<i>Escherichia coli</i>
食品	オメガ 3 脂肪酸	細菌	<i>Escherichia coli</i>
食品	オメガ 6 脂肪酸	細菌	<i>Escherichia coli</i>
食品	ビタミン B12	細菌	<i>Escherichia coli</i>
食品	ビタミン B2	細菌	<i>Escherichia coli</i>
食品	ビタミン B2	細菌	<i>Escherichia coli</i>
食品	エリスリトール	細菌	<i>Escherichia coli</i>
食品	エリスリトール	細菌	<i>Escherichia coli</i>
食品	エリスリトール	細菌	<i>Escherichia coli</i>
食品	ステビオールグリコシド	細菌	<i>Escherichia coli</i>
親水コロイド	ジウタンガム	細菌	<i>Escherichia coli</i>
親水コロイド	ゲランガム	細菌	<i>Escherichia coli</i>

10

20

30

【表 2 - 3】

産物カテゴリー	産物	宿主カテゴリー	宿主
親水コロイド	キサンタンガム	細菌	<i>Escherichia coli</i>
中間体	1, 3-PDO	細菌	<i>Escherichia coli</i>
中間体	1, 4-BDO	細菌	<i>Escherichia coli</i>
中間体	ブタジエン	細菌	<i>Escherichia coli</i>
中間体	n-ブタノール	細菌	<i>Escherichia coli</i>
有機酸	クエン酸	細菌	<i>Escherichia coli</i>
有機酸	クエン酸	細菌	<i>Escherichia coli</i>
有機酸	グルコン酸	細菌	<i>Escherichia coli</i>
有機酸	イタコン酸	細菌	<i>Escherichia coli</i>
有機酸	乳酸	細菌	<i>Escherichia coli</i>
有機酸	乳酸	細菌	<i>Escherichia coli</i>
有機酸	LCDAs - DDDA	細菌	<i>Escherichia coli</i>
ポリケチド/Ag	スピノサド	細菌	<i>Escherichia coli</i>
ポリケチド/Ag	スピネトラム	細菌	<i>Escherichia coli</i>

【0566】

目的の異種遺伝子

一実施形態では、微生物宿主細胞において異種遺伝子を発現させるための方法が本明細書に提供される。異種遺伝子を、本明細書に提供され、かつ／または当技術分野で公知の方法を使用し、微生物宿主細胞が異種遺伝子を使用して目的の産物を産生するように微生物宿主細胞に導入することができる。一実施形態では、微生物宿主細胞は、*E. coli* の株である。*E. coli* の株は、当技術分野で公知であり、かつ／または本明細書に提供される *E. coli* の任意の株であり得る。異種遺伝子は、前記遺伝子の野生型バージョンまたはその変異体であり得る。異種遺伝子は、プロモーター、ターミネーター、タンパク質可溶性タグ、タンパク質分解タグ、またはこれらの任意の組合せに作動可能に連結することができる。異種遺伝子をプロモーター、ターミネーター、タンパク質可溶性タグまたはタンパク質分解タグに作動可能に連結することは、本開示全体を通して提供されるプロモータースワップ、ターミネータースワップ、可溶性タグスワップおよび／または分解スワップ方法を使用して達成することができる。

【0567】

一実施形態では、異種遺伝子を表 1 から選択されるプロモーターに作動可能に連結する。一実施形態では、異種遺伝子を、60～90 bp のキメラ合成プロモーター配列に作動可能に連結し、ここで、キメラ合成プロモーターは、ラムダファージ p_R プロモーターの遠位部分、ラムダファージ p_L および p_R プロモーターの可変-35 および-10 領域、

ラムダファージ p_L および p_R プロモーターのコア部分、ならびにラムダファージ p_R プロモーターの 5' UTR/リボソーム結合部位 (RBS) 部分または *E. coli* *acc* 遺伝子のプロモーターの 5' UTR/リボソーム結合部位 (RBS) 部分のいずれかからなる。キメラ合成プロモーターに使用するためのラムダファージ p_R プロモーターの遠位部分、ラムダファージ p_L および p_R プロモーターの可変 -35 および -10 領域、ラムダファージ p_L および p_R プロモーターのコア部分、ならびにラムダファージ p_R プロモーターの 5' UTR/リボソーム結合部位 (RBS) 部分または *E. coli* *acc* 遺伝子のプロモーターの 5' UTR/リボソーム結合部位 (RBS) 部分のいずれかの核酸配列は、表 1. 5に見出される核酸配列から選択することができる。一実施形態では、異種遺伝子を、表 1. 4に見出される配列番号 132~207から選択される核酸配列を有するキメラ合成プロモーターに作動可能に連結することができる。

【0568】

一実施形態では、異種遺伝子を、表 1. 2から選択されるターミネーターに作動可能に連結する。別の実施形態では、異種遺伝子を、表 19から選択されるターミネーター配列に作動可能に連結する。

【0569】

一実施形態では、異種遺伝子を、表 17から選択される可溶性タグに作動可能に連結する。

【0570】

一実施形態では、異種遺伝子を、表 18から選択される分解タグ配列に作動可能に連結する。

【0571】

上記の実施形態に付け加えて、異種遺伝子は、表 2に見出される目的の産物を生成するために必要な遺伝子のいずれか、または目的の産物を産生させるために微生物宿主細胞 (例えば、*E. coli*) において異種遺伝子として発現させることができる当技術分野で公知の任意の遺伝子であり得る。

【0572】

一実施形態では、異種遺伝子は、図 19に示すリシン生合成経路の一部である遺伝子である。本実施形態に付け加えて、異種遺伝子は、*asd* 遺伝子、*ask* 遺伝子、*hom* 遺伝子、*dapA* 遺伝子、*dapB* 遺伝子、*dapD* 遺伝子、*ddh* 遺伝子、*argD* 遺伝子、*dapE* 遺伝子、*dapF* 遺伝子、*lysA* 遺伝子、*lysE* 遺伝子、*zwf* 遺伝子、*pgi* 遺伝子、*ktk* 遺伝子、*fbp* 遺伝子、*ppc* 遺伝子、*pcK* 遺伝子、*ddx* 遺伝子、*pyc* 遺伝子または *icd* 遺伝子から選択することができる。一実施形態では、本明細書で提供されるリシン経路の一部である異種遺伝子を配列番号 132~207から選択される核酸配列を有するキメラ合成プロモーターに作動可能に連結する。

【0573】

一実施形態では、異種遺伝子は、例えば図 59に示すリコピン生合成経路の一部である遺伝子である。本実施形態に付け加えて、異種遺伝子は、*dxs* 遺伝子、*ispC* 遺伝子、*ispE* 遺伝子、*ispD* 遺伝子、*ispF* 遺伝子、*ispG* 遺伝子、*ispH* 遺伝子、*idi* 遺伝子、*ispA* 遺伝子、*ispB* 遺伝子、*crtE* 遺伝子、*crtB* 遺伝子、*crtI* 遺伝子、*crtY* 遺伝子、*ymgA* 遺伝子、*dxr* 遺伝子、*elbA* 遺伝子、*gdhA* 遺伝子、*appY* 遺伝子、*elbB* 遺伝子、または *ymgB* 遺伝子から選択することができる。一実施形態では、本明細書で提供されるリコピン経路の一部である異種遺伝子を配列番号 132~207から選択される核酸配列を有するキメラ合成プロモーターに作動可能に連結する。

【0574】

一実施形態では、異種遺伝子は、生物学的製剤をコードする遺伝子または生物学的製剤を生成するための経路内の遺伝子である。一実施形態では、微生物宿主細胞は *E. coli* であり、生物学的製剤は *E. coli* において産生されることが示されている任意の生物学的製剤である。生物学的製剤は、ヒューマリン (*rh* インスリン)、イントロン A (

インターフェロンアルファ 2 b)、ロフェロン (インターフェロンアルファ 2 a)、ヒューマトロップ (ソマトロピン r h 成長ホルモン)、ニューボジェン (フィルグラスチム)、ベタフェロン (インターフェロンベーター 1 b)、リスプロ (速効型インスリン)、ラピリシン (レテプラゼ)、インフェルゲン (インターフェロンアルファコンー 1)、グルカゴン、ベロムン (タソネルミン)、オンタック (デニロイキンジフチトクス)、ランタス (長時間作用型インスリングラルギン)、キネレット (アナキンラ)、ナトレコール (ネシリチド)、ソマバート (ペグビソマント)、カルシトニン (組換えカルシトニンサケ)、ルセンティス (ラニビズマブ)、プレオタクト (ヒト副甲状腺ホルモン)、クリステクサ (r h 尿酸オキシダーゼ、P E G 化)、ニベスチム (フィルグラスチム、r h G C S F)、ボラキサゼ (グルカルピダーゼ)、またはプレオス (副甲状腺ホルモン) から選択することができる。一実施形態では、生物学的製剤をコードする異種遺伝子または本明細書で提供される生物学的製剤を生成する経路内の遺伝子を配列番号 1 3 2 ~ 2 0 7 から選択される核酸配列を有するキメラ合成プロモーターに作動可能に連結する。

10

【0575】

選択判定基準および目標

本開示の方法に適用される選択判定基準は、株改良プログラムの具体的な目標によって変動する。本開示は、任意のプログラム目標を満たすように適応させることができる。例えば、一部の実施形態では、プログラム目標は、差し迫った期限を伴わずに反応の単一バッチ収率を最大化することであり得る。他の実施形態では、プログラム目標は、生合成収率をリバランスして、特異的な産物を産生させ、または特定の比の産物を産生させることであり得る。他の実施形態では、プログラム目標は、産物の化学構造を改変すること、例えば、ポリマーの炭素鎖を長くすることであり得る。一部の実施形態では、プログラム目標は、性能特性、例えば、収率、力価、生産性、副生成物排除、プロセスエクスカージョンに対する寛容性、最適な増殖温度、および増殖速度を改善することであり得る。一部の実施形態では、プログラム目標は、微生物によって産生される目的の産物の体積生産性、比生産性、収率、または力価によって測定した宿主性能の改善である。

20

【0576】

他の実施形態では、プログラム目標は、入力量当たりの最終産物収率 (例えば、スクロース 1 ポンド当たりに産生されるエタノールの合計量) の観点から市販の株の合成効率を最適化することであり得る。他の実施形態では、プログラム目標は、例えば、バッチ完了率、または連続培養システムにおける歩留まり率の観点から測定した、合成速度を最適化することであり得る。他の実施形態では、プログラム目標は、特定のフェージに対する株の耐性を増大させること、またはさもなければ培養条件下で株の活力/ロバスト性を増大させることであり得る。

30

【0577】

一部の実施形態では、株改良プロジェクトは、1つを超える目標を対象とする場合がある。一部の実施形態では、株プロジェクトの目標は、品質、信頼性、または全体的な収益性次第であり得る。一部の実施形態では、本開示は、上述した株性質の 1つまたは複数に有する関連した選択された変異または変異の群の方法を教示する。

【0578】

当業者は、特定のプロジェクト目標を満たすように株選択判定基準をどのように合わせるかを認識する。例えば、反応が飽和しているときの株の単一バッチ最大収量の選択は、単一バッチ収量が高い株を同定するのに適切であり得る。一連の温度および条件にわたる収率の一貫性に基づく選択は、ロバスト性および信頼性が増大した株を同定するのに適切であり得る。

40

【0579】

一部の実施形態では、初期のハイスループットフェーズおよびタンクベースの検証の選択判定基準は、同一となる。他の実施形態では、タンクベースの選択は、追加のおよび/または異なる選択判定基準下で働き得る。例えば、一部の実施形態では、ハイスループット株選択は、単一バッチ反応完了収率に基づき得る一方、タンクベースの選択は、反応速

50

度に対する収率に基づく選択を含むように拡張することができる。

【0580】

シーケンシング

一部の実施形態では、本開示は、本明細書に記載の生物体の全ゲノムシーケンシングを教示する。他の実施形態では、本開示は、本開示の方法の品質対照としてのプラスミド、PCR産物、および他のオリゴのシーケンシングも教示する。大および小プロジェクトのためのシーケンシング法は、当業者に周知である。

【0581】

一部の実施形態では、核酸をシーケンシングするための任意のハイスループット技法を、本開示の方法において使用することができる。一部の実施形態では、本開示は、全ゲノムシーケンシングを教示する。他の実施形態では、本開示は、遺伝的バリエーションを同定するためのアンプリコンシーケンシング、ウルトラディープシーケンシングを教示する。一部の実施形態では、本開示は、タグメンテーションを含めたライブラリー調製の新規方法も教示する（WO／2016／073690を参照）。DNAシーケンシング技法としては、標識ターミネーターまたはプライマー、およびスラブまたはキャピラリーにおけるゲル分離を使用する古典的ジデオキシシーケンシング反応（サンガー法）；可逆的に終止される標識ヌクレオチドを使用する合成によるシーケンシング、ピロシーケンシング；454シーケンシング；標識オリゴヌクレオチドプローブのライブラリーへの対立遺伝子特異的ハイブリダイゼーション；ライゲーションが後に続く標識クロンのライブラリーへの対立遺伝子特異的ハイブリダイゼーションを使用する合成によるシーケンシング；重合ステップ中の標識ヌクレオチドの組み入れのリアルタイムモニタリング；ポロニーシーケンシング；およびSOLiDシーケンシングが挙げられる。

【0582】

本開示の一態様では、固体表面上の個々の分子を空間的に単離するステップであって、個々の分子が並行してシーケンシングされる、ステップを含むシーケンシングのハイスループット方法が使用される。このような固体表面として、非多孔性表面（Sollexaシーケンシング、例えば、Bentleyら、Nature、456巻：53～59頁（2008年）またはComplete Genomicsシーケンシング、例えば、Drmanacら、Science、327巻：78～81頁（2010年）などにおける）、ビーズ結合鋳型または粒子結合鋳型を含み得るウェルのアレイ（454、例えば、Marguliesら、Nature、437巻：376～380頁（2005年）、またはイオントレントシーケンシング、米国特許公開2010／0137143号もしくは同第2010／0304982号を用いてなど）、微細加工膜（SMRTシーケンシング、例えば、Eidら、Science、323巻：133～138頁（2009年）を用いてなど）、またはビーズアレイ（SOLiDシーケンシングまたはポロニーシーケンシング、例えば、Kimら、Science、316巻：1481～1414頁（2007年）を用いるような）を挙げることができる。

【0583】

別の実施形態では、本開示の方法は、単離分子が固体表面上で空間的に単離される前または後のいずれかに単離分子を増幅するステップを含む。事前の増幅は、エマルジョンPCRなどのエマルジョンに基づく増幅、またはローリングサークル増幅を含み得る。Sollexaに基づくシーケンシングも教示されており、この場合、Bentleyら（上記に引用した）に、かつ製造者の指示（例えば、TruSeq（商標）試料調製キットおよびデータシート、Illumina, Inc.、San Diego, Calif.、2010年）に；かつさらに、参照により組み込まれている以下の参考文献：米国特許第6,090,592号；同第6,300,070号；同第7,115,400号；およびEP0972081B1に記載されている通り、個々の鋳型分子が固体表面上で空間的に単離され、その後これらがブリッジPCRによって並行して増幅されて別個のクロン集団またはクラスターを形成し、次いでシーケンシングされる。

【0584】

一実施形態では、固体表面上で配置および増幅された個々の分子は、1 cm² 当たり少

10

20

30

40

50

なくとも 10^5 クラスターの密度で；または 1 cm^2 当たり少なくとも 5×10^5 の密度で；または 1 cm^2 当たり少なくとも 10^6 クラスターの密度でクラスターを形成する。一実施形態では、相対的に高い誤り率を有するシーケンシング化学反応が使用される。このような実施形態では、このような化学反応によって生じる平均品質スコアは、配列リード長の単調低下関数である。一実施形態では、このような低下は、配列リードの 0.5 パーセントが 1～75 位に少なくとも 1 つの誤差を有し；配列リードの 1 パーセントが 76～100 位に少なくとも 1 つの誤差を有し；配列リードの 2 パーセントが 101～125 位に少なくとも 1 つの誤差を有することに対応する。

【0585】

ゲノムワイド遺伝子設計判定基準の効果のコンピューター分析および予測

10

一部の実施形態では、本開示は、所与の宿主株内に組み入れている特定の遺伝子変更の効果を予測する方法を教示する。さらなる態様では、本開示は、所与の宿主株内に、前記宿主が特定の表現型形質または株パラメータを有するために組み入れるべきである提案された遺伝子変更を生成するための方法を提供する。所与の態様では、本開示は、新規宿主株を設計するのに利用することができる予測モデルを提供する。

【0586】

一部の実施形態では、本開示は、各ラウンドのスクリーニングの性能結果を分析する方法、およびその後のラウンドのスクリーニングで株性能を増強すると予測される新しく提案されたゲノムワイド配列修飾をおこすための方法を教示する。

【0587】

20

一部の実施形態では、本開示は、システムが、先のスクリーニング結果に基づいて宿主株に提案された配列修飾をおこすことを教示する。一部の実施形態では、本システムの推奨は、直前のスクリーニングからの結果に基づく。他の実施形態では、本システムの推奨は、先行するスクリーニングのうちの 1 つまたは複数の累積的結果に基づく。

【0588】

一部の実施形態では、本システムの推奨は、先に開発された HTP 遺伝子設計ライブラリーに基づく。例えば、一部の実施形態では、本システムは、同じまたは異なる宿主生物において、先のスクリーニングからの結果を保存し、異なるプロジェクトにこれらの結果を適用するように設計される。

【0589】

30

他の実施形態では、本システムの推奨は、科学的洞察に基づく。例えば、一部の実施形態では、推奨は、遺伝子の公知の性質（アノテートされた遺伝子データベースおよび関連した文献などの源からの）、コドン最適化、転写スリップ、uORF、または他の仮説主導型配列、および宿主最適化に基づく。

【0590】

一部の実施形態では、システムまたは予測モデルが推奨した宿主株への提案された配列修飾は、（1）プロモータスワップ、（2）SNP スワップ、（3）開始／停止コドン交換、（4）配列最適化、（5）Stop スワップ、（6）可溶性タグスワップ、（7）分解タグスワップ、および（8）上位性マッピングを含む開示した分子ツールセットの 1 つまたは複数の利用によって実行される。

40

【0591】

本明細書に記載の HTP 遺伝子操作プラットフォームは、任意の特定の微生物または表現型形質（例えば、特定の化合物の産生）に関してとらわれない。すなわち、本明細書に教示のプラットフォームおよび方法は、任意の宿主細胞とともに利用して、任意の所望の表現型形質を有するように前記宿主細胞を操作することができる。さらに、1 つの新規宿主細胞を生み出すのに使用される所与の HTP 遺伝子操作プロセスから学習した教訓を、教示された方法の間に行われる無数のプロセスパラメータの記憶、特徴付け、および分析の結果として、任意の数の他の宿主細胞に適用することができる。

【0592】

上位性マッピングセクションにおいて言及したように、HTP 遺伝子設計ライブラリー

50

からの変異のコレクションをある好適な予測モデルを介して特定のバックグラウンド内に統合することによって得られる仮定の株の性能（別名スコア）を推定することが可能である。このような予測モデルを考慮すると、コンビナトリアル統合を介して変異ライブラリーにアクセス可能なすべての仮定の株をスコア付けおよびランク付けすることが可能である。以下のセクションは、本H T Pプラットフォームにおいて利用される特定のモデルを概説する。

【0593】

予測的株設計

遺伝子変化および株性能を記述し、株内の変化の組成に基づいて株性能を予測し、高い予測された性能を有する候補設計を推奨し、予測を選別して二次の考慮事項、例えば、既存の株に対する類似度、上位性、または予測における信頼度を最適化する方法を含む、予測的株設計の手法が本明細書に記載されている。

10

【0594】

株設計モデルへの入力

一実施形態では、例示を容易にする目的で、入力データは、2つのコンポーネント：（1）遺伝子変化のセットおよび（2）相対的株性能を含み得る。当業者は、このモデルは、オーバーフィッティングの対抗する考慮事項に留意しながら、容易に拡張して多種多様な入力を考慮することができることを認識する。遺伝子変化に加えて、調整され得る入力パラメータ（独立変数）の一部は、細胞型（属、種、株、系統発生的特徴付けなど）、および発酵が細胞を用いて行われている下でのプロセスパラメータ（例えば、環境条件、装置のハンドリング、修飾技法など）である。

20

【0595】

遺伝子変化のセットは、H T P 遺伝子設計ライブラリーと呼ばれる遺伝子攪乱の先に論じたコレクションに由来し得る。相対的株性能は、任意の所与のパラメータまたは目的の表現型形質（例えば、目的の化合物、低分子、または産物の産生）に基づいて評価することができる。

【0596】

細胞型は、一般的なカテゴリー、例えば、原核生物システムおよび真核生物システム、属、種、株、組織培養物（分散細胞に対して）などにおいて指定され得る。調整され得るプロセスパラメータは、温度、圧力、反応器構成、および培地組成が挙げられる。反応器構成の例には、プロセスがバッチであっても連続であっても、反応器の体積、および連続である場合、体積流量などが含まれる。もしあれば、細胞が存在する、支持構造体を指定することもできる。培地組成の例としては、電解質の濃度、栄養分、老廃物、酸、p Hなどが挙げられる。

30

【0597】

予測的株設計モデルを生み出すのに次に使用される最初の線形回帰モデルにおいて利用される選択されたH T P 遺伝子設計ライブラリーからの遺伝子変化のセット

*Corynebacterium*における遺伝子変化の表からのエントリーの一例示的セットを以下の表3に示す。各行は、株7000051473における遺伝子変化、および変化の機構についてのメタデータ、例えば、プロモータースワップまたはS N P スワップを示す。*aceE*、*zwf*、および*pyc*は、すべて、クエン酸サイクルに関連する。

40

【0598】

この場合、株7000051473は、合計7つの変化を有する。「最後の変化」は、この株内の変化がこの株系統における最近の修飾を表すことを意味する。よって、この株の性能をその親の性能と比較することは、「最後の変化」の変異の性能に関するデータ点を表す。

【表 3】

表3ー株 7000051473 の株設計エントリー表

株	名称	ライブラリー	変化	変化元	変化先	最後の変化
7000051473	dlc19_42	proswp	pcg3121	cg1144	pcg3121_cg1144	1
7000051473	dlc19_42	scswp	acee atg>ttg	ttg	acee_atg	0
7000051473	dlc19_42	snpswp	dss_033	NA	na	0
7000051473	dlc19_42	snpswp	dss_084	NA	t	0
7000051473	dlc19_42	snpswp	dss_316	NA	na	0
7000051473	dlc19_42	proswp	pcg0007_39	zwf	pcg0007_39_zwf	0
7000051473	dlc19_42	proswp	pcg1860	pyc	pcg1860_pyc	0

【 0 5 9 9 】

構築された株 (built strain) 性能評価

教示したモデルの目標は、株に導入される遺伝子変化の組成に基づいて株性能を予測することである。比較のための標準を構築するために、株性能をアッセイプレート 1 つ当たり、株 1 つ当たりの性能の中央値を最初に算出することによって一般的な参照株と比べて算出する。次いで、相対的性能を、同じプレート内の操作された株と一般的な参照株との間の平均性能の差異として計算する。プレート内の比較に算出を制限することにより、検討中の試料がすべて同じ実験条件を受けたことを保証する。

【 0 6 0 0 】

図 2 3 は、検討中の入力データについての *Corynebacterium* の相対的株性能の分布を示す。ゼロの相対的性能は、操作された株がプレート内基準または「参照」株と同等によく働いたことを示す。ゼロを有意に上回って働きやすい株を同定する予測モデルの能力が目的のものである。さらに、かつより一般的に、いくつかの判定基準によって任意の所与の株がその親より性能が優れているかどうか、目的のものである。実際には、判定基準は、親レベルを超えるある閾値を満たす、または超える産物力価であり得るが、所望の方向で親からの統計的に有意な差異を有することも、代替としてまたは追加的に使用することができる。基準株または「参照」株の役割は、プレート内またはプレート間で比較を行うための付加された正規化因子として単に機能を果たすことである。

【 0 6 0 1 】

留意すべき概念は、親株と参照株との間の差異のものである。親株は、現在のラウンドの変異誘発について使用したバックグラウンドである。参照株は、特にプレート間で比較を促進するためにあらゆるプレートで実行される対照株であり、典型的には、上記に言及した「基準株」である。しかし、基準株（例えば、全体的な性能をベンチマークするのに使用される野生型株または工業用株）は、所与のラウンドの株改良において変異誘発標的であるという意味において必ずしも「基準」である必要はないので、より記述的な用語は、「参照株」である。

【 0 6 0 2 】

要約すると、基準／参照株は、一般に、構築された株の性能をベンチマークするのに使用され、一方、親株は、関連した遺伝的バックグラウンドにおける特異的な遺伝子変化の

性能をベンチマークするのに使用される。

【0603】

線形回帰を用いた構築された株の性能のランク付け

開示したモデルの目標は、構築された株内に導入される遺伝子変化の組成の関数として相対的株性能を記述することによって構築された株の性能をランク付けることである。本開示全体にわたって論じられるように、様々なHTP遺伝子設計ライブラリーは、操作された株内に導入される可能な遺伝子変化（例えば、遺伝子攪乱／変更）のレパートリーを提供する。線形回帰は、現在記載されている例示的な予測モデルの基礎である。

【0604】

以下の表（すなわち、表4）は、回帰に基づくモデル化のための例示的な入力を含む。株性能は、株内に含有される遺伝子変化の組成の関数として、一般的な基準株と比べてランク付けされる。

【0605】

各列の見出しは、遺伝子変化を表し、「1」は、変化の存在を表し、「0」は、変化の非存在を表す。「DSS」は、特定のライブラリー（相対的性能（relative _perf）後の最初の3列）からのSNPスワップを指す。最後の3列は、プロモータースワップであり、この場合、pcgXXXXは、特定のプロモーターを表し、最後の3文字は、プロモーターが適用されている遺伝子を表す。遺伝子は、中央代謝に関連している。プロモーターは、*Corynebacterium glutamicum*由来である（したがって「cg」表示）。利用されるプロモーターについてのさらなる情報は、プロモーターP1～P8を列挙する表1、および本願の配列表に見出すことができる。さらに、各プロモーターP1～P8についての詳細な情報は、2015年12月7日出願され、「Promoters from *Corynebacterium glutamicum*」という表題の米国仮出願第62/264,232号に見出すことができ、これは、参照により本明細書に組み込まれている。参照の容易さのために、以下の表において、pcg3121=P8；pcg0755=P4；およびpcg1860=P3である。

【表4】

表4—遺伝子変化の概要および相対的性能に対するこれらの効果

相対的性能	dss_033	dss_034	dss_056	pcg3121_pgi	pcg0755_zwf	pcg1860_pyc
0.1358908	0	0	0	0	0	1
-1.8946985	1	0	0	1	0	1
-0.0222045	0	0	0	1	0	0
0.6342183	1	0	1	0	0	0
-0.0803285	1	1	0	0	0	0
2.6468117	0	0	0	1	0	0

【0606】

構築された株を特徴付けるための線形回帰

線形回帰は、インプリメンテーションおよび解釈の容易さのために、記載されたHTPゲノム操作プラットフォームにとって魅力的な方法である。結果として生じる回帰係数は、各遺伝子変化の存在に帰することができる相対的株性能の平均の増大または低下として

解釈することができる。

【0607】

例えば、図24で分かるように、本技法は、本発明者らに、p g iプロモーターをp c g 3 1 2 1に変化させると、任意の負の上位性相互作用の非存在下で、相対的株性能が平均でおよそ5ユニット改善され、よって潜在的に高度に望ましい変化であると結論付けることを可能にする（注意：入力は、無単位正規化値である）。

【0608】

したがって、教示した方法は、様々な教示したライブラリーからの自己のゲノム内に導入された様々な遺伝子攪乱を有する構築された株を記述し／特徴付け、ランク付けるための線形回帰モデルを使用する。

10

【0609】

予測的設計モデル化

構築した株からのデータを利用した上述した線形回帰モデルは、まだ構築されていない株についての性能予測を行うのに使用することができる。

【0610】

手順は、以下の通り要約することができる：遺伝子変化のすべての可能な構成をi n s i l i c oで生成する、それに続き、回帰モデルを使用して相対的株性能を予測する、それに続き、性能による候補株設計を注文する。よって、未だ構築されていない株の性能を予測するための回帰モデルを利用することによって、本方法は、より少ない実験を同時に行いながらより高い性能の株の産生を可能にする。

20

【0611】

構成の生成

まだ構築されていない株の性能を予測するモデルを構築するとき、第1のステップは、設計候補の配列を生成することである。これは、株内の遺伝子変化の総数を固定し、次いで遺伝子変化のすべての可能な組合せを明確にすることによって行われる。例えば、潜在的な遺伝子変化／攪乱の総数を29（例えば、29の可能なSNP、もしくは29の異なるプロモーター、または遺伝子攪乱のユニバースが29である限り、これらの任意の組合せ）に設定し、次いで29の潜在的な遺伝子変化のすべての可能な3メンバー組合せを設計する決定を下すことができ、3, 654の候補株設計がもたらされる。

【0612】

上述の3, 654候補株にコンテキストを提供するために、 $n! / ((n-r)! * r!)$ を使用してnの可能なメンバーからサイズrの非重複グループの数を算出することができることを考慮されたい。r = 3である場合、n = 29は、3, 654を与える。よって、29の潜在的な変化のすべての可能な3メンバー組合せを設計する場合、結果は、3, 654候補株である。29の潜在的な遺伝子変化は、図25のx軸に存在する。

30

【0613】

新しい株設計の性能の予測

入力としてコンビナトリアル構成を用いて上で構築した線形回帰を使用して、次いで、各候補設計の予期される相対的性能を予測することができる。図25は、C o r y n e b a c t e r i u mについての上位100の予測された株設計についての変化の組成を要約する。x軸は、潜在的な遺伝子変化（29の可能な遺伝子変化）のプールを列挙し、y軸は、ランク順序を示す。黒色セルは、候補設計における特定の変化の存在を示し、一方、白色セルは、その変化の非存在を示す。この特定の例では、上位100設計のすべては、変化p c g 3 1 2 1 _ p g i、p c g 1 8 6 0 _ p y c、d s s _ 3 3 9、およびp c g 0 0 0 7 _ 3 9 _ l y s aを含有する。さらに、最上位候補設計は、変化d s s _ 0 3 4、d s s _ 0 0 9を含有する。

40

【0614】

予測精度は、新しい知見がモデルを反復して保持および再フィットするにつれて経時的に増大するはずである。本発明者らによるスタディからの結果は、予測モデルを反復して保持および改善することができる方法を例示する。図46は、モデル予測を観測された測

50

定値と比較するものである。モデル予測の品質は、予測値と観測値との間の関連性の強度を示す相関係数、または平均モデル誤差の尺度である二乗平均平方根誤差を含めたいくつかの方法によって評価することができる。モデル評価のための選択された測定基準を使用して、システムは、モデルが保持されるべきときの規則を定義することができる。

【0615】

上記モデルに対して2つの述べられていない仮定は、(1) 上位性相互作用が存在しないこと、および(2) 予測モデルを構築する(例えば、図24に例示された構築された株のデータ、またはモデルを構築するための参照として使用されるあらゆるデータセットから) 利用される遺伝子変化／攪乱はすべて、遺伝子変化の提案された組合せとして、同じ *Corynebacterium* バックグラウンドにおいて作製されたこと(例えば、図25に例示された)を含む。

10

【0616】

二次フィーチャのフィルタリング

上記実例は、予測された宿主細胞性能に基づく線形回帰予測に着目した。一部の実施形態では、本線形回帰方法は、非生体分子因子、例えば、飽和バイオマス、耐性、または他の測定可能な宿主細胞フィーチャに適用することもできる。よって、本開示の方法は、構築するための候補に優先順位を付けるとき、予測された性能以外の他のフィーチャを考慮することも教示する。追加の関連したデータが存在すると仮定して、非線形項も、回帰モデルに含められる。

20

【0617】

既存の株との密接な関係

すでに構築されたものと同様である予測された株は、最上位の予測された候補でないにもかかわらず、時間およびコスト削減をもたらすことができる。

【0618】

変化の多様性

上述のモデルを構築するとき、遺伝子変化は、上位性相互作用の存在に起因して真に相加的となる(線形回帰によって仮定し、上記仮定として述べた通り)と確認することができる。したがって、遺伝子変化相違点の知識は、正の相加性を見込みを増大させるのに使用することができる。例えば、上記の最上位にランク付けされた株由来の変化 *ds s* __ 034 および *ds s* __ 009 (これらはSNPスワップである)が同じ代謝経路上にあり、同様の性能特性を有すると分かると、その情報を使用して、変化の同じでない組成を有する別の最上位ランクの株を選択することができる。上位性マッピングに関して上記セクションに記載した通り、予測された最良の遺伝子変化を選別して十分に同じでない応答プロファイルを有する変異に選択を制限することができる。代替として、線形回帰は、予測に重み付けするために類似度行列を使用する重み付き最小二乗回帰であり得る。

30

【0619】

予測された性能の多様性

最後に、予測モデルを検証し、引き続いて改善するために、中程度のまたは不満足な予測された性能を有する株を設計するように選択することができる。

【0620】

反復株設計最適化

上記例について記載したように、上位100の株設計のすべては、変化 *p c g* 3121 __ *p g i*、*p c g* 1860 __ *p y c*、*ds s* __ 339、および *p c g* 0007 __ 39 __ *y s a* を含有する。さらに、最上位の候補株設計は、変化 *ds s* __ 034、*ds s* __ 009 を含有する。

40

【0621】

諸実施形態では、発注エンジン208は、最上位の候補変異を組み入れている微生物株を製造するために工場注文(factory order)をファクトリー210に出す。フィードバックループ様式では、結果を分析装置214によって分析して、どの微生物が所望の表現型性質を呈するかを決定することができる(314)。分析フェーズ中に、修飾された

50

株培養物が評価されて、これらの性能、すなわち、工業規模で産生される能力を含めて、これらの所望の表現型性質の発現が決定される。例えば、分析フェーズでは、とりわけ、プレートの画像データが使用されて、コロニー健康の指標として微生物コロニー増殖が測定される。分析装置 214 が使用されて、遺伝子変化が表現型性能と相関付けられ、結果として生じる遺伝子型－表現型相関データをライブラリー内に保存して（それは、ライブラリー 206 内に記憶され得る）、将来の微生物産生を知らせる。

【0622】

特に、十分に高い測定された性能を実際にはもたらす候補変化を、上記表 4 などの表にデータベース中の行として加えることができる。このようにして、最良に働く変異が、教師付き機械学習様式で予測的株設計モデルに加えられる。

10

【0623】

LIMS は、先のファクトリーランから発生した相関に基づいて、設計／構築／試験／分析サイクルを反復する。後続のサイクル中に、分析装置 214 は、単独で、または人のオペレーターと併せて、相関データを使用して遺伝子改変を微調整することによってより微細な粒度を有するより良好な表現型性能を達成して、入力インターフェース 202 に入力して戻すための基準株としての最良候補を選択することができる。このようにして、本開示の実施形態の実験室情報管理システム (laboratory information management system) は、品質改善フィードバックループをインプリメントする。

【0624】

20

つまり、図 33 のフロー図を参照して、反復予測的株設計ワークフローを以下の通り記述することができる。

- ・入出力変数、例えば、入力として遺伝子変化および出力として性能フィーチャの訓練セットを生成する (3302)。生成は、先の遺伝子変化およびこれらの遺伝子変化を組み入れている微生物株の対応する測定された性能に基づいて分析装置 214 によって実施することができる。

- ・訓練セットに基づいて初期モデル（例えば、線形回帰モデル）を開発する (3304)。これは、分析装置 214 によって実施することができる。

- ・設計候補株を生成する (3306)

○一実施形態では、分析装置 214 は、変化の組合せの形態でバックグラウンド株に行われる遺伝子変化の数を固定することができる。これらの変化を表すために、分析装置 214 は、インタープリター 204 に、変化のこれらの組合せを表す 1 つまたは複数の DNA の仕様を提供することができる。（これらの遺伝子変化またはこれらの変化を組み入れる微生物株は、「試験入力」と呼ばれる場合がある）。インタープリター 204 は、1 つまたは複数の DNA の仕様を解釈し、エグゼキューションエンジン 207 は、DNA の仕様を実行してこれらの変化について個々の候補設計株を表す分解された出力を有する DNA の仕様をポピュレートする。

30

- ・モデルに基づいて、分析装置 214 は、各候補設計株の予期される性能を予測する (3308)。

- ・分析装置 214 は、最高の予測された性能を有する限られた数の候補設計、例えば 100 を選択する (3310)。

40

○上位性マッピングに関して本明細書の他の場所で記載したように、分析装置 214 は、例えば、上位性効果について上位の設計を選別し、または予測モデル内に、上位性を計算に入れることによって上位性などの二次効果について説明することができる。

- ・発注エンジン 208 によって作成された工場注文に基づいて選別された候補株を構築する（ファクトリー 210 で） (3312)。

- ・分析装置 214 は、選択された株の実際の性能を測定し、これらの優れた実際の性能に基づいて限られた数のこれらの選択された株を選択し (3314)、設計変化およびこれらの結果として生じる性能を予測モデルに加える (3316)。線形回帰の例において、設計変化およびこれらの関連した性能のセットを表 4 中の新しい行として加える。

50

・次いで分析装置 2 1 4 は、新しい設計候補株の生成に戻って反復し（3 3 0 6）、停止条件が満たされるまで反復し続ける。停止条件は、例えば、性能測定基準、例えば、収率、増殖速度、または力価を満たす少なくとも 1 つの微生物株の測定された性能を含み得る。

【0 6 2 5】

上記例では、株設計の反復最適化は、フィードバックおよび線形回帰を使用して機械学習をインプリメントする。一般に、機械学習は、限られた数の標識データの例を使用し、次いで未知のデータに対して同じタスクを実施して、情報タスク（分類または回帰など）の性能における性能判定基準、例えば、パラメータ、技法、または他のフィーチャの最適化として記述することができる。上記線形回帰の例のものなどの教師付き機械学習では、機械（例えば、演算デバイス）は、例えば、訓練データが呈するパターン、カテゴリー、統計的關係、または他の属性を同定することによって学習する。次いで学習の結果が使用されて、新しいデータが、同じパターン、カテゴリー、統計的關係、または他の属性を呈するかどうかは予測される。

【0 6 2 6】

本開示の実施形態は、訓練データが入手可能であるとき、他の教師付き機械学習技法を使用することができる。訓練データの非存在下で、諸実施形態では、教師なし機械学習が使用される場合がある。代替として、諸実施形態では、少量の標識データおよび大量の未標識データを使用して、半教師付き機械学習が使用され得る。諸実施形態では、フィーチャ選択が使用されて最も関連したフィーチャのサブセットが選択されることによって機械学習モデルの性能が最適化される場合もある。選択される機械学習手法のタイプに応じて、線形回帰の代替として、またはそれに加えて、諸実施形態では、例えば、ロジスティック回帰、ニューラルネットワーク、サポートベクターマシン（SVM）、デシジョンツリー、隠れマルコフモデル、ベイジアンネットワーク、グラムシュミット、強化ベースの学習、階層的クラスタリングを含めたクラスターに基づく学習、遺伝的アルゴリズム、および当技術分野で公知の任意の他の適当な学習機械が使用され得る。特に、諸実施形態では、ロジスティック回帰が使用されて分類自体とともに分類（例えば、異なる機能群への遺伝子の分類）の確率が提供され得る。例えば、Shevade、A simple and efficient algorithm for gene selection using sparse logistic regression、Bioinformatics、19 巻、17 号、2003 年、2246～2253 頁、Lengら、Classification using functional data analysis for temporal gene expression data、Bioinformatics、22 巻、1 号、Oxford University Press（2006 年）、68～76 頁を参照。これらのすべては、その全体が参照により本明細書に組み込まれている。

【0 6 2 7】

諸実施形態では、グラフィックスプロセッシングユニット（GPU）加速アーキテクチャが使用される場合があり、これは、特に、ディープニューラルネットワーク（DNN）として公知の形態で、機械学習タスクの実施において評判が増大していることが見出されている。本開示の実施形態では、GPU に基づく機械学習、例えば、GPU-Based Deep Learning Inference: A Performance and Power Analysis、Nvidia Whitepaper、2015 年 11 月、Dahlら、Multi-task Neural Networks for QSAR Predictions、コンピューターサイエンス学科、トロント大学、2014 年 6 月（arXiv:1406.1231 [stat.ML]）に記載されているものが使用され得る。これらの文献のすべては、本明細書にその全体が参照により組み込まれている。本開示の実施形態に適用可能な機械学習技法は、他の参考文献の中でも、Libbrechtら、Machine learning applications in genetics and genomics、Nature Reviews: Genetics、16 巻、2015 年 6 月、Kashyapら、Big Data Analytics in Bioinformatics: A Machine Learning Perspective、Journal of Latex Class Files、13 巻、9 号、2014 年 9 月、Prompramoteら、Machine Learning in Bioinformatics、Bioinformatics Technologies の第 5 章、117～153 頁、Springer Berlin Heidelberg、2005 年にも見出すことができ、これらのすべては、本明細書にその全体が参照により組み込まれている。

【0628】

反復予測的株設計：例

以下は、上記に概説した反復予測的株設計ワークフローの例示的用途を提供するものである。

【0629】

訓練入力変数および出力変数の初期セットを調製した。このセットは、確定した遺伝子組成を有する1864のユニークな操作された株を含んでいた。各株は、5から15の間の操作された変化を含有していた。合計336のユニークな遺伝子変化が該訓練に存在した。

【0630】

初期の予測コンピューターモデルを開発した。インプリメンテーションは、一般化線形モデル（4次多項式カーネルを用いたカーネルリッジ回帰）を使用した。インプリメンテーションは、2つの個別の表現型（収率および生産性）をモデル化する。これらの表現型を、以下に示す通り、加重和として組み合わせてランク付けのための単一スコアを得た。様々なモデルパラメータ、例えば、正則化係数を、指定された訓練データに対してk倍交差検証を介してチューニングした。

【0631】

インプリメンテーションは、上記上位性マッピングセクションに記載した相互作用効果のいずれの明確な分析も組み入れない。しかし、当業者が理解するはずであるように、インプリメントされる一般化線形モデルは、カーネルの2、3、および4次項を通じて暗黙的に相互作用効果を捕捉することができる。

【0632】

モデルを訓練セットに対して訓練した。フィットされたモデルは、収率に関して0.52の R^2 値（決定の係数）および生産性に関して0.67の R^2 値を有する。図46は、収率モデルの訓練データへの有意な品質フィッティングを実証する。

【0633】

候補株を生成した。本例は、親株への新しい遺伝子変化の導入に関連した連続構築制約（serial build constraint）を含む（本例では、ただ1つの新しい変異を操作して同時に株内に入れた）。ここでは、候補は、所望の数の変化の関数として単に見なされない。代わりに、分析装置214は、出発点として、高性能測定基準を有すると分かっている先に設計された株（「シード株」）のコレクションを選択した。分析装置214は個々に、遺伝子変化をシード株のそれぞれに適用した。導入された遺伝子変化は、シード株内に既に存在するものを含まなかった。様々な技術的、生物学的、または他の理由で、ある特定の変異、例えば、opca_4が明示的に要求され、または例えば、dss_422が明示的に除外された。166の入手可能なシード株およびモデルによって特徴付けられた336の変化を使用して、6239の新規候補株を設計した。

【0634】

モデルに基づいて、分析装置214は、候補株設計の性能を予測した。分析装置214は、目的の2つの表現型（収率および生産性）に関する予測された性能に基づいて「最良」から「最悪」に候補をランク付けた。具体的には、分析装置214は、加重和を使用して候補株をスコア付けした：

【0635】

スコア = $0.8 * yield / \max(yields) + 0.2 * prod / \max(prods)$ 、

式中、yieldは、候補株について予測された収率を表し、

$\max(yields)$ は、すべての候補株に対する最大収率を表し、

prodは、候補株についての生産性を表し、

$\max(prods)$ は、すべての候補株に対する最大収率を表す。

【0636】

分析装置214は、能力の制約およびオペレーションの制約の両方を課すことによって

候補のランク付けされたリストから推奨の最終セットを生成した。本例では、能力限界は、48のコンピューター生成候補設計株で設定した。オペレーションの制約に起因して、本例では、ただ1つのシード株を96ウェルプレートの1列あたりに使用した。これは、シード株を選択した後、その株に対して最大で8つの変化を構築することができたが、わずか6のシード株しか任意の所与の週内に選択することができなかったことを意味する。

【0637】

訓練されたモデル（上述した）を使用して各候補株の予期される性能（収率および生産性についての）を予測した。分析装置214は、上記に与えたスコア付関数を使用して候補株をランク付けた。能力およびオペレーションの制約を適用して48の候補株の選別されたセットを得た。選別された候補株のこのセットを図47に表す。

10

【0638】

選別された候補株を、発注エンジン208が作成した工場注文に基づいて構築した（ファクトリー210において）（3312）。注文は、候補株に対応するDNAの仕様に基づいた。

【0639】

実際には、ビルドプロセスは、予期された失敗率を有し、それによって株のランダムセットを構築しない。このビルドサイクルについて、候補株のおおよそ20%が構築するのに失敗し、37の構築された株（built strain）をもたらした。

【0640】

分析装置214を使用して選択された株の実際の収率および生産性性能を測定した。分析装置214は、3つの判定基準：モデル精度；株性能の改善；および人間の専門家が作成した設計との同等性（または改善）に基づいてモデルおよび推奨株を評価した。

20

【0641】

収率および生産性表現型を、推奨株について測定し、モデルが予測した値と比較した。図48に示した通り、モデルは、有用な予測的有用性を実証する。特に、推奨株について予測された収率値は、対応する観察結果と0.59のピアソン-r相関係数を有する。

【0642】

次に、分析装置214は、推奨株のそれぞれの親株からのパーセンテージ性能変化を計算した。このデータを図49に（薄灰色で）示す。本発明者らは、予測された株の多くが、実際には、これらの直接の親に対して予期された性能獲得を呈することを見出した。特に、最良の予測された株は、その直接の親に対して6%の収率の改善を示した。

30

【0643】

上述したモデルに基づく株設計プロセスと並行して、48の株のコレクションを、人間の専門家が独立に設計した。これらの株のうち、37を順調に構築し、試験した。このデータは、モデルに基づく株設計が、人間の専門家が設計した株と同等に働くことを実証した。これらの専門家は、本発明の譲受人が雇用した、または別段に従事させている本開示の実施形態に精通している高度に熟練した（例えば、Ph.D.レベル）科学者である。2つの方法を比較するために、本発明者らは、各群の性能分布を最初に検査した（図50）。この実験では、モデルに基づく株の平均収率は、人間の専門家が作成した設計に対して1%の増大を示した。

40

【0644】

次いで本発明者らは、バックグラウンドによってグループ化された人間の専門家が設計した、およびコンピューターモデルが設計した株、すなわち、同じ親をもつ新しい株を比較した（図51）。やはり、本発明者らは、コンピューターが作成した設計は、人間の専門家が作成した設計と同等に、かつ一部の 경우에는、それより良好に働き、さらに、より少ないばらつきを生じる傾向があることを見出した。最後に、本発明者らは、人間の専門家およびモデルが設計した株の親株に対する変化率を比較した（図49）。やはり、これらの集団は、同等の獲得を示した。

【0645】

表にした要約統計については表4.1を参照。

50

【表 4 . 1 - 1】

表4.1ー予測モデルおよびヒト専門家の参照によって設計された株の測定された性能統計

		収率 [AU]	親からの 収率変化 [%]	生産性 [AU]	親からの生産性 変化 [%]
設計方法					
コンピューターモデル	カウント	37	37	37	37
	平均	1.058068108	0.3578340	0.737928919	-2.5428848
	std	0.017811031	1.8293665	0.083619804	9.6743873
	最小	1.015310000	-4.5346677	0.572780000	-23.3626353
	中央値	1.058710000	0.005007939	0.766870000	-1.1824159

【表 4 . 1 - 2】

	最大	1.093510000	6.0097309	0.872790000	26.6124119
ヒト専門家	カウント	37	37	37	37
	平均	1.038804595	-0.0005237	0.748320811	-1.6126436
	std	0.032053625	1.9227716	0.120527468	9.8530758
	最小	0.964910000	-3.1043233	0.535980000	-21.4589256
	中央値	1.045530000	0.0449168	0.760300000	-1.9241048
	最大	1.094790000	7.8487174	0.984110000	21.7335193

【0646】

予測、それに続く構築、それに続く試験サイクルの各ラウンドの最後に、本発明者らは、モデル予測の品質の評価および新しいデータの先のモデルへの反復組み入れに興味を持った。前者、モデル評価について、本発明者らは、モデル予測を実験的測定値と比較することによって予測精度を測定することに着目した。予測精度は、予測値と観測値との間の関連性の強度を示す相関係数、または平均モデル誤差の尺度である二乗平均平方根誤差を含めたいくつかの方法によって評価することができる。

【0647】

実験の多くのラウンドにわたって、モデル予測がずれる場合があり、新しい遺伝子変化を訓練入力に加えて予測精度を改善することができる。本例に関して、設計変化およびこれらの結果として生じる性能を予測モデルに加えた(3316)。

10

【0648】

サービスとしてのゲノム設計および操作

本開示の実施形態では、図32のLIMSシステムソフトウェア3210を、図32のクラウドコンピューティングシステム3202において実装して、複数のユーザーが本開示の実施形態によって微生物株を設計および構築するのを可能にすることができる。図32は、本開示の実施形態によるクラウドコンピューティング環境3204を例示する。図32に例示したものなどのクライアントコンピューター3206は、インターネットなどのネットワーク3208を介してLIMSシステムにアクセスする。諸実施形態では、LIMSシステムアプリケーションソフトウェア3210は、クラウドコンピューティングシステム3202内に存在する。LIMSシステムは、図32に例示したタイプの1つまたは複数のプロセッサを使用して1つまたは複数のコンピューターシステムを使用することができる。クラウドコンピューティングシステム自体は、ネットワーク3208を介してLIMSシステムアプリケーション3210をクライアントコンピューター3206にインターフェースをとるためのネットワークインターフェース3212を含む。ネットワークインターフェース3212は、アプリケーションプログラミングインターフェース(API)を含むことによって、クライアントコンピューター3206におけるクライアントアプリケーションがLIMSシステムソフトウェア3210にアクセスするのを可能にすることができる。特に、APIを通じて、クライアントコンピューター3206は、限定することなく、入力インターフェース202、インタープリター204、エグゼキューションエンジン207、発注エンジン208、ファクトリー210をランさせるソフトウェア、ならびに試験装置212および分析装置214を含めたLIMSシステム200のコンポーネントにアクセスすることができる。サービス型ソフトウェア(SaaS)のソフトウェアモジュール3214は、クライアントコンピューター3206へのサービス型LIMSシステムソフトウェア3210を提供する。クラウド管理モジュール3216は、クライアントコンピューター3206によるLIMSシステム3210へのアクセスを管理する。クラウド管理モジュール3216は、マルチテナントアプリケーションを使用するクラウドアーキテクチャ、当技術分野で公知の仮想化または他のアーキテクチャが、複数のユーザーにサービスすることを可能にし得る。

20

30

【0649】

ゲノム自動化

本開示の方法を自動化すると、同時に複数の試験株バリエーションからの標的産物のハイスループット表現型スクリーニングおよび同定が可能になる。

40

【0650】

上述のゲノム操作予測モデルリングプラットフォームは、数百および数千の変異体株がハイスループット様式で構築される事実を前提としている。以下に記載のロボットシステムおよびコンピューターシステムは、このようなハイスループットプロセスを実行することができる構造機構である。

【0651】

一部の実施形態では、本開示は、宿主細胞生産性を改善し、または工業用株を再建する

50

方法を教示する。このプロセスの一部として、本開示は、DNAをアセンブルし、新しい株を構築し、プレート中の培養物をスクリーニングし、タンク発酵のモデルにおける培養物をスクリーニングする方法を教示する。一部の実施形態では、本開示は、新しい宿主株を創製および試験する上述の方法の1つまたは複数は、自動化ロボットによって助けられることを教示する。

【0652】

一部の実施形態では、本開示は、図6A～Bまたは図26に表したハイスループット株操作プラットフォームを教示する。

【0653】

H T Pロボットシステム

一部の実施形態では、本開示の自動化された方法は、ロボットシステムを含む。本明細書に概説したシステムは一般に、96または384ウェルマイクロタイタープレートの使用を対象とするが、当業者が十分に理解するように、任意の数の異なるプレートまたは構成を使用することができる。さらに、本明細書に概説したステップのいずれか、またはすべてを自動化することができ、よって例えば、システムは、完全または部分的に自動化され得る。

