

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号

特許第7229923号

(P7229923)

(45)発行日 令和5年2月28日(2023.2.28)

(24)登録日 令和5年2月17日(2023.2.17)

(51)国際特許分類

F I

C 1 2 N 15/09 (2006.01)

C 1 2 N 15/09 1 1 0

C 1 2 M 1/34 (2006.01)

C 1 2 M 1/34 Z

C 1 2 Q 1/6869(2018.01)

C 1 2 Q 1/6869 Z

C 1 2 Q 1/6844(2018.01)

C 1 2 Q 1/6844 Z Z N A

C 1 2 Q 1/6876(2018.01)

C 1 2 Q 1/6876 Z

請求項の数 55 (全70頁)

(21)出願番号 特願2019-536933(P2019-536933)

(86)(22)出願日 平成30年1月5日(2018.1.5)

(65)公表番号 特表2020-503056(P2020-503056
A)

(43)公表日 令和2年1月30日(2020.1.30)

(86)国際出願番号 PCT/US2018/012652

(87)国際公開番号 WO2018/129368

(87)国際公開日 平成30年7月12日(2018.7.12)

審査請求日 令和3年1月5日(2021.1.5)

(31)優先権主張番号 62/443,212

(32)優先日 平成29年1月6日(2017.1.6)

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/502,434

(32)優先日 平成29年5月5日(2017.5.5)

最終頁に続く

(73)特許権者 517372508

エディタス・メディシン、インコーポレ
イテッドアメリカ合衆国、マサチューセッツ州
0 2 1 4 1、ケンブリッジ、ハーレー・
ストリート 1 1

(74)代理人 100078282

弁理士 山本 秀策

(74)代理人 100113413

弁理士 森下 夏樹

(74)代理人 100181674

弁理士 飯田 貴敏

(74)代理人 100181641

弁理士 石川 大輔

(74)代理人 230113332

最終頁に続く

(54)【発明の名称】ヌクレアーゼ切断を評価する方法

(57)【特許請求の範囲】

【請求項1】

方法であって、

(a) 部位特異的ヌクレアーゼと接触させた細胞または組織から得られたゲノムDNA試料を、トランスポゾンの5'末端に第1の検出配列を含む前記トランスポゾンと、前記トランスポゾンが前記ゲノムDNA試料に挿入され、かつ前記ゲノムDNA試料が断片化されて、複数のタグメンテーションされた二本鎖核酸断片であって、前記核酸断片の5'末端に付着された前記トランスポゾンを含む二本鎖核酸断片になる条件下で接触させることと；

(b)

(i) 前記ゲノムDNA内のあらかじめ決定された位置に相補的なヌクレオチド配列を含み、かつその5'末端に第2の検出配列を含む第1の固定プライマーと、

(ii) 前記第1の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第2の選択的プライマーと

を使用して、前記タグメンテーションされた核酸断片を増幅させて、前記第1の検出配列、前記核酸断片の前記5'末端に付着された前記トランスポゾン、および前記第2の検出配列を含む、増幅された核酸断片を形成させることと；

(c) 前記増幅された核酸断片を配列決定することとを含む方法。

【請求項2】

前記第1の検出配列が、第1の配列決定タグを含む、請求項1に記載の方法。

10

20

【請求項 3】

ステップ(c)が、前記増幅された核酸断片を、前記第1の配列決定タグとハイブリッド形成する第1の配列決定プライマーと接触させることを含む、請求項2に記載の方法。

【請求項 4】

前記第2の検出配列が、第2の配列決定タグを含む、請求項1～3のいずれか一項に記載の方法。

【請求項 5】

ステップ(c)が、前記増幅された核酸断片を、前記第2の配列決定タグとハイブリッド形成する第2の配列決定プライマーと接触させることを含む、請求項4に記載の方法。

【請求項 6】

前記ゲノムDNA試料が、単一の細胞から得られる、請求項1～5のいずれか一項に記載の方法。

【請求項 7】

ステップ(b)が、それぞれが前記ゲノムDNA内の異なるあらかじめ決定された位置に相補的なヌクレオチド配列を含む、複数の第1の固定プライマーを使用することを含む、請求項1～6のいずれか一項に記載の方法。

【請求項 8】

ステップ(b)が、それぞれの反応が異なる第1の固定プライマーを使用する、複数の増幅反応を実施することを含む、請求項7に記載の方法。

【請求項 9】

前記複数の第1の固定プライマーのそれぞれが、(i)同じ配列決定タグと(ii)固有のバーコードとを含む検出配列を含む、請求項7または8に記載の方法。

【請求項 10】

前記二本鎖核酸断片が、400～6000bpを含む、請求項1～9のいずれか一項に記載の方法。

【請求項 11】

前記部位特異的ヌクレアーゼが、Cas9である、請求項1～10のいずれか一項に記載の方法。

【請求項 12】

前記部位特異的ヌクレアーゼが、Cpf1である、請求項1～10のいずれか一項に記載の方法。

【請求項 13】

方法であって、

(a)細胞または組織から得られたゲノムDNA試料を、RNAガイド型ヌクレアーゼと、前記ヌクレアーゼにより前記ゲノムDNAが切断される条件下で接触させることと；
(b)前記切断されたゲノムDNAを、トランスポゾンの5'末端に第1の検出配列を含む前記トランスポゾンと、前記切断されたゲノムDNAに前記トランスポゾンが挿入され、かつ前記切断されたゲノムDNAが断片化されて、複数のタグメンテーションされた二本鎖核酸断片であって、前記核酸断片の5'末端に付着された前記トランスポゾンを含む二本鎖核酸断片になる条件下で接触させることと；

(c)

(i)前記ゲノムDNA内のあらかじめ決定された位置に相補的なヌクレオチド配列を含み、かつその5'末端に第2の検出配列を含む第1の固定プライマーと、

(ii)前記第1の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第2の選択的プライマーと

を使用して、前記タグメンテーションされた核酸断片を増幅させて、前記第1の検出配列、前記核酸断片の前記5'末端に付着された前記トランスポゾン、および前記第2の検出配列を含む、増幅された核酸断片を形成させることと；

(d)増幅された核酸断片を配列決定することとを含む方法。

10

20

30

40

50

【請求項 14】

前記第 1 の検出配列が、第 1 の配列決定タグを含む、請求項 13 に記載の方法。

【請求項 15】

ステップ (d) が、前記増幅された核酸断片を、前記第 1 の配列決定タグとハイブリッド形成する第 1 の配列決定プライマーと接触させることを含む、請求項 14 に記載の方法。

【請求項 16】

前記第 2 の検出配列が、第 2 の配列決定タグを含む、請求項 13 ~ 15 のいずれか一項に記載の方法。

【請求項 17】

ステップ (d) が、前記増幅された核酸断片を、前記第 2 の配列決定タグとハイブリッド形成する第 2 の配列決定プライマーと接触させることを含む、請求項 16 に記載の方法。

【請求項 18】

前記ゲノム DNA 試料が、単一の細胞から得られる、請求項 13 ~ 17 のいずれか一項に記載の方法。

【請求項 19】

ステップ (c) が、それぞれが前記ゲノム DNA 内の異なるあらかじめ決定された位置に相補的なヌクレオチド配列を含む、複数の第 1 の固定プライマーを使用することを含む、請求項 13 ~ 18 のいずれか一項に記載の方法。

【請求項 20】

ステップ (c) が、それぞれの反応が異なる第 1 の固定プライマーを使用する、複数の増幅反応を実施することを含む、請求項 19 に記載の方法。

【請求項 21】

前記複数の第 1 の固定プライマーのそれぞれが、(i) 同じ配列決定タグと (i i) 固有のバーコードとを含む検出配列を含む、請求項 19 または 20 に記載の方法。

【請求項 22】

前記二本鎖核酸断片が、400 ~ 6000 bp を含む、請求項 13 ~ 21 のいずれか一項に記載の方法。

【請求項 23】

前記 RNA ガイド型ヌクレアーゼが、Cas9 である、請求項 13 ~ 22 のいずれか一項に記載の方法。

【請求項 24】

前記 RNA ガイド型ヌクレアーゼが、Cpf1 である、請求項 13 ~ 22 のいずれか一項に記載の方法。

【請求項 25】

方法であって、

(a) 細胞または組織から得られたゲノム DNA 試料を、トランスポゾンの 5' 末端に第 1 の検出配列を含む前記トランスポゾンと、前記トランスポゾンが前記ゲノム DNA に挿入され、かつ前記ゲノム DNA が断片化されて、複数のタグメンテーションされた二本鎖核酸断片であって、前記核酸断片の 5' 末端に付着された前記トランスポゾンを含む二本鎖核酸断片になる条件下で接触させることであって、前記ゲノム DNA は、部位特異的ヌクレアーゼによって切断された少なくとも 1 つの部位に組み込まれた二本鎖オリゴヌクレオチドを含む、ことと；

(b)

(i) 前記二本鎖オリゴヌクレオチドの少なくとも一部分に相補的なヌクレオチド配列を含み、かつその 5' 末端に第 2 の検出配列を含む第 1 の固定プライマーと、

(i i) 前記第 1 の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第 2 の選択的プライマーと

を使用して、前記タグメンテーションされた核酸断片を増幅させて、前記第 1 の検出配列、前記核酸断片の前記 5' 末端に付着された前記トランスポゾン、および前記第 2 の検出配列を含む、増幅された核酸断片を形成させることと；

10

20

30

40

50

(c) 前記増幅された核酸断片を配列決定することとを含む方法。

【請求項 26】

前記部位特異的ヌクレアーゼが、Cas9である、請求項 25 に記載の方法。

【請求項 27】

前記部位特異的ヌクレアーゼが、Cpf1である、請求項 25 に記載の方法。

【請求項 28】

(a) 細胞または組織から得られたゲノムDNA試料を、RNAガイド型ヌクレアーゼおよび二本鎖オリゴヌクレオチドと、前記ヌクレアーゼによりゲノムDNAが切断され、および前記二本鎖オリゴヌクレオチドが少なくとも1つの切断部位に組み込まれる条件下で接触させることであって、それによって、組み込まれた二本鎖オリゴヌクレオチドを含むタグ付けされたゲノムDNAが産生される、ことと；

(b) 前記タグ付けされたゲノムDNAを、トランスポゾンの5'末端に第1の検出配列を含む前記トランスポゾンと、前記タグ付けされたゲノムDNAに前記トランスポゾンが挿入され、かつ前記タグ付けされたゲノムDNAが断片化されて、複数のタグメンテーションされた二本鎖核酸断片であって、前記核酸断片の5'末端に付着された前記トランスポゾンを含む二本鎖核酸断片になる条件下で接触させることと；

(c)

(i) 前記二本鎖オリゴヌクレオチドの少なくとも一部分に相補的なヌクレオチド配列を含み、かつその5'末端に第2の検出配列を含む第1の固定プライマーと、

(ii) 前記第1の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第2の選択的プライマーと

を使用して、前記タグメンテーションされた核酸断片を増幅させて、前記第1の検出配列、前記核酸断片の前記5'末端に付着された前記トランスポゾン、および前記第2の検出配列を含む、増幅された核酸断片を形成させることと；

(d) 前記増幅された核酸断片を配列決定することとを含む方法。

【請求項 29】

前記RNAガイド型ヌクレアーゼが、Cas9である、請求項 28 に記載の方法。

【請求項 30】

前記RNAガイド型ヌクレアーゼが、Cpf1である、請求項 28 に記載の方法。

【請求項 31】

前記二本鎖オリゴヌクレオチドのどちらの鎖も、前記ゲノムDNAに対してオルソログスである、請求項 25 ~ 30 のいずれか一項に記載の方法。

【請求項 32】

前記二本鎖オリゴヌクレオチドの各5'末端が、リン酸基を含む、請求項 25 ~ 31 のいずれか一項に記載の方法。

【請求項 33】

前記二本鎖オリゴヌクレオチドの各3'末端が、ホスホロチオエート結合を含む、請求項 25 ~ 32 のいずれか一項に記載の方法。

【請求項 34】

前記二本鎖オリゴヌクレオチドの各5'末端が、ホスホロチオエート結合を含む、請求項 25 ~ 31、および、請求項 32 を引用しない場合の請求項 33 のいずれか一項に記載の方法。

【請求項 35】

前記二本鎖オリゴヌクレオチドが、30 ~ 35 個のヌクレオチドまたは60 ~ 65 個のヌクレオチドを含む、請求項 25 ~ 34 のいずれか一項に記載の方法。

【請求項 36】

前記二本鎖オリゴヌクレオチドが、平滑末端化される、請求項 25 ~ 35 のいずれか一項に記載の方法。

10

20

30

40

50

【請求項 37】

前記二本鎖オリゴヌクレオチドが、1、2、3、または4塩基の突出ヌクレオチドを有する5'末端を含む、請求項25～35のいずれか一項に記載の方法。

【請求項 38】

前記第1の検出配列が、第1の配列決定タグを含む、請求項25～37のいずれか一項に記載の方法。

【請求項 39】

配列決定することが、前記増幅された核酸断片を、前記第1の配列決定タグとハイブリッド形成する第1の配列決定プライマーと接触させることを含む、請求項38に記載の方法。

10

【請求項 40】

前記第2の検出配列が、第2の配列決定タグを含む、請求項25～39のいずれか一項に記載の方法。

【請求項 41】

配列決定することが、前記増幅された核酸断片を、前記第2の配列決定タグとハイブリッド形成する第2の配列決定プライマーと接触させることを含む、請求項40に記載の方法。

【請求項 42】

前記二本鎖核酸断片が、400～6000bpを含む、請求項20～41のいずれか一項に記載の方法。

20

【請求項 43】

前記配列決定された核酸断片が、参照ゲノムDNA試料に対する転座を含む、請求項1～42のいずれか一項に記載の方法。

【請求項 44】

前記ゲノムDNA試料が、ウイルスによって導入された配列を含む、請求項1～43のいずれか一項に記載の方法。

【請求項 45】

前記ゲノムDNAを前記トランスポゾンと接触させる前に、ウイルスベクターを使用して、配列を前記ゲノムDNAに導入するステップをさらに含む、請求項1～12および25～27のいずれか一項に記載の方法。

30

【請求項 46】

前記配列決定された核酸断片を参照ゲノムDNA試料と比較することをさらに含む、請求項1～45のいずれか一項に記載の方法。

【請求項 47】

少なくとも1つの配列決定された核酸断片が、前記参照ゲノムDNA試料に対する転座を含む、請求項46に記載の方法。

【請求項 48】

前記配列決定された核酸断片の1%未満が、前記参照ゲノムDNA試料に対する転座を含む、請求項47に記載の方法。

【請求項 49】

前記ゲノムDNA試料が、生きている細胞または組織を含有しない試料から得られる、請求項1～48のいずれか一項に記載の方法。

40

【請求項 50】

前記ゲノムDNA試料が、痕跡量の細胞および/または組織を含有する試料から得られる、請求項1～48のいずれか一項に記載の方法。

【請求項 51】

前記ゲノムDNA試料が、真核細胞から得られる、請求項1～48または50のいずれか一項に記載の方法。

【請求項 52】

前記ゲノムDNA試料が、原核細胞から得られる、請求項1～48または50のいずれ

50

か一項に記載の方法。

【請求項 5 3】

前記ゲノム DNA 試料が、1 つまたは複数の微生物叢から得られる、請求項 1 ~ 5、1 3 ~ 1 7 および 2 5 ~ 4 1 のいずれか一項に記載の方法。

