

(19) 日本国特許庁 (JP)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2018-516563

(P2018-516563A)

(43) 公表日 平成30年6月28日 (2018. 6. 28)

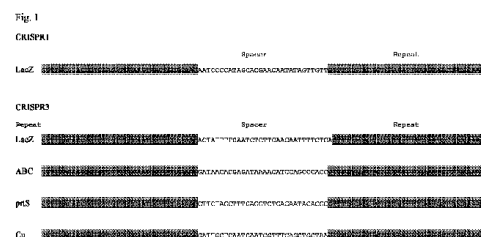
(51) Int. Cl.	F I	テーマコード (参考)
C 1 2 N 15/09 (2006. 01)	C 1 2 N 15/09 1 1 O	4 B O 6 3
C 1 2 Q 1/02 (2006. 01)	C 1 2 Q 1/02 Z N A	4 B O 6 5
C 1 2 N 1/16 (2006. 01)	C 1 2 N 1/16 Z	
C 1 2 N 1/19 (2006. 01)	C 1 2 N 1/19	
C 1 2 N 1/12 (2006. 01)	C 1 2 N 1/12 Z	
審査請求 未請求 予備審査請求 未請求 (全 91 頁) 最終頁に続く		

(21) 出願番号	特願2017-559860 (P2017-559860)	(71) 出願人	517366013
(86) (22) 出願日	平成28年5月27日 (2016. 5. 27)		ノース カロライナ ステート ユニバー シティ
(85) 翻訳文提出日	平成29年12月22日 (2017. 12. 22)		NORTH CAROLINA STAT E UNIVERSITY
(86) 国際出願番号	PCT/US2016/034812		アメリカ合衆国、ノースカロライナ州 2 7 6 0 6、ラレー、セカンド フロアー、 メイン キャンパス ドライブ 1 0 2 1
(87) 国際公開番号	W02016/196361	(74) 代理人	100113376
(87) 国際公開日	平成28年12月8日 (2016. 12. 8)		弁理士 南条 雅裕
(31) 優先権主張番号	62/168, 355	(74) 代理人	100179394
(32) 優先日	平成27年5月29日 (2015. 5. 29)		弁理士 瀬田 あや子
(33) 優先権主張国	米国 (US)	(74) 代理人	100185384
(31) 優先権主張番号	62/296, 853		弁理士 伊波 興一朗
(32) 優先日	平成28年2月18日 (2016. 2. 18)		
(33) 優先権主張国	米国 (US)		
		最終頁に続く	

(54) 【発明の名称】 C R I S P R核酸を用いて、細菌、古細菌、藻類、および、酵母をスクリーニングする方法

(57) 【要約】

本発明は、細菌、古細菌、藻類および／または酵母における必須および非必須の遺伝子および消費されてよいゲノムアイランドをスクリーニングするため、細菌、古細菌、藻類および／または酵母を死滅させるため、遺伝子（単数または複数）の表現型を同定するため、および／または、細菌、古細菌、藻類および／または酵母における縮小されたゲノムサイズおよび／または遺伝子欠失をスクリーニングするための、C R I S P R核酸の使用に関する。



【特許請求の範囲】

【請求項 1】

必須遺伝子、非必須遺伝子、および／または消費されてよい (expendable) ゲノムアイランドに関して、細菌細胞の集団をスクリーニングする方法であって、

細菌細胞の前記集団に、(5' から 3' に) リピートースペーサーーリピート配列または少なくとも 1 つのリピートースペーサー配列を含む C R I S P R アレイを含む異種核酸構築物を導入するステップ(ここで、前記リピートースペーサーーリピート配列または前記の少なくとも 1 つのリピートースペーサー配列の前記スペーサーは、前記集団の前記細菌細胞の前記ゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、それにより、形質転換された細菌細胞の集団を生産する) ; および

形質転換された細菌細胞の前記集団における欠失の存在または不存在を判定するステップ(ここで、形質転換された細菌細胞の前記集団における欠失の存在は、前記標的領域が非必須遺伝子および／または消費されてよいゲノムアイランド内に含まれることを示し、および、前記集団における欠失の不存在は、前記標的領域が必須遺伝子内に含まれることを意味する)

を含む、

方法。

【請求項 2】

必須遺伝子、非必須遺伝子、および／または消費されてよいゲノムアイランドに関して、細菌、古細菌、藻類または酵母細胞の集団をスクリーニングする方法であって、

細菌、古細菌、藻類または酵母細胞の前記集団に、(a) トランスコード化 C R I S P R (t r a c r) 核酸を含む異種核酸構築物、(b) (5' から 3' に) リピートースペーサーーリピート配列または少なくとも 1 つのリピートースペーサー配列を含む C R I S P R アレイを含む異種核酸構築物(ここで、前記リピートースペーサーーリピート配列または前記の少なくとも 1 つのリピートースペーサー配列の前記スペーサーは、前記集団の前記細菌、古細菌、藻類または酵母細胞の前記ゲノム(染色体および／またはプラスミド)内の標的領域に実質的に相補なヌクレオチド配列を含む)、および、(c) C a s 9 ポリペプチドまたは C a s 9 ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップ(それにより、形質転換された細菌、古細菌、藻類または酵母細胞の集団を生産する) ; および

形質転換された細菌、古細菌、藻類または酵母細胞の前記集団における欠失の存在または不存在を判定するステップ(ここで、形質転換された細菌、古細菌または酵母細胞の前記集団における欠失の存在は、前記標的領域が非必須遺伝子および／または消費されてよいゲノムアイランド内に含まれることを意味し、形質転換された細菌、古細菌、藻類または酵母細胞の前記集団における欠失の不存在は、前記標的領域が必須遺伝子内に含まれることを意味する)、

を含む、

方法。

【請求項 3】

細菌細胞の集団内の 1 つまたは複数の細菌細胞を死滅させる方法であって、

細菌細胞の前記集団に、(5' から 3' に) リピートースペーサーーリピート配列または少なくとも 1 つのリピートースペーサー配列を含む C R I S P R アレイ (c r R N A、c r D N A) を含む異種核酸構築物を導入するステップを含み、

ここで、前記リピートースペーサーーリピート配列または前記の少なくとも 1 つのリピートースペーサー配列の前記スペーサーは、前記集団の前記細菌細胞の前記ゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、それにより、細菌細胞の前記集団内の前記標的領域を含む 1 つまたは複数の細菌細胞を死滅させる、

方法。

【請求項 4】

細菌、古細菌、藻類または酵母細胞の集団内の 1 つまたは複数の細胞を死滅させる方法

であって、

細菌、古細菌、藻類または酵母細胞の前記集団に、(a) トランスコード化 C R I S P R (t r a c r) 核酸を含む異種核酸構築物、(b) (5' から 3' に) リピートースペーサーーリピート配列または少なくとも 1 つのリピートースペーサー配列を含む C R I S P R アレイを含む異種核酸構築物 (ここで、前記リピートースペーサーーリピート配列または前記の少なくとも 1 つのリピートースペーサー配列の前記スペーサーは、前記集団の前記の細菌、古細菌、藻類または酵母細胞の前記ゲノム (染色体および/またはプラスミド) 内の標的領域に実質的に相補なヌクレオチド配列を含む)、および、(c) C a s 9 ポリペプチドおよび/または C a s 9 ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップを含み、

10

それにより、それらのゲノム内の前記標的領域を含む細菌、古細菌、藻類または酵母細胞の集団内の 1 つまたは複数の細胞を死滅させる、方法。

【請求項 5】

請求項 3 または 4 の方法であって、

前記標的領域は、必須遺伝子または非必須遺伝子内にある、方法。

【請求項 6】

細菌遺伝子と関連する表現型を同定する方法であって、

細菌細胞の集団に、(5' から 3' に) リピートースペーサーーリピート配列または少なくとも 1 つのリピートースペーサー配列を含む C R I S P R アレイを含む異種核酸構築物を導入するステップ (ここで、前記の少なくとも 1 つのリピートースペーサー配列およびリピートースペーサーーリピート配列の前記スペーサーは、前記集団の前記細菌細胞の前記ゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、ここで、前記標的領域は、ポリペプチドまたは機能性核酸をコードするオープンリーディングフレームの少なくとも一部を含み、それにより、前記標的領域を含む前記細胞を死滅させ、そして、前記標的領域を含まない形質転換された細菌細胞の集団を生産する) ; および、

20

前記集団の表現型を分析するステップ、

を含む、

方法。

30

【請求項 7】

請求項 6 の方法であって、

分析するステップの前に、形質転換された細菌細胞の前記集団から個々の細菌コロニーを増殖させるステップ ; および

前記の個々のコロニーの表現型を分析するステップ、

を含む、

方法。

【請求項 8】

細菌、古細菌、藻類または酵母遺伝子の表現型を同定する方法であって、

細菌、古細菌、藻類または酵母細胞の集団に、(a) トランスコード化 C R I S P R (t r a c r) 核酸を含む異種核酸構築物、(b) (5' から 3' に) リピートースペーサーーリピート配列または少なくとも 1 つのリピートースペーサー配列を含む C R I S P R アレイを含む異種核酸構築物 (ここで、前記リピートースペーサーーリピート配列または前記の少なくとも 1 つのリピートースペーサー配列の前記スペーサーは、前記集団の前記の細菌、古細菌、藻類または酵母細胞の前記ゲノム (染色体および/またはプラスミド) 内の標的領域に実質的に相補なヌクレオチド配列を含む)、および、(c) C a s 9 ポリペプチドおよび/または C a s 9 ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップ (それにより、前記標的領域を含む前記の細菌、古細菌または酵母細胞を死滅させ、そして、前記標的領域を含まない形質転換された細菌、古細菌、藻類または酵母細胞の集団を生産する) ; および

40

50

形質転換された細菌、古細菌、藻類または酵母細胞の前記集団の表現型を分析するステップ、および／または、(i) 形質転換された細菌、古細菌、藻類または酵母細胞の前記集団から、個々の細菌、古細菌、藻類または酵母コロニーを増殖させるステップ、および、(ii) 前記の個々のコロニーの表現型を分析するステップ、を含む、

方法。

【請求項 9】

請求項 1、3、6、または 7 のいずれか一項の方法であって、

前記 CRISPR アレイは、I 型、II 型、III 型、IV 型、V 型 CRISPR アレイである、

10

方法。

【請求項 10】

請求項 9 の方法であって、

前記リピートスペーサーリピート配列または前記の少なくとも 1 つのリピートスペーサー配列は、野生型 I 型 CRISPR アレイ、野生型 II 型 CRISPR アレイ、野生型 III 型 CRISPR アレイ、野生型 IV 型 CRISPR アレイ、または、野生型 V 型 CRISPR アレイ由来のリピートと同一のリピートを含む、

方法。

【請求項 11】

請求項 2、4、5 または 8 のいずれか一項の方法であって、

20

前記リピートスペーサーリピート配列または前記の少なくとも 1 つのリピートスペーサー配列は、野生型 II 型 CRISPR アレイ由来のリピートと同一のリピートを含む、

方法。

【請求項 12】

請求項 1 から 11 のいずれか一項の方法であって、

前記標的領域は、無作為に選択される、または、特異的に選択される、

方法。

【請求項 13】

請求項 12 の方法であって、

30

無作為に選択される標的領域は、細菌、古細菌または酵母ゲノム内の PAM 配列に隣接して配置された、任意の少なくとも 10 個の連続するヌクレオチドから選択される、

方法。

【請求項 14】

請求項 12 の方法であって、

特異的に選択される標的領域は、細菌、古細菌、藻類または酵母ゲノム内の PAM 配列に隣接する少なくとも 10 個の連続するヌクレオチドを含む、遺伝子、オープンリーディングフレームまたは推定のオープンリーディングフレームから選択される、

方法。

【請求項 15】

40

請求項 2、4、5、8、11 から 14 のいずれか一項の方法であって、

トランスコード化 CRISPR (tracr) 核酸を含む前記異種核酸構築物および CRISPR アレイを含む前記異種核酸構築物は、場合により Cas9 ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物をさらに含む CRISPR ガイド (gRNA、gDNA) 内に含まれる、

方法。

【請求項 16】

請求項 15 の方法であって、

前記 CRISPR ガイドは、プロモーターに動作可能に連結される、

方法。

50

【請求項 17】

細菌細胞の集団から、縮小されたゲノムサイズを有する 1 つまたは複数の細菌細胞を選択する方法であって、

細菌細胞の集団に、（5' から 3' に）リピートスペーサーーリピート配列または少なくとも 1 つのリピートスペーサー配列を含む C R I S P R アレイを含む異種核酸構築物を導入するステップ（ここで、前記リピートスペーサーーリピート配列または前記の少なくとも 1 つのリピートスペーサー配列の前記スペーサーは、前記集団の 1 つまたは複数の細菌細胞の前記ゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、ここで、前記標的領域を含む細胞は死滅して、それにより、細菌細胞の前記集団から、前記標的領域を含まない縮小されたゲノムサイズを有する 1 つまたは複数の細菌細胞を選択する）、

を含む、
方法。

【請求項 18】

細菌細胞の集団から、縮小されたゲノムサイズを有する 1 つまたは複数の細菌細胞を選択する方法であって、

細菌細胞の集団に、

（a）（i）前記細菌細胞の前記ゲノム内に存在する少なくとも 300 個の連続したヌクレオチドと少なくとも 80% の同一性を有するヌクレオチド配列を含む、1 つまたは複数の異種核酸構築物、または、（ii）少なくとも 1 つのトランスポゾンを含む 2 以上の異種核酸構築物（それにより、前記ゲノム中に組み込まれた前記の 1 つまたは複数の異種核酸構築物および前記ゲノム内に存在する前記の少なくとも 300 個の連続したヌクレオチドの間、または、前記ゲノム中に組み込まれた第 1 および第 2 のトランスポゾンの間の、相同組み換えに関する非天然部位を含むトランスジェニック細菌細胞の集団を生産する）；および

（b）（5' から 3' に）リピートスペーサーーリピート配列または少なくとも 1 つのリピートスペーサー配列を含む C R I S P R アレイを含む異種核酸構築物、を導入するステップ（ここで、前記リピートスペーサーーリピート配列または前記の少なくとも 1 つのリピートスペーサー配列の前記スペーサーは、前記集団の 1 つまたは複数の細菌細胞の前記ゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、ここで、前記標的領域は、前記ゲノム中に導入された前記の 1 つまたは複数の異種核酸構築物および前記ゲノム内に存在する前記の少なくとも 300 個の連続したヌクレオチドの間、および／または、前記の第 1 のトランスポゾンおよび第 2 のトランスポゾンの間に位置し、前記標的領域を含む細胞は死滅して、それにより、トランスジェニック細菌細胞の前記集団から、前記標的領域を含まない縮小されたゲノムサイズを有する 1 つまたは複数の細菌細胞を選択する）、

を含む、
方法。

【請求項 19】

細菌、古細菌、藻類または酵母細胞の集団から、縮小されたゲノムサイズを有する 1 つまたは複数の細菌、古細菌、藻類または酵母細胞を選択する方法であって、

細菌、古細菌、藻類または酵母細胞の集団に、

（a）トランスコード化 C R I S P R （t r a c r）核酸を含む異種核酸構築物、

（b）リピートスペーサーーリピート配列または少なくとも 1 つのリピートスペーサー配列を含む C R I S P R アレイを含む異種核酸構築物（ここで、前記の少なくとも 1 つのリピートスペーサー配列および前記の少なくとも 1 つのリピートスペーサーーリピート配列の前記スペーサーは、前記集団の前記の細菌、古細菌、藻類または酵母細胞の前記ゲノム（染色体および／またはプラスミド）内の標的領域に実質的に相補なヌクレオチド配列を含む）、および

（c）C a s 9 ポリペプチドおよび／または C a s 9 ポリペプチドをコードするポリヌ

クレオチドを含む異種核酸構築物、を導入するステップ（ここで、前記標的領域を含む細胞は死滅して、それにより、細菌、古細菌、藻類または酵母細胞の前記集団から、前記標的領域を含まない縮小されたゲノムサイズを有する1つまたは複数の細菌、古細菌、藻類または酵母細胞を選択する）、

を含む、

方法。

【請求項20】

細菌、古細菌または酵母細胞の集団から、縮小されたゲノムサイズを有する1つまたは複数の細菌、古細菌、藻類または酵母細胞を選択する方法であって、

細菌、古細菌、藻類または酵母細胞の集団に、

(a) (i) 前記の細菌、古細菌、藻類または酵母細胞の前記ゲノム内に存在する少なくとも300個の連続したヌクレオチドと少なくとも80%の同一性を有するヌクレオチド配列を含む、1つまたは複数の異種核酸構築物、または、(ii) 少なくとも1つのトランスポゾンを含む2以上の異種核酸構築物（それにより、前記ゲノム中に組み込まれた前記の1つまたは複数の異種核酸構築物および前記ゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、または、前記ゲノム中に組み込まれた第1および第2のトランスポゾンの間の、相同組み換えに関する非天然部位を含むトランスジェニック細菌、古細菌、藻類または酵母細胞の集団を生産する）；および

(b) (i) トランスコード化CRISPR (tracr) 核酸を含む異種核酸構築物、(ii) リピートースペーサーリピート配列または少なくとも1つのリピートースペーサー配列を含むCRISPRアレイを含む異種核酸構築物（ここで、前記の少なくとも1つのリピートースペーサー配列および前記の少なくとも1つのリピートースペーサーリピート配列の前記スペーサーは、前記集団の1つまたは複数の細菌、古細菌、藻類または酵母細胞の前記ゲノム（染色体および／またはプラスミド）内の標的領域に実質的に相補なヌクレオチド配列を含む）、および、(iii) Cas9ポリペプチドおよび／またはCas9ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップ（ここで、前記標的領域は、前記ゲノム中に組み込まれた前記の1つまたは複数の異種核酸構築物および前記ゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、および／または、前記の第1のトランスポゾンおよび第2のトランスポゾンの間に位置し、および、前記標的領域を含む細胞は死滅して、それにより、トランスジェニック細菌、古細菌、藻類または酵母細胞の前記集団から、前記標的領域を含まない縮小されたゲノムサイズを有する1つまたは複数の細菌、古細菌、藻類または酵母細胞を選択する）、

を含む、

方法。

【請求項21】

細菌の集団において、そのゲノム内に欠失を有する少なくとも1つの単離物を同定する方法であって、

細菌細胞の集団に、(5'から3'に) リピートースペーサーリピート配列または少なくとも1つのリピートースペーサー配列を含むCRISPRアレイを含む異種核酸構築物を導入するステップ（ここで、前記リピートースペーサーリピート配列または前記の少なくとも1つのリピートースペーサー配列の前記スペーサーは、前記集団の1つまたは複数の細菌細胞の前記ゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、ここで、前記標的領域を含む細胞は死滅して、それにより、前記標的領域を含まない形質転換された細菌細胞の集団を生産する）；および

形質転換された細菌細胞の前記集団から個々の細菌コロニーを増殖させるステップ（それにより、形質転換された細菌の前記集団から、そのゲノム内に欠失を有する少なくとも1つの単離物を同定する）、

を含む、

方法。

【請求項 2 2】

細菌の集団において、そのゲノム内に欠失を有する少なくとも 1 つの単離物を同定する方法であって、

細菌細胞の前記集団に、

(a) (i) 前記細菌細胞の前記ゲノム内に存在する少なくとも 300 個の連続したヌクレオチドと少なくとも 80 % の同一性を有するヌクレオチド配列を含む、1 つまたは複数の異種核酸構築物、または、(ii) 少なくとも 1 つのトランスポゾンを含む 2 以上の異種核酸構築物（それにより、前記ゲノム中に組み込まれた前記の 1 つまたは複数の異種核酸構築物および前記ゲノム内に存在する前記の少なくとも 300 個の連続したヌクレオチドの間、または、前記ゲノム中に組み込まれた第 1 および第 2 のトランスポゾンの間の、相同組み換えに関する非天然部位を含むトランスジェニック細菌細胞の集団を生産する）；および

b) (5' から 3' に) リピートスペーサーリピート配列または少なくとも 1 つのリピートスペーサー配列を含む CRISPR アレイを含む異種核酸構築物、を導入するステップ（ここで、前記リピートスペーサーリピート配列または前記の少なくとも 1 つのリピートスペーサー配列の前記スペーサーは、前記集団の 1 つまたは複数の細菌細胞の前記ゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、ここで、前記標的領域は、前記ゲノム中に導入された前記の 1 つまたは複数の異種核酸構築物および前記ゲノム内に存在する前記の少なくとも 300 個の連続したヌクレオチドの間、および／または、前記の第 1 のトランスポゾンおよび第 2 のトランスポゾンの間に位置し、前記標的領域を含む細胞は死滅して、それにより、前記標的領域を含まない形質転換された細菌細胞の集団を生産する）；および

形質転換された細菌細胞の前記集団から個々の細菌コロニーを増殖するステップ（それにより、細菌の前記集団から、そのゲノム内に欠失を有する少なくとも 1 つの単離物を同定する）、

を含む、

方法。

【請求項 2 3】

細菌、古細菌、藻類または酵母細胞の集団において、そのゲノム内に欠失を有する少なくとも 1 つの単離物を同定する方法であって、

細菌、古細菌、藻類または酵母細胞の集団に、

(a) トランスコード化 CRISPR (t r a c r) 核酸を含む異種核酸構築物、

(b) (5' から 3' に) リピートスペーサーリピート配列または少なくとも 1 つのリピートスペーサー配列を含む CRISPR アレイを含む異種核酸構築物（ここで、前記リピートスペーサーリピート配列または前記の少なくとも 1 つのリピートスペーサー配列の前記スペーサーは、前記集団の前記の細菌、古細菌、藻類または酵母細胞の前記ゲノム（染色体および／またはプラスミド）内の標的領域に実質的に相補なヌクレオチド配列を含む）、および

(c) Cas9 ポリペプチドまたは Cas9 ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップ（ここで、前記標的領域を含む細胞は死滅して、それにより、前記標的領域を含まない形質転換された細菌、古細菌、藻類または酵母細胞の集団を生産する）；および

形質転換された細菌、古細菌、藻類または酵母細胞の前記集団から、個々の細菌、古細菌または酵母コロニーを増殖させるステップ（それにより、形質転換された細菌、古細菌、藻類または酵母細胞の前記集団から、そのゲノム内に欠失を有する少なくとも 1 つの単離物を同定する）、

を含む、

方法。

【請求項 2 4】

細菌、古細菌、藻類または酵母細胞の集団において、そのゲノム内に欠失を有する少な

くとも1つの単離物を同定する方法であって、

細菌、古細菌、藻類または酵母細胞の前記集団に、

(a) (i) 前記の細菌、古細菌、藻類または酵母細胞の前記ゲノム内に存在する少なくとも300個の連続したヌクレオチドと少なくとも80%の同一性を有するヌクレオチド配列を含む、1つまたは複数の異種核酸構築物、または、(ii) 少なくとも1つのトランスポゾンを含む2以上の異種核酸構築物(それにより、前記ゲノム中に組み込まれた前記の1つまたは複数の異種核酸構築物および前記ゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、または、前記ゲノム中に組み込まれた第1および第2のトランスポゾンの間の、相同組み換えに関する非天然部位を含むトランスジェニック細菌、古細菌、藻類または酵母細胞の集団を生産する) ; および

(b) (i) トランスコード化CRISPR (t r a c r) 核酸を含む異種核酸構築物、(ii) (5'から3'に) リピートースペーサーーリピート配列または少なくとも1つのリピートースペーサー配列を含むCRISPRアレイを含む異種核酸構築物(ここで、前記リピートースペーサーーリピート配列または前記の少なくとも1つのリピートースペーサー配列の前記スペーサーは、前記集団の1つまたは複数の細菌、古細菌、藻類または酵母細胞の前記ゲノム(染色体および/またはプラスミド)内の標的領域に実質的に相補なヌクレオチド配列を含む)、および、(iii) Cas9ポリペプチドおよび/またはCas9ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップ(ここで、前記標的領域は、前記ゲノム中に組み込まれた前記の1つまたは複数の異種核酸構築物および前記ゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、および/または、前記の第1のトランスポゾンおよび第2のトランスポゾンの間に位置し、前記標的領域を含む細胞は死滅して、それにより、前記標的領域を含まない形質転換された細菌、古細菌、藻類または酵母細胞の集団を生産する) ; および形質転換された細菌、古細菌、藻類または酵母細胞の前記集団から、個々の細菌、古細菌または酵母コロニーを増殖させるステップ(それにより、前記集団から、そのゲノム内に欠失を有する少なくとも1つの単離物を同定する)、

を含む、

方法。

【請求項25】

請求項17、18、21、または22のいずれか一項の方法であって、

前記CRISPRアレイは、I型、II型、III型、IV型、V型CRISPRアレイである、

方法。

【請求項26】

請求項25の方法であって、

前記リピートースペーサーーリピート配列または前記の少なくとも1つのリピートースペーサー配列は、野生型I型CRISPRアレイ、野生型II型CRISPRアレイ、野生型III型CRISPRアレイ、野生型IV型CRISPRアレイ、または、野生型V型CRISPRアレイ由来のリピートと同一のリピートを含む、

方法。

【請求項27】

請求項19、20、23、または24のいずれか一項の方法であって、

前記リピートースペーサーーリピート配列または前記の少なくとも1つのリピートースペーサー配列は、野生型II型CRISPRアレイ由来のリピートと同一のリピートを含む、

方法。

【請求項28】

請求項15から27のいずれか一項の方法であって、

前記標的領域は、無作為に選択される、または、特異的に選択される、

方法。

【請求項 29】

請求項 28 の方法であって、

無作為に選択される標的領域は、細菌、古細菌または酵母ゲノム内の PAM 配列に隣接して配置された任意の少なくとも 10 個の連続するヌクレオチドから選択される、方法。

【請求項 30】

請求項 28 の方法であって、

特異的に選択される標的領域は、細菌、古細菌、藻類または酵母ゲノム内の PAM 配列に隣接する少なくとも 10 個の連続するヌクレオチドを含む、遺伝子、オープンリーディングフレームまたは推定のオープンリーディングフレームから選択される、方法。

10

【請求項 31】

請求項 3 から 5 または請求項 9 から 15 のいずれか一項の方法であって、

前記の導入される CRISPR アレイは、死滅される前記の 1 つまたは複数の細菌細胞における CRISPR-Cas システムと適合し、細菌細胞の前記集団内の少なくとも 1 つまたは複数の細菌細胞の前記 CRISPR-Cas システムと適合しない、方法。

【発明の詳細な説明】

【技術分野】

【0001】

20

〔優先権の陳述〕

本出願は、2015 年 5 月 29 日に提出された米国仮出願第 62/168,355 号および 2016 年 2 月 18 日に提出された米国仮出願第 62/296,853 号の合衆国法典第 35 巻第 119 条 (e) の下での利益を主張し、それぞれの内容全体は、参照により本明細書中に援用される。

【0002】

〔本発明の分野〕

本発明は、細菌、古細菌、藻類および／または酵母における必須および非必須の遺伝子および消費されてよい (expendable) ゲノムアイランドをスクリーニングするため、細菌、古細菌、藻類および／または酵母を死滅させるため、遺伝子 (単数または複数) の表現型を同定するため、および／または、細菌、古細菌、藻類および／または酵母において縮小されたゲノムサイズおよび／または遺伝子欠失をスクリーニングするための、CRISPR 核酸の使用に関する。

30

【背景技術】

【0003】

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) は、関連配列 (cas) と組み合わせて、CRISPR-Cas システムを構成し、多くの細菌において適応免疫を与える。CRISPR が仲介する免疫付与は、プラスミドおよびファージなどの侵襲性遺伝因子由来の DNA の、新規の「スペーサー」としての取り込みを介して生じる。

40

【0004】

CRISPR-Cas システムは、超可変配列により隔てられた短い DNA リピートのアレイからなり、cas 遺伝子が隣接し、配列特異的ターゲティングおよび干渉を介して、ファージおよびプラスミドなどの侵襲性遺伝因子に対する適応免疫を提供する (Barrangou et al. 2007. Science. 315:1709-1712; Brouns et al. 2008. Science 321:960-4; Horvath and Barrangou. 2010. Science. 327:167-70; Marraffini and Sontheimer. 2008. Science. 322:1843-1845; Bhaya et al. 2011. Annu. Rev. Genet. 45:273-297; Terns and Terns. 2011.

50

Curr. Opin. Microbiol. 14:321-327; Westra et al. 2012. Annu. Rev. Genet. 46:311-339; Barrangou R. 2013. RNA. 4:267-278)。

【0005】

典型的に、侵襲性DNA配列は、新規の「スパーサー」として獲得され (Barrangou et al. 2007. Science. 315:1709-1712)、それぞれがCRISPRリピートと対になり、CRISPR遺伝子座において新規のリピータースパーサーユニットとして挿入される。続いて、リピータースパーサーアレイは、長いpre-CRISPR RNA (pre-crRNA)として転写され (Brouns et al. 2008. Science 321:960-4)、配列特異的認識を駆動する低分子干渉CRISPR RNA (crRNA)へ処理される。具体的には、crRNAは、Casエンドヌクレアーゼにより仲介される配列特異的核酸切断のために、ヌクレアーゼを相補標的に向かってガイドする (Garneau et al. 2010. Nature. 468:67-71; Haurwitz et al. 2010. Science. 329:1355-1358; Saprunauskas et al. 2011. Nucleic Acid Res. 39:9275-9282; Jinek et al. 2012. Science. 337:816-821; Gasiunas et al. 2012. Proc. Natl. Acad. Sci. 109:E2579-E2586; Magadan et al. 2012. PLoS One. 7:e40913; Karvelis et al. 2013. RNA Biol. 10:841-851)。

これらの広範なシステムは、細菌のほぼ半数 (約46%) および大多数の古細菌 (約90%) で生じる。

【0006】

一般論として、CRISPR-Casシステムの2つの主要なクラスが存在し (Makarova et al. Nat Rev Microbiol. 13:722-736 (2015))、cas遺伝子含有量、casオペロン構造、Casタンパク質配列、および処理ステップに基づき、5つの主要なタイプおよび16の異なるサブタイプを包含する (Makarova et al. Biol Direct. 6:38 (2011); Makarova and Koonin Methods Mol Biol. 1311:47-75 (2015); Barrangou, R. Genome Biology 16:247 (2015))。I型およびI II型では、特別なCasエンドヌクレアーゼがpre-crRNAを処理し、それから、crRNAに相補の核酸を認識および切断することのできる巨大な多重-Casタンパク質複合体に集合する。I型システムは最も頻繁であり普及したシステムであり、Cascadeによって駆動されてPAMに依存性の様式でDNAを標的化して、シグネチャタンパク質Cas 3を用いて標的核酸を破壊する。異なるプロセスは、I II型CRISPR-Casシステムに関する。ここで、pre-CRNAは、トランス活性化crRNA (tracrRNA)がcrRNAのリピーター領域にハイブリダイズするメカニズムにより処理される。ハイブリダイズしたcrRNA-tracrRNAは、RNase IIIにより切断され、各スパーサーの5'末端を除去する第2のイベントが続き、tracrRNAとCas 9の両方と結合したままの成熟crRNAが生産される。成熟複合体は、それから、複合体内のスパーサー配列と相補の標的dsDNA配列 (「プロトスパーサー」配列)を探して両鎖を切断する。I II型システムでの複合体による標的の認識および切断は、crRNA-tracrRNA複合体上のスパーサー配列と標的「プロトスパーサー」配列との間で相補の配列を必要とするだけでなく、プロトスパーサー配列の3'末端に位置するプロトスパーサー隣接モチーフ (PAM)配列も必要とする。

【発明の概要】

【課題を解決するための手段】

【0007】

本発明の一態様は、必須遺伝子、非必須遺伝子、および／または消費されてよいゲノム

アイランドに関して、細菌細胞の集団をスクリーニングする方法を提供し、細菌細胞の前記集団に、（５'から３'に）リピートスペーサーーリピート配列または少なくとも１つのリピートスペーサー配列を含むＣＲＩＳＰＲアレイを含む異種核酸構築物を導入するステップ（ここで、前記リピートスペーサーーリピート配列または前記の少なくとも１つのリピートスペーサー配列のスペーサーは、前記集団の細菌細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、それにより、形質転換された細菌細胞の集団を生産する）；形質転換された細菌、古細菌、藻類または酵母細胞の集団における欠失の存在または不存在を判定するステップ（ここで、形質転換された細菌、古細菌または酵母細胞の集団における欠失の存在は、標的領域が非必須遺伝子および／または消費されてよいゲノムアイランド内に含まれることを意味し、形質転換された細菌、古細菌、藻類または酵母細胞の集団における欠失の不存在は、標的領域が必須遺伝子内に含まれることを意味する）を含む。本発明で有用なＣＲＩＳＰＲアレイは、Ⅰ型、Ⅱ型、Ⅲ型、Ⅳ型またはⅤ型ＣＲＩＳＰＲアレイであり得る。

10

【０００８】

本発明の第２の態様は、必須遺伝子、非必須遺伝子、および／または消費されてよいゲノムアイランドに関して、細菌、古細菌、藻類または酵母細胞の集団をスクリーニングする方法を提供し、細菌、古細菌、藻類または酵母細胞の集団に、（ａ）トランスコード化ＣＲＩＳＰＲ（*tracr*）核酸を含む異種核酸構築物、（ｂ）（５'から３'に）リピートスペーサーーリピート配列または少なくとも１つのリピートスペーサー配列を含むＣＲＩＳＰＲアレイを含む異種核酸構築物（ここで、前記リピートスペーサーーリピート配列または前記の少なくとも１つのリピートスペーサー配列のスペーサーは、前記集団の細菌、古細菌、藻類または酵母細胞のゲノム（染色体および／またはプラスミド）内の標的領域に実質的に相補なヌクレオチド配列を含む）、および（ｃ）*Cas9*ポリペプチドまたは*Cas9*ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物を導入するステップ（それにより、形質転換された細菌、古細菌、藻類または酵母細胞の集団を生産する）；および、形質転換された細菌、古細菌、藻類または酵母細胞の集団における欠失の存在または不存在を判定するステップ（ここで、形質転換された細菌、古細菌または酵母細胞の集団における欠失の存在は、標的領域が非必須遺伝子および／または消費されてよいゲノムアイランド内に含まれることを意味し、形質転換された細菌、古細菌、藻類または酵母細胞の集団における欠失の不存在は、標的領域が必須遺伝子内に含まれることを意味する）、を含む。

20

30

【０００９】

本発明の第３の態様は、細菌細胞の集団内の１つまたは複数の細菌細胞を死滅させる方法を提供し、細菌細胞の集団に、（５'から３'に）リピートスペーサーーリピート配列または少なくとも１つのリピートスペーサー配列を含むＣＲＩＳＰＲアレイ（*crRNA*、*crDNA*）を含む異種核酸構築物を導入するステップを含み、ここで、前記リピートスペーサーーリピート配列または前記の少なくとも１つのリピートスペーサー配列のスペーサーは、前記集団の細菌細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、それにより、細菌細胞の集団内の標的領域を含む１つまたは複数の細菌細胞を死滅させる。本発明で有用なＣＲＩＳＰＲアレイは、Ⅰ型、Ⅱ型、Ⅲ型、Ⅳ型またはⅤ型ＣＲＩＳＰＲアレイであり得る。

