

(19)日本国特許庁(JP)

## (12)特許公報(B2)

(11)特許番号

特許第7093310号

(P7093310)

(45)発行日 令和4年6月29日(2022.6.29)

(24)登録日 令和4年6月21日(2022.6.21)

(51)国際特許分類

F I

C 1 2 N 15/90 (2006.01)

C 1 2 N

15/90

1 0 0 Z

C 1 2 N 1/19 (2006.01)

C 1 2 N

1/19

Z N A

C 1 2 N 1/15 (2006.01)

C 1 2 N

1/15

C 1 2 N 1/21 (2006.01)

C 1 2 N

1/21

C 1 2 N 5/10 (2006.01)

C 1 2 N

5/10

請求項の数 14 (全80頁)

(21)出願番号 特願2018-560571(P2018-560571)

(86)(22)出願日 平成29年5月18日(2017.5.18)

(65)公表番号 特表2019-516378(P2019-516378  
A)

(43)公表日 令和1年6月20日(2019.6.20)

(86)国際出願番号 PCT/US2017/033369

(87)国際公開番号 WO2017/201311

(87)国際公開日 平成29年11月23日(2017.11.23)

審査請求日 令和2年5月7日(2020.5.7)

(31)優先権主張番号 62/338,412

(32)優先日 平成28年5月18日(2016.5.18)

(33)優先権主張国・地域又は機関  
米国(US)

(73)特許権者 511289736

アミリス, インコーポレイテッド  
アメリカ合衆国 カリフォルニア 9 4 6  
0 8, エミリービル, ホリス ストリ  
ート 5 8 8 5, スイート 1 0 0

(74)代理人 100079108

弁理士 稲葉 良幸

(74)代理人 100109346

弁理士 大貫 敏史

(74)代理人 100117189

弁理士 江口 昭彦

(74)代理人 100134120

弁理士 内藤 和彦

(72)発明者 ジョージ, ケビン

アメリカ合衆国, カリフォルニア州 9  
最終頁に続く

(54)【発明の名称】 外因性ランディングパッドへの核酸のゲノム組込みのための組成物および方法

## (57)【特許請求の範囲】

## 【請求項 1】

宿主細胞のゲノムに1つまたは複数の外因性ドナー核酸を組み込むための方法であって、  
(a 1)または(a 2)のいずれか：

(a 1) 宿主細胞のゲノム中に組み込まれた1つまたは複数の外因性ランディングパッド  
を含む宿主細胞であって、各外因性ランディングパッドは、上流ランディングパッド相同  
配列(ULP)と下流ランディングパッド相同配列(DLP)との間に位置するヌクレアーゼ  
標的配列(NTS)を含む、宿主細胞を、

(i) 1つまたは複数の第1のコンポーネントポリヌクレオチドであって、各第1のコン  
ポーネントポリヌクレオチドは、5'から3'方向に、

(1) 1つまたは複数の外因性ランディングパッドのいずれかの任意の(ULP)と相同組  
換えが可能な上流ライブラリー配列(UL)；

(2) 第1の目的の核酸；および

(3) 第1のリンカー配列

を含む、第1のコンポーネントポリヌクレオチド；および

(ii) 1つまたは複数の最後のコンポーネントポリヌクレオチドであって、各最後のコン  
ポーネントポリヌクレオチドは、5'から3'方向に、

(1) 最後のリンカー配列；

(2) 最後の目的の核酸；および

(3) 1つまたは複数の外因性ランディングパッドのいずれかの任意の(DLP)で、相同

組換えが可能な下流ライブラリー配列（DL）、  
を含む、最後のコンポーネントポリヌクレオチド  
と接触させる工程であって、

1つまたは複数の第1のコンポーネントポリヌクレオチドの任意の第1のリンカー配列は、  
1つまたは複数の最後のコンポーネントポリヌクレオチドの任意の最後のリンカー配列  
と相同組換えが可能である、工程；または

（a2）宿主細胞のゲノム中に組み込まれた1つまたは複数の外因性ランディングパッド  
を含む宿主細胞であって、各外因性ランディングパッドは、上流ランディングパッド相同  
配列（ULP）と下流ランディングパッド相同配列（DLP）との間に位置するヌクレアーゼ  
標的配列（NTS）を含む、宿主細胞を、

10

（i）1つまたは複数の第1のコンポーネントポリヌクレオチドであって、各第1のコン  
ポーネントポリヌクレオチドは、5'から3'方向に、1つまたは複数の外因性ランディ  
ングパッドの任意の（ULP）と相同組換えが可能な上流ライブラリー配列（UL）、目的の核  
酸、およびリンカー配列を含む、第1のコンポーネントポリヌクレオチド；

（ii）1つまたは複数の中間コンポーネントポリヌクレオチドであって、各中間コン  
ポーネントポリヌクレオチドは、5'から3'方向に、第1のリンカー配列、目的の核酸、お  
よび第2のリンカー配列を含む、中間コンポーネントポリヌクレオチド；および

（iii）1つまたは複数の最後のコンポーネントポリヌクレオチドであって、各最後の  
コンポーネントポリヌクレオチドは、5'から3'方向に、最後のリンカー配列、目的の核  
酸、および1つまたは複数の外因性ランディングパッドの任意の（DLP）と相同組換えが  
可能な下流ライブラリー配列（DL）を含む、最後のコンポーネントポリヌクレオチド  
と接触させる工程であって、

20

各リンカー配列は、別のリンカー配列と相同組換えが可能である、工程；

（b）前記宿主細胞を、（NTS）に結合し、1つまたは複数の外因性ランディングパッド  
内の部位を切断することが可能な、1つまたは複数のヌクレアーゼ（N）と接触させる工程  
；および

（c）接触させた宿主細胞から生成した宿主細胞を回収する工程；

を含み、

（a1）について、1つまたは複数の第1のコンポーネントポリヌクレオチドと、1つま  
たは複数の最後のコンポーネントポリヌクレオチドとの任意の組合せが、インビボでそれ  
らのリンカー配列を介して相同組換えされ、該組合せが、各ランディングパッド周辺のゲ  
ノム配列から独立して、1つまたは複数の外因性ランディングパッドのいずれかにおいて  
組み込まれ；

30

（a2）について、1つまたは複数の第1のコンポーネントポリヌクレオチド、1つま  
たは複数の中間コンポーネントポリヌクレオチド、および1つまたは複数の最後のコン  
ポーネントポリヌクレオチドの任意の組合せが、インビボで上流のリンカー配列、第1のリン  
カー配列、第2のリンカー配列および/または最後のリンカー配列を介して相同組換えさ  
れ、該組合せが、各ランディングパッド周辺のゲノム配列から独立して、1つまたは複数  
の外因性ランディングパッドのいずれかにおいて組み込まれる、方法。

#### 【請求項2】

40

（a1）について、2つ以上の第1のコンポーネントポリヌクレオチドが、互いに同一な  
上流ライブラリー配列（UL）、および互いに同一な第1のリンカー配列を含み；1つま  
たは複数の最後のコンポーネントポリヌクレオチドが、互いに同一な最後のリンカー配列、  
および互いに同一な下流ライブラリー配列（DL）を含み；

（a2）について、2つ以上の第1のコンポーネントポリヌクレオチドが、互いに同一な  
上流ライブラリー配列（UL）、および互いに同一なリンカー配列を含み；2つ以上の中間  
コンポーネントポリヌクレオチドが、互いに同一な第1のリンカー配列、および互いに同  
一な第2のリンカー配列を含み；2つ以上の最後のコンポーネントポリヌクレオチドが、  
互いに同一なリンカー配列、および互いに同一な下流ライブラリー配列（DL）を含む、請  
求項1に記載の方法。

50

## 【請求項 3】

(ULP) および (DLP) の各々が、20ヌクレオチド～5,000ヌクレオチド長、25ヌクレオチド～1000ヌクレオチド長、25ヌクレオチド～500ヌクレオチド長、または100ヌクレオチド～500ヌクレオチド長を含む、請求項1または2に記載の方法。

## 【請求項 4】

宿主細胞が、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19または20個の外因性ランディングパッドを含む、請求項1～3のいずれか一項に記載の方法。

## 【請求項 5】

1つまたは複数の外因性ランディングパッドが、宿主細胞のゲノム中の選択されたニュートラルな遺伝子座において組み込まれている、請求項1～4のいずれか一項に記載の方法。

10

## 【請求項 6】

1つまたは複数の外因性ランディングパッドが、宿主細胞のゲノム中の遺伝子間領域において組み込まれている、請求項1～5のいずれか一項に記載の方法。

## 【請求項 7】

宿主細胞が、宿主細胞のゲノム中に組み込まれた少なくとも1つの二次ランディングパッド、二次上流ランディングパッド相同配列と二次下流ランディング相同性パッド配列との間に位置する二次ヌクレアーゼ標的配列を含む二次ランディングパッドをさらに含み、  
(a) 二次上流ランディングパッド配列は、1つまたは複数の外因性ランディングパッドの (ULP) とは異なっているか；

20

(b) 二次下流ランディングパッド配列は、1つまたは複数の外因性ランディングパッドの (DLP) とは異なっているか；

(c) 二次ヌクレアーゼ標的配列は、1つまたは複数の外因性ランディングパッドの (NTS) とは異なっているか；または

(d) それらの任意の組合せである、請求項1～6のいずれか一項に記載の方法。

## 【請求項 8】

(ULP) または (DLP) のいずれかまたは両方のヌクレオチド配列が、宿主細胞のゲノムの内因性ゲノム配列に実質的な相同性を有さないランダムに生成したヌクレオチド配列から得られる、請求項1～7のいずれか一項に記載の方法。

## 【請求項 9】

30

(ULP) または (DLP) のいずれかまたは両方のヌクレオチド配列が、宿主細胞のゲノム中に存在しない、請求項1～8のいずれか一項に記載の方法。

## 【請求項 10】

外因性ランディングパッドの各々が、上流内因性ゲノム配列と (ULP) の5'領域との間に位置するインシュレーター配列、(DLP) の3'領域と下流内因性ゲノム配列との間に位置するインシュレーター配列、または両方の配置に位置するインシュレーター配列を含む、請求項1～9のいずれか一項に記載の方法。

## 【請求項 11】

各第1のコンポーネントポリヌクレオチド、中間コンポーネントポリヌクレオチド、および最後のコンポーネントポリヌクレオチドが、追加の機能的なエレメントを含む、請求項1～10のいずれか一項に記載の方法。

40

## 【請求項 12】

追加の機能的なエレメントが、バーコード、(NTS) と異なる二次ヌクレアーゼ標的部位、シス調節エレメントのためのDNA結合部位、またはそれらの任意の組合せからなる群から選択される、請求項11に記載の方法。

## 【請求項 13】

宿主細胞が、真菌細胞、細菌細胞、植物細胞、および動物細胞からなる群から選択される、請求項1～12のいずれか一項に記載の方法。

## 【請求項 14】

真菌細胞が、サッカロマイセス・セレビシエ細胞である、請求項13に記載の方法。

50

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

[0001] 本明細書で提供される組成物および方法は、一般的に、分子生物学および遺伝子工学の分野に関する。

## 【背景技術】

## 【0002】

[0002] 遺伝子工学技術は、様々な分野で宿主細胞のゲノムに改変を導入するのに使用されてきた。合成生物学の分野において、微生物を操作して標的分子を生産することはしばしば、宿主細胞のゲノムへの多数の異なる核酸の導入および組込みを伴う。発酵のための工業用微生物は、重度に遺伝子改変されており、新たな標的分子生産に複雑な生合成経路を操作するための標準的な技術を使用することは、膨大な量の時間と資源を消費する可能性がある。さらに、標的分子を生産するように株が一旦操作されたら、追加の性能の獲得はしばしば、単一の標的設計ではなく、複数のパラメーターの同時の改変（例えば、複数の遺伝子の過剰発現を、それ以外の下方制御と組み合わせること）を必要とする。

## 【0003】

[0003] 一般的に、複数のパラメーターを改変するためのコンビナトリアル戦略は、株のエンジニア（strain engineer）によって予め決定されることになる。例えば、株のエンジニアは、宿主細胞に導入しようとする核酸コンストラクトのどの組合せがその表現型を改善するのかを手作業で決定することになる。しかしながら、大規模なコンビナトリアルなサーチの余地をカバーするために、何百または何千もの予め決定された核酸コンストラクトの組合せを構築し、宿主細胞に形質転換し、スクリーニングする必要があると予想される。このような大規模な核酸コンストラクトの組合せをスクリーニングして経験的に最適な組合せを決定することは、時間がかかると予想される。例えば、スクリーニングプロセスは、宿主細胞を何百または何千もの核酸コンストラクトの組合せで形質転換すること、各形質転換をPCRで検証すること、および得られた形質転換宿主細胞の表現型を分析することを包含すると予想される。多くの場合において、異なる宿主細胞のバックグラウンドは、組換えのための好適な相同性領域がない可能性があり、または宿主細胞の機能に負の影響を与えることなく核酸コンストラクトを組み込むことができる「オープンな」ニュートラルなゲノム遺伝子座の利用可能性によって制限されると予想されることから、核酸コンストラクトを再使用することは、そのようなバックグラウンドにおいて問題を生じる可能性がある。

## 【発明の概要】

## 【発明が解決しようとする課題】

## 【0004】

[0004] それゆえに、核酸コンストラクトの宿主細胞への導入および組込みのための改善された方法および組成物への必要性がある。さらに、宿主細胞の表現型および/または標的分子の産生をモジュレートするための核酸コンストラクトのコンビナトリアルな組込みのための改善された方法および組成物に対する必要性もある。

## 【課題を解決するための手段】

## 【0005】

[0005] 宿主細胞媒介組換えを介した外因性ドナー核酸のゲノム組込みは、典型的には、各外因性ドナー核酸が、ゲノムの部位特異的な相同配列を有することを必要とする。これらの配列は、外因性ドナー核酸を特異的なゲノム部位で相同組換え(homologously recombine)を可能にする。異なるゲノム部位で、または異なる宿主細胞のバックグラウンドで外因性ドナー核酸を組み込むことが望ましい場合、外因性ドナー核酸は、組込みのための新しいゲノム部位に適合性を有する相同配列が包含されるように改変され再カスタマイズされることを必要とする。このようなカスタマイゼーションは、時間と資源を消費する。

## 【0006】

[0006] したがって、一態様において、宿主細胞のゲノムへの組込みのための外因性ドナ

10

20

30

40

50

ー核酸の再カスタマイゼーションを最小化する組成物および方法が本明細書で提供される。一実施形態において、宿主細胞のゲノムは、標準化された相同配列を含む外因性ドナー核酸の組込みをさらに容易にするのに使用できるランディングパッドを含むように改変される。宿主細胞のゲノム中に操作により作製されたランディングパッドは、外因性ドナー核酸の標準化された相同配列の相同組換えが可能なランディングパッド相同配列を含む。ある特定の実施形態において、外因性ドナー核酸は、ゲノム組込みのためにゲノムの部位特異的な相同配列を頼らない標準化された相同配列を含むことから、外因性ドナー核酸は、再カスタマイズする必要がなく、適合性を有するランディングパッドが存在する限り、異なる宿主細胞のバックグラウンドで、または異なるゲノム遺伝子座で再使用することができる。

10

**【 0 0 0 7 】**

[0007] 特定の実施形態において、宿主細胞のゲノム中の各ランディングパッドは、上流ランディングパッド相同配列（ULP）と下流ランディングパッド相同配列（DLP）との間に位置するヌクレアーゼ標的配列（NTS）をさらに含む。特定の実施形態において、ヌクレアーゼ標的配列（NTS）は、部位特異的なヌクレアーゼによって認識され結合できるヌクレオチド配列を含む。外因性ランディングパッドを有する宿主細胞は、ランディングパッド相同配列に適合性を有する相同配列を含む外因性ドナー核酸（ES）と同時に、部位特異的なヌクレアーゼと接触させることができる。ヌクレアーゼは、ランディングパッドで標的化された二本鎖破断を引き起こすと予想され、これは、ヌクレアーゼと接触していない宿主細胞と比較して（ES）組込み効率の有意な増加をもたらすと予想される。

20

**【 0 0 0 8 】**

[0008] 特定の実施形態において、外因性ドナー核酸は、モジュラー式のDNAパーツである（コンポーネントポリヌクレオチドとも称される）。ある特定の実施形態において、単一のコンポーネントポリヌクレオチドそれだけでは、宿主細胞のゲノム中のランディングパッドでの相同組換えのための全ての相同配列を有していない。しかしながら、宿主細胞を複数のコンポーネントポリヌクレオチドと接触させる場合、適合性リンカー配列を有するこれらのコンポーネントポリヌクレオチドの2つ以上が、インビボで相同組換えして、組み立てられたコンポーネントポリヌクレオチドを生成することができる。組み立てられたコンポーネントポリヌクレオチドは、このようにしてランディングパッド相同配列に適合性を有する相同配列を含むこととなり、ランディングパッドでの相同組換えが可能になる。コンポーネントポリヌクレオチドは、コンビナトリアル的に互いに組み換えることができることから、コンポーネントポリヌクレオチドの使用は、特異的なゲノム遺伝子座に標的化される単一型の外因性ドナー核酸と比較して、外因性ドナー核酸のハイスループットな組込みおよびスクリーニングのために多大な分子多様性をもたらすことができる。

30

**【 0 0 0 9 】**

[0009] 別の態様において、宿主細胞は、宿主細胞のゲノム中に操作により作製された複数の標準化されたランディングパッドと共に提供される。用語「標準化された」ランディングパッドは、本明細書において、各ランディングパッドが、他のランディングパッド中のものと同じまたは実質的に同一な一対のランディングパッド相同配列（ULP）および（DLP）を含む複数のランディングパッドを指すものとして使用される。ある特定の実施形態において、外因性ドナー核酸もそれらの5'および3'領域において適合性を有する相同性領域で標準化される場合、それらは、それら各々のランディングパッド相同配列との相同組換えが可能である。外因性ドナー核酸は、標準化されたランディングパッドのいずれかで宿主細胞のゲノム中にランダムに組込みを起すことができる。標準化されたランディングパッドのいずれかへの外因性ドナー核酸の組込みのランダムさは、大量の外因性ドナー核酸または外因性ドナー核酸のライブラリーを迅速かつコンビナトリアル的に組み込むのに有効な手段を提供することができる。

40

**【 0 0 1 0 】**

[0010] 例えば、宿主細胞のゲノム中に組み込まれた30個の外因性ドナー核酸と3つの標準化されたランディングパッドとを用いれば、30個の外因性ドナー核酸のいずれかは

50

、3つの標準化されたランディングパッドのいずれかへの組込みを起こすことができ、その結果として、宿主細胞のゲノムに組み込まれた外因性ドナー核酸の27,000個の異なる組合せを含む可能性がある宿主細胞の集団が生じる。一方で、3つのゲノムの特異的な組込み部位への組込みを起こすように設計された30個の外因性ドナー核酸を用いれば、宿主細胞のゲノムに組み込まれた外因性ドナー核酸の1000個の異なる組合せを含む宿主細胞の集団を生成することができる。

#### 【0011】

[0011] コンビナトリアルな組込み多様性は、宿主細胞のゲノムに組み込もうとする外因性ドナー核酸としてコンポーネントポリヌクレオチドを利用することによってさらに強化することができる。例えば、上記で例示された例において、各外因性ドナー核酸が2つのコンポーネントポリヌクレオチドとして宿主細胞に導入される場合、それは、遺伝子型の216,000個の異なる組合せを含む可能性がある宿主細胞の集団を生成することができる。標準化されたランディングパッドのいずれかへの(特異的なゲノム部位においてではなく)組み立てられたコンポーネントポリヌクレオチドの組込みのランダムさは、宿主細胞のゲノムに組み込まれたドナー核酸の分子多様性をさらに強化する。それゆえに、このようなランダムさは、改変された宿主細胞によって呈される望ましい特色を包含するより一層大規模な表現型のバリエーションをもたらすことができる。

10

#### 【0012】

[0012] したがって、一態様において、宿主細胞のゲノム中に組み込まれた1つまたは複数の(x)外因性ランディングパッドを含む宿主細胞であって、各外因性ランディングパッドは、上流ランディングパッド相同配列(ULP)と下流ランディングパッド相同配列(DLP)との間に位置するヌクレアーゼ標的配列(NTS)を含む、宿主細胞が本明細書で提供される。ある特定の実施形態において、宿主細胞のゲノム中の(x)外因性ランディングパッドの数は、約1~約100、約2~約50、約3~約20などのあらゆる好適な数であり得る。ある特定の実施形態において、適合性を有する相同配列を有する外因性ドナー核酸が、標準化されたランディングパッドのいずれかで宿主細胞媒介相同組換えを介して組込みを起こすことができるように、全てのランディングパッドは、標準化されたランディングパッド相同配列を含む。ある特定の実施形態において、ランディングパッドは、宿主細胞のゲノム中の選択されたニュートラルな遺伝子座において組み込まれている。特定の実施形態において、ランディングパッドは、宿主細胞のゲノム中の遺伝子間領域(intergenic region)において組み込まれている。

20

30

#### 【0013】

[0013] 別の態様において、1つまたは複数の外因性ドナー核酸(ES)であって、各外因性ドナー核酸は、上流ライブラリー配列(UL)と下流ライブラリー配列(DL)との間に位置する目的の核酸(D)を含み、各(UL)は、(x)外因性ランディングパッドのいずれかの(ULP)で相同組換えが可能であり、各(DL)は、(x)外因性ランディングパッドのいずれかの(DLP)で相同組換えが可能である、外因性ドナー核酸(ES)が本明細書で提供される。ある特定の実施形態において、外因性ドナー核酸(ES)は、互いに同一な上流ライブラリー配列(UL)および互いに同一な下流ライブラリー配列(DL)を含む。ある特定の実施形態において、外因性ドナー核酸(ES)は、各ランディングパッド周辺の内因性ゲノム配列から独立して、外因性ランディングパッドのいずれかにおいて組込みを起こすことが可能である。

40

#### 【0014】

[0014] 別の態様において、宿主細胞のゲノムに外因性ドナー核酸を組み込むための方法が本明細書で提供される。ある特定の実施形態において、本方法は、  
(a) 宿主細胞のゲノム中に組み込まれた複数の(x)外因性ランディングパッドを含む宿主細胞であって、各外因性ランディングパッドは、上流ランディングパッド相同配列(ULP)と下流ランディングパッド相同配列(DLP)との間に位置するヌクレアーゼ標的配列(NTS)を含む、宿主細胞を、  
(i) 1つまたは複数の外因性ドナー核酸(ES)であって、各(ES)は、上流ライブラリ

50

ー配列 (UL) と下流ライブラリー配列 (DL) との間に位置する目的の核酸 (D) を含み、各 (UL) は、(x) 外因性ランディングパッドのいずれかの任意の (ULP) で相同組換えが可能であり、各 (DL) は、(x) 外因性ランディングパッドのいずれかの任意の (DLP) で相同組換えが可能である、外因性ドナー核酸 (ES) ; および (i i) (NTS) に結合し、1 つまたは複数の (x) 外因性ランディングパッド内の部位を切断することが可能な、1 つまたは複数のヌクレアーゼ (N)

と接触させる工程 ; および

(b) 接触させた宿主細胞から生成した宿主細胞を回収する工程を含み、

外因性ドナー核酸 (ES) のいずれかは、各ランディングパッド周辺のゲノム配列から独立して、(x) 外因性ランディングパッドのいずれかにおいて組み込まれる。一部の実施形態において、x は、少なくとも 2 の整数である。ある特定の実施形態において、少なくとも 2 つ以上の外因性ドナー核酸は、互いに同一な上流ライブラリー配列 (UL) および互いに同一な下流ライブラリー配列 (DL) を含む。

#### 【0015】

[0015] ある特定の実施形態において、上流ランディングパッド相同配列 (ULP) および下流ランディングパッド相同配列 (DLP) の各々は、各々上流ライブラリー配列 (UL) および下流ライブラリー配列 (DL) のそれぞれと相同性を有する約 100、約 200、または約 500 塩基対を含む。ある特定の実施形態において、外因性ドナー核酸の上流ライブラリー配列 (UL) および下流ライブラリー配列 (DL) の各々は、約 500 塩基対の長さである。

#### 【0016】

[0016] ある特定の実施形態において、外因性ドナー核酸の 2 つ以上は、宿主細胞に共に形質転換されるコンポーネントポリヌクレオチドである。一実施形態において、本方法は、(a) 宿主細胞のゲノム中に組み込まれた 1 つまたは複数の (x) 外因性ランディングパッドを含む宿主細胞であって、各外因性ランディングパッドは、上流ランディングパッド相同配列 (ULP) と下流ランディングパッド相同配列 (DLP) との間に位置するヌクレアーゼ標的配列 (NTS) を含む、宿主細胞を、

(i) 1 つまたは複数の第 1 のコンポーネントポリヌクレオチドであって、各第 1 のコンポーネントポリヌクレオチドは、5' から 3' 方向に、

(1) 1 つまたは複数の (x) 外因性ランディングパッドの任意の (ULP) と相同組換えが可能な上流ライブラリー配列 (UL) ;

(2) 第 1 の目的の核酸、および

(3) 第 1 のリンカー配列

を含む、第 1 のコンポーネントポリヌクレオチド ;

(ii) 1 つまたは複数の最後のコンポーネントポリヌクレオチドであって、各最後のコンポーネントポリヌクレオチドは、5' から 3' 方向に、

(1) 最後のリンカー配列、

(2) 最後の目的の核酸、および

(3) 1 つまたは複数の (x) 外因性ランディングパッドのいずれかの任意の (DLP) と相同組換えが可能な下流ライブラリー配列 (DL)

を含み、

1 つまたは複数の第 1 のコンポーネントポリヌクレオチドの任意の第 1 のリンカー配列は、1 つまたは複数の最後のコンポーネントポリヌクレオチドの任意の最後のリンカー配列と相同組換えが可能である、最後のコンポーネントポリヌクレオチド ; および

(iii) (NTS) に結合し、1 つまたは複数の (x) 外因性ランディングパッド内の部位を切断することが可能な、1 つまたは複数のヌクレアーゼ (N)

と接触させる工程 ; および

(b) 接触させた宿主細胞から生成した宿主細胞を回収する工程を含み、

10

20

30

40

50

1つまたは複数の第1のコンポーネントポリヌクレオチドからの第1のコンポーネントポリヌクレオチドと、1つまたは複数の最後のコンポーネントポリヌクレオチドからの最後のコンポーネントポリヌクレオチドとの任意の組合せが、インピボでそれらのリンカー配列を介して相同組換えされ、該組合せが、各ランディングパッド周辺のゲノム配列から独立して、1つまたは複数の(x)外因性ランディングパッドのいずれかにおいて組み込まれる。ある特定の実施形態において、xは、少なくとも1の整数である。ある特定の実施形態において、xは、少なくとも2の整数である。

#### 【0017】

[0017] 一部の実施形態において、3つ以上のコンポーネントポリヌクレオチドは、ランディングパッドでの組込みのために宿主細胞に共に形質転換される。一実施形態において、コンポーネントポリヌクレオチドは、(a) 1つまたは複数の第1のコンポーネントポリヌクレオチドであって、各第1のコンポーネントポリヌクレオチドは、5'から3'方向に、上流ランディングパッド相同配列(ULP)と相同組換えが可能な上流ライブラリー配列、D<sub>0</sub>群から選択される任意のDNAセグメント、リンカー配列LB<sub>0</sub>を含む、第1のコンポーネントポリヌクレオチド；1つまたは複数の中間(intermediate)コンポーネントポリヌクレオチドであって、各中間コンポーネントポリヌクレオチドは、5'から3'方向に、第1のリンカー配列LA<sub>n</sub>、D<sub>n</sub>群から選択される任意のDNAセグメント、第2のリンカー配列LB<sub>n</sub>を含み、nは、1から中間コンポーネントポリヌクレオチドの数までの整数を表す、中間コンポーネントポリヌクレオチド；および(c) 1つまたは複数の最後のコンポーネントポリヌクレオチドであって、各最後のコンポーネントポリヌクレオチドは、5'から3'方向に、リンカー配列LA<sub>m</sub>、D<sub>m</sub>群から選択される任意のDNAセグメント、および下流ランディングパッド相同配列(DLP)と相同組換えが可能な下流ライブラリー配列を含む、最後のコンポーネントポリヌクレオチドを含む。このような実施形態において、各リンカー配列LB<sub>(p-1)</sub>は、リンカー配列LA<sub>p</sub>と相同組換えが可能であり、nは、1~(m-1)の様々な整数であり、pは、1~mの整数を表し、各D<sub>0</sub>群、...D<sub>n</sub>群、...D<sub>m</sub>群、は、独立して、1つまたは複数のDNAセグメントからなる。この実施形態において、1つまたは複数の第1のコンポーネントポリヌクレオチドからの第1のコンポーネントポリヌクレオチド、1つまたは複数の中間コンポーネントポリヌクレオチドからの中間コンポーネントポリヌクレオチド、および1つまたは複数の最後のコンポーネントポリヌクレオチドからの最後のコンポーネントポリヌクレオチドの任意の組合せが、インピボでそれらのリンカー配列を介して相同組換えされ、該組合せが、各ランディングパッド周辺のゲノム配列から独立して、1つまたは複数の(x)外因性ランディングパッドのいずれかにおいて組み込まれる。一部の実施形態において、xは、少なくとも1の整数である。一部の実施形態において、xは、少なくとも2の整数である。

#### 【0018】

[0018] 別の態様において、ゲノム組込みの方法は、本明細書に記載される、(a1) 1つまたは複数の外因性ドナー核酸(ES)、(a2) 1つまたは複数の第1のコンポーネントポリヌクレオチドおよび1つまたは複数の最後のコンポーネントポリヌクレオチド、ならびに(a3) 1つまたは複数の第1のコンポーネントポリヌクレオチド、1つまたは複数の中間コンポーネントポリヌクレオチド、および1つまたは複数の最後のコンポーネントポリヌクレオチドの任意の組合せと、細胞を接触させることを含む。

#### 【0019】

[0019] 別の態様において、宿主細胞のゲノム中に組み込まれたランディングパッドの一部は、特異的なゲノム遺伝子座を標的化するように設計される。例えば、標的化されたランディングパッドは、そのネイティブの遺伝子座で、内因性遺伝子に隣接して組み込まれていてもよい。このような標的化されたランディングパッドは、プロモーターライブラリーをスクリーニングし、内因性遺伝子の発現の適正量を決定して(titrate)、最適なプロモーターを決定するのに使用することができる。別の例において、標的化されたランディングパッドを、目的の遺伝子のオープンリーディングフレームの3'に組み込んで、ターミネーターライブラリーまたはデグロンライブラリーをスクリーニングすることができる。こ

10

20

30

40

50



これらの実施形態において、標的化されたランディングパッドに適合性を有する標準化されたライブラリーまたはリンカー配列を含むコンポーネントポリヌクレオチドを使用することができる。ある特定の実施形態において、標的化されたランディングパッドは、1つまたは複数の標準化された外因性ランディングパッドと組み合わせて使用することができ、これは、標準化された外因性ランディングパッド周辺のゲノム配列から独立して、あらゆる外因性ドナー核酸を組み込むのに使用することができる。

#### 【0020】

[0020] 別の態様において、外因性ドナー核酸による表現型の寄与を決定する方法が本明細書で提供される。ある特定の実施形態において、本方法は、(a) 宿主細胞を、複数の外因性核酸(ES)と接触させる工程であって、各(ES)は、バーコード配列でタグ付けされており、宿主細胞のゲノムは、宿主細胞媒介相同組換えを介して、宿主細胞のゲノム中の複数の(ES)の任意の1つまたは組合せを組み込むように設計されている、工程；(b) 接触させた宿主細胞から生成した、特異的な表現型を呈する宿主細胞をスクリーニングする工程；および(c) 各(ES)に関連付けられたバーコード配列を使用して、宿主細胞のゲノム中に組み込まれた(ES)のアイデンティティ(identity)を決定する工程を含む。ある特定の実施形態において、複数の外因性ドナー核酸(ES)の少なくとも2つは、適合性リンカー配列を有するコンポーネントポリヌクレオチドである。コンポーネントポリヌクレオチドは、インビボで相同組換えして、宿主細胞媒介相同組換えを介してランディングパッドにおいて組み込まれている組み立てられたコンポーネントポリヌクレオチドを生成する。各コンポーネントポリヌクレオチドは、個々にバーコード配列でタグ付けされ、これは、どのコンポーネントポリヌクレオチドの組合せがランディングパッド中に組み込まれているかを決定するのに使用することができる。

#### 【0021】

[0021] 別の態様において、本明細書に記載される1つまたは複数のランディングパッドを含む改変された宿主細胞が本明細書で提供される。

#### 【0022】

[0022] 別の態様において、外因性核酸のゲノム組込みを実行するのに有用なキットが本明細書で提供される。一部の実施形態において、キットは、(a) 本明細書に記載される1種または複数の宿主細胞；(b) 本明細書に記載される1つまたは複数の外因性ドナー核酸；および/または(c) 本明細書に記載される1つまたは複数のヌクレアーゼを含む。一部の実施形態において、キットは、複数のプライマー対をさらに含み、各プライマー対は、各外因性ドナー核酸に関連付けられたバーコード配列を同定することが可能である。

#### 【0023】

[0023] これらのおよび他の実施形態を、その利点および特徴の多くと共に、以下の文章および添付された図面と共により詳細に記載する。

#### 【図面の簡単な説明】

#### 【0024】

【図1-1】[0024]図1Aは、ALG1遺伝子座(「iALG1」)の遺伝子間領域に位置する、「UL」と表記される上流ランディングパッド配列(ULP)、「DL」と表記される下流ランディングパッド配列(DLP)、および「X」と表記されるヌクレアーゼ標的配列(NTS)を有する外因性ランディングパッドの例示的な実施形態を例示する図である。外因性ドナー核酸(ES)との相同組換えに使用される外因性ランディングパッドの(ULP)および(DLP)は、全ての図面にわたり点描を含むボックスとして示される。[0025] 図1Bは、「UL」と表記される上流ライブラリー配列(UL)および「DL」と表記される下流ライブラリー配列(DL)を有する3つの外因性ドナー核酸のライブラリーを例示する図であり、各外因性ドナー核酸のULおよびDLは、それぞれ外因性ランディングパッドの(ULP)および(DLP)と相同組換えが可能である。3つの外因性ドナー核酸の各々は、目的の核酸として、オープンリーディングフレームと、リンカー配列(「A」)を介してプロモーター(矢印として示される)に作動可能に連結したターミネーターを含む。目的の核酸は、各外因性ドナー核酸中で(UL)と(DL)との間に位置する。[0026]図1

Cは、外因性ドナー核酸のマーカースレグノム組込みの例示的な実施形態を例示する図である。外因性ランディングパッドは、「X」と表記される(NTS)で、部位特異的ヌクレアーゼ(一丁のはさみとして示される)によって切断される。

【図1-2】[0027]図1Dは、外因性ランディングパッドの別の実施形態を例示する図であり、図1Dにおいて、外因性ランディングパッドは、「UL」と表記される(ULP)の5'末端に隣接するインシュレーター配列(例えば、転写ターミネーター)および「DL」と表記される(DLP)の3'末端に隣接する別のインシュレーター配列をさらに含む。ランディングパッド中に組み込まれたインシュレーター配列は、宿主細胞のゲノム中の外因性ランディングパッドの外側に配置された別のプロモーターからの予期せぬ「リードスルー」転写を防いだりまたは低減させたりすることができる。[0028]図1Eは、ランディングパッドに組み込むことができる3つの異なるタイプの外因性ドナー核酸を例示する図である。図1Eの一番上は、単一のDNAパーツで作製される外因性ドナー核酸を例示する。図1Eの中央は、インビボで組み立てられてランディングパッドに組み込むことができる2つの別個のコンポーネントポリヌクレオチドを例示する。図1Eの一番下は、インビボで組み立てられてランディングパッドに組み込むことができる3つの別個のコンポーネントポリヌクレオチドを例示する。[0029]図1Fは、第1のコンポーネントポリヌクレオチドとしてのプロモーターのライブラリー、中間コンポーネントポリヌクレオチドとしてのオープンリーディングフレームのライブラリー、および最後のコンポーネントポリヌクレオチドとしてのターミネーター配列のライブラリーを例示する図であり、これらは、コンビナトリアル的に組み合わせ、適合性を有するランディングパッド相同配列と共にランディングパッドに組み込むことができる。

【図2】[0030]図2Aは、宿主細胞のゲノム中に挿入された3つの標準化された外因性ランディングパッドの例示的な実施形態を例示する図である。標準化された外因性ランディングパッドは、宿主細胞のゲノム中のALG1遺伝子座、MGA1遺伝子座、およびYCT1遺伝子座の遺伝子間領域において組み込まれた、同じ上流ランディングパッド相同配列(ULP、「UL」として示される)および同じ下流ランディングパッド相同配列(DLP、「DL」として示される)を含む。[0031]図2B~2Dは、様々なDNA「ステッチ(stitch)」設計における外因性ドナー核酸の異なるライブラリーを例示する図である。各外因性ドナー核酸は、標準化された外因性ランディングパッドに相同組換えするための少なくとも1つの標準化された相同配列(UL、DLまたはULおよびDLの両方)を含む。[0032]図2Bは、外因性ドナー核酸のライブラリーの例示的な実施形態を例示する図であり、このようなライブラリーの各々は、上流ライブラリー配列(UL)と下流ライブラリー配列(DL)との間に位置する、プロモーターに作動可能に連結した(すなわち、「ロックドプロモーター」設計での)単一のオープンリーディングフレーム(ORF)を含む。[0033]図2Cは、コンポーネントポリヌクレオチドである外因性ドナー核酸のライブラリーの例示的な実施形態を例示する図である。宿主細胞中で、適合性リンカー配列を有する2つのコンポーネントポリヌクレオチドがインビボで相同組換えを介して組み立てられる場合、それらは、様々なマルチオープンリーディングフレームをロックドプロモーター設計で生成する。図2Cにおいて、適合性リンカー配列は、「SNAP」として示される。[0034]図2Dは、コンポーネントポリヌクレオチドである外因性ドナー核酸のライブラリーの別の例示的な実施形態を例示する図である。コンポーネントポリヌクレオチドがインビボで相同組換えを介して組み立てられる場合、それらは、プロモーターに作動可能に連結した単一のORFの異なる組合せを生成する。別個のコンポーネントポリヌクレオチドにおいて、プロモーターはORFから割裂することから、このDNAステッチ設計は、「スプリットプロモーター」設計と称される。

【図3】[0035]図3Aは、追加の機能的なエレメント、すなわち、外因性ランディングパッド中に存在するヌクレアーゼ標的配列(NTS)と異なるバーコード配列およびヌクレアーゼ標的配列を含む例示的な外因性ドナー核酸(これは、コンポーネントポリヌクレオチドである)を例示する図である。図3Aに示される第1のコンポーネントポリヌクレオチドは、「Y-カッター」ヌクレアーゼ標的配列を含む。図3Bは、追加の機能的なエレ

