

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第6854340号
(P6854340)

(45) 発行日 令和3年4月7日 (2021. 4. 7)

(24) 登録日 令和3年3月17日 (2021. 3. 17)

(51) Int. Cl.	F I
C 1 2 N 15/11 (2006. 01)	C 1 2 N 15/11 Z
C 4 O B 40/06 (2006. 01)	C 4 O B 40/06 Z N A
C 1 2 N 1/21 (2006. 01)	C 1 2 N 1/21
C 1 2 N 1/15 (2006. 01)	C 1 2 N 1/15
C 1 2 N 1/19 (2006. 01)	C 1 2 N 1/19

請求項の数 20 (全 80 頁) 最終頁に続く

(21) 出願番号 特願2019-510673 (P2019-510673)	(73) 特許権者 516040017 ツイスト バイオサイエンス コーポレー ション アメリカ合衆国 9 4 1 5 8 カリフォル ニア州 サンフランシスコ ミッション・ ベイ・ブールバード サウス 4 5 5 ス イート 5 4 5
(86) (22) 出願日 平成29年4月5日 (2017. 4. 5)	(74) 代理人 100082072 弁理士 清原 義博
(65) 公表番号 特表2019-528070 (P2019-528070A)	(72) 発明者 コックス, アンソニー アメリカ合衆国 9 4 0 4 0 カリフォル ニア州 マウンテンビュー プラザ・コー ト 1 7 6 8
(43) 公表日 令和1年10月10日 (2019. 10. 10)	
(86) 国際出願番号 PCT/US2017/026232	
(87) 国際公開番号 W02018/038772	
(87) 国際公開日 平成30年3月1日 (2018. 3. 1)	
審査請求日 令和2年4月3日 (2020. 4. 3)	
(31) 優先権主張番号 62/378, 134	
(32) 優先日 平成28年8月22日 (2016. 8. 22)	
(33) 優先権主張国・地域又は機関 米国 (US)	
(31) 優先権主張番号 62/382, 191	
(32) 優先日 平成28年8月31日 (2016. 8. 31)	
(33) 優先権主張国・地域又は機関 米国 (US)	最終頁に続く

(54) 【発明の名称】 デノボ合成された核酸ライブラリ

(57) 【特許請求の範囲】

【請求項 1】

核酸ライブラリであって、該核酸ライブラリは、少なくとも 5 0 0 の非同一の D N A 分子を含み、各非同一の D N A 分子は異なる g R N A 配列をコードし、各 g R N A 配列は哺乳類遺伝子に相補的な標的ドメインを含み、および、少なくとも 5 0 0 の非同一の D N A 分子の少なくとも 8 0 % が、核酸ライブラリ中の非同一の D N A 分子の各々について平均頻度の 2 倍以内の量で核酸ライブラリにそれぞれ存在する、核酸ライブラリ。

【請求項 2】

各非同一の D N A 分子は、2 0 % ~ 8 5 % の G C 塩基含有量を有する、請求項 1 に記載の核酸ライブラリ。

【請求項 3】

各非同一の D N A 分子は、3 0 % ~ 7 0 % の G C 塩基含有量を有する、請求項 1 に記載の核酸ライブラリ。

【請求項 4】

少なくとも 5 0 0 の非同一の D N A 分子の少なくとも 9 0 % は、核酸ライブラリ中の非同一の D N A 分子の各々について平均頻度の 2 倍以内の量で核酸ライブラリにそれぞれ存在する、請求項 1 に記載の核酸ライブラリ。

【請求項 5】

少なくとも 5 0 0 の非同一の D N A 分子は、少なくとも 2 0 0 0 の非同一の D N A 分子を含む、請求項 1 に記載の核酸ライブラリ。

【請求項 6】

少なくとも 500 の非同一の DNA 分子は、少なくとも 3500 の非同一の DNA 分子を含む、請求項 1 に記載の核酸ライブラリ。

【請求項 7】

各非同一の DNA 分子は、最大 200 の塩基の長さを含む、請求項 1 に記載の核酸ライブラリ。

【請求項 8】

少なくとも 500 の非同一の DNA 分子は、生物学的経路において、遺伝子を標的とする gRNA 配列をコードする非同一の DNA 分子を含む、請求項 1 に記載の核酸ライブラリ。

10

【請求項 9】

少なくとも 500 の非同一の DNA 分子は、全ゲノムにおいて、遺伝子を標的とする gRNA 配列をコードする非同一の DNA 分子を含む、請求項 1 に記載の核酸ライブラリ。

【請求項 10】

gRNA は単一の gRNA または二重の gRNA である、請求項 1 に記載の核酸ライブラリ。

【請求項 11】

核酸ライブラリであって、該核酸ライブラリは少なくとも 2000 の非同一の核酸を含み、各非同一の核酸は異なる sgRNA 配列をコードし、各 sgRNA 配列は真核生物の遺伝子に相補的な標的ドメインを含み、および、少なくとも 2000 の非同一の核酸の少なくとも 80% は、核酸ライブラリ中の非同一の核酸の各々について平均頻度の 2 倍以内の量で核酸ライブラリに存在する、核酸ライブラリ。

20

【請求項 12】

各非同一の核酸は、20% ~ 85% の GC 塩基含有量を有する、請求項 11 に記載の核酸ライブラリ。

【請求項 13】

少なくとも 2000 の非同一の核酸の少なくとも 90% は、核酸ライブラリ中の非同一の核酸の各々について平均頻度の 2 倍以内の量で核酸ライブラリ中にそれぞれ存在する、請求項 11 に記載の核酸ライブラリ。

【請求項 14】

各非同一の核酸は最大で 200 の塩基の長さを含む、請求項 11 に記載の核酸ライブラリ。

30

【請求項 15】

少なくとも 2000 の非同一の核酸が、生物学的経路において、遺伝子を標的とする sgRNA 配列をコードする非同一の核酸を含む、請求項 11 に記載の核酸ライブラリ。

【請求項 16】

少なくとも 2000 の非同一の核酸が、全ゲノムにおいて、遺伝子を標的とする sgRNA 配列をコードする非同一の核酸を含む、請求項 11 に記載の核酸ライブラリ。

【請求項 17】

各非同一の核酸は DNA 分子または RNA 分子を含む、請求項 11 に記載の核酸ライブラリ。

40

【請求項 18】

アンブリコンライブラリであって、該アンブリコンライブラリは複数の非同一の DNA 分子を含み、各非同一の DNA 分子は増幅産物の集団中に存在し、各非同一の DNA 分子は異なる gRNA 配列をコードし、各 gRNA 配列は真核生物の遺伝子に相補的な標的ドメインを含み、および複数の非同一の DNA 分子の少なくとも 80% は、アンブリコンライブラリ中の非同一の DNA 分子の各々について平均頻度の 2 倍以内の量でアンブリコンライブラリにそれぞれ存在する、アンブリコンライブラリ。

【請求項 19】

各非同一の DNA 分子は、30% ~ 70% の GC 塩基含有量を有する、請求項 18

50

に記載のアンプリコンライブラリ。

【請求項 20】

gRNA は単一の gRNA または二重の gRNA である、請求項 18 に記載のアンプリコンライブラリ。

【発明の詳細な説明】

【技術分野】

【0001】

本出願は、2016 年 11 月 9 日に出願された米国仮出願第 62 / 419, 881 号；2016 年 10 月 21 日に出願された米国仮出願第 62 / 411, 388 号；2016 年 9 月 13 日に
10 出願された米国仮出願第 62 / 393, 948 号；2016 年 8 月 31 日に
出願された米国仮出願第 62 / 382, 191 号；および、2016 年 8 月 22 日に
出願された米国仮出願第 62 / 378, 134 号の利益を主張し、その各々は、参照によって
その全体が本明細書に組込まれる。

【0002】

配列表

【0003】

本出願は、ASCII フォーマットで電子的に提出され、参照によってその全体が本明
細書に組み込まれる配列表を含んでいる。2017 年 4 月 4 日に作成された上記 ASCII
I のコピーは、44854 - 727__601__SL.txt とファイル名であり、13,
567 バイトのサイズである。
20

【背景技術】

【0004】

合成生物学の基礎は設計、構築、および試験プロセス、すなわち、これらの特定用途向
けの経路と生物体の迅速かつ手軽な生成と最適化のために DNA がアクセス可能であるこ
とを要求する反復するプロセスである。設計段階では、DNA を構成する A、C、T、お
よび G のヌクレオチドは、対象の領域を含む様々な配列へと組み立てられ、各配列変異体
が試験される特定の仮説を表す。こうした変異体配列は、進化生物学から始まり、遺伝子
、ゲノム、トランスクリプトーム、およびプロテオームを構築する配列の全体性に関係す
る概念である、配列空間の部分集合を表す。標的ゲノム編集の文脈において、遺伝子、遺
伝子クラスタ、経路、または全ゲノムの酵素編集を特異的に方向付けるための、非常に正
30 確かつ均一な核酸ライブラリの迅速な生成への必要性が存在する。

【0005】

< 引用による組み込み >

本明細書で言及される全ての刊行物、特許、および特許出願は、あたかも個々の刊行物
、特許、または特許出願が参照により組み込まれるように具体的かつ個々に指示される程
度に、参照により本明細書に組み込まれる。

【発明の概要】

【0006】

本明細書で提供されるのは、非常に正確な核酸ライブラリの効率的なデノボ合成および
スクリーニングのためのシステム、方法、および組成物である。本明細書で記載される核
酸ライブラリは、特異的に遺伝子、遺伝子クラスタ、生物学的経路、または全ゲノムを標
40 的とし編集するための核酸を含む。

【0007】

本明細書で提供されるのは核酸ライブラリであり、ここで、核酸ライブラリは、少なく
とも 500 の非同一の DNA 分子を含み、各非同一の DNA 分子は異なる gRNA 配列を
コードし、および、少なくとも 500 の非同一の DNA 分子の少なくとも約 80 % が、ラ
イブラリ中の非同一の DNA 分子の各々について平均頻度の 2 倍以内の量で核酸ライブラ
リにそれぞれ存在する。本明細書で提供されるのは、各々の非同一の DNA 分子が約 20
% から約 85 % の GC 塩基含有量を有する核酸ライブラリである。本明細書で提供され
るのは、各々の非同一の DNA 分子が約 30 % から約 70 % の GC 塩基含有量を有する核酸
50

ライブラリである。本明細書で提供されるのは、核酸ライブラリであり、ここで、少なくとも500の非同一のDNA分子の少なくとも約90%は、ライブラリ中の非同一のDNA分子の各々について平均頻度の2倍以内の量で核酸ライブラリにそれぞれ存在する。本明細書で提供されるのは、核酸ライブラリであり、ここで、少なくとも500の非同一のDNA分子の少なくとも99%は、ライブラリ中の非同一のDNA分子の各々について平均頻度の2倍以内の量で核酸ライブラリにそれぞれ存在する。本明細書で提供されるのは、少なくとも500の非同一のDNA分子が少なくとも2000の非同一のDNA分子を含む核酸ライブラリである。本明細書で提供されるのは、少なくとも500の非同一のDNA分子は少なくとも3500の非同一のDNA分子を含む核酸ライブラリである。本明細書で提供されるのは、少なくとも500の非同一のDNA分子が少なくとも100,000の非同一のDNA分子を含む核酸ライブラリである。本明細書で提供されるのは、各非同一のDNA分子が最大200の塩基の長さを含む核酸ライブラリである。本明細書で提供されるのは、各非同一のDNA分子が約100~約200の塩基の長さを含む核酸ライブラリである。本明細書で提供されるのは、少なくとも500の非同一のDNA分子は、生物学的経路において、遺伝子を標的とするgRNA配列をコードする非同一のDNA分子を含む、核酸ライブラリである。本明細書で提供されるのは、少なくとも500の非同一のDNA分子は、全ゲノムにおいて、遺伝子を標的とするgRNA配列をコードする非同一のDNA分子を含む、核酸ライブラリである。本明細書で提供されるのは、gRNAは単一のgRNAまたは二重のgRNAである核酸ライブラリである。

10

【0008】

20

本明細書で提供されるのは、核酸ライブラリあり、ここで、核酸ライブラリは少なくとも2000の非同一の核酸を含み、各非同一の核酸は異なるsgRNA配列をコードし、各sgRNA配列は真核生物の遺伝子に相補的な標的ドメインを含み、および、少なくとも2000の非同一の核酸の少なくとも約80%は、ライブラリ中の非同一の核酸の各々について平均頻度の2倍以内の量で核酸ライブラリに存在する。本明細書で提供されるのは、各非同一の核酸が約20%から約85%のGC塩基含有量を有する核酸ライブラリである。本明細書で提供されるのは、各非同一の核酸が約30%から約70%のGC塩基含有量を有する核酸ライブラリである。本明細書で提供されるのは核酸ライブラリであり、少なくとも2000の非同一の核酸の少なくとも約90%が、ライブラリ中の非同一の核酸の各々について平均頻度の2倍以内の量で核酸ライブラリに存在する。本明細書で提供されるのは、核酸ライブラリであり、少なくとも2000の非同一の核酸の少なくとも99%が、ライブラリ中の非同一の核酸の各々について平均頻度の2倍以内の量で核酸ライブラリに存在する。本明細書で提供されるのは、各非同一の核酸が最大200の塩基の長さを含む核酸ライブラリである。本明細書で提供されるのは、各々非同一の核酸が約100~約200の塩基の長さを含む核酸ライブラリである。本明細書で提供されるのは、少なくとも2000の非同一の核酸が、生物学的経路において、遺伝子を標的とするsgRNA配列をコードする非同一の核酸を含む核酸ライブラリである。本明細書で提供されるのは、少なくとも2000の非同一の核酸が、全ゲノムにおいて、遺伝子を標的とするsgRNA配列をコードする非同一の核酸を含む核酸ライブラリである核酸ライブラリである。本明細書で提供されるのは、各非同一の核酸がDNAまたはRNA分子を含む核酸ライブラリである。

30

40

【0009】

本明細書で提供されるのは、アンブリコンライブラリであり、ここで、アンブリコンライブラリは複数の非同一のDNA分子を含み、各非同一のDNAは増幅産物の集団中に存在し、各非同一のDNA分子は異なるgRNA配列をコードし、および複数の非同一のDNA分子の少なくとも約80%は、ライブラリ中の非同一のDNA分子の各々について平均頻度の2倍以内の量でアンブリコンライブラリにそれぞれ存在する。本明細書で提供されるのは、各非同一のDNA分子が約30%から約70%のGC塩基含有量を有するアンブリコンライブラリである。本明細書で提供されるのは、gRNAが単一のgRNAまたは二重のgRNAであるアンブリコンライブラリである。

50

【 0 0 1 0 】

本明細書で提供されるのは、細胞ライブラリであり、ここで、細胞ライブラリは複数の細胞集団を含み、細胞集団の各々は異なる gRNA 配列をコードする DNA 分子を含み、各 gRNA 配列は、遺伝子を結合するための標的領域を含み、細胞集団の少なくとも 15 % が遺伝子の発現において少なくとも 2 倍の枯渇を有する。本明細書で提供されるのは、少なくとも 45 % の細胞集団が、遺伝子の発現において少なくとも 2 倍の枯渇を有する細胞ライブラリである。本明細書で提供されるのは、gRNA が単一の gRNA または二重の gRNA である細胞ライブラリである。本明細書で提供されるのは、複数の細胞集団が単一の遺伝子当たり、少なくとも 3 つの様々な gRNA 配列をコードする DNA 分子を含む細胞ライブラリである。本明細書で提供されるのは、複数の細胞集団が単一の遺伝子当たり、少なくとも 5 つの様々な gRNA 配列をコードする DNA 分子を含む細胞ライブラリである。本明細書で提供されるのは、複数の細胞集団が少なくとも 2000 の細胞集団を含む細胞ライブラリである。本明細書で提供されるのは、複数の細胞集団が生物学的経路において、gRNA 配列をコードする DNA 分子を含む細胞ライブラリである。本明細書で提供されるのは、複数の細胞集団が全ゲノムにおいて、gRNA 配列をコードする DNA 分子を含む細胞ライブラリである。本明細書で提供されるのは、ゲノムが、シロイヌナズナ、線虫、イヌ科、コナミドリムシ、ゼブラフィッシュ、細胞性粘菌、キイロショウジョウバエ、大腸菌、ヒト、アカゲザル、ハツカネズミ、アナウサギ、ドブネズミ、出芽酵母、またはイノシシである、細胞ライブラリである。本明細書で提供されるのは、細胞集団の各々が原核細胞を含む細胞ライブラリである。本明細書で提供されるのは、細胞集団の各々が真核細胞を含む細胞ライブラリである。本明細書で提供されるのは、細胞集団の各々が哺乳類細胞を含む細胞ライブラリである。本明細書で提供されるのは、細胞集団の各々が外因性のヌクレアーゼ酵素をさらに含む細胞ライブラリである。本明細書で提供されるのは、DNA 分子がベクター配列をさらに含む細胞ライブラリである。

10

20

【 0 0 1 1 】

本明細書で提供されるのは、細胞ライブラリであり、ここで、細胞ライブラリは複数の細胞集団を含み、細胞集団の各々は異なる gRNA 配列をコードする DNA 分子を含み、各 gRNA 配列は、遺伝子を結合するための標的領域を含み、および、細胞集団の最大で 20 % が遺伝子の発現においてゼロまたは負の枯渇を有する。本明細書で提供されるのは、gRNA が単一の gRNA または二重の gRNA である細胞ライブラリである。本明細書で提供されるのは、複数の細胞集団が単一の遺伝子当たり、少なくとも 3 つの様々な gRNA 配列をコードする DNA 分子を含む細胞ライブラリである。本明細書で提供されるのは、複数の細胞集団が単一の遺伝子当たり、少なくとも 5 つの様々な gRNA 配列をコードする DNA 分子を含む細胞ライブラリである。本明細書で提供されるのは、複数の細胞集団が少なくとも 2000 の細胞集団を含む細胞ライブラリである。本明細書で提供されるのは、複数の細胞集団が少なくとも 10000 の細胞集団を含む細胞ライブラリである。

30

【 0 0 1 2 】

本明細書で提供されるのは、gRNA ライブラリを合成するための方法であって：その方法は、少なくとも 500 の非同一の DNA 分子のために予め決められた配列を提供する工程であって、各非同一の DNA 分子が gRNA をコードする工程；少なくとも 500 の非同一の DNA 分子を合成する工程；および、gRNA ライブラリを生成するために少なくとも 500 の非同一の DNA 分子を転写する工程であって、ここで、gRNA ライブラリ中の gRNA の少なくとも約 75 % が、少なくとも 500 の非同一の DNA 分子について予め決められた配列と比較して誤りがない工程を含む、方法。本明細書で提供されるのは、gRNA ライブラリを合成するための方法であって、ここで、転写工程の前に、少なくとも 500 の非同一の DNA 分子を細胞に移す工程をさらに含む。本明細書で提供されるのは、gRNA ライブラリを合成するための方法であって、ここで、少なくとも 500 の非同一の DNA 分子によってコードされた gRNA の少なくとも 96 % が、gRNA ライブラリ中に存在する。本明細書で提供されるのは、gRNA ライブラリを合成するため

40

50

の方法であって、ここで、gRNAライブラリ中のgRNAの少なくとも87%が、少なくとも500の非同一のDNA分子について予め決められた配列と比較して誤りがない。本明細書で提供されるのは、少なくとも500の非同一のDNA分子をベクターに挿入する工程をさらに含む、gRNAライブラリを合成するための方法である。本明細書で提供されるのは、少なくとも500の非同一のDNA分子を生物の細胞に移す工程をさらに含む、gRNAライブラリを合成するための方法である。本明細書で提供されるのは、生物が、シロイヌナズナ、線虫、イヌ科、コナミドリムシ、ゼブラフィッシュ、細胞性粘菌、キイロショウジョウバエ、大腸菌、ヒト、アカゲザル、ハツカネズミ、アナウサギ、ドブネズミ、出芽酵母、またはイノシシである、gRNAライブラリを合成するための方法である。本明細書で提供されるのは、各非同一のDNA分子が単一のgRNAまたは二重のgRNAをコードする、gRNAライブラリを合成するための方法である。

10

【0013】

本明細書で提供されるのは、gRNAライブラリを合成するための方法であって：その方法は、複数の非同一のDNA分子について予め決められた配列を提供する工程であって、ここで、各非同一のDNA分子がgRNAをコードする工程；核酸伸長反応のための遺伝子座のクラスタを含む表面を提供する工程；各非同一のDNA分子が表面から伸長している、複数の非同一のDNA分子を合成する工程；および、複数の非同一のDNA分子を細胞に移す工程を含む、方法。本明細書で提供されるのは、各々のクラスタが約50～約500の遺伝子座を含む、gRNAライブラリを合成するための方法である。本明細書で提供されるのは、各非同一のDNA分子が最大約200の塩基の長さを含む、gRNAライブラリを合成するための方法である。本明細書で提供されるのは、各非同一のDNA分子が単一のgRNAまたは二重のgRNAをコードする、gRNAライブラリを合成するための方法である。本明細書で提供されるのは、細胞が原核細胞である、gRNAライブラリを合成するための方法である。本明細書で提供されるのは、細胞が真核細胞である、gRNAライブラリを合成するための方法である。本明細書で提供されるのは、真核生物が哺乳動物細胞である、gRNAライブラリを合成するための方法である。本明細書で提供されるのは、細胞の各々が外因性のヌクレアーゼ酵素である、gRNAライブラリを合成するための方法である。

20

【図面の簡単な説明】

【0014】

30

【図1A】以下の構成要素を含んでいるクラスタ化して規則的な配置の短い回文配列リポート(CRISPR)複合体を例示する：PAM、標的配列、CAS9酵素、ガイドRNA(gRNA)、およびドナーDNA。

【図1B】以下の構成要素を含んでいるクラスタ化して規則的な配置の短い回文配列リポート(CRISPR)複合体を例示する：非相同末端結合修復(NHEJ)経路のためのPAM、標的配列、CAS9酵素、ガイドRNA(gRNA)、およびドナーDNA。

【図2】gRNAライブラリの設計、合成、クローニング、パッケージング、スクリーニング、および分析を含む、gRNAライブラリ・スクリーニング・ワークフローを例示する。

【図3】次のものを含むライブラリを構築するための、gRNAライブラリ・スクリーニング・ワークフローを例示する：アレイ上のオリゴ核酸ライブラリを合成、オリゴ核酸を増幅およびベクター内への移動、およびgRNA発現のための発現ライブラリの形成。

40

【図4A】様々なgRNAの図である。塩基対合領域、dCas9ハンドル(dCas9 handle)、および化膿性連鎖球菌の終結領域を有する、sgRNA配列(SEQ ID NO: 40)の図である。

【図4B】様々なgRNAの図である。sgRNAのみの図である。

【図4C】様々なgRNAの図である。dgRNAのみの図である。

【図5A】鋳型鎖標的構成におけるsgRNA配列の図である。

【図5B】非鋳型鎖標的構成におけるsgRNA配列の図である。

【図6A】転写される時、ヘアピン二次構造(hairpin secondary s

50

structure)を形成するgRNA配列をもたらす、T7プロモーターを有するgRNA配列の図である。

【図6B】転写される時、ヘアピン二次構造を形成しないgRNA配列をもたらす、T7プロモーターを有するgRNA配列の図である。

【図7】標的DNAのインビトロのCas9媒介切断のためのワークフローを示す。

【図8】コンピューターシステムの例を例示する。

【図9】コンピューターシステムの例示的構造を示すブロック図である。

【図10】複数のコンピューターシステム、複数の携帯電話および個人用携帯情報端末、ならびにネットワーク接続ストレージ(NAS)を組み込むように構成されたネットワークを実証する図である。

10

【図11】共有仮想アドレスメモリ空間を用いる、マルチプロセッサコンピューターシステムのブロック図である。

【図12】4つのsgRNA設計を示す。図12は、出現順に、SEQ ID NO 20、15、21、15、22、15、41、15、42、および43をそれぞれ開示する。

【図13】X軸上にヌクレオチド塩基およびY軸上に蛍光性の単位を有する、BioAnalyzer読み取り値のプロットである。

【図14A】X軸上にヌクレオチド塩基およびY軸上に蛍光性の単位を有する、BioAnalyzer読み取り値のプロットである。

【図14B】X軸上にヌクレオチド塩基およびY軸上に蛍光性の単位を有する、BioAnalyzer読み取り値のプロットである。

20

【図14C】X軸上にヌクレオチド塩基およびY軸上に蛍光性の単位を有する、BioAnalyzer読み取り値のプロットである。

【図14D】X軸上にヌクレオチド塩基およびY軸上に蛍光性の単位を有する、BioAnalyzer読み取り値のプロットである。

【図14E】X軸上にヌクレオチド塩基およびY軸上に蛍光性の単位を有する、BioAnalyzer読み取り値のプロットである。

【図14F】X軸上にヌクレオチド塩基およびY軸上に蛍光性の単位を有する、BioAnalyzer読み取り値のプロットである。

【図14G】X軸上にヌクレオチド塩基およびY軸上に蛍光性の単位を有する、BioAnalyzer読み取り値のプロットである。

30

【図14H】X軸上にヌクレオチド塩基およびY軸上に蛍光性の単位を有する、BioAnalyzer読み取り値のプロットである。

【図14I】X軸上にヌクレオチド塩基およびY軸上に蛍光性の単位を有する、BioAnalyzer読み取り値のプロットである。

【図14J】X軸上にヌクレオチド塩基およびY軸上に蛍光性の単位を有する、BioAnalyzer読み取り値のプロットである。

【図15】256のクラスタの画像であり、各クラスタは、そこから伸長するオリゴ核酸を有する121の遺伝子座を有する。

【図16A】各クラスタが121のオリゴ核酸を有する、240のクラスタからの29,040の特有のオリゴ核酸の合成からのプレートにわたるオリゴ核酸の表示(オリゴ核酸の頻度vs吸収度)のプロットである。

40

【図16B】ボックスによって同定された対照クラスタ(control clusters)を有する各個々のクラスタにわたるオリゴ核酸頻度vs吸収度の測定値のプロットである。

【図17】4つの個々のクラスタにわたる、オリゴ核酸頻度vs吸収度の測定値のプロットである。

【図18A】各クラスタが121のオリゴ核酸を有する、240のクラスタからの29,040の特有のオリゴ核酸の合成からのプレートにわたる、エラー率vs頻度のプロットである。

50

【図 18B】ボックスによって同定された対照クラスタを有する各個々のクラスタにわたる、オリゴ核酸エラー率 v s 頻度の測定値のプロットである。

【図 19】4つのクラスタにわたる、オリゴ核酸エラー率 v s 頻度の測定値のプロットである。

【図 20】オリゴ核酸当たりのパーセントの測定としての GC 含有量 v s オリゴ核酸の数のプロットである。

【図 21】2つの様々なポリメラーゼでの PCR からの結果のプロットを提供する。各図は、「観察頻度」(100, 000 当たりの測定された 0 ~ 35) v s オリゴ核酸の数 (0 ~ 2000) を示す。

【図 22】記録された増幅後のオリゴ核酸の集団均一性の定量化を伴う図を提供する。

10

【図 23】配列ドロップアウト (sequence dropouts) に対する過度の増幅の影響のプロットを示す。

【図 24】10, 000 の sgRNA オリゴ核酸の CRISPR ライブラリからの、回復されたオリゴ核酸の配列決定からの結果を示す。

【図 25】101, 000 の sgRNA オリゴ核酸の CRISPR ライブラリからの、回復されたオリゴ核酸の配列決定からの結果を示す。

【図 26A】少なくとも 2 倍の枯渇を伴う sgRNA のパーセンテージのグラフを示す。

【図 26B】ゼロまたは負の枯渇を伴う sgRNA のパーセンテージのグラフを示す。

【発明を実施するための形態】

【0015】

20

本明細書で提供されるのは、非常に正確なガイド RNA (「gRNA」) ライブラリの効率的な合成およびスクリーニングのための、システム、方法、および組成物である。本明細書にて記載されるデノボ合成方法は、標的遺伝子編集用の酵素システムに取り込むために、gRNA の大規模なライブラリの迅速で非常に正確な生成を提供する。

【0016】

定義

本開示の全体にわたって、様々な実施形態は範囲形式で提示される。範囲形式での記載は単に利便性と簡潔さのためのものに過ぎず、任意の実施形態の範囲に対する確固たる限定として解釈されてはならないということを理解されたい。これに応じて、範囲の記載は、文脈で別段の定めのない限り、すべての可能性のある下位範囲と、下限の単位の小数第 2 位までのその範囲内の個々の数値を具体的に開示したと考えられなければならない。例えば、1 乃至 6 などの範囲の記載は、1 乃至 3、1 乃至 4、1 乃至 5、2 乃至 4、2 乃至 6、3 乃至 6 などの下位範囲と、例えば、1.1、2.2、3.3、5.5、および 5.9 のその範囲内の個々の数値を具体的に開示したと考えられなければならない。これは、範囲の広さにかかわらず適用される。これらの介入する範囲の上限および下限は、より小さな範囲内に独立して含まれてもよく、また、定められた範囲内のあらゆる具体的に除外された限度に従って、本発明内に包含される。定められた範囲が上限および下限の 1 つ又はその両方を含む場合、これらの含まれた上限および下限のいずれかまたは両方を除く範囲もまた、文脈が明らかに他に指示しない限り、本発明内に包含される。

30

【0017】

40

本明細書で使用される用語は、特定の実施形態のみを記載するためにあり、任意の実施形態を制限することを意図していない。本明細書で使用されるように、単数形「1つ (a)」、「1つ (an)」、および「その (the)」は、文脈上他の意味を明白に示すものでない限り、同様に複数形を含むことを意図している。用語「含む (comprises)」および/または「含むこと (comprising)」は、本明細書で使用されるとき、明示される特徴、整数、工程、操作、要素、および/または構成要素の存在を指定するが、1つ以上の他の特徴、整数、工程、操作、要素、構成要素、そのおよび/またはその群の存在や添加を妨げない。本明細書で使用されるように、用語「および/または」は、関連する列挙された項目の 1 つ以上のあらゆる組み合わせを含む。

【0018】

50

別段の定めのない限り、あるいは文脈から明らかでない限り、本明細書で使用されるように、数あるいは数の範囲に関連して用語「約」とは、明示された数とその数 + / - 10 %、あるいはある範囲の列挙された値について列挙された下限の 10 % 以下と列挙された 10 % 以上を意味する。

【0019】

本明細書において言及される用語「gRNA」は、ガイドRNAは配列を指し、単一および二重のガイドRNA配列の両方を包含する。本明細書では、具体的に明示されるか文脈から明らかでない限り、本明細書において言及される用語「dgRNA」は二重のガイドRNA配列を指す：部分的に相補的RNAである、crRNA（標的配列への相補的なシード領域を含むスペーサー配列）および個別のtracrRNA（トランス活性化配列）。本明細書では、具体的に明示されるか文脈から明らかでない限り、本明細書において言及される用語「sgRNA」は、縮合crRNAおよびtracrRNAの両方を含む単一ガイドRNA配列を指す。

10

【0020】

本明細書では、具体的に明示されるか文脈から明らかでない限り、用語「オリゴ核酸」および「核酸」は、一本鎖の分子と同様に、二本または三本鎖の核酸も包含する。二本または三本鎖の核酸において、核酸鎖は同一の広がりをもつ必要はない（つまり、二本鎖の核酸は、両方の鎖の全長に沿って二本鎖である必要はない）。特に明記されない限り、核酸配列は提供される時、5'から3'の方向へと列挙される。本明細書に記載される方法は、単離された核酸の生成を提供する。本明細書に記載される方法はさらに、単離精製された核酸の生成を提供する。本明細書で言及される「オリゴ核酸」および「核酸」は、少なくとも、5、10、20、30、40、50、60、70、80、90、100、125、150、175、200、225、250、275、300、325、350、375、400、425、450、475、500、600、700、800、900、1000、1100、1200、1300、1400、1500、1600、1700、1800、1900、2000、またはより多くの塩基の長さを含むことができる。

20

【0021】

本明細書では、具体的に明示されるか文脈から明らかでない限り、本明細書で言及される用語「アンプリコン」は、増幅反応物を指す。

【0022】

本開示は、別段の定めのない限り、当該技術の範囲内である従来の分子生物学的技術を採用する。別段の定めのない限り、本明細書で使用されるすべての技術的および科学的な用語は、これらの発明が属する分野の当業者によって一般に理解されるのと同じ意味を有する。

30

【0023】

gRNAライブラリ・スクリーニング

本明細書で提供されるのは、クラスタ化して規則的な配置の短い回文配列リピート（CRISPR）- 酵素複合体中での取り込みのための、非常に正確なgRNAライブラリを設計、構築、およびスクリーニングするための方法である。例えば、図1A-ABを参照する。本明細書に記載された方法を用いて生成されたgRNAライブラリは、sgRNAライブラリおよびdgRNAライブラリの両方を含む。本明細書で提供されるのは、結果としてもたらされるライブラリで、予め決められたgRNAの高度な表示をもたらす非常に均一な合成のための方法である。設計段階では、gRNAが設計される。図2を参照する。設計戦略は、遺伝子に及ぶgRNAの設計が挙げられるが、これに限定されない。所望のワークフローに応じて、デノボ合成された核酸はDNAまたはRNA塩基である。

40

【0024】

デノボ合成されたDNAの場合、核酸を含むライブラリは合成され、ここで合成された各核酸は、転写産物としてのgRNA（例えばsgRNA）配列をコードするDNA配列である。いくつかの例では、合成された核酸はその後、発現ベクターへと挿入される。1つの例示的なワークフローでは、合成された核酸はウイルスベクターへと挿入され、細胞

50

へと形質導入するためにパッケージされ、続いてスクリーニングおよび分析される。図2を参照。典型的な細胞は、限定することなく、原核細胞および真核細胞を含む。典型的な真核細胞は、限定することなく、動物、植物および真菌の細胞を含む。典型的な動物細胞は、限定することなく、昆虫、魚、および哺乳動物の細胞を含む。典型的な哺乳動物細胞は、マウス、ヒト、および霊長類の細胞を含む。試験される典型的な細胞機能は、限定することなく、細胞増殖における変化、遊走/接着、代謝活性、および細胞シグナル伝達活性を含む。デノボ合成されたRNAの場合、gRNAそれ自身が合成され、細胞へのトランスフェクションなどの下流の適用に利用可能である。