40

【００１０】

本発明の第４の態様は、細菌、古細菌、藻類または酵母細胞の集団内の１つまたは複数の細胞を死滅させる方法を提供し、細菌、古細菌、藻類または酵母細胞の集団に、（ａ）トランスコード化ＣＲＩＳＰＲ（*tracr*）核酸を含む異種核酸構築物、（ｂ）（５'から３'に）リピートスペーサーーリピート配列または少なくとも１つのリピートスペーサー配列を含むＣＲＩＳＰＲアレイを含む異種核酸構築物（ここで、前記リピートスペーサーーリピート配列または前記の少なくとも１つのリピートスペーサー配列のスペーサーは、前記集団の細菌、古細菌、藻類または酵母細胞のゲノム（染色体および／またはプラスミド）内の標的領域に実質的に相補なヌクレオチド配列を含む）、および、（

50

c) Cas9ポリペプチドおよび／またはCas9ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物を導入するステップを含み、それにより、それらのゲノム内に標的領域を含む細菌、古細菌、藻類または酵母細胞の集団内の1つまたは複数の細胞を死滅させる。

【0011】

本発明の第5の態様は、細菌遺伝子と関連する表現型を同定する方法を提供し、細菌細胞の集団に、(5'から3'に)リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイを含む異種核酸構築物を導入するステップ(ここで、少なくとも1つのリピートスペーサー配列およびリピートスペーサーリピート配列のスペーサーは、前記集団の細菌細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、ここで、標的領域は、ポリペプチドまたは機能性核酸をコードするオープンリーディングフレームの少なくとも一部を含み、それにより、標的領域を含む細胞を死滅させ、そして、標的領域を含まない形質転換された細菌細胞の集団を生産する)；および、集団の表現型を分析するステップ、を含む。本発明で有用なCRISPRアレイは、I型、II型、III型、IV型またはV型CRISPRアレイであり得る。

【0012】

本発明の第6の態様は、細菌、古細菌、藻類または酵母遺伝子の表現型を同定する方法を提供し、細菌、古細菌、藻類または酵母細胞の集団に、(a)トランスコード化CRISPR(transcr)核酸を含む異種核酸構築物、(b)(5'から3'に)リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイを含む異種核酸構築物(ここで、前記リピートスペーサーリピート配列または前記の少なくとも1つのリピートスペーサー配列のスペーサーは、前記集団の細菌、古細菌、藻類または酵母細胞のゲノム(染色体および／またはプラスミド)内の標的領域に実質的に相補なヌクレオチド配列を含む)、および(c)Cas9ポリペプチドおよび／またはCas9ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップ(それにより、標的領域を含む細菌、古細菌または酵母細胞を死滅させ、そして、標的領域を含まない形質転換された細菌、古細菌、藻類または酵母細胞の集団を生産する)；および、形質転換された細菌、古細菌、藻類または酵母細胞の集団の表現型を分析するステップ、および／または、(i)形質転換された細菌、古細菌、藻類または酵母細胞の集団から、個々の細菌、古細菌、藻類または酵母コロニーを増殖させるステップ、および、(ii)個々のコロニーの表現型を分析するステップ、を含む。

【0013】

本発明の第7の態様は、細菌細胞の集団から、縮小されたゲノムサイズを有する1つまたは複数の細菌細胞を選択する方法を提供し、細菌細胞の集団に、(5'から3'に)リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイを含む異種核酸構築物を導入するステップを含み、ここで、前記リピートスペーサーリピート配列または前記の少なくとも1つのリピートスペーサー配列のスペーサーは、前記集団の1つまたは複数の細菌細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、ここで、標的領域を含む細胞は死滅して、それにより、細菌細胞の集団から、標的領域を含まない縮小されたゲノムサイズを有する1つまたは複数の細菌細胞を選択する。本発明で有用なCRISPRアレイは、I型、II型、III型、IV型またはV型CRISPRアレイであり得る。

【0014】

本発明の第8の態様は、細菌細胞の集団から、縮小されたゲノムサイズを有する1つまたは複数の細菌細胞を選択する方法を提供し、細菌細胞の集団に、(a)(i)前記細菌細胞のゲノム内に存在する少なくとも300個の連続したヌクレオチドと少なくとも80%の同一性を有するヌクレオチド配列を含む、1つまたは複数の異種核酸構築物、または、(ii)少なくとも1つのトランスポゾンを含む2以上の異種核酸構築物(それにより、ゲノム中に組み込まれた前記の1つまたは複数の異種核酸構築物およびゲノム内に存在

する前記の少なくとも300個の連続したヌクレオチドの間、または、ゲノム中に組み込まれた第1および第2のトランスポゾンの中の、相同組み換えに関する非天然部位を含むトランスジェニック細菌細胞の集団を生産する)；および(b)(5'から3'に)リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイを含む異種核酸構築物、を導入するステップを含み、ここで、前記リピートスペーサーリピート配列または前記の少なくとも1つのリピートスペーサー配列のスペーサーは、前記集団の1つまたは複数の細菌細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、ここで、標的領域は、ゲノム中に導入された前記の1つまたは複数の異種核酸構築物およびゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、および/または、第1のトランスポゾンおよび第2のトランスポゾンの中に位置し、および、標的領域を含む細胞は死滅して、それにより、トランスジェニック細菌細胞の集団から、標的領域を有さない縮小されたゲノムサイズを有する1つまたは複数の細菌細胞を選択する。本発明で有用なCRISPRアレイは、I型、II型、III型、IV型またはV型CRISPRアレイであり得る。

【0015】

本発明の第9の態様は、細菌、古細菌、藻類または酵母細胞の集団から、縮小されたゲノムサイズを有する1つまたは複数の細菌、古細菌、藻類または酵母細胞を選択する方法を提供し、細菌、古細菌、藻類または酵母細胞の集団に、(a)トランスコード化CRISPR(tracr)核酸を含む異種核酸構築物、(b)リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイを含む異種核酸構築物(ここで、前記の少なくとも1つのリピートスペーサー配列および前記の少なくとも1つのリピートスペーサーリピート配列のスペーサーは、前記集団の細菌、古細菌、藻類または酵母細胞のゲノム(染色体および/またはプラスミド)内の標的領域に実質的に相補なヌクレオチド配列を含む)、および、(c)Cas9ポリペプチドおよび/またはCas9ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップを含み、ここで、標的領域を含む細胞は死滅して、それにより、細菌、古細菌、藻類または酵母細胞の集団から、標的領域を有さない縮小されたゲノムサイズを有する1つまたは複数の細菌、古細菌、藻類または酵母細胞を選択する。

【0016】

本発明の第10の態様は、細菌、古細菌または酵母細胞の集団から、縮小されたゲノムサイズを有する1つまたは複数の細菌、古細菌、藻類または酵母細胞を選択する方法を提供し、細菌、古細菌、藻類または酵母細胞の集団に、(a)(i)前記の細菌、古細菌、藻類または酵母細胞のゲノム内に存在する少なくとも300個の連続したヌクレオチドと少なくとも80%の同一性を有するヌクレオチド配列を含む、1つまたは複数の異種核酸構築物、または(ii)少なくとも1つのトランスポゾンを含む2以上の異種核酸構築物(それにより、ゲノム中に組み込まれた前記の1つまたは複数の異種核酸構築物およびゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、または、ゲノム中に組み込まれた第1および第2のトランスポゾンの中の、相同組み換えに関する非天然部位を含むトランスジェニック細菌、古細菌、藻類または酵母細胞の集団を生産する)；および(b)(i)トランスコード化CRISPR(tracr)核酸を含む異種核酸構築物、(ii)リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイを含む異種核酸構築物(ここで、前記の少なくとも1つのリピートスペーサー配列および前記の少なくとも1つのリピートスペーサーリピート配列のスペーサーは、前記集団の1つまたは複数の細菌、古細菌、藻類または酵母細胞のゲノム(染色体および/またはプラスミド)内の標的領域に実質的に相補なヌクレオチド配列を含む)、および(iii)Cas9ポリペプチドおよび/またはCas9ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップを含み、ここで、標的領域は、ゲノム中に取り込まれた前記の1つまたは複数の異種核酸構築物およびゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、および/または、第一のトランスポゾンおよび第二のトランスポゾンの中に位置

し、標的領域を含む細胞は死滅して、それにより、トランスジェニック細菌、古細菌、藻類または酵母細胞の集団から、標的領域を有さない縮小されたゲノムサイズを有する1つまたは複数の細菌、古細菌、藻類または酵母細胞を選択する。

【0017】

本発明の第11の態様は、細菌の集団において、そのゲノム内に欠失を有する少なくとも1つの単離物を同定する方法を提供し、細菌細胞の集団に、(5'から3'に)リピートスペーサーーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイを含む異種核酸構築物を導入するステップ(ここで、前記リピートスペーサーーリピート配列または前記の少なくとも1つのリピートスペーサー配列のスペーサーは、前記集団の1つまたは複数の細菌細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、ここで、標的領域を含む細胞は死滅して、それにより、標的領域を含まない形質転換された細菌細胞の集団を生産する)；および、形質転換された細菌細胞の集団から個々の細菌コロニーを増殖させるステップを含み、それにより、形質転換された細菌の集団から、そのゲノム内に欠失を有する少なくとも1つの単離物を同定する。本発明で有用なCRISPRアレイは、I型、II型、III型、IV型またはV型CRISPRアレイであり得る。

【0018】

本発明の第12の態様は、細菌の集団において、そのゲノム内に欠失を有する少なくとも1つの単離物を同定する方法を提供し、細菌細胞の集団に、(a)(i)前記細菌細胞のゲノム内に存在する少なくとも300個の連続したヌクレオチドと少なくとも80%の同一性を有するヌクレオチド配列を含む、1つまたは複数の異種核酸構築物、または、(ii)少なくとも1つのトランスポゾンを含む2以上の異種核酸構築物(それにより、ゲノム中に組み込まれた前記の1つまたは複数の異種核酸構築物およびゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、または、ゲノム中に組み込まれた第1および第2のトランスポゾンの間の、相同組み換えに関する非天然部位を含むトランスジェニック細菌細胞の集団を生産する)；およびb)(5'から3'に)リピートスペーサーーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイを含む異種核酸構築物を導入するステップ(ここで、前記リピートスペーサーーリピート配列または前記の少なくとも1つのリピートスペーサー配列のスペーサーは、前記集団の1つまたは複数の細菌細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、ここで、標的領域は、ゲノム中に導入された前記の1つまたは複数の異種核酸構築物およびゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、および/または、第1のトランスポゾンおよび第2のトランスポゾンの間に位置し、標的領域を含む細胞は死滅して、それにより、標的領域を含まない形質転換された細菌細胞の集団を生産する)；および、形質転換された細菌細胞の集団から個々の細菌コロニーを増殖させるステップを含み、それにより、細菌の集団から、そのゲノム内に欠失を有する少なくとも1つの単離物を同定する。本発明で有用なCRISPRアレイは、I型、II型、III型、IV型またはV型CRISPRアレイであり得る。

【0019】

本発明の第13の態様は、細菌、古細菌、藻類または酵母細胞の集団において、そのゲノム内に欠失を有する少なくとも1つの単離物を同定する方法を提供し、細菌、古細菌、藻類または酵母細胞の集団に：(a)トランスコード化CRISPR(transcript)核酸を含む異種核酸構築物、(b)(5'から3'に)リピートスペーサーーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイを含む異種核酸構築物(ここで、前記リピートスペーサーーリピート配列または前記の少なくとも1つのリピートスペーサー配列のスペーサーは、前記集団の細菌、古細菌、藻類または酵母細胞のゲノム(染色体および/またはプラスミド)内の標的領域に実質的に相補なヌクレオチド配列を含む)、および(c)Cas9ポリペプチドまたはCas9ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップ(ここで、標的領域を含む細胞は死滅して、それにより、標的領域を含まない形質転換された細菌、

古細菌、藻類または酵母細胞の集団を生産する) ; および、形質転換された細菌、古細菌、藻類または酵母細胞の集団から、個々の細菌、古細菌または酵母コロニーを増殖させるステップを含み、それにより、形質転換された細菌、古細菌、藻類または酵母細胞の集団から、そのゲノム内に欠失を有する少なくとも1つの単離物を同定する。

【0020】

本発明の第14の態様は、細菌、古細菌、藻類または酵母細胞の集団において、そのゲノム内に欠失を有する少なくとも1つの単離物を同定する方法を提供し、細菌、古細菌、藻類または酵母細胞の集団に (a) (i) 前記の細菌、古細菌、藻類または酵母細胞のゲノム内に存在する少なくとも300個の連続したヌクレオチドと少なくとも80%の同一性を有するヌクレオチド配列を含む、1つまたは複数の異種核酸構築物、または (ii) 少なくとも1つのトランスポゾンを含む2以上の異種核酸構築物 (それにより、ゲノム中に組み込まれた前記の1つまたは複数の異種核酸構築物およびゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、または、ゲノム中に組み込まれた第1および第2のトランスポゾンの間の、相同組み換えに関する非天然部位を含むトランスジェニック細菌、古細菌、藻類または酵母細胞の集団を生産する) ; および (b) (i) トランスコード化CRISPR (tracr) 核酸を含む異種核酸構築物、(ii) (5' から3' に) リピートスパーサーリピート配列または少なくとも1つのリピートスパーサー配列を含むCRISPRアレイを含む異種核酸構築物 (ここで、前記リピートスパーサーリピート配列または前記の少なくとも1つのリピートスパーサー配列のスパーサーは、前記集団の1つまたは複数の細菌、古細菌、藻類または酵母細胞のゲノム (染色体および/またはプラスミド) 内の標的領域に実質的に相補なヌクレオチド配列を含む) 、および、(iii) Cas9ポリペプチドおよび/またはCas9ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップ (ここで、標的領域は、ゲノム中に組み込まれた前記の1つまたは複数の異種核酸構築物およびゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、および/または、第一のトランスポゾンおよび第二のトランスポゾンの間に位置し、標的領域を含む細胞は死滅して、それにより、標的領域を有さない形質転換された細菌、古細菌、藻類または酵母細胞の集団を生産する) ; および、形質転換された細菌、古細菌、藻類または酵母細胞の集団から、個々の細菌、古細菌または酵母コロニーを増殖させるステップを含み、それにより、集団から、そのゲノム内に欠失を有する少なくとも1つの単離物を同定する。

【0021】

本明細書においてさらに提供されるのは、発現カセット、本発明の核酸構築物、核酸アレイ、核酸分子および/またはヌクレオチド配列を含む細胞およびキットである。

【0022】

本発明のこれらおよび他の態様は、以下の本発明の説明において詳細に示される。

【図面の簡単な説明】

【0023】

【図1】それぞれの混成トランスポゾンを標的化するための、St hCRISPR1およびSt hCRISPR3アレイの配列を示す。

【図2】標的化プラスミドの構築のためのスプライシングオーバーラップエクステンション (SOE) 法の概略を示す。

【図3A-3E】必須遺伝子、挿入配列およびゲノムアイランドのマップを示す。(A) 推定の必須ORF (一番上の矢印)、挿入配列 (2番目の矢印) および推定のゲノムアイランド (3番目の矢印) の位置および分配。CRISPR-Cas (4番目の矢印) が仲介する欠失に関する潜在的標的を、Streptococcus thermophilus LMD-9のゲノム内の様々なファミリーの転位性エレメントをマッピングすることにより同定した。推定のゲノムアイランドおよび互いに対応するプロトスパーサー/PAMの組み合わせの遺伝的構成。(B) オリゴペプチド輸送システムをコードする、ゲノムアイランド1。(C) 細胞-エンベローププロテイナーゼPrtSを含む、ゲノムアイランド2。(D) ATPアーゼ銅-流出タンパク質をコードする、ゲノムアイランド3。

(E) *Lac* オペロンを含む選択された ORF をコードする、ゲノムアイランド。

【図 4】 *S. thermophilus* LMD-9 のゲノム全体に分配されたトランスポゾンコード配列の樹状図を提供する。アライメントは、Geneious (登録商標) ソフトウェアを用いて作成された。ファミリー指定は、www-is.bioutil.fr. を用いて割り当てられた。文字は、それぞれのファミリーに関して図 5 におけるアライメントに対応する。

【図 5 A-5 D】 *S. thermophilus* LMD-9 ゲノムに見られるそれぞれの主な IS ファミリーに関する、トランスポゾンコード配列のアライメント (Geneious (登録商標) ソフトウェアを用いた) を示す。異なるファミリーは、異なるレベルの長さの保存およびヌクレオチド同一性を示した。(A) *St h 6* トランスポゾンは、非常に多型であり、一部のコピー内の内部欠失に起因して見かけ上縮重する。対照的に、IS 1167 (B) および IS 1191 (C) は、より少ないコピーを有したが、長さおよび同一性の高い忠実性を維持した。IS 1193 (D) は、高い忠実性のコピーを有したが、長さのファミリー間多様性を最も高く示した。

【0024】

【図 6 A-6 C】 CRISPR-Cas9 による DNA ターゲティングの、見かけ上の細胞毒性および構造基礎を示す。*lacZ* を標的化するための、*St h CRISPR1* (A) および *St h CRISPR3* (B) に関するスペーサー配列。Cas9 は、DNA を調べて、*tracrRNA*:*crRNA* デュプレックスの形成を介して生じる標的における Cas9 の安定化を伴って PAM 配列に可逆的に結合する。Cas9 の活性化は、黒のウェッジによって示されるように、RuvC および HNHドメインによってそれぞれの鎖の同時の切断を生じさせる。コントロールおよび自己標的化プラスミドのエレクトロポレーションの後に回収された、形質転換体 (C)。試験されたそれぞれのプラスミドに関して非依存性の形質転換実験 ($n=4$) にわたってスクリーニングされた、平均クローン $\pm$ SD。

【図 7 A-7 D】 *Lac* - クローンのゲノムシーケンシングおよび表現型分析を示す。配列データは、CRISPR3 システムを用いて (A) 5' 末端、および、(B) *lacZ* のカチオン結合残基コード配列を標的化することにより独立に作成された 2 つの突然変異体において、*lacZ* をコードする染色体セグメントの不存在を明らかにした。欠失のサイズは、101, 865~102, 146 bp の長さの範囲であり、*S. thermophilus* のゲノムのおよそ 5.5% を構成する。(C) 3 つの非依存性の生物学的再現の平均 $\pm$ SD OD 600 nm として表された、半合成の *Elliker* 培地における、野生型と比較した大きな欠失株の増殖。(D) スキムミルク中での *S. thermophilus* 株の酸性化能力。

【0025】

【図 8 A-8 E】 挿入配列 (IS) の間の組み換えイベントの図解を提供する。(A) 形質転換体から回収した gDNA の PCR 分析によって生産された、大きな欠失アンプリコンのゲル電気泳動写真。推定の欠失部位の上流および下流の IS 1193 エlement に隣接するプライマーを用いてスクリーニングを行なった。 Δ で示されるレーンは、CRISPR-Cas が仲介する *lacZ* の標的化の後に回収された *Lac* - クローンの gDNA から増幅されたが、一方で、WT は野生型由来である。(B) 予測の組み換え部位の配列を、上流 (灰色) または下流 (黒) IS エlement のいずれかに対応する一塩基多型をマッピングすることによって決定した。3 つの部位は、両方の IS エlement において保存されている配列に基づいて予測される (薄い灰色)。示される部位は、非依存性クローン由来の遺伝子型を示し、9 個の異なる組み換え部位で観察される代表的な *Lac* - 現象である。キメラ IS エlement のフットプリントは、欠失ジャンクションのそれぞれのゲノムアイランド遺伝子座に同様に見られた。(C) *lacZ* をコードするアイランドの染色体欠失の間で組み換えることが予測される IS のスキーム。(D) 欠失を確認するために、ゲノムアイランド 1、2、および 3 と隣接するプライマーから生産されたアンプリコン。(E) それぞれの、CRISPR によって誘導される欠失培養物における、野生型配列

の不存在を確認するための、内部プライマーから生産されたアンプリコン。Δで示されるレーンは、CRISPR-Casが仲介する標的化の後に回収されたクローンのgDNAから増幅された。WTは野生型である。

【図9】致死性の標的を提供し、*Streptococcus thermophilus* LMD-9のゲノム標的化を介したII型CRISPR-Casシステムに基づく致死性を評価するための規定された遺伝子座の使用を示す。両方のオルトゴナルII型システム（CRISPR1およびCRISPR3）を試験した；CRISPR1標的は濃い灰色において標的化し、CRISPR3は薄い灰色において標的化する。特異的な遺伝的特徴を選択して、(i) 遺伝子間領域（INT）、(ii) 可動遺伝要素（ISSth7、oppC-GEI1、prtS-GEI2、copA-GEI3、cI、lacZ-GEI4、epsU）、(iii) 必須遺伝子（dltA、ltas）、(iv) レプリコアのポール（poles of the replichore）（OriC、xerS）、および、DNAのフォワード、対、リバーズ鎖（外側標的、対、内側標的）を試験した。

【図10】図9で定義された領域を標的化することによって得られる、CRISPRに基づく致死性を示す。CFU（細胞形成単位）のLog減少（Log reduction）を、非標的化プラスミドコントロール；pORI28の形質転換に関して計算した。致死性は、染色体位置、コード配列、または必須性にかかわらず、試験されたすべての標的に関して、2〜3 log減少の範囲だった。ISSth7-挿入配列エレメント、ltas-リポタイコ酸シンターゼ；prtS-ゲノムアイランド2；INT-遺伝子間領域；dltA-D-アラニンリガーゼ；rheB-chi部位欠失遺伝子座；oppC-ゲノムアイランド1；comS-chi部位密集遺伝子座；複製のxerS末端；copA-ゲノムアイランド3；cI-プロファージのレムナント；OriC-複製起点；Cas9-CRISPR3 Cas9コード配列；epsU-エキソポリサッカライドカセット。

【図11】CRISPRが仲介するゲノムアイランド欠失株の転写プロファイルを示す。

【図12】ゲノムアイランド欠失株、GEI1、GEI2、GEI3、およびGEI4の、log₂形質転換されたRNA配列解読範囲（read coverage）を示す。

【図13】ゲノムアイランド欠失株発現値（X軸）、対、野生型発現値（Y軸）のXYプロットを示す。ゲノムアイランド欠失株（GEI1〜GEI4）のそれぞれに関して、標的アイランド（黒）のそれぞれの上にコードされた遺伝子の発現は最小限であった。GEI1内にコードされた遺伝子を上の左パネルに示し、GEI2内にコードされた遺伝子を上の右パネルに示し、GEI3内にコードされた遺伝子を下の左パネルに示し、GEI4内にコードされた遺伝子を下の右パネルに示す。

【図14A-14B】*Streptococcus thermophiles*において、プログラムされた細胞死に関する内因性システムを利用（co-opt）するための、CRISPRアレイ（II型システム）をコードする外因性のファージ、プラスミドまたはファージミドの導入を示す。

【図15】*Lactobacillus casei*のII型ガイド。上の構造は予測のガイドである。下の左の図は、RNAシーケンシングによって確認した、正確なデュアルガイドcrRNA：tracrRNAである。下の右の図は、予測の人工的な単一ガイドの例である。

【図16】*Lactobacillus gasserii*の例示的なII型ガイドを提供する。上の構造は予測のガイドである。下の左の図は、RNAシーケンシングによって確認した、正確なデュアルガイドcrRNA：tracrRNAである。下の右の図は、予測の人工的な単一ガイドの例である。

【図17】*Lactobacillus pentosus*の例示的なII型ガイドを提供する。上の構造は予測のガイドである。下の左の図は、RNAシーケンシングによって確認した、正確なデュアルガイドcrRNA：tracrRNAである。下の右の図は、予測の人工的な単一ガイドの例である。

10

20

30

40

50

【図18】*Lactobacillus jensenii*の例示的なII型ガイドを提供する。左のパネルは、RNAシーケンスによって確認した、正確なデュアルガイドcrRNA：tracrRNAである。右のパネルは、予測の人工的な単一ガイドの例を提供する。

【図19】ネイティブの*L. gasseri* II型CRISPRアレイにおける、転写crRNAと最も高くマッチする(matches the most highly transcribe crRNA)プロトスペーサーを含むプラスミドの形質転換の結果を示す。左から右に、4つの異なるプラスミドを*L. gasseri*に形質転換した：空のpTRK563ベクター、正確なプロトスペーサーであるが不正確なPAMを有する構築物、正確なPAMであるがアレイ内でないプロトスペーサー、および、最も干渉の標的化および細胞死を示したPAMを有する正確なプロトスペーサー。報告される値は、3つの非依存性の再現の平均±SEMを表わす。

10

【図20】ネイティブの*L. pentosus* II型CRISPRアレイにおける、転写crRNAと最も高くマッチするプロトスペーサーを含むプラスミドの形質転換を示す。左から右に、4つの異なるプラスミドを*L. pentosus*に形質転換した：正確なプロトスペーサーであるが不正確なPAMを有する構築物(Lpe4 ctGttt)、正確なPAMであるがアレイ内でないプロトスペーサー(Lpe8 noSPCR)、空のpTRK563ベクター(pTRK563)、および、正確なプロトスペーサーおよび正確なPAMを有するプラスミド(Lpe1 gttaat)。報告される値は、3つの非依存性の再現の平均±SEMを表わす。

20

【図21】*Lactobacillus casei*の例示的なI型CRISPR-Casガイドを提供する。提供された配列は、*Lactobacillus casei* NCK 125に見られるネイティブのI型リーダーおよびリピートである。この人工的なアレイは、宿主ゲノム内の16s rDNA遺伝子を標的化するスペーサーを含む。

【図22】ネイティブの*L. jensenii* II型CRISPRアレイにおける、転写crRNAと最も高くマッチするプロトスペーサーを含むプラスミドの形質転換を示す。左から右に、4つの異なるプラスミドを*L. jensenii*に形質転換した：空のpTRK563ベクター、正確なプロトスペーサーであるが不正確なPAMを有する構築物、正確なPAMであるがアレイ内でないプロトスペーサー、および、最も干渉の標的化および細胞死が示されたPAMを有する正確なプロトスペーサー。

30

【図23】*Lactobacillus casei* NCK 125におけるネイティブのI型システムを用いた、標的化自殺を示す。2つの標的を16s rDNA遺伝子内に設計した。PAM5'-YAA-3'は、生物内のネイティブのスペーサー配列を用いて予測した。ネイティブのI型リーダー、リピート、および、選択スペーサーを含む人工的なアレイを、pTRK870にクローン化した。導入された構築物は、空のベクター(pTRK563)および2つの異なる人工的アレイを含んだ：一方は、16s遺伝子内の+鎖を標的化する単一スペーサーを含み(1-2 alt)、他方のアレイは、+鎖を標的化するオリジナルのスペーサーを含むが、16s遺伝子内の-鎖を標的化するさらなるスペーサーを含む(1, 2-3)。報告される値は、3つの非依存性の再現の平均±SEMを表わす。

40

【発明を実施するための形態】

【0026】

ここで、本発明は、以降に、付随する図面および実施例に関して説明されて、そこに本発明の実施態様を示される。この説明は、本発明が実施され得る全ての異なる手段または本発明に加えられ得る全ての特徴の、詳細な一覧であることを意図しない。例えば、一実施態様に関して説明される特徴は、他の実施態様に取り込まれ得て、特定の実施態様に関して説明される特徴は、その実施態様から取り除かれ得る。したがって、本発明は、本発明の一部の実施態様では、本明細書において示される任意の特徴または特徴の組み合わせを除外または省略することができることを検討する。加えて、本明細書において示唆される様々な実施態様に対する非常に多くの変更および追加は、本開示に照らして当業者に明

50

らかであり、本発明から逸脱しない。それ故に、以下の説明は、本発明の一部の特定の実施態様を説明することを意図し、それらの全ての並べ替え、組み合わせ、および変更を徹底的に特定することを意図しない。

【0027】

別段の定義がされない限り、本明細書で用いられる全ての技術的および科学的用語は、本発明が属する当業者の一人により一般に理解されるのと同じの意味を有する。本発明において、本発明の説明に用いられる専門用語は、特定の実施態様を説明する目的のためのものであり、本発明の限定を意図しない。

【0028】

本明細書において引用される全ての刊行物、特許出願、特許および他の参考文献は、参照が提供される文章および／または段落に関連のある教示に関して、それらの全体で参照により援用される。

【0029】

文脈が別段の指示をしない限り、本明細書に説明される本発明の様々な特徴は、任意の組み合わせで用いることができることが明確に意図される。さらにまた、本発明は、本発明の一部の実施態様では、本明細書において示される任意の特徴または特徴の組み合わせは、除外または省略することができることも検討する。説明すると、明細書が、組成物は構成要素A、BおよびCを含む、と述べる場合、A、BまたはCのいずれか、またはそれらの組み合わせが、単独でまたは任意の組み合わせで、省略および放棄され得ることが明確に意図される。

【0030】

本発明の説明および添付の特許請求の範囲において用いられる、単数形「a」、「an」および「the」は、文脈が明確に別段の指示をしない限り、複数形を同様に含むことが意図される。

【0031】

また、本明細書において用いられる「および／または」は、関係がある列挙された項目の1つまたは複数の任意および全てのあり得る組み合わせ、ならびに、選択肢で（「または」）解釈された場合は組み合わせの欠失を指し、および包含する。

【0032】

本明細書において用いられる用語「約」は、例えば用量または期間のような測定可能な値を言及する場合、規定の量の $\pm 20\%$ 、 $\pm 10\%$ 、 $\pm 5\%$ 、 $\pm 1\%$ 、 $\pm 0.5\%$ 、または実に $\pm 0.1\%$ の変動を指す。

【0033】

本明細書において用いられる、「XおよびYの間」および「約XおよびYの間」のような語句は、XおよびYを含むと解釈されるべきである。本明細書において用いられる「約XおよびYの間」のような語句は、「約Xおよび約Yの間」を意味し、「約XからYまで」のような語句は、「約Xから約Yまで」を意味する。

【0034】

本明細書において用いられる用語「含む（comprise）」、「含む（comprises）」、および、「含む（comprising）」は、述べられた特徴、整数、ステップ、操作、エレメント、および／または、構成要素の存在を特定するが、1つまたは複数の他の特徴、整数、ステップ、操作、エレメント、構成要素、および／または、それらの群の存在または追加を排除しない。

【0035】

本明細書において用いられる、移行句「から本質的になる」は、クレームの範囲が、クレーム内に挙げられた規定の材料またはステップおよびクレームされた発明の基本的および新規の特徴（単数または複数）に実質的に影響を及ぼさないものを包含すると解釈すべきであることを意味する。したがって、用語「から本質的になる」は、本発明のクレームにおいて用いられる場合、「含む（comprising）」と同等であると解釈されることを意図しない。

10

20

30

40

50

【0036】

「Cas9ヌクレアーゼ」は、CRISPR-Casシステムにおいて、二本鎖DNA切断を触媒するエンドヌクレアーゼの大きなグループを指す。これらのポリペプチドは、当技術分野でよく知られていて、それらの構造（配列）の多くが特徴付けられている（例えば、WO2013/176772；WO/2013/188638を参照）。二本鎖DNAの切断を触媒するドメインは、RuvCドメインおよびHNHドメインである。RuvCドメインは、（－）鎖にニックを入れるのを担い、HNHドメインは、（＋）鎖にニックを入れるのを担う（例えば、Gasiunas et al., PNAS 109（36）：E2579－E2586（September 4, 2012）を参照）。

【0037】

本明細書において用いられる「キメラ」は、少なくとも2つの構成要素が、異なる供給源（例えば、異なる生物、異なるコード領域）に由来する、核酸分子またはポリペプチドを指す。

【0038】

本明細書において用いられる「相補体」は、コンパレータヌクレオチド配列との100%相補性または同一性を意味することができ、または、100%未満の相補性（例えば、約70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%などの相補性）を意味することができる。

【0039】

本明細書において用いられる用語「相補」または「相補性」は、塩基対合による、許容的な塩および温度条件下での、ポリヌクレオチドの自然な結合を指す。例えば、配列「A－G－T」は、相補配列「T－C－A」と結合する。2つの一本鎖分子間の相補性は、ヌクレオチドの一部のみが結合する「部分的」であってよく、または、一本鎖分子間に全体的な相補性が存在する場合は完全であってよい。核酸鎖間の相補性の程度は、核酸鎖間のハイブリダイゼーションの効率および強度に重大な効果を有する。

【0040】

本明細書において用いられる「接触する（contact）」、「接触する（contacting）」、「接触される（contacted）」、および、それらの文法的変形は、所望の反応（例えば、組込み、形質転換、スクリーニング、選択、死滅、同定、増幅など）を行なうのに適切な条件下に、所望の反応の構成要素と一緒に置くことを指す。そのような反応を行なうための方法および条件は、当技術分野でよく知られている（例えば、Gasiunas et al., (2012) Proc. Natl. Acad. Sci. 109：E2579－E2586；M. R. Green and J. Sambrook (2012) Molecular Cloning：A Laboratory Manual, 4th Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NYを参照）。

【0041】

本明細書において用いられる「欠失」は、制限されないが、染色体またはプラスミドの一部の欠失、染色体またはプラスミド由来の遺伝子または遺伝子の一部の欠失を含む、遺伝物質の喪失または欠失を含むことができる。一部の実施態様では、欠失は、1つの遺伝子または1つよりも多い遺伝子を含むことができる。一部の実施態様では、欠失は、小さな非コードRNAをコードしてよい非タンパク質コード領域の喪失を含んでもよい。一部の実施態様では、欠失は、プラスミド全体または可動遺伝要素全体の喪失を含んでもよい。一部の実施態様では、可動遺伝要素の喪失は、例えば、複製または存続の不可能性として定義され得る。

【0042】

一部の実施態様では、本発明のファスミドは、I型CRISPR-Casシステム、II型CRISPR-Casシステム、III型CRISPR-Casシステム、IV型C

10

20

30

40

50

R I S P R-C a s システム、および／または、V型C R I S P R-C a s システム由来の、C R I S P Rアレイを含んでよい (Makarova et al. Nature Reviews Biotechnology 13:722736 (2015) を参照)。

【0043】

したがって、一部の実施態様では、I型c r R N Aに加えて、本発明のファスミドは、I型ポリペプチドおよび／またはI型C a s c a d eポリペプチド（すなわち、I型C R I S P R-C a s システム）を含んでよい。

【0044】

本明細書において用いられる「I型ポリペプチド」は、C a s 3ポリペプチド、C a s 3'ポリペプチド、C a s 3''ポリペプチド、それらの融合変異体、および、任意の1つまたは複数の、抗ウイルス防御のためのI型C l u s t e r e d R e g u l a r l y I n t e r s p a c e d S h o r t P a l i n d r o m i c R e p e a t s (C R I S P R) 関連複合体（「C a s c a d e」）ポリペプチドのいずれかを指す。したがって、用語「I型ポリペプチド」は、I-A型C R I S P R-C a s システム、I-B型C R I S P R-C a s システム、I-C型C R I S P R-C a s システム、I-D型C R I S P R-C a s システム、I-E型C R I S P R-C a s システム、I-F型C R I S P R-C a s システム、および／または、I-U型C R I S P R-C a s システムを作り上げるポリペプチドを指す。それぞれのI型C R I S P R-C a s システムは、少なくとも1つのC a s 3ポリペプチドを含む。C a s 3ポリペプチドは、一般に、ヘリカーゼドメインおよびH Dドメインの両方を含む。しかしながら、一部のI型C R I S P R-C a s システムでは、ヘリカーゼおよびH Dドメインは、別個のポリペプチド、C a s 3'およびC a s 3''内に見られる。具体的には、C a s 3'はヘリカーゼドメインをコードし、一方で、C a s 3''はH Dドメインをコードする。その結果として、両方のドメインがC a s 3機能に必要とされるので、I型サブタイプは、C a s 3 (I-C、I-D、I-E、I-F、I-U) またはC a s 3'およびC a s 3'' (I-A、I-B) のいずれかをコードする。