10

20

30

40

50

ント、すなわち、外因性ランディングパッド中に存在するヌクレアーゼ標的配列（NTS）と異なるバーコード配列およびヌクレアーゼ標的配列を含む例示的な外因性ドナー核酸（これは、コンポーネントポリヌクレオチドである）を例示する図である。図3Bに示される最後のコンポーネントポリヌクレオチドは、「Z-カッター」ヌクレアーゼ標的配列を含む。[0036]図3Cは、宿主細胞のゲノム中に操作により作製された3つの標準化された外因性ランディングパッドを例示する図である。[0037]図3Dは、図3Cに示される標準化された外因性ランディングパッドに組み込まれたバーコードが付いたコンポーネントポリヌクレオチドの異なる組合せを例示する図である。図3Dはまた、宿主細胞のゲノム中に組み込まれたバーコード配列の周りのプライマー結合セグメントにアニールするように設計された、一対の標準オリゴヌクレオチドプライマーの位置（矢印により示される）も例示する。

10

【図4】[0038]図4は、ヌクレアーゼ、X-カッター（すなわち、CphI）をコードするプラスミドと共に、第1のコンポーネントポリヌクレオチドおよび最後のコンポーネントポリヌクレオチドのコンビナトリアルライブラリーのゲノム組込みのための例示的なワークフローを例示する図である。例示的なワークフローは、コロニーの選択、ハイスループットスクリーニングアッセイ、および標準化されたランディングパッドに組み込まれた遺伝子およびプロモーターを同定するための標準プライマーを用いた有望な株の遺伝子型の決定をさらに例示する。ハイスループット組込みおよびスクリーニングアッセイから決定された遺伝子型と表現型との関係によれば、ライブラリーの組成を調整して、宿主細胞の表現型をさらに改善することができる。

20

【図5】[0039]図5Aは、3つの標準化された外因性ランディングパッド（ULおよびDLランディングパッド相同配列を含む）と、下流ライブラリー配列としてDL<sub>1</sub>を含む最後のコンポーネントポリヌクレオチドの標的化された組込みのために使用することができる遺伝子座<sub>1</sub>におけるアドレス指定可能なランディングパッド（ULおよびDL<sub>1</sub>ランディングパッド相同配列を含む）とを含む宿主細胞のゲノムの例示的な実施形態を例示する。[0039]図5Bは、3つの標準化された外因性ランディングパッド（ULおよびDLランディングパッド相同配列を含む）と、下流ライブラリー配列としてDL<sub>1</sub>を含む最後のコンポーネントポリヌクレオチドの標的化された組込みのために使用することができる遺伝子座<sub>1</sub>におけるアドレス指定可能なランディングパッド（ULおよびDL<sub>1</sub>ランディングパッド相同配列を含む）とを含む宿主細胞のゲノムの例示的な実施形態を例示する。

30

【図6】[0040]図6Aは、アドレス指定可能であり任意選択のランディングパッドをさらに取り入れることによって、標準化された外因性ランディングパッドのフレキシビリティ（ALG1遺伝子座、MGA1遺伝子座、およびYCT1遺伝子座の遺伝子間領域における）を増加させることの例示的な実施形態を例示する図である。図6Aは、宿主細胞に共に形質転換されることにより図6Bに示される全ての5つのランディングパッドへのDNAコンストラクトの同時のコンビナトリアルな組込みが可能な、プールされたDNAコンストラクト（すなわち、コンポーネントポリヌクレオチド）およびプールされたヌクレアーゼを例示する図である。図6Bは、アドレス指定可能であり任意選択のランディングパッドをさらに取り入れることによって、標準化された外因性ランディングパッドのフレキシビリティ（ALG1遺伝子座、MGA1遺伝子座、およびYCT1遺伝子座の遺伝子間領域における）を増加させることの例示的な実施形態を例示する図である。外因性ランディングパッドは、遺伝子座<sub>1</sub>であり、遺伝子座<sub>2</sub>は、標準化された外因性ランディングパッドの下流ランディングパッド相同配列（DL）と比較して、異なる下流ランディングパッド相同配列（それぞれDL<sub>1</sub>およびDL<sub>2</sub>）を含む。適合性を有するライブラリー配列（DL<sub>1</sub>およびDL<sub>2</sub>）を有する外因性ドナー核酸の特定のサブセットのみを組み込むことができるため、これらの遺伝子座におけるこれらの外因性ランディングパッドは、アドレス指定可能である。これらのアドレス指定可能な外因性ランディングパッドはまた、3つの標準化されたランディングパッドを切断するためにX-カッターが提供される場合に切断されないヌクレアーゼ標的配列「G」および「H」を含むため、それらは任意選択でもある。それらは、それぞれG-カッターおよびH-カッターが提供される場合にのみ切断される。

40

50

【図 7 - 1】[0041]図 7 A は、宿主細胞のゲノム中のネイティブの遺伝子座でオープンリーディングフレームの 5' 末端にプロモータースワップランディングパッドを取り込むのに使用することができる例示的なプロモータースワップランディングパッドビルディングコンストラクトを例示する図である。図 7 B に、宿主細胞のゲノム中に組み込まれたプロモータースワップランディングパッドを示す。図 7 B において、外因性ドナー核酸の相同組換えに使用されるプロモータースワップランディングパッド中の相同配列（UL および A）は、点描を含む領域で示される。図 7 C は、プロモーターのコンビナトリアルライブラリーおよびヌクレアーゼを用いたプロモータースワップランディングパッドを有する親株での形質転換をさらに例示する図である。図 7 D は、形質転換後、コロニーをピックアップし、望ましい表現型に関してスクリーニングし、遺伝子型解析して、プロモーターライブラリーからのどのプロモーターが宿主細胞中において望ましい表現型に寄与するのかを決定できることを例示する図である。

10

【図 7 - 2】[0042]図 7 E は、宿主細胞のゲノム中のネイティブの遺伝子座でオープンリーディングフレームの 3' 末端にターミネータースワップランディングパッドを組み込むのに使用できる例示的なターミネータースワップランディングパッドビルディングコンストラクトを例示する図である。図 7 F に、宿主細胞のゲノム中に組み込まれたターミネータースワップランディングパッドを示す。図 7 G および図 7 H は、デグロンライブラリーまたはターミネーターライブラリーのいずれかでのターミネータースワップランディングパッドを含む親株の形質転換を例示する図である。

【図 8】[0043]図 8 は、プロモータースワップランディングパッド（1）を、外因性の標準化されたランディングパッド（2）と組み合わせて使用して、これらのランディングパッドにコンポーネントポリヌクレオチドを同時に組み込んで、2 つの別個の遺伝子発現の適正量を決定する例示的な実施形態を例示する図である。

20

【図 9】[0044]図 9 は、プロモーターライブラリーを使用して 2 つのオープンリーディングフレーム（ORF1 および ORF2）の発現と、標的化された組込みとの釣り合いをとった例示的な実施形態を例示する図である。図 9 において、ORF1 を含む最後のコンポーネントポリヌクレオチドは、ALG1 遺伝子座におけるランディングパッド 1 でのみ組込みを起こすと予想され、ORF2 を含む最後のコンポーネントポリヌクレオチドは、iYCT1 遺伝子座におけるランディングパッド 2 でのみ組込みを起こすと予想される。

【図 10】[0045]図 10 A は、単一の ORF 形質転換コンポーネントを例示する図である。単一の ORF 形質転換は、UL/DL 相同性を端に有する等モル濃度の 4 つのコンストラクト：pGAL1 > GFP、pGAL1 > RFP、pGAL > RFP、およびプロモーターレスの非蛍光性「スペーサー」コンストラクトからなっていた。[0046]図 10 B は、スプリット ORF 形質転換コンポーネントを例示する図である。スプリット形質転換は、各組込み部位での組換えを容易にするために「A」リンカー相同性を含有する等モル濃度の 9 つのスプリットプロモーターおよび 4 つのプロモーターレス ORF からなっていた。pGAL1 の 1% から pGAL1 の 75% までの予測された発現レベルの範囲にかかるようにプロモーターを選んだ。[0047]図 10 C は、単一の ORF 形質転換（A）対スプリット形質転換（B）における特異的な GFP 測定頻度（ウェルの数）を示すヒストグラムを例示する図である。

30

40

【図 11】[0048]図 11 は、異なるランディングパッド相同配列を含む親株の組込み効率を例示する図である。

【図 12】[0049]図 12 は、ライブラリーコンポーネントおよびランディングパッド、ライブラリーのハイスループットスクリーニング、およびランディングパッド中に組み込まれたコンビナトリアルライブラリーコンポーネントの変異誘発と比較した性能を例示する図である。

【図 13】[0050]図 13 は、ピロリン酸イソペンテニル（IPP）産生、ならびに IPP およびジメチルアリルピロリン酸（DMAPP）のゲラニルピロリン酸（GPP）、ファルネシルピロリン酸（FPP）、およびゲラニルゲラニルピロリン酸（GGPP）への変換、ならびに様々なイソプレノイド合成のためのメバロン酸（MEV）経路の概略図を例示する図であ

50

る。

【発明を実施するための形態】

【0025】

#### 5.1 定義

【0051】用語「ランディングパッド」は、本明細書で使用される場合、外因性ドナー核酸の挿入をさらに容易にするための、宿主細胞のゲノム中に操作により作製された組換え標的部位を指す。ある特定の実施形態において、ランディングパッドは、外因性ドナー核酸の相同組換えに使用されるランディングパッド相同配列の間に位置するヌクレアーゼ標的配列（NTS）を含む。ヌクレアーゼ標的配列（NTS）は、部位特異的ヌクレアーゼによって認識され、ヌクレアーゼは、（NTS）に結合し、外因性ランディングパッド内の部位を切断する。当業者は、ヌクレアーゼが、NTSの内側または外側で切断することが可能であることを認識していると予想される。

10

【0026】

【0052】用語「外因性」ランディングパッドは、通常天然に見出されないランディングパッドを指す。ある特定の実施形態において、外因性ランディングパッドは、それらの挿入の前に宿主細胞のゲノム中に存在しないランディングパッド相同配列（例えば、上流ランディングパッド相同配列および下流ランディングパッド相同配列）を含み、および/またはそのランディングパッド相同配列は、宿主細胞のゲノム中のその天然の遺伝子座の外側に（すなわち、非ネイティブの遺伝子座に）位置する。

【0027】

20

【0053】用語「異種」または「外因性」は、通常天然に見出されないものを指す。例えば、用語「異種ヌクレオチド配列」または「外因性核酸」は、自然界で所与の細胞に通常見出されないヌクレオチド配列を指す。したがって、異種ヌクレオチド配列または外因性ヌクレオチド配列は、（a）その宿主細胞にとって外来であるもの（すなわち、細胞にとって「外因性」であるもの）；（b）宿主細胞中に天然に見出されるが（すなわち、「内因性」であるが）、細胞中で天然にはない量（すなわち、宿主細胞中に天然に見出される量より多いまたは少ない量）で存在するもの；または（c）宿主細胞中に天然に見出されるが、その天然の遺伝子座の外側に位置するものであり得る。

【0028】

【0054】用語「外因性ドナー核酸」は、宿主細胞のゲノムに組み込むために供与される外因性核酸を指す。ある特定の実施形態において、外因性ドナー核酸は、外因性ランディングパッドへの組込みに必要な相同配列を有する単一パーツのDNA断片であり得る。他の実施形態において、外因性ドナー核酸は、外因性ランディングパッドへの組込みを起こすように他のコンポーネントポリヌクレオチドと一般的に組み合わせられるコンポーネントポリヌクレオチドであり得る。

30

【0029】

【0055】用語「内因性核酸」は、本明細書で使用される場合、自然界で所与の細菌、生物、または細胞に通常見出される、および/またはそれらによって産生される核酸を指す。また「内因性核酸」は、「ネイティブの核酸」、または所与の細菌、生物、または細胞にとって「ネイティブ」であり、宿主細胞のゲノム中のその天然の遺伝子座に位置する核酸とも称される。

40

【0030】

【0056】用語「接触させる」は、本明細書で使用される場合、核酸を取り込んで宿主細胞のゲノムに組み込むことを可能にするために、核酸が、宿主細胞に十分に近接して、内部または外部に置かれることを指す。

【0031】

【0057】用語「ライブラリー」は、本明細書で使用される場合、それらの5'および3'末端に公知の共通の配列または実質的に共通の配列を含有する核酸の集合体を指す。ある特定の実施形態において、ライブラリー中の外因性ドナー核酸は、宿主細胞のゲノム中の外因性ランディングパッドへの相同組換えに使用される、その5'領域における上流ライブラ

50

リー配列およびその3'領域における下流ライブラリー配列を含む。ある特定の実施形態において、ライブラリー中のコンポーネントポリヌクレオチドは、その5'領域に上流ライブラリー配列、およびその3'領域にリンカーを含んでいてもよい。ある特定の実施形態において、ライブラリー中のコンポーネントポリヌクレオチドは、その5'領域にリンカー、およびその3'領域に下流ライブラリー配列を含んでいてもよい。ある特定の実施形態において、コンポーネントポリヌクレオチドのライブラリーは、5'および3'領域の両方にリンカーを含んでいてもよい。

#### 【0032】

[0058] 用語「切断する」、「切断」および/または「切断すること」は、本明細書で使用される場合、ヌクレアーゼ、例えばホーミングエンドヌクレアーゼ、ジンクフィンガーヌクレアーゼまたはTALE-エフェクターヌクレアーゼに関して、特定の核酸において二本鎖破断(DSB)を作り出す作用を指す。DSBは、当業者によって理解されている通り、平滑末端または付着末端(すなわち、5'または3'オーバーハング)を残すことができる。

10

#### 【0033】

[0059] 用語「操作された宿主細胞」は、本明細書で使用される場合、遺伝子工学技術(すなわち、組換えテクノロジー)を使用して親細胞を遺伝学的に改変することによって生成される宿主細胞を指す。操作された宿主細胞は、親細胞のゲノムに対してヌクレオチド配列の付加、欠失、および/または改変を含んでいてもよい。

#### 【0034】

20

[0060] 用語「相同性」は、本明細書で使用される場合、2つ以上の核酸配列、または2つ以上のアミノ酸配列間の同一性を指す。配列同一性は、同一性(または類似性もしくは相同性)のパーセンテージに関して測定することができる。パーセンテージが高いほど、配列が互いに同一により近くなる。核酸またはアミノ酸配列の相同体またはオースログは、標準的方法を使用して並べた場合、比較的高い程度の配列同一性を有する。比較のための配列のアライメント方法は当業界において周知である。様々なプログラムおよびアライメントアルゴリズムが、Smith & Waterman, Adv. Appl. Math. 2:482, 1981; Needleman & Wunsch, J. Mol. Biol. 48:443, 1970; Pearson & Lipman, Proc. Natl. Acad. Sci. USA 85:2444, 1988; Higgins & Sharp, Gene, 73:237-44, 1988; Higgins & Sharp, CABIOS 5:151-3, 1989; Corpet et al., Nuc. Acids Res. 16:10881-90, 1988; Huang et al. Computer Appls. Biosc. 8, 155-65, 1992; and Pearson et al., Meth. Mol. Bio. 24:307-31, 1994. Altschul et al., J. Mol. Biol. 215:403-10, 1990に記載されており、配列のアライメント方法および相同性の計算の詳細な考察が提示されている。配列分析プログラムblastp、blastn、blastx、tblastnおよびtblastxと関連して使用するための、NCBIベーシックローカルアライメントサーチツール(BLAST)(Altschul et al., J. Mol. Biol. 215:403-10, 1990)は、国立バイオテクノロジー情報センター(NCBI, National Library of Medicine、Building 38A, Room 8N805, Bethesda, Md. 20894)などの数々のソースから、さらにインターネットで入手可能である。追加の情報は、NCBIのウェブサイトで見出すことができる。

30

40

#### 【0035】

[0061] 用語「配列同一性」または「パーセント同一性」は、本明細書で使用される場合、そのコンテキスト中、または2つ以上の核酸もしくはタンパク質配列において、2つ以上の配列または部分配列が同じであること、または2つ以上の配列または部分配列が、同じであるアミノ酸残基またはヌクレオチドの特定のパーセンテージを有することを指す。例えば、配列は、比較ウィンドウにわたり、または配列比較アルゴリズムを使用して、もしくは手作業でのアライメントと目視検査によって測定した場合に指定された領域にわたり、比較して最大限一致するように並べられた場合、特定の領域にわたり、参照配列に対して、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

50

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%、少なくとも99%のパーセントの同一性、またはそれより高い同一性を有していてもよい。例えば、同一性のパーセントは、配列中の同一なヌクレオチド（またはアミノ酸残基）の数を、全てのギャップの長さを引いた総ヌクレオチド（またはアミノ酸残基）の長さで割った比率を計算することによって決定される。

#### 【0036】

[0062] 便宜上、2つの配列間の同一性の程度は、当業界において公知のコンピュータプログラムおよび数学的なアルゴリズムを使用して確認することができる。このようなパーセント配列同一性を計算するアルゴリズムは、一般的に、比較領域にわたり配列ギャップおよびミスマッチを考慮に入れる。配列を比較し並べるプログラム、例えば、C l u s t a l W (Thompson et al., (1994) Nucleic Acids Res., 22: 4673-4680)、A l l i g n (Myers et al., (1988) CABIOS, 4: 11-17)、F A S T A (Pearson et al., (1988) PNAS, 85:2444-2448; Pearson (1990), Methods Enzymol., 183: 63-98) および g a p p e d B L A S T (Altschul et al., (1997) Nucleic Acids Res., 25: 389-3402) は、この目的のために有用である。配列分析プログラム B L A S T P、B L A S T N、B L A S T X、T B L A S T N および T B L A S T X と関連して使用するための、B L A S T または B L A S T 2 . 0 (Altschul et al., J. Mol. Biol. 215:403-10, 1990) は、国立バイオテクノロジー情報センター (NCBI) などの数々のソースから、さらにインターネットで入手可能である。追加の情報は、N C B I のウェブサイトで見出すことができる。

#### 【0037】

[0063] ある特定の実施形態において、配列アライメントおよびパーセント同一性の計算は、その標準的なデフォルトパラメーターを使用する B L A S T プログラムを使用して決定することができる。ヌクレオチド配列のアライメントおよび配列同一性の計算の場合、B L A S T N プログラムは、そのデフォルトパラメーター（ギャップ開始ペナルティー = 5、ギャップ伸長ペナルティー = 2、核酸マッチ = 1、核酸ミスマッチ = - 3、期待値 = 10.0、文字サイズ = 11）と共に使用される。ポリペプチド配列のアライメントおよび配列同一性の計算の場合、B L A S T P プログラムは、そのデフォルトパラメーター（ギャップ開始 = 11、ギャップ伸長ペナルティー = 2；核酸マッチ = 1；核酸ミスマッチ = - 3、期待値 = 10.0；文字サイズ = 11；行列 Blosum 62）と共に使用される。代替として、以下のプログラムおよびパラメーターが使用される：C l o n e M a n a g e r S u i t e、バージョン5の A l i g n P l u s ソフトウェア (Sci-Ed Software)；DNA 比較：グローバル比較、標準的なリニアスコア行列、ミスマッチペナルティー = 2、開始ギャップペナルティー = 4、伸長ギャップペナルティー = 1。アミノ酸比較：グローバル比較、B L O S U M 62 スコア行列。

#### 【0038】

[0064] 用語「マーカーレス」は、本明細書で使用される場合、選択可能マーカーの組込みを伴わない宿主細胞のゲノム内の外因性ランディングパッドへのドナーDNA（例えば、外因性ドナー核酸）の組込みを指す。一部の実施形態において、この用語はまた、宿主細胞のゲノムへの選択可能マーカーの組込みに頼る選択スキームを利用せずに、このような宿主細胞を回収することも指す。例えば、ある特定の実施形態において、エピソームまたは染色体外の選択マーカーが、ゲノム標的部位を切断することが可能なヌクレアーゼをコードするプラスミドを含む細胞に関して選択するのに利用される場合がある。このような使用は、選択可能マーカーが宿主細胞のゲノムに組み込まれない限り、「マーカーレス」とみなされることになる。

#### 【0039】

[0065] 用語「ポリヌクレオチド」は、本明細書で使用される場合、当業者によって理解される通り、ヌクレオチド単位で構成されるポリマーを指す。好ましいヌクレオチド単位としては、これらに限定されないが、アデニン (A)、グアニン (G)、シトシン (C)、チミン (T)、およびウラシル (U) を含むものが挙げられる。有用な改変されるヌクレオ

チド単位としては、これらに限定されないが、4 - アセチルシチジン、5 - (カルボキシヒドロキシルメチル)ウリジン、2 - O - メチルシチジン、5 - カルボキシメチルアミノメチル - 2 - チオウリジン、5 - カルボキシメチルアミノ - メチルウリジン、ジヒドロウリジン、2 - O - メチルプソイドウリジン、2 - O - メチルグアノシン、イノシン、N 6 - イソペンチルアデノシン、1 - メチルアデノシン、1 - メチルプソイドウリジン、1 - メチルグアノシン、1 - メチルイノシン、2 , 2 - ジメチルグアノシン、2 - メチルアデノシン、2 - メチルグアノシン、3 - メチルシチジン、5 - メチルシチジン、N 6 - メチルアデノシン、7 - メチルグアノシン、5 - メチルアミノメチルウリジン、5 - メトキシアミノメチル - 2 - チオウリジン、5 - メトキシウリジン、5 - メトキシカルボニルメチル - 2 - チオウリジン、5 - メトキシカルボニルメチルウリジン、2 - メチルチオ - N 6 - イソペンチルアデノシン、ウリジン - 5 - オキシ酢酸 - メチルエステル、ウリジン - 5 - オキシ酢酸、ワイプトキソシン、ワイプトシン、プソイドウリジン、クエオシン、2 - チオシチジン、5 - メチル - 2 - チオウリジン、2 - チオウリジン、4 - チオウリジン、5 - メチルウリジン、2 - O - メチル - 5 - メチルウリジン、2 - O - メチルウリジンなどを含むものが挙げられる。ポリヌクレオチドとしては、天然に存在する核酸、例えばデオキシリボ核酸(「DNA」)およびリボ核酸(「RNA」)、加えて核酸類似体が挙げられる。核酸類似体としては、天然に存在しない塩基を包含するもの、天然に存在するホスホジエステル結合以外の他のヌクレオチドとの連結で係合するヌクレオチド、またはホスホジエステル結合以外の連結を介して結合する塩基を包含するヌクレオチドが挙げられる。したがって、ヌクレオチド類似体としては、これらに限定されないが、例えば、ホスホリチオエート、ホスホリジチオエート、ホスホトリエステル、ホスホリアミデート、ボラノホスフェート、メチルホスホネート、キラル - メチルホスホネート、2 - O - メチルリボヌクレオチド、ペプチド核酸(PNA)などが挙げられる。

#### 【0040】

[0066] 本明細書において、ポリヌクレオチド配列を記載するのに従来の表記法が使用される。一本鎖ポリヌクレオチド配列の左側の末端は5'末端であり、二本鎖ポリヌクレオチド配列の左側方向は5'方向と称される。

#### 【0041】

[0067] 用語「DNAセグメント」は、その代わりに以下の例では「ビット」とも称されるが、本明細書で使用される場合、DNAのあらゆる単離されたまたは分離可能な分子を指す。有用な例としては、これらに限定されないが、タンパク質コード配列、レポーター遺伝子、蛍光マーカーコード配列、プロモーター、エンハンサー、ターミネーター、イントロン、エクソン、ポリAテール、多重クローニング部位、核局在化シグナル、mRNA安定化シグナル、選択可能マーカー、組込み遺伝子座、エピトープタグコード配列、分解シグナル、スパーサー配列、もしくはバーコード配列、または他のあらゆる天然に存在するもしくは合成DNA分子が挙げられる。一部の実施形態において、DNAセグメントは、天然起源であってもよい。代替として、DNAセグメントは、インビトロで生産された、完全に合成起源であってもよい。さらに、DNAセグメントは、単離された天然に存在するDNA分子の任意の組合せ、または単離された天然に存在するDNA分子と合成DNA分子との任意の組合せを含んでもよい。例えば、DNAセグメントは、タンパク質コード配列に作動可能に連結した異種プロモーター、ポリAテールに連結されたタンパク質コード配列、フレーム内でエピトープタグコード配列と連結されたタンパク質コード配列などを含んでもよい。

#### 【0042】

[0068] 「プライマー」は、ポリヌクレオチドテンプレート配列、例えばプライマー結合セグメントに特異的にハイブリダイズすることが可能であり、合成に好適な条件下で、すなわちヌクレオチドおよび合成反応を触媒する物質(例えば、DNAポリメラーゼ)の存在下で、相補的ポリヌクレオチドを合成するための始点を提供することが可能なポリヌクレオチド配列を指す。プライマーは、ポリヌクレオチドテンプレート配列に相補的であるが、ポリヌクレオチドテンプレート配列の正確な相補物でなくてもよい。例えば、プライ



マーは、ポリヌクレオチドテンプレート配列の相補物に、少なくとも約 80、85、90、95、96、97、98、または 99% 同一であってもよい。プライマーは、可変長を有していてもよいが、一般的には少なくとも 15 塩基である。一部の実施形態において、プライマーは、15 塩基 ~ 35 塩基の長さである。一部の実施形態において、プライマーは、35 塩基より長い長さである。他の実施形態において、プライマーは、少なくとも 50 の融解温度 ( $T_m$ )、すなわち DNA 二重鎖の半分が解離して一本鎖になると予想される温度を有する。他の実施形態において、プライマーは、約 50 ~ 70 の  $T_m$  を有する。さらに他の実施形態において、プライマーは、ポリヌクレオチドテンプレート配列へのハイブリダイゼーションの効率に影響を与えないように、明らかな DNA または RNA 2 次構造を形成しない。

10

【0043】

[0069] 用語「プライマー結合セグメント」は、本明細書で使用される場合、合成に好適な条件下で相補的ポリヌクレオチドを合成するための始点が提供されるようにプライマーに結合するポリヌクレオチド配列である。

【0044】

[0070] 用語「リンカー配列」は、本明細書で使用される場合、インピボでの相同組換えを介して別のコンポーネントポリヌクレオチド中のリンカー配列と相同組換えが可能な、コンポーネントポリヌクレオチド中のポリヌクレオチド配列を指す。ある特定の実施形態において、コンポーネントポリヌクレオチドのリンカー配列は、外因性ランディングパッドのランディングパッドリンカー配列と相同組換えが可能である。

20

【0045】

[0071] 用語「同時に」は、本明細書で使用される場合、複数の組込みに関して使用される場合、宿主細胞が、ヌクレアーゼ、例えばヌクレアーゼをコードするプラスミドと、宿主細胞のゲノムに組み込もうとする 1 つより多くのドナー DNA とで共に形質転換される時点から始まり、形質転換宿主細胞またはそのクローン集団が、ランディングパッドにおけるドナー DNA 組込みの成功に関してスクリーニングされる時点で終わる期間を包含する。一部の実施形態において、「同時に」に包含される期間は、少なくともヌクレアーゼが宿主細胞の染色体内のその標的配列に結合してそれを切断するのに必要な時間である。一部の実施形態において、「同時に」に包含される期間は、宿主細胞が、ヌクレアーゼ、例えばヌクレアーゼをコードするプラスミドと 1 つより多くのドナー DNA とで共に形質転換される時点から始まる、少なくとも 6、12、24、36、48、60、72、96 または 96 時間より長い時間である。

30

【0046】

[0072] 用語「デグロン配列」は、デグロンタンパク質にフレーム内で融合した別のタンパク質に、融合したタンパク質の分解速度を変化させることによって不安定さを付与するタンパク質をコードする核酸を指す。

【0047】

[0073] 用語「ターミネーター」は、転写単位（例えば遺伝子）の終わりを規定し、転写機構から新たに合成された RNA を放出するプロセスを開始させる転写ターミネーター配列を指す。

40

【0048】

[0074] 用語「インシュレーター配列」は、外因性ランディングパッド中に挿入されたヌクレオチド配列であって、宿主細胞のゲノム中でランディングパッドの上流に配置された別のプロモーターからの予期せぬ「リードスルー」転写からランディングパッドを隔離する配列を指す。インシュレーター配列の例は、転写ターミネーター、または上流プロモーターからの予期せぬリードスルー転写を空間配置的に妨げる構造を形成する核酸である。

【0049】

[0075] 用語「ニュートラルな遺伝子座」は、公知の機能または遺伝子転写がない、またはそこへの外因性ランディングパッドの組込みおよび/または外因性ドナー核酸の組込みが、対照細胞と比較して細胞の生存率または機能に有意に影響を与えない宿主細胞のゲノ

50

ムにおける配置を指す。用語「ニュートラルな遺伝子座」は、複数の外因性ランディングパッドのコンテキストにおいて、タンパク質をコードする同じDNAコンストラクトのそこへの組込みが類似したタンパク質発現レベルをもたらすゲノム遺伝子座を指す可能性がある。

【0050】

[0076] 用語「遺伝子間領域」は、公知の機能がない遺伝子間に配置されたDNA配列のストレッチを指し、非コードDNAのサブセットである。

【0051】

5.2 宿主細胞のゲノム中に外因性ランディングパッドを含む宿主細胞の生成

[0077] 宿主細胞のゲノム中に組み込まれた1つまたは複数の外因性ランディングパッドを含むように改変された宿主細胞、およびこのような宿主細胞を生成する方法が本明細書で提供される。外因性ランディングパッドは、本明細書で使用される場合、相同組換えを介した目的の外因性ドナー核酸の挿入をさらに容易にするために、安定して宿主細胞のゲノムに組み込まれる組換えによって生成した標的部位を指す。外因性ランディングパッドは、適合性を有する相同配列を含む外因性ドナー核酸の相同組換えを容易にするために、その5'および3'末端にランディングパッド相同性領域を含む。ある特定の実施形態において、その5'および3'末端におけるランディングパッドの相同性領域（ランディングパッド相同配列とも称される）は、宿主細胞のゲノムにとって外因性であり、宿主細胞のゲノムにおいて提示されず、外因性ドナー核酸の組込みは、宿主細胞における相同組換えのためにいかなる内因性ゲノム配列にも頼らない。したがって、ある特定の実施形態において、ランディングパッドにおける外因性ドナー核酸の組込み事象は、外因性ランディングパッドが配置されているところの周辺の内因性ゲノム配列から独立している。

【0052】

[0078] ある特定の実施形態において、ランディングパッドは、ランディングパッド内でDNA二本鎖破断（DSB）を生成することが可能な部位特異的ヌクレアーゼによって認識可能であり、それと結合可能なヌクレアーゼ標的配列（NTS）をさらに含む。ランディングパッド内のDSBは、外因性ドナー核酸の組込み効率を増加させ、DSBを修復できる形質転換宿主細胞が富化される。特定の実施形態において、宿主細胞のゲノム中に組み込まれたランディングパッドは、上流ランディングパッド相同配列（ULP）、下流ランディングパッド相同配列（DLP）、および（ULP）と（DLP）との間に位置するヌクレアーゼ標的配列（NTS）を含む。

【0053】

[0079] 図1Aは、宿主細胞中のALG1遺伝子座下流の遺伝子間領域において組み込まれた1つのランディングパッドを含む宿主細胞の例示的な実施形態を例示する。この実施形態において、上流ランディングパッド相同配列（ULP）は、「UL」と表記され、下流ランディングパッド相同配列（DLP）は、「DL」と表記される。ヌクレアーゼ標的配列（NTS）は、「X」と表記される。ある特定の実施形態において、ランディングパッドは、予め決定された遺伝子座、すなわち宿主細胞中で下流の遺伝子間領域ALG1遺伝子座中に組換え操作により作製される。ある特定の実施形態において、（ULP）および（DLP）の少なくとも1つは、宿主細胞のゲノム中に存在しない、組換えによって生成した合成配列である。ある特定の実施形態において、（ULP）および（DLP）の少なくとも1つは、宿主細胞のゲノム中に内因的に存在するが、それらの天然の遺伝子座の外側に（すなわち、ランディングパッド中の非ネイティブの遺伝子座に）配置されている配列であってもよい。

【0054】

[0080] 図1Bは、外因性ドナー核酸のライブラリーを例示しており、各外因性ドナー核酸（ES）は、上流ライブラリー配列（UL）と下流ライブラリー配列（DL）との間に位置する目的の核酸（すなわち、曲がった矢印として示されるプロモーターに作動可能に連結したターミネーター配列を有するオープンリーディングフレーム）を含む。図1Bに示される例示的な実施形態において、各プロモーターは、リンカー配列「A」を介してORF

に作動可能に連結している。しかしながら、リンカー配列 A は、ある特定の実施形態において省略されてもよく、外因性ドナー核酸は、プロモーターがリンカー配列を介さずオープンリーディングフレーム (ORF) に直接作動可能に連結するように操作されてもよい。

【 0 0 5 5 】

[0081] 図 1 A で例示された実施形態において、ランディングパッドの上流ランディングパッド相同配列 (ULP) は、外因性ドナー核酸の (UL) と同じヌクレオチド配列を含み、ランディングパッドの下流ランディングパッド相同配列 (DLP) は、外因性ドナー核酸の (DL) と同じヌクレオチド配列を含む。それゆえに、ライブラリー中のいずれの外因性ドナー核酸の (UL) も、ランディングパッドの (UL) と相同組換えが可能であり、ライブラリー中のいずれの外因性ドナー核酸 (DL) も、図 1 A に示されるランディングパッドの (DL) と相同組換えが可能である。ランディングパッド中のヌクレアーゼ標的配列は、ヌクレアーゼ X - カッター (例えば、C p h I であり、図 1 A では一丁のはさみとして示される) によって認識可能であり、それと結合可能である。図 1 C で示されるように、ランディングパッド中に一旦組み込まれたら、目的の核酸 (例えば、プロモーターに作動可能に連結した O R F 1) の端部に、外因性ランディングパッド相同配列 (すなわち、UL および DL) が配置され、これらは、宿主細胞のゲノム中の内因性ゲノム配列 (例えば、i A L G 1) に入れ子になっている (nested within)。しかしながら、この実施形態において、内因性ゲノム配列は、相同組換えに使用されない。

10

【 0 0 5 6 】

[0082] 図 1 D は、外因性ランディングパッドの別の実施形態を例示する。i A L G 1 遺伝子座において組み込まれた外因性ランディングパッドは、図 1 A に示されるランディングパッドに実質的に類似しているが、ランディングパッドは、ランディングパッドの 5' 領域に「UL」と表記される上流ランディングパッド配列に隣接して、さらにランディングパッドの 3' 領域に「DL」と表記される下流ランディングパッド配列に隣接してインシュレーター配列をさらに含む。図 1 D は、ランディングパッドの 5' および 3' 領域の両方に存在するインシュレーター配列を例示するが、ある特定の実施形態において、ランディングパッドは、ランディングパッドの 5' 領域にのみ、3' 領域にのみインシュレーター配列を有するか、または 5' 領域にも 3' 領域にもインシュレーター配列を有さない。ある特定の実施形態において、インシュレーター配列は、転写ターミネーター配列である。

20

【 0 0 5 7 】

[0083] 図 1 E は、ランディングパッドに組み込むことができる 3 つの異なるタイプの外因性ドナー核酸を例示する。図 1 E の一番上は、単一の DNA パーツで作製される外因性ドナー核酸を例示する。外因性ドナー核酸は、上流ライブラリー配列 (UL) と下流ライブラリー配列 (DL) との間に位置する目的の核酸を含む。目的の核酸は、プロモーター (曲がった矢印として示される) とターミネーター (「term」) との間に位置するオープンリーディングフレーム (「ORF」) である。プロモーターおよび O R F は、リンカー配列「A」を介して連結される。

30

【 0 0 5 8 】

[0084] 図 1 E の中央は、インピボで組み立てられ、ランディングパッドへの組込みを起こすことができる 2 つの別個のコンポーネントポリヌクレオチドとしての外因性ドナー核酸を例示する。左側に示される第 1 のコンポーネントポリヌクレオチドは、目的の核酸としてプロモーター配列を有し、右側に示される最後のコンポーネントポリヌクレオチドは、目的の核酸として O R F を有する。左側のコンポーネントポリヌクレオチドおよび右側のコンポーネントポリヌクレオチドが、インピボで一致するリンカー配列「A」を介して組み立てられる場合、それらは、図 1 D に示されるランディングパッドへの組込みを起こすことが可能である。

40

【 0 0 5 9 】

[0085] 図 1 E の一番下は、インピボで組み立てられ、組み立てられたパーツとしてランディングパッドへの組込みを起こすことができる、3 つの別個のコンポーネントポリヌクレオチドの外因性ドナー核酸を例示する。左側に示される第 1 のコンポーネントポリヌク

50

レオチドは、目的の核酸としてプロモーター配列を有し、中央に示される中間コンポーネントポリヌクレオチドは、目的の核酸としてORFを有し、右側に示される最後のコンポーネントポリヌクレオチドは、目的の核酸としてターミネーターを有する。第1のコンポーネントポリヌクレオチドおよび中間コンポーネントポリヌクレオチドの各々は、整合リンカー配列「A」を含み、中間コンポーネントポリヌクレオチドおよび最後のコンポーネントポリヌクレオチドの各々は、整合リンカー配列「B」を含む。コンポーネントポリヌクレオチドが、インピボでそれら各々の整合リンカー配列「A」および「B」を介して組み立てられる場合、それらは、図1Dに示されるランディングパッドへの組込みを起こすことが可能である。

#### 【0060】

[0086] 図1Eは、ランディングパッドに核酸を組み込むための、単一パーツの外因性ドナー核酸、2パーツのコンポーネントポリヌクレオチド、または3パーツのコンポーネントポリヌクレオチドの使用を例示するが、ランディングパッドへの組込みに利用できるコンポーネントポリヌクレオチドの数は、3より多くてもよい。以下のセクション5.3.3で、ゲノム組込みへのコンポーネントポリヌクレオチドの使用をさらに説明する。

#### 【0061】

[0087] 図1Fは、図1Dに示されるコンビナトリアル的に組み合わせてランディングパッドに組み込むことができる3パーツのコンポーネントポリヌクレオチドを例示する。図1Fにおいて、第1のコンポーネントポリヌクレオチドのプロモーターライブラリーが左側に示され、これらの各々は、上流ライブラリー配列(UL)と第1のリンカー配列(A)との間に位置する異なるプロモーター配列を含む。図1Fの中央に、中間コンポーネントポリヌクレオチドのORFライブラリーが示され、これらの各々は、上流中間リンカー配列「A」と下流中間リンカー配列「B」との間に位置する異なるORFを含む。図1Fの右側に、最後のコンポーネントポリヌクレオチドのターミネーターライブラリーが示され、これらの各々は、最後のリンカー配列「B」と下流ライブラリー配列(DL)との間に位置する異なるターミネーターを含む。各ライブラリーからの各コンポーネントポリヌクレオチドは、他のライブラリー中の他のコンポーネントポリヌクレオチドとコンビナトリアル的に組み合わせることができる。それゆえに、図1Fに示される9つのコンポーネントポリヌクレオチドをコンビナトリアル的に組み合わせて、単一のランディングパッドにおいて組み込まれたコンポーネントポリヌクレオチドの27個の異なる組合せを生成することができ、これらは、27種の異なる表現型を有する宿主細胞の集団を生成する可能性がある。