【0025】

オリゴ核酸は、アレイ(301)上で伸長するために、クラスタ(303)の位置(「遺伝子座」)内で合成されることができる。図3を参照する。そのような構成は、クラスタ化された遺伝子座構成のないプレート全体にわたって合成されたオリゴ核酸の増幅産物と比較した時、合成されたオリゴ核酸のオリゴ核酸からの改善された産物のオリゴ核酸表示 - “アンプリコン”と呼ばれる - を提供する。いくつかの例では、一般的に「ドリフト」と呼ばれる、重度のGC含有量を有するオリゴ核酸を有する多くのオリゴ核酸集団の繰り返された合成により、増幅反応物におけるGCの低いまたはGCの高いアンプリコンの提示不足により、単一のクラスタ内で合成されたオリゴ核酸の増幅(310)は、表示における負の効果に対抗する。いくつかの例では、本明細書に記載される単一のクラスタは、50 - 1000、75 - 900、100 - 800、125 - 700、150 - 600、200 - 500、または300 - 400の別々の遺伝子座を含む。いくつかの例では、単一のクラスタは、50 - 500別々の遺伝子座を含む。いくつかの例では、遺伝子座は、スポット、ウェル、マイクロウェル、チャンネル、またはポストである。いくつかの例では、各クラスタは、同一の配列を有するオリゴ核酸の伸長を支持する別個の特徴の、少なくとも1倍、2倍、3倍、4倍、5倍、6倍、7倍、8倍、9倍、10倍、またはより多くの冗長性を有する。

【0026】

本明細書で提供されるのは、発現ベクターに挿入するためのgRNAライブラリである。図3におけるワークフローを続けて見ると、アレイ(301)は、オリゴ核酸合成および伸長のための遺伝子座の多数のクラスタ(303)を含む。デノボDNAは合成されプレートから取り除かれて、オリゴ核酸の集団(305)(例えばsgRNAをコードするDNA)を形成し、これは、増幅(310)を受けてベクター(330)の挿入のための増幅したオリゴ核酸のライブラリ(320)を形成し、合成されたDNA(335)を含むベクターのライブラリを形成する。一旦細胞内に入ると、DNAはgRNA(例えばsgRNA)に転写され、ゲノム編集レジーム(genomic editing regime)(例えば、Cas9に基づいたシステム)との結合に利用可能となる。細胞は、編集酵素(例えばCas9)の自然発現または異所的発現を有してもよい。編集酵素(例えばCas9)は、二本鎖DNA切断活性、あるいは、ニッキング、塩基交換、または配列交換活性などの修飾された活性を有してもよい。ベクターへの挿入のために合成されたDNAは、sgRNA、dgRNA、またはそれらの断片を含み得る。

【0027】

本明細書に開示される核酸ライブラリを挿入するための発現ベクターは、真核生物または原核生物の発現ベクターを含む。典型的な発現ベクターは、限定することなく、哺乳動物の発現ベクター：pSF-CMV-NEO-NH2-PPT-3XFLAG、pSF-CMV-NEO-COOH-3XFLAG、pSF-CMV-PURO-NH2-GST-TEV、pSF-OBX20-COOH-TEV-FLAG(R)-6His、pCEP4 pDEST27、pSF-CMV-Ub-KrYFP、pSF-CMV-FMDV-daGFP、pEF1a-mCherry-N1ベクター、pEF1a-tdTomatoベクター、pSF-CMV-FMDV-Hygro、pSF-CMV-PGK-Puro、pMCP-タグ(m)、およびpSF-CMV-PURO-NH2-CMYC；細菌性の発現ベクター：pSF-OBX20-BetaGal、pSF-OBX20-Fl

uc、pSF-OB20、およびpSF-Tac；植物の発現ベクター：pRI 101-AN DNAおよびpCambia2301；および、酵母発現ベクター：pTYB21およびpKLAC2、および昆虫のベクター：pAc5.1/V5-His AおよびpDEST8を含む。

【0028】

本明細書に記載される方法によって合成されたデノボオリゴ核酸ライブラリは、細胞内で発現され得る。いくつかの例では、細胞は疾患状態に関連している。例えば、疾患状態に関連付けられる細胞として、これらに限定されないが、被験体からの細胞株、組織サンプル、初代細胞、被験体からの増殖された培養細胞、またはモデルシステムにおける細胞が挙げられる。いくつかの例では、モデルシステムは、植物または動物のシステムである。いくつかの例では、デノボオリゴ核酸ライブラリは、細胞活動における変化について評価するために細胞内で発現され得る。典型的な細胞活性として、これらに限定されないが、増殖、周期進行、細胞死、接着、遊走、繁殖、細胞シグナル伝達、エネルギー産生、酸素利用、代謝活性、および老化、遊離基損傷に対する反応、またはそれらの任意の組み合わせが挙げられる。

【0029】

本明細書で提供されるのは、遺伝子あたり複数の非同一のgRNAを含むgRNAライブラリ（または、転写された時にgRNAライブラリをもたらすDNAライブラリ）を合成する方法である。gRNAは、sgRNAまたはdgRNAをコードし得る。いくつかの例では、gRNAライブラリは、遺伝子あたり少なくとも、3、4、5、6、7、8、9、10、または10よりも多い非同一のgRNAを含む。いくつかの例では、gRNAライブラリは、1以上の遺伝子を標的にする。いくつかの例では、gRNAライブラリは、約1、5、10、20、30、40、50、60、70、80、90、100、または100よりも多い遺伝子を標的にする。いくつかの例では、gRNAライブラリは、約1-100、2-95、5-90、10-85、15-80、20-75、25-70、30-65、35-60、または40-50の遺伝子を標的にする。いくつかの例では、本明細書に記載されるgRNAライブラリは、経路内の遺伝子を標的にする。典型的な経路として、これらに限定されないが、代謝、細胞死、細胞周期進行、免疫細胞活性化、炎症反応、血管新生、リンパ球新生、低酸素ストレス応答、酸化ストレス応答、または細胞接着、および細胞遊走経路が挙げられる。

【0030】

本明細書で記載されるようなgRNAライブラリを合成するための方法は、ゲノムの一部への相補的な塩基対合領域、ゲノム標的領域を有する非同一のgRNAの合成を提供し得る。ゲノム標的領域は、エクソン、イントロン、コーディング、または非コーディングの配列を含んでもよい。いくつかの例では、gRNAライブラリは、全ゲノム中の遺伝子の少なくともまたは約5%に相補的な塩基対合領域を相対的に有する非同一のgRNAを含む。いくつかの例では、gRNAライブラリは、全ゲノム中の遺伝子の少なくともまたは約80%に相補的な塩基対合領域を相対的に有する非同一のgRNAを含む。いくつかの例では、gRNAライブラリは、全ゲノム中の遺伝子の少なくともまたは約90%に相補的な塩基対合領域を相対的に有する非同一のgRNAを含む。いくつかの例では、gRNAライブラリは、全ゲノム中の遺伝子の少なくともまたは約95%に相補的な塩基対合領域を相対的に有する非同一のgRNAを含む。いくつかの例では、gRNAライブラリは、全ゲノム中の遺伝子の少なくともまたは約100%に相補的な塩基対合領域を相対的に有する非同一のgRNAを含む。

【0031】

本明細書で提供されるのは、様々な細胞にわたる遺伝子の少なくとも2倍の枯渇を有するgRNAをもたらす、本明細書に記載される方法によって合成されたgRNAライブラリである。いくつかの例では、gRNAライブラリは、細胞内または複数の細胞集団内に存在する時に、少なくともあるいは、約10%、12%、15%、16%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、またはより多

くの少なくとも2倍の遺伝子の枯渇を提供するgRNAを含む。いくつかの例では、遺伝子は必須遺伝子、つまり、細胞生存のために重大な遺伝子である。例示的な必須遺伝子は、これらに限定されないが、PCNA、PSMA7、RPP21、およびSF3B3が挙げられる。いくつかの例では、gRNAライブラリは、細胞内に存在する時に、少なくとも2倍、3倍、4倍、5倍、6倍、または6倍よりも多い遺伝子の枯渇を提供するgRNAを含む。いくつかの例では、gRNAライブラリは、細胞内または複数の細胞集団内に存在する時に、最大で5%、10%、12%、15%、または20%の遺伝子のゼロまたは負の枯渇を有するgRNAを含む。いくつかの例では、複数の細胞集団は、少なくともまたは約2000、3000、4000、5000、6000、7000、8000、9000、10000、12000、15000、20000、25000、30000、または30000よりも多い細胞集団を含む。いくつかの例では、gRNAライブラリは、複数の遺伝子について、少なくとも2倍、3倍、4倍、5倍、6倍、または6倍よりも多い枯渇を有するgRNAを含む。いくつかの例では、gRNAライブラリは、複数の遺伝子について少なくとも2倍の枯渇を提供する平均で少なくともまたは約40%、45%、50%、60%、65%、70%、75%、80%、または90%よりも多いgRNAを含む。そのような遺伝子欠失プロファイルを提供するgRNAは、sgRNAまたはdgRNAでありえる。

10

【0032】

本明細書で提供されるのは、オリゴ核酸の非常に均一なライブラリを合成する方法である。場合によっては、90%よりも多い合成されたオリゴ核酸(RNAまたはDNA)が、核酸ライブラリのためのオリゴ核酸表示について、平均の4倍以内で表わされる。場合によっては、90%よりも多いオリゴ核酸が、ライブラリのためのオリゴ核酸表示について平均の2倍以内で表わされる。場合によっては、90%よりも多いオリゴ核酸が、ライブラリのためのオリゴ核酸表示について平均の1.5倍以内で表わされる。場合によっては、80%よりも多いオリゴ核酸が、ライブラリのためのオリゴ核酸表示について平均の1.5倍以内で表わされる。

20

【0033】

本明細書に記載される方法によってデノボ合成されたオリゴ核酸ライブラリは、予め決められた配列と比較して、高い割合の正確な配列を含む。いくつかの例では、本明細書に開示されるデノボオリゴ核酸ライブラリは、オリゴ核酸について予め決められた配列と比較して、70%を超える正確な配列を有する。いくつかの例では、本明細書に開示されるデノボオリゴ核酸ライブラリは、オリゴ核酸について予め決められた配列と比較して、75%を超える正確な配列を有する。いくつかの例では、本明細書に開示されるデノボオリゴ核酸ライブラリは、オリゴ核酸について予め決められた配列と比較して、80%を超える正確な配列を有する。いくつかの例では、本明細書に開示されるデノボオリゴ核酸ライブラリは、オリゴ核酸について予め決められた配列と比較して、85%を超える正確な配列を有する。いくつかの例では、本明細書に開示されるデノボオリゴ核酸ライブラリは、オリゴ核酸について予め決められた配列と比較して、90%を超える正確な配列を有する。いくつかの例では、本明細書に開示されるデノボオリゴ核酸ライブラリは、オリゴ核酸について予め決められた配列と比較して、95%を超える正確な配列を有する。いくつかの例では、本明細書に開示されるデノボオリゴ核酸ライブラリは、オリゴ核酸について予め決められた配列と比較して、100%を超える正確な配列を有する。

30

40

【0034】

いくつかの例では、本明細書に開示されるデノボ合成されたオリゴ核酸ライブラリは、増幅反応後のオリゴ核酸についての予め決められた配列と比較して、70%を超える正確な配列を有する。いくつかの例では、本明細書に開示されるデノボ合成されたオリゴ核酸ライブラリは、増幅反応後のオリゴ核酸についての予め決められた配列と比較して、75%を超える正確な配列を有する。いくつかの例では、本明細書に開示されるデノボ合成されたオリゴ核酸ライブラリは、増幅反応後のオリゴ核酸についての予め決められた配列と比較して、80%を超える正確な配列を有する。いくつかの例では、本明細書に開示され

50

るデノボ合成されたオリゴ核酸ライブラリは、増幅反応後のオリゴ核酸についての予め決められた配列と比較して、85%を超える正確な配列を有する。いくつかの例では、本明細書に開示されるデノボ合成されたオリゴ核酸ライブラリは、増幅反応後のオリゴ核酸についての予め決められた配列と比較して、90%を超える正確な配列を有する。いくつかの例では、本明細書に開示されるデノボ合成されたオリゴ核酸ライブラリは、増幅反応後のオリゴ核酸についての予め決められた配列と比較して、95%を超える正確な配列を有する。いくつかの例では、本明細書に開示されるデノボ合成されたオリゴ核酸ライブラリは、増幅反応に続くオリゴ核酸のための予め決められた配列と比較すると、100%の正確な配列を有する。

【0035】

いくつかの例では、本明細書に開示されるデノボ合成されたオリゴ核酸ライブラリは、細胞内に移された時、オリゴ核酸についての予め決められた配列と比較して、80%を超える正確な配列をもたらす。いくつかの例では、本明細書に開示されるデノボ合成されたオリゴ核酸ライブラリは、細胞内に移された時、オリゴ核酸についての予め決められた配列と比較して、85%を超える正確な配列をもたらす。いくつかの例では、本明細書に開示されるデノボ合成されたオリゴ核酸ライブラリは、細胞内に移された時、オリゴ核酸についての予め決められた配列と比較して、90%を超える正確な配列をもたらす。いくつかの例では、本明細書に開示されるデノボ合成されたオリゴ核酸ライブラリは、細胞内に移された時、オリゴ核酸についての予め決められた配列と比較して、95%を超える正確な配列をもたらす。いくつかの例では、本明細書に開示されるデノボ合成されたオリゴ核酸ライブラリは、細胞内へ移された時、オリゴ核酸についての予め決められた配列と比較して、100%正確な配列をもたらす。

【0036】

いくつかの例では、本明細書に開示されるデノボ合成されたオリゴ核酸ライブラリは、細胞内へ移された時、80%を超える配列表示をもたらす。いくつかの例では、本明細書に開示されるデノボ合成されたオリゴ核酸ライブラリは、細胞内へ移された時、90%を超える配列表示をもたらす。いくつかの例では、本明細書に開示されるデノボ合成されたオリゴ核酸ライブラリは、細胞内へ移された時、95%を超える配列表示をもたらす。いくつかの例では、本明細書に開示されるデノボ合成されたオリゴ核酸ライブラリは、細胞内へ移された時、100%の配列表示をもたらす。

【0037】

本明細書に記載されるデノボオリゴ核酸ライブラリは、ポリメラーゼ酵素および増幅試薬（例えば緩衝剤、リン酸塩、およびデオキシヌクレオチド三リン酸）の添加を有する増幅反応を受けてもよい。いくつかの例では、デノボオリゴ核酸ライブラリは、少なくともまたは約6、8、10、15、20、または20よりも多いサイクルのPCRによって増幅される。いくつかの例では、デノボオリゴ核酸ライブラリは、約6~20、7~18、8~17、9~16または10~15のサイクルの範囲でPCRによって増幅される。いくつかの例では、デノボオリゴ核酸ライブラリは、約15のサイクルのPCRによって増幅される。

【0038】

いくつかの例では、デノボオリゴ核酸ライブラリの増幅は、DNA分子のアンプリコンライブラリを提供する。いくつかの例では、アンプリコンライブラリは、gRNA配列をコードする非同一の核酸を含む。いくつかの実例では、gRNA配列はsgRNAまたはdgRNAである。

【0039】

いくつかの例では、デノボオリゴ核酸ライブラリは非同一の核酸を含み、各非同一の核酸はDNA分子を含む。いくつかの例では、DNA分子の数は、約500、2000、3500、または3500よりも多い分子である。いくつかの例では、DNA分子の数は、少なくともまたは約、250、500、1000、1250、2000、2500、3000、3500、4000、4500、5000、5500、6000、6500、700

10

20

30

40

50

0、7500、8000、8500、9000、9500、10000、15000、20000、50000、100000、250000、750000、1000000、または1000000より多い分子である。いくつかの例では、DNA分子の数は最大で、250、500、1000、1250、2000、2500、3000、3500、4000、4500、5000、5500、6000、6500、7000、7500、8000、8500、9000、9500、10000、15000、20000、50000、100000、250000、500000、750000、1000000、または1000000よりも多い分子である。いくつかの例では、DNA分子はgRNA配列をコードする。いくつかの実例では、gRNA配列はsgRNAまたはdgRNAである。

10

【0040】

いくつかの例では、デノボオリゴ核酸ライブラリは非同一の核酸を含み、各非同一の核酸はRNA分子を含む。いくつかの例では、RNA分子の数が約2000の分子である。いくつかの例では、RNA分子の数は、少なくともまたは約、250、500、1000、1250、2000、2500、3000、3500、4000、4500、5000、5500、6000、6500、7000、7500、8000、8500、9000、9500、10000、15000、20000、50000、100000、250000、750000、1000000、または1000000よりも多い分子である。いくつかの例では、RNA分子の数は、最大で、250、500、1000、1250、2000、2500、3000、3500、4000、4500、5000、5500、6000、6500、7000、7500、8000、8500、9000、9500、10000、15000、20000、50000、100000、250000、500000、750000、1000000、または1000000よりも多い分子である。いくつかの例では、RNA分子はgRNA配列をコードする。いくつかの実例では、gRNA配列はsgRNAまたはdgRNAである。

20

【0041】

本明細書で提供されるのは、増幅後の高い均一性を有するデノボオリゴ核酸ライブラリである。いくつかの例では、本明細書に記載されるデノボオリゴ核酸ライブラリ内の80%より多くのオリゴ核酸は、増幅後のライブラリ全体について、平均表示の少なくとも約1.5倍以内で表わされる。いくつかの例では、本明細書に記載されるデノボオリゴ核酸ライブラリ内の90%より多くのオリゴ核酸は、増幅後のライブラリ全体について、平均表示の少なくとも約1.5倍以内で表わされる。いくつかの例では、本明細書に記載されるデノボオリゴ核酸ライブラリ内の80%より多くのオリゴ核酸は、増幅後のライブラリ全体について、平均表示の少なくとも約2倍以内で表わされる。いくつかの例では、本明細書に記載されるデノボオリゴ核酸ライブラリ内の80%より多くのオリゴ核酸は、増幅後のライブラリ全体について、平均表示の少なくとも約2倍以内で表わされる。

30

【0042】

本明細書に記載され方法を用いるデノボ合成されたオリゴ核酸の非増幅集団は、非同一のオリゴ核酸配列の数を変えることができる。いくつかの例では、非同一のオリゴ核酸配列の数は、約2000-1000000、3000-900000、4000-800000、5000-700000、6000-600000、7000-500000、8000-400000、9000-300000、10000-200000、11000-100000、12000-75000、14000-60000、および20000-50000の配列の範囲である。場合によっては、非同一のオリゴ核酸配列の数は、約50-2000、75-1800、100-1700、150-1600、200-1500、250-1400、300-1300、400-1200、500-1100、600-1000、700-900の配列の範囲である。いくつかの例では、非同一のオリゴ核酸配列の数は2000の配列である。いくつかの例では、非同一のオリゴ核酸配列の数は1000000よりも多い配列である。いくつかの例では、非同一のオリゴ核酸配列の数は、少なくとも、50、100、200、400、600、800、1000、1

40

50

200、1400、1600、1800、2000、3000、5000、7000、10000、20000、30000、50000、100000、500000、700000、1000000、10000000、100000000、1000000000、または10000000000より多くの配列である。いくつかの例では、非同一のオリゴ核酸配列の数は、最大で、50、100、200、400、600、800、1000、1200、1400、1600、1800、2000、3000、5000、7000、10000、20000、30000、50000、100000、500000、700000、1000000、または10000000より多くの配列である。いくつかの例では、非同一のオリゴ核酸配列の数は、最大で、50、100、200、400、600、800、1000、1200、1400、1600、1800、2000、3000、5000、7000、10000、20000、30000、50000、100000、500000、700000、および10000000の配列である。

10

【0043】

非増幅集団のオリゴ核酸は、様々な量で存在する。いくつかの例では、非増幅集団のオリゴ核酸は、少なくともまたは約0.25 fmol/mgの量で存在する。いくつかの例では、非増幅集団のオリゴ核酸は、少なくともまたは約1 fmol/mgの量で存在する。いくつかの例では、非増幅集団のオリゴ核酸は、少なくとも0.25、1、10、20、30、40、50、100、250、500、750、1000、または1000 fmol/mgよりも多い量で存在する。いくつかの例では、非増幅集団のオリゴ核酸は、最大で、0.25、1、10、20、30、40、50、100、250、500、750、および1000 fmol/mgの量で存在する。

20

【0044】

本明細書で提供されるのは、非同一のオリゴ核酸のライブラリを合成するための方法であり、非同一のオリゴ核酸の配列長または平均の配列長は様々である。場合によっては、非同一のオリゴ核酸の配列長または平均配列長は、最大で150塩基である。場合によっては、非同一のオリゴ核酸の配列長または平均配列長は、約100～約200の塩基の範囲である。いくつかの例では、非同一のオリゴ核酸の配列長または平均配列長は、少なくとも、30、50、100、125、150、200、250、300、350、400、450、500、あるいは500よりも多い塩基である。いくつかの例では、非同一のオリゴ核酸の配列長または平均配列長は、最大で、150、200、250の、300、350、400、450、あるいは500の塩基である。非同一のオリゴ核酸の例示的な配列長は、約25～約150、あるいは約50～約200の塩基の範囲である。場合によっては、非同一のオリゴ核酸の配列長または平均配列長は、約125～約200、あるいは約150～約200の塩基の範囲である。

30

【0045】

ガイドRNA配列

【0046】

本明細書で提供されるのは、特定の標的核酸配列にゲノム配列編集酵素（例えばCas9）を向けるための、単一ガイドRNA（sgRNA）配列である。Cas9酵素と複合したsgRNAの例が図4Aで例示され、単独の例が図4Bで例示される。sgRNAは、図4Cに例示されるような二重のガイドRNAであってもよい。本明細書に開示されるガイド配列は、塩基対合領域を含む。塩基対合領域は、標的配列に結合するためのシード領域、および随意にスペーサー領域を含む。塩基対合領域は、長さが変わってもよい。例えば、塩基対合領域は、約1～10、1～20、20～25、または1～30の塩基の長さを含んでもよい。いくつかの例において、塩基対合領域は、少なくとも、10、15、20、25、30、または30よりも多くの塩基の長さを含む。いくつかの例では、塩基対合領域は、少なくとも10の塩基の長さのシード領域を含む。シード領域は、約8～20の塩基の長さを含み得る。いくつかの例では、シード領域は約12の塩基の長さである。いくつかの例では、本明細書に記載される塩基対合領域は、転写中に鋳型鎖を標的とするように設計される（図5A）。いくつかの例では、本明細書に記載される塩基対合領域は、

40

50

転写中に非鋳型鎖を標的とするように設計される（図 5 B）。

【 0 0 4 7 】

いくつかの例では、s g R N A の塩基対合領域の 3' は、C a s 9 に結合するための C a s 9 ハンドル領域である。いくつかの例では、C a s 9 ハンドル領域は、d C a s 9 酵素に結合するための d C a s 9 ハンドル領域である。ハンドル領域は、長さが変わってもよい。例えば、ハンドル領域は、約 1 ~ 5 0、2 0 ~ 4 5、または 1 5 ~ 6 0 の塩基の長さを含み得る。いくつかの例において、ハンドル領域は、少なくとも、3 5、4 0、4 5、5 0、または 5 0 より多くの塩基の長さを含む。ハンドル領域は、約 4 2 の塩基の長さを含み得る。

【 0 0 4 8 】

いくつかの例では、s g R N A のハンドル領域の 3' は、終結領域である。いくつかの例では、終結領域は化膿性連鎖球菌の終結領域である。いくつかの例では、終結領域は約 4 0 の塩基の長さを含む。いくつかの例では、終結領域は約 1 0 ~ 5 0、2 0 ~ 6 0、または 3 0 ~ 5 5 の塩基の長さを含む。

【 0 0 4 9 】

本明細書に記載される g R N A 配列のための設計スキームは、g R N A 配列をコードする D N A の D N A 依存性 R N A ポリメラーゼ・プロモーター領域 5' 上流の包含を含み得る。例示的な D N A 依存性 R N A ポリメラーゼ・プロモーター領域として、これらに限定されないが、T 3 および T 7 R N A ポリメラーゼ・プロモーター配列が挙げられる。例えば、図 6 A は、T 7 プロモーター領域が g R N A の 5' 上流にあり、結果としての転写された g R N A が生成され、ここで、g R N A がヘアピンを含む構成を示す。いくつかの構成では、g R N A は、ヘアピン二次構造を形成する配列を欠くように設計される（図 6 B）。ヘアピン二次構造は、C a s 9 ハンドルおよび/または終結領域内で欠いていることもある。

【 0 0 5 0 】

本明細書で提供されるのは、ゲノム配列編集酵素（例えば C a s 9）を特定の標的核酸配列に向けるための d g R N A である。いくつかの例では、ライブラリは d g R N A のための配列をコードするオリゴ核酸配列を含む。いくつかの例では、ライブラリは核酸を含み、ここで、合成された各核酸は、転写産物としての d g R N A 配列をコードする D N A 配列である。いくつかの例では、ライブラリは核酸を含み、ここで、合成された各核酸は R N A 配列であり、d g R N A 自体は合成される。いくつかの例では、d g R N A ライブラリは、別個のオリゴ核酸として合成される c r R N A および t r a c r R N A のためのオリゴ核酸配列を含む。いくつかの例では、オリゴ核酸は、c r R N A および t r a c r R N A を別々にコードする。いくつかの例では、オリゴ核酸は、転写された時に個別の c r R N A 配列および個別の t r a c r R N A 配列をもたらす単一の配列をコードする。c r R N A および t r a c r R N A のための例示的な配列は、表 1 で見られる。

【 0 0 5 1 】

【表 1】

表 1.

SEQ ID NO	名称	配列
5	crRNA-sp2	5'ATAACTCAATTTGTAAAAAAGTTTTAGAGCTATGCTGTTTTG3'
6	tracrRNA	5'GGAACCATTCAAAACAGCATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAAGTGGC ACCGAGTCGGTGTCTTTTTT3'

【 0 0 5 2 】

本明細書に記載される g R N A ライブラリは、インビトロスクリーニングおよび分析の

ために使用され得る。そのような構成の例が図7で示され、標的の二本鎖のDNA配列がgRNA配列およびCas9酵素と共にインキュベートする。その混合物は、二重鎖DNA切断をもたらす。DNA切断は、ゲノム要素の機能または発現において測定可能な変化をもたらす。本明細書に記載されるgRNA、またはgRNAをコードするDNAは、当技術分野において既知の様々な方法で細胞に加えられ、これらの方法としては、限定されないが、トランスフェクション、形質導入、またはエレクトロポレーションが挙げられる。

【0053】

いくつかの例では、本明細書に記載されるgRNAライブラリは、インビボまたはエクスピボのスクリーニングおよび分析のために使用される。スクリーニングのための細胞は、生物体または細胞株から取られた初代細胞を含む。細胞は、原核生物（例えば、バクテリアおよび菌類）または真核生物（例えば、動植物）からのものであってもよい。例示的な動物細胞としては、これらに限定されないが、マウス、ウサギ、霊長類、および昆虫からのものが挙げられる。いくつかの例では、本明細書に記載されるgRNAライブラリは、多細胞生物に提供されてもよい。例示的な多細胞生物としては、これらに限定されないが、植物、マウス、ウサギ、霊長類、および昆虫が挙げられる。

【0054】

ゲノム編集

【0055】

本明細書で提供されるのは、特定の標的核酸配列のヌクレアーゼ標的化のための核酸を含むライブラリである。いくつかの例では、本明細書に記載されるライブラリは合成された核酸を含み、ここで、核酸は、DNA、RNA、任意のアナログ、またはそれらの誘導体である。いくつかの例では、標的核酸配列は、DNA、RNA、任意のアナログ、あるいはそれらの誘導体を含む。いくつかの例では、ヌクレアーゼは標的核酸配列を切断する。いくつかの例では、ヌクレアーゼは標的核酸を結合するが、それを切断しない。ヌクレアーゼの型としては、これらに限定されないが、転写活性化因子様エフェクターヌクレアーゼ（TALEN）、ジンクフィンガーヌクレアーゼ（ZFN）、メガヌクレアーゼ、アルゴノート、およびクラスタ化して規則的な配置の短い回文配列リピート（CRISPR）-関連の（Cas）タンパク質が挙げられる。いくつかの例では、ヌクレアーゼは野生型、遺伝子組換え型、または組み換え型である。

【0056】

標的遺伝子編集のためのモデルシステムは、Cas9に基づくアプローチを含む。gRNAと共に細胞内に発現されたか移された時、Cas9は、mRNAのクラスタ化して規則的な配置の短い回文配列リピート（CRISPR）配列との複合体を介して、遺伝子情報の標的導入または欠失を可能にする。図1A-1Bに例示されるようなCas9複合体は、ガイドRNA（「gRNA」）転写産物と係合するCas9タンパク質を含む。例示されるgRNAは、標的配列領域、PAM領域、およびヘアピン領域を含む。CRISPR/Cas9プロセスでは、gRNAは、Cas9酵素を特異的なひと続きのDNAに誘導する。示されたgRNAはsgRNA（一本鎖のガイドRNA）であるが、複合体はdgRNA（二本鎖のガイドRNA）と共に形成されてもよい。その後、Cas9は、遺伝子を不能にするか修復するためにDNAを切断する。このプロセスへの例示的な修正の非限定的なリストが本明細書で記載される。CRISPR/dCas9プロセスにおいて、不能にされたまたは「死んでいる」Cas9（「dCas9」）は、もはやスプライシング機能を有さないが、別の酵素活性の添加によって、異なる標的分子修正機能を果たす。例えば、シチジンデアミナーゼのdCas9への繫留は、C-G DNA塩基対をT-A塩基対へと変換する。他のdCas9プロセスでは、dCas9に繫留された異なる酵素は、標的DNA内で塩基CをTへ、または塩基GをAへと変更することをもたらす。あるいは、dCas9プロセスは、RNAポリメラーゼ活性を遮断するか活性化するために、転写因子の融合によって修飾することができ、結果として遺伝子転写の遮断（CRISPRi）または活性化（CRISPRa）をもたらす、したがって遺伝子発現を調節する

。例えば、d C a s 9 プロセスは転写抑制因子との融合によって修飾される。いくつかの例では、d C a s 9 プロセスは、転写活性化因子との融合によって修飾される。いくつかの例では、d C a s 9 プロセスは、複数の転写抑制因子または転写活性化因子との融合によって修飾される。他の構成では、g R N A は切断のための複数の部位を有し、遺伝子編集のための複数の領域を有する g R N A を結果としてもたらす。C a s 9 n、すなわち「ニッキング C a s 9」の場合には、R u v C または H N H の切断ドメインのいずれかが不活性になるように修正される。この不活性化は、C a s 9 が、二本鎖切断ではなく、D N A における一本鎖切断（ニック）を生成することしかできないようにさせる。いくつかの構成では、2 つの C a s 9 n 酵素、各鎖に 1 本ずつ、が二本鎖切断を生成するために使用される。それらが、切断部位の上流および下流の領域の両方を認識することができるように、オフターゲット効果が除去される。h f C a s 9 の場合には、オフターゲット効果のない C a s 9 切断を生成するために、二重の C a s 9 n タンパク質を使用する代わりに、修飾された C a s 9 酵素が、結合標的特異性のストリンジェンシーを緩めて酵素活性に先立って完全未満の適合を可能にする。いくつかの例では、d C a s 9 プロセスは標的核酸を検知するためのラベルまたはタグとの融合によって、修飾される。例えば、ラベルは、標的核酸を検知するための蛍光性のマーカー（例えば G F P）である。いくつかの例では、d C a s 9 はエピトープタグと融合し、g R N A によって特定された標的核酸の精製のために使用される。

【 0 0 5 7 】

本明細書で提供されるのは、ヌクレアーゼを特定の標的核酸配列に向けるための核酸を含むライブラリ。いくつかの例では、標的核酸配列は D N A を含む。いくつかの例では、標的核酸配列は R N A を含む。例えば、C 2 c 2 を方向付けるための核酸を含むライブラリは、R N A 配列を標的とするために生成される。いくつかの例では、D N A または R N A は一本鎖のまたは二本鎖である。

【 0 0 5 8 】

本明細書で提供されるのは、特定の標的核酸配列のヌクレアーゼ標的化のための核酸を含むライブラリであり、ここで、ヌクレアーゼは、これらに限定されないが、連鎖球菌属、カンピロバクター属、ニトラティフラクター属（*Nitratifractor*）、ブドウ球菌属、パービバキュラム属（*Parvibaculum*）、ロゼブリア、ナイセリア属、グルコンアセトバクター属（*Glucanacetobacter*）、アゾスピリルム属（*Azospirillum*）、フェロケタ属（*Sphaerochaeta*）、乳酸桿菌属、ユーバクテリウム属、コリネバクテリウム属（*Corynebacter*）、カルノバクテリウム属（*Carnobacterium*）、ロドバクター属（*Rhodobacter*）、リステリア菌、パルディバクター属（*Paludibacter*）、クロストリジウム属、ラクノスピラ科（*Lachnospiraceae*）、クロストリジアリジウム属（*Clostridiaridium*）、レプトトリキア属（*Leptotrichia*）、フランシセラ属、レジオネラ属、アリシクロバチルス属（*Alicyclobacillus*）、メタノメチオフィラス属（*Methanomethyophilus*）、ポルフィロモナス属（*Porphyromonas*）、プレボテラ属（*Prevotella*）、バクテロイデス門（*Bacteroidetes*）、ヘルココッカス属（*Helcococcus*）、レプトスピラ属（*Letospira*）、デサルフォピブリオ属、デサルフォナトロナム属（*Desulfonatronum*）、デスルフロコッカス属（*Desulfurococcus*）、オピウトウス科（*Opitutaceae*）、テュベリバシラス属（*Tuberibacillus*）、かん菌、ブレビバチルス属（*Brevibacillus*）、メチロバクテリウム属（*Methylobacterium*）、ナトロノバクテリウム属（*Natronobacterium*）、フラボバクテリウム属、サッカロミケス属、クラミドモナス属、サーマス属、パイロコッカス属（*Pyrococcus*）、マイコプラズマ属、またはアシダミノコッカス属の種に由来する。例示的なヌクレアーゼは、表 2 A において列挙される。本明細書に記載される g R N A は、上記の種のいずれかに由来のヌクレアーゼ、または酵素がゲノム編集機能を可能

にするさらなる種の由来のヌクレアーゼのターミネーター配列に結合することができる。
例示的なターミネーター配列は、限定されないが、表 2 B に列挙されたものが挙げられる。
例示的な P A M 配列は、限定されないが、表 2 C において列挙されるものが挙げられる。

【 0 0 5 9 】

【表 2 A 】

表 2A.