【0045】

本明細書において用いられる「I型C a s c a d eポリペプチド」は、I型C R I S P R-C a s システムにおける、P r e-c r R N Aのプロセッシングおよびその後の標的D N Aに対する結合に関与するポリペプチドの複合体を指す。これらのポリペプチドは、限定されないが、I型サブタイプI-A、I-B、I-C、I-D、I-EおよびI-FのC a s c a d eポリペプチドを含む。I-A型C a s c a d eポリペプチドの非限定の例は、C a s 7 (C s a 2)、C a s 8 a 1 (C s x 1 3)、C a s 8 a 2 (C s x 9)、C a s 5、C s a 5、C a s 6 a、C a s 3'および／またはC a s 3''を含む。I-B型C a s c a d eポリペプチドの非限定の例は、C a s 6 b、C a s 8 b (C s h 1)、C a s 7 (C s h 2) および／またはC a s 5を含む。I-C型C a s c a d eポリペプチドの非限定の例は、C a s 5 d、C a s 8 c (C s d 1)、および／またはC a s 7 (C s d 2)を含む。I-D型C a s c a d eポリペプチドの非限定の例は、C a s 1 0 d (C s c 3)、C s c 2、C s c 1、および／またはC a s 6 dを含む。I-E型C a s c a d eポリペプチドの非限定の例は、C s e 1 (C a s A)、C s e 2 (C a s B)、C a s 7 (C a s C)、C a s 5 (C a s D) および／またはC a s 6 e (C a s E)を含む。I-F型C a s c a d eポリペプチドの非限定の例は、C y s 1、C y s 2、C a s 7 (C y s 3) および／またはC a s 6 f (C s y 4)を含む。I-U型C a s c a d eポリペプチドの非限定の例は、C a s 8 c、C a s 7、C a s 5、C a s 6 および／またはC a s 4を含む。

【0046】

一部の実施態様では、本発明のファスミドは、I I型c r R N Aに加えて、I I型C R I S P R-C a s システムを含んでよい。I I型C R I S P R-C a s システムは、3つのサブタイプ：I I-A型、I I-B型およびI I-C型を含み、それぞれ、多重ドメイ

ンタンパク質、Cas 9を、適合ポリペプチド、Cas 1、Cas 2および場合により、Csn 2および／またはCas 4に加えて含む。また、ほとんどのII型遺伝子座は、tracrRNAもコードする。例示的なII型CRISPR-Casシステムを含む生物は、Legionella pneumophila str. Paris、Streptococcus thermophilus CNRZ1066およびNeisseria lactamica 020-06を含む。

【0047】

さらなる実施態様では、本発明のファスミドは、III型crRNAに加えて、III型CRISPR-Casシステムを含んでよい。I型CRISPR-Casシステムと同様に、III型システムでは、プロセッシングおよび干渉は、多重タンパク質CRISPR RNA (crRNA) -エフェクター複合体によって仲介される (Makarova et al. Nature Reviews Biotechnology 13: 722736 (2015)) - I型における「CASCADE」およびIII型における「Csm」または「Cmr」。したがって、一部の実施態様では、III型CRISPR-Casシステムは、Csm複合体 (例えば、III-A型Csm) および／またはCmr複合体 (例えば、III-B型Cmr)、および場合により、Cas 6ポリペプチドを含むことができる。代表的な実施態様では、Csm複合体は、Cas 10 (またはCsm1)、Csm2、Csm3、Csm4、Csm5、およびCsm6ポリペプチドを含んでよく、Cmr複合体は、Cmr1、Cas 10 (またはCmr2)、Cmr3、Cmr4、Cmr5、およびCmr6ポリペプチドを含んでよい。Csm複合体またはCmr複合体に加えて、III型CRISPR-Casシステムは、Cas 7ポリペプチドをさらに含んでよい。III型CRISPR-Casシステムの4つのサブタイプ、III-A、III-B、III-C、III-Dが特徴付けられている。一部の実施態様では、III-A型CRISPR-Casシステムは、Cas 6、Cas 10、Csm2、Cas 7 (Csm3)、Cas 5 (Csm4)、Cas 7 (Csm5)、およびCsm6ポリペプチドを含む。一部の実施態様では、III-B型CRISPR-Casシステムは、Cas 7 (Cmr1)、Cas 10、Cas 5 (Cmr3)、Cas 7 (Cmr4)、Cmr5、Cas 6、およびCas 7 (Cmr6) ポリペプチドを含む。一部の実施態様では、III-C型CRISPR-Casシステムは、Cas 7 (Cmr1)、Cas 7 (Cmr6)、Cas 10、Cas 7 (Cmr4)、Cmr5およびCas 5 (Cmr3)、ポリペプチドを含む。一部の実施態様では、III-D型CRISPR-Casシステムは、Cas 10、Cas 7 (Csm3)、Cas 5 (Cs×10)、Csm2、Cas 7 (Csm3)、およびall1473ポリペプチドを含む。

【0048】

一部の実施態様では、本発明のファスミドは、IV型crRNAに加えて、IV型CRISPR-Casシステムを含んでよい。IV型CRISPR-Casシステムは、Csf4ポリペプチド (dinG) および／またはCsf1、Cas 7 (Csf2) および／またはCas 5 (csf3) ポリペプチドを含むことができる (Makarova et al. Nature Reviews Microbiology 13: 722-736 (2015))。

【0049】

一部の実施態様では、本発明のファスミドは、V型crRNAに加えて、V型CRISPR-Casシステムをさらに含む。V型CRISPR-Casシステムは、Cpf1ポリペプチドおよび／またはCas 1、Cas 2および／またはCas 4ポリペプチドを含むことができる (Makarova et al. Nature Reviews Microbiology 13: 722-736 (2015))。

【0050】

本発明のヌクレオチド配列の「フラグメント」または「部分」は、参照核酸またはヌクレオチド配列に対して縮小された長さのヌクレオチド配列を意味すると理解され (例えば、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、

10

20

30

40

50

17、18、19、20個、またはそれよりも多くのヌクレオチドが縮小される）、参照核酸またはヌクレオチド配列と同一または実質的に同一（例えば、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%同一）の連続したヌクレオチドのヌクレオチド配列を含む、から本質的になる、および／またはからなる。本発明に係るそのような核酸フラグメントまたは部分は、適切な場合、それが構成成分である、より大きなポリヌクレオチド内に含まれてよい。したがって、例えば、標的DNA（例えば、ゲノム内の標的領域）の少なくとも一部にハイブリダイズすること（hybridizing）（または、ハイブリダイズする（hybridizes）、および、それらの他の文法的変形）は、標的DNAの連続したヌクレオチドの長さと同じまたは実質的に同一であるヌクレオチド配列に対するハイブリダイゼーションを指す。一部の実施態様では、リピートスペーサー配列またはリピートスペーサーリピート配列のリピートは、野生型CRISPR遺伝子座のリピート配列または合成のCRISPRアレイのリピート配列のフラグメントを含むことができ、ここで、リピートのフラグメントは、tracr核酸とハイブリダイズするCRISPRアレイにおけるリピートの機能を維持する。

【0051】

一部の実施態様では、本発明は、Cas9、Cas3、Cas3'、Cas3''、またはCpf1ヌクレアーゼの機能性フラグメントを含んでよい。Cas9機能性フラグメントは、制限されないが、HNHヌクレアーゼ活性、RuvCヌクレアーゼ活性、DNA、RNAおよび／またはPAM認識および結合活性を含む、ネイティブのCas9ヌクレアーゼの活性の1つまたは複数を維持する。Cas9ヌクレアーゼの機能性フラグメントは、Cas9ポリヌクレオチドのフラグメントによってコードされてよい。Cas3、Cas3'またはCas3''機能性フラグメントは、制限されないが、ニッカーゼ活性、エキソヌクレアーゼ活性、DNA結合、および／またはRNA結合を含む、ネイティブのCas9ヌクレアーゼの活性の1つまたは複数を維持する。Cas3、Cas3'またはCas3''ヌクレアーゼの機能性フラグメントは、Cas3、Cas3'またはCas3''ポリヌクレオチドのフラグメントによって、それぞれコードされてよい。

【0052】

本明細書において用いられる用語「遺伝子」は、mRNA、アンチセンスRNA、RNAi（miRNA、siRNA、shRNA）、抗マイクロRNAアンチセンスオリゴデオキシリボヌクレオチド（AMO）などを生産するために用いることが可能な核酸分子を指す。遺伝子は、機能性タンパク質または遺伝子産物を生産するための用いることが可能であってよく、または可能でなくてよい。遺伝子は、コードおよび非コード領域の両方（例えば、イントロン、調節エレメント、プロモーター、エンハンサー、終結配列および／または5'および3'非翻訳領域）を含むことができる。遺伝子は、「単離」されてよく、それによって、その天然の状態と関連して通常見られる構成要素から実質的または本質的にフリーである核酸が意味される。そのような構成要素は、他の細胞物質、組み換え生産による培養培地、および／または、核酸の化学合成で用いられる様々な化学物質を含む。

【0053】

本明細書において用いられる用語「ゲノム」は、生物の染色体／核ゲノムならびに任意のミトコンドリアおよび／またはプラスミドゲノムを含む。

【0054】

本明細書において用いられる「ヘアピン配列」は、ヘアピンを含むヌクレオチド配列（例えば、1つまたは複数のヘアピン構造を形成するもの）である。ヘアピン（例えば、ステムループ、フォールドバック）は、一本鎖領域によっていずれか側にさらに隣接される二本鎖を形成する相補ヌクレオチドの領域を含む二次構造を有する核酸分子を指す。そのような構造は、当技術分野でよく知られている。当技術分野で知られているように、二本鎖領域は、塩基対合内にいくつかのミスマッチを含むことができ、または、完全に相補で

あることができる。本開示の一部の実施態様では、核酸構築物のヘアピン配列は、t r a c r 核酸の3'末端に位置することができる。

【0055】

「異種」または「組み換え」ヌクレオチド配列は、それが導入される宿主細胞と天然に関連しないヌクレオチド配列であり、非天然起源の、多コピーの天然起源ヌクレオチド配列を含む。

【0056】

相同性を有する異なる核酸またはタンパク質は、本明細書において「相同体」と呼ぶ。相同体という用語は、同一および他の種由来の相同配列および同一および他の種由来のオルソログス配列を含む。「相同性」は、位置同一性（すなわち配列類似性または同一性）のパーセントの観点からの、2以上の核酸および／またはアミノ酸配列の間の類似性のレベルを指す。また、相同性は、異なる核酸またはタンパク質間の類似の機能特性の概念を指す。したがって、本発明の組成物および方法は、本発明のヌクレオチド配列およびポリペプチド配列に対する相同体をさらに含む。本明細書において用いられる「オルソログス」は、種分化の間に共通の祖先遺伝子から生じる異なる種における相同ヌクレオチド配列および／またはアミノ酸配列を指す。本発明のヌクレオチド配列の相同体は、本発明の前記ヌクレオチド配列と実質的な配列同一性（例えば、少なくとも約70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、および／または100%）を有する。したがって、例えば、I型、II型、III型、IV型、またはV型ポリヌクレオチドまたはポリペプチドの相同体は、任意の公知または後に同定されるI型、II型、III型、IV型、またはV型ポリヌクレオチドまたはポリペプチドのいずれか1つと、約70%相同またはそれよりも高くてもよい。

【0057】

本明細書において用いられる、ハイブリダイゼーション（hybridization）、ハイブリダイズする（hybridize）、ハイブリダイズする（hybridizing）、およびそれらの文法的変形は、2つの完全に相補のヌクレオチド配列またはいくつかのミスマッチの塩基対が存在してよい実質的に相補の配列の結合を指す。ハイブリダイゼーションの条件は当該分野でよく知られていて、ヌクレオチド配列の長さおよびヌクレオチド配列間の相補性の程度に基づき変化する。一部の実施態様では、ハイブリダイゼーションの条件は、高ストリンジェンシーであってよく、または、それらは、ハイブリダイズされる配列の長さおよび相補性の量に応じて、中ストリンジェンシーまたは低ストリンジェンシーであってよい。ヌクレオチド配列間のハイブリダイゼーションの目的に関して低、中および高ストリンジェンシーを構成する条件は、当該分野でよく知られている（例えば、Gasiunas et al. (2012) Proc. Natl. Acad. Sci. 109:E2579-E2586; M. R. Green and J. Sambrook (2012) Molecular Cloning: A Laboratory Manual, 4th Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NYを参照）。

【0058】

本明細書において用いられる用語「増大する（increase）」、「増大する（increasing）」、「増大した（increased）」、「増強する（enhance）」、「増強した（enhanced）」、「増強する（enhancing）」および「増強（enhancement）」（およびそれらの文法的変形）は、コントロールと比べて、少なくとも約25%、50%、75%、100%、150%、200%、300%、400%、500%またはそれよりも高い上昇を表す。

【0059】

「ネイティブ」または「野生型」の核酸、ヌクレオチド配列、ポリペプチドまたはアミノ酸配列は、天然起源または内因性の核酸、ヌクレオチド配列、ポリペプチドまたはアミ

ノ酸配列を指す。したがって、例えば、「野生型 mRNA」は、生物において天然起源または内因性である mRNA である。「相同」核酸配列は、それが導入される宿主細胞と天然に関連するヌクレオチド配列である。

【0060】

また、本明細書において用いられる用語「核酸」、「核酸分子」、「核酸構築物」、「ヌクレオチド配列」および「ポリヌクレオチド」は、直鎖または分岐の一本鎖または二本鎖である RNA または DNA、またはそれらのハイブリッドを指す。また、用語は、RNA/DNA ハイブリッドも包含する。dsRNA が合成的に生産される場合は、一般的でない塩基、例えば、イノシン、5-メチルシトシン、6-メチルアデニン、ヒポキサンチンおよびその他もまた、アンチセンス、dsRNA、およびリボザイム対合に用いることができる。例えば、シチジンおよびウリジンの C-5 プロピナアナログを含むポリヌクレオチドは、高アフィニティで RNA と結合し、遺伝子発現の強力なアンチセンス阻害剤となることが示されている。他の修飾、例えば、ホスホジエステル主鎖、または RNA のリボース糖基内の 2'-ヒドロキシに対する修飾もすることができる。本開示の核酸構築物は、DNA または RNA であってよいが、好ましくは DNA である。したがって、本発明の核酸構築物は、DNA の形態で記載され用いられ得るが、使用目的に応じて、RNA の形態で記載され用いられてもよい。

【0061】

明細書において用いられる用語「ヌクレオチド配列」は、ヌクレオチドのヘテロポリマーまたは核酸分子の 5' から 3' 末端のこれらのヌクレオチドの配列を指し、DNA または RNA 分子を含み、cDNA、DNA フラグメントまたは部分、ゲノム DNA、合成（例えば化学合成）DNA、プラスミド DNA、mRNA、およびアンチセンス RNA を含み、そのいずれかは一本鎖または二本鎖であってよい。用語「ヌクレオチド配列」「核酸」、「核酸分子」、「オリゴヌクレオチド」および「ポリヌクレオチド」は、本明細書において交換可能に用いられて、ヌクレオチドのヘテロポリマーも指す。別段の指示がされる場合を除き、本明細書において提供される核酸分子および/またはヌクレオチド配列は、本明細書において、5' から 3' の方向で左から右に示され、米国の配列規則（連邦規則法典第 37 巻セクション 1.821~1.825）および世界知的所有権機関（WIPO）標準 ST.25 に示されるヌクレオチド記号を表すための標準的なコードを用いて表される。

【0062】

本明細書において用いられる用語「パーセント配列同一性」または「パーセント同一性」は、2つの配列が最適に並べられた場合の、試験（「対象」）ポリヌクレオチド分子（またはその相補鎖）と比較した、参照（「クエリー」）ポリヌクレオチド分子（またはその相補鎖）の直鎖ポリヌクレオチド配列内の同一のヌクレオチドのパーセンテージを指す。一部の実施態様では、「パーセント同一性」は、アミノ酸配列内の同一のアミノ酸のパーセンテージを指し得る。

【0063】

「プロトスペーサー配列」は、標的二本鎖 DNA を指し、具体的には、CRISPR リピートスペーサー配列、CRISPR リピートスペーサーリピート配列、および/または CRISPR アレイの、スペーサー配列と完全または実質的に相補（およびハイブリダイズする）標的 DNA（例えば、または、ゲノム内の標的領域）の部分を目指す。

【0064】

本明細書において用いられる用語「縮小する（reduce）」、「縮小された（reduced）」、「縮小する（reducing）」、「縮小（reduction）」、「減少する（diminish）」、「抑制する（suppress）」、および、「低減する（decrease）」（およびそれらの文法的変形）は、例えば、コントロールと比較して、少なくとも約 5%、10%、15%、20%、25%、35%、50%、75%、80%、85%、90%、95%、97%、98%、99%、または 100% の低減を説明する。特定の実施態様では、縮小は、測定される構成要素（例えば、ゲノムサ

イズまたは細胞の集団)の検出可能な活性または量をもたらさなくてよく、または本質的にもたらさなくてよい(すなわち、ほんのわずかな量、例えば、約10%未満または実に5%)。したがって、例えば、縮小されたゲノムサイズは、コントロールと比較して、少なくとも約5%、10%、15%、20%、25%、35%、50%、75%、80%、85%、90%、95%、97%、98%、99%、または100%のゲノムのサイズの縮小を意味することができる。

【0065】

本明細書において用いられるコントロールは、例えば、本発明の異種核酸構築物によって形質転換されていない、細菌、古細菌、藻類または酵母細胞の集団であってよい。一部の実施態様では、コントロールは、細菌、古細菌、藻類または酵母細胞の野生型集団であってよく、または、リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイを含む異種構築物によって形質転換された細菌、古細菌または酵母細胞の集団であってよく、ここで、スペーサーは、前記集団の細菌、古細菌または酵母細胞のゲノム内の標的領域に相補でないヌクレオチド配列を含む(すなわち、非自己標的化/「スクランブルスペーサー」)。さらなる態様では、コントロールは、例えば、細菌、古細菌、藻類または酵母細胞の野生型集団、または、リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイを含む異種構築物によって形質転換された細菌、古細菌、藻類または酵母細胞の集団であってよく、ここで、スペーサーは、プロトスペーサー隣接モチーフ(PAM)に隣接して配置されない前記集団の細菌、古細菌、藻類または酵母細胞のゲノム内の標的領域に実質的に相補であるヌクレオチド配列を含む。

【0066】

本明細書において用いられる「リピート配列」は、例えば、野生型CRISPR遺伝子座の任意のリピート配列、または、「スペーサー配列」によって分離された合成のCRISPRアレイのリピート配列(例えば、本発明のリピートスペーサー配列またはリピートスペーサーリピート配列)を指す。本発明で有用なリピート配列は、CRISPR遺伝子座の任意の公知または後に同定されるリピート配列であってよい。したがって、一部の実施態様では、リピートスペーサー配列またはリピートスペーサーリピートは、野生型II型CRISPRアレイ由来のリピートと実質的に同一の、(例えば、少なくとも約70%同一(例えば、少なくとも約70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、またはそれよりも高い))リピートを含む。一部の実施態様では、リピート配列は、野生型I型CRISPRアレイ、野生型II型CRISPRアレイ、野生型III型CRISPRアレイ、野生型IV型CRISPRアレイ、または野生型V型CRISPRアレイ由来のリピートと100%同一である。さらなる実施態様では、本発明で有用なリピート配列は、I型crRNA、II型crRNA、III型crRNA、IV型crRNA、またはV型crRNAのいずれかのCRISPR遺伝子座または合成のCRISPRアレイのリピート配列の連続したヌクレオチドのフラグメントまたは一部である部分的リピートを含むヌクレオチド配列を含むことができる。

【0067】

本明細書において用いられる、I型、II型、III型、IV型、またはV型CRISPR-Casシステムの「CRISPRアレイ」は、5'から3'に、リピートスペーサーリピート配列を含む、または、5'から3'に、少なくとも1つのリピートスペーサー配列(例えば、約1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、または25のリピートスペーサー配列、およびその中の任意の範囲または値)を含む、核酸構築物を指す。1つよりも多いリピートスペーサーがCRISPRアレイに含まれる場合は、前の(5'から3')リピートスペーサー配列のスペーサーは、後のリピートスペーサーのリピートに連結することができる(例えば、第一のリピートスペーサー配列のスペ

ーサーは、第二のリピートスパーサー配列のリピートに連結される)。一部の実施態様では、CRISPRアレイは、スパーサーによって分離された2つのリピート(または2つの部分的リピート)(例えば、リピートスパーサーリピート配列)を含むことができる。

【0068】

本明細書において用いられる「配列同一性」は、2つの最適に並べられたポリヌクレオチドまたはペプチド配列が、構成要素(例えば、ヌクレオチドまたはアミノ酸)のアライメントのウィンドウ全体で不変である程度を指す。「同一性」は、制限されないが: Computational Molecular Biology (Lesk, A. M., ed.) Oxford University Press, New York (1988); Biocomputing: Informatics and Genome Projects (Smith, D. W., ed.) Academic Press, New York (1993); Computer Analysis of Sequence Data, Part I (Griffin, A. M., and Griffin, H. G., eds.) Humana Press, New Jersey (1994); Sequence Analysis in Molecular Biology (von Heinje, G., ed.) Academic Press (1987); and Sequence Analysis Primer (Gribskov, M. and Devereux, J., eds.) Stockton Press, New York (1991)に記載のものを含む公知の方法により、容易に計算することができる。

【0069】

本明細書において用いられる「スパーサー配列」は、標的DNA(すなわち、プロトスパーサー隣接モチーフ(PAM)配列に隣接するゲノム内の標的領域または「プロトスパーサー配列」)に相補のヌクレオチド配列である。スパーサー配列は、標的DNAと完全に相補または実質的に相補であることができる(例えば、少なくとも約70%相補(例えば、約70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、またはそれよりも高い))。代表的な実施態様では、スパーサー配列は、標的DNAと100%相補性を有する。さらなる実施態様では、標的DNAに対するスパーサー配列の3'領域の相補性は、100%であるが、スパーサーの5'領域では100%未満であり、したがって、標的DNAに対するスパーサー配列の全体的な相補性は100%未満である。したがって、例えば、20個のヌクレオチドスパーサー配列の3'領域における最初の7、8、9、10、11、12、13、14、15、16個などのヌクレオチド(シード配列)は、標的DNAと100%相補であることができるが、一方で、スパーサー配列の5'領域における残りのヌクレオチドは、標的DNAと実質的に相補(例えば、少なくとも約70%相補)である。一部の実施態様では、スパーサー配列の最初の7~12個のヌクレオチドは、標的DNAと100%相補であることができるが、一方で、スパーサー配列の5'領域における残りのヌクレオチドは、標的DNAと実質的に相補(例えば、少なくとも約70%相補)である。他の実施態様では、スパーサー配列の最初の7~10個のヌクレオチドは、標的DNAと100%相補であることができるが、一方で、スパーサー配列の5'領域における残りのヌクレオチドは、標的DNAと実質的に相補(例えば、少なくとも約70%相補)である。代表的な実施態様では、スパーサー配列の最初の7個のヌクレオチド(シード内)は、標的DNAと100%相補であることができるが、一方で、スパーサー配列の5'領域における残りのヌクレオチドは、標的DNAと実質的に相補(例えば、少なくとも約70%相補)である。

【0070】

本明細書において用いられる、「標的DNA」、「標的領域」、または、「ゲノム内の標的領域」は、リピートスパーサー配列またはリピートスパーサーリピート配列内

のスペーサー配列と、完全に相補または実質的に相補の（例えば、少なくとも70%相補の（例えば、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、またはそれよりも高い））、生物のゲノムの領域を指す。一部の実施態様では、標的領域は、生物のゲノムにおいて、PAM配列（標的領域のすぐ3'に配置されたPAM配列）にすぐ隣接して配置された、約10～約40個の連続した長さのヌクレオチドであってよい（例えば、I型CRISPR-CasシステムおよびII型CRISPR-Casシステム）。一部の実施態様では、例えばI型システムでは、PAMは、プロトスペーサーの片側にある（5'末端）。III型システムについて公知のPAMは存在しない。Makarova et al. は、CRISPRシステムの全てのクラス、タイプおよびサブタイプに関する命名法を記載する（Nature Reviews Microbiology 13:722-736 (2015)）。ガイド構造およびPAMは、R. Barrangou (Genome Biol. 16:247 (2015)) に記載される。

10

【0071】

一部の実施態様では、本発明で有用な標的領域は、必須遺伝子または非必須遺伝子内に配置される。

【0072】

代表的な実施態様では、標的領域は、無作為に選択することができ、または、特異的に選択することができる。一部の実施態様では、無作為に選択される標的領域は、細菌、古細菌、藻類または酵母ゲノム内において、PAM配列にすぐ隣接して配置された、任意の少なくとも10個の連続したヌクレオチド（例えば、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50個など、およびその中の任意の範囲または値）から選択してよい。一部の実施態様では、標的領域は、細菌、古細菌、藻類または酵母ゲノム内において、プロトスペーサー隣接モチーフ（PAM）配列にすぐ隣接して配置された、約10～約20個の連続したヌクレオチド、約10～約30個の連続したヌクレオチド、および／または、約10～約40個の連続したヌクレオチドなど、または、その中の任意の範囲または値であることができる。一部の実施態様では、標的領域を特異的に選択するステップは、前記の1つまたは複数の細菌、古細菌、藻類または酵母細胞のゲノムにおいて、互いから、約100個の全てのヌクレオチド～約1000個の全てのヌクレオチド、約100個の全てのヌクレオチド～約2000個全て、約100個の全てのヌクレオチド～約3000個全て、約100個の全てのヌクレオチド～約4000個全て、および／または、約100個の全てのヌクレオチド～約5000個の全てのヌクレオチドなどに配置された2以上の標的領域を選択するステップを含むことができる。特定の実施態様では、標的領域を特異的に選択するステップは、細菌、古細菌、藻類または酵母ゲノム内において、PAM配列にすぐ隣接する少なくとも約10～約40個の連続したヌクレオチドを含む、遺伝子、オープンリーディングフレーム、推定のオープンリーディングフレームまたは遺伝子間領域から、標的領域を特異的に選択するステップを含む。

20

30

40

【0073】

本明細書において用いられる「トランス活性化CRISPR (tracr) 核酸」または「tracr核酸」は、任意のtracr RNA（またはそのコードDNA）を指す。tracr核酸は、5'から3'に、下側ステム、上側ステム、バルジ、ネクサスヘアピンおよび末端ヘアピンを含む（Briner et al. (2014) Molecular Cell. 56(2):333-339を参照）。トランス活性化CRISPR (tracr) 核酸は、成熟または未熟crRNAのリピート部分へのハイブリダイズにおいて機能し、Cas9タンパク質を標的部位に動員し、そして、構造的再配置を誘導することによってCas9の触媒活性を促進し得る。tracrRNA分子の機能性成分は

50

上記に挙げられる。t r a c r R N Aに関する配列は、C R I S P R－C a sシステムに特有であり、可変であってよい。公知または後に同定される任意のt r a c r核酸を、本発明で用いることができる。

【0074】

本明細書において用いられる、2つの核酸分子、ヌクレオチド配列またはタンパク質配列の関連における語句「実質的に同一」または「実質的な同一性」は、以下の配列比較アルゴリズムの1つを用いて測定して、または目視検査により、最大一致に関して比較およびアライメントした場合、少なくとも約70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、および／または100%のヌクレオチドまたはアミノ酸残基の同一性を有する2以上の配列またはサブシーケンスを指す。本発明の一部の実施態様では、実質的な同一性は、少なくとも約50残基～約150残基の長さである配列の領域にわたって存在する。したがって、本発明の一部の実施態様では、実質的な同一性は、少なくとも約3～約15（例えば、3、4、5、6、7、8、9、10、11、12、13、14、15残基の長さなど、またはその中の任意の値または任意の範囲）、少なくとも約5～約30、少なくとも約10～約30、少なくとも約16～約30、少なくとも約18～少なくとも約25、少なくとも約18、少なくとも約22、少なくとも約25、少なくとも約30、少なくとも約40、少なくとも約50、約60、約70、約80、約90、約100、約110、約120、約130、約140、約150、またはそれよりも多くの残基の長さ、および、その中の任意の範囲である配列の領域にわたって存在する。代表的な実施態様では、配列は、少なくとも約22個のヌクレオチドにわたって実質的に同一であることができる。一部の特定の実施態様では、配列は、少なくとも約150個の残基にわたって実質的に同一である。一部の実施態様では、本発明の配列は、少なくとも約16個のヌクレオチド～約25個のヌクレオチドにわたって約70%～約100%同一であることができる。一部の実施態様では、本発明の配列は、少なくとも約16個のヌクレオチド～約25個のヌクレオチドにわたって約75%～約100%同一であることができる。さらなる実施態様では、本発明の配列は、少なくとも約16個のヌクレオチド～約25個のヌクレオチドにわたって約80%～約100%同一であることができる。さらなる実施態様では、本発明の配列は、少なくとも約7個のヌクレオチド～約25個のヌクレオチドにわたって約80%～約100%同一であることができる。一部の実施態様では、本発明の配列は、少なくとも約18個のヌクレオチドにわたって約70%同一であることができる。他の実施態様では、配列は、約22個のヌクレオチドにわたって約85%同一であることができる。さらに他の実施態様では、配列は、約16個のヌクレオチドにわたって100%同一であることができる。さらなる実施態様では、配列は、コード領域の全長にわたって実質的に同一である。さらに、例示的な実施態様では、実質的に同一のヌクレオチドまたはポリペプチド配列は、実質的に同一の機能（例えば、c r R N A、t r a c r核酸、リピート配列、C a s 9ヌクレアーゼ（ニックアーゼ、DNA、RNAおよび／またはPAM認識および結合）、C a s 3、C a s 3'、C a s 3''または任意の他のC R I S P R－C a sポリヌクレオチドまたはポリペプチドの機能または活性）を発揮する。

【0075】

配列比較のために、典型的に、1つの配列は、試験配列が比較される参照配列としての役割がある。配列比較アルゴリズムを用いる場合、試験および参照配列がコンピュータに入力され、必要な場合はサブシーケンス座標が指定されて、配列アルゴリズムプログラムパラメーターが指定される。それから、配列比較アルゴリズムは、指定されたプログラムパラメーターに基づき、参照配列に対する試験配列（単数または複数）のパーセント配列同一性を計算する。

【0076】

比較ウィンドウを並べるための最適な配列アライメントは当業者に広く知られ、S m i

10

20

30

40

50

th and Watermanのローカル相同性アルゴリズム、Needleman and Wunschの相同性アライメントアルゴリズム、Pearson and Lipmanの類似性検索方法などのツールにより、および場合により、GCG（登録商標）Wisconsin Package（登録商標）（Accelrys Inc., San Diego, CA）の一部として利用可能なGAP、BESTFIT、FASTA、および、TFASTAなどのアルゴリズムのコンピューター実施により、行なつてよい。試験配列および参照配列のアライメントされたセグメントの「同一性分数」は、2つの並べられた配列が共有する同一の構成要素の数を、参照配列セグメント（すなわち、参照配列全体または参照配列のより小さな規定部分）内の構成要素の全数で割った数である。パーセント配列同一性は、同一性分数に100をかけて表される。1つまたは複数のポリヌクレオチド配列の比較は、全長ポリヌクレオチド配列またはその一部に対して、または、より長いポリヌクレオチド配列に対して、され得る。本発明の目的に関し、「パーセント同一性」は、翻訳ヌクレオチド配列のためのBLASTX version 2.0およびポリヌクレオチド配列のためのBLASTN version 2.0を用いて決定されてもよい。

10

【0077】

BLAST解析を行なうためのソフトウェアは、国立生物工学情報センターを通じて一般に利用可能である。このアルゴリズムは、データベース配列内の同じ長さのワードと並べた場合に、一部の正の値の閾値スコアTと一致する、または満たす、クエリー配列内の長さWのショートワードを特定することにより、高スコア配列ペア（HSP）を最初に特定するステップを含む。Tは、隣接ワードスコア閾値と呼ばれる（Altschul et al., 1990）。これらの最初の隣接ワードヒットは、それらを含むより長いHSPを見つけるためのサーチを始める種として作用する。それから、ワードヒットは、累積アライメントスコアが増加し得る限り、各配列に沿って両方向に伸びる。累積スコアは、ヌクレオチド配列について、パラメーターM（一致する残基のペアに関する恩恵スコア；常に >0 ）およびN（ミスマッチ残基に関するペナルティスコア；常に <0 ）を用いて計算される。アミノ酸配列については、スコアマトリックスを用いて累積スコアを計算する。各方向でのワードヒットの伸長は、累積アライメントスコアがその最大達成値から量X低下した場合は停止し、1つまたは複数の負のスコアの残基アライメントの蓄積、またはいずれかの配列が終わりに到達したことに起因して、累積スコアは0以下になる。BLASTアルゴリズムパラメーターW、T、およびXは、アライメントの感度および速度を決定する。BLASTNプログラム（ヌクレオチド配列）は、デフォルトとして、ワード長（W）11、期待値（E）10、カットオフ100、 $M=5$ 、 $N=-4$ 、および両方の鎖の比較を使用する。アミノ酸配列に関しては、BLASTPプログラムは、デフォルトとして、ワード長（W）3、期待値（E）10、およびBLOSUM62スコアマトリックスを使用する（Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89:10915 (1989)を参照）。

20

30

【0078】

パーセント配列同一性の計算に加え、BLASTアルゴリズムは、2つの配列間の類似性の統計分析も行う（例えば、Karlin & Altschul, Proc. Natl. Acad. Sci. USA 90:5873-5877 (1993)を参照）。BLASTアルゴリズムにより提供される類似性の1つの評価基準は最小和確率（ $P(N)$ ）であり、2つのヌクレオチドまたはアミノ酸配列の間で一致が偶然に生じる確率の示度を提供する。例えば、参照ヌクレオチド配列に対する試験ヌクレオチド配列の比較における最小和確率が約0.1未満～約0.001未満である場合に、試験核酸配列は参照配列と類似すると考えられる。したがって、本発明の一部の実施態様では、参照ヌクレオチド配列に対する試験ヌクレオチド配列の比較における最小和確率は、約0.001未満である。

40

【0079】

また、2つのヌクレオチド配列は、2つの配列がストリンジェントな条件下で互いにハイブリダイズする場合に、実質的に相補であると考えることができる。一部の代表的な実

50

施態様では、実質的に相補であると考えられる2つのヌクレオチド配列は、高ストリンジェントな条件下で、互いにハイブリダイズする。

【0080】

サザンおよびノーザンハイブリダイゼーションなどの核酸ハイブリダイゼーション実験の関連における、「ストリンジェントなハイブリダイゼーション条件」および「ストリンジェントなハイブリダイゼーション洗浄条件」は配列依存性であり、異なる環境パラメーターの下で異なる。核酸のハイブリダイゼーションに対する広範なガイドは、Tijssen Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes part I chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays" Elsevier, New York (1993)に見られる。一般に、高ストリンジェントなハイブリダイゼーションおよび洗浄条件は、規定のイオン強度およびpHでの特異的配列に関する融点 (T_m) よりも約5℃低くなるように選択される。