【0654】

一部の実施形態では、本開示の自動化システムは、1つまたは複数の作業モジュールを含む。例えば、一部の実施形態では、本開示の自動化システムは、DNA合成モジュール、ベクタークローニングモジュール、株形質転換モジュール、スクリーニングモジュール、およびシーケンシングモジュールを含む（図7を参照）。

【0655】

当業者が十分に理解するとおり、自動化システムは、これらに限定されないが、液体ハンドラー；1つまたは複数のロボットアーム；マイクロプレートを適所配置するためのプレートハンドラー；プレートシーラー、プレートピアサー、非交差汚染プレート上のウェルの蓋を除去および交換するための自動化蓋ハンドラー；使い捨てチップを用いて試料分配するための使い捨てチップアセンブリー；試料分配のための洗浄可能なチップアセンブリー；96ウェルローディングブロック；一体型サーマルサイクラー；冷却された試薬ラック；マイクロタイタープレートピペットポジション（任意選択で冷却された）；プレートおよびチップのためのスタッキングタワー；磁気ビーズ処理ステーション；濾過システム；プレートシェーカー；バーコードリーダーおよびアプリケーションター；ならびにコンピューターシステムを含めた多種多様なコンポーネントを含み得る。

【0656】

一部の実施形態では、本開示のロボットシステムは、ハイスループットピペット操作を可能にして遺伝子標的化および組換え用途のプロセスにおけるすべてのステップを実施する自動化液体および粒子ハンドリングを含む。これは、液体および粒子マニピュレーション、例えば、吸引、分注、混合、希釈、洗浄、正確な容量移送；ピペットチップの取戻しおよび廃棄；ならびに単一試料吸引からの複数の送達についての同一体積の繰り返しピペット操作などを含む。これらのマニピュレーションは、交差汚染のない液体、粒子、細胞、および生物体移送である。機器は、フィルター、膜、および／または娘プレートへのマイクロプレート試料の自動化複製、高密度移送、フルプレート連続希釈、ならびに大容量オペレーションを遂行する。

【0657】

一部の実施形態では、本開示の特注自動化液体ハンドリングシステムは、TECAN機械（例えば、特注TECAN Freedom Evo）である。

【0658】

一部の実施形態では、本開示の自動化システムは、マルチウェルプレート、ディープウェルプレート、スクウェアウェルプレート、試薬トラフ、試験管、ミニチューブ、マイクロチューブ、クライオバイアル、フィルター、マイクロアレイチップ、光ファイバー、ビーズ、アガロースゲルおよびアクリルアミドゲルのためのプラットフォームに適合し、他

の固相マトリックスまたはプラットフォームは、アップグレード可能なモジュラーデッキ上に收容されている。一部の実施形態では、本開示の自動化システムは、ソースおよび産出試料、試薬、試料および試薬希釈、アッセイプレート、試料および試薬リザーバー、ピペットチップ、ならびにアクティブチップ洗浄ステーションを配置するための多位置作業面のための少なくとも1つのモジュラーデッキを含有する。

【0659】

一部の実施形態では、本開示の自動化システムは、ハイスループット電気穿孔システムを含む。一部の実施形態では、ハイスループット電気穿孔システムは、96または384ウェルプレート内で細胞を形質転換することができる。一部の実施形態では、ハイスループット電気穿孔システムとしては、VWR（登録商標）ハイスループット電気穿孔システム、BTX（商標）、Bio-Rad（登録商標）Gene Pulser MXcell（商標）、または他のマルチウェル電気穿孔システムが挙げられる。

10

【0660】

一部の実施形態では、一体型サーマルサイクラーおよび／または温度調節器が、制御ブロックまたはプラットフォームなどの熱交換器の温度を安定化させるのに使用されて、0℃～100℃に試料をインキュベートする正確な温度制御をもたらす。

【0661】

一部の実施形態では、本開示の自動化システムは、液体、粒子、細胞、および多細胞生物をロボット制御でマニピュレートすることができる、単一もしくは複数の磁気プローブ、親和性プローブ、レプリケーター、またはピペッタを有する交換可能なマシンヘッド（単一またはマルチチャンネル）に適合する。マルチウェルまたはマルチチューブ磁気分離器および濾過ステーションは、単一または複数の試料フォーマットにおいて液体、粒子、細胞、および生物体をマニピュレートする。

20

【0662】

一部の実施形態では、本開示の自動化システムは、カメラビジョンおよび／または分光計システムに適合する。よって、一部の実施形態では、本開示の自動化システムは、進行中の細胞培養における色および吸収の変化を検出およびロギングすることができる。

【0663】

一部の実施形態では、本開示の自動化システムは、複数のハードウェアアドオンに柔軟であり、適応できることによって、システムが複数のアプリケーションを実行することを可能にするように設計される。ソフトウェアプログラムモジュールは、方法の創製、改変、および試行を可能にする。システムの診断モジュールは、セットアップ、機器整列、およびモータ動作を可能にする。特注のツール、実験機器、ならびに液体および粒子移送パターンは、異なるアプリケーションをプログラムおよび実施することを可能にする。データベースは、方法およびパラメータ記憶を可能にする。ロボットインターフェースおよびコンピューターインターフェースは、機器間の通信を可能にする。

30

【0664】

よって、一部の実施形態では、本開示は、図26に表したハイスループット株操作プラットフォームを教示する。

【0665】

当業者は、本開示のHTP操作方法を実行することができる様々なロボットプラットフォームを認識する。以下の表5は、図26に記載した本開示のHTP操作ステップの各ステップを実行することができる科学的装置の非排他的リストを提供する。

40

【表 5 - 1】

表 5-本開示の HTP 操作方法に適合する科学的装置の非排他的なリスト。

	装置タイプ	実施したオペレーション	適合する装置 メーカー (Make) /モデル/構成
DNA 片の取得および構築	液体ハンドラー	DNA 部分の PCR 増幅のプライマー/鋳型のヒットピッキング(移行による組合せ)	Hamilton Microlab STAR, Labcyte Echo 550, Tecan EVO 200, Beckman Coulter Biomek FX, または等価物
	サーマルサイクラー	DNA 部分の PCR 増幅	Inheco Cyclor, ABI 2720, ABI Proflex 384, ABI Veriti, または等価物
QC DNA 部分	断片分析器(キャピラリー電気泳動)	適切なサイズの PCR 産物を確認するためのゲル電気泳動	Agilent Bioanalyzer, AATI Fragment Analyzer, または等価物
	シーケンサー (サンガー:Beckman)	部分/鋳型の配列の検証	Beckman Ceq-8000, Beckman GenomeLab(商標), または等価物
	NGS(次世代シーケンシング) 機器	部分/鋳型の配列の検証	Illumina MiSeq シリーズシーケンス, illumina Hi-Seq, Ion torrent, pac bio または他の等価物
	ナノドロップ/プレートリーダー	DNA 試料の濃度の評価	Molecular Devices SpectraMax M5, Tecan M1000, または等価物

10

20

30

40

【表 5 - 2】

DNA アセンブリの生 成	液体ハンドラ ー	クローニングベクターと ともにアセンブルするた めの DNA 部分のヒット ピッキング(移行による組 合せ)、アセンブリー反応 /プロセスのための試薬の 添加	Hamilton Microlab STAR, Labcyte Echo 550, Tecan EVO 200, Beckman Coulter Biomek FX, または等価物	10
	コロニーピッ カー	液体培地にコロニーを接 種するため	Scirobotics Pickolo, Molecular Devices QPix 420	
QC DNA アセンブリー	液体ハンドラ ー	プライマー/鋳型のヒット ピッキング、試料の希釈	Hamilton Microlab STAR, Labcyte Echo 550, Tecan EVO 200, Beckman Coulter Biomek FX, または等価物	20
	断片分析器(キ ャピラリー電 気泳動)	適切なサイズのアセンブ ルされた産物を確認する ためのゲル電気泳動	Agilent Bioanalyzer, AATI Fragment Analyzer	
	シーケンサー (サンガ ー:Beckman)	アセンブルされたプラス ミドの配列の検証	ABI3730 Thermo Fisher, Beckman Ceq-8000, Beckman GenomeLab(商標), または等 価物	
	NGS(次世代シ ーケンシング) 機器	アセンブルされたプラス ミドの配列の検証	Illumina MiSeq シリーズシー ケンス, illumina Hi-Seq, Ion torrent, pac bio または他の等 価物	30
基準株および DNA アセンブリの 調製	遠心分離機	細胞のスピンング/ペレッ ト化	Beckman Avanti 床遠心分離 機, Hettich Centrifuge	40

【表 5 - 3】

DNA の基準株への形質転換	エレクトロポレーター	細胞のエレクトロポレーションによる形質転換	BTX Gemini X2, BIO-RAD MicroPulser Electroporator
	弾道形質転換	細胞の弾道形質転換	BIO-RAD PDS1000
	インキュベーター、サーマルサイクラー	化学変換/熱ショックのため	Inheco Cyclor, ABI 2720, ABI Proflex 384, ABI Veriti, または等価物
	液体ハンドラー	DNA、細胞、緩衝液を組み合わせるため	Hamilton Microlab STAR, Labcyte Echo 550, Tecan EVO 200, Beckman Coulter Biomek FX, または等価物
基準株のゲノム内への DNA の組み込み	コロニーピッカー	液体培地にコロニーを接種するため	SciRobotics Pickolo, Molecular Devices QPix 420
	液体ハンドラー	寒天培地への細胞の移動、培養プレートから異なる培養プレートへの移動のため(他の選択培地への接種)	Hamilton Microlab STAR, Labcyte Echo 550, Tecan EVO 200, Beckman Coulter Biomek FX, または等価物
	プラットフォームシェーカー-インキュベーター	マイクロタイタープレート培養物の振とうによるインキュベーション	Kuhner Shaker ISF4-X, Inforsht Multitron Pro

10

20

30

【表 5 - 4】

QC 形質転換株	コロニーピッカー	液体培地にコロニーを接種するため	Scirobotics Pickolo, Molecular Devices QPix 420	10
	液体ハンドラー	プライマー/鋳型のヒットピッキング、試料の希釈	Hamilton Microlab STAR, Labcyte Echo 550, Tecan EVO 200, Beckman Coulter Biomek FX, または等価物	
	サーマルサイクラー	株の cPCR 検証	Inheco Cyclor, ABI 2720, ABI Proflex 384, ABI Veriti, または等価物	
	断片分析器(キャピラリー電気泳動)	適切なサイズの cPCR 産物を確認するためのゲル電気泳動	Infors-ht Multitron Pro, Kuhner Shaker ISF4-X	20
	シーケンサー(サンガー:Beckman)	導入された修飾の配列検証	Beckman Ceq-8000, Beckman GenomeLab(商標), または等価物	
	NGS(次世代シーケンシング)機器	導入された修飾の配列検証	Illumina MiSeq シリーズシーケンス, illumina Hi-Seq, Ion torrent, pac bio または他の等価物	
試験プレートへの QC された株の選択および統合	液体ハンドラー	培養プレートから異なる培養プレートへの移動のため(産生培地への接種)	Hamilton Microlab STAR, Labcyte Echo 550, Tecan EVO 200, Beckman Coulter Biomek FX, または等価物	30
	コロニーピッカー	液体培地にコロニーを接種するため	Scirobotics Pickolo, Molecular Devices QPix 420	
	プラットフォームシェーカー-インキュベーター	マイクロタイタープレート培養物の振とうによるインキュベーション	Kuhner Shaker ISF4-X, Infors-ht Multitron Pro	40

【表 5 - 5】

シードプレート中の株の培養	液体ハンドラー	培養プレートから異なる培養プレートへの移動のため(産生培地への接種)	Hamilton Microlab STAR, Labcyte Echo 550, Tecan EVO 200, Beckman Coulter Biomek FX, または等価物
	プラットフォームシェーカー-インキュベーター	マイクロタイタープレート培養物の振とうによるインキュベーション	Kuhner Shaker ISF4-X, Infors-ht Multitron Pro
	液体ディスペンサー	液体培養基のマイクロタイタープレートへの分注	Well mate (Thermo), Benchcel2R (velocity 11), plateloc (velocity 11)
	マイクロプレートラベラー	バーコーダーのプレートへの適用	Microplate labeler (a2+ cab - agilent), benchcell 6R (velocity11)
株からの産物の産生	液体ハンドラー	培養プレートから異なる培養プレートへの移動のため(産生培地への接種)	Hamilton Microlab STAR, Labcyte Echo 550, Tecan EVO 200, Beckman Coulter Biomek FX, または等価物
	プラットフォームシェーカー-インキュベーター	マイクロタイタープレート培養物の振とうによるインキュベーション	Kuhner Shaker ISF4-X, Infors-ht Multitron Pro
	液体ディスペンサー	液体培養基の複数のマイクロタイタープレートおよびシールプレートへの分注	well mate (Thermo), Benchcel2R (velocity 11), plateloc (velocity 11)
	マイクロプレートラベラー	バーコーダーのプレートへの適用	microplate labeler (a2+ cab - agilent), benchcell 6R (velocity11)

10

20

30

40

【表 5 - 6】

性能の評価	液体ハンドラ ー	下流分析のための培養ブ ロスの処理のため	Hamilton Microlab STAR, Labcyte Echo 550, Tecan EVO 200, Beckman Coulter Biomek FX, または等価物	10
	UHPLC, HPLC	前駆体および標的化合物 の定量分析	Agilent 1290 Series UHPLC ならびに UV および RI 検出 器を有する 1200 Series HPLC、または等価物、また 任意の LC/MS	
	LC/MS	前駆体および標的化合物 ならびに副産物および分 解産物の高度に特異的な 分析	1290 Series UHPLC にカップ リングした Agilent 6490 QQQ および 6550 QTOF	
	分光光度計	分光光度計に基づくアッ セイを使用する異なる化 合物の定量	Tecan M1000, spectramax M5, Genesys 10S	
培養中の株の培 養 プラットフォーム	発酵槽:	振とうでのインキューペ ーション	Sartorius, DASGIPs (Eppendorf), BIO-FLOs (Sartorius-stedim). Applikon	20
	プラットフォ ームシェーカ ー		innova 4900, または任意の等 価物	
株からの産物の産生	発酵槽: DASGIPs (Eppendorf), BIO-FLOs (Sartorius-stedim)			30

【表 5－7】

性能の評価	液体ハンドラー	培養プレートから異なる培養プレートへの移動のため(産生培地への接種)	Hamilton Microlab STAR, Labcyte Echo 550, Tecan EVO 200, Beckman Coulter Biomek FX, または等価物
	UHPLC, HPLC	前駆体および標的化合物の定量分析	Agilent 1290 Series UHPLC ならびに UV および RI 検出器を有する 1200 Series HPLC、または等価物、また任意の LC/MS
	LC/MS	前駆体および標的化合物ならびに副産物および分解産物の高度に特異的な分析	1290 Series UHPLC にカップリングした Agilent 6490 QQQ および 6550 QTOF
	フローサイトメーター	株性能の特徴付け(生存能の測定)	BD Accuri, Millipore Guava
	分光光度計	株性能の特徴付け(バイオマスの測定)	Tecan M1000, Spectramax M5, または他の等価物

10

20

【0666】

コンピューターシステムハードウェア

図 34 は、本開示の実施形態による非一時的コンピューター可読媒体（例えば、メモリー）内に記憶されたプログラムコードを実行するのに使用され得るコンピューターシステム 800 の一例を例示する。コンピューターシステムは、入力／出力サブシステム 802 を含み、これは、アプリケーションに応じて人間のユーザーおよび／または他のコンピューターシステムとインターフェースで接続するのに使用することができる。I/O サブシステム 802 は、例えば、キーボード、マウス、グラフィカルユーザーインターフェース、タッチスクリーン、または入力のための他のインターフェース、および例えば、LED もしくは他のフラットスクリーンディスプレイ、あるいはアプリケーションプログラムインターフェース（API）を含めた出力のための他のインターフェースを含み得る。LIMS システムのコンポーネントなどの本開示の実施形態の他のエレメントは、コンピューターシステム 800 のもののようなコンピューターシステムで実装することができる。

30

【0667】

プログラムコードは、二次メモリー 810 もしくはメインメモリー 808 または両方内の永続記憶装置などの非一時的媒体内に記憶することができる。メインメモリー 808 として、ランダムアクセスメモリー（RAM）などの揮発性メモリー、またはリードオンリーメモリー（ROM）などの不揮発性メモリー、ならびに指示およびデータにより速くアクセスするための異なるレベルのキャッシュメモリーを挙げることができる。二次メモリーとして、永続記憶装置、例えば、ソリッドステートドライブ、ハードディスクドライブ、または光ディスクなどを挙げることができる。1 つまたは複数のプロセッサ 804 は、1 つまたは複数の非一時的媒体からのプログラムコードを読み込み、コードを実行してコンピューターシステムが本明細書の実施形態によって実施される方法をなしとげることができる。当業者は、プロセッサ（複数可）は、ソースコードを取り込み、ソースコードを解釈またはコンパイルして、プロセッサ（複数可）804 のハードウェアゲートレベルで理解可能であるマシンコードにすることができることを理解する。プロセッサ（複数可）804 は、計算集約型タスクを扱うためのグラフィックスプロセッシングユニット（GPU）を含み得る。特に、機械学習では、1 つまたは複数の CPU 804 は、

40

50

大量のデータの処理を1つまたは複数のGPU 804にオフロードすることができる。

【0668】

プロセッサ（複数可）804は、1つまたは複数の通信インターフェース807、例えば、ネットワークインターフェースカード、WiFiトランシーバーなどを介して外部ネットワークと通信することができる。バス805は、I/Oサブシステム802、プロセッサ（複数可）804、周辺デバイス806、通信インターフェース807、メモリー808、および永続記憶装置810を通信可能にカップルする。本開示の実施形態は、この代表的なアーキテクチャに限定されない。代替の実施形態は、異なる配置およびタイプのコンポーネント、例えば、入出力コンポーネントおよびメモリーサブシステムのための別個のバスを使用することができる。

10

【0669】

当業者は、本開示の実施形態のエレメントの一部またはすべて、およびこれらの付随するオペレーションは、コンピューターシステム800のもののような1つまたは複数のプロセッサおよび1つまたは複数のメモリーシステムを含む1つまたは複数のコンピューターシステムによって完全または部分的にインプリメントされ得ることを理解する。特に、本明細書に記載のLIMSシステム200ならびに任意のロボットおよび他の自動化システムまたはデバイスのエレメントは、コンピューター実装することができる。一部のエレメントおよび機能性は、ローカルにインプリメントすることができ、他のものは、例えば、異なるサーバーを通じたネットワークにわたって分散様式で、例えば、クライアント-サーバー様式でインプリメントすることができる。特に、図32に示したように、サーバーサイドオペレーションを、サービス型ソフトウェア（SaaS）様式で複数のクライアントに利用可能にすることができる。

20

【0670】

本文脈における用語のコンポーネントは、ソフトウェア、ハードウェア、もしくはファームウェア（またはこれらの任意の組合せ）コンポーネントを広く指す。コンポーネントは、典型的には、指定された入力（複数可）を使用して有用なデータまたは他の出力を生成することができる機能コンポーネントである。コンポーネントは、自己充足型であってもよく、または自己充足型でなくてもよい。アプリケーションプログラム（「アプリケーション」とも呼ばれる）は、1つもしくは複数のコンポーネントを含み得、またはコンポーネントは、1つもしくは複数のアプリケーションプログラムを含むことができる。

30

【0671】

一部の実施形態は、他のモジュールまたはアプリケーションコンポーネントとともにコンポーネントの一部、すべてを含み、またはどれも含まない。なおさらに、様々な実施形態は、これらのコンポーネントの2つもしくはそれ超を単一モジュール内に組み入れ、かつ/またはこれらのコンポーネントの1つもしくは複数の機能性の一部を異なるコンポーネントと結びつけることができる。

【0672】

用語「メモリー」は、情報を記憶させるのに使用される任意のデバイスまたは機構であり得る。本開示の一部の実施形態によれば、メモリーは、これらに限定されないが、揮発性メモリー、不揮発性メモリー、およびダイナミックメモリーの任意のタイプを包含するように意図されている。例えば、メモリーは、ランダムアクセスメモリー、メモリー記憶デバイス、光学メモリーデバイス、磁気媒体、フロッピー（登録商標）ディスク、磁気テープ、ハードドライブ、SIMM、SDRAM、DIMM、RDRAM、DDR RAM、SODIMMS、消去可能プログラム可能リードオンリーメモリー（EPROM）、電氣的消去可能プログラム可能リードオンリーメモリー（EEPROM）、コンパクトディスク、DVDなどであり得る。一部の実施形態によれば、メモリーは、1つまたは複数のディスクドライブ、フラッシュドライブ、データベース、ローカルキャッシュメモリー、プロセッサキャッシュメモリー、リレーショナルデータベース、フラットデータベース、サーバー、クラウドベースプラットフォームなどを含み得る。さらに、当業者は、情報を記憶するための多くの追加のデバイスおよび技法をメモリーとして使用することができ

40

50

ることを十分に理解する。

【0673】

メモリーは、プロセッサ上の1つまたは複数のアプリケーションまたはモジュールをランさせるための指示を記憶させるのに使用され得る。例えば、メモリーは、一部の実施形態では、本開示の開示のモジュールおよび／またはアプリケーションの1つまたは複数の機能性を実行するのに必要な指示のすべてまたは一部を収納するのに使用することができる。

【0674】

遺伝子設計予測に基づくHTP微生物株操作：例示的なワークフロー

一部の実施形態では、本開示は、本開示のコンピューター分析システムの推奨に基づく新しい宿主生物の指向操作を教示する。

10

【0675】

一部の実施形態では、本開示は、すべての遺伝子設計およびクローニング方法に適合する。すなわち、一部の実施形態では、本開示は、伝統的なクローニング技法、例えば、ポリメラーゼ連鎖反応、制限酵素消化、ライゲーション、相同組換え、RT-PCR、および当技術分野で一般に公知であり、例えば、参照により本明細書に組み込まれている、Sambrookら（2001年）、Molecular Cloning: A Laboratory Manual（3版、Cold Spring Harbor Laboratory Press、Plainview、New York）に開示されているその他のものなどの使用を教示する。

【0676】

20

一部の実施形態では、クローン化された配列は、本明細書に教示のHTP遺伝子設計ライブラリーのいずれか由来の可能なもの、例えば、プロモータースワップライブラリー由来のプロモーター、SNPスワップライブラリー由来のSNP、開始／停止コドン交換ライブラリー由来の開始もしくは停止コドン、STOPスワップライブラリー由来のターミネーター、SOLUBILITY TAGスワップライブラリー由来のタンパク質可溶性タグ、DEGRADATION TAGスワップライブラリー由来のタンパク質分解タグまたは配列最適化ライブラリー由来の配列最適化を含み得る。

【0677】

さらに、特定の構築物内に含まれるべき正確な配列組合せは、上位性マッピング機能によって知らされ得る。

30

【0678】

他の実施形態では、クローン化された配列は、合理的な設計（仮説主導型）に基づく配列および／または科学刊行物などの他のソースに基づく配列も含むことができる。

【0679】

一部の実施形態では、本開示は、i) カスタムメイドSNP特異的DNAを生成するステップと、ii) SNP特異的プラスミドをアセンブルするステップと、iii) SNP特異的DNAで標的宿主細胞を形質転換するステップと、iv) 任意の選択マーカーをループアウトさせるステップとを含む、指向操作の方法を教示する（図2を参照）。

【0680】

図6Aは、DNAを取得およびアセンブルステップと、ベクターをアセンブルするステップと、宿主細胞を形質転換するステップと、選択マーカーを除去するステップとを含む、本開示の株操作法の一般的なワークフローを表す。

40

【0681】

特異的DNAオリゴヌクレオチドを構築する

一部の実施形態では、本開示は、宿主細胞生物体のDNAセグメントを挿入することおよび／または置き換えることおよび／または変更することおよび／または欠失させることを教示する。一部の態様では、本明細書に教示の方法は、宿主生物のゲノム内に組み入れる目的のオリゴヌクレオチド（すなわち、標的DNAセグメント）を構築することを伴う。一部の実施形態では、本開示の標的DNAセグメントは、公知の鋳型からのコピーもしくはカット、変異、またはDNA合成を含む、当技術分野で公知の任意の方法を介して得

50

ることができる。一部の実施形態では、本開示は、標的DNA配列を生成するための市販の遺伝子合成産物（例えば、GeneArt（商標）、GeneMaker（商標）、GenScript（商標）、Anagen（商標）、Blue Heron（商標）、Entelechon（商標）、GenOsys, Inc.、またはQiagen（商標））に適合する。

【0682】

一部の実施形態では、標的DNAセグメントは、宿主生物の選択されたDNA領域内にSNPを組み入れる（例えば、有益なSNPを付加する）ように設計されている。他の実施形態では、DNAセグメントは、宿主生物のDNAからSNPを除去する（例えば、有害なまたは中立のSNPを除去する）ように設計されている。

10

【0683】

一部の実施形態では、本発明の方法で使用されるオリゴヌクレオチドは、当技術分野で公知の酵素合成または化学合成の方法のいずれかを使用して合成することができる。オリゴヌクレオチドは、固体支持体、例えば、制御細孔ガラス（CPG）、ポリスチレンビーズ、またはCPGを含有し得る熱可塑性ポリマーから構成される膜上で合成することができる。オリゴヌクレオチドは、アレイ上で、マイクロフルイディクスを使用して並列のマイクロスケールで（Tianら、Mol. BioSyst.、5巻、714～722頁（2009年））、または両方の組合せを提供する公知の技術で（Jacobsenら、米国特許出願第2011/0172127号を参照）合成することもできる。

【0684】

アレイ上の、またはマイクロフルイディクスによる合成は、より低い試薬使用を通じてコストを低減することによって、慣例的な固体支持体合成に対して利点を提供する。遺伝子合成に要求される規模は低く、したがって、アレイから、またはマイクロフルイディクスを通じて合成されるオリゴヌクレオチド産物の規模は、許容される。しかし、合成されるオリゴヌクレオチドは、固体支持体合成を使用するときより低い品質のものである（Tian、以下を参照；Staehlerら、米国特許出願第2010/0216648号も参照）。

20

【0685】

多数の進歩が、それが1980年代に最初に記載されて以来、伝統的な4ステップホスホロミダイト化学において達成されている（例えば、ペルオキシアニオンの脱保護を使用するSierzchalaら、J. Am. Chem. Soc.、125巻、13427～13441頁（2003年）；代替の保護基についてHayakawaら、米国特許第6,040,439号；ユニバーサル支持体についてAzhayevら、Tetrahedron、57巻、4977～4986頁（2001年）；大孔CPGの使用によるより長いオリゴヌクレオチドの改善された合成についてKozlovら、Nucleosides, Nucleotides, and Nucleic Acids、24巻（5～7号）、1037～1041頁（2005年）；および改善された誘導体化についてDamhaら、NAR、18巻、3813～3821頁（1990年）を参照）。

30

【0686】

合成のタイプにかかわらず、次いで結果として生じるオリゴヌクレオチドは、より長いオリゴヌクレオチドのためのより小さいビルディングブロックを形成することができる。一部の実施形態では、より小さいオリゴヌクレオチドは、当技術分野で公知のプロトコール、例えば、ポリメラーゼ鎖アセンブリー（PCA）、リガーゼ連鎖反応（LCR）、および熱力学的にバランスの取れたインサイドアウト合成（thermodynamically balanced inside-out synthesis）（TBIO）を使用して一緒に繋ぐことができる（Czarら、Trends in Biotechnology、27巻、63～71頁（2009年）を参照）。PCAでは、所望のより長い産物の全長にまたがるオリゴヌクレオチドが、複数のサイクル（典型的には約55サイクル）においてアニールおよび伸長されて完全長産物を最終的に達成する。LCRは、リガーゼ酵素を使用して2つのオリゴヌクレオチドを繋ぎ、これらはともにアニールされて第3のオリゴヌクレオチドになる。TBIO合成は、所望の産物の中心から開始し、遺伝子の5'末端における順方向鎖と、かつ遺伝子の3'末端における逆方向鎖と相同である重複オリゴヌクレオチドを使用することによって両方向で漸次伸長される

40

50

。

【0687】

より大きい二本鎖DNA断片を合成する別の方法は、トップストランドPCR (TSP) によってより小さいオリゴヌクレオチドを組み合わせることである。本方法では、複数のオリゴヌクレオチドが所望の産物の全長にわたり、隣接するオリゴヌクレオチド（複数可）に対して重複領域を含有する。増幅は、ユニバーサル順方向プライマーおよび逆方向プライマーを用いて、かつ複数のサイクルの増幅を介して実施することができ、全長の二本鎖DNA産物が形成される。次いでこの産物は、任意選択の誤差補正およびさらなる増幅を受けることができ、それにより、所望の二本鎖DNA断片最終産物がもたらされる。

【0688】

TSPの一方法では、全長の所望の産物を形成するのに組み合わせられるより小さいオリゴヌクレオチドのセットは、40～200塩基の間の長さであり、互いに少なくとも約15～20塩基重複する。実用的な目的のために、重複領域は、オリゴヌクレオチドの特異的アニーリングを保証するために最低でも十分長く、使用される反応温度でアニールするために十分高い融解温度 (T_m) を有すべきである。重複は、所与のオリゴヌクレオチドが隣接するオリゴヌクレオチドによって完全に重複されるポイントまで伸長することができる。重複の量は、最終産物の品質に対していずれの効果も有さないようである。アセンブリー内の最初および最後のオリゴヌクレオチドビルディングブロックは、順方向および逆方向の増幅プライマーのための結合部位を含有すべきである。一実施形態では、最初および最後のオリゴヌクレオチドの末端側終端配列は、同じ相補性を有する配列を含有してユニバーサルプライマーの使用を可能にする。

【0689】

特注プラスミドアセンブリ／クローニング

一部の実施形態では、本開示は、宿主生物のゲノム内に所望の標的DNA切片（例えば、特定のSNPを含有する）を挿入することができるベクターを構築するための方法を教示する。一部の実施形態では、本開示は、標的DNA、相同アーム、および少なくとも1つの選択マーカーを含むベクターをクローニングする方法を教示する（図3を参照）。

【0690】

一部の実施形態では、本開示は、宿主生物内への形質転換に適した任意のベクターに適合する。一部の実施形態では、本開示は、宿主細胞に適合するシャトルベクターの使用を教示する。一実施形態では、本明細書に提供される方法において使用するためのシャトルベクターは、*E. coli* および／または *Corynebacterium* 宿主細胞に適合するシャトルベクターである。本明細書に提供される方法において使用するためのシャトルベクターは、本明細書に記載の選択および／または対抗選択のためのマーカーを含み得る。マーカーは、当技術分野で公知のかつ／または本明細書に提供される任意のマーカーであり得る。シャトルベクターは、任意の調節配列（複数可）および／または当技術分野で公知の前記シャトルベクターのアセンブリーにおいて有用な配列をさらに含み得る。シャトルベクターは、例えば、*E. coli* または *C. glutamicum* など本明細書に提供される宿主細胞内での繁殖に必要とされ得る任意の複製起点をさらに含むことができる。調節配列は、当技術分野で公知の、または本明細書に提供される任意の調節配列、例えば、宿主細胞の遺伝機構によって使用される、プロモーター、開始、停止、シグナル、分泌、および／または終結配列などであり得る。ある特定の事例では、標的DNAは、任意のリポジトリまたはカタログ製品から入手可能なベクター、コンストラクト、またはプラスミド、例えば、市販のベクター（例えば、DNA 2.0 特注または GATEWAY（登録商標）ベクターを参照）内に挿入することができる。ある特定の事例では、標的DNAは、任意のリポジトリまたはカタログ製品から入手可能なベクター、コンストラクト、またはプラスミド、例えば、市販のベクター（例えば、DNA 2.0 特注または GATEWAY（登録商標）ベクターを参照）内に挿入することができる。

【0691】

一部の実施形態では、本開示のアセンブリー／クローニング方法は、以下のアセンブリ

ーストラテジー：i) I I 型慣習的クローニング、ii) I I S 型媒介または「ゴールデンゲート」クローニング（例えば、Engler, C., R. Kandzia、およびS. Marillonnet、2008年、「A one pot, one step, precision cloning method with high-throughput capability」、Plos One、3巻：e3647頁；Kotera, I.およびT. Nagai、2008年、「A high-throughput and single-tube recombination of crude PCR products using a DNA polymerase inhibitor and type IIS restriction enzyme」、J Biotechnol、137巻：1～7頁；Weber, E., R. Gruetzner, S. Werner、C. Engler、およびS. Marillonnet、2011年、Assembly of Designer TAL Effectors by Golden Gate Cloning、PloS One、6巻：e19722頁を参照）、iii) G A T E W A Y（登録商標）組換え、iv) T O P O（登録商標）クローニング、エキソヌクレアーゼ媒介アセンブリー（Aslanidisおよびde Jong、1990年、「Ligation-independent cloning of PCR products (LIC-PCR)」、Nucleic Acids Research、18巻、20号、6069頁）、v) 相同組換え、vi) 非相同末端結合、vii) ギブソンアセンブリー（Gibsonら、2009年、「Enzymatic assembly of DNA molecules up to several hundred kilobases」、Nature Methods、6巻、343～345頁）、またはこれらの組合せの少なくとも1つを使用することができる。モジュラーI I S 型ベースアセンブリーストラテジーは、P C T 公開WO 2011/154147に開示されており、その開示は、参照により本明細書に組み込まれている。

10

【0692】

一部の実施形態では、本開示は、少なくとも1つの選択マーカーを有するベクターをクローニングすることを教示する。選択圧下で原核細胞（例えば、アンピシリン、カナマイシン、テトラサイクリン、クロラムフェニコール、ゼオシン、スペクチノマイシン／ストレプトマイシンに対する）、または真核細胞（例えば、ジェネテシン、ネオマイシン、ハイグロマイシン、ピューロマイシン、ブラストサイジン、ゼオシン）内の選択について抗生物質耐性機能をコードすることが多い様々な選択マーカー遺伝子が当技術分野で公知である。他のマーカーシステム、例えば、順調に形質導入された宿主細胞において発現されるX-galまたは緑色もしくは赤色蛍光タンパク質などの蛍光レポーターの存在下で陽性クローンを選択するのに細菌において使用される周知のブルー／ホワイトスクリーニングシステムは、望まれるまたは望まれない細胞のスクリーニングおよび同定を可能にする。そのほとんどが原核生物系内でのみ機能的である選択マーカーの別のクラスは、産生細胞を死滅させる毒性遺伝子産物を発現する、「デス遺伝子 (death gene)」とも呼ばれることが多い対抗選択可能マーカー遺伝子に関する。このような遺伝子の例としては、sacB、rpsL (strA)、tetAR、pheS、thyA、gata-1、またはccdBが挙げられ、その機能は、(Reyratら、1998年、「Counterselectable Markers: Untapped Tools for Bacterial Genetics and Pathogenesis」、Infect Immun.、66巻(9号)：4011～4017頁)に記載されている。

20

30

【0693】

DNAベクターアセンブリー、増幅、およびゲノム編集

一部の実施形態では、本開示は、E. coliに特異的なHTPゲノム操作ステップを教示する。したがって、一部の実施形態では、本開示は、E. coliにおいてコンストラクトを構築し、増幅させる方法、およびE. coliを操作する方法を教示する。

40

【0694】

一部の実施形態では、本開示のDNAベクターは、i) 条件付き複製起点(R6K)、ii) 抗生物質耐性遺伝子、iii) 1つまたは複数の対抗選択遺伝子(複数可)（例えば、sacBおよび/またはPheS）およびiv) S. cerevisiaeの複製起点を含む。

【0695】

一部の実施形態では、本開示は、栄養要求性S. cerevisiaeにおいてDNAコンストラクトをアセンブルする方法を教示する。よって、一部の実施形態では、本開示のベクターは、S. cerevisiaeの複製起点を含む。これにより、アセンブリー

50

中に *S. cerevisiae* においてベクターが複製することが可能になる。

【0696】

一部の実施形態では、本開示は、*p i r* タンパク質を含有する *E. coli* に存在するアセンブルされた DNA を増殖させる方法を教示する。よって、一部の実施形態では、本開示のベクターは、R 6 K 複製起点を含む。一部の実施形態では、R 6 K 複製起点は、*p i r* タンパク質の存在を条件とする。すなわち、一部の実施形態では、R 6 K 複製起点を含む本開示のベクターは、*p i r* 遺伝子を含む宿主細胞においてのみ増幅する。これにより、研究者が本開示のベクターをベクターの構築および増幅ステップの間に増幅させると同時に、宿主細胞操作ステップの間のベクターの余分の染色体発現を防ぐことも可能になる。

10

【0697】

一部の実施形態では、本開示のベクターは、*P h e S* 遺伝子を含む。*E s c h e r i c h i a c o l i* フェニルアラニル-tRNA 合成酵素 (*P h e S*) は、その A 2 9 4 G バリエーションでは翻訳の間に 4-クロロフェニルアラニン (4 C P) が細胞タンパク質中に誤って組み入れられ、それにより、細胞死が引き起こされるので、対抗選択マーカーとして有用であり得る。一部の実施形態では、*P h e S* 遺伝子を、宿主細胞のゲノム中に一時的に組み入れられるように設計する。一部の実施形態では、本開示は、4 C P を伴う最小培地中で宿主細胞を増殖させるステップを含む対抗選択方法を教示する。それでもなおベクターを含む細胞は、4 C P をタンパク質中に組み入れ、死滅する。*P h e S* 配列がループアウトされた細胞は生存する。

20

【0698】

一部の実施形態では、本開示は、ベクターを構築し、アセンブルし、宿主細胞中に組み込む方法を教示する。一部の実施形態では、本開示は、相同アーム (*h o m L* および *h o m R*) を PCR によって増幅することを教示する。一部の実施形態では、所望の遺伝子変化 (*h o m L* と *h o m R* の間の黒の棒) は、*h o m L* の逆方向プライマーおよび *h o m R* の順方向プライマーに存在する (図 3 を参照されたい)。一部の実施形態では、*h o m L* の順方向プライマーおよび *h o m R* の逆方向プライマーは、骨格プラスミドとの配列同一性を有する。ループインおよびループアウトプロセスのさらなる図解を図 4 5 に示す。

【0699】

一部の実施形態では、本開示のベクターは、1 つまたは複数のインスレーター配列を含む。インスレーター配列は、当技術分野で公知の任意のインスレーター配列であり得る。一実施形態では、インスレーター核酸配列は、本明細書に提供されるインスレーター 1 配列 (配列番号 2 1 8)、インスレーター 2 配列 (配列番号 2 1 9) またはインスレーター 1 および 2 配列の両方である。一実施形態では、本開示のベクターは、相同アーム (*h o m L* および *h o m R*) に隣接するインスレーター配列を含む。一実施形態では、本開示のベクターは、相同アーム (*h o m L* および *h o m R*) に隣接するインスレーター配列ならびにターミネーター配列 (複数可) を含む。インスレーター配列は、制限エンドヌクレアーゼ配列を含まないように生成することができる。

30

【0700】

一部の実施形態では、本開示のベクターは、本明細書に提供されるエレメントの組合せを含む。一部のケースでは、本明細書に提供される方法において使用するためのベクターは、例えば、配列番号 2 1 4 の核酸配列を有するベクター 1 (図 5 5 を参照されたい) など、R 6 K 複製起点、*S a c B* 遺伝子、対抗選択マーカーとして *P h e S* 遺伝子および *U R A 3* 酵母栄養要求性マーカーを含み得る。一部のケースでは、本明細書に提供される方法において使用するためのベクターは、配列番号 2 1 5 の核酸配列を有する、例えばベクター 2 (図 5 6 を参照されたい) などの、ベクター 1 の変更バージョンであり得る。ベクター 1 に関して先に記載したエレメントに加えて、ベクター 2 は、表 1 5 に見出されるエレメントをさらに含み得る。本明細書に提供される方法において使用するための追加のベクターは、ベクター 3 (核酸配列番号 2 1 6 ; 図 5 7) およびベクター 4 (核酸配列番号 2 1 7 ; 図 5 8) を含む。ベクター 3 および 4 はどちらもベクター 2 バックグラウンドから作

40

50

り直した。しかし、ベクター3では、sacBのプロモーター配列がP2-MCD2プロモーターを含有するプロモーター（Mutalikら、Nat Methods.、2013年4月；10巻（4号）：354～60頁を参照されたい）およびT251A/A294G変異を含有する生来のpheS遺伝子のコドン最適化バージョン（Miyazaki, K. Biotechniques.、2015年2月1日；58巻（2号）：86～8頁を参照されたい）で置き換えられており、ベクター4は、ベクター3バックグラウンドを含み、URA3選択マーカーがTRP1マーカーで置き換えられている。

【表15】

表15.ベクター2の配列エレメントを選択する

エレメント名	部分配列
インスレーター1	TATTCACACGCAATCAACAGGCAGGATAATCGCTGGTAAGGTCAGTGC TTTCTTCAGGTAGTAGAGATAACAATAGTTCCCAACGATAGGTGGCAGA TTTCACTTTACAGACCGACTGGTTCAGAAGCGTAGATAATAGCCCGTGT TTTCCAATAAGGGATAGTGTAGGTAAGTCAACTCCTCCGTCAGAGCCA ACCGTTT (配列番号 218)
インスレーター2	GCACTAGGACTTGCCGCGGATACTGCCCCATTACATGAATTGCAGCCT CAGGGACGTCAGTAGATCATGGAGGTAGGGCATATGTCCTCTGTTGTT AAAATGTGAGTTCTCAACGAAGCACGAATCGGTCAGAACCTACACTAA GGAGATTTGGTAGGTGCACGGTTTCTGTCGCATAGACCAGTTCATTTCA GATGTCT (配列番号 219)
T1	GGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTT TTATCTGTTGTTTGTCTGGTGAACGCTCTCCTGAGTAGGACAAATCCGCC GCCCTAGA (配列番号 220)
B0015	CCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTC GTTTTATCTGTTGTTTGTCTGGTGAACGCTCTCTACTAGAGTCACACTGG CTCACCTTCGGGTGGGCCTTCTGCGTTTATA (配列番号 221)
SacB プロモーター	TTGACAATTAATCATCCGGCTCGTAATTTATGTGGATCTTAATCATGCT AAGGAGGTTTTCTAATG (配列番号 222)
PheS プロモーター & 配列	GGCGGTGTTGACATAAATACCACTGGCGGTGATACTGAGCACATCAGC AGGTCACACAGGAAAGTACTAGATGTCTGCATCTTGCAGAATTAGTAGC TTCAGCGAAGGCCGCGATTCTCAGGCGAGTGACGTCGCAGCACTGGA TAATGTACGTGTTGAGTACCTGGGAAAGAAGGGACACCTTACTCTTCA AATGACAACCCTGCGCGAACTGCCGCCGAGGAACGCCCCGCAGCAG GAGCGGTAATCAATGAGGCAAAGGAGCAAGTACAACAGGCACTGAAC GCCCCGTAAGGCTGAGTTGGAATCCGCCGCATTAAACGCGCGCCTTGCT GCGGAAACCATTGATGTCTCGCTGCCCGGGCGCCGCATTGAGAATGGA GGCTTACACCCAGTGACTCGTACCATCGACCGTATCGAATCTTTCTTTG GCGAACTTGCTTCACTGTGGCAACTGGACCGGAGATTGAGGACGACT ACCACAATTTTCGATGCCTTGAACATTCCCGGTCATCATCCTGCACGCGC CGATCATGATACATTCTGGTTTGATAACACCCGTTTGCTTCGTACCCAG ACAAGCGGTGTCCAAATCCGTACGATGAAGGCTCAGCAACCACCGATC CGTATCATTGCTCCAGGGCGCGTGTACCGTAACGATTATGACCAGACA CATAACCCGATGTTTCACCAAATGGAAGGGTTGATTGTGGATACGAAT ATCTCTTTACGAATCTGAAGGGCACCTTACATGATTTCTTACGCAACT TTTTTCGAGGAGGACCTTCAAATTCGCTTTTCGTCCATCGTACTTCCCTTT GCAGAACCTTCGGCTGAAGTGGATGTAATGGGGAAAAACGGTAAGTG GCTGGAGGTTTTAGGTTGCGGGATGGTTCATCCAAATGTGCTTCGCAAC GTCGGCATCGACCCCGAAGTCTACAGTGGATTCCGATTCCGGATGGGA ATGGAACGTCTGACTATGCTTCGTTACGGCGTAACGGATTTGCGCTCCT TTTTTGAGAACGATCTTCGTTTTCTGAAGCAATTCAAATAA (配列番号 223)

10

20

30

40

50

【0701】

本開示の一実施形態によるDNAアセンブリーに関連する一般的なワークフローを図35に示す。

【0702】

一部の実施形態では、骨格および挿入断片(homLおよびhomR)の間のDNAアセンブリーは、酵母において酵母ギャップ修復組換えによって実施される。一部の実施形態では、最小培地中でのアセンブルされたDNAの選択のために、栄養要求性マーカー(TRPまたはURA)を骨格プラスミド中に存在させる。次いで、アセンブルされたプラスミドを酵母培養物から抽出する。

【0703】

一部の実施形態では、次に、抽出されたプラスミドを、pir遺伝子を含有するE. coli株中に形質転換して、以下の目的の株の形質転換のための所望のプラスミドを増殖させる。形質転換体を所与の抗生物質に対する耐性について選択する。形質転換体をシーケンシングのために選び取って、正確にアセンブルされた目的のプラスミドを選択する。

【0704】

一部の実施形態では、正確にアセンブルされたプラスミドを、ゲノム操作のために電気穿孔によって宿主細胞中に形質転換する。標的宿主細胞はpirタンパク質を含有しないので、抗生物質選択培地上に形成されたコロニーは、ゲノム内の所望の遺伝子座におけるプラスミドの組み込みを有することが予測される。それぞれhomLおよびhomRの外側のプライマー、およびプラスミドの内側に結合するプライマーを用いたPCRによってプラスミドの正確な組み込みを検証することができる。

【0705】

一部の実施形態では、本開示は、プラスミドの骨格を宿主のゲノムから除去するループアウト方法を教示する。よって、一部の実施形態では、本開示は、対抗選択培地(スクロース)および/または4CPにおける後続の選択を使用して、プラスミドの骨格部分を含有しないクローンを単離できることを教示する。したがって、一部の実施形態では、正確な組み込み体を含む単離された宿主細胞をLB培地中に播種し、培養物をスクロースおよび4-p-クロロフェニルアラニンを含有するLB寒天プレート(LB+suc+4CP)上に蒔いた。sacB遺伝子を発現する細胞のスクロースに対する感度、およびPhesを発現する細胞の4CPに対する感度に起因して、LB+suc+4CP寒天プレート上に形成されるコロニーは、目的の遺伝子の変異体または野生型のいずれかを有することが予測される。標的ヌクレオチド(複数可)のPCR増幅およびPCR産物のシーケンシングにより、所望のゲノム改変を有する新しいクローンを単離することが可能になる。

【0706】

一部の実施形態では、結果として生じるクローンをシーケンシングして、所望のヌクレオチドの変化(複数可)を有するクローンを見つけ出す。一部の実施形態では、上記のすべてのプロセスを液体ハンドラーで実施することができる。

【0707】

宿主細胞の形質転換

一部の実施形態では、本開示のベクターは、形質転換、トランスフェクション、形質導入、ウイルス感染、遺伝子銃、またはTi媒介遺伝子移入を含めた様々な技法のいずれかを使用して宿主細胞内に導入することができる(Christie, P.J.およびGordon, J.E., 2014年、「The Agrobacterium Ti Plasmids」、Microbiol Spectr., 2014年; 2巻(6号); 10.1128頁を参照)。特定の方法としては、リン酸カルシウムトランスフェクション、DEAE-デキストラン媒介トランスフェクション、リポフェクション、または電気穿孔が挙げられる(Davis, L., Dibner, M., Battey, I., 1986年、「Basic Methods in Molecular Biology」)。形質転換の他の方法としては、例えば、酢酸リチウム形質転換および電気穿孔が挙げられる。例えば、Gietzら、Nucleic Acids Res., 27巻: 69~74頁(1992年); Itoら、J. Bacteriol., 153巻: 163~168頁(1983年); ならびにBeckerおよびGuarente、Methods in Enz

10

20

30

40

50

ymology、194巻：182～187頁（1991年）を参照。一部の実施形態では、形質転換宿主細胞は、組換え宿主株と呼ばれる。

【0708】

一部の実施形態では、本開示は、本開示の96-ウェルプレートロボットプラットフォームおよび液体ハンドリング機械を使用する細胞のハイスループット形質転換を教示する。

【0709】

一部の実施形態では、本開示は、上述した1種または複数の選択マーカーを有する形質転換細胞をスクリーニングすることを教示する。このような一実施形態では、カナマイシン耐性マーカー（Kan^R）を含むベクターで形質転換された細胞を、有効量のカナマイシン抗生物質を含有する培地に蒔く。カナマイシンが入った培地上で目に見えるコロニー形成単位は、ベクターカセットをこれらのゲノム内に組み入れたと推定される。所望の配列の挿入は、関連した挿入部位のPCR、制限酵素分析、および／またはシーケンシングを介して確認することができる。

【0710】

選択された配列のループアウト

一部の実施形態では、本開示は、宿主生物からDNAの選択された領域をループアウトする方法を教示する。ループアウト法は、Nakashimaら、2014年、「Bacterial Cellular Engineering by Genome Editing and Gene Silencing」、Int. J. Mol.