名称	受入番号	種
Cas9	Q99ZW2	化膿性連鎖球菌
Cas9	J7RUA5	黄色ブドウ球菌
Cas9	G3ECR1	ストレプトコッカス・サーモフィルス
C2c2	P0DOC6	レプトトリキア・シャヒイ
Cpf1	U2UMQ6	アシダミノコッカス <i>sp</i> (<i>Acidaminococcus sp.</i>)
C2c1	T0D7A2	アリシクロバチルス・アシドテレストリス
FokI	P14870	フラボバクテリウム・オケアノコイテス
AciI	A0A0C5GQT3	ラクトバチルス・アシドフィルス
I-SceI	P03882	出芽酵母
I-CreI	P05725	コナミドリムシ
I-DmoI	1B24_A	デスルフロコッカス・モビリス
TtAgo	Q746M7	高度好熱菌
PfAgo	Q8U3D2	パイロコッカス・フリオサス
NgAgo	A0A172MAH6	ナトロノバクテリウム・グレゴリー (<i>Natronobacterium gregoryi</i>)

【 0 0 6 0 】

【表 2 B 】

表 2B

SEQ ID NO	種	配列
7	ストレプトコッカス・サーモフィルス	5' TACTCAACTTGAAAAGGTGGCACCGATTTCGGTGTTTTT 3'
8	ミュータンス菌	5' TACACAACTTGAAAAAGTGCGCACCGATTTCGGTGCTTTT 3'
9	リステリア・イノキュア	5' TTATCAACTTTTAATTAAGTAGCGCTGTTTCGGCGCTTTT 3'
10	マイコプラズマ・モービレ	5' TATGCCGTAACCTACTACTTATTTTCAAAATAAGTAGTTTT 3'
11	カンピロバクター・ジェジュニ	5' GACTCTGCGGGGTTACAATCCCCTAAAACCGCTTTT 3'

【 0 0 6 1 】

【表 2 C】

表 2C

Cas9 の種/変異体	PAM 配列
化膿性連鎖球菌 / SpCas9	NGG
化膿性連鎖球菌 / SpCas9 D1135E variant	NGG
化膿性連鎖球菌 / SpCas9 VRER variant	NGCG
化膿性連鎖球菌 / SpCas9 EQR variant	NGAG
化膿性連鎖球菌 / SpCas9 VQR variant	NGAN or NGNG
黄色ブドウ球菌 / SaCas9	NNGRRT or NNRR(N)
髄膜炎菌	NNNNGATT
ストレプトコッカス・サーモフィルス	NNAGAAW
トレポネーマ・デンティコラ	NAAAAC

10

【 0 0 6 2 】

20

本明細書で提供されるのは、特定の核酸配列に 1 以上のヌクレアーゼを標的とするための核酸を含むライブラリである。いくつかの例では、ヌクレアーゼは、T A L E N、Z F N、メガヌクレアーゼ、アルゴノート、および C a s タンパク質の少なくとも 1 つである。例えば、1 を超えるヌクレアーゼは、大きなゲノム欠失を生成するため、直ちに複数の配列を修飾するため、またはニッカーゼなどの他の酵素と合わせて使用されるために多重化することができる。いくつかの例では、ヌクレアーゼの数は、標的核酸配列に対して少なくとも 2 のヌクレアーゼである。いくつかの例では、ヌクレアーゼの数は、標的核酸配列に対して、約 2 ~ 3、2 ~ 4、2 ~ 5、2 ~ 6、2 ~ 7、2 ~ 8、2 ~ 9、または 2 ~ 10 のヌクレアーゼの範囲である。

【 0 0 6 3 】

30

本明細書で提供されるのは、特定の核酸配列のニッカーゼ標的化のための合成された核酸を含むライブラリである。ニッカーゼは、核酸配列内で一本鎖の切断を生成する酵素である。いくつかの例では、合成された核酸は、DNA、RNA、任意のアナログ、またはそれらの誘導体である。いくつかの例では、特定の核酸配列は、含む DNA、RNA、任意のアナログ、またはそれらの誘導体を含む。いくつかの例では、ニッカーゼは特定の核酸配列を切断する。いくつかの例では、ニッカーゼは特定の核酸を結合するが、それを切断しない。いくつかの例では、ニッカーゼは変更されたヌクレアーゼであり、ここで、ヌクレアーゼが、T A L E N、Z F N、メガヌクレアーゼ、アルゴノート、または C a s タンパク質である。いくつかの例では、ニッカーゼは、T A L E N、Z F N、メガヌクレアーゼ、アルゴノート、または C a s タンパク質のヌクレアーゼドメインを変更することによって生成される。いくつかの例では、ニッカーゼは、C a s 9 のヌクレアーゼドメインを変更することによって生成される。

40

【 0 0 6 4 】

いくつかの例では、ライブラリは、特定の核酸配列の 1 以上のニッカーゼ標的化のための核酸を含む。いくつかの例では、ニッカーゼの数は、特定の核酸配列に対して少なくとも 2 つのニッカーゼである。いくつかの例では、ニッカーゼの数は、特定の核酸配列に対して、約 2 ~ 3、2 ~ 4、2 ~ 5、2 ~ 6、2 ~ 7、2 ~ 8、2 ~ 9、または 2 ~ 10 のニッカーゼの範囲である。いくつかの例では、ライブラリは、1 以上のニッカーゼおよび 1 以上のヌクレアーゼを特定の核酸配列に向けるための核酸を含む。

【 0 0 6 5 】

50

本明細書で提供される特定の核酸配列に対してヌクレアーゼを標的とするための核酸を含むライブラリは、特定の核酸配列の切断を結果としてもたらし得る。いくつかの例では、ヌクレアーゼは、T A L E N、Z F N、メガヌクレアーゼ、アルゴノート、およびC a sタンパク質の少なくとも1つである。いくつかの例では、ヌクレアーゼは、切断以外の、特定の核酸配列の修飾を提供するキメラヌクレアーゼである。例えば、キメラヌクレアーゼは、メチル化、脱メチル反応、ポリアデニル化、脱アデニル化、脱アミノ化、またはポリウリジニル化 (p o l y u r i d i n y l a t i o n) をもたらす。

【 0 0 6 6 】

転写活性化因子様エフェクターヌクレアーゼ

本明細書で提供されるのは、特定の核酸配列の転写活性化因子様エフェクターヌクレアーゼ (T A L E N) 標的化のための核酸を含む、核酸ライブラリを合成する方法である。T A L E N は、特定の標的配列で二本鎖切断を誘発するために使用することができる、操作された配列の特異的なヌクレアーゼの類である。T A L E N は、ヌクレアーゼの触媒ドメインに、転写活性化因子様 (T A L) エフェクターDNA結合ドメイン、またはそれらの機能部分を融合させることによって生成され得る。T A L エフェクターDNA結合ドメインは、一連のT A L 反復を含み、それは、各々がR V D (r e p e a t v a r i a b l e d i r e s i d u e) として知られる非常に可変的な12番目と13番目のアミノ酸を含む一般的に高度に保存された33または34のアミノ酸配列セグメントである。各R V D は、特異的なヌクレオチドを認識し、それに結合することができる。したがって、T A L エフェクター結合ドメインは、適切なR V D を含むT A L 反復を組み合わせることによってヌクレオチドの特異的なシーケンスを認識するように操作され得る。

【 0 0 6 7 】

本明細書で提供されるのは、T A L エフェクターDNA結合ドメインをコードする非同一の核酸を含む核酸ライブラリを合成する方法である。いくつかの例では、T A L エフェクターDNA結合ドメインは、特定の標的核酸配列を認識し、特別の部位で二本鎖切断を誘発するようにするように設計される。いくつかの例では、T A L エフェクターDNA結合ドメインは、特定の核酸配列を認識してそれに結合するように設計される多くのT A L 反復を含む。いくつかの例では、T A L 反復の数は、少なくともまたは約、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、またはそれより多くのT A L 反復である。

【 0 0 6 8 】

いくつかの例では、T A L エフェクターDNA結合ドメインをコードする非同一の核酸を含む核酸ライブラリは合成される。いくつかの例では、本明細書に記載されるような核酸ライブラリは、翻訳される時にタンパク質ライブラリをコードする。いくつかの例では、核酸ライブラリが細胞内で発現し、タンパク質ライブラリが生成される。いくつかの例では、合成された核酸ライブラリは、発現ベクターへと挿入される。いくつかの例では、合成された核酸ライブラリは、発現ベクターへと挿入され、細胞内で発現する。

【 0 0 6 9 】

本明細書に記載される方法によって生成されたT A L エフェクターDNA結合ドメインをコードする核酸を含む核酸ライブラリは、T A L E N を生成するために用いられる。いくつかの例では、これは、クローン化され、ヌクレアーゼを有するベクター内で発現するT A L エフェクター結合ドメインのライブラリを混合することによって達成される。例示的なヌクレアーゼとしては、これらに限定されないが、A c i I、A c u I、A l w I、B b v I、B c c I、B c e A I、B c i V I、B f u A I、B m g B I、B m r I、B p m I、B p u E I、B s a I、B s m A I、B s m F I、B s e R I、B s p C N I、B s r I、B s g I、B s m I、B s p M I、B s r B I、B s r D I、B t g Z I、B t s I、B t s C I、E a r I、E c i l、F o k I、H g a I、H p h I、H p y A V、M b o I I、M l y I、M m e I、M n l I、N m e A I I I、P l e I、S f a N I、B b v C I、B p u 1 0 I、B s p Q I、S a p I、B a e I、B s a X I、またはC

s p C I が挙げられる。いくつかの例では、混合は、ライゲーションによって生じる。例示的なリガーゼとしては、これらに限定されないが、大腸菌リガーゼ、T4リガーゼ、哺乳類のリガーゼ（例えば、DNAリガーゼI、DNAリガーゼII、DNAリガーゼIII、DNAリガーゼIV）、熱安定性リガーゼ、およびファストリガーゼ（fast ligases）が挙げられる。本明細書に記載される方法で生成されたTALLENは、発現ベクターに挿入され得る。いくつかの例では、TALLENは、発現ベクターに挿入され、細胞内で発現する。

【0070】

本明細書で提供されるのは、原核生物または真核生物のゲノム中の遺伝子について、非同一の核酸配列を含むTALエフェクターDNA結合ドメインのライブラリを合成するための方法である。いくつかの例では、TALエフェクターDNA結合ドメインのライブラリは、ゲノムの少なくとも5%について、遺伝子のための少なくとも3、4、5、6、7、8、9、10、または10よりも多い非同一の核酸配列を含む。いくつかの例では、TALエフェクターDNA結合ドメインのライブラリは、ゲノムの少なくとも5%について、1以上の遺伝子のための非同一の核酸配列を含む。いくつかの例では、TALエフェクターDNA結合ドメインのライブラリは、ゲノムの少なくとも5%について、約1、5、10、20、30、40、50、60、70、80、90、100、または100よりも多い遺伝子のための、非同一の核酸配列を含む。いくつかの例では、TALエフェクターDNA結合ドメインのライブラリは、ゲノムの少なくとも5%について、約1 - 100、2 - 95、5 - 90、10 - 85、15 - 80、20 - 75、25 - 70、30 - 65、35 - 60、または40 - 50の遺伝子のための、非同一の核酸配列を含む。

【0071】

ジンクフィンガーヌクレアーゼ

【0072】

本明細書で提供されるのは、特定の核酸配列のジンクフィンガーヌクレアーゼ（ZFN）標的化のための核酸を含む核酸ライブラリを合成する方法である。ZFNは、ヌクレアーゼとDNA結合ジンクフィンガードメイン（ZFD）との融合によって生成され得る。ZFDは、1以上のジンクフィンガーを介して標的核酸配列に結合することができる。いくつかの例では、ZFDは、少なくとも2、3、4、5、6、7、8、9、10、または10以上のジンクフィンガーを含む。いくつかの例では、ZFDは、最大で2、3、4、5、6、7、8、9、10、または10以上のジンクフィンガーを含む。いくつかの例では、ZFDは、特定の標的核酸配列を認識し、特別の部位で二本鎖切断を誘発するようにするように設計される。

【0073】

本明細書で提供されるのは、転写され翻訳された時、ZFDをコードする核酸を含む核酸ライブラリを合成する方法である。いくつかの例では、核酸ライブラリが翻訳される時、タンパク質ライブラリをコードする。いくつかの例では、核酸ライブラリが細胞内で発現され、タンパク質ライブラリが生成される。いくつかの例では、合成された核酸ライブラリは、発現ベクターに挿入される。いくつかの例では、合成された核酸ライブラリは、発現ベクターに挿入され、細胞内で発現する。

【0074】

本明細書に記載される方法によって生成されたZFDをコードする核酸を含む核酸ライブラリは、ZFNを生成するために用いられる。いくつかの例では、これは、クローン化され、ヌクレアーゼを有するベクター内で発現するZFDを混合することによって達成される。例示的なヌクレアーゼとしては、これらに限定されないが、AciI、AclI、AlwI、BbvI、BccI、BceAI、BciVI、BfuAI、BmgBI、BmrI、BpmI、BpuEI、BsaI、BsmAI、BsmFI、BseRI、BspCNI、BsrI、BsgI、BsmI、BspMI、BsrBI、BsrDI、BtgZI、BtsI、BtsCI、EarI、EciI、FokI、HgaI、HphI、HpyAV、MboII、MlyI、MmeI、MnlI、NmeAIII、PleI、

S f a N I、B b v C I、B p u 1 0 I、B s p Q I、S a p I、B a e I、B s a X I、またはC s p C Iが挙げられる。いくつかの例では、混合は、ライゲーションによって生じる。例示的なリガーゼとしては、これらに限定されないが、大腸菌リガーゼ、T4リガーゼ、哺乳類のリガーゼ（例えば、DNAリガーゼI、DNAリガーゼII、DNAリガーゼIII、DNAリガーゼIV）、熱安定性リガーゼ、およびファストリガーゼ（fast ligases）が挙げられる。本明細書に記載される方法によって生成されたT A L E Nは、発現ベクターに挿入され得る。いくつかの例では、Z F Nは、発現ベクターに挿入され、細胞内で発現する。

【0075】

本明細書で提供されるのは、原核生物または真核生物のゲノム中の遺伝子について、非同一の核酸配列を含むZ F Dライブラリを合成するための方法である。いくつかの例では、Z F Dライブラリは、ゲノムの少なくとも5%について、遺伝子のための少なくとも3、4、5、6、7、8、9、10、または10よりも多い非同一の核酸配列を含む。いくつかの例では、Z F Dライブラリは、ゲノムの少なくとも5%について、1以上の遺伝子のための、非同一の核酸配列を含む。いくつかの例では、Z F Dライブラリは、ゲノムの少なくとも5%について、約1、5、10、20、30、40、50、60、70、80、90、100、または100よりも多い遺伝子のための、非同一の核酸配列を含む。いくつかの例では、Z F Dライブラリは、ゲノムの少なくとも5%について、約1 - 100、2 - 95、5 - 90、10 - 85、15 - 80、20 - 75、25 - 70、30 - 65、35 - 60、または40 - 50の遺伝子のための、非同一の核酸配列を含む。

【0076】

メガヌクレアーゼ

【0077】

本明細書で提供されるのは、特定の核酸配列のメガヌクレアーゼ標的化のための核酸を含む核酸ライブラリを合成する方法である。メガヌクレアーゼは、塩基対の長い（例えば12 - 40の塩基対）DNA標的を認識し切断することができる酵素である。いくつかの例では、メガヌクレアーゼは、標的核酸配列について特異性を与える他の酵素のドメインを含むように操作される。例えば、メガヌクレアーゼはT A L エフェクターDNA結合ドメインを含むように操作される。

【0078】

本明細書で提供されるのは、転写され翻訳された時、メガヌクレアーゼと共に使用するための結合ドメインをコードする核酸を含む核酸ライブラリを合成する方法である。いくつかの例では、核酸ライブラリが翻訳される時、タンパク質ライブラリをコードする。いくつかの例では、核酸ライブラリが細胞内で発現され、タンパク質ライブラリが生成される。いくつかの例では、合成された核酸ライブラリは、発現ベクターに挿入される。いくつかの例では、合成された核酸ライブラリは、発現ベクターに挿入され、細胞内で発現する。

【0079】

本明細書に記載される方法によって生成されたドメインをコードする核酸を含む核酸ライブラリは、特定の核酸配列を標的とするためのメガヌクレアーゼを操作するために使用される。いくつかの例では、これは、クローン化され、メガヌクレアーゼを有するベクター内で発現するT A L エフェクター結合ドメインのライブラリのような結合ドメインのライブラリを混合することによって達成される。本明細書で提供される方法と共に使用するための例示的なメガヌクレアーゼとしては、これら限定されないが、I - S c e l、I - S c e l l、I - S c e I I I、I - S c e I V、I - S c e V、I - S c e V I、I - S c e V I I、I - C e u l、I - C e u A I I P、I - C r e l、I - C r e p s b I P、I - C r e p s b l l P、I - C r e p s b I I I P、I - C r e p s b I V P、I - T l i l、I - P p o l、P I - P s p I、F - S c e l、F - S c e l l、F - S u v l、F - T e v I、F - T e v I I、I - A m a l、I - A n i l、I - C h u l、I - C m o e l、I - C p a l、I - C p a I I、I - C s m l、I - C v u l、I - C

v u A I P、I - D d i l、I - D d i I I、I - D i r l、I - D m o l、I - H m u l、I - H m u I I、I - H s N I P、I - L l a l、I - M s o l、I - N a a l、I - N a n l、I - N c I I P、I - N g r I P、I - N i t l、I - N j a l、I - N s p 2 3 6 I P、I - P a k l、I - P b o I P、I - P c u I P、I - P c u A I、I - P c u V I、I - P g r l P、I - P o b I P、I - P o r l、I - P o r I I P、I - P b p I P、I - S p B e t a I P、I - S c a l、I - S e x I P、I - S n e I P、I - S p o m l、I - S p o m C P、I - S p o m I P、I - S p o m I I P、I - S q u I P、I - S s p 6 8 0 3 I、I - S t h P h i J P、I - S t h P h i S T 3 P、I - S t h P h i S T e 3 b P、I - T d e I P、I - T e v l、I - T e v I I、I - T e v I I I、I - U a r A P、I - U a r H G P A I P、I - U a r H G P A 1 3 P、I - V i n I P、I - Z b i I P、P I - M t u I、P I - M t u H I P、P I - M t u H I I P、P I - P f u I、P I - P f u I I、P I - P k o I、P I - P k o I I、P I - R m a 4 3 8 1 2 I P、P I - S p B e t a I P、P I - S c e I、P I - T f u I、P I - T f u I I、P I - T h y l、P I - T l i l、P I - T l i I I、またはそれらの断片が挙げられる。いくつかの例では、混合は、ライゲーションによって生じる。例示的なリガーゼとしては、これらに限定されないが、大腸菌リガーゼ、T4リガーゼ、哺乳類のリガーゼ（例えば、DNAリガーゼI、DNAリガーゼII、DNAリガーゼIII、DNAリガーゼIV）、熱安定性リガーゼ、およびファストリガーゼ（fast ligases）が挙げられる。本明細書に記載される方法によって操作されたメガヌクレアーゼは、発現ベクターに挿入され得る。いくつかの例では、操作されたメガヌクレアーゼは、発現ベクターに挿入され、細胞内で発現する。

【0080】

本明細書で提供されるのは、原核生物または真核生物のゲノム中の遺伝子について非同一の核酸配列を含むメガヌクレアーゼと共に使用するための結合ドメインライブラリを合成するための方法である。いくつかの例では、ドメインライブラリは、ゲノムの少なくとも5%のゲノムについて、遺伝子のための少なくとも3、4、5、6、7、8、9、10、または10よりも多い非同一の核酸配列を含む。いくつかの例では、ドメインライブラリは、ゲノムの少なくとも5%について、1以上の遺伝子のための、非同一の核酸配列を含む。いくつかの例では、ドメインライブラリは、ゲノムの少なくとも5%について、約1、5、10、20、30、40、50、60、70、80、90、100、または100よりも多い遺伝子のための、非同一の核酸配列を含む。いくつかの例では、ドメインライブラリは、ゲノムの少なくとも5%について、約1 - 100、2 - 95、5 - 90、10 - 85、15 - 80、20 - 75、25 - 70、30 - 65、35 - 60、または40 - 50の遺伝子のための、非同一の核酸配列を含む。

【0081】

アルゴノート

【0082】

本明細書で提供されるのは、特定の核酸配列のメガヌクレアーゼ標的化のための核酸を含む核酸ライブラリを合成する方法である。アルゴノートは、RNAまたはDNAのガイドヌクレアーゼのファミリーである。いくつかの例では、アルゴノートは、標的核酸を識別するためにガイド核酸を使用する。いくつかの例では、ガイド核酸は、単一ガイドRNA（sgRNA）である。いくつかの例では、ガイド核酸は、ガイドDNA（gDNA）である。例示的なアルゴノートとしては、これらに限定されないが、TtAgo、PfAgo、およびNgAgoが挙げられる。いくつかの実施形態では、アルゴノートはNgAgoである。

【0083】

本明細書で提供されるのは、原核生物または真核生物のゲノム中の遺伝子について、非同一の核酸配列を含むガイド核酸ライブラリを合成するための方法である。幾つかの事例では、ガイド核酸ライブラリはsgRNAライブラリである。幾つかの事例では、ガイド核酸ライブラリはdgRNAライブラリである。いくつかの例では、ガイド核酸ライブラ

リは、ゲノムの少なくとも5%について、遺伝子のための少なくとも3、4、5、6、7、8、9、10、または10よりも多い非同一の核酸配列を含む。いくつかの例では、ガイド核酸ライブラリは、ゲノムの少なくとも5%について、1以上の遺伝子のための、非同一の核酸配列を含む。いくつかの例では、ガイド核酸ライブラリは、ゲノムの少なくとも5%について、約1、5、10、20、30、40、50、60、70、80、90、100、または100よりも多い遺伝子のための、非同一の核酸配列を含む。いくつかの例では、ガイド核酸ライブラリは、ゲノムの少なくとも5%について、約1 - 100、2 - 95、5 - 90、10 - 85、15 - 80、20 - 75、25 - 70、30 - 65、35 - 60、または40 - 50の遺伝子のための、非同一の核酸配列を含む。

【0084】

CRISPR関連タンパク質

【0085】

本明細書で提供されるのは、特定の核酸配列のCRISPR関連(Cas)タンパク質標的化のためのgRNAをコードする核酸を含む核酸ライブラリを合成する方法である。いくつかの例では、Casタンパク質は、少なくとも、Cpf1、C2c1、C2c2、Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas6、Cas7、Cas8、Cas9(Csn1またはCsx12)、Cas10、Csy1、Csy2、Csy3、Cse1、Cse2、Csc1、Csc2、Csa5、Csn2、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Csb1、Csb2、Csb3、Csx17、Csx14、Csx10、Csx16、CsaX、Csx3、Csx1、Csx15、CsF1、CsF2、CsF3、CsF4、それらの相同体、およびそれらの修飾バージョンの1つである。いくつかの例では、Casタンパク質はCas9である。

【0086】

本明細書で提供されるのは、原核生物または真核生物のゲノム中の遺伝子について、非同一の核酸配列を含むgRNAライブラリを合成するための方法である。いくつかの例では、gRNAライブラリは、ゲノムの少なくとも5%について、遺伝子のための少なくとも3、4、5、6、7、8、9、10、または10よりも多い非同一の核酸配列を含む。いくつかの例では、gRNAライブラリは、ゲノムの少なくとも5%について、1以上の遺伝子のための、非同一の核酸配列を含む。いくつかの例では、gRNAライブラリは、ゲノムの少なくとも5%について、約1、5、10、20、30、40、50、60、70、80、90、100、または100よりも多い遺伝子のための、非同一の核酸配列を含む。いくつかの例では、gRNAライブラリは、ゲノムの少なくとも5%について、約1 - 100、2 - 95、5 - 90、10 - 85、15 - 80、20 - 75、25 - 70、30 - 65、35 - 60、または40 - 50の遺伝子のための、非同一の核酸配列を含む。gRNAライブラリは、sgRNAまたはdgRNAをコードし得る。

【0087】

変異体ライブラリ合成

【0088】

本明細書で提供されるのは、gRNAおよびヌクレアーゼ、例えば、Cas9酵素またはCas9変異酵素を有する、完全または一部の遺伝子配列をコードする核酸の組み合わせによって生成された変異体核酸ライブラリを合成するための方法である。断片は、遺伝子の全領域に相対的に間隔をあけてもよい。場合によっては、ライブラリがDNAまたはRNAをコードする。場合によっては、ライブラリが、単一の遺伝子または最大で全ゲノムをコードする。例えば、約20,000の遺伝子を含むゲノムについて、遺伝子あたり、5つのgRNAをコードするgRNAライブラリは、約100,000のgRNAを結果としてもたらす。そのようなライブラリは、単一遺伝子、遺伝子の経路、または単一ゲノム内の全遺伝子を、選択的にサイレンシングするか修飾するために使用することができる。ある構成では、gRNAは相同配列を欠き、無作為な末端結合が生じる。そのようなプロセスは、非相同末端結合(「NHEJ」)を結果としてもたらす。いくつかの例では、

10

20

30

40

50

N H E J に続いて、挿入、欠失、フレームシフト、または一塩基スワッピング (s i n g l e b a s e s w a p p i n g) が生じる。図 2 の B を参照する。

【 0 0 8 9 】

本明細書に記載される合成されたライブラリは、C R I S P R - C a s 9 機能における適用のために使用することができ、ここで、生成された g R N A 配列が、細胞内または標的 D N A および C a s 9 酵素を含む混合物内で標的 D N A 配列の発現を破壊するか発現産物配列を変更するために使用される。いくつかの実施形態では、各変異体は、翻訳中に異なるアミノ酸をもたらすコドンをコードする。表 3 は、変異部位にとって可能性のあるそれぞれのコドン (と代表的なアミノ酸) のリストを提供する。

【 0 0 9 0 】

【表 3 - 1】

表 3. コドンおよびアミノ酸の表

アミノ酸	1 文 字コ ード	3 文 字コ ード	コドン					
アラニン	A	Ala	GCA	GCC	GCG	GCT		
システイン	C	Cys	TGC	TGT				
アスパラギ ン酸	D	Asp	GAC	GAT				
グルタミン 酸	E	Glu	GAA	GAG				
フェニルア ラニン	F	Phe	TTC	TTT				
グリシン	G	Gly	GGA	GGC	GGG	GGT		
ヒスチジン	H	His	CAC	CAT				
イソロイシ ン	I	Iso	ATA	ATC	ATT			
リジン	K	Lys	AAA	AAG				
ロイシン	L	Leu	TTA	TTG	CTA	CTC	CTG	CTT
メチオニン	M	Met	ATG					
アスパラギ ン	N	Asn	AAC	AAT				
プロリン	P	Pro	CCA	CCC	CCG	CCT		
グルタミン	Q	Gln	CAA	CAG				

【表 3 - 2】

アルギニン	R	Arg	AGA	AGG	CGA	CGC	CGG	CGT
セリン	S	Ser	AGC	AGT	TCA	TCC	TCG	TCT
スレオニン	T	Thr	ACA	ACC	ACG	ACT		
バリン	V	Val	GTA	GTC	GTG	GTT		
トリプトファン	W	Trp	TGG					
チロシン	Y	Tyr	TAC	TAT				

10

【 0 0 9 2 】

本明細書で提供されるのは、ヌクレアーゼを有する完全なまたは一部の遺伝子配列をコードする核酸の組み合わせによって生成された変異核酸ライブラリを合成するための方法であり、ここで、ヌクレアーゼは T A L E N、Z F N、または操作されたメガヌクレアーゼである。いくつかの例では、ヌクレアーゼを有する s g R N A などのガイド核酸を有する完全なまたは一部の遺伝子配列をコードする核酸の組み合わせによって生成された、変異体核酸ライブラリを合成するための方法であり、ここで、ヌクレアーゼはアルゴノートまたは C a s タンパク質である。本明細書に記載される合成されたライブラリは、ヌクレアーゼ機能における適用のために使用することができ、ここで、生成された核酸配列が、細胞内または標的 D N A およびヌクレアーゼを含む混合物内で標的 D N A 配列の、発現を破壊するか発現産物配列を変更するために使用される。いくつかの実施形態では、各変異体は、翻訳中に異なるアミノ酸をもたらすコドンにコードする。

20

30

【 0 0 9 3 】

本明細書に記載されるような変異核酸ライブラリは、合成された核酸によってコードされる後続タンパク質において、単一の残基の複数の様々な変異体が標準的な翻訳プロセスによって生成されるように、少なくとも単一のコドンにコードする標的核酸配列を変えるための s g R N A または d g R N A を含む。いくつかの事例では、各変異体は、タンパク質ドメインの異なるアミノ酸をもたらすコドンにコードする。例えば、タンパク質ドメインは、保存ドメインまたは触媒ドメインである。いくつかの実施形態では、タンパク質ドメインとしては、これらに限定されないが、キナーゼドメイン、A T P 結合ドメイン、G T P 結合ドメイン、グアニンヌクレオチド交換因子 (G E F) ドメイン G T P アーゼ活性化タンパク質ドメイン、ヒドロラーゼドメイン、エンドヌクレアーゼドメイン、エキソヌクレアーゼドメイン、プロテアーゼドメイン、ホスファターゼドメイン、ホスホリパーゼドメイン、プレクストリン相同ドメイン、S r c 相同ドメイン、およびユビキチン結合ドメインが挙げられる。いくつかの例では、変異体核酸ライブラリは、少なくとも 1、2、3、4、5、6、7、8、9、10、または 10 よりも多いタンパク質ドメインにおける変異をコードする核酸配列を標的とするために、s g R N A または d g R N A を含む。

40

【 0 0 9 4 】

いくつかの実施形態では、変異体は、特定の活性を有するタンパク質についてアミノ酸を体コードする。例えば、変異体は、それはメチルトランスフェラーゼ活性、デメチラーゼ活性、ジスムターゼ活性、アルキル化活性、デプリネーション活性、酸化活性、ピリミジンダイマー形成活性、インテグラーゼ活性、トランスポサーゼ活性、レコンピナーゼ活

50

性、ポリメラーゼ活性、リガーゼ活性、ヘリカーゼ活性、ホトリアーゼ活性、グリコシラーゼ活性、アセチルトランスフェラーゼ活性、デアセチラーゼ活性、キナーゼ活性、ホスファターゼ活性、ユビキチンリガーゼ活性、脱ユビキチン活性、アデニル化活性、脱アデニル化活性、SUMO化活性、脱SUMO化活、リボシル化活性、脱リボシル化活性、ミリストイル化活性、リモデリング活性、プロテアーゼ活性、オキシドレダクターゼ活性、トランスフェラーゼ活性、ヒドロラーゼ活性、リアーゼ活性、イソメラーゼ活性、シンターゼ活性、シンテターゼ活性、または脱ミリストイル化活性を含むタンパク質をエンコードする。

【0095】

相同組換え修復 (HDR) によって生成される変異

変異核酸ライブラリ生成のための例示的なプロセスにおいて、Cas9切断および相同組換えが、標的DNAライブラリ内に変型を生成するために組込まれる。まず、(RNAのデノボ合成またはDNAのデノボ合成のいずれか、その後(インビボまたはインビトロの)転写をしてgRNAを生成することによって)gRNAライブラリが合成され、ここで、ライブラリは、遺伝子あたり複数のgRNA分子を含む。例えば、gRNAライブラリは、遺伝子あたり、1、2、3、4、5、6、7、8、9、10、またはそれより多くのgRNAを含み得る。gRNAライブラリは、Cas9酵素および標的DNAライブラリと混合され、ここで、標的DNAライブラリは少なくとも1つの遺伝子断片または少なくとも1つの遺伝子をコードする核酸配列を含む。例えば、標的DNAライブラリは1、2、3、4、5、6、7、8、9、10、またはそれより多くの遺伝子または遺伝子断片を含み得る。いくつかの例では、標的DNAライブラリは、経路内の同義遺伝子からの配列または生物体内の全遺伝子からの配列を含む。さらに、変異が標的DNA鎖に導入されるように、相性配列および変異核酸配列を含む置換配列が混合物に加えられる。結果として生じる標的DNAライブラリは、複数の変異DNA配列を含むだろう。いくつかの例では、変異は、欠失、フレームシフト、または標的DNA配列への挿入を導入する。いくつかの例では、変異DNA配列は、遺伝子または遺伝子の断片あたり、少なくとも1のコドンについての変異を結果としてもたらす。いくつかの例では、遺伝子の一部が標的DNAに挿入され、あるいは、標的DNA配列の一部(つまり、遺伝子の断片または全遺伝子)は、標的DNAから取り除かれる。いくつかの例では、変異DNA配列は、遺伝子または遺伝子の断片と関連付けられる少なくとも1つの転写調節配列、例えば、プロモーター、UTR、あるいはターミネーター配列についての変異を結果としてもたらす。

【0096】

変異核酸ライブラリ生成のためのいくつかの例では、ヌクレアーゼ切断および相同組換えは、標的DNAライブラリ内に変型を生成するために組込まれ、ここで、ヌクレアーゼは、TALEN、ZFN、メガヌクレアーゼ、Cas、またはアルゴノートである。いくつかの例では、ヌクレアーゼがTALENである場合、(RNAのデノボ合成またはDNAのデノボ合成のいずれか、その後、(インビボまたはインビトロの)転写および翻訳によって)TALEフェクターDNA結合ドメインのライブラリが合成され、ここで、ライブラリは、遺伝子あたり複数のTALEフェクターDNA結合ドメイン分子を含む。例えば、TALEフェクターDNA結合ドメインのライブラリは、遺伝子あたり、1、2、3、4、5、6、7、8、9、10、またはそれ以上のTALEフェクターDNA結合ドメイン分子を含み得る。その後、TALEフェクターDNA結合ドメインのライブラリは、TALENを生成するためにヌクレアーゼ酵素と混合され得る。いくつかの例では、TALENは標的DNAライブラリと結合され、ここで、標的DNAライブラリは、少なくとも1つの遺伝子断片または少なくとも1つの遺伝子をコードする核酸配列を含む。例えば、標的DNAライブラリは1、2、3、4、5、6、7、8、9、10、またはそれ以上の遺伝子または遺伝子断片を含み得る。いくつかの例では、標的DNAライブラリは、経路内の同義遺伝子からの配列または生物体内の全遺伝子からの配列を含む。いくつかの例では、変異が標的DNA鎖に導入されるように、相性配列および変異核酸配列を含む置換配列が混合物に加えられる。結果として生じる標的DNAライブラリは、複数の変異DNA