【0081】

T_m は、標的配列の50%が完全に一致するプローブにハイブリダイズする、(規定のイオン強度およびpHでの) 温度である。非常にストリンジェントな条件は、特異的なプローブに関する T_m と等しくなるように選択される。サザンまたはノーザンブロットでのフィルターにおける、100よりも多い相補残基を有する相補ヌクレオチド配列のハイブリダイゼーションに関するストリンジェントなハイブリダイゼーション条件の例は、1mgのヘパリンを含む50%ホルムアミド(42℃)であり、ハイブリダイゼーションは一晩行われる。高ストリンジェントな洗浄条件の例は、0.15M NaCl、72℃、約15分である。ストリンジェントな洗浄条件の例は、0.2×SSC洗浄、65℃、15分である(SSCバッファーの説明については、Sambrook(以下)を参照)。バックグラウンドのプローブシグナルを除去するために、高ストリンジェンシーな洗浄の前に低ストリンジェンシーな洗浄を行うことが多い。例えば、100個を超えるヌクレオチドのデュプレックスに関する、中ストリンジェンシーな洗浄の例は、1×SSC、45℃、15分である。例えば、100個を超えるヌクレオチドのデュプレックスに関する、低ストリンジェンシーな洗浄の例は、4~6×SSC、40℃、15分である。短いプローブ(例えば、約10~50個のヌクレオチド)に関しては、ストリンジェントな条件は典型的に、約1.0M未満のNaイオンの塩濃度、典型的に、約0.01~1.0MのNaイオン濃度(または他の塩)(pH7.0~8.3)を含み、温度は典型的に、少なくとも約30℃である。また、ストリンジェントな条件は、ホルムアミドなどの不安定化剤の添加によっても達成することができる。一般に、特定のハイブリダイゼーションアッセイにおいて、関係のないプローブについて見られるものより2倍の(またはそれより高い)シグナル・ノイズ比は、特異的なハイブリダイゼーションの検出を示す。ストリンジェントな条件下で互いにハイブリダイズしないヌクレオチド配列は、それらがコードするタンパク質が実質的に同一であるならば、依然として実質的に同一である。これは、例えば、遺伝子コードにより許容される最大コドン縮退を用いてヌクレオチド配列のコピーが作られる場合に生じ得る。

【0082】

以下は、本発明の参照ヌクレオチド配列と実質的に同一である相同ヌクレオチド配列をクローニングするために用いられ得る、ハイブリダイゼーション/洗浄条件のセットの例である。一実施態様では、参照ヌクレオチド配列は、7%ドデシル硫酸ナトリウム(SDS)、0.5M NaPO₄、1mM EDTA(50℃)中、2×SSC、0.1% SDSの洗浄(50℃)で、「試験」ヌクレオチド配列とハイブリダイズする。別の実施態様では、参照ヌクレオチド配列は、7%ドデシル硫酸ナトリウム(SDS)、0.5M NaPO₄、1mM EDTA(50℃)中、1×SSC、0.1% SDSの洗浄(50℃)で、または、7%ドデシル硫酸ナトリウム(SDS)、0.5M NaPO₄、1m

M E D T A (5 0 ℃) 中、0. 5 × S S C、0. 1 % S D S の洗浄 (5 0 ℃) で、「試験」ヌクレオチド配列とハイブリダイズする。また、さらなる実施態様では、参照ヌクレオチド配列は、7 % ドデシル硫酸ナトリウム (S D S)、0. 5 M N a P O ₄、1 m M E D T A (5 0 ℃) 中、0. 1 × S S C、0. 1 % S D S の洗浄 (5 0 ℃) で、または 7 % ドデシル硫酸ナトリウム (S D S)、0. 5 M N a P O ₄、1 m M E D T A (5 0 ℃) 中、0. 1 × S S C、0. 1 % S D S の洗浄 (6 5 ℃) で、「試験」ヌクレオチド配列とハイブリダイズする。

【0083】

本発明の任意のヌクレオチド配列および／または異種核酸構築物は、目的の任意の種の発現のためにコドン最適化され得る。コドン最適化は当技術分野で広く知られていて、種特異的なコドン使用頻度を用いたコドン使用バイアスに関するヌクレオチド配列の変更を含む。コドン使用頻度は、目的の種に関して最も高く発現される遺伝子の配列解析に基づいて作られる。ヌクレオチド配列が核内で発現される場合、コドン使用頻度は、目的の種に関して高発現の核遺伝子の配列解析に基づいて作られる。ヌクレオチド配列の変更は、種特有のコドン使用頻度を、天然のポリヌクレオチド配列に存在するコドンと比較することにより決定される。当該分野において理解されるように、ヌクレオチド配列のコドン最適化は、ネイティブのヌクレオチド配列と 1 0 0 % 未満の同一性 (例えば、7 0 %、7 1 %、7 2 %、7 3 %、7 4 %、7 5 %、7 6 %、7 7 %、7 8 %、7 9 %、8 0 %、8 1 %、8 2 %、8 3 %、8 4 %、8 5 %、8 6 %、8 7 %、8 8 %、8 9 %、9 0 %、9 1 %、9 2 %、9 3 %、9 4 %、9 5 %、9 6 %、9 7 %、9 8 %、9 9 % など) を有するが、依然としてオリジナルのネイティブのヌクレオチド配列によりコードされるものと同一の機能を有するポリペプチドをコードする、ヌクレオチド配列を生じさせる。したがって、本発明の代表的な実施態様では、本発明のヌクレオチド配列および／または異種核酸構築物は、特定の目的の種での発現のためにコドン最適化され得る。

【0084】

一部の実施態様では、本発明の異種または組み換え核酸分子、ヌクレオチド配列および／またはポリペプチドは、「単離される」。「単離された」核酸分子、「単離された」ヌクレオチド配列または「単離された」ポリペプチドは、人の手によりその天然環境から離れて存在し、したがって天然の産物ではない、核酸分子、ヌクレオチド配列またはポリペプチドである。単離された核酸分子、ヌクレオチド配列またはポリペプチドは、天然起源の生物またはウイルスの他の構成要素の少なくとも一部、例えば、細胞またはウイルスの構造上の構成要素、またはポリヌクレオチドと関連して一般に見られる他のポリペプチドまたは核酸から、少なくとも部分的に切り離された、精製された形態で存在し得る。代表的な実施態様では、単離された核酸分子、単離されたヌクレオチド配列および／または単離されたポリペプチドは、少なくとも約 1 %、5 %、1 0 %、2 0 %、3 0 %、4 0 %、5 0 %、6 0 %、7 0 %、8 0 %、9 0 %、9 5 %、またはそれよりも高い純度である。

【0085】

他の実施態様では、単離された核酸分子、ヌクレオチド配列またはポリペプチドは、非天然の環境 (例えば組み換え宿主細胞) 内に存在し得る。したがって、例えばヌクレオチド配列に関しては、用語「単離された」は、それが天然に生じる染色体および／または細胞から切り離されることを意味する。また、ポリヌクレオチドは、それが天然に生じる染色体および／または細胞から切り離され、それから、それが天然に生じない遺伝的状況、染色体および／または細胞 (例えば、異なる宿主細胞、異なる制御配列、および／または天然に見られるのとは異なるゲノム内の位置) に挿入される場合もまた、単離される。したがって、異種核酸構築物、ヌクレオチド配列およびそれらのコードされるポリペプチドは、人の手により、それらの天然環境から離れて存在し、したがって、天然の産物ではない点で「単離される」が、一部の実施態様では、組み換え宿主細胞内に導入され存在することができる。

【0086】

一部の実施態様では、本発明の異種または組み換え核酸構築物は、「合成」である。「

合成」の核酸分子、「合成」のヌクレオチド配列または「合成」のポリペプチドは、天然に見られないが人の手によって作られて、したがって、天然の産物ではない、核酸分子、ヌクレオチド配列またはポリペプチドである。

【0087】

本明細書に記載される任意の実施態様では、本発明のヌクレオチド配列および／または異種核酸構築物は、様々な生物細胞における発現に関する様々なプロモーターおよび他の調節エレメントと作動可能に関連することができる。したがって、代表的な実施態様では、本発明の核酸構築物は、1つまたは複数のヌクレオチド配列に動作可能に連結された1つまたは複数のプロモーターをさらに含むことができる。

【0088】

本明細書において用いられる「動作可能に連結」または「動作可能に関連」は、示された因子が互いに機能的に関連し、また一般に、物理的にも関連することを意味する。したがって、本明細書において用いられる用語「動作可能に連結」または「動作可能に関連」は、機能的に関連がある単一核酸分子上のヌクレオチド配列を指す。したがって、第二のヌクレオチド配列に動作可能に連結された第一のヌクレオチド配列は、第一のヌクレオチド配列が、第二のヌクレオチド配列と機能的関係で配置される状況の意味する。例えば、プロモーターがヌクレオチド配列の転写または発現を果たす場合に、プロモーターは前記ヌクレオチド配列と動作可能に関連する。当業者は、コントロール配列がその発現を方向付けるように機能する限り、コントロール配列（例えばプロモーター）は、それが動作可能に関連するヌクレオチド配列と隣接している必要はないことを理解するだろう。したがって、例えば、介在する非翻訳でなお転写される配列が、プロモーターとヌクレオチド配列との間に存在してよく、プロモーターは依然としてヌクレオチド配列に「動作可能に連結」していると考えることができる。

【0089】

「プロモーター」は、プロモーターと動作可能に関連するヌクレオチド配列（すなわち、コード配列）の転写を制御または調節するヌクレオチド配列である。コード配列は、ポリペプチドおよび／または機能性RNAをコードし得る。典型的に、「プロモーター」は、RNAポリメラーゼIIに関する結合部位を含むヌクレオチド配列を指し、転写の開始を管理する。一般に、プロモーターは、5'または対応するコード配列のコード領域のスタートに対して上流に見られる。プロモーター領域は、遺伝子発現のレギュレーターとして作用する他のエレメントを含んでよい。これらは、TATAボックスコンセンサス配列を含み、そして、CAATボックスコンセンサス配列を含むことが多い（Breathnach and Chambon, (1981) Annu. Rev. Biochem. 50:349）。植物では、CAATボックスはAGGAボックスで置き換えられ得る（Messing et al., (1983) in Genetic Engineering of Plants, T. Kosuge, C. Meredith and A. Hollaender (eds.), Plenum Press, pp. 211-227）。

【0090】

プロモーターは、異種核酸構築物（すなわち、「キメラ遺伝子」または「キメラポリヌクレオチド」）の調製で用いるための、例えば、構成的、誘導性、時間的に調節された、発生学的に調節された、化学的に調節された、組織一好適および／または組織一特異的なプロモーターを含んでよい。これらの様々なタイプのプロモーターは、当該分野で知られている。

【0091】

プロモーターの選択は、発現のための時間的および空間的要求に応じて変化し、および、形質転換される宿主細胞にも応じて変化する。多くの異なる生物のためのプロモーターは、当該分野でよく知られている。当該分野に存在する広範囲の知識に基づき、適切なプロモーターを、目的の特定の宿主生物のために選択することができる。したがって、例えば、モデル生物において、高度に構成的に発現された遺伝子上流のプロモーターについ

10

20

30

40

50

て多くが知られ、そのような知識は、必要に応じて、他のシステムで容易に利用および実施することができる。

【0092】

一部の実施態様では、本発明の核酸構築物は、「発現カセット」であってよく、または、発現カセット内に構成することができる。本明細書において用いられる「発現カセット」は、目的のヌクレオチド配列（例えば、本発明の核酸構築物（例えば、合成の *tracr* 核酸構築物、合成の *CRISPR* 核酸構築物、合成の *CRISPR* アレイ、キメラ核酸構築物；目的のポリペプチド、*I* 型ポリペプチド、*II* 型ポリペプチド、*III* 型ポリペプチド、*IV* 型ポリペプチド、および／または *V* 型ポリペプチドをコードするヌクレオチド配列）を含む異種核酸構築物を意味し、ここで、前記ヌクレオチド配列は、少なくとも

10

【0093】

目的のヌクレオチド配列を含む発現カセットは、キメラであってよく、その構成要素の少なくとも1つは、その他の構成要素の少なくとも1つに対して異種であることを意味する。また、発現カセットは、天然起源であるが異種発現に有用な組み換え体で得られたものであってもよい。

【0094】

また、発現カセットは、場合により、選択された宿主細胞内で機能性の転写および／または翻訳終結領域（すなわち終結領域）を含んでもよい。様々な転写終結因子が発現カセットでの使用に利用可能であり、目的の異種ヌクレオチド配列を超えた (*beyond*) 転写の終結および正しい (*correct*) *mRNA* のポリアデニル化に参与する。終結領域は、転写開始領域に対してネイティブであり得て、動作可能に連結された目的のヌクレオチド配列に対してネイティブであり得て、宿主細胞に対してネイティブであり得て、または、別の起源に由来し得る（すなわち、プロモーターに対して、目的のヌクレオチド配列に対して、宿主に対して、外因または異種、またはそれらの任意の組み合わせ）。

20

【0095】

また、発現カセットは、形質転換された宿主細胞を選択するために用いることのできる選択可能なマーカーに関するヌクレオチド配列を含むこともできる。本明細書において用いられる「選択可能なマーカー」は、発現した場合に、マーカーを発現している宿主細胞に、明白な表現型を与え、したがって、そのマーカーを有さないものからそのような形質転換された細胞が区別されるのを可能にする、ヌクレオチド配列を意味する。そのようなヌクレオチド配列は、マーカーが、化学的手段により、例えば選択剤（例えば抗生物質など）を用いることにより、選択することのできる形質を与えるかどうか、または、マーカーが単に、例えばスクリーニング（例えば蛍光）による観察または試験を介して同定され得る形質であるかどうかに応じて、選択可能なマーカーまたはスクリーニング可能なマーカーのいずれかをコードし得る。もちろん、適切な選択可能なマーカーの多くの例が当該分野で知られていて、本明細書に記載の発現カセットにおいて用いることができる。

30

【0096】

発現カセットに加えて、本明細書に記載の核酸分子およびヌクレオチド配列は、ベクターとの関連で用いることができる。用語「ベクター」は、核酸（単数）（または核酸（複数））を細胞に移行、送達または導入するための構成物を指す。ベクターは、移行、送達または導入されるヌクレオチド配列（単数または複数）を含む核酸分子を含む。宿主生物の形質転換での使用のためのベクターは、当該分野でよく知られている。ベクターの全般的なクラスの非限定的な例は、限定されないが、二本鎖または一本鎖の直鎖または環状型の、ウイルスベクター、プラスミドベクター、ファージベクター、ファージミドベクター、コスミドベクター、フォスミドベクター、バクテリオファージ、人工染色体、またはアグロバクテリウムバイナリーベクターを含み、自己で伝染または動員が可能であってよく、または不可能であってよい。本明細書に定義されるベクターは、細胞ゲノムへの組み

40

50

により原核生物または真核生物宿主を形質転換することができ、または、染色体外に存在することができる（例えば複製起点を有する自己複製プラスミド）。天然にまたは設計により、放線菌および関連種、細菌および真核生物（例えば、より高度の植物、哺乳類、酵母または真菌細胞）から選択され得る2つの異なる宿主生物での複製が可能であるDNAビヒクルを意味する、シャトルベクターがさらに含まれる。一部の代表的な実施態様では、ベクター内の核酸は、宿主細胞での転写のための適切なプロモーターまたは他の調節エレメントの制御下であり、作動可能に連結されている。ベクターは、複数の宿主内で機能する二機能性発現ベクターであり得る。ゲノムDNAの場合は、これは、それ自身のプロモーターまたは他の調節エレメントを含み得て、cDNAの場合は、これは、宿主細胞での発現のための適切なプロモーターまたは他の調節エレメントの制御下であり得る。したがって、本発明の核酸分子、および／または、発現カセットは、本明細書に記載され、当該分野で知られているように、ベクター内に含まれ得る。

10

【0097】

目的のポリヌクレオチドの関連における、「導入する（Introducing）」、「導入する（introduce）」、「導入した（introduced）」（およびその文法的変形）は、目的のポリヌクレオチドを、宿主生物または前記生物の細胞（例えば、宿主細胞）に、ポリヌクレオチドが細胞内部へのアクセスを獲得する様式で、提供することを意味する。1よりも多いポリヌクレオチドが導入される場合は、これらのポリヌクレオチドは、単一のポリヌクレオチドまたは核酸構築物の一部として、または、別個のポリヌクレオチドまたは核酸構築物として集めることができ、そして、同一または異なる発現構築物または形質転換ベクター上に配置することができる。

20

【0098】

したがって、これらのポリヌクレオチドは、単一の形質転換イベントで、別個の形質転換／トランスフェクションイベントで、細胞に導入することができ、または、例えば、それらは、従来の繁殖プロトコルによって生物に組み込むことができる。したがって、一部の態様では、本発明の1つまたは複数の核酸構築物は、単一または組み合わせで、宿主生物または前記宿主生物の細胞に導入することができる。細胞の集団の関連における「導入する」は、集団を、本発明の異種核酸構築物と、本発明の異種核酸構築物が集団の1つまたは複数の細胞の内部へのアクセスを獲得する条件下で接触させて、それにより、集団の前記の1つまたは複数の細胞を形質転換することを意味する。

30

【0099】

本明細書において用いられる用語「形質転換」または「トランスフェクション」は、細胞への異種核酸の導入を指す。細胞の形質転換は、安定的または一過的であってよい。したがって、一部の実施態様では、宿主細胞または宿主生物は、本発明の核酸構築物によって安定的に形質転換される。他の実施態様では、宿主細胞または宿主生物は、本発明の核酸構築物によって一過的に形質転換される。したがって、代表的な実施態様では、本発明の異種核酸構築物は、安定的および／または一過的に細胞に導入することができる。

【0100】

ポリヌクレオチドの関連における「一過的な形質転換」は、ポリヌクレオチドが細胞に導入されて、細胞のゲノム中に組み込まれないことを意味する。

40

【0101】

ポリヌクレオチドの関連における「安定的に導入する」または「安定的に導入される」は、導入されたポリヌクレオチドが、細胞のゲノム中に安定的に組み込まれて、したがって、細胞が、ポリヌクレオチドによって安定的に形質転換されることを意味する。

【0102】

本明細書において用いられる「安定した形質転換」または「安定的に形質転換された」は、核酸構築物が細胞内に導入されて、細胞のゲノム内に組み込まれることを意味する。したがって、組み込まれた核酸構築物は、その子孫により、より具体的には、複数の後に続く世代の子孫により、受け継がれることが可能である。本明細書において用いられる「ゲノム」は、核、ミトコンドリアおよびプラスチドゲノムを含み、したがって、例えばプ

50

ラスミドまたはミトコンドリアゲノムへの核酸構築物の組込みを含んでよい。また、本明細書において用いられる安定した形質転換は、染色体外に例えばミニ染色体またはプラスミドとして維持される導入遺伝子も指し得る。

【0103】

一過的な形質転換は、生物内に導入された1つまたは複数の導入遺伝子によりコードされるペプチドまたはポリペプチドの存在を検出することのできる、酵素結合免疫吸着測定法（ELISA）またはウエスタンブロットなどにより、検出され得る。細胞の安定した形質転換は、生物（例えば、細菌、古細菌、酵母、藻類など）に導入された導入遺伝子のヌクレオチド配列と特異的にハイブリダイズする核酸配列を用いた、細胞のゲノムDNAのサザンブロットハイブリダイゼーションアッセイなどにより、検出することができる。細胞の安定した形質転換は、植物または他の生物に導入された導入遺伝子のヌクレオチド配列と特異的にハイブリダイズする核酸配列を用いた、細胞のRNAのノーザンブロットハイブリダイゼーションアッセイなどにより、検出することができる。また、細胞の安定した形質転換は、例えば、導入遺伝子の標的配列（単数または複数）とハイブリダイズする特異的なプライマー配列を用いて、ポリメラーゼ連鎖反応（PCR）または当技術分野でよく知られている他の増幅反応により検出することもでき、導入遺伝子配列の増幅をもたらす、標準的な方法に従って検出することができる。また、形質転換は、当技術分野でよく知られているダイレクトシーケンスおよび／またはハイブリダイゼーションのプロトコルにより検出することもできる。

【0104】

したがって、一部の実施態様では、ヌクレオチド配列、構築物、発現カセットは、一過的に発現することができ、および／または、それらは宿主生物のゲノム内に安定的に取り込まれ得る。

【0105】

本発明の異種核酸構築物は、当業者に公知の任意の方法によって細胞中に導入することができる。本発明の一部の実施態様では、細胞の形質転換は、核の形質転換を含む。また、さらなる実施態様では、本発明の異種核酸構築物（単数または複数）は、従来の繁殖技術を介して細胞中に導入することができる。

【0106】

真核生物および原核生物は両方とも、形質転換の手順は広く知られていて当該分野においてルーチンであり、文献全体を通して記載される（例えば、Jiang et al. 2013. Nat. Biotechnol. 31: 233-239; Ran et al. Nature Protocols 8: 2281-2308 (2013)を参照）。

【0107】

したがって、ヌクレオチド配列は、当該分野でよく知られている任意の数の方法で、宿主生物またはその細胞に導入することができる。本発明の方法は、それらが生物の少なくとも1つの細胞の内部に接近しさえすれば、1つまたは複数のヌクレオチド配列が生物内に導入されるための特定の方法に依存しない。1個よりも多いヌクレオチド配列が導入される場合は、それらは、単一の核酸構築物の一部として、または別個の核酸構築物として、集まることができ、同一または異なる核酸構築物上に位置することができる。したがって、ヌクレオチド配列は、単一の形質転換イベントで、または別個の形質転換イベントで、目的の細胞内に導入することができ、あるいは、関連する場合は、ヌクレオチド配列は、繁殖プロトコルの一部として生物内に取り込むことができる。

【0108】

可動遺伝要素（MGE）は、ゲノムの安定性に対する連続的なチャレンジを細菌に与えて、遺伝子水平伝播を通じた進化を促進させる。用語MGEは、プラスミド、バクテリオファージ、転位性エレメント、ゲノムアイランド、および多くの他の特別な遺伝因子を含む（1）。MGEは、高率の普及（high rates of dissemination）、宿主に対する適応性の利点、および、ゲノムの安定性を与える遺伝子を含むし、細菌ゲノムにおけるそれらのほぼ普遍的な存在をもたらす。捕食性のバクテリオフ

ファージおよび利己的な遺伝因子の永続的な脅威に立ち向かうために、細菌は、外因的な遺伝因子を標的化する先天性および適応性の両方の免疫系を進化させている。先天性の免疫は、細胞壁改変、制限／改変システム、および、未発達のファージ感染を含む(2)。Clustered regularly interspaced short palindromic repeats (CRISPR) およびCRISPR関連遺伝子(Cas)は、細菌における侵襲性遺伝因子に対して標的化された、適応性の免疫系である(3)。CRISPR-Casが仲介する免疫は、獲得(acquisition)、発現(expression)、および干渉(interference)として分類される別個の分子プロセスに依拠する(3)。獲得は、外来遺伝因子の分子「サンプリング」を介して生じて、そこからスペーサーと呼ばれる短い配列が、二極化された様式でCRISPRアレイ中に組み込まれる(4)。CRISPRアレイの発現は構成的であり、その前のリーダー配列内のプロモーターエレメントによって誘導可能である(5～6)。干渉は、それぞれのリピート配列において選択的に処理される対応する転写から生じて、スペーサーに相補の標的DNAの配列特異的な認識および切断のためにCasタンパク質をガイドするCRISPR RNA(crRNA)を形成する(7)。CRISPR-Cas技術は、株のタイピングおよび検出(8～10)、可動遺伝要素に対する天然／操作された免疫性の開発(11)、多様なバックグラウンドにおける、プログラム可能なゲノム編集(12)、転写制御(13～14)、および、規定されたコンソーシアム内の微生物集団の操作(15)での適用を有する。

10

【0109】

20

様々なCRISPRシステムが当技術分野で知られている。例えば、Makarova et al. を参照し、それは、CRISPRシステムの全てのクラス、タイプおよびサブタイプに関する命名法を記載する(Nature Reviews Microbiology 13: 722-736 (2015)); R. Barrangou (Genome Biol. 16: 247 (2015) も参照)。

【0110】

CRISPRアレイに対応する配列特徴が多く、生物で以前に記載されたが(16～17)、Streptococcus thermophilusは、特定のcas遺伝子およびCRISPRアレイ構成要素の役割が解明された最初の微生物だった(4)。S. thermophilusは、ヨーグルトおよび様々なチーズの栄養共生的な(syntrophic)生産においてラクトースを乳酸に異化するスターター培養として用いられる非病原の好熱性グラム陽性細菌である(18)。S. thermophilusは、4つまでのCRISPR-Casシステムをコードし、それらのうちの2つ(SthCRISPR1およびSthCRISPR3)は、獲得および干渉の両方において先天的に活性であるII-A型システムとして分類される(4、19)。したがって、S. thermophilusおよびそのバクテリオファージのゲノム解析は、ファージ／DNA保護に関するCRISPR-Casシステムと同様のメカニズムを確立した。S. thermophilusにおけるCRISPR-Casシステムの研究は、スペーサー起源の生物情報学的分析(4、20)、プロトスペーサー隣接モチーフ(PAM)配列の発見(19; 21)、ファージ宿主動態の理解(22～23)、Cas9エンドヌクレアーゼ活性の証明(7、24～25)、および、最近になって、II型システムのtracrRNA構造モチーフ管理機能および直交性の決定(26)をもたらした。S. thermophilusのゲノム解析は、炭水化物異化作用の喪失を通じたミルクに対する進化的適応および病原性Streptococcusに見られる病原性遺伝子を明らかにした(18)。また、S. thermophilusは、例えば、低温ショックタンパク質、銅耐性タンパク質、プロテイナーゼ、バクテリオシン、およびラクトース異化作用タンパク質をコードするものを含むニッチ関連遺伝子の重大な獲得を受けた(18)。挿入配列(IS)は、S. thermophilusゲノムにおいて高度に普及していて、そして、乳製品適応遺伝子と関連するアイランドの普及を促進することによって、株間の遺伝的異質性に寄与する(18)。付随する、S. thermophilusにおける機能性CRISP

30

40

50

R-Cas システムおよび MGE の存在は、ゲノムホメオスタシスが、これらの動的外力の相互作用によって少なくとも一部管理されることを示唆する。したがって、*S. thermophilus* は、ゲノムアイランドの CRISPR-Cas 標的化の遺伝的成果を研究するための理想的な宿主を構成する。

【0111】

CRISPR-Cas システムは、最近になって、ゲノム編集の適用において熱烈的な研究対象となっているが (12)、最も内因性の微生物システムの進化的役割は未知のままである (27)。外来 DNA 取り込みの防止を超えて (beyond) 活性の CRISPR-Cas システムをハウジングする進化的成果 (7)、スパーサー獲得イベント (4)、および、染色体の自己標的化に起因する突然変異 (28~32) に関してさえ、ほとんど知られていない。したがって、本発明者らは、内因性 II 型 CRISPR-Cas システムを用いて、組み込まれた MGE を標的化する成果を決定することを追求した。4 つのアイランドが *S. thermophilus* LMD-9 において同定されて、長さは 8~102 kbp の範囲であり、合計およそ 132 kbp であり、またはゲノムの 7% であった。ゲノムアイランドを標的化するために、自己標的化スパーサーを有する操作された CRISPR アレイの、プラスミドに基づく発現を、*S. thermophilus* LMD-9 に形質転換した。まとめると、我々の結果は、自己標的化イベントの基礎的な遺伝的成果を解明し、そして、CRISPR-Cas システムが、細菌集団レベルで、ゲノム進化を指示することができることを示す。

【0112】

これらの知見を利用して、本発明者らは、細菌、古細菌、藻類または酵母細胞の集団を、必須遺伝子、非必須遺伝子、および／または消費されてよい (expendible) ゲノムアイランドに関してスクリーニングするための；細菌、古細菌、藻類または酵母細胞の集団内の 1 つまたは複数の細胞を死滅させるための；細菌、古細菌、藻類または酵母細胞の表現型を特定するための；細菌、古細菌または酵母細胞の集団から、縮小されたゲノムサイズを有する 1 つまたは複数の細菌、古細菌、藻類または酵母細胞を選択するための；および／または、細菌、古細菌、藻類または酵母細胞の集団において、そのゲノム内に突然変異 (例えば、欠失) を有する少なくとも 1 つの単離物を同定するための、新規の方法を開発した。

【0113】

したがって、一態様では、本発明者らは、標的化を免れる能力をそれらに与える変更された遺伝的内容を有する集団における遺伝的変異を特定するための方法を開発した。ここで、標的配列は改変されていて、その改変を有する残存物を探す (one looks for survivors that have that modification)。一部の態様では、改変 (すなわち、突然変異) は、欠失である。さらに、標的配列が改変されている場合、その結果、野生型遺伝子型は必須ではない。

【0114】

したがって、本発明の一態様では、必須遺伝子、非必須遺伝子、および／または消費されてよいゲノムアイランドに関して、細菌細胞の集団をスクリーニングする方法が提供され：細菌細胞の前記集団に、(5' から 3' に) リピートスパーサーリピート配列または少なくとも 1 つのリピートスパーサー配列を含む CRISPR アレイを含む異種核酸構築物を導入するステップ (ここで、前記リピートスパーサーリピート配列または前記の少なくとも 1 つのリピートスパーサー配列のスパーサーは、前記集団の細菌細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、それにより、形質転換された細菌細胞の集団を生産する)；および、形質転換された細菌細胞の集団における欠失の存在または不存在を判定するステップを含む (ここで、形質転換された細菌細胞の集団における欠失の存在は、標的領域が、非必須遺伝子および／または消費されてよいゲノムアイランド内に含まれることを示し、集団における欠失の不存在は、標的領域が必須遺伝子内に含まれることを意味する)。本発明で有用な CRISPR アレイは、I 型、II 型、III 型、IV 型または V 型 CRISPR アレイであり得る。

【0115】

さらなる態様では、本発明は、必須遺伝子、非必須遺伝子、および／または消費されてよいゲノムアイランドに関して、細菌、古細菌、藻類または酵母細胞の集団をスクリーニングする方法を提供し、細菌、古細菌、藻類または酵母細胞の集団に：(a) トランス活性化CRISPR (tracr) 核酸を含む異種核酸構築物、(b) (5' から3' に) リピートスパーサーリピート配列または少なくとも1つのリピートスパーサー配列を含むCRISPRアレイ (crRNA、crDNA) を含む異種核酸構築物 (ここで、スパーサーは、前記集団の細菌、古細菌、藻類または酵母細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含む)、および、(c) Cas9ポリペプチドまたはCas9ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップ (それにより、形質転換された細菌、古細菌、藻類または酵母細胞の集団を生産する) ; および、形質転換された細菌、古細菌、藻類または酵母細胞の集団における欠失の存在または不存在を判定するステップを含む (ここで、形質転換された細菌、古細菌または酵母細胞の集団における欠失の存在は、標的領域が非必須遺伝子および／または消費されてよいゲノムアイランド内に含まれることを意味し、形質転換された細菌、古細菌、藻類または酵母細胞の集団における欠失の不存在は、標的領域が必須遺伝子内に含まれることを意味する) 。

10

【0116】

他の態様では、細菌細胞の集団に、リピートスパーサーリピート配列または少なくとも1つのリピートスパーサー配列を含むCRISPRアレイ (crRNA、crDNA) を含む異種核酸構築物を導入するステップを含む、細菌細胞の集団内の1つまたは複数の細菌細胞を死滅させる方法が提供され、ここで、前記リピートスパーサーリピート配列または少なくとも1つのリピートスパーサー配列のスパーサーは、前記集団の細菌細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、それにより、集団内の標的領域を含む1つまたは複数の細菌細胞を死滅させる。本発明で有用なCRISPRアレイは、I型、II型、III型、IV型またはV型CRISPRアレイであり得る。

20

【0117】

さらなる態様では、細菌、古細菌、藻類または酵母細胞の集団内の1つまたは複数の細胞を死滅させる方法が提供され、当該方法は：細菌、古細菌または酵母細胞の集団に (a) トランス活性化CRISPR (tracr) 核酸を含む異種核酸構築物、(b) リピートスパーサーリピート配列または少なくとも1つのリピートスパーサー配列を含むCRISPRアレイ (crRNA、crDNA) を含む異種核酸構築物 (ここで、前記の少なくとも1つのリピートスパーサー配列およびリピートスパーサーリピート配列のスパーサーは、前記集団の細菌、古細菌、藻類または酵母細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含む)、および、(c) Cas9ポリペプチドおよび／またはCas9ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップを含み、それにより、細菌、古細菌、藻類または酵母細胞の集団内のこれらのゲノム内の標的領域を含む1つまたは複数の細胞を死滅させる。

30

【0118】

細菌ゲノム標的化CRISPR RNAの形質転換を用いて、配列特異的な基準で細菌細胞を選択的に死滅させて、遺伝学的に別個の亜集団を差し引き、それにより、標的配列を欠く細菌集団を濃縮することができる。この区別は、遺伝学的に類似する集団内のオルトゴナルCRISPR-Casシステムの不均一分配に基づいて生じ得る。したがって、一部の実施態様では、細胞の集団に導入されるCRISPRアレイは、死滅される前記の1つまたは複数の細菌細胞におけるCRISPR-Casシステムと適合することができるが (すなわち、機能性)、集団内の少なくとも1つまたは複数の細菌細胞のCRISPR-Casシステムと適合しない (すなわち、機能性でない)。例えば、Escherichia coli および Klebsiella pneumoniae は、I-E型またはI-F型CRISPR-Casシステムのいずれかを示すことができ; Clostr

40

50

idium difficile は、I-B 型システムをコードし、および、*S. thermophilus* の異なる株は、II-A 型および I-E 型システムの両方または II-A 型システムのみを示す。細菌の混合物中に形質転換された特定の CRISPR RNA に応じて、その同起源のシステムとのその機能性の適合性に基づいて、集団のそのサブセットを特異的に標的化することができる。これは、内因性 CRISPR-Cas システムを含む多様な種、例えば、制限されないが：*Pseudomonas* spp. (例えば：*P. aeruginosa*)、*Escherichia* spp. (例えば：*E. coli*)、*Enterobacter* spp. (例えば：*E. cloacae*)、*Staphylococcus* spp. (例えば：*S. aureus*)、*Enterococcus* spp. (例えば：*E. faecalis*、*E. faecium*)、*Streptomyces* spp. (例えば：*S. somaliensis*)、*Streptococcus* spp. (例えば：*S. pyogenes*)、*Vibrio* spp. (例えば：*V. cholerae*)、*Yersinia* spp. (例えば：*Y. pestis*)、*Francisella* spp. (例えば：*F. tularensis*、*F. novicida*)、*Bacillus* spp. (例えば：*B. anthracis*、*B. cereus*)、*Lactobacillus* spp. (例えば：*L. casei*、*L. reuteri*、*L. acidophilus*、*L. rhamnosus*)、*Burkholderia* spp. (例えば：*B. mallei*、*B. pseudomallei*)、*Klebsiella* spp. (例えば：*K. pneumoniae*)、*Shigella* spp. (例えば：*S. dysenteriae*、*S. sonnei*)、*Salmonella* spp. (例えば：*S. enterica*)、*Borrelia* spp. (例えば：*B. burgdorferi*)、*Neisseria* spp. (例えば：*N. meningitidis*)、*Fusobacterium* spp. (例えば：*F. nucleatum*)、*Helicobacter* spp. (例えば：*H. pylori*)、*Chlamydia* spp. (例えば：*C. trachomatis*)、*Bacteroides* spp. (例えば：*B. fragilis*)、*Bartonella* spp. (例えば：*B. quintana*)、*Bordetella* spp. (例えば：*B. pertussis*)、*Brucella* spp. (例えば：*B. abortus*)、*Campylobacter* spp. (例えば：*C. jejuni*)、*Clostridium* spp. (例えば：*C. difficile*)、*Bifidobacterium* spp. (例えば：*B. infantis*)、*Haemophilus* spp. (例えば：*H. influenzae*)、*Listeria* spp. (例えば：*L. monocytogenes*)、*Legionella* spp. (例えば：*L. pneumophila*)、*Mycobacterium* spp. (例えば：*M. tuberculosis*)、*Mycoplasma* spp. (例えば：*M. pneumoniae*)、*Rickettsia* spp. (例えば：*R. rickettsii*)、*Acinetobacter* spp. (例えば：*A. calcoaceticus*、*A. baumannii*)、*Ruminococcus* spp. (例えば：*R. albus*)、*Propionibacterium* spp. (例えば：*P. freudenreichii*)、*Corynebacterium* spp. (例えば：*C. diphtheriae*)、*Propionibacterium* spp. (例えば：*P. acnes*)、*Brevibacterium* spp. (例えば：*B. iodinum*)、*Micrococcus* spp. (例えば：*M. luteus*)、および／または、*Prevotella* spp. (例えば：*P. histocola*) に適用することができる。