#### 【0062】

[0088] 特定の実施形態において、宿主細胞は、宿主細胞のゲノム中に組み込まれた2つ以上の外因性ランディングパッドを含む。ある特定の実施形態において、2つ以上の外因性ランディングパッドは、それらの上流ランディングパッド相同配列(ULP)が互いに同一または実質的に同一であり、それらの下流ランディングパッド相同配列(DLP)が互いに同一または実質的に同一であるという点で標準化される。標準化されたランディングパッド相同配列は、外因性ドナー核酸が、その5'および3'領域に適合性を有する相同性領域を含む限り、あらゆる外因性ドナー核酸のあらゆるランディングパッドへの宿主細胞媒介相同組換えを容易にする。したがって、ある特定の実施形態において、ライブラリー中の全ての外因性ドナー核酸は、その5'末端に、ランディングパッドのいずれかで任意の(ULP)と相同組換えできる標準化された上流ライブラリー配列(UL)、およびその3'末端に、ランディングパッドのいずれかで任意の(DLP)と相同組換えできる標準化された下流ライブラリー配列(DL)を含む。

#### 【0063】

[0089] ある特定の実施形態において、ランディングパッドは、宿主細胞のゲノムの「ニュートラルな」遺伝子座であることが予め決定されているゲノム遺伝子座に組み込まれている。ある特定の実施形態において、ランディングパッド組込みのためのゲノム遺伝子座は、RNAシーケンシングによってゲノム遺伝子座の周りの領域に遺伝子転写が見出され

10

20

30

40

50

ない場合、ニュートラルな遺伝子座とみなされる。ある特定の実施形態において、RNA配列によってゲノム遺伝子座の周りの領域に遺伝子転写が見出されない場合、さらに、ゲノム遺伝子座が、他の近縁種（例えば、サッカロマイセス・セレビシエ（*Saccharomyces cerevisiae*）およびサッカロマイセス・パラドクス（*Saccharomyces paradoxus*））において、それらのゲノム配列を並べたときに類似の位置を有するゲノム遺伝子座に相同な配列を含まない場合、ランディングパッド組込みのためのゲノム遺伝子座は、ニュートラルな遺伝子座とみなされる。近縁種のゲノム配列は、あらゆる好適な方法を使用して、例えば視覚的に、または例えばBLASTn、genome.ucsc.eduからのUCSCゲノムブラウザなどのアルゴリズムを使用して比較することができる。一部の実施形態において、1つまたは複数のランディングパッドは、公知の機能がない宿主細胞のゲノム中の遺伝子間領域において組み込まれている。一部の実施形態において、1つまたは複数のランディングパッドは、宿主細胞のゲノムの非コード領域において組み込まれている。一部の実施形態において、1つまたは複数のランディングパッドは、宿主細胞のゲノムのコード領域において組み込まれるかまたはそれに隣接している。

#### 【0064】

[0090] 一部の実施形態において、配列分析によってニュートラルであると決定されたこれらのゲノム遺伝子座はさらに、宿主細胞の機能に基づき検証される。例えば、標的分子産生、バイオマス収量、細胞増殖速度などのパラメーターに対するランディングパッド組込みの作用を測定して、ランディングパッドの組込みがない対照宿主細胞と比較することができる。一部の実施形態において、1つのランディングパッド中に組み込まれたレポーター遺伝子の発現を、宿主細胞のゲノム中の他のランディングパッド中に組み込まれた場合の発現と比較して、異なるランディングパッドにおける組込み効率またはレポーター遺伝子の発現のあらゆる変動を決定することができる。

#### 【0065】

[0091] 用語「ニュートラルな遺伝子座」は、複数の外因性ランディングパッドのコンテキストにおいて、そこにタンパク質をコードする同じDNAコンストラクトが組み込まれると、タンパク質発現のレベルの類似が起こるゲノム遺伝子座を指す場合がある。例えば、複数の外因性ランディングパッドにおいて組み込まれたレポーター遺伝子の発現レベルは、外因性ランディングパッドがニュートラルな遺伝子座において組み込まれる場合、50%未満、40%未満、30%未満、20%未満、10%未満、または5%未満変化する可能性がある。それゆえに、ある特定の実施形態において、ニュートラルな遺伝子座での外因性ランディングパッドの組込みは、外因性ランディングパッドにおいて組み込まれた外因性核酸の発現のためのあらゆるゲノム配置のコンテキストを除去する。

#### 【0066】

[0092] 図2Aは、宿主細胞のゲノム中に組み込まれた3つの標準化された外因性ランディングパッドを含む宿主細胞の例示的な実施形態を例示する。図2Aに示される例示的な実施形態において、3つの標準化されたランディングパッドは、酵母細胞ゲノム中の遺伝子座ALG1、MGA1、およびYCT1下流（すなわち、それぞれiALG1、iMGA1、およびiYCT1）の遺伝子間領域中に組み込まれている。これらの遺伝子間のゲノム遺伝子座は、ニュートラルな遺伝子座として検証されている。図2Aで例示された実施形態において、各ランディングパッドは、ランディングパッドの上流ランディングパッド相同配列（ULP）と下流ランディングパッド相同配列（DLP）との間に位置する、ヌクレアーゼ「X-カッター」（例えば、CphI）によって認識可能であり、それと結合可能なヌクレアーゼ標的配列（NTS）を含む。図2Aにおいて、3つのランディングパッドの（ULP）は、外因性ドナー核酸のライブラリーの上流ライブラリー配列（UL）と同じ配列を含み、3つのランディングパッドの（DLP）は、外因性ドナー核酸のライブラリーの下流ライブラリー配列（DL）と同じ配列を含む。図2A～2Dで例示された実施形態において、3つのランディングパッドの各々が、ヌクレアーゼX-カッター（例えば、CphI）によって結合できる同じヌクレアーゼ標的配列を含むことから、これらのランディングパッドは、ヌクレアーゼ標的配列を特異的に認識し、ランディングパッド内の部位を切断する単

10

20

30

40

50

一のヌクレアーゼだけを使用して、複数の外因性ドナー核酸の同時の組み込みを適合させることができる。

#### 【0067】

[0093] あらゆる好適な数のランディングパッドを、宿主細胞のゲノムに組み込まれていてもよい。ある特定の実施形態において、宿主細胞のゲノムに組み込まれたランディングパッドの数は、1～500である。ある特定の実施形態において、宿主細胞のゲノム中に組み込まれたランディングパッドの数は、2～250である。ある特定の実施形態において、宿主細胞のゲノム中に組み込まれたランディングパッドの数は、3～200である。ある特定の実施形態において、宿主細胞のゲノム中に組み込まれたランディングパッドの数は、3～50である。ある特定の実施形態において、宿主細胞のゲノム中に組み込まれたランディングパッドの数は、3～20である。ある特定の実施形態において、宿主細胞のゲノム中に組み込まれたランディングパッドの数は、3～10である。ある特定の実施形態において、宿主細胞のゲノム中に組み込まれたランディングパッドの数は、1、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、60、70、80、90、100、200、300、400、500、またはこれらの数の間のあらゆる数である。

10

#### 【0068】

[0094] 図2Aは、3つの別個のゲノム遺伝子座中の3つのランディングパッドを例示するが、一部の実施形態において、2つ以上のランディングパッドは、単一のゲノム遺伝子座中に組み込まれていてもよい。例えば、タンデム式の2つのランディングパッドは、ALG1遺伝子座から下流の遺伝子間領域に組み込まれていてもよい。このような実施形態において、相同組換えを介してランディングパッドをループ化する代わりに、ランディングパッドを保持する宿主細胞の選択圧力を付加するために、必須の遺伝子を2つのランディングパッド間に挿入してもよい。代替として、組換え事象により生存不可能な細胞が生じるように、染色体転移に適合性を有する染色体位置を有する遺伝子座を選択することができる。

20

#### 【0069】

[0095] 各ランディングパッド配列は、当業界において公知のあらゆる技術を使用して生成することができる。一部の実施形態において、ランディングパッド配列の全体を、化学合成および/または組換え分子生物学技術を使用して新たに合成することができる。例えば、KosuriおよびChurch、Nature Methods 11:499～507(2014)を参照されたい。ある特定の実施形態において、各ランディングパッドのパーツ(例えば、ULP、DLP、およびNTS)を別個のパーツとして別々に合成して、次いで一緒にライゲートしてもよい。代替として、ランディングパッド配列は、商業的な核酸合成サービスから得ることができる。これらとしては、例えば、Twist Bioscience(San Francisco, CA)、Biomatik(Wilmington, DE)、Genescript(Piscataway, NJ)、およびwww.introgen.com経由で利用可能なGeneArt遺伝子合成サービスが挙げられる。

30

#### 【0070】

[0096] ランディングパッドは、当業界において公知のあらゆる好適な方法を使用して宿主細胞のゲノムに組み込むことができる。例えば、選択可能マーカースを使用する相同組換えによる従来の遺伝子標的化方法を使用して、ランディングパッドを宿主細胞のゲノムに導入し組み込むことができる。一部の実施形態において、それに続き、選択可能マーカース遺伝子を宿主細胞のゲノムから切り出してもよい。例えば米国特許番号第7,919,605号及び第9,018,364号を参照。一部の実施形態において、デザイナーヌクレアーゼ、例えばCRISPR-Cas系などを使用して、望ましい特異的な標的ゲノム遺伝子座への1つまたは複数のランディングパッドの組み込みを容易にすることができる。例えば、米国特許公開公報第2015/0184199号; Horwitz et al., Cell Systems 1, 88-96, Jul 29, 2015; Cong et al., 2013, Science 339: 819-823; Jao et al., 2013, Proc. Natl. Acad. Sci USA 110, 13904-13909; Wang et al. 2013, Cell 15

40

50

3, 910-918; Jacobs et al., 2014, Nat. Commun. 5, 5344を参照されたい。

【0071】

5.2.1. 上流および下流ランディングパッド相同配列

[0097] ある特定の実施形態において、宿主細胞のゲノム中に操作により作製された外因性ランディングパッドの各々は、その5'領域に上流ランディングパッド相同配列(ULP)およびその3'領域に下流ランディングパッド相同配列(DLP)を含む。これらの配列は、宿主細胞媒介組換えおよび外因性ドナー核酸の組み込みのための相同性領域を提供する。特定の実施形態において、ランディングパッド相同配列は、それらが組み込まれるゲノム遺伝子座にとって「外因性」である。言い換えると、これらのランディングパッド相同配列は、それらが組み込まれるゲノム遺伝子座の内因性配列の一部ではない。他の実施形態において、ランディングパッド相同配列は、宿主細胞のゲノム中に存在していてもよいが、それらは、ランディングパッド中のそれらの非ネイティブの遺伝子座に配置される。

10

【0072】

[0098] ある特定の実施形態において、ランディングパッド相同配列は、いずれのゲノムDNAとも配列相同性を共有しない。ある特定の実施形態において、(ULP)および/または(DLP)のヌクレオチド配列は、合成である。ある特定の実施形態において、(ULP)および/または(DLP)のヌクレオチド配列は、コンピュータアルゴリズムによってランダムに生成する。好適なランディングパッド相同配列は、望ましくない特徴を有するランダムに生成した配列、例えば、宿主細胞のゲノム中の、タンDEMおよび/または逆方向反復配列、DNAの2次構造、ならびに公知の生物学的配列へのあらゆる有意な相同性を有する配列を除去することによって選択される。ある特定の実施形態において、ランダムに生成したヌクレオチドをさらにフィルタリングして、約30%~約70%、典型的には約40%~60%、典型的には約45%~約55%、または約50%のGC含量(またはグアニン-シトシン含量)を含むものを選択する。

20

【0073】

[0099] 一部の実施形態において、(ULP)および(DLP)は、いずれのランディングパッドでの外因性核酸(ES)のゲノム組み込みを可能にする十分な長さおよび配列同一性を有するあらゆるヌクレオチド配列を含んでいてもよい。ある特定の実施形態において、(ULP)および(DLP)の各々は、独立して、約20~5,000ヌクレオチドからなる。ある特定の実施形態において、(ULP)および(DLP)の各々は、独立して、約20~2,500ヌクレオチドからなる。ある特定の実施形態において、(ULP)および(DLP)の各々は、独立して、約25~1,000ヌクレオチドからなる。ある特定の実施形態において、(ULP)および(DLP)の各々は、独立して、約50~500ヌクレオチドからなる。ある特定の実施形態において、(ULP)および(DLP)の各々は、独立して、約100~約500ヌクレオチドからなる。ある特定の実施形態において、(ULP)および(DLP)の各々は、独立して、約20、25、50、75、100、150、200、250、300、350、400、450、500、600、700、800、900、1000、1100、1200、1300、1400、1500、1600、1700、1800、1900、2000、2100、2200、2300、2400、2500、2600、2700、2800、2900、3000、3100、3200、3300、3400、3500、3600、3700、3800、3900、4000、4100、4200、4300、4400、4500、4600、4700、4800、4900、5000ヌクレオチドからなるか、またはこれらの数の間のあらゆる数である。一部の実施形態において、(ULP)および(DLP)の各々は、独立して、約500ヌクレオチドからなる。一部の実施形態において、(ULP)および(DLP)の各々は、独立して、約200ヌクレオチドからなる。一部の実施形態において、(ULP)および(DLP)の各々は、独立して、約100ヌクレオチドからなる。

30

40

【0074】

[00100] 宿主細胞のゲノム中に組み込まれた2つ以上のランディングパッドが存在するある特定の実施形態において、ランディングパッド中の全ての(ULP)のヌクレオチド配

50

列は、互いに実質的に同一であり、ランディングパッド中の全ての (DLP) 中のヌクレオチド配列は、互いに実質的に同一である。例えば、ランディングパッド中の全ての (ULP) は、互いに少なくとも 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% または 100% 同一であるヌクレオチド配列を含み、ランディングパッド中の全ての (DLP) は、互いに少なくとも 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% または 100% 同一であるヌクレオチド配列を含む。一部の実施形態において、ランディングパッド中の全ての (ULP) のヌクレオチド配列は、互いに同一であり、ランディングパッド中の全ての (DLP) 中のヌクレオチド配列は、互いに同一である。このような例示的な実施形態は、3つのランディングパッドが「UL」と表記される3つの同一な (ULP) および「DL」と表記される3つの同一な (DLP) を含む図2Aで例示される。宿主細胞のゲノム中に組み込まれた2つ以上のランディングパッドが、実質的に同一な (ULP) および (DLP) を含む場合、これらのランディングパッドは、標準化されたランディングパッドまたは一次ランディングパッドと称される。

10

#### 【0075】

[00101] ある特定の実施形態において、宿主細胞のゲノムは、一次ランディングパッドとは異なる1つまたは複数の代替ランディングパッドを含んでいてもよい。これらの代替ランディングパッドの例示的な実施形態は図5B、6B、7A~7C、および8で例示されており、これらは以下のセクションでさらに詳細に説明される。これらの代替ランディングパッドは、一次ランディングパッドに追加して取り込んでもよいし、またはその代わりに取り込んでもよい。

20

#### 【0076】

##### 5.2.2.ヌクレアーゼ標的配列

[00102] 各ランディングパッド中のヌクレアーゼ標的配列 (NTS) は、ヌクレアーゼによって認識されるヌクレオチド配列を含む。ヌクレアーゼが各ランディングパッド中のヌクレアーゼ標的配列 (NTS) を認識してそれに結合すると、ヌクレアーゼは、ランディングパッド中の (NTS) 内またはそのすぐ近くの部位を切断することができる。ある特定の実施形態において、ヌクレアーゼ標的配列は、ランディングパッド配列の少なくとも1つに隣接して配置される。ある特定の実施形態において、ヌクレアーゼ標的配列 (NTS) は、上流ランディングパッド相同配列 (ULP) と下流ランディングパッド相同配列 (DLP) との間に位置する。ヌクレアーゼがヌクレアーゼ標的配列を有するランディングパッドを含む宿主細胞に導入されると、ヌクレアーゼは、ランディングパッド内の部位で二本鎖切断を引き起こすことができ、それにより切断部位で、またはその近くで相同組換えの頻度が大きく増加する。

30

#### 【0077】

[00103] 一部の実施形態において、複数のランディングパッドは、同じヌクレアーゼ標的配列を含み、それによってヌクレアーゼ標的配列を特異的に認識する単一のヌクレアーゼのみを使用した同時の複数の組込み事象が容易になる。一部の実施形態において、宿主細胞のゲノム中の複数のランディングパッドの一部は、異なるヌクレアーゼによって認識可能であり、それと結合可能である異なるヌクレアーゼ標的配列を有していてもよい。このような実施形態において、異なるヌクレアーゼを同時にまたは逐次的に使用して、宿主細胞のゲノム中の同時のまたは逐次的な外因性ドナー核酸の組込みのために異なるヌクレアーゼ標的配列を切断することができる。

40

#### 【0078】

[00104] 一部の実施形態において、ランディングパッド中のヌクレアーゼ標的配列は、宿主細胞のネイティブのゲノム中に特に提示されないヌクレアーゼ標的配列を含む。一部の実施形態において、ヌクレアーゼによって認識されるヌクレアーゼ標的配列は、ランディングパッド内でのみ宿主細胞のゲノム中に存在し、それによってヌクレアーゼによるあらゆるオフターゲットのゲノム結合および切断が最小化される。ある特定の実施形態において、特異的なヌクレアーゼは、少なくとも12塩基対の長さの、または一部の場

50



いて14～55塩基対の長さのポリヌクレオチド認識部位を有する希少切断エンドヌクレアーゼであってもよい。このようなエンドヌクレアーゼは、エンドヌクレアーゼ活性を有する天然タンパク質由来のもの、例えばホーミングエンドヌクレアーゼ(WO2004/067736)、または様々な核酸結合ポリペプチドをヌクレアーゼコンポーネント、例えばFok-1またはIv-1触媒ドメイン(WO2012138927)に融合することによるもののいずれかであり得る。この点において操作することができる適切な核酸結合ドメインは、例えば、ジンクフィンガードメイン(Kim et al., 1994, Chimeric restriction endonuclease, PNAS, 91:883-887)、キサントモナス属関連の微生物由来のTALエフェクター(WO2011/072246)または内共生性のバクテロイデリア・リゾキシニカ(*Burkholderia rhizoxinica*)に由来するMBBD(モジュラー式の塩基対塩基結合ドメイン)である。加えて、ヌクレアーゼCas9相同体およびRNAアーゼIIを含む系(CRISPR/Cas9)は、細菌性微生物の免疫系から開発されている。この系において、エンドヌクレアーゼタンパク質複合体の特異性は、「ガイドRNA」(gRNA)と呼ばれる特異的な一本鎖RNAによってアドレス指定される。このガイドRNAは、ヌクレアーゼコンポーネントCas9によって切断される核酸標的配列をハイブリダイズする能力を有する(Le Cong et al., 2013, Science, 339(6121): 819-823)。これらのヌクレアーゼによって認識可能であるか、および/またはそれと結合できるヌクレアーゼ標的配列は周知であり、当業界で説明されている。以下のセクション5.6で、好適なヌクレアーゼおよびヌクレアーゼ標的配列の説明をさらに詳細に説明する。

【0079】

#### 5.2.3. 好適な宿主細胞

[00105] ランディングパッドを有する親株を生成するための好適な宿主細胞としては、目的の核酸または「ドナーDNA」が染色体のまたはエピソームの遺伝子座に組み込まれることが望ましいあらゆる細胞が挙げられる。一部の実施形態において、細胞は、相同組換えを起こす能力を有する生物の細胞である。例示的な実施形態のいくつかは酵母(*S. cerevisiae*)で実証されているが、本明細書で提供されるゲノム改変方法は、組換え系が酵母の場合ほどの能力がないとしても、機能的な組換え系を有する全ての生物で実施が可能であると考えられる。機能的な相同組換え系を有する他の細胞または細胞型としては、細菌、例えばバチルス・ズブチリス(*Bacillus subtilis*)および大腸菌(*E. coli*) (これは、RecE RecT組換え能を有する; Muirers et al., EMBO rep. 1: 239-243, 2000); 原生動物(例えば、プラスモジウム属、トキソプラズマ属); 他の酵母(例えば、シゾサッカロマイセス・ポンベ(*Schizosaccharomyces pombe*)); 糸状菌(例えば、アシュビヤ・ゴシッピー(*Ashbya gossypii*)); 植物、例えば蘚類のヒメツリガネゴケ(*Physcomitrella patens*) (Schaefer and Zryd, Plant J. 11: 1195-1206, 1997; および動物細胞、例えば哺乳類細胞およびニワトリDT40細胞(Dieken et al., Nat. Genet. 12:174-182, 1996)が挙げられる。

【0080】

[00106] 一部の実施形態において、宿主細胞は、原核細胞である。一部の実施形態において、宿主細胞は、真核細胞である。一部の実施形態において、細胞は、真菌細胞(例えば、酵母細胞)、細菌細胞、植物細胞、または動物細胞(例えば、ニワトリ細胞)である。一部の実施形態において、宿主細胞は、哺乳類細胞である。一部の実施形態において、宿主細胞は、チャイニーズハムスター卵巣(CHO)細胞、COS-7細胞、マウス線維芽細胞、マウス胚性癌腫細胞、またはマウス胚幹細胞である。一部の実施形態において、宿主細胞は、昆虫細胞である。一部の実施形態において、宿主細胞は、S2細胞、シュナイダー(Schneider)細胞、S12細胞、5B1-4細胞、Tn5細胞、またはSf9細胞である。一部の実施形態において、宿主細胞は、単細胞の真核生物細胞である。

【0081】

[00107] 特定の実施形態において、宿主細胞は、酵母細胞である。有用な酵母宿主細胞としては、微生物受託機関(例えばIFO、ATCCなど)に寄託された酵母細胞が挙げられ、なかでも、アシクロコニディウム属(*Aciculoconidium*)、アンプロシオザイマ属(*Am*

10

20

30

40

50

brosiozyma)、アルスロアスカス属(Arthroascus)、アルキシオザイマ属(Arxiozyma)、アシュビア属(Ashbya)、バブジェビア属(Babjevia)、ベンスिंगトニア属(Bensingtonia)、ボトリオアスカス属(Botryoascus)、ボトリオザイマ属(Botryozyma)、ブレッタノマイセス属(Brettanomyces)、ビュレラ属(Bullera)、ビュレロマイセス属(Bulleromyces)、カンジダ属(Candida)、シテロミセス属(Citeromyces)、クラビスポラ属(Clavispora)、クリプトコックス属(Cryptococcus)、シストフィロバシディウム属(Cystofilobasidium)、デバリオマイセス属(Debaryomyces)、デッケラ属(Dekkara)、ディボダスコプシス属(Dipodascopsis)、ディボダスカス属(Dipodascus)、エニエラ属(Eeniella)、エンドマイコブセラ属(Endomycopsella)、エレマスカス属(Eremascus)、エレモテシウム属(Eremothecium)、エリスロバシディウム属(Erythrobasidium)、フェロマイセス属(Fellomyces)、フィロバシディウム属(Filobasidium)、ガラクトマイセス属(Galactomyces)、ゲオトリクム属(Geotrichum)、ガイラーモンデラ属(Guilliermondella)、ハンセニアスポラ属(Hanseniaspora)、ハンセヌラ属(Hansenula)、ハセガワエア属(Hasegawaea)、ホルターマンニア属(Holtermannia)、ホルモアスカス属(Hormoascus)、ハイフォピキア属(Hyphopichia)、イッサチェンキア属(Issatchenkia)、クロエケラ属(Kloeckera)、クロエケラスポラ属(Kloeckeraspora)、クルイベロマイセス属(Kluyveromyces)、コンドア属(Kondoa)、クライシア属(Kuraishia)、クルツマノマイセス属(Kurtzmanomyces)、ロイコスポリジウム属(Leucosporidium)、リポマイセス属(Lipomyces)、ロデロマイセス属(Lodderomyces)、マラセジア属(Malassezia)、メトシュニコウィア属(Metschnikowia)、ムラキア属(Mrakia)、ミクソザイマ属(Myxozyma)、ナドソニア属(Nadsonia)、ナカザワエア属(Nakazawaea)、ネマトスポラ属(Nematospora)、オガタエア属(Ogataea)、オースポリディウム属(Oosporidium)、パチソレン属(Pachysolen)、ファチコスポラ属(Phachytichospora)、ファフィア属(Phaffia)、ピキア属(Pichia)、ロドスポリディウム属(Rhodosporidium)、ロドトルラ属(Rhodotorula)、サッカロマイセス属(Saccharomyces)、サッカロマイコーデス属(Saccharomycodes)、サッカロマイコプシス属(Saccharomycopsis)、サイトエラ属(Saitoella)、サカグチア属(Sakaguchia)、サターノスポラ属(Saturnospora)、シゾブラストスポリオン属(Schizoblastosporion)、シゾサッカロマイセス属(Schizosaccharomyces)、シュワニオミセス属(Schwanniomycetes)、スポリディオボラス属(Sporidiobolus)、スポロボロマイセス属(Sporobolomyces)、スポロパキデミア属(Sporopachydermia)、ステファノアスカス属(Stephanoascus)、ステリグマトマイセス属(Sterigmatomyces)、ステリグマトスポリディウム属(Sterigmatosporidium)、シンピオタフリナ属(Symbiotaphrina)、シンボディオマイセス属(Sympodiomyces)、シンボディオマイコプシス属(Sympodiomyopsis)、トルラスポラ属(Torulasporea)、トリコスボリエラ属(Trichosporiella)、トリコスボロン属(Trichosporon)、トリゴノプシス属(Trigonopsis)、ツチヤエア属(Tsuchiyaia)、ウデニオマイセス属(Udeniomyces)、ワルトマイセス属(Waltomyces)、ウィカーハミア属(Wickerhamia)、ウィカーハミエラ属(Wickerhamiella)、ウィカーハミエラ属(Williopsis)、ヤマダザイマ属(Yamadazyma)、ヤロウイア属(Yarrowia)、ザイゴアスカス属(Zygoascus)、チゴサッカロマイセス属(Zygosaccharomyces)、ザイゴウィリオプシス属(Zygowilliopsis)、およびザイゴザイマ属(Zygozyma)に属するものが挙げられる。

#### 【0082】

[00108] 一部の実施形態において、酵母宿主細胞は、サッカロマイセス・セレビスエ細胞、ピキア・パストリス(Pichia pastoris)細胞、シゾサッカロマイセス・ボンベ細胞、デッケラ・ブルクセレンシス(Dekkera bruxellensis)細胞、クルイベロマイセス・ラクティス(Kluyveromyces lactis)細胞、アークスラ・アデニニボランス(Arxula adenivorans)細胞、またはハンセヌラ・ポリモルファ(Hansenula polymorpha)(現在、ピキア・アンガスタ(Pichia angusta)として公知)細胞である。特定の実施形態

10

20

30

40

50

において、酵母宿主細胞は、サッカロマイセス・セレビスエ細胞である。一部の実施形態において、酵母宿主細胞は、サッカロマイセス・フラギリス (*Saccharomyces fragilis*) 細胞またはクルイペロマイセス・ラクティス (以前はサッカロマイセス・ラクティスと呼ばれていた) 細胞である。一部の実施形態において、酵母宿主細胞は、カンジダ属に属する細胞であり、例えばカンジダ・リポリチカ (*Candida lipolytica*)、カンジダ・ギリエルモンジイ (*Candida guilliermondii*)、カンジダ・クルセイ (*Candida krusei*)、カンジダ・プソイドトロピカリス (*Candida pseudotropicalis*)、またはカンジダ・ユチリス (*Candida utilis*) である。他の特定の実施形態において、酵母宿主細胞は、クリペロマイセス・マルキシアヌス (*Kluveromyces marxianus*) 細胞である。

#### 【0083】

[00109] 特定の実施形態において、酵母宿主細胞は、パン酵母、CBS 7959細胞、CBS 7960細胞、CBS 7961細胞、CBS 7962細胞、CBS 7963細胞、CBS 7964細胞、IZ-1904細胞、TA細胞、BG-1細胞、CR-1細胞、SA-1細胞、M-26細胞、Y-904細胞、PE-2細胞、PE-5細胞、VR-1細胞、BR-1細胞、BR-2細胞、ME-2細胞、VR-2細胞、MA-3細胞、MA-4細胞、CAT-1細胞、CB-1細胞、NR-1細胞、BT-1細胞およびAL-1細胞からなる群から選択されるサッカロマイセス・セレビスエ細胞である。一部の実施形態において、宿主細胞は、PE-2細胞、CAT-1細胞、VR-1細胞、BG-1細胞、CR-1細胞、およびSA-1細胞からなる群から選択されるサッカロマイセス・セレビスエ細胞である。特定の実施形態において、サッカロマイセス・セレビスエ宿主細胞は、PE-2細胞である。他の特定の実施形態において、サッカロマイセス・セレビスエ宿主細胞は、CAT-1細胞である。他の特定の実施形態において、サッカロマイセス・セレビスエ宿主細胞は、BG-1細胞である。

#### 【0084】

[00110] 一部の実施形態において、酵母宿主細胞は、工業的な発酵、例えばバイオエタノール発酵に好適な細胞である。特定の実施形態において、細胞は、工業的な発酵環境において認められるストレス条件である、高い溶媒濃度、高温、基質利用の拡張、栄養素の制限、浸透圧ストレス、酸性度、亜硫酸塩および細菌汚染、またはそれらの組合せのもとで存在するように調整される。

#### 【0085】

### 5.3 ランディングパッドに組み込まれる外因性ドナー核酸

[0001] 別の態様において、標準化されたランディングパッドがそのゲノム中に操作により作製されたあらゆる宿主細胞での再使用に適した、標準化された相同配列を含む外因性ドナー核酸が本明細書で提供される。ある特定の実施形態において、宿主細胞に導入された外因性ドナー核酸は、それらがランディングパッドにそれ自体で相同組換えできるように、上流および下流ライブラリー配列の両方を含む。ある特定の実施形態において、外因性ドナー核酸は、コンポーネントポリヌクレオチドであり、それらはいずれも、ランディングパッドで相同組換えするための上流および下流ライブラリー配列の両方を有していない。適合性リンカー配列を有するコンポーネントポリヌクレオチドが、インピボで相同組換えを介して組み立てられる場合、上流および下流ライブラリー配列の両方を有する組み立てられたコンポーネントポリヌクレオチドは、宿主細胞のゲノム中のランディングパッドで相同組換えが可能である。一部の実施形態において、各コンポーネントポリヌクレオチドは、上流ライブラリー配列 (UL) とリンカー配列との間に位置する目的の核酸を含む。他の実施形態において、各コンポーネントポリヌクレオチドは、リンカー配列と下流ライブラリー配列 (DL) との間に位置する目的の核酸を含む。ある特定の実施形態において、コンポーネントポリヌクレオチドは、2つのリンカー配列の間に位置する目的の核酸を含む。コンポーネントポリヌクレオチドのこれらの異なるセットが宿主細胞と接触すると、それら各々のリンカー配列は、インピボで相同組換えして、宿主細胞のゲノム中のランディングパッドのいずれか1つに組み立てられたコンポーネントポリヌクレオチドを組み込むことができる。ある特定の実施形態において、外因性ドナー核酸は、直鎖状DNA分

10

20

30

40

50

子として宿主細胞に導入される。ある特定の実施形態において、外因性ドナー核酸は、環状DNA分子として宿主に導入される。

#### 【0086】

[00111] 外因性ドナー核酸は、当業者にとって明白なあらゆる技術によって生成することができる。ある特定の実施形態において、組み込みポリヌクレオチドは、当業界において周知のポリメラーゼ連鎖反応 (PCR) および分子クローニング技術を使用して生成される。例えば、PCR Technology: Principles and Applications for DNA Amplification, ed. HA Erlich, Stockton Press, New York, N.Y. (1989); Sambrook et al., 2001, Molecular Cloning - A Laboratory Manual, 3<sup>rd</sup> edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; PCR Technology: Principles and Applications for DNA Amplification, ed. HA Erlich, Stockton Press, New York, N.Y. (1989); 米国特許番号第 8, 110, 360 号を参照されたい。

#### 【0087】

5.3.1. 外因性ドナー核酸における上流および下流ライブラリー配列

[00112] ある特定の実施形態において、外因性ドナー核酸 (ES) は、上流ライブラリー配列 (UL) と下流ライブラリー配列 (DL) との間に位置する目的の核酸 (D) を含み、(UL) および (DL) は、宿主細胞のゲノム中の 1 つまたは複数のランディングパッドで宿主細胞媒介相同組換えを開始させることが可能である。ある特定の実施形態において、外因性ドナー核酸のライブラリーに関して、それらの上流ライブラリー配列 (UL) は、同一または実質的に同一であり、それらの下流ライブラリー配列 (DL) は、同一または実質的に同一である。相同組換えによって外因性ドナー核酸をゲノムに組み込むために、外因性ドナーポリヌクレオチドは、一般的に、一方の末端に (UL) および他方の末端に (DL) を含む。一部の実施形態において、(UL) は、宿主細胞のゲノム中に組み込まれた 1 つまたは複数のランディングパッドの 5' 領域、例えば、上流ランディングパッド相同配列 (ULP) に相同である。ある特定の実施形態において、(DL) は、1 つまたは複数のランディングパッドの 3' 領域、例えば下流ランディングパッド相同配列 (DLP) に相同である。一部の実施形態において、(UL) は、1 つまたは複数のランディングパッドの 5' 領域、例えば、(ULP) に約 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% または 100% 相同または同一なヌクレオチド配列を含む。一部の実施形態において、(DL) は、1 つまたは複数のランディングパッドの 3' 領域、例えば、(DLP) に約 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% または 100% 相同または同一なヌクレオチド配列を含む。

#### 【0088】

[00113] ある特定の実施形態において、(UL) は、目的の核酸 (D) の 5' に位置する。一部の実施形態において、(UL) は、(D) の 5' 末端に直接隣接して位置する。一部の実施形態において、(UL) は、(D) の 5' の上流に位置する。ある特定の実施形態において、(DL) は、目的の核酸 (D) の 3' に位置する。一部の実施形態において、(DL) は、(D) の 3' 末端に直接隣接して位置する。一部の実施形態において、(DL) は、(D) の 3' の下流に位置する。

#### 【0089】

[00114] 典型的には、ランディングパッドにおける外因性ドナーポリヌクレオチドの組み込みに影響を与える可能性がある特性としては、これらに限定されないが、ゲノム組み込みに使用される上流および下流ライブラリー配列の長さ、外因性ドナー核酸コンストラクトの全体の長さ、およびランディングパッドのゲノム組み込み遺伝子座のヌクレオチド配列または配置が挙げられる。例えば、ライブラリー配列の 1 つの鎖と、宿主細胞のゲノム中の特定のランディングパッド相同配列の 1 つの鎖との間での有効なヘテロ二重鎖形成は、ランディングパッド相同配列に関して、外因性ドナー核酸の上流および下流ライブラリー配列の長さに依存する可能性がある。ゲノム組み込み配列として使用されるライブラリー配列の長さの有効な範囲は、一般的に 20 ~ 5,000 ヌクレオチド、典型的には 25 ~ 5,000 ヌクレオチド、より典型的には 50 ~ 5,000 ヌクレオチドである。ゲノム組込

み配列とゲノム遺伝子座との間の相同性の有効な長さの議論については、Hasty et al., Mol Cell Biol 11:5586-91 (1991)を参照されたい。

#### 【0090】

[00115] ある特定の実施形態において、(UL)および(DL)は、適合性を有するランディングパッド相同配列と共にいずれのランディングパッドでの外因性核酸(ES)のゲノム組込みを可能にする十分な長さおよび配列同一性を有するあらゆるヌクレオチド配列を含んでいてもよい。ある特定の実施形態において、(UL)および(DL)の各々は、独立して、約20~5,000ヌクレオチドからなる。ある特定の実施形態において、(UL)および(DL)の各々は、独立して、約100~2,500ヌクレオチドからなる。ある特定の実施形態において、(UL)および(DL)の各々は、独立して、約100~1,000ヌクレオチドからなる。ある特定の実施形態において、(UL)および(DL)の各々は、独立して、約250~750ヌクレオチドからなる。ある特定の実施形態において、(UL)および(DL)の各々は、独立して、約100、200、300、400、500、600、700、800、900、1000、1100、1200、1300、1400、1500、1600、1700、1800、1900、2000、2100、2200、2300、2400、2500、2600、2700、2800、2900、3000、3100、3200、3300、3400、3500、3600、3700、3800、3900、4000、4100、4200、4300、4400、4500、4600、4700、4800、4900または5,000ヌクレオチドからなる。一部の実施形態において、(UL)および(DL)の各々は、独立して、約500ヌクレオチドからなる。

#### 【0091】

[00116] ある特定の実施形態において、1つまたは複数の外因性ドナー核酸における(UL)のヌクレオチド配列は、互いに実質的に同一である。ある特定の実施形態において、1つまたは複数の外因性ドナー核酸における(DL)のヌクレオチド配列は、互いに実質的に同一である。例えば、1つまたは複数の外因性ドナー核酸における全ての(UL)は、互いに少なくとも70%、75%、80%、85%、90%、95%、96%、97%、98%、99%または100%同一であるヌクレオチド配列を含み、1つまたは複数の外因性ドナー核酸における全ての(DL)は、互いに少なくとも70%、75%、80%、85%、90%、95%、96%、97%、98%、99%または100%同一であるヌクレオチド配列を含む。一部の実施形態において、1つまたは複数の外因性ドナー核酸における全ての(DL)のヌクレオチド配列は、互いに同一であり、1つまたは複数の外因性ドナー核酸における全ての(DL)のヌクレオチド配列は、互いに同一である。

#### 【0092】

##### 5.3.2. 目的の核酸

[00117] 一部の実施形態において、外因性ドナー核酸は、目的の核酸(D)をさらに含む。目的の核酸は、当業者によって有用とみなされているあらゆるDNAセグメントであり得る。例えば、DNAセグメントは、宿主ゲノムに「ノックイン」可能な目的の遺伝子を含んでいてもよい。他の実施形態において、DNAセグメントは、宿主細胞のゲノムの標的部位へのコンストラクトの組込み時に標的遺伝子を特異的に崩壊させることにより、崩壊した遺伝子を機能できなくすることが可能な「ノックアウト」コンストラクトとして機能する。目的の核酸(D)の有用な例としては、これらに限定されないが、タンパク質コード配列、レポーター遺伝子、蛍光マーカーコード配列、プロモーター、エンハンサー、ターミネーター、転写活性化因子、転写リプレッサー、転写活性化因子結合部位、転写リプレッサー結合部位、イントロン、エクソン、ポリAテール、多重クロニング部位、核局在化シグナル、mRNA安定化シグナル、組込み遺伝子座、エピトープタグコード配列、分解シグナル、スペーサーもしくはスタッファー(stuffer)配列(例えば、公知の機能がないランダムに生成した配列)、リンカー配列、デグロン配列、融合パートナー配列、他のあらゆる天然に存在するもしくは合成DNA分子、またはそれらの組合せもしくは部分的組合せが挙げられる。一部の実施形態において、(D)は、天然の起点に由来するものでもよい。代替として、(D)は、完全にインビトロで生産された合成の起点に由来す

るものでもよい。

【 0 0 9 3 】

[00118] さらに、(D)は、単離された天然に存在するDNA分子の任意の組合せ、または単離された天然に存在するDNA分子と合成DNA分子との任意の組合せを含んでもよい。例えば、(D)は、タンパク質コード配列に作動可能に連結した異種プロモーター、ポリAテールに連結されたタンパク質コード配列、フレーム内でエピトープタグコード配列と連結されたタンパク質コード配列などを含んでもよい。目的の核酸(D)は、クローニングされたDNA(例えば、DNA「ライブラリー」)から当業界において公知の標準的手順により得てもよいし、化学合成により得てもよいし、cDNAクローニングにより得てもよいし、または望ましい細胞から精製されたゲノムDNAもしくはその断片のクローニングにより得てもよいし、またはPCR増幅およびクローニングにより得てもよい。例えば、Sambrook et al., Molecular Cloning, A Laboratory Manual, 3d. ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (2001); Glover, D.M. (ed.), DNA Cloning: A Practical Approach, 2d. ed., MRL Press, Ltd., Oxford, U.K. (1995)を参照されたい。