10

20

30

40

50

配列を含むだろう。いくつかの例では、変異は、欠失、フレームシフト、または標的 DNA 配列への挿入を導入する。いくつかの例では、変異 DNA 配列は、遺伝子または遺伝子の断片あたり、少なくとも 1 のコドンについての変異を結果としてもたらす。いくつかの例では、遺伝子の一部が標的 DNA に挿入され、あるいは、標的 DNA 配列の一部（つまり、遺伝子の断片または全遺伝子）は、標的 DNA から取り除かれる。いくつかの例では、変異 DNA 配列は、遺伝子または遺伝子の断片と関連付けられる少なくとも 1 つの転写調節配列、例えば、プロモーター、UTR、あるいはターミネーター配列についての変異を結果としてもたらす。

【0097】

修飾された Cas9 酵素によって生成される変異

10

【0098】

変異核酸ライブラリの生成のための 2 番目の例示的なプロセスは、修飾された Cas9 酵素が、変異標的 DNA ライブラリを生成するために組込まれる。まず、(RNA のデノボ合成または DNA のデノボ合成のいずれか、その後 (インビボまたはインビトロの) 転写をして gRNA を生成することによって) gRNA ライブラリが合成され、ここで、ライブラリは、遺伝子あたり複数の gRNA 分子を含む。例えば、gRNA ライブラリは、遺伝子あたり、1、2、3、4、5、6、7、8、9、10、またはそれ以上の gRNA を含み得る。gRNA ライブラリは、修飾された Cas9 酵素および標的 DNA ライブラリと混合され、ここで、標的 DNA ライブラリは少なくとも 1 つの遺伝子断片または少なくとも 1 つの遺伝子をコードする核酸配列を含む。例えば、標的 DNA ライブラリは 1、2、3、4、5、6、7、8、9、10、またはそれ以上の遺伝子または遺伝子断片を含み得る。いくつかの例では、標的 DNA ライブラリは、経路内の同義遺伝子からの配列または生物体内の全遺伝子からの配列を含む。修飾された Cas9 酵素は、核酸配列修飾能力を有する別の酵素に繫留される。例示的な修飾された Cas9 酵素は、dCas9 プロセスを含み、このプロセスは、不能にされたまたは「死んでいる」Cas9 (「dCas9」) は、もはやスプライシング機能を有しないが、別の酵素活性の添加によって、異なる標的分子修正機能を果たす。例えば、シチジンデアミナーゼの dCas9 への繫留は、C - G DNA 塩基対を T - A 塩基対へと変換する。他の dCas9 プロセスでは、dCas9 に繫留された異なる酵素は、標的 DNA 内で塩基 C を T へ、または塩基 G を A へと変更することをもたらす。結果として生じる標的 DNA ライブラリは、複数の変異標的 DNA 配列を含む。いくつかの例では、変異は、欠失、フレームシフト、または標的 DNA 配列への挿入を導入する。いくつかの例では、変異 DNA 配列は、遺伝子または遺伝子の断片あたり、少なくとも 1 のコドンについての変異を結果としてもたらす。いくつかの例では、変異 DNA 配列は、遺伝子または遺伝子の断片と関連付けられる少なくとも 1 つの転写調節配列、例えば、プロモーター、UTR、あるいはターミネーター配列についての変異を結果としてもたらす。

20

30

【0099】

修飾されたヌクレアーゼによって生成される変異

【0100】

本明細書で提供されるのは、変異標的 DNA ライブラリを生成するために組込まれる修飾されたヌクレアーゼ酵素を含む、変異核酸ライブラリの生成のための方法である。いくつかの例では、ヌクレアーゼは TALEN である。いくつかの例では、(RNA のデノボ合成または DNA のデノボ合成のいずれか、その後、転写および翻訳をして TALE 効果ター DNA 結合ドメインのライブラリを生成することによって) TALE 効果ター DNA 結合ドメインのライブラリが合成され、ここで、ライブラリが遺伝子あたり複数の非同一の核酸配列を含む。例えば、TALE 効果ター DNA 結合ドメインのライブラリは、遺伝子あたり、1、2、3、4、5、6、7、8、9、10、またはそれ以上の非同一の核酸配列を含み得る。その後、TALE 効果ター DNA 結合ドメインのライブラリは、ヌクレアーゼ酵素と混合され得る。いくつかの例では、TALEN は標的 DNA ライブラリと混合され、ここで、標的 DNA ライブラリは、少なくとも 1 つの遺伝子断片または少

40

50

なくとも1つの遺伝子をコードする核酸配列を含む。例えば、標的DNAライブラリは、1、2、3、4、5、6、7、8、9、10、またはそれ以上の遺伝子または遺伝子断片を含み得る。いくつかの例では、標的DNAライブラリは、経路内の同義遺伝子からの配列または生物体内の全遺伝子からの配列を含む。

【0101】

いくつかの例では、ヌクレアーゼはZFNである。いくつかの例では、(RNAのデノボ合成またはDNAのデノボ合成のいずれか、その後、転写および翻訳をしてZFDライブラリを生成することによって)ZFDライブラリが合成され、ここで、ライブラリが遺伝子あたり複数の非同一の核酸配列を含む。例えば、ZFDライブラリは、遺伝子あたり、1、2、3、4、5、6、7、8、9、10、またはそれ以上の非同一の核酸配列を含み得る。その後、ZFDライブラリは、ZFNを生成するためにヌクレアーゼ酵素と混合され得る。いくつかの例では、ZFNは標的DNAライブラリと混合され、ここで、標的DNAライブラリは、少なくとも1つの遺伝子断片または少なくとも1つの遺伝子をコードする核酸配列を含む。例えば、標的DNAライブラリは、1、2、3、4、5、6、7、8、9、10、またはそれ以上の遺伝子または遺伝子断片を含み得る。いくつかの例では、標的DNAライブラリは、経路内の同義遺伝子からの配列または生物体内の全遺伝子からの配列を含む。

10

【0102】

いくつかの例では、ヌクレアーゼは、メガヌクレアーゼである。いくつかの例では、特定の核酸配列に対してメガヌクレアーゼを標的とするための、TALエフェクターDNA結合ドメインのライブラリのような結合ドメインライブラリは、(RNAのデノボ合成またはDNAのデノボ合成のいずれか、その後、転写をして結合ドメインライブラリを生成することによって)合成され、ここで、結合ドメインライブラリは、遺伝子あたり複数の非同一の核酸配列を含む。例えば、結合ドメインライブラリは、遺伝子あたり、1、2、3、4、5、6、7、8、9、10、またはそれより多くの非同一の核酸配列を含み得る。その後、結合ドメインライブラリは、操作されたメガヌクレアーゼを生成するためにメガヌクレアーゼ酵素と混合され得る。いくつかの例では、操作されたメガヌクレアーゼは標的DNAライブラリと混合され、ここで、標的DNAライブラリは、少なくとも1つの遺伝子断片または少なくとも1つの遺伝子をコードする核酸配列を含む。例えば、標的DNAライブラリは、1、2、3、4、5、6、7、8、9、10、またはそれ以上の遺伝子または遺伝子断片を含み得る。いくつかの例では、標的DNAライブラリは、経路内の同義遺伝子からの配列または生物体内の全遺伝子からの配列を含む。

20

30

【0103】

いくつかの例では、ヌクレアーゼはアルゴノートである。いくつかの例では、(RNAのデノボ合成またはDNAのデノボ合成のいずれか、その後、転写をしてガイド核酸ライブラリを生成することによって)ガイド核酸ライブラリ (gRNAまたはgDNA) が合成され、ここで、ガイド核酸ライブラリが、遺伝子あたり複数の非同一の核酸配列を含む。例えば、ガイド核酸ライブラリは、遺伝子あたり、1、2、3、4、5、6、7、8、9、10、またはそれより多くの非同一の核酸配列を含み得る。ガイド核酸ライブラリは、修飾されたアルゴノート酵素および標的DNAライブラリと混合され、ここで、標的DNAライブラリは少なくとも1つの遺伝子断片または少なくとも1つの遺伝子をコードする核酸配列を含む。例えば、標的DNAライブラリは、1、2、3、4、5、6、7、8、9、10、またはそれ以上の遺伝子または遺伝子断片を含み得る。いくつかの例では、標的DNAライブラリは、経路内の同義遺伝子からの配列または生物体内の全遺伝子からの配列を含む。

40

【0104】

いくつかの例では、修飾されたヌクレアーゼ酵素は、核酸配列修飾能力を有する別の酵素に繫留される。例示的な修飾能力としては、これらに限定されないが、メチル化、脱メチル化、ポリアデニル化、脱アデニル化、脱アミノ化、およびポリウリジニル化 (polyuridylation) が挙げられる。いくつかの例では、複数の変異標的D

50

N A配列を含む標的D N Aライブラリは、変異を結果としてもたらず。いくつかの例では、変異は、欠失、フレームシフト、または標的D N A配列への挿入を導入する。いくつかの例では、変異D N A配列は、遺伝子または遺伝子の断片あたり、少なくとも1のコドンについての変異を結果としてもたらず。いくつかの例では、変異D N A配列は、遺伝子または遺伝子の断片に関連付けられる少なくとも1つの転写調節配列、例えば、プロモーター、U T R、あるいはターミネーター配列についての変異を結果としてもたらず。

【0105】

モデルシステムの遺伝子を標的とするためのg R N Aライブラリ合成

【0106】

本明細書で提供されるのは、本明細書に記載される核酸ライブラリを有するモデルシステムをスクリーニングための方法である。いくつかの事例では、核酸ライブラリは、本明細書に記載されるg R N Aライブラリである。いくつかの例では、核酸ライブラリは、転写された時にg R N A配列の転写をもたらす本明細書に記載されるD N Aライブラリである。モデル生物体の非限定的で例示的なリストが、表4において提供される。

【0107】

【表4】

表4.生物体および遺伝子数

モデルシステム	タンパク質コーディング遺伝子*
シロイヌナズナ	27000
線虫	20000
イヌ科	19000
コナミドリムシ	14000
ゼブラフィッシュ	26000
細胞性粘菌	13000
キイロショウジョウバエ	14000
大腸菌	4300
アカゲザル	22000
ハツカネズミ	20000
アナウサギ	27000
ドブネズミ	22000
出芽酵母	6600
イノシシ	21000
ヒト	21000

*ここでの数は、タンパク質コーディング遺伝子の数を反映し、転移R N Aおよび非コーディングRNAを除外する。Ron Milo & Rob Phillips, Cell Biology by the Numbers 286 (2015)。

【0108】

(R N Aのデノボ合成またはD N Aのデノボ合成のいずれか、その後、(インビボまたはインビトロ)転写をしてg R N Aを生成することによって) g R N Aライブラリが合成され、ここで、ライブラリは、遺伝子あたり複数のg R N A分子を含む。例えば、本明

細書に記載されるライブラリは、遺伝子あたり、1、2、3、4、5、6、7、8、9、10、またはそれ以上の gRNA を含む。いくつかの例では、デノボ合成されたライブラリ内の核酸は、単一遺伝子あたり、少なくともまたは約 3 の非同一の gRNA についての配列をコードする。いくつかの例では、核酸は、単一遺伝子あたり、約 1 ~ 約 10 の非同一の gRNA の範囲で配列をコード化する。いくつかの例では、核酸は、単一遺伝子あたり、少なくともまたは約 1 の非同一の gRNA についての配列をコードする。いくつかの例では、核酸は、単一遺伝子あたり、最大で 10 の非同一の gRNA についての配列をコードする。いくつかの例では、核酸は、単一遺伝子あたり、1 ~ 2、1 ~ 3、1 ~ 4、1 ~ 5、1 ~ 6、1 ~ 7、1 ~ 8、1 ~ 9 または 1 ~ 10 の非同一の gRNA についての配列をコードする。いくつかの例では、gRNA は s gRNA である。いくつかの例では、gRNA は d gRNA である。

10

【0109】

いくつかの例では、本明細書に記載される gRNA ライブラリは、遺伝子または生物体あたり、1 以上の非同一の gRNA を含む。いくつかの例では、gRNA ライブラリは、生物体のための遺伝子あたり、1、2、3、4、5、6、7、8、9、10、またはそれより多くの非同一の gRNA を含む。例示的な生物体としては、これらに限定されないが、シロイヌナズナ、線虫、イヌ科、コナミドリムシ、ゼブラフィッシュ、細胞性粘菌、キイロショウジョウバエ、大腸菌、アカゲザル、ハツカネズミ、アナウサギ、ドブネズミ、出芽酵母、イノシシ、およびヒトが挙げられる。いくつかの例では、gRNA は s gRNA である。いくつかの例では、gRNA は d gRNA である。いくつかの例では、gRNA ライブラリは、生物体の全ゲノムの少なくともまたは約 5 % について、非同一の gRNA を含む。いくつかの例では、gRNA ライブラリは、生物体の全ゲノムの約 5 % ~ 約 100 % について、非同一の gRNA を含む。いくつかの例では、gRNA ライブラリは、生物体の全ゲノムの少なくともまたは約 80 % について、非同一の gRNA を含む。いくつかの例では、s gRNA ライブラリは、生物体の全ゲノムの少なくともまたは約 90 % について、非同一の gRNA を含む。いくつかの例では、gRNA ライブラリは、生物体の全ゲノムの少なくともまたは約 95 % について、非同一の gRNA を含む。いくつかの例では、gRNA ライブラリは、生物体の全ゲノムの少なくともまたは約 100 % について、非同一の gRNA を含む。いくつかの例では、gRNA ライブラリは、生物体の全ゲノムの約 5 % ~ 10 %、5 % ~ 20 %、5 % ~ 30 %、5 % ~ 40 %、5 % ~ 50 %、5 % ~ 60 %、5 % ~ 70 %、5 % ~ 80 %、5 % ~ 90 %、5 % ~ 95 %、5 % ~ 100 %、10 % ~ 20 %、10 % ~ 30 %、10 % ~ 40 %、10 % ~ 50 %、10 % ~ 60 %、10 % ~ 70 %、10 % ~ 80 %、10 % ~ 90 %、10 % ~ 95 %、10 % ~ 100 %、20 % ~ 30 %、20 % ~ 40 %、20 % ~ 50 %、20 % ~ 60 %、20 % ~ 70 %、20 % ~ 80 %、20 % ~ 90 %、20 % ~ 95 %、20 % ~ 100 %、30 % ~ 40 %、30 % ~ 50 %、30 % ~ 60 %、30 % ~ 70 %、30 % ~ 80 %、30 % ~ 90 %、30 % ~ 95 %、30 % ~ 100 %、40 % ~ 50 %、40 % ~ 60 %、40 % ~ 70 %、40 % ~ 80 %、40 % ~ 90 %、40 % ~ 95 %、40 % ~ 100 %、50 % ~ 60 %、50 % ~ 70 %、50 % ~ 80 %、50 % ~ 90 %、50 % ~ 95 %、50 % ~ 100 %、60 % ~ 70 %、60 % ~ 80 %、60 % ~ 90 %、60 % ~ 95 %、60 % ~ 100 %、70 % ~ 80 %、70 % ~ 90 %、70 % ~ 95 %、70 % ~ 100 %、80 % ~ 90 %、80 % ~ 95 %、80 % ~ 100 %、90 % ~ 95 %、90 % ~ 100 %、または 95 % ~ 100 % の非同一の gRNA を含む。いくつかの例では、gRNA ライブラリは、経路内の同義遺伝子からの配列または生物体内の全遺伝子からの配列を含む。gRNA の数は、表 4 列挙される生物体において、遺伝子あたり、少なくとも 2 倍、3 倍、5 倍、または 10 倍を含み得る。いくつかの例では、gRNA ライブラリは、遺伝子、遺伝子の群（例えば 3 - 10 の遺伝子）、経路（例えば 10 - 100 の遺伝子）、またはシャーシ（例えば 100 - 1000 の遺伝子）のうち少なくとも 1 つを標的とする。

20

30

40

【0110】

高度に平行なデノボ核酸合成

50

【0111】

本明細書に記載されるのは、革新的な合成プラットフォームを作るために、シリコン上のナノウェル内のオリゴ核酸合成から遺伝子アセンブリへの末端間プロセスの小型化、並列化、および垂直的統合を利用するプラットフォームアプローチがされる。本明細書に記載される装置は、96ウェルのプレートと同じフットプリントを提供し、シリコン合成プラットフォームは、従来の合成方法と比較して、100~1,000倍スループットを増加させることができ、単一の高度に並列されたラン(run)で、およそ1,000,000までのオリゴ核酸を産生することができる。いくつかの例では、本明細書に記載される単一のシリコンプレートは、約6,100の非同一のオリゴ核酸の合成を提供する。いくつかの例では、非同一のオリゴ核酸の各々は、クラスタ内に位置づけられる。クラスタは、50~500の非同一のオリゴ核酸を含んでもよい。

10

【0112】

いくつかの例では、本明細書に記載されるgRNAライブラリをコードするDNAライブラリは、DNAについて予め決められた配列と比較すると、1:500未満のエラー率を有する。いくつかの例では、本明細書に開示されるデノボオリゴ核酸ライブラリは、DNAについて予め決められた配列と比較すると、1:500、1:1000、1:1500、1:2000、1:3000、1:5000以下、またはそれ未満の総エラー率を有する。いくつかの例では、DNAについて予め決められた配列と比較すると、総エラー率は1:1000未満である。エラー率は、総エラー率または平均エラー率であり得る。

【0113】

20

いくつかの例では、本明細書に記載されるgRNAライブラリをコードするRNAライブラリは、RNAについての予め決められた配列と比較すると、1:500より低いエラー率を有する。いくつかの例では、本明細書に開示されるデノボオリゴ核酸ライブラリは、核酸についての予め決められた配列と比較すると、1:500、1:1000、1:1500、1:2000、1:3000、1:5000、1:10,000、またはそれより低い総エラー率を有する。いくつかの例では、RNAについての予め決められた配列と比較すると、総エラー率は1:1000より低い。

【0114】

基質

【0115】

30

場合により、本明細書には、複数のクラスタを含む基質が記載され、ここで各クラスタはオリゴ核酸の付着及び合成を支持する複数の遺伝子座を含む。本明細書で使用されるような用語「遺伝子座」は、表面から伸長する単一の予め決められた配列をコードするオリゴ核酸に対する支持を提供する構造上の離散的領域を指す。幾つかの実施形態において、遺伝子座は、二次元表面、例えば、実質的に平らな表面上にある。幾つかの実施形態において、遺伝子座は、三次元表面、例えば、ウェル、微小ウェル、チャネル、又はポスト上にある。幾つかの実施形態において、遺伝子座の表面は、オリゴ核酸合成のための少なくとも1つのヌクレオチド、又は好ましくは、オリゴ核酸の集団の合成のための同一のヌクレオチドの集団に付着するために活発に官能化される物質を含む。幾つかの実施形態において、オリゴ核酸は、同じ核酸配列をコードするオリゴ核酸の集団を指す。場合により、基質の表面は、基質の1つ又は複数の表面を包含する。提供されるシステム及び方法を用いてライブラリ内で合成されるオリゴ核酸に関する平均エラー率は多くの場合、1000分の1未満、2000分の1未満、3000分の1未満、或いはそれよりも遥かに少ない。

40

【0116】

幾つかの実施形態において、基質は、共通の支持体上のアドレス可能位置に異なる予め決められた配列を有している複数のオリゴ核酸の合成を支持する表面を含む。幾つかの実施形態において、基質は、50、100、200、400、600、800、1000、1200、1400、1600、1800、2,000; 5,000; 10,000; 20,000; 50,000; 100,000; 200,000; 300,000; 400

50

, 0 0 0 ; 5 0 0 , 0 0 0 ; 6 0 0 , 0 0 0 ; 7 0 0 , 0 0 0 ; 8 0 0 , 0 0 0 ; 9 0 0 , 0 0 0 ; 1 , 0 0 0 , 0 0 0 ; 1 , 2 0 0 , 0 0 0 ; 1 , 4 0 0 , 0 0 0 ; 1 , 6 0 0 , 0 0 0 ; 1 , 8 0 0 , 0 0 0 ; 2 , 0 0 0 , 0 0 0 ; 2 , 5 0 0 , 0 0 0 ; 3 , 0 0 0 , 0 0 0 ; 3 , 5 0 0 , 0 0 0 ; 4 , 0 0 0 , 0 0 0 ; 4 , 5 0 0 , 0 0 0 ; 5 , 0 0 0 , 0 0 0 ; 1 0 , 0 0 0 , 0 0 0、又はそれ以上の非同一のオリゴ核酸の合成のための支持を提供する。場合により、基質は、5 0、1 0 0、2 0 0、4 0 0、6 0 0、8 0 0、1 0 0 0、1 2 0 0、1 4 0 0、1 6 0 0、1 8 0 0、2 , 0 0 0 ; 5 , 0 0 0 ; 1 0 , 0 0 0 ; 2 0 , 0 0 0 ; 5 0 , 0 0 0 ; 1 0 0 , 0 0 0 ; 2 0 0 , 0 0 0 ; 3 0 0 , 0 0 0 ; 4 0 0 , 0 0 0 ; 5 0 0 , 0 0 0 ; 6 0 0 , 0 0 0 ; 7 0 0 , 0 0 0 ; 8 0 0 , 0 0 0 ; 9 0 0 , 0 0 0 ; 1 , 0 0 0 , 0 0 0 ; 1 , 2 0 0 , 0 0 0 ; 1 , 4 0 0 , 0 0 0 ; 1 , 6 0 0 , 0 0 0 ; 1 , 8 0 0 , 0 0 0 ; 2 , 0 0 0 , 0 0 0 ; 2 , 5 0 0 , 0 0 0 ; 3 , 0 0 0 , 0 0 0 ; 3 , 5 0 0 , 0 0 0 ; 4 , 0 0 0 , 0 0 0 ; 4 , 5 0 0 , 0 0 0 ; 5 , 0 0 0 , 0 0 0 ; 1 0 , 0 0 0 , 0 0 0 を超える、又はそれ以上のオリゴ核酸の合成のための支持を提供する。幾つかの実施形態において、オリゴ核酸の少なくとも一部は、同一の配列を有しているか、又は同一の配列で合成されるように構成されている。幾つかの実施形態において、基質は、少なくとも 8 0、9 0、1 0 0、1 2 0、1 5 0、1 7 5、2 0 0、2 2 5、2 5 0、2 7 5、3 0 0、3 2 5、3 5 0、3 7 5、4 0 0、4 2 5、4 5 0、4 7 5、5 0 0、又はそれ以上の塩基を持つオリゴ核酸の成長のための表面環境を提供する。

【 0 1 1 7 】

幾つかの実施形態において、オリゴ核酸は、基質の別々の遺伝子座上で合成され、ここで各遺伝子座はオリゴ核酸の集団の合成を支持する。場合により、各遺伝子座は、別の遺伝子座上で成長させられたオリゴ核酸の集団とは異なる配列を有するオリゴ核酸の集団の合成を支持する。幾つかの実施形態において、基質の遺伝子座は複数のクラスタ内に位置する。幾つかの例において、基質は、少なくとも 1 0、5 0 0、1 0 0 0、2 0 0 0、3 0 0 0、4 0 0 0、5 0 0 0、6 0 0 0、7 0 0 0、8 0 0 0、9 0 0 0、1 0 0 0 0、1 1 0 0 0、1 2 0 0 0、1 3 0 0 0、1 4 0 0 0、1 5 0 0 0、2 0 0 0 0、3 0 0 0 0、4 0 0 0 0、5 0 0 0 0、又はそれ以上のクラスタを含む。幾つかの実施形態において、基質は、2 , 0 0 0 ; 5 , 0 0 0 ; 1 0 , 0 0 0 ; 1 0 0 , 0 0 0 ; 2 0 0 , 0 0 0 ; 3 0 0 , 0 0 0 ; 4 0 0 , 0 0 0 ; 5 0 0 , 0 0 0 ; 6 0 0 , 0 0 0 ; 7 0 0 , 0 0 0 ; 8 0 0 , 0 0 0 ; 9 0 0 , 0 0 0 ; 1 , 0 0 0 , 0 0 0 ; 1 , 1 0 0 , 0 0 0 ; 1 , 2 0 0 , 0 0 0 ; 1 , 3 0 0 , 0 0 0 ; 1 , 4 0 0 , 0 0 0 ; 1 , 5 0 0 , 0 0 0 ; 1 , 6 0 0 , 0 0 0 ; 1 , 7 0 0 , 0 0 0 ; 1 , 8 0 0 , 0 0 0 ; 1 , 9 0 0 , 0 0 0 ; 2 , 0 0 0 , 0 0 0 ; 3 0 0 , 0 0 0 ; 4 0 0 , 0 0 0 ; 5 0 0 , 0 0 0 ; 6 0 0 , 0 0 0 ; 7 0 0 , 0 0 0 ; 8 0 0 , 0 0 0 ; 9 0 0 , 0 0 0 ; 1 , 0 0 0 , 0 0 0 ; 1 , 2 0 0 , 0 0 0 ; 1 , 4 0 0 , 0 0 0 ; 1 , 6 0 0 , 0 0 0 ; 1 , 8 0 0 , 0 0 0 ; 2 , 0 0 0 , 0 0 0 ; 2 , 5 0 0 , 0 0 0 ; 3 , 0 0 0 , 0 0 0 ; 3 , 5 0 0 , 0 0 0 ; 4 , 0 0 0 , 0 0 0 ; 4 , 5 0 0 , 0 0 0 ; 5 , 0 0 0 , 0 0 0 ; 又は 1 0 , 0 0 0 , 0 0 0、或いはそれ以上の別々の遺伝子座を含む。幾つかの実施形態において、基質は約 1 0 , 0 0 0 の別々の遺伝子座を含む。単一のクラスタ内の遺伝子座の量は、異なる実施形態において変動する。場合により、各クラスタは、1、2、3、4、5、6、7、8、9、1 0、2 0、3 0、4 0、5 0、6 0、7 0、8 0、9 0、1 0 0、1 2 0、1 3 0、1 5 0、2 0 0、3 0 0、4 0 0、5 0 0、又はそれ以上の遺伝子座を含む。幾つかの実施形態において、各クラスタは約 5 0 - 5 0 0 の遺伝子座を含む。幾つかの実施形態において、各クラスタは約 1 0 0 - 2 0 0 の遺伝子座を含む。幾つかの実施形態において、各クラスタは約 1 0 9、1 2 1、1 3 0、又は 1 3 7 の遺伝子座を含む。幾つかの実施形態において、各クラスタは約 1 9、2 0、6 1、6 4、又はそれ以上の遺伝子座を含む。

【 0 1 1 8 】

本明細書には、シリコンプレート上で非同一のオリゴ核酸を合成する方法が提供される

。幾つかの例において、シリコンプレートは約 1 - 10、1 - 50、又は 50 - 500 のクラスタを含む。幾つかの例において、シリコンプレートは、約 50、100、250、500、2500、5000、6000、6150、10000、又はそれ以上のクラスタを含む。幾つかの例において、各クラスタは 121 の遺伝子座を含む。幾つかの例において、各クラスタは、約 50 - 5000、50 - 200、100 - 150 の遺伝子座を含む。幾つかの例において、各クラスタは、少なくとも約 50、100、150、200、500、1000、又はそれ以上の遺伝子座を含む。幾つかの例において、単一のプレートは、100、500、10000、20000、30000、50000、100000、500000、700000、1000000、又はそれ以上の遺伝子座を含む。

【0119】

幾つかの実施形態において、基質上で合成された別々のオリゴ核酸の数は、基質における利用可能な別々の遺伝子座の数に左右され得る。幾つかの実施形態において、基質のクラスタ内の遺伝子座の密度は、 1 mm^2 当たり少なくとも又は約 1 の遺伝子座、 1 mm^2 当たり 10 の遺伝子座、 1 mm^2 当たり 25 の遺伝子座、 1 mm^2 当たり 50 の遺伝子座、 1 mm^2 当たり 65 の遺伝子座、 1 mm^2 当たり 75 の遺伝子座、 1 mm^2 当たり 100 の遺伝子座、 1 mm^2 当たり 130 の遺伝子座、 1 mm^2 当たり 150 の遺伝子座、 1 mm^2 当たり 175 の遺伝子座、 1 mm^2 当たり 200 の遺伝子座、 1 mm^2 当たり 300 の遺伝子座、 1 mm^2 当たり 400 の遺伝子座、 1 mm^2 当たり 500 の遺伝子座、 1 mm^2 当たり 1,000 の遺伝子座、又はそれ以上である。場合により、基質は、 1 mm^2 から約 500 mm^2 当たり約 10 の遺伝子座、 1 mm^2 から約 400 mm^2 当たり約 25 の遺伝子座、 1 mm^2 から約 500 mm^2 当たり約 50 の遺伝子座、 1 mm^2 から約 500 mm^2 当たり約 100 の遺伝子座、 1 mm^2 から約 500 mm^2 当たり約 150 の遺伝子座、 1 mm^2 から約 250 mm^2 当たり約 10 の遺伝子座、 1 mm^2 から約 250 mm^2 当たり約 50 の遺伝子座、 1 mm^2 から約 200 mm^2 当たり約 10 の遺伝子座、 1 mm^2 から約 200 mm^2 当たり約 50 の遺伝子座を含む。幾つかの実施形態において、クラスタ内の 2 つの隣接した遺伝子座の中心間の距離は、約 $10\text{ }\mu\text{m}$ から約 $500\text{ }\mu\text{m}$ 、約 $10\text{ }\mu\text{m}$ から約 $200\text{ }\mu\text{m}$ 、又は約 $10\text{ }\mu\text{m}$ から約 $100\text{ }\mu\text{m}$ である。場合により、隣接した遺伝子座の 2 つの中心間の距離は、約 $10\text{ }\mu\text{m}$ 、 $20\text{ }\mu\text{m}$ 、 $30\text{ }\mu\text{m}$ 、 $40\text{ }\mu\text{m}$ 、 $50\text{ }\mu\text{m}$ 、 $60\text{ }\mu\text{m}$ 、 $70\text{ }\mu\text{m}$ 、 $80\text{ }\mu\text{m}$ 、 $90\text{ }\mu\text{m}$ 、又は $100\text{ }\mu\text{m}$ より長い。場合により、2 つの隣接した遺伝子座の中心間の距離は、約 $200\text{ }\mu\text{m}$ 、 $150\text{ }\mu\text{m}$ 、 $100\text{ }\mu\text{m}$ 、 $80\text{ }\mu\text{m}$ 、 $70\text{ }\mu\text{m}$ 、 $60\text{ }\mu\text{m}$ 、 $50\text{ }\mu\text{m}$ 、 $40\text{ }\mu\text{m}$ 、 $30\text{ }\mu\text{m}$ 、 $20\text{ }\mu\text{m}$ 、又は $10\text{ }\mu\text{m}$ 未満である。幾つかの例において、各遺伝子座の幅は、約 $0.5\text{ }\mu\text{m}$ 、 $1\text{ }\mu\text{m}$ 、 $2\text{ }\mu\text{m}$ 、 $3\text{ }\mu\text{m}$ 、 $4\text{ }\mu\text{m}$ 、 $5\text{ }\mu\text{m}$ 、 $6\text{ }\mu\text{m}$ 、 $7\text{ }\mu\text{m}$ 、 $8\text{ }\mu\text{m}$ 、 $9\text{ }\mu\text{m}$ 、 $10\text{ }\mu\text{m}$ 、 $20\text{ }\mu\text{m}$ 、 $30\text{ }\mu\text{m}$ 、 $40\text{ }\mu\text{m}$ 、 $50\text{ }\mu\text{m}$ 、 $60\text{ }\mu\text{m}$ 、 $70\text{ }\mu\text{m}$ 、 $80\text{ }\mu\text{m}$ 、 $90\text{ }\mu\text{m}$ 、又は $100\text{ }\mu\text{m}$ である。場合により、各遺伝子座の幅は、約 $0.5\text{ }\mu\text{m}$ から $100\text{ }\mu\text{m}$ 、約 $0.5\text{ }\mu\text{m}$ から $50\text{ }\mu\text{m}$ 、約 $10\text{ }\mu\text{m}$ から $75\text{ }\mu\text{m}$ 、又は約 $0.5\text{ }\mu\text{m}$ から $50\text{ }\mu\text{m}$ である。

【0120】

幾つかの実施形態において、基質内のクラスタの密度は、 100 mm^2 当たり少なくとも又は約 1 のクラスタ、 10 mm^2 当たり 1 のクラスタ、 5 mm^2 当たり 1 のクラスタ、 4 mm^2 当たり 1 のクラスタ、 3 mm^2 当たり 1 のクラスタ、 2 mm^2 当たり 1 のクラスタ、 1 mm^2 当たり 1 のクラスタ、 1 mm^2 当たり 2 のクラスタ、 1 mm^2 当たり 3 のクラスタ、 1 mm^2 当たり 4 のクラスタ、 1 mm^2 当たり 5 のクラスタ、 1 mm^2 当たり 10 のクラスタ、 1 mm^2 当たり 50 のクラスタ、又はそれ以上である。幾つかの実施形態において、基質は、 10 mm^2 当たり約 1 のクラスタから 1 mm^2 当たり約 10 のクラスタを含む。幾つかの実施形態において、2 つの隣接したクラスタの中心間の距離は、約 $50\text{ }\mu\text{m}$ 、 $100\text{ }\mu\text{m}$ 、 $200\text{ }\mu\text{m}$ 、 $500\text{ }\mu\text{m}$ 、 $1000\text{ }\mu\text{m}$ 、 $2000\text{ }\mu\text{m}$ 、又は $5000\text{ }\mu\text{m}$ 未満である。場合により、2 つの隣接したクラスタの中心間の距離は、約 $50\text{ }\mu\text{m}$ から約 $100\text{ }\mu\text{m}$ 、約 $50\text{ }\mu\text{m}$ から約 $200\text{ }\mu\text{m}$ 、約 $50\text{ }\mu\text{m}$ から約 $300\text{ }\mu\text{m}$ 、約 $50\text{ }\mu\text{m}$ から約 $500\text{ }\mu\text{m}$ 、及び約 $100\text{ }\mu\text{m}$ から約 $2000\text{ }\mu\text{m}$ である。場合により、2 つの隣接したクラスタの中心間の距離は、約 0.05 mm から約 50 mm 、約 0.05 mm