【0119】

CRISPR 標的化は、混合された集団内の別個の遺伝的内容に基づいて、特定の細菌のサブセットを取り除くことができる。このクレームに関するサポートは、実施例 4、5 に示され、そこでは、 Lac^- 細菌が選択されているが Lac^+ は集団から取り除かれている。 Lac^+ および Lac^- 株間の遺伝的区別は、実施例 8 および 10 に示され、そこ

では、残存しているクローンのシーケンシングは、参照の野生型 *S. thermophilus* 株と比較して、遺伝的内容に 5.5% までの違いを明らかにした。したがって、CRISPR-標的化スパーサーは、保存されたまたは異なった遺伝子配列を標的化することにより、様々なレベルの細菌関連性に調和され得る。したがって、一部の実施態様では、集団内の細菌細胞は、同一の CRISPR-Cas システムを含むことができ、導入された CRISPR アレイは、したがって、全体として細菌集団において機能性であり得るが、細菌集団を作り上げる異なる株または種の遺伝的内容は、導入された CRISPR アレイに関する標的領域が、死滅される集団の 1 つまたは複数の細菌種においてのみ見られるように十分に異なる。これは、内因性 CRISPR-Cas システムを含む多様な種、例えば、制限されないが：*Pseudomonas* spp. (例えば：*P. aeruginosa*)、*Escherichia* spp. (例えば：*E. coli*)、*Enterobacter* spp. (例えば：*E. cloacae*)、*Staphylococcus* spp. (例えば：*S. aureus*)、*Enterococcus* spp. (例えば：*E. faecalis*、*E. faecium*)、*Streptomyces* spp. (例えば：*S. somaliensis*)、*Streptococcus* spp. (例えば：*S. pyogenes*)、*Vibrio* spp. (例えば：*V. cholerae*)、*Yersinia* spp. (例えば：*Y. pestis*)、*Francisella* spp. (例えば：*F. tularensis*、*F. novicida*)、*Bacillus* spp. (例えば：*B. anthracis*、*B. cereus*)、*Lactobacillus* spp. (例えば：*L. casei*、*L. reuteri*、*L. acidophilus*、*L. rhamnosus*)、*Burkholderia* spp. (例えば：*B. mallei*、*B. pseudomallei*)、*Klebsiella* spp. (例えば：*K. pneumoniae*)、*Shigella* spp. (例えば：*S. dysenteriae*、*S. sonnei*)、*Salmonella* spp. (例えば：*S. enterica*)、*Borrelia* spp. (例えば：*B. burgdorferi*)、*Neisseria* spp. (例えば：*N. meningitidis*)、*Fusobacterium* spp. (例えば：*F. nucleatum*)、*Helicobacter* spp. (例えば：*H. pylori*)、*Chlamydia* spp. (例えば：*C. trachomatis*)、*Bacteroides* spp. (例えば：*B. fragilis*)、*Bartonella* spp. (例えば：*B. quintana*)、*Bordetella* spp. (例えば：*B. pertussis*)、*Brucella* spp. (例えば：*B. abortus*)、*Campylobacter* spp. (例えば：*C. jejuni*)、*Clostridium* spp. (例えば：*C. difficile*)、*Bifidobacterium* spp. (例えば：*B. infantis*)、*Haemophilus* spp. (例えば：*H. influenzae*)、*Listeria* spp. (例えば：*L. monocytogenes*)、*Legionella* spp. (例えば：*L. pneumophila*)、*Mycobacterium* spp. (例えば：*M. tuberculosis*)、*Mycoplasma* spp. (例えば：*M. pneumoniae*)、*Rickettsia* spp. (例えば：*R. rickettsii*)、*Acinetobacter* spp. (例えば：*A. calcoaceticus*、*A. baumannii*)、*Ruminococcus* spp. (例えば：*R. albus*)、*Propionibacterium* spp. (例えば：*P. freudenreichii*)、*Corynebacterium* spp. (例えば：*C. diphtheriae*)、*Propionibacterium* spp. (例えば：*P. acnes*)、*Brevibacterium* spp. (例えば：*B. iodinum*)、*Micrococcus* spp. (例えば：*M. luteus*)、および／または *Prevotella* spp. (例えば：*P. histocola*) に適用することができる。

【0120】

本発明の方法を用いた、集団内の死滅程度は、標的領域が非必須遺伝子、必須遺伝子ま

たは消費されてよいアイランド内に含まれるかどうかに加えて、形質転換に対する特定の集団の従順さによって影響を受け得る。細菌、古細菌または酵母細胞の集団内の死滅程度は、例えば、生物、属および種によって変化し得る。したがって、本明細書において用いられる「死滅」は、集団内の細胞 $2 \log$ またはそれよりも多くの除去を意味する（ 1% 生存またはそれよりも低い）。 $1 \log$ 未満の死滅は、集団内のわずかな縮小であり；一方で、 $2 \sim 3 \log$ の死滅は、集団の著しい縮小をもたらす；および、 $3 \log$ を超える死滅は、集団が実質的に全滅していることを示す。

【0121】

別の態様では、細菌遺伝子と関連する表現型を同定する方法が提供され、細菌細胞の集団に、リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイ（crRNA、crDNA）を含む異種核酸構築物を導入するステップ（ここで、前記の少なくとも1つのリピートスペーサー配列およびリピートスペーサーリピート配列のスペーサーは、前記集団の細菌細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、ここで、標的領域は、ポリペプチドまたは機能性核酸をコードするオープンリーディングフレームの少なくとも一部を含み、それにより、標的領域を含む細胞を死滅させ、そして、標的領域を含まない形質転換された細菌細胞の集団を生産する）（すなわち、生き残った細胞は、標的領域を含まない）；および、（i）細胞の集団の表現型を分析するステップ、または、（ii）形質転換された細菌細胞の集団から個々の細菌コロニーを増殖させるステップ、および、個々のコロニーの表現型を分析するステップ、を含む。本発明で有用なCRISPRアレイは、I型、II型、III型、IV型またはV型CRISPRアレイであり得る。

【0122】

別の態様では、細菌、古細菌、藻類、または酵母遺伝子の表現型を特定する方法が提供され、細菌、古細菌、藻類または酵母細胞の集団に、（a）トランス活性化CRISPR（tracr）核酸を含む異種核酸構築物、（b）（5'から3'に）リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイ（crRNA、crDNA）を含む異種核酸構築物（ここで、スペーサーは、前記集団の細菌、古細菌、藻類または酵母細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含む）、および、（c）Cas9ポリペプチドおよび／またはCas9ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップを含み（それにより、標的領域を含む細菌、古細菌、藻類または酵母細胞を死滅させて、そして、標的領域を有さない形質転換された細菌、古細菌、藻類または酵母細胞の集団を生産する）；および、（i）細胞の集団の表現型を分析するステップ、および／または、（ii）形質転換された細菌、古細菌、藻類または酵母細胞の集団から、個々の細菌、古細菌、または酵母コロニーを増殖させるステップ；および、個々のコロニーの表現型を分析するステップ、を含む。

【0123】

一部の実施態様では、当該分析は、PCR、光学ゲノムマッピング、ゲノムシーケンシング、制限マッピングおよび／または突然変異を特定および特徴付けるための制限分析、および相補性分析および／または表現型を分析するための表現型分析を含む。

【0124】

本発明の一部の実施態様では、集団における死滅または縮小の程度を判定するステップは、制限されないが、（1）細胞をプレーティングしてコロニーを計数するステップ、（2）光学濃度、（3）顕微鏡計数、（4）最確数、および／または、（5）メチレンブルー還元を含む、集団数を判定するための任意の方法を含むことができる。一部の実施態様では、 $16S$ rDNAシーケンシングを用いて、混合された集団の組成をプロファイルすることができる。これは、例えば、全体としてのサンプルからDNAを精製して、ハイスループット技術を用いた全ゲノムショットガンシーケンシングを行なうことにより、または、例えば、 $16S$ 遺伝子をPCR増幅して産物を同様の様式でシーケンシングすることによって、行なうことができる。それから、配列は、特定の細菌分類群にコンピュータ

一的に割り当てられ得る。他の実施態様では、定量PCR法を用いて細菌レベルを定量してもよい。そのような技術は当技術分野でよく知られている。例えば、qPCRのためのプライマーは、配列を共有する生物の株種、属、またはグループから特異的に増幅するように設計することができる。したがって、閾値数（ct）を用いて、前記生物または生物のグループを定量し得る。さらなる実施態様では、標的集団に特異的な任意の細菌活性（表現型）を測定基準として用いて、集団の枯渇を判定してもよい。

【0125】

さらなる実施態様では、細菌細胞の集団から、縮小されたゲノムサイズを有する1つまたは複数の細菌細胞を選択する方法が提供され、細菌細胞の集団に、（5'から3'に）リピートスパーサーリピート配列または少なくとも1つのリピートスパーサー配列を含むCRISPRアレイ（crRNA、crDNA）を含む異種核酸構築物を導入するステップを含み、ここで、スパーサーは、前記集団の1つまたは複数の細菌細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、ここで、標的領域を含む細胞は死滅して、それにより、細菌細胞の集団から、標的領域を有さない縮小されたゲノムサイズを有する1つまたは複数の細菌細胞を選択する。本発明で有用なCRISPRアレイは、I型、II型、III型、IV型またはV型CRISPRアレイであり得る。

【0126】

一部の実施態様では、細菌細胞の集団から、縮小されたゲノムサイズを有する1つまたは複数の細菌細胞を選択する方法は、細菌細胞の集団に：（a）（i）前記細菌細胞のゲノム内に存在する少なくとも300個の連続したヌクレオチドと少なくとも80%の同一性を有するヌクレオチド配列を含む、1つまたは複数の異種核酸構築物、または（ii）少なくとも1つのトランスポゾンを含む2以上の異種核酸構築物（それにより、ゲノム中に組み込まれた前記の1つまたは複数の異種核酸構築物およびゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、または、ゲノム中に組み込まれた第1および第2のトランスポゾンの間の、相同組み換えに関する非天然部位を含むトランスジェニック細菌細胞の集団を生産する）；および（b）（5'から3'に）リピートスパーサーリピート配列または少なくとも1つのリピートスパーサー配列を含むCRISPRアレイ（crRNA、crDNA）を含む異種核酸構築物、を導入するステップを含む（ここで、前記リピートスパーサーリピート配列または少なくとも1つのリピートスパーサー配列のスパーサーは、前記集団の1つまたは複数の細菌細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、標的領域は、ゲノムに導入された前記の1つまたは複数の異種核酸構築物およびゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、および／または、第一のトランスポゾンおよび第二のトランスポゾンの間に位置し、ここで、標的領域を含む細胞は死滅して、そして、標的領域を含まない細胞は生き残り、それにより、トランスジェニック細菌細胞の集団から、標的領域を有さない縮小されたゲノムサイズを有する1つまたは複数の細菌細胞を選択する）。本発明で有用なCRISPRアレイは、I型、II型、III型、IV型またはV型CRISPRアレイであり得る。

【0127】

当技術分野でよく知られているように、トランスポゾンは、例えば、PCR増幅を介して、または、設計されたDNA合成を通して作ることができ、そして、形質転換の任意の方法を介して導入され得る。

【0128】

一部の実施態様では、本発明は、細菌、古細菌、藻類または酵母細胞の集団から、縮小されたゲノムサイズを有する1つまたは複数の細菌、古細菌、藻類または酵母細胞を選択する方法を提供し、細菌、古細菌または酵母細胞の集団に、（a）トランス活性化CRISPR（tracr）核酸を含む異種核酸構築物、（b）リピートスパーサーリピート配列または少なくとも1つのリピートスパーサー配列を含むCRISPRアレイ（crRNA、crDNA）を含む異種核酸構築物（ここで、前記の少なくとも1つのリピートスパーサー配列および前記の少なくとも1つのリピートスパーサーリピート配列

のスペーサーは、前記集団の細菌、古細菌、藻類または酵母細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含む）、および、（c）Cas9ポリペプチドおよび／またはCas9ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップを含む（ここで、標的領域を含む細胞は死滅して、それにより、細菌、古細菌、藻類または酵母細胞の集団から、標的領域を有さない縮小されたゲノムサイズを有する1つまたは複数の細菌、古細菌、藻類または酵母細胞を選択する）。

【0129】

他の実施態様では、細菌、古細菌、藻類または酵母細胞の集団から、縮小されたゲノムサイズを有する1つまたは複数の細菌、古細菌、藻類または酵母細胞を選択する方法が提供され、細菌、古細菌、藻類または酵母細胞の集団に：（a）（i）前記の細菌、古細菌、藻類または酵母細胞のゲノム内に存在する少なくとも300個の連続したヌクレオチドと少なくとも80%の同一性を有するヌクレオチド配列を含む、1つまたは複数の異種核酸構築物、または（ii）少なくとも1つのトランスポゾンを含む2以上の異種核酸構築物（それにより、ゲノム中に組み込まれた前記の1つまたは複数の異種核酸構築物およびゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、または、ゲノム中に組み込まれた第1および第2のトランスポゾンの間の、相同組み換えに関する非天然部位を含むトランスジェニック細菌、古細菌、藻類または酵母細胞の集団を生産する）；および（b）（i）トランス活性化CRISPR（tracr）核酸を含む異種核酸構築物、（ii）リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイ（crRNA、crDNA）を含む異種核酸構築物（ここで、前記の少なくとも1つのリピートスペーサー配列または前記の少なくとも1つのリピートスペーサーリピート配列のスペーサーは、前記集団の1つまたは複数の細菌、古細菌、藻類または酵母細胞のゲノム内の標的領域に実質的に相補のヌクレオチド配列を含む）、および、（iii）Cas9ポリペプチドおよび／またはCas9ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップを含む（ここで、標的領域は、ゲノム中に組み込まれた前記の1つまたは複数の異種核酸構築物およびゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、および／または、第一のトランスポゾンおよび第二のトランスポゾンの間に位置し、ここで、標的領域を含む細胞は死滅して、そして、標的領域を含まない細胞は生き残り、それにより、トランスジェニック細菌、古細菌、藻類または酵母細胞の集団から、標的領域を有さない縮小されたゲノムサイズを有する1つまたは複数の細菌、古細菌、藻類または酵母細胞を選択する）。

【0130】

一部の態様では、縮小されたゲノムサイズは、コントロールと比較して縮小され得る。一部の態様では、コントロールは、細菌、古細菌、藻類または酵母細胞の野生型集団、または、リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイ（例えば、I型、II型、III型、IV型またはV型CRISPRアレイ）を含む異種構築物によって形質転換された細菌、古細菌、藻類または酵母細胞の集団であってよく、ここで、前記のリピートスペーサーリピート配列または前記の少なくとも1つのリピートスペーサー配列のスペーサーは、前記集団の細菌、古細菌、藻類または酵母細胞のゲノム内の標的領域に相補でないヌクレオチド配列を含む（すなわち、非自己標的化／「スクランブルスペーサー」）。さらなる態様では、コントロールは、リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイを含む異種構築物によって形質転換された細菌、古細菌、藻類または酵母細胞の集団であってよく、ここで、前記リピートスペーサーリピート配列または前記の少なくとも1つのリピートスペーサー配列のスペーサーは、前記集団の細菌、古細菌、藻類または酵母細胞のゲノム内の標的領域に実質的に相補であるがプロトスペーサー隣接モチーフ（PAM）を欠くヌクレオチド配列を含む。

【0131】

一部の実施態様では、細菌の集団において、そのゲノム内に欠失（例えば、染色体およ

び／またはプラスミド欠失)を有する少なくとも1つの単離物を同定する方法が提供され、細菌細胞の集団に、(5'から3'に)リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイ(c rRNA、c rDNA)を含む異種核酸構築物を導入するステップ(前記のリピートスペーサーリピート配列または前記の少なくとも1つのリピートスペーサー配列のスペーサーは、前記集団の1つまたは複数の細菌細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、そして、標的領域を含む細胞は死滅して、それにより、標的領域を含まない形質転換された細菌細胞の集団を生産する)；および、形質転換された細菌細胞の集団から個々の細菌コロニーを増殖させるステップを含み、それにより、形質転換された細菌の集団から、そのゲノム内に欠失を有する少なくとも1つの単離物を同定する。本発明で有用なCRISPRアレイは、I型、II型、III型、IV型またはV型CRISPRアレイであり得る。

10

【0132】

さらなる実施態様では、本発明は、細菌の集団において、そのゲノム内に欠失を有する少なくとも1つの単離物を同定する方法を提供し、細菌細胞の集団に：(a)(i)前記細菌細胞のゲノム内に存在する少なくとも300個の連続したヌクレオチドと少なくとも80%の同一性を有するヌクレオチド配列を含む、1つまたは複数の異種核酸構築物、または、(ii)少なくとも1つのトランスポゾンを含む2以上の異種核酸構築物(それにより、ゲノム中に組み込まれた前記の1つまたは複数の異種核酸構築物およびゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、または、ゲノム中に組み込まれた第1および第2のトランスポゾンの間の、相同組み換えに関する非天然部位を含むトランスジェニック細菌細胞の集団を生産する)；およびb)リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイ(c rRNA、c rDNA)を含む異種核酸構築物、を導入するステップ(ここで、スペーサーは、前記集団の1つまたは複数の細菌細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含み、ここで、標的領域は、ゲノム中に導入された前記の1つまたは複数の異種核酸構築物およびゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、および／または、第1のトランスポゾンおよび第2のトランスポゾンの間に位置し、標的領域を含む細胞は死滅して[標的領域を含まない細胞は生き残る]、それにより、標的領域を含まない形質転換された細菌細胞の集団を生産する)；および、形質転換された細菌細胞の集団から、個々の細菌コロニーを増殖させるステップを含み、それにより、細菌の集団から、そのゲノム内に欠失を有する少なくとも1つの単離物を同定する。本発明で有用なCRISPRアレイは、I型、II型、III型、IV型またはV型CRISPRアレイであり得る。

20

30

【0133】

さらなる実施態様では、細菌、古細菌、藻類または酵母細胞の集団において、そのゲノム内に欠失を有する少なくとも1つの単離物を同定する方法が提供され、細菌、古細菌、藻類または酵母細胞の集団に：(a)トランス活性化CRISPR(t r a c r)核酸を含む異種核酸構築物；(b)リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイ(c rRNA、c rDNA)を含む異種核酸構築物(ここで、前記リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列のスペーサーは、前記集団の細菌、古細菌、藻類または酵母細胞のゲノム(例えば、染色体、ミトコンドリアおよび／またはプラスミドのゲノム)内の標的領域に実質的に相補なヌクレオチド配列を含む)；および(c)Cas9ポリペプチドまたはCas9ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップ(ここで、標的領域を含む細胞は死滅して、それにより、標的領域を含まない形質転換された細菌、古細菌、藻類または酵母細胞の集団を生産する)；および形質転換された細菌、古細菌、藻類または酵母細胞の集団から、個々の細菌、古細菌または酵母コロニーを増殖させるステップを含み、それにより、形質転換された細菌、古細菌、藻類または酵母細胞の集団から、そのゲノム内に欠失を有する少なくとも1つの単

40

50

離物を同定する。

【0134】

また、さらなる実施態様では、本発明は、細菌、古細菌、藻類または酵母細胞の集団において、そのゲノム内に欠失を有する少なくとも1つの単離物を同定する方法を提供し、細菌、古細菌、藻類または酵母細胞の集団に：(a)(i)前記の細菌、古細菌、藻類または酵母細胞のゲノム内に存在する少なくとも300個の連続したヌクレオチドと少なくとも80%の同一性を有するヌクレオチド配列を含む、1つまたは複数の異種核酸構築物、または(ii)少なくとも1つのトランスポゾンを含む2以上の異種核酸構築物(それにより、ゲノム中に組み込まれた前記の1つまたは複数の異種核酸構築物およびゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、または、ゲノム中に組み込まれた第1および第2のトランスポゾンの間の、相同組み換えに関する非天然部位を含むトランスジェニック細菌、古細菌、藻類または酵母細胞の集団を生産する)；および、(b)(i)トランス活性化CRISPR(tracr)核酸を含む異種核酸構築物、(ii)リピートスペーサーリピート配列または少なくとも1つのリピートスペーサー配列を含むCRISPRアレイ(crRNA、crDNA)を含む異種核酸構築物(ここで、スペーサーは、前記集団の1つまたは複数の細菌、古細菌、藻類または酵母細胞のゲノム内の標的領域に実質的に相補なヌクレオチド配列を含む)；および、(iii)Cas9ポリペプチドおよび/またはCas9ポリペプチドをコードするポリヌクレオチドを含む異種核酸構築物、を導入するステップ(ここで、標的領域は、ゲノム中に組み込まれた前記の1つまたは複数の異種核酸構築物およびゲノム内に存在する前記の少なくとも300個の連続したヌクレオチドの間、および/または、第1のトランスポゾンおよび第2のトランスポゾンの間に位置し、標的領域を含む細胞は死滅して、標的領域を含まない細胞は生き残り、それにより、標的領域を含まない形質転換された細菌、古細菌、藻類または酵母細胞の集団を生産する)；および、形質転換された細菌、古細菌、藻類または酵母細胞の集団から、個々の細菌、古細菌または酵母コロニーを増殖させるステップを含み、それにより、集団から、そのゲノム内に欠失を有する少なくとも1つの単離物を同定する。

【0135】

一部の実施態様では、適合性(fitness)/増殖速度は、ゲノムサイズの縮小によって、または、転写および翻訳のために高いエネルギーインプットを必要とする選択遺伝子(ポリペプチドまたは機能性核酸(例えば、転写調節因子)をコードする)の欠失によって、増大され得る。したがって、一部の実施態様では、細菌、古細菌、藻類または酵母細胞の集団の適合性または増殖速度を増大させる方法が提供され、縮小されたゲノムサイズ(例えば、ゲノムの一部の不存在を選択する)および/または本明細書に記載の集団の細菌、古細菌、藻類または酵母細胞のゲノム内の欠失を選択するステップを含む。一部の実施態様では、欠失は、1つの遺伝子または1つよりも多い遺伝子を含んでよい。したがって、ゲノムサイズの縮小または特定の遺伝子(単数または複数)の欠失を通して、集団の細胞は、存在しないゲノムの部分または欠失した遺伝子(単数または複数)の転写/翻訳に、もはやエネルギーを消費せず、それにより、前記のゲノムの部分および/または前記の遺伝子(単数または複数)をまだ含んでいるコントロール集団と比較して、縮小されたエネルギー必要性および増大された適合性を有する。

【0136】

他の実施態様では、細菌、古細菌、藻類または酵母細胞の集団から生産される産物の量を増大させる方法が提供されて、本明細書に記載の細菌、古細菌、藻類または酵母細胞のゲノム内の欠失に関して選択することにより、細胞の適合性または増殖速度を増大させるステップを含む。一部の実施態様では、産物は、限定されないが、抗生物質、二次代謝物、ビタミン類、タンパク質、酵素、酸、および製剤を含むことができる。

【0137】

一部の実施態様では、本発明で有用なCRISPRアレイ(crRNA、crDNA)は、I型CRISPR-Casシステム、II型CRISPR-Casシステム、III

型CRISPR-Casシステム、IV型CRISPR-Casシステム、またはV型CRISPR-Casシステムのいずれか由来のアレイであってよい。

【0138】

先行する実施態様に関して、tracr核酸を含む異種核酸構築物およびCRISPRアレイを含む異種核酸構築物は、同一の構築物（例えば、発現カセットまたはベクター）中に、または、異なる構築物中に、含まれてよく、または導入されてよい。特定の実施態様では、tracr核酸を含む異種核酸構築物およびCRISPRアレイを含む異種核酸構築物は、場合によりCas9ポリペプチドをコードするポリヌクレオチドをさらに含んでよい単一の構築物（例えば、発現カセットおよび／またはベクター）内に含まれてよい。一部の実施態様では、tracr核酸を含む異種核酸構築物およびCRISPRアレイを含む異種核酸構築物は、単一のプロモーターおよび／または別個のプロモーターに動作可能に連結されてよい。

10

【0139】

一部の実施態様では、トランス活性化CRISPR (tracr) 核酸を含む異種核酸構築物およびCRISPRアレイ (crRNA、crDNA) を含む異種核酸構築物は、CRISPRガイド (gRNA、gDNA) 内に含まれてよい。一部の実施態様では、CRISPRガイドは、プロモーターに動作可能に連結されてよい。

【0140】

一部の実施態様では、本発明で有用なCas9ポリペプチドは、Cas9ヌクレアーゼのアミノ酸配列と、少なくとも70%の同一性を含む（例えば、約70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%など）。本発明で有用な、例示的なCas9ヌクレアーゼは、CRISPR-CasシステムにおいてDNA切断を触媒することが知られる任意のCas9ヌクレアーゼであってよい。当技術分野で知られているように、そのようなCas9ヌクレアーゼは、HNHモチーフおよびRuvCモチーフを含む（例えば、WO2013/176772；WO/2013/188638を参照）。一部の実施態様では、Cas9ヌクレアーゼの機能性フラグメントは、本発明で有用であり得る。

20

【0141】

CRISPR-CasシステムおよびCas9ヌクレアーゼのグルーピングは、当技術分野でよく知られていて、例えば、Cas9ヌクレアーゼのStreptococcus thermophilus CRISPR 1 (SthnCR1) グループ、Cas9ヌクレアーゼのStreptococcus thermophilus CRISPR 3 (Sth CR3) グループ、Cas9ヌクレアーゼのLactobacillus buchneri CD034 (Lb) グループ、および、Cas9ヌクレアーゼのLactobacillus rhamnosus GG (Lrh) グループを含む。さらなるCas9ヌクレアーゼは、限定されないが、Lactobacillus curvatus CRL 705のものを含む。さらに、本発明で有用なさらなるCas9ヌクレアーゼは、限定されないが、Lactobacillus animalis KCTC 3501およびLactobacillus farciminis WP 010018949.1由来のCas9を含む。

30

40

【0142】

さらに、特定の実施態様では、Cas9ヌクレアーゼは、標的DNAを含む生物に関してコドン最適化されたヌクレオチド配列によってコードされ得る。さらに他の実施態様では、Cas9ヌクレアーゼは、少なくとも1つの核局在化配列を含むことができる。

【0143】

一部の実施態様では、本発明で有用なI型ポリペプチドは、Cas3、Cas3'ヌクレアーゼ、Cas3''ヌクレアーゼ、それらの融合変異体のアミノ酸配列と、少なくとも70%の同一性を含む（例えば、約70%、71%、72%、73%、74%、75%

50

、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%など）。一部の実施態様では、本発明で有用なI型Cascadeポリペプチドは、Cas7 (Csa2)、Cas8a1 (Csx13)、Cas8a2 (Csx9)、Cas5、Csa5、Cas6a、Cas6b、Cas8b (Csh1)、Cas7 (Csh2)、Cas5、Cas5d、Cas8c (Csd1)、Cas7 (Csd2)、Cas10d (Csc3)、Csc2、Csc1、Cas6d、Cse1 (CasA)、Cse2 (CasB)、Cas7 (CasC)、Cas5 (CasD)、Cas6e (CasE)、Cys1、Cys2、Cas7 (Cys3)、Cas6f (Csy4)、Cas6および／またはCas4のアミノ酸配列と、少なくとも70%の同一性を含む（例えば、約70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%など）。

【0144】

I型CRISPR-Casシステムは、当技術分野でよく知られていて、例えば、例示的なI-A型CRISPR-Casシステムを含む*Archaeoglobus fulgidus*、例示的なI-B型CRISPR-Casシステムを含む*Clostridium kluyveri* DSM 555、例示的なI-C型CRISPR-Casシステムを含む*Bacillus halodurans* C-125、例示的なI-D型CRISPR-Casシステムを含む*Cyanothecce* sp. PCC 802、例示的なI-E型CRISPR-Casシステムを含む*Escherichia coli* K-12、例示的なI-U型CRISPR-Casシステムを含む*Geobacter sulfurreducens* および、例示的なI-F型CRISPR-Casシステムを含む*Yersinia pseudotuberculosis* YPIIIを含む。

【0145】

一部の実施態様では、本発明で有用なII型ポリペプチドは、Cas9のアミノ酸配列と、少なくとも70%の同一性を含む（例えば、約70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%など）。II型CRISPR-Casシステムは、当技術分野でよく知られていて、例えば、*Legionella pneumophila* str. Paris、*Streptococcus thermophilus* CNRZ1066、および、*Neisseria lactamica* 020-06を含む。

【0146】

一部の実施態様では、本発明で有用なIII型ポリペプチドは、Cas6、Cas10（またはCsm1）、Csm2、Csm3、Csm4、Csm5、およびCsm6、Cmr1、Cas10（またはCmr2）、Cmr3、Cmr4、Cmr5、およびCmr6、Cas7、Cas10、Cas7 (Csm3)、Cas5 (Csm4)、Cas7 (Csm5)、Csm6、Cas7 (Cmr1)、Cas5 (Cmr3)、Cas7 (Cmr4)、Cas7 (Cmr6)、Cas7 (Cmr6)、Cmr5、Cas5 (Cmr3)、Cas5 (Cs×10)、Csm2、Cas7 (Csm3)、およびall1473のアミノ酸配列と、少なくとも70%の同一性を含む（例えば、約70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%など）。III型CRISPR-Casシステムは、当技術分野でよく知られていて、例えば、例示的なIII-A型CRISPR-Casシステムを含む*Staphylococcus epidermidis* RP62A、例示的なIII-B型CRISPR-Casシステムを含むPy

rococcus furiosus DSM 3638、例示的なIII-C型CRISPR-Casシステムを含む*Methanothermobacter thermautotrophicus* str. Delta H、および、例示的なIII-D型CRISPR-Casシステムを含む*Roseiflexis* sp. Rs-1を含む。

【0147】

一部の実施態様では、本発明で有用なIV型ポリペプチドは、Cs f 4 (d i n G)、Cs f 1、Cas 7 (Cs f 2) および／またはCas 5 (c s f 3) のアミノ酸配列と、少なくとも70%の同一性を含む(例えば、約70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%など)。IV型CRISPR-Casシステムは、当技術分野でよく知られていて、例えば、*Acidithiobacillus ferrooxidans* ATCC 23270は、例示的なIV型CRISPR-Casシステムを含む。

【0148】

一部の実施態様では、本発明で有用なV型ポリペプチドは、Cp f 1、Cas 1、Cas 2、またはCas 4のアミノ酸配列と、少なくとも70%の同一性を含む(例えば、約70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%など)。V型CRISPR-Casシステムは、当技術分野でよく知られていて、例えば、例示的なV型CRISPR-Casシステムを含む*Francisella* cf. *novicida* Fx1を含む。

【0149】

さらに本明細書において提供されるのは、本発明の核酸構築物、核酸アレイ、核酸分子および／またはヌクレオチド配列を含む発現カセットおよびベクターであって、本開示の方法で使用することができる。

【0150】

さらなる態様では、本発明の核酸構築物、核酸アレイ、核酸分子、および／またはヌクレオチド配列は、宿主生物の細胞に導入することができる。本発明が有用である任意の細胞／宿主生物を使用することができる。例示的な宿主生物は、細菌、古細菌、藻類および菌類(例えば、酵母)を含む。

【0151】

ここで、本発明は、以下の実施例に関して説明される。これらの実施例は、本発明に対する特許請求の範囲を制限することを意図しないが、むしろ、特定の実施態様の例示であることを意図すると理解されるべきである。当業者が想到する例示された方法における任意の変型は、本発明の範囲内であることが意図される。

【実施例】

【0152】

実施例1. 細菌株

全ての細菌株を表1に挙げる。細菌培養物を、25%グリセロール(v o l / v o l)を含む適切な増殖培地中で凍結保存して、-80℃で保管した。*S. thermophilus*を、37℃にて静的好気的条件下で1%牛肉エキス(w t / v o l)および1.9% (w t / v o l) β-グリセロールリン酸(S i g m a) ブロスで補充されたE l l i k e r 培地(D i f c o)中で、または、1.5% (w t / v o l) 寒天(D i f c o)を含む固形培地上で繁殖させて37℃にて48時間嫌氣的にインキュベートした。2 μ g / m L の濃度のエリスロマイシン(E m) および5 μ g / m L のクロラムフェニコール(C m) (S i g m a) を、適切な場合、*S. thermophilus* におけるプラスミド選択に用いた。*E. coli* EC1000を、好氣的に37℃にてL u r i a - B e r t a n i (D i f c o) ブロス中、または、1.5% 寒天で補充された脳心臓浸出物(

BHI) (Difco) 固形培地上で繁殖させた。E. coli の抗生物質選択を、適切な場合、 $40 \mu\text{g}/\text{mL}$ カナマイシン (Kn) および $150 \mu\text{g}/\text{mL}$ の Em を用いて、組み換え E. coli について維持した。β-ガラクトシダーゼ活性に関する S. thermophilus 誘導体のスクリーニングを、合成の Elliker 培地を 1% ラクトース、1.5% 寒天、および 0.04% プロモークレゾールパープル (pH 指示薬として) で補充することによって、定性的に評価した。

【0153】

実施例 2. DNA 単離およびクローニング

全てのキット、酵素、および試薬は、製造元の説明書に従って用いた。DNA 精製およびクローニングを、以前に記載されたとおりに行なった (41)。手短には、S. thermophilus からゲノム DNA の精製は、ZR Fungal/Bacterial MiniPrep kit (Zymo) を用いた。プラスミド DNA を、E. coli から、Qiagen Spin miniprep kit (Qiagen) を用いて単離した。高い忠実度の、DNA の PCR 増幅を、PFU HS II DNA ポリメラーゼ (Stratagene) を用いて行なった。日常的な PCR は、Choice Taq Blue ポリメラーゼ (Denville) を用いて行なった。PCR 増幅に関するプライマーは、Integrated DNA Technologies (Coralville, IA) から購入した。アガロースゲルからの DNA 抽出は、Zymo Clean DNA ゲル回収キット (Zymo) を用いて行なった。制限エンドヌクレアーゼは、Roche Molecular Biochemicals から取得した。ライゲーションは、New England Biolabs quick T4 リガーゼを用いて行なった。シーケンシングは、Davis Sequencing Inc. (Davis, CA) によって行なった。凍結保存された塩化ルビジウムコンピテント E. coli 細胞は、前述したとおりに調製した (41)。それぞれ、連続して (1) St h C R I S P R 1 または St h C R I S P R 3 に特異的なネイティブのリーダー配列 (2) C R I S P R 1 または C R I S P R 3 に特異的なネイティブのリピーター (3) lac Z の 5' 末端に特異的なスペーサー配列 (4) 別のネイティブリピーター (図 1) からなる、lac Z 標的化アレイを有するプラスミドを構築した。それぞれのプラスミドを作るために、上記に挙げた配列の特徴を伸長プライマーとしてオーダーして (表 2)、オーバーラップエクステンション PCR によるスプライシングを用いて組み合わせて (42)、pORI 28 にクローニングした (図 2)。