【請求項 5 4】

前記切断されたゲノム DNA を前記トランスポゾンと接触させる前に、ウイルスベクターを使用して、配列を前記ゲノム DNA に導入するステップをさらに含む、請求項 1 3 ~ 2 4 のいずれか一項に記載の方法。

【請求項 5 5】

前記タグ付けされたゲノム DNA を前記トランスポゾンと接触させる前に、ウイルスベクターを使用して、配列を前記ゲノム DNA に導入するステップをさらに含む、請求項 2 8 ~ 3 0 のいずれか一項に記載の方法。

10

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

関連出願の相互参照

本願は、2017年10月10日に出願された米国仮特許出願第62/570,300号明細書、2017年5月5日に出願された同第62/502,434号明細書、および2017年1月6日に提出された同第62/443,212号明細書の利益を主張するものであり、これらの明細書のそれぞれの内容全体が参照によって本明細書に組み込まれる。

20

【背景技術】

【0 0 0 2】

ジンクフィンガーヌクレアーゼ (ZFN)、転写活性化因子様エフェクターヌクレアーゼ (TALEN)、およびクラスター化して規則的な配置の短い回文配列リピート (clustered regularly-interspersed short palindromic repeat) (CRISPR) 関連ヌクレアーゼなどのヌクレアーゼは、特定の DNA 配列を標的にするその能力のおかげで、ますます使用されるようになっている。遺伝性疾患の治療のための手段としてのこれらのものなどのヌクレアーゼの価値は、広く認識されている。例えば、アメリカ食品医薬品局 (U.S. Food and Drug Administration) (FDA) は、2016年11月15日に、こうしたシステムの使用、およびそれによって生じる法的問題点の可能性に取り組む科学委員会 (Science Board Meeting) を開催した。この委員会では、FDA は、Cas9/ガイド RNA (gRNA) リボ核タンパク質 (RNP) 複合体を、対象となる遺伝子座での正確な編集をもたらすようにカスタマイズすることができるが、この複合体は、他の「オフターゲット」位置とも相互作用し、切断する可能性があることを指摘した。オフターゲット切断 (「オフターゲット」) の可能性は、ひいては、少なくとも、これらのヌクレアーゼを利用する治療法の承認に関する法的問題点の可能性を生じる。

30

ゲノム編集の修復結果を解析するための一般的方法は、標的とされるゲノム領域の PCR 増幅、および塩基ミスマッチでのエンドヌクレアーゼ切断または配列決定によるその後の解析を含む (Mashal et al. Nat. Genet. 1995; 9:177-183; Vouilliot et al. G3 2015; 5:407-415; Hendel et al. Trends Biotechnol. 2015; 33:132-140; Tykco et al. Mol. Cell. 2016; 63:355-370; Brinkman et al. Nucleic Acids Res. 2014; 42:e168)。しかし、PCR 介在性の分析は、根本的に、小さいインデルと共に、ゲノムへの構造的変化 (例えば、欠失 (100 bp よりも大きいもの)、逆位、および転座) を測定することができない。予期せぬ転座、および他の構造的変化は、NIH の組換え DNA 諮問委員会 (Recombinant DNA Advisory Committee) (RAC) および FDA によるゲノム編集療法の研究について、具体的に挙げられている (Atkins et al. RAC Review 2016; Witten, Cell a

40

50

nd Gene Meeting on the Mesa, La Jolla, CA, 2016)。構造的変化を測定することは、発癌性の転座の臨床的検出を意図するAMP-seq (アンカードマルチプレックスPCR (Anchored Multiplex PCR) 配列決定) と呼ばれる方法、およびHTGTS (ハイスループット、ゲノムワイド転座配列決定 (High Throughput, Genome-wide Translocation Sequencing)) またはLAM-HTGTS (線形増幅を介するHTGTS (Linear Amplification-Mediated-HTGTS)) と呼ばれる類似の方法を使用して、最近、より実現可能になりつつある (Zheng et al. Nat. Med. 2014; 20: 1479 - 1484; Chiarle et al. Cell 2011; 147: 107 - 119; Hu et al. Nat. Protocol. 2016; 11: 853 - 871)。GUIDE-seq、すなわちAMP-seqの改変は、CRISPR RNAガイド型ヌクレアーゼによる新規オフターゲット編集を捉えるための強力なツールである (Tsai et al. Nat. Biotechnol. 2014; 33: 187 - 197)。これらの方法はすべて、「一方向性の」増幅、および配列決定を達成するために、切断されたDNA上のアダプター連結ユニバーサルプライミング部位に加えて、標的特異的なプライマーを利用する。しかし、DNA切断は、これらのすべての方法において使用されるライブラリ調製における厄介なステップである。DNA切断は、特殊化された設備を必要とし、また、多数の試料の研究を容易に行いやすくはない。したがって、DNA切断は、遺伝子編集分野において、広くは適用されていない。さらに、DNAを切断することは、低レベル (例えば: 1%未満) の遺伝子編集頻度を評価する場合に問題となり得る塩基のミスコールをもたらすDNA損傷を誘発することが示されている (Chen et al. Science 2017; 355: 752 - 6; Costello et al. Nucleic Acids Res. 2013; 41: 1 - 12)。

【先行技術文献】

【非特許文献】

【0003】

【文献】Mashal et al. Nat. Genet. 1995; 9: 177 - 183
 Vouilliot et al. G3 2015; 5: 407 - 415
 Hendel et al. Trends Biotechnol. 2015; 33: 132 - 140
 Tykco et al. Mol. Cell. 2016; 63: 355 - 370
 Brinkman et al. Nucleic Acids Res. 2014; 42: e168
 Atkins et al. RAC Review 2016
 Witten, Cell and Gene Meeting on the Mesa, La Jolla, CA, 2016
 Zheng et al. Nat. Med. 2014; 20: 1479 - 1484
 Chiarle et al. Cell 2011; 147: 107 - 119
 Hu et al. Nat. Protocol. 2016; 11: 853 - 871
 Tsai et al. Nat. Biotechnol. 2014; 33: 187 - 197
 Chen et al. Science 2017; 355: 752 - 6
 Costello et al. Nucleic Acids Res. 2013; 41: 1 - 12

【発明の概要】

【課題を解決するための手段】

【0004】

本開示は、特に、トランスポゾン組成物を使用してDNAを断片化およびタグ付けすることによるヌクレアーゼ活性の特徴付けのための方法およびシステムを提供する。一態様では、この開示は、(a) 部位特異的ヌクレアーゼと接触させた細胞または組織から得ら

れたゲノムDNA試料を、5'末端に第1の検出配列を含むトランスポゾンと、トランスポゾンがゲノムDNAに挿入され、かつゲノムDNAが断片化されて核酸断片の5'末端に付着されたトランスポゾンを含む複数のタグメンテーションされた(tagmented)二本鎖核酸断片になる条件下で接触させることと；(b)(i)ゲノムDNA内のあらかじめ決定された位置に相補的なヌクレオチド配列を含み、かつその5'末端に第2の検出配列を含む第1の固定のプライマーと、(ii)第1の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第2の選択的プライマーとを使用して、タグメンテーションされた核酸断片を増幅させて、第1の検出配列、核酸断片の5'末端に付着されたトランスポゾン、および第2の検出配列を含む、増幅された核酸断片を形成させることと；(c)増幅された核酸断片を配列決定することを含む方法の特徴とする。

10

【0005】

いくつかの実施態様では、第1の検出配列は、第1の配列決定タグを含む。いくつかの実施態様では、ステップ(c)は、増幅された核酸断片を、第1の配列決定タグとハイブリッド形成する第1の配列決定プライマーと接触させることを含む。いくつかの実施態様では、第2の検出配列は、第2の配列決定タグを含む。いくつかの実施態様では、ステップ(c)は、増幅された核酸断片を、第2の配列決定タグとハイブリッド形成する第2の配列決定プライマーと接触させることを含む。

【0006】

いくつかの実施態様では、ゲノムDNA試料は、単一の細胞から得られる。

【0007】

20

いくつかの実施態様では、ステップ(b)は、それぞれがゲノムDNA内の異なるあらかじめ決定された位置に相補的なヌクレオチド配列を含む、複数の第1の固定プライマーを使用することを含む。いくつかの実施態様では、ステップ(b)は、それぞれの反応が異なる第1の固定プライマーを使用する、複数の増幅反応を実施することを含む。いくつかの実施態様では、複数の第1の固定プライマーのそれぞれは、(i)同じ配列決定タグと(ii)固有のバーコードとを含む検出配列を含む。いくつかの実施態様では、複数の第1の固定プライマーが、多重化アッセイで使用される。

【0008】

いくつかの実施態様では、二本鎖核酸断片は、約500~約5000bp、例えば、約100~約4000bp、例えば、約500~約3000bpを含む。

30

【0009】

いくつかの実施態様では、部位特異的ヌクレアーゼは、Cas9である。

【0010】

別の態様では、この開示は、(a)RNAガイド型ヌクレアーゼと接触させた細胞または組織から得られたゲノムDNA試料を、ヌクレアーゼによるゲノムDNAの切断に好都合な条件下で接触させることと；(b)切断されたゲノムDNAを、5'末端に第1の検出配列を含むトランスポゾンと、切断されたゲノムDNAにトランスポゾンが挿入され、かつ切断されたゲノムDNAが断片化されて核酸断片の5'末端に付着されたトランスポゾンを含む複数のタグメンテーションされた二本鎖核酸断片になる条件下で接触させることと；(c)(i)ゲノムDNA内のあらかじめ決定された位置に相補的なヌクレオチド配列を含み、かつその5'末端に第2の検出配列を含む第1の固定プライマーと、(ii)第1の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第2の選択的プライマーとを使用して、タグメンテーションされた核酸断片を増幅させて、第1の検出配列、核酸断片の5'末端に付着されたトランスポゾン、および第2の検出配列を含む、増幅された核酸断片を形成させることと；(d)増幅された核酸断片を配列決定することを含む方法の特徴とする。

40

【0011】

いくつかの実施態様では、第1の検出配列は、第1の配列決定タグを含む。いくつかの実施態様では、ステップ(d)は、増幅された核酸断片を、第1の配列決定タグとハイブリッド形成する第1の配列決定プライマーと接触させることを含む。いくつかの実施態様

50

では、第2の検出配列は、第2の配列決定タグを含む。いくつかの実施態様では、ステップ(d)は、増幅された核酸断片を、第2の配列決定タグとハイブリッド形成する第2の配列決定プライマーと接触させることを含む。

【0012】

いくつかの実施態様では、ゲノムDNA試料は、単一の細胞から得られる。

【0013】

いくつかの実施態様では、ステップ(c)は、それぞれがゲノムDNA内の異なるあらかじめ決定された位置に相補的なヌクレオチド配列を含む、複数の第1の固定プライマーを使用することを含む。いくつかの実施態様では、ステップ(c)は、それぞれの反応が異なる第1の固定プライマーを使用する、複数の増幅反応を実施することを含む。いくつかの実施態様では、複数の第1の固定プライマーのそれぞれは、(i)同じ配列決定タグと(ii)固有のバーコードとを含む検出配列を含む。いくつかの実施態様では、複数の第1の固定プライマーが、多重化アッセイで使用される。

10

【0014】

いくつかの実施態様では、二本鎖核酸断片は、約500~約5000bp、例えば、約100~約4000bp、例えば、約500~約3000bpを含む。

【0015】

いくつかの実施態様では、RNAガイド型ヌクレアーゼは、Cas9である。

【0016】

別の態様では、この開示は、(a)細胞または組織から得られたゲノムDNA試料を、5'末端に第1の検出配列を含むトランスポゾンと、トランスポゾンがゲノムDNAに挿入され、かつゲノムDNAが断片化されて核酸断片の5'末端に付着されたトランスポゾンを含む複数のタグメンテーションされた二本鎖核酸断片になる条件下で接触させることであって、前記ゲノムDNAは、部位特異的ヌクレアーゼによって切断された少なくとも1つの部位に組み込まれた二本鎖オリゴヌクレオチドを含む、ことと；(b)(i)二本鎖オリゴヌクレオチドの少なくとも一部分に相補的なヌクレオチド配列を含み、かつその5'末端に第2の検出配列を含む第1の固定プライマーと、(ii)第1の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第2の選択的プライマーとを使用して、タグメンテーションされた核酸断片を増幅させて、第1の検出配列、核酸断片の5'末端に付着されたトランスポゾン、および第2の検出配列を含む、増幅された核酸断片を形成させることと；(c)増幅された核酸断片を配列決定することを含む方法の特徴とする。

20

30

【0017】

いくつかの実施態様では、部位特異的ヌクレアーゼは、Cas9である。

【0018】

別の態様では、この開示は、(a)細胞または組織から得られたゲノムDNA試料を、ヌクレアーゼによるゲノムDNAの切断、および二本鎖オリゴヌクレオチドの少なくとも1つの切断部位への組み込みに好都合な条件下で、RNAガイド型ヌクレアーゼおよび二本鎖オリゴヌクレオチドと接触させることとであって、それによって、組み込まれた二本鎖オリゴヌクレオチドを含むタグ付けされたゲノムDNAが産生される、ことと；(b)タグ付けされたゲノムDNAを、5'末端に第1の検出配列を含むトランスポゾンと、タグ付けされたゲノムDNAにトランスポゾンが挿入され、かつタグ付けされたゲノムDNAが断片化されて核酸断片の5'末端に付着されたトランスポゾンを含む複数のタグメンテーションされた二本鎖核酸断片になる条件下で接触させることと；(c)(i)二本鎖オリゴヌクレオチドの少なくとも一部分に相補的なヌクレオチド配列を含み、かつその5'末端に第2の検出配列を含む第1の固定プライマーと、(ii)第1の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第2の選択的プライマーとを使用して、タグメンテーションされた核酸断片を増幅させて、第1の検出配列、核酸断片の5'末端に付着されたトランスポゾン、および第2の検出配列を含む、増幅された核酸断片を形成させることと；(d)増幅された核酸断片を配列決定することを含む方法の特徴とする。

40

【0019】

50

いくつかの実施態様では、RNAガイド型ヌクレアーゼは、Cas9である。

【0020】

いくつかの実施態様では、二本鎖オリゴヌクレオチドのどちらの鎖も、ゲノムDNAに対してオルソログスである。

【0021】

いくつかの実施態様では、二本鎖オリゴヌクレオチドの各5'末端は、リン酸基を含む。いくつかの実施態様では、二本鎖オリゴヌクレオチドの各3'末端は、ホスホロチオエート結合を含む。いくつかの実施態様では、二本鎖オリゴヌクレオチドの各5'末端は、ホスホロチオエート結合を含む。いくつかの実施態様では、二本鎖オリゴヌクレオチドは、30~35個のヌクレオチドまたは60~65個のヌクレオチドを含む。いくつかの実施態様では、二本鎖オリゴヌクレオチドは、平滑末端化される。いくつかの実施態様では、二本鎖オリゴヌクレオチドは、1、2、3、または4塩基の突出ヌクレオチドを有する5'末端を含む。