10

20

30

40

50

mから約10 mm、約0.05 mmから約5 mm、約0.05 mmから約4 mm、約0.05 mmから約3 mm、約0.05 mmから約2 mm、約0.1 mmから10 mm、約0.2 mmから10 mm、約0.3 mmから約10 mm、約0.4 mmから約10 mm、約0.5 mmから10 mm、約0.5 mmから約5 mm、又は約0.5 mmから約2 mmである。場合により、各クラスは、約0.5から2 mm、約0.5から1 mm、又は約1から2 mmの断面を有している。場合により、各クラスは、約0.5、0.6、0.7、0.8、0.9、1、1.1、1.2、1.3、1.4、1.5、1.6、1.7、1.8、1.9、又は2 mmの断面を有している。場合により、各クラスは、約0.5、0.6、0.7、0.8、0.9、1、1.1、1.15、1.2、1.3、1.4、1.5、1.6、1.7、1.8、1.9、又は2 mmの内部断面を有している。

10

【0121】

幾つかの実施形態において、基質は、例えば、約100~200 mm×約50~150 mmの、およその標準の96ウェルプレートのサイズであり得る。幾つかの実施形態において、基質は、約1000 mm、500 mm、450 mm、400 mm、300 mm、250 mm、200 mm、150 mm、100 mm、又は50 mm以下の直径を有している。幾つかの実施形態において、基質の直径は、約25 mmから1000 mm、約25 mmから約800 mm、約25 mmから約600 mm、約25 mmから約500 mm、約25 mmから約400 mm、約25 mmから約300 mm、又は約25 mmから約200である。基質サイズの非限定的な例は、約300 mm、200 mm、150 mm、130 mm、100 mm、76 mm、51 mm、及び25 mmを含む。幾つかの実施形態において、基質は、少なくとも約100 mm²；200 mm²；500 mm²；1,000 mm²；2,000 mm²；5,000 mm²；10,000 mm²；12,000 mm²；15,000 mm²；20,000 mm²；30,000 mm²；40,000 mm²；50,000 mm²又はそれ以上の平面の表面積を有している。幾つかの実施形態において、基質の厚さは、約50 mmから約2000 mm、約50 mmから約1000 mm、約100 mmから約1000 mm、約200 mmから約1000 mm、又は約250 mmから約1000 mmである。基質の厚さの非限定的な例は、275 mm、375 mm、525 mm、625 mm、675 mm、725 mm、775 mm、及び925 mmを含む。場合により、基質の厚さは、直径によって変わり、基質の組成物に左右される。例えば、シリコン以外の物質を含む基質は、同じ直径のシリコン基質とは異なる厚さを有している。基質の厚さは、使用される物質の機械強度によって決定され、取り扱い中に割れることなくそれ自体の重量を支えるのに十分でなければならない。

20

30

【0122】

表面物質

【0123】

本明細書に提供される基質、装置、及びリアクターは、本明細書に記載される方法及び組成物に適したあらゆる様々な物質から作り上げられる。特定の実施形態において、低レベルのヌクレオチド結合を示すために、基質物質が作り上げられる。場合により、基質物質は、高レベルのヌクレオチド結合を示す別々の表面を生成するために修正される。幾つかの実施形態において、基質物質は、可視光及び/又はUV光に対して透過性である。幾つかの実施形態において、基質物質は、十分に導電性である、例えば、基質の全て又は一部にわたって均一な電場を形成することができる。幾つかの実施形態において、導電性材料は電気接地(electrical ground)に接続される。場合により、基質は、熱伝導性であるか又は断熱される。場合により、物質は、化学的又は生化学的な反応、例えばオリゴ核酸合成反応プロセスを支持するために耐薬品性及び耐熱性である。幾つかの実施形態において、基質は可撓性材料を含む。可撓性材料は、限定されないが、修飾ナイロン、非修飾ナイロン、ニトロセルロース、ポリプロピレンなどを含む。幾つかの実施形態において、基質は剛性材料を含む。剛性材料は、限定されないが、ガラス、石英ガラス(fused silica)、シリコン、二酸化ケイ素、窒化ケイ素、プラスチック(例えば、ポリテトラフルオロエチレン、ポリプロピレン、ポリスチレン、ポリカーボネート

40

50

、及びそれらの混合物など）、及び金属（例えば、金、白金など）を含む。幾つかの実施形態において、基質は、シリコン、ポリスチレン、アガロース、デキストラン、セルロース系ポリマー、ポリアクリルアミド、ポリジメチルシロキサン（PDMS）、ガラス、又はそれらの任意の組み合わせを含む材料から作り上げられる。場合により、基質は、本明細書にリストされた材料又は当該技術分野において既知の他の適切な材料の組み合わせで製造される。

【0124】

表面のアーキテクチャ

【0125】

様々な実施形態において、基質は隆起した及び／又は沈降した特徴を含む。そのような特徴を有している１つの利点は、オリゴ核酸合成を支持する表面積の増大である。幾つかの実施形態において、隆起した及び／又は沈降した特徴を有する基質は、三次元基質と呼ばれる。場合により、三次元基質は１つ以上のチャンネルを含む。場合により、１つ以上の遺伝子座はチャンネルを含む。場合により、チャンネルは、オリゴ核酸シンセサイザーなどの堆積装置を介した試薬の堆積に利用可能である。場合により、試薬及び／又は流体は、１つ以上のチャンネルと流体連通してより大きなウェルに集まる。例えば、基質は、クラスタを有する複数の遺伝子座に対応する複数のチャンネルを含み、複数のチャンネルは、クラスタの１つのウェルと流体連通している。

幾つかの方法において、オリゴ核酸のライブラリは、クラスタの複数の遺伝子座において合成される。

【0126】

幾つかの実施形態において、その構造は、表面上のオリゴ核酸合成に関する流れの及び物質移動経路の制御を可能にするように構成されている。幾つかの実施形態において、基質の構成は、オリゴ核酸合成中の物質移動経路、化学暴露時間、及び／又は洗浄効果の制御並びにその分布までも可能にする。幾つかの実施形態において、基質の構成は、例えば、成長しているオリゴ核酸による排除体積が、オリゴ核酸の成長に利用可能な又は適切な最初に利用可能な体積の５０、４５、４０、３５、３０、２５、２０、１５、１４、１３、１２、１１、１０、９、８、７、６、５、４、３、２、１％を超えて、又はそれ以下を占めないほど十分な、オリゴ核酸の成長に対する体積を提供することによって、奏効率の増大を可能にする。幾つかの実施形態において、三次元構造は、化学暴露の急速な交換を可能にするために流体の流れの管理を可能にする。

【0127】

幾つかの実施形態において、物理構造によって隔離が達成される。幾つかの実施形態において、オリゴ核酸合成に対する能動領域及び受動領域を発生させる表面の差次的な官能化によって、隔離が達成される。差次的な官能化はまた、基質表面にわたる疎水性を変え、それによって、堆積した試薬の水滴（beading）又は湿りを引き起こす水接触角の効果を作り出すことによって達成される。より大きな構造を利用することで、飛び散り（splashing）及び隣接するスポットの試薬による別々のオリゴ核酸合成位置の相互汚染を減らすことができる。場合により、オリゴ核酸シンセサイザーなどの装置が、別々のオリゴ核酸合成位置のための試薬を堆積させるために使用される。三次元の特徴を有する基質は、低いエラー率（例えば、約１：５００、１：１０００、１：１５００、１：２，０００、１：３，０００、１：５，０００、又は１：１０，０００未満）で多数のオリゴ核酸（例えば約１０，０００を超える）の合成を可能にする方法で構成される。幾つかの事例では、基質は、 1 mm^2 当たり約１、５、１０、２０、３０、４０、５０、６０、７０、８０、１００、１１０、１２０、１３０、１４０、１５０、１６０、１７０、１８０、１９０、２００、３００、４００又は５００、或いはそれらを超える特徴の密度を備えた特徴を含む。

【0128】

基質のウェルは、基質の別のウェルと同じ又は異なる幅、高さ、及び／又は体積を有し得る。基質のチャンネルは、基質の別のチャンネルと同じ又は異なる幅、高さ、及び／又は体

積を有し得る。幾つかの実施形態において、クラスタの直径、又はクラスタを含むウェルの直径、或いはその両方は、約 0.05 mm から約 50 mm、約 0.05 mm から約 10 mm、約 0.05 mm から約 5 mm、約 0.05 mm から約 4 mm、約 0.05 mm から約 3 mm、約 0.05 mm から約 2 mm、約 0.05 mm から約 1 mm、約 0.05 mm から約 0.5 mm、約 0.05 mm から約 0.1 mm、約 0.1 mm から 10 mm、約 0.2 mm から 10 mm、約 0.3 mm から約 10 mm、約 0.4 mm から約 10 mm、約 0.5 mm から 10 mm、約 0.5 mm から約 5 mm、又は約 0.5 mm から約 2 mm である。幾つかの実施形態において、クラスタ、ウェル、又はその両方の直径は、約 5 mm、4 mm、3 mm、2 mm、1 mm、0.5 mm、0.1 mm、0.09 mm、0.08 mm、0.07 mm、0.06 mm 又は 0.05 mm、或いはそれ未満である。幾つかの実施形態において、クラスタ、ウェル、又はその両方の直径は約 1.0 ~ 1.3 mm である。幾つかの実施形態において、クラスタ、ウェル、又はその両方の直径は約 1.150 mm である。幾つかの実施形態において、クラスタ、ウェル、又はその両方の直径は約 0.08 mm である。クラスタの直径は、二次元又は三次元の基質内のクラスタを指す。

【0129】

幾つかの実施形態において、ウェルの高さは、約 20 μm から約 1000 μm 、約 50 μm から約 1000 μm 、約 100 μm から約 1000 μm 、約 200 μm から約 1000 μm 、約 300 μm から約 1000 μm 、約 400 μm から約 1000 μm 、又は約 500 μm から約 1000 μm である。場合により、ウェルの高さは、約 1000 μm 未満、約 900 μm 未満、約 800 μm 未満、約 700 μm 未満、又は約 600 μm 未満である。

【0130】

幾つかの実施形態において、基質は、クラスタ内の複数の遺伝子座に対応する複数のチャネルを含み、チャネルの高さ又は深さは、約 5 μm から約 500 μm 、約 5 μm から約 400 μm 、約 5 μm から約 300 μm 、約 5 μm から約 200 μm 、約 5 μm から約 100 μm 、約 5 μm から約 50 μm 、又は約 10 μm から約 50 μm である。場合により、チャネルの高さは、100 μm 未満、80 μm 未満、60 μm 未満、40 μm 未満又は 20 μm 未満である。

【0131】

幾つかの実施形態において、チャネル、遺伝子座（例えば、実質的に平面の基板における）、又はチャネル及び遺伝子座の両方の直径（例えば、遺伝子座がチャネルに対応する三次元構造基質における）は、約 1 μm から約 1000 μm 、約 1 μm から約 500 μm 、約 1 μm から約 200 μm 、約 1 μm から約 100 μm 、約 5 μm から約 100 μm 、又は約 10 μm から約 100 μm 、例えば、約 90 μm 、80 μm 、70 μm 、60 μm 、50 μm 、40 μm 、30 μm 、20 μm 、又は 10 μm である。幾つかの実施形態において、チャネル、遺伝子座、又はチャネル及び遺伝子座の両方の直径は、約 100 μm 、90 μm 、80 μm 、70 μm 、60 μm 、50 μm 、40 μm 、30 μm 、20 μm 又は 10 μm 未満である。幾つかの実施形態において、2つの隣接チャネル、遺伝子座、又はチャネル及び遺伝子座の中心間の距離は、約 1 μm から約 500 μm 、約 1 μm から約 200 μm 、約 1 μm から約 100 μm 、約 5 μm から約 200 μm 、約 5 μm から約 100 μm 、約 5 μm から約 50 μm 、又は約 5 μm から約 30 μm 、例えば、約 20 μm である。

【0132】

表面修飾

【0133】

様々な実施形態において、基質表面又は基質表面の選択部位或いは領域の1つ以上の化学的及び/又は物理的な特性を変更する加算或いは減算のプロセスによる化学的及び/又は物理的な変更のために、表面修飾が利用される。例えば、表面修飾は、限定されないが、(1)表面の湿潤性を変更すること、(2)表面を官能化すること、つまり、表面官能基を提供する、修飾する、又は置換すること、(3)表面を脱官能基化すること、つまり

、表面官能基を除去すること、(4) そうでなければ、例えばエッチングによって、表面の化学組成を変更すること、(5) 表面粗さを増大又は低減すること、(6) 表面上にコーティング、例えば、表面の湿潤性とは異なる湿潤性を示すコーティングを提供すること、及び/又は(7) 表面上に粒子を堆積させることを含む。

【0134】

場合により、(接着プロモーターと呼ばれる)表面の上部の化学層の追加は、基質の表面上の遺伝子座の構造化されたパターン化を促進する。接着促進の適用のための典型的な表面は、限定されないが、ガラス、シリコン、二酸化ケイ素及び窒化ケイ素を含む。場合により、接着プロモーターは、高い表面エネルギーを有する化学物質である。幾つかの実施形態において、基質の表面上に第2の化学層が堆積される。場合により、第2の化学層は、低い表面エネルギーを有している。場合により、表面上にコーティングされた化学層の表面エネルギーは、表面上での液滴の局在化を支持する。選択されるパターン化の配置によって、遺伝子座の接近及び/又は遺伝子座での流体接触の領域は変更可能である。

【0135】

幾つかの実施形態において、例えばオリゴ核酸合成のために、核酸又は他の部分が堆積される、基質表面、或いは分解された遺伝子座は、滑らか又は実質的に平面である(例えば、二次元)か、或いは隆起した又は沈降した特徴などの不規則性を有している(例えば三次元の特徴)。幾つかの実施形態において、基質表面は、化合物の1つ以上の異なる層で修飾される。対象のそのような修飾層は、限定されないが、金属、金属酸化物、ポリマー、小さな有機分子などの無機層及び有機層を含む。非限定的なポリマー層は、ペプチド、タンパク質、核酸又はそれらの模倣物(例えば、ペプチド核酸など)、多糖類、リン脂質、ポリウレタン、ポリエステル、ポリカーボネート、ポリ尿素、ポリアミド、ポリエチレンアミン、ポリアリーレンスルフィド、ポリシロキサン、ポリイミド、ポリアセテート、及び本明細書に記載される又はそうでなければ当該技術分野に既知の他の適切な化合物を含む。場合により、ポリマーはヘテロポリマーである。場合により、ポリマーはホモポリマーである。場合により、ポリマーは、官能性部分を含むか又は結合される。

【0136】

幾つかの実施形態において、基質の分解された遺伝子座は、表面エネルギーを増大させる及び/又は低減させる1つ以上の部分で官能化される。場合により、部分は化学的に不活性である。場合により、部分は、望ましい化学反応、例えば、オリゴ核酸合成反応における1つ以上のプロセスを支持するように構成されている。表面の表面エネルギー、即ち疎水性は、表面上へと付着するヌクレオチドの親和性を決定するための因子である。幾つかの実施形態において、基質の官能化のための方法は、(a) 二酸化ケイ素を含む表面を有する基質を提供する工程; 及び(b) 本明細書に記載される又はそうでなければ当該技術分野で既知の適切なシラン化剤、例えば、有機官能性アルコキシシラン分子を使用して、表面をシラン処理する工程を含む。場合により、有機官能性アルコキシシラン分子は、ジメチルクロロ-オクトデシル-シラン、メチルジクロロ-オクトデシル-シラン、トリクロロ-オクトデシル-シラン、トリメチル-オクトデシル-シラン、トリエチル-オクトデシル-シラン、又はそれらの任意の組み合わせを含む。幾つかの実施形態において、基質表面は、ポリエチレン/ポリプロピレン(ヒドロキシアシル表面へのガンマ線照射又はクロム酸酸化、及び還元によって官能化)、高度に架橋されたポリスチレン-ジビニルベンゼン(クロロメチル化によって誘導体化され、ベンジルアミン官能面にアミノ化された)、ナイロン(末端のアミノヘキシル基は直接反応性である)で官能化されるか、又は還元されたポリテトラフルオロエチレンでエッチングされる。他の方法及び官能化剤は、米国特許第5,474,796号に記載され、これは参照によってその全体が本明細書に組み込まれる。

【0137】

幾つかの実施形態において、基質表面は、典型的に基質表面上に存在する反応性の親水性部分を介して、基質表面にシランを結合させるのに有効な反応条件下で、シランの混合物を含有している誘導体化組成物との接触によって官能化される。シラン処理は、一般に

、自己組織化を介して有機官能性アルコキシシラン分子で表面を覆う。当該技術分野において現在知られているように、例えば、表面エネルギーを低下させる又は増大させるための様々なシロキサン官能化試薬が更に使用され得る。有機官能性のアルコキシシランは、その有機官能基に応じて分類される。シロキサン官能化試薬の非限定的な例は、ヒドロキシアルキルシロキサン（シリル化表面（silylate surface）、ジボランで官能化し、過酸化水素によりアルコールを参加させる）、ジオール（ヒドロキシアルキル）シロキサン（シリル化表面、ジオールへと加水分解）、アミノアルキルシロキサン（アミンは中間の官能化工程を必要としない）、グリシドキシシラン（3 - グシドキシプロピル - ジメチル - エトキシシラン、グリシドキシ - トリメトキシシラン、メルカプトシラン（3 - メルカプトプロピル - トリメトキシシラン、3 - 4 エポキシシクロヘキシル - エチルトリメトキシシラン、又は3 - メルカプトプロピル - メチル - ジメトキシシラン）、ピシクロヘプタヘニル - トリクロロシラン、ブチル - アルデヒド（aldehyde） - トリメトキシシラン、又は二量体の第2のアミノアルキルシロキサンを含む。例示的なヒドロキシアルキルシロキサンは、3 - ヒドロキシプロピルへと変わるアリルトリクロロクロロシラン、又は8 - ヒドロキシオクチルへと変わる7 - オクタ - 1 - エニルトリクロロクロロシランを含む。ジオール（ジヒドロキシアルキル）シロキサンは、グリシジルトリメトキシシラン由来（2, 3 - ジヒドロキシプロピルオキシ）プロピル（GOPS）を含む。アミノアルキルシロキサンは、3 - アミノプロピル（3 - アミノプロピル - トリエトキシシラン、3 - アミノプロピル - ジエトキシ - メチルシラン、3 - アミノプロピル - ジメチル - エトキシシラン、又は3 - アミノプロピル - トリメトキシシラン）へと変わる3 - アミノプロピルトリメトキシシランを含む。例示的な二量体の第2のアミノアルキルシロキサンは、ビス（シリルオキシプロピル）アミンへと変わるビス（3 - トリメトキシシリルプロピル）アミンを含む。幾つかの実施形態において、官能化剤は、11 - アセトキシウンデシルトリエトキシシラン、n - デシルトリエトキシシラン、（3 - アミノプロピル）トリメトキシシラン、（3 - アミノプロピル）トリエトキシシラン、グリシジロキシプロピル/トリメトキシシラン、及びN - （3 - トリエトキシシリルプロピル） - 4 - ヒドロキシブチルアミドを含む。

【0138】

オリゴ核酸合成

【0139】

オリゴ核酸合成のための方法は、様々な実施形態において、ホスホロアミダイト化学物質を含むプロセスを含む。幾つかの実施形態において、オリゴ核酸合成は、塩基をホスホラミダイトと結合する工程を含む。幾つかの実施形態において、オリゴ核酸合成は、結合条件下でホスホラミダイトの堆積によって塩基を結合する工程を含み、ここで同じ塩基が、1回より多く、即ち、二重の結合でホスホラミダイトと共に随意に堆積される。幾つかの実施形態において、オリゴ核酸合成は、未反応部位のキャッピングを含む。場合により、キャッピングは随意である。幾つかの実施形態において、オリゴ核酸合成は酸化を含む。幾つかの実施形態において、オリゴ核酸合成はデプロッキング又は脱トリチル化を含む。幾つかの実施形態において、オリゴ核酸合成は硫化を含む。場合により、オリゴ核酸合成は、酸化又は硫化のいずれかを含む。幾つかの実施形態において、オリゴ核酸合成反応中の1つ又は各々の工程間で、基質は、例えば、テトラゾール又はアセトニトリルを使用して洗浄される。ホスホラミダイト合成方法における任意の1工程に対する時間枠は、約2分、1分、50秒、40秒、30秒、20秒及び10秒未満を含む。

【0140】

ホスホラミダイト方法を使用するオリゴ核酸合成は、亜リン酸塩トリエステル結合の形成のために成長しているオリゴ核酸鎖へのホスホラミダイト構築ブロック（例えば、ヌクレオシドホスホラミダイト）のその後の追加を含む。ホスホラミダイトオリゴ核酸合成は、3'方向から5'方向に進む。ホスホラミダイトオリゴ核酸合成は、1つの合成サイクル当たり、成長している核酸鎖への1つのヌクレオチドの追加の制御を可能にする。幾つかの実施形態において、各合成サイクルは結合工程を含む。ホスホラミダイト結合は、活

性化されたヌクレオシドホスホラミダイトと、例えばリンカーを介して基質に結合されたヌクレオシドとの間の亜リン酸塩トリエステル結合の形成を含む。幾つかの実施形態において、ヌクレオシドホスホラミダイトは、活性化された基質に提供される。幾つかの実施形態において、ヌクレオシドホスホラミダイトは、活性化因子と共に基質に提供される。幾つかの実施形態において、ヌクレオシドホスホラミダイトは、基質に結合されたヌクレオシドよりも1.5、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、50、60、70、80、90、100倍、又はそれ以上の過剰量で、基質に提供される。幾つかの実施形態において、ヌクレオシドホスホラミダイトの追加は、無水環境において、例えば、無水アセトニトリルにおいて実行される。ヌクレオシドホスホラミダイトの追加に続いて、基質は随意に洗浄される。幾つかの実施形態において、結合工程は、随意に、基質へのヌクレオシドホスホラミダイトの追加間の洗浄工程と共に、追加で1回以上繰り返される。幾つかの実施形態において、本明細書で使用されるオリゴ核酸合成方法は、1、2、3又はそれ以上の連続結合工程を含む。結合前に、多くの場合において、基質に結合されたヌクレオシドは、保護基の除去によって脱保護され、ここで、保護基は重合を防ぐように機能する。一般的な保護基は、4,4'-ジメトキシトリチル(DMT)である。

【0141】

結合に続いて、ホスホラミダイトオリゴ核酸合成方法は、随意にキャッピング工程を含む。キャッピング工程では、成長しているオリゴ核酸は、キャッピング剤で処置される。キャッピング工程は、更なる鎖伸長からの結合後に未反応の基質に結合した5'-OH基をブロックするのに有用であり、これによって、内部塩基欠失(internal base deletions)でのオリゴ核酸の形成を防ぐ。更に、1H-テトラゾールで活性化されたホスホラミダイトは、小さい程度まで、グアノシンのO6位置と反応し得る。理論に縛られることなく、I₂/水での酸化で、この副産物は、恐らくO6-N7遊走を介して、脱プリン化を受けることもある。脱プリン部位は、最終的に、オリゴ核酸の最終的な脱保護の間に切断され、これにより全長産物の収率が低下する。O6修飾は、I₂/水での酸化前にキャッピング試薬による処置によって除去され得る。幾つかの実施形態において、オリゴ核酸合成の間にキャッピング工程を含めることで、キャッピングなしでの合成と比較して、エラー率は低下する。一例として、キャッピング工程は、無水酢酸と1-メチルイミダゾールとの混合物で、基質に結合したオリゴ核酸を処置することを含む。キャッピング工程に続いて、基質は随意に洗浄される。

【0142】

幾つかの実施形態において、ヌクレオシドホスホラミダイトの追加に続いて、及び随意にキャッピングと1以上の洗浄工程の後に、基質に結合した成長している核酸が酸化される。酸化工程は、亜リン酸塩トリエステルが、自然発生のリン酸ジエステルのヌクレオシド間の結合の保護された前駆体である、四配位リン酸塩トリエステルへと酸化される。場合により、成長しているオリゴ核酸の酸化は、随意に弱塩基(例えば、ピリジン、ルチジン、コリジン)の存在下で、ヨウ素及び水による処置によって達成される。酸化は、例えば、tert-ブチルヒドロペルオキシド又は(1S)-(+)-(10-カンファースルホニル)-オキサジリジン(CSO)を使用して、無水条件下で実行され得る。幾つかの方法では、キャッピング工程は、酸化に続いて実行される。持続し得る酸化からの残留水が、後の結合を阻害することができるため、第2のキャッピング工程は基質の乾燥を可能にする。酸化に続いて、基質及び成長しているオリゴ核酸は随意に洗浄される。幾つかの実施形態において、酸化の工程は、オリゴヌクレオチドホスホロチオエートを得るための硫化工程と置き換えられ、ここでキャッピング工程は硫化後に実行され得る。限定されないが、3-(ジメチルアミノメチリデン)アミノ)-3H-1,2,4-ジチアゾール-3-チオン、DDTT、Beaucage試薬としても知られている3H-1,2-ベンゾジチオール-3-オン1,1-ジオキシド、及びN,N,N',N'-テトラエチルチウラムジスルフィド(TETD)を含む、多くの試薬が、効率的な硫黄移動を行うことができる。

【 0 1 4 3 】

ヌクレオシド取り込みの後のサイクルを、結合を介して生じさせるために、基質に結合した成長しているオリゴ核酸の保護された 5' 末端は除去され、その結果、一次ヒドロキシル基が次のヌクレオシドホスホラミダイトと反応する。幾つかの実施形態において、保護基はDMTであり、デブロッキングがジクロロメタン中でトリクロロ酢酸により生じる。長時間にわたり、又は推奨された酸の溶液よりも強力なもので、脱トリチル化を行うことは、固形支持体に結合したオリゴヌクレオチドの脱プリン化の増大につながり、故に望ましい全長産物の収率が低下することがある。本明細書に記載の本発明の方法及び組成物は、望ましくない脱プリン反応を制限する制御されたデブロッキングの条件を提供する。場合により、基質に結合したオリゴ核酸は、デブロッキング後に洗浄される。場合により、デブロッキング後の効率的な洗浄は、低いエラー率を有する合成オリゴ核酸に寄与する。

10

【 0 1 4 4 】

オリゴ核酸の合成のための方法には、典型的に下記工程の一連の繰り返し (i t e r a t i n g s e q u e n c e) が伴う：活性化された表面、リンカー、又は以前に脱保護された単量体と結合するために、保護された単量体の活発に官能化された表面 (例えば遺伝子座) への適用；後に適用される保護された単量体と反応するような、適用された単量体の脱保護；及び結合のための別の保護された単量体の適用。1 以上の中間工程は、酸化又は硫化を含む。場合により、1 以上の洗浄工程は、工程の 1 つ又は全てに先行する、又はそれらに続いて行われる。

20

【 0 1 4 5 】

ホスホラミダイトベースのオリゴ核酸合成のための方法は、一連の化学工程を含む。幾つかの実施形態において、合成方法の 1 以上の工程は、試薬のサイクリングを含み、ここで、方法の 1 以上の工程は、この工程に有用な試薬の基質への適用を含む。例えば、試薬は、一連の液体堆積及び真空乾燥の工程によって循環させられる。ウェル、マイクロウェル、チャンネルなどの三次元の特徴を含む基質のために、試薬は、随意にウェル及び/又はチャンネルを介して基質の 1 つ以上の領域に通される。

【 0 1 4 6 】

本明細書に記載される方法及び/又は基質を使用して合成されたオリゴ核酸は、様々な実施形態において、少なくとも約 20、30、40、50、60、70、75、80、90、100、120、150、又はそれ以上の塩基を含む。幾つかの実施形態において、少なくとも約 1 p m o l、10 p m o l、20 p m o l、30 p m o l、40 p m o l、50 p m o l、60 p m o l、70 p m o l、80 p m o l、90 p m o l、100 p m o l、150 p m o l、200 p m o l、300 p m o l、400 p m o l、500 p m o l、600 p m o l、700 p m o l、800 p m o l、900 p m o l、1 n m o l、5 n m o l、10 n m o l、100 n m o l、又はそれ以上のオリゴ核酸は、遺伝子座内で合成される。本明細書で提供される、表面上でのオリゴ核酸合成の方法は、高速での合成を可能にする。一例として、1 時間につき少なくとも 3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50、55、60、70、80、90、100、125、150、175、200 のヌクレオチド、又はそれ以上が合成される。ヌクレオチドは、アデニン、グアニン、チミン、シトシン、ウリジンの構築ブロック、又はそれらのアナログ/修飾されたバージョンを含む。幾つかの実施形態において、オリゴ核酸のライブラリは基質上で平行に合成される。例えば、約又は少なくとも約 100；1,000；10,000；100,000；1,000,000；2,000,000；3,000,000；4,000,000；又は 5,000,000 の分解された遺伝子座を含む基質は、少なくとも同じ数の別個のオリゴ核酸の合成を支援することができ、ここで、別個の配列をコードするオリゴ核酸が、分解された遺伝子座の上で合成される。幾つかの実施形態において、オリゴ核酸のライブラリは、約 3 か月、2 か月、1 か月、3 週、15 日、14 日、13 日、12 日、11 日、10 日、9 日、8 日

30

40

50

、7日、6日、5日、4日、3日、2日、24時間未満で、本明細書に記載される低いエラー率を伴う基質上で合成される。幾つかの実施形態において、本明細書に記載される基質及び方法を用いて低いエラー率で合成されるオリゴ核酸ライブラリから組み立てられる、より大きな核酸は、約3か月、2か月、1か月、3週、15日、14日、13日、12日、11日、10日、9日、8日、7日、6日、5日、4日、3日、2日、24時間未満で調製される。

【0147】

一旦、生成のためにより大きなオリゴ核酸が選択されると、オリゴ核酸の予め定められたライブラリは、デノボ合成のために設計される。高密度のオリゴ核酸アレイを生成するための様々な適切な方法が知られている。ワークフローの例において、基質表面層が提供される。この例において、表面の化学的性質(chemistry)は、オリゴ核酸合成プロセスを改善するために変更される。低い表面エネルギーの領域が液体を弾くために生成され、一方で高い表面エネルギーの領域が液体を引き付けるために生成される。表面自体は、平坦な表面の形であり、又は、表面積を増大させる突起又はマイクロウェルなどの形状の変形を含み得る。ワークフローの例において、全体において参照により本明細書に組み込まれる国際特許出願公開WO/2015/021080に開示されるように、選択された高い表面エネルギー分子は、DNAの化学的性質を支援する二元機能に役立つ。

【0148】

オリゴ核酸アレイのインサイツの調製は、固形支持体上で生成され、平行して複数のオリゴマーを伸長させるために単一のヌクレオチド伸長プロセスを利用する。オリゴ核酸シンセサイザーなどの堆積装置は、段階的な様式で試薬を放出するように設計され、その結果、複数のオリゴ核酸が平行して、一度に1つの残基を伸長させ、それにより予め定められた核酸配列を持つオリゴマーを生成する。場合により、オリゴ核酸はこの段階で表面から切断される。切断は、例えばアンモニア又はメチルアミンによる気相切断(gas cleavage)を含む。

【0149】

コンピューターシステム

【0150】

本明細書に記載されるシステムの何れかは、コンピューターに操作可能に連結され、且つコンピューターを介して局所的に又は遠隔で自動操作され得る。様々な実施形態において、本発明の方法及びシステムは更に、コンピューターシステム上のソフトウェアプログラム、及びその使用を含み得る。従って、材料堆積装置の動作、分配行為、及び減圧の作動を編成及び同期するなど、機能の分配/減圧/再充填の同期のためのコンピューター制御は、本発明の範囲内にある。コンピューターシステムは、ユーザーに指定された塩基配列と材料堆積装置の位置との間に干渉するようにプログラムされ、基質の指定された領域に正確な試薬を送達する。

【0151】

図8に例示されるコンピューターシステム(800)は、媒体(811)及び/又はネットワークポート(805)から命令を読み出すことが可能な、論理的な装置として理解され、固定された媒体(812)を持つサーバー(809)に随意に接続され得る。図8に示されるものなどのシステムは、CPU(801)、ディスクドライブ(803)、キーボード(815)及び/又はマウス(816)などの随意の入力装置、及び随意のモニター(807)を含み得る。データ通信は、局所又は遠隔の位置でサーバーに対して示された通信媒体を通じて達成され得る。通信媒体は、データを送信及び/又は受信する任意の手段を含み得る。例えば、通信媒体は、ネットワーク接続、無線接続、又はインターネット接続であり得る。そのような接続は、ワールド・ワイド・ウェブ上での通信を提供することができる。本開示に関するデータは、図8に例示されるように当事者(822)による受理及び/又は検討のためにそのようなネットワーク又は接続によって伝達され得ることが想定される。

【0152】

図9は、本発明の実施形態の例と関連して使用され得るコンピューターシステム(900)の第1の例のアーキテクチャを示すブロック図である。図9に表されるように、コンピューターシステムの例は、命令を処理するためのプロセッサ(902)を含み得る。プロセッサの限定されない例は、以下を含む: Intel Xeon(商標)プロセッサ、AMD Opteron(商標)プロセッサ、Samsung 32-bit RISC ARM 1176JZ(F)-S v1.0(商標)プロセッサ、ARM Cortex-A8 Samsung S5PC100(商標)プロセッサ、ARM Cortex-A8 Apple A4(商標)プロセッサ、Marvell PXA 930(商標)プロセッサ、又は機能的に同等なプロセッサ。複数のスレッドの実行が並列処理のために使用され得る。幾つかの実施形態において、複数のプロセッサ、又は複数のコアを持つプロセッサはまた、単一のコンピューターシステム中でも、クラスタの中でも、又は、複数のコンピューター、携帯電話、及び/又は個人用携帯情報端末装置を含むネットワーク上のシステムにわたって分布されても、使用され得る。