Sci.、15巻（2号）、2773～2793頁に記載された通りのものであり得る。一部の実施形態では、本開示は、陽性形質転換体から選択マーカーをループアウトすることを教示する。ループアウト欠失技法は、当技術分野で公知であり、（Tearら、2014年、「Excision of Unstable Artificial Gene-Specific inverted Repeats Mediates Scar-Free Gene Deletions in Escherichia coli.」、Appl. Biochem. Biotech.、175巻：1858～1867頁）に記載されている。本明細書に提供される方法で使用されるループアウト法は、シングルクロスオーバー相同組換えまたはダブルクロスオーバー相同組換えを使用して実施することができる。一実施形態では、本明細書に記載の選択された領域のループアウトは、本明細書に記載のシングルクロスオーバー相同組換えを使用することを必然的に伴い得る。

【0711】

第1に、ループアウトベクター（loop out vector）が、宿主生物のゲノム内の選択された標的領域内に挿入される（例えば、相同組換え、CRISPR、ラムダレッド媒介組換え、または他の遺伝子編集技法を介して）。一実施形態では、シングルクロスオーバー相同組換えが、図3に表したもののなどの環状プラスミドまたはベクターをループインするために、環状プラスミドまたはベクターと宿主細胞ゲノムとの間に使用される。挿入されるベクターは、既存のまたは導入される近傍の宿主配列のダイレクトリピートである配列を用いて設計することができ、その結果、ダイレクトリピートがループングおよび欠失を予定しているDNAの領域に隣接する。挿入されると、ループアウトプラスミドまたはベクターを含有する細胞を、選択領域の欠失について対抗選択することができる（例えば、図4；選択遺伝子に対する耐性の欠如を参照）。ループインプロセスおよびループアウトプロセスのさらなる例示は、図45に記載される。

【0712】

当業者は、ループアウト手順の記載は、ゲノムから望まれない領域を欠失させるための一例示的方法を代表するにすぎないことを認識する。実際に、本開示の方法は、これらに限定されないが、ラムダレッド、CRISPR、TALENs、FOK、または他のエンドヌクレアーゼを介した遺伝子編集を含めて、ゲノム欠失の任意の方法に適合する。当業者は、相同組換え技法を介してゲノムの望まれない領域を置き換える能力も認識する。

【0713】

ラムダレッド媒介性遺伝子編集

本明細書に提供されるように、本明細書に記載の遺伝子編集を、その内容全体が参照に

10

20

30

40

50

よって本明細書に組み込まれるDatsenkoおよびWanner、PNAS USA、97巻：6640～6645頁（2000年）に記載のラムダレッド媒介性相同組換えを使用して実施することができる。

【0714】

ラムダレッドシステムは、ラムダレッドバクテリオファージに由来し、その遺伝子操作ツールとしての使用は、相同組換え媒介性遺伝子操作の略であるリコンビニアリングと称することができる。ラムダレッドシステムを使用して、改変の組合せ：選択可能および非選択可能配列の挿入および欠失、点変異または他の小さな塩基対変化、およびタンパク質タグの付加をなすことができる。ラムダレッドシステムはまた、E. coli 染色体、プラスミドDNAまたはBAC DNAを改変する柔軟性も有する。ラムダレッドリコンビニアリングシステムを使用して標的DNAを改変するために、直鎖ドナーDNA基質（dsDNAまたはssDNAのいずれか—以下を参照されたい）を、ラムダレッド酵素を発現するE. coli中に電気穿孔することができる。次いで、これらの酵素により、標的DNA配列を有する基質の相同組換えが触媒される。これは、遺伝子変化が試験管内で起こる制限酵素クロニングと比較して、クロニングがin vivoにおいて起こることを意味する。ドナーDNA基質には、組換えの標的部位との約50ヌクレオチドの相同性のみが要求される。

【0715】

ラムダレッドリコンビニアリングシステムは、3つのコンポーネント：1) Exo、2) Beta、および3) Gamを有する。dsDNA基質を用いたリコンビニアリングのためには3つすべてが必要であるが、ssDNA基質を用いて改変を生成するときにはBetaのみが必要になる。

【0716】

Gam：Gamは、内在性RecBCDおよびSbcCDヌクレアーゼの両方が、E. coli 宿主細胞に導入された直鎖DNAを消化することを妨げる。

【0717】

Exo：Exoは、5'→3' dsDNA依存性エキソヌクレアーゼである。Exoは、直鎖dsDNAを5'末端から開始して分解し、2つの可能性のある産物：1) 一本鎖3'突出部を有する部分的dsDNA二重鎖または2) dsDNAが十分に短い場合、相補鎖全体が分解されたssDNAを生成することができる。

【0718】

Beta：Betaは、Exoによって創製されたssDNAを保護し、該ssDNAの、細胞の相補的ssDNA標的とのアニーリングを促進することができる。ssDNAオリゴ基質を用いたリコンビニアリングのためにはBeta発現のみが必要である。

【0719】

本発明における使用に関しては、ラムダレッドリコンビニアリング方法は、基質DNAを設計し、生成するステップと；ラムダレッド組換え遺伝子を発現させるステップと；基質DNAを形質転換（例えば、電気穿孔）するステップと；形質転換体を増殖させるステップと；組換えクローンを選択し、確認するステップとを伴い得る。

【0720】

基質DNAの設計および生成

直鎖dsDNAを使用するかまたはssDNA基質を使用するかは、実験の目的に依存し得る。dsDNA基質は、およそ20ヌクレオチドを超える挿入または欠失のために最良であり得、一方ssDNA基質は、ほんの数塩基対の点変異または変化のために最良であり得る。

【0721】

dsDNA基質

dsDNA挿入断片は、目的のDNA配列を増幅させ、目的のDNA配列に、標的にされる挿入断片部位と50塩基対の相同性で隣接するプライマーを使用したPCRによって作製することができる。プライマーの長さは約70ヌクレオチドであり得る（目的のDN

10

20

30

40

50

A配列にアニールする20ヌクレオチドおよび標的部に隣接する領域と相同な50ヌクレオチド)。dsDNA挿入断片は、抗生物質耐性遺伝子などの選択可能DNA断片、ならびに遺伝子置換えおよびタグなどの非選択可能DNA断片を含めた大きな挿入または欠失を含み得る。

【0722】

ssDNA基質

ssDNA基質は、合成オリゴヌクレオチドまたは短いPCR産物であり得る。どちらにしても、基質の長さは、約70～100ヌクレオチドであるべきであり、所望の変更（複数可）が配列の中央に位置する。ラムダレッドは、高い組換え頻度を有するので、DNAのラギング鎖を標的にするときには、目的の標的領域を通る複製の方向を決定し、ラギング鎖と相補的なオリゴを設計することが最良である。一部の場合では、両方の鎖を標的にするオリゴを設計する。2つのオリゴのうち一方が他方よりも高い効率で組み換えられ、これがラギング鎖の同定の助けになり得る。

【0723】

ssDNA基質はdsDNAよりも効率的であり得、組換え頻度は0.1%～1%であり、メチル指向ミスマッチ修復（MMR）システムの活性化を回避するオリゴを設計することによって25～50%までの高さに上昇し得る。MMRの役目は、DNA複製の間に生じるDNAミスマッチを補正することである。MMRの活性化は、1）ノックアウトされる重要なMMRタンパク質を有する細菌の株を使用すること、または2）MMRを回避するためにssDNAオリゴを特別に設計することによって回避することができる：1）不活化MMRを有するE. coli：不活性MMRを有するE. coliの使用は、2つの選択肢のうち明確に容易な方であるが、これらの細胞は、変異を起こしやすく、それらのゲノムに対する意図されない変化がより多くなり得る。2）MMR活性化を回避するssDNAオリゴの設計：一実施形態では、編集部位の6塩基対または6塩基対以内にC/Cミスマッチを導入する。別の実施形態では、所望の変化はゆらぎコドン内の4～5のサイレント変化に隣接する、すなわち、ヌクレオチド配列は変化させるが翻訳されるタンパク質のアミノ酸配列は変化させない隣接する4～5コドンの第3の塩基対を変化させる。これらの変化は、所望の変化の5'側または3'側であり得る。

【0724】

ラムダレッド組換え遺伝子の発現

ラムダレッドリコンビニアリングシステムは、宿主細胞において、1）欠損プロファージが組み込まれた細菌から、2）プラスミドから、3）ミニλから、または4）ラムダレッドファージ自体から発現させることができる。Redタンパク質の発現を制御することは、Gam発現の毒作用を最小化するため、およびRedが構成的に発現されたときに生じる自然変異を限定するために極めて重要である。いずれの組換え系を使用するかは、編集したいDNAのタイプに依存するが、BAC DNAは以下に記載の手法のいずれかを用いて改変することができる。

【0725】

欠損プロファージが組み込まれた細菌株：

【0726】

欠損ラムダレッドファージの組込みに起因してラムダレッドリコンビニアリング遺伝子を安定に発現するいくつかのE. coli株が存在する。1つのこのような株はDY380であり、これは、DH10B E. coli株に由来する。リコンビニアリングのために一般に使用されるいくつかの他の細菌株は、Thomasonら（Recombineering: Genetic Engineering in Bacteria Using Homologous Recombination. Current Protocols in Molecular Biology.、106：V：1.16：1.16.1～1.16.39）およびSharanら（Recombineering: A Homologous Recombination-Based Method of Genetic Engineering. Nature protocols.、2009年；4巻（2号）：206～223頁）に見出すことができる。

【0727】

これらの株の一部では、*exo*、*beta*および*gam*の発現は、内在性ファージプロモーター*pL*およびリプレッサー*CI*によって緊密に調節される。リコンビニアリング目的で、温度感受性バージョンのリプレッサー遺伝子、*CI857*を使用する。この変異体リプレッサーは、低温（30～34℃）で組換え遺伝子の発現を妨げる。細菌を42℃にシフトして15分置くことにより、リプレッサーが迅速に不活化し、組換え遺伝子の発現が可能になる。この後、温度を低下させることにより、リプレッサーが再生し、再度*exo*、*beta*および*gam*の発現を抑止する。ラムダレッド発現にこの方法を使用することの主要な利点は、この方法では、リコンビニアリングシステムの発現を維持するための抗生物質による選択が必要ないことである。この構成を使用して、染色体遺伝子を改変することもできる。最初の編集事象後、欠損プロファージを第2のラムダレッド組換え事象によって宿主*E. coli*の染色体から除去することができる。あるいは、改変された対立遺伝子が選択可能である場合、*P1*形質導入によって異なる遺伝的バックグラウンドに移すことができる。

10

【0728】

プラスミド：

【0729】

プラスミドからラムダレッド遺伝子を発現させることにより、可動リコンビニアリングシステムが可能になるが、実験の成功のためには発現の緊密な調節が必要になる。*Red*の発現を制御するために一般に使用されるプロモーターとしては、*IPTG*誘導性*lac*プロモーター、アラビノース誘導性*pBAD*プロモーターおよび内在性ファージ*pL*プロモーターが挙げられる。一部の場合では、*Red*システムの漏出性発現を限定するために、これらのプロモーターに関連するリプレッサー（*lacI*、*araC*、*CI857*）も発現するプラスミドを使用することができる。組換えクロンが生成されたらリコンビニアリングシステムを除去することが容易であるので、ラムダリコンビニアリングシステムを発現させるためのプラスミドの使用を、細菌染色体DNAを編集するために使用することができる。これを行うための単純なやり方は、熱感受性複製起点を有するプラスミドからラムダレッド遺伝子を発現させることである。リコンビニアリングシステムがなくなったら、細菌を42℃で増殖させることによって細菌を「回復」させることができる。

20

【0730】

ミニーλ：

30

【0731】

プラスミドの使用と欠損プロファージの安定な組込みのハイブリッドは、細菌に導入されるとゲノム中に組み込まれるファージDNAの欠損性、非複製性の環状片であるミニーλを使用することである。ミニーλは、内在性赤色プロモーター*pL*および*CI857*リプレッサーを使用して発現を調節する。抗生物質を使用して陽性クロンを選択することができるが、ミニーλは安定に組み込まれるので、維持のための薬物選択は必要ない。42℃への温度シフトにより、リコンビニアリングに必要な*red*遺伝子の活性化が可能になるだけでなく、宿主の染色体からのミニーλの切り出しに参与する*int*および*xis*遺伝子の発現ももたらされる。この後、ミニーλをプラスミドと同様に細菌から容易に精製することができる。

40

【0732】

ファージ：

【0733】

*Red*システムを発現させるための別の選択肢は、テトラサイクリン耐性遺伝子およびラムダレッドリプレッサー*CI857*を有するラムダレッドファージ、*λTetR*を使用することであり得る。導入されたら、プロファージは安定であり、もはや薬物選択は必要ない。この手法の1つの欠点は、一般的な分子生物学技法ではないバクテリオファージの生成が必要になることである。しかし、この方法の利点は、*Red*システムを目的の*yoa*株に安定に組み込むことができ、必要であれば、*P1*形質導入を使用して改変を異なるバックグラウンドに移すことができることである。この手法は、*E. coli*ゲノムへの

50

ファージの安定な組込みをもたらすので、プラスミドまたはBACの改変に最も良く適する。

【0734】

組換えクローンの選択および確認

抗生物質耐性遺伝子が挿入されたら、組換え体をまず抗生物質耐性によって選択することができるが、すべてのクローンをさらに試験して、所望の改変の存在を確認すべきである。ほとんどの場合、陽性クローンをスクリーニングするためにコロニーPCRを使用することができ、プラスミドを適切な変異についてスクリーニングするために制限酵素消化を使用することができる。点変異および他のわずかな変化はシーケンシングによって確認することができ、シーケンシングは、どのタイプのDNA：E. coli 染色体、プラスミド、またはBACが改変の標的にされるかにかかわらずすべてのクローンについて確認するためにも使用することができる。

10

【0735】

CRISPR媒介性遺伝子編集

本明細書に提供される一態様では、宿主細胞のゲノムをCRISPRによって改変することができる。E. coliにおける遺伝子編集のためにCRISPR/Cas9システムを利用することに関する例示的な実施形態は、実施例18および19に見出すことができる。

【0736】

CRISPR/Casシステムは、プラスミドおよびファージ内に存在するものなどの外来遺伝子エレメントに耐性を付与し、また、獲得免疫の一形態をもたらす原核生物免疫系である。CRISPRは、クラスター化された規則的な配置の短い回文配列リピート (Clustered Regularly Interspaced Short Palindromic Repeat) を表し、casは、CRISPR関連システムを表し、CRISPR複合体に関連する小さなcas遺伝子を指す。

20

【0737】

CRISPR-Casシステムは、クラス1またはクラス2システムのいずれかとして最も広範に特徴付けられている。これらの2つのシステムを区別する主要な特徴は、Cas-エフェクターモジュールの特質である。クラス1システムでは、干渉を媒介するために多数のCasタンパク質を複合体 (「カスケード複合体」と呼ばれる) にアセンブリする必要があるが、クラス2システムでは、大きな単一のCas酵素を使用して干渉を媒介する。クラス1およびクラス2システムはそれぞれ、特定のCasタンパク質の存在に基づいて多数のCRISPR-Casタイプにさらに分類される。例えば、クラス1システムは、以下の3つのタイプに分類される：Cas3タンパク質を含有するタイプIシステム；Cas10タンパク質を含有するタイプIIIシステム；および、Cas8様タンパク質であるCsf1タンパク質を含有する推定タイプIVシステム。クラス2システムは、概して、クラス1システムより一般的でなく、以下の3つのタイプにさらに分類される：Cas9タンパク質を含有するタイプIIシステム；Cas12aタンパク質 (以前はCpf1として公知であり、本明細書ではCpf1と呼ばれる)、Cas12b (以前はC2c1として公知であった)、Cas12c (以前はC2c3として公知であった)、Cas12d (以前はCasYとして公知であった)、およびCas12e (以前はCasXとして公知であった) を含有するタイプVシステム；ならびにCas13a (以前はC2c2として公知であった)、Cas13b、およびCas13cを含有するタイプVIシステム。Pyzochaら、ACS Chemical Biology、13巻 (2号)、347~356頁。一実施形態では、本明細書に提供される方法において使用するためのCRISPR-Casシステムは、クラス2システムである。一実施形態では、本明細書に提供される方法において使用するためのCRISPR-Casシステムは、タイプII、タイプVまたはタイプVIクラス2システムである。一実施形態では、本明細書に提供される方法において使用するためのCRISPR-Casシステムは、Cas9、Cas12a、Cas12b、Cas12c、Cas12d、Cas12e、Cas13a、Cas13b、C

30

40

50

a s 1 3 c またはそのホモログ、オルソログもしくはパラログから選択される。

【0738】

本明細書に開示される方法において使用される C R I S P R システムは、本明細書では C a s エフェクタータンパク質と呼ばれる1つまたは複数の核酸ガイド型 C R I S P R 関連 (C a s) ヌクレアーゼを含む C a s エフェクターモジュールを含む。一部の実施形態では、C a s タンパク質は、1つまたは多数のヌクレアーゼドメインを含み得る。C a s エフェクタータンパク質は、一本鎖または二本鎖核酸分子 (例えば、DNA または RNA 核酸) を標的とすることができ、二本鎖または一本鎖切断を生成することができる。一部の実施形態では、C a s エフェクタータンパク質は、野生型または天然に存在する C a s タンパク質である。一部の実施形態では、C a s エフェクタータンパク質は、変異体 C a s タンパク質であり、これは、親 C a s タンパク質と比較して1つまたは複数の変更された特性を有する C a s タンパク質が生じるように、WT または天然に存在する C a s タンパク質 (例えば、親 C a s タンパク質) に1つまたは複数の変異、挿入、または欠失がなされたものである。

10

【0739】

一部の例では、C a s タンパク質は、野生型 (W T) ヌクレアーゼである。本開示において使用するための適当な C a s タンパク質の限定されない例としては、C 2 c 1、C 2 c 2、C 2 c 3、C a s 1、C a s 1 B、C a s 2、C a s 3、C a s 4、C a s 5、C a s 6、C a s 7、C a s 8、C a s 9 (C s n 1 および C s x 1 2 としても公知)、C a s 1 0、C p f 1、C s y 1、C s y 2、C s y 3、C s e 1、C s e 2、C s c 1、C s c 2、C s a 5、C s n 2、C s m 1、C s m 2、C s m 3、C s m 4、C s m 5、C s m 6、C m r 1、C m r 3、C m r 4、C m r 5、C m r 6、C s b 1、C s b 2、C s b 3、C s x 1 7、C s x 1 4、C s x 1 0 0、C s x 1 6、C s a X、C s x 3、C s x 1、C s x 1 5、C s f 1、C s f 2、C s f 3、C s f 4、M A D 1-2 0、S m C s m 1、そのホモログ、そのオルソログ、そのバリエーション、その変異体、またはその改変されたバージョンが挙げられる。適当な核酸ガイド型ヌクレアーゼ (例えば、C a s 9) は、これらに限定されないが、*Thiomicrospira*、*Succinivibrio*、*Candidatus*、*Porphyromonas*、*Acidomonococcus*、*Prevotella*、*Smithella*、*Moraxella*、*Synergistes*、*Francisella*、*Leptospira*、*Catenibacterium*、*Kandleria*、*Clostridium*、*Dorea*、*Coprococcus*、*Enterococcus*、*Fructobacillus*、*Weissella*、*Pediococcus*、*Corynebacter*、*Sutterella*、*Legionella*、*Treponema*、*Roseburia*、*Filifactor*、*Eubacterium*、*Streptococcus*、*Lactobacillus*、*Mycoplasma*、*Bacteroides*、*Flaviivola*、*Flavobacterium*、*Sphaerochaeta*、*Azospirillum*、*Gluconacetobacter*、*Neisseria*、*Roseburia*、*Parvibaculum*、*Staphylococcus*、*Nitratifractor*、*Mycoplasma*、*Alicyclobacillus*、*Brevibacillus*、*Bacillus*、*Bacteroidetes*、*Brevibacillus*、*Carnobacterium*、*Clostridiaridium*、*Clostridium*、*Desulfonatronum*、*Desulfovibrio*、*Helcococcus*、*Leptotrichia*、*Listeria*、*Methanomethyophilus*、*Methylobacterium*、*Opitutaceae*、*Paludibacter*、*Rhodobacter*、*Sphaerochaeta*、*Tuberibacillus*、および *Campylobacter* を含む属の生物体由来し得る。このような属の生物体の種は、他で本明細書において論じられる通りであり得る。

20

30

40

【0740】

適当な核酸ガイド型ヌクレアーゼ (例えば、C a s 9) は、これらに限定されないが、

50

Firmicute、Actinobacteria、Bacteroidetes、Proteobacteria、Spirochaetes、およびTenericutesを含む門の生物体に由来し得る。適当な核酸ガイド型ヌクレアーゼは、これらに限定されないが、Erysipelotrichia、Clostridia、Bacilli、Actinobacteria、Bacteroidetes、Flavobacteri-
 10 a、Alphaproteobacteria、Betaproteobacteri-
 a、Gammaproteobacteria、Deltaproteobacteri-
 a、Epsilonproteobacteria、Spirochaetes、およびMollicutesを含む綱の生物体に由来し得る。適当な核酸ガイド型ヌクレアーゼは、これらに限定されないが、Clostridiales、Lactobacilla-
 20 lles、Actinomycetales、Bacteroidales、Flavobacteriales、Rhizobiales、Rhodospirillales、Burkholderiales、Neisseriales、Legionellales、Nautiliales、Campylobacterales、Spirochaetales、Mycoplasmatales、およびThiotrichalesを含む目の生物体に由来し得る。適当な核酸ガイド型ヌクレアーゼは、これらに限定されないが、Lachnospiraceae、Enterococcaceae、Leuc-
 onostocaceae、Lactobacillaceae、Streptococ-
 caceae、Peptostreptococcaceae、Staphylococ-
 caceae、Eubacteriaceae、Corynebacterineae、
 20 Bacteroidaceae、Flavobacterium、Cryomorph-
 aceae、Rhodobiaceae、Rhodospirillaceae、Ace-
 tobacteraceae、Sutterellaceae、Neisseriace-
 ae、Legionellaceae、Nautiliaceae、Campyloba-
 cteraceae、Spirochaetaceae、Mycoplasmatace-
 ae、およびFrancisellaceaeを含む科の生物体に由来し得る。

【0741】

本開示の方法、システム、および組成物における使用に適した他の核酸ガイド型ヌクレ-
 アーゼ（例えば、Cas9）としては、例えば、これらに限定されないが、Thiomicrospira sp. XS5、Eubacterium rectale、Succinivibrio dextrinosolvens、Candidatus Methanoplasma termitum、Candidatus Methanome-
 30 thylophilus alvus、Porphyromonas crevioricanis、Flavobacterium branchiophilum、Acidomonococcus sp.、Lachnospiraceae bacterium COE1、Prevotella brevis ATCC 19188、Smit-
 hellia sp. SCADC、Moraxella bovoculi、Synergistes jonesii、Bacteroidetes oral taxon 274、Francisella tularensis、Leptospira in-
 40 dai serovar Lyme str. 10、Acidomonococcus sp. crystal structure (5B43) S. mutans、S. agalactiae、S. equisimilis、S. sanguinis、S. pneumoniae; C. jejuni、C. coli; N. salmuginis、N. ter-
 garcus; S. auricularis、S. carnosus; N. meningitides、N. gonorrhoeae; L. monocytogenes、L. ivanovii; C. botulinum、C. difficile、C. tetani、
 C. sordellii; Francisella tularensis 1、Prevotella albensis、Lachnospiraceae bacterium MC2017 1、Butyrivibrio proteoclasticus、Peregrinibacteria bacterium GW2011_GWA
 50

2_33_10、Parcubacteria bacterium GW2011_GWC2_44_17、Smithella sp. SCADC、Microgenomates、Acidaminococcus sp. BV3L6、Lachnospiraceae bacterium MA2020、Candidatus Methanoplasma termitum、Eubacterium eligens、Moraxella bovoculi 237、Leptospira inadai、Lachnospiraceae bacterium ND2006、Porphyromonas crevioricanis 3、Prevotella disiens、Porphyromonas macacae、Catenibacterium sp. CAG:290、Kandleria vitulina、Clostridiales bacterium KA00274、Lachnospiraceae bacterium 3-2、Dorea longicatena、Coproccoccus catus GD/7、Enterococcus columbae DSM 7374、Fructobacillus sp. EFB-N1、Weissella halotolerans、Pediococcus acidilactici、Lactobacillus curvatus、Streptococcus pyogenes、Lactobacillus versmoldensis、および Filifactor alocis ATCC 35896 などの生物体に由来するものが挙げられる。それぞれが参照により本明細書に組み込まれている米国特許第8,697,359号；同第8,771,945号；同第8,795,965号；同第8,865,406号；同第8,871,445号；同第8,889,356号；同第8,895,308号；同第8,906,616号；同第8,932,814号；同第8,945,839号；同第8,993,233号；同第8,999,641号；同第9,822,372号；同第9,840,713号；米国特許出願第13/842,859号（US2014/0068797 A1）；同第9,260,723号；同第9,023,649号；同第9,834,791号；同第9,637,739号；米国特許出願第14/683,443号（US2015/0240261 A1）；米国特許出願第14/743,764号（US2015/0291961 A1）；同第9,790,490号；同第9,688,972号；同第9,580,701号；同第9,745,562号；同第9,816,081号；同第9,677,090号；同第9,738,687号；米国特許出願第15/632,222号（US2017/0369879 A1）；米国特許出願第15/631,989号；米国特許出願第15/632,001号；および米国特許第9,896,696号を参照されたい。

【0742】

一部の実施形態では、Casエフェクタータンパク質は、以下の活性の1つまたは複数を含む：

【0743】

ニッカーゼ活性、すなわち、核酸分子の一本鎖を切断する能力；

【0744】

二本鎖ヌクレアーゼ活性、すなわち、二本鎖核酸の両方の鎖を切断し、二本鎖切断を生じさせる能力；

【0745】

エンドヌクレアーゼ活性；

【0746】

エキソヌクレアーゼ活性；および／または

【0747】

ヘリカーゼ活性、すなわち、二本鎖核酸のヘリックス構造を巻き戻す能力。

【0748】

本開示の諸態様では、用語「ガイド核酸」は、1) 標的配列にハイブリダイズすることができるガイド配列（本明細書では、「標的化セグメント」と呼ばれる）および2) 本明

細書に記載の核酸ガイド型ヌクレアーゼと相互作用することができる足場配列（単独で、または t r a c r R N A 分子と組み合わせてのいずれかで）（本明細書では、「足場セグメント」と呼ばれる）を含むポリヌクレオチドを指す。ガイド核酸は、D N A であり得る。ガイド核酸は、R N A であり得る。ガイド核酸は、D N A および R N A のどちらも含み得る。ガイド核酸は、修飾された天然に存在しないヌクレオチドを含み得る。ガイド核酸が R N A を含む場合、R N A ガイド核酸は、本明細書に開示されるプラスミド、直鎖状コンストラクト、または編集カセットなどのポリヌクレオチド分子上の D N A 配列によってコードされ得る。

【0749】

一部の実施形態では、本明細書に記載のガイド核酸は R N A ガイド核酸（「ガイド R N A」または「g R N A」）であり、標的化セグメントおよび足場セグメントを含む。一部の実施形態では、g R N A の足場セグメントが 1 つの R N A 分子に含まれ、標的化セグメントは別の独立した R N A 分子に含まれる。このような実施形態は、本明細書では、「二重分子 g R N A」または「2 分子 g R N A」または「二重 g R N A」と呼ばれる。一部の実施形態では、g R N A は単一の R N A 分子であり、本明細書では、「単一のガイド R N A」または「s g R N A」と呼ばれる。用語「ガイド R N A」または「g R N A」は包括的であり、2 分子ガイド R N A および s g R N A のどちらも指す。

【0750】

g R N A の D N A 標的化セグメントは、標的核酸配列内の配列と相補的なヌクレオチド配列を含む。したがって、g R N A の標的化セグメントは、標的核酸と配列特異的にハイブリダイゼーション（すなわち、塩基対合）によって相互作用し、標的化セグメントのヌクレオチド配列により、g R N A が結合する標的 D N A 内の場所が決定される。ガイド配列とその対応する標的配列の間の相補性の度合いは、適当な整列アルゴリズムを使用して最適に整列したとき、約 50 % であるもしくはそれよりも大きい、約 60 % であるもしくはそれよりも大きい、約 75 % であるもしくはそれよりも大きい、約 80 % であるもしくはそれよりも大きい、約 85 % であるもしくはそれよりも大きい、約 90 % であるもしくはそれよりも大きい、約 95 % であるもしくはそれよりも大きい、約 97.5 % であるもしくはそれよりも大きい、約 99 % であるもしくはそれよりも大きい、またはそれよりも大きい。最適な整列は、配列を整列するための任意の適当なアルゴリズムを使用して決定することができる。一部の実施形態では、ガイド配列の長さは、約 5 ヌクレオチドであるもしくはそれよりも長い、約 6 ヌクレオチドであるもしくはそれよりも長い、約 7 ヌクレオチドであるもしくはそれよりも長い、約 8 ヌクレオチドであるもしくはそれよりも長い、約 9 ヌクレオチドであるもしくはそれよりも長い、約 10 ヌクレオチドであるもしくはそれよりも長い、約 11 ヌクレオチドであるもしくはそれよりも長い、約 12 ヌクレオチドであるもしくはそれよりも長い、約 13 ヌクレオチドであるもしくはそれよりも長い、約 14 ヌクレオチドであるもしくはそれよりも長い、約 15 ヌクレオチドであるもしくはそれよりも長い、約 16 ヌクレオチドであるもしくはそれよりも長い、約 17 ヌクレオチドであるもしくはそれよりも長い、約 18 ヌクレオチドであるもしくはそれよりも長い、約 19 ヌクレオチドであるもしくはそれよりも長い、約 20 ヌクレオチドであるもしくはそれよりも長い、約 21 ヌクレオチドであるもしくはそれよりも長い、約 22 ヌクレオチドであるもしくはそれよりも長い、約 23 ヌクレオチドであるもしくはそれよりも長い、約 24 ヌクレオチドであるもしくはそれよりも長い、約 25 ヌクレオチドであるもしくはそれよりも長い、約 26 ヌクレオチドであるもしくはそれよりも長い、約 27 ヌクレオチドであるもしくはそれよりも長い、約 28 ヌクレオチドであるもしくはそれよりも長い、約 29 ヌクレオチドであるもしくはそれよりも長い、約 30 ヌクレオチドであるもしくはそれよりも長い、約 31 ヌクレオチドであるもしくはそれよりも長い、約 32 ヌクレオチドであるもしくはそれよりも長い、約 33 ヌクレオチドであるもしくはそれよりも長い、約 34 ヌクレオチドであるもしくはそれよりも長い、約 35 ヌクレオチドであるもしくはそれよりも長い、約 36 ヌクレオチドであるもしくはそれよりも長い、約 37 ヌクレオチドであるもしくはそれよりも長い、約 38 ヌクレオチドであるもしくはそれよりも長い、

10

20

30

40

50

約 39ヌクレオチドであるもしくはそれよりも長い、約 40ヌクレオチドであるもしくはそれよりも長い、約 45ヌクレオチドであるもしくはそれよりも長い、約 50ヌクレオチドであるもしくはそれよりも長い、約 75ヌクレオチドであるもしくはそれよりも長い、またはそれよりも長い。一部の実施形態では、ガイド配列の長さは、約 75ヌクレオチド未満、約 50ヌクレオチド未満、約 45ヌクレオチド未満、約 40ヌクレオチド未満、約 35ヌクレオチド未満、約 30ヌクレオチド未満、約 25ヌクレオチド未満、約 20ヌクレオチド未満である。諸態様では、ガイド配列の長さは、10～30ヌクレオチドである。ガイド配列の長さは、15～20ヌクレオチドであり得る。ガイド配列の長さは15ヌクレオチドであり得る。ガイド配列の長さは16ヌクレオチドであり得る。ガイド配列の長さは17ヌクレオチドであり得る。ガイド配列の長さは18ヌクレオチドであり得る。ガイド配列の長さは19ヌクレオチドであり得る。ガイド配列の長さは20ヌクレオチドであり得る。

10

【0751】

ガイドRNAの足場セグメントは、1つまたは複数のCasエフェクタータンパク質と相互作用して、リボ核タンパク質複合体（本明細書では、CRISPR-RNPまたはRNP-複合体と呼ばれる）を形成する。ガイドRNAにより、結合したポリペプチドが標的核酸配列内の特定のヌクレオチド配列に上記の標的化セグメントを介して向きつけられる。ガイドRNAの足場セグメントは、互いに相補的であり、二本鎖RNA二重鎖を形成するヌクレオチドの2つのストレッチ含む。標的化可能なヌクレアーゼ複合体の形成を促進するのに十分な足場配列内の配列は、足場配列内の2つの配列領域、例えば二次構造の形成に關与する1つまたは2つの配列領域などの長さに沿ってある程度の相補性を含み得る。一部の場合では、1つまたは2つの配列領域は、同じポリヌクレオチドに含まれるまたはそれにコードされる。一部の場合では、1つまたは2つの配列領域は、別のポリヌクレオチドに含まれるまたはそれらにコードされる。最適な整列は、任意の適当な整列アルゴリズムによって決定することができ、1つまたは2つの配列領域のいずれか内の自己相補性などの二次構造をさらに説明することができる。一部の実施形態では、1つまたは2つの配列領域間の、2つのうち短い方の長さに沿った相補性の度合いは、最適に整列したとき、約25%であるもしくはそれよりも大きい、約30%であるもしくはそれよりも大きい、約40%であるもしくはそれよりも大きい、約50%であるもしくはそれよりも大きい、約60%であるもしくはそれよりも大きい、約70%であるもしくはそれよりも大きい、約80%であるもしくはそれよりも大きい、約90%であるもしくはそれよりも大きい、約95%であるもしくはそれよりも大きい、約97.5%であるもしくはそれよりも大きい、約99%であるもしくはそれよりも大きい、またはそれよりも大きい。一部の実施形態では、2つの配列領域の少なくとも1つの長さは、約5ヌクレオチドであるもしくはそれよりも長い、約6ヌクレオチドであるもしくはそれよりも長い、約7ヌクレオチドであるもしくはそれよりも長い、約8ヌクレオチドであるもしくはそれよりも長い、約9ヌクレオチドであるもしくはそれよりも長い、約10ヌクレオチドであるもしくはそれよりも長い、約11ヌクレオチドであるもしくはそれよりも長い、約12ヌクレオチドであるもしくはそれよりも長い、約13ヌクレオチドであるもしくはそれよりも長い、約14ヌクレオチドであるもしくはそれよりも長い、約15ヌクレオチドであるもしくはそれよりも長い、約16ヌクレオチドであるもしくはそれよりも長い、約17ヌクレオチドであるもしくはそれよりも長い、約18ヌクレオチドであるもしくはそれよりも長い、約19ヌクレオチドであるもしくはそれよりも長い、約20ヌクレオチドであるもしくはそれよりも長い、約25ヌクレオチドであるもしくはそれよりも長い、約30ヌクレオチドであるもしくはそれよりも長い、約40ヌクレオチドであるもしくはそれよりも長い、約50ヌクレオチドであるもしくはそれよりも長い、またはそれよりも長い。

20

30

40

【0752】

対象gRNAの足場配列は、二次構造を含み得る。二次構造は、シュードノット領域またはステムループ構造を含み得る。一部の例では、ガイド核酸および核酸ガイド型ヌクレアーゼの適合性は、少なくとも部分的に、ガイドRNAの二次構造領域内の配列または二

50

次構造領域に隣接する配列によって決定される。一部の場合では、ガイド核酸の核酸ガイド型ヌクレアーゼに対する結合カイネティクスは、部分的に、足場配列内の二次構造によって決定される。一部の場合では、ガイド核酸の核酸ガイド型ヌクレアーゼに対する結合カイネティクスは、部分的に、足場配列を有する核酸配列によって決定される。

【0753】

gRNA-Casエフェクタータンパク質の組合せに適合する足場配列は、生来のCasヌクレアーゼ遺伝子座に隣接する配列をスキャンすることによって見出すことができる。言い換えれば、生来のCasヌクレアーゼは、対応する適合するガイド核酸または足場配列の近傍にあるゲノム上にコードされ得る。

【0754】

核酸ガイド型ヌクレアーゼは、ヌクレアーゼが内在する宿主内では見出されないガイド核酸と適合し得る。このような直交性ガイド核酸は、経験的試験によって決定することができる。直交性ガイド核酸は、異なる細菌種に由来し得るまたは合成もしくは他のやり方で操作されて天然には存在しないものであり得る。一般的な核酸ガイド型ヌクレアーゼと適合する直交性ガイド核酸は、1つまたは複数の一般的な特徴を含み得る。一般的な特徴は、シュードノット領域の外側の配列を含み得る。一般的な特徴は、シュードノット領域を含み得る。一般的な特徴は、一次配列または二次構造を含み得る。

【0755】

ガイド核酸は、ガイド配列が標的配列と相補的になり、それにより、ガイド配列と標的配列のハイブリダイゼーションが可能になるようにガイド配列を変更することによって、所望の標的配列を標的にするように操作することができる。操作されたガイド配列を有するガイド核酸は操作されたガイド核酸と呼ぶことができる。操作されたガイド核酸は、多くの場合、天然に存在せず、天然には見出されない。

【0756】

一部の実施形態では、本開示は、gRNAをコードするポリヌクレオチドを提供する。一部の実施形態では、gRNAコード核酸は、発現ベクター、例えば、組換え発現ベクター中に含まれる。一部の実施形態では、本開示は、部位指向変異性ポリペプチドをコードするポリヌクレオチドを提供する。一部の実施形態では、部位指向変異性ポリペプチドをコードするポリヌクレオチドは、発現ベクター、例えば、組換え発現ベクター中に含まれる。

【実施例】

【0757】

以下の実施例は、本開示の種々の実施形態を例証するために提供するのであって、それらには、いかなる場合であっても、本開示を制限する意図はない。特許請求の範囲により定義される、本開示の精神の範囲内に包含されるその変化および他の使用は、当業者により認識される。

【0758】

特に、実施例1～9は、CoynebacteriumにおけるHTPゲノム操作プラットフォームの実証である。しかし、同様の手順がE. coliのためにカスタマイズされ、本発明者らにより首尾よく実施されている。

【0759】

簡潔な目次を、下記に提供するが、これはもっぱら、読者を支援するためである。この目次により、本出願の実施例または開示の範囲を制限する意図は全くない。

10

20

30

40

【表 5. 1 - 1】

表5.1-実施例のセクションの目次

実施例 の番号	表題	簡潔な説明
1	<i>Corynebacterium</i> の HTP 形質転換および SNP ライブラリー創製の実証	本開示のハイスループット遺伝子操作法の実施形態を記載する。
2	HTP ゲノム操作-工業用微生物株を再建/改良するための、SNP ライブラリーのインプリメンテーション	工業用生物を、本開示の SNP スワップ方法により再建するためのアプローチについて記載する。
3	HTP ゲノム操作- <i>Corynebacterium</i> 中でのリシンの産生における株性能を改善するための、SNP スワップライブラリーのインプリメンテーション	リシンを産生する <i>Corynebacterium</i> の株性能を改善するための SNP スワップ技法のインプリメンテーションについて記載する。また、選択された二次および三次の変異の統合も開示する。
4	HTP ゲノム操作-工業用微生物株を改良するための、プロモータースワップライブラリーのインプリメンテーション	宿主生物の株性能を、本開示の PRO スワップ遺伝子設計ライブラリーにより改善するための方法について記載する。
5	HTP ゲノム操作-リシンの産生に関して、株性能を改善するための PRO ス	リシンを産生する <i>Corynebacterium</i> 株の性能を改善するための PRO スワップ

10

20

30

40

【表 5. 1 - 2】

実施例 の番号	表題	簡潔な説明
	ワップライブラリーのインプリメンテーション	技法のインプリメンテーションについて記載する。
6	上位性マッピング-変異の有益な統合を予測するためのアルゴリズムツール	遺伝子変異の有益な統合を予測するための、本開示の自動化されたツール/アルゴリズムの実施形態を記載する。
7	HTP ゲノム操作-PRO スワップ変異の統合および多因子のコンビナトリアルな試験	本開示の HTP 方法により、遺伝子/遺伝子設計ライブラリーを複数組み合わせたコンビナトリアルな統合により創製される大きな解空間を有効に探索することが可能になることを記載し、例示する。
8	HTP ゲノム操作-工業用宿主株を改良するための、ターミネーターライブラリーのインプリメンテーション	本開示の STOP スワップ 遺伝子設計ライブラリーの適用について記載し、例示する。
9	HTP ツールセットと従来の UV 変異との比較	本開示の HTP 遺伝子設計法と従来の変異による株改良プログラムとを比較した実験結果 を提供する。
10	HTP ゲノム操作-Escherichia coli の株性能を改善するための、トラ	<i>E. coli</i> 細胞に適用される、本開示のハイスループット

10

20

30

40

【表 5. 1 - 3】

実施例 の番号	表題	簡潔な説明
	ンスポゾン変異誘発ライブラリーのインプリメンテーション	遺伝子操作方法を用いたトランスポゾン変異誘発ライブラリーの使用の諸実施形態を記載する。
11	HTP ゲノム操作- <i>Escherichia coli</i> における HTP ゲノム操作に使用するためのベクター骨格の生成	<i>E. coli</i> 細胞に適用される、本開示のハイスループット遺伝子操作方法に使用するためのベクター骨格の生成の諸実施形態を記載する。
12	HTP ゲノム操作-工業用微生物株の改良において使用するための追加のプロモータースワップライブラリーの生成および試験	本開示の PRO スワップ遺伝子設計ライブラリーを設計し、生成し、試験するための方法を記載する。
13	HTP ゲノム操作-ベクター2 骨格を使用した表 1.4 のプロモータースワップライブラリーの <i>E. coli</i> のゲノムへの組込みの試験	単一コピーの異種プロモーター-遺伝子コンストラクトの <i>E. coli</i> のゲノムへの組込みを駆動するための、実施例 11 からのベクター2 骨格と表 1.4 のプロモーターライブラリーからのプロモーターの組合せの使用の概念実証について記載する。

10

20

30

40

【表 5. 1 - 4】

実施例 の番号	表題	簡潔な説明
14	HTP ゲノム操作-表 1. 4 から得られるプロモーターライブラリーを使用した PROSWP 方法のインプリメンテーション。	本開示の表 1. 4 から得られる PRO スワップ遺伝子設計ライブラリーを通じて宿主生物の株性能を改善するための方法を記載する。
15	HTP ゲノム操作-リコピンの産生に関して、株性能を改善するための、ターミネータースワップライブラリーのインプリメンテーション	本開示のターミネータースワップ遺伝子設計ライブラリーを通じて宿主生物の株性能を改善するための方法を記載する。
16	リコピンの産生に関して、株性能を改善するための、HTP ゲノム操作-ターミネータースワップライブラリーまたは PRO スワップライブラリーの、SOLUBILITY TAG スワップライブラリーまたは DEGRADATION Tag スワップライブラリーのいずれかと組み合わせたインプリメンテーション	本開示のターミネータースワップ、溶解性タグスワップおよび分解タグスワップ遺伝子設計ライブラリーを通じて宿主生物の株性能を改善するための方法を記載する。

【0760】

(実施例 1)

Corynebacterium の HTP 形質転換および SNP ライブラリー創製の実証
この実施例は、本開示の HTP 遺伝子操作法の実施形態を示す。宿主細胞を、異なるサイズの多様な SNP 配列を用いて形質転換し、すべて、ゲノムの異なるエリアを標的化する。結果から、本開示の方法により、任意の種類の迅速な遺伝子変化を、宿主細胞のゲノム全体にわたり生成することが可能であることが示されている。

【0761】

A. 形質転換ベクターのクローニング

10

20

30

40

50

「カスタムプラスミドアセンブリー／クローニング」のセクションにおける上記の記載の通り、および図3に示すように、多様なSNPを、*Corynebacterium glutamicum* (ATCC 21300) からランダムに選び、*Corynebacterium* クローニングベクター中に、酵母相同組換えクローニングの技法を使用してクローニングして、各SNPがダイレクトリピート領域に隣接する、ベクターをアセンブルした。

【0762】

この実施例のためのSNPカセットは、0.5 Kb、1 Kb、2 Kbおよび5 Kbの範囲のダイレクトリピートホモロジーアームの様々な長さを含むように設計した。さらに、SNPカセットは、下記でより詳細に記載する通り、ゲノムの種々の明確に異なる領域を標的とする相同組換えのために設計した。

10

【0763】

C. glutamicum のゲノムのサイズは、3,282,708 bpである(図9を参照されたい)。ゲノムを、24個の等しいサイズの遺伝子領域に恣意的に分け、SNPカセットを、24個の領域のそれぞれを標的とするように設計した。したがって、この実施例のために、全部で96個の明確に異なるプラスミドをクローニングした(4つの異なる挿入断片サイズ×24個の明確に異なるゲノム領域)。

【0764】

各DNA挿入断片を、相同領域のPCR増幅により、商業的に供給されているオリゴ、および鋳型としての、上に記載した宿主株のゲノムDNAを使用して産生した。ゲノム中に導入しようとするSNPを、オリゴのテールにコードした。酵母において、PCR断片を、ベクターの骨格中に相同組換えを使用してアセンブルした。

20

【0765】

ベクター中への各SNPおよびホモロジーアームのクローニングを、図6A～B、図3および表5に記載するHTP操作のワークフローに従って実施した。

【0766】

B. E. coli 中へのアセンブルされたクロンの形質転換

最初に、標準的な熱ショック形質転換の技法を使用して、ベクターをE. coli 中へ形質転換して、正確にアセンブルされたクロンを同定し、*Corynebacterium* の形質転換のためのベクターDNAを増幅した。

30

【0767】

例えば、形質転換したE. coli 細菌を、アセンブリーの成功について試験した。各E. coli 形質転換プレートからの4つのコロニーを培養し、正確なアセンブリーについてPCRにより試験した。このプロセスを、24個の形質転換の場所のそれぞれおよび4つの異なる挿入断片サイズのそれぞれ(すなわち、この実施例の96個の形質転換体のすべて)について繰り返した。この実験からの結果を、各処理(挿入断片サイズおよびゲノムの場所)について試験した4つのコロニーから同定した正確なコロニーの数として表した(図12を参照されたい)。より長い5 kbの挿入断片は、より短い対応物と比較して、アセンブリー効率の減少を示した($n=96$)。

【0768】

C. *Corynebacterium* 中へのアセンブルされたクロンの形質転換

バリデートしたクロンを、電気穿孔により*Corynebacterium glutamicum* 宿主細胞に形質転換した。各形質転換について、1 μ g のDNA当たりのコロニー形成単位(CFU)の数を、挿入断片サイズの関数として決定した(図13を参照されたい)。また、Coryneのゲノムへの組込みを、ホモロジーアームの長さの関数としても分析し、結果は、より短いアームは、より低い効率を有することを示した(図13を参照されたい)。

40

【0769】

また、C. *glutamicum* 形質転換体中の標的化したゲノムの場所に関するゲノムへの組込み効率も分析した。ゲノムの1位および2位は、ゲノムの残りの部分と比較し

50

て、若干より低い組込み効率を示した（図10を参照されたい）。

【0770】

D. 選択マーカーのループアウト

挿入断片カセットの組込みに成功したことが同定された *Corynebacterium* の培養物を、5%のスクロースを含有する培地上で培養し、対抗選択して、*sacB* 選択遺伝子をループアウトした。種々のダイレクトリピートホモロジーアームについてのスクロース耐性の頻度は、アームの長さに関して有意には変化しなかった（図14を参照されたい）。これらの結果から、ループアウト効率が、0.5 kb～5 kbのホモロジーアームの長さにわたり一様に維持されることが示唆された。

【0771】

ループアウト事象をさらにバリデーションするために、スクロース耐性を示すコロニーを培養し、シーケンシングにより分析した。

【0772】

挿入断片のゲノム領域のシーケンシングについての結果を、下記の表6に要約する。

【表6】

表6-ループアウトのバリデーション頻度

結果	頻度(試料採取誤差/95%信頼性)
成功したループアウト	13% (9%/20%)
ループが依然として存在する	42% (34%/50%)
混合型の読取り	44% (36%/52%)

【0773】

シーケンシングの結果から、ループアウトの10～20%の効率が示された。実際のループアウトはおそらく、挿入断片配列に若干依存する。しかし、10～20個のスクロース耐性のコロニーを選ぶことによって、高い成功率がもたらされる。

【0774】

E. 要約

下記の表7に、本発明のHTPゲノム操作法の効率の定量的な評価を示す。酵母ホモロジーの方法論についての構築物のアセンブリーの率として、期待されるDNA構築物を、10個の試験したコロニーのうちの約9個において得た。2 kbのホモロジーアームを有するSNP構築物による *Coryne* の形質転換により、平均して1マイクログラムのDNA当たり51コロニー形成単位(CFU/μg)が得られ、前記コロニーの98%が、正確に組み込まれたSNP挿入断片を示した（標的化効率）。ループアウト効率として、スクロースに曝露する場合に、耐性となる細胞が0.2%に留まり、これらの13%が、正確にループアウトされた配列を示した。

【表 7】

表 7-Corynebacterium glutamicum 株の操作についての結果の要約

QC ステップ	2 kb のホモロジーアームについての結果
構築物のアセンブリーの成功	87%
Coryne の形質転換の効率	51 CFU/ μ g DNA (+/- 15)
標的化効率	98%
ループアウト効率	0.2% (+/- 0.03%)

10

【0775】

20

(実施例 2)

H T P ゲノム操作—工業用微生物株を再建／改良するための、S N P ライブラリーのインプリメンテーション

この実施例は、本開示の H T P 株改良プログラムの S N P スワップライブラリーのいくつかの態様を例示する。具体的には、実施例は、現時点で存在する工業用株を再建するためのいくつかの想定されるアプローチを示す。この実施例は、「基準」株と「中間」株と工業用株との間に存在し得る遺伝子の複数の差により生み出される表現型の解空間を探索する、ウェーブアップおよびウェーブダウンのアプローチについて記載する。

【0776】

30

A. 多様性プール中の S N P の同定

本開示の方法を使用する例示的な株改良プログラムを、工業生産用微生物株に対して実施した；これを、本明細書では「C」と呼ぶ。このプログラムについての多様性プールの株を、A、B および C により表す。株 A は、あらゆる変異誘発の前の、元々の、産生のための宿主株を表した。株 C は、現在の工業用株を表し、従来の株改良プログラムによる長年の変異誘発および選択を受けている。株 B は、「中位」の株を表し、何らかの変異誘発を受けており、株 C の先行株であった。（図 17 A を参照されたい）。

【0777】

40

株 A、株 B および株 C について、シーケンシングし、それらのゲノムを、株間の遺伝子の差について分析した。全部で 332 個の非同義 S N P を同定した。これらのうち、133 個の S N P が C にユニークであり、加えて、153 個が B および C により共有され、46 個が株 B にユニークであった（図 17 B を参照されたい）。これらの S N P を、下流の株改良サイクルのために多様性プールとして使用する。

【0778】

B. S N P スワッピングの分析

実施例 2 のパート A の多様性プールから同定した S N P を分析して、宿主細胞の性能に対するそれらの効果を決定する。株性能の最初の「学習」のラウンドを、下に記載し、図 18 に図示した通り、6 つのステップに分解する。

【0779】

第 1 に、C からの S N P をすべて、基準株 A 中に個々および／またはコンビナトリアルにクローニングする。これは、最少 286 個の個別の形質転換体に相当する。これらの形

50

質転換体の目的は、有益な S N P を同定することである。

【0780】

第2に、CからのS N Pをすべて、市販株Cから個々および／またはコンビナトリアルに除去する。これは、最少286個の個別の形質転換体に相当する。これらの形質転換体の目的は、中立のS N Pおよび有害なS N Pを同定することである。また、追加の任意選択のステップ3～6についても、下に記載する。2つの遺伝的な時点（genetic time point）（基準株Aおよび工業用株C）からのS N Pを付加および減じる、第1および第2のステップを、本明細書では「ウェーブ」と呼び、ウェーブは、「ウェーブアップ」（基準株へのS N Pの付加、第1のステップ）および「ウェーブダウン」（工業用株からのS N Pの除去、第2のステップ）を含む。ウェーブの概念を、S N Pのさらなる付加／削減に拡張する。 10

【0781】

第3に、BからのS N Pをすべて、基準株A中に個々および／またはコンビナトリアルにクローニングする。これは、最少199個の個別の形質転換体に相当する。これらの形質転換体の目的は、有益なS N Pを同定することである。形質転換体のうちのいくつかはまた、第1のステップで産生する形質転換体についてのバリデーションデータとしても役立つ。

【0782】

第4に、BからのS N Pをすべて、市販株Bから個々および／またはコンビナトリアルに除去する。これは、最少199個の個別の形質転換体に相当する。これらの形質転換体の目的は、中立のS N Pおよび有害なS N Pを同定することである。形質転換体のうちのいくつかはまた、第2のステップで産生する形質転換体についてのバリデーションデータとしても役立つ。 20

【0783】

第5に、CにユニークなS N P（すなわち、B中には存在しない）をすべて、個々および／またはコンビナトリアルに市販株B中にクローニングする。これは、最少46個の個別の形質転換体に相当する。これらの形質転換体の目的は、有益なS N Pを同定することである。形質転換体のうちのいくつかはまた、第1および第3のステップで産生する形質転換体についてのバリデーションデータとしても役立つ。

【0784】

第6に、CにユニークなS N Pをすべて、市販株Cから個々および／またはコンビナトリアルに除去する。これは、最少46個の個別の形質転換体に相当する。これらの形質転換体の目的は、中立のS N Pおよび有害なS N Pを同定することである。形質転換体のうちのいくつかはまた、第2および第4のステップで産生する形質転換体についてのバリデーションデータとしても役立つ。 30

【0785】

これらのステップのそれぞれから収集するデータを使用して、各S N Pを、疎明の、有益、中立または有害として分類する。

【0786】

C. S N Pの有益な組合せを決定するための、上位性マッピングの利用 40

実施例2のパートBで同定した有益なS N Pを、組み合わせた場合に、宿主の性能を改善する可能性があるS N Pを同定するために、本開示の上位性マッピング法により分析する。

【0787】

実施例1の操作法を使用して、新たに操作された株バリエーションを創製して、S N Pの組合せを、上位性マッピングによる予測に従って試験する。S N Pの統合は、逐次的に行うことができ、または代替的に、複数の分枝にわたって行うこともでき、したがって、1つ超の改良された株が、有益なS N Pのサブセットを伴い存在し得る。S N Pの統合は、株改良の複数のラウンドにわたり、有益なS N Pの最適な組合せを含有する最終の株を産生するまで続け、最終の株は、中立のS N Pまたは有害なS N Pの荷物をいずれも有さない 50

。

【0788】

(実施例3)

H T Pゲノム操作-C o r y n e b a c t e r i u m中でのリシンの産生における株性能を改善するための、S N Pスワップライブラリーのインプリメンテーション

この実施例では、実施例2のS N PスワップH T P設計株改良プログラムの一部を例示的に実行し、目標は、C o r y n e b a c t e r i u m中でのリシンの産生の収率および生産性の改善をもたらすことである。

【0789】

この実施例のセクションBは、本開示のH T P株改良プログラムの、変異を統合するステップをさらに示す。したがって、この実施例は、本開示のH T P株改良法の第1、第2および第3のラウンドの統合についての実験結果を提供する。

10

【0790】

第2および第3のラウンドの統合のための変異は、別個の遺伝子ライブラリースワップに由来する。したがってまた、これらの結果により、H T P株プログラムの、多分枝平行トラック (multi-branch parallel tracks) を実施する能力、および本開示の種々の形態の遺伝子設計ライブラリーに伴うメタデータ中に埋め込むことができる有益な変異の「メモリー」も示される。