10

【 0 0 9 4 】

[00119] 特定の実施形態において、目的の核酸(D)は、選択可能マーカーをコードする核酸を含まない。これらの実施形態において、本明細書に記載される方法によってもたらされる組込みの高い効率、選択培地で形質転換した細胞を増殖させる必要なく組込み事象のスクリーニングおよび同定を可能にする。しかしながら、それにもかかわらず選択培地での増殖が望ましい他の実施形態において、目的の核酸(D)は、宿主ゲノムへの外因性核酸の組込みを選択するのに使用できる選択可能マーカーを含んでもよい。

20

【 0 0 9 5 】

[00120] 多種多様の選択可能マーカーが当業界において公知である

(例えば、Kaufinan, Meth. Enzymol. 185:487 (1990); Kaufman, Meth. Enzymol., 185:537 (1990); Srivastava and Schlessinger, Gene, 103:53 (1991); Romanos et al., in DNA Cloning 2: Expression Systems, 2nd Edition, pages 123-167 (IRL Press 1995); Markie, Methods Mol. Biol., 54:359 (1996); Pfeifer et al., Gene, 188:183 (1997); Tucker and Burke, Gene, 199:25 (1997); Hashida-Okado et al., FEBS Letters, 425:117 (1998)を参照)。一部の実施形態において、選択可能マーカーは、薬物耐性マーカーである。薬物耐性マーカーは、そうでなければ細胞を殺していたと予想される外因性薬物を細胞が無毒化することを可能にする。薬物耐性マーカーの説明に役立つ例としては、これらに限定されないが、例えばアンピシリン、テトラサイクリン、カナマイシン、ブレオマイシン、ストレプトマイシン、ハイグロマイシン、ネオマイシン、Zeoцин(商標)などの抗生物質に対する耐性を付与するものが挙げられる。他の実施形態において、選択可能マーカーは、栄養要求性マーカーである。栄養要求性マーカーは、必須成分が欠如した培地中で増殖しながら、細胞が必須成分(通常はアミノ酸)を合成することを可能にする。選択可能な栄養要求性遺伝子配列としては、例えば、hisDが挙げられ、これは、ヒスチジノールの存在下でのヒスチジン非含有培地中における増殖を可能にするものである。他の選択可能マーカーとしては、ブレオマイシン耐性遺伝子、メタロチオネイン遺伝子、ハイグロマイシンB-ホスホトランスフェラーゼ遺伝子、AURI遺伝子、アデノシンデアミナーゼ遺伝子、アミノグリコシドホスホトランスフェラーゼ遺伝子、ジヒドロ葉酸レダクターゼ遺伝子、チミジンキナーゼ遺伝子、キサンチン-グアニンホスホリボシルトランスフェラーゼ遺伝子などが挙げられる。他の実施形態において、選択可能マーカーは、栄養要求性突然変異をレスキューするもの以外のマーカーである。例えば、宿主細胞株は、突然変異を公知の選択方法によって同定できる限りは、栄養要求性突然変異以外の突然変異、例えば、宿主にとって致死性ではない突然変異や、株の意図した使用、例えば工業的発酵に対して有害作用も引き起こさない突然変異を含んでもよい。

30

40

【 0 0 9 6 】

50

[00121] 染色体に組み込まれたポリヌクレオチドを含む宿主細胞形質転換株はまた、個々のDNAセグメント（例えば、コンポーネントポリヌクレオチド）またはDNAセグメントの組合せによってコードされた他の特色、例えば発光するペプチドの発現を呈する宿主細胞形質転換株を選択することによって、または個々の宿主細胞コロニーの分子分析によって、例えば、単離された組み立てられたポリヌクレオチドまたは染色体の組み込み部位の制限酵素マッピング、PCR増幅、または配列分析によって同定することもできる。ある特定の実施形態において、各外因性ドナー核酸に関連付けられたバーコード配列は、個々のDNAセグメントまたは宿主細胞のゲノム中に組み込まれたDNAセグメントの組合せを同定するのに使用することができる。

【0097】

5.3.3. コンポーネントポリヌクレオチドおよびそのランディングパッドでの組み込み

[00122] ほとんどのラージスケールの宿主細胞の操作は、既存のDNAパーツの使用および再使用を伴う。ある特定の実施形態において、外因性ドナー核酸は、組み立てられたポリヌクレオチドの様々な組合せを生成するためのモジュラーパーツとして共に形質転換することができるDNAパーツ（コンポーネントポリヌクレオチドとも称される）である。組み立てられたポリヌクレオチドの様々な組合せが、ランディングパッド中に組み込まれると、それらは、ランディングパッド中の組み込みのために5'および3'領域の両方に相同配列を有する単一の断片のポリヌクレオチドと比較して、操作された宿主細胞においてより大きい分子多様性をもたらす。分子多様性が大きければ大きいほど、操作された宿主細胞が呈する表現型をより多様にすることができる。このような結果は、以下のさらなる詳細に記載される図10で実証される。さらに、コンポーネントポリヌクレオチドは、インピボでコンビナトリアル的に組み立てられ、宿主細胞のゲノムにコンビナトリアル的に組み込まれることから、モジュラーパーツとしてのコンポーネントポリヌクレオチドの使用は、宿主細胞の生成において時間とコストを節約することができる。また、コンポーネントポリヌクレオチドは、宿主細胞が、適合性を有する相同配列を有するランディングパッドを含む場合、異なる宿主細胞のバックグラウンドまたは異なるゲノム遺伝子座での組み込みのために再カスタマイズする必要もない。

【0098】

[00123] したがって、相同組換えを介してインピボで組み立てて、組み立てられたコンポーネントポリヌクレオチドを生成することができるコンポーネントポリヌクレオチドが本明細書で提供される。ある特定の実施形態において、2つのコンポーネントポリヌクレオチド、すなわち第1のコンポーネントポリヌクレオチドおよび最後のコンポーネントポリヌクレオチドは、インピボで宿主細胞において相同組換えされて、宿主細胞のゲノム中のランディングパッドのいずれかに組み込まれる。この実施形態において、1つまたは複数の第1のコンポーネントポリヌクレオチドの各々は、5'から3'方向に、上流ランディングパッド相同配列（ULP）と相同組換えが可能な上流ライブラリー配列（UL）；第1の目的の核酸、および第1のリンカー配列を含む。1つまたは複数の最後のコンポーネントポリヌクレオチドの各々は、5'から3'方向に、最後のリンカー配列、最後の目的の核酸、および下流ランディングパッド相同配列（DLP）と相同組換えが可能な下流ライブラリー配列（DL）を含む。1つまたは複数の第1のコンポーネントポリヌクレオチドの3'末端における各第1のリンカー配列は、1つまたは複数の最後のコンポーネントポリヌクレオチドの5'末端で各最後のリンカー配列と相同組換えが可能であり、コンポーネントポリヌクレオチドの様々な組合せを生成することができる。ランディングパッド内で部位特異的ヌクレアーゼによって二本鎖破断が誘導されると、内因性相同組換え機構は、ランディングパッド中の切断された部位で組み立てられたコンポーネントポリヌクレオチドを高い組み込み効率で組み込む。図1Eの中央に、2つのコンポーネントポリヌクレオチドの例を例示する。

【0099】

[00124] ある特定の実施形態において、3つ以上の異なるタイプのコンポーネントポリヌクレオチド（例えば、第1のコンポーネントポリヌクレオチド、中間コンポーネントポ

10

20

30

40

50

リヌクレオチド、および最後のコンポーネントポリヌクレオチド)は、宿主細胞に共に形質転換される。この実施形態において、1つまたは複数の第1のコンポーネントポリヌクレオチドの各々は、5'から3'方向に、上流ランディングパッド相同配列(ULP)と相同組換えが可能な上流ライブラリー配列(UL)、D<sub>0</sub>群から選択される任意のDNAセグメント、リンカー配列LB<sub>0</sub>を含む。1つまたは複数の中間コンポーネントポリヌクレオチドの各々は、5'から3'方向に、第1のリンカー配列LA<sub>n</sub>、D<sub>n</sub>群から選択される任意のDNAセグメント、第2のリンカー配列LB<sub>n</sub>を含む。この実施形態において、nは、1から中間コンポーネントポリヌクレオチドの数までの整数を表す。1つまたは複数の最後のコンポーネントポリヌクレオチドの各々は、5'から3'方向に、リンカー配列LA<sub>m</sub>、D<sub>m</sub>群から選択される任意のDNAセグメント、および下流ランディングパッド相同配列(DLP)と相同組換えが可能な下流ライブラリー配列(DL)を含む。この実施形態において、各リンカー配列LB<sub>(p-1)</sub>は、リンカー配列LA<sub>p</sub>と相同組換えが可能であり、nは、1~(m-1)までの様々な整数であり、pは、1~mの整数を表し、各D<sub>0</sub>群、...D<sub>n</sub>群、...D<sub>m</sub>群、は、独立して、1つまたは複数のDNAセグメントからなる。

#### 【0100】

[00125] ある特定の実施形態において、DNAセグメント(D<sub>0</sub>群、D<sub>n</sub>群、またはD<sub>m</sub>群における)は、あらゆる目的の核酸であり得る。例えば、セクション5.2.2に記載されるあらゆる目的の核酸は、コンポーネントポリヌクレオチド中にDNAセグメントとして取り込むことができる。

#### 【0101】

[00126] ある特定の実施形態において、3つのコンポーネントポリヌクレオチドを組み立てて、ランディングパッドに組み込むことができる。3つのコンポーネントポリヌクレオチド(第1のコンポーネントポリヌクレオチド、中間コンポーネントポリヌクレオチド、および最後のコンポーネントポリヌクレオチド)の例は、図1Eの一番下に例示される。

#### 【0102】

[00127] ある特定の実施形態において、インピボでコンポーネントポリヌクレオチドを互いに相同組換えするリンカー配列は、そのカウンターパートのリンカー配列への相同組換えを可能にする十分な長さおよび配列同一性を有するあらゆるヌクレオチド配列を含んでいてもよい。ある特定の実施形態において、各リンカー配列は、約20~5,000ヌクレオチドからなる。ある特定の実施形態において、各リンカー配列は約24~2,500ヌクレオチドからなる。ある特定の実施形態において、各リンカー配列は、約24~1,000ヌクレオチドからなる。ある特定の実施形態において、各リンカー配列は約24~500ヌクレオチドからなる。ある特定の実施形態において、各リンカー配列は、約24~約100ヌクレオチドからなる。ある特定の実施形態において、短いリンカー配列(例えば、約24~約36ヌクレオチド)は、それらが2つのコンポーネントポリヌクレオチドを近接させる必要があるDNAセグメント(例えば、プロモーターおよびオープンリーディングフレーム)と共に組み立てるのに使用される場合、使用される。例えば、プロモーターをオープンリーディングフレームに接続するのに使用される図1Eに示されるリンカー配列「A」は、比較的短い配列(例えば、約24~約36ヌクレオチド)であり得る。一方で、2つのオープンリーディングフレームを組み立てるのに使用される図2Cに示されるリンカー配列「SNAP」は、比較的長い配列(例えば、500~5000ヌクレオチド)であり得る。

#### 【0103】

[00128] ある特定の実施形態において、インピボでコンポーネントポリヌクレオチドを相同組換えする2つのリンカー配列は、互いに約70%、75%、80%、85%、90%、95%、96%、97%、98%、99%または100%相同または同一なヌクレオチド配列を含んでいてもよい。例えば、第1のコンポーネントポリヌクレオチドおよび最後のコンポーネントポリヌクレオチドがインピボで組み立てられ、ランディングパッドへの組込みを起こす場合、第1のコンポーネントポリヌクレオチドの3'末端における第1の

10

20

30

40

50



リンカー配列および最後のコンポーネントポリヌクレオチドの 5' 末端における最後のリンカー配列は、互いに約 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% または 100% 相同であるかまたは同一であるヌクレオチド配列を含む。1 つまたは複数の中間コンポーネントポリヌクレオチドが、1 つまたは複数の第 1 のコンポーネントポリヌクレオチドおよび 1 つまたは複数の最後のコンポーネントポリヌクレオチドと共に導入される場合、リンカー配列  $L_B(p-1)$  およびリンカー  $L_{Ap}$  は、互いに約 70%、75%、80%、85%、90%、95%、96%、97%、98%、99% または 100% 相同であるかまたは同一であるヌクレオチド配列を含む。

【0104】

[00129] 好適なコンポーネントポリヌクレオチドおよびコンポーネントポリヌクレオチドに関する他の特徴は、その全体が参照により本明細書に組み入れられる米国特許第 8, 110, 360 号でさらに詳細に説明されている。

【0105】

5. 3. 4. 追加の機能的なエレメント：バーコード配列および追加のヌクレアーゼ標的配列

[00130] 本明細書で提供される外因性ドナー核酸は、追加の機能的なエレメントを含んでいてもよい。ある特定の実施形態において、追加の機能的なエレメントは、宿主細胞のゲノム中の外因性ランディングパッド中に存在するものとは異なるヌクレアーゼ標的配列を包含していてもよい。例えば、図 3 A で示されるように、第 1 のコンポーネントポリヌクレオチドは、上流ライブラリー配列 (UL) と第 1 のリンカー配列 (「A」) との間に位置するプロモーター (曲がった矢印として示される) を含む。図 3 A に示される第 1 のコンポーネントポリヌクレオチドは、宿主細胞のゲノム中のランディングパッド内の標的ヌクレアーゼ配列「X」(図 3 C に示される) とは異なる、ヌクレアーゼ Y - カッターによって結合できるヌクレアーゼ標的配列「Y」をさらに含む。図 3 B に示される最後のコンポーネントポリヌクレオチドは、最後のリンカー配列 (「A」) と下流ライブラリー配列 (DL) との間に位置するオープンリーディングフレーム (ORF) を含む。最後のコンポーネントポリヌクレオチドは、宿主細胞のゲノム中のランディングパッド内の標的ヌクレアーゼ配列「X」とは異なる、ヌクレアーゼ Z - カッターによって結合できるヌクレアーゼ標的配列「Z」をさらに含む。コンポーネントポリヌクレオチドがランディングパッドに組換えおよび組込みを起こすと、これらの追加のヌクレアーゼ標的配列 (例えば、Y および Z 部位) は、後続の形質転換中に宿主細胞のゲノム中に追加の外因性ドナー核酸をさらに導入するために、後で利用することができる。

【0106】

[00131] ある特定の実施形態において、公知の配列の固有なバーコードは、各外因性ドナー核酸に関連付けられている。バーコード配列は、宿主細胞のゲノム中のランディングパッド内に組み込まれる外因性ドナー核酸の同定をさらに容易にするために使用することができる。一般的に、バーコード配列は、宿主細胞のゲノム中のランディングパッド中に組み込まれた外因性ドナー核酸またはコンポーネントポリヌクレオチドの同定が可能になるように十分な長さを有し、互いに十分に異なっている。ある特定の実施形態において、各バーコード配列は、少なくとも 4、5、6、7、8、9、10、11、12、13、14、15、16、またはそれより多くのヌクレオチド長を有していてもよい。一実施形態において、バーコード配列は、7 ヌクレオチド長を有していてもよい。

【0107】

[00132] バーコード配列は、各外因性ドナー核酸またはコンポーネントポリヌクレオチド中のあらゆる好適な配置に取り込むことができる。コンポーネントポリヌクレオチド中のバーコード配列の例示的な配置は、図 3 A および 3 B において一連の水平のバーライン (「NNNNNNN の 7nt バーコード」とも表記される) として示される。ある特定の実施形態において、第 1 のコンポーネントポリヌクレオチドは、上流ライブラリー配列 (UL) と目的の核酸 (例えば、曲がった矢印として示されるプロモーター) との間に位置するバーコード配列を含んでいてもよい。ある特定の実施形態において、最後のコンポーネントポ

10

20

30

40

50

リヌクレオチドは、下流ライブラリー配列（DL）と目的の核酸（例えば、ORF）との間に位置するバーコード配列を含んでいてもよい。特定の実施形態において、バーコード配列は、上流ライブラリー配列（UL）と追加の機能的なエレメント、例えばヌクレアーゼ標的配列（例えば、Y - カッター部位）との間に挿入することができる。ある特定の実施形態において、ヌクレアーゼ標的配列（例えば、Y - カッター部位）は、所与のライブラリーにおいて、全ての第1のコンポーネントポリヌクレオチドに共通である。特定の実施形態において、バーコード配列は、下流ライブラリー配列と追加のエレメント、例えばヌクレアーゼ標的配列（例えば、Z - カッター部位）との間に挿入することができる。ある特定の実施形態において、ヌクレアーゼ標的配列（例えば、Z - カッター部位）は、所与のライブラリーにおいて、全ての最後のコンポーネントポリヌクレオチドに共通である。「Y」および「Z」ヌクレアーゼ標的配列を外因性ドナー核酸に組み込むことによって、それらは、バーコード増幅のための共通のプライマー結合部位として役立たせることができ、さらに追加のラウンドまたは切断およびヌクレオチド組込みを可能にする。図3Cに示される実施形態において、iALG1 遺伝子座でランディングパッド中に組み込まれた第1のコンポーネントポリヌクレオチドのアイデンティティは、一対の以下の標準プライマー：iALG1のゲノム配列に結合する「UL」と表記される（ULP）の5'の上流に位置するプライマー、および「Y」ヌクレアーゼ標的配列に結合するプライマーPYを使用することによって決定することができる。iALG1 遺伝子座におけるランディングパッドにおいて組み込まれた最後のコンポーネントポリヌクレオチドのアイデンティティは、一対の以下の標準プライマー：「Z」ヌクレアーゼ標的配列に結合するプライマーPZ、および「DL」と表記される（DLP）の3'の下流に位置するiALG1のゲノム配列に結合するプライマーを使用することによって決定することができる。

【0108】

#### 5.4 外因性ドナー核酸のゲノム組込みの方法

[00133] 宿主細胞のゲノム中において操作により作製された1つまたは複数のランディングパッドに1つまたは複数の外因性ドナー核酸をゲノム組込みのための方法が本明細書で提供される。上述したように、ある特定の実施形態において、宿主細胞のゲノムに組み込まれた各ランディングパッドは、外因性ドナー核酸の宿主細胞媒介相同組換えを容易にするために、標準化された相同配列を含む。このような実施形態において、標準化されたランディングパッド相同配列は、外因性ドナー核酸それ自体が、ランディングパッド相同配列に適合性を有する相同配列を含む限り、あらゆる外因性ドナー核酸の、宿主細胞のゲノム中のあらゆるランディングパッドでの組込みを可能にする。それゆえに、ある特定の実施形態において、外因性ドナー核酸の組込み事象は、ランディングパッド周辺のゲノム遺伝子座から独立しており、組換えのために内因性ゲノム配列との相同性を頼らない。特定の実施形態において、本方法は、宿主細胞を、1つまたは複数の外因性ドナー核酸、およびランディングパッド中に配置されたヌクレアーゼ標的配列を切断することが可能な1つまたは複数のヌクレアーゼと接触させる工程を含む。ランディングパッド内での切断は、切断部位で、またはその近くで相同組換えの頻度を大きく増加させる。

【0109】

#### 5.4.1 ランディングパッドにおける外因性ドナー核酸の組込み

[00134] 図1A～1Cは、部位特異的ヌクレアーゼを使用した単一のランディングパッドにおける外因性ドナー核酸のライブラリーからの単一の外因性ドナー核酸のマーカーレズゲノム組込み方法の例示的な実施形態を例示する。外因性ドナー核酸は、宿主細胞に導入され、ここで外因性ドナー核酸は、上流ライブラリー配列（UL）と下流ライブラリー配列（DL）と端部に有する目的の核酸を含む（例えば、プロモーターに作動可能に連結したORF）。図1A～1Cに示される実施形態において、外因性ドナー核酸の（UL）および（DL）は、それぞれランディングパッドの上流ランディングパッド相同配列（ULP）および下流ランディングパッド相同配列（DLP）と配列同一性を共有する。図1Bに示される実施形態において、図1Aに示されるランディングパッド中に組込みを起こすことができる4つの異なる外因性ドナー核酸がある。部位特異的ヌクレアーゼ（一丁のはさみとして示

される)は、ランディングパッド中のヌクレアーゼ標的配列(NTS、「X」として示される)を認識し切断する。部位特異的ヌクレアーゼによってヌクレアーゼ標的配列内の二本鎖破断が誘導されると、内因性相同組換え機構は、切断された部位で目的の核酸を組み込む。図1Cで示されるように、宿主細胞のゲノムに一旦組み込まれたら、目的の核酸は、上流および下流ランディングパッド相同配列(図1Cに示されるULおよびDL)の間に配置され、ニュートラルなゲノム遺伝子座(例えば、酵母ゲノム中のALG1遺伝子座下流の遺伝子間領域)に入れ子になっている。図1A~1Cで例示されているように、外因性ドナー核酸のゲノム組込み事象は、相同組換えのために内因性ゲノム配列を頼らない。図1D~1Fは、ランディングパッド中の組込みに好適な、ランディングパッドおよび外因性ドナー核酸およびコンポーネントポリヌクレオチドの代替の実施形態を例示する。

10

#### 【0110】

[00135] 図1A~1Fで例示された実施形態において、内因性相同組換え機構は、二本鎖破断を含まないゲノム部位と比較してより高い頻度で、切断された部位に目的の核酸を組み込む。この組込み頻度の増加は、組換え事象を受けた形質転換株を選択するために選択可能マーカーを共に組み込む必要性を取り除く。例えば操作された微生物の構築中の選択可能マーカーへの必要性をなくすことによって、完全で機能的な生合成経路を含む株を構築するのに必要な時間は大幅に少なくなる。加えて、所与の宿主生物にとって利用可能なマーカーの蓄えが限定されているために選択可能マーカーを再利用する必要性があることにより、操作戦略がもはや制限されることはない。

#### 【0111】

20

[00136] 一部の実施形態において、うまく組み込まれた外因性核酸を含む形質転換した細胞のマーカーレス回収は、1000、900、800、700、600、500、400、300、200または100個のスクリーニングされた接触させた宿主細胞またはそれらのクローン集団毎に約1の頻度内で起こる。特定の実施形態において、うまく組み込まれた外因性核酸を含む形質転換した細胞のマーカーレス回収は、90、80、70、60、50、40、30、20または10個のスクリーニングされた接触させた宿主細胞またはそれらのクローン集団毎に約1の頻度内で起こる。より特定の実施形態において、うまく組み込まれた外因性核酸を含む形質転換した細胞のマーカーレス回収は、9、8、7、6、5、4、3または2個のスクリーニングされた接触させた宿主細胞またはそれらのクローン集団毎に約1の頻度内で起こる。ある特定の実施形態において、宿主細胞のゲノム中に操作により作製された全てのランディングパッド(例えば、3つのランディングパッド)への外因性ドナー核酸組込みの成功率は、90%より大きい。ある特定の実施形態において、宿主細胞のゲノム中に操作により作製された全てのランディングパッド(例えば、3つのランディングパッド)への外因性ドナー核酸組込みの成功率は、95%より大きい。より特定の実施形態において、宿主細胞は、酵母細胞であり、他の宿主細胞型に比べて酵母の相同組換え能力が増加したことに起因して、組込み頻度が増加している。

30

#### 【0112】

[00137] 選択可能マーカーを使用せずに外因性ドナー核酸がランディングパッド中に組み込まれた細胞を同定するための様々な方法が利用可能である。一部の実施形態において、このような方法は、ランディングパッド中のあらゆる変化を検出しようとするものであり、これらに限定されないが、PCR方法、シーケンシング方法、ヌクレアーゼ消化、例えば、制限マッピング、サザンブロット、およびそれらの任意の組合せが挙げられる。ある特定の実施形態において、外因性ドナー核酸に関連付けられたバーコード配列は、ランディングパッドにおいて組み込まれた外因性ドナー核酸のアイデンティティを決定するのに使用することができる。ランディングパッド中の外因性ドナー核酸(またはコンポーネントポリヌクレオチド)のゲノム組込みの成功率は高いことから、形質転換宿主細胞の表現型は、宿主細胞遺伝子型解析せずに目的の外因性ドナー核酸を有する細胞を同定するのに使用することができる。

40

#### 【0113】

[00138] 特定の実施形態において、宿主細胞のゲノムの複数の外因性ランディングパッ

50

ド中に外因性核酸を同時にゲノム組込みのための方法が本明細書で提供される。本方法は、  
( a ) 宿主細胞のゲノム中に組み込まれた複数の ( x ) 外因性ランディングパッドを含む宿主細胞であって、各外因性ランディングパッドは、上流ランディングパッド相同配列 ( ULP ) と下流ランディングパッド相同配列 ( DLP ) との間に位置するヌクレアーゼ標的配列 ( NTS ) を含む、宿主細胞を、

( i ) 1 つまたは複数の外因性ドナー核酸 ( ES ) であって、各 ( ES ) は、上流ライブラリー配列 ( UL ) と下流ライブラリー配列 ( DL ) との間に位置する目的の核酸 ( D ) を含み、各 ( UL ) は、( x ) 外因性ランディングパッドのいずれかの任意の ( ULP ) で相同組換えが可能であり、各 ( DL ) は、( x ) 外因性ランディングパッドのいずれかの任意の ( DLP ) で相同組換えが可能である、外因性ドナー核酸 ( ES ) ; および

( i i ) 複数の ( x ) 外因性ランディングパッドのいずれかの ( NTS ) に結合し、( x ) 外因性ランディングパッド内の部位を切断することが可能な、1 つまたは複数のヌクレアーゼ ( N )

と接触させる工程 ; および

( b ) 接触させた宿主細胞から生成した宿主細胞を回収する工程を含む、

外因性ドナー核酸 ( ES ) のいずれかは、各ランディングパッド周辺のゲノム配列から独立して、複数の ( x ) 外因性ランディングパッドのいずれかにおいて組み込まれる。ある特定の実施形態において、x は、少なくとも 2 の整数である。ある特定の実施形態において、x は、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19 または 20 の整数である。

#### 【 0 1 1 4 】

[00139] 図 2 A および図 2 B は、複数のランディングパッドへの複数の外因性ドナー核酸の同時のゲノム組込みの例示的な実施形態を例示する。上述したように、図 2 A は、ゲノム遺伝子座 A L G 1、M G A 1、および Y C T 1 下流の遺伝子間領域に 3 つのランディングパッドを含む例示的な宿主細胞を例示する。この実施形態において、各ランディングパッドは、上流ランディングパッド相同配列 ( ULP ) と下流ランディングパッド相同配列 ( DLP ) との間に位置する、ヌクレアーゼ X - カッター (例えば、CphI) によって切断できる「X」として示されるヌクレアーゼ標的配列 ( NTS ) を含む。図 2 A に示される実施形態において、全ての 3 つのランディングパッドは、( ULP ) に関して同じヌクレオチド配列「U L」、( DLP ) に関して同じヌクレオチド配列「D L」、および同じヌクレアーゼ標的配列 X を有する。それゆえに、図 2 A に示される例示的な実施形態において、全ての 3 つのランディングパッドは、ランディングパッドのいずれかへの組込みのために適合性を有する相同配列を含むあらゆる外因性ドナー核酸を組み込むように標準化されている。

#### 【 0 1 1 5 】

[00140] 図 2 B は、ランディングパッド相同配列との宿主細胞媒介相同組換えに適合性を有する相同配列 (ライブラリー配列と称される) を含む外因性ドナー核酸のライブラリーを例示する。図 2 B で例示された実施形態において、各外因性ドナー核酸は、プロモーター (ORF 上流の曲がった矢印として示される) に作動可能に連結した遺伝子またはオープンリーディングフレーム (「ORF」) である目的の核酸を含む。図 2 B において、各目的の核酸 ( D ) は、上流ライブラリー配列 (図 2 B では「U L」と表記される) と下流ライブラリー配列 (図 2 B では「D L」と表記される) との間に位置する。宿主細胞に導入されると、宿主細胞媒介相同組換えを介して、外因性ドナー核酸の各上流ライブラリー配列 ( UL ) は、あらゆる上流ランディングパッド相同配列 ( ULP ) で相同組換えが可能であり、外因性ドナー核酸の各下流ライブラリー配列 ( DL ) は、あらゆる下流ランディングパッド相同配列 ( DLP ) で相同組換えが可能である。

#### 【 0 1 1 6 】

[00141] 一部の実施形態において、図 2 B に示される 2 つ以上の異なるタイプの外因性ドナー核酸は、外因性ドナー核酸の異なる組合せを宿主細胞のゲノム中のランディングパッドに組み込むことができるように、宿主細胞と接触させることができる。一部の実施形

10

20

30

40

50

態において、異なるタイプの外因性ドナー核酸のモル比は、各タイプの外因性ドナー核酸の組み込み率に互いに相対的に影響を与えるように変更することができる。他の実施形態において、同じ外因性ドナー核酸の複数のコピーが宿主細胞のゲノム中の標準化されたランディングパッドにおいて組み込むことができるように、単一タイプの外因性ドナー核酸を宿主細胞に導入することができる。

【0117】

[00142] 一部の実施形態において、外因性ドナー核酸は、「スタッパー」または「スペーサー」配列（例えば、ランダムに生成した、機能しない配列）を含む別の外因性ドナー核酸と組み合わせて導入されてもよい。例えば、図4を参照されたい。スタッパー配列を含む外因性ドナー核酸を使用することによって、宿主細胞のゲノム中に組み込まれた他の機能的な核酸のコピー数を変更することができる。例えば、これらの2つのDNAコンストラクトで形質転換した宿主細胞の一部は、3つのランディングパッド中に組み込まれた、ORF1の単一のコピーおよびスペーサー配列の2つのコピーを包含していてもよいし、ORF1の2つのコピーおよびスペーサー配列の1つのコピーを包含していてもよいし、ORF1の3つのコピーおよび0個のスペーサー配列などを包含していてもよい。ある特定の実施形態において、ランディングパッド中に組み込まれた異なるコピー数の各タイプの外因性ドナー核酸を有する宿主細胞を生成するために、スタッパー配列を含む外因性ドナー核酸は、2つ以上の異なるタイプの外因性ドナー核酸と組み合わせて共に導入することができる。

10

【0118】

[00143] 図2Aおよび図2Bで例示された実施形態において、ヌクレアーゼX-カッター（例えば、CphI）も宿主細胞に導入され、ここでヌクレアーゼX-カッターは、その対応するヌクレアーゼ標的配列（NTS）内の固有な配列を認識し切断することが可能である。ヌクレアーゼX-カッターによってヌクレアーゼ標的配列（NTS）が切断されたら、内因性相同組換え機構は、二本鎖破断を含まないランディングパッドと比較してより高い頻度で、いずれかの3つの標準化されたランディングパッドに、外因性ドナー核酸（例えば、ORF1、ORF2、ORF3、またはORF4を含む）のいずれかを同時に組み込むことを容易にする。4つのオープンリーディングフレームのいずれかが3つのランディングパッドのいずれかにおいて組み込みを起こすことができるため、本明細書で提供される方法は、ランディングパッド中に組み込まれたオープンリーディングフレームの様々な組合せを有する多様な宿主細胞の集団を生成する。

20

30

【0119】

[00144] ある特定の実施形態において、宿主細胞と接触する異なるタイプの外因性核酸（ES）の数は、宿主細胞のゲノム中に存在するランディングパッド（x）の数によって限定されない。一実施形態において、宿主細胞に導入された異なるタイプの外因性ドナー核酸の数は、宿主細胞のゲノム中に存在するランディングパッド（x）の数を超えていてもよい。他の実施形態において、宿主に導入された異なるタイプの外因性ドナー核酸の数は、宿主細胞のゲノム中に存在するランディングパッド（x）の数未満であってもよい。

【0120】

[00145] 図1A～図1Cで例示される単一のランディングパッドにおける単一の外因性核酸の組み込みと同様に、複数の外因性ドナー核酸の同時の複数の組み込みは、ランディングパッドを二本鎖破断を誘導することが可能なヌクレアーゼと接触させない場合と比較してより実質的に高い頻度で起こる。一部の実施形態において、この組み込み頻度の増加は、複数の組換え事象を同定するために1つまたは複数の選択可能マーカーを共に組み込む必要性を取り除く。一部の実施形態において、複数のうまく組み込まれた外因性核酸を含む形質転換した細胞のマーカーレス回収は、1000、900、800、700、600、500、400、300、200または100個のスクリーニングされた接触させた宿主細胞またはそれらのクローン集団毎に約1の頻度内で起こる。特定の実施形態において、マーカーレス回収は、90、80、70、60、50、40、30、20または10個のスクリーニングされた接触させた宿主細胞またはそれらのクローン集団毎に約1の頻度内で

40

50

起こる。より特定の実施形態において、マーカーレス回収は、9、8、7、6、5、4、3または2個のスクリーニングされた接触させた宿主細胞またはそれらのクローン集団毎に約1の頻度内で起こる。ある特定の実施形態において、宿主細胞のゲノム中に操作により作製された全てのランディングパッド（例えば、3つのランディングパッド）への外因性ドナー核酸組込みの成功率は、90%より大きい。ある特定の実施形態において、宿主細胞のゲノム中に操作により作製された全てのランディングパッド（例えば、3つのランディングパッド）への外因性ドナー核酸組込みの成功率は、95%より大きい。より特定の実施形態において、宿主細胞は、酵母細胞であり、他の宿主細胞型に比べて酵母の相同組換え能力が増加したことに起因して、組込み頻度が増加している。

【0121】

5. 4. 1. 1 ランディングパッドにおけるコンポーネントポリヌクレオチドの組込み [00146] 別の態様において、コンポーネントポリヌクレオチドである外因性ドナー核酸を組み込むための方法が本明細書で提供される。コンポーネントポリヌクレオチドは、宿主細胞に導入されると、互いに相同組換えして、1つまたは複数の組み立てられたコンポーネントポリヌクレオチドを生成する。これらの組み立てられたコンポーネントポリヌクレオチドは、その5'および3'末端に相同性領域を含み、それにより、宿主細胞のゲノム中のランディングパッドのいずれかでそれらが相同組換えおよび組込みを起こすことを可能にする。ある特定の実施形態において、各コンポーネントポリヌクレオチドは、少なくとも1つのリンカー配列を含み、それにより、インビボでリンカー配列を使用して各コンポーネントポリヌクレオチドを別のコンポーネントポリヌクレオチドと相同組換えすることを可能にする。ある特定の実施形態において、コンポーネントポリヌクレオチドは、単一の形質転換反応で、宿主細胞に共に導入される。

【0122】

[00147] したがって、一態様において、宿主細胞のゲノム中の1つまたは複数のランディングパッド中にコンポーネントポリヌクレオチドをゲノムで組み込む(genomically integrating)方法が本明細書で提供される。一実施形態において、本方法は、

(a) 宿主細胞のゲノム中に組み込まれた1つまたは複数の(x)外因性ランディングパッドを含む宿主細胞であって、各外因性ランディングパッドは、上流ランディングパッド相同配列(ULP)と下流ランディングパッド相同配列(DLP)との間に位置するヌクレアーゼ標的配列(NTS)を含む、宿主細胞を、

(i) 1つまたは複数の第1のコンポーネントポリヌクレオチドであって、各第1のコンポーネントポリヌクレオチドは、5'から3'方向に、

(1) 上流ランディングパッド相同配列(ULP)と相同組換えが可能な上流ライブラリー配列(UL)；

(2) 第1の目的の核酸、および

(3) 第1のリンカー配列

を含む、第1のコンポーネントポリヌクレオチド；

(ii) 1つまたは複数の最後のコンポーネントポリヌクレオチドであって、各最後のコンポーネントポリヌクレオチドは、5'から3'方向に、

(1) 最後のリンカー配列、

(2) 最後の目的の核酸、および

(3) 下流ランディングパッド相同配列(DLP)と相同組換えが可能な下流ライブラリー配列(DL)

を含み、

1つまたは複数の第1のコンポーネントポリヌクレオチドの任意の第1のリンカー配列は、1つまたは複数の最後のコンポーネントポリヌクレオチドの任意の最後のリンカー配列と相同組換えが可能である、最後のコンポーネントポリヌクレオチド；および

(iii) 1つまたは複数の(x)外因性ランディングパッドのいずれかにおいて任意の(NTS)を切断することが可能なヌクレアーゼ(N)

と接触させる工程；および

10

20

30

40

50

(b) 接触させた宿主細胞から生成した宿主細胞を回収する工程を含み、

1つまたは複数の第1のコンポーネントポリヌクレオチドからの第1のコンポーネントポリヌクレオチドと、1つまたは複数の最後のコンポーネントポリヌクレオチドからの最後のコンポーネントポリヌクレオチドとの任意の組合せが、インビボでそれらのリンカー配列を介して相同組換えされ、該組合せが、各ランディングパッド周辺のゲノム配列から独立して、1つまたは複数の(x) 外因性ランディングパッドのいずれかにおいて組み込まれる。一部の実施形態において、xは、少なくとも1の整数である。

#### 【0123】

[00148] この実施形態において、コンポーネントポリヌクレオチドが、1つまたは複数のランディングパッド中の上流ランディングパッド相同配列(ULP)に相同組換えおよび組込みを起こすように、その5'末端に相同性領域を含む場合、そのコンポーネントポリヌクレオチドは、第1のコンポーネントポリヌクレオチドと称される。同様に、コンポーネントポリヌクレオチドが、1つまたは複数のランディングパッド中の下流ランディングパッド相同配列(DLP)と相同組換えするように、その3'末端に相同性領域を含む場合、そのコンポーネントポリヌクレオチドは、最後のコンポーネントポリヌクレオチドと称される。

#### 【0124】

[00149] ある特定の実施形態において、宿主細胞のゲノム中に存在するランディングパッド(x)の数は、少なくとも2つであり、第1および最後のコンポーネントポリヌクレオチドの任意の組合せは、その数のランディングパッド(x)のいずれかにおいて組み込まれる。ある特定の実施形態において、宿主細胞のゲノム中に存在するランディングパッド(x)の数は、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19または20である。ある特定の実施形態において、第1および最後のコンポーネントポリヌクレオチドの任意の組合せは、これらの複数の(x)ランディングパッドのいずれかにおいて組み込まれる。