【0154】

実施例 3. C R I S P R スペーサーの選択および設計

染色体切断のプログラム可能な特異性は、標的対立遺伝子に特有の所望のスペーサー配列の選択で定まる (hinges upon)。特異性は、標的配列内のプロトスペーサーに近位でなければならない必須 PAM、短い保存配列によってさらに作り上げられる (compounded by) (21, 43)。したがって、スペーサーの選択および設計に関する厳格な基準は、コンセンサス PAM 配列の位置および外来ゲノムの遺伝子座に対する偶発的な配列同一性であった。推定のプロトスペーサーを、lac Z のセンス鎖およびアンチセンス鎖内の全ての推定の PAM 配列の位置を最初に規定することにより構築した。3,081 nt 遺伝子内には、22 個の C R I S P R 1 (AGAAW) および 39 個の C R I S P R 3 (GGNG) PAM 部位が存在し、それらの生物情報学的に得られたコンセンサス配列と同一であった (21)。潜在的スペーサーを同定した後に、完全なプロトスペーサー、シード、および PAM 配列を、非特異的な遺伝子座のさらなる標的化を防ぐために、S. thermophilus LMD-9 のゲノムに対する BLAST 分析に供した。C R I S P R 1 および C R I S P R 3 に関するスペーサーは、配列および対応する PAM 部位が全く異なったが、lac Z の 5' 末端を標的化するように設計されて、6 nt 離れて存在する予測された切断部位をもたらした。したがって、それぞれのプラスミド上のリーダー配列、リピーター、およびスペーサーは、それぞれ、C R I S P R 1 または C R I S P R 3 に特有のオルトゴナル特性を示した。標的遺伝子座依存性の突然

変異を評価するために、さらなるCRISPR3プラスミドを、 β -ガラクトシダーゼ活性に必須の金属カチオン結合残基に対するスパーサーを用いて作製した。非自己スパーサーを含むCRISPR1アレイプラスミドをコントロールとして用いて、自己標的化の致死性を定量化した。

【0155】

実施例4．形質転換

プラスミドを、温度感受性ヘルパープラスミド、pTRK669を含むコンピテントS．thermophilusに、以前に記載された方法(44)に従ってエレクトロポレーションした。手短には、S．thermophilusの一晩培養物を、1%牛肉エキス、1.9% β -グリセロリン酸およびCm選択で補充された50mLのElliker培地中に、1%(vol/vol)で接種した。培養物がOD_{600nm}=0.3を達成したときに、エレクトロポレーション効率を促進するために(45)、ペニシリンGを添加して、10 μ g/mLの終濃度を得た。遠心分離によって細胞を回収して、10mLの冷却エレクトロポレーションバッファー(1Mスクロースおよび3.5mM MgCl₂)中で3回洗浄した。細胞をエレクトロポレーションバッファー中に100倍濃縮して、40 μ Lの懸濁液を、0.1mmエレクトロポレーションキュベット中に等分した。それぞれの懸濁液を、700ngのプラスミドと組み合わせた。エレクトロポレーション条件を、2,500V、25 μ Fd静電容量、および、200 Ohms抵抗に設定した。時定数を記録して、4.4~4.6msの範囲だった。懸濁液を950 μ Lの回復培地と直ちに組み合わせて、37℃にて8時間インキュベートした。細胞懸濁液を選択培地上にプレーティングして、エレクトロポレーションキュベットを培地で洗浄して細胞の回復を確実にした。

【0156】

実施例5． β -ガラクトシダーゼ表現型の確認

CRISPR1およびCRISPR3の両方から生産された形質転換体を、はじめに、固形炭水化物源として1%ラクトースで補充された半合成のElliker培地上にコロニーを再ストリークすることによって、 β -ガラクトシダーゼ欠失表現型についてスクリーニングした。 β -ガラクトシダーゼ活性の喪失を、Millerアッセイを行なうことによって確認した(オ-ニトロフェニル- β -D-ガラクトシド(ONPG)(46))。手短には、培養物を、5mLの培地中で対数期後期(OD_{600nm}=1.2)まで繁殖させて、遠心分離(4,000xg、10分間)によって回収した。細胞を洗浄して、0.5mLのリン酸緩衝生理食塩水(Gibco-Invitrogen)中に再懸濁した。それぞれの懸濁液を、100 μ Lの0.1mmガラスビーズ(Biospec)と合わせて、それから、Mini-Beadbeater(Biospec)中で、5回の60秒サイクルの均質化に供した。それからサンプルを遠心分離して(15,000xg、5分間)、デブリおよび無傷の細胞を除去した。細胞溶解物(10 μ Lアリコート)を、600 μ Lの基質溶液(60mM Na₂HPO₄；40mM NaH₂PO₄；1mg/mL ONPG；2.7 μ L/mL β -メルカプトエタノール)と組み合わせて、室温にて5分間インキュベートして、その時点で、700 μ Lのストップ溶液を添加した(1M NaCO₃)。420nmにおける吸光度を記録して、 β -ガラクトシダーゼの活性をMiller単位として報告して、前述のとおり計算した(46)。

【0157】

実施例6．増殖および活性評価

培養物を、ラクトースを含まない半合成のElliker培地中で、12世代継代培養することによって、増殖アッセイのために事前に整えた。新鮮な培地に、一晩培養物を1%(vol/vol)で接種して、37℃で静的にインキュベートした。培養が定常期に達するまで、OD₆₀₀を1時間ごとにモニターした。ミルクの酸性化を、10⁸cfu/mLのレベルまでの一晩培養物をスキムミルクに接種して42℃でインキュベートすることによって評価した。続いて、pHをMettler Toledo Seven Easy pHメーターおよびAccumetプローブを用いてモニターした。スキムミルク

クは、NC SU乳製品工場から取得して、80℃で30分間低温殺菌した。

【0158】

実施例7. 消費されてよいゲノム領域の同定

CRISPR-Cas 標的化に関する可動および消費されてよい遺伝子座のインシリコ予測を、i) ISエレメントの位置、方位、およびヌクレオチド同一性、およびii) 必須ORFの位置に基づいて行なった。Bacillus subtilisでは、271個の必須ORFが、ゲノムワイドの遺伝子ノックアウトの致死性を判定することによって同定された(33)。S. thermophilusゲノムを、アミノ酸配列に関するデフォルトのスコアリングマトリックスの下でBLASTp検索ツールを用いて、B. subtilis由来の各必須遺伝子に対する相同体に関して検索要求した。約239個の必須ORFに対する相同体がS. thermophilusにおいて同定されて、その全てが、染色体上にコードされていた(表4)。DNA複製/ホメオスタシス、翻訳機構、およびコア代謝経路を含む、保存されている細胞プロセスに関与するタンパク質は、容易に同定された。シクロム生合成/呼吸に対応する相同体は、発酵性細菌の代謝プロファイルに従って観察されなかった。それぞれの推定の必須ORFを、SnapGeneソフトウェアを用いて参照ゲノムにマッピングして、S. thermophilus LMD-9におけるそれらの位置および分配の可視化を容易にした(図3A)。

【0159】

S. thermophilusゲノム内のISエレメントを、Geneious(登録商標)ソフトウェアを用いてトランスポゾンコード配列を並べることによってグループ化した(図4)。ファミリー指定を、ISエレメントデータベース内のBLAST分析に従って決定した(www-is.biotoul.fr/)。組み換えによって仲介される染色体セグメントの切断に関する潜在性を予測するために、関連するISエレメントの相対的位置をS. thermophilusゲノムに対してマッピングした(図3A)。ISエレメントのIS1193およびSth6ファミリーは、ゲノムにおいて最も頻出であり、Streptococcus pneumoniaeおよびStreptococcus mutansにおいて共通して見られる(34)。IS1193エレメントの普及性にかかわらず、これらの遺伝子座の多くは、いくらかの多型および縮退を示す小さなフラグメントであることが示されたが、高レベルの配列同一性で存在するいくつかのコピーも存在した(図5A)。対照的に、Sth6ファミリーは、かなりの多型および高い縮退を示し、一部のコピーは、著しい内部欠失を保有した(図5B)。IS1167およびIS1191エレメントは頻度がより低かったが、ゲノムにおいて同定されたコピー間でほぼ完全な忠実性を示した(図5Cおよび図5D)。S. thermophilusのIS1167およびIS1191エレメントの長さおよび配列の保存、および、ミルク適応遺伝子に対するそれらの相対的近接に基づいて、我々は、これらの保存された/高い忠実性トランスポゾンは、ゲノムにおいて最近になって獲得されたと仮定する。

【0160】

予測された必須ORFおよびISエレメントの位置を組み合わせることによって、高い忠実性のISエレメントが隣接する消費されてよいアイランドを同定した(図3A)(表3)。第一のアイランドは、未知の特異性を有する推定のATP-依存性オリゴヌクレオチド輸送システムをコードする、S. thermophilus LMD-9に特有のオペロンを含んだ(図3B)(35)。第二は、S. thermophilusの迅速な酸性化の表現型に寄与する細胞-エンベローププロテイナーゼPrtSを保有する(図3C)(36)。とりわけ、prtSは、S. thermophilusゲノムにおいてユビキタスでないが、prtSをコードするゲノムアイランドは、自然な能力を用いて株間で移行可能であることが示されている(36)。第三のアイランドは、推定のATP-依存性銅流出タンパク質を含み、全てのシーケンスされたS. thermophilus株に存在する(図3D)。第四のアイランドは、長さが102kbpであり、lacオペロンを含む102個の予測されたORFを有する遺伝子内容という観点から、はるかに最も大きい(図3E)。このアイランドは、S. thermophilusの全ての株において

見られるが、特定の遺伝子内容および長さは、株間で変化する。両方の内因性 I I 型システムを用いて大きなゲノムアイランドを標的化する成果を判定するために、リピートスペーサーアレイを *lacZ* コード配列に関して生産して (図 3 E)、*pORI28* にクローニングした (図 2)。第四のアイランドを、そのサイズ、*S. thermophilus* 株における遍在性、および、 β -ガラクトシダーゼ陰性の表現型に基づいて *lacZ* 突然変異に関してスクリーニングする能力に起因して、CRISPR-Cas 標的化のために選択した。

【0161】

実施例 8. *lacZ* の CRISPR-Cas 標的化は、大きな欠失イベントを選択する

I I 型システムでは、Cas 9 は DNA を調べて、*tracrRNA::crRNA* デュプレックスの形成を介して生じる標的で、Cas 9 の活性化によって、可逆的に PAM 配列に結合して (37)、最終的に dsDNA 切断をもたらす (図 6 A および 図 6 B) (25)。図 14 A および 図 14 B は、標的化された死滅のために内因性 CRISPR システムを利用する一般的なアプローチを示す概略である。具体的には、これらの図 14 A および 図 14 B は、標的化された死滅のために *Streptococcus thermophilus* における内因性 I I 型システムを利用するアプローチを示す。したがって、*S. thermophilus* では、プログラムされた細胞死は、発現がネイティブまたは合成のプロモーターによって作動される、ネイティブのリピートが隣接したゲノム標的化スペーサー配列を設計することにより、CRISPR-*StH1* (A) または CRISPR-*StH3* (B) I I 型システムを用いて達成された。転写されたリピートスペーサーアレイは、宿主がコードする *RNAase III* および Cas 9 を介して処理されて、成熟 *crRNA* を生じさせて、Cas 9 をゲノムに動員して、細胞死をもたらす二本鎖 DNA 切断を引き起こす。

【0162】

CRISPR-Cas システムによって染色体自己標的化を誘発するプラスミドを用いた形質転換は、非自己標的化プラスミドと比較した、残存する形質転換体の相対的な縮小によって測定されるように、細胞傷害性を示した (15、29)。*S. thermophilus* における *lacZ* 遺伝子の標的化は、回収された形質転換体の約 2.5-log 縮小をもたらした (図 6 C)、形質転換効率の限界に近づいた。二本鎖 DNA 切断 (DSB) は、生物の生存に対して著しい脅威を継続する。対応する修復経路は、平滑末端 DNA を修復するために末端切除を要求することが多い。Cas 9 が影響するヌクレオチド鎖切断は、野生型への標的遺伝子座の復元が、その後の CRISPR 標的化を回避しないので、DSB に起因する突然変異が生じる圧力をさらに悪化させる。乳酸菌内のスペーサー起源の特定は、22% のスペーサーが自己に対して相補性を示すこと、および、対応するゲノム遺伝子座が変更されて、自然発生の自己標的化イベントの生存を促進している可能性があることを明らかにした (28)。

【0163】

標的遺伝子座が、Cas 9 によって誘導された切断に応答して突然変異されたかどうか判定するために、形質転換体を、 β -ガラクトシダーゼ活性の喪失について最初にスクリーニングした。活性が欠失したクローンは、*lacZ* 遺伝子座において遺伝子型決定された。古典的または代替の末端ジャンクションに起因する突然変異も、いかなる自然発生的な一塩基多型も、シーケンスされたどのクローンにおいても見られなかった。一塩基多型の不存在は、点突然変異の低発生率によって作り上げられる (compounded by) 低い形質転換効率、および、非相同末端結合の不存在と相関する Ku および *Ligase IV* 相同体の不存在に起因し得る (38)。PCR スクリーニングは、野生型 *lacZ* が存在しないことを示したが、PCR アンプリコンはネイティブの *lacZ* 遺伝子座に対応せず；むしろ、別のゲノム遺伝子座における IS エlement 隣接配列が増幅された。 β -ガラクトシダーゼ活性の喪失の要因である遺伝子型を調べるために、Single Molecule Real Time シーケンシングを、2 つのクローンについて行なった (1 つは、*lacZ* の 5' 末端を標的化する CRISPR3 から生産され、1 つは、

β -ガラクトシダーゼ触媒作用に必要なイオン結合ポケットをコードする配列を標的化するCRISPR3から生産される) (図7Aおよび図7B)。このシーケンシング戦略を、その長い読長 (read length) について用いて、ゲノム内のISエレメントの多い数に起因する、正確な遺伝子座に対して確実にマッピングする読長の困難性を回避した (35)。読長は、Geneiousソフトウェアを用いて参照ゲノム配列に対してマッピングされて、そして、lacZオープンリーディングフレームをコードする大きなセグメント (約102 kbp) の不存在を明らかにした (図7Aおよび図7B)。シーケンスされた両方の株は、大きな欠失の境界 (boundaries) の再現性が確認されて、欠失は、lacZスパー配列または標的化に用いられるCRISPR-Casシステムとは独立に起きることが示された。しかしながら、シーケンシングデータは、欠失の正確なジャンクションを確実に示さなかった。

10

【0164】

欠失した102 kbpセグメントは、*S. thermophilus* の1.86 Mbpゲノムのおよそ5.5%を構成する。領域は、ABCトランスポーター、2成分制御システム、バクテリオシン合成遺伝子、ファージ関連タンパク質、ラクトース異化作用遺伝子、および、注釈付けられた (annotated) 機能を持たないいくつかのクリプティック (cryptic) 遺伝子をコードする、102個の推定ORF (STER_1278-1379) を含んだ (35)。増殖の表現型に対する、欠失の効果を、経時的にOD600 nmを測定することによって、ブロス培養において評価した (図7C)。欠失クローンは、より長い遅滞期 (lag phase)、より低い最終OD ($p < 0.01$) を有することが分かり、野生型の62分と比較して、対数期の間に、平均103分という、有意により長い世代時間を示した ($p < 0.001$)。欠失誘導体は、世代あたり、複製するために、5.5%少ないゲノム (5.5% less of the genome to replicate per generation) を有し、除去されたORFの転写または翻訳において資源を消費しないが、野生型と比較して適合性の明らかな増大は観察されなかった。 β -ガラクトシダーゼ活性は、乳酸菌の産業用途に関するホールマーク特性であり、酸性化を通じた食品システムの保存に必須である。したがって、lacZ欠失*S. thermophilus*株がミルクを酸化する能力を、pHをモニタリングすることによって評価した (図7D)。予想通りに、欠失株は、野生型において見られる迅速な酸性化表現型と明らかに対照的に、実験全体においてミルクを酸化することができなかった。

20

30

【0165】

実施例9. ゲノムの欠失は、相同ISエレメント間の組み換えを通して生じる

欠失のメカニズムを調べるために、セグメントに隣接するヌクレオチド配列を決定した。ジャンクションに見られる唯一の相同配列は、全体的に727 bpにわたって91%のヌクレオチド配列同一性を示す2つのランケットされたIS1193挿入配列であった。したがって、2つのISエレメントに隣接するプライマーペアを、欠失を示している生き残ったクローンのゲノムDNAを増幅するように設計した。欠失株のそれぞれは、予測されたサイズ (約1.2 kb) の強いバンドを示し、大きなゲノムの欠失イベントが確認された (図8A)。興味深いことに、染色体欠失に対応するわずかなアンプリコンが野生型において観察され、この領域は、野生型集団内で低い割合でゲノムから自然に切除され得ることを示した。ジャンクションアンプリコンのシーケンスを、CRISPR3による染色体自己標的化によって生産された20クローンについて行なった。遺伝子座の遺伝子型決定は、それぞれのクローン内に1つのキメラISエレメントの存在を明らかにし、したがって、それぞれのクローンに関してキメラ内での上流エレメントから下流配列への移行を明らかにした (図8B)。観察された欠失のサイズは、101, 865~102, 146 bpの範囲だった。移行の正確な遺伝子座は可変であったが、クローン内ではランダムではなく、欠失メカニズムの潜在的傾向を暗示した。*S. thermophilus* は、RecA (STER_0077)、デュアルのATP-依存性DNAエキソヌクレアーゼとして機能するAddAB相同体 (STER_1681およびSTER_1682)、

40

50

および、RecDファミリーのヘリカーゼ（STER_1742）としてコードされる典型的な組み換え機構を保有した。隣接するISエレメント間の高いヌクレオチド同一性、および、*S. thermophilus*が部位特異的な組み換えを行なう能力（4）は、RecAが仲介する組み換えが、ゲノムのセグメントの切断を仲介する潜在性を確認する（図8C）。

【0166】

次に、CRISPR-Cas標的化を、同一の遺伝的構造を有するそれぞれの遺伝子座に関する欠失の単離を促進する能力について評価した。この目的のために、3つのCRISPR3リピートスペーサーアレイ（第一の遺伝子座内のオリゴヌクレオチドトランスポーター、第二の遺伝子座由来のprtS、および、第三の遺伝子座由来のATPアーゼ銅流出遺伝子を標的化するもの）を生産して、pORI28にクローニングした（図5）。欠失についてスクリーニングするために、それぞれの遺伝子座でISエレメントに隣接するプライマーを、それぞれの欠失ジャンクションが増幅されるように設計した（図8D）。また、野生型遺伝子座の不存在も、それぞれのゲノムアイランドに関する内部プライマーを設計することによって、それぞれの場合において確認された（図8E）。標的化プラスミドを用いた形質転換の後に、それぞれの遺伝子座における欠失を単離して、野生型の不存在を確認した。欠失ジャンクションアンプリコンのシーケンシングは、単一のカメラISエレメントフットプリントの残存を確認し、それぞれの遺伝子座における欠失に関する一般的なメカニズムを示した。興味深いことに、ISエレメントに隣接するプライマーは、野生型gDNAからも増幅されて、さらに驚くべきことに、それぞれの遺伝子座で自然に生じた集団異質性は、自然発生的なゲノムの欠失に起因した。これらの結果は、配列特異的なCas9切断が、標的化に必要なプロトスペーサーおよびPAMの組み合わせを欠く変異を選択することを暗示する。したがって、自然発生的なゲノムの欠失は、これらのゲノムアイランドを既に失っている微生物変異に関する強い選択として、CRISPR-Cas標的化を用いて単離することができる。

【0167】

実施例10. 集団スクリーニング

本研究では、*S. thermophilus*に保有されるネイティブのIIA型システムを、大きなゲノムアイランドの自然発生的な欠失を定義するために再利用した。*S. thermophilus*内の4つのアイランドを独立に標的化することによって、合わせてゲノムの合計7%を喪失している安定した突然変異体が生産された。欠失ジャンクションの特性化は、IS-依存性の組み換えメカニズムが集団異質性に寄与することを示唆し、8~102kbpの範囲の欠失イベントを明らかにした。カメラISエレメントの正確なマッピングは、天然の組み換えイベントが、おそらく、*S. thermophilus*における大きな染色体欠失の要因であり、標的化されたゲノム編集に潜在的に活用され得ることを示した。

【0168】

我々の結果は、野生型クローンが集団から除去され、一方で、CRISPR-Casによって標的化される機構のない突然変異体は生き残ったことを示す。したがって、適応性のアイランドが同定および確認されて、内因性Cas9による正確な標的化を、混合された集団における大きな欠失変異株を単離するために活用することができることを示した。

【0169】

細菌のゲノム進化は、遺伝子水平伝播、内因性の突然変異、および、ゲノム再構築を通して生じる。*S. thermophilus*株のゲノムシーケンシングおよび比較分析は、有意なゲノム減衰を明らかにしているが、栄養豊富な食料環境に対する適応がニッチ特異的な遺伝子獲得を通して生じたことも示す（18；35）。*S. thermophilus*ゲノム内の、組込みおよびコンジュゲートエレメント、プロファージ、およびISエレメントを含む、MGEの存在は、乳製品環境に対する迅速な進化を示す（38~39）。可動遺伝機構は、遺伝子獲得、および反対に、非必須配列の不活性化または喪失を促進する。その結果として、MGEは、適合性を増大または生態学的生活様式を変更する手段

として、ゲノムの可塑性を与える。我々の結果は、これらのエレメントのCRISPR-Cas標的化は染色体再配列およびホメオスタシスに影響し得ることを強く示す。これは、必須の機構を標的化する実験とは対照的であり、不活化されたCRISPR-Cas機構を有する変異体の選択をもたらす(Jiang 2013)。必須ORFの突然変異は、CRISPR-Cas標的化の回避のための生存可能な道ではなく、したがって、不活化されたCRISPR-Casシステムを有するクローンのみが残る。高頻度可変性および消費されてよいと予測される遺伝因子を標的化して設計することによって、変更された遺伝子座を有する変異体が生存可能であることが示されて、活性のCRISPR-Casシステムは、自己標的化イベントのあいだ維持された。

【0170】

細菌ゲノム内でのISエレメントのほぼユビキタスな分配にもかかわらず、それらは、機能におけるそれらの多様性および可塑性に主に起因して、不可解な遺伝的物質を維持する(34)。我々の結果は、関連するISエレメント間の組み換えを、それらの位置、方位、および配列保存を分析することによって予測することが可能であることを示唆する(図4および図5)。CRISPR-Cas標的化を用いて、それぞれの予測された遺伝子座における集団異質性を実験的に確認することができ、および同時に、低い発生率の突然変異体の回収を増大させることができる。乳酸菌、および特に*S. thermophilus*におけるMGEの高い普及性は、ゲノム進化を通じたこれらの過剰に適合された細菌の種分化におけるそれらの役割に従う(39~40)。さらに、CRISPR-Cas標的化を用いたゲノム欠失突然変異体の回収は、機能未知の遺伝子の表現型特性化を促進することができる。*lac*オペロンをコードする102 kbアイランドの欠失を示す突然変異体は、野生型と比較して有意に増大した世代時間を有し、より低い最終ODを達成した。その中に102個の予測されたORFがあるので、さらなる表現型が影響を受けて、遺伝子の多くは注釈付けられた機能を有さないと考えられる。*prtS*のようなニッチ特異的な遺伝子の産業的関連性を考慮すると、この方法は、アイランドによってコードされる遺伝子がミルクにおいて増殖するための適合にどのように寄与するかについて、直接的な評価を可能にする。さらに、それらはおそらく分離したアイランドとして獲得されたのだろうと考えられるので、それは、これらの水平的に獲得された形質の天然のゲノム的および生態学的コンテキスト内にある。これらの結果は、転位、DNA修復メカニズム、およびゲノム可塑性を調べるための細菌における自己標的化CRISPR-Cas9システムの適用に関する新たな道を確立する。

【0171】

CRISPR-Casシステムは、一般に、遺伝因子との干渉を通して遺伝的多様性を制限するが、獲得されたMGEは、宿主細菌に対して適応性の利点を与えることもできる。したがって、ゲノム的に組み込まれたMGEを維持する利点は、CRISPR-Cas標的化にかかわらず、ゲノムホメオスタシスの重要なドライバーである。まとめると、我々の結果は、インシリコでのGEIの予測を、CRISPR-Cas標的化と結び付けて、大きなゲノム欠失を示しているクローンを単離することができることを確立した。それぞれの欠失ジャンクションにおけるキメラ挿入配列フットプリントは、4つのアイランド全てに関する欠失の共通したメカニズムを示した。ゲノム遺伝子座に対する同一性を示している自己標的化スパーサーの高い普及性は、ゲノム変化の実験的論証と組み合わせ、CRISPR-Cas自己標的化は、細菌のゲノム進化に著しく寄与し得ることを示唆する(28; 30)。まとめると、CRISPR-Casによって誘導される大きな欠失に関する研究は、注釈を欠いている遺伝子クラスターの必須性および機能性を評価するための迅速かつ効果的な手段としてこのアプローチを具体化し、および、転位性エレメントを通して生じる染色体欠失に基づいて、最小細菌ゲノムを定義する。

【0172】

図9は、*Streptococcus thermophilus* LMD-9のゲノムの標的化を介してII型CRISPR-Casシステムに基づく致死性を評価するための規定された遺伝子座を示す。この分析を行なうための方法は、当技術分野で知られてい

10

20

30

40

50

る。Selle and Barrangou PNAS. 112 (26) : 8076 – 8081 (2015) を参照。

【0173】

オルトゴナル I I 型システム (CRISPR1 および CRISPR3) の両方とも試験した; CRISPR1 は濃い灰色において標的化し、CRISPR3 は薄い灰色において標的化する。特異的な遺伝的特徴を選択して、(i) 遺伝子間領域 (INT)、(ii) 可動遺伝要素 (IS、GEI1–GEI3、PRO、lacZ、EPS)、(iii) 必須遺伝子 (dltA、LTA)、(iv) レプリコアのポール (ORI、TER)、および、DNA のフォワード、対、リバーズ鎖 (外側標的、対、内側標的) を試験した。

【0174】

図 10 は、図 9 で定義された領域を標的化することによって得られた、CRISPR に基づく致死性を示す。CFU の Log 減少を、非標的化プラスミドコントロール; pORI 28 の形質転換に関して計算した。致死性は、染色体位置、コード配列、または必須性にかかわらず、試験されたすべての標的に関して、2 ~ 3 log 減少の範囲だった。

【0175】

図 11 は、CRISPR が仲介するゲノムアイランド欠失株の転写プロファイルを示す。ゲノムアイランド 1 ~ 4 の CRISPR 標的化を生き残る細胞の回収および遺伝子型決定は、それぞれの実験において、標的化されたゲノムアイランドを欠いている安定的な非依存性突然変異体の同定をもたらした。続いて、細胞を繁殖させて、それらの全 RNA を単離してシーケンシングした。このアプローチを用いて、転写プロファイルを、参照ゲノムに対してシーケンシング解読をマッピングすることによって生産した。それぞれの場合において、予測されたゲノムアイランド遺伝子座に対するシーケンシング解読の不存在は、残りのゲノムの至る所に、コア遺伝子の発現に対して最小限の影響を有しながら、標的の遺伝的物質の喪失をさらに示唆した。

【0176】

さらに、RNA シーケンシングデータは、解読範囲および転写値比較を用いることにより、欠失の境界をサポートし、そして、同一のデータセットを用いて生産された比較トランスクリプトミクスを用いて、表現型の識別をさらにサポートする。具体的には、ハイスループット RNA シーケンシングを用いた、予測された欠失領域に存在する転写活性の喪失は、図 12 に示されるように確認される。図 12 は、ゲノムアイランド欠失株の、log₂ 形質転換された RNA シーケンシング解読範囲を示し、それぞれのゲノムアイランド株 (GEI1 ~ GEI4) に関して、予測されたゲノムアイランド遺伝子座に対するシーケンシング解読の不存在は、残りのゲノムの至る所に、コア遺伝子の発現に対して最小限の影響を有しながら、標的の遺伝的物質の喪失をさらに示唆した。

【0177】

図 13 は、欠失した遺伝子に関する転写活性の喪失をさらに確認する。ゲノムアイランド欠失株のそれぞれについて (GEI1 ~ GEI4)、標的アイランド (黒) のそれぞれの上にコードされる遺伝子の発現は、最小限であった。GEI1 内にコードされる遺伝子を上の左パネルに示し、GEI2 内にコードされる遺伝子を上の右パネルに示し、GEI3 内にコードされる遺伝子を下の左パネルに示し、GEI4 内にコードされる遺伝子を下の右パネルに示す。概して、ゲノムアイランド欠失 1 および 2 は、他の遺伝子の転写に最小限の影響を有し (灰色)、一方で、ゲノムアイランド 3 および 4 は、アイランド上にコードされない他の遺伝子の転写に影響を及ぼすようだった。

【0178】

加えて、RNA シーケンシングデータを用いて、欠失したアイランド (GEI4) 上にコードされない遺伝子、すなわち、染色体にまだ存在する他の遺伝子の転写レベルを比較して、ゲノムの欠失と関連する表現型を特定した。欠失株において示差的に転写される遺伝子は、失われた遺伝子によって細胞プロセスが影響を受けたこと、または、これらの遺伝子の活性の喪失に関する代償が存在することを示唆する。したがって、これらの遺伝子またはゲノム領域が関連する経路について推論することができる。表 5 は、欠失株 GEI

10

20

30

40

50

4において同定された示差的に発現される遺伝子のリストを提供する。示差的に発現されることが観察された遺伝子の多くは、芳香族アミノ酸およびプリン生合成を含む、*Streptococcus thermophilus*の生合成能力に関する。

【0179】

実施例11. II型CRISPR-Casシステムを用いた、*Lactobacillus casei*の標的化した死滅

*L. Casei*の例示的なCRISPR-Cas II型ガイドを図15に提供する。上のパネルは、予測されたガイドを提供し、一方で、下の左パネルは、例示的なデュアルガイド構造を示し、下の右パネルは、例示的な単一ガイド構造を示す。

【0180】

実施例12. II型CRISPR-Casシステムを用いた、*Lactobacillus gasserii*の標的化した死滅

*L. gasserii*の標的化した死滅のための例示的なII型ガイドを図16に提供する。上のパネルは、予測されたガイドを提供し、一方で、下の左パネルは、正確なデュアルガイドcrRNA：tracrRNAを提供し（RNAシーケンスによって確認した）、下の右パネルは、例示的な予測された単一ガイドを提供する。

【0181】

プラスミドを、それぞれ以下の異なる構築物を保有する*L. gasserii*に形質転換した：空のpTRK563ベクター、正確なプロトスペーサーであるが不正確なPAMを有する構築物、正確なPAMであるがアレイ内でないプロトスペーサー、および、PAMを有する正確なプロトスペーサー。結果を図19に示す。正確なプロトスペーサーおよび正確なPAMを有するプラスミドは、有意により多くの干渉の標的化および細胞死を示した。

【0182】

実施例13. II型CRISPR-Casシステムを用いた、*Lactobacillus pentosus*の標的化した死滅

*L. pentosus*の標的化した死滅のための例示的なII型ガイドを図17に提供する。上のパネルは、予測されたガイドを示す。下の左のパネルは、正確なデュアルガイドcrRNA：tracrRNAであり（RNAシーケンシングによって確認した）、下の右パネルは、例示的な予測された人工の単一ガイドである。

【0183】

プラスミドを*L. pentosus*に形質転換して、それぞれのプラスミドは、以下の異なる構築物を保有した：正確なプロトスペーサーであるが不正確なPAMを有する構築物、正確なPAMであるがアレイ内でないプロトスペーサーを有する構築物、空のpTRK563ベクター、および、正確なPAMを有する正確なプロトスペーサー。結果を図20に示す。正確なプロトスペーサーおよび正確なPAMを有するプラスミド（Lpelgttaat）は、有意により多くの干渉の標的化および細胞死を示した。

【0184】

実施例14. II型CRISPR-Casシステムを用いた、*Lactobacillus jensenii*の標的化した死滅

図18は、例示的なCRISPR-Cas II型ガイドを提供する。左のパネルは、RNAシーケンスによって確認された正確なデュアルガイドcrRNA：tracrRNAであり、右のパネルは、例示的な予測された人工の単一ガイドである。

【0185】

プラスミドを、それぞれ以下の異なる構築物を保有する*L. jensenii*に形質転換した：空のpTRK563ベクターを含む構築物、正確なプロトスペーサーであるが不正確なPAMを有する構築物、正確なPAMであるがアレイ内でないプロトスペーサーを有する構築物、および、正確なPAMと正確なプロトスペーサーを有する構築物。結果を図22に示す。正確なプロトスペーサーおよび正確なPAMを有するプラスミドは、実質的により多くの干渉の標的化および細胞死を示した。

10

20

30

40

50

【0186】

実施例15. I型CRISPR-Casシステムを用いた、Lactobacillus casei NCK 125の標的化した死滅

図21は、L. caseiの例示的なI型CRISPR-Casガイドを提供し、それは、L. casei NCK 125に見られるネイティブのI型リーダーおよびリピートの配列を含む。PAM5'-YAA-3'は、生物におけるネイティブのスペーサー配列を用いて予測された。人工的なアレイは、宿主ゲノム内の16srDNA遺伝子を標的化するスペーサーを含む。結果を図22に提供し、それは、空のベクターおよび2つの異なる人工的アレイの間の有意な縮小を示す：その1つは、16s遺伝子の+鎖を標的化する単一スペーサーを含み（1-2 alt）、他方は、+鎖を標的化するオリジナルのスペーサーを含むが、16s遺伝子の-鎖を標的化するさらなるスペーサーを含む（1、2-3）。

10

【0187】

本明細書に記載のCRISPR-Casシステムは、例えば、（i）いずれかの医学的介入（例えば、病原体、制限されないが、菌類、線虫類、原虫類（例えば、マラリア）、条虫類、球虫類（微孢子虫類）、吸虫類、舌形動物、鉤頭虫類、節足動物などを含む）の場合における病原体の標的化された縮小のため；（ii）消耗品（食品システム、動物、作物）の保護のため；（iii）産業的発酵プロセス（生の材料、処理装置、スターター培養）からの好ましくない生物の制御および／または除去のため、および、（iv）エコシステムおよび／または化学サイクルに影響を与える環境微生物コンソーシアムの制御のため、ならびに、治療のために、用いてよい。

20

【0188】

【表 1】

表 1. 細菌株およびプラスミド

株指定	説明	オリジナル の参照
<i>E. coli</i> EC1000	pORI プラスミドのための宿主、染色体 $repA^+$ (pWVO1)、 Km^R pTRK935のための宿主	47
<i>S. thermophilus</i> LMD-9 pTRK669を 有する <i>S.</i> <i>thermophilus</i> LMD-9	野生型 野生型、pTRK669によって与えられる $RepA^+$ および Cm^R	40 本研究
プラスミド		
pORI28	広範非複製ベクター、 Em^R	47
pTRK669	Ts-ヘルパープラスミド $repA^+$ 、 Cm^R	44
pCRISPR1::lacZ	pORI28::CRISPR1-リーダー-RSR- lacZ N末スペーサー	本研究
pCRISPR3::lacZ	pORI28::CRISPR1-リーダー-RSR- lacZ活性の部位スペーサー	本研究
pCRISPR3::ABC	pORI28::CRISPR1-リーダー-RSR- ABCスペーサー	本研究
pCRISPR3::Cu	pORI28::CRISPR1-リーダー-RSR- Cu流出スペーサー	本研究
pCRISPR3::prtS	pORI28::CRISPR1-リーダー-RSR- prtSスペーサー	本研究
pCRISPR3::lacZ	pORI28::CRISPR3-リーダー-RSR- lacZ N末スペーサー	本研究
pCRISPR1::非自 己	pORI28::CRISPR1-リーダー-RSR- 非自己スペーサー	本研究