10

【0022】

いくつかの実施態様では、第1の検出配列は、第1の配列決定タグを含む。いくつかの実施態様では、配列決定することは、増幅された核酸断片を、第1の配列決定タグとハイブリッド形成する第1の配列決定プライマーと接触させることを含む。いくつかの実施態様では、第2の検出配列は、第2の配列決定タグを含む。いくつかの実施態様では、配列決定することは、増幅された核酸断片を、第2の配列決定タグとハイブリッド形成する第2の配列決定プライマーと接触させることを含む。

20

【0023】

いくつかの実施態様では、二本鎖核酸断片は、約500~約5000bpを含む。

【0024】

いくつかの実施態様では、配列決定された核酸断片は、参照ゲノムDNA試料に対する転座を含む。いくつかの実施態様では、ゲノムDNA試料は、ウイルスによって導入された配列を含む。

【0025】

いくつかの実施態様では、方法は、ゲノムDNA、切断されたゲノムDNA、またはタグ付けされたゲノムDNAをトランスポゾンと接触させる前に、ウイルスベクターを使用して、配列をゲノムDNAに導入するステップをさらに含む。

30

【0026】

いくつかの実施態様では、方法は、配列決定された核酸断片を参照ゲノムDNA試料と比較するステップをさらに含む。

【0027】

いくつかの実施態様では、少なくとも1つの配列決定された核酸は、参照ゲノムDNA試料に対する転座を含む。いくつかの実施態様では、配列決定された核酸断片の1%未満は、参照ゲノムDNA試料に対する転座を含む。

【0028】

別の態様では、この開示は、本明細書に記載したタグメンテーションされた核酸断片のライブラリを特徴とする。別の態様では、この開示は、本明細書に記載したトランスポザーゼ、プライマー、および/または捕捉プローブを含む、標的核酸、例えばDNA、例えばゲノムDNAを配列決定するためのシステムを特徴とする。

40

【0029】

限定はされないが、ヌクレアーゼ切断を評価するためのコンピュータによって実行されるプロセスを含めた、本明細書に記載したプロセス、方法、システム、および技術の少なくとも一部は、1つまたは複数の非一過性の機械可読記憶媒体上に保存されるおよび1つまたは複数の処理装置上で遂行可能である命令を含む、コンピュータプログラム製品として実行することができる。非一過性の機械可読記憶媒体の例としては、例えば、読み出し専用メモリ、光ディスクドライブ、メモリディスクドライブ、ランダムアクセスメモリなどが挙げられる。本明細書に記載したプロセス、方法、システム、および技術の少なくとも

50

も一部は、定められた操作を実施するための、1つまたは複数の処理装置、および1つまたは複数の処理装置によって遂行可能である命令を保存するメモリを含む、装置、方法、またはシステムとして実行することができる。

【図面の簡単な説明】

【0030】

【図1】ヌクレアーゼによって改変されたゲノムDNAについての例示的な配列決定方法のステップを示す模式図を示す。

【図2】(図2A)固有分子インデックス(Unique Molecular Index)(UMI)バーコード、プーリング(pooling)バーコード、および5配列タグを伴う、例示的なカスタムな高活性なTn5酵素(Tn5*)を示す模式図を示す。(図2B)ヌクレアーゼによって改変されたゲノムDNAについての例示的な配列決定方法のステップを示す模式図を示す。

10

【図3】配列決定ライブラリを作成するための、二本鎖オリゴヌクレオチドインサート(dsODN)に相補的なプライマーを使用する、例示的な配列決定方法のステップを示す模式図を示す。

【図4-1】例示的な方法に使用されるプライマーを図示する。Illuminaテイルには下線を引いてある。

【図4-2】例示的な方法に使用されるプライマーを図示する。Illuminaテイルには下線を引いてある。

【図5】UDiTASを使用する、G1およびG2標的についてのMiSeqリードの結果を図示する。

20

【図6】UDiTASにおける、G1固定プライマーおよびG2固定プライマーを使用する、G1およびG2遺伝子のリード被覆プロットを図示する。

【図7】UDiTASを使用するインデルの検出の結果を図示する。

【図8】デュアルガイド編集による可能な染色体再構成の模式図である。

【図9】安定なHEK293TクロンのゲノムDNA混合物を使用する、染色体再構成の検出の検証を図示する。

【図10】ゲノムDNA混合物を使用する、染色体再構成の検出の追加の検証を図示する。

【図11】この開示の例示的な方法の模式図である。

30

【図12】この開示のある種の方法に関する例示的なバイオインフォマティクス・パイプラインの模式図である。

【図13】小さいインデル、大きい欠失(>100bpを超えるもの)、および逆位を含めた遺伝子編集事象の検出を示す模式図を図示する。

【図14AB】(図14A)約1.1kb離れた2つのSaCas9ガイドを用いるGene3 IVS遺伝子座でのU-2 OS細胞の編集後の、例示的な編集の型を図示する。(図14B)改変されていない親系統と様々な割合で混合されたHEK293細胞株の、例示的な線形性および検出下限(Lower Limit of Detection)(LLoD)を図示する。

【図14C】1つはGene1を、もう1つはGene2を標的にする2つのSpCas9 RNPを用いるヌクレオフェクション(nucleofection)後に測定された、T細胞における例示的な染色体間転座を図示する。この模式図および表は、すべての結果の可能性、および適用できる場合には測定された結果を示す。縦線は、編集部位を示し、矢印は、プライマーを示す(1=OLI6259および2=OLI6256)。

40

【図15A】Integrated Genome Viewer(IGV)(Thorvaldsdottir et al. Brief. Bioinform. 2013;14:178-192; Robinson et al. Nat. Biotechnol. 2011;29:24-26)において示される、U-2 OS編集研究における例示的な編集事象を図示する。各図の最上部の模式図は、観察された編集事象を示す。赤色/青色に色付けしたリードを、上部/下部のゲノム参照DNA配列に対して整列させた。図15Aは

50

、ガイド2の部位の小さいインデルを図示する。図15Bは、ガイド1および2を使用する大きい1.1kbの欠失接合部を図示する。図15Cは、ガイド1および2を使用する大きい1.1kbの逆位接合部を図示する。図15Dは、ガイド2を使用する相同の接合部を図示する。

【図15B】Integrated Genome Viewer (IGV) (Thorvaldsdottir et al. Brief. Bioinform. 2013; 14: 178-192; Robinson et al. Nat. Biotechnol. 2011; 29: 24-26)において示される、U-2 OS編集研究における例示的な編集事象を図示する。各図の最上部の模式図は、観察された編集事象を示す。赤色/青色に色付けしたリードを、上部/下部のゲノム参照DNA配列に対して整列させた。図15Aは、ガイド2の部位の小さいインデルを図示する。図15Bは、ガイド1および2を使用する大きい1.1kbの欠失接合部を図示する。図15Cは、ガイド1および2を使用する大きい1.1kbの逆位接合部を図示する。図15Dは、ガイド2を使用する相同の接合部を図示する。

10

【図15C】Integrated Genome Viewer (IGV) (Thorvaldsdottir et al. Brief. Bioinform. 2013; 14: 178-192; Robinson et al. Nat. Biotechnol. 2011; 29: 24-26)において示される、U-2 OS編集研究における例示的な編集事象を図示する。各図の最上部の模式図は、観察された編集事象を示す。赤色/青色に色付けしたリードを、上部/下部のゲノム参照DNA配列に対して整列させた。図15Aは、ガイド2の部位の小さいインデルを図示する。図15Bは、ガイド1および2を使用する大きい1.1kbの欠失接合部を図示する。図15Cは、ガイド1および2を使用する大きい1.1kbの逆位接合部を図示する。図15Dは、ガイド2を使用する相同の接合部を図示する。

20

【図15D】Integrated Genome Viewer (IGV) (Thorvaldsdottir et al. Brief. Bioinform. 2013; 14: 178-192; Robinson et al. Nat. Biotechnol. 2011; 29: 24-26)において示される、U-2 OS編集研究における例示的な編集事象を図示する。各図の最上部の模式図は、観察された編集事象を示す。赤色/青色に色付けしたリードを、上部/下部のゲノム参照DNA配列に対して整列させた。図15Aは、ガイド2の部位の小さいインデルを図示する。図15Bは、ガイド1および2を使用する大きい1.1kbの欠失接合部を図示する。図15Cは、ガイド1および2を使用する大きい1.1kbの逆位接合部を図示する。図15Dは、ガイド2を使用する相同の接合部を図示する。

30

【図16】UDiTASに適用される例示的な二項の検出力計算を図示する。所与の数の予想された観察(1、2、または3)についての、固有分子識別子(例えば:試行)の数に対して編集頻度(例えば:成功の確率)をプロットする、シミュレーションされた二項分布。左のグラフは、95%の信頼度を示し、右のグラフは、99%の信頼度を示す。

【図17】UDiTASに対する例示的なゲノムマッピング率を図示する。Gene3ガイド2に対する遺伝子特異的プライマー(OL16062)を使用して分析した10本の別個の試料を、x軸上にプロットした。y軸は、各試料についての予想された参照アンプリコンに対してマッピングされたリードの割合を示す。

40

【図18】Gene3構造多型(図18A)またはGene1/Gene2均衡型転座(図18B)を含有するプラスミド標準を使用する、例示的なUDiTAS特徴付け、およびAMP-Seqとの比較を図示する。

【図19】担体DNAを伴わないプラスミド標準の例示的なUDiTAS特徴付けを図示する。

【発明を実施するための形態】

【0031】

定義

50

この明細書全体を通して、以下の段落で定義されるいくつかの用語が用いられる。この明細書の本文内では、他の定義も参照することができる。本出願では、文脈からそうではないと明らかなでない限り、(i) 用語「 a 」は、「少なくとも 1 つ」を意味すると理解することができ；(i i) 用語「または」は、「および/または」を意味すると理解することができ；(i i i) 用語「～を含むこと」および「～を含めて」は、それらだけで示されるか、1 つまたは複数の追加の構成要素またはステップと共に示されるかにかかわらず、列挙された構成要素またはステップを包含すると理解することができ；(i v) 用語「約」および「およそ」は、当業者によって理解されるであろう標準偏差を許容すると理解することができ；(v) 範囲が提供される場合、両端が含まれる。

【 0 0 3 2 】

本明細書で使用する場合、数に関する用語「約」および「およそ」は、本明細書では、そうではないと記述がないまたは文脈からそうではないと明らかなでない限り、その数のいずれかの（それよりも大きいまたは小さい）方向の 2 0 %、1 0 %、5 %、または 1 % の範囲内にある数を含めるために使用される（こうした数が可能な値の 1 0 0 % を超えるであろう場合を除く）。

【 0 0 3 3 】

「バーコード」は、本明細書で使用する場合、オリゴヌクレオチド配列を指す。いくつかの実施態様では、ランダムなバーコードは、いかなる特定の配列とも関係がなく、例えば、品質管理目的で使用する事ができる。例えば、ランダムなバーコードを含む連結産物を分析する際に、ある特定の連結産物の増幅バイアスを評価するために、ランダムなバーコードを使用することができる。増幅産物間の所与のランダムなバーコードの過剰なまたは低い提示によって、増幅バイアスを示すことができる。いくつかの実施態様では、こうした偏りのある増幅産物に伴うデータは除外される。ランダムなバーコードの好適なサイズは、実施態様に応じて変動し得る。非限定的な例として、いくつかの実施態様では、ランダムなバーコードは、長さが 2、3、4、5、6、7、8、9、10、11、または 12 ヌクレオチドである。

【 0 0 3 4 】

いくつかの実施態様では、バーコードは、ハミングコード、すなわち、誤り訂正バーコードである。ハミングコードは、例えば、試料が多重化されている場合にある特定の試料を特定するために、使用することができる配列である。いくつかの実施態様では、定義された数の可能なハミングコード、例えば、非限定的な例として、最大 2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、または 20 個の可能なハミングコードなどが、集合的に存在する。

【 0 0 3 5 】

いくつかの実施態様では、バーコードは、試料バーコードである。いくつかの実施態様では、試料バーコードは、ある特定の試料に対して固有であり、その結果、その特定の試料由来の配列決定ライブラリのすべてのメンバーは、同じバーコードを含む。いくつかの実施態様では、試料バーコードは、例えば、複数の試料と一緒に分析される場合に、配列解析の結果を特定および/または分類するために、使用される。いくつかの実施態様では、試料バーコードは、プーリングバーコードと呼ばれる。

【 0 0 3 6 】

「切断」は、本明細書で使用する場合、DNA 分子の共有結合骨格の破壊を指す。切断は、限定はされないが、ホスホジエステル結合の酵素的または化学的加水分解を含めた様々な方法によって開始することができる。一本鎖切断と二本鎖切断のどちらも可能であり、二本鎖切断は、2 つの別個の一本鎖切断事象の結果として起こる可能性がある。DNA 切断は、平滑末端または付着末端のいずれかの産生をもたらす可能性がある。

【 0 0 3 7 】

本明細書で使用する場合、用語「縮重した」は、オリゴヌクレオチドまたはヌクレオチド配列を指すために使用される場合、複数の異なる塩基のいずれかを含有することができ 1 つまたは複数の位置を指す。オリゴヌクレオチドまたはヌクレオチド配列内の縮重し

10

20

30

40

50

た残基は、以下に示した通りの標準の I U P A C 核酸表記によって表される：

【 0 0 3 8 】

【 表 A 】

文字	縮重した塩基
K	GまたはT/U
M	AまたはC
R	AまたはG
Y	CまたはT/U
S	CまたはG
W	AまたはT/U
B	C、G、またはT/U
V	A、C、またはG
H	A、C、またはT/U
D	A、G、またはT/U
N	A、C、G、またはT/U

10

【 0 0 3 9 】

そうではないと指定されない限り、縮重した残基は、可能な塩基のランダムなまたは均等な分布を意味するわけではなく、例えば、「N」残基は、A、C、G、および/または T / U 残基の均等な分布を表すわけではない。

【 0 0 4 0 】

本明細書で使用する場合、核酸分子またはその断片を「検出すること」という用語は、典型的には、核酸分子またはその断片が、試料または組成物の他の構成要素から完全にまたは部分的に分離されている場合に、核酸分子の存在を決定することを指し、核酸分子またはその断片の電荷質量比、質量、量、吸光度、蛍光、または他の特性を決定することも含めることができる。