【0791】

上記のように、提供された参照の基準参照株 (株A) および第2の「操作された」株 (株C) のゲノムについてシーケンシングし、遺伝子の差をすべて同定した。基準株は、U V変異誘発を経験したことがないC o r y n e b a c t e r i u m g l u t a m i c u mバリエーションであった。操作された株もまた、C . g l u t a m i c u m株であり、これは、従来の変異改善プログラムの数ラウンド後に基準株から産生されていた。この実施例は、186個の明確に異なる非同義S N Pの、株Aと株Cとの間で同定した差についてのS N Pスワップの結果を提供する。

20

【0792】

A. H T P操作およびハイスループットスクリーニング

本開示のクローニングおよび形質転換の方法に従って、186個の同定したS N Pのそれぞれを個々に、基準株中に付加して戻した。単一のS N Pを含む新たに創製した株をそれぞれ、リシンの収率について、産物力価での性能を評価するように設計した小規模の培養において試験した。小規模の培養は、工業規模の培養からの培地を使用して実施した。標準的な比色分析アッセイを用いて、産物力価を光学的に炭素消耗時に測定した (すなわち、単一バッチの収率を表す)。手短に述べると、濃縮アッセイ混合物を調製し、これを発酵試料に、試薬の最終濃度が、160 mMのリン酸ナトリウム緩衝液、0.2 mMのA m p l e x R e d、0.2 U / m Lの西洋ワサビペルオキシダーゼ、および0.005 U / m Lのリシンオキシダーゼとなるように添加した。反応を終点まで進行させ、T e c a n M 1 0 0 0 プレート分光光度計を使用して、光学密度を560 nmの波長において測定した。実験の結果を、下記の表8に要約し、図38に示す。

30

【表 8 - 1】

表8ーリシンの生産に関する、SNPスワップ株の操作についての結果の要約

SNP	N	リシンの平均収率 (参照株と比較した A ₅₆₀ における変化)	標準誤差	参照に対する%変化	%変化の誤差
DSS_033	4	0.1062	0.00888	11.54348	2.895652
DSS_311	2	0.03603	0.01256	3.916304	4.095652
DSS_350	1	0.03178	0.01777	3.454348	5.794565
DSS_056	3	0.02684	0.01026	2.917391	3.345652

【表 8 - 2】

SNP	N	リシンの平均収率 (参照株と比較した A ₅₆₀ における変化)	標準誤差	参照に対する%変化	%変化の誤差
DSS_014	4	0.02666	0.00888	2.897826	2.895652
DSS_338	3	0.02631	0.01026	2.859783	3.345652
DSS_128	1	0.02584	0.01777	2.808696	5.794565
DSS_038	4	0.02467	0.00888	2.681522	2.895652
DSS_066	4	0.02276	0.00888	2.473913	2.895652
DSS_108	2	0.02216	0.01256	2.408696	4.095652
DSS_078	4	0.02169	0.00888	2.357609	2.895652
DSS_017	3	0.02102	0.01026	2.284783	3.345652
DSS_120	3	0.01996	0.01026	2.169565	3.345652
DSS_064	4	0.01889	0.00888	2.053261	2.895652
DSS_380	4	0.01888	0.00888	2.052174	2.895652
DSS_105	3	0.0184	0.01026	2	3.345652
DSS_407	1	0.01831	0.01777	1.990217	5.794565
DSS_018	2	0.01825	0.01256	1.983696	4.095652
DSS_408	3	0.01792	0.01026	1.947826	3.345652
DSS_417	3	0.01725	0.01026	1.875	3.345652

10

20

30

40

【表 8 - 3】

SNP	N	リシンの平均収率 (参照株と比較した A ₅₆₀ における変化)	標準誤差	参照に対する%変化	%変化の誤差
DSS_130	3	0.01724	0.01026	1.873913	3.345652
DSS_113	4	0.0172	0.00888	1.869565	2.895652
DSS_355	3	0.01713	0.01026	1.861957	3.345652
DSS_121	3	0.01635	0.01026	1.777174	3.345652
DSS_097	2	0.0162	0.01256	1.76087	4.095652
DSS_107	3	0.01604	0.01026	1.743478	3.345652
DSS_110	2	0.01524	0.01256	1.656522	4.095652
DSS_306	4	0.01501	0.00888	1.631522	2.895652
DSS_316	1	0.01469	0.01777	1.596739	5.794565
DSS_325	4	0.01436	0.00888	1.56087	2.895652
DSS_016	4	0.01416	0.00888	1.53913	2.895652
DSS_324	4	0.01402	0.00888	1.523913	2.895652
DSS_297	4	0.01391	0.00888	1.511957	2.895652
DSS_118	2	0.01371	0.01256	1.490217	4.095652
DSS_100	2	0.01326	0.01256	1.441304	4.095652
DSS_019	1	0.01277	0.01777	1.388043	5.794565

10

20

30

40

【表 8 - 4】

SNP	N	リシンの平均収率 (参照株と比較した A ₅₆₀ における変化)	標準誤差	参照に対する%変化	%変化の誤差
DSS_131	3	0.01269	0.01026	1.379348	3.345652
DSS_394	4	0.01219	0.00888	1.325	2.895652
DSS_385	3	0.01192	0.01026	1.295652	3.345652
DSS_395	1	0.01162	0.01777	1.263043	5.794565
DSS_287	4	0.01117	0.00888	1.21413	2.895652
DSS_418	2	0.01087	0.01256	1.181522	4.095652
DSS_290	3	0.01059	0.01026	1.151087	3.345652
DSS_314	2	0.01036	0.01256	1.126087	4.095652
DSS_073	4	0.00986	0.00888	1.071739	2.895652
DSS_040	4	0.00979	0.00888	1.06413	2.895652
DSS_037	4	0.00977	0.00888	1.061957	2.895652
DSS_341	1	0.00977	0.01777	1.061957	5.794565
DSS_302	4	0.00939	0.00888	1.020652	2.895652
DSS_104	4	0.00937	0.00888	1.018478	2.895652
DSS_273	2	0.00915	0.01256	0.994565	4.095652
DSS_322	4	0.00906	0.00888	0.984783	2.895652

10

20

30

40

【表 8 - 5】

SNP	N	リシンの平均収率 (参照株と比較した A ₅₆₀ における変化)	標準誤差	参照に対する%変化	%変化の誤差
DSS_271	3	0.00901	0.01026	0.979348	3.345652
DSS_334	2	0.00898	0.01256	0.976087	4.095652
DSS_353	4	0.00864	0.00888	0.93913	2.895652
DSS_391	4	0.00764	0.00888	0.830435	2.895652
DSS_372	1	0.00737	0.01777	0.801087	5.794565
DSS_007	1	0.00729	0.01777	0.792391	5.794565
DSS_333	2	0.0072	0.01256	0.782609	4.095652
DSS_402	4	0.00718	0.00888	0.780435	2.895652
DSS_084	1	0.0069	0.01777	0.75	5.794565
DSS_103	3	0.00676	0.01026	0.734783	3.345652
DSS_362	1	0.00635	0.01777	0.690217	5.794565
DSS_012	2	0.00595	0.01256	0.646739	4.095652
DSS_396	2	0.00574	0.01256	0.623913	4.095652
DSS_133	3	0.00534	0.01026	0.580435	3.345652
DSS_065	3	0.00485	0.01026	0.527174	3.345652
DSS_284	2	0.00478	0.01256	0.519565	4.095652

10

20

30

40

【表 8 - 6】

SNP	N	リシンの平均収率 (参照株と比較した A ₅₆₀ における変化)	標準誤差	参照に対する%変化	%変化の誤差
DSS_301	3	0.00465	0.01026	0.505435	3.345652
DSS_281	4	0.00461	0.00888	0.501087	2.895652
DSS_405	2	0.00449	0.01256	0.488043	4.095652
DSS_361	3	0.00438	0.01026	0.476087	3.345652
DSS_342	4	0.00434	0.00888	0.471739	2.895652
DSS_053	3	0.00422	0.01026	0.458696	3.345652
DSS_074	4	0.00422	0.00888	0.458696	2.895652
DSS_079	4	0.00375	0.00888	0.407609	2.895652
DSS_381	3	0.0036	0.01026	0.391304	3.345652
DSS_294	1	0.00336	0.01777	0.365217	5.794565
DSS_313	2	0.00332	0.01256	0.36087	4.095652
DSS_388	2	0.00305	0.01256	0.331522	4.095652
DSS_392	4	0.00287	0.00888	0.311957	2.895652
DSS_319	4	0.00282	0.00888	0.306522	2.895652
DSS_310	4	0.00263	0.00888	0.28587	2.895652
DSS_344	3	0.00259	0.01026	0.281522	3.345652

10

20

30

40

【表 8 - 7】

SNP	N	リシンの平均収率 (参照株と比較した A ₅₆₀ における変化)	標準誤差	参照に対する%変化	%変化の誤差
DSS_025	4	0.00219	0.00888	0.238043	2.895652
DSS_412	1	0.00204	0.01777	0.221739	5.794565
DSS_300	3	0.00188	0.01026	0.204348	3.345652
DSS_299	2	0.00185	0.01256	0.201087	4.095652
DSS_343	4	0.00184	0.00888	0.2	2.895652
DSS_330	3	0.00153	0.01026	0.166304	3.345652
DSS_416	4	0.00128	0.00888	0.13913	2.895652
DSS_034	3	0.00128	0.01026	0.13913	3.345652
DSS_291	2	0.00102	0.01256	0.11087	4.095652
DSS_115	4	0.00063	0.00888	0.068478	2.895652
DSS_288	4	0.00044	0.00888	0.047826	2.895652
DSS_309	4	0.00008	0.00888	0.008696	2.895652
DSS_125	3	0	0.01026	0	3.345652
DSS_358	3	-0.00015	0.01026	-0.0163	3.345652
DSS_099	2	-0.00015	0.01256	-0.0163	4.095652
DSS_111	4	-0.00017	0.00888	-0.01848	2.895652

10

20

30

40

【表 8 - 8】

SNP	N	リシンの平均収率 (参照株と比較した A ₅₆₀ における変化)	標準誤差	参照に対する%変化	%変化の誤差
DSS_359	3	-0.00022	0.01026	-0.02391	3.345652
DSS_015	4	-0.00043	0.00888	-0.04674	2.895652
DSS_060	3	-0.0007	0.01026	-0.07609	3.345652
DSS_098	2	-0.00088	0.01256	-0.09565	4.095652
DSS_379	4	-0.00089	0.00888	-0.09674	2.895652
DSS_356	4	-0.0009	0.00888	-0.09783	2.895652
DSS_278	4	-0.00095	0.00888	-0.10326	2.895652
DSS_368	4	-0.001	0.00888	-0.1087	2.895652
DSS_351	1	-0.0015	0.01777	-0.16304	5.794565
DSS_296	1	-0.0015	0.01777	-0.16304	5.794565
DSS_119	3	-0.00156	0.01026	-0.16957	3.345652
DSS_307	3	-0.00163	0.01026	-0.17717	3.345652
DSS_077	4	-0.00167	0.00888	-0.18152	2.895652
DSS_030	3	-0.00188	0.01026	-0.20435	3.345652
DSS_370	2	-0.00189	0.01256	-0.20543	4.095652
DSS_375	2	-0.00212	0.01256	-0.23043	4.095652

10

20

30

40

【表 8 - 9】

SNP	N	リシンの平均収率 (参照株と比較した A ₅₆₀ における変化)	標準誤差	参照に対する%変化	%変化の誤差
DSS_280	3	-0.00215	0.01026	-0.2337	3.345652
DSS_345	4	-0.00225	0.00888	-0.24457	2.895652
DSS_419	1	-0.00234	0.01777	-0.25435	5.794565
DSS_298	2	-0.00249	0.01256	-0.27065	4.095652
DSS_367	3	-0.0026	0.01026	-0.28261	3.345652
DSS_072	3	-0.00268	0.01026	-0.2913	3.345652
DSS_366	4	-0.00272	0.00888	-0.29565	2.895652
DSS_063	4	-0.00283	0.00888	-0.30761	2.895652
DSS_092	3	-0.00292	0.01026	-0.31739	3.345652
DSS_347	4	-0.0033	0.00888	-0.3587	2.895652
DSS_114	4	-0.0034	0.00888	-0.36957	2.895652
DSS_303	3	-0.00396	0.01026	-0.43043	3.345652
DSS_276	4	-0.00418	0.00888	-0.45435	2.895652
DSS_083	1	-0.00446	0.01777	-0.48478	5.794565
DSS_031	2	-0.00456	0.01256	-0.49565	4.095652
DSS_328	3	-0.00463	0.01026	-0.50326	3.345652

10

20

30

40

【表 8 - 1 0】

SNP	N	リシンの平均収率 (参照株と比較した A ₅₆₀ における変化)	標準誤差	参照に対する%変化	%変化の誤差
DSS_039	4	-0.00475	0.00888	-0.5163	2.895652
DSS_331	4	-0.00475	0.00888	-0.5163	2.895652
DSS_117	4	-0.00485	0.00888	-0.52717	2.895652
DSS_382	4	-0.00506	0.00888	-0.55	2.895652
DSS_323	4	-0.00507	0.00888	-0.55109	2.895652
DSS_041	2	-0.00527	0.01256	-0.57283	4.095652
DSS_069	4	-0.00534	0.00888	-0.58043	2.895652
DSS_308	3	-0.00534	0.01026	-0.58043	3.345652
DSS_365	3	-0.00536	0.01026	-0.58261	3.345652
DSS_403	3	-0.00594	0.01026	-0.64565	3.345652
DSS_376	1	-0.00648	0.01777	-0.70435	5.794565
DSS_293	3	-0.00652	0.01026	-0.7087	3.345652
DSS_286	1	-0.00672	0.01777	-0.73043	5.794565
BS.2C	139	-0.00694	0.00151	-0.75435	0.492391
DSS_410	1	-0.00724	0.01777	-0.78696	5.794565
DSS_312	2	-0.00725	0.01256	-0.78804	4.095652

10

20

30

40

【表 8 - 1 1】

SNP	N	リシンの平均収率 (参照株と比較した A ₅₆₀ における変化)	標準誤差	参照に対する%変化	%変化の誤差
DSS_336	1	-0.00747	0.01777	-0.81196	5.794565
DSS_327	2	-0.00748	0.01256	-0.81304	4.095652
DSS_127	4	-0.00801	0.00888	-0.87065	2.895652
DSS_332	3	-0.0085	0.01026	-0.92391	3.345652
DSS_054	2	-0.00887	0.01256	-0.96413	4.095652
DSS_024	2	-0.00902	0.01256	-0.98043	4.095652
DSS_106	3	-0.0096	0.01026	-1.04348	3.345652
DSS_400	4	-0.00964	0.00888	-1.04783	2.895652
DSS_346	3	-0.00976	0.01026	-1.06087	3.345652
DSS_320	1	-0.01063	0.01777	-1.15543	5.794565
DSS_275	4	-0.01066	0.00888	-1.1587	2.895652
DSS_371	3	-0.01111	0.01026	-1.20761	3.345652
DSS_277	1	-0.01315	0.01777	-1.42935	5.794565
DSS_282	3	-0.01326	0.01026	-1.4413	3.345652
DSS_393	3	-0.01379	0.01026	-1.49891	3.345652
DSS_378	3	-0.01461	0.01026	-1.58804	3.345652

10

20

30

40

【表 8 - 1 2】

SNP	N	リシンの平均収率 (参照株と比較した A ₅₆₀ における変化)	標準誤差	参照に対する%変化	%変化の誤差
DSS_289	3	-0.01563	0.01026	-1.69891	3.345652
DSS_317	1	-0.01565	0.01777	-1.70109	5.794565
DSS_062	4	-0.01626	0.00888	-1.76739	2.895652
DSS_340	1	-0.01657	0.01777	-1.80109	5.794565
DSS_109	2	-0.01706	0.01256	-1.85435	4.095652
DSS_011	2	-0.0178	0.01256	-1.93478	4.095652
DSS_089	4	-0.01844	0.00888	-2.00435	2.895652
DSS_059	1	-0.01848	0.01777	-2.0087	5.794565
DSS_112	2	-0.01959	0.01256	-2.12935	4.095652
DSS_043	2	-0.0213	0.01256	-2.31522	4.095652
DSS_413	1	-0.02217	0.01777	-2.40978	5.794565
DSS_305	4	-0.0227	0.00888	-2.46739	2.895652
DSS_045	4	-0.02289	0.00888	-2.48804	2.895652
DSS_082	2	-0.0231	0.01256	-2.51087	4.095652
DSS_272	1	-0.02311	0.01777	-2.51196	5.794565
DSS_390	4	-0.02319	0.00888	-2.52065	2.895652

10

20

30

40

【表 8 - 1 3】

SNP	N	リシンの平均収率 (参照株と比較した A ₅₆₀ における変化)	標準誤差	参照に対する%変化	%変化の誤差
DSS_010	3	-0.02424	0.01026	-2.63478	3.345652
DSS_357	2	-0.02525	0.01256	-2.74457	4.095652
DSS_085	4	-0.03062	0.00888	-3.32826	2.895652
DSS_044	3	-0.04088	0.01026	-4.44348	3.345652
DSS_315	2	-0.0501	0.01256	-5.44565	4.095652
DSS_080	2	-0.13519	0.01256	-14.6946	4.095652

10

20

【0793】

B. 第2のラウンドのHTP操作およびハイスループットスクリーニング-SNPスワップライブラリーと選択されたPROスワップヒットとの統合

本開示のHTP方法の強みの1つは、HTP遺伝子設計ライブラリーを、宿主細胞の表現型に対する各SNP/プロモーター/ターミネーター/開始コドンの効果と関連がある情報と一緒に保存する能力である。本発明者らは先に、プロモータースワップの実験を実施したことがあり、この実験で、生合成の収率に対してプラス効果を有する、C. glutamicum中でのいくつかのzwfプロモータースワップを同定した（例えば、図2

30

【0794】

本発明者らは、実施例5から、先に同定したzwfプロモータースワップのうちの1つもまた含むように、この実施例の基準株Aを改変した。表8中の、上に記載した最初のスクリーニングから同定した上位の176個のSNPを、この新しい基準株中に再導入して、新しいSNPスワップ遺伝子設計微生物ライブラリーを創製した。これまでのステップと同様に、単一のSNPを含む新たに創製した各株も、リシンの収率について試験した。また、選択されたSNP変異体株を、上に記載した比色分析法を使用して、24時間時のリシンの産生を測定することによって、生産性のプロキシについても試験した。このステップからの結果を、下記の表9に要約し、図39に示す。

40

【表 9 - 1】

表9ーリシンの生産に関する、SNPスワップ株操作についての第2のラウンドのスクリーニング

株の ID	SNP	N、 24時間	N、 96時間	平均、 24時間 (A ₅₆₀)	平均、 96時間 (A ₅₆₀)	標準誤差、 24時間	標準誤差、 96時間
7000006318	BS2C_P00 07_39zwf	20	2	0.49	0.82	0.00	0.02
7000008538	DSS_002	4	2	0.53	0.78	0.01	0.02
7000008539	DSS_003	4		0.56		0.01	
7000008541	DSS_005	4		0.27		0.01	
7000008542	DSS_006	4		0.49		0.01	
7000008547	DSS_011	4		0.55		0.01	
7000008548	DSS_012	4		0.58		0.01	
7000008549	DSS_013	4		0.56		0.01	
7000008550	DSS_014	4		0.52		0.01	
7000008551	DSS_015	4		0.54		0.01	
7000008552	DSS_016	4	2	0.50	0.84	0.01	0.02
7000008553	DSS_017	4		0.44		0.01	
7000008555	DSS_019	4	4	0.46	0.84	0.01	0.01
7000008557	DSS_021	4	4	0.46	0.86	0.01	0.01

10

20

30

40

【表 9 - 2】

株の ID	SNP	N、 24時間	N、 96時間	平均、 24時間 (A ₅₆₀)	平均、 96時間 (A ₅₆₀)	標準誤差、 24時間	標準誤差、 96時間
7000008559	DSS_023	4	2	0.55	0.86	0.01	0.02
7000008561	DSS_025	4		0.54		0.01	
7000008562	DSS_026	2		0.46		0.01	
7000008564	DSS_028	4		0.51		0.01	
7000008565	DSS_029	4	4	0.48	0.87	0.01	0.01
7000008566	DSS_030	4	4	0.47	0.85	0.01	0.01
7000008567	DSS_031	4		0.56		0.01	
7000008569	DSS_033	4	4	0.46	0.86	0.01	0.01
7000008570	DSS_034	2	2	0.53	0.85	0.01	0.02
7000008573	DSS_037	4		0.54		0.01	
7000008574	DSS_038	4		0.53		0.01	
7000008575	DSS_039	4		0.55		0.01	
7000008576	DSS_040	4		0.57		0.01	
7000008577	DSS_041	4		0.45		0.01	
7000008578	DSS_042	4	4	0.52	0.87	0.01	0.01
7000008579	DSS_043	4	4	0.45	0.87	0.01	0.01

10

20

30

40

【表 9 - 3】

株の ID	SNP	N、 24時間	N、 96時間	平均、 24時間 (A ₅₆₀)	平均、 96時間 (A ₅₆₀)	標準誤差、 24時間	標準誤差、 96時間
7000008580	DSS_044	4	2	0.50	0.85	0.01	0.02
7000008581	DSS_045	4		0.47		0.01	
7000008582	DSS_046	4	2	0.61	0.85	0.01	0.02
7000008583	DSS_047	4	2	0.61	0.82	0.01	0.02
7000008586	DSS_050	4		0.57		0.01	
7000008587	DSS_051	4		0.56		0.01	
7000008588	DSS_052	4	2	0.49	0.85	0.01	0.02
7000008589	DSS_053	4	4	0.45	0.85	0.01	0.01
7000008590	DSS_054	4	4	0.45	0.88	0.01	0.01
7000008592	DSS_056	4		0.42		0.01	
7000008596	DSS_060	4	2	0.55	0.87	0.01	0.02
7000008597	DSS_061	4	2	0.37	0.86	0.01	0.02
7000008598	DSS_062	4	4	0.45	0.87	0.01	0.01
7000008601	DSS_065	4	4	0.47	0.88	0.01	0.01
7000008602	DSS_066	4		0.47		0.01	
7000008604	DSS_068		2		0.51		0.02

10

20

30

40

【表 9 - 4】

株の ID	SNP	N、 24時間	N、 96時間	平均、 24時間 (A ₅₆₀)	平均、 96時間 (A ₅₆₀)	標準誤差、 24時間	標準誤差、 96時間
7000008605	DSS_069	4	4	0.47	0.88	0.01	0.01
7000008606	DSS_070	4		0.55		0.01	
7000008607	DSS_071	4	2	0.56	0.84	0.01	0.02
7000008608	DSS_072	4	2	0.54	0.83	0.01	0.02
7000008609	DSS_073	4	2	0.47	0.84	0.01	0.02
7000008610	DSS_074	4	2	0.51	0.83	0.01	0.02
7000008612	DSS_076	4	4	0.48	0.76	0.01	0.01
7000008613	DSS_077	4	4	0.46	0.87	0.01	0.01
7000008614	DSS_078	4	2	0.44	0.87	0.01	0.02
7000008615	DSS_079	4	2	0.47	0.90	0.01	0.02
7000008616	DSS_080	4	2	0.48	0.81	0.01	0.02
7000008619	DSS_083	4	2	0.59	0.86	0.01	0.02
7000008620	DSS_084	4	2	0.70	0.89	0.01	0.02
7000008621	DSS_085	4	4	0.49	0.89	0.01	0.01
7000008622	DSS_086	4	2	0.48	0.82	0.01	0.02
7000008624	DSS_088	4	2	0.47	0.88	0.01	0.02

10

20

30

40

【表 9 - 5】

株の ID	SNP	N、 24時間	N、 96時間	平均、 24時間 (A ₅₆₀)	平均、 96時間 (A ₅₆₀)	標準誤差、 24時間	標準誤差、 96時間
7000008625	DSS_089	4	4	0.45	0.89	0.01	0.01
7000008626	DSS_090	4	4	0.47	0.87	0.01	0.01
7000008627	DSS_091	4		0.46		0.01	
7000008629	DSS_093	4	4	0.50	0.87	0.01	0.01
7000008630	DSS_094	4	2	0.57	0.86	0.01	0.02
7000008634	DSS_098	4	2	0.53	0.85	0.01	0.02
7000008636	DSS_100	4		0.52		0.01	
7000008637	DSS_101	4	2	0.49	0.85	0.01	0.02
7000008640	DSS_104	4	2	0.51	0.84	0.01	0.02
7000008645	DSS_109	4		0.51		0.01	
7000008646	DSS_110	4	2	0.57	0.86	0.01	0.02
7000008648	DSS_112	4	2	0.54	0.86	0.01	0.02
7000008651	DSS_115	4		0.49		0.01	
7000008652	DSS_116	4	2	0.52	0.82	0.01	0.02
7000008653	DSS_117	4	2	0.50	0.84	0.01	0.02
7000008657	DSS_121	4	2	0.78	0.88	0.01	0.02

10

20

30

40

【表 9 - 6】

株の ID	SNP	N、 24時間	N、 96時間	平均、 24時間 (A ₅₆₀)	平均、 96時間 (A ₅₆₀)	標準誤差、 24時間	標準誤差、 96時間
7000008659	DSS_123	4		0.54		0.01	
7000008663	DSS_127	4		0.58		0.01	
7000008665	DSS_129	4		0.48		0.01	
7000008666	DSS_130	4		0.56		0.01	
7000008669	DSS_133	4		0.50		0.01	
7000008670	DSS_271	4	2	0.52	0.86	0.01	0.02
7000008672	DSS_273	4		0.56		0.01	
7000008677	DSS_278	2		0.46		0.01	
7000008678	DSS_279	4		0.55		0.01	
7000008681	DSS_282	4		0.51		0.01	
7000008683	DSS_284	4		0.59		0.01	
7000008684	DSS_285	4		0.51		0.01	
7000008685	DSS_286	4		0.56		0.01	
7000008687	DSS_288	4		0.46		0.01	
7000008688	DSS_289	4		0.57		0.01	
7000008689	DSS_290	4		0.47		0.01	

10

20

30

40

【表 9 - 7】

株の ID	SNP	N、 24時間	N、 96時間	平均、 24時間 (A ₅₆₀)	平均、 96時間 (A ₅₆₀)	標準誤差、 24時間	標準誤差、 96時間
7000008693	DSS_294	4	2	0.52	0.63	0.01	0.02
7000008696	DSS_297	4	2	0.52	0.86	0.01	0.02
7000008697	DSS_298	4		0.58		0.01	
7000008699	DSS_300	4		0.48		0.01	
7000008700	DSS_301	4		0.58		0.01	
7000008701	DSS_302	4		0.47		0.01	
7000008702	DSS_303	3		0.46		0.01	
7000008703	DSS_304	3		0.48		0.01	
7000008705	DSS_306	4	2	0.53	0.80	0.01	0.02
7000008708	DSS_309	4		0.56		0.01	
7000008709	DSS_310	4		0.56		0.01	
7000008711	DSS_312	4		0.55		0.01	
7000008712	DSS_313	4		0.51		0.01	
7000008718	DSS_319	4	2	0.50	0.82	0.01	0.02
7000008720	DSS_321	4		0.56		0.01	
7000008722	DSS_323	2	2	0.48	0.85	0.01	0.02

10

20

30

40

【表 9 - 8】

株の ID	SNP	N、 24時間	N、 96時間	平均、 24時間 (A ₅₆₀)	平均、 96時間 (A ₅₆₀)	標準誤差、 24時間	標準誤差、 96時間
7000008723	DSS_324	4		0.55		0.01	
7000008724	DSS_325	4		0.50		0.01	
7000008725	DSS_326	3		0.46		0.01	
7000008726	DSS_327	3		0.47		0.01	
7000008730	DSS_331	4		0.56		0.01	
7000008731	DSS_332	4	4	0.47	0.89	0.01	0.01
7000008732	DSS_333	4	4	0.47	0.87	0.01	0.01
7000008733	DSS_334	4		0.45		0.01	
7000008734	DSS_335	2		0.47		0.01	
7000008735	DSS_336	4		0.47		0.01	
7000008739	DSS_340	4		0.46		0.01	
7000008740	DSS_341	4	2	0.46	0.89	0.01	0.02
7000008741	DSS_342	4		0.56		0.01	
7000008742	DSS_343	4		0.55		0.01	
7000008743	DSS_344	4	4	0.48	0.87	0.01	0.01
7000008746	DSS_347	4	4	0.48	0.85	0.01	0.01

10

20

30

40

【表 9 - 9】

株の ID	SNP	N、 24時間	N、 96時間	平均、 24時間 (A ₅₆₀)	平均、 96時間 (A ₅₆₀)	標準誤差、 24時間	標準誤差、 96時間
7000008747	DSS_348	4	4	0.46	0.86	0.01	0.01
7000008749	DSS_350	4	2	0.29	0.74	0.01	0.02
7000008752	DSS_353	4	2	0.46	0.85	0.01	0.02
7000008753	DSS_354	4	4	0.45	0.87	0.01	0.01
7000008755	DSS_356	4	4	0.46	0.86	0.01	0.01
7000008756	DSS_357	4	4	0.46	0.86	0.01	0.01
7000008758	DSS_359	2	2	0.45	0.85	0.01	0.02
7000008760	DSS_361	4	2	0.46	0.84	0.01	0.02
7000008761	DSS_362	4		0.44		0.01	
7000008763	DSS_364	4		0.44		0.01	
7000008764	DSS_365	4		0.46		0.01	
7000008765	DSS_366	4		0.55		0.01	
7000008766	DSS_367	4		0.55		0.01	
7000008767	DSS_368	4	2	0.44	0.86	0.01	0.02
7000008770	DSS_371	4	2	0.47	0.88	0.01	0.02
7000008771	DSS_372	4	2	0.46	0.83	0.01	0.02

10

20

30

40

【表 9 - 1 0】

株の ID	SNP	N、 24時間	N、 96時間	平均、 24時間 (A ₅₆₀)	平均、 96時間 (A ₅₆₀)	標準誤差、 24時間	標準誤差、 96時間
7000008772	DSS_373	4	2	0.46	0.88	0.01	0.02
7000008774	DSS_375	4		0.45		0.01	
7000008776	DSS_377	4		0.45		0.01	
7000008777	DSS_378	4		0.57		0.01	
7000008778	DSS_379	4		0.54		0.01	
7000008779	DSS_380	4	2	0.46	0.87	0.01	0.02
7000008781	DSS_382	4	2	0.46	0.84	0.01	0.02
7000008782	DSS_383	4		0.48		0.01	
7000008783	DSS_384	4	2	0.47	0.82	0.01	0.02
7000008784	DSS_385	4	2	0.46	0.83	0.01	0.02
7000008786	DSS_387	3		0.43		0.01	
7000008787	DSS_388	3		0.47		0.01	
7000008788	DSS_389	4	2	0.46	0.89	0.01	0.02
7000008790	DSS_391	4		0.57		0.01	
7000008791	DSS_392	4		0.44		0.01	
7000008795	DSS_396	4	2	0.46	0.82	0.01	0.02

10

20

30

40

【表 9 - 1 1】

株の ID	SNP	N、 24時間	N、 96時間	平均、 24時間 (A ₅₆₀)	平均、 96時間 (A ₅₆₀)	標準誤差、 24時間	標準誤差、 96時間
7000008799	DSS_400	4		0.47		0.01	
7000008800	DSS_401	4	2	0.46	0.86	0.01	0.02
7000008801	DSS_402	4		0.54		0.01	
7000008805	DSS_406	4	2	0.47	0.85	0.01	0.02
7000008807	DSS_408	4		0.45		0.01	
7000008810	DSS_411	4	2	0.46	0.87	0.01	0.02
7000008812	DSS_413	3		0.47		0.01	
7000008813	DSS_414	4	2	0.45	0.84	0.01	0.02
7000008815	DSS_416	4	2	0.45	0.87	0.01	0.02
7000008816	DSS_417	4		0.46		0.01	
7000008818	DSS_419	4	2	0.47	0.84	0.01	0.02
7000008820	DSS_421	4	2	0.45	0.79	0.01	0.02
7000008821	DSS_422	4		0.44		0.01	

【0795】

S N Pスワップのこの第2のラウンドからの結果から、z w fプロモータースワップ変異を含む基準株において、基準株のリシンの収率および生産性を増加させることが可能ないくつかのS N Pが同定された（例えば、図39の右上部の隅にあるS N P 0 8 4およびS N P 1 2 1を参照されたい）。

【0796】

C. タンク培養のバリデーション

上記のH T Pのステップの間に同定した上位のS N Pを含有する株を、中型の試験発酵タンク中に培養した。手短に述べると、各株の小型の1 0 0 m lの培養物を、一晚成長させ、次いで、等しい播種量を有する試験発酵タンクに5リットルの培養物を播種するために使用した。播種物は、O D 6 0 0の測定値に従って同じ細胞密度を含有するように基準に合わせた（normalized）。

【0797】

結果として得られたタンク培養物は、培養を3日間進行させてから収集した。収率およ

び生産性の測定値を、発酵全体を通して種々の点でタンクから採取した試料中の基質および産物の力価から計算した。試料を、特定の低分子の濃度について、高圧液体クロマトグラフィーにより、適切な標準物質を使用して分析した。この実験についての結果を、下記の表 10 に要約し、図 40 に示す。

【表 10】

表10ーSNPスワップ微生物のタンクバリデーション

株	N	平均収率 (%)(産出されたリシンの g/ 消費されたグルコースの g)	標準 誤差	平均生産性 (g/L/時間)	標準 誤差
基準株	1	41.1502	0.5940 1	3.29377	0.2450 8
基準株 + zwf	7	48.2952	0.2245 1	2.73474	0.1000 5
基準株 + zwf + SNP121	2	50.325	0.4200 3	4.51397	0.1733
基準株 + zwf + pyc + lysA	5	52.191	0.2656 5	4.15269	0.1225 4

【0798】

小規模のハイスループット培養物により予測されるように、zwf プロモータースワップと SNP 121 との組合せを含む株についてのより大きなタンク培養物は、参照の基準株に対して、収率および生産性の有意な増加を示した。例えば、この株の生産性は、基準株の生産性、3.29 g/L/時間と比較して、4.5 g/L/時間に急増した（2 ラウンドの SNP スワップだけで、生産性が 37.0 % 増加した）。

【0799】

（実施例 4）

H T P ゲノム操作ー工業用微生物株を改良するための、プロモータースワップライブラリーのインプリメンテーション

これまでの実施例は、工業用株を再建するために、本開示の H T P 株改良プログラムの力を実証してきた。実施例 2 および 3 は、種々の基準株、中間株および工業用株における、現存する遺伝子の多様性を探索する、SNP スワップの技法およびライブラリーのインプリメンテーションについて記載した。

【0800】

この実施例は、本開示の P R O スワップの技法を使用する H T P 株改良プログラムの実施形態を示す。実施例 3 とは異なり、この実施例は、P R O スワップライブラリーの生成による変異のデノボ生成のための方法を教示する。

【0801】

A. プロモータースワッピングのための標的の同定

前述のように、プロモータースワッピングは、標的とするための「n 個」の遺伝子のセ

ットを選択するステップを含む、多段階ステップのプロセスである。

【0802】

この実施例では、発明者らは、潜在的な経路遺伝子23個の群を同定して、本開示のプロモーターラダー法により、モジュレートした（分子リシンを産生する例示的な代謝経路における、過剰発現に対して19個の遺伝子および下方制御に対して4+個の転用遺伝子（diverting gene））（図19を参照されたい）。

【0803】

B. プロモーターラダーの創製

プロモータースワップのプロセスのインプリメンテーションにおける別のステップは、「ラダー」として働く「x個」のプロモーターのセットの選択である。これらのプロモーターが複数のゲノム遺伝子座にわたって高度に変動する発現をもたらすことが示されている理想的であるが、それらが遺伝子の発現を何らかの様式で攪乱させることが唯一の要件である。

【0804】

特定の実施形態では、目的の標的遺伝子と関連がある天然の、生来のまたは野生型のプロモーターを同定し、次いで、前記プロモーターを変異させて、複数の変異プロモーター配列を得ることによって、これらのプロモーターラダーを生み出す。これらの変異プロモーターのそれぞれは、標的遺伝子の発現に対する効果について試験される。一部の実施形態では、編集されたプロモーターは、様々な条件にわたって発現活性について試験され、その結果、各プロモーターバリエーションの活性が記録され／特徴付けられ／アノテートされ、かつデータベース内に記憶される。結果として生じる編集されたプロモーターバリエーションは、引き続いて組織化されて、これらの発現の強度に基づいて配列された「ラダー」にされる（例えば、頂部付近に高度発現バリエーションおよび底部付近に減弱された発現を有し、したがって、用語「ラダー」をもたらす）。

【0805】

例示的な本実施形態では、発明者らは、図19中の同定した標的遺伝子のそれぞれについて、プロモーターラダー：ORFの組合せを創製した。

【0806】

C. ラダーからのプロモーターと標的遺伝子との関連性

プロモータースワップのプロセスのインプリメンテーションにおける別のステップは、種々の株のHTP操作であり、これらの株は、特定の標的遺伝子と関連があるプロモーターラダーからの所与のプロモーターを含む。

【0807】

生来のプロモーターが標的遺伝子nの前に存在し、その配列が分かっている場合には、生来のプロモーターを、ラダー中のxプロモーターのそれぞれでの置換えを実施することができる。生来のプロモーターが存在しないかまたはその配列が分かっていない場合には、遺伝子nの前に、ラダー中のxプロモーターのそれぞれの挿入を実施することができる。このようにして、株のライブラリーを構築し、この場合、ライブラリーの各メンバーが、n標的に作動可能に連結しているxプロモーターの一例であり、は、他の点では同一の遺伝子状況である（例えば、図20を参照されたい）。

【0808】

D. 株のHTPスクリーニング

プロモータースワップのプロセスにおける最後のステップは、上記のライブラリーにおける株のHTPスクリーニングである。得られた株のそれぞれが、n標的に連結しているxプロモーターの一例を表し、他の点では同一の遺伝的バックグラウンドである。

【0809】

発明者らは、1つまたは複数の測定基準に対するそれらの性能を特徴付けるシナリオにおいて、各株のHTPスクリーニングをインプリメントすることによって、所与の測定基準（例えば、目的の分子の産生の最適化）に関して、プロモーター／標的遺伝子のどの関連性が最も有益であるかを決定することができる。図20（プロモーターP1～P8の、

10

20

30

40

50

目的の遺伝子に対する効果)を参照されたい。

【0810】

図19～22に示す例示的な実施形態では、発明者らは、プロモータースワップのプロセスを利用して、リシンの産生を最適化した。上に記載したProスワップ方法の適用例を、下の実施例5に記載する。

【0811】

(実施例5)

HTPゲノム操作ーリシンの産生に関して、株性能を改善するためのPROスワップライブラリーのインプリメンテーション

以下のセクションは、実施例4に記載の、本開示のPROスワップHTP設計株改良プログラムツールの例示的なインプリメンテーションを示す。この実施例では、宿主細胞のリシンの収率を増加させるために、Corynebacterium株を、本開示のPROスワップ方法に供した。

10

【0812】

A. プロモータースワップ

プロモータースワップを、実施例4の記載の通り実施した。図19中のリシン生合成経路から選択した遺伝子を、プロモーターP1～P8を使用してプロモータースワップに対して標的化した。

【0813】

B. HTP操作およびハイスループットスクリーニング

20

プロモータースワップのHTP操作を、実施例1および3に記載の通り実施した。結果として得られたプロモータースワップ株のHTPスクリーニングを、実施例3に記載の通り実施した。全部で145個のPROスワップを実施した。実験の結果を、下記の表11に要約し、図41に示す。

【表 1 1 - 1】

表11ーリシンPROスワップライブラリーのHTPスクリーニング

株	プロモーター-標的	N	平均 (A ₅₆₀)	標準誤差	基準からの%収率変化
7000007713	Pcg1860-asd	8	0.84595	0.00689	3.927615
7000007736	Pcg0755-asd	4	0.84036	0.00974	3.240866
7000007805	Pcg0007_119-asd	8	0.82493	0.00689	1.345242
7000007828	Pcg3121-asd	8	0.8246	0.00689	1.3047
7000007759	Pcg0007_265-asd	8	0.81155	0.00689	-0.29853
7000007782	Pcg3381-asd	8	0.8102	0.00689	-0.46438
7000007712	Pcg1860-ask	8	0.83958	0.00689	3.14504
7000007735	Pcg0755-ask	8	0.81673	0.00689	0.337846
7000007827	Pcg3121-ask	8	0.81498	0.00689	0.122853
7000007804	Pcg0007_119-ask	8	0.81492	0.00689	0.115482

10

20

30

【表 1 1 - 2】

株	プロモーター-標的	N	平均 (A ₅₆₀)	標準誤差	基準からの%収率変化
7000007758	Pcg0007_265-ask	8	0.80381	0.00689	-1.24942
7000007781	Pcg3381-ask	8	0.80343	0.00689	-1.2961
7000007780	Pcg3381-aspB	8	0.84072	0.00689	3.285093
7000007803	Pcg0007_119- aspB	8	0.82106	0.00689	0.8698
7000007809	Pcg0007_119- cg0931	8	0.83446	0.00689	2.516032
7000007717	Pcg1860-cg0931	4	0.83129	0.00974	2.126588
7000007763	Pcg0007_265- cg0931	4	0.82628	0.00974	1.511094
7000007671	Pcg0007_39- cg0931	8	0.82554	0.00689	1.420182
7000007740	Pcg0755-cg0931	8	0.81921	0.00689	0.642522
7000007694	Pcg0007-cg0931	8	0.80444	0.00689	-1.17202
7000007691	Pcg0007-dapA	8	0.8299	0.00689	1.955822
7000007783	Pcg3381-dapA	8	0.80951	0.00689	-0.54915
7000007760	Pcg0007_265- dapA	8	0.76147	0.00689	-6.45102

10

20

30

40

【表 1 1 - 3】

株	プロモーター-標的	N	平均 (A ₅₆₀)	標準誤差	基準からの%収率変化
7000007806	Pcg0007_119- dapA	8	0.35394	0.00689	-56.5174
7000007761	Pcg0007_265- dapB	8	0.84157	0.00689	3.389518
7000007738	Pcg0755-dapB	4	0.84082	0.00974	3.297378
7000007692	Pcg0007-dapB	8	0.83088	0.00689	2.076218
7000007784	Pcg3381-dapB	8	0.82474	0.00689	1.3219
7000007715	Pcg1860-dapB	8	0.82232	0.00689	1.024595
7000007830	Pcg3121-dapB	8	0.81236	0.00689	-0.19902
7000007807	Pcg0007_119- dapB	4	0.69622	0.00974	-14.4672
7000007762	Pcg0007_265- dapD	8	0.84468	0.00689	3.771591
7000007808	Pcg0007_119- dapD	8	0.83869	0.00689	3.035701
7000007785	Pcg3381-dapD	8	0.83397	0.00689	2.455834
7000007670	Pcg0007_39- dapD	8	0.81698	0.00689	0.368559
7000007831	Pcg3121-dapD	4	0.8155	0.00974	0.186737

10

20

30

40

【表 1 1 - 4】

株	プロモーター-標的	N	平均 (A ₅₆₀)	標準誤差	基準からの%収率変化
7000007693	Pcg0007-dapD	8	0.8117	0.00689	-0.28011
7000007716	Pcg1860-dapD	8	0.79044	0.00689	-2.89196
7000007739	Pcg0755-dapD	8	0.78694	0.00689	-3.32195
7000007787	Pcg3381-dapE	8	0.83814	0.00689	2.968132
7000007833	Pcg3121-dapE	8	0.83721	0.00689	2.853878
7000007741	Pcg0755-dapE	8	0.83263	0.00689	2.291211
7000007810	Pcg0007_119- dapE	8	0.83169	0.00689	2.175729
7000007718	Pcg1860-dapE	8	0.81855	0.00689	0.561439
7000007672	Pcg0007_39- dapE	8	0.80932	0.00689	-0.5725
7000007765	Pcg0007_265- dapF	8	0.8327	0.00689	2.299811
7000007788	Pcg3381-dapF	8	0.82942	0.00689	1.896853
7000007811	Pcg0007_119- dapF	8	0.82926	0.00689	1.877196
7000007696	Pcg0007-dapF	8	0.82099	0.00689	0.861201
7000007719	Pcg1860-dapF	8	0.82067	0.00689	0.821888

10

20

30

40

【表 1 1 - 5】

株	プロモーター-標的	N	平均 (A ₅₆₀)	標準誤差	基準からの%収率変化
7000007673	Pcg0007_39- dapF	8	0.82062	0.00689	0.815745
7000007789	Pcg3381-ddh	8	0.84817	0.00689	4.200349
7000007835	Pcg3121-ddh	8	0.82141	0.00689	0.912799
7000007812	Pcg0007_119- ddh	8	0.82093	0.00689	0.853829
7000007674	Pcg0007_39-ddh	8	0.81494	0.00689	0.117939
7000007720	Pcg1860-ddh	8	0.81473	0.00689	0.09214
7000007766	Pcg0007_265- ddh	8	0.81427	0.00689	0.035627
7000007743	Pcg0755-ddh	8	0.80655	0.00689	-0.9128
7000007697	Pcg0007-ddh	8	0.80621	0.00689	-0.95457
7000007779	Pcg3381-fbp	8	0.85321	0.00689	4.819529
7000007802	Pcg0007_119-fbp	4	0.81425	0.00974	0.03317
7000007710	Pcg1860-fbp	4	0.40253	0.00974	-50.5479
7000007687	Pcg0007-fbp	8	0.14881	0.00689	-81.7182
7000007825	Pcg3121-fbp	4	0.12471	0.00974	-84.679
7000007733	Pcg0755-fbp	4	0.08217	0.00974	-89.9052

10

20

30

40

【表 1 1 - 6】

株	プロモーター-標的	N	平均 (A ₅₆₀)	標準誤差	基準からの%収率変化
7000007746	Pcg0755-hom	8	0.81925	0.00689	0.647436
7000007792	Pcg3381-hom	4	0.77674	0.00974	-4.57505
7000007723	Pcg1860-hom	8	0.71034	0.00689	-12.7325
7000007838	Pcg3121-hom	8	0.559	0.00689	-31.3251
7000007800	Pcg0007_119-icd	8	0.83236	0.00689	2.258041
7000007823	Pcg3121-icd	8	0.83155	0.00689	2.15853
7000007777	Pcg3381-icd	8	0.82844	0.00689	1.776456
7000007708	Pcg1860-icd	8	0.82384	0.00689	1.211332
7000007662	Pcg0007_39-icd	12	0.82008	0.00562	0.749404
7000007685	Pcg0007-icd	8	0.81257	0.00689	-0.17322
7000007754	Pcg0007_265-icd	4	0.81172	0.00974	-0.27765
7000007698	Pcg0007-lysA	4	0.8504	0.00974	4.474311
7000007675	Pcg0007_39- lysA	8	0.84414	0.00689	3.705251
7000007836	Pcg3121-lysA	4	0.83545	0.00974	2.637657
7000007767	Pcg0007_265- lysA	8	0.83249	0.00689	2.274012

10

20

30

40

【表 1 1 - 7】

株	プロモーター-標的	N	平均 (A ₅₆₀)	標準誤差	基準からの%収率変化
7000007813	Pcg0007_119-lysA	8	0.83096	0.00689	2.086046
7000007790	Pcg3381-lysA	8	0.8118	0.00689	-0.26782
7000007676	Pcg0007_39-lysE	8	0.84394	0.00689	3.68068
7000007699	Pcg0007-lysE	4	0.83393	0.00974	2.45092
7000007768	Pcg0007_265-lysE	8	0.83338	0.00689	2.383351
7000007837	Pcg3121-lysE	4	0.83199	0.00974	2.212585
7000007791	Pcg3381-lysE	8	0.81476	0.00689	0.095825
7000007814	Pcg0007_119-lysE	8	0.81315	0.00689	-0.10197
7000007775	Pcg3381-odx	8	0.82237	0.00689	1.030738
7000007752	Pcg0007_265-odx	8	0.81118	0.00689	-0.34399
7000007729	Pcg0755-odx	8	0.81103	0.00689	-0.36242
7000007683	Pcg0007-odx	8	0.80507	0.00689	-1.09462
7000007706	Pcg1860-odx	4	0.79332	0.00974	-2.53815
7000007660	Pcg0007_39-odx	8	0.79149	0.00689	-2.76297

10

20

30

40

【表 1 1 - 8】

株	プロモーター-標的	N	平均 (A ₅₆₀)	標準誤差	基準からの%収率変化
7000007798	Pcg0007_119- odx	8	0.77075	0.00689	-5.31094
7000007821	Pcg3121-odx	4	0.74788	0.00974	-8.12059
7000007822	Pcg3121-pck	8	0.85544	0.00689	5.093491
7000007776	Pcg3381-pck	8	0.8419	0.00689	3.43006
7000007799	Pcg0007_119- pck	8	0.83851	0.00689	3.013588
7000007753	Pcg0007_265- pck	8	0.82738	0.00689	1.646232
7000007730	Pcg0755-pck	4	0.81785	0.00974	0.475442
7000007661	Pcg0007_39-pck	8	0.80976	0.00689	-0.51844
7000007684	Pcg0007-pck	8	0.79007	0.00689	-2.93742
7000007707	Pcg1860-pck	8	0.71566	0.00689	-12.0789
7000007840	Pcg3121-pgi	4	1.01046	0.00974	24.13819
7000007817	Pcg0007_119-pgi	7	0.99238	0.00736	21.917
7000007794	Pcg3381-pgi	7	0.99008	0.00736	21.63444
7000007771	Pcg0007_265-pgi	8	0.94665	0.00689	16.29893
7000007725	Pcg1860-pgi	8	0.85515	0.00689	5.057864

10

20

30

40

【表 1 1 - 9】

株	プロモーター-標的	N	平均 (A ₅₆₀)	標準誤差	基準からの%収率変化
7000007702	Pcg0007-pgi	4	0.8056	0.00974	-1.02951
7000007658	Pcg0007_39-ppc	4	0.85221	0.00974	4.696676
7000007750	Pcg0007_265- ppc	8	0.84486	0.00689	3.793705
7000007727	Pcg0755-ppc	8	0.84166	0.00689	3.400575
7000007773	Pcg3381-ppc	4	0.82883	0.00974	1.824369
7000007796	Pcg0007_119- ppc	8	0.82433	0.00689	1.27153
7000007704	Pcg1860-ppc	8	0.81736	0.00689	0.415244
7000007819	Pcg3121-ppc	8	0.79898	0.00689	-1.8428
7000007732	Pcg0755-ptsG	8	0.84055	0.00689	3.264208
7000007709	Pcg1860-ptsG	8	0.81075	0.00689	-0.39682
7000007663	Pcg0007_39- ptsG	8	0.80065	0.00689	-1.63763
7000007778	Pcg3381-ptsG	8	0.23419	0.00689	-71.229
7000007801	Pcg0007_119- ptsG	8	0.17295	0.00689	-78.7525
7000007824	Pcg3121-ptsG	8	0.16035	0.00689	-80.3005
7000007705	Pcg1860-pyc	8	0.85143	0.00689	4.60085

10

20

30

40

【表 1 1 - 1 0】

株	プロモーター-標的	N	平均 (A ₅₆₀)	標準誤差	基準からの%収率変化
7000007728	Pcg0755-pyc	8	0.79803	0.00689	-1.95951
7000007659	Pcg0007_39-pyc	8	0.75539	0.00689	-7.19797
7000007751	Pcg0007_265- pyc	8	0.73664	0.00689	-9.50146
7000007682	Pcg0007-pyc	4	0.73142	0.00974	-10.1428
7000007774	Pcg3381-pyc	4	0.66667	0.00974	-18.0975
7000007797	Pcg0007_119- pyc	4	0.52498	0.00974	-35.5046
7000007820	Pcg3121-pyc	8	0.52235	0.00689	-35.8277
7000007841	Pcg3121-tkt	8	0.82565	0.00689	1.433696
7000007818	Pcg0007_119-tkt	8	0.81674	0.00689	0.339075
7000007749	Pcg0755-tkt	8	0.81496	0.00689	0.120396
7000007703	Pcg0007-tkt	4	0.76763	0.00974	-5.69424
7000007795	Pcg3381-tkt	8	0.72213	0.00689	-11.2841
7000007772	Pcg0007_265-tkt	8	0.68884	0.00689	-15.3738
7000007701	Pcg0007-zwf	4	0.95061	0.00974	16.78542
7000007747	Pcg0755-zwf	8	0.92595	0.00689	13.75587