【0153】

図9に例示されるように、高速キャッシュ(904)は、プロセッサ(902)に接続するか、又はその中に組み込まれることで、プロセッサ(902)により近年使用されてきた又は頻繁に使用されている命令又はデータのための高速メモリを提供することができる。プロセッサ(902)は、プロセッサバス(908)によりノースブリッジ(906)に接続される。ノースブリッジ(906)は、メモリバス(912)によりランダムアクセスメモリ(RAM)(910)に接続され、プロセッサ(902)によりRAM(910)へのアクセスを管理する。ノースブリッジ(906)はまた、チップセットバス(916)によりサウスブリッジ(914)に接続される。サウスブリッジ(914)は次に、周辺バス(918)に接続される。周辺バスは、例えばPCI、PCI-X、PCI Express、又は他の周辺バスであり得る。ノースブリッジ及びサウスブリッジは頻繁に、プロセッサチップセットと称され、周辺バス(918)上でプロセッサと、RAMと、周辺コンポーネントとの間のデータ転送を管理する。幾つかの代替的なアーキテクチャにおいて、ノースブリッジの機能性は、別個のノースブリッジチップを使用する代わりにプロセッサに組み込まれ得る。幾つかの実施形態において、システム(900)は、周辺バス(918)に付けられたアクセラレータカード(922)を含むことができる。アクセラレータは、特定の処理を促進するためのフィールドプログラマブルゲートアレイ(FPGA)又は他のハードウェアを含み得る。例えば、アクセラレータは、適応データの再構築のために、又は、拡張設定処理に使用される代数式を評価するために使用され得る。

【0154】

ソフトウェアとデータは、外部記憶装置(924)に記憶され、プロセッサによる使用のためにRAM(910)及び/又はキャッシュ(904)へとロードされ得る。システム(900)は、システムリソースの管理のためのオペレーティングシステムを含み; オペレーティングシステムの限定されない例は、以下を含む: Linux(登録商標)、Windows(商標)、MAC OS(商標)、BlackBerry OS(商標)、iOS(商標)、及び他の機能的に同等なOS、同様に、本発明の実施形態の例に従ってデータの記憶と最適化を管理するためのオペレーティングシステム上で実行するアプリケーションソフトウェア。この例において、システム(900)はまた、ネットワーク接続ストレージ(NAS)などの外部記憶装置、及び分散並列処理に使用され得る他のコンピューターシステムに対するネットワークインターフェースを提供するために、周辺バスに接続されるネットワークインターフェースカード(NIC)(920)及び(921)を含む。

【0155】

図10は、複数のコンピューターシステム(1002a)及び(1002b)、複数の携帯電話及び個人用携帯情報端末(1002c)、並びにネットワーク接続ストレージ(NAS)(1004a)及び(1004b)を含むネットワーク(1000)を示す略図

10

20

30

40

50

である。実施形態の例において、システム(1002a)、(1002b)、及び(1002c)は、データ記憶を管理し、ネットワーク接続ストレージ(NAS)(1004a)及び(1004b)に記憶されたデータに対するデータアクセスを最適化することができる。数学モデルはこのデータに対して使用され、コンピューターシステム(1002a)及び(1002b)、並びに携帯電話及び個人用携帯情報端末システム(1002c)にわたって分散並列処理を使用して評価され得る。コンピューターシステム(1002a)及び(1002b)、並びに携帯電話及び個人用携帯情報端末システム(1002c)はまた、ネットワーク接続ストレージ(NAS)(1004a)及び(1004b)に記憶されたデータの適応データ再構築に対して並列処理を提供することができる。図10は一例のみを例示しており、様々な他のコンピューターのアーキテクチャ及びシステムは、本発明の様々な実施形態と共に使用され得る。例えば、ブレードサーバーは並列処理を提供するために使用され得る。プロセッサブレードは、並列処理を提供するためにバックプレーンを通じて接続され得る。ストレージはまた、別個のネットワークインターフェースを通してバックプレーンに、又はネットワーク接続ストレージ(NAS)として接続され得る。幾つかの実施形態の例において、プロセッサは、別個のメモリ空間を維持し、ネットワークインターフェース、バックプレーン、又は他のプロセッサによる並列処理のための他のコネクタを通じてデータを伝達することができる。他の実施形態において、プロセッサの幾つか又は全てが、共有仮想アドレスメモリ空間を使用することができる。

【0156】

図11は、実施形態の例に従って共有仮想アドレスメモリ空間を使用するマルチプロセッサコンピューターシステムのブロック図である。システムは、共有メモリサブシステム(1104)にアクセス可能な複数のプロセッサ(1102a-f)を含む。システムは、メモリサブシステム(1104)に複数のプログラマブルハードウェアのメモリアルゴリズムプロセッサ(MAP)(1106a-f)を組み込む。MAP(1106a-f)は各々、メモリ(1108a-f)及び1以上のフィールドプログラマブルゲートアレイ(FPGA)(1110a-f)を含み得る。MAPは設定可能な機能ユニットを提供し、特定のアルゴリズム又はその一部は、各プロセッサと密接に協働して処理を行うためにFPGA(1110a-f)に提供され得る。例えば、MAPは、データモデルに関する代数式を評価し、且つ実施形態の例における適応データの再構築を行なうために使用され得る。この例において、MAPは各々、このような目的のためのプロセッサ全てにより世界的にアクセス可能である。1つの構成において、MAPは各々、関連するメモリ(1108a-f)にアクセスするためにダイレクトメモリアccess(DMA)を使用することができ、それにより、各マイクロプロセッサ(1102a-f)とは別個に、且つこれらから非同期的にタスクを実行することが可能となる。この構成において、MAPは、パイプライン処理(pipelining)及びアルゴリズムの並列の実行のために別のMAPに直接結果を供給することができる。

【0157】

上述のコンピューターのアーキテクチャ及びシステムは単なる例であり、様々な他のコンピューター、携帯電話、個人用携帯情報端末のアーキテクチャ及びシステムは、一般的なプロセッサ、コプロセッサ、FPGA、及び他のプログラム可能論理回路の任意の組み合わせを使用するシステム、システムオンチップ(SOC)、特定用途向け集積回路(ASIC)、及び他の処理要素と論理素子を含む実施形態の例と共に使用され得る。幾つかの実施形態において、コンピューターシステムの全て又は一部は、ソフトウェア又はハードウェアに実装され得る。様々なデータ記憶媒体が、例えばランダムアクセスメモリ、ハードドライブ、フラッシュメモリ、テープドライブ、ディスクアレイ、ネットワーク接続ストレージ(NAS)、及び他のローカル又は分散データ記憶装置及びシステムを含む実施形態の例と共に使用され得る。

【0158】

実施形態の例において、コンピューターシステムは、上述の又は他のコンピューターのアーキテクチャ及びシステムの何れかにて実行するソフトウェアモジュールを使用して実

施され得る。他の実施形態において、システムの機能は、ファームウェア、図11で言及されるようなフィールドプログラマブルゲートアレイ（FPGA）などのプログラム可能論理回路、システムオンチップ（SOC）、特定用途向け集積回路（ASIC）、又は他の処理要素及び論理素子において部分的又は完全に実施され得る。例えば、セットプロセッサ及びオブティマイザは、図9に示されるアクセラレータカード（922）などのハードウェアアクセラレータカードの使用によるハードウェアアクセラレーションで実施され得る。

【0159】

以下の実施例は、当業者に対し本明細書に開示される実施形態の原理及び実践をより明白に示すために説明されるものであり、任意の請求された実施形態の範囲を制限するものとして解釈されるものではない。他に明示されない限り、全ての部分及びパーセンテージは重量基準である。

【実施例】

【0160】

以下の実施例は、本発明の様々な実施形態を例証する目的で与えられ、あらゆる様式で本発明を制限することを意図していない。本明細書に記載される方法とともに、本実施例は、好ましい実施形態の代表例であり、典型的なものであり、且つ、本発明の範囲を限定するものとして意図されない。請求項の範囲によって定義される本発明の趣旨内に含まれるその変化及び他の使用が、当業者に想定される。

【0161】

実施例1：基質表面の機能化

【0162】

基質を機能化し、オリゴ核酸のライブラリの付着及び合成を支援した。基質表面を先ず、20分間、90%の H_2SO_4 及び10%の H_2O_2 を含むピラニア溶液を使用して湿式洗浄した。基質を、DI水を含む幾つかのビーカーの中ですすぎ、5分間DI水のグーズネック形状の蛇口の下で保持して、 N_2 で乾燥させた。基質を5分間、 NH_4OH （1：100；3mL：300mL）に浸し、ハンドガン（handgun）を使用してDI水ですすぎ、DI水を含む3つの連続するビーカーの中で各々1分間浸し、その後ハンドガンを使用してDI水で再びすすいだ。その後、基質表面を O_2 にさらすことにより基質をプラズマ洗浄した。SAMCO PC-300機器を使用して、下流モードで1分間、250ワットで O_2 をプラズマエッチングした。

【0163】

以下のパラメータを有するYES-1224P気相蒸着オープンシステムを使用して、清潔になった基質表面をN-（3-トリエトキシシリルプロピル）-4-ヒドロキシブチルアミドを含む溶液で活動的に官能化した：0.5から1トル、60分、70、135の気化器。Brewer Science 200Xスピンコータを使用して、基質表面をレジストコートした（resist coated）。SPR（商標）3612フォトレジストを、40秒間2500rpmで基質上でスピンコートした。基質を予め、Brewerホットプレート上で、90で30分間焼いた。Karl Suss MA6マスクアライナー機器を使用して、基質をフォトリソグラフィーにさらした。基質を2.2秒間さらし、MSF 26Aの中で1分間展開させた（developed）。残りの展開物（developer）をハンドガンですすぎ、基質を5分間水に浸した。基質をオープンの中、100で30分間焼き、その後、Nikon L200を使用してリソグラフィーの欠損に対する目視検査を行った。洗浄処理を使用し、SAMCO PC-300機器を用いて残りのレジストを取り除き、1分間250ワットで O_2 プラズマエッチングした。

【0164】

基質表面を、10 μ Lの軽油と混合したパーフルオロオクチルトリクロロシランの100 μ Lの溶液で受動的に官能化した。基質をチャンバに入れ、10分間ポンプでくみ出し、その後、バルブを閉じてポンプを止め、10分間放置した。チャンバを通気した。最大

パワー（Crestシステム上で9）での超音波処理による70で500mLのNMPの中で5分間の2回の浸漬を行なうことにより、基質をレジスト剥離した。その後、最大パワーでの超音波処理により室温で500mLのイソプロパノールの中で5分間、基質を浸した。基質を、300mLの200プルーフエタノール（200 proof ethanol）に漬けて、N₂で送風乾燥した。官能化した表面は活性化され、オリゴ核酸合成のための支持体として役立った。

【0165】

実施例2：オリゴ核酸合成装置上での50merの配列の合成

【0166】

二次元オリゴ核酸合成装置をフローセルに組み入れ、フローセル（Applied Biosystems「ABI394 DNA Synthesizer」）に接続させた。N-（3-トリエトキシシリルプロピル）-4-ヒドロキシブチルアミド（Gelbst）で二次元オリゴ核酸合成装置を均一に官能化し、これを使用して、本明細書に記載されるオリゴ核酸合成方法を用いて50bpの例示的なオリゴ核酸（「50merのオリゴ核酸」）を合成した。

【0167】

50merの配列は、SEQ ID NO：1に記載されている通りであった。5' AGACAATCAACCATTTGGGGTGGACAGCCTTGACCTCTAGACCTTCGGCAT##TTTTTTTTTTT3'（SEQ ID NO：1）、ここで、#は、チミジン-スクシニルヘキサミドCEDホスホラミダイト（ChemGenesのCLP-2244）を表わし、これは、脱保護中に表面からのオリゴ核酸の放出を可能にする切断可能なリンカーである。

【0168】

表5のプロトコル及びABIシンセサイザーに従い標準DNA合成化学（結合、カップリング、酸化、及び非ブロック化）を使用して、合成を行った。

【0169】

【表 5 - 1】

表 5:

表 5		
一般的な DNA 合成プロセスの名称	プロセスの工程	時間 (秒)
洗浄 (アセトニトリル洗浄フロー)	アセトニトリル系フラッシュ	4
	アセトニトリルからフローセル	23
	N2 系フラッシュ	4
	アセトニトリル系フラッシュ	4
DNA 塩基付加 (ホスホロアミダイト + 活性化因子フロー)	活性化因子マニホールドフラッシュ	2
	活性化因子からフローセル	6
	活性化因子 + ホスホロアミダイトからフローセル	6
	活性化因子からフローセル	0.5
	活性化因子 + ホスホロアミダイトからフローセル	5
	活性化因子からフローセル	0.5
	活性化因子 + ホスホロアミダイトからフローセル	5
	活性化因子からフローセル	0.5
	活性化因子 + ホスホロアミダイトからフローセル	5
	25 秒間インキュベート	25
洗浄 (アセトニトリル洗浄フロー)	アセトニトリル系フラッシュ	4
	アセトニトリルからフローセル	15
	N2 系フラッシュ	4
	アセトニトリル系フラッシュ	4
DNA 塩基付加 (ホスホロアミダイト + 活性化因子フロー)	活性化因子マニホールドフラッシュ	2
	活性化因子からフローセル	5
	活性化因子 + ホスホロアミダイトからフローセル	18
	25 秒間インキュベート	25

【表 5 - 2】

一般的な DNA 合成プロセスの名称	表 5	
	プロセスの工程	時間 (秒)
洗浄 (アセトニトリル洗浄フロー)	アセトニトリル系フラッシュ	4
	アセトニトリルからフローセル	15
	N2 系フラッシュ	4
	アセトニトリル系フラッシュ	4
キャッピング (CapA+B, 1:1, フロー)	CapA+B からフローセル	15
洗浄 (アセトニトリル洗浄フロー)	アセトニトリル系フラッシュ	4
	アセトニトリルからフローセル	15
	アセトニトリル系フラッシュ	4
酸化 (オキシダイザーフロー)	オキシダイザーからフローセル	18
洗浄 (アセトニトリル洗浄フロー)	アセトニトリル系フラッシュ	4
	N2 系フラッシュ	4
	アセトニトリル系フラッシュ	4
	アセトニトリルからフローセル	15
	アセトニトリル系フラッシュ	4
	アセトニトリルからフローセル	15
	N2 系フラッシュ	4
	アセトニトリル系フラッシュ	4
	アセトニトリルからフローセル	23
	N2 系フラッシュ	4
	アセトニトリル系フラッシュ	4
デブロッキング (デブロックフロー)	デブロックからフローセル	36
洗浄 (アセトニトリル洗浄フロー)	アセトニトリル系フラッシュ	4
	N2 系フラッシュ	4

【表 5 - 3】

一般的な DNA 合成プロセスの名称	表 5	
	プロセスの工程	時間 (秒)
	アセトニトリル系フラッシュ	4
	アセトニトリルからフローセル	18
	N ₂ 系フラッシュ	4.13
	アセトニトリル系フラッシュ	4.13
	アセトニトリルからフローセル	15

【0172】

ホスホラミダイト / 活性化因子の組み合わせを、フローセルを通じたバルク試薬の送達と同様に送達した。環境が時間全体にわたり試薬により「湿った」ままであると、乾燥工程を行わなかった。

【0173】

フローリストリクターをABI 394 シンセサイザーから取り除き、より速い流れを可能した。フローリストリクターなしで、アミダイト (ACN 中で 0.1 M)、活性化因子 (ACN 中で 0.25 M のベンゾイルチオテトラゾール (「BT T」; Glen Research の 30-3070-xx))、及び Ox (20% のピリジン、10% の水、及び 70% の THF 中の 0.02 M の I₂) の流量は、(フローリストリクターで全ての試薬に対し ~ 50 uL/sec と比較して) およそ ~ 100 uL/sec、アセトニトリル (「ACN」) 及びキャッピング試薬 (Cap A と Cap B の 1:1 の混合物、ここで、Cap A は THF / ピリジン中の無水酢酸であり、Cap B は THF 中の 16% の 1-メチルイミダゾール (methylimidazole)) についてはおよそ ~ 200 uL/sec、及び、Deblock (トルエン中の 3% のジクロロ酢酸) についてはおよそ ~ 300 uL/sec であった。Oxidizer を完全に押し出す時間を観察し、化学フロー時間のタイミングを適宜調整し、余分な ACN 洗浄を異なる化学物質間に導入した。オリゴヌクレオチド合成の後、75 psi で一晩、ガス状のアンモニア中でチップを脱保護した。表面に水を 5 滴加えて、オリゴ核酸を再生した。その後、再生したオリゴ核酸を、BioAnalyzer の小さな RNA チップ上で分析した (データは示さない)。

【0174】

実施例 3: オリゴ核酸合成装置上での 100mer の配列の合成

【0175】

50mer の配列の合成について実施例 2 に記載されるのと同じプロセスを、2 つの異なるシリコンチップ上で 100mer のオリゴ核酸 (「100mer のオリゴ核酸」; 5' CGGGATCCTTATCGTCA TCGTCTCGTACAGATCCCGACCCATTTGCTGTCCACCA GTCA TGTAGCCATACCATGATGATGATGATGATGAGAACCCCGCAT ##TTTTTTTTTTT3、ここで、# はチミジン - スクシニルヘキサミド CED ホスホラミダイト (ChemGenes の CLP-2244) を表わす; SEQ ID NO: 2) の合成に使用し、一方のシリコンチップは N-(3-トリエトキシシリルプロピル)-4-ヒドロキシブチルアミドで均一に官能化され、他方のシリコンチップは 11-アセトキシウンデシルトリエトキシシランと n-デシルトリエトキシシランの 5/95 の混合物で官能化されており、表面から抽出されたオリゴ核酸を、BioAnalyzer 機器の上で分析した (データは示さない)。

【 0 1 7 6 】

以下の熱サイクルプログラムを使用して、50 μ LのPCR混合物(25 μ LのNEB Q5 master mix、2.5 μ Lの10 μ Mフォワードプライマー、2.5 μ Lの10 μ Mリバースプライマー、表面から抽出した1 μ Lのオリゴ核酸、及び最大50 μ Lの水)の中で、フォワードプライマー(5' ATGCGGGGTTCATCATC3'; SEQ ID NO:3)及びリバースプライマー(5' CGGGATCCTTATCGTCAATCG3'; SEQ ID NO:4)を使用して、2つのチップからの10のサンプル全てを更にPCR増幅した:

98 C、30 秒

98 C、10 秒; 63 、10 秒; 72 、10 秒; 12のサイクルを繰り返す

72 、2 分

10

【 0 1 7 7 】

PCR産物をBio Analyzer上でも実行して(データは示さず)、100merの位置で急なピークを実証した。次に、PCR増幅サンプルをクローン化し、Sanger配列決定を行った(Sanger sequence)。表6は、チップ1からのスポット1 - 5から得たサンプル、及びチップ2からのスポット6 - 10から得たサンプルに対するSanger配列決定から生じる結果を要約する。

【 0 1 7 8 】

【表6】

20

表 6:

スポット	エラー率	サイクル効率
1	1/763 bp	99.87%
2	1/824 bp	99.88%
3	1/780 bp	99.87%
4	1/429 bp	99.77%
5	1/1525 bp	99.93%
6	1/1615 bp	99.94%
7	1/531 bp	99.81%
8	1/1769 bp	99.94%
9	1/854 bp	99.88%
10	1/1451 bp	99.93%

30

【 0 1 7 9 】

故に、高品質及び高均一性の合成されたオリゴ核酸を、異なる界面化学的性質を持つ2つのチップ上で繰り返した。配列決定された100merの262のうち233に対応している、全体の89%が、エラーのない完全な配列であった。

40

【 0 1 8 0 】

最後に、表7は、スポット1 - 10からのオリゴ核酸サンプルから得た配列に対するエラー特徴を要約する。

【 0 1 8 1 】

【表 7】

表 7:

サンプル ID/スポット番号	OSA_0046/1	OSA_0047/2	OSA_0048/3	OSA_0049/4	OSA_0050/5	OSA_0051/6	OSA_0052/7	OSA_0053/8	OSA_0054/9	OSA_0055/10
合計の配列	32	32	32	32	32	32	32	32	32	32
配列決定の質	28 の内 25	27 の内 27	30 の内 26	23 の内 21	26 の内 25	30 の内 29	31 の内 27	31 の内 29	29 の内 28	28 の内 25
オリゴの質	25 の内 23	27 の内 25	26 の内 22	21 の内 18	25 の内 24	29 の内 25	27 の内 22	29 の内 28	28 の内 26	25 の内 20
ROI 一致の数	2500	2698	2561	2122	2499	2666	2625	2899	2798	2348
ROI 突然変異	2	2	1	3	1	0	2	1	2	1
ROI 複数の塩基の欠失	0	0	0	0	0	0	0	0	0	0
ROI 小さな挿入	1	0	0	0	0	0	0	0	0	0
ROI 単一の塩基の欠失	0	0	0	0	0	0	0	0	0	0
大きな欠失数	0	0	1	0	0	1	1	0	0	0
突然変異: G>A	2	2	1	2	1	0	2	1	2	1
突然変異: T>C	0	0	0	1	0	0	0	0	0	0
ROI エラー数	3	2	2	3	1	1	3	1	2	1
ROI エラー率	Err: 834 において~1	Err:1350 において~1	Err: 1282 において~1	Err: 708 において~1	Err: 2500 において~1	Err: 2667 において~1	Err: 876 において~1	Err: 2900 において~1	Err: 1400 において~1	Err: 2349 において~1
ROI を引くプライマーのエラー率	MP Err: 763 において~1	MP Err: 824 において~1	MP Err: 780 において~1	MP Err: 429 において~1	MP Err: 1525 において~1	MP Err: 1615 において~1	MP Err: 531 において~1	MP Err: 1769 において~1	MP Err: 854 において~1	MP Err: 1451 において~1

10

【 0 1 8 2 】

実施例 4: s g R N A 設計

20

【 0 1 8 3 】

5'末端の可変領域を有するキメラ s g R N A 配列を、C a s 9 タンパク質による直接の配列特異的な切断のために設計した。図 4 A を参照。s g R N A 配列は、12 の塩基の種子領域を含む特異的な D N A 結合に関する 20 の塩基の塩基対合領域を有していた。塩基対合領域の 5'末端は、転写開始部位となるように設計された。塩基対合領域に近接する 3'は、長さが 42 の塩基であった、C a s 9 結合のために d C a s 9 処理された領域であった。d C a s 9 処理された領域に近接する 3'は、長さが 40 の塩基であった、化膿レンサ球菌終結領域であった。d C a s 9 処理された領域及び終結領域はそれぞれ、U 字型の構造をもたらす配列を含むように設計された。

【 0 1 8 4 】

30

s g R N A も、鋳型 (T) 又は非鋳型 (N T) D N A 鎖を標的とするように設計された。図 5 A - 5 B を参照。鋳型 D N A 鎖を標的とするために設計された s g R N A は、転写された配列と同じ配列同一性を持つ s g R N A の塩基対合領域を含んでいた。非鋳型 D N A 鎖を標的とするために設計された s g R N A は、転写された配列の逆補体である s g R N A の塩基対合領域を含んでいた。

【 0 1 8 5 】

追加の構成において、T7 プロモーターは、可変塩基対合領域のすぐ上流に設計された。

図 6 A - 図 6 B を参照。T7 プロモーター領域を加えて、T7 ポリメラーゼを有する s g R N A のインビトロの産生を可能にした。

40

【 0 1 8 6 】

実施例 5: s g R N A をコードする D N A の合成 - 設計及びポリメラーゼ分析

【 0 1 8 7 】

D N A オリゴ核酸を、接合時に s g R N A 配列をコードする断片として設計した。図 12 を参照。s g R N A を、可変配列領域 (1233) のすぐ上流での T7 プロモーターの封入のため設計した。D N A オリゴ核酸のデノボ合成に続いて、増幅反応を行い、重なる断片を接合且つ伸長させた。

【 0 1 8 8 】

(1201) における D N A オリゴ核酸の転写は、D N A 鋳型からの T7 ポリメラーゼを有する s g R N A のインビトロの産生をもたらした。

50

【 0 1 8 9 】

設計 1 (1 2 2 0)、設計 2 (1 2 2 2)、設計 3 (1 2 2 4)、及び設計 4 (1 2 2 6) の配列を、表 8 に示されるように設計した。設計 1、設計 2、設計 3、及び設計 4 それぞれの配列は、T 7 プロモーター、可変配列部分、及び一定の配列領域 (ハンドル及びターミネーター) を含む (表 8)。具体的には、図 1 2 に見られるような一定の配列領域は、塩基対領域 (1 2 1 1)、(1 2 1 3)、(1 2 1 5)、(1 2 1 7)、(1 2 2 3)、及び (1 2 2 5) を含む Cas 9 ハンドルヘアピン、及び、塩基対領域 (1 2 1 9) 及び (1 2 2 1) を含むターミネーターヘアピンを備えている。

【 0 1 9 0 】

【表 8】

表 8.

SEQ ID NO	名称	配列
12	設計 1	TAATACGACTCACTATAGGGGATGCGCGCAGTT GTCCGTTTTAGAGCTAGAAATAGCAAGTTAAAAT AAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCA CCGAGTCGGTGCTTTT
13	設計 1 の T7 プロ モーター	TAATACGACTCACTATA
14	設計 1 の可変配列	GGATGCGCGCAGTTGTCC
15	設計 1 のハンドル とターミネーター	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG CTAGTCCGTTATCAACTTGAAAAAGTGGCACCGA GTCGGTGCTTTT
44	設計 2	GAAATTAATACGACTCACTATAGGGGATGCGCG CAGTTGTCCGTTTTAGAGCTAGAAATAGCAAGTT AAAATAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGCTTTT
13	設計 2 の T7 プロ モーター	TAATACGACTCACTATA
14	設計 2 の可変配列	GGATGCGCGCAGTTGTCC
15	設計 2 のハンドル とターミネーター	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG CTAGTCCGTTATCAACTTGAAAAAGTGGCACCGA GTCGGTGCTTTT
16	設計 3	GAGCTAATACGACTCACTATAGGGGATGCGCGC AGTTGTCCGTTTTAGAGCTAGAAATAGCAAGTTA AAATAAGGCTAGTCCGTTATCAACTTGAAAAAGT GGCACCGAGTCGGTGCTTTT
13	設計 3 の T7 プロ モーター	TAATACGACTCACTATA
14	設計 3 の可変配列	GGATGCGCGCAGTTGTCC
15	設計 3 のハンドル とターミネーター	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG CTAGTCCGTTATCAACTTGAAAAAGTGGCACCGA GTCGGTGCTTTT
17	設計 4	CGAGCTAATACGACTCACTATAGGGGATGCGCG CAGTTGTCCGTTTTAGAGCTAGAAATAGCAAGTT AAAATAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGCTTTT
13	設計 4 の T7 プロ モーター	TAATACGACTCACTATA
14	設計 4 の可変配列	GGATGCGCGCAGTTGTCC
15	設計 4 のハンドル とターミネーター	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG CTAGTCCGTTATCAACTTGAAAAAGTGGCACCGA GTCGGTGCTTTT

【 0 1 9 1 】

インピトロの分析に関しては、T7 RNAポリメラーゼ・プロモーター領域がT7 RNAポリメラーゼによる認識のための二本鎖であるべきであることに注目されたい。アンチセンスオリゴ核酸をハイブリダイゼーションのために使用した：5' - T A A T A C G A C T C A C T A T A G G - 3' (SEQ ID NO : 18)。加えて、表9は、鋳型オリゴ核酸及び増幅オリゴ核酸の4つの異なるセットの分析のために使用されたプライ

10

20

30

40

50

マーのリストを提供する。図 12 を参照する。

【 0 1 9 2 】

【 表 9 】

表 9.

SEQ ID NO	名称 (# 塩基 - 融解温度)	配列
19	SgR1-R1 (80bp)	AAAAGCACCGACTCGGTGCCACTTTTTCAAGT TGATAACGGACTAGCCTTATTTTAACTTGCTAT TTCTAGCTCTAAAAC
20	SgR1-F1 (58bp - 51°C/54°C)	TAATACGACTCACTATAGGGGATGCGCGCAGT TGTCCGTTTTAGAGCTAGAAATAGCA
21	SgR1-F2 (65bp - 56°C/56°C)	GAAATTAATACGACTCACTATAGGGGATGCGC GCAGTTGTCCGTTTTAGAGCTAGAAATAGCAA G
22	SgR1-F3 (66bp - 59°C/58°C)	GAGCTAATACGACTCACTATAGGGGATGCGCG CAGTTGTCCGTTTTAGAGCTAGAAATAGCAAG TT
23	SgR1-F4 (78bp - 62°C/62°C)	GAGCTAATACGACTCACTATAGGGGATGCGCG CAGTTGTCCGTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGG
24	SgR1-AR1 (14bp - 52°C)	AAAAGCACCGACTC
25	SgR1-AR2 (15bp - 56°C)	AAAAGCACCGACTCG
26	SgR1-AR3 (16bp - 60°C)	AAAAGCACCGACTCGG
27	SgR1-AR4 (17bp - 62°C)	AAAAGCACCGACTCGGT
18	sgR1-AF1 (19bp - 51°C)	TAATACGACTCACTATAGG
28	SgR1-AF2 (24bp - 56°C)	GAAATTAATACGACTCACTATAGG
29	SgR1-AF3 (25bp - 59°C)	GAGCTAATACGACTCACTATAGG
30	SgR1-AF4 (25bp - 62°C)	GCGAGCTAATACGACTCACTATAGG

10

20

30

【 0 1 9 3 】

鋳型オリゴ核酸及び増幅オリゴ核酸の4つの異なるセットは、完全長の鋳型の純度及び収量を最適化する様々な条件の下で分析された。10 µl のPCR反応を、100 fMolでの鋳型オリゴ核酸 (SgR1-R1 & SgR1-F1、SgR1-F2、SgR1-F3、SgR1-F4)、及び、600 nMolのポリメラーゼ-1 PCR濃度での増幅プライマーの各セットにより実行した。Eppendorf Mastercycler 上で勾配を使用して、3つのアニーリング温度 (50、55、60) を、2つの高忠実度DNAポリメラーゼ (ポリメラーゼ1及び3) 並びに標準的なDNAポリメラーゼ (ポリメラーゼ2) を使用して、25のサイクルのPCRにおいて評価した。表10は反応条件の概要を提供し、表11は増幅プロトコルを提供する。

40

【 0 1 9 4 】

【表 10】

表 10:

試薬 (10ul rxns)		Pol. 1			Pol. 2			Pol. 3	
		1x	15x		1x	15x		1x	15x
Pol		0.1	1.5		0.1	1.5		0.1	1.5
緩衝液		2	30		1	15		2	30
dNTP's		0.2	3		0.2	3		0.2	3
増幅プライマー(10 uM)		0.6			0.6			0.6	
鋳型オリゴ(100 nM)		1			1			1	
H2O		6.1	91.5		7.1	106.5		6.1	91.5
合計		10	150		10	150		10	150

10

【 0 1 9 5 】

【表 11】

表 11.

ポリメラーゼ 1, 2			ポリメラーゼ 3		
98	30 秒		95	3 分	
98	10 秒	25x	98	10 秒	25x
50 / 55 / 60	15 秒		50 / 55 / 60	15 秒	
72	10 秒		72	10 秒	
72	30 秒		72	30 秒	
4	保持		4	保持	

20

30

【 0 1 9 6 】

ポリメラーゼ - 1 の PCR 反応からの結果を、収量を評価するために Bio Analyzer (図示せず) 上で実行し、表 12 において要約した。DNA 収量を ng / ul で表す (表 12)。オリゴ核酸の設計 3 と 4 はそれぞれ、オリゴ核酸の設計 1 と 2 よりも高い DNA 収量をもたらした。より高いアニーリング温度は同様に収量の増加をもたらし、60 ではより高い収量であった。

【 0 1 9 7 】

【表 12】

表 12.

ポリメラーゼ 1	50 °C	55 °C	60 °C
オリゴ設計 1	0	0	0
オリゴ設計 2	2.9	4.5	3.7
オリゴ設計 3	6.8	9.2	10
オリゴ設計 4	9.9	13	15.3

10

収量を ng/ul で列挙する。

【0198】

ポリメラーゼ - 2 の P C R 反応からの結果を、収量を評価するために B i o A n a l y z e r (図示せず) 上で実行し、表 13 において要約した。DNA 収量を ng / ul で表す (表 13)。再度、オリゴ核酸の設計 3 と 4 はそれぞれ、オリゴ核酸の設計 1 と 2 よりも高い DNA 収量をもたらした。より高いアニーリング温度は同様に収量の増加をもたらし、60 °C ではより高い収量であった。

【0199】

20

【表 13】

表 13.

ポリメラーゼ 2	50 °C	55 °C	60 °C
オリゴ設計 1	0	0	0
オリゴ設計 2	7.6	5.9	6.9
オリゴ設計 3	6.1	8.5	10.5
オリゴ設計 4	7.4	11.1	19.4

30

収量を ng/ul で列挙する。

【0200】

ポリメラーゼ - 3 の P C R 反応からの結果を、収量を評価するために B i o A n a l y z e r (図示せず) 上で実行し、表 14 において要約した。DNA 収量を ng / ul で表す (表 14)。オリゴ核酸の設計 3 と 4 はそれぞれ、オリゴ核酸の設計 1 と 2 よりも高い DNA 収量をもたらした。より高いアニーリング温度は同様に収量の増加をもたらし、60 °C ではより高い収量であった。

40

【0201】

【表 14】

表 14.

ポリメラーゼ 3	50 °C	55 °C	60 °C
オリゴ設計 1	10	13	12.1
オリゴ設計 2	12.4	14.3	15.9
オリゴ設計 3	13.2	26.1	28.8
オリゴ設計 4	16.1	13.2	18.5

10

収量を ng/ul で列挙する。

【0202】

要するに、オリゴ核酸の設計 3 と 4 は、3 つのポリメラーゼ全てを含む DNA 収量の増加をもたらした。加えて、60 °C のより高いアニーリング温度は DNA 収量の増加をもたらした。

【0203】

実施例 6 : CRISPR sgRNA の合成 - 温度分析

【0204】

20

実施例 5 からのオリゴ核酸プライマーを使用して、アニーリング温度条件の増加の影響を、実施例 5 に記載されるような PCR 反応の実行後に分析した。増幅産物を、収量を評価するために BioAnalyzer (図示せず) で実行し、表 15 において要約した。DNA 収量を ng/ul で表す (表 15)。要するに、ポリメラーゼ 3 は DNA 収量の増加をもたらし、60 °C のアニーリング温度は DNA 収量の増加をもたらした。

【0205】

【表 15】

表 15.