30

【 0 0 4 1 】

本明細書で使用する場合、用語「検出可能な標識」は、検出可能であるあらゆる元素、分子、官能基、化合物、断片、または部分を指す。いくつかの実施態様では、検出可能な実体は、単独で提供されるまたは利用される。いくつかの実施態様では、検出可能な実体は、別の薬剤に伴って（例えば、連結させて）提供されるおよび/または利用される。検出可能な実体の例としては、限定はされないが：様々なリガンド、放射性核種（例えば、 ^3H 、 ^{14}C 、 ^{18}F 、 ^{19}F 、 ^{32}P 、 ^{35}S 、 ^{135}I 、 ^{125}I 、 ^{123}I 、 ^{64}Cu 、 ^{187}Re 、 ^{111}In 、 ^{90}Y 、 $^{99\text{m}}\text{Tc}$ 、 ^{177}Lu 、 ^{89}Zr など）、蛍光染料（具体的な例示的な蛍光染料については以下を参照）、化学発光剤（例えば、アクリジニウムエステル、安定化ジオキセタンなど）、生物発光剤、スペクトルの分解可能な無機蛍光半導体ナノ結晶（すなわち、量子ドット）、金属ナノ粒子（例えば、金、銀、銅、白金など）のナノクラスター、常磁性金属イオン、酵素（酵素の具体例については以下を参照）、比色定量標識（例えば、染料、コロイド金など）、ピオチン、ジオキシゲニン（*dioxigenin*）、ハプテン、および抗血清またはモノクローナル抗体が利用可能なタンパク質が挙げられる。

40

【 0 0 4 2 】

本明細書で使用する場合、用語「連結」、「連結すること」、およびその文法的等価物は、典型的には鋳型によって推進される反応において、2つ以上の核酸、例えばオリゴヌクレオチドおよび/またはポリヌクレオチドの末端間の共有結合またはつながりを形成することを指す。結合またはつながりの性質は、大きく変動する可能性があり、また、連結

50

は、酵素的または化学的に行われる可能性がある。本明細書で使用する場合、連結は通常、酵素的に行われて、あるオリゴヌクレオチドの5'炭素末端ヌクレオチドと別のヌクレオチドの3'炭素との間のホスホジエステル結合が形成される。用語「連結」はまた、ホスホジエステル結合の非酵素的形成、ならびにオリゴヌクレオチドの末端間の非ホスホジエステル共有結合、例えばホスホロチオエート結合、ジスルフィド結合などの形成も包含する。

【0043】

本明細書で使用する場合、用語「ヌクレアーゼ」は、核酸のヌクレオチドサブユニット間のホスホジエステル結合を切断することが可能なポリペプチドを指し；用語「エンドヌクレアーゼ」は、ポリヌクレオチド鎖内のホスホジエステル結合を切断することが可能なポリペプチドを指す。

【0044】

本明細書で使用する場合、用語「核酸」、「核酸分子」、または「ポリヌクレオチド」は、本明細書では互換的に使用される。これらは、一本鎖または二本鎖形態のデオキシリボヌクレオチドまたはリボヌクレオチドのポリマーを指し、そうではないと記述がない限り、天然起源のヌクレオチドと類似の方式で機能することができる天然のヌクレオチドの公知の類似体を包含する。この用語は、合成の骨格を有する核酸様の構造、ならびに増幅産物を包含する。DNAとRNAは、どちらもポリヌクレオチドである。このポリマーには、天然のヌクレオチド（すなわち、アデノシン、チミジン、グアノシン、シチジン、ウリジン、デオキシアデノシン、デオキシチミジン、デオキシグアノシン、およびデオキシシチジン）、ヌクレオチド類似体（例えば、2-アミノアデノシン、2-チオチミジン、イノシン、ピロロ-ピリミジン、3-メチルアデノシン、C5-プロピニルシチジン、C5-プロピニルウリジン、C5-プロモウリジン、C5-フルオロウリジン、C5-ヨードウリジン、C5-メチルシチジン、7-デアザアデノシン、7-デアザグアノシン、8-オキソアデノシン、8-オキソグアノシン、O(6)-メチルグアニン、および2-チオシチジン）、化学的に修飾された塩基、生物学的に修飾された塩基（例えば、メチル化された塩基）、挿入を受けた（*intercalated*）塩基、修飾された糖（例えば、2'-フルオロリボース、リボース、2'-デオキシリボース、アラビノース、およびヘキソース）、または修飾されたリン酸基（例えば、ホスホロチオエートおよび5'-N-ホスホロアミダイト結合）が含まれ得る。

【0045】

本明細書で使用する場合、用語「オリゴヌクレオチド」は、ヌクレオチドまたはその類似体の鎖を指す。オリゴヌクレオチドは、例えば、化学合成、制限酵素消化、またはPCRを含めた、いくつかの方法によって得ることができる。当業者によって認識される通り、オリゴヌクレオチドの長さ（すなわち、ヌクレオチドの数）は、しばしばオリゴヌクレオチドの意図される機能または使用に応じて、大きく変動し得る。一般に、オリゴヌクレオチドは、約5～約300ヌクレオチド、例えば、約15～約200ヌクレオチド、約15～約100ヌクレオチド、または約15～約50ヌクレオチドを含む。この明細書全体を通して、オリゴヌクレオチドが、（4つの塩基の文字：A、C、G、およびT（これらはそれぞれ、アデノシン、シチジン、グアノシン、およびチミジンを表す）から選択される）文字の連続によって表される時はいつでも、ヌクレオチドは、左から右に、5'～3'の順序で示される。ある種の実施態様では、オリゴヌクレオチドの配列には、本明細書に記載した1つまたは複数の縮重した残基が含まれる。

【0046】

本明細書で使用する場合、用語「参照」は、比較が実施されるものに対する標準または対照を表現する。例えば、いくつかの実施態様では、対象となる薬剤、動物、個人、集団、試料、配列、または値が、参照または対照の薬剤、動物、個人、集団、試料、配列、または値と比較される。いくつかの実施態様では、参照または対照は、対象となる試験または決定と実質的に同時に試験および/または決定される。いくつかの実施態様では、参照または対照は、任意選択で有形の媒体に組み込まれた歴史的な参照または対照である。典型的には、当業者によって理解されるであろう通り、参照または対照は、評価を受けるも

10

20

30

40

50

のに匹敵する条件または状況下で決定されるまたは特徴付けられる。当業者は、ある特定の可能な参照または対照に対する信頼性および/または比較を正当化するための、十分な類似性が存在する時を認識するであろう。

【0047】

本明細書で使用する場合、用語「標的部位」は、結合が存在するのに十分な条件ならば、結合分子が結合することとなる核酸の部分定義する核酸配列を指す。いくつかの実施態様では、標的部位は、本明細書に記載したヌクレアーゼが結合するおよび/またはこうしたヌクレアーゼによって切断される核酸配列である。いくつかの実施態様では、標的部位は、本明細書に記載したガイドRNAが結合する核酸配列である。標的部位は、一本鎖または二本鎖であり得る。二量体化するヌクレアーゼ、例えばFokI DNA切断ドメインを含むヌクレアーゼという状況では、標的部位は、典型的には、左半分部位（ヌクレアーゼの1つのモノマーが結合する）、右半分部位（ヌクレアーゼの第2のモノマーが結合する）、および切断が行われる半分の部位の間のスペーサー配列を含む。いくつかの実施態様では、左半分部位および/または右半分部位は、10～18ヌクレオチド長である。いくつかの実施態様では、半分の部位のいずれかまたは両方は、より短いまたはより長い。いくつかの実施態様では、左半分部位と右半分部位は、異なる核酸配列を含む。ジンクフィンガーヌクレアーゼという状況では、標的部位は、いくつかの実施態様では、4～8bpの長さである特定されないスペーサー領域に隣接する、それぞれ6～18bpの長さである2つの半分の部位を含むことができる。TALENという状況では、標的部位は、いくつかの実施態様では、10～30bpの長さである特定されないスペーサー領域に隣接する、それぞれ10～23bpの長さである2つの半分の部位を含むことができる。RNAガイド型（例えば、RNA-プログラム可能）ヌクレアーゼという状況では、標的部位は、典型的には、RNA-プログラム可能ヌクレアーゼのガイドRNAに相補的なヌクレオチド配列、およびガイドRNA相補的配列に隣接する3'末端または5'末端のプロトスペーサー隣接モチーフ（PAM）を含む。RNAガイド型ヌクレアーゼCas9については、標的部位は、いくつかの実施態様では、16～24塩基対+3～6塩基対（PAM）（例えばNNN（ここでは、Nは、あらゆるヌクレオチドに相当する））であり得る。Cas9などのRNAガイド型ヌクレアーゼについての例示的な標的部位は、当業者に公知であり、限定はされないが、NNG、NGN、NAG、およびNGG（ここでは、Nは、あらゆるヌクレオチドに相当する）が含まれる。さらに、異なる種（例えば、化膿レンサ球菌（*S. pyogenes*）の代わりにストレプトコッカス・サーモフィルス（*S. thermophilus*））由来のCas9ヌクレアーゼは、配列NGGNGを含むPAMを認識する。限定はされないがNNAGAAWおよびNAARを含めた、追加のPAM配列が公知である（例えば、その内容全体が参照により本明細書に組み込まれる、Esvelt and Wang, Molecular Systems Biology, 9:641(2013)を参照のこと）。例えば、例えばCas9などのRNAガイド型ヌクレアーゼの標的部位は、構造[Nz]-[PAM]（ここでは、各Nは、独立に、あらゆるヌクレオチドであり、zは、1～50の整数である）を含むことができる。いくつかの実施態様では、zは、少なくとも2、少なくとも3、少なくとも4、少なくとも5、少なくとも6、少なくとも7、少なくとも8、少なくとも9、少なくとも10、少なくとも11、少なくとも12、少なくとも13、少なくとも14、少なくとも15、少なくとも16、少なくとも17、少なくとも18、少なくとも19、少なくとも20、少なくとも25、少なくとも30、少なくとも35、少なくとも40、少なくとも45、または少なくとも50である。いくつかの実施態様では、zは、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、または50である。いくつかの実施態様では、Zは、20である。

【0048】

本明細書で使用する場合、「タグ」は、それが連結された核酸断片を突き止める手段を

提供する、非標的核酸構成要素、一般にDNAを指す。例えば、いくつかの実施態様では、タグは、（例えば、DNAポリメラーゼによる伸長のためのプライマーなどのオリゴヌクレオチド、または捕捉のためのもしくは連結反応のためのオリゴヌクレオチドをアニーリングさせるための部位を提供することによって）タグが付着されたDNAの同定、認識、および/または分子的または生化学的操作を可能にするヌクレオチド配列であるまたは当該ヌクレオチド配列を含む。タグをDNA分子に連結させるためのプロセスは、時として、本明細書では「タギング」と呼ばれ、タギングを受けるまたはタグを含有するDNAは、「タグ付けされた」（例えば、「タグ付けされたDNA」）と呼ばれる。

【0049】

本明細書で使用する場合、「トランスポザーゼ」は、トランスポゾン含有組成物（例えばトランスポゾン）との機能複合体を形成し、インビトロ転移反応においてそれと共にインキュベートされる二本鎖標的核酸、例えば、DNAへのトランスポゾンの挿入または転移を触媒することが可能な酵素を意味する。

10

【0050】

本明細書で使用する場合、「トランスポゾン」は、インビトロ転移反応において機能性であるトランスポザーゼまたはインテグラーゼ酵素との複合体を形成するのに必要であるヌクレオチド配列（「トランスポゾン配列」）を含む二本鎖核酸、例えばDNAを意味する。トランスポゾンは、トランスポゾンを認識し結合するトランスポザーゼまたはインテグラーゼと共に、「複合体」または「シナプス複合体」または「トランスポソーム複合体」または「トランスポソーム組成物」を形成し、その複合体は、トランスポゾンを、インビトロ転移反応においてそれと共にインキュベートされる標的DNAに挿入または転移させることが可能である。トランスポゾンは、「転移トランスポゾン配列」または「転移鎖」および「非転移トランスポゾン配列」または「非転移鎖」からなる2つの相補的配列を呈する。転移鎖の3'末端は、インビトロ転移反応において、標的DNAに連結または転移される。転移トランスポゾン配列に相補的なトランスポゾン配列を呈する非転移鎖は、インビトロ転移反応において、標的DNAに連結も転移もされない。いくつかの実施態様では、トランスポゾンは、検出配列を含む。いくつかの実施態様では、検出配列は、追加の配列、例えば、UMIバーコード、プーリングもしくは試料バーコード、および/または配列タグを含む。

20

【0051】

本明細書で使用する場合、「トランスポゾン連結末端」は、挿入部位で標的DNAと連結する二本鎖トランスポゾンDNAの末端を意味する。トランスポゾン連結末端は、標的DNAと実際に連結する鎖を指さない。例えば、トランスポゾンMuA連結末端の3'末端は、標的DNAの3'末端とトランスポゾン連結末端の5'末端との間のギャップを残して、標的DNAの5'末端と連結する。

30

【0052】

本明細書で使用する場合、用語「トランスポザーゼ」は、レトロトランスポゾンおよびレトロウイルス由来のインテグラーゼを含めた、転移反応に必要とされる1つまたは複数のトランスポゾンとの機能複合体を形成することが可能な酵素を意味する。

【0053】

本明細書で使用する場合、「転移反応」は、トランスポゾンが標的DNAのランダムなまたはほぼランダムな部位に挿入する反応である。転移反応における不可欠な構成要素は、トランスポゾンおよびトランスポザーゼもしくはインテグラーゼ酵素、または機能性の転移複合体を形成するために必要とされる何らかの他の構成要素である。ランダムなまたはほぼランダムな方式で増幅プライマー部位を挿入することが可能なすべての転移システムが、本明細書に記載した方法において有用である（Craig, Science 270:253-254, 1995）。こうしたシステムの例は、Ty1（Devine and Boeke, Nucleic Acids Research, 1994, 22(18):3765-3772、および国際公開第95/23875号パンフレット）、トランスポゾンTn7（Craig, Curr. Top. Microbiol. Immuno

40

50

1. 204:27-48, 1996)、Tn₁₀およびIS10 (Kleckner et al. Curr. Top. Microbiol. Immunol., 1996, 204: 49-82)、マリナートランスポザーゼ (Lampe et al., EMBO J., 1996, 15(19): 5470-5479)、Tc1 (Vos et al., Genes Dev., 1996, 10(6): 755-61)、Tn5 (Park et al., J. Korean Soc. Microbiol. 27: 381-389 (1992))、P因子 (Kaufman and Rio, Cell, 1992, 69(1): 27-39) および Tn3 (Ichikawa and Ohtsubo, J. Biol. Chem., 1990, 265(31): 18829-32)、細菌性挿入配列 (Ohtsubo and Sekine, Curr. Top. Microbiol. Immunol., 1996, 204: 1-26)、レトロウイルス (Varmus Retroviruses in Mobile DNA. Berg D.E. and Howe M.M. eds. American society for microbiology, Washington D.C. pp. 53-108 (1989))、および酵母のレトロトランスポゾン (Boeke, Berg D.E. and Howe M.M. eds. American society for microbiology, Washington D.C. pp. 335-374 (1989)) である。

【0054】

本明細書で使用する場合、用語「配列決定タグ」は、そのタグが連結される核酸の配列決定を容易にする目的のための（例えば、合成による配列決定のためのプライミング部位を提供するための、または連結による配列決定のためのアニーリング部位を提供するための、またはハイブリダイゼーションによる配列決定のためのアニーリング部位を提供するための）配列であるまたは当該配列を含むタグを意味する。例えば、いくつかの実施態様では、配列決定タグは、DNA断片またはDNA断片の相補体のDNA合成を開始するための部位を提供する。いくつかの実施態様では、配列決定タグは、全長のIllumina フォワードアダプター (i5) である。いくつかの実施態様では、配列決定タグは、全長のIllumina リバースアダプター (i7) である。