10

20

30

40

【表 1 1 - 1 1】

株	プロモーター-標的	N	平均 (A ₅₆₀)	標準誤差	基準からの%収率変化
7000007770	Pcg0007_265-zwf	8	0.9029	0.00689	10.9241
7000007724	Pcg1860-zwf	8	0.79309	0.00689	-2.5664
7000007839	Pcg3121-zwf	4	0.13379	0.00974	-83.5635
7000000017	-	11 6	0.92115	0.00181	13.16617
7000006284	-	12 8	0.81398	0.00172	0
7000005754	-	64	0.79489	0.00243	-2.34527

10

20

【0 8 1 4】

プロモータースワップライブラリーのスクリーニングの結果は、可視化すると、測定している性能の測定基準と最も密接に相関する遺伝子標的を同定するのに役立つ。この場合、遺伝子標的 p g i、z w f、p p c、p c k、f b p および d d h は、プロモータースワップにより、基準株に対して、収率の大きなゲインがもたらされる遺伝子と同定された。

【0 8 1 5】

表 1 1 から選択した株を、上記の通り、小型のプレート中で再培養し、リシンの収率について試験した。この二次スクリーニングからの結果を、図 2 2 に示す。

30

【0 8 1 6】

(実施例 6)

上位性マッピング変異の有益な統合を予測するためのアルゴリズムツール

この実施例は、本開示の H T P 株改良プログラムの一部として利用する予測モデリング技法の実施形態を記載する。本開示は、(上に記載の、遺伝子設計ライブラリーの使用による)潜在的に有益な変異の最初の同定の後の、H T P 株改良の第 2、第 3、第 4 のおよびさらにそれらに続くラウンドにおける、有益な変異を統合する方法を教示する。一部の実施形態では、本開示は、変異の統合が、前記変異それぞれの個々の性能に基づき得ることを教示する。他の実施形態では、本開示は、2 つまたはそれ超の変異が、単一の宿主細胞中に統合する場合に、相加効果または相乗効果を示す可能性を予測するための方法を教示する。下記の実施例は、本開示の予測ツールの実施形態を示す。

40

【0 8 1 7】

実施例 3 および 5 の S N P スワップライブラリーおよびプロモータースワッピング (P R O スワップ) ライブラリーから選択した変異を分析して、宿主の株性能の改善をもたらす可能性が最も高い S N P / P R O スワップの組合せを同定した。

【0 8 1 8】

本開示の「上位性マッピング」のセクションの記載の通り、S N P スワップライブラリーの配列を、コサイン類似性マトリックスを使用して相互に比較した。分析の結果から、

50

各 SNP / PRO スワップの組合せについて、機能類似性スコアを得た。すべての SNP / PRO スワップの間の機能類似性の視覚的表現を、図 15 にヒートマップに示す。また、結果として得られた、機能類似性スコアを使用して、それぞれの SNP / PRO スワップ間の類似性の距離を示す樹状図も展開した（図 16 A）。

【0819】

同じまたは類似の機能群（すなわち、高い機能類似性を有する SNP / PRO スワップ）からの変異は、同じ機構により働く可能性が高く、したがって、組み合わせた場合に、全般的な宿主の性能に対して、負または中立の上位性を示す可能性が高い。対照的に、異なる機能群からの変異は、独立の機構により働く可能性がより高く、したがって、宿主の性能に対して、有益な相加的またはコンビナトリアルな効果をもたらす可能性が高い。

10

【0820】

上位性に対する生物学的経路の効果を示すために、多様な機能類似性を示す SNP スワップと PRO スワップとを組み合わせ、宿主株において試験した。実施例 1 の記載の通り、SNP / PRO スワップの 3 つの組合せ：i) P c g 0 0 0 7 : : z w f の PRO スワップ + P c g 1 8 6 0 : : p y c の PRO スワップ、ii) P c g 0 0 0 7 : : z w f の PRO スワップ + S N P 3 0 9、および iv) P c g 0 0 0 7 : : z w f の PRO スワップ + P c g 0 0 0 7 : : l y s A の PRO スワップを、C o r y n e b a c t e r i u m g l u t a m i c u m のゲノム中に操作した（機能類似性の関係については、図 15 および図 16 A を参照されたい）。

【0821】

20

SNP / PRO スワップの組合せを含有するそれぞれの宿主細胞の性能を、実施例 3 の記載の通り試験し、z w f の PRO スワップのみを含有する対照の宿主細胞の性能と比較した。下記の表 1 2 および 1 3 に、それぞれの株について、宿主細胞の収率（9 6 時間における測定）および生産性（2 4 時間における測定）の結果を要約する。

【表 1 2】

表12ー上位性マッピング実験の24時間時のリシンの蓄積

SNP/ PRO スワップ	平均リシン (A ₅₆₀)	標準偏差
6318 (zwf)	0.51	0.03
8126 (zwf+ lysA)	0.88	0.06
8156 (zwf+ pyc)	0.53	0.01
8708 (zwf+ SNP 309)	0.56	0.00

30

40

【表 1 3】

表13ー上位性マッピング実験の96時間時のリシンの蓄積

SNP/ PRO スワップ	平均リシン (A ₅₆₀)	標準偏差
6318 (zwf)	0.83	0.01
8126 (zwf + lysA)	0.94	0.02
8156 (zwf+ pyc)	0.83	0.06

10

【0 8 2 2】

各 SNP / PRO スワップの組合せについての宿主の収率による性能の結果はまた、図 1 6 B にも示す。2 4 時間時の測定および 9 6 時間時の測定の両方で、より低い機能類似性を示す SNP / PRO スワップを組み合わせた宿主株が、組み合わせた SNP がより高い機能類似性を示す株より性能が優れていた。

20

【0 8 2 3】

したがって、上位性マッピングの手順は、設計した遺伝子変化の有効なおよび／またはプラスの統合について予測／プログラミング／情報提供を行うのに有用である。上位性マッピングの手順による分析的洞察から、微生物株の開発のその後のラウンドを導くことができる予測規則のセットの創製が可能になる。上位性ライブラリーから得られる予測のための洞察は、微生物のタイプおよび標的分子のタイプにわたって使用することができる。

【0 8 2 4】

(実施例 7)

H T P ゲノム操作ー P r o スワップ変異の統合および多因子のコンビナトリアル試験

30

これまで実施例は、少数の事前選択した PRO スワップ変異と SNP スワップライブラリーとを統合するための方法を示してきた（実施例 3）。他の実施例は、相加的または相乗的に有益な宿主細胞の特性をもたらす可能性が最も高い変異の統合を選択するための上位性による方法を示してきた（実施例 6）。この実施例は、本開示の H T P 方法により、複数の遺伝子／遺伝子設計ライブラリーの組合せ（例えば、PRO スワップライブラリー × SNP ライブラリー、または PRO スワップライブラリー内の組合せ）のコンビナトリアルな統合により創製される大きな解空間を有効に探索することが可能になることを示す。

【0 8 2 5】

本開示の H T P 株改良法のこの例示的な適用例では、実施例 5 で宿主の性能に対してプラス効果を有することが同定されているプロモータースワップを元々の PRO スワップライブラリーと二次の組合せで統合する。PRO スワップ変異を統合する決定は、収率または生産性に対するそれぞれの変異の全般的な効果、および 2 つの変異の組合せが相加的または相乗的な効果をもたらす可能性に基づいた。

40

【0 8 2 6】

例えば、出願人らは、実施例 6 の上位性マッピングの結果に基づいて、P c g 0 0 0 7 : : z w f および P c g 0 0 0 7 : : l y s A の組み合わせを選択したことについて言及する。

【0 8 2 7】

A. P R O スワップ株操作のための統合のラウンド

50

先の実施例 1 の記載の通り、株を形質転換した。手短に述べると、所望の P R O スワップ変異を 1 つすでに含有する株をもう一度、第 2 の所望の P R O スワップ変異を用いて形質転換した。全部で 1 4 5 個の、実施例 5 からの試験済の P R O スワップを、5 3 個の第 2 のラウンドの統合株へと統合して、それらはそれぞれ、有益な相加効果または相乗効果を示すことが期待される 2 つの P R O スワップ変異を含んだ。

【 0 8 2 8 】

結果として得られた第 2 のラウンドの株を、実施例 3 の記載の通りもう一度スクリーニングした。この実験からの結果を、下記の表 1 4 に要約し、図 1 1 に示す。

【表 1 4 - 1】

表14－第2のラウンドの統合されたりシンPROスワップライブラリーのHTPスクリーニング

株の ID	数	PRO スワップ 1	PRO スワップ 2	平均収率 (A ₅₆₀)	標準偏差
7000008489	4	Pcg0007-lysA	Pcg3121-pgi	1.17333	0.020121
7000008530	8	Pcg1860-pyc	Pcg0007-zwf	1.13144	0.030023
7000008491	7	Pcg0007-lysA	Pcg0007-zwf	1.09836	0.028609
7000008504	8	Pcg3121-pck	Pcg0007-zwf	1.09832	0.021939
7000008517	8	Pcg0007_39- ppc	Pcg0007-zwf	1.09502	0.030777
7000008502	4	Pcg3121-pck	Pcg3121-pgi	1.09366	0.075854
7000008478	4	Pcg3381-ddh	Pcg0007-zwf	1.08893	0.025505
7000008465	4	Pcg0007_265- dapB	Pcg0007-zwf	1.08617	0.025231
7000008535	8	Pcg0007-zwf	Pcg3121-pgi	1.06261	0.019757
7000008476	6	Pcg3381-ddh	Pcg3121-pgi	1.04808	0.084307
7000008510	8	Pcg3121-pgi	Pcg1860-pyc	1.04112	0.021087
7000008525	8	Pcg1860-pyc	Pcg0007_265- dapB	1.0319	0.034045
7000008527	8	Pcg1860-pyc	Pcg0007-lysA	1.02278	0.043549
7000008452	5	Pcg1860-asd	Pcg0007-zwf	1.02029	0.051663

10

20

30

40

【表 1 4 - 2】

株の ID	数	PRO スワップ 1	PRO スワップ 2	平均収率 (A ₅₆₀)	標準偏差
7000008463	4	Pcg0007_265- dapB	Pcg3121-pgi	1.00511	0.031604
7000008524	8	Pcg1860-pyc	Pcg1860-asd	1.00092	0.026355
7000008458	4	Pcg3381-aspB	Pcg1860-pyc	1.00043	0.020083
7000008484	8	Pcg3381-fbp	Pcg1860-pyc	0.99686	0.061364
7000008474	8	Pcg3381-ddh	Pcg3381-fbp	0.99628	0.019733
7000008522	8	Pcg0755-ptsG	Pcg3121-pgi	0.99298	0.066021
7000008528	8	Pcg1860-pyc	Pcg3121-pck	0.99129	0.021561
7000008450	4	Pcg1860-asd	Pcg3121-pgi	0.98262	0.003107
7000008448	8	Pcg1860-asd	Pcg3381-fbp	0.97814	0.022285
7000008494	8	Pcg0007_39- lysE	Pcg3381-fbp	0.97407	0.027018
7000008481	8	Pcg3381-fbp	Pcg0007-lysA	0.9694	0.029315
7000008497	8	Pcg0007_39- lysE	Pcg1860-pyc	0.9678	0.028569
7000008507	8	Pcg3121-pgi	Pcg3381-fbp	0.96358	0.035078
7000008501	8	Pcg3121-pck	Pcg0007-lysA	0.96144	0.018665

10

20

30

40

【表 1 4 - 3】

株の ID	数	PRO スワップ 1	PRO スワップ 2	平均収率 (A ₅₆₀)	標準偏差
7000008486	8	Pcg0007-lysA	Pcg0007_265- dapB	0.94523	0.017578
7000008459	8	Pcg0007_265- dapB	Pcg1860-asd	0.94462	0.023847
7000008506	2	Pcg3121-pgi	Pcg0007_265- dapD	0.94345	0.014014
7000008487	8	Pcg0007-lysA	Pcg3381-ddh	0.94249	0.009684
7000008498	8	Pcg3121-pck	Pcg1860-asd	0.94154	0.016802
7000008485	8	Pcg0007-lysA	Pcg1860-asd	0.94135	0.013578
7000008499	8	Pcg3121-pck	Pcg0007_265- dapB	0.93805	0.013317
7000008472	8	Pcg3381-ddh	Pcg1860-asd	0.93716	0.012472
7000008511	8	Pcg0007_39- ppc	Pcg1860-asd	0.93673	0.015697
7000008514	8	Pcg0007_39- ppc	Pcg0007-lysA	0.93668	0.027204
7000008473	8	Pcg3381-ddh	Pcg0007_265- dapB	0.93582	0.030377

10

20

30

40

【表 1 4 - 4】

株の ID	数	PRO スワップ 1	PRO スワップ 2	平均収率 (A ₅₆₀)	標準偏差
7000008461	7	Pcg0007_265- dapB	Pcg3381-fbp	0.93498	0.037862
7000008512	8	Pcg0007_39- ppc	Pcg0007_265- dapB	0.93033	0.017521
7000008456	8	Pcg3381-aspB	Pcg3121-pck	0.92544	0.020075
7000008460	8	Pcg0007_265- dapB	Pcg0007_265- dapD	0.91723	0.009508
7000008492	8	Pcg0007_39- lysE	Pcg3381-aspB	0.91165	0.012988
7000008493	8	Pcg0007_39- lysE	Pcg0007_265- dapD	0.90609	0.031968
7000008453	8	Pcg3381-aspB	Pcg0007_265- dapB	0.90338	0.013228
7000008447	8	Pcg1860-asd	Pcg0007_265- dapD	0.89886	0.028896
7000008455	8	Pcg3381-aspB	Pcg0007-lysA	0.89531	0.027108
7000008454	6	Pcg3381-aspB	Pcg3381-ddh	0.87816	0.025807
7000008523	8	Pcg0755-ptsG	Pcg1860-pyc	0.87693	0.030322
7000008520	8	Pcg0755-ptsG	Pcg3381-fbp	0.87656	0.018452

10

20

30

40

【表 1 4 - 5】

株の ID	数	PRO スワップ 1	PRO スワップ 2	平均収率 (A ₅₆₀)	標準偏差
7000008533	4	Pcg0007-zwf	Pcg3381-fbp	0.84584	0.017012
7000008519	8	Pcg0755-ptsG	Pcg0007_265- dapD	0.84196	0.025747

10

【0 8 2 9】

上位性モデルにより予測されたように、P c g 0 0 0 7 : : z w f 変異および P c g 0 0 0 7 : : l y s A 変異を含む、第 2 のラウンドの P R O スワップ株は、最高の収率改善のうちの 1 つを示し、収率を、P c g 0 0 0 7 : : l y s A 単独に対してほぼ 3 0 % 改善し、基準株に対して 3 5 . 5 % 改善した（図 1 1 上の円で囲んだデータ点を参照されたい）。

【0 8 3 0】

20

単一の変異および統合した二重の変異の解空間を探索するための H T P 方法はまた、第 3、第 4 およびその後の変異の統合にも適用することができる。また、例えば、開示する、z w f、p y c および l y s a に対応する 3 つの変化の統合株にも注目されたい；この株は、上記の表 1 4 に示す 2 つの変化の統合において同定した上位のヒットの中から作製し、これらのヒットは、本開示の上位性による方法によって同定されたものである。この 3 つの変化の統合株について、タンク中で妥当性をさらにバリデートしたところ、親、または親 + z w f と比較して有意に改善されていた（上記の表 1 0 および図 4 0 を参照されたい）。

【0 8 3 1】

（実施例 8）

30

H T P ゲノム操作—工業用宿主株を改善するための、ターミネーターライブラリーのインプリメンテーション

本実施例では、本開示の H T P 方法を、S T O P スワップを含む追加の H T P 遺伝子設計ライブラリーに適用する。実施例は、本開示によれば、（例えば、P R O スワップ、S N P スワップ、S T O P スワップ等の）基本の遺伝子設計ライブラリーからのエレメントを組み合わせて、より複雑な遺伝子設計ライブラリー（例えば、プロモーターおよびターミネーターの両方が組み入れられている P R O - S T O P のスワップライブラリー）を生み出すことができることをさらに示す。本開示は、一部の実施形態では、これまでに開示した遺伝子設計ライブラリーのうちのいずれかの組合せから得られるライブラリーを含めた、一切の可能な遺伝子設計ライブラリーを教示する。

40

【0 8 3 2】

この実施例では、小規模の実験を実施して、遺伝子の発現に対する本発明の S T O P スワップ方法の効果を実証した。下に記載するように、本開示のターミネーター T 1 ~ T 8 と、C o r y n e b a c t e r i u m g l u t a m i c u m の 2 つの生来のプロモーターのうちの 1 つとの対を形成し、蛍光タンパク質の発現に影響を及ぼすそれらの能力について分析した。

【0 8 3 3】

A. D N A 構築物のアセンブリー

ターミネーター T 1 ~ T 8 と C o r y n e b a c t e r i u m g l u t a m i c u m の 2 つの生来のプロモーターのうちの 1 つ（例えば、P c g 0 0 0 7 または P c g 0 0 4

50

7) との、黄色蛍光タンパク質 (YFP) を発現する対を形成した。DNA の増幅およびアセンブリーを促進するために、プロモーター-YFP-ターミネーターの最終の配列を、2つの部分に分けて合成した。第1の部分は、(5' から 3' へ) i) ベクターのホモロジーアーム、ii) 選択されたプロモーター、および iii) YFP 遺伝子の 2/3 をコードした。第2の部分は、(5' から 3' へ) iv) YFP 遺伝子の次の 2/3、v) 選択されたターミネーター、および vi) 第2の、ベクターのホモロジーアームをコードした。各部分を、合成オリゴヌクレオチドを使用して増幅し、ゲル上で精製した。酵母相同組換えを使用して、ゲル上で精製したアンプリコンをベクターの骨格とアセンブルした。

【0834】

B. E. coli 中へのアセンブルされたクロンの形質転換

正確にアセンブルされたクロンを同定し、Corynebacterium の形質転換のためのベクター-DNA を増幅するために、プロモーター-YFP-ターミネーターの配列を含有するベクターのそれぞれを個々に、E. coli 中へ形質転換した。正確にアセンブルされたベクターを、制限酵素消化およびサンガーシーケンシングにより確認した。陽性クロンを、後に使用するために -20 °C で保存した。

【0835】

C. Corynebacterium 中へのアセンブルされたクロンの形質転換

Corynebacterium glutamicum 宿主細胞中へ、ベクターの検証したクロンを個々に、電気穿孔により形質転換した。各ベクターを、異種の黄色蛍光タンパク質の発現を可能にするが、宿主細胞にとって有害にならないように経験的に決定された Corynebacterium glutamicum ゲノム内の中立の組み込み部位に組み込まれるように設計した。組み込みを促進するために、発現ベクターは、所望の組み込み部位に相同な約 2 kbp の配列 (すなわち、ホモロジーアーム) をさらに含み、それによって、上に記載した各遺伝子カセットが、ホモロジーアーム (homology arm) の下流に挿入された。ゲノム中への組み込みは、シングルクロスオーバーによる組み込みにより行われた。次いで、形質転換した Corynebacterium を、正確な組み込みについて PCR により試験した。各遺伝子構築物について実施した形質転換のそれぞれについて、このプロセスを繰り返した。

【0836】

D. Corynebacterium 中の個々のターミネーター構築物の評価

次いで、プロモーター-YFP-ターミネーターの構築物を含有する各 Corynebacterium 形質転換体の表現型を、発現を評価するために、培地の2つのタイプ (ブレインハートインフュージョン-BHI 培地およびHTP 試験培地) 中で、2つの時点において試験した。手短に述べると、PCR により確認した形質転換体 4~6 つを選び、96 ウェルフォーマット中の選択培地中で培養した。次いで、最初の培養物を、選択 BHI 培地または選択種培地中に分割した。48 時間時に、種培地中の培養物を、選択 HTP 試験培地または BHI 培地中に播種し、2つの時点において分析し、これらの時点は、成長曲線の異なる部分を表した。HTP 試験培地の培養物の場合の時点は、播種してから 48 および 96 時間後であった。選択 BHI 培地中の培養物は、播種してから 48 および 72 時間後に分析した。

【0837】

培養物の分析は、ベンチトップ型フローサイトメーターを使用して実施した。手短に述べると、培養物を、200 µl のリン酸緩衝食塩水 (PBS) 中に 1:100 で希釈した。各培養物に関して、3000~5000 個の個別事象 (すなわち、細胞) の間で、黄色の蛍光について分析した。ベンチトップ型フローサイトメーターが、各「事象」の黄色の蛍光のヒストグラムをプロットし、各ウェル内の蛍光の中央値を計算する。図 36 は、(4~6 つの生物学的複製物にわたる) 各構築物についての蛍光の中央値の平均を示す。エラーバーは、各データ点の 95 % 信頼区間を示す。条件 A~D はそれぞれ、単一の培地および単一の時点を示す。したがって、条件 A および条件 B は、BHI 培地についての 2 つ

10

20

30

40

50

の時点を表し、一方、C点およびD点は、H T P試験培地についての2つの時点を表す。任意単位（例えば、A U）は、ベンチトップ型フローサイトメーターが記録した蛍光の中央値を表すことに留意されたい。

【0838】

結果から、S T O Pスワップ遺伝子設計ライブラリーのターミネーター1～8は、Y F P発現を連続的な範囲でもたらしことが示されている。したがって、これらのターミネーターにより、ターミネーターラダーが形成され、このラダーは、本開示のH T P方法に従って、後の遺伝子設計ライブラリー中にインプリメントすることができる。

【0839】

（実施例9）

H T Pツールセットと従来のU V変異との比較

この実施例は、本開示のH T P遺伝子設計ライブラリーが、従来の変異による株改良プログラムよりも有益であることを実証する。本明細書のこの部分における実験では、本開示のH T P方法によりなしとげる表現型の改善について、改善の大きさおよびスピードを、従来のU V変異誘発に対して定量する。

【0840】

本開示は、宿主細胞の株改良プログラムを加速させるための新しい方法を教示する。一部の実施形態では、本開示のH T P株改良プログラムは、遺伝子の攪乱を生成し、同定するH T Pツールセットの能力に依存する。本発明者らは、本開示のプロモータースワップ技法と従来のU V変異アプローチとを比較する小型の平行追跡型の株改良プログラムを実施することによって、H T Pツールセットの利益を定量することを試みた。

【0841】

U Vおよびプロモータースワップの両方による遺伝子の攪乱についての出発点として、目的の生化学的代謝産物を産生する参照基準株を選んだ。

【0842】

A. U V変異

基準株の培養物を、10のO D₆₀₀に基準に合わせたO Dである培養物においてB H I培地中で成長させた。この培養物を、無菌のペトリ皿中にアリコートし、小型の磁気攪拌棒を使用してかき混ぜた。次いで、254nmの波長のU Vトランスイルミネーターを、U V曝露の5および9分時に培養物上で反転させ、一定分量を採取した。これらの試料を10倍段階希釈して、各希釈物を、B H I培地のQ - t r a y上に蒔いた。自動化コロニーピッキング機器を使用して、これらのQ - t r a yから、各U V曝露点からのコロニーをおよそ2500個選び取り、下記の通り性能を評価した。

【0843】

B. プロモータースワップ

15個の遺伝子標的について、基準株中に、P R Oスワップ構築物を、表1に記載したP 1、P 3、P 4およびP 8から選択されるプロモーターの全部またはサブセットのいずれかを使用して生成した。目的の産物の生合成における最後のステップは、律速補助因子の可能性があるS - アデノシルメチオニンを利用するO - メチルトランスフェラーゼ酵素によって触媒される。したがって、P R Oスワップのための遺伝子標的は、それらがこの補助因子または上流の代謝産物の生合成に直接関与するということに基づいて選択された。

【0844】

C. U Vおよびプロモータースワップライブラリーの評価

この実施例のために開発した各C o r y n e b a c t e r i u m株の表現型を、選択された生体分子を産生するその能力について試験した。手短に述べると、各P R Oスワップ株からの、配列が確認されたコロニー4～6つ、および各U V株についての単一のコロニーを選び、液体産生培地中、96ウェルフォーマット中の選択培地中で増殖した。

【0845】

96ウェルマイクロウェルプレート中でバイオマスを増殖させた後に、細胞塊を、96

10

20

30

40

50

ウェルマイクロウェルプレート中の基質を含有する発酵培地に添加し、生物変換を24時間進行させた。各株について、24時間時に採取した試料から、産物の力価を、高速液体クロマトグラフィーを使用して決定した。各遺伝子の攪乱（UVおよびPROスワップ）についての力価の結果を分析した。各複製物についての結果を平均し、前記株の全般的な性能を表すために割り当てた。次いで、基準株の収率に対する比として表現する、測定した収率に対する各変異の効果に基づいて、株をカテゴリーに分けた。

【0846】

図37に、この実験の結果を要約し、i) 収率を変化させなかった、ii) 1.2~1.4倍の収率の改善をもたらした、iii) 1.4~1.6倍の収率の改善をもたらした、iv) 1.6~1.8倍の収率の改善をもたらした、またはv) 1.8~2倍の収率の改善をもたらした各株改良の技法について、株の数を提示する。

10

【0847】

結果は、本開示のHTPツールセットは、従来のUV変異誘発アプローチに対して有益であることを示している。例えば、PROスワップ株は、収率のプラスの変化をより高い率で示し、したがって、株を有意に改良することができる変異を提供する可能性がより高かったことを、図37の結果は実証している。最も特筆すべきは、PROスワップライブラリーにおいては、1.6、1.8および2倍の増加を示す優れた改良株の発生率が高く、UVライブラリーにおいては、改善がほとんど~全く同定されなかったことである。

【0848】

また、結果により、本開示のPROスワップ方法の加速された改善速度が強調されることから、結果は重要である。実際に、PROスワップライブラリーについての結果は、100個未満の、プロモーター：遺伝子の攪乱に基づき、一方、UV変異の結果は、4,000個超の明確に異なる変異体株のスクリーニングを含んだ。したがって、本開示の方法は、性能における大きなゲインを株に付与することが可能な遺伝子の攪乱を同定する前にスクリーニングしなければならない変異体の数を劇的に低下させる。

20

【0849】

(実施例10)

HTPゲノム操作—*Escherichia coli*の株性能を改善するための、トランスポゾン変異誘発ライブラリーのインプリメンテーション

これまでの実施例は、*Corynebacterium*に対するHTP株改良プログラムの適用を示している。この実施例は、同じ技法を*E. coli*細胞に適用できることを実証する。

30

【0850】

この実施例は、株改良の目的で、*Escherichia coli*ランダム株ライブラリーを生成するための、トランスポゾン変異誘発の適用について記載する。これらの株ライブラリーをトリプトファンの収率などの所望の表現型に対してスクリーニングして、性能が改善されたバリエーションを同定することができる。

【0851】

本開示は、*Escherichia coli*においてEZ-Tn5トランスポゾンシステム（*Epicer Bio*）を適用することによって変異体のライブラリーを生成するための方法について記載する。EZ-Tn5トランスポザーゼを、モザイク型エレメント配列が隣接するペイロードDNAと一緒にインキュベートする。インキュベーションすると、EZ-Tn5トランスポザーゼがDNAと複合体化して、トランスポソームを形成する。次いで、DNA/タンパク質トランスポソーム複合体を電気穿孔によって*Escherichia coli*中に形質転換する。EZ-Tn5トランスポザーゼは、ペイロードDNAの*Escherichia coli*ゲノムへのランダムな組込みを触媒し、したがって、株バリエーションのランダムなライブラリーを生じさせる。

40

【0852】

ペイロードDNAの特定の配列をさらに変動させて、標的ゲノムへのトランスポゾン挿入の機能喪失（LoF）または機能獲得（GoF）効果のいずれかに偏向させることがで

50

きる。

【0853】

L o F は、抗生物質選択マーカーをDNAペイロードに含めることによって達成することができる。抗生物質マーカーにより、生産的トランスポゾン挿入を伴う細胞を選択することが可能になる。DNAペイロードの挿入により、それが挿入されたDNAの機能を、これに限定されないが、破壊された遺伝子の翻訳を妨げる、オープンリーディングフレームの破壊を含めた様々なやり方で破壊することができる。

【0854】

G o F は、抗生物質マーカーおよび強力なプロモーターをDNAペイロードに含めることによって達成することができる。抗生物質マーカーにより、生産的トランスポゾン挿入を伴う細胞を選択することが可能になる。DNAペイロードの挿入により、挿入部位の近位の遺伝子の発現を、強力なプロモーターの作用を通じて増大させることができる。

【0855】

L o F DNAペイロードまたはG o F DNAペイロードのいずれも、マーカー再利用、したがって、さらなるラウンドの操作を可能にするために、選択マーカーに加えて対抗選択マーカーをさらに含有し得る。

【0856】

上記のトランスポゾン変異誘発方法によって生成された株バリエーションのライブラリーを所望の表現型に対してスクリーニングすることができる。株をハイスループットに培養し、試験して、親株と比べて所望の表現型が改善された株を同定することができる。

【0857】

改善された株バリエーションを周期的操作の追加のラウンドに供して、トリプトファンの収率などの所望の表現型をさらに改善することができる。追加のラウンドの操作は、トランスポゾン変異誘発またはSNPSWP、PROSWPなどの他のライブラリータイプ、またはランダム変異誘発と一貫し得る。改善された株を、改善された表現型を呈する他の株バリエーションと統合して、別個の有益な変異の相加効果によってさらに改良された株を生じさせることもできる。

【0858】

本明細書に記載の方法により、周期的操作でスクリーニングするための高品質のライブラリーを構築するための費用が減少する。トランスポゾン変異誘発を*Escherichia coli*に適用することにより、何千もの全ゲノムL o FまたはG o F変異体を単一反応で生成することが可能になる。代替の方法は、何千もの割り当てられたプラスミドを構築して、株を単一のクロスオーバー相同組換え（SCHR）によって操作することである。別の代替の方法は、何千もの割り当てられた直鎖状断片を構築して、株をラムダレッドリコンビニアリングによって操作することである。どちらの方法も、組換えを標的ゲノム上の特定の場所に向きつけるための意図されたペイロードDNAおよび配列相同性を含有する各変異体についてユニークなDNA断片を生成することが必要になるので、高価である。トランスポゾン変異誘発では、単一のDNAペイロードを使用し、標的ゲノムへのランダムな組込みによって多様性を生じさせる。

【0859】

（実施例11）

HTPゲノム操作—*Escherichia coli*におけるHTPゲノム操作に使用するためのベクター骨格の生成

この実施例は、*Escherichia coli*におけるリコンビニアリングのためのHTPゲノム操作に使用するためのベクターであって、効率的な形質転換およびプラスミド組込みを付与するベクターの生成について記載する。

【0860】

ベクター1（核酸配列番号214）を生成した。これは、R6K複製起点、SacB遺伝子、対抗選択マーカーとしてPheS遺伝子およびURA3酵母栄養要求性マーカーを含む。効率および単一クロスオーバー相同組換えを改善するために、ベクター1の骨格を

、表15のエレメントを含有するように改変し、その結果、ベクター2（核酸配列番号215）が生じた。図55に示されているプラスミドマップは、ベクター1の一般的なコンポーネントを示す。ベクター2では、ランダムなインスレーター配列であるインスレーター1（配列番号218）およびインスレーター2（配列番号219）を相同アームに隣接して付加し、ターミネーター配列T1（核酸配列番号220；Oroszら、Eur J Biochem、1991年11月1日；201巻（3号）：653～9頁を参照されたい）およびB0015（核酸配列番号221）を付加して、ゲノムの挿入部位における転写リードスルーを排除した。図56に示されているプラスミドマップは、ベクター2の一般的なコンポーネントを示す。

【0861】

ベクターまたはプラスミドの有用性をノックアウト実験で試験した。要約すると、E. coliをLBブロス中に播種し、37℃で振とうしながら8時間にわたって増殖させた。その後、終夜培養物の一定分量を使用して、より大きな体積のLBブロスに播種し、18℃で振とうしながら16時間にわたって増殖させた。形質転換のために、試験プラスミド100～400ngをコンピテント細胞に添加し、電気穿孔によって形質転換を行った。細胞をSOC培地中に回収し、37℃で3時間インキュベートした後、カナマイシンを伴うLB寒天に蒔いた。プレートを37℃でインキュベートして、試験プラスミドを有するコロニーを増殖させた。

【0862】

ノックアウトさせる遺伝子標的はE. coli aroA遺伝子であった。したがって、E. coli aroA遺伝子に対する相同アームを、それぞれベクター1の骨格（バージョン1）およびベクター2の骨格（バージョン2）に、相同アームがカナマイシン耐性遺伝子に隣接して、E. coli宿主細胞における単一クロスオーバー相同組換えが可能になるように挿入することによって「バージョン1におけるaroA-KO」（すなわち、ベクター1）および「バージョン2におけるaroA-KO」（すなわち、ベクター2）の試験プラスミドを構築した。これらの試験プラスミドの形質転換およびカナマイシンでの選択により、「バージョン2におけるaroA-KO」が形質転換およびプラスミド組込みの効率の改善を示すことが検証された（図53）。

【0863】

ベクター2におけるベクター1骨格のさらなる改変により、PheS配列を付加することによって、スクロースおよび4クロロフェニルアラニンを含有する培地での効率的な対抗選択が可能になった（表15）。ベクター2におけるPheSプロモーター配列は、KincadeおよびdeHasethにより同定されたλファージPLプロモーター（KinacadeおよびdeHaseth Gene. 1991年1月2日；97巻（1号）：7～12頁を参照されたい）、そのすぐ後の、iGEMに由来するB0032と呼ばれるRBS配列からなることに留意すべきである。さらに、ベクター2では、sacB遺伝子のプロモーター配列をP5-MCD2プロモーター（Mutalikら、Nat Methods、2013年4月；10巻（4号）：354～60頁）および追加のATGを含有するプロモーターで置き換えた。この改変により、バージョン2骨格が組み込まれた株を用いたスクロースでの効率的な対抗選択が可能になった。ベクター3を生成するために、骨格ベクター1におけるC. glutamicum pheS*遺伝子のプロモーター配列およびCDSを新しいプロモーター配列およびCDS、具体的には必要な変異を含有する、生来のE. coli pheSのコドン最適化バージョンで置き換えた（T251A/A294G、Miyazaki, K. Biotechniques、2015年2月1日；58巻（2号）：86～8頁を参照されたい）（表15）。この改変により、ベクター3骨格が組み込まれた株を用いた4クロロフェニルアラニンでの対抗選択の改善が可能になった。図57に示されているプラスミドマップは、ベクター3の一般的なコンポーネントを示す。

【0864】

HTPゲノム操作のために使用される骨格は、プラスミドアセンブリーにおいて使用するための様々な酵母選択マーカーを含有し得る。本開示では、ベクター3骨格の改変によ

10

20

30

40

50

り、U R A 3 酵母選択マーカーをT R P 1 マーカーで置き換えてベクター 4 を生じさせた。図 5 8 に示されているプラスミドマップは、ベクター 4 の一般的なコンポーネントを示す。

【0865】

(実施例 12)

H T P ゲノム操作—工業用微生物株の改良において使用するための追加のプロモーター スワップライブラリーの生成および試験

この実施例は、工業用株性能を改善するための取り組みにおいて微生物宿主細胞（例えば、E s c h e r i c h i a c o l i ）を遺伝子操作するための本明細書に提供される H T P ゲノム操作方法に後で使用するための追加の P R O S W P ライブラリーの生成について記載する。

10

【0866】

本実施例では、いくつかの生来の E . c o l i プロモーターおよび合成プロモーターをコンパイルして、表 1 . 4 に見出されるプロモーター スワップライブラリーを生成した。生来のプロモーターに関しては、長さが 6 0 ~ 9 0 b p の一連のプロモーター配列を E . c o l i K - 1 2 株（すなわち、E . c o l i W 3 1 1 0 ）のゲノムから選択した。特に、多数の増殖条件にわたるマイクロアレイに基づく発現データに従って、関連する遺伝子の発現に関して最小のバリエーションを示したプロモーターを選択した（Lewisら、M o l S y s t B i o l . , 2 0 1 0 年；6 巻：3 9 0 頁）。生来のプロモーター配列は、推定転写開始点の前の 5 0 b p および推定開始コドン含まずに、推定開始コドンまでの配列からなった（表 1 . 4 参照）。さらに、公知のラムダファージプロモーターの一部である p_L および p_R、E . c o l i 遺伝子 a c s の前のプロモーター、ならびに - 3 5 および - 1 0 領域を構成する可変 6 b p 配列からなる一連のキメラ合成プロモーター配列を創製した（図 5 4、表 1 . 5）。合成プロモーターはそれぞれ長さ 6 0 ~ 9 0 b p であった。

20

【0867】

表 1 . 4 に見出されるプロモーターのそれぞれの、それに作動可能に連結した遺伝子の発現を駆動する能力を試験するために、それぞれが表 1 . 4 に列挙されているプロモーターのうちの 1 つによって駆動される R F P 遺伝子を含む一連の低コピー複製性プラスミドを構築した。選択された低コピー複製性プラスミドは、核酸配列配列番号 2 1 3 を有する O r i _ P l s m d 2 7 と呼ばれるプラスミドであった。このベクターを選択した理由は、プロモーターライブラリーをできるだけ早く構築し、評価するために複製性プラスミドが望ましく、また、O r i _ P l s m d 2 7 などの低コピープラスミドは単一コピーのみがゲノムに組み込まれるシナリオにより厳密に近づくからであった。O r i _ P l s m d 2 7 は、E . c o l i 複製起点 p 1 5 A を有するので、低コピーである。p 1 5 A 複製起点は、典型的には、各細胞においておよそ 1 0 コピーのプラスミドをもたらす。これは、細胞当たり > 2 0 またはさらには数百のプラスミドコピーを維持し得る他の一般的なプラスミドと比較して「低コピー」である。

30

【0868】

プラスミドは、標準の分子生物学技法を使用して構築した。具体的には、R F P 遺伝子とアニールする配列；導入されるプロモーター配列；および O r i _ P l s m d 2 7 と重複する配列からなる順方向 P C R プライマーを購入した。E C K 1 2 0 0 3 3 7 3 7 ターミネーター（生来の E . c o l i ターミネーター）とアニールする配列および O r i _ P l s m d 2 7 と重複する配列でからなる単一の逆方向 P C R プライマーを得た。R F P 遺伝子を順方向プライマーおよび逆方向プライマーを用いて P C R によって増幅して、それぞれが R F P 遺伝子および表 1 . 4 に列挙されているプロモーターのうちの 1 つを含む一連の P C R アンプリコンを生成した。市販の D N A アセンブリー酵素ミックスを使用して O r i _ P l s m d 2 7 を X h o I 制限酵素で消化し、対応する P C R アンプリコンを挿入することによってプラスミドを構築した。陰性対照として、R F P 遺伝子の上流に位置する C . g l u t a m i c u m T s o d ターミネーター（表 1 6 の核酸配列番号 2 2 4）を含むコンストラクトを生成した。

40

50

【0869】

プラスミドを電気穿孔によってE. coli W3110中に形質転換した。評価される各プロモーターについて、4つのコロニーを選び取り、96ウェル培養プレート中、25 $\mu\text{g}/\text{mL}$ のカナマイシンを含有するLBブロス1 mLに播種した。培養物を37℃で終夜、1000 rpmで振とうしながら増殖させた。培養物10 μL を使用して、96ウェル培養プレート中、25 $\mu\text{g}/\text{mL}$ のカナマイシンを含有する培地1（グルコース、酵母抽出物、塩、およびリン酸緩衝剤を含有する富栄養培地）1 mLに播種した。培養物を37℃で24時間にわたって1000 rpmで振とうしながら増殖させた。培養物をblack wall clear bottom 96ウェルプレート中、水に希釈し、分光光度計で2つの測定値：OD₆₀₀（600 nmにおける光学濃度）および蛍光（554 nm励起、590 nm放出）を取得した。培地1中の培養物10 μL を使用して、96ウェル培養プレート中、25 $\mu\text{g}/\text{mL}$ のカナマイシンを含有する培地2（培地1よりも高グルコースを含有するがほんの少量の酵母抽出物を含有し、その代わりに、硫酸アンモニウムを窒素源として微量元素と一緒に含有する富栄養培地）1 mLに播種した。培地2中の培養物を同様に37℃で終夜増殖させ、24時間後に測定した。

10

【0870】

OD₆₀₀測定値を、ブランクウェル（培地のみを含有するウェル）の値を差し引くことによって補正した。表16は、補正されたOD₆₀₀について正規化された蛍光測定値を示す。表16において見ることができるよう、結果として生じた株は2つの異なる培地中で有効に増殖し、約5000倍の範囲にわたって蛍光タンパク質発現を可能にした。

20

【表16-1】

表16. (2つの)異なる培地におけるプロモーター-RFPコンストラクトについてのRFP発現レベル。

プロモーター名	配列番号	タイプ*	培地1における蛍光	培地2における蛍光
空ベクター		対照	0.59310262	1.038945766
プロモーターなし		対照	0.372576311	0.759290158
ターミネーター	224	対照	0.41480568	0.638974345
b0904_プロモーター	71	生来の	797.7465067	5344.268555
b2405_プロモーター	72	生来の	181.125065	918.6698873
b0096_プロモーター	73	生来の	124.2236508	404.6536222
b0576_プロモーター	74	生来の	22.74750166	231.1578591
b2017_プロモーター	75	生来の	43.07006732	214.1483559

30

【表 1 6 - 2】

b1278_プロモーター	76	生来の	24.51965879	167.5502652
b4255_プロモーター	77	生来の	35.05859344	159.6253961
b0786_プロモーター	78	生来の	34.42969151	136.7168488
b0605_プロモーター	79	生来の	19.80356197	116.9009802
b1824_プロモーター	80	生来の	24.81069692	111.2950744
b1061_プロモーター	81	生来の	23.595542	104.971453
b0313_プロモーター	82	生来の	20.68806067	75.34101875
b0814_プロモーター	83	生来の	16.46706854	70.36934251
b4133_プロモーター	84	生来の	18.1034447	67.11707926
b4268_プロモーター	85	生来の	12.80314528	60.91103948
b0345_プロモーター	86	生来の	9.57124889	56.38701501
b2096_プロモーター	87	生来の	9.852378505	44.58545057
b1277_プロモーター	88	生来の	7.821796935	43.50459312
b1646_プロモーター	89	生来の	18.43768693	33.66381279
b4177_プロモーター	90	生来の	12.51676981	30.71734482
b0369_プロモーター	91	生来の	6.614285367	28.3955664
b1920_プロモーター	92	生来の	6.196134846	28.12617522
b3742_プロモーター	93	生来の	5.646106239	27.85277232
b3929_プロモーター	94	生来の	9.120314333	27.26180409
b3743_プロモーター	95	生来の	5.416134854	26.11889953
b1613_プロモーター	96	生来の	6.653798388	24.49553744
b1749_プロモーター	97	生来の	4.14048349	24.04592462
b2478_プロモーター	98	生来の	5.387025634	20.49940077
b0031_プロモーター	99	生来の	3.227415659	18.8898903
b2414_プロモーター	100	生来の	5.054709408	17.84254467
b1183_プロモーター	101	生来の	5.109245812	14.42155472
b0159_プロモーター	102	生来の	3.662900763	14.16361794
b2837_プロモーター	103	生来の	3.427953441	13.22830567
b3237_プロモーター	104	生来の	3.238274954	11.27938911
b3778_プロモーター	105	生来の	1.778132048	8.89448301
b2349_プロモーター	106	生来の	1.918547188	8.854115649
b1434_プロモーター	107	生来の	3.05979593	7.982575199

10

20

30

40

【表 1 6 - 3】

b3617_プロモーター	108	生来の	1.631428047	7.872487968
b0237_プロモーター	109	生来の	2.524629463	6.00416496
b4063_プロモーター	110	生来の	1.75938278	5.338598994
b0564_プロモーター	111	生来の	1.099810487	5.27271929
b0019_プロモーター	112	生来の	1.407515599	5.174702317
b2375_プロモーター	113	生来の	0.758754127	4.329319889
b1187_プロモーター	114	生来の	1.042830698	4.119645144
b2388_プロモーター	115	生来の	1.068375672	4.008318722
b1051_プロモーター	116	生来の	2.595570861	3.770122087
b4241_プロモーター	117	生来の	0.86986247	3.520447288
b4054_プロモーター	118	生来の	1.166071321	3.009684331
b2425_プロモーター	119	生来の	0.684355198	2.847357334
b0995_プロモーター	120	生来の	0.783460564	2.701492023
b1399_プロモーター	121	生来の	0.67920166	2.675743061
b3298_プロモーター	122	生来の	0.693842895	2.248914103
b2114_プロモーター	123	生来の	0.680005978	2.166569672
b2779_プロモーター	124	生来の	0.516911046	1.708793442
b1114_プロモーター	125	生来の	0.477678475	1.423398279
b3730_プロモーター	126	生来の	0.527510171	1.40174991
b3025_プロモーター	127	生来の	0.528069061	1.375386866
b0850_プロモーター	128	生来の	0.580500872	1.365244168
b2365_プロモーター	129	生来の	0.600099606	1.293746086
b4117_プロモーター	130	生来の	0.465598493	1.292314074
b2213_プロモーター	131	生来の	0.502863501	1.066458143
pMB029_プロモーター	132	合成	3244.159758	6397.50355
pMB023_プロモーター	133	合成	2716.103558	5324.355721
pMB025_プロモーター	134	合成	2529.260134	5315.845766
pMB019_プロモーター	135	合成	2632.199487	4730.881793
pMB008_プロモーター	136	合成	3183.229982	4685.405472
pMB020_プロモーター	137	合成	1719.381656	4331.834596
pMB022_プロモーター	138	合成	1432.727416	4224.994617
pMB089_プロモーター	139	合成	1588.496061	4173.815385

10

20

30

40

【表 1 6 - 4】

pMB001_プロモーター	140	合成	1344.15431	3710.844718
pMB051_プロモーター	141	合成	1485.624649	3614.713536
pMB070_プロモーター	142	合成	1148.371596	3494.100223
pMB074_プロモーター	143	合成	1372.63931	3423.518954
pMB046_プロモーター	144	合成	1070.100822	3325.635875
pMB071_プロモーター	145	合成	829.6337908	2753.919683
pMB013_プロモーター	146	合成	649.3261186	2255.268459
pMB080_プロモーター	147	合成	584.6548742	2064.645666
pMB038_プロモーター	148	合成	580.7846278	2043.19218
pMB060_プロモーター	149	合成	575.1167059	1916.349295
pMB064_プロモーター	150	合成	376.0013091	1390.767111
pMB058_プロモーター	151	合成	292.4875916	1093.376645
pMB085_プロモーター	152	合成	288.8178401	1085.653214
pMB081_プロモーター	153	合成	278.7856475	670.8190328
pMB091_プロモーター	154	合成	296.1836771	655.9453316
pMB027_プロモーター	155	合成	262.5866089	601.1559107
pMB048_プロモーター	156	合成	235.0853436	567.7578783
pMB055_プロモーター	157	合成	189.6983907	499.3358783
pMB006_プロモーター	158	合成	205.2252242	497.2257006
pMB012_プロモーター	159	合成	118.144338	480.5859474
pMB014_プロモーター	160	合成	119.9861606	474.4330596
pMB028_プロモーター	161	合成	204.8853226	454.9894988
pMB059_プロモーター	162	合成	101.7399612	387.0993978
pMB061_プロモーター	163	合成	128.6404482	381.8073111
pMB043_プロモーター	164	合成	139.0031576	380.0980602
pMB066_プロモーター	165	合成	95.64407537	328.5708555
pMB079_プロモーター	166	合成	105.9768914	312.5115498
pMB032_プロモーター	167	合成	89.71910699	311.1818243
pMB068_プロモーター	168	合成	97.56292371	276.8593245
pMB082_プロモーター	169	合成	84.20594634	262.0127742
pMB030_プロモーター	170	合成	92.28874802	255.1052708
pMB067_プロモーター	171	合成	49.74078463	243.9642233

10

20

30

40

【表 16-5】

pMB050_プロモーター	172	合成	69.03704925	223.6129199
pMB069_プロモーター	173	合成	66.09043115	213.5258967
pMB017_プロモーター	174	合成	50.43740198	213.0316985
pMB039_プロモーター	175	合成	52.58639745	209.9432421
pMB011_プロモーター	176	合成	63.05192023	194.8104082
pMB072_プロモーター	177	合成	47.21948722	165.6730417
pMB016_プロモーター	178	合成	47.20169015	163.6018648
pMB077_プロモーター	179	合成	43.98192292	158.5866563
pMB047_プロモーター	180	合成	30.55918215	121.2753356
pMB052_プロモーター	181	合成	34.86596074	114.4169675
pMB090_プロモーター	182	合成	14.01229264	55.03412277
pMB035_プロモーター	183	合成	14.7011564	50.95895734
pMB073_プロモーター	184	合成	11.79654115	47.47608915
pMB004_プロモーター	185	合成	8.701914436	33.9085268
pMB054_プロモーター	186	合成	8.934971498	32.06510507
pMB024_プロモーター	187	合成	6.760860897	27.59246488
pMB007_プロモーター	188	合成	11.5557009	25.41738039
pMB005_プロモーター	189	合成	5.354276069	23.77742451
pMB003_プロモーター	190	合成	6.389447374	22.7246451
pMB088_プロモーター	191	合成	5.873205603	22.42214302
pMB065_プロモーター	192	合成	5.44173707	20.60442307
pMB037_プロモーター	193	合成	4.809680789	19.26108105
pMB009_プロモーター	194	合成	3.591881545	13.6498423
pMB041_プロモーター	195	合成	4.78150005	13.03957332
pMB036_プロモーター	196	合成	2.508122856	11.98560161
pMB049_プロモーター	197	合成	2.916151498	11.02944113
pMB044_プロモーター	198	合成	2.719080697	10.47251128
pMB042_プロモーター	199	合成	2.661753833	8.883733663
pMB086_プロモーター	200	合成	2.33212712	8.316649305
pMB053_プロモーター	201	合成	2.336160735	7.262981543
pMB057_プロモーター	202	合成	1.518801989	6.27420435
pMB018_プロモーター	203	合成	1.261000991	4.714509042

10

20

30

40

【表 1 6 - 6】

pMB002_プロモーター	204	合成	1.290587949	4.473711828
pMB015_プロモーター	205	合成	0.994888783	4.328541658
pMB087_プロモーター	206	合成	0.94621781	2.900396821
pMB063_プロモーター	207	合成	0.425085793	1.459419227

* *Escherichia coli* 由来の生来のプロモーター

【0 8 7 1】

(実施例 1 3)

H T P ゲノム操作—ベクター 2 骨格を使用した表 1 . 4 のプロモータースワップライブラリーの *E. coli* のゲノムへの組込みの試験

この実施例は、単一コピーの異種プロモーター—遺伝子コンストラクトの *E. coli* のゲノムへの組込みを駆動するための、実施例 1 1 からのベクター 2 骨格と表 1 . 4 のプロモーターライブラリーからのプロモーターのサブセットの組合せの使用の概念実証について記載する。

【0 8 7 2】

この実施例のために、蛍光遺伝子 R F P および G F P を *E. coli* における 2 つの遺伝子座 (n u p G および a s l) に挿入するために、表 1 . 4 におけるセットからの 1 4 種のプロモーターのサブセットを用いてプラスミドのセットを構築する。これにより、これらのプロモーターを、低コピー複製性プラスミドではなくゲノムに組み込まれた単一コピーとして評価することが可能になる (実施例 1 2 を参照されたい)。

【0 8 7 3】

プラスミドは、本開示全体を通して提供される「ルーブイン」による *E. coli* のゲノムへの組込みを促進するために、R F P または G F P 遺伝子に隣接する相同アームを含む。結果として生じる株を蛍光について試験し、これにより、この表 1 . 4 からの 1 4 種のプロモーターのサブセットが実施例 1 1 に記載のベクター骨格を使用して試験され、本開示に記載されている方法を使用して異種遺伝子を *E. coli* のゲノムに挿入するために使用できることが実証される。

【0 8 7 4】

(実施例 1 4)

H T P ゲノム操作—表 1 . 4 から得られるプロモーターライブラリーを使用した P R O S W P 方法のインプリメンテーション。

以下のセクションは、実施例 4 および 5 に記載の通り、本開示の P R O スワップ H T P 設計株改良プログラムツールの例示的なインプリメンテーションを提供する。本実施例では、*E. coli* ゲノム内の遺伝子の発現をモジュレートするために、*E. coli* 株を本開示の P R O スワップ方法に供した。この実施例は、本開示の P R O S W P 方法における、表 1 . 4 からのプロモーターを含むプロモーターライブラリーの使用を示すという点で、この実施例は実施例 1 2 および 1 3 の結果に基づく。