	オリゴ設計 3				オリゴ設計 4			
	ポリメラーゼ 3		ポリメラーゼ 1		ポリメラーゼ 3		ポリメラーゼ 1	
	60 °C	65 °C	60 °C	65 °C	60 °C	65 °C	60 °C	65 °C
25 サイクル	27.3	15.2	11.6	3.9	28.4	29.5	13.8	10.2

30

収量を ng/ul で列挙する。

40

【0206】

実施例 7 : sgRNA の生成 - 構造の無い RNA

【0207】

2 つのオリゴ核酸の集合体を設計し、T7 プロモーター配列とターミネーターを用いるが、tracrRNA を用いることなく、修飾された sgRNA 鋳型 (120 bp) を生成した。表 16 を参照する。

【0208】

【表 16】

表 16.

SEQ ID NO	名称	配列決定
31	sgR2-鋳型	CGAGCTAATACGACTCACTATAGGGGCACAACGT GGAGGATGGCAGCGTGCAGCTGGCTGATCACTAC CAGCAAAACACTCCAATCGGTGATGGTCCTGTTGC ACCGAGTCGGTGCTTTT
32	sgR2-F	CGAGCTAATACGACTCACTATAGGGGCACAACGT GGAGGATGGCAGCGTGCAGCTGGCTGATCACTAC CAG
33	sgR2-R	AAAGCACCGACTCGGTGCAACAGGACCATCACCG ATTGGAGTGTTTTGCTGGTAGTGATCAGCCAGCTG

10

【0209】

オリゴ核酸の集合体を、実施例 5 の sgRNA の増幅に使用される同じオリゴ核酸を用いて増幅した。使用される反応条件を表 17 に要約する。

【0210】

【表 17】

表 17.

20

成分	1x		PCR 条件	時間	サイクル
Pol 1	0.1		95°C	3 分	
緩衝液	2		98 °C	10 秒	25 x
dNTP's	0.2		65 °C	15 秒	
増幅プライマー (10um)	0.6		72 °C	10 秒	
鋳型オリゴ (100nm)	1		72 °C	30 秒	
H2O	6.1		4 °C	保持	
合計	10				

30

40

【0211】

T7 RNAポリメラーゼでの転写は、二次構造を欠いた 80bp の RNA 産物をもたらすと推測された。増幅産物の転写をインビトロの転写キット (NEB HiScribe) により行った。反応混合物を BioAnalyzer 上で分析した。図 13A - 13B を参照する。修飾された sgRNA 産物は、tracrRNA ヘアピン含有配列を持つ sgRNA (図 13A) よりも、構造の無い設計により清潔であった (図 13B)。

【0212】

実施例 8 : sgRNA 配向 Cas9 切断

【0213】

50

3つのsgRNA配列をT7プロモーター領域で設計し、各々には720bp GFPコード配列の領域のための異なる認識配列が備わっていた。sgRNA配列の各々を2つのオリゴ核酸のPCRから組み立てた。sgRNAバックボーン及びプライマーを表18に提供する。

【0214】

【表18】

表 18.

SEQ ID NO	名称	配列
34	sgRNAバックボーン	CGAGCTAATACGACTCACTATAGgNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTATGCTGAAAAGCATA GCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTG AAAAAGTGGCACCGAGTCGGTGCTTTT
35	GFP	ATGcgtAAAggcGAAgagCTGttcACTggtGTCgtcCCTatt CTGgtgGAActgGATggtGATgtcAACggtCATaagTTTtcc GTGcgtGGCgagGGTgaaGGTgacGCAactAATggtAAAct gACGctgAAGttcATCtgtACTactGGTaaaCTGccgGTAcc tTGccgACTctgGTAacgACGctgACTtatGGTggtCAGtg cTTTgctCGTtatCCGgacCATatgAAGcagCATgacTTcTt cAAGtccGCCatgCCGgaaGGCtatGTGcagGAacgcACGa ttTCctttAAGgatGACggcACgtacAAAacgCGTgagGAAG tgAAAtttGAaggcGATaccCTGgtaAACcgcatTTgagCTGa aaGGCattGACtttAAAgaaGACggcAATatcCTGggcCAT aagCTGgaaTACaatTTTaaAGCcacAATggtTACatcAC CgccGATAaaaCAaaaaAATggcATTaaaGCGaatTTTaaaA TTcgcACAacGTGgagGATggcAGCgtgCAGctgGCTgat CActacCAGcaaAACactCCAatcGGTgatGGTcctGTTctg CTGccaGACaatCActatCTGagcACGcaaAGCgttCTGtct AAAgatCCGaacGAGaaaCGCgatCATatgGTTctgCTGg agTTCgtaACCGcaGCGggcATCacgCATggtATGgatGAA ctgTACaaaTGAtaa
36	sgR35-F	CGAGCTAATACGACTCACTATAGGAacgcACGattTC CtttAGTTTTAGAGCTATGCTGAAAAGCATAGC
37	sgR36-F	CGAGCTAATACGACTCACTATAGGCattGACtttAAA gaaGAGTTTTAGAGCTATGCTGAAAAGCATAGC
38	sgR37-F	CGAGCTAATACGACTCACTATAGGagGATggcAGCgt gCAGcGTTTTAGAGCTATGCTGAAAAGCATAGC
39	sgR3-R	AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGA TAACGGACTAGCCTTATTTTAACCTTGCTATGCTTT TCAGCATAGCTCTAAAC
30	SgR1-AF4	GCGAGCTAATACGACTCACTATAGG
27	SgR1-AR4	AAAAGCACCGACTCGGT

【0215】

オリゴ核酸集合体を、表19に要約された反応条件下で増幅した。

【0216】

10

20

30

40

【表 19】

表 19.

成分	1x		PCR 条件	時間	サイク ル
ポリメラーゼ 3	0.1		95°C	3 分	
緩衝液	2		98 °C	10 秒	25 x
dNTP's	0.2		95 °C	15 秒	
増幅プライマー (10 um)	0.6		60 °C	15 秒	
鋳型オリゴ (100 nm)	1		72 °C	10 秒	
H2O	6.1		72 °C	30 秒	
合計	10		4 °C	保持	

10

20

【0217】

各 sgRNA 集合体反応からのサンプルを Bio Analyzer 上で分析した (図 14A - 14C)。T7 RNA ポリメラーゼ PCR 増幅産物を使用する転写反応を行った。各反応からのサンプルを Bio Analyzer 上で分析した (図 14D - 14F)。

【0218】

Cas9 消化物を、GFP 増幅産物、Cas9、及び転写された sgRNA を使用して調製した。対照の単一のピークと比較して、2つのピークを3つ全ての消化物に観察した。(図 14G - 14J)。3つの合成された sgRNA を使用する Cas9 切断から予想され且つ結果として生じた断片を、表 20 に列挙する。

30

【0219】

【表 20】

表 20.

sgRNA	予測された 断片 1	結果として 生じる断片 1	予測された 断片 2	結果として 生じる断片 2
sgR35	321	324	439	451
sgR36	342	350	418	430
sgR37	208	137	552	560

40

【0220】

Cas9 消化サンプルを精製し、Bio Analyzer 上で再び分析した (図示せず)。

精製されたサンプルからの結果を表 21 に要約する。

【0221】

【表 2 1】

表 21.

sgRNA	予測された 断片 1	結果として 生じる断片 1	予測された 断片 2	結果として 生じる断片 2
sgR35	321	323	439	451
sgR36	342	353	418	427
sgR37	208	220	552	560

10

【 0 2 2 2 】

実施例 9：29，040 の固有のオリゴ核酸の平行的な集合体

【 0 2 2 3 】

平坦なシリコン上に各々 12 の遺伝子座を含む、256 のクラスタ（1505）を含む構造を、図 15 に示すように製造した。クラスタの拡大図を 121 の遺伝子座と共に（1510）で示す。256 のクラスタのうち 240 からの遺伝子座は、別個の配列を持つオリゴ核酸の合成の付着と支持をもたらした。オリゴ核酸合成を、実施例 3 の一般法を使用するホスホロアミダイト化学物質によって実行した。256 のクラスタのうち 16 からの遺伝子座は対照クラスタであった。合成された 29，040 の固有のオリゴ核酸の全体的な分布（1つのクラスタ当たり 240 の非対照クラスタ × 121 のオリゴ核酸集団）を図 16A に示す。NGS 配列決定は、合成のために選択される設計されたオリゴ核酸の 100% の表示を確認した。図 16B に示されるように、分布を各クラスタについて測定した。4つの代表的なクラスタにおいて合成された固有のオリゴ核酸の分布を、図 17 に示す。全体的なレベルでは、合成のために設計された全てのオリゴ核酸が存在し、オリゴ核酸の 99% が平均の 2 倍以内にある存在量を備えており、このことは合成の高い均一性を示している。この同じ観察はクラスタ当たりのレベル上で一貫していた。

20

【 0 2 2 4 】

各オリゴ核酸のエラー率を、Illumina MiSeq 遺伝子シーケンサーを使用して判定した。29，040 の固有のオリゴ核酸に関するエラー率分布を図 18A に示し、平均は 500 の塩基において約 1 であり、幾つかエラー率は 800 の塩基におけるものほど低いものであった。図 18B に示されるように、分布を各クラスタについて測定した。4つの代表的なクラスタにおける固有のオリゴ核酸に関するエラー率分布を図 19 に示す。29，040 の固有のオリゴ核酸のライブラリを 20 時間未満で合成した。29，040 の固有のオリゴ核酸の全てにわたるオリゴ核酸表示に対する GC 割合の分析は、合成が GC 含有量（1つのオリゴ核酸当たりおよそ 20% ~ 85%）にもかかわらず一定であったことを示した（図 20）。

30

【 0 2 2 5 】

実施例 10：sgRNA をコードするデノボ合成された DNA ライブラリの PCR 増幅分析

40

【 0 2 2 6 】

20 - 80% の様々な GC 含有量を含む無作為化された配列の長さ 100 の塩基を有する 9，996 のオリゴ核酸を設計し、実施例 9 に記載される同様の構成を持つ構造上で合成した。GC の表示に対する PCR 増幅の効果を判定するために、オリゴ核酸集団を、高忠実度 DNA ポリメラーゼ（DNA ポリメラーゼ 1）を伴う 6 又は 20 のサイクルのために増幅した。代替的に、オリゴ核酸集団を、6、8、10、又は 15 のサイクルのために他の 2 つの高忠実度 PCR 酵素を使用して増幅し、ポリメラーゼ選択が増幅後の全体的な配列の表示に対し効果があったかどうかを判定した。PCR 増幅の後、サンプルを次世代配列決定のために調製し、Illumina MiSeq プラットフォーム上で配列決定

50

した。150の塩基対のSEリードを、100Xのおよそのリード適用範囲に生成した。生のFASTQファイルを分析した。6、10、又は15のサイクルのために何れかのポリメラーゼによるオリゴ核酸の表示を、図21に図示する。配列決定リードにおける表現の頻度によって測定されたオリゴ核酸の均一性を、様々な条件について評価し、表22に要約した。

【0227】

【表22】

表 22.

	サイクル	1.5x 内の%	2x 内の%
ポリメラーゼ 1	6	72.1%	92.6%
	8	76.1%	90.3%
	10	70.9%	86.6%
	15	64.1%	82.7%
ポリメラーゼ 2	6	91.9%	98.9%
	8	89.9%	98.1%
	10	90.1%	98.4%
	15	89.2%	97.9%

【0228】

増幅されたオリゴ核酸集団それぞれのドロップアウトの数を、図22、即ち増幅サイクル vs 平均閾値の10%より下の集団の分画に示されるように定量化した。ポリメラーゼ1のドロップアウトは早く成長した一方、ポリメラーゼ2のドロップアウトは比較的一定に留まった。

【0229】

GC分布に対する増幅の影響を、図23において評価した。一般的に、30%から70%のGC含有量を有するオリゴ核酸は、傾向線である $Y = X$ に従い、より多くのサイクルにより頻度を増大させた。70%より大きいGC含有量を有するオリゴ核酸は一般的に、20のサイクルの後にわずかに多く頻繁であったが、30%未満のGC含有量を有するオリゴ核酸は一般的に、6つのサイクルの後にわずかに多く頻繁であった。

【0230】

実施例 11：ヒトの後成的なCRISPRスクリーン

【0231】

sgRNAスクリーンを行い、CRISPR-Cas9を使用して機能ドメインをコードするエクソンへの突然変異を導入した。約10,000のDNAオリゴ核酸を、実施例3に記載されるものと同様の方法を使用して、実施例1に記載されるシリコンチップ上で、実施例9に記載される同様の構成を持つ構造で、デノボ合成した。総じて、オリゴ核酸には約1:500以下の総合エラー率があった。各オリゴ核酸は長さ最大200の塩基であり、オリゴ核酸種1つ当たり少なくとも1fmoleが生成された。オリゴ核酸をPCR増幅し、ベクターへとクローン化し、sgRNA転写のために培養細胞へと電気穿孔した。核酸を細胞から単離し、次世代配列決定を使用して配列決定した。

【0232】

配列決定の結果は、sgRNAの最小のバイアス及び高忠実度産生と共に、非常に正確で均一なライブラリ合成を示した。下流のスクリーニングのために正確な配列を持つsgRNAの30%以上の最小の配列決定の回復を伴うガイド配列1つ当たりのより多くのリ

ードが、競合プール (competitor pool) と比較される。表 23 を参照する。プールされた配列決定の結果は、競合プールに基づいたアレイを持つ 6 つの対数と比較して、ガイド配列当たりより多くのリード、及びリードのはるかにタイトな分布 (4 つの対数) を示した。図 24 A - 24 B を参照する。クロンの配列決定の検証は、100 % の sgRNA 回復 (図 24 A)、及び市販のアレイベースのプールに比べて高い配列の正確さ (図 24 B) を示した。配列決定されたクロンのうち、著しく多くが正確な sgRNA 配列により回復された。表 23 を参照する。予め決められた配列の 100 % がオリゴ核酸集団に表わされた。sgRNA クロンの NGS ベースの検証は、100 % の sgRNA 回復、及び競合集団 (図示せず) と比較して 13 % 高い 1 つのクローン当たりの正確さを示した。

10

【 0 2 3 3 】

【表 23】

表 23.

	合成されたオリゴ集団	市販のオリゴ集団
回復された sgRNA オリゴ	100%	>95.5%
正確な配列の割合 (MiSeq)	約 87%	約 74%
正確な配列 (サンガー10 のクローン)	約 100%	約 70%
クローン化したオリゴ集団における 1 つの sgRNA 当たりの平均リード (100x の標準化)	約 256	約 1024

20

【 0 2 3 4 】

実施例 12 : 全体的なゲノム sgRNA ライブラリ

【 0 2 3 5 】

DNA ライブラリを、101,000 の異なるオリゴ核酸のためにクローンを生成するために、sgRNA (20200 の遺伝子標的当たり 5 つの sgRNAs) をコードする DNA を含むように設計した。101,000 のオリゴ核酸を、実施例 3 に記載されるものと同様の方法を使用して、実施例 1 に記載されるシリコンチップ上で、実施例 9 に記載される同様の構成を持つ構造で、デノボ合成した。合成されたオリゴ核酸を PCR 増幅し、消化し、レンチウイルスベクターへとクローン化し、細胞へと形質転換した。核酸を細胞から単離し、次世代配列決定を使用して配列決定した。代替的に、合成されたオリゴ核酸を PCR 増幅して、アンプリコンベースのライブラリを形成し、配列決定した。

30

【 0 2 3 6 】

回復された sgRNA の数に対する次世代配列決定のプロットは、オリゴ核酸プールのサイズが増加すると、オリゴ核酸集団がライブラリ全体にわたりリードのより均一でタイトな分布を維持し、最小の尾部が市販のアレイベースの参照オリゴ核酸集団と比較されることを示した。図 25 を参照。

40

【 0 2 3 7 】

実施例 13 : 標的化と活性が改善した sgRNA ライブラリの設計

【 0 2 3 8 】

sgRNA ライブラリを設計し、実施例 3 に記載されるものと同様の方法を使用して、実施例 1 に記載されるシリコンチップ上で、実施例 9 に記載される同様の構成を持つ構造で、デノボ合成した。

合成されたオリゴ核酸を PCR 増幅し、消化し、ベクターへとクローン化し、スクリーニング及び分析を含む下流の適用のための使用のために細胞へと転移させた。

【 0 2 3 9 】

NAG PAM、NGG PAM、高活性、低いオフターゲットで特徴されるライブラ

50

りを含む、異なる s g R N A 設計のパラメータを比較し、フィルタ処理した。本明細書に記載される方法により設計される s g R N A は、他の市販の g R N A システムに比べて高い割合の s g R N A をもたらし、結果として、遺伝子発現の少なくとも 2 倍の枯渇、約 16 % の s g R N A をもたらし。図 26 A を参照。s g R N A はまた、他の市販の g R N A システムに比べて低い割合の s g R N A をもたらし、結果として、遺伝子発現のゼロ又は負の枯渇、約 17 % をもたらし。図 26 B を参照。

【0240】

s g R N A 媒介性の枯渇を、同様に実質的な遺伝子発現レベルについて評価し、そこでは以下の遺伝子を s g R N A により標的化した：P C N A、P S M A 7、R P P 2 1、及び S F 3 B 3。少なくとも 2 倍の枯渇を示した s g R N A の数を分析すると、s g R N A ライブラリには、競合 1、競合 2、及び競合 3 と比較して高い割合の、実質的な遺伝子を枯渇する s g R N A があった。表 24 を参照。

【0241】

【表 24】

表 24.

遺伝子名	競合 1	競合 2	競合 3	CRISPR ライブラリ
PCNA	1/5	2/6	5/9	5/5
PSMA7	1/5	0/6	2/9	3/5
RPP21	3/5	1/6	4/9	2/5
SF3B3	0/5	0/6	4/9	3/5
平均 (%)	25%	12.5%	42%	65%

【0242】

実施例 14：M S 2 に関する s g R N A ライブラリ

【0243】

s g R N A をコードする非同一 D N A 配列を含む D N A ライブラリを、C 2 c 2 タンパク質による配列特異的な切断のために設計した。ライブラリは、バクテリオファージ M S 2 ゲノムの C 2 c 2 標的化のための全ての可能なスペーサー配列を含んでいた。レプトトリキア・シャヒイからの C 2 c 2 の成熟 c r R N A が 28 のヌクレオチドの最大のスペーサー長を含むため、バクテリオファージゲノムに全ての可能な 28 のヌクレオチド標的部位のタイリング (t i l l i n g) は結果として約 3500 のスペーサー配列のライブラリをもたらした。

【0244】

約 3,500 の非同一オリゴ核酸を、実施例 3 に記載されるものと同様の方法を使用して、実施例 1 に記載されるシリコンチップ上で、実施例 9 に記載される同様の構成を持つ構造で、デノボ合成した。約 3500 の配列のライブラリをベクターへと挿入し、大腸菌に形質転換させた。大腸菌細胞を、M S 2 の 3 つの希釈を使用して M S 2 バクテリオファージで感染させた。その後、ライブラリを、M S 2 感染に対する大腸菌耐性を与えた配列についてスクリーニングした。

【0245】

多数のスペーサー配列は耐性を与えることが見出された。スペーサー表現 (c r R N A 頻度) を比較すると、多くのスペーサー配列が M S 2 感染の 3 つの希釈において 1.25 を超える対数 2 倍の富化を示したが、非標的スペーサー配列は富化されることが分かった。

【0246】

実施例 15：ゼブラフィッシュに関する s g R N A ライブラリ

【0247】

D N A ライブラリを、約 130,000 の s g R N A をコードする配列により設計する。平均して、約 5 の s g R N A 鋳型を各ゼブラフィッシュ遺伝子のために設計する。オリ

ゴ核酸を、実施例 3 に記載されるものと同様の方法を使用して、実施例 1 に記載されるシリコンチップ上で、実施例 9 に記載される同様の構成を持つ構造で、デノボ合成した。デノボ合成は、130,000 のオリゴ核酸を賛成し、各々、シリコンプレートの表面上で異なる遺伝子座から伸長している。オリゴ核酸をプレートから取り除き、PCR によって増幅し、発現ベクターへとクローン化した。各鋳型を配列決定にさらす。sgRNA ライブラリをゼブラフィッシュの胚へと注入する。ゼブラフィッシュを成体まで育てる。その後、精子を凍結保存し、生殖系列伝達された挿入と欠失を同定するために配列決定することによりスクリーニングする。生殖系列のスクリーンの後、競合の対立遺伝子に特異的な PCR によって精子の遺伝子型を同定する。

【0248】

10

実施例 16：マウスのための gRNA ライブラリ

【0249】

DNA ライブラリを、約 100,000 の sgRNA をコードする配列により設計する。平均して、約 5 の sgRNA 鋳型を 1 つのマウス遺伝子ごとに設計する。オリゴ核酸を、実施例 3 に記載されるものと同様の方法を使用して、実施例 1 に記載されるシリコンチップ上で、実施例 9 に記載される同様の構成を持つ構造で、デノボ合成した。sgRNA 配列をコードする sgRNA ライブラリをデノボ合成して、100,000 のオリゴ核酸を生成する。デノボ合成は、100,000 のオリゴ核酸を産生し、各々、シリコンプレートの表面上で異なる遺伝子座から伸長している。オリゴ核酸をプレートから取り除き、PCR によって増幅し、ベクターへとクローン化した。各鋳型を配列決定にさらす。sgRNA のオンターゲット効率を、監督者によるヌクレアーゼアッセイ又は配列決定によって検証する。その後、望ましい遺伝的バックグラウンドを持つマウスの接合子において sgRNA を微量注入する。代替的に、sgRNA 効率の検証の後、sgRNA をアデノ関連性ウイルス (AAV) などのウイルスベクターへとパッケージングする。その後、sgRNA を、望ましい位置にてマウスへと定位的に送達する。予め選択された標的遺伝子の発現レベルを、マウスから採取した組織において観察する。

20

【0250】

実施例 17：受容体チロシンキナーゼのための gRNA ライブラリ

【0251】

DNA オリゴ核酸ライブラリを、表 25 に列挙される 58 のヒト受容体チロシンキナーゼのための遺伝子を標的とする 5 つの sgRNA をコードする配列で設計し、異なる DNA のオリゴ核酸は合計 290 であった。オリゴ核酸を、実施例 3 に記載されるものと同様の方法を使用して、実施例 1 に記載されるシリコンチップ上で、実施例 9 に記載される同様の構成を持つ構造で、デノボ合成した。オリゴ核酸をプレートから取り除き、PCR により増幅し、ベクターへとクローン化し、細胞の予め選択された集団へと移入した。表 25 に列挙された予め選択された遺伝子の発現レベルは、キナーゼ特異的 sgRNA の無い対照にさらされる細胞の対照集団に対する細胞の予め選択された集団において比較される。

30

【0252】

【表 25 - 1】

表 25.

遺伝子	Hs NT ACC #	Hs PROT ACC#
ALK	NM_004304	NP_004295
LTK	NM_002344	NP_002335
AXL	NM_001699	NP_001690
MER	NM_006343	NP_006334
TYRO3	NM_006293	NP_006284
DDR1	NM_013993	NP_001945
DDR2	NM_006182	NP_006173
EGFR	NM_005228	NP_005219
ERBB2	NM_004448	NP_004439
ERBB3	NM_001982	NP_001973
ERBB4	NM_005235	NP_005226
EPHA1	NM_005232	NP_005223
EPHA2	NM_004431	NP_004422
EPHA3	NM_005233	NP_005224
EPHA4	NM_004438	NP_004429
EPHA5	L36644	P54756
EPHA6	AL133666	
EPHA7	NM_004440	NP_004431
EPHA8	AB040892	CAB81612
EPHB1	NM_004441	NP_004432
EPHB2	AF025304	AAB94602
EPHB3	NM_004443	NP_004434
EPHB4	NM_004444	NP_004435
EPHB6	NM_004445	NP_004436
EPHX		
FGFR1	M34641	AAA35835
FGFR2	NM_000141	NP_000132
FGFR3	NM_000142	NP_000133
FGFR4	NM_002011	NP_002002
IGF1R	NM_000875	NP_000866
INSR	NM_000208	NP_000199
INSRR	J05046	AAC31759
MET	NM_000245	NP_000236
RON	NM_002447	NP_002438
MUSK	NM_005592	NP_005583
CSF1R	NM_005211	NP_005202
FLT3	NM_004119	NP_0041110
KIT	NM_000222	NP_000213
PDGFRA	NM_006206	NP_006197
PDGFRB	NM_002609	NP_002600
PTK7	NM_002821	NP_002812
RET	X12949	P07949
ROR1	NM_005012	NP_005003
ROR2	NM_004560	NP_004551

【表 2 5 - 2】

ROS1	NM_002944	NP_002935
RYK	S59184	AAB26341
TEK	NM_000459	NP_000450
TIE	NM_005424	NP_005415
NTRK1	NM_002529	NP_002520
NTRK2	NM_006180	NP_006171
NTRK3	NM_002530	NP_002521
VEGFR1	NM_002019	NP_002010
VEGFR2		AAB88005
VEGFR3	NM_002020	NP_002011
AATYK	NM_004920	NP_004911
AATYK2	NM_014916	NP_055731
AATYK3		
DKFZp761P1010	NM_018423	NP_060893

10

【 0 2 5 4】

実施例 18：ヒトキノームのための gRNA ライブラリ

【 0 2 5 5】

20

DNA オリゴ核酸ライブラリを、518 のヒトキナーゼのための遺伝子を標的とする 5 つの sgRNA をコードする配列で設計し、異なる DNA のオリゴ核酸は合計 2590 であった。オリゴ核酸をプレートから取り除き、PCR により増幅し、ベクターへとクローン化し、細胞の予め選択された集団へと移入した。予め選択された 518 の遺伝子の発現レベルは、キナーゼ特異的 sgRNA の無い対照にさらされる細胞の対照集団に対する細胞の予め選択された集団において比較される。

【 0 2 5 6】

実施例 19：ヒトホスファターゼのための gRNA ライブラリ

【 0 2 5 7】

30

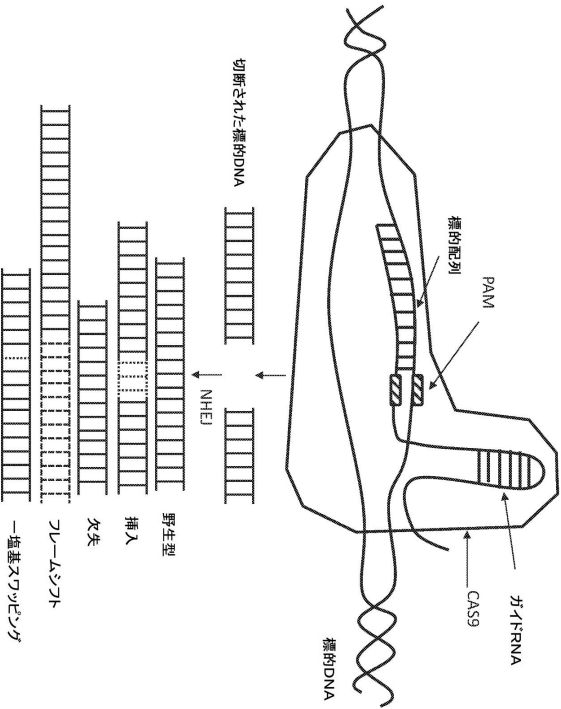
DNA オリゴ核酸ライブラリを、200 のヒトホスファターゼのための遺伝子を標的とする 5 つの sgRNA をコードする配列で設計し、異なる DNA のオリゴ核酸は合計 1000 であった。オリゴ核酸をプレートから取り除き、PCR により増幅し、ベクターへとクローン化し、細胞の予め選択された集団へと移入した。200 の予め選択された遺伝子の発現レベルは、キナーゼ特異的 sgRNA の無い対照にさらされる細胞の対照集団に対する細胞の予め選択された集団において比較される。

【 0 2 5 8】

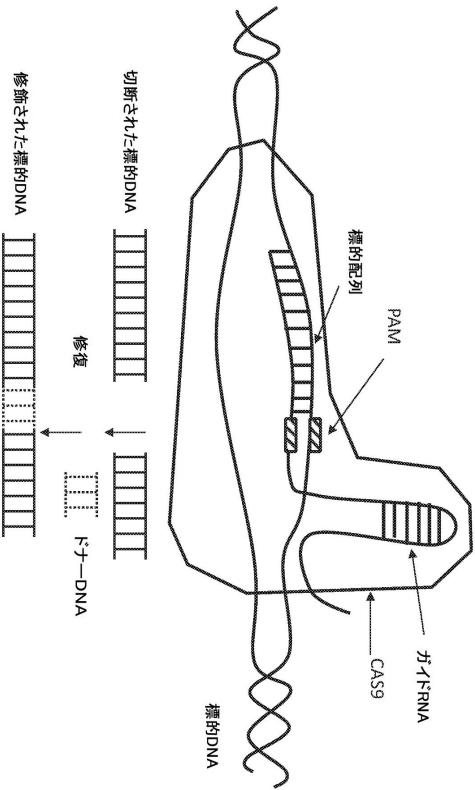
本発明の好ましい実施形態が本明細書中で示され且つ記載されてきたが、このような実施形態はほんの一例として提供されるものであることは、当業者に明らかであろう。多数の変形、変更、及び置き換えは、本発明から逸脱することなく、当業者によって現在想到されるものである。本明細書に記載される本発明の実施形態の様々な代案が、本発明の実施において利用されるかもしれないことを理解されたい。以下の特許請求の範囲は本発明の範囲を定義するものであり、この特許請求の範囲及びその同等物の範囲内の方法及び構造は、それにより包含されることが、意図されている。

40

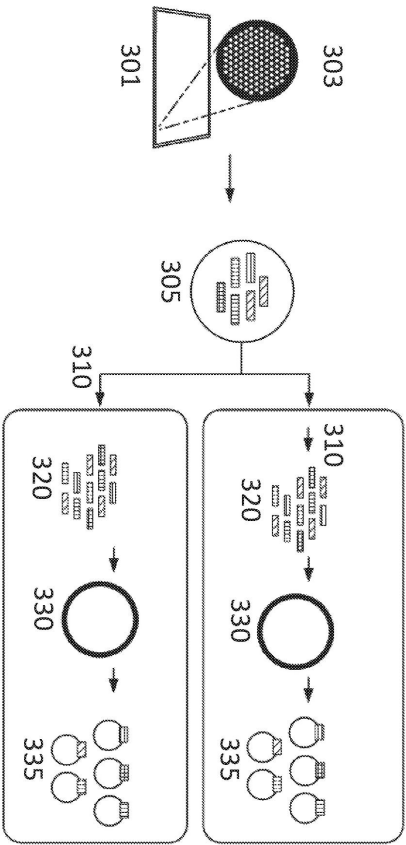
【図 1 B】



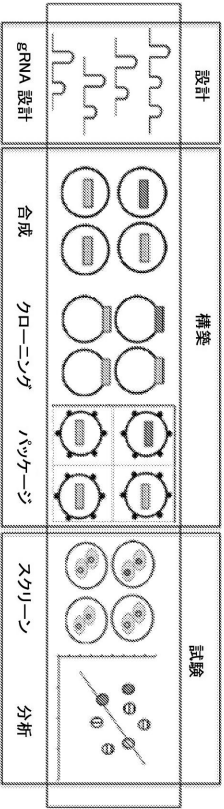
【図 1 A】



【図 3】



【図 2】



【 図 4 B 】

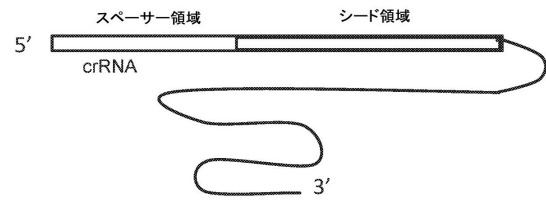
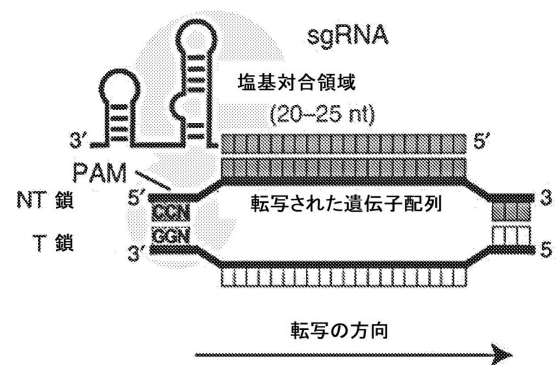
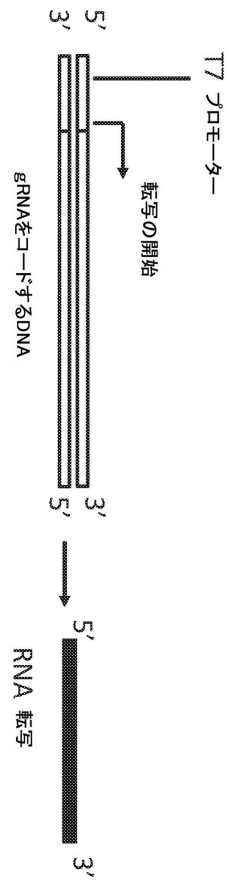


Diagram illustrating the interaction between crRNA and tracrRNA. The crRNA (CRISPR RNA) is shown as a horizontal bar with a 'Spacer domain' (スペーサー領域) and a 'Seed domain' (シード領域). The tracrRNA (CRISPR-associated RNA) is shown as a looped structure below it, with its 5' and 3' ends labeled. The crRNA is also labeled with 5' and 3' ends.