【0055】

本明細書で使用する場合、用語「増幅タグ」は、前記タグが付加される核酸の増幅を容易にする目的のための配列であるまたは当該配列を含むタグを意味する。例えば、いくつかの実施態様では、増幅タグは、DNAポリメラーゼを使用する核酸増幅反応（例えば、PCR増幅反応もしくは鎖置換増幅反応、またはローリングサークル増幅反応）のためのプライミング部位、または核酸増幅反応（例えば、ライゲーション連鎖反応 (ligation chain reaction)）における鋳型依存性リガーゼを使用するプローブの連結のための連結鋳型を提供する。

【0056】

本明細書で使用する場合、核酸または核酸反応に関して使用される場合の用語「増幅する」、「増幅された」、または「増幅すること」は、標的核酸、または例えば本明細書に記載した方法によって生成されたタグ付けされた核酸などの、ある特定の核酸のコピーを作製するインビトロ方法を指す。核酸を増幅する多数の方法が、当技術分野で公知であり、増幅反応には、ポリメラーゼ連鎖反応、リガーゼ連鎖反応、鎖置換増幅反応、ローリングサークル増幅反応、NASBAなどの転写増幅法（例えば、米国特許第5,409,818号明細書）、ループ介在増幅法（例えば、ループ形成配列を使用する「LAMP」増幅、例えば、米国特許第6,410,278号明細書に記載されている通り）が含まれる。増幅される核酸は、改変されたDNAおよび/またはRNAを含めた、DNAまたはRNAまたはDNAとRNAの混合物を含む、またはこれらからなる、またはこれらに由来するDNAであり得る。1つまたは複数の核酸分子の増幅から得られる産物（すなわち「増幅産物」）は、出発核酸が、DNA、RNA、またはその両方であるかにかかわらず、DNAまたはRNA、またはDNAとRNA（ヌクレオシドまたはヌクレオチド）の両方の混合物であり得る。または、これらは、改変されたDNAまたはRNA（ヌクレオシド

またはヌクレオチド)を含むことができる。「コピー」は、必ずしも、標的配列に対する完全な配列相補性または同一性を意味しない。例えば、コピーには、ヌクレオチド類似体、例えばデオキシイノシンまたはデオキシウリジン、意図的な配列変更(標的配列に対してハイブリッド形成可能であるが相補的ではない配列を含むプライマーを介して導入された配列変更など)、および/または増幅中に起こる配列の誤りが含まれる可能性がある。

【0057】

本明細書で使用する場合、「プライマー」は、核酸ポリメラーゼによって伸長することができる、一般に遊離の3'-OH基を有するオリゴヌクレオチドである。鋳型依存性ポリメラーゼについては、一般に、プライマーの少なくとも3'部分は、鋳型核酸の一部分に相補的であり、その鋳型核酸の一部分に、オリゴヌクレオチドが、水素結合および他の分子力によって「結合」(または「複合体形成」、「アニーリング」、もしくは「ハイブリッド形成」)して、DNAポリメラーゼによる合成の開始のためのプライマー/鋳型複合体が与えられ、これは、DNA合成のプロセスにおいて、鋳型に相補的であるその3'末端に連結される共有結合された塩基の付加によって伸長(すなわち「プライマー伸長」)される。その結果は、プライマー伸長生成物である。鋳型依存性DNAポリメラーゼ(逆転写酵素が含まれる)は、一般に、DNA合成(「プライミング」)を開始するために、オリゴヌクレオチドプライマーの一本鎖鋳型に対する複合体形成を必要とするが、RNAポリメラーゼは一般に、DNA鋳型に相補的であるRNAの合成(転写)のためのプライマーを必要としない。

【0058】

「固定プライマー」は、本明細書で使用する場合、標的または鋳型核酸、例えばDNA内の公知の配列に相補的な配列を有するプライマーを意味する。

【0059】

「選択的プライマー」は、本明細書で使用する場合、トランスポゾンDNAの配列に相補的な(連結部の5'末端を有する鎖に相補的な)配列を有するプライマーを意味する。選択的プライマーは、標的または鋳型核酸、例えばDNAに対して5'→3'方向である。

【0060】

詳細な説明

本開示は、特に、トランスポゾン組成物を使用してDNAを断片化およびタグ付けすることによるヌクレアーゼ活性の特徴付けのための方法およびシステムを提供する。本明細書に記載した方法、組成物、およびキットは、本明細書に記載したヌクレアーゼでの切断を受けている標的または鋳型核酸(例えばDNA、例えばゲノムDNA)から核酸断片を産生および増幅するのに有用である。いくつかの実施態様では、こうした増幅された核酸断片は、例えばヌクレアーゼ活性を特徴付けるために、配列決定することができる。

【0061】

本明細書に記載した方法、組成物、およびキットは、標的または鋳型核酸内の大きい挿入、大きい欠失、逆位、および/または転座を評価するのに特に有用である。例えば、開示した方法、組成物、および/またはキットは、試料(例えばゲノムDNA試料)の5%未満、1%未満、0.5%未満、0.3%未満、0.1%未満、0.05%未満、および/または0.01%未満の存在量で存在する、染色体再構成(例えば、大きい挿入、大きい欠失、逆位、および/または転座)を検出することができる。この開示の方法は、例えば、アンカードマルチプレックスPCR配列決定(AMP-seq)(例えば、Zheng et al., Nat Med. 2014 Dec; 20(12): 1479-84; Adey et al., Genome Biology, 11(12), R119(2010)を参照のこと)を含めた公知の方法よりも好都合である。

【0062】

いくつかの実施態様では、この開示は、一方向性の標的配列決定(Uni-Directional Targeted Sequencing)(「UDiTAS」)を提供する。本明細書に記載したある種の実施態様では、こうした方法は、単一プライマーターゲティング手法、例えば、トランスポザーゼと複合体化された単一のアダプター、アダプタ

10

20

30

40

50

ー上の固有分子インデックス (UMI) (例えば、UMI バーコード) および / またはブーリングもしくは試料バーコード、タグメンテーション中の UMI を直接的に用いた DNA の標識、および PCR タグメンテーションされた産物を利用する。いくつかの実施態様では、第 2 ラウンドの本発明者らの PCR は、アダプター (例えば、リバース Illumina アダプター、例えば i7) および追加の試料バーコードを追加する。ある種の実施態様では、標的部位を通してゲノム DNA 断片の末端までの単一プライマー伸長が、転座事象を特定することができる。こうした方法はまた、本明細書に記載した通りの多重化にも適合できる。

【0063】

本明細書で提供した戦略、方法、および試薬は、あらゆる部位特異的ヌクレアーゼ、例えば、ジンクフィンガーヌクレアーゼ (ZFN)、転写活性化因子様エフェクターヌクレアーゼ (TALEN)、ホーミングエンドヌクレアーゼ、有機化合物ヌクレアーゼ、およびエンジン抗生物質 (例えば、ダイネミシン (dynemycin)、ネオカルジノスタチン、カリチアマイシン、エスペラミシン、プレオマイシン) に対する切断および / または修復事象を分析するために利用することができる。本明細書に記載したものに加えて、好適なヌクレアーゼは、この開示に基づいて、当業者には明らかであろう。

【0064】

さらに、本明細書で提供した戦略、方法、および試薬は、当業者が、特異性が高められたヌクレアーゼを特定、設計、および / または選択すること、および / またはあらゆる所与のヌクレアーゼ (例えば、部位特異的ヌクレアーゼ、例えば ZFN、および TALEN、ならびに RNA - プログラム可能ヌクレアーゼ、例えば Cas9) のオフターゲット効果を最小限にすることを可能にする。特に DNA および DNA 切断ヌクレアーゼに関しては、本明細書で提供した発明の概念、方法、戦略、および試薬は、この点では制限されないが、あらゆる核酸：ヌクレアーゼ対に適用することができる。

【0065】

一方向性の標的配列決定 (Uni - Directional Targeted Sequencing) (UDiTAS)

いくつかの態様では、本開示は、ヌクレアーゼおよび / またはヌクレアーゼバリエーションを、ある特定の標的部位を切断する能力について評価する方法を提供する。いくつかの実施態様では、UDiTAS は、可変長の PCR 増幅に伴うバイアスを除去し、単一の反応で、小さな挿入および欠失事象 (インデル) に加えて、構造的変化を測定する。いくつかの実施態様では、UDiTAS 方法は、全長の Illumina フォワードアダプター (i5)、試料バーコード、および / または固有分子識別子 (UMI) を含むカスタム設計された Tn5 トランスポゾンを使用する (図 2A)。一般に、こうした方法は、核酸鋳型 (例えば DNA、例えばゲノム DNA、例えば、本明細書に記載したヌクレアーゼを使用して切断、改変、および / または編集されているゲノム DNA) を、トランスポゾンと、トランスポゾンが核酸鋳型に挿入される条件下で接触させることを含む。

【0066】

この条件は、例えば本明細書に記載したトランスポザラーゼの存在下で、転移反応が起こるのに十分である。こうした転移反応および条件は、当技術分野で公知である (例えば、米国特許第 6,593,113 号明細書および米国特許第 9,080,211 号明細書を参照のこと)。いくつかの実施態様では、転移条件は、所望される断片サイズを念頭において選択される。転移反応は、核酸鋳型の、複数のタグメンテーションされた二本鎖核酸断片への断片化をもたらし、ここでは、トランスポゾンの転移鎖の 3' 末端は、核酸断片の 5' 末端に付着される。トランスポゾンの転移鎖は、5' 末端に第 1 の検出配列を含む。いくつかの実施態様では、第 1 の検出配列は、UMI バーコード、ブーリングバーコード、および / またはアダプター (全長の Illumina フォワードアダプター (i5)) を含む。

【0067】

転移反応の後に、タグメンテーションされた核酸断片を、例えば、プライマーのセット

10

20

30

40

50

を使用するPCRを使用して、増幅させる。第1のプライマーは、ゲノムDNA内のあらかじめ決定された位置に相補的なヌクレオチド配列を含む、固定プライマーであり得る。第1のプライマーはまた、本明細書に記載した通りの二本鎖オリゴヌクレオチドの少なくとも一部分に相補的なヌクレオチド配列を含む、固定プライマーであり得る。第1のプライマーはまた、その5'末端に第2の検出配列を含む。第2のプライマーは、第1の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む、選択的プライマーである。増幅は、(5'から3'方向に)第1の検出配列、核酸断片の5'末端に付着されたトランスポゾンの転移鎖、および第2の検出配列を含む、増幅された核酸断片を形成する。

【0068】

いくつかの実施態様では、第1の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第2の選択的プライマーと、第2の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第3の選択的プライマーを使用して、第2ラウンドのPCRが実施される。いくつかの実施態様では、第3の選択的プライマーは、バーコード(例えば、プーリングバーコード)および/または配列タグ(例えば、リバースILLUMINAアダプター(i7))を含む。例示的な方法を、図1および2A~2Cに示す。

【0069】

次いで、増幅された核酸断片を、配列決定することができる。例えば、第1および第2の検出配列は、配列決定を容易にするための、本明細書に記載した配列決定タグを含むことができる。いくつかの実施態様では、この方法は、タグメンテーション後かつPCR前に、サイズ分離ステップを含むことができる。

【0070】

いくつかの実施態様では、ゲノムDNA試料中の特定のアレル(例えば、1つまたは複数のインデルを有する、および/または1つまたは複数の染色体再構成を有する野生型)の相対存在量は、増幅された核酸断片中のアレルの相対存在量と同じである。

【0071】

いくつかの実施態様では、複数のタグメンテーションされた二本鎖核酸断片は、こうした断片のライブラリであり得る。いくつかの実施態様では、ライブラリは、約 10^2 、 10^3 、 10^4 、 10^5 、 10^6 、 10^7 、 10^8 、 10^9 、 10^{10} 個、またはそれ以上の核酸断片を含む。いくつかの実施態様では、本明細書に記載した方法は、500bp、1000bp、1500bp、2000bp、2500bp、3000bp、3500bp、4000bp、5000bp、6000bp、7000bp、8000bp、9000bp、10000bp、またはそれ以上を配列決定することができる。

【0072】

いくつかの実施態様では、この方法は、(a)核酸鋳型を、5'末端に第1の検出配列を含むトランスポゾンと、トランスポゾンが核酸鋳型に挿入され、かつ核酸鋳型が断片化されて核酸断片の5'末端に付着されたトランスポゾンを含む複数のタグメンテーションされた二本鎖核酸断片になる条件下で接触させることと；(b)(i)核酸鋳型内のあらかじめ決定された位置に相補的なヌクレオチド配列を含み、かつその5'末端に第2の検出配列を含む第1の固定プライマーと、(ii)第1の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第2の選択的プライマーとを使用して、タグメンテーションされた核酸断片を増幅させて、第1の検出配列、核酸断片の5'末端に付着されたトランスポゾン、および第2の検出配列を含む、増幅された核酸断片を形成させることと；(c)増幅された核酸断片を配列決定することとを含み、ここでは、核酸鋳型は、RNAガイド型ヌクレアーゼを使用して改変されている。

【0073】

いくつかの実施態様では、ステップ(b)の増幅された核酸断片は、第1の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第2の選択的プライマーと、第2の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第3の選択的プライマーとを使用して増幅される。いくつかの実施態様では、第3の選択的プライマーは、バーコード(例えば、プーリングバーコード)および/または配列タグ(例えば、リバースILLUMINAアダプター(i7))を含む。

l u m i n a アダプター (i 7)) を含む。

【 0 0 7 4 】

いくつかの実施態様では、この方法は、(a) 細胞または組織から得られたゲノム DNA 試料を、ヌクレアーゼによるゲノム DNA の切断、および二本鎖オリゴヌクレオチドの少なくとも 1 つの切断部位への組み込みに好都合な条件下で、RNA ガイド型ヌクレアーゼおよび二本鎖オリゴヌクレオチド (d s O D N) と接触させることとであって、それによって、組み込まれた二本鎖オリゴヌクレオチドを含むタグ付けされたゲノム DNA が産生される、ことと；(b) タグ付けされたゲノム DNA を、5 ' 末端に第 1 の検出配列を含むトランスポゾンと、タグ付けされたゲノム DNA にトランスポゾンが挿入され、かつタグ付けされたゲノム DNA が断片化されて核酸断片の 5 ' 末端に付着されたトランスポゾンを含む複数のタグメンテーションされた二本鎖核酸断片になる条件下で接触させることと；(c) (i) 二本鎖オリゴヌクレオチドの少なくとも一部分に相補的なヌクレオチド配列を含み、かつその 5 ' 末端に第 2 の検出配列を含む第 1 の固定プライマーと、(i i) 第 1 の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第 2 の選択的プライマーとを使用して、タグメンテーションされた核酸断片を増幅させて、第 1 の検出配列、核酸断片の 5 ' 末端に付着されたトランスポゾン、および第 2 の検出配列を含む、増幅された核酸断片を形成させることと；(d) 増幅された核酸断片を配列決定することを含む。