【0 8 7 5】

A. プロモータースワップ

プロモータースワップを実施例 4 に記載の通り行う。*E. coli* ゲノム由来の遺伝子を、表 1 . 4 からの 1 4 種のプロモーターのサブセットを含む実施例 1 3 に記載のプロモーターライブラリーを使用したプロモータースワップに供す。本実施例において使用される 1 4 種のプロモーターのサブセットを、実施例 1 2 および 1 3 において決定された遺伝子発現に対するそれらの効果に基づいて選択する。

【0 8 7 6】

B. H T P 操作およびハイスループットスクリーニング

プロモータースワップの H T P 操作を実施例 1 および 3 に記載の通り行う。結果として生じるプロモータースワップ株の H T P スクリーニングを実施例 3 に記載の通り行う。全

10

20

30

40

50

部で、14回のPROスワップを行う。最後に、これらの改変の、目的の産物の産生に対する影響を試験する。

【0877】

(実施例15)

HTPゲノム操作ーリコピンの産生に関して、株性能を改善するための、ターミネータースワップライブラリーのインプリメンテーション

以下のセクションは、本開示のターミネータースワップHTP設計株改良プログラムツールの例示的なインプリメンテーションを提供する。本実施例では、リコピンの宿主細胞による収率に影響を及ぼすために、E. coli株を本開示のターミネータースワップ方法に供した。

【0878】

図59に示されるリコピン生合成経路内の遺伝子を標的とするターミネータースワップを、本開示全体を通して示されるターミネータースワップ方法を使用して行った。コンストラクトを下記の通り設計し、組換えをCRISPR/Cas9システムで媒介した。本実施例におけるターミネータースワップのために使用したターミネーターは、表19に見出されるターミネーターであった。

【表19】

表19-リコピン生合成経路内の遺伝子を標的にするための、本実施例において使用されるターミネーター

名称	説明	長さ (bp)
Spy	ターミネーター (配列番号 225)	90
pheA	ターミネーター (配列番号 226)	51
osmE	ターミネーター (配列番号 227)	42
rpoH	ターミネーター (配列番号 228)	41
vibE	ターミネーター (配列番号 229)	71
Thrl_ABC	ターミネーター (配列番号 230)	57

【0879】

コンストラクト設計

ゲノムを所望の位置で切断するために、標的挿入配列の付近にあり、NGG PAM配列に隣接する20塩基ガイドRNAを同定した。ゲノムに挿入されることが意図された配列は、両末端に90塩基の相同性で隣接し、したがって、相同性により、生来の配列が所望の通り欠失するまたは保持されるように向きつけられた。本実施例では組換えをCRISPR/Cas9システムによって促進したが、株はすべて、従来の単一および二重クロスオーバー相同組換え方法ならびに本開示全体を通して記載されているラムダレッドシステムによって構築することもできることに留意するべきである。したがって、ターミネータースワップライブラリータイプのそれぞれは、構築/組換え方法に関してとらわれない。

【0880】

シード培養の播種

ペトリプレートから編集基準株 (W3110 pKD46-cas9 pLYC4) のコロニーを選び取り、より大きな体積のLB clon100 cmp25中に播種し、培養物を30℃で振とうしながら約16時間にわたって増殖させた。

【0881】

コンピテント細胞の調製および形質転換

終夜培養物の1:10希釈物を調製し、OD600を測定した。LB clin100 cmp25をOD600が0.05になるまで播種し、30℃で振とうしながら約2時間にわたって増殖させた。

【0882】

2時間後、誘導標的ODに達するまで定期的にOD600を測定し、標的に達したら、20%アラビノースを最終濃度0.2%まで添加した。

【0883】

培養物を4℃、5,000×Gで5分遠心分離する。上清を流し出し、元の培養物体積と同等の最終体積まで再懸濁させる。 10

【0884】

ステップ7を繰り返して、3回目に細胞を洗浄する。

【0885】

3回洗浄後、細胞をペレット化し、10%グリセロールに元の培養物体積の約1/250になるまで再懸濁した。再懸濁した細胞の1:500希釈物を調製し、形質転換当たり細胞40μLに対して適切な体積を用いて所望のOD600まで再懸濁させた。

【0886】

Framestar PCRプレート（または微量遠心管）中、細胞40μLを、ガイドRNAプラスミド100ng、および精製されたPCR産物修復鋳型約4μLと混合した。修復鋳型にオリゴを使用する場合、オリゴを最終濃度2μMまで添加した。 20

【0887】

細胞に電気穿孔を行い、すぐにLBに再懸濁させ、ディープウェルプレート中に回収し、30℃で振とうしながら1時間置いた。

【0888】

回収された細胞を希釈し、LB寒天clin100 kan50 cmp25に蒔き、30℃で24~36時間置いた。コロニーをコロニーPCR、シーケンシング、または表現型スクリーニングによってスクリーニングした。pKD46-cas9プラスミドを37℃またはそれよりも高い温度で増殖させることによってキュアリングすることができ、pCRISPR2を10%スクロースで増殖させることによってキュアリングすることができる。 30

【0889】

HTP操作およびハイスループットスクリーニング

スワップ組合せのHTP操作を、コンストラクトのE. coliゲノム中への相同組換えを促進するためにCRISPR/Cas9を使用したことを除いて実施例1および3に記載の通り行った。結果として生じたプロモータースワップ/ターミネータースワップ株、プロモータースワップ/分解タグスワップ株、プロモータースワップ/可溶性タグスワップ株およびプロモータースワップ/ターミネータースワップ/分解タグスワップ/可溶性タグスワップ株のHTPスクリーニングを実施例3に記載の通り行った。実験の結果を図60および61に示す。 40

【0890】

図60に示されている通り、ターミネーターTyj bEを使用したリコピン経路標的idiおよびymgAにおけるターミネータースワップにより、対照と比べて株性能の低下が実証され、したがって、極めて重要な経路標的の同定に関するこれらのライブラリータイプの有用性が強調される。この結論は、多数のリコピン経路標的に対してターミネータースワップを行った図61に示されている結果によってさらに裏付けられた。

【0891】

（実施例16）

リコピンの産生に関して、株性能を改善するための、HTPゲノム操作-ターミネータースワップライブラリーまたはPROスワップライブラリーの、SOLUBILITY 50

T A GスワップライブラリーまたはD E G R A D A T I O N T a gスワップライブラリーのいずれかと組み合わせたインプリメンテーション

以下のセクションは、S O L U B I L I T Y T A Gスワップおよび本開示のターミネータースワップH T P設計株改良プログラムツールならびに本開示のP R OスワップおよびD E G R A D A T I O N T A Gスワップ設計株改良プログラムツールの例示的なインプリメンテーションを提供する。本実施例では、リコピンの宿主細胞による収率に影響を及ぼすために、E . c o l i株を、P R Oスワップと、本開示のD E G R A D A T I O N

T A Gスワップ方法ならびに本開示のS O L U B I L I T Y T A Gスワップおよびターミネータースワップ方法の組合せに供した。

【0892】

プロモータースワップ／ターミネータースワップ／可溶性タグスワップ／分解タグスワップ

ターミネータースワップは実施例15に記載の通り行い、可溶性タグスワップおよびプロモータースワップと分解タグスワップの組合せは基本的に実施例4および5に記載の通り行った。コンストラクトを下記の通り設計し、組換えをC R I S P R／C a s 9システムで媒介した。本実施例におけるプロモータースワップのために使用したバイシストロン性プロモーターは、Mutalikら、Nat Methods.、2013年4月；10巻（4号）：354～60頁からのものであり、表20に見出すことができる。本明細書に提供されるプロモーターのいずれかを下記の方法に使用できることに留意するべきである。

【表20】

表20-本実施例におけるプロモータースワップ組合せに使用したプロモーター

名称	長さ	配列番号
P3_BCD1	133	255
P4_BCD22	121	256
P7_BCD19	132	257
P8_BCD15	121	258
P11_BCD17	121	259
P13_BCD8	121	260

【0893】

コンストラクト設計

ゲノムを所望の位置で切断するために、標的挿入配列の付近にあり、N G G P A M配列に隣接する20塩基ガイドRNAを同定した。ゲノムに挿入されることが意図された配列は、両末端に90塩基の相同性で隣接し、したがって、相同性により、生来の配列が所望の通り欠失するまたは保持されるようになりつけられた。本実施例では組換えをC R I S P R／C a s 9システムによって促進したが、株はすべて、従来の単一および二重クロスオーバー相同組換え方法ならびに本開示全体を通して記載されているラムダレッドシステムによって構築することもできることに留意するべきである。したがって、これらのライブラリータイプ（プロモータースワップ、タンパク質可溶性タグスワップ、タンパク質分解タグスワップ、およびターミネータースワップ）はそれぞれ、単独でまたは組合せて、構築／組換え方法にとらわれない。

【0894】

シード培養の播種

ペトリプレートから編集基準株（W 3 1 1 0 p K D 4 6 - c a s 9 p L Y C 4）のコロニーを選び取り、より大きな体積のL B c l i n 1 0 0 c m p 2 5中に播種し、培養物を30℃で振とうしながら約16時間にわたって増殖させた。

【0895】

コンピテント細胞の調製および形質転換

終夜培養物の1：10希釈物を調製し、OD600を測定した。L B c l i n 1 0 0 c m p 2 5をOD600が0.05になるまで播種し、30℃で振とうしながら約2時

10

20

30

40

50

間にわたって増殖させた。

【0896】

2時間後、誘導標的ODに達するまで定期的にOD600を測定し、標的に達したら、20%アラビノースを最終濃度0.2%まで添加した。

【0897】

培養物を4℃、5,000×Gで5分遠心分離する。上清を流し出し、元の培養物体積と同等の最終体積まで再懸濁させる。

【0898】

ステップ7を繰り返して、3回目に細胞を洗浄する。

【0899】

3回洗浄後、細胞をペレット化し、10%グリセロールに元の培養物体積の約1/250になるまで再懸濁した。再懸濁した細胞の1:500希釈物を調製し、形質転換当たり細胞40μLに対して適切な体積を用いて所望のOD600まで再懸濁させた。

【0900】

Framestar PCRプレート（または微量遠心管）中、細胞40μLを、ガイドRNAプラスミド100ng、および精製されたPCR産物修復鋳型約4μLと混合した。修復鋳型にオリゴを使用する場合、オリゴを最終濃度2μMまで添加した。

【0901】

細胞に電気穿孔を行い、すぐにLBに再懸濁させ、ディープウェルプレート中に回収し、30℃で振とうしながら1時間置いた。

【0902】

回収された細胞を希釈し、LB寒天clin100 kan50 cmp25に蒔き、30℃で24~36時間置いた。コロニーをコロニーPCR、シーケンシング、または表現型スクリーニングによってスクリーニングした。pKD46-cas9プラスミドを37℃またはそれよりも高い温度で増殖させることによってキュアリングすることができ、pCRI-SPR2を10%スクロースで増殖させることによってキュアリングすることができる。

【0903】

HTP操作およびハイスループットスクリーニング

スワップ組合せのHTP操作を、コンストラクトのE.coliゲノム中への相同組換えを促進するためにCRISPR/Cas9を使用したことを除いて実施例1および3に記載の通り行った。結果として生じたプロモータースワップ/ターミネータースワップ株、プロモータースワップ/分解タグスワップ株、プロモータースワップ/可溶性タグスワップ株およびプロモータースワップ/ターミネータースワップ/分解タグスワップ/可溶性タグスワップ株のHTPスクリーニングを実施例3に記載の通り行った。

【0904】

例えばP4__BCD22など、異なることが特に示されない限り、dxs遺伝子座における改変を試験中であるすべての株にプロモーターP3__BCD1を使用したことに留意すべきである。dxs以外の任意の遺伝子座には、特に断りのない限り、生来のプロモーター配列を使用した。これは、例えば、dxs遺伝子座におけるssrA__LAAと記載される株は、P3__BCD1も含有したが、gdhA遺伝子座におけるssrA__LAAと記載される株には、生来のプロモーター配列を使用したことを意味する。以下の表21における試験した株の全内容。

10

20

30

40

【表 2 1】

表 21-実施例 16 において生成された株の内容。

dxs 遺伝子座における改変						
	*対照-P	P4_BCD22	ssrA_ASV	ssrA_LAA	Tspy	TyjbE
プロモーター	P3_BCD1	P4_BCD22	P3_BCD1	P3_BCD1	P3_BCD1	P3_BCD1
溶解性タグ	なし	なし	なし	なし	なし	なし
分解タグ	なし	なし	ssrA_ASV	ssrA_LAA	なし	なし
ターミネーター	生来の配列	生来の配列	生来の配列	生来の配列	Tspy	TyjbE
gdhA 遺伝子座における改変						
	*対照-P	FH8	GB1	P4_BCD22	Tspy	TyjbE
プロモーター	生来の配列	生来の配列	生来の配列	P4_BCD22	生来の配列	生来の配列
溶解性タグ	なし	FH8	GB1	なし	なし	なし
分解タグ	なし	なし	なし	なし	なし	なし
ターミネーター	生来の配列	生来の配列	生来の配列	生来の配列	Tspy	TyjbE

【0905】

実験の結果を要約し、図 6 2 および 6 3 に示す。

【0906】

図 6 2 に示されている通り、ssrA_LAA 分解タグにより、対照と比べて株性能の改善が実証される。この株は単一の経路標的における PROSWP と分解タグの組合せであるので、これは予想外である。最初の PROSWP によりタンパク質存在量が増大することが予想され、分解タグによりタンパク質存在量が減少することが予想され、したがって、最適な株性能の調整に関するライブラリータイプの組合せの有用性が実証される。図 6 3 に示されている通り、可溶性タグ FH8 により、対照と比べて株性能の改善が実証されるが、GB1 可溶性タグではこれは実証されず、よって、各改変タイプのライブラリーを評価することの必要性が実証される。

【0907】

全体として、この実施例が実証することは、本発明のコンポーネントは系統的な株改良に関して個別に有用であり得るが、他の手法と相乗的に使用することもできるということである。例えば、ターミネーター改変によって mRNA 安定性を改善した後、強力なプロモーターを挿入して、タンパク質産生を、いずれかの手法単独でのレベルを超えてさらに増大させることができる。同様に、この新しいタンパク質産生レベルの上昇を、タンパク質可溶性タグと他の改変を併せてさらに改善することができる。本発明のコンポーネントを一緒に、および先の手法とともに利用すれば、標的分子の産生のためのよりロバストかつ有効な株改良が可能になり得る。

【0908】

さらなる番号を付けた本開示の実施形態

本開示が意図する他の主題を、以下の番号を付けた実施形態において示す：

実施形態 1. 所望の表現型を獲得するために E. coli 微生物を進化させるゲノム操作のハイスループット (HTP) 方法であって、

a. 同じゲノム株バックグラウンドを有する最初の複数の E. coli 微生物のゲノムを攪乱することによって、ユニークな遺伝的バリエーションを有する個々の E. coli 株を含む最初の HTP 遺伝子設計 E. coli 株ライブラリーを創製するステップと；

b. 前記最初の HTP 遺伝子設計 E. coli 株ライブラリーの個々の株を、前記所望の表現型についてスクリーニングおよび選択するステップと；

10

20

30

40

50

c. それぞれが遺伝的バリエーションのユニークな組合せを含む後続の複数の E. coli 微生物を提供することによって、後続の HTP 遺伝子設計 E. coli 株ライブラリーを創製するステップであって、前記遺伝的バリエーションは、先行するステップにおいてスクリーニングされた少なくとも 2 つの個々の E. coli 株に存在する遺伝的バリエーションから選択される、ステップと；

d. 前記後続の HTP 遺伝子設計 E. coli 株ライブラリーの個々の E. coli 株を、前記所望の表現型についてスクリーニングおよび選択するステップと；

e. E. coli 微生物が前記所望の表現型を獲得するまで、線形または非線形様式でステップ c) ~ d) を 1 回または複数回繰り返すステップであって、各後続の反復により、先行する HTP 遺伝子設計 E. coli 株ライブラリーの少なくとも 2 つの個々の E. coli 株の中から選択される遺伝的バリエーションの組合せであるユニークな遺伝的バリエーションを内包する個々の E. coli 株を含む新しい HTP 遺伝子設計 E. coli 株ライブラリーが創製される、ステップとを含む、方法。

実施形態 2. 前記最初の HTP 遺伝子設計 E. coli 株ライブラリーが、プロモータースワップ微生物株ライブラリー、SNP スワップ微生物株ライブラリー、開始/停止コドン微生物株ライブラリー、最適化配列微生物株ライブラリー、ターミネータースワップ微生物株ライブラリー、タンパク質可溶性タグ微生物株ライブラリー、タンパク質分解タグ微生物株ライブラリー、およびこれらの任意の組合せからなる群から選択される少なくとも 1 つのライブラリーを含む、実施形態 1 に記載のゲノム操作の HTP 方法。

実施形態 3. 前記最初の HTP 遺伝子設計 E. coli 株ライブラリーが、プロモータースワップ微生物株ライブラリーを含む、実施形態 1 に記載のゲノム操作の HTP 方法。

実施形態 4. 前記最初の HTP 遺伝子設計 E. coli 株ライブラリーが、少なくとも 1 つのバイシストロン設計 (BCD) 調節配列を含有するプロモータースワップ微生物株ライブラリーを含む、実施形態 1 に記載のゲノム操作の HTP 方法。

実施形態 4. 1 前記 BCD 調節配列が、

- a. 作動可能に連結したプロモーター；
- b. 第 1 のリボソーム結合部位 (SD1)；
- c. 第 1 のシストロン配列 (Cis1)；
- d. 第 2 のリボソーム結合部位 (SD2)

を順に含み、

前記 BCD 配列が、標的遺伝子配列 (Cis2) に作動可能に連結している、実施形態 4 に記載のゲノム操作の HTP 方法。

実施形態 4. 2 SD1 および SD2 がそれぞれ NNNGGANNN の配列を含む、実施形態 4. 1 に記載のゲノム操作の HTP 方法。

実施形態 4. 3 SD1 および SD2 が異なる、実施形態 4. 1 または 4. 2 に記載のゲノム操作の HTP 方法。

実施形態 4. 4 Cis1 が停止コドンを含み、Cis2 が開始コドンを含み、Cis1 停止コドンおよび Cis2 開始コドンが少なくとも 1 ヌクレオチド重複する、実施形態 4. 1 から 4. 3 までのいずれか 1 つに記載のゲノム操作の HTP 方法。

実施形態 4. 5 SD2 が、Cis1 中に完全に組み込まれている、実施形態 4. 1 から 4. 4 までのいずれか 1 つに記載のゲノム操作の HTP 方法。

実施形態 5. 最初の HTP 遺伝子設計 E. coli 株ライブラリーが、SNP スワップ微生物株ライブラリーを含む、実施形態 1 から 4. 5 までのいずれか 1 つに記載のゲノム操作の HTP 方法。

実施形態 6. 前記最初の HTP 遺伝子設計 E. coli 株ライブラリーが、

a. キメラ生合成酵素をコードする少なくとも 1 つのポリヌクレオチドであって、前記キメラ生合成酵素は、

- i. E. coli において調節経路に関与する酵素を含み、

i i. 前記酵素に、DNA結合部位に結合することができるDNA結合ドメインが翻訳的に融合する、

少なくとも1つのポリヌクレオチド；および

b. 前記キメラ生合成酵素の前記DNA結合ドメインに対応する前記DNA結合部位を含む少なくとも1つのDNA足場配列

を含む微生物株ライブラリーを含む、実施形態1から5までのいずれか1つに記載のゲノム操作のHTP方法。

実施形態6. 1 最初のHTP遺伝子設計E. coli株ライブラリーが、

a. キメラ生合成酵素をコードする少なくとも1つのポリヌクレオチドであって、前記キメラ生合成酵素は、

i. E. coliにおいて調節経路に参与する酵素を含み；

i i. 前記酵素に、動員ペプチドに結合することができるタンパク質結合ドメインが翻訳的に融合する、

少なくとも1つのポリヌクレオチド、および

b. キメラ生合成酵素のタンパク質結合ドメインに対応する動員ペプチドを含む少なくとも1つのタンパク質足場配列

を含む微生物株ライブラリーを含む、実施形態1から5までのいずれか1つに記載のゲノム操作のHTP方法。

実施形態7. 前記後続のHTP遺伝子設計E. coli株ライブラリーが、前記最初のHTP遺伝子設計E. coli株ライブラリーにおける前記遺伝的バリエーションに由来する完全コンビナトリアル株ライブラリーである、実施形態1から6. 1までのいずれか1つに記載のゲノム操作のHTP方法。

実施形態8. 前記後続のHTP遺伝子設計E. coli株ライブラリーが、前記最初のHTP遺伝子設計E. coli株ライブラリーにおける前記遺伝的バリエーションに由来する完全コンビナトリアル株ライブラリーのサブセットである、実施形態1から6. 1までのいずれか1つに記載のゲノム操作のHTP方法。

実施形態9. 前記後続のHTP遺伝子設計E. coli株ライブラリーが、先行するHTP遺伝子設計E. coli株ライブラリーにおける前記遺伝的バリエーションに由来する完全コンビナトリアル株ライブラリーである、実施形態1から6. 1までのいずれか1つに記載のゲノム操作のHTP方法。

実施形態10. 前記後続のHTP遺伝子設計E. coli株ライブラリーが、先行するHTP遺伝子設計E. coli株ライブラリーにおける前記遺伝的バリエーションに由来する完全コンビナトリアル株ライブラリーのサブセットである、実施形態1から6. 1までのいずれか1つに記載のゲノム操作のHTP方法。

実施形態11. 前記ゲノムを攪乱するステップが、ランダム変異誘発、標的配列挿入、標的配列欠失、標的配列の置換、およびこれらの任意の組合せからなる群から選択される少なくとも1つの方法を利用することを含む、実施形態1から10までのいずれか1つに記載のゲノム操作のHTP方法。

実施形態12. 前記最初の複数のE. coli微生物が、工業生産E. coli株に由来するユニークな遺伝的バリエーションを含む、実施形態1から11までのいずれか1つに記載のゲノム操作のHTP方法。

実施形態13. 前記最初の複数のE. coli微生物が、S₁Gen₁で表される工業生産株微生物およびS_nGen_nで表されるこれらに由来する任意の数の後続の微生物世代を含む、実施形態1から12までのいずれか1つに記載のゲノム操作のHTP方法。

実施形態14. SNPスワップE. coli株ライブラリーを生成するための方法であって、

a. 参照E. coli株および第2のE. coli株を提供するステップであって、前記第2のE. coli株は、前記参照E. coli株内に存在しない一塩基多型、DNA挿入、およびDNA欠失から選択される複数の同定された遺伝的バリエーションを含む、ステップと；

10

20

30

40

50

b. 前記参照 E. coli 株または前記第 2 の E. coli 株のいずれかのゲノムを攪乱することによって、複数の個々の E. coli 株を含む最初の SNP スワップ E. coli 株ライブラリーを創製するステップであって、前記複数の個々の株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記参照 E. coli 株と前記第 2 の E. coli 株との間の前記複数の同定された遺伝的バリエーションから選択される単一の遺伝的バリエーションに対応する、ステップと

を含む、方法。

実施形態 15. 前記参照 E. coli 株の前記ゲノムが攪乱されて、前記第 2 の E. coli 株内に見出される前記同定された一塩基多型、DNA 挿入、または DNA 欠失の 1 つまたは複数が付加される、実施形態 14 に記載の SNP スワップ E. coli 株ライブラリーを生成するための方法。

実施形態 16. 前記第 2 の E. coli 株の前記ゲノムが攪乱されて、前記参照 E. coli 株内に見出されない前記同定された一塩基多型、DNA 挿入、または DNA 欠失の 1 つまたは複数が除去される、実施形態 14 または 15 に記載の SNP スワップ E. coli 株ライブラリーを生成するための方法。

実施形態 17. ユニークな遺伝的バリエーションを有する結果として生じる前記複数の個々の E. coli 株が、前記参照 E. coli 株と前記第 2 の E. coli 株との間のすべての前記同定された遺伝的バリエーションの完全コンビナトリアルライブラリーと一緒に含む、実施形態 14 から 16 までのいずれか 1 つに記載の SNP スワップ E. coli 株ライブラリーを生成するための方法。

実施形態 18. ユニークな遺伝的バリエーションを有する結果として生じる前記複数の個々の E. coli 株が、前記参照 E. coli 株と前記第 2 の E. coli 株との間のすべての前記同定された遺伝的バリエーションの完全コンビナトリアルライブラリーのサブセットと一緒に含む、実施形態 14 から 16 までのいずれか 1 つに記載の SNP スワップ E. coli 株ライブラリーを生成するための方法。

実施形態 19. 産生 E. coli 株の表現型性能を再建および改善するための方法であって、

a. 親系統 E. coli 株およびそれに由来する産生 E. coli 株を提供するステップであって、前記産生 E. coli 株は、前記親系統株内に存在しない一塩基多型、DNA 挿入、および DNA 欠失から選択される複数の同定された遺伝的バリエーションを含む、ステップと；

b. 前記親系統 E. coli 株または前記産生 E. coli 株のいずれかのゲノムを攪乱して、最初の E. coli 株のライブラリーを創製するステップであって、前記最初のライブラリーにおける各株は、前記親系統 E. coli 株と前記産生 E. coli 株との間の前記複数の同定された遺伝的バリエーションからのユニークな遺伝的バリエーションを含む、ステップと；

c. 参照 E. coli 株に対する表現型性能改善について前記最初のライブラリーの個々の株をスクリーニングおよび選択することによって、表現型性能改善をもたらすユニークな遺伝的バリエーションを同定するステップと；

d. 前記先行するステップにおいてスクリーニングされた少なくとも 2 つの個々の E. coli 株に存在する遺伝的バリエーションからのユニークな遺伝的バリエーションの組合せをそれぞれが含む後続の複数の E. coli 微生物を提供することによって、E. coli 株の後続のライブラリーを創製するステップと；

e. 前記参照 E. coli 株に対する表現型性能改善について前記後続のライブラリーの個々の株をスクリーニングおよび選択することによって、追加の表現型性能改善をもたらす遺伝的バリエーションのユニークな組合せを同定するステップと；

f. E. coli 株が、前記産生 E. coli 株の前記表現型性能と比較して、所望のレベルの表現型性能の改善を呈するまで、線形または非線形様式でステップ d) ~ e) を 1 回または複数回繰り返すステップであって、各後続の反復により、微生物株の新しいラ

10

20

30

40

50

イブラリーが創製され、前記新しいライブラリーにおける各株は、先行するライブラリーの少なくとも2つの個々のE. coli株の中から選択される遺伝的バリエーションの組合せである遺伝的バリエーションを含む、ステップとを含む、方法。

実施形態20. 前記最初のE. coli株のライブラリーが、前記親系統E. coli株と前記産生E. coli株との間の前記同定された遺伝的バリエーションのすべてを含む完全コンビナトリアルライブラリーである、実施形態19に記載の産生E. coli株の表現型性能を再建および改善するための方法。

実施形態21. 前記最初のE. coli株のライブラリーが、前記親系統E. coli株と前記産生E. coli株との間の前記同定された遺伝的バリエーションのサブセットを含む完全コンビナトリアルライブラリーのサブセットである、実施形態19に記載の産生E. coli株の表現型性能を再建および改善するための方法。

実施形態22. 前記後続のE. coli株のライブラリーが、前記最初のライブラリーの完全コンビナトリアルライブラリーである、実施形態19に記載の産生E. coli株の表現型性能を再建および改善するための方法。

実施形態23. 前記後続のE. coli株のライブラリーが、前記最初のライブラリーの完全コンビナトリアルライブラリーのサブセットである、実施形態19に記載の産生E. coli株の表現型性能を再建および改善するための方法。

実施形態24. 前記後続のE. coli株のライブラリーが、先行するライブラリーの完全コンビナトリアルライブラリーである、実施形態19に記載の産生E. coli株の表現型性能を再建および改善するための方法。

実施形態25. 前記後続のE. coli株のライブラリーが、先行するライブラリーの完全コンビナトリアルライブラリーのサブセットである、実施形態19に記載の産生E. coli株の表現型性能を再建および改善するための方法。

実施形態26. 前記親系統E. coli株の前記ゲノムが攪乱されて、前記産生E. coli株内に見出される前記同定された一塩基多型、DNA挿入、またはDNA欠失の1つまたは複数が付加される、実施形態19から25までのいずれか1つに記載の産生E. coli株の表現型性能を再建および改善するための方法。

実施形態27. 前記産生E. coli株の前記ゲノムが攪乱されて、前記親系統E. coli株内に見出されない前記同定された一塩基多型、DNA挿入、またはDNA欠失の1つまたは複数が除去される、実施形態19から25までのいずれか1つに記載の産生E. coli株の表現型性能を再建および改善するための方法。

実施形態28. 前記ゲノムを攪乱するステップが、ランダム変異誘発、標的配列挿入、標的配列欠失、標的配列の置換、およびこれらの組合せからなる群から選択される少なくとも1つの方法を利用することを含む、実施形態19から27までのいずれか1つに記載の産生E. coli株の表現型性能を再建および改善するための方法。

実施形態29. 後続のライブラリーのE. coli株の前記表現型性能が、前記産生E. coli株の前記表現型性能と比較して、測定された表現型変数における少なくとも10%の増加を呈するまで、ステップd)～e)を繰り返す、実施形態19から28までのいずれか1つに記載の産生E. coli株の表現型性能を再建および改善するための方法。

実施形態30. 後続のライブラリーのE. coli株の前記表現型性能が、前記産生E. coli株の前記表現型性能と比較して、測定された表現型変数における少なくとも1倍の増加を呈するまで、ステップd)～e)を繰り返す、実施形態19から28までのいずれか1つに記載の産生E. coli株の表現型性能を再建および改善するための方法。

実施形態31. ステップf)の表現型性能の改善が、目的の産物の体積生産性、目的の産物の比生産性、目的の産物の収率、目的の産物の力価、およびこれらの組合せからなる群から選択される、実施形態19、29および30のいずれか1つに記載の産生E. coli株の表現型性能を再建および改善するための方法。

10

20

30

40

50

実施形態 3 2. ステップ f) の表現型性能の改善が、目的の産物の増大したまたはより効率的な産生であり、前記目的の産物は、低分子、酵素、ペプチド、アミノ酸、有機酸、合成化合物、燃料、アルコール、一次細胞外代謝産物、二次細胞外代謝産物、細胞内コンポーネント分子、およびこれらの組合せからなる群から選択される、実施形態 1 9、2 9 および 3 0 のいずれか 1 つに記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

実施形態 3 3. 前記同定された遺伝的バリエーションが、プロモータースワップライブラリー由来の人工プロモータースワップ遺伝的バリエーションをさらに含む、実施形態 1 9 から 3 2 までのいずれか 1 つに記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

10

実施形態 3 4. 前記最初の E. coli 株のライブラリー、または
後続の E. coli 株のライブラリー

のいずれかの少なくとも 1 つの微生物株のゲノムを、内在性 E. coli 標的遺伝子に作動可能に連結したプロモーターラダー由来の 1 つまたは複数のプロモーターを含むように操作するステップ
をさらに含む、実施形態 1 9 から 3 3 までのいずれか 1 つに記載の産生 E. coli 株の表現型性能を再建および改善するための方法。

実施形態 3 5. プロモータースワップ E. coli 株ライブラリーを生成するための方法であって、

a. 基準 E. coli 株に内在する複数の標的遺伝子、およびプロモーターラダーを提供するステップであって、前記プロモーターラダーは、前記基準 E. coli 株において異なる発現プロファイルを呈する複数のプロモーターを含む、ステップと；

20

b. 前記基準 E. coli 株のゲノムを操作することによって、複数の個々の E. coli 株を含む最初のプロモータースワップ E. coli 株ライブラリーを創製するステップであって、前記複数の個々の E. coli 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記基準 E. coli 株に内在する前記標的遺伝子の 1 つに作動可能に連結した前記プロモーターラダー由来の前記プロモーターの 1 つまたは複数を含む、ステップと
を含む、方法。

実施形態 3 6. 前記複数のプロモーターの少なくとも 1 つが、バイシストロン設計 (BCD) 調節配列を含む、実施形態 3 5 に記載のプロモータースワップ E. coli 株ライブラリーを生成するための方法。

30

実施形態 3 7. 産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法であって、

a. 基準 E. coli 株に内在する複数の標的遺伝子、およびプロモーターラダーを提供するステップであって、前記プロモーターラダーは、前記基準 E. coli 株において異なる発現プロファイルを呈する複数のプロモーターを含む、ステップと；

b. 前記基準 E. coli 株のゲノムを操作することによって、複数の個々の E. coli 株を含む最初のプロモータースワップ E. coli 株ライブラリーを創製するステップであって、前記複数の個々の E. coli 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記基準 E. coli 株に内在する前記標的遺伝子の 1 つに作動可能に連結した前記プロモーターラダー由来の前記プロモーターの 1 つまたは複数を含む、ステップと；

40

c. 参照 E. coli 株に対する表現型性能改善について前記最初のプロモータースワップ E. coli 株ライブラリーの個々の E. coli 株をスクリーニングおよび選択することによって、表現型性能改善をもたらすユニークな遺伝的バリエーションを同定するステップと；

d. 前記先行するステップにおいてスクリーニングされた少なくとも 2 つの個々の E. coli 株に存在する遺伝的バリエーションからのユニークな遺伝的バリエーションの組合せをそれぞれが含む後続の複数の E. coli 微生物を提供することによって、後続の

50

プロモータースワップ E. coli 株ライブラリーを創製するステップと；

e. 前記参照 E. coli 株に対する表現型性能改善について前記後続のプロモータースワップ E. coli 株ライブラリーの個々の E. coli 株をスクリーニングおよび選択することによって、追加の表現型性能改善をもたらす遺伝的バリエーションのユニークな組合せを同定するステップと；

f. E. coli 株が、前記産生 E. coli 株の前記表現型性能と比較して、所望のレベルの表現型性能の改善を呈するまで、線形または非線形様式でステップ d) ~ e) を 1 回または複数回繰り返すステップであって、各後続の反復により、微生物株の新しいプロモータースワップ E. coli 株ライブラリーが創製され、前記新しいライブラリーにおける各株は、先行するライブラリーの少なくとも 2 つの個々の E. coli 株の中から

10

選択される遺伝的バリエーションの組合せである遺伝的バリエーションを含む、ステップと

を含む、方法。
実施形態 37. 1 複数のプロモーターの少なくとも 1 つが、バイシストロン設計 (BCD) 調節配列を含む、実施形態 37 に記載の産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法。

実施形態 38. 前記後続のプロモータースワップ E. coli 株ライブラリーが、前記最初のプロモータースワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである、実施形態 37 または 37. 1 に記載の産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法。

20

実施形態 39. 前記後続のプロモータースワップ E. coli 株ライブラリーが、前記最初のプロモータースワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである、実施形態 37 または 37. 1 に記載の産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法。

実施形態 40. 前記後続のプロモータースワップ E. coli 株ライブラリーが、先行するプロモータースワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである、実施形態 37 または 37. 1 に記載の産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法。

実施形態 41. 前記後続のプロモータースワップ E. coli 株ライブラリーが、先行するプロモータースワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである、実施形態 37 または 37. 1 に記載の産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法。

30

実施形態 42. 後続のプロモータースワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 10 % の増加を呈するまで、ステップ d) ~ e) を繰り返す、実施形態 37 から 41 までのいずれか 1 つに記載の産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法。

実施形態 43. 後続のプロモータースワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 1 倍の増加を呈するまで、ステップ d) ~ e) を繰り返す、実施形態 37 から 41 までのいずれか 1 つに記載の産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法。

40

実施形態 44. ステップ f) の表現型性能の改善が、目的の産物の体積生産性、目的の産物の比生産性、目的の産物の収率、目的の産物の力価、およびこれらの組合せからなる群から選択される、実施形態 37、42 および 43 のいずれか 1 つに記載の産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法。

実施形態 45. ステップ f) の表現型性能の改善が、目的の産物の増大したまたはより効率的な産生であり、前記目的の産物は、低分子、酵素、ペプチド、アミノ酸、有機酸、合成化合物、燃料、アルコール、一次細胞外代謝産物、二次細胞外代謝産物、細胞内コンポーネント分子、およびこれらの組合せからなる群から選択される、実施形態 37、4

50

2 および 4 3 のいずれか 1 つに記載の産生 E. coli 株の表現型性能を改善するためのプロモータースワップ方法。

実施形態 4 6. ターミネータースワップ E. coli 株ライブラリーを生成するための方法であって、

a. 基準 E. coli 株に内在する複数の標的遺伝子、およびターミネーターラダーを提供するステップであって、前記ターミネーターラダーは、前記基準 E. coli 株において異なる発現プロファイルを呈する複数のターミネーターを含む、ステップと；

b. 前記基準 E. coli 株のゲノムを操作することによって、複数の個々の E. coli 株を含む最初のターミネータースワップ E. coli 株ライブラリーを創製するステップであって、前記複数の個々の E. coli 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記基準 E. coli 株に内在する前記標的遺伝子の 1 つに作動可能に連結した前記ターミネーターラダー由来の前記ターミネーターの 1 つまたは複数を含む、ステップとを含む、方法。

10

実施形態 4 7. 産生 E. coli 株の表現型性能を改善するためのターミネータースワップ方法であって、

a. 基準 E. coli 株に内在する複数の標的遺伝子、およびターミネーターラダーを提供するステップであって、前記ターミネーターラダーは、前記基準 E. coli 株において異なる発現プロファイルを呈する複数のターミネーターを含む、ステップと；

b. 前記基準 E. coli 株のゲノムを操作することによって、複数の個々の E. coli 株を含む最初のターミネータースワップ E. coli 株ライブラリーを創製するステップであって、前記複数の個々の E. coli 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記基準 E. coli 株に内在する前記標的遺伝子の 1 つに作動可能に連結した前記ターミネーターラダー由来の前記ターミネーターの 1 つまたは複数を含む、ステップと；

20

c. 参照 E. coli 株に対する表現型性能改善について前記最初のターミネータースワップ E. coli 株ライブラリーの個々の E. coli 株をスクリーニングおよび選択することによって、表現型性能改善をもたらすユニークな遺伝的バリエーションを同定するステップと；

d. 前記先行するステップにおいてスクリーニングされた少なくとも 2 つの個々の E. coli 株に存在する遺伝的バリエーションからのユニークな遺伝的バリエーションの組合せをそれぞれが含む後続の複数の E. coli 微生物を提供することによって、後続のターミネータースワップ E. coli 株ライブラリーを創製するステップと；

30

e. 前記参照 E. coli 株に対する表現型性能改善について前記後続のターミネータースワップ E. coli 株ライブラリーの個々の E. coli 株をスクリーニングおよび選択することによって、追加の表現型性能改善をもたらす遺伝的バリエーションのユニークな組合せを同定するステップと；

f. E. coli 株が、前記産生 E. coli 株の前記表現型性能と比較して、所望のレベルの表現型性能の改善を呈するまで、線形または非線形様式でステップ d) ~ e) を 1 回または複数回繰り返すステップであって、各後続の反復により、微生物株の新しいターミネータースワップ E. coli 株ライブラリーが創製され、前記新しいライブラリーにおける各株は、先行するライブラリーの少なくとも 2 つの個々の E. coli 株の中から選択される遺伝的バリエーションの組合せである遺伝的バリエーションを含む、ステップとを含む、方法。

40

実施形態 4 8. 前記後続のターミネータースワップ E. coli 株ライブラリーが、前記最初のターミネータースワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである、実施形態 4 7 に記載の産生 E. coli 株の表現型性能を改善するためのターミネータースワップ方法。

実施形態 4 9. 前記後続のターミネータースワップ E. coli 株ライブラリーが、

50

前記最初のターミネータースワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである、実施形態 47 に記載の産生 E. coli 株の表現型性能を改善するためのターミネータースワップ方法。

実施形態 50. 前記後続のターミネータースワップ E. coli 株ライブラリーが、先行するターミネータースワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである、実施形態 47 に記載の産生 E. coli 株の表現型性能を改善するためのターミネータースワップ方法。

実施形態 51. 前記後続のターミネータースワップ E. coli 株ライブラリーが、先行するターミネータースワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである、実施形態 47 に記載の産生 E. coli 株の表現型性能を改善するためのターミネータースワップ方法。 10

実施形態 52. 後続のターミネータースワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 10% の増加を呈するまで、ステップ d) ~ e) を繰り返す、実施形態 47 から 51 までのいずれか 1 つに記載の産生 E. coli 株の表現型性能を改善するためのターミネータースワップ方法。

実施形態 53. 後続のターミネータースワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 1 倍の増加を呈するまで、ステップ d) ~ e) を繰り返す、実施形態 47 から 51 までのいずれか 1 つに記載の産生 E. coli 株の表現型性能を改善するためのターミネータースワップ方法。 20

実施形態 54. ステップ f) の表現型性能の改善が、目的の産物の体積生産性、目的の産物の比生産性、目的の産物の収率、目的の産物の力価、およびこれらの組合せからなる群から選択される、実施形態 47 から 51 までのいずれか 1 つに記載の産生 E. coli 株の表現型性能を改善するためのターミネータースワップ方法。

実施形態 55. ステップ f) の表現型性能の改善が、目的の産物の増大したまたはより効率的な産生であり、前記目的の産物は、低分子、酵素、ペプチド、アミノ酸、有機酸、合成化合物、燃料、アルコール、一次細胞外代謝産物、二次細胞外代謝産物、細胞内コンポーネント分子、およびこれらの組合せからなる群から選択される、実施形態 47 から 51 までのいずれか 1 つに記載の産生 E. coli 株の表現型性能を改善するためのターミネータースワップ方法。 30

実施形態 56. E. coli 宿主細胞における生合成経路由来の生合成酵素を共局在化させるためのシステムであって、

a. 酵素反応に関与する 2 種またはそれよりも多くのキメラ酵素タンパク質であって、各キメラ酵素タンパク質は、DNA 結合ドメイン部分とカップリングした酵素部分を含む、2 種またはそれよりも多くのキメラ酵素タンパク質；

b.

i. 少なくとも 1 つの核酸スパーサーによって分離された 2 つまたはそれよりも多くの異なる DNA 結合部位をそれぞれが含む 1 つまたは複数のサブユニットを含む DNA 足場 40

を含む、前記キメラ酵素タンパク質は、そのそれぞれが、前記 DNA 足場内の少なくとも 1 つの DNA 結合部位に結合する、該キメラ酵素タンパク質とカップリングした DNA 結合ドメイン部分によって前記 DNA 足場に補充される、システム。

実施形態 57. 前記キメラ酵素タンパク質の前記 DNA 結合ドメイン部分が、ジンクフィンガー DNA 結合ドメインを含み、前記 DNA 足場の前記 DNA 結合部位が、対応するジンクフィンガー結合配列を含む、実施形態 56 に記載のシステム。

実施形態 58. 前記 2 種またはそれよりも多くのキメラ酵素タンパク質のそれぞれの前記酵素部分が、そのそれぞれの DNA 結合ドメイン部分と、ポリペプチドリンカー配列 50

を介してカップリングしている、実施形態 56 または 57 に記載のシステム。

実施形態 59. 前記 2 種またはそれよりも多くのキメラ酵素タンパク質のそれぞれの前記酵素部分が、そのそれぞれの DNA 結合ドメイン部分と、そのアミノ末端またはそのカルボキシ末端を介してカップリングしている、実施形態 56 から 58 までのいずれか 1 つに記載のシステム。

実施形態 60. 前記 2 種またはそれよりも多くのキメラ酵素タンパク質が、アミノ酸生合成経路の酵素を含む、実施形態 56 から 59 までのいずれか 1 つに記載のシステム。

実施形態 61. a. 作動可能に連結したプロモーター；

b. 第 1 のリボソーム結合部位 (SD1)；

c. 第 1 のシストロン配列 (Cis1)；

d. 第 2 のリボソーム結合部位 (SD2)

を順に含むバイシストロン設計調節 (BCD) 配列であって、

標的遺伝子配列 (Cis2) に作動可能に連結している、BCD 配列。

実施形態 62. SD1 および SD2 がそれぞれ N N N G G A N N N の配列を含む、実施形態 61 に記載の BCD。

実施形態 63. SD1 と SD2 が異なる、実施形態 61 または 62 に記載の BCD。

実施形態 64. Cis1 が停止コドンを含み、Cis2 が開始コドンを含み、前記 Cis1 停止コドンと前記 Cis2 開始コドンが少なくとも 1 ヌクレオチド重複している、実施形態 61 から 63 までのいずれか 1 つに記載の BCD。

実施形態 65. SD2 が Cis1 内に完全に組み込まれている、実施形態 61 から 63 までのいずれか 1 つに記載の BCD。

実施形態 66. 2 種の標的遺伝子タンパク質を宿主生物において発現させるための方法であって、

a. 前記宿主生物に、第 1 の標的遺伝子タンパク質をコードする第 1 のポリヌクレオチドを導入するステップであって、前記第 1 のポリヌクレオチドは、第 1 の実施形態 61 から 65 までのいずれか 1 つに記載のバイシストロン設計調節 (BCD) 配列に作動可能に連結している、ステップと；

b. 前記宿主生物に、第 2 の標的遺伝子タンパク質をコードする第 2 のポリヌクレオチドを導入するステップであって、前記第 2 のポリヌクレオチドは、第 2 の実施形態 61 から 65 までのいずれか 1 つに記載の BCD に作動可能に連結している、ステップとを含む、

前記第 1 および第 2 の BCD は、それらのそれぞれの Cis1 配列以外は同一であり、前記標的遺伝子タンパク質は、前記宿主生物においてそれぞれ第 1 および第 2 の発現レベルで発現される、方法。

実施形態 67. 前記第 1 の発現レベルが前記第 2 の発現レベルの 1.5 倍以内である、実施形態 66 に記載の方法。

実施形態 68. 前記第 1 および第 2 のポリヌクレオチドが、前記宿主細胞において、前記第 1 および第 2 のポリヌクレオチドが同一の BCD によって発現される対照宿主細胞と比較して低レベルの相同組換えを受ける、実施形態 66 または 67 に記載の方法。

実施形態 69. タンパク質可溶性タグスワップ E. coli 株ライブラリーを生成するための方法であって、

a. 基準 E. coli 株に内在する複数の標的遺伝子、および可溶性タグラダーを提供するステップであって、前記可溶性タグラダーは、前記基準 E. coli 株において異なる可溶性プロファイルを呈する複数の可溶性タグを含む、ステップと；

b. 前記基準 E. coli 株のゲノムを操作することによって、複数の個々の E. coli 株を含む最初の可溶性タグスワップ E. coli 株ライブラリーを創製するステップであって、前記複数の個々の E. coli 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記基準 E. coli 株に内在する前記標的遺伝子の 1 つに作動可能に連結した前記可溶性タグラダー由来の前記可溶性タグの 1 つまたは複数を含む、ステップと

10

20

30

40

50

を含む、方法。

実施形態 70. 産生 E. coli 株の表現型性能を改善するためのタンパク質可溶性タグスワップ方法であって、

a. 基準 E. coli 株に内在する複数の標的遺伝子、および可溶性タグラダーを提供するステップであって、前記可溶性タグラダーは、前記基準 E. coli 株において異なる発現プロファイルを呈する複数の可溶性タグを含む、ステップと；

b. 前記基準 E. coli 株のゲノムを操作することによって、複数の個々の E. coli 株を含む最初の可溶性タグスワップ E. coli 株ライブラリーを創製するステップであって、前記複数の個々の E. coli 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記基準 E. coli 株に内在する前記標的遺伝子の 1 つに作動可能に連結した前記可溶性タグラダー由来の前記可溶性タグの 1 つまたは複数を含む、ステップと；

c. 参照 E. coli 株に対する表現型性能改善について前記最初の可溶性タグスワップ E. coli 株ライブラリーの個々の E. coli 株をスクリーニングおよび選択することによって、表現型性能改善をもたらすユニークな遺伝的バリエーションを同定するステップと；

d. 前記先行するステップにおいてスクリーニングされた少なくとも 2 つの個々の E. coli 株に存在する遺伝的バリエーションからのユニークな遺伝的バリエーションの組合せをそれぞれが含む後続の複数の E. coli 微生物を提供することによって、後続の可溶性タグスワップ E. coli 株ライブラリーを創製するステップと；

e. 前記参照 E. coli 株に対する表現型性能改善について前記後続の可溶性タグスワップ E. coli 株ライブラリーの個々の E. coli 株をスクリーニングおよび選択することによって、追加の表現型性能改善をもたらす遺伝的バリエーションのユニークな組合せを同定するステップと；

f. E. coli 株が、前記産生 E. coli 株の前記表現型性能と比較して、所望のレベルの表現型性能の改善を呈するまで、線形または非線形様式でステップ d) ~ e) を 1 回または複数回繰り返すステップであって、各後続の反復により、微生物株の新しい可溶性タグスワップ E. coli 株ライブラリーが創製され、前記新しいライブラリー内の各株は、先行するライブラリーの少なくとも 2 つの個々の E. coli 株の中から選択される遺伝的バリエーションの組合せである遺伝的バリエーションを含む、ステップとを含む、方法。

実施形態 71. 前記後続の可溶性タグスワップ E. coli 株ライブラリーが、前記最初の可溶性タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである、実施形態 70 に記載の産生 E. coli 株の表現型性能を改善するための可溶性タグスワップ方法。

実施形態 72. 前記後続の可溶性タグスワップ E. coli 株ライブラリーが、前記最初の可溶性タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである、実施形態 70 に記載の産生 E. coli 株の表現型性能を改善するための可溶性タグスワップ方法。

実施形態 73. 前記後続の可溶性タグスワップ E. coli 株ライブラリーが、先行する可溶性タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである、実施形態 70 に記載の産生 E. coli 株の表現型性能を改善するための可溶性タグスワップ方法。

実施形態 74. 前記後続の可溶性タグスワップ E. coli 株ライブラリーが、先行する可溶性タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである、実施形態 70 に記載の産生 E. coli 株の表現型性能を改善するための可溶性タグスワップ方法。

実施形態 75. 後続の可溶性タグスワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 10 % の増加を呈するまで、ステップ d) ~ e) を繰り返

す、実施形態 70 から 74 までのいずれか 1 つに記載の産生 E. coli 株の表現型性能を改善するための可溶性タグスワップ方法。

実施形態 76. 後続の可溶性タグスワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 1 倍の増加を呈するまで、ステップ d) ~ e) を繰り返す、実施形態 70 から 74 までのいずれか 1 つに記載の産生 E. coli 株の表現型性能を改善するための可溶性タグスワップ方法。

実施形態 77. ステップ f) の表現型性能の改善が、目的の産物の体積生産性、目的の産物の比生産性、目的の産物の収率、目的の産物の力価、およびこれらの組合せからなる群から選択される、実施形態 70、75 および 76 のいずれか 1 つに記載の産生 E. coli 株の表現型性能を改善するための可溶性タグスワップ方法。

実施形態 78. ステップ f) の表現型性能の改善が、目的の産物の増大したまたはより効率的な産生であり、前記目的の産物は、低分子、酵素、ペプチド、アミノ酸、有機酸、合成化合物、燃料、アルコール、一次細胞外代謝産物、二次的細胞外代謝産物、細胞内コンポーネント分子、およびこれらの組合せからなる群から選択される、実施形態 70、75 および 76 のいずれか 1 つに記載の産生 E. coli 株の表現型性能を改善するための可溶性タグスワップ方法。

実施形態 79. タンパク質分解タグスワップ E. coli 株ライブラリーを生成するための方法であって、

a. 基準 E. coli 株に内在する複数の標的遺伝子、および分解タグラダーを提供するステップであって、前記分解タグラダーは、前記基準 E. coli 株において異なる溶解性プロファイルを呈する複数の分解タグを含む、ステップと；

b. 前記基準 E. coli 株のゲノムを操作することによって、複数の個々の E. coli 株を含む最初の分解タグスワップ E. coli 株ライブラリーを創製するステップであって、前記複数の個々の E. coli 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記基準 E. coli 株に内在する前記標的遺伝子の 1 つに作動可能に連結した前記分解タグラダー由来の前記分解タグの 1 つまたは複数を含む、ステップとを含む、方法。