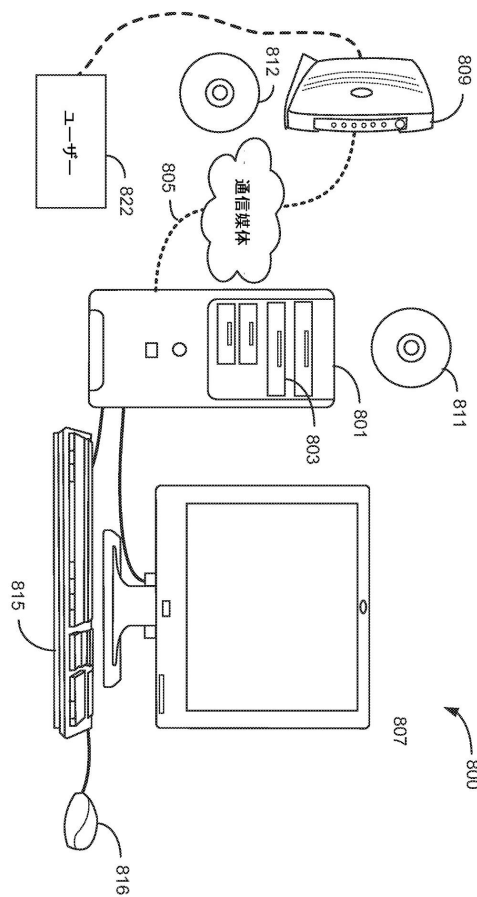
【 図 5 B 】



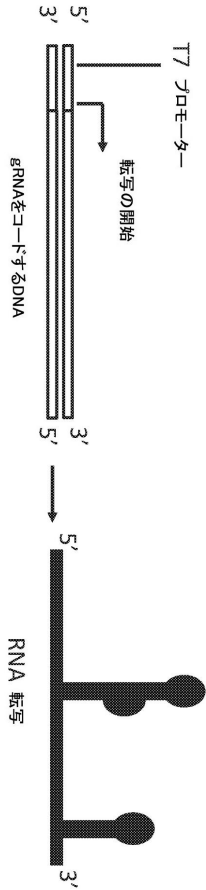
【図 6 B】



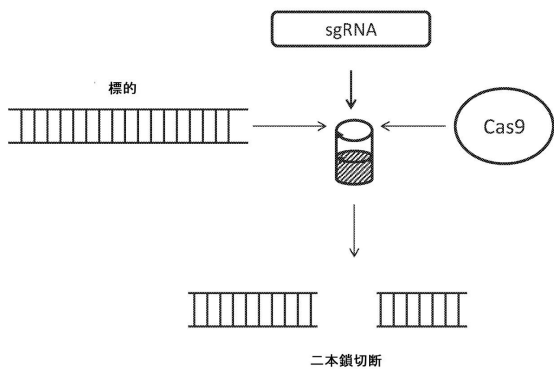
【図 8】



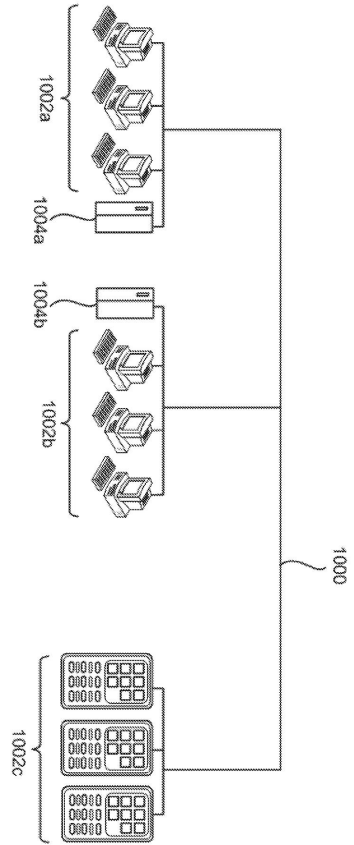
【図 6 A】



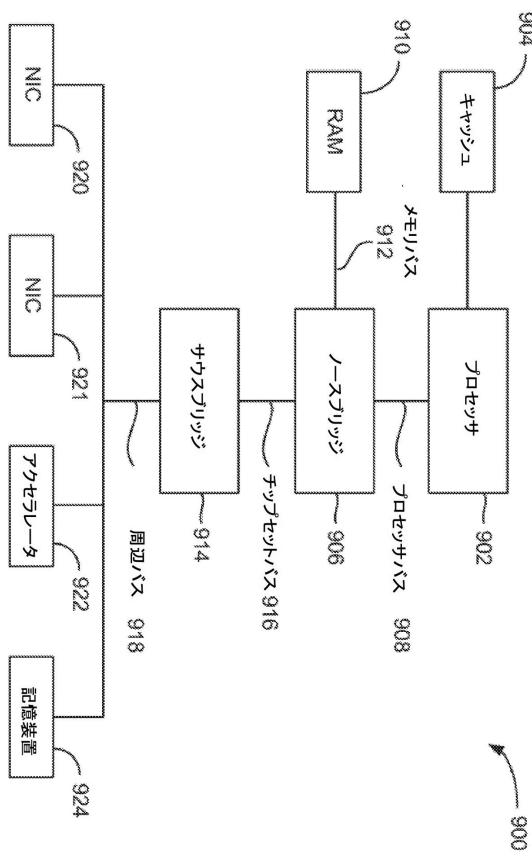
【図 7】



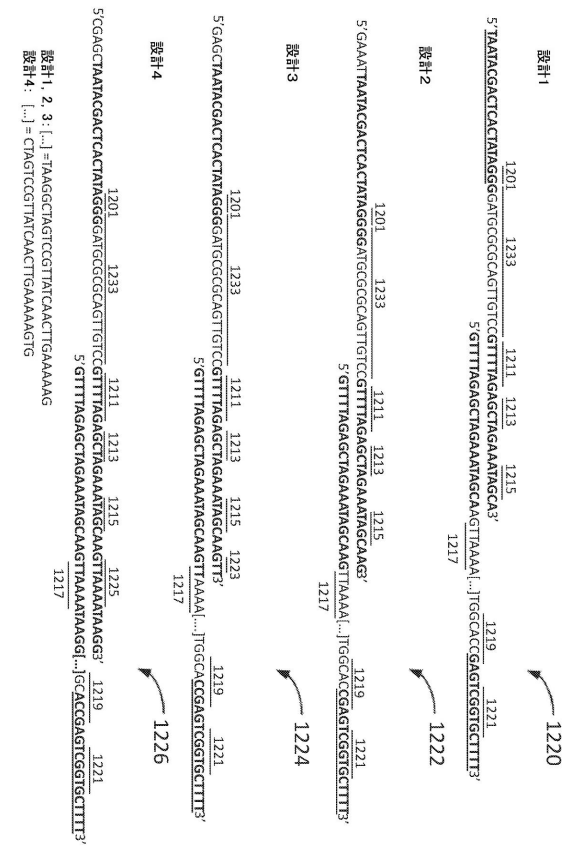
【 図 1 0 】



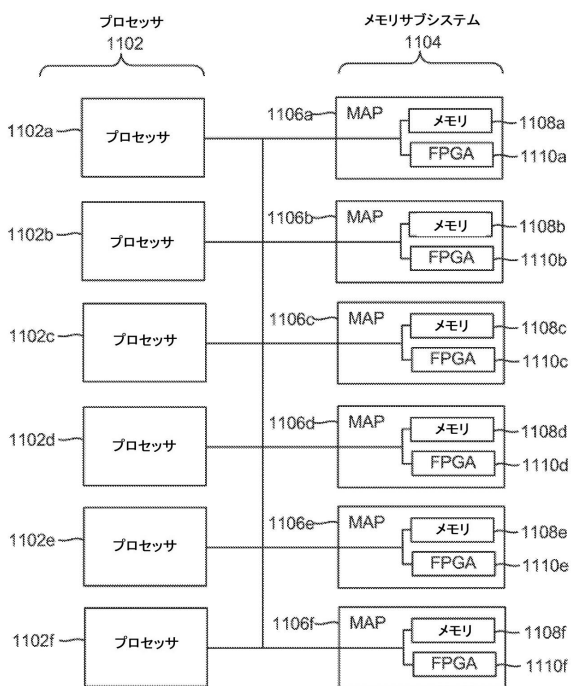
【 図 9 】



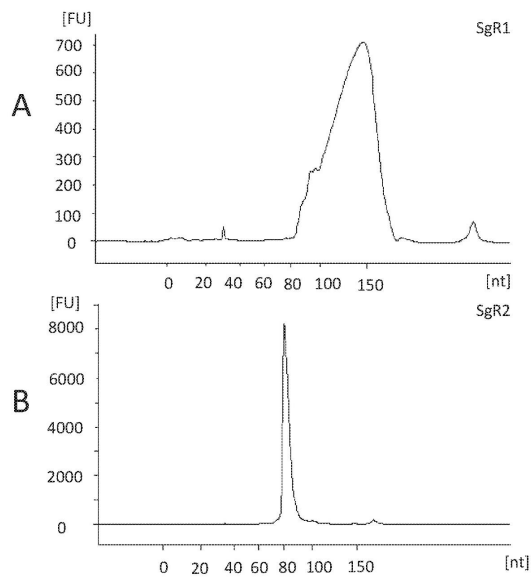
【 図 1 2 】



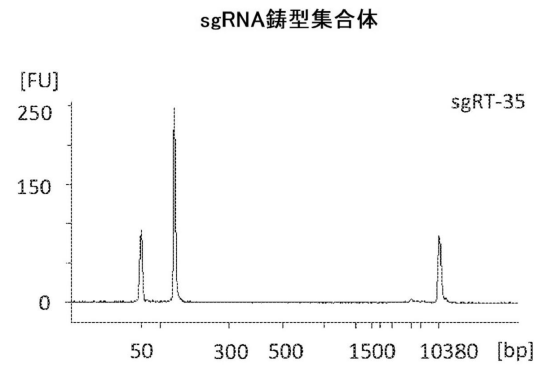
【 図 1 1 】



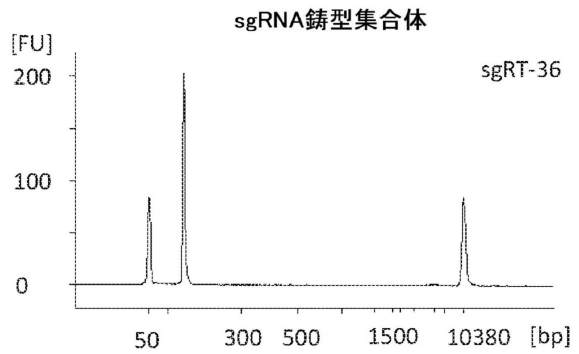
【図 1 3】



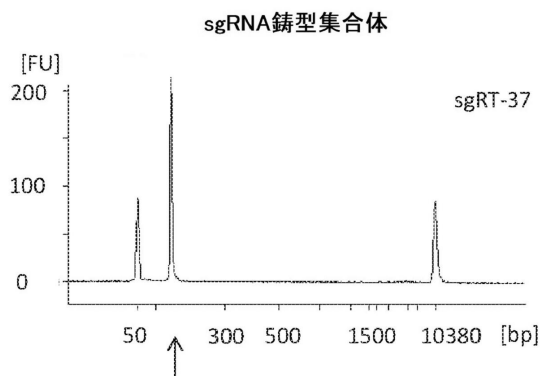
【図 1 4 A】



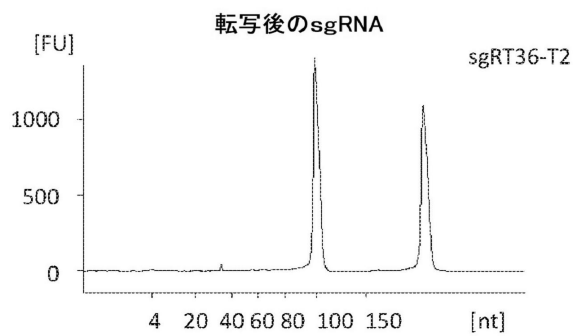
【図 1 4 B】



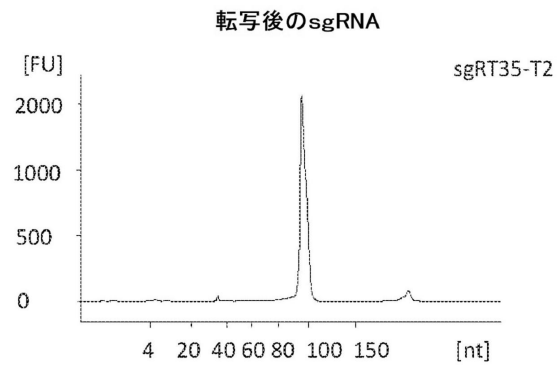
【図 1 4 C】



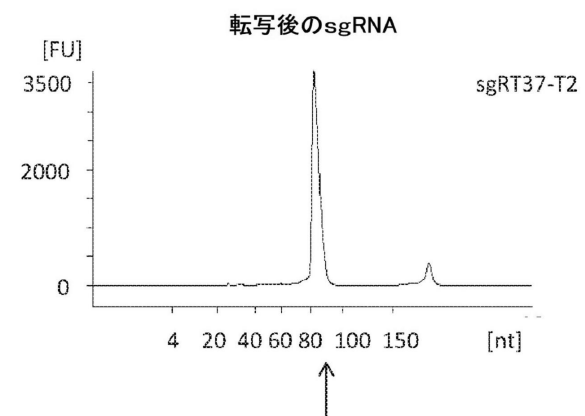
【図 1 4 E】



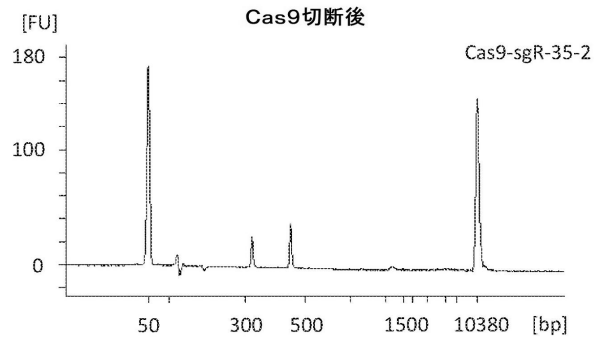
【図 1 4 D】



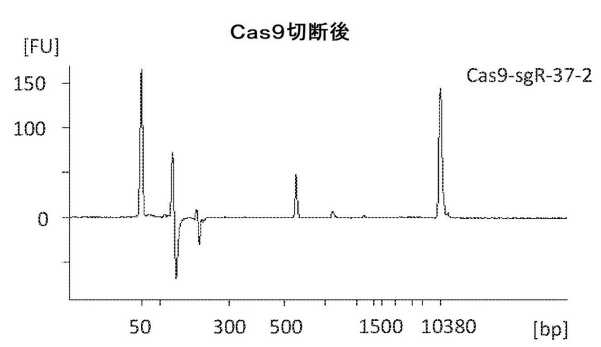
【図 1 4 F】



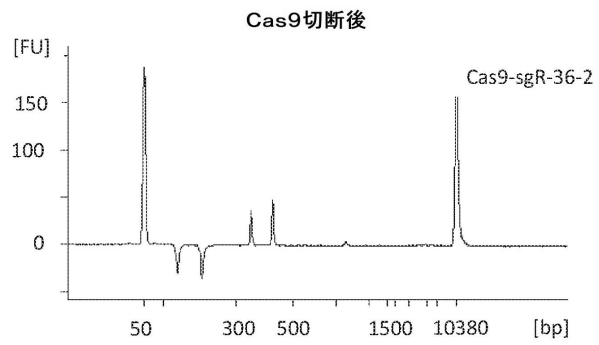
【図 1 4 G】



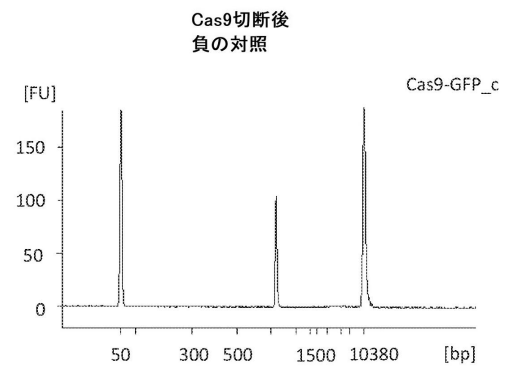
【図 1 4 I】



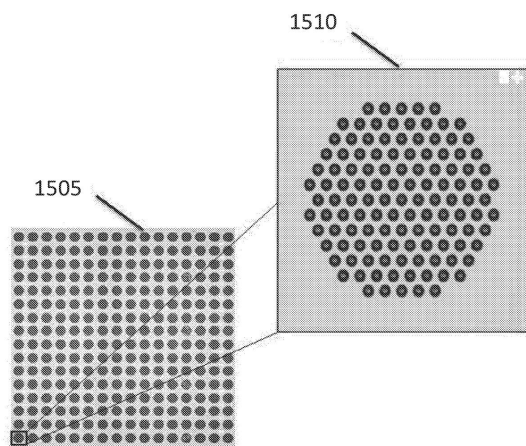
【図 1 4 H】



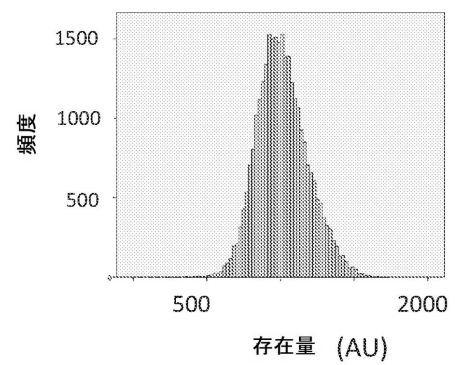
【図 1 4 J】



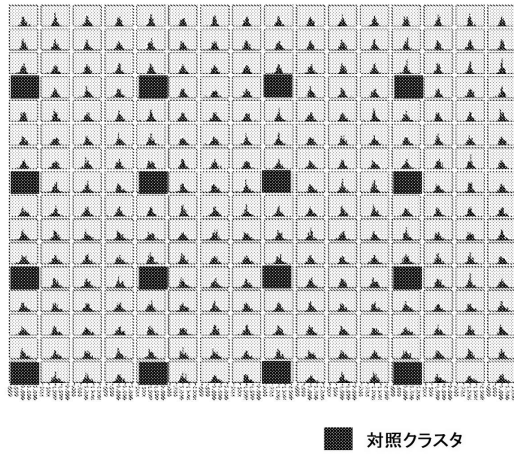
【図 1 5】



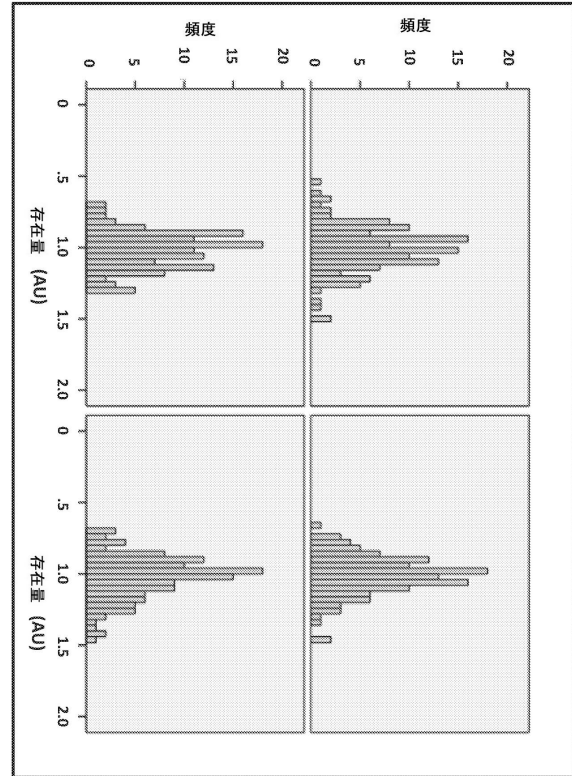
【図 1 6 A】



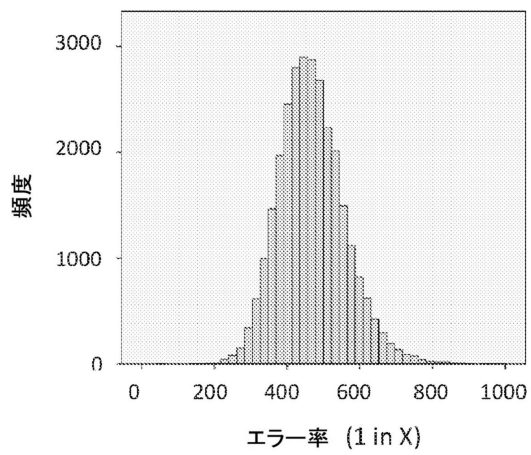
【図 16 B】



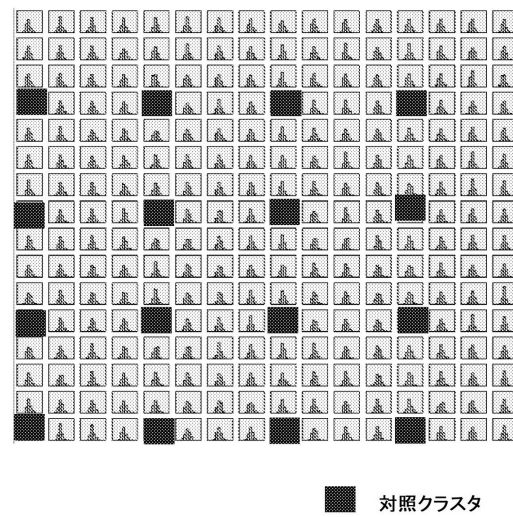
【図 17】



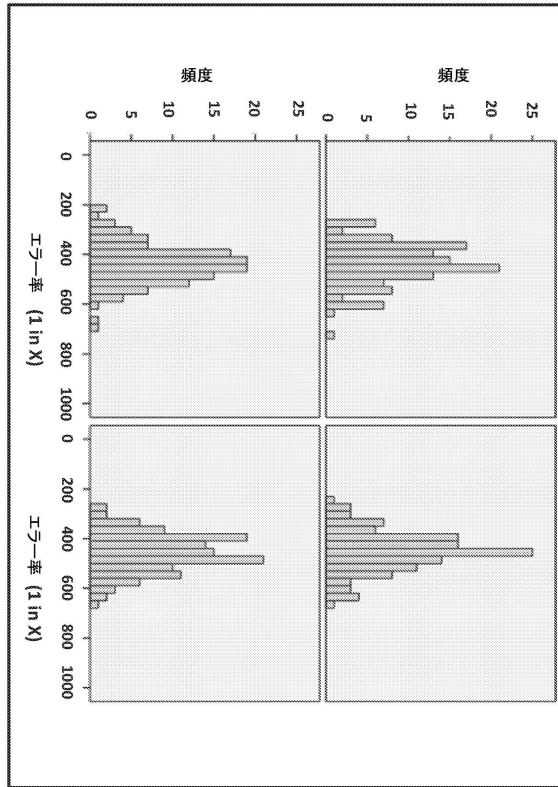
【図 18 A】



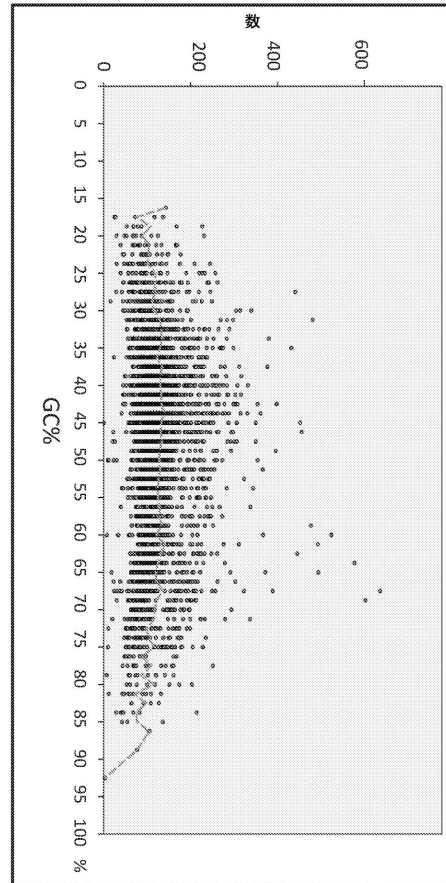
【図 18 B】



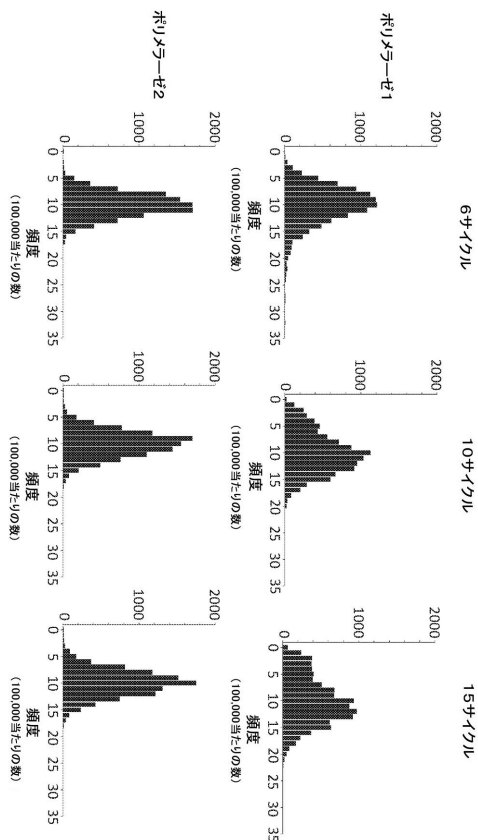
【図 19】



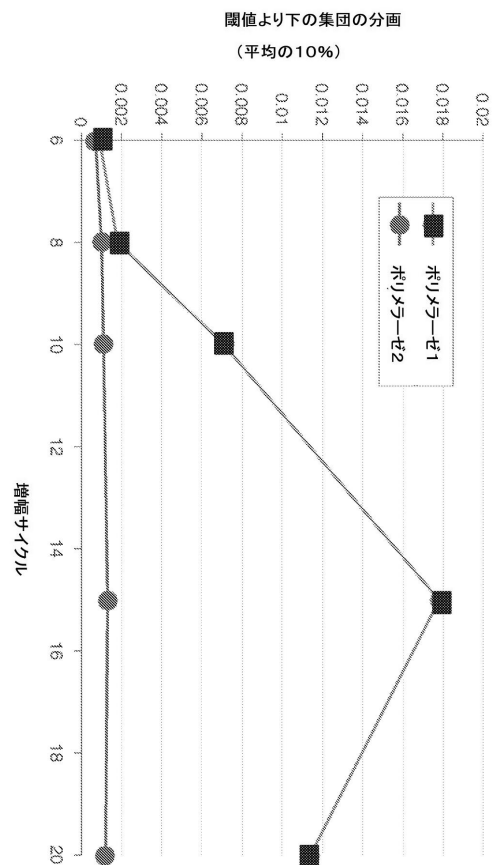
【図 20】



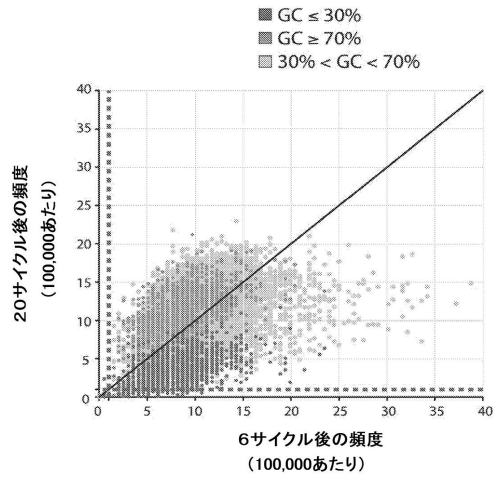
【図 21】



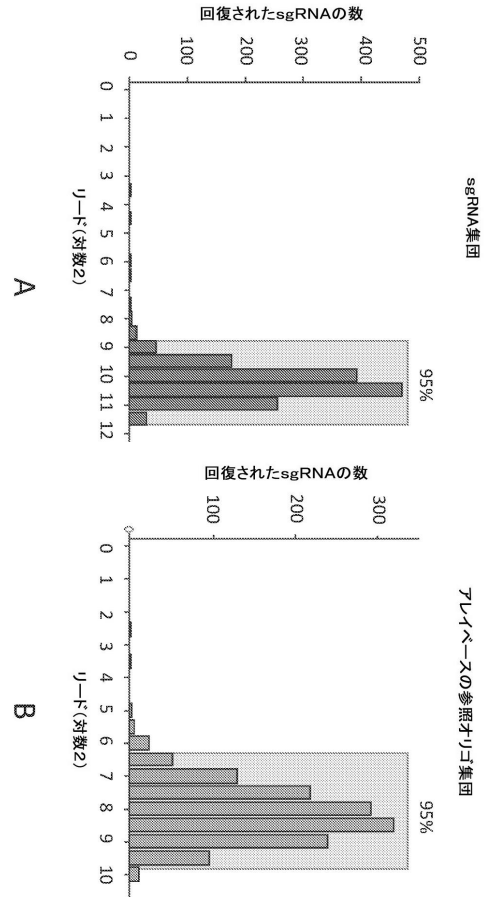
【図 22】



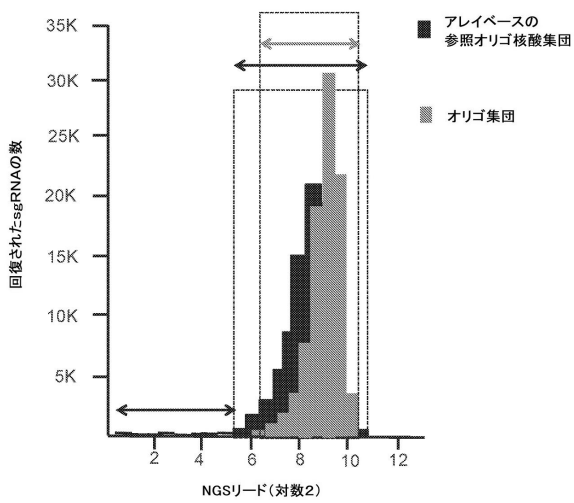
【図 2 3】



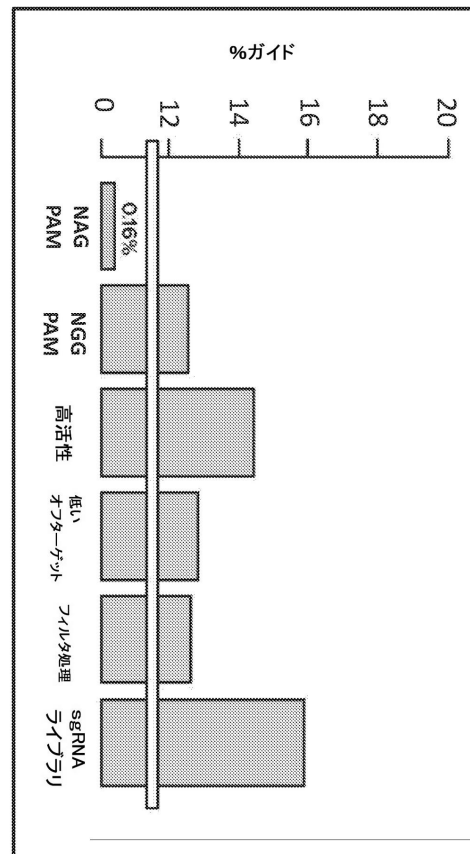
【図 2 4】



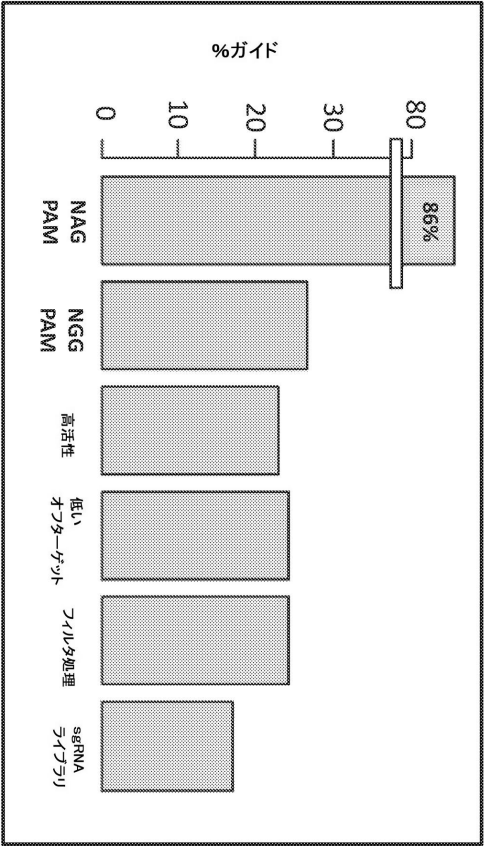
【図 2 5】



【図 2 6 A】



【図 2 6 B】



【配列表】

0006854340000001.app

フロントページの続き

(51)Int.Cl.			F I		
C 1 2 N	5/10	(2006.01)	C 1 2 N	5/10	
C 1 2 N	9/16	(2006.01)	C 1 2 N	9/16	Z
C 1 2 N	15/63	(2006.01)	C 1 2 N	15/63	Z
C 1 2 Q	1/6872	(2018.01)	C 1 2 Q	1/6872	Z
C 1 2 N	1/11	(2006.01)	C 1 2 N	1/11	
C 1 2 Q	1/6844	(2018.01)	C 1 2 Q	1/6844	Z
C 1 2 N	15/55	(2006.01)	C 1 2 N	15/55	

(31)優先権主張番号 62/393,948

(32)優先日 平成28年9月13日(2016.9.13)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/411,388

(32)優先日 平成28年10月21日(2016.10.21)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/419,881

(32)優先日 平成28年11月9日(2016.11.9)

(33)優先権主張国・地域又は機関

米国(US)

早期審査対象出願

(72)発明者 チェン, スーユアン

アメリカ合衆国 9 4 4 0 2 カリフォルニア州 サンマテオ アパディーン・ドライブ 1 3 0
2

審査官 松原 寛子

(56)参考文献 国際公開第2 0 1 5 / 0 2 1 0 8 0 (WO, A 1)

国際公開第2 0 1 6 / 0 1 1 0 8 0 (WO, A 1)

(58)調査した分野(Int.Cl., DB名)

C 1 2 N 1 5 / 1 1

C 4 0 B 4 0 / 0 6