【 0 0 7 5 】

いくつかの実施態様では、ステップ (c) の増幅された核酸断片は、第 1 の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第 2 の選択的プライマーと、第 2 の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第 3 の選択的プライマーとを使用して増幅される。いくつかの実施態様では、第 3 の選択的プライマーは、バーコード (例えば、プーリングバーコード) および / または配列タグ (例えば、リバース I l l u m i n a アダプター (i 7)) を含む。例示的な方法を、図 3 に示す。

【 0 0 7 6 】

いくつかの実施態様では、配列決定のステップから得られた配列が、参照ゲノム DNA 試料 (例えば、対照試料) からの 1 つまたは複数の配列と比較される。

【 0 0 7 7 】

例えば、本明細書によって開示された方法は、診断ツールとして使用することができる。

【 0 0 7 8 】

例えば、配列決定のステップから得られた配列に基づいて、および / またはこうした配列の、参照ゲノム DNA 試料からの配列との比較に基づいて、決定を行うことができる。

【 0 0 7 9 】

いくつかの実施態様では、核酸鋳型は、標的 DNA への複数の挿入が起こり、これらがそれぞれ、転移鎖を含むまたは転移鎖からなる第 1 のタグの、標的 DNA 内のヌクレオチドの 5 ' 末端への連結をもたらし、それによって、標的 DNA が断片化され、アニーリングされた 5 ' - タグ付けされた DNA 断片 (これらはそれぞれ、5 ' 末端上に第 1 のタグを有する) の集団がもたらされる条件下かつ十分な時間、トランスポゾンと接触される。

【 0 0 8 0 】

いくつかの実施態様では、インビトロ転移反応において使用されるトランスポザーゼおよびトランスポゾンの、またはトランスポソーム組成物の量は、50 マイクロリットルの反応あたり、50 ナノグラムの標的 DNA あたり、約 1 ピコモル ~ 約 25 ピコモルである。いくつかの実施態様では、インビトロ転移反応において使用されるトランスポザーゼおよびトランスポゾン組成物の、またはトランスポソーム組成物の量は、50 マイクロリットルの反応あたり、50 ナノグラムの標的 DNA あたり、約 5 ピコモル ~ 約 50 ピコモルである。トランスポザーゼが高活性な T n 5 トランスポザーゼまたは M u A トランスポザーゼであるいくつかの実施態様では、インビトロ転移反応において使用されるトランスポザーゼおよびトランスポゾンの、またはトランスポソームの最終濃度は、少なくとも 250 n M であり；いくつかの他の実施態様では、高活性な T n 5 トランスポザーゼまたは M u A トランスポザーゼの、また、そのそれぞれのトランスポゾン最終組成物またはトラン

10

20

30

40

50

スポソーム組成物の最終濃度は、少なくとも500 nMである。

【0081】

いくつかの実施態様では、インビトロ転移反応のための反応時間は、2時間以下、1時間以下、30分以下、または15分以下である。いくつかの実施態様では、インビトロ転移反応のための反応時間は、5分以下である。トランススポソーム組成物が高活性なTn5トランスポザゼを含むいくつかの実施態様では、インビトロ転移反応のための反応時間は、5分以下である。

【0082】

トランスポゾン、検出可能な標識で標識することができる。こうした標識は、例えば、トランスポゾンを含む核酸を捉える、精製する、および/または濃縮するために有用であり得る。例えば、ビオチン、蛍光標識、抗体標識、抗体断片標識、および/または本明細書に言及した他の検出可能な標識を含めた、非常に様々な検出可能な標識のいずれかが、この目的のために好適であり得る。

【0083】

いくつかの実施態様では、配列決定タグは、ある特定の配列決定プラットフォームのための次世代配列決定のための鋳型の産生のために使用される(例えば、ROCHE 454 Aまたは454 B配列決定プラットフォームのための; ILLUMINA SOLEXA配列決定プラットフォームのための; APPLIED BIOSYSTEMS SOLID(商標)配列決定プラットフォームのための; PACIFIC BIOSCIENCE S' SMRT(商標)配列決定プラットフォームのための; POLLONATOR POLLONY配列決定プラットフォームのための; HELICOS配列決定プラットフォームのための; COMPLETE GENOMICS配列決定プラットフォームのための; INTELLIGENT BIOSYSTEMS配列決定プラットフォームのための; または他のあらゆる配列決定プラットフォームのための配列決定タグ)。いくつかの実施態様では、配列決定タグは、全長のIlluminaフォワードアダプター(i5)である。いくつかの実施態様では、配列決定タグは、全長のIlluminaリバースアダプター(i7)である。

【0084】

PCRによる増幅反応を実施するために、非常に様々な酵素およびキットが利用可能である。例えば、いくつかの実施態様では、PCR増幅は、EPICENTRE Biot technologies, Madison, Wis. のFAILSAFE(商標)PCRシステムまたはMASTERAMP(商標)Extra-Long PCRシステムを、製造業者によって記載されている通りに使用して実施される。これらのシステムは、特定の鋳型およびプライマー対に最適なPremixを特定するための各システムと共に提供される一連の2xPCR Premixを使用して、PCR反応条件の迅速な最適化を可能にする。しかし、この開示は、増幅反応のためのこれらの生成物または条件の使用に限定されず、標的配列とアニーリングするプライマーと、トランスポゾンとアニーリングするプライマーとの間の配列の増幅を可能にする、あらゆる好適な熱安定性DNAポリメラーゼおよび反応混合物を使用することができる。本明細書に記載した方法のいずれかにおいて、非鎖置換型ポリメラーゼまたは鎖置換型ポリメラーゼを使用することができる。

【0085】

核酸の増幅のためのプロトコルは、当技術分野で公知であり、例えば、1ラウンドの増幅または複数ラウンドの増幅が含まれ得る。例えば、第1の増幅ラウンドの後に、増幅の2つのラウンドの間に1つまたは複数の処理(例えば、精製、濃縮など)ステップを伴ってまたは伴わずに、第2の増幅ラウンドが続くことができる。いくつかの実施態様では、間に1つまたは複数の処理ステップを伴ってまたは伴わずに、追加のラウンドの増幅を使用することができる。

【0086】

増幅サイクルの数は、実施態様に依って様々であり得る。例えば、増幅の各ラウンドは、少なくとも8、10、12、14、16、18、20、22、24、26、28、30

10

20

30

40

50

、または32サイクルを含むことができる。2ラウンド以上の増幅を含む実施態様では、各ラウンドにおいて使用される増幅サイクルの数は、同じである可能性も、異なる可能性もある。2ラウンドの増幅が使用される場合、2つのラウンドにおける増幅サイクルの数は異なり、第1のラウンドは、第2のラウンドよりも多いサイクルを含むこともできるし、第2のラウンドよりも少ないサイクルを含むこともできる。一例として、増幅の第1のラウンドは、12サイクルを含むことができ、増幅の第2のラウンドは、15サイクルを含むことができる。別の例として、増幅の第1のラウンドは、10サイクルを含むことができ、増幅の第2のラウンドは、12サイクルを含むことができる。

【0087】

2ラウンド以上の増幅が使用される場合、同じまたは異なるセットのプライマーを使用することができる。例えば、ネステッド増幅スキーム（例えばネステッドPCR）を使用することができる。

【0088】

2ラウンドの増幅を含むいくつかの実施態様では、あるラウンドにおけるフォワードプライマーは、別のラウンドにおけるフォワードプライマーとして使用されるが、2つのラウンドにおけるリバースプライマーは異なる。あるいは、フォワードプライマーとリバースプライマーが、2つのラウンドにおいて異なる可能性がある。あるいは、一方のラウンドにおけるリバースプライマーは、他方のラウンドにおけるリバースプライマーとして使用されるが、2つのラウンドにおけるフォワードプライマーは異なる。

【0089】

増幅反応におけるプライマーの最終濃度は、実施態様に応じて変動し得る。例えば、単一ラウンドの増幅において、または多ラウンドの増幅の同じラウンドにおいて、2つ以上のプライマーと一緒に使用される場合、プライマーの濃度は、異なる可能性も同じである可能性もある。例えば、プライマーは、増幅反応において、約100nM～約500nMの範囲の濃度、例えば、約100nM、約150nM、約200nM、約250nM、約300nM、約350nM、約400nM、約450nM、または約400nMで存在することができる。

【0090】

いくつかの実施態様では、同じラウンドの増幅中に存在する別のプライマーに対して過剰量の、あるプライマーが使用される。いくつかの実施態様では、少なくとも1ラウンドの増幅における少なくとも1種のプライマーは、500nMを超える濃度、例えば、1μM、2μM、3μM、4μM、5μMなどで存在する。

【0091】

いくつかの実施態様では、反応は、マグネシウム、例えばMgCl₂を含有する緩衝液中で行われる。マグネシウム（例えばMgCl₂）の好適な濃度の例としては、限定はされないが、約1mM、約1.5mM、約2mM、約2.5mM、約3mM、約3.5mM、約4mM、約4.5mM、または約5mMが挙げられる。

【0092】

いくつかの実施態様では、dNTPの混合物は、約0.2mM～約0.8mMまでの範囲の濃度、例えば、約0.2mM、約0.3mM、約0.4mM、約0.5mM、約0.6mM、約0.7mM、または約0.8mMで使用される。

【0093】

増幅反応の総体積は、実施態様に応じて変動し得る。例えば、約10μL～約200μL、例えば、約10μL、約15μL、約20μL、約25μL、約30μL、約35μL、約40μL、約45μL、約50μL、約55μL、約60μL、約65μL、約70μL、約75μL、約80μL、約85μL、約90μL、約95μL、約100μL、約110μL、約120μL、約130μL、約140μL、約150μL、約160μL、約170μL、約180μL、約190μL、または約200μLの反応体積を使用することができる。

【0094】

10

20

30

40

50

一般に、増幅のサイクルは、変性段階（その間、核酸の鎖が分離される）、アニーリング段階（その間、プライマーが核酸鋳型にアニーリングする）、および伸長段階（その間、アニーリングされたプライマーが、核酸鋳型を使用して伸長する）を含む。

【0095】

各段階の温度は異なり、変性段階中に使用される温度は、伸長段階の温度よりも高い傾向があり、伸長段階の温度は、アニーリング段階中に使用される温度よりも高い傾向がある。

【0096】

例えば、変性段階に適した温度は、約90～約100までの範囲、例えば、約90、約91、約92、約93、約94、約95、約96、約97、約98、約99、または約100であり得る。

10

【0097】

例えば、アニーリング段階に適した温度は、約50～約65までの範囲、例えば、約50、約51、約52、約53、約54、約55、約56、約57、約58、約59、約60、約61、約62、約63、約64、または約65であり得る。

【0098】

例えば、伸長段階に適した温度は、約68～約80までの範囲、例えば、約68、約69、約70、約71、約72、約73、約74、約75、約76、約77、約78、約79、または約80であり得る。

20

【0099】

単一のサイクル中の各段階の期間は、変動し得る。一般に、アニーリングおよび伸長段階の期間は、変性段階の期間よりも長い可能性がある。

【0100】

例えば、変性段階は、各サイクルにおいて、約4秒～約20秒、例えば、約4秒、約5秒、約6秒、約7秒、約8秒、約9秒、約10秒、約11秒、約12秒、約13秒、約14秒、約15秒、約16秒、約17秒、約18秒、約19秒、または約20秒、持続することができる。

【0101】

アニーリングおよび伸長段階の期間は、同じまたは異なる可能性がある。例えば、アニーリングまたは伸長段階は、各サイクルにおいて、約20秒～約2分、例えば、約20秒、約25秒、約30秒、約35秒、約40秒、約45秒、約50秒、約55秒、約60秒、約65秒、約70秒、約75秒、約80秒、約85秒、約90秒、約95秒、約100秒、約105秒、約110秒、または約120秒、持続することができる。

30

【0102】

2ラウンド以上の増幅が使用される実施態様では、ラウンド間に、1つまたは複数の処理ステップを使用することができる。例えば、1つまたは複数の精製および/または濃縮ステップを使用することができる。好適な精製ステップには、限定はされないが、プロテイナーゼKでの処理、タンパク質の化学変性剤（例えば尿素）での処理、ビーズに基づく方法（例えば、可逆的固相固定化（Solid Phase Reversible Immobilization）技術を使用するもの）、カラム上DNA精製法などが含まれる。DNA精製および/または濃縮装置のキットは、市販品として入手可能であり、例えば、Zymo Research、Thermo Fisher Scientific、Sigma Aldrichなどからのキットが挙げられる。

40

【0103】

例えば、反応の構成成分を不活性化するための1つまたは複数のステップ、例えば、加熱、pHの変化、および/または化学物質によるポリメラーゼなどの酵素の失活を使用することができる。

【0104】

この開示はまた、DNA断片を増幅するためのPCRの使用に限定されない。同じ配列

50

を増幅し、かつ意図される目的に適した組成および量の増幅産物を産生する、あらゆる好適な増幅方法（例えば、ローリングサークル増幅、リボプライマー増幅（例えば、米国特許第7,413,857号明細書）、ICAN、UCAN、ribospia、末端タギング（米国特許出願第20050153333号明細書）、Eberwine型aRNA増幅、または鎖置換増幅）を、本開示の実施態様において使用することができる。例えば、使用することができるいくつかの鎖置換方法は、宝酒造株式会社（日本、京都）の国際公開第02/16639号パンフレット；国際公開第00/56877号パンフレット；およびAU 00/29742号明細書；Becton Dickinson and Companyの米国特許第5,523,204号；同第5,536,649号；同第5,624,825号；同第5,631,147号；同第5,648,211号；同第5,733,752号；同第5,744,311号；同第5,756,702号；および同第5,916,779号明細書；Nanogen/Becton Dickinson Partnershipの米国特許第6,238,868号；同第6,309,833号；および同第6,326,173号明細書；Bio Merieuxの米国特許第5,849,547号；同第5,874,260号；および同第6,218,151号明細書；Gen-Probe, Inc.の米国特許第5,786,183号；同第6,087,133号；および同第6,214,587号明細書；Wicket al.の米国特許第6,063,604号明細書；Kurnの米国特許第6,251,639号明細書；栄研化学株式会社（日本、東京）の米国特許第6,410,278号明細書；および国際公開第00/28082号パンフレット；Auerbachの米国特許第5,591,609号；同第5,614,389号；同第5,773,733号；同第5,834,202号；および同第6,448,017号明細書；ならびにLizardiの米国特許第6,124,120号；および同第6,280,949号明細書に記載されている。

【0105】

ゲノムDNA

本明細書に開示した方法において使用されるゲノムDNA試料は、一般に、あらゆる種類の細胞および/または組織、例えば、初代細胞および/または組織、少なくともある期間培養した細胞または組織、または初代と培養細胞および/または組織の組み合わせから得ることができる。