【0189】

【表 2】

表 2. プライマー

プライマー名	配列	機能
C1_N-term_F	CAAGAACAGTTATTGATTITATAATCACTATGTGGGTATGAAATCTCAAAATCATTTGAGGTTTT TGTACTCTCAAGATTAAAGTAACTGTACACATTAAAGAGATTGTCTTAACTT	SOE-PCRのためのテンプレート
C3_N-term	AGCGGATAACAATTTCACGTTTTGGAACCATTCGAAACAACACAGCTCTAAAACTCAGAAAATCTT CAAGAGATTCAAAATACTGTTTTGGAACCATTCGAAACAACACAGCTCTAAAACTCTAGGATATC TTTTCTAC	SOE-PCRのためのテンプレート
C1_N-term_R	AGCGGATAACAATTTCACGTTGTACAGTTACTTAAATCTTGMAGTACAAAAACAGGGAGATGAAG TTAAGACAATCTCTTAATGT	SOE-PCRのためのテンプレート
C3_A-site	AGCGGATAACAATTTCACGTTTTGGAACCATTCGAAACAACACAGCTCTAAAAACGAAGCTTTGGTCT TCCAACAGCTTGTGTAGTTTTGGAACCATTCGAAACAACACAGCTCTAAAACTCTAGGATATC TTTTCTAC	SOE-PCRのためのテンプレート
C3_ABC	AGCGGATAACAATTTCACGTTTTGGAACCATTCGAAACAACACAGCTCTAAAAACGATAACACGAGATAA AACATCCAGCCACCGTTTTGGAACCATTCGAAACAACACAGCTCTAAAACTCTAGGATATCTTTCT tac	
C3_prtS	AGCGGATAACAATTTCACGTTTTGGAACCATTCGAAACAACACAGCTCTAAAAACGTTGTAGCTTTGAGG TCTGAGATACACGCTTTTGGGAACCATTCGAAACAACACAGCTCTAAAACTCTAGGATATCTTTCT tac	
C3_Cu	AGCGGATAACAATTTCACGTTTTGGAACCATTCGAAACAACACAGCTCTAAAAACGATTGCTCAATCA ATCGTTTCAGCTCTCAAGTTTTGGAACCATTCGAAACAACACAGCTCTAAAACTCTAGGATATCT TTCTAC	
C3LF	AGCAAGGATCCTGGTAATAAGTATAGTAGTCTTG	gDNAからSt h3リーダーを増幅する
C3LR	CTCGTAGGATATCTTTCTAC	gDNAからSt h3リーダーを増幅する
C1F	AGCAGGGATCCCAAGACAGTTATTGATTITATAATC	CRISPR1 SOE-PCRフォワード
C3F	AGCAGGGATCCTGGTAATAAGTATAGTAGTCTTG	CRISPR3 SOE-PCRフォワード
C1C3R	TGCTGGAGCTCGTGAATTTGTTATCCGCT	CRISPR1およびCRISPR3 SOE-PCR リバー
1193F	TTGAACCTAGGAACCTCATA	欠失ジャンクション増幅
1193R	CGTAAGGTTTGTGACTCAAG	欠失ジャンクション増幅
pORI28F	TTGGTTGATAATGAAGTGTCTG	pORI28のMCSのシーケンシング
pORI28R	TTGTTGTTTATGATTACAAAGTGA	pORI28のMCSのシーケンシング

【0190】

【表 3】

Table 3. 推定の消費されてよいゲノムアイランド

ゲノム アイ ランド	ORF領域	長さ k b p	GC含量 %	注目すべき遺伝子	IS ファミリー
1	STER_139-STER_148	7.81	37.1	オリゴペプチドトランスポーター	IS6
2	STER_840-STER_848	10.29	39.9	プロテイナーゼPrt S	ISSt h16/IS1 167
3	STER_881-STER_888	8.71	39.3	銅流出	IS1191
4	STER_1277-STER_1380	101.76	37.2	ラクトース異化、二成分制御、 バクテリオシン合成、ABC トランスポーター	IS1193

【0191】

【表 4】

表 4. *S. thermophilus* において同定された約 239 個の必須 ORF に対する相合体。

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	<i>Bacillus subtilis</i> 注釈
CSTER	1	101	1465	+	染色体複製開始タンパク質 DnaA	98%	4.00E-124	44%	染色体複製イニシエータータンパク質 DnaA
CSTER	2	1620	2756	+	DNAポリメラーゼ III サブユニット β	99%	1.00E-88	39%	DNAポリメラーゼ III サブユニット β dnan
CSTER	6	5818	6387	+	ペプチジル- tRNA ヒドロラーゼ	97%	1.00E-61	51%	
CSTER	9	10331	10702	+	細胞周期タンパク質	94%	1.00E-05	22%	細胞分裂タンパク質 DivIC
CSTER	12	12122	13387	+	tRNA (Ile) -ライシジン合成酵素、MesJ	92%	3.00E-40	32%	tRNA (Ile) -ライシジンシンターゼ tils
CSTER	13	13469	14011	+	ヒポキサンチン-グアニンホスホリボシルトランスフェラーゼ	99%	2.00E-82	63%	ヒポキサンチン-グアニンホスホリボシルトランスフェラーゼ hprt
CSTER	40	25595	26425	+	細胞形状決定タンパク質 MreC	94%	1.00E-22	28%	細胞形状決定タンパク質 MreC
CSTER	43	28617	29582	+	リボース-リン酸ピロホスホキナーゼ	97%	3.00E-145	65%	リボース-リン酸ピロホスホキナーゼ prs

10

20

30

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	<i>Bacillus subtilis</i> 注釈
CSTER	47	32842	33846	+	推定のグリセロール-3-ホスフェートアシルトランスフェラーゼ	94%	9.00E-124	54%	ホスフェートアシルトランスフェラーゼ PlsX
CSTER	48	33846	34091	+	アシル担体タンパク質	76%	1.00E-08	40%	アシル担体タンパク質
CSTER	65	53036	54727	-	アルギニル- tRNA 合成酵素	98%	6.00E-52	28%	
CSTER	95	72798	77192	+	DNAポリメラーゼ III P _o IC	99%	0.00E+00	51%	DNAポリメラーゼ III P _o IC 型
CSTER	105	82956	83723	+	30S リボソームタンパク質 S2	96%	1.00E-118	69%	
CSTER	106	83841	84881	+	翻訳伸長因子 Ts	97%	8.00E-72	44%	
CSTER	117	97363	98706	+	システイニル- tRNA 合成酵素	99%	1.00E-178	54%	
CSTER	127	104406	104852	+	50S リボソームタンパク質 L13	99%	9.00E-57	57%	
CSTER	128	104880	105272	+	30S リボソームタンパク質 S9	100%	3.00E-59	68%	
CSTER	193	159363	160967	+	CTP 合成酵素	99%	0.00E+00	68%	CTP シンターゼ pyr g
CSTER	199	166439	166897	-	保存された仮定のタンパク質	81%	2.00E-07	28%	タンパク質 NrdI

ゲノム部位	STER	開始	終了	方向		ク エ リー カバ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	208	176203	176472	+	30Sリボソームタンパク質S15	100%	6.00E-38	63%	
CSTER	217	183028	184185	+	ウンデカブレニルピロリン酸ホスファターゼ	98%	2.00E-68	38%	可能性の (probable) ウンデカブレニルホスフェートN-アセチルグルコサミニル1-ホスフェートトランスフェラーゼ tagO
CSTER	218	184346	185116	+	ABCトランスポーターATPアーゼ	95%	8.00E-130	71%	植物性 (Vegetative) タンパク質296 sufC
CSTER	219	185153	186415	+	仮定のタンパク質	97%	4.00E-111	42%	FeSクラスター集合タンパク質SufD
CSTER	220	186469	187701	+	推定のアミノトランスフェラーゼ (V型)	98%	0.00E+00	60%	システイン脱硫酵素SufS
CSTER	221	187688	188122	+	NifU fam. タンパク質	95%	3.00E-48	51%	亜鉛-依存性スルフルトランスフェラーゼSufU
CSTER	245	208199	208993	+	フォスファチジン酸シチジリルトランスフェラーゼ	98%	3.00E-67	41%	フォスファチジン酸シチジリルトランスフェラーゼ
CSTER	247	210343	212205	+	プロリル-tRNA合成酵素	98%	0.00E+00	50%	

10

ゲノム部位	STER	開始	終了	方向		ク エ リー カバ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	252	215735	216022	+	コ-シャペロニンGroES (HSP10)	97%	1.00E-18	43%	
CSTER	253	216071	217690	+	シャペロニンGroEL (HSP60ファミリー)	98%	0.00E+00	75%	
CSTER	261	224055	224204	+	50Sリボソームタンパク質L33	97%	9.00E-12	50%	
CSTER	262	224216	224392	+	ブレタンパク質トランスロカーゼサブユニットSecE	81%	1.40E+00	21%	タンパク質トランスロカーゼサブユニットSecE
CSTER	268	227616	230117	+	ロイシル-tRNA合成酵素	99%	0.00E+00	70%	
CSTER	273	232342	233796	+	ニコチン酸ホスホリボシルトランスフェラーゼ	97%	0.00E+00	63%	ニコチン酸ホスホリボシルトランスフェラーゼpncb
CSTER	274	233808	234629	+	NAD合成酵素	100%	9.00E-109	59%	NH(3)-依存性NAD(+)合成酵素nade
CSTER	286	248252	249580	+	UDP-N-アセチルムラミン酸-アラニンリガーゼ	99%	3.00E-170	55%	UDP-N-アセチルムラミン酸-アラニンリガーゼmurC
CSTER	302	260849	261664	+	グルタメートラセマーゼ	94%	7.00E-83	48%	グルタメートラセマーゼ1 racE
CSTER	307	264328	265041	+	分離および濃縮 タンパク質A	93%	5.00E-37	39%	分離および濃縮 タンパク質A

20

30

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	308	265034	265615	+	分離および濃縮 タンパク質 B	94%	9.00E-36	40%	分離および濃縮 タンパク質 B
CSTER	313	269788	270297	+	r RNA メチルトランスフェラーゼ	96%	3.00E-58	51%	r RNA メチルトランスフェラーゼ
CSTER	349	303983	305959	+	トランスケトララーゼ	98%	0.00E+00	58%	トランスケトララーゼ t k t
CSTER	357	312857	314032	+	染色体複製開始/メンブレン付着タンパク質 D n a B	63%	3.00E-02	18%	複製開始およびメンブレン付着タンパク質
CSTER	358	314036	314938	+	ブライモソームタンパク質 D n a I	99%	7.00E-58	35%	ブライモソームタンパク質 D n a I
CSTER	359	315044	316354	+	G T P-結合タンパク質 E n g A	100%	0.00E+00	67%	G T Pアーゼ D e r
CSTER	368	321054	322394	-	セリルー t RNA 合成酵素	100%	0.00E+00	63%	
CSTER	376	329673	330116	+	保存された仮定のタンパク質	89%	6.00E-33	43%	t RNA スレオニルカルバモイルアデノシン生成タンパク質 T s a E
CSTER	380	332545	333732	+	転写伸長因子	90%	2.00E-113	46%	転写終結/抗終結タンパク質 N u s A
CSTER	383	334363	337194	+	翻訳開始因子 I F-2	99%	0.00E+00	55%	

10

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	387	339864	341309	-	UDP-N-アセチルムラモイルアラニル-D-グルタメート-2, 6-ジアミノピメレートリガーゼ	92%	3.00E-52	29%	UDP-N-アセチルムラモイルアラニル-D-グルタメート-2, 6-ジアミノピメレートリガーゼ m u r E
CSTER	419	364962	365894	+	推定のマンガン-依存性無機ピロホスファターゼ	99%	3.00E-125	58%	マンガン-依存性無機ピロホスファターゼ p p a c
CSTER	430	375355	375579	+	アシル担体タンパク質	93%	5.00E-14	49%	アシル担体タンパク質
CSTER	432	376667	377593	+	アシル-担体-タンパク質 S-マロニルトランスフェラーゼ	95%	3.00E-86	47%	マロニル CoA-アシル担体タンパク質 トランスアシラーゼ
CSTER	433	377606	378340	+	3-ケトアシル- (アシル-担体-タンパク質) レダクターゼ	99%	6.00E-79	47%	3-オキソアシル- [アシル-担体-タンパク質] レダクターゼ F a b G
CSTER	434	378401	379633	+	3-オキソアシル- (アシル担体タンパク質) シンターゼ I I	99%	1.00E-133	48%	3-オキソアシル- [アシル-担体-タンパク質] シンターゼ 2
CSTER	435	379637	380125	+	アセチル-CoA カルボキシラーゼ ピオチンカルボキシル担体タンパク質サブユニット	98%	9.00E-26	37%	ピオチンカルボキシル担体タンパク質 o f アセチル-CoA カルボキシラーゼ

20

30

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	437	380667	382037	+	アセチル-CoAカルボキシラ ーゼピオチンカルボキシラーゼ サブユニット	98%	0.00E+00	60%	ピオチンカルボキシラーゼ 1
CSTER	438	382043	382909	+	アセチル-CoAカルボキシラ ーゼサブユニット β	94%	8.00E-98	51%	アセチル-CoAコエンザイムAカル ボキシラーゼカルボキシルトラ ンスフェラーゼサブユニット β
CSTER	439	382906	383676	+	アセチル-CoAカルボキシラ ーゼサブユニット α	76%	2.00E-81	53%	アセチル-CoAコエンザイムAカル ボキシラーゼカルボキシルトラ ンスフェラーゼサブユニット α
CSTER	442	385819	387417	+	保存された仮定のタンパク質	98%	0.00E+00	58%	リボスクレアーゼ Y ymda
CSTER	455	402156	402470	+	50Sリボソームタンパク質L 21	99%	2.00E-44	66%	
CSTER	456	402507	402797	+	50Sリボソームタンパク質L 27	95%	5.00E-47	79%	
CSTER	460	405038	405805	+	ジヒドロジピコリネートレダク ターゼ	98%	7.00E-95	53%	4-ヒドロキシテトラヒドロ ジピコリネートレダクターゼ d apb
CSTER	461	405802	407010	+	tRNA CCA-ピロホスホ リラーゼ	98%	6.00E-86	40%	CCA-付加酵素

10

20

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	475	422256	422813	+	リボソームリサイクリング因子	100%	2.00E-68	56%	
CSTER	485	430547	432550	+	メチオニル-tRNA合成酵素	98%	0.00E+00	58%	
CSTER	492	438748	439890	+	プロテアーゼ成熟タンパク質前 駆体	91%	6.00E-11	26%	フォルダーゼタンパク質 Prs A
CSTER	493	440232	442850	+	アラニル-tRNA合成酵素	99%	0.00E+00	52%	
CSTER	513	460453	463104	+	バリル-tRNA合成酵素	99%	0.00E+00	63%	
CSTER	523	469888	471168	+	細胞分裂タンパク質 FtsW	93%	4.00E-59	34%	推定の脂質 II フリッパーゼ F tsW
CSTER	524	471406	472602	+	伸長因子 Tu	99%	0.00E+00	76%	
CSTER	525	472851	473609	+	トリオースリン酸イソメラーゼ	97%	8.00E-109	62%	トリオースリン酸イソメラーゼ tpia
CSTER	526	473848	474477	+	チミジレートキナーゼ	96%	4.00E-73	55%	チミジレートキナーゼ tmk
CSTER	527	474486	475361	+	DNAポリメラーゼ III サブ ユニット δ'	96%	8.00E-41	33%	DNAポリメラーゼ III サブ ユニット δ'
CSTER	539	484827	485744	+	グリシル-tRNA合成酵素、 α サブユニット	97%	4.00E-169	74%	
CSTER	540	486028	488064	+	グリシル-tRNA合成酵素、 β サブユニット	97%	0.00E+00	45%	

30

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	567	512413	512916	+	50Sリボソームタンパク質L10	100%	7.00E-61	57%	
CSTER	568	512991	513359	+	50Sリボソームタンパク質L7/L12	100%	4.00E-38	67%	
CSTER	603	548926	550308	+	N-アセチルグルコサミン-1-ホスフェートウリジルトランスフェラーゼ	97%	3.00E-164	52%	二官能性タンパク質GlmU
CSTER	623	566278	566781	+	ジヒドロ葉酸レダクターゼ	94%	8.00E-34	38%	ジヒドロ葉酸レダクターゼdfra
CSTER	626	568210	568809	+	GTPアーゼEngB	96%	7.00E-92	63%	可能性のGTP-結合タンパク質EngB
CSTER	632	573613	574254	-	推定のグリセロール-3-ホスフェートアシルトランスフェラーゼPisY	90%	3.00E-40	46%	グリセロール-3-ホスフェートアシルトランスフェラーゼpIsy
CSTER	633	574392	576341	+	DNAトポイソメラーゼIVサブユニットB	97%	0.00E+00	71%	DNAトポイソメラーゼ4サブユニットB
CSTER	634	576969	579434	+	DNAトポイソメラーゼIVサブユニットA	95%	0.00E+00	54%	DNAトポイソメラーゼ4サブユニット
ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	660	598871	599725	+	メチレンテトラヒドロ葉酸デヒドロゲナーゼ/ メチルテトラヒドロ葉酸シクロヒドラーゼ	99%	2.00E-109	56%	二官能性タンパク質FolD
CSTER	668	605570	606469	+	GTPアーゼEra	98%	5.00E-143	64%	GTPアーゼEra
CSTER	670	607313	607927	+	デホスホ-COAキナーゼ	95%	2.00E-50	45%	デホスホ-CoAキナーゼcoa
	672	609192	609338		50Sリボソームタンパク質L33	97%	1.00E-16	52%	
CSTER	684	620012	621316	+	エノラーゼ	100%	0.00E+00	70%	エノラーゼeno
CSTER	731	664470	665690	-	推定のシトシン-C5特異的なDNAメチラーゼ	60%	5.00E-16	31%	可能性のBsuMI改変メチラーゼサブユニットYdiO
CSTER	733	666939	668429	-	リシル- tRNA合成酵素 (II型)	97%	0.00E+00	64%	
CSTER	761	694548	695582	+	DNAポリメラーゼIIIDサブユニット	84%	2.00E-37	33%	性質不明のタンパク質YqeN
CSTER	773	704704	706056	+	UDP-N-アセチルムラモイル-L-アラニル-D-グルタメート合成酵素	98%	8.00E-147	49%	UDP-N-アセチルムラモイルアラニン-D-グルタメートリガーゼmurD

10

20

30

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	774	706060	707130	+	N-アセチルグルコサミニルトランスフェラーゼ	95%	3.00E-30	27%	UDP-N-アセチルグルコサミン-N-アセチルムラミル(ペンタペプチド)ピロホスホリル-ウンデカプレノールN-アセチルグルコサミニルトランスフェラーゼmurG
CSTER	775	707140	708264	+	細胞分裂タンパク質ftsQ	79%	4.00E-15	26%	細胞分裂タンパク質DivIB
CSTER	776	708388	709764	+	細胞分裂タンパク質ftsA	84%	4.00E-103	44%	細胞分裂タンパク質ftsA
CSTER	777	709793	711115	+	細胞分裂タンパク質ftsZ	92%	2.00E-123	53%	細胞分裂タンパク質ftsZ
CSTER	783	714803	717592	+	イソロイシル-tRNA合成酵素	99%	0.00E+00	58%	
CSTER	787	720193	720459	-	50Sリボソームタンパク質L31	100%	1.00E-22	50%	
CSTER	793	725315	726394	+	ペプチド鎖放出因子1	99%	6.00E-158	59%	
CSTER	796	727911	729161	+	セリンヒドロキシメチルトランスフェラーゼ	97%	2.00E-171	60%	セリンヒドロキシメチルトランスフェラーゼglya
CSTER	833	760678	762396	-	推定のホスホグルコムターゼ	98%	0.00E+00	47%	ホスホグルコムターゼpgm
CSTER	850	783500	783736	-	30Sリボソームタンパク質S20	94%	1.00E-13	46%	

10

20

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	864	797961	799307	+	アスパラギニル-tRNA合成酵素	99%	4.00E-180	56%	
CSTER	903	833853	835661	+	D-フルクトース-6-ホスフェートアミドトランスフェラーゼ	100%	0.00E+00	59%	グルタミン-フルクトース-6-ホスフェートアミノトランスフェラーゼ[異性化]glmS
CSTER	915	845349	846911	+	シグナル認識粒子タンパク質	95%	2.00E-178	57%	シグナル認識粒子タンパク質
CSTER	919	848991	849845	+	リボソーム生合成GTPアーゼ	98%	8.00E-101	50%	リボソーム生合成GTPアーゼrbga
CSTER	923	852605	854749	+	DNAトポイソメラーゼI	98%	0.00E+00	64%	DNAトポイソメラーゼ1
CSTER	994	915311	917623	+	ATP-依存性DNAヘリカーゼPcrA	99%	0.00E+00	54%	ATP-依存性DNAヘリカーゼPcrA
CSTER	1034	959843	961231	-	分岐鎖 α -ケト酸デヒドロゲナーゼサブユニットE2	99%	5.00E-69	34%	2-オキソグルタレートデヒドロゲナーゼ複合体のジヒドロリポイルリシン-残基スクニルトランスフェラーゼ構成要素odhb
CSTER	1036	962412	963383	-	アセトインデヒドロゲナーゼ複合体、E1構成要素、 α サブユニット	81%	3.00E-33	28%	ビルビン酸デヒドロゲナーゼE1構成要素サブユニット α pdha

30

40

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	1087	1007529	1007888	-	50Sリボソームタンパク質L20	85%	2.00E-56	80%	
CSTER	1088	1007945	1008145	-	50Sリボソームタンパク質L35	100%	1.00E-23	62%	
CSTER	1089	1008184	1008714	-	翻訳開始因子3 (IF-3)	98%	2.00E-63	60%	
CSTER	1102	1017867	1018880	-	ペプチド鎖放出因子2	91%	2.00E-137	57%	
CSTER	1115	1029796	1031136	-	センサーヒスチジinkinase	97%	4.00E-117	71%	転写調節タンパク質YycF
CSTER	1116	1031129	1031836	-	二成分応答調節因子	68%	2.00E-120	47%	センサーヒスチジinkinaseYycG
CSTER	1123	1036555	1037826	-	UDP-N-アセチルグルコサミン1-カルボキシビニルトランスフェラーゼ	96%	2.00E-172	59%	UDP-N-アセチルグルコサミン1-カルボキシビニルトランスフェラーゼ1 muraa
CSTER	1135	1049456	1050649	-	S-アデノシルメチオニン合成酵素	97%	0.00E+00	67%	S-アデノシルメチオニンシンターゼmetk
CSTER	1136	1050929	1051867	+	ビオチン- (アセチル-CoAカルボキシラーゼ) リガーゼ	92%	4.00E-44	33%	二官能性リガーゼ/リプレッサーBirA
CSTER	1138	1052266	1053918	-	DNAポリメラーゼIII、 γ/ϵ サブユニット	93%	3.00E-137	44%	DNAポリメラーゼIIIサブユニット γ/ϵ DnaX

10

20

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	1144	1056546	1056893	-	50Sリボソームタンパク質L19	100%	2.00E-61	77%	
CSTER	1164	1073225	1074244	-	6-ホスホフルクトキナーゼ	93%	4.00E-124	61%	ATP-依存性6-ホスホフルクトキナーゼpfkA
CSTER	1165	1074337	1077447	-	DNAポリメラーゼIII、 α サブユニットDnaE	96%	0.00E+00	35%	DNAポリメラーゼIIIサブユニット α dnaE
CSTER	1166	1077600	1078385	-	推定の翻訳因子 (SUA5)	42%	3.00E-04	24%	推定のRNA結合モチーフを有するywlC未知の保存されたタンパク質 TW
CSTER	1175	1085950	1086225	-	ヒストン様DNA結合タンパク質	95%	4.00E-40	72%	DNA-結合タンパク質HU1
CSTER	1182	1091629	1092501	-	ゲラニルゲラニルピロリン酸シンターゼ	89%	5.00E-58	44%	ファルネシルニリン酸シンターゼispayqid
CSTER	1188	1095764	1096462	-	DNA複製タンパク質DnaD	85%	1.00E-07	22%	DNA複製タンパク質DnaD
CSTER	1193	1101272	1102201	-	リボヌクレアーゼZ	99%	2.00E-105	50%	リボヌクレアーゼZ rnz
CSTER	1197	1105450	1106877	-	細胞分裂タンパク質	88%	2.00E-57	38%	Rod形状決定タンパク質RodA
CSTER	1230	1131390	1132742	-	ホスホグルコサミンムターゼ	100%	0.00E+00	63%	ホスホグルコサミンムターゼybbt glnM

30

40

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	1248	1154457	1156616	+	リボヌクレオチドニリン酸レダクターゼ、 α サブユニット	98%	0.00E+00	47%	リボヌクレオチドニリン酸レダクターゼサブユニット α n r d e
CSTER	1249	1156775	1157737	+	リボヌクレオチドニリン酸レダクターゼ、 β サブユニット	97%	8.00E-98	49%	リボヌクレオチドニリン酸レダクターゼサブユニット β n r d f
CSTER	1256	1161831	1164284	-	DNA ジャイレース、A サブユニット	98%	0.00E+00	62%	DNA ジャイレースサブユニット A
CSTER	1267	1174041	1176449	-	フェニルアラニル-tRNA 合成酵素、 β サブユニット	99%	0.00E+00	47%	
CSTER	1270	1177285	1178328	-	フェニルアラニル-tRNA 合成酵素 α サブユニット	98%	3.00E-161	62%	
CSTER	1271	1178950	1182483	-	染色体分離 ATP アーゼ、SMC タンパク質	99%	0.00E+00	37%	染色体分配タンパク質 S m c
CSTER	1272	1182486	1183175	-	リボヌクレアーゼ I I I	89%	4.00E-68	47%	リボヌクレアーゼ 3
CSTER	1273	1183332	1184267	-	ジヒドロジピコリネートシンターゼ	93%	5.00E-88	47%	4-ヒドロキシ-テトラヒドロジピコリネートシンターゼ d a p A

10

20

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	1274	1184605	1185681	-	アスパラギン酸-セミアルデヒドデヒドロゲナーゼ	99%	1.00E-143	58%	アスパラギン酸-セミアルデヒドデヒドロゲナーゼ a s d
CSTER	1382	1293065	1294063	-	チオレドキシンレダクターゼ	97%	7.00E-113	50%	フェレドキシン-NADP レダクターゼ 2 y u m c
CSTER	1383	1294065	1294784	-	tRNA (グアニン-N(1)-)メチルトランスフェラーゼ	100%	6.00E-87	56%	tRNA (グアニン-N(1)-)メチルトランスフェラーゼ
CSTER	1395	1303326	1304261	-	メチオニル-tRNA ホルミルトランスフェラーゼ	96%	5.00E-98	49%	
CSTER	1396	1304279	1306675	-	ブライモソーム集合タンパク質 P r i A	99%	0.00E+00	48%	ブライモソームタンパク質 N P r i A
CSTER	1398	1307153	1307782	-	グアニル酸キナーゼ	98%	5.00E-87	59%	グアニル酸キナーゼ g m k
CSTER	1399	1308015	1309406	-	細胞分裂タンパク質 F t s Y	97%	8.00E-121	55%	シグナル認識粒子受容体 F t s Y
CSTER	1422	1330503	1331339	-	無機ポリリン酸/ATP-NAD キナーゼ	100%	5.00E-78	44%	NAD キナーゼ 1 p p n k
CSTER	1425	1333019	1333990	+	リボース-リン酸ピロホスホキナーゼ	100%	4.00E-113	54%	リボース-リン酸ピロホスホキナーゼ p r s
CSTER	1426	1333994	1335106	+	推定のシステイン脱硫酵素	98%	2.00E-115	47%	推定のシステイン脱硫酵素 I s c S

30

40

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	1448	1355769	1356878	-	RNAポリメラーゼの因子RpoD	93%	6.00E-167	70%	RNAポリメラーゼの因子SigA
CSTER	1449	1356882	1358693	-	DNAプライマーゼ	68%	9.00E-86	37%	DNAプライマーゼDnaG
CSTER	1451	1359076	1359252	-	30Sリボソームタンパク質S21	98%	3.00E-29	89%	
CSTER	1464	1371306	1372619	-	GTPアーゼObgE	100%	0.00E+00	67%	GTPアーゼObgE
CSTER	1480	1387053	1389005	-	DNAジャイレースサブユニットB	100%	0.00E+00	68%	DNAジャイレースサブユニットB
CSTER	1497	1401241	1402143	-	UDP-N-アセチレノールビルボイルグルコサミンレダクターゼ	98%	1.00E-66	40%	UDP-N-アセチレノールビルボイルグルコサミンレダクターゼmurB
CSTER	1506	1409996	1410268	-	30Sリボソームタンパク質S16	100%	5.00E-42	67%	
CSTER	1512	1415065	1416078	-	推定の脂質キナーゼ	95%	1.00E-96	51%	ジアシルグリセロールキナーゼdagK
CSTER	1513	1416088	1418034	-	NAD-依存性DNAリガーゼ	97%	0.00E+00	55%	DNAリガーゼLigA
CSTER	1516	1419659	1420519	-	メチオニンアミノペプチダーゼ	99%	6.00E-56	37%	

10

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	1519	1422447	1423709	-	UDP-N-アセチルグルコサミン1-カルボキシビニルトランスフェラーゼ	96%	6.00E-125	46%	UDP-N-アセチルグルコサミン1-カルボキシビニルトランスフェラーゼ1 muraa
CSTER	1522	1426828	1427583	-	1-アシル-sn-グリセロール-3-ホスフェートアシルトランスフェラーゼ	92%	1.00E-29	32%	1-アシル-sn-グリセロール-3-ホスフェートアシルトランスフェラーゼplsC
CSTER	1534	1439747	1441120	-	UDP-N-アセチルムラモイルトリペプチド-D-アラニル-D-アラニンリガーゼ	98%	2.00E-98	38%	UDP-N-アセチルムラモイルトリペプチド-D-アラニル-D-アラニンリガーゼmurF
CSTER	1544	1447537	1448583	-	D-アラニル-アラニン合成酵素A	98%	2.00E-95	42%	D-アラニン-D-アラニンリガーゼddl
CSTER	1583	1483475	1484107	-	ニコチン酸モノヌクレオチドアデニルルトランスフェラーゼ	98%	2.00E-62	44%	ニコチン酸-ヌクレオチドアデニルルトランスフェラーゼnad
CSTER	1584	1484211	1484528	-	推定のRNA-結合タンパク質	95%	1.00E-24	50%	可能性のRNA-結合タンパク質YqeI
CSTER	1585	1484767	1485885	-	GTP-結合タンパク質YqeH	100%	6.00E-163	59%	性質不明のタンパク質YqeH

20

30

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	1590	1488021	1489463	-	アスパルチル/グルタミル- tRNAアミドトランスフェラーゼサブユニットB	99%	0.00E+00	63%	
CSTER	1591	1489463	1490929	-	アスパルチル/グルタミル- tRNAアミドトランスフェラーゼサブユニットA	99%	0.00E+00	59%	
CSTER	1592	1490929	1491231	-	アスパルチル/グルタミル- tRNAアミドトランスフェラーゼサブユニットC	97%	1.00E-18	41%	
CSTER	1615	1511378	1512298	-	チオレドキシンレダクターゼ	95%	3.00E-126	58%	チオレドキシンレダクターゼ t r x b
CSTER	1665	1554189	1555211	-	ホスホ-N-アセチルムラモイル-ペンタペプチド-トランスフェラーゼ	92%	1.00E-66	41%	ホスホ-N-アセチルムラモイル-ペンタペプチド-トランスフェラーゼ m r a y
CSTER	1666	1555213	1557480	-	推定のペニシリン-結合タンパク質 2 X	90%	9.00E-88	32%	ペニシリン-結合タンパク質 2 B
CSTER	1667	1557484	1557804	-	細胞分裂タンパク質 F t s L				
CSTER	1701	1591284	1592387	-	アラニンラセマーゼ	93%	1.00E-98	43%	アラニンラセマーゼ 1 a l r

10

20

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	1702	1592408	1592767	-	推定の4'-ホスホパンテチエニルトランスフェラーゼ	97%	6.00E-29	41%	ホロ- [アシル-担体-タンパク質] シンターゼ
CSTER	1705	1594955	1597504	-	タンパク質トランスロカーゼサブユニット S e c A	100%	0.00E+00	55%	タンパク質トランスロカーゼサブユニット S e c A
CSTER	1726	1616337	1616576	-	30Sリボソームタンパク質 S 18	100%	5.00E-32	65%	
CSTER	1727	1616618	1617136	-	単鎖DNA-結合タンパク質	93%	1.00E-41	56%	一本鎖DNA-結合タンパク質 B
CSTER	1728	1617148	1617438	-	30Sリボソームタンパク質 S 6	96%	4.00E-40	61%	
CSTER	1745	1633808	1634821	-	推定のO-シアロ糖タンパク質エンドペプチダーゼ	97%	6.00E-130	55%	tRNA N6-アデノシンスレオニルカルバモイルトランスフェラーゼ t s a d
CSTER	1747	1635236	1635922	-	推定の糖タンパク質エンドペプチダーゼ	100%	6.00E-44	38%	tRNAスレオニルカルバモイルアデノシン生成タンパク質 T s a B
CSTER	1749	1636421	1638103	+	mRNA分解リボヌクレアーゼ s J 1 / J 2	99%	0.00E+00	62%	リボヌクレアーゼ J 1 y q k c

30

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	1755	1642501	1643700	-	ホスホグリセレートキナーゼ	100%	2.00E-124	50%	ホスホグリセレートキナーゼ p g k
CSTER	1762	1648240	1650321	-	翻訳伸長因子 G	100%	0.00E+00	77%	
CSTER	1763	1650542	1651012	-	30S リボソームタンパク質 S 7	100%	1.00E-86	75%	
CSTER	1764	1651031	1651444	-	30S リボソームタンパク質 S 12	98%	3.00E-80	86%	
CSTER	1770	1656078	1656950	-	リボソーム関連 GTPアーゼ	99%	5.00E-105	51%	推定のリボソーム生合成 GTP アーゼ R s g A
CSTER	1776	1660099	1660413	-	推定のチオレドキシン	92%	1.00E-37	61%	チオレドキシン t r x a
CSTER	1787	1670058	1670192	-	50S リボソームタンパク質 L 34	100%	1.00E-15	70%	
CSTER	1790	1673060	1673395	-	リボスクレアーゼ P	95%	1.00E-31	47%	リボスクレアーゼ P タンパク質 構成要素
CSTER	1793	1677296	1678750	-	グルタミル- t RNA 合成酵素	100%	0.00E+00	55%	
CSTER	1797	1681106	1681795	-	50S リボソームタンパク質 L 1	98%	4.00E-99	65%	50S リボソームタンパク質 L 1