#### 【0125】

[00150] 図2Cは、コンビナトリアル的に組み合わせられて、宿主細胞のゲノム中に存在する3つのランディングパッドのいずれかにコンビナトリアル的に組み込むことができるコンポーネントポリヌクレオチドの例示的な実施形態を例示する。図2Cで示されるように、各々それ自体のプロモーターに作動可能に連結した2つのオープンリーディングフレームは、コンビナトリアル的に組み合わせられて、図2Aに示されるランディングパッドのいずれかに組み立てられたコンポーネントポリヌクレオチドを組み込む。図2Cにおいて、第1のコンポーネントポリヌクレオチドは、5'から3'方向に、上流ライブラリー配列(UL)、プロモーターに作動可能に連結したオープンリーディングフレームORF1、および第1のリンカー配列「SNAP」を含む。第1のコンポーネントポリヌクレオチドの上流ライブラリー配列(UL)は、図2Aに示される実施形態において上流ライブラリー配列(UL)と同じ配列を有する上流ランディングパッド相同配列(ULP)と相同組換えが可能である。最後のコンポーネントポリヌクレオチドは、5'から3'方向に、最後のリンカー配列SNAP、プロモーターに作動可能に連結したオープンリーディングフレーム、および下流ランディングパッド相同配列(DLP)と相同組換えが可能な下流ライブラリー配列DLを含む。図2Aに示される例示的な実施形態において、(DLP)は、下流ライブラリー配列(DL)と同じ配列を有する。第1のコンポーネントポリヌクレオチドの第1のリンカー配列(すなわち、図2CにおけるSNAP)は、インビボで最後のコンポーネントポリヌクレオチドの最後のリンカー配列(すなわち、図2CにおけるSNAP)と相同組換えが可能であり、組み立てられたコンポーネントポリヌクレオチドを生成することができる。相同組換えする場合、組み立てられたコンポーネントポリヌクレオチドは、各オープンリーディングフレームがそれ自体のプロモーター(「ロックドプロモーター」設計と称される)に作動可能に連結した複数のオープンリーディングフレームを生成する。第1のコンポーネントポリヌクレオチドにおける上流ライブラリー配列(UL)および最後のコンポ

10

20

30

40

50

ーメントポリヌクレオチドにおける下流ライブラリー配列 (DL) は、図 2 A に示される宿主細胞のゲノム中の 3 つのランディングパッドのいずれかにおける組み立てられたコンポーネントポリヌクレオチドの組み込みを可能にする。組み立てられたコンポーネントポリヌクレオチドはいずれも、複数のランディングパッドのいずれかでコンビナトリアル的に組み込むことができる。

【 0 1 2 6 】

[00151] 図 2 C に示されるコンポーネントポリヌクレオチドは、それ自体のプロモーターに作動可能に連結した各オープンリーディングフレームを例示するが、一部の実施形態において、プロモーターは、オープンリーディングフレームから分離されており、2 つの異なるコンポーネントポリヌクレオチドとして宿主細胞に共に導入されてもよい。例えば、図 2 C において、(UL) および ORF 1 を含むコンポーネントポリヌクレオチドは、2 つの別個のコンポーネントポリヌクレオチド (リンカー配列 A を有するプロモーターコンポーネントポリヌクレオチドおよびリンカー配列 A を有するプロモーターレス ORF 1) に分離していてもよいし、(DL) および ORF 2、ORF 3、または ORF 4 を含むコンポーネントポリヌクレオチドは、同様に 2 つの別個のコンポーネントポリヌクレオチドに分離していてもよい。宿主細胞のゲノムにコンビナトリアル的に組み込むことができる追加のコンポーネントポリヌクレオチドで宿主細胞を形質転換することによって、株におけるより大きいコンビナトリアルな多様性をもたらすことができる。

【 0 1 2 7 】

[00152] 図 2 D は、一緒に組み立てられて宿主細胞のゲノム中のランディングパッドに組み込まれるコンポーネントポリヌクレオチドの別の例示的な実施形態を例示する。図 2 D に示される例示的な実施形態において、「スプリット - プロモーター」設計のコンポーネントポリヌクレオチドが例示され、第 1 のコンポーネントポリヌクレオチドは、様々なプロモーター強度を有する異なるプロモーターを含み、最後のコンポーネントポリヌクレオチドは、異なるオープンリーディングフレームを含む。より具体的には、第 1 のコンポーネントポリヌクレオチドは、5' から 3' の方向で、宿主細胞のゲノム中のランディングパッド中の上流ランディング配列 (ULP) と相同組換えが可能な上流ライブラリー配列 (UL)、プロモーター；および「A」で表される第 1 のリンカー配列を含む。最後のコンポーネントポリヌクレオチドは、5' から 3' の方向で、「A」で表される最後のリンカー配列、オープンリーディングフレーム (例えば、ORF 1、ORF 2、ORF 3、および ORF 4)、および宿主細胞のゲノム中のランディングパッド中の下流ランディング配列 (DLP) と相同組換えが可能な下流ライブラリー配列 (DL) を含む。1 つまたは複数の第 1 のコンポーネントポリヌクレオチドが、1 つまたは複数の最後のコンポーネントポリヌクレオチドと組み合わせる宿主細胞に導入される場合、第 1 のコンポーネントポリヌクレオチドの第 1 のリンカー配列 A は、最後のコンポーネントポリヌクレオチドの最後のリンカー配列 A と相同組換えして、組み立てられたコンポーネントポリヌクレオチドの異なる組合せを生成する。4 つの異なる第 1 のコンポーネントポリヌクレオチドが、図 2 D で示されるように 4 つの異なる最後のコンポーネントポリヌクレオチドと共に宿主細胞に導入される場合、宿主細胞は、16 個の異なる組み立てられたコンポーネントポリヌクレオチドを生成する可能性があり、図 2 A に示される 3 つのランディングパッドに組み込むことができる。異なるコピー数の組み立てられたコンポーネントポリヌクレオチドは各々、3 つのランディングパッド中に組み込まれる可能性があるため、標準化されたランディングパッドへの組み込みのランダムさはさらに、宿主細胞中で追加の遺伝学的変異をもたらすことができる。

【 0 1 2 8 】

[00153] 図 2 C および図 2 D は、組み立てられたコンポーネントポリヌクレオチドの異なる組合せを生成するのに 2 つのコンポーネントポリヌクレオチドの組合せ (すなわち、第 1 のコンポーネントポリヌクレオチドおよび最後のコンポーネントポリヌクレオチド) が使用される例示的な実施形態を例示するが、インビボで組み立てられたコンポーネントポリヌクレオチドを生成するために、3 つ以上のコンポーネントポリヌクレオチドの組合



せを使用することができる。例えば、ある特定の実施形態において、1つまたは複数の中間コンポーネントポリヌクレオチドを、第1のコンポーネントポリヌクレオチドと最後のコンポーネントポリヌクレオチドとの間に組み立てて、コンポーネントポリヌクレオチドのDNA集合体を生成することができる。さらに、図2Dは、(UL)に隣接してプロモーターを含むコンポーネントポリヌクレオチドを例示するが、一部の実施形態において、(DL)とリンカー配列との間にプロモーターを取り込むことができる。

#### 【0129】

[00154] したがって、一態様において、1種または複数の宿主細胞のゲノム中に操作により作製されたランディングパッドのいずれかに組み込むための、インビボで組み立てられた3つ以上のコンポーネントポリヌクレオチドをゲノムで組み込む方法が本明細書で提供される。一実施形態において、本方法は、

10

(a) 宿主細胞のゲノム中に組み込まれた1つまたは複数の(x)外因性ランディングパッドを含む宿主細胞であって、1つまたは複数の外因性ランディングパッドの各々は、上流ランディングパッド相同配列(ULP)と下流ランディングパッド相同配列(DLP)との間に位置するヌクレアーゼ標的配列(NTS)を含む、宿主細胞を、

(i) 1つまたは複数の第1のコンポーネントポリヌクレオチドであって、各第1のコンポーネントポリヌクレオチドは、5'から3'方向に、上流ランディングパッド相同配列(ULP)と相同組換えが可能な上流ライブラリー配列(UL)、D<sub>0</sub>群から選択される任意のDNAセグメント、リンカー配列LB<sub>0</sub>を含む、第1のコンポーネントポリヌクレオチド；

20

(ii) 1つまたは複数の中間コンポーネントポリヌクレオチドであって、各中間コンポーネントポリヌクレオチドは、5'から3'方向に、第1のリンカー配列LA、D<sub>n</sub>群から選択される任意のDNAセグメント、第2のリンカー配列LB<sub>n</sub>を含み、nは、1から中間コンポーネントポリヌクレオチドの数までの整数を表す、中間コンポーネントポリヌクレオチド；

(iii) 1つまたは複数の最後のコンポーネントポリヌクレオチドであって、各最後のコンポーネントポリヌクレオチドは、5'から3'方向に、リンカー配列LA、D<sub>m</sub>群から選択される任意のDNAセグメント、および下流ランディングパッド相同配列(DLP)と相同組換えが可能な下流ライブラリー配列(DL)を含む、最後のコンポーネントポリヌクレオチド

と接触させる工程であって、

30

各リンカー配列LB<sub>(p-1)</sub>は、リンカー配列LA<sub>p</sub>と相同組換えが可能であり、nは、1~(m-1)の様々な整数であり、pは、1~mの整数を表し、各D<sub>0</sub>群、...D<sub>n</sub>群、...D<sub>m</sub>群、は、独立して、1つまたは複数のDNAセグメントからなる、工程；および

(b) 接触させた宿主細胞から生成した宿主細胞を回収する工程を含み、

1つまたは複数の第1のコンポーネントポリヌクレオチドからの第1のコンポーネントポリヌクレオチド、1つまたは複数の中間コンポーネントポリヌクレオチドからの中間コンポーネントポリヌクレオチド、および1つまたは複数の最後のコンポーネントポリヌクレオチドからの最後のコンポーネントポリヌクレオチドの任意の組合せが、インビボでそれらのリンカー配列を介して相同組換えされ、該組合せが、各ランディングパッド周辺のゲノム配列から独立して、1つまたは複数の(x)外因性ランディングパッドのいずれかににおいて組み込まれる。一部の実施形態において、xは、少なくとも1の整数である。一部の実施形態において、xは、2の整数である。一部の実施形態において、xは、3の整数である。一部の実施形態において、xは、4、5、6、7、8、9または10の整数である。

40

#### 【0130】

[00155] ある特定の実施形態において、宿主細胞のゲノム中に存在するランディングパッド(x)の数は、少なくとも2つであり、第1の、中間、および最後のコンポーネントポリヌクレオチドの任意の組合せは、ランディングパッドのいずれかににおいて組み込まれる。ある特定の実施形態において、宿主細胞のゲノム中に存在するランディングパッド(x)

50

の数は、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19または20である。ある特定の実施形態において、第1の、中間、および最後のコンポーネントポリヌクレオチドの任意の組合せは、これらの複数のランディングパッドのいずれかにおいて組み込まれる。

【0131】

[00156] コンポーネントポリヌクレオチドに関する組成物および方法の追加の詳細は、参照によりその全体が本明細書に組み入れられる米国特許第8,110,360号に見出すことができる。

【0132】

5.4.2. 外因性ドナー核酸の遺伝子型 / 表現型の関係の追跡

10

[00157] 本明細書で提供される組成物および方法は、外因性ドナー核酸の何百および何千ものコンビナトリアルな組込みを生成する迅速で効率的な方法を提供する。外因性ドナー核酸の多様なコンビナトリアルな組込みは、ある特定の観察可能な特色または表現型（例えば、細胞機能、標的分子産生など）を有する操作された宿主細胞を生成することができる。このような多様なコンビナトリアルな組込みによって生成した異なる表現型を追跡し、外因性ドナー核酸の様々な組合せによる遺伝子型の寄与と共にマッピングすることができる。

【0133】

[00158] したがって、外因性ドナー核酸による表現型の寄与をスクリーニングする方法が本明細書で提供される。一実施形態において、本方法は、

20

(a) 宿主細胞を、複数の外因性核酸(ES)と接触させる工程であって、各(ES)は、バーコード配列でタグ付けされており、宿主細胞のゲノムは、宿主細胞媒介相同組換えを介して、宿主細胞のゲノム中の複数の(ES)の任意の1つまたは組合せを組み込むように設計されている、工程；

(b) 接触させた宿主細胞から生成した、特異的な表現型を呈する宿主細胞をスクリーニングする工程；および

(c) 各(ES)に関連付けられたバーコード配列を使用して、宿主細胞のゲノムに組み込まれた(ES)のアイデンティティを決定する工程を含む。

【0134】

30

[00159] ある特定の実施形態において、宿主細胞のゲノム中に組み込まれた複数の(x)外因性ランディングパッドを提供することによって、宿主細胞のゲノム中の複数の(ES)の任意の1つまたは組合せを組み込むように設計されている。ある特定の実施形態において、各外因性ランディングパッドは、各ランディングパッド中に、上流ランディングパッド相同配列(ULP)と下流ランディングパッド相同配列(DLP)との間に位置するヌクレアーゼ標的配列(NTS)を含む。ある特定の実施形態において、宿主細胞は、ヌクレアーゼまたはヌクレアーゼをコードするプラスミドおよび複数の(ES)と同時に接触させる。

【0135】

[00160] ある特定の実施形態において、複数の外因性ドナー核酸(ES)の少なくとも2つは、インビボで宿主細胞に共に形質転換され、相同組換えを介して組み立てられてランディングパッドにおいて組込みを起こすコンポーネントポリヌクレオチドである。このような実施形態において、各コンポーネントポリヌクレオチドは、各コンポーネントポリヌクレオチドを他のコンポーネントポリヌクレオチドと区別するその固有なバーコード配列と連結されていてもよい。ある特定の実施形態において、1つまたは複数の第1のコンポーネントポリヌクレオチドおよび1つまたは複数の最後のコンポーネントポリヌクレオチドが宿主細胞と接触して、宿主細胞のゲノム中の複数のランディングパッドにおいて組み込まれた組み立てられたコンポーネントポリヌクレオチドが生成される。ある特定の実施形態において、1つまたは複数の第1のコンポーネントポリヌクレオチド、1つまたは複数の中間コンポーネントポリヌクレオチド、および1つまたは複数の最後のコンポーネントポリヌクレオチドが宿主細胞と接触して、宿主細胞のゲノム中の複数のランディングパ

40

50

ッドにおいて組み込まれた組み立てられたコンポーネントポリヌクレオチドが生成される。

【 0 1 3 6 】

[00161] 図 4 は、外因性ドナー核酸によって表現型の寄与をスクリーニングする方法の 1 つの例示的な実施形態を例示する。図 2 A に関して上述したように、図 4 に示される宿主細胞は、宿主細胞のゲノム中に 3 つのランディングパッドを含み、各ランディングパッドは、上流ランディングパッド相同配列 (ULP) と下流ランディングパッド相同配列 (DLP) との間に位置するヌクレアーゼ標的配列 (NTS) を含む。宿主細胞は、コンポーネントポリヌクレオチド (例えば、様々な強度のプロモーターを含む 9 つの第 1 のコンポーネントポリヌクレオチド、および図 4 に示されるオープンリーディングフレームを含む 10 個の最後のコンポーネントポリヌクレオチド) と接触し、インビボで相同組換えされると、組み立てられたコンポーネントポリヌクレオチドの様々な組合せが生成する。図 4 に示される実施形態において、宿主細胞は、ランディングパッド中のヌクレアーゼ標的配列 X を切断することが可能なヌクレアーゼ X (例えば、F-CphI) をコードするプラスミドと同時に形質転換される。

10

【 0 1 3 7 】

[00162] 形質転換した細胞は、あらゆる好適な培地で増殖させることができる。例えば、図 4 で示されるように、形質転換宿主細胞を選択培地で平板培養して、コロニーを回収する。コロニーをピックアップし、ハイスループットスクリーニングを使用して、「ヒット」、または表現型について関心のあるコロニーに関してスクリーニングする。例えば、図 4 で示されるように、形質転換宿主細胞の表現型は、2 つの異なるアッセイ、すなわちアッセイ 1 およびアッセイ 2 でスクリーニングすることができる。アッセイ 1 は、例えばある条件下での標的分子の産生であってもよいし、アッセイ 2 は、別の条件下での標的分子の産生であってもよい。次いで、ある特定の表現型 (例えば、両方のアッセイにおける高いレベルの標的分子産生) を呈する形質転換宿主細胞を選択することができる。次いで、宿主細胞のゲノム中のランディングパッドに組み込まれたコンポーネントポリヌクレオチド (例えば、遺伝子およびプロモーター) を同定するために、標準プライマーを用いて選択された宿主細胞を遺伝子型解析する。図 4 は、標準プライマー結合部位の配置を示す。これらの標準化されたプライマーを使用して、ランディングパッド中に組み込まれたコンポーネントポリヌクレオチドだけでなく、各コロニーにおける (ES) 組込みのゲノム配置も決定することができる。外因性ドナー核酸の様々なコンビナトリアルな組込みの遺伝子型 - 表現型の関係を決定した後、遺伝子改変された宿主細胞に関する望ましい表現型 (例えば、標的分子産生) をさらに改善するために、外因性ドナー核酸の組成を調整することができる。

20

30

【 0 1 3 8 】

[00163] 図 4 は、バーコード配列を使用してある特定の表現型を有する個々の株を選択し、遺伝子型解析する例示的な実施形態を例示するが、バルクのコンビナトリアル的に操作された宿主細胞の集団の遺伝子型は、バーコード配列を使用して追跡することができる。言い換えれば、個々の形質転換株中に組み込まれた外因性ドナー核酸のアイデンティティを決定する代わりに、ある特定の表現型 (例えば、望ましいまたは望ましくない表現型のいずれか) を有する複数の形質転換株と一緒にプールして、プールされた集団の遺伝子型を全体として決定することができる。例えば、様々な株のプールされた集団において、どの外因性ドナー核酸が集団中で富化されるかを決定することができる。ある特定の実施形態において、集団全体の遺伝子型を、最初にスクリーニングを介してそれらの表現型を決定せずとも経時的に追跡することができる。例えば、発酵容器に株の混合集団を接種することができ、どの遺伝子型が経時的に「継承されるか」を決定することができる。ある特定の実施形態において、PCR 反応を実行して各ランディングパッド領域を増幅することにより、プールされた集団中の富化された外因性ドナー核酸のアイデンティティを決定することができる。あらゆる好適なシーケンシング方法 (例えば、次世代シーケンシング) が、PCR 反応から生成したアンプリコンに関連付けられたバーコード配列を決定するのに使用することができる。バーコード情報に基づき、プールされた株の集団中のどの外

40

50

因性ドナー核酸がある特定の表現型に寄与するのかを決定することができる。この情報に基づき、望ましくない表現型に関連する外因性ドナー核酸の優先順位を下げて、コンビナトリアル設計の次のラウンドから除去することができ、望ましい表現型を有するものを優先させて、コンビナトリアル設計の次のラウンドで繰り返すことができる。

【0139】

5.4.3. アドレス指定可能なまたは任意選択のランディングパッドでのゲノム組込み [00164] 別の態様において、図2Aに示される標準化された一次ランディングパッドに加えて、またはその代わりに任意選択のまたはアドレス指定可能なランディングパッドを含む宿主細胞のゲノム中の外因性ドナー核酸をゲノム組込み方法が本明細書で提供される。ある特定の実施形態において、標準化された一次ランディングパッドに加えて、またはその代わりにアドレス指定可能なランディングパッドを作り出して、外因性ドナー核酸のサブセットに標的化された組込みを実行することが望ましい場合がある。これらのランディングパッドは、「アドレス指定可能な」ランディングパッドと称されるが、これはなぜなら、外因性ドナー核酸のサブセットが、標準化された一次ランディングパッドではなくこれらのランディングパッドへの組込みを起こすようにアドレス指定または指示され得るためである。

【0140】

[00165] 図5Bは、二次ランディングパッドが図2Aに示される標準化された一次ランディングパッドとは異なる例示的な実施形態を例示する。図2Aで例示された実施形態において、全ての3つの一次ランディングパッドは、全ての(ULP)につき同じ上流相同性ランディングパッド配列「UL」および全ての(DLP)につき同じ下流相同性ランディングパッド配列「DL」を含む。宿主細胞のゲノム中の遺伝子座<sub>1</sub>において組み込まれた第4のランディングパッドは、上流ランディングパッド相同配列(UL)を含み、これは、標準化された一次ランディングパッドと同じであるが異なる下流ランディングパッド相同配列(DL<sub>1</sub>)を含む。(UL)および(DL<sub>1</sub>)を有する二次ランディングパッドを使用して、特異的な外因性ドナー核酸(例えば、図5Aに示される最後のコンポーネントポリヌクレオチドA-ORF5-DL<sub>1</sub>)を、異なるプロモーターを有する第1のコンポーネントポリヌクレオチドのいずれかと組み合わせる組み込むことができる。これらのコンポーネントポリヌクレオチドは、第1のコンポーネントポリヌクレオチドのいずれかとコンビナトリアル的に組み合わせられて、ALG1、iMGA1、およびiYCT1遺伝子座で標準化された一次ランディングパッドに組み込まれるように設計された他の最後のコンポーネントポリヌクレオチド(A-ORF1-DL、A-ORF2-DL、A-ORF3-DLおよびA-ORF4-DL)と共に形質転換することができる。

【0141】

[00166] 図6Bは、標準化された一次ランディングパッドとは異なる2つの任意選択のアドレス指定可能なランディングパッドをさらに含む別の例示的な実施形態を例示する。図6Bに示される宿主細胞は、宿主細胞のゲノム中の遺伝子間領域iALG1、iMGA1、およびiYCT1において組み込まれている3つの一次ランディングパッドを含む。これらの一次ランディングパッドは、上流ランディングパッド相同配列(UL)の同じセット、下流ランディングパッド相同配列(DL)の同じセット、(UL)と(DL)との間に位置するヌクレアーゼ認識部位Xの同じセットを含む。図6Bで示されるように、宿主細胞は、遺伝子座<sub>1</sub>および遺伝子座<sub>2</sub>における2つのアドレス指定可能なランディングパッドをさらに含む。これらの2つのアドレス指定可能なランディングパッドを使用して、特異的な外因性ドナー核酸(例えば、遺伝子座<sub>1</sub>における最後のコンポーネントポリヌクレオチドA-ORF4-DL<sub>1</sub>、および遺伝子座<sub>2</sub>における最後のコンポーネントポリヌクレオチドA-ORF5-DL<sub>2</sub>)を、異なるプロモーターを有する第1のコンポーネントポリヌクレオチドのいずれかと組み合わせる組み込むことができる。これらのコンポーネントポリヌクレオチドは、第1のコンポーネントポリヌクレオチドのいずれかとコンビナトリアル的に組み合わせられ、標準化されたランディングパッドへの組込みを起こすように設計された他の最後のコンポーネントポリヌクレオチド(A-ORF1-DL、A-ORF2-DL、A-O

RF3-DL)と共に形質転換することができる。

【0142】

[00167] 遺伝子座1および遺伝子座2におけるアドレス指定可能なランディングパッドはまた、任意選択のランディングパッドでもあり、これはなぜなら、それらが、X - カッターが導入されて一次ランディングパッドを切断するときに切断されないが、任意選択で別個のヌクレアーゼによって「開裂」または切断される標的ヌクレアーゼ配列を含むためである。遺伝子座1における任意選択のランディングパッドは、G - カッターによって認識される標的ヌクレアーゼ配列「G」を含む。遺伝子座2における任意選択のランディングパッドは、H - カッターによって認識および/または結合できる標的ヌクレアーゼ配列「H」を含む。これらの任意選択のランディングパッドは、一次ランディングパッドの切断と同時に、またはその後逐次的に、それ自体のヌクレアーゼ（例えば、G - カッターおよびH - カッター）によって切断することができる。

10

【0143】

[00168] 図6Aは、プールされた「カッター」試薬（例えば、ヌクレアーゼをコードするプラスミド）と組み合わせて使用できるプールされたDNAコンストラクト（すなわち、コンポーネントポリヌクレオチド）を例示する。図6Bに示される全てのランディングパッドが上流ランディングパッド相同配列として（UL）を有するため、上流ライブラリー配列として（UL）を含むコンポーネントポリヌクレオチドの第1のセットは、5つのランディングパッドのいずれか1つで（UL）と相同組換えするのに適合性を有する。一方で、ORF1、ORF2、またはORF3を含むコンポーネントポリヌクレオチドの第2のセットは、iALG1遺伝子座、iMGA1遺伝子座、およびiYCT1遺伝子座で一次ランディングパッドのいずれかにコンビナトリアル的への組込みを起こすことができる。しかしながら、ORF4を含むコンポーネントポリヌクレオチドのみが、遺伝子座1における（DL1）に対するその（DL1）の相同性のために、遺伝子座1において組込みを起こすことができる。同様に、ORF5を含むコンポーネントポリヌクレオチドのみが、遺伝子座2における（DL2）に対するその（DL2）相同性のために、遺伝子座2において組込みを起こすことができる。

20

【0144】

[00169] 任意選択のおよび/またはアドレス指定可能なランディングパッドには他の多くのバリエーションがある。例えば、宿主細胞は、上流ランディングパッド相同配列および下流ランディングパッド相同配列の両方に関して、標準化されたランディングパッドの一次セットと比較して異なる相同配列を含むランディングパッドの追加のセットを含んでもよい。ある特定の実施形態において、ランディングパッドは、ゲノム配列にとって外因性の上流ランディングパッド相同配列（例えば、ランダムに生成したヌクレオチド配列）、およびゲノム遺伝子座にとって内因性の下流ランディングパッド相同配列（例えば、ゲノム遺伝子座のオープンリーディングフレーム）を含んでもよい。ある特定の実施形態において、ランディングパッドは、ゲノム遺伝子座にとって内因性の上流ランディングパッド相同配列（例えば、ゲノム遺伝子座のオープンリーディングフレーム）、およびゲノム遺伝子座にとって外因性の下流ランディングパッド相同配列を含んでもよい。

30

【0145】

5.4.4. 標的化されたランディングパッドへの標的化されたランディングパッドおよび外因性ドナー核酸の組込み

40

[00170] 本明細書で提供される組成物および方法は、特異的なゲノム部位への外因性ドナー核酸の標的化された組込み、および遺伝子改変された宿主細胞の表現型に対するその作用のスクリーニングにおいて適用される可能性がある。ある特定の実施形態において、特異的なゲノム部位における外因性ドナー核酸の標的化された組込みは、外因性ドナー核酸を宿主細胞に導入する前に、そのようなゲノム部位にランディングパッドを組み込むことによって達成することができる。目的の遺伝学的エレメント（例えば、オープンリーディングフレーム）に隣接する特異的なゲノム部位で一旦組み込まれたら、これらのランディングパッドは、外因性ドナー核酸を特異的なゲノム部位において組み込むことを可能に

50

する。このような特異的なゲノム部位に構築された外因性ランディングパッドは、外因性ドナー核酸を特異的なゲノム部位で標的化してランディングパッドに組み込むことを可能にすることから、「標的化されたランディングパッド」と称される。したがって、ある特定の実施形態において、目的の核酸（例えば、プロモーター、ターミネーター、またはデグロン配列）を含む外因性ドナー核酸を、標的化されたランディングパッドに組み込んで、オープンリーディングフレームの転写および/またはそのタンパク質の発現もしくは安定性に対するそれらの作用をスクリーニングすることができる。

#### 【0146】

[00171] ある特定の実施形態において、標的化されたランディングパッドは、パーツ（例えば、図1Eおよび図2B～2Dに示される（UL）、（DL）、（A））の既存のライブラリー中の外因性ドナー核酸の相同配列に適合性を有する相同配列を含む。したがって、標準的なライブラリー配列および/または標準化されたリンカー配列を有する外因性ドナー核酸のライブラリーは、再カスタマイゼーション不要の特異的なゲノム部位における標的化されたランディングパッドへの標的化された組込みに利用することができる。

10

#### 【0147】

[00172] したがって、一態様において、宿主細胞のゲノムへの核酸の標的化された組込みのための方法であって、

（a）オープンリーディングフレームの5'に組み込まれた外因性ランディングパッドを含む宿主細胞であって、外因性ランディングパッドは、上流ランディングパッド相同配列（ULP）と下流ランディングパッドリンカー配列（DLPL）との間に位置するヌクレアーゼ標的配列（NTS）を含む、宿主細胞を、

20

（i）1つまたは複数の第1のコンポーネントポリヌクレオチドであって、各第1のコンポーネントポリヌクレオチドは、5'から3'方向に、

（1）（ULP）と相同組換えが可能な上流ライブラリー配列（UL）；

（2）第1の目的の核酸；および

（3）第1のリンカー配列

を含む、第1のコンポーネントポリヌクレオチド；

（ii）（NTS）に結合し、外因性ランディングパッド内の部位を切断することが可能な、ヌクレアーゼ（N）

と接触させる工程

30

（b）接触させた宿主細胞から生成した宿主細胞を回収する工程を含む、

（DLPL）は、1つまたは複数の第1のコンポーネントポリヌクレオチドのいずれかの第1のリンカー配列と相同組換えして、外因性ランディングパッドで、1つまたは複数の第1のコンポーネントポリヌクレオチドからの第1のコンポーネントポリヌクレオチドを外因性ランディングパッドに組み込むことが可能である、方法が本明細書で提供される。

#### 【0148】

[00173] ある特定の実施形態において、オープンリーディングフレームは、そのネイティブの遺伝子座における内因性遺伝子のオープンリーディングフレームである。他の実施形態において、オープンリーディングフレームは、オープンリーディングフレームが配置されている特異的なゲノム部位に標的化されたランディングパッドを導入する前の非ネイティブの遺伝子座に組み込まれた外因性遺伝子のオープンリーディングフレームである。

40

#### 【0149】

[00174] ある特定の実施形態において、本方法は、（a）回収された宿主細胞の表現型をスクリーニングして、特定の表現型を有する宿主細胞を選択する工程；および（b）宿主細胞中の外因性ランディングパッド中に組み込まれた第1のコンポーネントポリヌクレオチドのアイデンティティを決定する工程をさらに含む。ある特定の実施形態において、各コンポーネントポリヌクレオチドは、固有なバーコード配列を含み、バーコード配列は、標的化されたランディングパッド中に組み込まれた第1のコンポーネントポリヌクレオチドを同定するのに使用することができる。ある特定の実施形態において、標的化された

50

ランディングパッドは、上記のセクション 5 . 2 に記載される 1 つまたは複数の標準化されたランディングパッドと共に、宿主細胞のゲノム中に存在する。

【 0 1 5 0 】

[00175] ある特定の実施形態において、本方法は、調節配列（例えば、プロモーター）を含む第 1 のコンポーネントポリヌクレオチドのライブラリーを含む目的の内因性遺伝子の適正量を決定すること（例えば、下方制御または上方制御すること）に適用することができる。図 7 A は、オープンリーディングフレームのネイティブの遺伝子座で「プロモータースワップランディングパッド」を組み込むのに使用することができるプロモータースワップランディングパッドビルディングコンストラクトの概略図を例示する。宿主細胞のゲノム中に一旦組み込まれたら、プロモータースワップランディングパッドは、ライブラリー中のあらゆるプロモーターを目的のオープンリーディングフレームの 5 ' に組み込むことができる。

10

【 0 1 5 1 】

[00176] 図 7 A を参照すると、プロモータースワップランディングパッドビルディングコンストラクトは、第 1 のコンポーネントポリヌクレオチドの相同組換えのための、上流ランディングパッド相同配列（ULP、「UL」として示される）、および下流ランディングパッドリンカー配列（DLPL、「A」として示される）を含む。プロモータースワップランディングパッドビルディングコンストラクトは、ヌクレアーゼ標的配列（NTS、「X」として示される）、および（ULP）と（DLPL）との間に位置する一時的なプロモーター p A C T 1 をさらに含む。プロモーターランディングパッドビルディングコンストラクトは、O R F のネイティブの遺伝子座における相同組換えのために、O R F の上流配列（「US\_ORF」）、およびそれぞれ相同的に 5 ' および 3 ' 領域のための O R F の一部をさらに含む。図 7 B に示したように、プロモータースワップランディングパッドが宿主細胞のゲノム中に一旦組み込まれると、一時的なプロモーター p A C T 1 がオープンリーディングフレームの発現を作動させる。

20

【 0 1 5 2 】

[00177] プロモータースワップランディングパッドは、あらゆる好適な方法を使用して望ましいゲノム部位において組み込むことができる。例えば、図 7 A で示されるように、部位特異的ヌクレアーゼ（例えば、C R I S P R / C a s 9）は、オープンリーディングフレームのためのネイティブのプロモーター p O R F を切断するのに使用することができる。ネイティブのプロモーター p O R F が一旦切断されると、プロモータースワップランディングパッドビルディングコンストラクトは、オープンリーディングフレームの上流配列（US\_ORF）および O R F の 5 ' 領域と宿主細胞媒介相同組換えを介して相同組換えされる。図 7 B に、O R F の 5 ' において組み込まれたプロモータースワップランディングパッドで形質転換したこのような宿主細胞を示す。この宿主細胞を親宿主細胞として使用することにより、O R F の転写、そのタンパク質発現または安定性を改変できる様々な調節配列を含む第 1 のコンポーネントポリヌクレオチドをさらに組み込んでスクリーニングすることができる。

30

【 0 1 5 3 】

[00178] 図 7 C を参照すると、プロモータースワップランディングパッドを有する親宿主細胞は、プロモーターライブラリーと接触させる。プロモーターライブラリーは、目的の核酸として様々な強度のプロモーターを含む第 1 のコンポーネントポリヌクレオチドを含む。ライブラリー中の各第 1 のコンポーネントポリヌクレオチドは、上流ライブラリー配列（UL）と第 1 のリンカー配列 A との間に位置する異なるプロモーター（曲がった矢印として示される）を含む。ある特定の実施形態において、ライブラリー中の各第 1 のコンポーネントポリヌクレオチドは、図 3 A で示されるように、バーコード配列および / または追加のヌクレアーゼ標的配列をさらに含む。プロモータースワップランディングパッドを有する親宿主細胞は、プロモータースワップランディングパッド中の標的ヌクレアーゼ配列 X を切断するために、ヌクレアーゼ X - カッター（例えば、CphI）をコードするプラスミドと接触させることができ、それにより、第 1 のコンポーネントポリヌクレオチドが

40

50

プロモータースワップランディングパッドにおいて組込みを起こすことが可能になる。

#### 【 0 1 5 4 】

[00179] 図 7 D において、概略図は、プロモーターライブラリーで形質転換した宿主細胞のコロニーのピックアップ、および望ましい表現型を有する宿主細胞のスクリーニングをさらに例示する。コロニーの遺伝子型（すなわち、プロモーター配列）は、あらゆる好適な方法を使用して同定することができる。例えば、第 1 のコンポーネントポリヌクレオチドの各々がバーコード配列を含む場合、第 1 のコンポーネントポリヌクレオチドに関連付けられたバーコード配列は、回収された宿主細胞を遺伝子型解析するのに使用することができる。バーコード配列に隣接する領域に結合する標準プライマーは、バーコード配列を含むアンプリコンを生成するのに使用することができる。アンプリコンは、次世代シーケンシングなどのあらゆる好適な技術を使用してシーケンシングすることができる。ある特定の実施形態において、標準プライマーを使用して生成したアンプリコンは、次世代シーケンシング（例えば、IlluminaからのMiSeqシステムを介した）の前に断片化する必要性をなくする特定の断片サイズ（例えば、約 5 0 0 ~ 約 7 0 0 塩基対）を有する。

10

#### 【 0 1 5 5 】

[00180] 図 7 A ~ 7 D に示される内因性遺伝子発現（例えば、そのネイティブの遺伝子座におけるオープンリーディングフレーム）の適正量を決定する方法は、単なる例示であり、他の多くのバリエーションがある。例えば、第 1 のコンポーネントポリヌクレオチドは、他の機能的なエレメント、例えば、遺伝子改変された宿主細胞の様々な細胞機能または表現型の改変を引き起こすことができる、転写レギュレーター配列、デグロン配列、または融合パートナー配列を含んでいてもよい。

20

#### 【 0 1 5 6 】

[00181] 別の態様において、標的化されたランディングパッドを、特異的なゲノム部位に操作により作製してオープンリーディングフレームの 3' に位置させることができる。したがって、宿主細胞のゲノムに核酸を標的化組込みするための方法であって、

（ a ）オープンリーディングフレームの 3' に組み込まれた外因性ランディングパッドを含む宿主細胞であって、外因性ランディングパッドは、上流ランディングパッドリンカー配列（ULPL）と下流ランディングパッド相同配列（DLP）との間に位置するヌクレアーゼ標的配列（NTS）を含む、宿主細胞を、

（ i ） 1 つまたは複数の最後のコンポーネントポリヌクレオチドであって、各最後のコンポーネントポリヌクレオチドは、5' から 3' 方向に、

30

（ 1 ）最後のリンカー配列；

（ 2 ）最後の目的の核酸；および

（ 3 ）（ DLP ）と相同組換えが可能な下流ライブラリー配列（ DL ）

を含む、最後のコンポーネントポリヌクレオチド；

（ i i ）（ NTS ）に結合し、外因性ランディングパッド内の部位を切断することが可能な、ヌクレアーゼ（ N ）

と接触させる工程、

（ c ）接触させた宿主細胞から生成した宿主細胞を回収する工程であって、

（ ULPL ）は、 1 つまたは複数の最後のコンポーネントポリヌクレオチドのいずれかの最後のリンカー配列と相同組換えして、外因性ランディングパッドで、 1 つまたは複数の最後のコンポーネントポリヌクレオチドからの最後のコンポーネントポリヌクレオチドを組み込むことが可能である、工程

40

を含む方法も本明細書で提供される。

#### 【 0 1 5 7 】

[00182] 図 7 E ~ 図 7 H に、オープンリーディングフレームの 3' に位置する標的化されたランディングパッドを利用する例示的な実施形態を例示する。図 7 E は、上流ランディングパッドリンカー配列（ ULPL 、 「 A 」 として示される ）と下流ランディングパッド相同配列（ DLP 、 「 DL 」 として示される ）との間に位置するヌクレアーゼ標的配列（ NTS 、 「 X 」 として示される ）を含む、ターミネータースワップランディングパッドビルディング

50



コンストラクトを例示する。ターミネータースワップランディングパッドビルディングコンストラクトは、その5'領域におけるオープンリーディングフレーム(「ORF」)の一部およびその3'領域におけるORFの下流配列(「DS\_ORF」)をさらに含む。これらの5'および3'領域における配列は、ORFのネイティブの遺伝子座へのターミネータースワップランディングパッドの宿主細胞媒介相同組換えに使用することができる。図7Eに示される実施形態において、ターミネータースワップランディングパッドビルディングコンストラクトは、(ULPL、「A」として示される)と(NTS、「X」として示される)との間に位置する一時的なターミネーターであるtACT1をさらに含む。宿主細胞は、ターミネータースワップランディングパッドビルディングコンストラクトおよびORFのネイティブのターミネーター(「tORF」として示される)を標的化しそれを切断することができるヌクレアーゼで形質転換される。図7Fに、それにより得られた宿主細胞のゲノム中の組み込まれたターミネータースワップランディングパッドを示す。宿主細胞のゲノム中でORFの3'に組み込まれたターミネータースワップランディングパッドを有するこの親宿主細胞は、標準ライブラリーリンカー配列および下流ライブラリー配列を有する最後のコンポーネントポリヌクレオチドの様々なライブラリーをスクリーニングするのに使用することができる。

10

#### 【0158】

[00183] ある特定の実施形態において、図7Gで示されるように、ターミネータースワップランディングパッドは、デグロンライブラリーを組み込むのに使用することができる。このような実施形態において、(ULPL、「A」として示される)に隣接するオープンリーディングフレームの3'末端は、ターミネータースワップランディングパッド中の「\*」と表記される位置に終止コドンを含まない。図7Gにおいて、デグロンライブラリーは、4つの最後のコンポーネントポリヌクレオチドを含み、各最後のコンポーネントポリヌクレオチドは、5'から3'方向に、最後のリンカー配列(「A」として示される)、最後の目的の核酸(ターミネーターに融合したデグロン配列)、および下流ライブラリー配列(DL)を含む。4つの最後のコンポーネントポリヌクレオチドの各々は、異なるデグロン配列(デグロン1、デグロン2、デグロン3、およびデグロン4)を含む。組み込まれたターミネータースワップランディングパッドを有する親宿主細胞は、(NTS)に結合してランディングパッド内の部位を切断することが可能なヌクレアーゼで形質転換される場合、デグロンライブラリー中の最後のコンポーネントポリヌクレオチドのいずれかが、ランディングパッドにおいて組み込まれたデグロン配列は、フレーム内でORFに融合され、このようなコンストラクトから発現された融合タンパク質は、融合タンパク質の分解をモジュレートすることができる。融合タンパク質の分解速度または安定性をモジュレートすることが望ましい場合、図7Gで示されるように、最後のコンポーネントポリヌクレオチドのデグロンライブラリーは、ターミネータースワップランディングパッドを使用してスクリーニングすることができる。