【0106】

いくつかの実施態様では、UDiTASは、単一の細胞由来のゲノムDNA上で実施される。例えば、単一の細胞由来のゲノムDNAは、UDiTASを実施する前に増幅することができる。全ゲノム増幅方法は、当技術分野で公知である。様々なプロトコルおよび/または市販品として入手可能なキットのいずれかを使用することができる。市販品として入手可能なキットの例としては、限定はされないが、QIAGENのREPLI-g Single Cell Kit、Sigma AldrichのGENOMEPLEX（登録商標）Single Cell Whole Genome Amplification Kit、Silicon BiosystemsのAmpli1（商標）WGA Kit、およびGE Healthcare Life Sciencesのillustra Single Cell GenomiPhi DNA Amplification Kitが挙げられる。したがって、いくつかの実施態様では、UDiTASが、単一の細胞由来のゲノムDNA上で実施されるとしても、そのゲノムDNA上で、複数の反応、例えば、それぞれの反応が本明細書にさらに記載する通りの異なる固定プライマーを使用する複数の反応を実施することが可能である。

【0107】

この開示の方法はまた、ウイルスによって導入された配列を含む、および/または本明細書に記載した二本鎖オリゴヌクレオチドを含む、ゲノムDNA試料を使用して実施することができる。例えば、この開示の方法は、ウイルスベクターを使用して配列をゲノムDNAに導入する、および/または二本鎖オリゴヌクレオチドを本明細書に記載した通りにゲノムDNAに挿入するステップを含むことができる。配列を導入するステップは、一般

に、ゲノムDNAをトランスポゾンと接触させるステップの前に実施される。

【0108】

いくつかの実施態様では、ゲノムDNA試料は、生きている細胞および/または組織を含有しない、または、痕跡量の細胞および/または組織を含有する試料から得られる。例えば、本開示の方法は、限定はされないが、血液、骨、爪、毛髪、歯、スワブ（例えば、頬側スワブ、生殖器スワブ、子宮頸部スワブ、肛門スワブなど）、および尿試料などの、法医学的DNA試料に対して使用することができる。使用されたガーゼまたは創傷被覆材からの試料も使用することができる。

【0109】

あるいはまたはさらに、本明細書に開示した方法は、真核細胞由来のゲノムDNA試料および/または原核細胞由来の組織と共に使用することができる。例えば、本開示の方法は、微生物由来の、および/または患者（例えば、抗生物質を受けている患者）からの分離物由来のゲノムDNAを使用して実施することができる。いくつかの実施態様では、例えばメタゲノムマイニングのために、微生物群集および/または1つまたは複数の微生物叢由来のゲノムDNAが使用される。（例えば、Delmont et al, "Metagenomic mining for microbiologists," ISME J. 2011 Dec; 5(12): 1837-43を参照のこと）。

【0110】

ゲノムDNAは、例えば、本明細書に記載した細胞および/または組織に対するある種の操作を含めた、様々な好適な方法のいずれかを使用して調製することができる。例示的な非限定的な操作には、細胞および/または組織をヌクレアーゼ（例えば、部位特異的ヌクレアーゼおよび/またはRNAガイド型ヌクレアーゼ）と接触させること、またはこうしたヌクレアーゼを含むゲノム編集システムが含まれる。

【0111】

検出配列

検出配列は、本明細書で使用する場合、本明細書に記載した核酸鋳型、トランスポゾン、および/またはプライマー上に存在することができる、また、それを含有する核酸、または核酸断片の回収および/または検出を容易にする配列要素を指す。いくつかの実施態様では、1つまたは複数の検出配列が、オリゴヌクレオチドアレイによる捕捉を容易にするまたは媒介する、および/または配列決定、例えば、本明細書に記載した連結産物の配列決定を容易にするまたは媒介する。

【0112】

いくつかの実施態様では、検出配列は、増幅および/または配列決定を容易にする。いくつかの実施態様では、検出配列は、増幅および/または配列決定プライマーによって認識されることが可能な1つまたは複数の配列を含む。

【0113】

いくつかの実施態様では、検出配列は、UMIを含む。いくつかの実施態様では、検出配列は、バーコードを含む。いくつかの実施態様では、検出配列は、試料バーコードを含む。

【0114】

例えば、いくつかの実施態様では、検出配列は、配列決定方法に使用するための配列アダプターを含む。いくつかの実施態様では、こうした配列アダプターは、増幅プライマー結合部位および配列決定プライマー結合部位を含む。いくつかの実施態様では、こうした配列アダプターは、増幅プライマー結合部位と配列決定プライマー結合部位の両方として働くプライマー結合部位を含む。いくつかの実施態様では、増幅プライマー結合部位は、配列決定プライマー結合部位と重なる。

【0115】

いくつかの実施態様では、増幅プライマー結合部位は、長距離の増幅のために使用される。

【0116】

10

20

30

40

50

いくつかの実施態様では、配列アダプターは、アダプターの一端を標識するマーカー配列をさらに含む。

【0117】

いくつかの実施態様では、配列アダプターは、バーコード配列をさらに含む。本明細書に記載した方法において使用することができる検出配列は、当技術分野で公知である。例えば、配列決定アダプター（例えば、MiSeqアダプター）（Illuminaから入手可能）を、検出配列として使用することができる（例えば、i5、i7）。

【0118】

二本鎖オリゴヌクレオチド

本明細書に記載した通り、ゲノムDNAは、二本鎖オリゴヌクレオチドを含むことができる。いくつかの実施態様では、（ゲノムDNA試料が得られる）細胞または組織が、二本鎖オリゴヌクレオチドおよびヌクレアーゼ（例えば、部位特異的ヌクレアーゼおよび/またはRNAガイド型ヌクレアーゼ）と、例えば、ヌクレアーゼによるゲノムDNAの切断に好都合な条件下で接触される。二本鎖オリゴヌクレオチドが使用される場合、二本鎖オリゴヌクレオチドを、ゲノムDNA内の1つまたは複数のヌクレアーゼ切断部位に組み込むことができる。例えば、ヌクレアーゼは、ゲノムDNA中の1つまたは複数の二本鎖切断を誘発することができ、二本鎖オリゴヌクレオチドが、1つまたは複数の二本鎖切断の少なくとも1つに組み込まれることができる。

10

【0119】

ゲノムDNAが、二本鎖オリゴヌクレオチドを含む場合、本明細書に記載した方法は、例えば増幅ステップにおいて、二本鎖オリゴヌクレオチドの少なくとも一部分に相補的なヌクレオチド配列を含む固定プライマーを使用することを含む。

20

【0120】

好適な二本鎖オリゴヌクレオチドを設計し、こうしたオリゴヌクレオチドを組み込む方法は、当技術分野で公知であり、例えば、その内容全体が参照により本明細書に組み込まれる、国際公開第2015/200378A1号パンフレットに記載されている。

【0121】

いくつかの実施態様では、二本鎖オリゴヌクレオチドのどちらの鎖も、ゲノムDNAに対してオルソログスである。例えば、ゲノムDNAがヒトゲノムDNAである実施態様では、二本鎖オリゴヌクレオチドのどちらの鎖も、存在する場合には、ヒトゲノムDNAに対してオルソログスであり得る。

30

【0122】

いくつかの実施態様では、二本鎖オリゴヌクレオチドは、化学修飾を含む。例えば、二本鎖オリゴヌクレオチドの各5'末端は、リン酸基を含むことができる。あるいはまたはさらに、二本鎖オリゴヌクレオチドは、ホスホロチオエート結合を含有することができる。例えば、二本鎖オリゴヌクレオチドの各3'末端および/または各5'末端は、ホスホロチオエート結合を含むことができる。

【0123】

いくつかの実施態様では、二本鎖オリゴヌクレオチドは、両方の5'末端でリン酸化され、両方の5'末端上にホスホロチオエート結合を含有し、両方の3'末端上にホスホロチオエート結合を含有する。

40

【0124】

いくつかの実施態様では、二本鎖オリゴヌクレオチドは、30～35個のヌクレオチドまたは60～65個のヌクレオチドを含む。

【0125】

二本鎖オリゴヌクレオチドの末端は、平滑であるまたは突出を含む（例えば、1、2、3、または4塩基の突出ヌクレオチドを有する5'末端を含む）こともできるし、一方の末端が平滑であり、他方が突出を含むこともできる。

【0126】

増幅ステップを含む方法において二本鎖オリゴヌクレオチドが使用されるいくつかの実

50

施態様では、増幅ステップは、二本鎖オリゴヌクレオチド内のあらかじめ決定された配列に相補的なヌクレオチド配列を含む固定プライマーを使用して、ローリングサークルPCRを実施することを含む。

【0127】

検出可能な標識

いくつかの実施態様では、本明細書に記載したトランスポゾン、二本鎖オリゴヌクレオチド、および/またはプライマーは、検出可能な標識で標識される。こうした標識は、例えば、トランスポゾン、二本鎖オリゴヌクレオチド、および/またはプライマーを含有する核酸を捉える、精製する、および/または濃縮するために有用であり得る。例えば、ビオチン、蛍光標識、抗体標識、抗体断片標識、および/または本明細書に言及した他の検出可能な標識を含めた、非常に様々な検出可能な標識のいずれかが、この目的のために好適であり得る。

10

【0128】

固有分子識別子

いくつかの実施態様では、本明細書に記載した核酸鋳型、トランスポゾン、および/またはプライマーは、固有分子識別子（本明細書では「UMI」と省略される）を含むことができる。UMIは、核酸鋳型またはその一部分に関する情報を検索するために使用することができる配列を指す。例えば、複数のプライマーを含む開示の方法では、各UMIは、特定のプライマー、例えば、特異的な標的配列と結合するプライマーと関係がある可能性がある。いくつかの実施態様では、UMIは、UMIバーコードと呼ばれる。

20

【0129】

いくつかの実施態様では、検出配列の検出は、特定のUMIを特定する、例えば、その特定のUMIが伴われる標的、トランスポゾン、またはプライマーを特定するために使用することができる。

【0130】

いくつかの実施態様では、UMIは、ランダムに作成される配列である。いくつかの実施態様では、UMIは、長さが8～20ヌクレオチド、例えば、長さが10～16ヌクレオチド、例えば長さが10、11、12、13、14、15、および16ヌクレオチドである。様々な状況でのUMIの生成および使用は、当技術分野で公知である。

【0131】

ヌクレアーゼ

本開示の方法は、周知のヌクレアーゼとそれほど特徴づけられていないヌクレアーゼとの両方を含めた、様々なヌクレアーゼの切断プロフィールを評価するのに適している。一般に、ヌクレアーゼは、ヌクレアーゼの「標的部位」と呼ばれる特定の配列または配列のセットでのみ切断することが知られているまたは予想される点において、部位特異的である。

30

【0132】

本明細書によって開示された方法では、ヌクレアーゼとのインキュベーションステップは、一般に、ヌクレアーゼによる切断に好都合な条件下で行われる。すなわち、所与の候補標的部位またはバリエーション標的部位が、ヌクレアーゼによって実際には切断されないとしても、インキュベーション条件は、ヌクレアーゼが、その公知の標的部位を含有する鋳型の少なくともかなりの部分（例えば、少なくとも10%、少なくとも20%、少なくとも30%、少なくとも40%、少なくとも50%、少なくとも60%、少なくとも70%、少なくとも80%、少なくとも90%、または少なくとも95%）を切断するようなものである。公知のおよび一般に良く特徴付けられたヌクレアーゼについては、こうした条件は、当技術分野で一般に公知である、および/または容易に発見または最適化することができる。新しく発見されたヌクレアーゼについては、こうした条件は、一般に、より特徴付けられている関連するヌクレアーゼ（例えば、相同体およびオルソログ）に関する情報を使用して概算することができる。

40

【0133】

50

いくつかの実施態様では、ヌクレアーゼは、エンドヌクレアーゼである。いくつかの実施態様では、ヌクレアーゼは、部位特異的エンドヌクレアーゼ（例えば、制限エンドヌクレアーゼ、メガヌクレアーゼ、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）、ジンクフィンガーヌクレアーゼなど）である。

【0134】

いくつかの実施態様では、部位特異的ヌクレアーゼの部位特異性は、補助的な分子によって与えられる。例えば、C R I S P R 関連（C a s）ヌクレアーゼは、本明細書に記載した通りの「ガイドRNA」またはgRNAによって、特定の部位に誘導される。いくつかの実施態様では、ヌクレアーゼは、RNAガイド型ヌクレアーゼである。いくつかの実施態様では、ヌクレアーゼは、C R I S P R 関連ヌクレアーゼである。

10

【0135】

いくつかの実施態様では、ヌクレアーゼは、これまでに知られているヌクレアーゼの相同体またはオルソログ、例えば、新しく発見された相同体またはオルソログである。

【0136】

RNAガイド型ヌクレアーゼ

本開示によるRNAガイド型ヌクレアーゼには、限定はされないが、天然起源のクラス2 C R I S P R ヌクレアーゼ、例えばC a s 9、およびC p f 1、ならびにそれに由来するまたはそこから得られる他のヌクレアーゼが含まれる。RNAガイド型ヌクレアーゼは、機能面では、(a) gRNAと相互作用（例えば複合体形成）する；および(b) gRNAと共に、(i) gRNAのターゲティングドメインに相補的な配列、および場合によっては(ii)「プロトスペーサー隣接モチーフ」または「PAM」と呼ばれる追加の配列（これは、以下により詳細に記載される）を含むDNAの標的領域と結合し、場合によっては切断するまたは改変するヌクレアーゼと定義される。以下の実施例が説明する通り、RNAガイド型ヌクレアーゼは、同じPAM特異性または切断活性を有する個々のRNAガイド型ヌクレアーゼの間に差異が存在する可能性があるとしても、大まかに、そのPAM特異性および切断活性によって定義することができる。当業者は、本開示のいくつかの態様が、ある種のPAM特異性および/または切断活性を有するあらゆる好適なRNAガイド型ヌクレアーゼを使用して実行することができるシステム、方法、および組成物に関することを認識するであろう。この理由から、そうではないと指定されない限り、RNAガイド型ヌクレアーゼという用語は、包括的な用語と理解されるべきであり、RNAガイド型ヌクレアーゼの、いずれかの特定の型（例えば、C p f 1に対するC a s 9）、種（例えば、黄色ブドウ球菌（S . a u r e u s）に対する化膿レンサ球菌（S . p y o g e n e s））、または差異（例えば、切り詰めまたは分断に対する全長；操作されたPAM特異性に対する天然起源のPAM特異性など）に限定されない。

20

30

【0137】

PAM配列は、gRNAターゲティングドメイン（または「スペーサー」）に相補的である「プロトスペーサー」配列に対するその連続的な関係から、その名前が付けられている。プロトスペーサー配列と共に、PAM配列は、特定のRNAガイド型ヌクレアーゼ/gRNA組み合わせのための標的領域または配列を定義する。

【0138】

様々なRNAガイド型ヌクレアーゼが、PAMとプロトスペーサーとの間の異なる連続的な関係を必要とする可能性がある。一般に、C a s 9は、ガイドRNAターゲティングドメインに対して可視化されているプロトスペーサーの3'であるPAM配列を認識する。他方で、C p f 1は、一般に、プロトスペーサーの5'であるPAM配列を認識する。

40

【0139】

RNAガイド型ヌクレアーゼは、PAMおよびプロトスペーサーの特定の連続的な方向性を認識することに加えて、特定のPAM配列を認識することもできる。黄色ブドウ球菌（S . a u r e u s）C a s 9は、例えば、N N G R R TまたはN N G R R V（ここでは、N残基は、gRNAターゲティングドメインによって認識される領域に隣接する3'である）のPAM配列を認識する。化膿レンサ球菌（S . p y o g e n e s）C a s 9は、N G