10

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	1809	1690036	1691058	+	グリセロール-3-ホスフェートデヒドロゲナーゼ	94%	4.00E-122	53%	グリセロール-3-ホスフェートデヒドロゲナーゼ [NAD(P)+] g s p A
CSTER	1813	1693298	1694431	-	金属依存性アミダーゼ/アミノアシラーゼ/カルボキシペプチダーゼ	95%	6.00E-120	47%	N-アセチルジアミノピメレートデアセチラーゼ y k u r
CSTER	1814	1694510	1695208	-	推定の 2, 3, 4, 5-テトラヒドロピリジン-2-カルボキシレート N-スクシニルトランスフェラーゼ	98%	4.00E-86	60%	2, 3, 4, 5-テトラヒドロピリジン-2, 6-ジカルボキシレート N-アセチルトランスフェラーゼ d a p H
CSTER	1821	1699209	1699601	-	単鎖結合タンパク質	93%	1.00E-25	40%	一本鎖 DNA-結合タンパク質 B
CSTER	1844	1713286	1716924	-	DNA-依存性 RNA ポリメラーゼサブユニット β'	97%	0.00E+00	68%	DNA-依存性 RNA ポリメラーゼサブユニット β
CSTER	1845	1717025	1720606	-	DNA-依存性 RNA ポリメラーゼサブユニット β	98%	0.00E+00	71%	DNA-依存性 RNA ポリメラーゼサブユニット β'
CSTER	1847	1723493	1724749	+	チロシル- t RNA 合成酵素	99%	0.00E+00	58%	

20

30

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	1876	1751169	1752050	-	フルクトースービスホスフェートアルドラーゼ	100%	4.00E-83	45%	可能性のフルクトースービスホスフェートアルドラーゼ f b a a
CSTER	1880	1755102	1755488	-	50S リボソームタンパク質 L 17	100%	2.00E-62	73%	
CSTER	1881	1755506	1756444	-	DNA-依存性RNAポリメラーゼαサブユニット	99%	7.00E-139	62%	DNA-依存性RNAポリメラーゼサブユニットα
CSTER	1882	1756493	1756876	-	30S リボソームタンパク質 S 11	93%	7.00E-69	87%	
CSTER	1883	1756904	1757269	-	30S リボソームタンパク質 S 13	100%	2.00E-58	73%	
CSTER	1884	1757290	1757403	-	50S リボソームタンパク質 L 36	100%	4.00E-16	84%	
CSTER	1885	1757429	1757647	-	翻訳開始因子1 (IF-1)	100%	1.00E-38	78%	
CSTER	1886	1757765	1758421	-	アデニル酸キナーゼ	99%	3.00E-89	59%	アデニル酸キナーゼ a d k
CSTER	1887	1758553	1759848	-	プレタンパク質トランスロカーゼサブユニット S e c Y	99%	4.00E-140	49%	タンパク質トランスロカーゼサブユニット S e c Y
CSTER	1888	1759865	1760305	-	50S リボソームタンパク質 L 15	100%	1.00E-70	72%	

10

20

ゲノム部位	STER	開始	終了	方向		ク エ リ ー カ バ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	1889	1760433	1760615	-	50S リボソームタンパク質 L 30	98%	3.00E-20	62%	
CSTER	1890	1760630	1761124	-	30S リボソームタンパク質 S 5	93%	8.00E-70	76%	
CSTER	1891	1761143	1761499	-	50S リボソームタンパク質 L 18	100%	4.00E-55	73%	
CSTER	1892	1761589	1762125	-	50S リボソームタンパク質 L 6	100%	3.00E-69	60%	50S リボソームタンパク質 L 6
CSTER	1893	1762252	1762650	-	30S リボソームタンパク質 S 8	100%	1.00E-73	77%	
CSTER	1894	1762773	1762958	-	30S リボソームタンパク質 S 14	100%	1.00E-31	77%	
CSTER	1895	1762976	1763518	-	50S リボソームタンパク質 L 5	99%	3.00E-100	77%	
CSTER	1896	1763545	1763850	-	50S リボソームタンパク質 L 24	99%	7.00E-38	64%	
CSTER	1897	1763931	1764299	-	50S リボソームタンパク質 L 14	100%	5.00E-74	87%	

30

ゲノム部位	STER	開始	終了	方向		ク エ リー カバ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	1898	1764324	1764584	-	30Sリボソームタンパク質S 17	97%	4.00E-46	84%	
CSTER	1899	1764612	1764818	-	50Sリボソームタンパク質L 29	78%	1.00E-14	58%	
CSTER	1900	1764828	1765241	-	50Sリボソームタンパク質L 16	93%	1.00E-76	83%	
CSTER	1901	1765245	1765898	-	30Sリボソームタンパク質S 3	100%	1.00E-119	75%	
CSTER	1902	1765911	1766255	-	50Sリボソームタンパク質L 22	96%	7.00E-49	64%	
CSTER	1903	1766271	1766549	-	30Sリボソームタンパク質S 19	100%	2.00E-54	83%	
CSTER	1904	1766643	1767476	-	50Sリボソームタンパク質L 2	100%	5.00E-159	76%	50Sリボソームタンパク質L 2
CSTER	1905	1767494	1767790	-	50Sリボソームタンパク質L 23	94%	2.00E-31	60%	
CSTER	1906	1767790	1768413	-	50Sリボソームタンパク質L 4	100%	9.00E-92	60%	50Sリボソームタンパク質L 4

10

20

ゲノム部位	STER	開始	終了	方向		ク エ リー カバ ー	e 値	aa id	Bacillus subtilis 注釈
CSTER	1907	1768438	1769064	-	50Sリボソームタンパク質L 3	99%	4.00E-115	75%	50Sリボソームタンパク質L 3
CSTER	1908	1769181	1769489	-	30Sリボソームタンパク質S 10	100%	1.00E-58	78%	
CSTER	1936	1795296	1795484	+	50Sリボソームタンパク質L 28	100%	7.00E-26	69%	
CSTER	1948	1805428	1807179	-	アスパルチル- tRNA合成酵 素	98%	0.00E+00	57%	
CSTER	1950	1807719	1808999	-	ヒスチジル- tRNA合成酵素	99%	1.00E-140	49%	
CSTER	1953	1810651	1810833	+	50Sリボソームタンパク質L 32	91%	1.00E-10	48%	
CSTER	1954	1810849	1810998	+	50Sリボソームタンパク質L 33	100%	1.00E-14	55%	
CSTER	1973	1823718	1824329	-	30Sリボソームタンパク質S 4	100%	1.00E-95	69%	
CSTER	1975	1824852	1826213	-	複製DNAヘリカーゼ	95%	8.00E-176	56%	複製DNAヘリカーゼDnaC
CSTER	1976	1826257	1826715	-	50Sリボソームタンパク質L 9	100%	2.00E-39	55%	

30

40

ゲノム部位	STER	開始	終了	方向		クエリー カバー	e 値	aaid	Bacillus subtilis 注釈
CSTER	1979	1830931	1832052	-	tRNA (5-メチルアミノメチル-2-チオウリジレート)-メチルトランスフェラーゼ	98%	9.00E-177	70%	tRNA-特異的2-チオウリジラーゼMnmA
CSTER	1986	1838183	1838725	-	ホスファチジルグリセロールリン酸シンターゼ	93%	2.00E-45	52%	CDP-ジアシルグリセロール-グリセロール-3-ホスフェート3-ホスファチルトランスフェラーゼpgsA
CSTER	1992	1843910	1845391	-	イノシトール-5-モノホスフェートデヒドロゲナーゼ	99%	0.00E+00	68%	イノシン-5'-モノホスフェートデヒドロゲナーゼguaB
CSTER	1993	1845568	1846590	-	トリプトファニール-tRNA合成酵素II	98%	8.00E-42	35%	

10

【0192】

【表5】

表5. 欠失株GEI4において同定された示唆的に発現される遺伝子のリスト。

オペロン	推定の経路/機能	調節因子	遺伝子	log ₂ 変化	p 値
STER_0390-STER_0393	システイン代謝	Cmbr/Homr/Htar	STER_0390	-1.1	2.7E-23
			STER_0391	-1.3	2.2E-22
			STER_0392	-1.4	1.2E-32
			STER_0393	-1.2	2.6E-15
STER_1016-STER_1017	マルトデキストリン代謝		STER_1016	-1.0	0
			STER_1017	-0.6	7.50E-36
STER_1548-STER_1555	芳香族アミノ酸生合成	T-box RNA (Tfp)	STER_1548	-1.3	2.3E-37
			STER_1549	-1.0	1.5E-26
			STER_1550	-1.3	7.9E-38
			STER_1551	-1.3	1.2E-35
			STER_1552	-1.3	0
			STER_1553	-1.4	0
			STER_1554	-1.3	0
STER_1960-STER_1963	膜たんぱく質		STER_1555	-1.6	7.8E-73
			STER_1960	-0.9	0
			STER_1961	-1.0	0
			STER_1962	-1.0	8.3E-317
			STER_1963	-1.1	3.2E-965
STER_0049-STER_0054	プリン生合成	PurR	STER_0049	1.2	5.6E-34
			STER_0050	1.1	1.4E-29
			STER_0051	1.1	0
			STER_0052	1.1	0
			STER_0053	1.1	0
			STER_0054	1.1	6.6E-40
STER_0699-STER_0701	エタノールアミン代謝		STER_0699	1.6	0
			STER_0700	2.1	9.5E-200
			STER_0701	2.2	7.8E-329
STER_1020-STER_1024	ツインアルギニントランスロカーゼ	TatA	STER_1020	1.3	5.3E-14
			STER_1021	1.3	8.8E-12
			STER_1022	1.4	3.3E-12
			STER_1023	1.0	9.6E-6
			STER_1024	0.9	0.0022
STER_1026-STER_1028	鉄ホメオスタシス	PerR	STER_1026	1.1	1.9E-09
			STER_1027	0.9	2.5E-6
			STER_1028	1.2	3E-12
			STER_1029	1.2	1.5E-17
STER_1405-STER_1409	ABCペプチド輸送	Cody	STER_1405	1.6	9.6E-90
			STER_1406	1.6	0
			STER_1407	1.6	3.7E-59
			STER_1408	1.5	0
			STER_1409	1.4	4.2E-54
STER_1821-STER_1823	ストレス		STER_1821	2.0	0
			STER_1822	2.2	5.4E-98
			STER_1823	2.0	0

20

30

【0193】

参考文献

40

1. Darmon E, Leach DF. (2014) Bacterial Genome Instability. *Microbiol. Mol. Biol. Rev.* 78, 1–39.
2. Labrie SJ, Samson JE, and Moineau S. (2010) Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* 8, 317–327.
3. Barrangou R, Marraffini LA (2014) CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity. *Mol Cell* 54(2):234–244.
4. Barrangou R, et al. (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315(5819):1709–1712. 10
5. Brouns SJJ, et al. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321(5891):960–964.
6. Young JC et al. (2012) Phage-induced expression of CRISPR-associated proteins is revealed by shotgun proteomics in *Streptococcus thermophilus*. *PLoS ONE* 7(5):e38077.
7. Garneau JE, et al. (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468(7320):67–71. 20
8. Groenen PM, Bunschoten AE, Van Soolingen D, Van Embden JD (1993) Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. *Mol Microbiol* 10(5):1057–1065.
9. Yin S, et al. (2013) The evolutionary divergence of Shiga toxin-producing *Escherichia coli* is reflected in clustered regularly interspaced short palindromic repeat (CRISPR) spacer composition. *Appl Environ Microbiol* 79(18):5710–5720. 30
10. Liu F, et al. (2011) Novel virulence gene and clustered regularly interspaced short palindromic repeat (CRISPR) multilocus sequence typing scheme for subtyping of the major serovars of *Salmonella enterica* subsp. *enterica*. *Appl Environ Microbiol* 77(6):1946–1956.
11. Barrangou R, Horvath P (2012) CRISPR: new horizons in phage resistance and strain identification. *Annu Rev Food Sci Technol* 3, 143–162.
12. Sander JD, and Joung JK. (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* 32, 347–355. 40
13. Bikard D, et al. (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res* 41(15):7429–7437

14. Qi LS, et al. (2013) Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* 152(5):1173–1183.
15. Goma AA, et al. (2014) Programmable Removal of Bacterial Strains by Use of Genome-Targeting CRISPR-Cas Systems. *mBio* 5(1):e00928–13.
16. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A (1987) Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol* 169(12):5429–5433. 10
17. Jansen R, Embden JDA van, Gaastra W, Schouls LM (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* 43(6):1565–1575.
18. Bolotin A, et al. (2004) Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nat Biotechnol* 22(12):1554–1558. 20
19. Horvath P, et al. (2008) Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J Bacteriol* 190(4):1401–1412.
20. Bolotin A, Quinguis B, Sorokin A, Ehrlich SD (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 151(Pt 8):2551–2561.
21. Deveau H, et al. (2008) Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J Bacteriol* 190(4):1390–1400.
22. Paez-Espino D, et al. (2013) Strong bias in the bacterial CRISPR elements that confer immunity to phage. *Nat Commun* 4, 1430. 30
23. Sun CL, et al. (2013) Phage mutations in response to CRISPR diversification in a bacterial population. *Environ Microbiol* 15(2):463–470.
24. Sapranas R, et al. (2011) The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res* 39(21):9275–9282.
25. Gasiunas G, Barrangou R, Horvath P, Siksnys V (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci USA* 109(39):E2579–2586. 40
26. Briner AE, et al. (2014) Guide RNA Functional Modules Direct Cas9 Activity and Orthogonality. *Molecular Cell*. 56(2):333–339
27. Bondy-Denomy J, Davidson AR. (2014) To acquire or resist: the complex biological effects of CRISPR-Cas systems. *Trends Microbiol.* 22, 218–225.

28. Horvath P, et al. (2009) Comparative analysis of CRISPR loci in lactic acid bacteria genomes. *Int J Food Microbiol* 131(1):62-70.
29. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. (2013a) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol* 31(3):233-239.
30. Vercoe RB, et al. (2013) Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. *PLoS Genet* 9(4):e1003454.
31. Oh JH, van Pijkeren JP. (2014) CRISPR-Cas9-assisted recombineering in *Lactobacillus reuteri*. *Nucleic Acids Res* 10.1093/nar/gku623
32. Selle K, Barrangou R. (2015) Harnessing CRISPR-Cas systems for bacterial genome editing. *Trends Microbiol*. In press
33. Kobayashi K, et al. (2003) Essential *Bacillus subtilis* genes. *Proc. Natl. Acad. Sci. U.S.A.* 100, 4678-4683.
34. Mahillon J, Chandler M. (1998) Insertion sequences. *Microbiol Mol Biol Rev* 62(3):725-774.
35. Goh YJ, Goin C, O'Flaherty S, Altermann E, Hutkins R (2011) Specialized adaptation of a lactic acid bacterium to the milk environment; the comparative genomics of *Streptococcus thermophilus* LMD-9. *Microbial Cell Factories* 10 (Suppl 1):S22
36. Dandoy D, et al. (2011) The fast milk acidifying phenotype of *Streptococcus thermophilus* can be acquired by natural transformation of the genomic island encoding the cell-envelope proteinase PrtS. *Microb. Cell Fact.* 10 Suppl 1, S21.
37. Deltcheva E, et al. (2011) CRISPR RNA maturation by trans-activating small RNA and host factor RNase III. *Nature* 471, 602-607.
38. Aravind L, Koonin EV (2001) Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. *Genome Res* 11(8):1365-1374.
39. Koonin EV, Makarova KS (2007) Evolutionary genomics of lactic acid bacteria. *J Bacteriol* 189(4):1199-1208
40. Makarova KS, et al. (2006) Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci USA* 103(42):15611-15616.
41. Goh YJ, et al. (2009) Development and application of a upp-based counterselective gene replacement system for the study of the S-layer protein SlpX of *Lactobacillus acidophilus* NCFM. *Appl Environ Microbiol* 75(10):3093-3105.

10

20

30

40

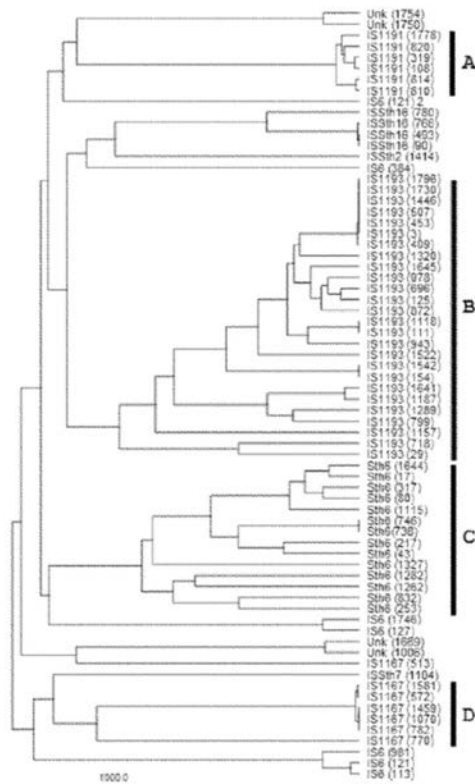
42. Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR (1989) Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 77(1):61–68.
43. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Almendros C (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155(Pt 3):733–740.
44. Russell WM, Klaenhammer TR (2001) Efficient System for Directed Integration into the *Lactobacillus acidophilus* and *Lactobacillus gasseri* Chromosomes via Homologous Recombination. *Appl Environ Microbiol* 67(9):4361–4364. 10
45. Wei M-Q, et al. (1995) An improved method for the transformation of *Lactobacillus* strains using electroporation. *Journal of Microbiological Methods* 21(1):97–109.
46. Zhang X, Bremer H (1995) Control of the *Escherichia coli* rrnB p1 promoter strength by ppGpp. *Journal of Biological Chemistry* 270(19):11181–11189. 20
47. Law J, et al. (1995) A system to generate chromosomal mutations in *Lactococcus lactis* which allows fast analysis of targeted genes. *J Bacteriol* 177(24):7011–7018.

【 0 1 9 4 】

前述は本発明の例示であり、その制限と解釈されるべきでない。本発明は、以下の特許請求の範囲によって定義され、特許請求の範囲の同等物はそれに含まれる。

【図 4】

Fig. 4



【図 5 A - 5 D】

Fig. 5A

IS1191

Fig. 5B

Sth6

Fig. 5C

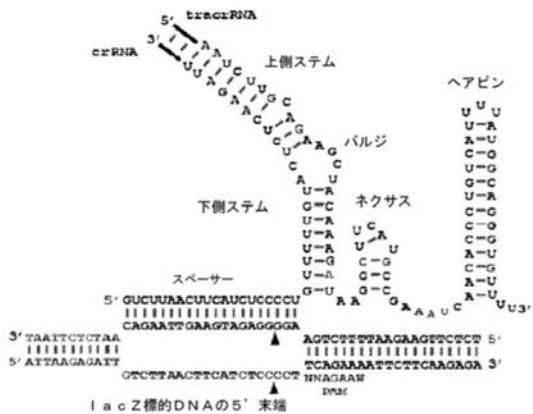
IS1193

Fig. 5D

IS1167

【図 6 A】

Fig. 6A



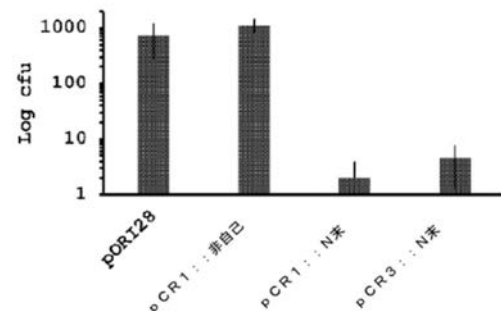
【図 6 B】

Fig. 6B

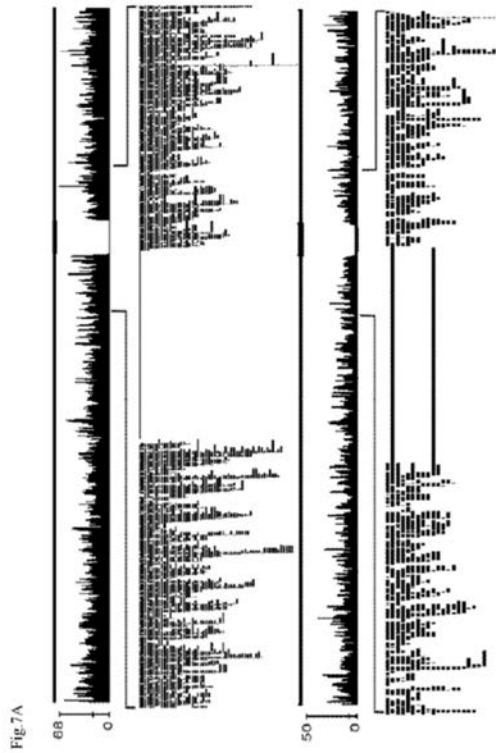


【図 6 C】

Fig. 6C

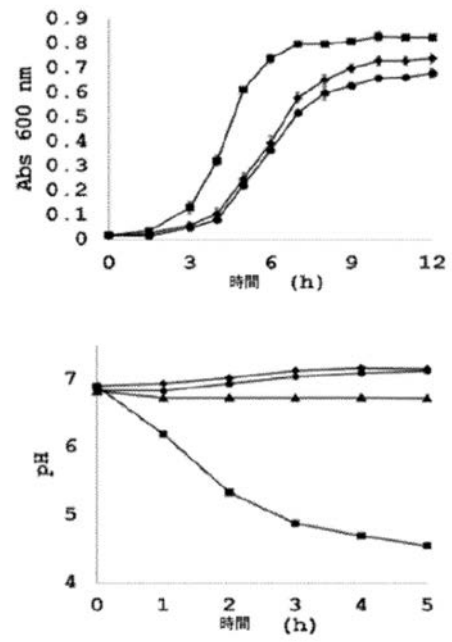


【図 7 A】

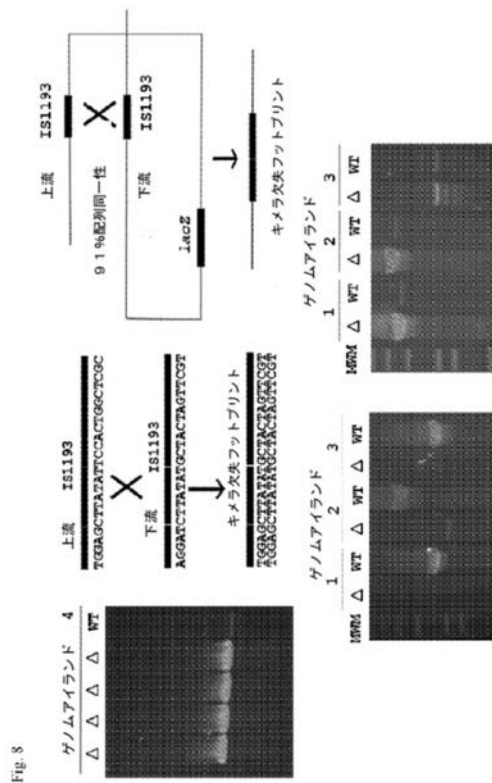


【図 7 B】

Fig. 7B

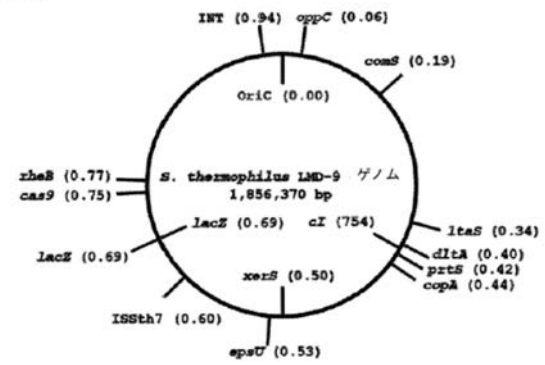


【図 8】



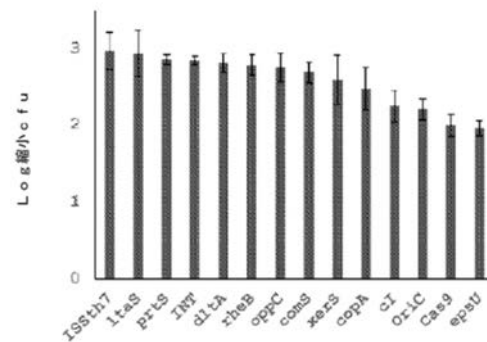
【図 9】

Fig. 9

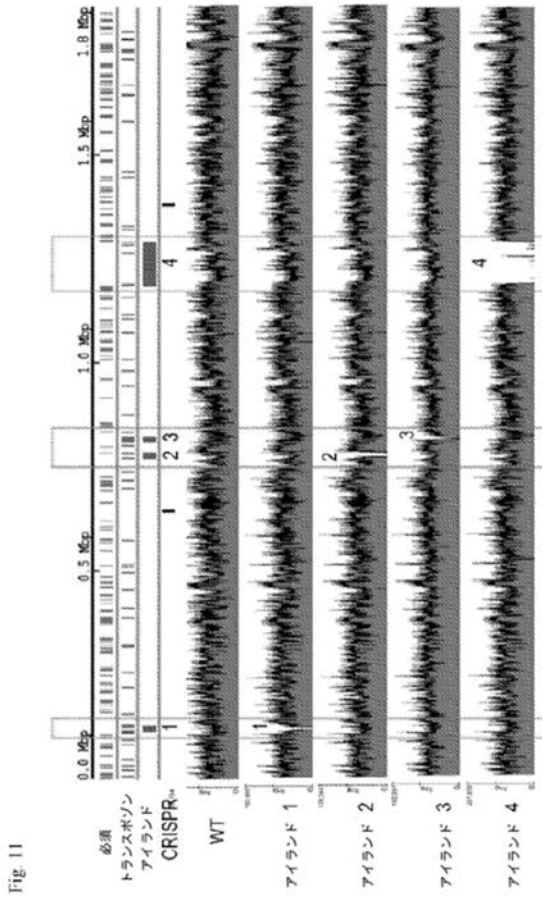


【図 10】

Fig. 10

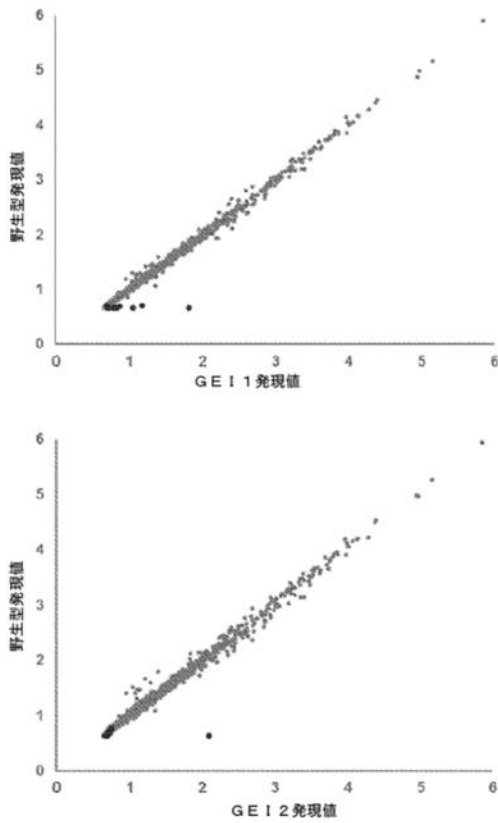


【図 1 1】

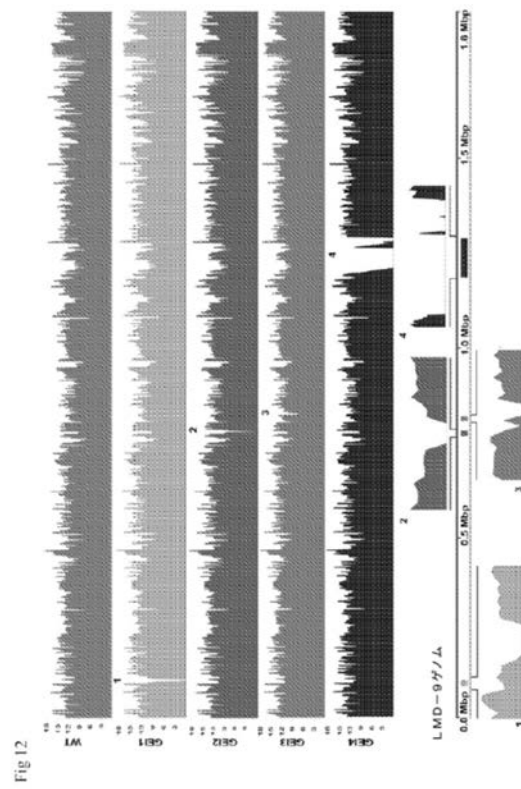


【図 1 3 - 1】

Fig. 13

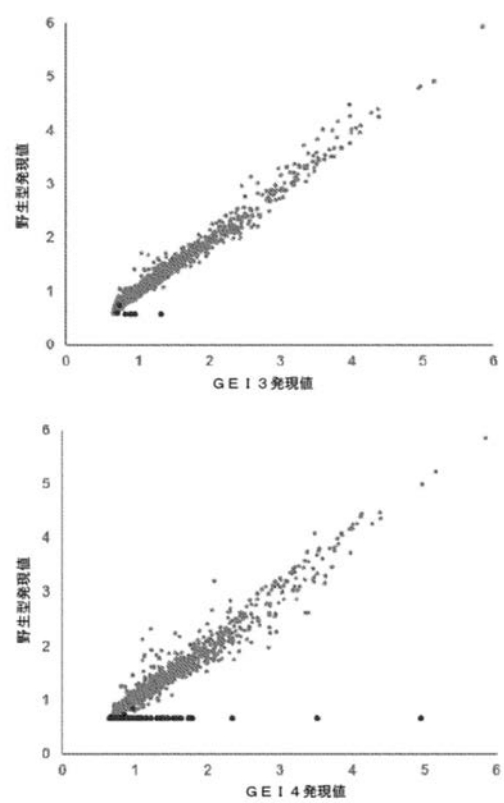


【図 1 2】

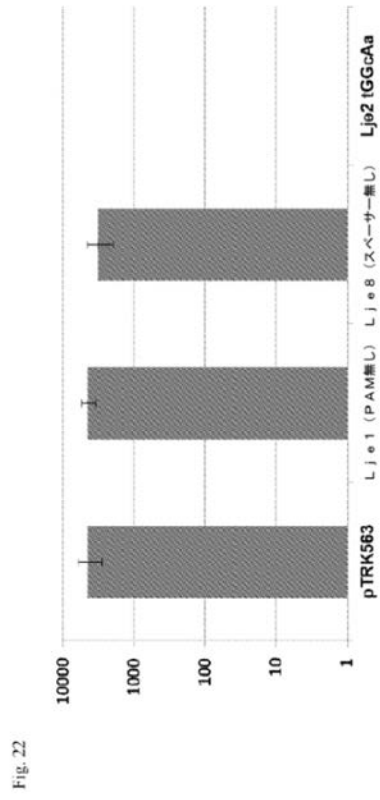


【図 1 3 - 2】

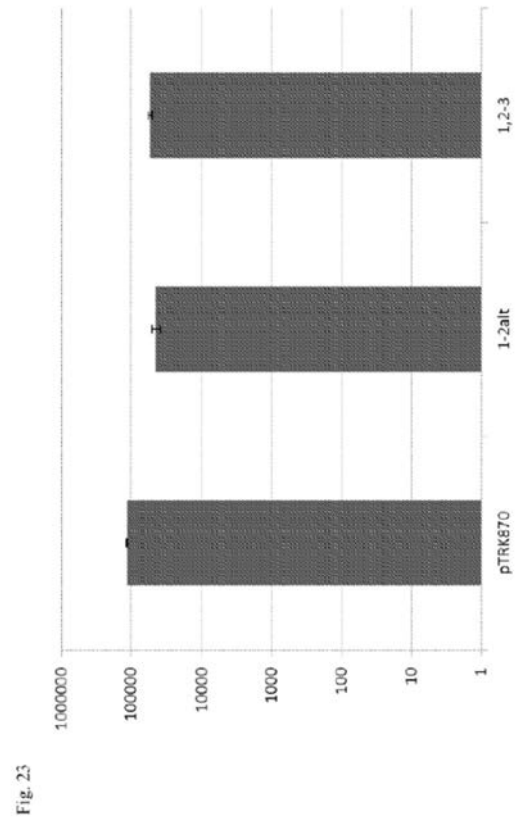
Fig. 13, cont'd.



【図 2 2】



【図 2 3】



【配列表】

2018516563000001.app

【国際調査報告】

INTERNATIONAL SEARCH REPORT	International application No. PCT/US 2016/034812
Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)	
This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:	
1.	☐ Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:
2.	☐ Claims Nos.: because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3.	☒ Claims Nos.: 12-16, 28-31 because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).
Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)	
This International Searching Authority found multiple inventions in this international application, as follows:	
1.	☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.	☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3.	☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4.	☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
Remark on Protest	☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee. ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation. ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT		International application No. PCT/US 2016/034812
A. CLASSIFICATION OF SUBJECT MATTER <i>C12Q 1/02 (2006.01)</i> <i>C12Q 1/68 (2006.01)</i> <i>C12N 1/21 (2006.01)</i> <i>C12N 15/63 (2006.01)</i> According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols)		
C12Q 1/02, 1/68, C12N 1/21, 15/63, 7/01		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
DB WIPO (Patentscope), DB Espacenet, PatFT & AppFT USPTO, RUPAT, EAPATIS, CIPO, PAJ, KIPRIS, K-PION, SIPO, DWPI; PubMed, Scopus		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y A Y A	GOMAA AHMED A. et al. "Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems." mBio, 2014, Vol.5, Issue 1: e00928-13. doi: 10.1128/mBio.00928-13, p.1-9, especially abstract, p.1-7 WO 2014/204727 A1 (THE BROAD INSTITUTE INC. et al.) 24.12.2014, abstract, paragraphs [0095], [00130]-[00141], example 1 EP 2 860 267 A1 (DUPONT NUTRITION BIOSCIENCES APS) 15.04.2015, abstract	1, 3, 5-7, 17, 21 2, 4, 8-11, 19, 23, 25-27 18, 20, 22, 24 2, 4, 8-11, 19, 23, 25-27 1-11, 17-27
☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 30 August 2016 (30.08.2016)		Date of mailing of the international search report 15 September 2016 (15.09.2016)
Name and mailing address of the ISA/RU: Federal Institute of Industrial Property, Berezhkovskaya nab., 30-1, Moscow, G-59, GSP-3, Russia, 125993 Facsimile No: (8-495) 531-63-18, (8-499) 243-33-37		Authorized officer T.Babakova Telephone No. 495 531 65 15

フロントページの続き

(51)Int.Cl.	F I	テーマコード (参考)
C 1 2 N 1/13 (2006.01)	C 1 2 N 1/13	
C 1 2 N 1/20 (2006.01)	C 1 2 N 1/20	Z
C 1 2 N 1/21 (2006.01)	C 1 2 N 1/21	

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

(74)代理人 100137811

弁理士 原 秀貢人

(72)発明者 バランゴウ ロドルフ

アメリカ合衆国 2 7 6 0 6 ノースカロライナ, ローリー メイン キャンパス ドライブ
1 0 2 1 セカンド フロア

(72)発明者 セル クルト エム.

アメリカ合衆国 2 7 6 0 6 ノースカロライナ, ローリー メイン キャンパス ドライブ
1 0 2 1 セカンド フロア

Fターム(参考) 4B063 QA08 QQ05 QQ06 QQ07 QQ13 QQ34 QQ43 QQ44 QR14 QR33

QR75 QR76 QR80 QS05 QS38 QX01

4B065 AA01X AA01Y AA72X AA83X AB01 AC20 BA01 CA46