実施形態 80. 産生 E. coli 株の表現型性能を改善するためのタンパク質分解タグスワップ方法であって、

a. 基準 E. coli 株に内在する複数の標的遺伝子、および分解タグラダーを提供するステップであって、前記分解タグラダーは、前記基準 E. coli 株において異なる発現プロファイルを呈する複数の分解タグを含む、ステップと；

b. 前記基準 E. coli 株のゲノムを操作することによって、複数の個々の E. coli 株を含む最初の分解タグスワップ E. coli 株ライブラリーを創製するステップであって、前記複数の個々の E. coli 株の各株内にユニークな遺伝的バリエーションが見出され、前記ユニークな遺伝的バリエーションのそれぞれは、前記基準 E. coli 株に内在する前記標的遺伝子の 1 つに作動可能に連結した前記分解タグラダー由来の前記分解タグの 1 つまたは複数を含む、ステップと；

c. 参照 E. coli 株に対する表現型性能改善について前記最初の分解タグスワップ E. coli 株ライブラリーの個々の E. coli 株をスクリーニングおよび選択することによって、表現型性能改善をもたらすユニークな遺伝的バリエーションを同定するステップと、

d. 前記先行するステップにおいてスクリーニングされた少なくとも 2 つの個々の E. coli 株に存在する遺伝的バリエーションからのユニークな遺伝的バリエーションの組合せをそれぞれが含む後続の複数の E. coli 微生物を提供することによって、後続の分解タグスワップ E. coli 株ライブラリーを創製するステップと；

e. 前記参照 E. coli 株に対する表現型性能改善について前記後続の分解タグスワップ E. coli 株ライブラリーの個々の E. coli 株をスクリーニングおよび選択す

10

20

30

40

50

ることによって、追加の表現型性能改善をもたらす遺伝的バリエーションのユニークな組合せを同定するステップと；

f. E. coli 株が、前記産生 E. coli 株の前記表現型性能と比較して、所望のレベルの表現型性能の改善を呈するまで、線形または非線形様式でステップ d) ~ e) を 1 回または複数回繰り返すステップであって、各後続の反復により、微生物株の新しい分解タグスワップ E. coli 株ライブラリーが創製され、前記新しいライブラリー内の各株は、先行するライブラリーの少なくとも 2 つの個々の E. coli 株の中から選択される遺伝的バリエーションの組合せである遺伝的バリエーションを含む、ステップとを含む、方法。

実施形態 81. 前記後続の分解タグスワップ E. coli 株ライブラリーが、前記最初の分解タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである、実施形態 80 に記載の産生 E. coli 株の表現型性能を改善するための分解タグスワップ方法。

実施形態 82. 前記後続の分解タグスワップ E. coli 株ライブラリーが、前記最初の分解タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである、実施形態 80 に記載の産生 E. coli 株の表現型性能を改善するための分解タグスワップ方法。

実施形態 83. 前記後続の分解タグスワップ E. coli 株ライブラリーが、先行する分解タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーである、実施形態 80 に記載の産生 E. coli 株の表現型性能を改善するための分解タグスワップ方法。

実施形態 84. 前記後続の分解タグスワップ E. coli 株ライブラリーが、先行する分解タグスワップ E. coli 株ライブラリーの完全コンビナトリアルライブラリーのサブセットである、実施形態 80 に記載の産生 E. coli 株の表現型性能を改善するための分解タグスワップ方法。

実施形態 85. 後続の分解タグスワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 10 % の増加を呈するまで、ステップ d) ~ e) を繰り返す、実施形態 80 から 84 までのいずれか 1 つに記載の産生 E. coli 株の表現型性能を改善するための分解タグスワップ方法。

実施形態 86. 後続の分解タグスワップ E. coli 株ライブラリーの E. coli 株の表現型性能が、前記産生 E. coli 株の前記表現型性能と比較して、測定された表現型変数における少なくとも 1 倍の増加を呈するまで、ステップ d) ~ e) を繰り返す、実施形態 80 から 84 までのいずれか 1 つに記載の産生 E. coli 株の表現型性能を改善するための分解タグスワップ方法。

実施形態 87. ステップ f) の表現型性能の改善が、目的の産物の体積生産性、目的の産物の比生産性、目的の産物の収率、目的の産物の力価、およびこれらの組合せからなる群から選択される、実施形態 80、85 および 86 のいずれか 1 つに記載の産生 E. coli 株の表現型性能を改善するための分解タグスワップ方法。

実施形態 88. ステップ f) の表現型性能の改善が、目的の産物の増大したまたはより効率的な産生であり、前記目的の産物は、低分子、酵素、ペプチド、アミノ酸、有機酸、合成化合物、燃料、アルコール、一次細胞外代謝産物、二次的細胞外代謝産物、細胞内コンポーネント分子、およびこれらの組合せからなる群から選択される、実施形態 80、85 および 86 までのいずれか 1 つに記載の産生 E. coli 株の表現型性能を改善するための分解タグスワップ方法。

実施形態 89. 微生物宿主細胞における発現のための、異種遺伝子に作動可能に連結したキメラ合成プロモーターであって、長さが 60 ~ 90 ヌクレオチドであり、ラムダファージ p_R プロモーターの遠位部分、それぞれ長さが 6 ヌクレオチドであるラムダファージ p_L および p_R プロモーターの可変 -35 および -10 領域、ラムダファージ p_L および p_R プロモーターのコア部分、ならびにラムダファージ p_R プロモーターの 5' UTR

10

20

30

40

50

／リボソーム結合部位（RBS）部分からなる、キメラ合成プロモーター。

実施形態90． 前記ラムダファージ p_R プロモーターの前記遠位部分、前記ラムダファージ p_L および p_R プロモーターの前記可変-35および-10領域、前記ラムダファージ p_L および p_R プロモーターの前記コア部分、ならびに前記ラムダファージ p_R プロモーターの前記5'UTR／リボソーム結合部位（RBS）部分の核酸配列が、表1.5に見出される核酸配列から選択される、実施形態89に記載のキメラ合成プロモーター。

実施形態91． 微生物宿主細胞における発現のための、異種遺伝子に作動可能に連結したキメラ合成プロモーターであって、長さが60～90ヌクレオチドであり、ラムダファージ p_R プロモーターの遠位部分、それぞれ長さが6ヌクレオチドであるラムダファージ p_L および p_R プロモーターの可変-35および-10領域、ラムダファージ p_L および p_R プロモーターのコア部分ならびにE. coli *accs* 遺伝子のプロモーターの5'UTR／リボソーム結合部位（RBS）部分からなる、キメラ合成プロモーター。

実施形態92． 前記ラムダファージ p_R プロモーターの前記遠位部分、前記ラムダファージ p_L および p_R プロモーターの前記可変-35および-10領域、前記ラムダファージ p_L および p_R プロモーターの前記コア部分、ならびに前記E. coli *accs* 遺伝子の前記プロモーターの前記5'UTR／リボソーム結合部位（RBS）部分の核酸配列が、表1.5に見出される核酸配列から選択される、実施形態91に記載のキメラ合成プロモーター。

実施形態93． 配列番号132～152、159～160、162、165、174～175、188、190、199～201または207から選択される核酸配列からなる、実施形態89から90までのいずれかに記載のキメラ合成プロモーター。

実施形態94． 配列番号153～158、161、163～164、166～173、176～187、189、191～198または202～206から選択される核酸配列からなる、実施形態91から92までのいずれかに記載のキメラ合成プロモーター。

実施形態95． 前記微生物宿主細胞がE. coliである、実施形態89から94までのいずれかに記載のキメラ合成プロモーター。

実施形態96． 前記異種遺伝子が、表2に見出される目的のタンパク質産物をコードする、実施形態95に記載のキメラ合成プロモーター。

実施形態97． 前記異種遺伝子が、リシン生合成経路の一部である遺伝子である、実施形態95に記載のキメラ合成プロモーター。

実施形態98． 前記異種遺伝子が、*asd* 遺伝子、*ask* 遺伝子、*hom* 遺伝子、*dapA* 遺伝子、*dapB* 遺伝子、*dapD* 遺伝子、*ddh* 遺伝子、*argD* 遺伝子、*dapE* 遺伝子、*dapF* 遺伝子、*lysA* 遺伝子、*lysE* 遺伝子、*zwf* 遺伝子、*pgi* 遺伝子、*ktk* 遺伝子、*fbp* 遺伝子、*ppc* 遺伝子、*pcK* 遺伝子、*ddx* 遺伝子、*pyc* 遺伝子または*icd* 遺伝子から選択される、実施形態97に記載のキメラ合成プロモーター。

実施形態99． 前記異種遺伝子が、リコピン生合成経路の一部である遺伝子である、実施形態95に記載のキメラ合成プロモーター。

実施形態100． 前記異種遺伝子が、*dxs* 遺伝子、*ispC* 遺伝子、*ispE* 遺伝子、*ispD* 遺伝子、*ispF* 遺伝子、*ispG* 遺伝子、*ispH* 遺伝子、*idi* 遺伝子、*ispA* 遺伝子、*ispB* 遺伝子、*crtE* 遺伝子、*crtB* 遺伝子、*crtI* 遺伝子、*crtY* 遺伝子、*ymgA* 遺伝子、*dxr* 遺伝子、*elbA* 遺伝子、*gdhA* 遺伝子、*appY* 遺伝子、*elbB* 遺伝子、または*ymgB* 遺伝子から選択される、実施形態99に記載のキメラ合成プロモーター。

実施形態101． 前記異種遺伝子が、生物学的製剤をコードするまたは生物学的製剤を生成するための経路内の遺伝子である、実施形態95に記載のキメラ合成プロモーター。

実施形態102． 前記生物学的製剤が、ヒューマリン（rhインスリン）、イントロンA（インターフェロンアルファ2b）、ロフェロン（インターフェロンアルファ2a）、ヒューマトロップ（ソマトロピンrh成長ホルモン）、ニューボジェン（フィルグラス

10

20

30

40

50

チム)、ベタフェロン(インターフェロンベーター1b)、リスプロ(速効型インスリン)、ラピリシン(レテプラゼ)、インフェルゲン(インターフェロンアルファコンー1)、グルカゴン、ベロムン(タソネルミン)、オンタック(デニロイキンジフチトクス)、ランタス(長時間作用型インスリングラルギン)、キネレット(アナキンラ)、ナトレコール(ネシリチド)、ソマバート(ペグビソマント)、カルシトニン(組換えカルシトニン サケ)、ルセンチス(ラニビズマブ)、プレオタクト(ヒト副甲状腺ホルモン)、クリステクサ(rh尿酸オキシダーゼ、PEG化)、ニベスチム(フィルグラスチム、rhGCSF)、ボラキサゼ(グルカルピダーゼ)、またはプレオス(副甲状腺ホルモン)から選択される、実施形態101に記載のキメラ合成プロモーター。

実施形態103. 配列番号132~207から選択される核酸配列を有するキメラ合成プロモーターに作動可能に連結した異種遺伝子。

10

実施形態104. 表2に見出される目的のタンパク質産物をコードする、実施形態103に記載の異種遺伝子。

実施形態105. リシン生合成経路の一部である遺伝子である、実施形態103に記載の異種遺伝子。

実施形態106. asd遺伝子、ask遺伝子、hom遺伝子、dapA遺伝子、dapB遺伝子、dapD遺伝子、ddh遺伝子、argD遺伝子、dapE遺伝子、dapF遺伝子、lysA遺伝子、lysE遺伝子、zwf遺伝子、pgi遺伝子、ktk遺伝子、fbp遺伝子、ppc遺伝子、pcK遺伝子、ddx遺伝子、pyc遺伝子またはicd遺伝子から選択される、実施形態105に記載の異種遺伝子。

20

実施形態107. リコピン生合成経路の一部である遺伝子である、実施形態103に記載の異種遺伝子。

実施形態108. dxs遺伝子、ispC遺伝子、ispE遺伝子、ispD遺伝子、ispF遺伝子、ispG遺伝子、ispH遺伝子、idi遺伝子、ispA遺伝子、ispB遺伝子、crtE遺伝子、crtB遺伝子、crtI遺伝子、crtY遺伝子、ymgA遺伝子、dxr遺伝子、elbA遺伝子、gdhA遺伝子、appY遺伝子、elbB遺伝子、またはymgB遺伝子から選択される、実施形態107に記載の異種遺伝子。

実施形態109. 生物学的製剤をコードするまたは生物学的製剤を生成するための経路内の遺伝子である、実施形態103に記載の異種遺伝子。

30

実施形態110. 前記生物学的製剤が、ヒューマリン(rhインスリン)、イントロンA(インターフェロンアルファ2b)、ロフェロン(インターフェロンアルファ2a)、ヒューマトロップ(ソマトロピンrh成長ホルモン)、ニューボジェン(フィルグラスチム)、ベタフェロン(インターフェロンベーター1b)、リスプロ(速効型インスリン)、ラピリシン(レテプラゼ)、インフェルゲン(インターフェロンアルファコンー1)、グルカゴン、ベロムン(タソネルミン)、オンタック(デニロイキンジフチトクス)、ランタス(長時間作用型インスリングラルギン)、キネレット(アナキンラ)、ナトレコール(ネシリチド)、ソマバート(ペグビソマント)、カルシトニン(組換えカルシトニン サケ)、ルセンチス(ラニビズマブ)、プレオタクト(ヒト副甲状腺ホルモン)、クリステクサ(rh尿酸オキシダーゼ、PEG化)、ニベスチム(フィルグラスチム、rhGCSF)、ボラキサゼ(グルカルピダーゼ)、またはプレオス(副甲状腺ホルモン)から選択される、実施形態109に記載の異種遺伝子。

40

【表 2 2 - 1】

配列番号識別子を伴う本開示の配列

名称(省略名)	源	核酸配列番号	アミノ酸配列番号	説明
Pcg0007_lib_39 (P1)		1		
Pcg0007 (P2)		2		
Pcg1860 (P3)		3		
Pcg0755 (P4)		4		
Pcg0007_265 (P5)		5		
Pcg3381 (P6)		6		
Pcg0007_119 (P7)		7		
Pcg3121 (P8)		8		
cg0001 (T1)		9		
cg0007 (T2)		10		
cg0371 (T3)		11		
cg0480 (T4)		12		
cg0494 (T5)		13		
cg0564 (T6)		14		
cg0610 (T7)		15		
cg0695 (T8)		16		
Cis1 ヌクレオチド配列		17		
セリンリッチリンカー		18		
セリンリッチリンカー		19		
セリンリッチリンカー		20		
アラニンリッチリンカー		21		
リンカー		22		

10

20

30

【表 2 2 - 2】

リンカー		23		
リンカー		24		
リンカー		25		
Zif268		26		DNA 結合ドメイン配列
PBSII		27		DNA 結合ドメイン配列
ZFa		28		DNA 結合ドメイン配列
ZFb		29		DNA 結合ドメイン配列
ZFc		30		DNA 結合ドメイン配列
Tyr123		31		DNA 結合ドメイン配列
Tyr456		32		DNA 結合ドメイン配列
Blues ジンクフィンガー		33		DNA 結合ドメイン配列
Jazz ジンクフィンガー		34		DNA 結合ドメイン配列
Bagly ジンクフィンガー		35		DNA 結合ドメイン配列

10

20

【表 2 2 - 3】

Bagly ジンクフィンガー結合部位		36		DNA 結合配列 (5' → 3')
Gli1		37		DNA 結合ドメイン配列
Gli1 結合部位		38		DNA 結合配列 (5' → 3')
HIVC ジンクフィンガー		39		DNA 結合ドメイン配列
B3 ジンクフィンガー		40		DNA 結合ドメイン配列
N1 ジンクフィンガー		41		DNA 結合ドメイン配列
Sp-1 第 1 ジンクフィンガー		42		DNA 結合ドメイン配列
Sp-1 第 2 フィンガー		43		DNA 結合ドメイン配列
クラス I SH3 タンパク質結合ドメイン			44	
クラス II SH3 タンパク質結合部位			45	
SH3 タンパク質結合部位			46	
SH3 動員ペプチド			47	
SH3 動員ペプチド			48	
SH3 動員ペプチド			49	
PDZ タンパク質結合ドメイン			50	
PDZ タンパク質結合ドメイン			51	

10

20

【表 2 2 - 4】

PDZ タンパク質結合ドメイン			52	
PDZ タンパク質結合ドメイン			53	
PDZ タンパク質結合ドメイン			54	
PDZ 動員ペプチド			55	
PDZ 動員ペプチド			56	
PDZ 動員ペプチド			57	
PDZ 動員ペプチド			58	
PDZ 動員ペプチド			59	
PDZ 動員ペプチド			60	
PDZ 動員ペプチド			61	
GBD タンパク質結合ドメイン			62	
GBD タンパク質結合ドメイン			63	
GBD タンパク質結合ドメイン			64	
GBD 動員ペプチド			65	
ロイシンジッパー結合ドメイン			66	
ロイシンジッパー結合ドメイン			67	
ロイシンジッパー結合ドメイン			68	
バイシストロン設計 (BCD) 調節配列			69	図 43 参照
バイシストロン設計 (BCD) 調節配列			70	図 43 参照
b0904_プロモーター	<i>E. coli</i>	71		
b2405_プロモーター	<i>E. coli</i>	72		
b0096_プロモーター	<i>E. coli</i>	73		
b0576_プロモーター	<i>E. coli</i>	74		
b2017_プロモーター	<i>E. coli</i>	75		
b1278_プロモーター	<i>E. coli</i>	76		
b4255_プロモーター	<i>E. coli</i>	77		
b0786_プロモーター	<i>E. coli</i>	78		
b0605_プロモーター	<i>E. coli</i>	79		
b1824_プロモーター	<i>E. coli</i>	80		
b1061_プロモーター	<i>E. coli</i>	81		

10

20

30

【表 2 2 - 5】

b0313_プロモーター	<i>E. coli</i>	82		
b0814_プロモーター	<i>E. coli</i>	83		
b4133_プロモーター	<i>E. coli</i>	84		
b4268_プロモーター	<i>E. coli</i>	85		
b0345_プロモーター	<i>E. coli</i>	86		
b2096_プロモーター	<i>E. coli</i>	87		
b1277_プロモーター	<i>E. coli</i>	88		
b1646_プロモーター	<i>E. coli</i>	89		
b4177_プロモーター	<i>E. coli</i>	90		
b0369_プロモーター	<i>E. coli</i>	91		
b1920_プロモーター	<i>E. coli</i>	92		
b3742_プロモーター	<i>E. coli</i>	93		
b3929_プロモーター	<i>E. coli</i>	94		
b3743_プロモーター	<i>E. coli</i>	95		
b1613_プロモーター	<i>E. coli</i>	96		
b1749_プロモーター	<i>E. coli</i>	97		
b2478_プロモーター	<i>E. coli</i>	98		
b0031_プロモーター	<i>E. coli</i>	99		
b2414_プロモーター	<i>E. coli</i>	100		
b1183_プロモーター	<i>E. coli</i>	101		
b0159_プロモーター	<i>E. coli</i>	102		
b2837_プロモーター	<i>E. coli</i>	103		
b3237_プロモーター	<i>E. coli</i>	104		
b3778_プロモーター	<i>E. coli</i>	105		
b2349_プロモーター	<i>E. coli</i>	106		
b1434_プロモーター	<i>E. coli</i>	107		
b3617_プロモーター	<i>E. coli</i>	108		
b0237_プロモーター	<i>E. coli</i>	109		
b4063_プロモーター	<i>E. coli</i>	110		
b0564_プロモーター	<i>E. coli</i>	111		
b0019_プロモーター	<i>E. coli</i>	112		
b2375_プロモーター	<i>E. coli</i>	113		

10

20

30

【表 2 2 - 6】

b1187_プロモーター	<i>E. coli</i>	114		
b2388_プロモーター	<i>E. coli</i>	115		
b1051_プロモーター	<i>E. coli</i>	116		
b4241_プロモーター	<i>E. coli</i>	117		
b4054_プロモーター	<i>E. coli</i>	118		
b2425_プロモーター	<i>E. coli</i>	119		
b0995_プロモーター	<i>E. coli</i>	120		
b1399_プロモーター	<i>E. coli</i>	121		
b3298_プロモーター	<i>E. coli</i>	122		
b2114_プロモーター	<i>E. coli</i>	123		
b2779_プロモーター	<i>E. coli</i>	124		
b1114_プロモーター	<i>E. coli</i>	125		
b3730_プロモーター	<i>E. coli</i>	126		
b3025_プロモーター	<i>E. coli</i>	127		
b0850_プロモーター	<i>E. coli</i>	128		
b2365_プロモーター	<i>E. coli</i>	129		
b4117_プロモーター	<i>E. coli</i>	130		
b2213_プロモーター	<i>E. coli</i>	131		
pMB029_プロモーター	合成	132		
pMB023_プロモーター	合成	133		
pMB025_プロモーター	合成	134		
pMB019_プロモーター	合成	135		
pMB008_プロモーター	合成	136		
pMB020_プロモーター	合成	137		
pMB022_プロモーター	合成	138		
pMB089_プロモーター	合成	139		
pMB001_プロモーター	合成	140		
pMB051_プロモーター	合成	141		
pMB070_プロモーター	合成	142		
pMB074_プロモーター	合成	143		
pMB046_プロモーター	合成	144		
pMB071_プロモーター	合成	145		

10

20

30

【表 2 2 - 7】

pMB013_プロモーター	合成	146		
pMB080_プロモーター	合成	147		
pMB038_プロモーター	合成	148		
pMB060_プロモーター	合成	149		
pMB064_プロモーター	合成	150		
pMB058_プロモーター	合成	151		
pMB085_プロモーター	合成	152		
pMB081_プロモーター	合成	153		
pMB091_プロモーター	合成	154		
pMB027_プロモーター	合成	155		
pMB048_プロモーター	合成	156		
pMB055_プロモーター	合成	157		
pMB006_プロモーター	合成	158		
pMB012_プロモーター	合成	159		
pMB014_プロモーター	合成	160		
pMB028_プロモーター	合成	161		
pMB059_プロモーター	合成	162		
pMB061_プロモーター	合成	163		
pMB043_プロモーター	合成	164		
pMB066_プロモーター	合成	165		
pMB079_プロモーター	合成	166		
pMB032_プロモーター	合成	167		
pMB068_プロモーター	合成	168		
pMB082_プロモーター	合成	169		
pMB030_プロモーター	合成	170		
pMB067_プロモーター	合成	171		
pMB050_プロモーター	合成	172		
pMB069_プロモーター	合成	173		
pMB017_プロモーター	合成	174		
pMB039_プロモーター	合成	175		
pMB011_プロモーター	合成	176		
pMB072_プロモーター	合成	177		

10

20

30

【表 2 2 - 8】

pMB016_プロモーター	合成	178		
pMB077_プロモーター	合成	179		
pMB047_プロモーター	合成	180		
pMB052_プロモーター	合成	181		
pMB090_プロモーター	合成	182		
pMB035_プロモーター	合成	183		
pMB073_プロモーター	合成	184		
pMB004_プロモーター	合成	185		
pMB054_プロモーター	合成	186		
pMB024_プロモーター	合成	187		
pMB007_プロモーター	合成	188		
pMB005_プロモーター	合成	189		
pMB003_プロモーター	合成	190		
pMB088_プロモーター	合成	191		
pMB065_プロモーター	合成	192		
pMB037_プロモーター	合成	193		
pMB009_プロモーター	合成	194		
pMB041_プロモーター	合成	195		
pMB036_プロモーター	合成	196		
pMB049_プロモーター	合成	197		
pMB044_プロモーター	合成	198		
pMB042_プロモーター	合成	199		
pMB086_プロモーター	合成	200		
pMB053_プロモーター	合成	201		
pMB057_プロモーター	合成	202		
pMB018_プロモーター	合成	203		
pMB002_プロモーター	合成	204		
pMB015_プロモーター	合成	205		
pMB087_プロモーター	合成	206		
pMB063_プロモーター	合成	207		
合成プロモーターの遠位部分	λ ファージ	208		R _h プロモーター

10

20

30

【表 2 2 - 9】

合成プロモーターのコア部分	λ ファージ	209		R_K プロモーター
合成プロモーターの遠位部分	λ ファージ	210		R_L プロモーター
合成プロモーターの 5' UTR/RBS 部分	λ ファージ	211		R_K プロモーター
合成プロモーターの 5' UTR/RBS 部分	<i>E. coli</i>	212		<i>acs</i> 遺伝子のプロモーター
Ori_Plsmd27		213		
ベクター骨格 1		214		
ベクター骨格 2		215		
ベクター骨格 3		216		
ベクター骨格 4		217		
インスレーター 1		218		表 15 参照
インスレーター 2		219		表 15 参照
T1		220		Orosz ら、Eur J Biochem、 1991 年 11 月 1 日:201 巻(3 号):653~9 頁 より;表 15 参照
B0015		221		表 15 参照
SacB プロモーター		222		表 15 参照
PheS プロモーター&配列		223		表 15 参照
Tsod ターミネーター	<i>C. glutamicum</i>	224		
Spy	<i>E. coli</i>	225		ターミネーター配列
pheA	<i>E. coli</i>	226		ターミネーター配列
osmE	<i>E. coli</i>	227		ターミネーター配列
rpoH	<i>E. coli</i>	228		ターミネーター配列
vibE	<i>E. coli</i>	229		ターミネーター配列

10

20

30

【表 2 2 - 1 0】

ThrI_ABC	<i>E. coli</i>	230		ターミネータ ー配列
GB1 (PST1)	<i>Streptococcus</i> sp.	231	235	Protein G の IgG ドメイン B1
FH8 (PST2)	<i>F. hepatica</i>	232	236	<i>Fasciola</i> <i>hepatica</i> 8-kDa 抗原
ユビキチン (PST3)		233	237	
SUMO (PST4)	<i>Homo sapiens</i>	234	238	低分子ユビキ チン修飾
ssrA_LAA (PDT1)	<i>E. coli</i>	239	247	生来の
ssrA_LVA (PDT2)	<i>E. coli</i>	240	248	変異体
ssrA_AAV (PDT3)	<i>E. coli</i>	241	249	変異体
ssrA_ASV (PDT4)	<i>E. coli</i>	242	250	変異体
ftsH-cll189-97 (PDT5)	<i>E. coli</i>	243	251	生来の
cll108 (PDT6)	<i>E. coli</i>	244	252	生来の
sul20 (PDT7)	<i>E. coli</i>	245	253	生来の
β 20 (PDT8)	<i>E. coli</i>	246	254	生来の
P3_BCD1	人工	255		
P4_BCD22	人工	256		
P7_BCD19	人工	257		
P8_BCD15	人工	258		
P11_BCD17	人工	259		
P13_BCD8	人工	260		

10

20

参照による組込み

本明細書に引用したすべての参考文献、論文、刊行物、特許、特許公開、および特許出願は、すべての目的に関してその全体が参照により組み込まれている。しかし、本明細書に引用した任意の参考文献、論文、刊行物、特許、特許公開、および特許出願の言及は、これらが有効な先行技術を構成するか、または世界のいずれかの国における共通の一般的知識の一部を形成するという承認または何らかの形式の示唆として解釈されず、かつそのように解釈されるべきでない。

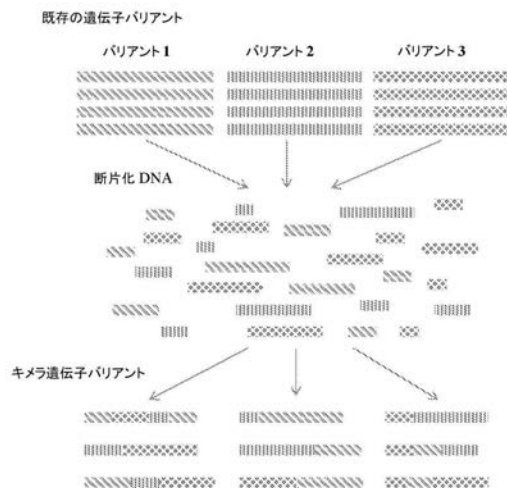
30

さらに、以下の特定の出願が参照によって本明細書に組み込まれる：20016年12月30日に米国出願第15/396,230号（米国公開第US2017/0159045号A1）；2016年12月07日に米国出願されたPCT/US2016/065465（WO2017/100377A1）；2016年4月27日に米国出願された米国出願第15/140,296号（US2017/0316353A1）；2017年4月26日に米国出願されたPCT/US2017/029725（WO2017/189784A1）；2016年12月07日に米国出願されたPCT/US2016/065464（WO2017/100376A2）；2016年12月07日に米国出願された米国仮出願第62/431,409号；2015年12月07日に米国出願された米国仮出願第62/264,232号；および2016年7月29日に米国出願された米国仮出願第62/368,786号。

40

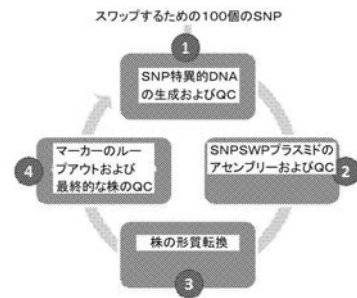
【図 1】

FIGURE 1



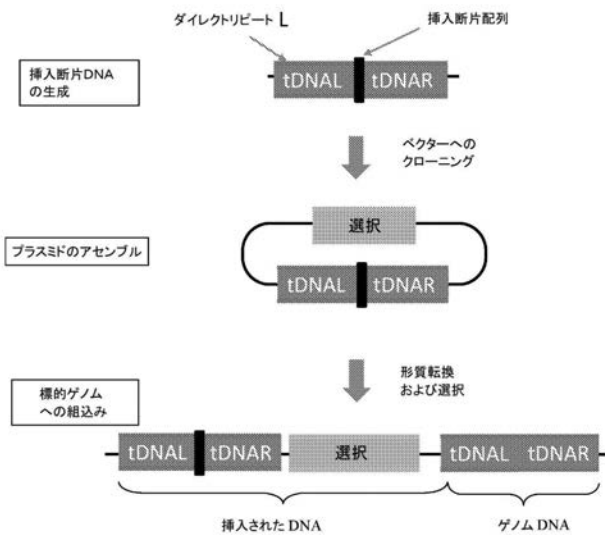
【図 2】

FIGURE 2



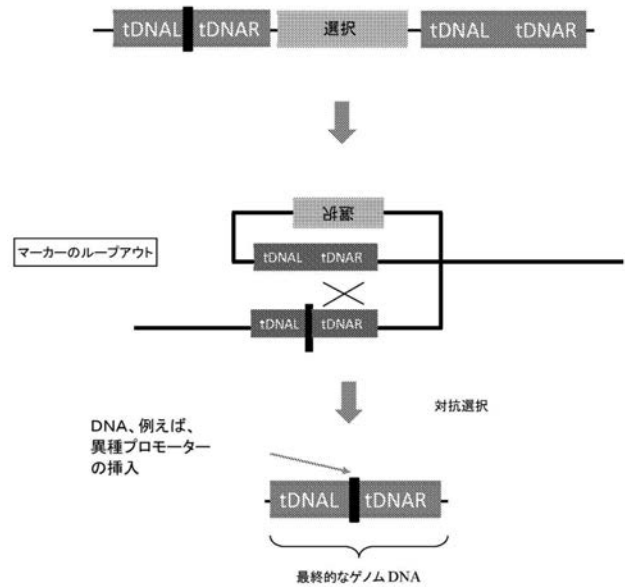
【図 3】

FIGURE 3



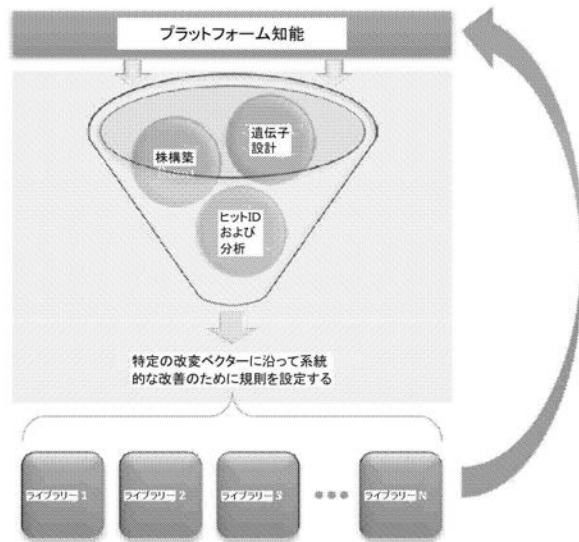
【図 4】

FIGURE 4



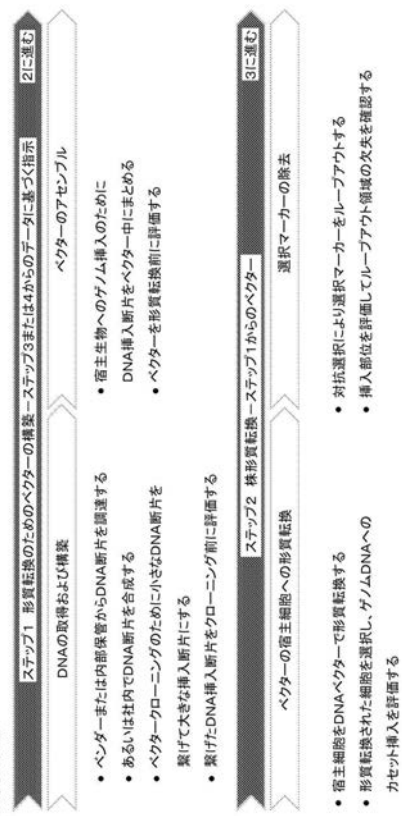
【図 5】

FIGURE 5



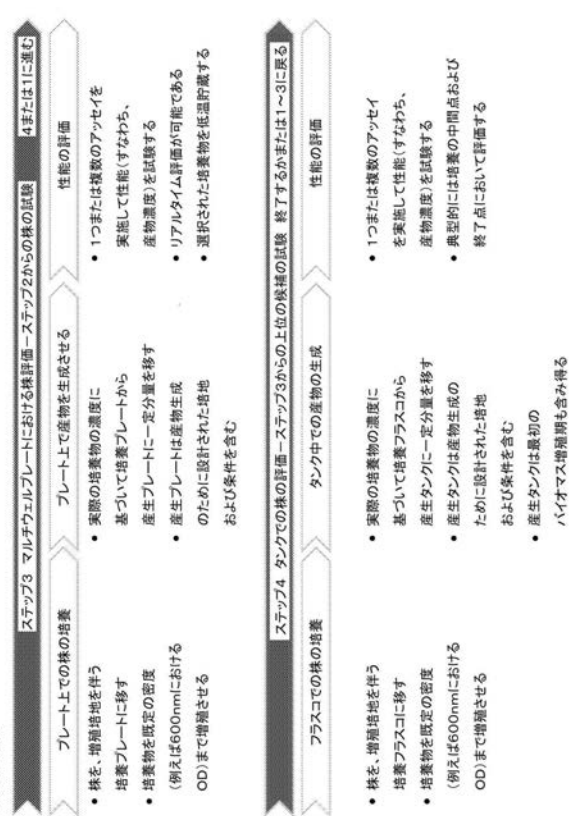
【図 6 A】

FIGURE 6A



【図 6 B】

FIGURE 6B



【図 7】

FIGURE 7

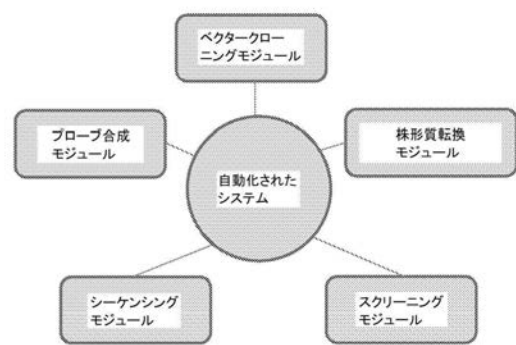
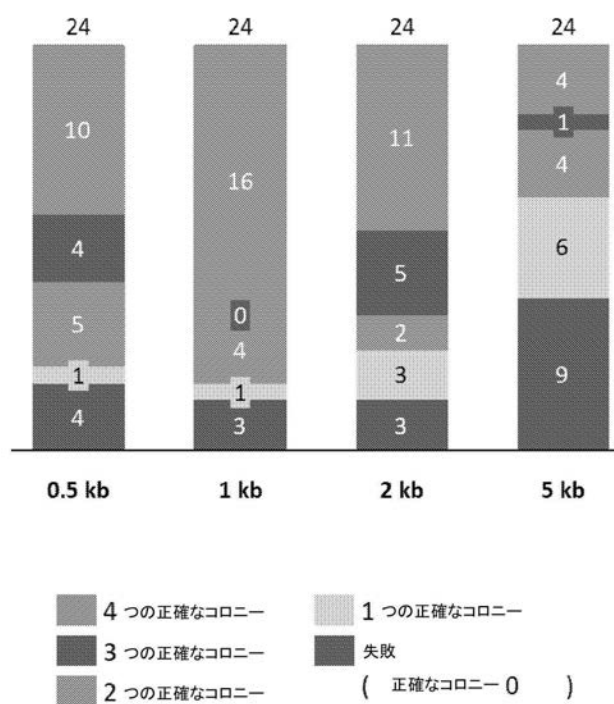
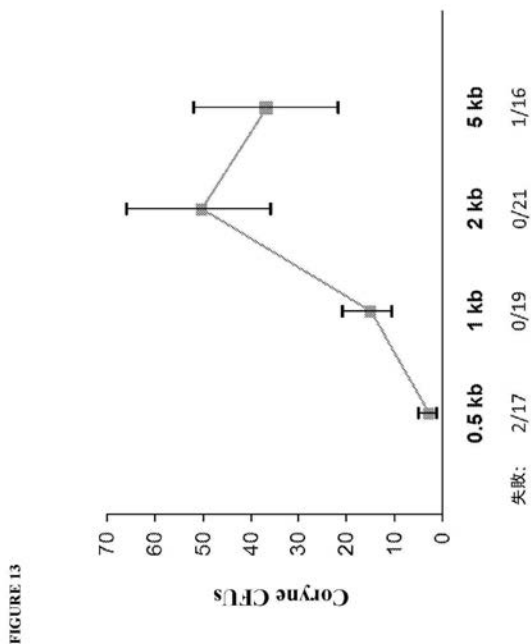