20

30

#### 【0159】

[00184] 他の実施形態において、ターミネータースワップランディングパッドは、オープンリーディングフレームの3'末端に終止コドンを含む。このようなターミネータースワップランディングパッドは、最後のコンポーネントポリヌクレオチドのターミネーターライブラリーを組み込むことにおいて有用である。ターミネーターは、RNAプロセッシングにおいて重要な役割を果たし、RNA半減期の変動性、最終的には遺伝子発現に寄与する。図7Hにおいて、ターミネーターライブラリーは、4つのコンポーネントポリヌクレオチドを含み、各最後のコンポーネントポリヌクレオチドは、5'から3'方向に、最後のリンカー(「A」として示される)、最後の目的の核酸(ターミネーター)、および下流ライブラリー配列(DL)を含む。4つのコンポーネントポリヌクレオチドの各々は、異なるターミネーター配列(term1、term2、term3、およびterm4)を含む。組み込まれたターミネータースワップランディングパッドを有する親宿主細胞が、(NTS)に結合しランディングパッド内の部位を切断することが可能なヌクレアーゼで形質転換される場合、ターミネーターライブラリー中の最後のコンポーネントポリヌクレオチドのいずれかが、タ

40

50

ーミネータースワップランディングパッドにおいて組込みを起こすことができる。RNA半減期、したがってオープンリーディングフレームからの遺伝子発現を改変することが望ましい場合、ターミネーターライブラリーのスクリーニングは特に有用である。

【0160】

[00185] 別の態様において、標的化されたランディングパッドに外因性ドナー核酸の標的化された組込みを組み合わせること、および外因性ドナー核酸のゲノム遺伝子座に頼らない別のランディングパッドへの組込みの方法が本明細書で提供される。例えば、宿主細胞は、図7に関して説明したような標的化されたランディングパッド、および1つまたは複数の二次ランディングパッドを含み、各二次ランディングパッドは、二次上部ランディングパッド相同配列と二次下流ランディングパッド相同配列との間に位置するヌクレアーゼ標的配列を含む。この実施形態において、二次上部ランディングパッド相同配列および二次下流ランディングパッド相同配列の少なくとも1つが、標的化されたランディング相同配列とは異なっている。

10

【0161】

[00186] ある特定の実施形態において、標的化されたランディングパッドへの外因性ドナー核酸の標的化された組込みを、標準化されたランディングパッド（例えば、図2Aに示される）へのランダムな外因性ドナー核酸の組込みと組み合わせて、宿主細胞の表現型に対するそれらのコンビナトリアルな作用を決定することができる。図8は、このような例を例示する。標的化されたランディングパッドとは対照的に、宿主細胞のゲノム中のニュートラルな遺伝子座における標準の外因性ランディングパッドは、ランディングパッドが配置されるゲノム部位のいかなるコンテキストも要せずに、そこに組み込まれた外因性ドナー核酸を発現させることを可能にする。それゆえに、ニュートラルな遺伝子座における外因性ランディングパッドは、ゲノム遺伝子座に頼らない組込みランディングパッドとも称される。

20

【0162】

[00187] 図8で例示された実施形態において、親宿主細胞は、2つの異なるランディングパッド、ERG9遺伝子座における標的化されたランディングパッド、およびゲノム遺伝子座に頼らない組込みランディングパッドである二次ランディングパッドを含む。ERG9遺伝子座における標的化されたランディングパッドは、ネイティブのERG9遺伝子の発現の適正量を決定するための様々なプロモーターの組込みを可能にする。iALG1遺伝子座の遺伝子間領域における二次ランディングパッドは、様々なプロモーターをファルネセンシンターゼ遺伝子（「FS」）とコンビナトリアル的に組み合わせるのに使用される。

30

【0163】

[00188] ネイティブのERG9は、様々な生物においてステロール生合成の第1の工程を触媒するスクアレニンターゼをコードする。ステロールの産生は、細胞バイオマス産生などの様々な細胞機能において重要である。しかしながら、細胞バイオマス産生は、発酵中の工業用微生物における標的分子産生に対して釣り合いをとることが必要である。図8に示される例示的な実施形態において、生合成経路（例えば、メバロン酸経路）は、ファルネセンシンターゼを介して望ましい標的分子（例えば、ファルネセン）に、またはERG9を介してバイオマスに変換できる中間体（例えば、FPPまたはファルネシルリニン酸）を生産する。標的分子（例えば、ファルネセン）の産生を増加させるために、ERG9の発現を減少させ、ファルネセンシンターゼ遺伝子を過剰発現することが望ましい。

40

【0164】

[00189] ERG9の発現とファルネセンシンターゼの最適な釣り合いを達成するために、2つの遺伝子の発現は、図8で示されるように、プロモーターライブラリーの異なるセットを使用して宿主細胞で同時に適正量を決定することができる。比較的弱いpSNL1プロモーターのライブラリーは、pSLN1ライブラリーの上流ランディングパッド相同配列（ULP、「UL<sub>1</sub>」として示される）および上流ライブラリー配列（UL<sub>1</sub>）の相同組換えを介して、ERG9遺伝子座で標的化されたプロモータースワップランディングパッド

50

に組み込むことができる。比較的強い p G A L プロモーターのライブラリーは、p G A L ライブラリーの上流ランディングパッド相同配列 (ULP、「UL」として示される) および上流ライブラリー配列 (UL) の相同組換えを介して、二次ランディングパッドに組み込むことができる。プロモーターライブラリーのこれらの 2 つのセットで同時に形質転換される宿主細胞をスクリーニングして、細胞バイオマス産生への必要性和釣り合いがとられた最適な標的分子の産生を有する株を同定することができる。ある特定の実施形態において、最適な株に組み込まれたプロモーターは、コンポーネントポリヌクレオチド中の各プロモーターに関連付けられたバーコード配列を使用して同定することができる。

#### 【 0 1 6 5 】

[00190] 別の態様において、外因性ドナー核酸の標的化された組込みのために宿主細胞のゲノムにおいて標準化されたランディングパッドを利用する方法が本明細書で提供される。一部の場合において、標的化されたランディングパッドを作り出す代わりに、特異的な遺伝子の標的化された組込みは、ランディングパッド相同配列に隣接する内因性ゲノム配列を使用して達成することができる。したがって、宿主細胞のゲノムへの外因性ドナー核酸の標的化された組込みの方法が本明細書において提供され、本方法は、

( a ) 宿主細胞のゲノム中に組み込まれた 1 つまたは複数の ( x ) 外因性ランディングパッドを含む宿主細胞であって、各外因性ランディングパッドは、上流ランディングパッド相同配列 (ULP) と下流ランディングパッド相同配列 (DLP) との間に位置するヌクレアーゼ標的配列 (NTS) を含み、各ランディングパッドは、上部の内因性ゲノム配列 (UEG)  $\times$  と下流内因性ゲノム配列 (DEG)  $\times$  とに入れ子になっている、宿主細胞を、

( i ) 1 つまたは複数の ( x ) 外因性ランディングパッドを切断することが可能なヌクレアーゼ (NTS) ; および

( i i ) 1 つまたは複数の第 1 のコンポーネントポリヌクレオチドであって、各第 1 のコンポーネントポリヌクレオチドは、5' から 3' 方向に、

( 1 ) 1 つまたは複数の外因性ランディングパッドのいずれかの任意の (ULP) と相同組換えが可能な上流ライブラリー配列 (UL) ;

( 2 ) 第 1 の目的の核酸 ;

( 3 ) 第 1 のリンカー配列

を含む、第 1 のコンポーネントポリヌクレオチド ; および

( i i i ) 1 つまたは複数の最後のコンポーネントポリヌクレオチドであって、各最後のコンポーネントポリヌクレオチドは、5' から 3' 方向に、

( 1 ) 最後のリンカー配列 ;

( 2 ) 最後の目的の核酸 ; および

( 3 ) 1 つまたは複数の ( x ) 外因性ランディングパッドの (DEG)  $\times$  と相同組換えが可能な下流ライブラリー配列 (DL)

を含む、最後のコンポーネントポリヌクレオチド

と接触させる工程、

( b ) 接触させた細胞から生成した宿主細胞を回収する工程

を含み、

1 つまたは複数の第 1 のコンポーネントポリヌクレオチドからの第 1 のコンポーネントポリヌクレオチドと、1 つまたは複数の最後のコンポーネントポリヌクレオチドからの最後のコンポーネントポリヌクレオチドとの任意の組合せが、インピボでそれらのリンカー配列を介して相同組換えされ、該組合せが、(DEG)  $\times$  を有する外因性ランディングパッドにおいて組み込まれ、 $\times$  は、少なくとも 1 の整数である。

#### 【 0 1 6 6 】

[00191] 図 9 に、例示的な実施形態を例示する。図 9 に示される親宿主細胞は、2 つのランディングパッドを含む。第 1 および第 2 のランディングパッドは両方とも、上流ランディングパッド相同配列 (UL) と下流ランディングパッド相同配列 (DL) との間に位置するヌクレアーゼ標的配列 (X) を含む標準化されたランディングパッドである。2 つのランディングパッドは、親宿主細胞において、それぞれ A L G 1 遺伝子座および Y C T 1

遺伝子座下流の遺伝子間領域中に操作により作製される。外因性ドナー核酸の標的化された組込みのために、親株は、以下：１）ヌクレアーゼ標的配列 X を切断するヌクレアーゼ；２）様々なプロモーターを含むプロモーターライブラリー（すなわち、様々なプロモーター強度を有するプロモーターを含む 9 つの第 1 のコンポーネントポリヌクレオチド）；３）常にランディングパッド 1 において組込みを起こすように設計される O R F 1 を含む最後のコンポーネントポリヌクレオチド；および 2 ）常にランディングパッド 2 において組込みを起こすように設計される O R F 2 を含む最後のコンポーネントポリヌクレオチドと接触させる。O R F 1 を含む最後のコンポーネントポリヌクレオチドは、i A L G 1 遺伝子座の下流内因性ゲノム配列（DEG）に同一な下流ライブラリー配列を含むことから、常にランディングパッド 1 において組込みを起こす。O R F 2 を含む最後のコンポーネントポリヌクレオチドは、i Y C T 1 遺伝子座の下流内因性ゲノム配列（DEG）に同一な下流ライブラリー配列を含むことから、ランディングパッド 2 で常に組込みを起こす。これらの O R F を含む外因性ドナー核酸は、コンビナトリアル的に組み合わせられ、プロモーターを含む第 1 のコンポーネントポリヌクレオチドのいずれかと組み込まれる。様々なプロモーターに作動可能に連結した O R F 1 および O R F 2 で形質転換した宿主細胞をスクリーニングして、望ましい表現型を有する宿主細胞を選択することによって、これらの O R F の最適に釣り合った共発現を決定することができる。ある特定の実施形態において、望ましい表現型を有する選択された宿主細胞は、各コンポーネントポリヌクレオチドに関連付けられたバーコード配列を使用して遺伝子型解析することができる。

10

【 0 1 6 7 】

20

[00192] 別の態様において、宿主細胞のゲノムへの外因性ドナー核酸の標的化された組込みの方法は、相同組換えのための上流ライブラリー配列として（UEG）を含む、第 1 のコンポーネントポリヌクレオチドのライブラリーを相同組換えするための上流内因性ゲノム配列（UEG）を利用する。本方法は、

（ a ）宿主細胞のゲノム中に組み込まれた 1 つまたは複数の（ x ）外因性ランディングパッドを含む宿主細胞であって、各外因性ランディングパッドは、上流ランディングパッド相同配列（ULP）と下流ランディングパッド相同配列（DLP）との間に位置するヌクレアーゼ標的配列（NTS）を含み、各ランディングパッドは、上部の内因性ゲノム配列（UEG）<sub>x</sub> と下流内因性ゲノム配列（DEG）<sub>x</sub> とに入れ子になっている、宿主細胞を、

（ i ） 1 つまたは複数の（ x ）外因性ランディングパッドを切断することが可能なヌクレアーゼ（NTS）；および

30

（ i i ） 1 つまたは複数の第 1 のコンポーネントポリヌクレオチドであって、各第 1 のコンポーネントポリヌクレオチドは、5' から 3' 方向に、

（ 1 ） 1 つまたは複数の（ x ）外因性ランディングパッドの（UEG）<sub>x</sub> と相同組換えが可能な上流ライブラリー配列（UL）；

（ 2 ）第 1 の目的の核酸；

（ 3 ）第 1 のリンカー配列

を含む、第 1 のコンポーネントポリヌクレオチド；および

（ i i i ） 1 つまたは複数の最後のコンポーネントポリヌクレオチドであって、各最後のコンポーネントポリヌクレオチドは、5' から 3' 方向に、

40

（ 1 ）最後のリンカー配列；

（ 2 ）最後の目的の核酸；および

（ 3 ） 1 つまたは複数の（ x ）外因性ランディングパッドのいずれかの任意の D L P と相同組換えが可能な下流ライブラリー配列（DL）

を含む、最後のコンポーネントポリヌクレオチド

と接触させる工程

（ b ）接触させた宿主細胞から生成した宿主細胞を回収する工程

を含み、

1 つまたは複数の第 1 のコンポーネントポリヌクレオチドからの第 1 のコンポーネントポリヌクレオチドと、1 つまたは複数の最後のコンポーネントポリヌクレオチドからの最後

50

のコンポーネントポリヌクレオチドとの任意の組合せが、インピボでそれらのリンカー配列を介して相同組換えされ、該組合せが、 $(UEG)_x$ を有する外因性ランディングパッドにおいて組み込まれ、 $x$ は、少なくとも1の整数である。

#### 【0168】

##### 5.5 代謝経路を操作するための方法

[00193] 本明細書に記載される方法および組成物は、例えば生物燃料、医薬品またはバイオマテリアルへのバイオマス変換に最適化された生合成経路を含む組換え生物を構築するための特定の利点を提供する。機能的な非ネイティブの生物学的経路が、抗マラリア薬であるアルテミシニン（例えば、Martin et al., Nat Biotechnol 21:796-802 (2003) 参照）脂肪酸由来の燃料および化学物質（例えば、脂肪エステル、脂肪族アルコールおよびワックス；例えば、Steen et al., Nature 463:559-562 (2010) 参照；ハロゲン化メチル由来の燃料および化学物質（例えば、Bayer et al., J Am Chem Soc 131:6508-6515 (2009) 参照）；コレステロール降下薬を生成するポリケチドシンターゼ（例えば、Ma et al., Science 326:589-592 (2009) 参照）；およびポリケチド（例えば、Kodumal, Proc Natl Acad Sci USA 101:15573-15578 (2004) 参照）の前駆体の産生のために微生物宿主においてうまく構築されてきた。

#### 【0169】

[00194] 従来、代謝を操作すること、特定には生合成経路の構築、または生合成経路を介した代謝フラックスの最適化は、1つずつの連続的な様式で進められてきたが、それにより経路コンポーネントが導入され、すなわち単一の遺伝子座に同時に、宿主細胞のゲノムに組み込まれてきた。本明細書で提供される組み込み方法を利用して、典型的に宿主細胞、例えば微生物細胞を、新しい代謝経路、すなわち宿主細胞によって内因的に生産されない代謝産物を産生する代謝経路の酵素をコードする1つまたは複数の異種ヌクレオチド配列を含むように操作するのに必要とされる時間を低減することができる。他の特定の実施形態において、本明細書で提供される組み込み方法は、宿主細胞にとって内因性の代謝経路、すなわち宿主細胞によって内因的に産生される代謝産物を産生する代謝経路の酵素をコードする1つまたは複数の異種ヌクレオチド配列を含むように宿主細胞を効率的に操作するのに使用することができる。他の特定の実施形態において、本明細書で提供される組み込み方法は、生成物を産生する代謝経路の酵素をコードする2つ以上の異種ヌクレオチド配列の発現の釣り合いをとることによって、代謝フラックスを最適化し、高い生成物のタイターを達成するのに使用することができる。例えば、図8に関して上述したように、標準化された相同配列を有するプロモーターライブラリーを利用することにより代謝フラックスを最適化して、バイオマスおよび目的の最終生成物の産生の釣り合いをとることができる。さらに、本方法は、1つの宿主細胞シャーシ中の複数のランディングパッドに組み込まれた最適化された経路コンポーネントを含むDNAアセンブリを、第2の宿主細胞シャーシ中の類似のランディングパッドに移すことを可能にする。望ましい遺伝子型を操作するのに必要なラウンド数を低減することによって、代謝経路を最適化する構築のペースが実質的に増加する。

#### 【0170】

##### 5.5.1 イソプレノイド経路の操作

[00195] 一部の実施形態において、本明細書で提供される方法は、操作された宿主細胞の生成物プロファイルを変更するために、生合成経路の1つまたは複数のコンポーネントを同時に導入する、置き換える、または適正量を決定するのに利用することができる。一部の実施形態において、生合成経路は、イソプレノイド経路である。

#### 【0171】

[00196] テルペンは、多くの生物で産生される炭化水素の大分類である。テルペンが化学的に改変されると（例えば、炭素骨格の酸化または転位を介して）、得られた化合物は、一般的にテルペノイドと称され、これはまた、イソプレノイドとしても公知である。イソプレノイドは、例えば、電子伝達鎖におけるキノンとして、タンパク質プレニル化を介した細胞内の標的化および制御における膜のコンポーネントとして、カロチノイド、クロ

10

20

30

40

50

ロフィルなどの光合成色素として、ホルモンおよび補因子として、ならびに様々なモノテルペン、セスキテルペン、およびジテルペンを含む植物防御化合物として、多くの重要な生物学的な役割を果たす。それらは、抗生物質、ホルモン、抗がん薬、防虫剤、および化学物質として工業的に有用である。

#### 【0172】

[00197] テルペンは、イソプレン単位 ( $C_5H_8$ ) を連結することによって誘導され、存在するイソプレン単位の数によって分類される。ヘミテルペンは、単一のイソプレン単位からなる。イソプレンそれ自体が、唯一のヘミテルペンと考えられている。モノテルペンの例は、ゲラニオール、リモネン、およびテルピネオールである。セスキテルペンは、3つのイソプレン単位で構成され、分子式  $C_{15}H_{24}$  を有する。セスキテルペンの例は、ファルネセンおよびファルネソールである。ジテルペンは、4つのイソプレン単位で作られており、分子式  $C_{20}H_{32}$  を有する。ジテルペンの例は、カフェストール、カーウエオール、センブレン、およびタキサジエンである。セスタテルペンは、5つのイソプレン単位で作られており、分子式  $C_{25}H_{40}$  を有する。セスタテルペンの例は、ゲラニルファルネソールである。トリテルペンは、6つのイソプレン単位からなり、分子式  $C_{30}H_{48}$  を有する。テトラテルペンは、8つのイソプレン単位を含有し、分子式  $C_{40}H_{64}$  を有する。生物学的に重要なテトラテルペンとしては、非環式リコピン、単環式ガンマ-カロテン、および二環式アルファおよびベータ-カロテンが挙げられる。ポリテルペンは、多くのイソプレン単位の長鎖からなる。天然ゴムは、二重結合が *cis* であるポリイソプレンからなる。

#### 【0173】

[00198] テルペンは、ピロリン酸イソペンテニル (イソペンテニルニリン酸またはIPP) およびその異性体であるジメチルアリルピロリン酸 (ジメチルアリルニリン酸またはDMAPP) の縮合を介して生合成される。IPPおよびDMAPPを生成する2つの経路、すなわち真核生物のメバロン酸依存性 (MEV) 経路 (図13)、および原核生物のメバロン酸非依存性またはデオキシキシルロース-5-リン酸 (DXP) 経路が公知である。植物は、MEV経路およびDXP経路の両方を使用する。順にIPPおよびDMAPPが、プレニルニリン酸シンターゼ (例えば、それぞれGPPシンターゼ、FPPシンターゼ、およびGGPPシンターゼ) の作用を介して、ポリプレニルニリン酸 (例えば、ゲラニルニリン酸またはGPP、ファルネシルニリン酸またはFPP、およびゲラニルゲラニルニリン酸またはGGPP) に縮合される。ポリプレニルニリン酸中間体は、テルペンシンターゼによってより複雑なイソプレノイド構造に変換される。

#### 【0174】

[00199] テルペンシンターゼは、複数の生成物を形成する大きい遺伝子ファミリーに組織化されている。テルペンシンターゼの例としては、GPPをモノテルペンに変換するモノテルペンシンターゼ; GGPPをジテルペンに変換するジテルペンシンターゼ; およびFPPをセスキテルペンに変換するセスキテルペンシンターゼが挙げられる。セスキテルペンシンターゼの例は、FPPをファルネセンに変換するファルネセンシンターゼである。テルペンシンターゼは、代謝の分岐点で作動し、細胞により産生されるイソプレノイドのタイプを指図するため、イソプレノイドへの経路フラックスの制御において重要である。さらに、テルペンシンターゼは、このようなテルペンの高収量の産生の鍵を握っている。したがって、異種イソプレノイド産生のために操作された宿主における経路フラックスを改善する1つの戦略は、テルペンシンターゼをコードする核酸の複数のコピーを導入することである。例えば、ファルネセンなどのセスキテルペンの産生が望ましいMEV経路を含む操作された微生物において、セスキテルペンシンターゼ、例えばファルネセンシンターゼが経路の末端酵素として利用され、ファルネセン産生に最適化された株の生成のために、ファルネセンシンターゼ遺伝子の複数のコピーが宿主細胞に導入されてもよい。

#### 【0175】

[00200] あらゆるイソプレノイドの生合成がプレニルニリン酸シンターゼおよびテルペ

ンシンターゼ上流の同じ経路コンポーネントを頼るため、これらの経路コンポーネントは、宿主「プラットフォーム」株中に一旦操作により作製されたら、あらゆるセスキテルペンの産生に利用することができ、セスキテルペンのアイデンティティは、宿主細胞に導入された特定のセスキテルペンシンターゼによって決定することができる。さらに、異なるイソプレニル単位を有するテルペンの産生、例えばセスキテルペンの代わりにモノテルペンの産生が望ましい場合、それでもなお経路の上流コンポーネントを利用しながら、異なるテルペンを産生するようにプレニルニリン酸シンターゼとテルペンシンターゼの両方を置き換えることができる。

#### 【0176】

[00201] したがって、本明細書で提供される方法および組成物は、望ましいイソプレノイドおよび/またはイソプレノイド産生レベルの強化がもたらされるように、イソプレノイド産生経路、例えばMEV経路を含む宿主細胞を効率的に改変するのに利用することができる。一部の実施形態において、宿主細胞は、MEV経路を含み、本明細書で提供される同時の複数の組込みの方法は、宿主細胞のテルペン生成物プロファイルを定義するためにプレニルニリン酸シンターゼおよび/またはテルペンシンターゼの複数のコピーを同時に導入するのに利用することができる。一部の実施形態において、プレニルニリン酸シンターゼは、GPPシンターゼであり、テルペンシンターゼは、モノテルペンシンターゼである。一部の実施形態において、プレニルニリン酸シンターゼは、FPPシンターゼであり、テルペンシンターゼは、セスキテルペンシンターゼである。一部の実施形態において、プレニルニリン酸シンターゼは、GGPPシンターゼであり、テルペンシンターゼは、ジテルペンシンターゼである。他の実施形態において、宿主細胞は、MEV経路、ならびに第1のタイプのテルペン、例えばファルネセンの産生のためのプレニルニリン酸シンターゼおよび/またはテルペンシンターゼを含み、本明細書で提供される同時の複数の組込みの方法は、第2のタイプのテルペン、例えばアモルファジエンを産生するためにプレニルニリン酸シンターゼおよび/またはテルペンシンターゼの1つまたは複数のコピーを同時に置き換えるのに利用することができる。本明細書で提供される方法は、経路コンポーネントの複数のコピーを利用するあらゆる生合成経路の構築および/または改変のために同様に利用することができ、単一の経路コンポーネントの複数のコピーの付加またはスワップによってその生成物プロファイルを容易に改変できる宿主細胞操作することに特に有用である。

#### 【0177】

5.5.2. 生合成経路のためのコンビナトリアルな組込みライブラリーを生成する方法  
[00202] 生合成経路が一旦構築されたら、代謝フラックスが最適化されて高い生成物のタイターが達成されるように、全てのコンポーネントの発現レベルを計画する必要がある。フラックスを最適化するための一般的なアプローチとしては、経路コンポーネント遺伝子のアイデンティティを変更すること、遺伝子のコドン最適化、溶解性タグの使用、短縮化または公知の突然変異の使用、および遺伝子の発現コンテキスト（すなわちプロモーターおよびターミネーターの選択）が挙げられる。従来の方法を使用した株構築の過程におけるこの変動性をサンプリングすることは、非現実的に多くの数の株の生成と保管を必要とする。例えば、株のエンジニアが3つの遺伝子座にコンストラクトを組み込むことを計画し、各遺伝子座につき10個のバリエーションを考案したとすれば、コンビナトリアルな多様性を完全にサンプリングするには1,000個の株の生成が必要になる。経路の遺伝子が協同して働いており、必ずしも全ての代謝産物の中間体を容易にスクリーニングできるとは限らないため、各組込みサイクル後に経路遺伝子の個々の寄与を評価することは、しばしば不可能である。したがって、株のエンジニアは、新規の代謝経路を構築するとき技術者がサンプリングする設計の余地を著しく制限する選択をすることが慣例になっている。

#### 【0178】

[00203] 最適な経路設計をよりよく同定するために、本明細書で提供されるゲノム改変方法は、外因性ドナー核酸のコンビナトリアルライブラリーを含む株を生成するのに利用

することができる。一部の実施形態において、本方法は、宿主細胞のゲノムを、1つまたは複数のヌクレアーゼおよび1つまたは複数のドナーDNAアセンブリ（例えば、コンポーネントポリヌクレオチド）と接触させて、ゲノム中の標準化されたランディングパッドにおけるドナーDNAの複数の同時組み込みを容易にすることに頼る。操作された株の多様性をもたらすために、本方法は、宿主株のコンビナトリアルな組み込みライブラリーが生成できるように、ドナーDNAのライブラリー、すなわち、1つまたは複数の標準化されたランディングパッドおよび/または標的化されたランディングパッドのための組み込みコンストラクトの混合物を共に形質転換することを含む。高い頻度の複数の組み込みの達成は、大規模なゲノムの品質管理を要せずとも直接的に生成物に関して得られた株を合理的にスクリーニングできること、さらに、スクリーニング後に、上位の株のアイデンティティを例えばシーケンシングによって決定できることを意味する。この方法は、個々の株生成、品質管理および保管の負担をなくし、技術者が1つの試験管中で多様な組み込み組合せを生成して、例えば経路の最終生成物に関するスクリーニングによって最良の性能を示した株を選別することを可能にする。

#### 【0179】

[00204] したがって、一部の実施形態において、本明細書で提供される宿主細胞のゲノムに複数の外因性核酸を組み込むための方法は、本明細書に記載される1つまたは複数のランディングパッドを含む宿主細胞を複数のライブラリーと接触させる工程を含み、各ライブラリー（L）は、複数の外因性核酸を含み、選択された外因性核酸は、5'から3'方向に、上流ライブラリー配列（UL）、（D）群から選択されるあらゆる目的の核酸、および下流ライブラリー配列（DL）を含み、ここで（UL）および（DL）は、宿主細胞のゲノム中の1つまたは複数のランディングパッドにおける選択された外因性核酸の宿主細胞媒介相同組換えを開始させることが可能であり；さらに（NTS）で切断することが可能なヌクレアーゼ（N）を含み、その結果、前記切断は、ランディングパッドの相同組換えが起こる。

#### 【0180】

[00205] 一部の実施形態において、各ライブラリー（L）は、共通の生合成経路の酵素をコードする外因性核酸を含む。一部の実施形態において、（D）群は、少なくとも10<sup>1</sup>、10<sup>2</sup>、10<sup>3</sup>、10<sup>4</sup>、10<sup>5</sup>、10<sup>6</sup>または10<sup>6</sup>より多くの固有な目的の核酸を含む。一部の実施形態において、各ライブラリー（L）は、生合成経路の酵素のバリエーションをコードする複数の外因性核酸を含む。用語「バリエーション」は、本明細書で使用される場合、選択された酵素と比較して異なるヌクレオチドまたはアミノ酸配列を有する生合成経路の酵素を指す。例えば、一部の実施形態において、ライブラリー（L）は、セスキテルペンシンターゼバリエーションを含み、選択されたセスキテルペンシンターゼの野生型バージョンと比較して、セスキテルペンシンターゼバリエーションは、ヌクレオチドの付加、欠失、および/または置換を含んでいてもよく、このような付加、欠失、および/または置換は、対応するアミノ酸配列に変化をもたらす場合もあるし、またはもたらさない場合もある。他の実施形態において、酵素バリエーションは、参照酵素、例えば野生型バージョンと比べてアミノ酸の付加、欠失および/または置換を含む。

#### 【0181】

[00206] 一部の実施形態において、宿主細胞は、前記接触の前に、生合成経路の1つまたは複数の酵素をコードする1つまたは複数の異種ヌクレオチド配列を含む。一部の実施形態において、生合成経路の1つまたは複数の酵素をコードする1つまたは複数の異種ヌクレオチド配列は、ランディングパッドにゲノムで組み込まれる。

#### 【0182】

#### 5.6ヌクレアーゼ

[00207] 本明細書に記載される方法の一部の実施形態において、宿主細胞のゲノムは、ランディングパッド中のヌクレアーゼ標的配列（NTS）で、またはそのすぐ近くで、二本鎖を切断すること、すなわち二本鎖破断を引き起こすことが可能な1つまたは複数のヌクレアーゼと接触させる。一部の実施形態において、二本鎖破断誘導物質は、特異的なポリ



ヌクレオチド認識配列を認識し、および/またはそれに結合して、認識配列で、またはその近くで破断を引き起こすあらゆる物質である。二本鎖破断誘導物質の例としては、これらに限定されないが、エンドヌクレアーゼ、部位特異的リコンビナーゼ、トランスポザーゼ、トポイソメラーゼ、およびジンクフィンガーヌクレアーゼが挙げられ、さらにそれらの改変された誘導体、バリエーション、および断片が挙げられる。

【0183】

[00208] 本明細書に記載される方法の一部の実施形態において、宿主細胞のゲノムは、ランディングパッド内の指定された領域で切断すること、すなわち破断を引き起こすことが可能な1つまたは複数のヌクレアーゼと接触させる。一部の実施形態において、破断は、ヌクレアーゼ標的配列のDNA鎖の両方ではなく一方が切断される(すなわち、「ニックが入られる」)一本鎖破断である。一部の実施形態において、破断は、二本鎖破断である。一部の実施形態において、破断誘導物質は、特異的なポリヌクレオチド認識配列を認識し、および/またはそれに結合して、認識配列で、またはその近くで破断を引き起こすあらゆる物質である。破断誘導物質の例としては、これらに限定されないが、エンドヌクレアーゼ、部位特異的リコンビナーゼ、トランスポザーゼ、トポイソメラーゼ、およびジンクフィンガーヌクレアーゼが挙げられ、さらにそれらの改変された誘導体、バリエーション、および断片が挙げられる。

10

【0184】

[00209] 一部の実施形態において、1つまたは複数のヌクレアーゼの各々は、選択されたヌクレアーゼ標的配列(NTS)内の指定された領域で破断を引き起こすことが可能である。一部の実施形態において、ヌクレアーゼは、ランディングパッド中の(NTS)の5'領域と3'領域との間に位置する領域で破断を引き起こすことが可能である。他の実施形態において、ヌクレアーゼは、(NTS)の5'および3'領域の上流または下流に位置する領域で破断を引き起こすことが可能である。

20

【0185】

[00210] ヌクレアーゼ標的配列(NTS)は、破断誘導物質を特異的に認識し、および/またはそれと結合するあらゆるポリヌクレオチド配列である認識配列を含む。認識部位配列の長さは、様々であってもよく、例えば、少なくとも10、12、14、16、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ヌクレオチドまたはそれより長い長さを有する配列が挙げられる。

30

【0186】

[00211] 一部の実施形態において、認識配列は、パリンドロームであり、すなわち1つの鎖の配列が相補鎖で逆方向で同じに読める。一部の実施形態において、ニック/切断部位は、認識配列内である。他の実施形態において、ニック/切断部位は、認識配列の外側である。一部の実施形態において、切断は、平滑末端を生じる。他の実施形態において、切断は、一本鎖オーバーハング、すなわち「付着末端」を生じ、これは、5'オーバーハング、または3'オーバーハングのいずれであってもよい。

40

【0187】

[00212] 一部の実施形態において、ランディングパッド内の認識配列は、宿主細胞のゲノムにとって内因性であってもよいし、または外因性であってもよい。認識部位が内因性配列である場合、認識部位は、天然に存在する、またはネイティブの破断誘導物質によって認識される認識配列であり得る。代替として、内因性認識部位は、内因性認識配列を特異的に認識して破断をもたらすように設計または選択された、改変または操作された破断誘導物質によって認識され、および/またはそれと結合することができる。一部の実施形態において、改変された破断誘導物質は、ネイティブの天然に存在する破断誘導物質から誘導される。他の実施形態において、改変された破断誘導物質は、人工的に作り出されるかまたは合成される。このような改変または操作された破断誘導物質を選択するための方

50

法は、当業界において公知である。例えば、タンパク質のアミノ酸配列バリエーションは、DNA中の突然変異によって調製することができる。変異誘発およびヌクレオチド配列の変更のための方法としては、例えば、Kunkel, (1985) *Proc Natl Acad Sci USA* 82:488-92; Kunkel, et al., (1987) *Meth Enzymol* 154:367-82; 米国特許第 4, 873, 192 号; Walker and Gaastra, eds. (1983) *Techniques in Molecular Biology* (Mac Millan Publishing Company, New York) およびそこで引用された文献が挙げられる。タンパク質の生物活性に影響を与える可能性が低いアミノ酸置換に関する指針は、例えば、Dayhoff, et al. (1978) *Atlas of Protein Sequence and Structure* (Natl Biomed Res Found, Washington, D.C.) のモデルで見出される。保存的置換、例えば 1 つのアミノ酸を類似の特性を有する別のアミノ酸でスワップすることが好ましい場合がある。保存的な欠失、挿入、およびアミノ酸置換は、タンパク質の特徴における大幅な変化を引き起こすことが予期されず、あらゆる置換、欠失、挿入、またはそれらの組合せの作用は、慣例的なスクリーニングアッセイによって評価することができる。二本鎖破断を誘導する活性に関するアッセイが公知であり、一般的には、認識部位を含有する DNA 基質に対するそのような物質の全体的な活性および特異性を測定するものである。

#### 【0188】

5.6.1. クラスター化された規則的に間隔があいた短いパ lindロームの反復 (CRISPR; Clustered Regulatory Interspaced Short Palindromic Repeats)

[00213] 本明細書で提供される方法の一部の実施形態において、ヌクレアーゼは、CRISPR/Cas 由来の RNA ガイドエンドヌクレアーゼである。CRISPR は、タイプ II の原核性 CRISPR (クラスター化された規則的に散在した短いパ lindロームの反復) 適応免疫系に基づくゲノム編集ツールである。真正細菌および古細菌における CRISPR 系は、外から侵入してきた DNA を標的化し切断するのに小型 RNA および CRISPR 関連 (Cas) エンドヌクレアーゼを使用する。Bhaya et al., *Annu Rev Genet* 45:273-297 (2011); Terns et al., *Curr Opin Microbiol* 14(3):321-327 (2011); および Wiedenheft et al., *Nature* 482(7385):331-338 を参照されたい。細菌において、CRISPR 遺伝子座は、スペーサーと呼ばれる外因性 DNA (長さ約 30 bp) のセグメントによって分離した一連の反復で構成されている。反復 - スペーサーアレイは長い前駆体として転写され、反復配列内でプロセシングされて、CRISPR ニュクレアーゼによって切断された標的配列を特定する小型 crRNA (プロトスペーサーとしても公知) を生成する。次いで CRISPR スペーサーを使用して、RNA または DNA レベルで外因性遺伝学的エレメントを認識しサイレンシングする。切断には、プロトスペーサー隣接モチーフ (PAM) として公知の標的領域の 3' 末端のすぐ下流にある配列モチーフが必須である。PAM は、標的 DNA 中に存在するが、それを標的化する crRNA には存在しない。

#### 【0189】

[00214] 最も簡単な CRISPR 系の 1 つは、化膿性連鎖球菌 (*Streptococcus pyogenes*) 由来のタイプ II CRISPR 系である。CRISPR 関連 Cas9 エンドヌクレアーゼおよび 2 つの小型 RNA、標的相補的 CRISPR RNA (crRNA); およびトランス作用性 crRNA (tracrRNA) が、外来 DNA の RNA ガイド切断にとって十分である。Cas9 タンパク質は、タイプ II CRISPR - Cas 系の特徴的なタンパク質であり、これは、RuvC および HNH ニュクレアーゼに相同な 2 つのヌクレアーゼドメインを含有する大きい単量体 DNA ニュクレアーゼである。Cas9 は、PAM (プロトスペーサー隣接モチーフ) 配列モチーフに隣接する DNA 標的配列に crRNA : tracrRNA 複合体によってガイドされる。tracrRNA への成熟 crRNA 塩基対は、Cas9 を標的 DNA に方向付ける 2 つの RNA 構造を形成する。crRNA ガイド配列に相補的な部位で、Cas9 HNH ニュクレアーゼドメインは相補鎖を切断し、それに対して Cas9 RuvC 様ドメインは非相補鎖を切断し、結果として標的 DNA において二本鎖破断が生じる。例えば、Deltcheva et al., *Nature* 47(7340):602-607 (2011) を参照されたい。

#### 【0190】

10

20

30

40

50

[00215] 近年の研究によれば、C a s 9をガイドしてインビトロで部位特異的DNA二本鎖破断を導入するためのゲノム編集ツールとして、c r R N A : t r a c r R N A複合体を模擬する単一のガイドRNA (gRNA) キメラが、C a s 9と共に利用できることが示されている。標的ゲノム内の切断の特異性は、ネイティブのc r R N A : t r a c r R N A複合体を模擬するキメラガイドRNA分子 (gRNA) のスパーサー様の部分によって決定される。したがって、機能的なC R I S P R / C a s系における最小限の数のコンポーネントは、2つ、すなわちC a s 9およびs g R N Aである。その5'末端に配置されたs g R N Aガイド配列は、DNA標的特異性を付与する。それゆえに、ガイド配列を改変することによって、異なる標的特異性を有するs g R N Aを作り出すことが可能になる。ガイド配列の標準的な長さは、20bpである。結果として、DNA標的も、20bpとそれに続くコンセンサスNGGであり、それに続いてPAM配列が存在する。この改変されたC R I S P R系の使用は、インビトロで(例えば、Jinek et al., Science 337(6096):816-821 (2012)参照)、哺乳類細胞株において(例えば、Mali et al., Science 339(6121):823-826 (2013), Jinek et al., Elife 2:e00417 (2013); Cong et al., Science 339(6121):819-823 (2013); and Cho et al., Nat Biotechnol 31(3):230-232 (2013) 参照)、細菌において(例えば、Jiang et al., Nat Biotechnol 31(3):233-239 (2013); and Gasiunas et al., Proc Natl Acad Sci USA 109(39):E2579-E2586. (2012) 参照)、酵母において(例えば、DiCarlo et al., Nucleic Acid Res 41(7):4336-4343 (2013)参照)、ゼブラフィッシュにおいて(例えば、Hwang et al., Nat Biotechnol 31(3):227-229 (2013); およびChang et al., Cell Res 23(4):465-472 (2013) 参照)、マウスにおいて(例えば、Wang et al., Cell 153(4):910-918 (2013) 参照) および植物において(例えば、Belhaj et al., Plant Methods 9:39 (2013) 参照) 実証されている。

#### 【0191】

[00216] C a s 9ヌクレアーゼは、(1)異種宿主内での発現を増加させるためのコドン最適化;(2)適した区画化のための核局在化シグナル(NLS)への融合;および(3)ヌクレアーゼを鎖特異的ニッカーゼに変換するためのHNHまたはRuvCドメインのいずれかの部位特異的変異誘発によって改変することができる。RuvCまたはHNH-モチーフのいずれかにおけるC a s 9の部位特異的変異誘発は鎖切断特異性を示したことから、野生型エンドヌクレアーゼに加えて2つの鎖に特異的なニッカーゼが提供され、DNAの標的化された一本鎖破断を可能にする。例えば、Jinek et al., Science 337(6096):816-821 (2012)およびGasiunas et al., Proc Natl Acad Sci USA 109(39):E2579-E2586. (2012)を参照されたい。ジンクフィンガーヌクレアーゼおよびTALENに関して報告されているように、1つの鎖のみを破断するニッカーゼとして機能するようにヌクレアーゼを改変することは、オフターゲット切断による毒性を低減し、さらに非HRメカニズム、例えば、NHEJを介した破断修復速度も低くすることができる。例えば、Jinek et al., Science 337(6096):816-821 (2012)参照。

#### 【0192】

[00217] 当業界において公知のあらゆるC R I S P R / C a s系は、本明細書で提供される方法および組成物におけるヌクレアーゼとして用途がある。高度に多様なC R I S P R - C a s系は、3つの主要なタイプに類別され、これはさらにコアエレメント含量および配列に基づき10のサブタイプに分割される(例えば、Makarova et al., Nat Rev Microbiol 9:467-77 (2011)参照)。外来核酸のc r R N A媒介サイレンシングに關与する核タンパク質複合体の構造的な構成および機能は、個々のC R I S P R / C a sタイプ間で異なっている(Wiedenheft et al., Nature 482:331-338 (2012)参照)。タイプI-E系において、大腸菌(Escherichia coli)によって例示されるように、c r R N Aは、カスケードと呼ばれるマルチサブユニットエフェクター複合体(抗ウイルス性防御のためのC R I S P R関連複合体)に取り込まれ(Brouns et al., Science 321:960-4 (2008))、これが標的DNAに結合して、シグネチャーC a s 3タンパク質による分解を引き起こす(Sinkunas et al., EMBO J 30:1335-1342 (2011); Beloglazova et al., EM

BO J 30:616-27 (2011))。スルフォロブス・ソルファタリカス (*Sulfolobus solfataricus*) およびパイロコッカス・フリオサス (*Pyrococcus furiosus*) のタイプ I I I の C R I S P R / C a s 系において、C a s R A M P モジュール (Cmr) および c r R N A 複合体は、インビトロで合成 R N A を認識してそれを切断するが (Hale et al., Mol Cell 45:292-302 (2012); Zhang et al., Mol Cell, 45:303-13 (2012))、スタフィロコッカス・エピデルミディス (*Staphylococcus epidermidis*) の C R I S P R / C a s 系は、インビボで D N A を標的化する (Marraffini & Sontheimer, Science. 322:1843-5 (2008))。タイプ I I の C R I S P R / C a s 系による、より具体的にはストレプトコッカス・サーモフィラス (*Streptococcus thermophilus*) D G C C 7 7 1 0 の C R I S P R 3 / C a s 系における D N A サイレンシングに關与する R N P 複合体 (Horvath & Barrangou, Science 327:167-70 (2010)) は、12 個の反復 - スペーサー単位の上流に配置される 4 つの c a s 遺伝子: c a s 9、c a s 1、c a s 2、および c s n 2 からなる。C a s 9 (以前の名称は cas5 または csn1) は、タイプ I I 系のシグネチャー遺伝子である (Makarova et al., Nat Rev Microbiol 9:467-77 (2011))。

【0193】

[00218] 本明細書で提供される方法および組成物において用途がある C R I S P R 系としてはさらに、その内容が本明細書にその全体として取り込まれる国際公報番号 W O 2 0 1 3 / 1 4 2 5 7 8 A 1 および W O 2 0 1 3 / 0 9 8 2 4 4 A 1 に記載されるものも挙げられる。

【0194】

5.6.2. 転写活性化因子様エフェクターヌクレアーゼ (TALEN)

[00219] 本明細書で提供される方法の一部の実施形態において、ヌクレアーゼの 1 種または複数は、T A L - エフェクター D N A 結合ドメイン - ヌクレアーゼ融合タンパク質 (TALEN) である。キサントモナス属における植物病原菌の T A L エフェクターは、宿主 D N A と結合してエフェクター特異的な宿主遺伝子を活性化することによって、疾患において重要な役割を果たすか、または防御を引き起こす。例えば、Gu et al. (2005) Nature 435:1122-5; Yang et al., (2006) Proc. Natl. Acad. Sci. USA 103:10503-8; Kay et al., (2007) Science 318:648-51; Sugio et al., (2007) Proc. Natl. Acad. Sci. USA 104:10720-5; Romer et al., (2007) Science 318:645-8; Boch et al., (2009) Science 326(5959):1509-12; および Moscou and Bogdanove, (2009) 326(5959):1501 参照。T A L エフェクターは、1 つまたは複数のタンデムリピートドメインを介して配列特異的な方式で D N A と相互作用する D N A 結合ドメインを含む。繰り返しの配列は、典型的には 3 4 アミノ酸を含み、リピートは、典型的には互いに 9 1 ~ 1 0 0 % 相同である。リピートの多型は通常、1 2 および 1 3 位に配置され、1 2 および 1 3 位における反復可変二残基 (repeat variable-diresidue) のアイデンティティと、T A L - エフェクターの標的配列における隣接するヌクレオチドのアイデンティティとが 1 対 1 で対応すると考えられる。

【0195】

[00220] T A L - エフェクター D N A 結合ドメインは、望ましい標的配列に結合し、例えばタイプ I I 制限エンドヌクレアーゼからのヌクレアーゼドメイン、典型的には F o k I などのタイプ I I 制限エンドヌクレアーゼからの非特異的切断ドメインに融合するように操作することができる (例えば、Kim et al. (1996) Proc. Natl. Acad. Sci. USA 93:1156-1160 参照)。他の有用なエンドヌクレアーゼとしては、例えば、H h a I、H i n d I I I、N o d、B b v C I、E c o R I、B g l I および A l w I が挙げられる。したがって、好ましい実施形態において、T A L E N は、組み合わされて、T A L E N が特異的なヌクレオチド配列内またはそれに隣接する標的 D N A を切断するように標的 D N A 配列中の特異的なヌクレオチド配列に結合する複数の T A L エフェクターリピート配列を含む T A L エフェクタードメインを含む。本明細書で提供される方法に有用な T A L E N S としては、W O 1 0 / 0 7 9 4 3 0 および米国特許公開公報第 2 0 1 1 / 0 1 4 5 9 4 0 号に記載されるものが挙げられる。

10

20

30

40

50

## 【 0 1 9 6 】

[00221] 一部の実施形態において、標的DNA内の特異的なヌクレオチド配列に結合するTALエフェクタードメインは、10個以上のDNA結合リピート、好ましくは15個以上のDNA結合リピートを含んでいてもよい。一部の実施形態において、各DNA結合リピートは、標的DNA配列中の塩基対の認識を決定する反復可変二残基(RVD)を含み、この場合、各DNA結合リピートは、標的DNA配列中の1つの塩基対を認識することに関与し、RVDは、Cを認識するためのHD; Tを認識するためのNG; Aを認識するためのNI; GまたはAを認識するためのNN; AまたはCまたはGまたはTを認識するためのNS; CまたはTを認識するためのN<sup>\*</sup>、ここで、<sup>\*</sup>は、RVDの第2の位置におけるギャップである; Tを認識するためのHG; Tを認識するためのH<sup>\*</sup>、ここで、<sup>\*</sup>は、RVDの第2の位置におけるギャップである; Tを認識するためのIG; Gを認識するためのNK; Cを認識するためのHA; Cを認識するためのND; Cを認識するためのHI; Gを認識するためのHN; Gを認識するためのNA; GまたはAを認識するためのSN; およびTを認識するためのYGの1つまたは複数を含む。本明細書で提供される方法の一部の実施形態において、ヌクレアーゼの1種または複数は、部位特異的リコンビナーゼである。部位特異的リコンビナーゼは、リコンビナーゼとも称され、これは、その適合性を有する組換え部位間の保存的部位特異的組換えを触媒するポリペプチドであり、ネイティブのポリペプチドに加えて、活性を保持する誘導体、バリエーションおよび/または断片、ならびにネイティブのポリヌクレオチド、活性を保持するリコンビナーゼをコードする誘導体、バリエーションおよび/または断片を包含する。部位特異的リコンビナーゼおよびその認識部位の総論に関しては、Sauer(1994)Curr Op Biotechnol 5:521-7;およびSadowski、(1993)FASEB 7:760-7を参照されたい。一部の実施形態において、リコンビナーゼは、セリンリコンビナーゼまたはチロシンリコンビナーゼである。一部の実施形態において、リコンビナーゼは、インテグラーゼまたはレゾルパーゼファミリー由来である。一部の実施形態において、リコンビナーゼは、FLP、Cre、ラムダインテグラーゼ、およびRからなる群から選択されるインテグラーゼである。インテグラーゼファミリーの他のメンバーに関しては、例えば、Esposito, et al., (1997) Nucleic Acids Res 25:3605-14およびAbremski, et al., (1992) Protein Eng 5:87-91を参照されたい。反応速度、補因子相互作用および必要条件、発現、最適条件、ならびに/または認識部位特異性、ならびにリコンビナーゼおよびバリエーションの活性に関するスクリーニングを改変するための方法は公知であり、例えば、Miller, et al., (1980) Cell 20:721-9; Lange-Gustafson and Nash, (1984) J Biol Chem 259:12724-32; Christ, et al., (1998) J Mol Biol 288:825-36; Lorbach, et al., (2000) J Mol Biol 296:1175-81; Vergunst, et al., (2000) Science 290:979-82; Dorgai, et al., (1995) J Mol Biol 252:178-88; Dorgai, et al., (1998) J Mol Biol 277:1059-70; Yagu, et al., (1995) J Mol Biol 252:163-7; Scrimente, et al., (2001) Nucleic Acids Res 29:5044-51; Santoro and Schultze, (2002) Proc Natl Acad Sci USA 99:4185-90; Buchholz and Stewart, (2001) Nat Biotechnol 19:1047-52; Voziyanov, et al., (2002) Nucleic Acids Res 30:1656-63; Voziyanov, et al., (2003) J Mol Biol 326:65-76; Klippel, et al., (1988) EMBO J 7:3983-9; Arnold, et al., (1999) EMBO J 18:1407-14; WO 03 / 08045; WO 99 / 25840; およびWO 99 / 25841を参照されたい。認識部位は、約30ヌクレオチドの最小の部位から数百ヌクレオチドまでの範囲である。天然に存在する部位、およびバリエーションなどのリコンビナーゼのためのあらゆる認識部位を使用することができる。バリエーション認識部位は公知であり、例えば、Hoess, et al., (1986) Nucleic Acids Res 14:2287-300; Albert, et al., (1995) Plant J 7:649-59; Thomson, et al., (2003) Genesis 36:162-7; Huang, et al., (1991) Nucleic Acids Res 19:443-8; Siebler and Bode, (1997) Biochemistry 36:1740-7; Schlake and Bode, (1994) Biochemistry 33:12746-51; Thygarajan, et al., (2001) Mol Cell Biol 21:3926-34; Umlauf and Cox, (1988) EMBO J 7:1845-52; Lee and Saito, (1998) Gene 216:55-65; WO 01 / 23545; WO 99 / 25821; WO 99 / 25851; WO 01 / 11058; WO

10

20

30

40

50

01/07572および米国特許第5,888,732号を参照されたい。

#### 【0197】

[00222] 本明細書で提供される方法の一部の実施形態において、ヌクレアーゼの1種または複数は、トランスポザゼである。トランスポザゼは、ゲノム中の1つの配置から別の配置へのトランスポゾンの転移を媒介するポリペプチドである。トランスポザゼは、典型的には、二本鎖破断を誘導して、トランスポゾンを切り出し、末端近くのリピートを認識し、切り出されトランスポゾンの末端を一緒にするが、一部の系では、転移中に末端を一緒にするのに他のタンパク質も必要である。トランスポゾンおよびトランスポザゼの例としては、これらに限定されないが、トウモロコシ由来のAc/Ds、Dt/rDt、Mu-M1/MnおよびSpm(En)/dSpmエレメント、キンギョソウ由来のTamエレメント、バクテリオファージ由来のMuトランスポゾン、細菌トランスポゾン(Tn)および挿入配列(IS)、酵母のTyエレメント(レトロトランスポゾン)、シロイヌナズナ由来のTa1エレメント(レトロトランスポゾン)、ショウジョウバエ由来のPエレメントトランスポゾン(Gloor, et al., (1991) Science 253:1110-1117)、ショウジョウバエ由来のCopial、MarinerおよびMinosエレメント、イエバエ由来のHermesエレメント、イラクサギンウワバ(Trichoplusia ni)由来のPiggyBackエレメント、C.エレガンス(C. elegans)由来のTc1エレメント、およびマウス由来のIAPエレメント(レトロトランスポゾン)が挙げられる。

10

#### 【0198】

5.6.3. ジンクフィンガーヌクレアーゼ(ZFN)

20

[00223] 本明細書で提供される方法の一部の実施形態において、ヌクレアーゼの1つまたは複数は、ジンクフィンガーヌクレアーゼ(ZFN)である。ZFNは、ジンクフィンガー-DNA結合ドメインおよび破断誘導物質ドメインで構成される操作された破断誘導物質である。操作されたZFNは、2つのジンクフィンガーアレイ(ZFA)からなり、それらの各々は、二量体化されると活性になるFokI酵素由来のヌクレアーゼドメインなどの非特異的なエンドヌクレアーゼの単一のサブユニットに融合している。典型的には、単一のZFAは、3または4つのジンクフィンガードメインからなり、それらの各々は、特異的なヌクレオチドの三つ組み(GGC、GATなど)を認識するように設計されている。したがって、2つの「3フィンガー」ZFAで構成されるZFNは、18塩基対の標的部位を認識することが可能である。18塩基対の認識配列は、一般的に、ヒトや植物のゲノムなどの大きいゲノム内であっても固有である。2つのFokIヌクレアーゼ単量体の共局在化および二量体化を方向付けることによって、ZFNは、標的化された遺伝子座におけるDNAで破断を作り出す機能的な部位特異的なエンドヌクレアーゼを生成する。

30

#### 【0199】

[00224] 有用なジンクフィンガーヌクレアーゼとしては、公知のもの、および本明細書に記載される1種または複数のヌクレアーゼ標的配列(NTS)に特異性を有するように操作されたものが挙げられる。ジンクフィンガードメインは、例えば宿主細胞のゲノムの標的部位内の選択されたポリヌクレオチド認識配列に特異的に結合するポリペプチドを設計するのに適している。ZFNは、非特異的なエンドヌクレアーゼドメイン、例えばHOまたはFokIなどのタイプIIエンドヌクレアーゼ由来のヌクレアーゼドメインに連結された、操作されたDNA結合ジンクフィンガードメインからなる。代替として、操作されたジンクフィンガー-DNA結合ドメインは、DNAニック導入/切断活性を保持する他の破断誘導物質またはその誘導体に融合していてもよい。例えば、このタイプの融合は、破断誘導物質を異なる標的部位に方向付ける、ニックまたは切断部位の配置を変更する、誘導物質をより短い標的部位に方向付ける、または誘導物質をより長い標的部位に方向付けるのに使用することができる。一部の例において、ジンクフィンガー-DNA結合ドメインは、部位特異的リコンビナーゼ、トランスポザゼ、またはDNAニック導入および/または切断活性を保持するその誘導体に融合していてもよい。追加の官能基が、ジンクフィンガー結合ドメイン、例えば転写活性化因子ドメイン、転写リプレッサードメイン、およびメチラーゼなどに融合していてもよい。一部の実施形態において、ヌクレアーゼドメイ

40

50

ンの二量体化が、切断活性に必要である。

#### 【 0 2 0 0 】

[00225] 各ジンクフィンガーは、標的DNA中の3つの連続した塩基対を認識する。例えば、3フィンガードメインは、ヌクレアーゼの二量体化の要件を満たした場合に9つの隣接したヌクレオチドの配列を認識したが、18ヌクレオチドの認識配列と結合するには、2セットのジンクフィンガー三つ組みが使用される。有用なデザイナージンクフィンガーモジュールとしては、様々なGNNおよびANN三つ組みを認識するもの (Dreier, et al., (2001) J Biol Chem 276:29466-78; Dreier, et al., (2000) J Mol Biol 303:489-502; Liu, et al., (2002) J Biol Chem 277:3850-6), as well as those that recognize various CNN or TNN triplets (Dreier, et al., (2005) J Biol Chem 280:35588-97; Jamieson, et al., (2003) Nature Rev Drug Discov 2:361-8). See also, Durai, et al., (2005) Nucleic Acids Res 33:5978-90; Segal, (2002) Methods 26:76-83; Porteus and Carroll, (2005) Nat Biotechnol 23:967-73; Pabo, et al., (2001) Ann Rev Biochem 70:313-40; Wolfe, et al., (2000) Ann Rev Biophys Biomol Struct 29:183-212; Segal and Barbas, (2001) Curr Opin Biotechnol 12:632-7; Segal, et al., (2003) Biochemistry 42:2137-48; Beerli and Barbas, (2002) Nat Biotechnol 20:135-41; Carroll, et al., (2006) Nature Protocols 1:1329; Ordiz, et al., (2002) Proc Natl Acad Sci USA 99:13290-5; Guan, et al., (2002) Proc Natl Acad Sci USA 99:13296-301; WO 2 0 0 2 0 9 9 0 8 4 ; WO 0 0 / 4 2 2 1 9 ; WO 0 2 / 4 2 4 5 9 ; WO 2 0 0 3 0 6 2 4 5 5 ; US 2 0 0 3 0 0 5 9 7 6 7 ; 米国特許出願公開番号 2 0 0 3 / 0 1 0 8 8 8 0 ; 米国特許第 6 , 1 4 0 , 4 6 6 号、第 6 , 5 1 1 , 8 0 8 号および第 6 , 4 5 3 , 2 4 2 号が挙げられる。有用なジンクフィンガーヌクレアーゼとしてはまた、WO 0 3 / 0 8 0 8 0 9 ; WO 0 5 / 0 1 4 7 9 1 ; WO 0 5 / 0 8 4 1 9 0 ; WO 0 8 / 0 2 1 2 0 7 ; WO 0 9 / 0 4 2 1 8 6 ; WO 0 9 / 0 5 4 9 8 5 ; および WO 1 0 / 0 6 5 1 2 3 に記載されるものも挙げられる。

#### 【 0 2 0 1 】

##### 5 . 6 . 4 . エンドヌクレアーゼ

[00226] 本明細書で提供される方法の一部の実施形態において、ヌクレアーゼの1つまたは複数は、エンドヌクレアーゼである。エンドヌクレアーゼは、ポリヌクレオチド鎖内のホスホジエステル結合を切断する酵素であり、塩基を損傷することなく特異的な部位としてDNAを切断する制限エンドヌクレアーゼが挙げられる。制限エンドヌクレアーゼは、タイプI、タイプII、タイプIII、およびタイプIVエンドヌクレアーゼを包含し、これらはさらに、サブタイプを包含する。制限エンドヌクレアーゼは、例えばREBASEデータベース (rebase.neb.comのウェブページ; Roberts, et al., (2003) Nucleic Acids Res 31:418-20), Roberts, et al., (2003) Nucleic Acids Res 31:1805-12, and Belfort, et al., (2002) in Mobile DNA II, pp. 761-783, Eds. Craigie, et al., ASM Press, Washington, D.C.でさらに説明され分類されている。

#### 【 0 2 0 2 】

[00227] 本明細書で使用される場合、エンドヌクレアーゼはまた、制限エンドヌクレアーゼと同様に、特異的な認識配列と結合し切断するホーミングエンドヌクレアーゼも包含する。しかしながら、ホーミングエンドヌクレアーゼの認識部位は、典型的にはより長く、例えば約18bpまたはそれより長い。ホーミングエンドヌクレアーゼは、メガヌクレアーゼとしても公知であり、保存された配列モチーフに基づき以下のファミリー: LAGLIDADG (配列番号: 1) ホーミングエンドヌクレアーゼ、HNHホーミングエンドヌクレアーゼ、His-Cysボックスホーミングエンドヌクレアーゼ、GIY-YIG (配列番号2) ホーミングエンドヌクレアーゼ、およびシアノバクテリアのホーミングエンドヌクレアーゼに分類されている。例えば、Stoddard, Quarterly Review of Biophysics 38(1):49-95(2006)を参照されたい。これらのファミリーは、それらの保存されたヌクレアーゼ活性部位のコアモチーフおよび触媒メカニズム、生物学的およびゲノム分布、ならびに非ホーミングヌクレアーゼ系とのより広い関係において大きく異なっている。

例えば、Guhan and Muniyappa (2003) Crit Rev Biochem Mol Biol 38:199-248; Lucas, et al., (2001) Nucleic Acids Res 29:960-9; Jurica and Stoddard, (1999) Cell Mol Life Sci 55:1304-26; Stoddard, (2006) Q Rev Biophys 38:49-95;およびMoure, et al., (2002) Nat Struct Biol 9:764を参照されたい。これらのファミリー由来の有用な特異的ホーミングエンドヌクレアーゼの例としては、これらに限定されないが、I - C r e I (Rochaix et al., Nucleic Acids Res. 13: 975-984 (1985)参照)、I - M s o I (Lucas et al., Nucleic Acids Res. 29: 960-969 (2001)参照)、I - S c e I (Foury et al., FEBS Lett. 440: 325-331 (1998)参照)、I - S c e I V (Moran et al., Nucleic Acids Res. 20: 4069-4076 (1992)参照)、H - D r e I (Chevalier et al., Mol. Cell 10: 895-905 (2002)参照)、I - H m u I (Goodrich-Blair et al., Cell 63: 417-424 (1990); Goodrich-Blair et al., Cell 84: 211-221 (1996)参照)、I - P p o I (Muscarella et al., Mol. Cell. Biol. 10: 3386-3396 (1990)参照)、I - D i r I (Johansen et al., Cell 76: 725-734 (1994); Johansen, Nucleic Acids Res. 21: 4405 (1993)参照)、I - N j a I (Elde et al., Eur. J. Biochem. 259: 281-288 (1999); De Jonckheere et al., J. Eukaryot. Microbiol. 41: 457-463 (1994)参照)、I - N a n I (Elde et al., S. Eur. J. Biochem. 259: 281-288 (1999); De Jonckheere et al., J. Eukaryot. Microbiol. 41: 457-463 (1994)参照)、I - N i t I (De Jonckheere et al., J. Eukaryot. Microbiol. 41: 457-463 (1994); Elde et al., Eur. J. Biochem. 259: 281-288 (1999)参照)、I - T e v I (Chu et al., Cell 45: 157-166 (1986)参照)、I - T e v I I (Tomaschewski et al., Nucleic Acids Res. 15: 3632-3633 (1987)参照)、I - T e v I I I (Eddy et al., Genes Dev. 5: 1032-1041 (1991)参照)、F - T e v I (Fujisawa et al., Nucleic Acids Res. 13: 7473-7481 (1985)参照)、F - T e v I I (Kadyrov et al., Dokl. Biochem. 339: 145-147 (1994); Kaliman, Nucleic Acids Res. 18: 4277 (1990)参照)、F - C p h I (Zeng et al., Curr. Biol. 19: 218-222 (2009)参照)、P I - M g a I (Saves et al., Nucleic Acids Res. 29:4310-4318 (2001)参照)、I - C s m I (Colleaux et al., Mol. Gen. Genet. 223:288-296 (1990)参照)、I - C e u I (Turmel et al., J. Mol. Biol. 218: 293-311 (1991)参照)およびP I - S c e I (Hirata et al., J. Biol. Chem. 265: 6726-6733 (1990)参照)が挙げられる。

### 【0203】

[00228] 本明細書に記載される方法の一部の実施形態において、ホーミングエンドヌクレアーゼの天然に存在するバリエーションおよび/または操作された誘導体を使用される。反応速度、補因子相互作用、発現、最適条件、および/または認識部位特異性、ならびに活性にするスクリーニングを改変するための方法は公知であり、例えば、Epinat, et al., (2003) Nucleic Acids Res 31:2952-62; Chevalier, et al., (2002) Mol Cell 10:895-905; Gimble, et al., (2003) Mol Biol 334:993-1008; Seligman, et al., (2002) Nucleic Acids Res 30:3870-9; Sussman, et al., (2004) J Mol Biol 342:31-41; Rosen, et al., (2006) Nucleic Acids Res 34:4791-800; Chames, et al., (2005) Nucleic Acids Res 33:e178; Smith, et al., (2006) Nucleic Acids Res 34:e149; Gruen, et al., (2002) Nucleic Acids Res 30:e29; Chen and Zhao, (2005) Nucleic Acids Res 33:e154; WO 2 0 0 5 1 0 5 9 8 9 ; WO 2 0 0 3 0 7 8 6 1 9 ; WO 2 0 0 6 0 9 7 8 5 4 ; WO 2 0 0 6 0 9 7 8 5 3 ; WO 2 0 0 6 0 9 7 7 8 4 ; および WO 2 0 0 4 0 3 1 3 4 6 参照。有用なホーミングエンドヌクレアーゼは、WO 0 4 / 0 6 7 7 3 6 ; WO 0 4 / 0 6 7 7 5 3 ; WO 0 6 / 0 9 7 7 8 4 ; WO 0 6 / 0 9 7 8 5 3 ; WO 0 6 / 0 9 7 8 5 4 ; WO 0 7 / 0 3 4 2 6 2 ; WO 0 7 / 0 4 9 0 9 5 ; WO 0 7 / 0 4 9 1 5 6 ; WO 0 7 / 0 5 7 7 8 1 ; WO 0 7 / 0 6 0 4 9 5 ; WO 0 8 / 1 5 2 5 2 4 ; WO 0 9 / 0 0 1 1 5 9 ; WO 0 9 / 0 9 5 7 4 2 ; WO 0 9 / 0 9 5 7 9 3 ; WO 1 0 / 0 0 1 1 8 9 ; WO 1 0 / 0 1 5 8 9 9 ; および WO 1 0 / 0 4 6 7 8 6 に記載のものも含む。



## 【0204】

[00229] 二本鎖破断誘導物質としてあらゆるホーミングエンドヌクレアーゼを使用することができ、これらに限定されないが、H - Dre I、I - Sce I、I - Sce II、I - Sce III、I - Sce IV、I - Sce V、I - Sce VI、I - Sce VII、I - Ceu I、I - CeuA I I P、I - Cre I、I - Creps b I P、I - Creps b I I P、I - Creps b I I I P、I - Creps b I V P、I - Tli I、I - Ppo I、Pi - Psp I、F - Sce I、F - Sce II、F - Suv I、F - Cph I、F - Tev I、F - Tev II、I - Ama I、I - Ani I、I - Chu I、I - Cmo e I、I - Cpa I、I - Cpa I I、I - Csm I、I - Cvu I、I - CvuA I P、I - Ddi I、I - Ddi I I、I - Dir I、I - Dmo I、I - Hmu I、I - Hmu I I、I - HsN I P、I - Lla I、I - Mso I、I - Naa I、I - Nan I、I - Ncl I P、I - Ngr I P、I - Nit I、I - Nja I、I - Nsp 236 I P、I - Pak I、I - Pbo I P、I - Pcu I P、I - PcuA I、I - PcuV I、I - Pgr I P、I - Pob I P、I - Por I、I - Por I I P、I - Pbp I P、I - SpBeta I P、I - Sca I、I - Sex I P、I - Sne I P、I - Spom I、I - SpomCP、I - Spom I P、I - Spom I I P、I - Squ I P、I - Ssp 68031、I - SthPhiJP、I - SthPhiST3P、I - SthPhiSTe3bP、I - Tde I P、I - Tev I、I - Tev II、I - Tev I I I、I - UarAP、I - UarHGPA I P、I - UarHGPA13P、I - Vin I P、I - Zbi I P、PI - Mga I、PI - Mtu I、PI - MtuHI P、PI - MtuHI I P、PI - Pfu I、PI - Pfu I I、PI - Pko I、PI - Pko I I、PI - Rma 43812 I P、PI - SpBeta I P、PI - Sce I、PI - Tfu I、PI - Tfu I I、PI - Thy I、PI - Tli I または PI - Tli I I またはそれらのあらゆるバリエーションもしくは誘導体などが挙げられる。

10

20

## 【0205】

[00230] 一部の実施形態において、エンドヌクレアーゼは、ネイティブのまたは内因性の認識配列と結合する。他の実施形態において、エンドヌクレアーゼは、非ネイティブのまたは外因性の認識配列と結合し、ネイティブのまたは内因性の認識配列と結合しない改変されたエンドヌクレアーゼである。

## 【0206】

30

## 5.6.5. ランディングパッド中のヌクレアーゼ標的配列

[00231] 本明細書で提供される方法において、ランディングパッド中のヌクレアーゼ標的配列の近くまたはその内で二本鎖破断を引き起こすことが可能なヌクレアーゼが、宿主細胞に導入されると、それにより切断部位で、またはその近くで相同組換えの頻度が大きく増加する。好ましい実施形態において、ヌクレアーゼの認識配列を含むヌクレアーゼ標的配列 (NTS) は、ランディングパッドでのみ宿主細胞のゲノム中に存在し、それによってヌクレアーゼによるあらゆるオフターゲットのゲノム結合および切断が最小化される。

## 【0207】

[00232] 利用しようとするヌクレアーゼがCRISPR/Cas由来のRNAガイドエンドヌクレアーゼである場合、最適なヌクレアーゼ標的配列は、使用される特定のCRISPR-Casエンドヌクレアーゼの標的認識の必要条件に従って選択してもよい。例えばCas9標的認識は、標的DNA中の「プロトスペーサー」配列とcrRNA中の残りのスペーサー配列との間で相補性が検出されたときに起こる。Cas9は、3'末端に正しいプロトスペーサー隣接モチーフ (PAM) も存在する場合だけDNAを切断する。異なるタイプII系は、異なるPAM必要条件を有する。S. ピオゲネス (S. pyogenes) 系は、NGG配列を必要とし、式中Nはあらゆるヌクレオチドであり得る。S. サーマフィラストタイプII系は、NGGNGおよびNNAGAAWをそれぞれ必要とするが、異なるS. ミュータンス系はNGGまたはNAARを許容する。遺伝子情報科学分析により、新しいPAMを同定し、CRISPR標的可能な配列のセットを拡張するのに役立つ可能性がある、様々な細菌におけるCRISPR遺伝子座の大規模データベースが生成した。例え

40

50

ば、Rho et al., PLoS Genet. 8, e1002441 (2012); and D. T. Pride et al., Genome Res. 21, 126 (2011)参照。S. サーモフィラスにおいて、Cas9は、プロトスペーサーの3bp上流に平滑末端化された二本鎖破断を生成するが、これは、Cas9タンパク質中の2つの触媒ドメイン：DNAの相補鎖を切断するHNHドメインおよび非相補鎖を切断するRuvC様ドメインにより媒介されるプロセスである。

#### 【0208】

[00233] 利用しようとするヌクレアーゼがジンクフィンガーヌクレアーゼである場合、最適なヌクレアーゼ標的配列は、多数の公共的に利用可能なオンラインリソースを使用して選択してもよい。例えば、その全体が参照により組み入れられるReyon et al., BMC Genomics 12:83(2011)を参照されたい。例えば、オリゴマー化プール工学 (Oligomerized Pool Engineering; OPEN) は、高い特異性およびインピボにおける官能性を有するジンクフィンガーアレイを操作するための高度にロバストな公共的に利用可能なプロトコールであり、植物、ゼブラフィッシュ、およびヒトの体細胞および多能性幹細胞において効率的に機能するZFNを生成するのにうまく使用されてきた。OPENは、事前に構築された候補ZFAのランダム化されたプールをスクリーニングして、望ましい標的配列に高い親和性および特異性を有するものを同定する選択ベースの方法である。ZFN Genomeは、OPENによって生成したZFNの可能性のある標的部位を同定し可視化するためのGBrowseベースのツールである。ZFN Genomeは、モデル生物のシーケンシングされ注釈が付けられたゲノムにおける可能性のあるZFN標的部位の要約を提供する。現在のところZFN Genomeは、7つのモデル生物；S. セレビスエ、コナミドリムシ (C. reinhardtii)、シロイスナズナ (A. thaliana)、キイロショウジョウバエ (D. melanogaster)、ゼブラフィッシュ (D. rerio)、C. エレガンス、およびホモサピエンス (H. sapiens) の完全にシーケンシングされたゲノム内にマッピングされた合計で11.60万種を超える可能性のあるZFN標的部位を包含する。近い将来、追加のモデル生物、3つの植物種、Glycine max (ダイズ)、Oryza sativa (イネ)、Zea mays (トウモロコシ)、および3つの動物種、Tribolium castaneum (コクヌストモドキ)、Mus musculus (マウス)、Rattus norvegicus (ドブネズミ) などが追加されると予想される。ZFN Genomeは、転写開始部位に対するその染色体の配置および位置などの各々の可能性のあるZFN標的部位に関する情報を提供する。使用者は、数々の異なる基準 (例えば、遺伝子番号、転写番号、標的部位の配列) を使用してZFN Genomeに問い合わせることができる。

#### 【0209】

[00234] 利用しようとするヌクレアーゼがTALEN-エフェクターヌクレアーゼである場合、一部の実施形態において、最適なヌクレアーゼ標的配列は、その全体が参照により組み入れられるSanjana et al., Nature Protocols, 7:171-192(2012)によって記載された方法に従って選択してもよい。簡単に言えば、TALENは、二量体として機能し、左および右TALENと称される一対のTALENはDNAの逆の鎖における配列を標的化する。TALENは、TALEN DNA-結合ドメインおよび単量体のFokI触媒ドメインの融合体として操作により作製される。FokIの二量体化を容易にするために、左および右TALEN標的部位が、およそ14~20塩基の間隔をあけて選ばれる。それゆえに、一対のTALENの場合、各々の標的化される20bpの配列、すなわち最適な標的部位は、5'-TN19N14-20N19A-3'の形態を有することになり、ここで左TALENは5'-TN19-3'を標的化し、右TALENは5'-N19A-3'のアンチセンス鎖 (N=A, G, TまたはC) を標的化する。

#### 【0210】

[00235] 本明細書で提供される方法の他の実施形態において、ヌクレアーゼ標的配列は、宿主細胞にとって外因性である。一部の実施形態において、同じヌクレアーゼ標的配列の複数のコピーが、ランディングパッド中に操作により作製され、それによってヌクレアーゼ標的配列を特異的に認識する単一のヌクレアーゼのみを使用した同時の複数の組み

10

20

30

40

50

事象が容易になる。他の実施形態において、複数の異なるヌクレアーゼ標的配列は、異なるランディングパッドで、宿主細胞のゲノム中に操作により作製される。一部の実施形態において、操作により作製されたランディングパッドは、他の状況では宿主細胞のネイティブのゲノム中に提示されない標的ヌクレアーゼ配列を含む。例えば、ホーミングエンドヌクレアーゼは、通常イントロンまたはインティン中に埋め込まれている大きい認識部位（12～40bp）を標的化し、したがって、それらの認識部位は、極めてまれであり、これらの部位は哺乳類サイズのゲノム中に全く存在しないかまたはわずかしかな存在しない。したがって、一部の実施形態において、外因性ヌクレアーゼ標的配列は、ホーミングエンドヌクレアーゼの認識配列である。一部の実施形態において、ホーミングヌクレアーゼは、H-DreI、I-SceI、I-SceII、I-SceIII、I-SceIV、I-SceV、I-SceVI、I-SceVII、I-CeuI、I-CeuAII、I-CreI、I-CrepsbIP、I-CrepsbIIP、I-CrepsbIIIP、I-CrepsbIVP、I-TliI、I-PpoI、Pi-PspI、F-SceI、F-SceII、F-SuvI、F-CphI、F-TevI、F-TevII、I-AmaI、I-AniI、I-ChuI、I-CmoEI、I-CpaI、I-CpaII、I-CsmI、I-CvuI、I-CvuAIP、I-DdiI、I-DdiII、I-DirI、I-DmoI、I-HmuI、I-HmuII、I-HsNI、I-LlaI、I-MsoI、I-NaaI、I-NanI、I-NclIP、I-NgrIP、I-NitI、I-NjaI、I-Nsp236IP、I-PakI、I-PboIP、I-PcuIP、I-PcuAI、I-PcuVI、I-PgrIP、I-PobIP、I-PorI、I-PorIIP、I-PbpIP、I-SpBetaIP、I-ScaI、I-SexIP、I-SneIP、I-SpomI、I-SpomCP、I-SpomIP、I-SpomIIP、I-SquIP、I-Ssp68031、I-SthPhiJP、I-SthPhiST3P、I-SthPhiSTe3bP、I-TdeIP、I-TevI、I-TevII、I-TevIII、I-UarAP、I-UarHGPAIP、I-UarHGPA13P、I-VinIP、I-ZbiIP、PI-MgaI、PI-MtuI、PI-MtuHIP、PI-MtuHIIP、PI-PfuI、PI-PfuII、PI-PkoI、PI-PkoII、PI-Rma43812IP、PI-SpBetaIP、PI-SceI、PI-TfuI、PI-TfuII、PI-ThyI、PI-TliIもしくはPI-TliII、またはそれらのあらゆるバリエーションもしくは誘導体からなる群から選択される。特定の実施形態において、外因性ゲノムヌクレアーゼ標的配列は、I-SceI、VDE（PI-SceI）、F-CphI、PI-MgaIまたはPI-MtuIIの認識配列であり、これらの各々を以下に提供する。

【0211】

10

20

30

40

50

## 【表 1】

表 1: 選択されたホーミングエンドヌクレアーゼの認識および切断部位

| ヌクレアーゼ        | 認識配列  |
|---------------|---|
| I-SceI        | TAGGGATAACAGGGTAAT (配列番号 3)                       |
| VDE (PI-SceI) | TATGTCGGGTGCGGAGAAAGAGGTAATGAAA (配列番号 4)          |
| F-CphI        | GATGCACGAGCGCAACGCTCACAA (配列番号 5)                 |
| PI-MgaI       | GCGTAGCTGCCCAGTATGAGTCAG (配列番号 6)                 |
| PI-MtuII      | ACGTGCACTACGTAGAGGGTCGCACCGCACCGATCTACAA (配列番号 7) |