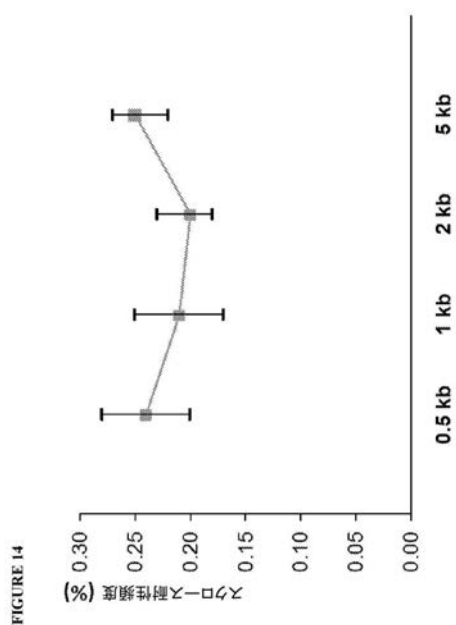
FIGURE 12



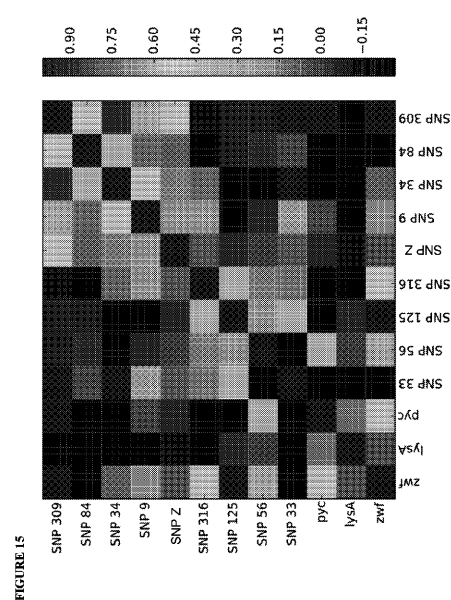
【 図 1 3 】



【 図 1 4 】

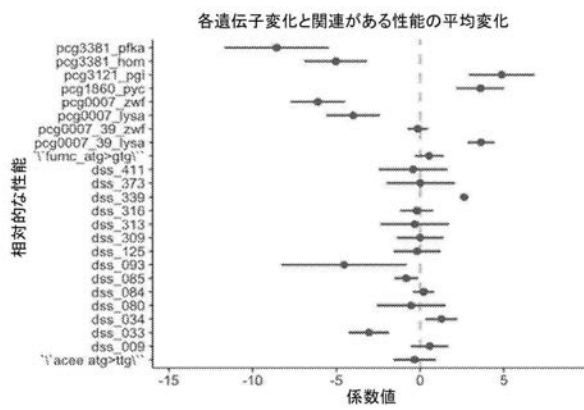


【 図 1 5 】



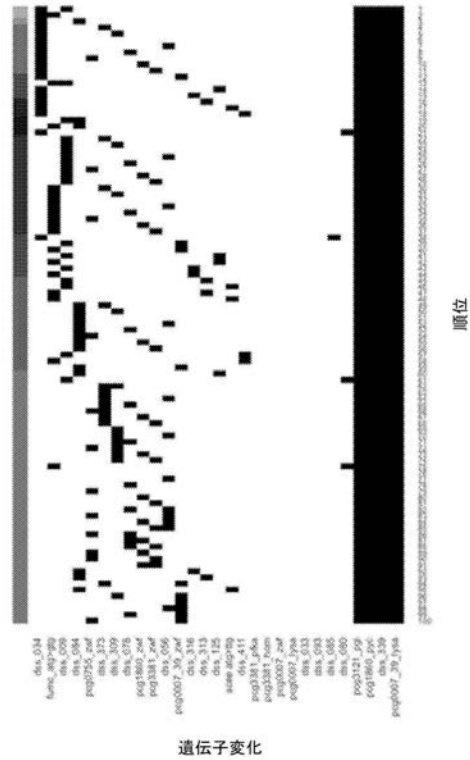
【図 2 4】

FIGURE 24



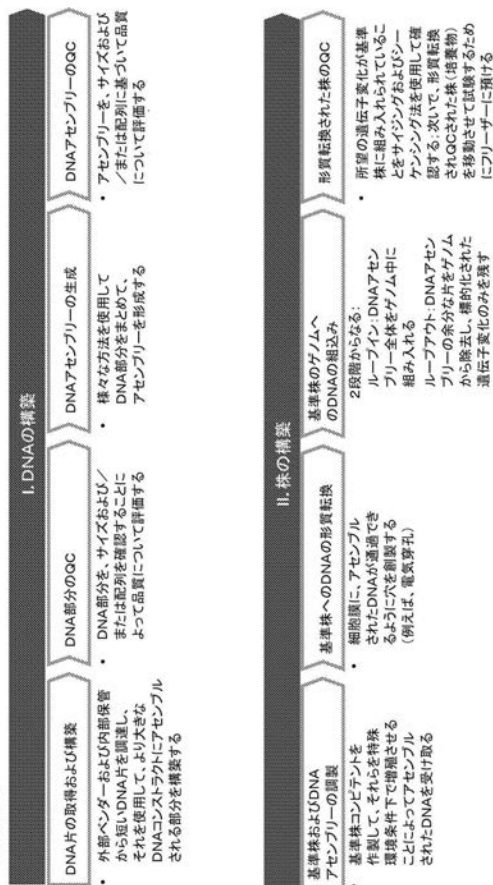
【図 2 5】

FIGURE 25



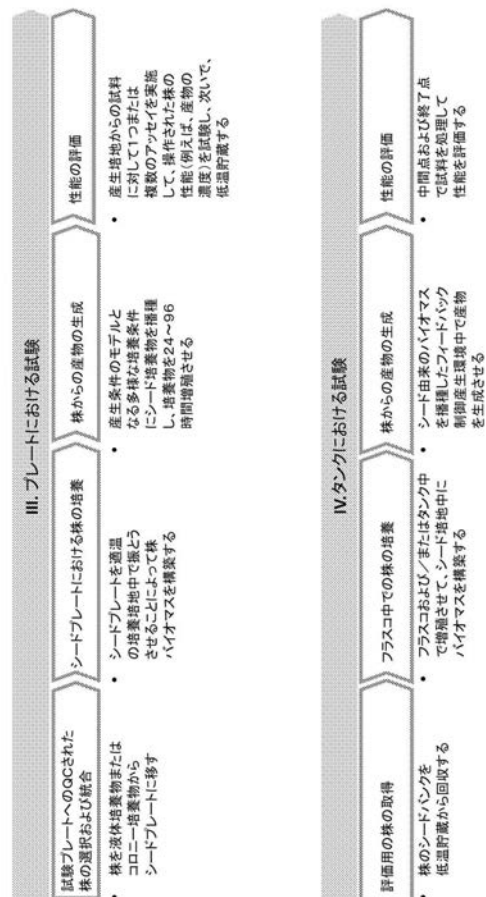
【図 2 6】

FIGURE 26

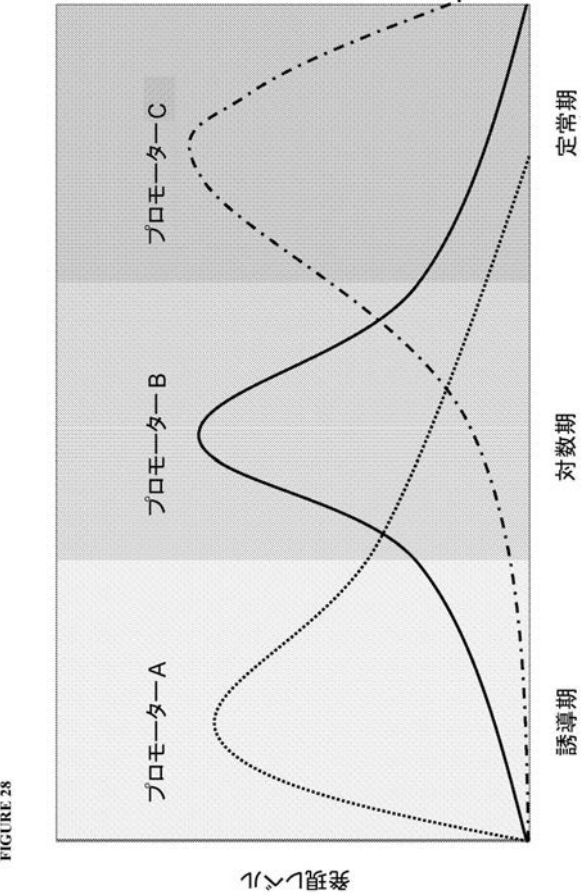


【図 2 7】

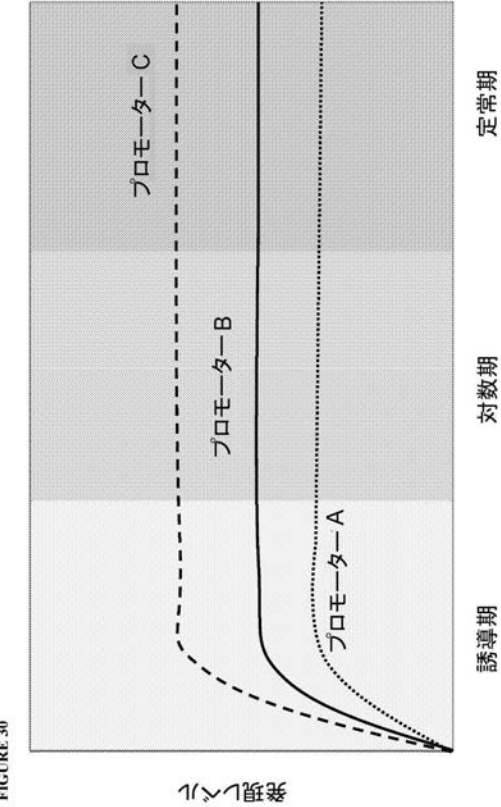
FIGURE 27



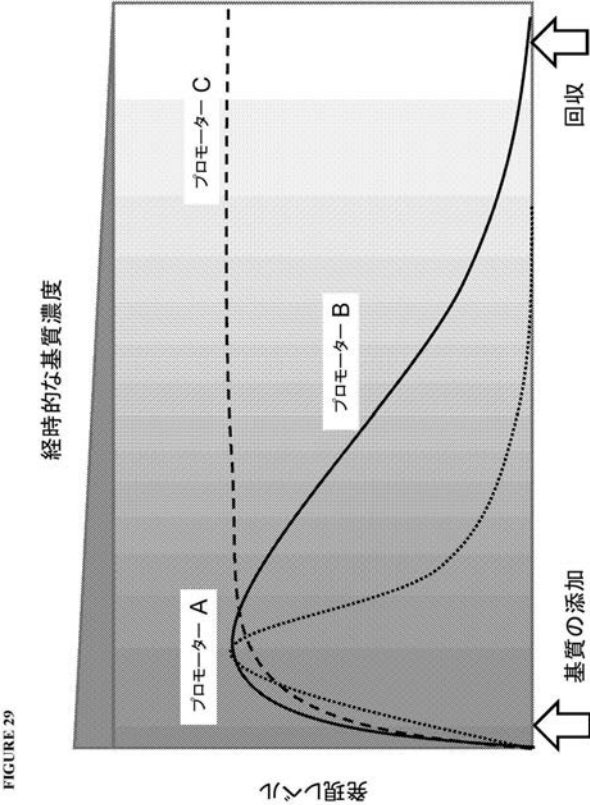
【図 28】



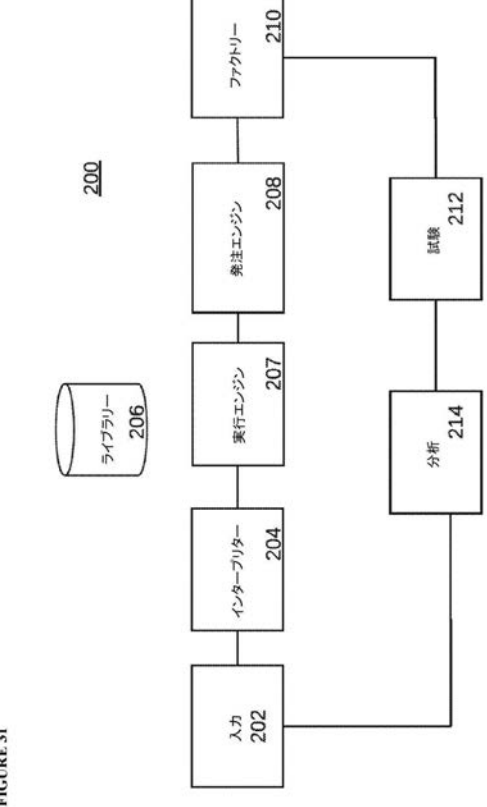
【図 30】



【図 29】

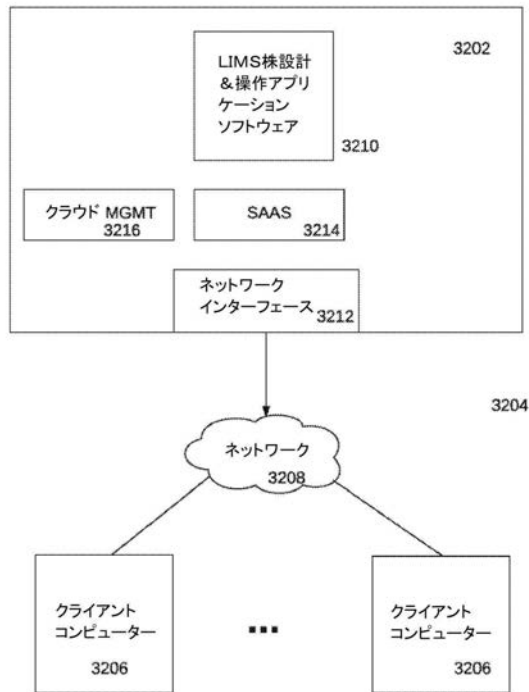


【図 31】



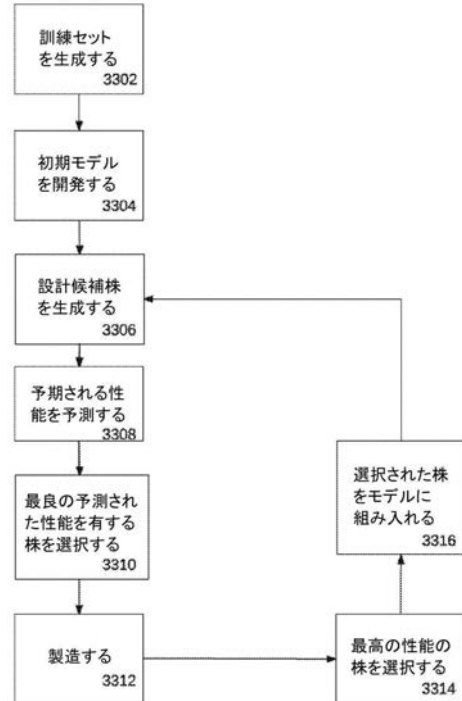
【図 3 2】

FIGURE 32



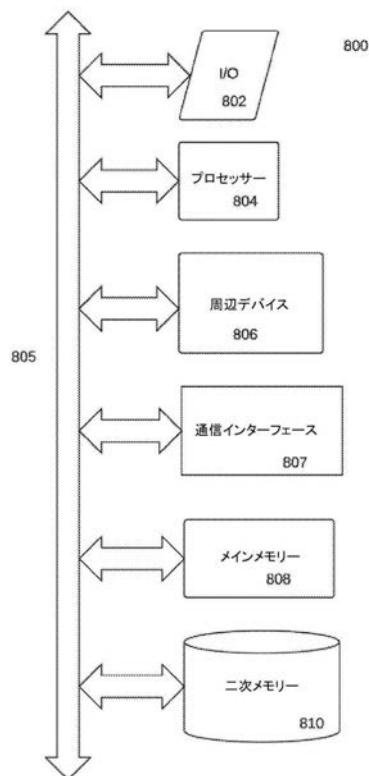
【図 3 3】

FIGURE 33



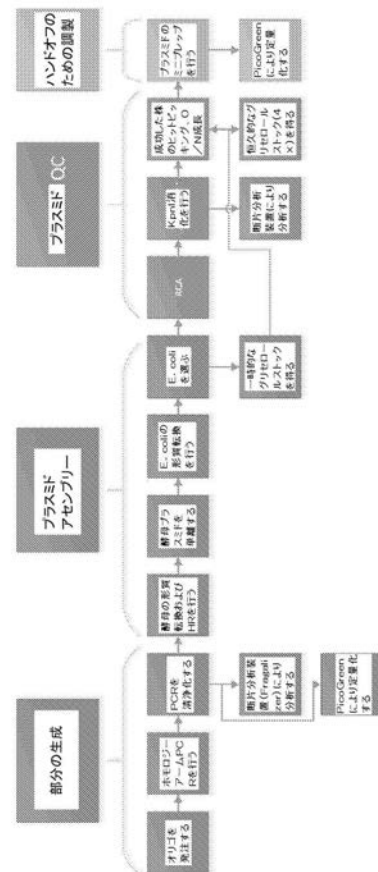
【図 3 4】

FIGURE 34



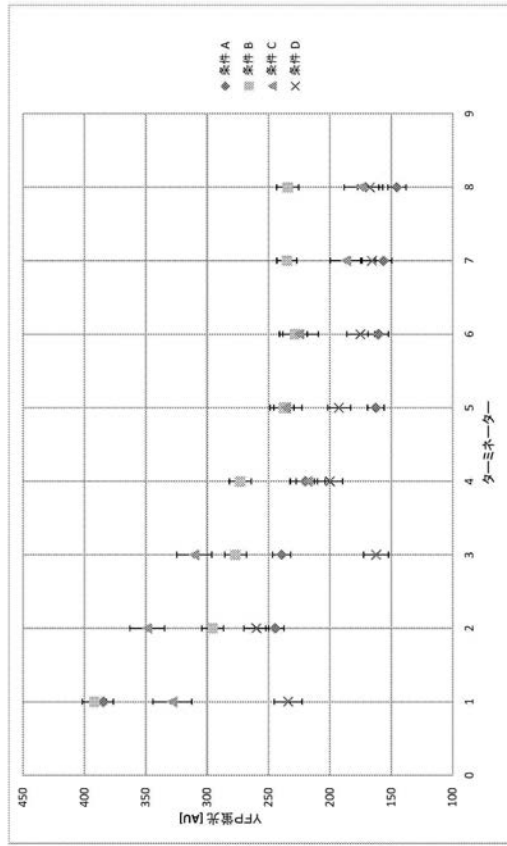
【図 3 5】

FIGURE 35



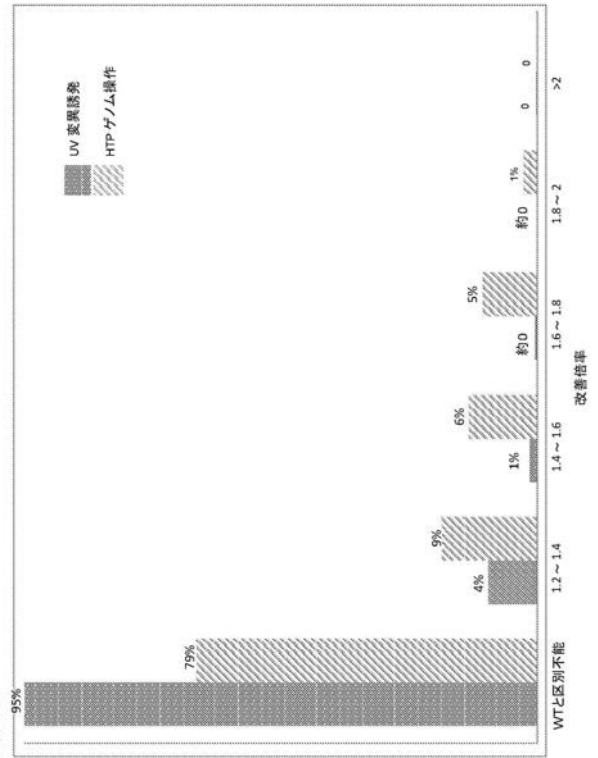
【図 36】

FIGURE 36



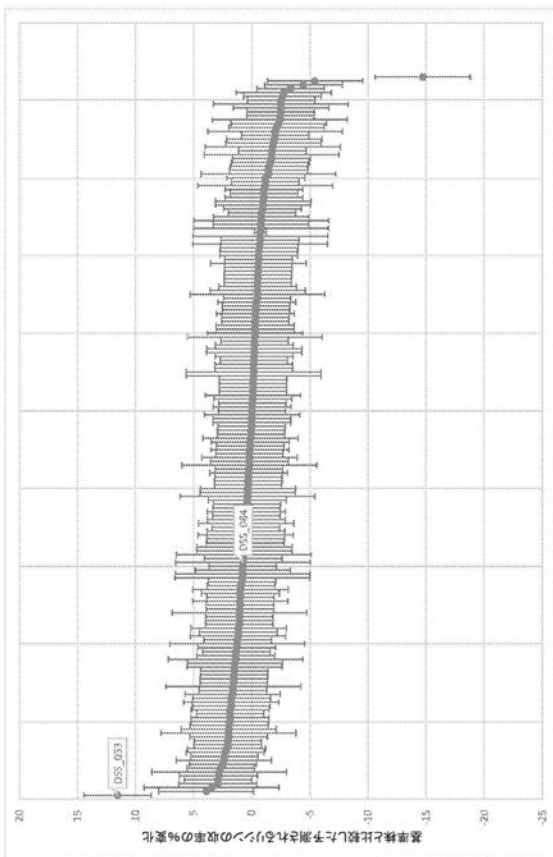
【図 37】

FIGURE 37



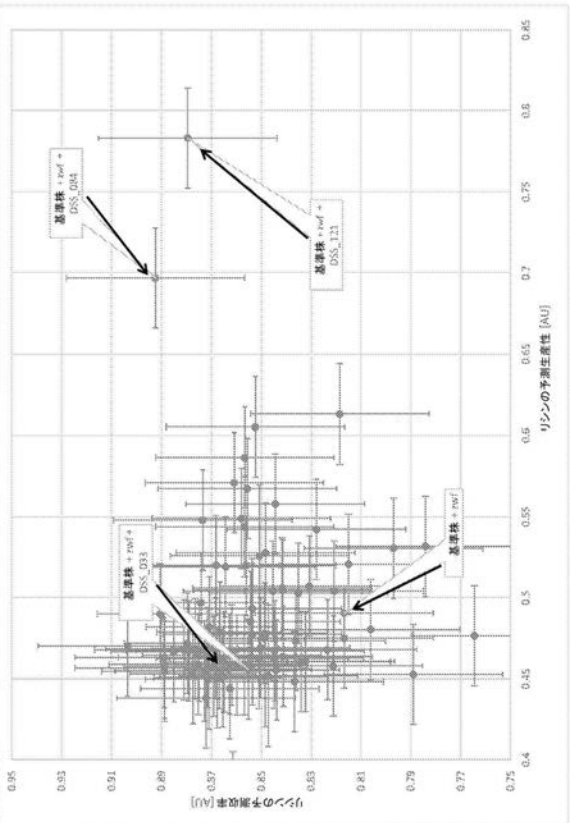
【図 38】

FIGURE 38

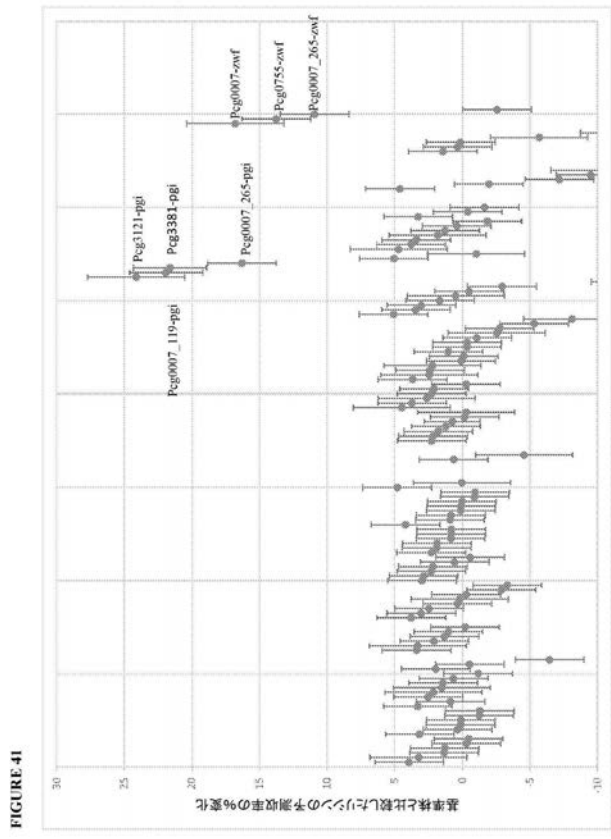


【図 39】

FIGURE 39



【図 4 1】



【图 4 3】

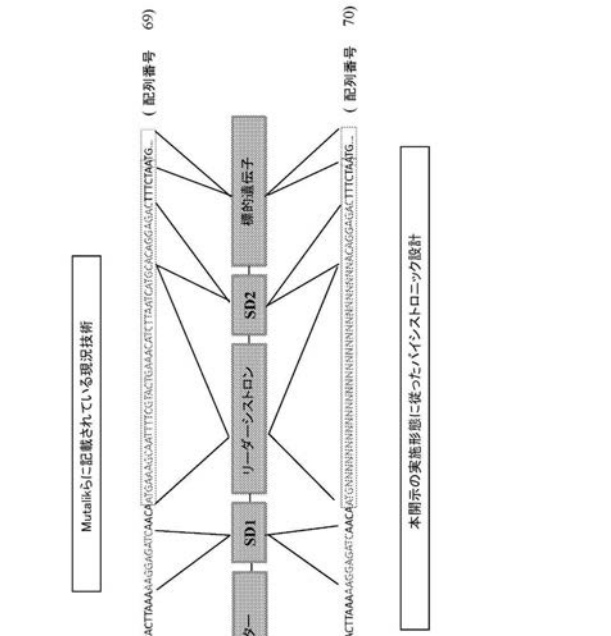
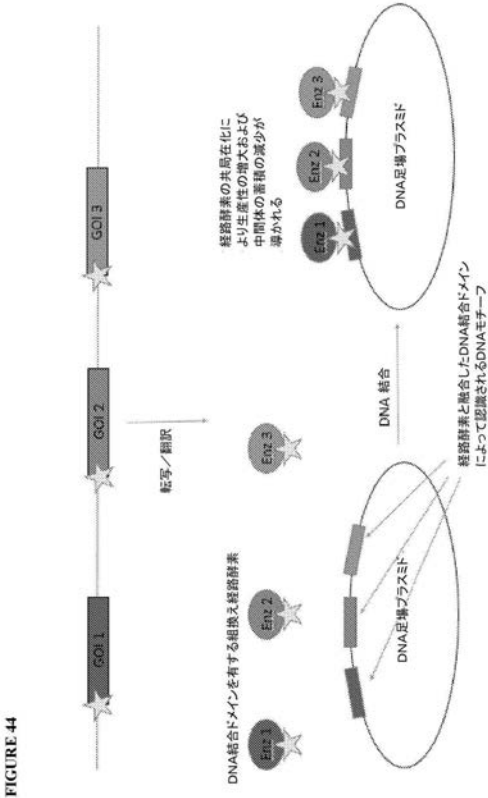
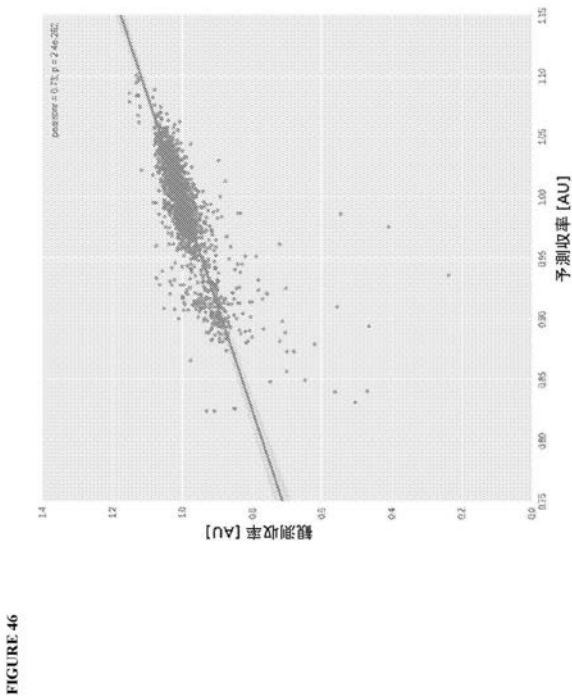


FIGURE 43

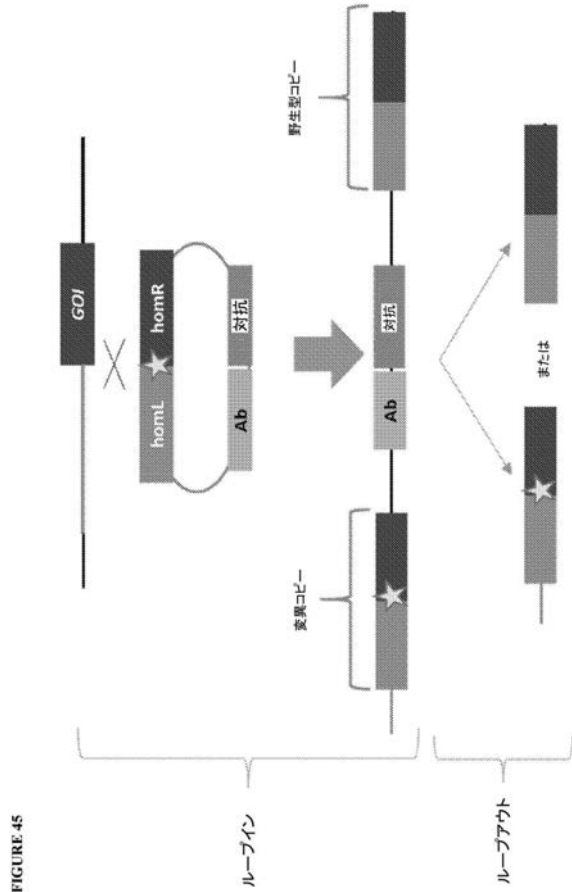
【 図 4 4 】



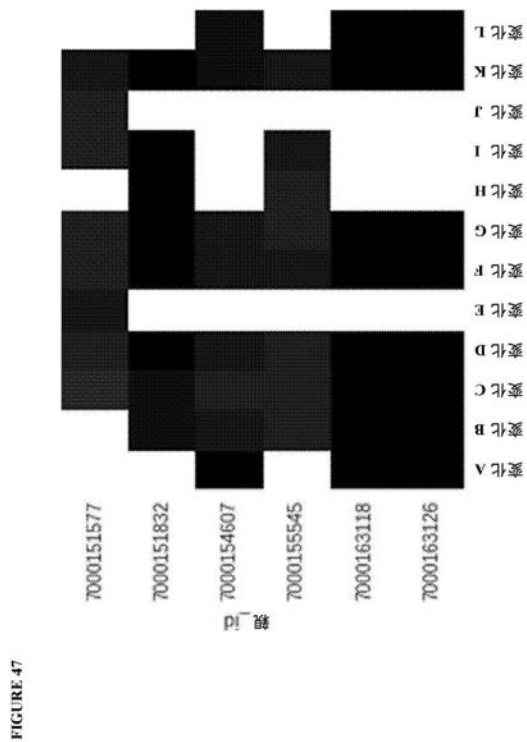
【 図 4 6 】



【 図 4 5 】



【 図 4 7 】



【図 4 8】

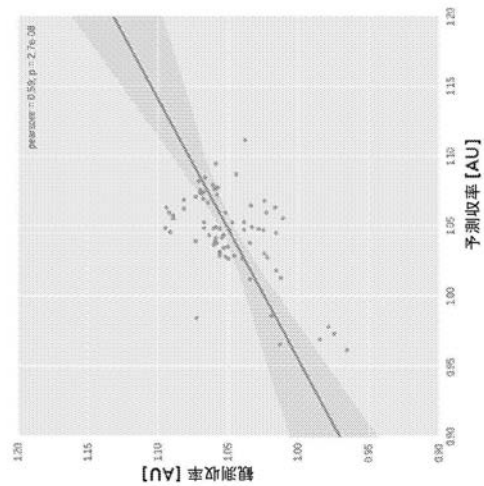


FIGURE 48

【図 4 9】

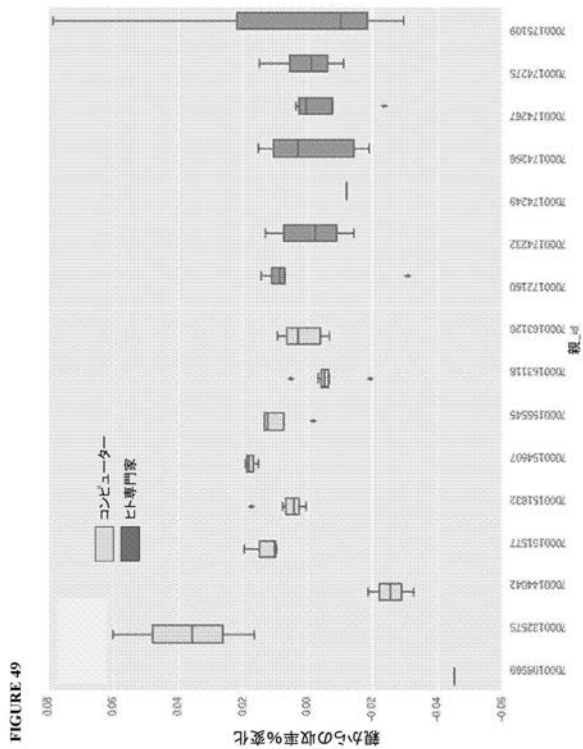


FIGURE 49

【図 5 0】

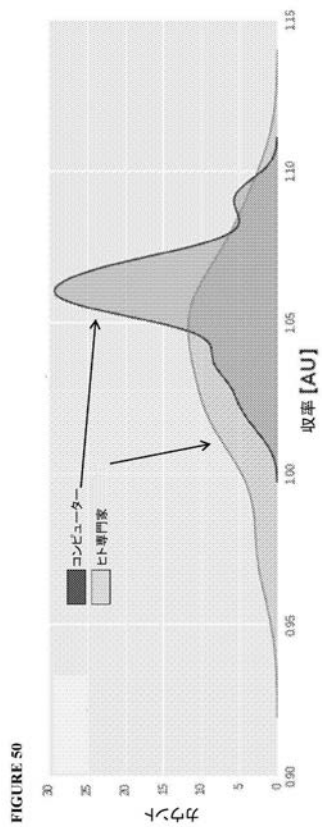


FIGURE 50

【図 5 1】

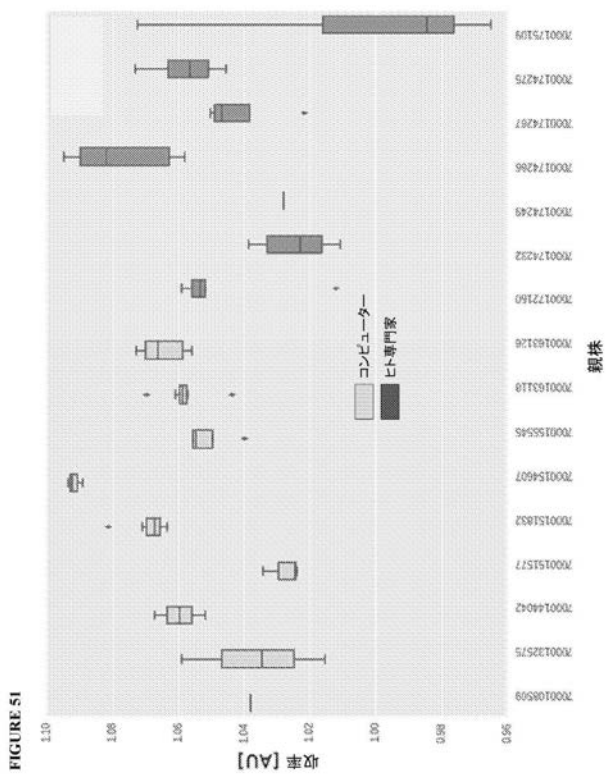
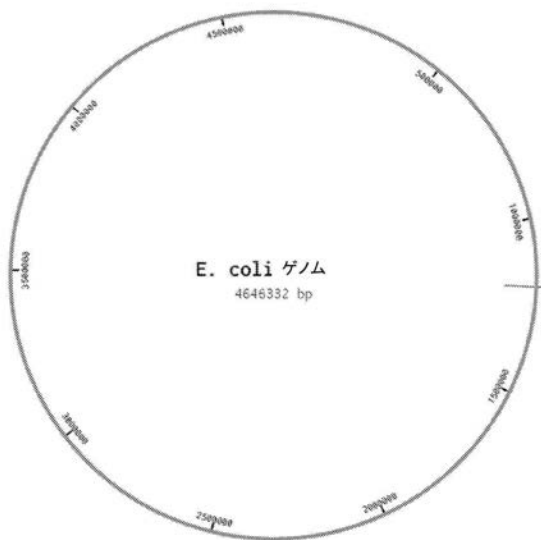


FIGURE 51

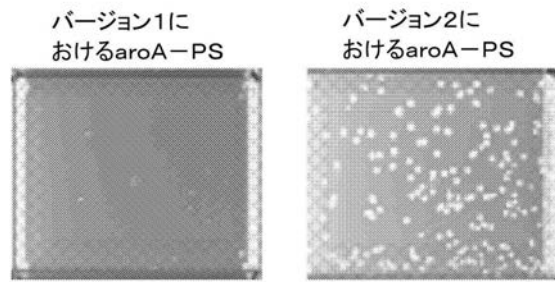
【図 5 2】

FIGURE 52



【図 5 3】

FIGURE 53



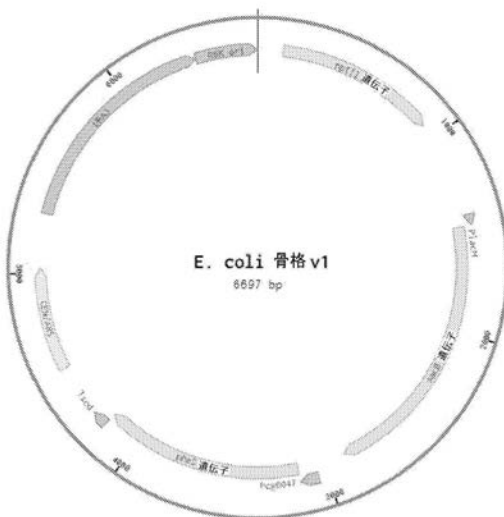
【図 5 4】

FIGURE 54



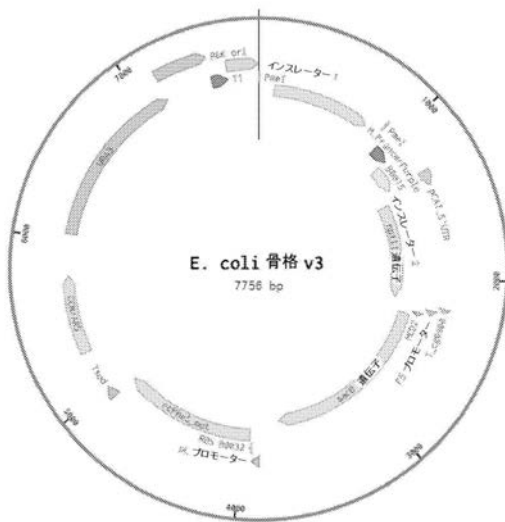
【図 5 5】

FIGURE 55



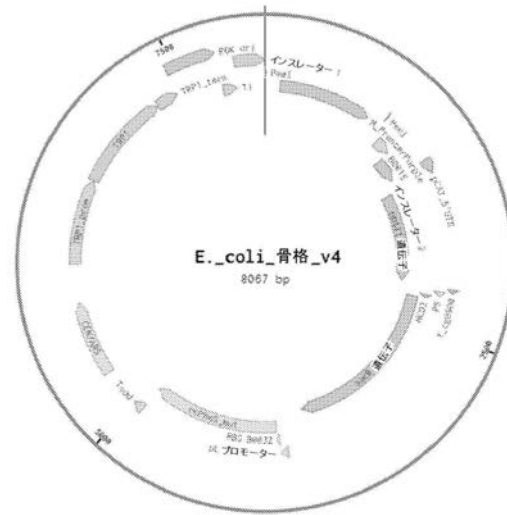
【図 57】

FIGURE 57



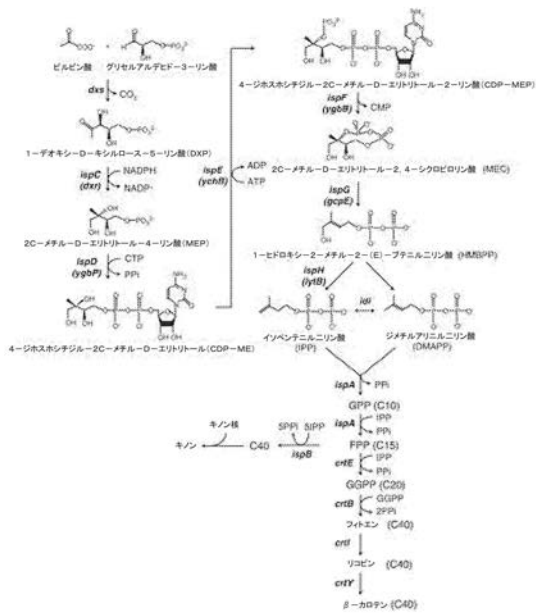
【図 58】

FIGURE 58



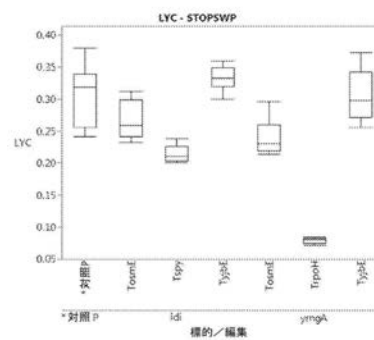
【図 59】

FIGURE 59



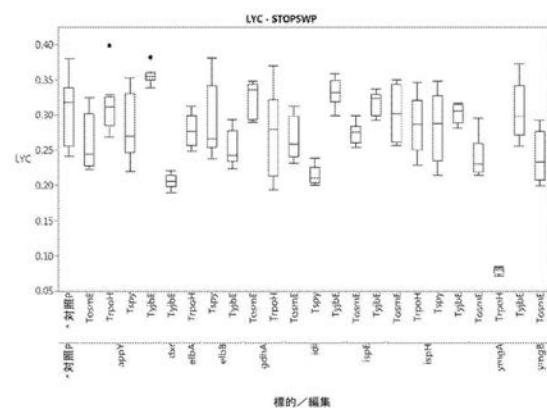
【図 60】

FIGURE 60



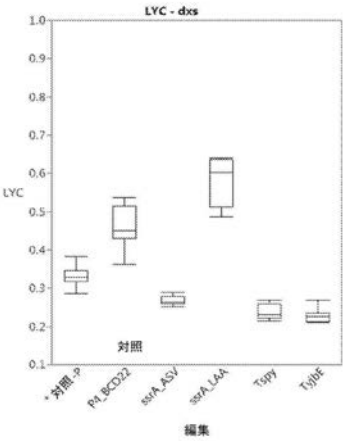
【図 61】

FIGURE 61



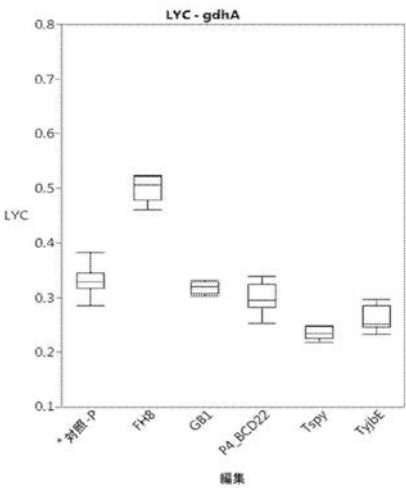
【図 6 2】

FIGURE 62



【図 6 3】

FIGURE 63



【配列表】

20205244900000001.app

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2018/036333

A. CLASSIFICATION OF SUBJECT MATTER INV. C12N15/10 C12N15/73 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) C12N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EP0-Internal, BIOSIS, CAB Data, Sequence Search, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 02/29032 A2 (DIVERSA CORP [US]; SHORT JAY M [US]; FU PENGCHENG [US]; LATTERICH MART) 11 April 2002 (2002-04-11) the whole document	1-13
X	ZHANG Y-X ET AL: "Genome shuffling leads to rapid phenotypic improvement in bacteria", NATURE, MACMILLAN JOURNALS LTD, LONDON, GB, vol. 415, no. 6872, 1 January 2002 (2002-01-01), pages 644-646, XP002228258, ISSN: 0028-0836, DOI: 10.1038/415644A the whole document ----- -/--	1-13
☒ Further documents are listed in the continuation of Box C. ☒ 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		Date of mailing of the international search report
31 October 2018		16/11/2018
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer
		Hornig, Horst

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

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2018/036333

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 00/04190 A1 (MAXYGEN INC [US]; DEL CARDAYRE STEPHEN [US]; TOBIN MATTHEW [US]; STEMM) 27 January 2000 (2000-01-27) the whole document	1-13
X	----- WO 98/31837 A1 (MAXYGEN INC [US]; DELCARDAYRE STEPHEN B [US]; TOBIN MATHEW B [US]; STE) 23 July 1998 (1998-07-23) the whole document	1-13
X	----- WO 97/20078 A1 (AFFYMAX TECH NV [NL]; STEMMER WILLEM P C [US]; CRAMERI ANDREAS [US]) 5 June 1997 (1997-06-05) the whole document	1-13
X	----- WO 97/35966 A1 (MAXYGEN INC [US]; MINSHULL JEREMY [US]; STEMMER WILLEM P C [US]) 2 October 1997 (1997-10-02) the whole document	1-13
X	----- WO 2010/002966 A2 (DOW GLOBAL TECHNOLOGIES INC [US]; RETALLACK DIANE [US]) 7 January 2010 (2010-01-07) the whole document	1-13
X	----- WANG HARRIS H ET AL: "Programming cells by multiplex genome engineering and accelerated evolution.", NATURE 13 AUG 2009, vol. 460, no. 7257, 13 August 2009 (2009-08-13), pages 894-898, XP55111854, ISSN: 1476-4687	1-13
A	the whole document	14-45
A	----- LEE S Y ET AL: "Systems biotechnology for strain improvement", TRENDS IN BIOTECHNOLOGY, ELSEVIER PUBLICATIONS, CAMBRIDGE, GB, vol. 23, no. 7, 1 July 2005 (2005-07-01), pages 349-358, XP027778087, ISSN: 0167-7799 [retrieved on 2005-07-01] the whole document	1-45
A	----- ALPER HAL ET AL: "Characterization of lycopene-overproducing E. coli strains by combining systematic and combinatorial gene knockout targets", NATURE BIOTECHNOLOGY MAY 2006, vol. 24, no. 5, October 2006 (2006-10), pages 612-616, XP2783819, ISSN: 0175-7598 the whole document	1-45
	----- -/--	

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2018/036333

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MAMIE Z LI ET AL: "MAGIC, an in vivo genetic method for the rapid construction of recombinant DNA molecules", NATURE GENETICS., vol. 37, no. 3, 30 January 2005 (2005-01-30), pages 311-319, XP55410043, NEW YORK, US ISSN: 1061-4036, DOI: 10.1038/ng1505 the whole document -----	1-45
A	ALPER HAL ET AL: "Tuning genetic control through promoter engineering", PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, NATIONAL ACADEMY OF SCIENCES, US, vol. 102, no. 36, 6 September 2005 (2005-09-06), pages 12678-12683, XP002428722, ISSN: 0027-8424, DOI: 10.1073/PNAS.0504604102 the whole document -----	1-45
A	MASATO IKEDA ET AL: "A genome-based approach to create a minimally mutated Corynebacterium glutamicum strain for efficient l-lysine production", JOURNAL OF INDUSTRIAL MICROBIOLOGY & BIOTECHNOLOGY ; OFFICIAL JOURNAL OF THE SOCIETY FOR INDUSTRIAL MICROBIOLOGY, SPRINGER, BERLIN, DE, vol. 33, no. 7, 28 February 2006 (2006-02-28), pages 610-615, XP019388133, ISSN: 1476-5535, DOI: 10.1007/S10295-006-0104-5 the whole document -----	1-45
X	WO 2012/142591 A2 (UNIV COLORADO [US]; GILL RYAN T [US]; LYNCH SEAN [US]; WARNER JOSEPH R) 18 October 2012 (2012-10-18) the whole document -----	1-13
X,P	US 2017/159045 A1 (SERBER ZACH [US] ET AL) 8 June 2017 (2017-06-08) the whole document -----	1-45
X,P	WO 2017/100377 A1 (ZYMERGEN INC [US]) 15 June 2017 (2017-06-15) the whole document -----	1-45
A,P	WO 2017/100376 A2 (ZYMERGEN INC [US]) 15 June 2017 (2017-06-15) the whole document -----	1-45
	----- -/--	

International application No
PCT/US2018/036333

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A, P	WO 2017/189784 A1 (ZYMERGEN INC [US]) 2 November 2017 (2017-11-02) the whole document	1-45
A	----- FOTINI A TRIKKA ET AL: "Iterative carotenogenic screens identify combinations of yeast gene deletions that enhance sclareol production", MICROBIAL CELL FACTORIES,, vol. 14, no. 1, 24 April 2015 (2015-04-24) , page 60, XP021214826, ISSN: 1475-2859, DOI: 10.1186/S12934-015-0246-0 the whole document	1-45
A	----- OHNISHI J ET AL: "A novel methodology employing Corynebacterium glutamicum genome information to generate a new L-lysine-producing mutant", APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, SPRINGER, DE, vol. 58, no. 2, 1 February 2002 (2002-02-01), pages 217-223, XP002289984, ISSN: 0175-7598, DOI: 10.1007/S00253-001-0883-6 the whole document	1-45
A	----- LEE MICHAEL E ET AL: "A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly.", ACS SYNTHETIC BIOLOGY 18 SEP 2015, vol. 4, no. 9, 18 September 2015 (2015-09-18), pages 975-986, XP002786200, ISSN: 2161-5063 the whole document	1-45
A	----- SNITKIN EVAN S ET AL: "Epistatic interaction maps relative to multiple metabolic phenotypes.", PLOS GENETICS 10 FEB 2011, vol. 7, no. 2, E1001294, 10 February 2011 (2011-02-10), pages 1-15, XP002786201, ISSN: 1553-7404 the whole document	1-45
X	----- RAINER KNAUS ET AL: "PL of coliphage lambda: an alternative solution for an efficient promoter", THE EMBO JOURNAL,, vol. 7, no. 9, 25 August 1988 (1988-08-25) , pages 2919-2923, XP001317078, the whole document	89,90
Y	----- the whole document	93-110

-/--

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2018/036333

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KINCADE J M ET AL: "Bacteriophage lambda promoters pL and pR sequence determinants of in vivo activity and of sensitivity to the DNA gyrase inhibitor, coumermycin", GENE, ELSEVIER, AMSTERDAM, NL, vol. 97, no. 1, 1 January 1991 (1991-01-01), pages 7-12, XP023545045, ISSN: 0378-1119, DOI: 10.1016/0378-1119(91)90003-T [retrieved on 1991-01-01] the whole document	93-110
A	EP 2 657 250 A2 (CJ CHEILJEDANG CORP [KR]) 30 October 2013 (2013-10-30) paragraph [0051] - paragraph [0052]	93-110
A	US 2006/141578 A1 (PAN JAE-GU [KR] ET AL) 29 June 2006 (2006-06-29) the whole document	93-110
A	WILLIAM R MCCLEARY: "Application of promoter swapping techniques to control expression of chromosomal genes", APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, SPRINGER, BERLIN, DE, vol. 84, no. 4, 25 July 2009 (2009-07-25), pages 641-648, XP019737788, ISSN: 1432-0614, DOI: 10.1007/S00253-009-2137-Y the whole document	35-45
A	PENGFEI GU ET AL: "One-step of tryptophan attenuator inactivation and promoter swapping to improve the production of L-tryptophan in Escherichia coli", MICROBIAL CELL FACTORIES,, vol. 11, no. 1, 2 March 2012 (2012-03-02), page 30, XP021118480, ISSN: 1475-2859, DOI: 10.1186/1475-2859-11-30 the whole document	35-45
A	RICE CHRISTOPHER D ET AL: "Employment of a promoter-swapping technique shows that PhoU modulates the activity of the PstSCAB2 ABC transporter in Escherichia coli.", APPLIED AND ENVIRONMENTAL MICROBIOLOGY FEB 2009, vol. 75, no. 3, February 2009 (2009-02), pages 573-582, XP002786202, ISSN: 1098-5336 the whole document	35-45
	----- -/--	

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2018/036333

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	GUEGUEN ERWAN ET AL: "Promoter swapping unveils the role of the <i>Citrobacter rodentium</i> CTS1 type VI secretion system in interbacterial competition.", APPLIED AND ENVIRONMENTAL MICROBIOLOGY JAN 2013, vol. 79, no. 1, January 2013 (2013-01), pages 32-38, XP002786203, ISSN: 1098-5336 the whole document -----	35-45
A	WO 2012/149470 A1 (AMYRIS INC [US]; SERBER ZACH [US]; HORWITZ ANDREW [US]) 1 November 2012 (2012-11-01) the whole document -----	1-45

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2018/036333

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0229032	A2	11-04-2002	NONE
WO 0004190	A1	27-01-2000	AT 317453 T 15-02-2006 AU 771511 B2 25-03-2004 CA 2331263 A1 27-01-2000 CA 2720311 A1 27-01-2000 DE 69929796 T2 12-10-2006 EP 1108058 A1 20-06-2001 EP 1707641 A2 04-10-2006 ES 2257866 T3 01-08-2006 JP 4674329 B2 20-04-2011 JP 2002520071 A 09-07-2002 JP 2010131014 A 17-06-2010 US 6287862 B1 11-09-2001 US 6326204 B1 04-12-2001 US 6335198 B1 01-01-2002 US 6528311 B1 04-03-2003 US 6716631 B1 06-04-2004 US 2012252681 A1 04-10-2012 WO 0004190 A1 27-01-2000
WO 9831837	A1	23-07-1998	AT 334225 T 15-08-2006 AU 743305 B2 24-01-2002 CA 2276064 A1 23-07-1998 DE 69835360 T2 16-08-2007 DK 1717322 T3 22-10-2012 EP 1007732 A1 14-06-2000 EP 1717322 A2 02-11-2006 EP 2261373 A2 15-12-2010 ES 2270505 T3 01-04-2007 JP 4062366 B2 19-03-2008 JP 2001508662 A 03-07-2001 KR 20000070258 A 25-11-2000 KR 20050043983 A 11-05-2005 US 6251674 B1 26-06-2001 US 6352859 B1 05-03-2002 US 6379964 B1 30-04-2002 WO 9831837 A1 23-07-1998
WO 9720078	A1	05-06-1997	AT 205877 T 15-10-2001 AT 205878 T 15-10-2001 AU 713952 B2 16-12-1999 CA 2239099 A1 05-06-1997 DE 876509 T1 05-04-2001 DE 911396 T1 05-04-2001 DE 69615406 D1 25-10-2001 DE 69615406 T2 20-06-2002 DE 69615411 D1 25-10-2001 DE 69615411 T2 20-06-2002 DK 0876509 T3 18-02-2002 DK 0911396 T3 18-02-2002 EP 0876509 A1 11-11-1998 EP 0911396 A2 28-04-1999 EP 1103606 A2 30-05-2001 EP 1138763 A2 04-10-2001 ES 2164929 T3 01-03-2002 ES 2165124 T3 01-03-2002 JP 4015193 B2 28-11-2007

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

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2018/036333

Patent document cited in search report	Publication date	Patent family member(s)	Publication date	
		JP 2000308490 A	07-11-2000	
		JP 2000500981 A	02-02-2000	
		JP 2004187694 A	08-07-2004	
		US 6117679 A	12-09-2000	
		US 6180406 B1	30-01-2001	
		US 6291242 B1	18-09-2001	
		US 6323030 B1	27-11-2001	
		US 6344356 B1	05-02-2002	
		US 6372497 B1	16-04-2002	
		US 6413774 B1	02-07-2002	
		US 6573098 B1	03-06-2003	
		WO 9720078 A1	05-06-1997	

WO 9735966	A1	02-10-1997	AU 2542697 A	17-10-1997
			EP 0906418 A1	07-04-1999
			JP 2000507444 A	20-06-2000
			JP 2001197895 A	24-07-2001
			US 5837458 A	17-11-1998
			US 2002025517 A1	28-02-2002
			WO 9735966 A1	02-10-1997

WO 2010002966	A2	07-01-2010	AU 2009266989 A1	07-01-2010
			CA 2729839 A1	07-01-2010
			EP 2313507 A2	27-04-2011
			NZ 590619 A	31-08-2012
			US 2011111977 A1	12-05-2011
			WO 2010002966 A2	07-01-2010

WO 2012142591	A2	18-10-2012	US 2015368639 A1	24-12-2015
			US 2017067046 A1	09-03-2017
			US 2018258421 A1	13-09-2018
			WO 2012142591 A2	18-10-2012

US 2017159045	A1	08-06-2017	US 2017159045 A1	08-06-2017
			US 2018216099 A1	02-08-2018
			US 2018216100 A1	02-08-2018
			US 2018216101 A1	02-08-2018

WO 2017100377	A1	15-06-2017	CA 3007840 A1	15-06-2017
			CN 108027849 A	11-05-2018
			EP 3387571 A1	17-10-2018
			JP 2018523968 A	30-08-2018
			KR 20180012280 A	05-02-2018
			WO 2017100377 A1	15-06-2017

WO 2017100376	A2	15-06-2017	CA 3007635 A1	15-06-2017
			KR 20180084756 A	25-07-2018
			WO 2017100376 A2	15-06-2017

WO 2017189784	A1	02-11-2017	US 2017316353 A1	02-11-2017
			WO 2017189784 A1	02-11-2017

EP 2657250	A2	30-10-2013	BR 112013015991 A2	10-07-2018
			CN 103797027 A	14-05-2014
			DK 2657250 T3	11-12-2017
			EP 2657250 A2	30-10-2013
			ES 2650451 T3	18-01-2018
			JP 5758499 B2	05-08-2015

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2018/036333

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		JP 2014502843 A	06-02-2014
		KR 20120070531 A	29-06-2012
		PL 2657250 T3	28-02-2018
		RU 2013133626 A	27-01-2015
		US 2013273615 A1	17-10-2013
		WO 2012087039 A2	28-06-2012

US 2006141578	A1	29-06-2006	NONE

WO 2012149470	A1	01-11-2012	AU 2012249390 A1
			12-12-2013
			AU 2017204456 A1
			20-07-2017
			CA 2834375 A1
			01-11-2012
			CN 103492578 A
			01-01-2014
			EP 2702160 A1
			05-03-2014
			JP 6158170 B2
			12-07-2017
			JP 2014513535 A
			05-06-2014
			JP 2017042173 A
			02-03-2017
			KR 20140027354 A
			06-03-2014
			MX 351043 B
			29-09-2017
			SG 194089 A1
			29-11-2013
			US 2012277120 A1
			01-11-2012
			US 2014186942 A1
			03-07-2014
			US 2017240923 A1
			24-08-2017
			WO 2012149470 A1
			01-11-2012
			ZA 201307225 B
			23-12-2014

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2018/036333

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. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

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:

see additional sheet

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.:

1-45, 89-110

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/ US2018/ 036333

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-13

A high-throughput (HTP) method of genomic engineering to evolve an E. coli microbe to acquire a desired phenotype, comprising: a. perturbing the genomes of an initial plurality of E. coli microbes having the same genomic strain background, to thereby create an initial HTP genetic design E. coli strain library comprising individual E. coli strains with unique genetic variations; b. screening and selecting individual strains of the initial HTP genetic design E. coli strain library for the desired phenotype; c. providing a subsequent plurality of E. coli microbes that each comprise a unique combination of genetic variation, said genetic variation selected from the genetic variation present in at least two individual E. coli strains screened in the preceding step, to thereby create a subsequent HTP genetic design E. coli strain library; d. screening and selecting individual E. coli strains of the subsequent HTP genetic design E. coli strain library for the desired phenotype; and e. repeating steps c)-d) one or more times, in a linear or non-linear fashion, until an E. coli microbe has acquired the desired phenotype, wherein each subsequent iteration creates a new HTP genetic design E. coli strain library comprising individual E. coli strains harboring unique genetic variations that are a combination of genetic variation selected from amongst at least two individual E. coli strains of a preceding HTP genetic design E. coli strain library.

2. claims: 14-18

A method for generating a SNP swap E. coli strain library, comprising the steps of: a. providing a reference E. coli strain and a second E. coli strain, wherein the second E. coli strain comprises a plurality of identified genetic variations selected from single nucleotide polymorphisms, DNA insertions, and DNA deletions, which are not present in the reference E. coli strain; and b. perturbing the genome of either the reference E. coli strain, or the second E. coli strain, to thereby create an initial SNP swap E. coli strain library comprising a plurality of individual E. coli strains with unique genetic variations found within each strain of said plurality of individual strains, wherein each of said unique genetic variations corresponds to a single genetic variation selected from the plurality of identified genetic variations between the reference E. coli strain and the second E. coli strain.

3. claims: 19-34

International Application No. PCT/ US2018/ 036333

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

A method for rehabilitating and improving the phenotypic performance of a production E. coli strain, comprising the steps of: a. providing a parental lineage E. coli strain and a production E. coli strain derived therefrom, wherein the production E. coli strain comprises a plurality of identified genetic variations selected from single nucleotide polymorphisms, DNA insertions, and DNA deletions, not present in the parental lineage strain; b. perturbing the genome of either the parental lineage E. coli strain, or the production E. coli strain, to create an initial library of E. coli strains, wherein each strain in the initial library comprises a unique genetic variation from the plurality of identified genetic variations between the parental lineage E. coli strain and the production E. coli strain; c. screening and selecting individual strains of the initial library for phenotypic performance improvements over a reference E. coli strain, thereby identifying unique genetic variations that confer phenotypic performance improvements; d. providing a subsequent plurality of E. coli microbes that each comprise a combination of unique genetic variations from the genetic variations present in at least two individual E. coli strains screened in the preceding step, to thereby create a subsequent library of E. coli strains; e. screening and selecting individual strains of the subsequent library for phenotypic performance improvements over the reference E. coli strain, thereby identifying unique combinations of genetic variation that confer additional phenotypic performance improvements; and f. repeating steps d)-e) one or more times, in a linear or non-linear fashion, until an E. coli strain exhibits a desired level of improved phenotypic performance compared to the phenotypic performance of the production E. coli strain, wherein each subsequent iteration creates a new library of microbial strains, where each strain in the new library comprises genetic variations that are a combination of genetic variations selected from amongst at least two individual E. coli strains of a preceding library.

4. claims: 35-45

A method for generating a promoter swap E. coli strain library; A promoter swap method for improving the phenotypic performance of a production E. coli strain.

5. claims: 46-55

A method for generating a terminator swap E. coli strain library; A terminator swap method for improving the phenotypic performance of a production E. coli strain.

6. claims: 56-60

International Application No. PCT/ US2018/ 036333

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

A system for colocalizing biosynthetic enzymes from a biosynthetic pathway in an E. coli host cell, said system comprising: a. two or more chimeric enzyme proteins involved in an enzymatic reaction, each chimeric enzyme protein comprising an enzyme portion coupled to a DNA binding domain portion; and b. a DNA scaffold comprising i. one or more subunits, each subunit comprising two or more different DNA binding sites separated by at least one nucleic acid spacer; wherein the chimeric enzyme proteins are recruited to the DNA scaffold by their coupled DNA binding domain portions, each of which bind at least one DNA binding site in the DNA scaffold.

7. claims: 61-68

A bicistronic design regulatory (BCD) sequence, said BCD sequence comprising in order: a. a promoter operably linked to; b. a first ribosomal binding site (SD1); c. a first cistronic sequence (Cis1); d. a second ribosome binding site (SD2); wherein said BCD sequence is operably linked to a target gene sequence (Cis2); A method for expressing two target gene proteins in a host organism, said method comprising the steps of:

a. introducing into the host organism a first polynucleotide encoding for a first target gene protein, wherein said first polynucleotide is operably linked to a first of said bicistronic design regulatory (BCD) sequence; and b. introducing into the host organism a second polynucleotide encoding for a second target gene protein, wherein said second polynucleotide is operably linked to a second of said BCD; wherein the first and second BCDs are identical except for their respective Cis 1 sequences, and wherein the target gene proteins are expressed in the host organism at a first and second expression level, respectively.

8. claims: 69-78

A method for generating a protein solubility tag swap E. coli strain library; A protein solubility tag swap method for improving the phenotypic performance of a production E. coli strain.

9. claims: 79-88

A method for generating a protein degradation tag swap E. coli strain library; A protein degradation tag swap method for improving the phenotypic performance of a production E. coli strain.

10. claims: 89, 90(completely); 93-110(partially)

International Application No. PCT/ US2018/ 036333

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

A chimeric synthetic promoter operably linked to a heterologous gene for expression in a microbial host cell, wherein the chimeric synthetic promoter is 60-90 nucleotides in length and consists of a distal portion of lambda phage pR promoter, variable -35 and -10 regions of lambda phage PL and pR promoters that are each six nucleotides in length, core portions of lambda phage PL and pR promoters and a 5' UTR/Ribosomal Binding Site (RBS) portion of lambda phage pR promoter.

11. claims: 91, 92(completely); 93-110(partially)

A chimeric synthetic promoter operably linked to a heterologous gene for expression in a microbial host cell, wherein the chimeric synthetic promoter is 60-90 nucleotides in length and consists of a distal portion of lambda phage pR promoter, variable -35 and -10 regions of lambda phage PL and pR promoters that are each six nucleotides in length, core portions of lambda phage PL and pR promoters and a 5' UTR/Ribosomal Binding Site (RBS) portion of the promoter of the E. coli acs gene.

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 15/52 (2006.01)	C 1 2 N 15/52	Z
C 1 2 N 15/62 (2006.01)	C 1 2 N 15/62	Z
C 1 2 N 15/113 (2010.01)	C 1 2 N 15/113	1 3 0 Z
C 1 2 N 15/70 (2006.01)	C 1 2 N 15/113	Z
	C 1 2 N 15/70	Z

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

- (72)発明者 デイビス, マシュー
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- (72)発明者 ウィスニュースキー, クリスティー
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- (72)発明者 ウェストフォール, パトリック
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- (72)発明者 サーバー, ザック
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- (72)発明者 ディーン, エリック ジェデディア
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- (72)発明者 マンチェスター, ショーン
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- (72)発明者 グラ, キャサリン
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- (72)発明者 シェルマン, エリン
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- (72)発明者 キンボール, アーロン
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- (72)発明者 シジカ, ショーン
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- (72)発明者 フレウエン, バーバラ
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付

- (72)発明者 トレイナー, トーマス
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- (72)発明者 フラッシュマン, マイケル
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- (72)発明者 ハウシャルター, ロバート
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- (72)発明者 モーガン, ステイシーアン
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- (72)発明者 ブレイズ, マイケル
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- (72)発明者 ラマクリシュナン, プラバ
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- (72)発明者 ロスチャイルドーマンチネッリ, カイル
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- (72)発明者 キム, ヨンニョン
アメリカ合衆国 カリフォルニア 94608, エメリービル, ホートン ストリート 59
80, スイート 105, ザイマージェン インコーポレイテッド 気付
- F ターム(参考) 4B063 QA20 QQ06 QQ42 QR32 QR75 QS02 QS38 QX01
4B065 AA26X AA26Y AB01 AC14 AC20 BA01 CA27 CA46