10

## 【0212】

## 5.6.6. 送達

[00236] 一部の実施形態において、本明細書に記載される方法に有用な 1 つまたは複数のヌクレアーゼは、精製されたタンパク質として宿主細胞に提供される、例えば送達される。他の実施形態において、1 つまたは複数のヌクレアーゼは、ヌクレアーゼをコードする核酸を含むポリヌクレオチドを介して提供される。他の実施形態において、1 つまたは複数のヌクレアーゼは、宿主の細胞核中で直接翻訳することができる精製された RNA として宿主細胞に導入される。

20

## 【0213】

[00237] ある特定の実施形態において、組込みポリヌクレオチド、上述したようにヌクレアーゼもしくは精製されたヌクレアーゼタンパク質をコードするポリヌクレオチド、またはそれらの任意の組合せが、当業界において公知の細胞に外因性タンパク質および/または核酸を導入するためのあらゆる従来の技術を使用して宿主細胞に導入することができる。このような方法としては、これらに限定されないが、溶液からの細胞による分子の直接的な取り込み、または例えば、リポソームまたはイムノリポソームを使用したりポフェクションを介して促進された取り込み；粒子媒介トランスフェクションなどが挙げられる。例えば、米国特許第 5,272,065 号；Goeddel et al., eds, 1990, Methods in Enzymology, vol. 185, Academic Press, Inc., CA；Krieger, 1990, Gene Transfer and Expression--A Laboratory Manual, Stockton Press, NY；Sambrook et al., 1989, Molecular Cloning--A Laboratory Manual, Cold Spring Harbor Laboratory, NY；および Ausubel et al., eds., Current Edition, Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley Interscience, NY を参照されたい。細胞を形質転換するための特定の方法は当業界において周知である。Hinnen et al., Proc. Natl. Acad. Sci. USA 75:1292-3 (1978)；Cregg et al., Mol. Cell. Biol. 5:3376-3385 (1985) を参照されたい。例示的な技術としては、これらに限定されないが、スフェロプラスチ化、エレクトロポレーション、PEG 1000 媒介形質転換、および酢酸リチウムまたは塩化リチウム媒介形質転換が挙げられる。

30

40

## 【0214】

[00238] 一部の実施形態において、宿主細胞、特定には、それ以外の状況で従来の技術を使用して形質転換/トランスフェクトすることが難しい宿主細胞、例えば植物に、組込みポリヌクレオチド、ヌクレアーゼ、精製されたヌクレアーゼタンパク質をコードするポリヌクレオチド、またはそれらの任意の組合せを導入するのに、遺伝子銃が利用される。遺伝子銃は、形質転換反応物を顕微鏡レベルの金粒子に結合させ、次いで圧縮ガスを使用

50

して標的細胞に粒子を押し出すことによって作用する。

【0215】

[00239] 一部の実施形態において、ヌクレアーゼをコードする核酸を含むポリヌクレオチドは、宿主細胞内でヌクレアーゼの発現を可能にする発現ベクターである。好適な発現ベクターとしては、これらに限定されないが、大腸菌、酵母、または哺乳類細胞における遺伝子発現での使用に関して公知のものが挙げられる。大腸菌発現ベクターの例としては、これらに限定されないが、pSCM525、pDIC73、pSCM351およびpSCM353が挙げられる。酵母発現ベクターの例としては、これらに限定されないが、pPEX7およびpPEX408が挙げられる。好適な発現ベクターの他の例としては、CEN・ARS配列および酵母選択可能マーカーを含むシャトルベクターの酵母 - 大腸菌のpRSシリーズ；および2 μプラスミドが挙げられる。一部の実施形態において、ヌクレアーゼをコードするポリヌクレオチドは、天然に存在するポリヌクレオチド配列と比較して宿主細胞中でより高い使用頻度を有するコドンで置換されるように改変されていてもよい。例えばヌクレアーゼをコードするポリヌクレオチドは、天然に存在するポリヌクレオチド配列と比較してS・セレピシエ中でより高い使用頻度を有するコドンで置換されるように改変されていてもよい。

10

【0216】

[00240] ヌクレアーゼが各単量体の別々の発現を必要とするヘテロ二量体として機能する一部の実施形態において、ジンクフィンガーヌクレアーゼおよびTAL-エフェクターヌクレアーゼの場合と同様に、ヘテロ二量体の各単量体は、同じ発現プラスミドから発現されてもよいし、または異なるプラスミドから発現されてもよい。異なる標的部位で二本鎖破断を行うために複数のヌクレアーゼが細胞に導入される実施形態において、ヌクレアーゼは、単一のプラスミドにコードされてもよいし、または別個のプラスミドにコードされてもよい。

20

【0217】

[00241] ある特定の実施形態において、ヌクレアーゼ発現ベクターは、発現ベクターを含む宿主細胞の選択を可能にする選択可能マーカーをさらに含む。このような選択は、十分な量のヌクレアーゼの発現が、例えば、12、24、36、48、60、72、84、96時間、または96時間より長い期間にわたり起こるのに必要な期間、宿主細胞中にベクターを保持するために有用である場合があり、その後、宿主細胞を、発現ベクターが保持されなくなる条件下で増殖させてもよい。ある特定の実施形態において、選択可能マーカーは、URA3、ハイグロマイシンBホスホトランスフェラーゼ、アミノグリコシドホスホトランスフェラーゼ、ゼオシン耐性、およびホスフィノトリシンN-アセチルトランスフェラーゼからなる群から選択される。一部の実施形態において、ヌクレアーゼ発現ベクターは、1つまたは複数のドナー核酸分子の組み込み後に発現ベクターを含有しない宿主細胞の選択を可能にする対抗選択可能なマーカー(counter-selectable marker)を含んでいてもよい。使用されるヌクレアーゼ発現ベクターはまた、選択マーカーを有さない一過性のベクターであってもよいし、または選択されないベクターである。特定の実施形態において、一過性のヌクレアーゼ発現ベクターを含む宿主細胞の後代は、時間経過に伴いベクターを失う。

30

40

【0218】

[00242] ある特定の実施形態において、発現ベクターは、ヌクレアーゼをコードするヌクレオチド配列に作動可能に連結した転写終結配列およびプロモーターをさらに含む。一部の実施形態において、プロモーターは、構成的プロモーターである。一部の実施形態において、プロモーターは、誘導性プロモーターである。酵母細胞で使用するのに好適なプロモーターの説明に役立つ例としては、これらに限定されないが、K・ラクティス(K.lactis)のTEF1遺伝子のプロモーター、サッカロマイセス・セレピシエのPGK1遺伝子のプロモーター、サッカロマイセス・セレピシエのTDH3遺伝子のプロモーター、抑制性プロモーター、例えば、サッカロマイセス・セレピシエのCTR3遺伝子のプロモーター、および誘導性プロモーター、例えば、サッカロマイセス・セレピシエのガラクトー

50

ス誘導性プロモーター（例えば、G A L 1、G A L 7 および G A L 1 0 遺伝子のプロモーター）が挙げられる。

#### 【 0 2 1 9 】

[00243] 一部の実施形態において、核局在化配列（NLS）を含む追加のヌクレオチド配列が、ヌクレアーゼをコードするヌクレオチド配列の 5' に連結される。NLS は、より大きいヌクレアーゼ（> 25 kD）の核局在化を容易にすることができる。一部の実施形態において、核局在化配列は、SV40 核局在化配列である。一部の実施形態において、核局在化配列は、酵母核局在化配列である。

#### 【 0 2 2 0 】

[00244] ヌクレアーゼ発現ベクターは、当業者にとって明白なあらゆる技術によって作製することができる。ある特定の実施形態において、ベクターは、当業界において周知のポリメラーゼ連鎖反応（PCR）および分子クローニング技術を使用して作製される。例えば、PCR Technology: Principles and Applications for DNA Amplification, ed. H A Erlich, Stockton Press, New York, N.Y. (1989); Sambrook et al., 2001, Molecular Cloning - A Laboratory Manual, 3rd edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 参照。

#### 【 0 2 2 1 】

##### 5.7 キット

[00245] 別の態様において、本明細書に記載される 1 つまたは複数の外因性核酸をゲノムで組み込むための方法を実行するのに有用なキットが本明細書で提供される。一部の実施形態において、キットは、

（a）宿主細胞のゲノム中に組み込まれた複数の外因性ランディングパッドを含む宿主細胞であって、各外因性ランディングパッドは、上流ランディングパッド相同配列（ULP）と下流ランディングパッド相同配列（DLP）との間に位置するヌクレアーゼ標的配列（NTS）を含む、宿主細胞；

（b）複数の外因性ドナー核酸であって、各外因性ドナー核酸（ES）は、上流ライブラリー配列（UL）と下流ライブラリー配列（DL）との間に位置する目的の核酸（D）を含み、各（UL）は、宿主細胞のゲノム中のランディングパッドの任意の上流ランディングパッド相同配列（ULP）で相同組換えが可能であり、各（DL）は、任意の下流ランディングパッド相同配列（DLP）で相同組換えが可能である、外因性ドナー核酸；および

（c）ランディングパッド中の任意の（NTS）を切断することが可能な、1 つまたは複数のヌクレアーゼ（N）を含む。

#### 【 0 2 2 2 】

[00246] 一部の実施形態において、（D）は、選択可能マーカー、プロモーター、エピトープタグをコードする核酸配列、目的の遺伝子、レポーター遺伝子、終止コドンにコードする核酸配列からなる群から選択される。一部の実施形態において、キットは、複数のプライマー対をさらに含み、各プライマー対は、PCR によって外因性ドナー核酸の組込みを同定することが可能である。

#### 【 0 2 2 3 】

[00247] 一部の実施形態において、キットは、宿主細胞のいずれかのランディングパッドに 1 つまたは複数の外因性ドナー核酸を組み込む方法を説明する使用説明書をさらに含む。

#### 【 実施例 】

#### 【 0 2 2 4 】

##### 6.1 実施例 1：CLiX 親株の構築および形質転換

[00248] CLiX（X-カッターを使用したコンビナトリアルライブラリー組込み）親株を生成するために、エンドヌクレアーゼ認識部位（F-CphI）の端にある「上流ライブラリー」（UL）および「下流ライブラリー」（DL）配列からなる DNA「ランディングパッド」を、サッカロマイセス・セレビスエ CEN. PK 2 の 3 つの別個のゲノム遺伝子座

10

20

30

40

50

に組み込んだ。具体的には、ランディングパッドを、 $ALG1(iALG1)$ 、 $MGA1(iMGA1)$ および $YCT1(iYCT1)$ の下流の遺伝子間領域に標的化した。これらの遺伝子座の各々におけるDNA組込みが、株表現型に測定可能な影響がないことを事前に確認された。ULおよびDL配列を、500bpの長さおよび50%のGC含量の制約でランダムな配列ジェネレーターから誘導した。これらの配列を手作業で選定して、ヌクレオチドおよび翻訳されたBLASTサーチ(それぞれBLASTnおよびBLASTx)に基づき、タンDEMおよび逆方向反復配列、DNA2次構造、および公知の生物学的な配列に対するあらゆる有意な相同性を除去した。遺伝子座の標的配列を包含する各ランディングパッドのヌクレオチド配列は、配列番号8~10として提供される。

#### 【0225】

[00249] 各ランディングパッドを、最適化した酢酸リチウム(LiAc)形質転換体で、標準的な分子生物学技術を用いてCEN.PK2に組み込んだ。簡単に言えば、30で振盪しながら(200rpm)酵母抽出物ペプトンデキストロース(YPD)培地中で細胞を一晩増殖し、YPD100mL中0.1のOD600に希釈し、0.6~0.8のOD600に増殖させた。各形質転換体につき、培養物5mLを遠心分離によって回収し、滅菌水5mL中で洗浄し、再度遠沈させ、100mMのLiAc 1mLに再懸濁し、マイクロ遠沈管に移した。細胞を30秒間遠沈させ(13,000×g)上清を除去し、細胞を、50%PEG 240μL、1MのLiAc 36μL、加熱処理済みサケ精子DNA 10μL、およびドナーDNA 74μLからなる形質転換ミックス中に再懸濁した。42で40分間のヒートショック後、細胞をYPD培地中で一晩回復させ、その後、選択培地上で平板培養した。ランディングパッド組込みを、5'および3'組込み端部を標的化するプライマーを用いたコロニーPCRによって確認した。

#### 【0226】

6.2 実施例2:「ロックドプロモーター」および「スプリットプロモーター」設計のゲノム組込みの比較

[00250] 多数の設計バリエーションが、CLiX組込みに適合性を有する。「ロックドプロモーター」設計は、ULおよびDL相同性を端部に有する個々のオープンリーディングフレーム(ORF)を駆動する固定されたプロモーターを含有する。「スプリット」設計は、小さい相同な「リンカー」配列において組換えを介して組み合わされた複数のパーツからなる。最も一般的には、1つのパーツは、UL相同配列、それに続くプロモーターおよびDNAリンカーからなる「スプリットプロモーター」であり、第2のパーツは、DNAリンカー、オープンリーディングフレーム、およびDL相同配列からなるプロモーターレスORFである。この設計バリエーションを使用して、複数の「スプリットプロモーター」のプールされた形質転換体は、コンビナトリアル的に各ORFの発現の適正量を決定するのに使用することができる。本発明者らは、スパーサー配列を含有する「スパーサー」コンストラクトと共に蛍光タンパク質(GFP、RFP、およびBFP)をコードするDNAコンストラクトを使用して、ロックドおよびスプリットプロモーター設計の両方を試験した。

#### 【0227】

[00251] ロックドプロモーター設計のために、各蛍光性遺伝子をpGAL1プロモーターによって転写し、ULおよびDL相同配列を端部に付して、各ランディングパッドでの組込みを可能にした。「スパーサー」コンストラクトは、ULおよびDL相同性を端部に有する1kbの「スパーサー」配列からなっていた。本発明者らは、等モル量のこれらの4種のコンストラクトと一緒にプールして(ロックドプロモーターGFP、RFP、BFP、および「スパーサー」、図10Aを参照)、1μgの総DNA濃度にし、CLiX親株を形質転換した。1μgのG418耐性F-CphIエンドヌクレアーゼをコードするプラスミドも形質転換に提供して、各CLiX組込み部位を切断した。形質転換体をG418を含有する寒天プレート上で平板培養して、CphIを発現し、提供されたドナーDNAで各二本鎖破断を修復した細胞を選択した。G418プレート上でドナーDNAおよびCphIを含有するプールされた形質転換体から100個のコロニーを回収したが、CphIのみを含有する対照形質転換体ではコロニーは存在しなかった。この結果から、C

10

20

30

40

50

p h I は、高い効率で切断し、非相同末端結合 (NHEJ) などの修復メカニズムは偽陽性を生じる可能性が低いことが示される。176 個のコロニーを B S M 培地 + 2 % スクロースを含有する 96 ウェルプレートにピックアップし、30 で一晩インキュベートした。翌朝、B S M 培地 + 4 % ガラクトースを含有する 96 ウェルプレートに継代接種を実行して、p G A L 1 から発現を誘導した。24 時間のインキュベーション後、各ウェルにつき G F P、R F P、および B F P 蛍光を測定した。

#### 【0228】

[00252] スプリット設計のために、本発明者らは、G F P、R F P、B F P、および「スパーサー」配列のためのプロモーターレス O R F と対を形成するための異なる発現強度を有する 9 種のスプリットプロモーターを構築した (図 10 B)。p G A L 1 は、試験された最も強いプロモーターであり、これを、単一の O R F 設計との比較を可能にするために提供した。他の 8 種の試験されたスプリットプロモーターは、G F P 蛍光で定量化したところ、p G A L 1 の発現レベルが 1 % (p S L N 1) ~ 75 % (p G A L 1\_v20) の範囲であった。全ての 13 種のコンストラクトを等モル量で一緒にプールして 1 μ g の最終濃度にし、F - C p h I と共に C L i X 親株に形質転換した。上述したような B S M 培地 + 4 % ガラクトース中での 24 時間のインキュベーション後、528 個のコロニーを G F P、R F P、および B F P の蛍光に関して査定した。

#### 【0229】

[00253] ロックドプロモーター設計のために、各固定プロモーター蛍光性コンストラクトは、コピー数 (すなわち、どれだけ多くの回数コンストラクトが 3 つの利用可能なランディングパッドに組み込まれたか) のみによって変更が可能である。したがって、本発明者らは、各測定ウェルからの蛍光値が 0 回、1 回、2 回、および 3 回組み込みに対応する 4 つの別個の「ピン」に分類されることを観察した。p G A L 1 > G F P 組み込みの回数に対応する G F P 蛍光に関する 1 回、2 回、および 3 回の「ピン」は、図 10 C において平行斜線で示されたカラムとしてはっきりと確認できる (より見やすくするために、< 300 R F U である 0 回のピンはこのプロットから除去される)。ここで、各蛍光タンパク質でこのパターンが観察されたことに留意されたい。

#### 【0230】

[00254] スプリット設計のために、各蛍光性コンストラクトは、プロモーター強度 (すなわち、スプリットプロモーターがプロモーターレス O R F と組換えされた) およびコピー数において変更が可能である。追加のプロモーターは p G A L 1 より弱いため、本発明者らの懸念は、スプリット設計からの蛍光測定が、各ロックドプロモーター蛍光性コンストラクトの 1 回、2 回、および 3 回コピー間の間隔に「埋もれる」と予想されることであった。図 10 C で観察されたように、スプリットプロモーターの R F U 値 (灰色のバー) が実際に拡張された範囲をカバーした。200 R F U によるビンニングによって、p G A L 1 > G F P の 0 回から 3 回コピーまでのほぼ全ての発現レベルを単一の形質転換でカバーした。

#### 【0231】

##### 6.3 実施例 3: C L i X 親株を使用したゲノム組込みの効率

[00255] この実施例は、宿主細胞のゲノムにおける標的化された二本鎖破断の誘導後の、S. セレビスエ宿主の 3 つの異なるランディングパッドにおける同時組込みを実証する結果を提供する。株 A、B、および C は同じ S. セレビスエ株由来であり、それら全ては、A L G 1、M G A 1 および Y C T 1 の下流の遺伝子間領域に配置された 3 つの C L i X ランディングパッドを含む。これらの株間の差は、上流および下流ランディングパッド相同配列 (U L / D L) の長さである。株 A におけるランディングパッドの U L および D L の各々の長さは、500 塩基対であり、株 B におけるランディングパッドの U L および D L の各々の長さは、200 塩基対であり、株 C におけるランディングパッドの U L および D L の各々の長さは、100 塩基対である。株 A を「ロックド」ステッチ設計 (すなわち、プロモーターに作動可能に連結した GFP) で形質転換し、別の実験で、株 A を「スプリット」プロモーター設計でも形質転換し、ここで 2 つの別々の D N A コンストラクトを使用し

10

20

30

40

50



て、実施例 2 に記載したものと類似した G F P および様々な強度の異なるプロモーターを導入した。株 B を「スプリット」プロモーター設計の D N A コンストラクトで形質転換した。株 C を「スプリット」プロモーター設計の D N A コンストラクトで形質転換した。「ロックド」ステッチ設計または「スプリット」プロモーター設計いずれかの全ての D N A コンストラクトにおいて、相同組換えに使用した上流ライブラリー配列および下流ライブラリー配列の各々は、5 0 0 塩基対である。c P C R、蛍光測定、およびサンガーシーケンシングによって組込みを検証した。

#### 【 0 2 3 2 】

[00256] 図 1 1 で示されるように、全ての 3 つのランディングパッドへの組込みの成功率は 9 5 % より大きかった。

#### 【 0 2 3 3 】

6 . 4 実施例 4 : イソプレノイド産生におけるライブラリーコンポーネントおよびランディングパッドのコンビナトリアルな組込みの性能

[00257] 図 1 2 で例示されているように、3 つのコンビナトリアルな「C L i X」ランディングパッドを、イソプレノイド化合物を高いタイターで産生するように操作されたサッカロマイセス・セレビシエ C E N . P K 2 株に組み込んだ。形質転換およびランディングパッド組込みのための方法は、実施例 1 および 2 に記載されている。ライブラリー組込みのために、U L 相同配列、バーコード配列、「Y」ヌクレアーゼ認識、プロモーター配列、およびリンカー配列からなる 6 種のコンストラクトを、リンカー配列、オープンリーディングフレーム、「Z」ヌクレアーゼ認識部位、バーコード配列、および D L 相同配列からなる 8 種のプロモーターレス O R F と対にした。イソプレノイドのタイターを改善すると仮定された O R F を選んだ。リンカー、ランダムな D N A ヌクレオチドからなる「スタッファー」配列、「Z」ヌクレアーゼ認識部位、バーコード配列、および D L 相同配列からなる類似の構造を有する第 9 の D N A コンストラクトを加えた。このコンストラクトの意図は、このランダムな配列の組込みは株の性能に影響を与えないと予想されるため、1 または 2 つの O R F の組込みを可能にすることであった。

#### 【 0 2 3 4 】

[00258] この 1 5 種のコンストラクトのライブラリーを等モル量で一緒にプールして 1  $\mu$  g の最終濃度にし、F - C p h I をコードする G 4 1 8 耐性プラスミド 1  $\mu$  g と共に親株に形質転換した。これまでに 1 . 1 で説明したような L B + G 4 1 8 からなる固体培地上でコロニーを回収した。5 7 6 0 個のコロニーを、2 % 糖と共に最小培地 ( B S M ) を含有する 9 6 ウェルプレートにピックアップした。一晚増殖させた後、4 % 糖 1 または 6 % 糖 2 のいずれかを含有する液状培地に培養物を継代接種した。増殖の 3 日後、培養物を、U V ベースのアッセイを使用してイソプレノイドのタイターに関してアッセイした。比較のために、「親」対照株 ( 3 つのランディングパッドを有するが組込みがないイソプレノイド産生株 ) を提供した。

#### 【 0 2 3 5 】

[00259] 糖 1、糖 2、または両方の糖 (「ヒット」として公知) において改善された性能を有していたライブラリーから 4 2 個の株を選択した。これらの株を、イソプレノイドのタイターに関して複製を増加させて (株 1 つ当たり 1 2 個の複製) 再度アッセイし、親株を U V によって変異誘発した別個のライブラリーからの 4 2 個の「ヒット」と比較した。コンビナトリアルライブラリーからの 4 2 個の株のうち 3 7 個は再度、イソプレノイドのタイターに関して親より有意に改善されていた。親からの改善の平均有効サイズまたは規模は変異誘発ライブラリーよりも劇的に増加した。

#### 【 0 2 3 6 】

[00260] 複製を増加させてアッセイした 4 2 個の株を、標準化されたバーコードを使用して遺伝子型解析して、各遺伝子座において組み込まれたプロモーターおよび O R F のアイデンティティを決定した。O R F 1 の場合、転写活性化因子がライブラリーの残りに比べて富化された。P 3 は、ライブラリー中の最も弱いプロモーターであるが、これは、組込みの 7 5 % 超において O R F 1 と対を形成した。これらの結果から、O R F 1 の過剰発

10

20

30

40

50

現は、イソプレノイドのタイターを増加させるのに有効な方法であったが、低いレベルで発現された場合のみであったことが示唆された。

【 0 2 3 7 】

[00261] 図 1 2 に示されるこれらの結果は、ライブラリーコンポーネント（例えば、様々な発現強度のプロモーターおよびORF）のコンビナトリアルな組込みが、遺伝子の共発現の最適な釣り合いを決定して標的分子産生を最大にするために、同時に様々な遺伝子の発現の適正量を決定することにおいて有効な手段であることを実証する。ランディングパッドへのライブラリーコンポーネントのコンビナトリアルな組込みはさらに、遺伝子型のより一層多くのバリエーションを生成する、および株につき得られた表現型をさらに改善する利点を提供することができる。

10

【 0 2 3 8 】

[00262] 本明細書または図面に記載されるあらゆる実施形態からの 1 つまたは複数の特徴は、本明細書や図面に記載されるあらゆる他の実施形態の 1 つまたは複数の特徴と、本発明の範囲から逸脱することなく組み合わせることができる。

【 0 2 3 9 】

[00263] 「1つの(a)」、「1つの(an)」、または「その(the)」の記載は、具体的にそうではないことが示されない限り、「1つまたは複数」を意味することが意図される。

【 0 2 4 0 】

[00264] 本明細書に引用された全ての公報および特許、出願は、各々個々の公報または特許出願が具体的かつ個々に参照により組み入れられることが示されたのと同様に、参照により本明細書に組み入れられる。特許請求された主題を様々な実施形態に関して説明してきたが、当業者は、様々な改変、置換、省略、および変更がそれらの本質から逸脱することなくなし得ることを理解しているものと予想される。したがって、主題の範囲は、それらの均等物を含め以下の特許請求の範囲によってのみ限定されることが意図される。

20

30

40

50



プールされた DNA コンストラクト

プールされたヌクレアーゼ

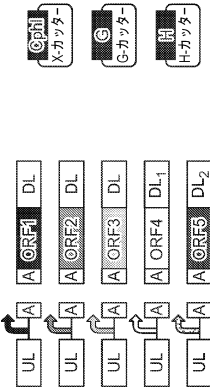


図 6A



G-カッターが提供された場合のみ、「開裂」と予想される  
H-カッターが提供された場合のみ、「開裂」と予想される

図 6B

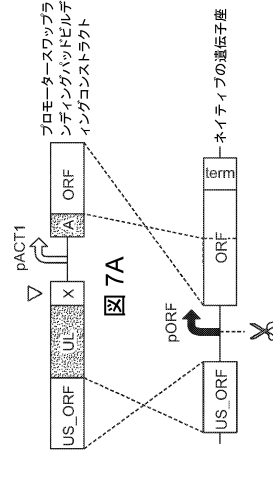


図 7A

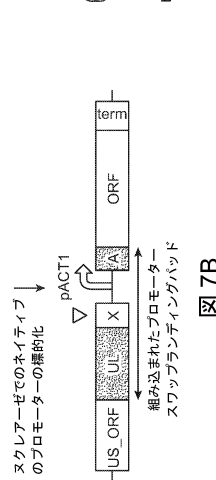


図 7B

【図 7 - 1】

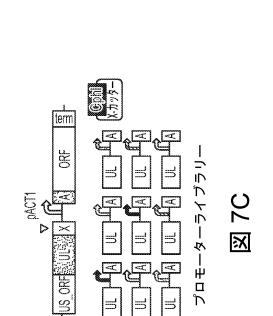


図 7C

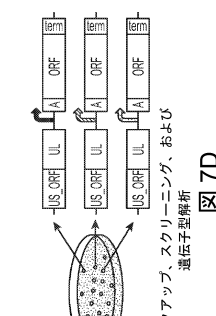


図 7D

【図 6】

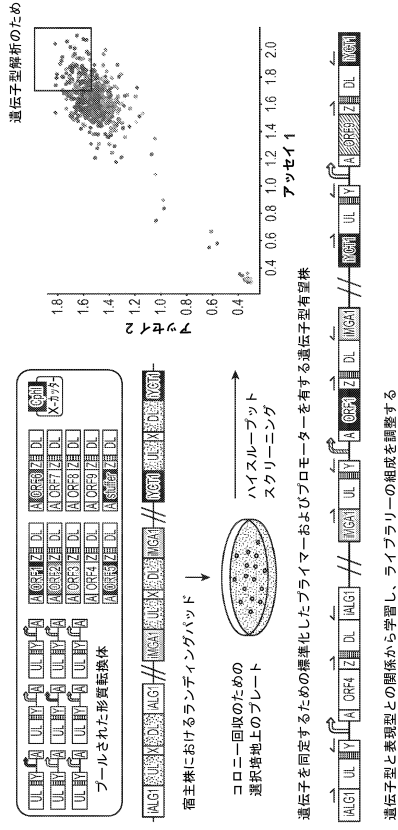


図 4

【図 5】

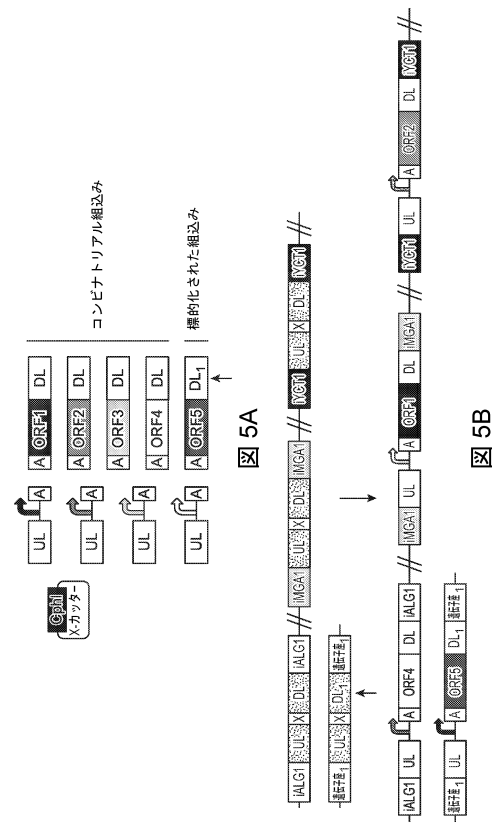
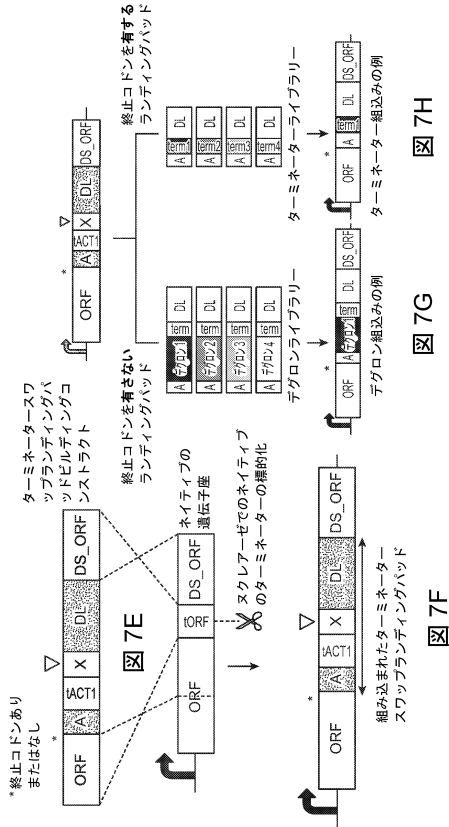


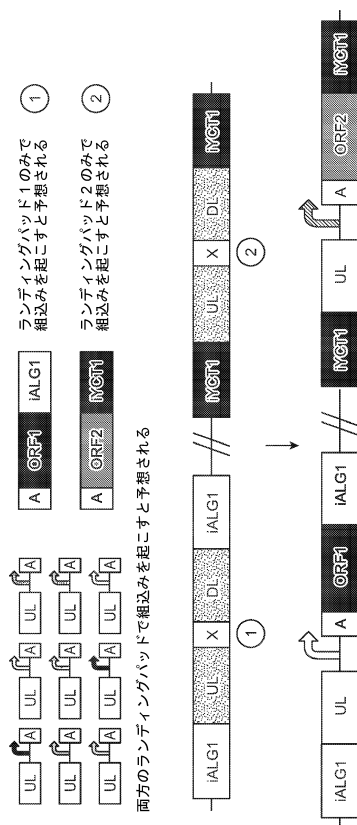
図 5A

図 5B

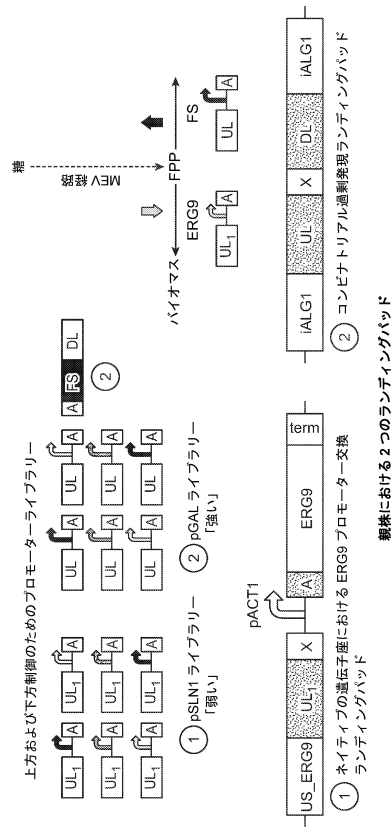
【圖 7 - 2】



【 図 9 】



【 図 8 】



【 図 1 0 】

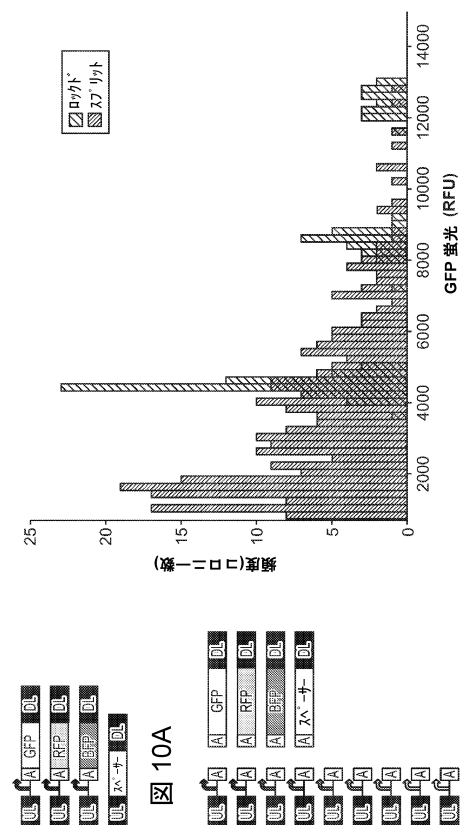
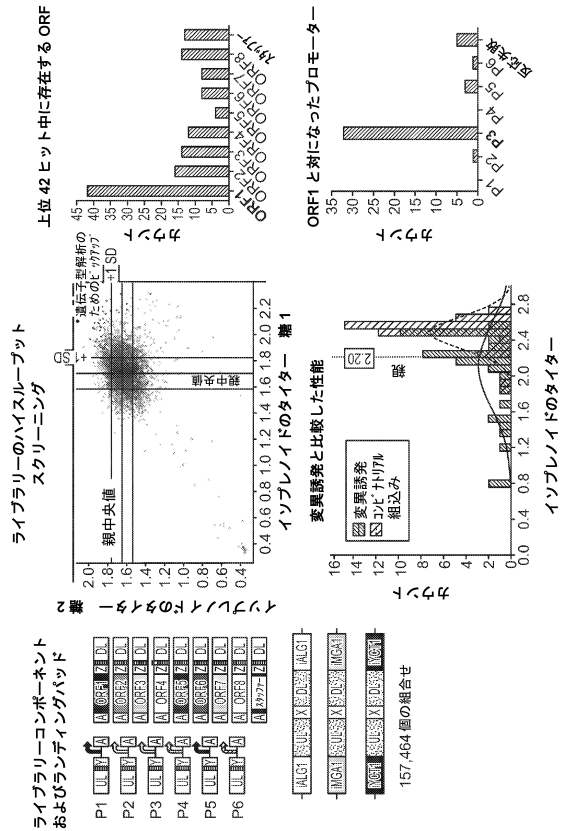
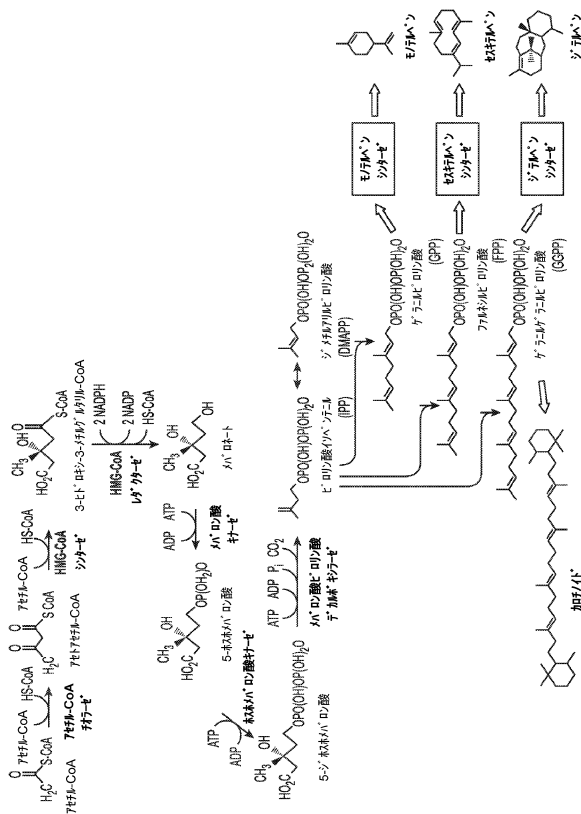


表1. CLIXを使用する組込み効率

| 株番号<br>UL/DL 長さ | 株 A<br>500 bp | 株 A<br>500 bp | 株 B<br>200 bp | 株 C<br>100 bp |
|-----------------|---------------|---------------|---------------|---------------|
| ステッチ設計          | 96            | 96            | 96            | 96            |
| コロニー数           | 288           | 288           | 288           | 288           |
| 遺伝子座数           |               |               |               |               |
| 組込み w/インデル      | 0             | 3             | 9             | 7             |
| 組込みなし           | 0             | 3             | 3             | 0             |
| 正しい組込み          | 288           | 282           | 276           | 281           |
| 成功率             | 100%          | 97.9%         | 95.8%         | 97.6%         |

株 A、株 B、および株 C は同じナীব株に由来し、全て ALG1、MGA1 および YCT1 の下流遺伝子間領域中に配置された 3 つの Clx ランディングパッドを有する。cPQR、蛍光測定、およびサンガーシーケンシングによって網込みを検証した。UL、上流ライブラリー-DL、下流ライブラリー



【配列表】

0007093310000001.app

10

20

30

40

50

## フロントページの続き

- 4 6 0 8 , エミリービル , ホリス ストリート 5 8 8 5 シー / オー , スイート 1 0 0  
(72)発明者 マイン , アンドリュー  
アメリカ合衆国 , カリフォルニア州 9 4 6 0 8 , エミリービル , ホリス ストリート 5 8 8 5  
シー / オー , スイート 1 0 0  
(72)発明者 ツァイ , チア - ホン  
アメリカ合衆国 , カリフォルニア州 9 4 6 0 8 , エミリービル , ホリス ストリート 5 8 8 5  
シー / オー , スイート 1 0 0  
審査官 竹内 祐樹  
(56)参考文献 特表 2 0 1 4 - 5 1 3 5 3 5 ( J P , A )  
特表 2 0 0 5 - 5 0 0 8 4 1 ( J P , A )  
NUCLEIC ACIDS RESEARCH , 2010年04月 , Vol.38, No.6 e92 , p.1-10  
(58)調査した分野 (Int.Cl. , D B 名)  
C 1 2 N 1 5 / 0 0 - 1 5 / 9 0  
J S T P l u s / J M E D P l u s / J S T 7 5 8 0 ( J D r e a m I I I )  
C A p l u s / M E D L I N E / E M B A S E / B I O S I S ( S T N )  
P u b M e d