50

G PAM配列を認識する。また、フランシセラ・ノビシダ (*F. novicida*) Cpfl は、TTN PAM配列を認識する。PAM配列は、様々なRNAガイド型ヌクレアーゼについて同定されており、新規のPAM配列を同定するための戦略は、Shmakov et al., 2015, *Molecular Cell* 60, 385-397, November 5, 2015によって記載されている。操作されたRNAガイド型ヌクレアーゼが、参照分子のPAM特異性とは異なるPAM特異性を有する可能性がある(例えば、操作されたRNAガイド型ヌクレアーゼの場合、参照分子は、RNAガイド型ヌクレアーゼが由来する天然起源のバリエーション、または操作されたRNAガイド型ヌクレアーゼに対する最大のアミノ酸配列相同性を有する天然起源のバリエーションであり得る)ことにも留意するべきである。

10

【0140】

RNAガイド型ヌクレアーゼは、そのPAM特異性に加えて、そのDNA切断活性で特徴付けることができる:天然起源のRNAガイド型ヌクレアーゼは、典型的には、標的核酸においてDSBを形成するが、SSBのみをもたらす(上記)参照により本明細書に組み込まれるRan&Hsu, et al., *Cell* 154(6), 1380-1389, September 12, 2013("Ran")), または全く切断しない、操作されたバリエーションが産生されている。

【0141】

Cas9

化膿レンサ球菌 (*S. pyogenes*) Cas9について(Jinek et al., *Science* 343(6176), 1247997, 2014("Jinek 2014")), また、単分子ガイドRNAと標的DNAとの複合体の黄色ブドウ球菌 (*S. aureus*) Cas9について(Nishimasu 2014; Anders et al., *Nature*, 2014 Sep 25; 513(7519): 569-73("Anders 2014")); およびNishimasu 2015)、結晶構造が決定されている。

20

【0142】

天然起源のCas9タンパク質は、2つのローブ:認識(REC)ローブとヌクレアーゼ(NUC)ローブを含む;これらはそれぞれ、特定の構造および/または機能ドメインを含む。RECローブは、アルギニンリッチブリッジヘリックス(BH)ドメイン、および少なくとも1つのRECドメイン(例えば、REC1ドメインと、場合によってはREC2ドメイン)を含む。RECローブは、他の公知のタンパク質との構造的類似性を有さず、これが固有の機能ドメインであることが示唆される。いかなる理論にも拘泥されることを望まないが、変異解析は、BHおよびRECドメインについての特定の機能的役割を示唆している: BHドメインが、gRNA:DNA認識における役割を果たすようであるのに対し、RECドメインは、gRNAのリビート:アンチリビート二本鎖と相互作用して、Cas9/gRNA複合体の形成を媒介すると考えられる。

30

【0143】

NUCローブは、RuvCドメイン、HNHドメイン、およびPAM相互作用(interacting)(PI)ドメインを含む。RuvCドメインは、レトロウイルスインテグラーゼスーパーファミリーメンバーとの構造的類似性を有し、標的核酸の非相補(すなわち、ボトム)鎖を切断する。これは、2つ以上のスプリットRuvCモチーフ(化膿レンサ球菌 (*S. pyogenes*) および黄色ブドウ球菌 (*S. aureus*) におけるRuvC I、RuvC II、およびRuvC IIIなど)から形成される可能性がある。一方、HNHドメインは、HNNエンドヌクレアーゼモチーフと構造的に類似であり、標的核酸の相補(すなわち、トップ)鎖を切断する。PIドメインは、その名前が示唆する通り、PAM特異性に寄与する。

40

【0144】

Cas9のある種の機能は、(必ずしもそれによって完全に決定されるわけではないが)上で説明した特定のドメインと関係があるが、これらの機能および他の機能は、他のC

50

a s 9 ドメインによって、またはいずれかのローブ上の複数のドメインによって媒介されるまたは影響を受ける可能性がある。例えば、化膿レンサ球菌 (*S . p y o g e n e s*) C a s 9 では、N i s h i m a s u 2 0 1 4 に記載されている通り、g R N A のリピート：アンチリピート二本鎖は、R E C と N U C ローブとの間の溝に入り、この二本鎖内のヌクレオチドは、B H、P I、および R E C ドメイン内のアミノ酸と相互作用する。第 1 のステム・ループ構造内のいくつかのヌクレオチドも、複数のドメイン (P I、B H、および R E C 1) 内のアミノ酸と相互作用し、第 2 および第 3 のステム・ループ内のいくつかのヌクレオチドも同様である (R u v C および P I ドメイン)。

【 0 1 4 5 】

C p f 1

c r R N A と二本鎖 (d s) D N A 標的 (T T T N P A M 配列を含む) との複合体のアシダミノコッカス (*A c i d a m i n o c o c c u s*) 種 C p f 1 の結晶構造は、Y a m a n o らによって解明されている (参照により本明細書に組み込まれる C e l l 1 . 2 0 1 6 M a y 5 ; 1 6 5 (4) : 9 4 9 - 9 6 2 (“ Y a m a n o ”)) 。 C p f 1 は、C a s 9 のように、2 つのローブ：R E C (認識) ローブ、および N U C (ヌクレアーゼ) ローブを有する。R E C ローブは、あらゆる公知のタンパク質構造との類似性がない R E C 1 および R E C 2 ドメインを含む。一方、N U C ローブは、3 つの R u v C ドメイン (R u v C - I、- I I、および - I I I)、ならびに B H ドメインを含む。しかし、C a s 9 とは対照的に、C p f 1 R E C ローブは、H N H ドメインがなく、やはり公知のタンパク質構造との類似性がない他のドメイン：構造的に固有の P I ドメイン、3 つのくさび (W e d g e) (W E D) ドメイン (W E D - I、- I I、および - I I I)、ならびにヌクレアーゼ (N u c) ドメインを含む。

【 0 1 4 6 】

C a s 9 と C p f 1 は、構造および機能において類似性を有するが、ある種の C p f 1 活性が、あらゆる C a s 9 ドメインと類似ではない構造ドメインによって媒介されることを認識するべきである。例えば、標的 D N A の相補鎖の切断は、C a s 9 の H N H ドメインとは配列的および空間的に異なる N u c ドメインによって媒介されるようである。さらに、C p f 1 g R N A の非ターゲティング部分 (ハンドル) は、C a s 9 g R N A においてリピート：アンチリピート二本鎖によって形成されるステム・ループ構造ではなく、シュードノット構造をとる。

【 0 1 4 7 】

R N A ガイド型ヌクレアーゼをコードする核酸

R N A ガイド型ヌクレアーゼ、例えば、C a s 9、C p f 1、またはその機能性断片をコードする核酸は、本明細書で提供される。R N A ガイド型ヌクレアーゼをコードする例示的な核酸は、以前に記載されている (例えば、C o n g e t a l . , S c i e n c e . 2 0 1 3 F e b 1 5 ; 3 3 9 (6 1 2 1) : 8 1 9 - 2 3 (“ C o n g 2 0 1 3 ”) ; W a n g e t a l . , P L o S O n e . 2 0 1 3 D e c 3 1 ; 8 (1 2) : e 8 5 6 5 0 (“ W a n g 2 0 1 3 ”) ; M a l i 2 0 1 3 ; J i n e k 2 0 1 2 を参照のこと)。

【 0 1 4 8 】

いくつかの場合では、R N A ガイド型ヌクレアーゼをコードする核酸は、合成の核酸配列であり得る。例えば、合成の核酸分子は、化学的に改変することができる。ある種の実施態様では、R N A ガイド型ヌクレアーゼをコードする m R N A は、次の性質の 1 つまたは複数 (例えばすべて) を有することとなる：キャップ形成され得る；ポリアデニル化され得る；および 5 - メチルシチジンおよび / またはシュードウリジンで置換され得る。

【 0 1 4 9 】

合成の核酸配列はまた、コドン最適化することができる。例えば、少なくとも 1 つの非共通コドンまたはそれほど共通でないコドンが、共通コドンによって置き換えられている。例えば、合成の核酸は、例えば本明細書に記載した、例えば哺乳類の発現系における発現のために最適化された、最適化されたメッセンジャー m R N A の合成を誘導することが

10

20

30

40

50

できる。コドン最適化された Cas 9 コード配列の例は、国際公開第 2016/073990 号パンフレット(“Cotta-Ramusino”)に示される。

【0150】

さらにまたはあるいは、RNA ガイド型ヌクレアーゼをコードする核酸は、核局在化配列(NLS)を含むことができる。核局在化配列は、当技術分野で公知である。

【0151】

ガイドRNA(gRNA)分子

用語「ガイドRNA」および「gRNA」は、Cas 9 または Cpf 1 などの RNA ガイド型ヌクレアーゼの、細胞内のゲノムまたはエピソーム配列などの標的配列に対する特異的な結合(または「ターゲティング」)を促進するあらゆる核酸を指す。gRNA は、単分子(単一の RNA 分子を含み、あるいはキメラとも呼ばれる)またはモジュール式(2 つ以上、典型的には 2 つの別の RNA 分子(crRNA および tracrRNA など)を含み、これらは通常、例えば二本鎖化によって互いに結合されている)であり得る。gRNA およびその構成部分は、文献を通して、例えば、Briner et al. (参照により組み込まれる Molecular Cell 56(2), 333-339, October 23, 2014 (“Briner”)) に、また(and) Cotta-Ramusino に記載されている。

【0152】

細菌および古細菌では、II 型 CRISPR 系は一般に、Cas 9 などの RNA ガイド型ヌクレアーゼタンパク質、外来配列に相補的である 5' 領域を含む CRISPR RNA(crRNA)、および crRNA の 3' 領域に相補的でありこれと二本鎖を形成する 5' 領域を含むトランス活性化型 crRNA(tracrRNA)を含む。いかなる理論に拘泥されることも意図しないが、この二本鎖が、Cas 9 / gRNA 複合体の形成を容易にする - Cas 9 / gRNA 複合体の活性に必要である - ことが考えられる。II 型 CRISPR 系は、遺伝子編集に使用するために適合されたので、ある非限定的な例では、crRNA(その 3' 末端)および tracrRNA(その 5' 末端)の相補的領域をつなぐ、4 つのヌクレオチド(例えば GAAA)「テトラループ」または「リンカー」配列を用いて、crRNA と tracrRNA が連結されて単一の単分子またはキメラガイドRNAになる可能性があることが発見された。(Mali et al. Science, 2013 Feb 15; 339(6121): 823-826 (“Mali 2013”); Jiang et al. Nat Biotechnol, 2013 Mar; 31(3): 233-239 (“Jiang”); および Jinek et al., 2012 Science Aug. 17; 337(6096): 816-821 (“Jinek 2012”)(これらをすべて参照により本明細書に組み込む)。

【0153】

ガイドRNAは、単分子かモジュール式かにかかわらず、編集が所望される細胞のゲノム内のDNA配列などの標的配列内の標的ドメインに完全にまたは部分的に相補的である「ターゲティングドメイン」を含む。ターゲティングドメインは、限定はされないが、「ガイド配列」(参照により本明細書に組み込まれる Hsu et al., Nat Biotechnol, 2013 Sep; 31(9): 827-832 (“Hsu”))、「相補性領域」(Cotta-Ramusino)、「スペーサー」(Briner)、および総称的な「crRNA」(Jiang)を含めて、文献中で様々な名前と呼ばれている。与えられた名前とは関係なく、ターゲティングドメインは、典型的には、長さが 10 ~ 30 ヌクレオチドであり、ある種の実施態様では、長さが 16 ~ 24 ヌクレオチド(例えば、長さが 16、17、18、19、20、21、22、23、または 24 ヌクレオチド)であり、Cas 9 gRNA の場合には 5' 末端にまたはその近くに、Cpf 1 gRNA の場合には 3' 末端にまたはその近くにある。

【0154】

ターゲティングドメインに加えて、gRNAは、典型的には(以下に論じる通り、必ずしもというわけではないが)、gRNA / Cas 9 複合体の形成または活性に影響を与え

10

20

30

40

50

る可能性がある複数のドメインを含む。例えば、上に言及した通り、gRNAの第1と第2の相補性ドメインによって形成された二本鎖化された構造(リピート:アンチリピート二本鎖とも呼ばれる)は、Cas9の認識(REC)ローブと相互作用し、Cas9/gRNA複合体の形成を媒介することができる。(どちらも参照により本明細書に組み込まれるNishimasu et al., Cell 156, 935-949, February 27, 2014 (“Nishimasu 2014”)およびNishimasu et al., Cell 162, 1113-1126, August 27, 2015 (“Nishimasu 2015”)。第1のおよび/または第2の相補性ドメインが、RNAポリメラーゼによって終結シグナルと認識される可能性がある1つまたは複数のポリ-Aトラクトを含有する可能性があることに留意すべきである。したがって、第1および第2の相補性ドメインの配列は、場合によっては、これらのトラクトを除去し、例えば、Brinerに記載されている通りのA-G交換、またはA-U交換の使用を介する、gRNAの完全なインビトロ転写を促進するために改変される。第1および第2の相補性ドメインに対するこれらのおよび他の同様の改変は、本開示の範囲内である。

【0155】

第1および第2の相補性ドメインと共に、Cas9 gRNAは典型的には、インビトロでは必ずしもというわけではないがインビボでのヌクレアーゼ活性に関与する、2つ以上の追加の二本鎖化された領域を含む(Nishimasu 2015)。第2の相補性ドメインの3'部分の近くの第1のステム・ループ第一(one)は、「近位ドメイン」(Cotta-Ramusino)、「ステム・ループ1」(Nishimasu 2014 and 2015)、および「ネクサス」(Briner)と様々に呼ばれる。gRNAの3'末端の近くに、1つまたは複数の追加のステム・ループ構造が一般に存在し、その数は、種によって様々である:化膿レンサ球菌(S. pyogenes) gRNAは典型的には、2つの3'ステム・ループ(リピート:アンチリピート二本鎖を含む合計4つのステム・ループ構造に対して)を含むのに対して、黄色ブドウ球菌(S. aureus)および他の種は、1つ(合計3つのステム・ループ構造に対して)しか有しない。種によって組織化された、保存されているステム・ループ構造(および、より一般的にはgRNA構造)の説明は、Brinerによって提供されている。

【0156】

前述の説明は、Cas9と共に使用するためのgRNAに焦点を合わせてきたが、この時点までに記載されているものといくつかの点で異なるgRNAを利用する、他のRNAガイド型ヌクレアーゼが、発見または発明されている(または将来、される可能性がある)ことを認識すべきである。例えば、Cpf1(「プレボテラ(Prevotella)およびフランシスセラ(Francisella)由来のCRISPR 1(CRISPR from Prevotella and Francisella 1)»)は、機能するためにtracrRNAを必要としない、最近発見されたRNAガイド型ヌクレアーゼである。(参照により本明細書に組み込まれるZetsche et al., 2015, Cell 163, 759-771 October 22, 2015 (“Zetsche I”)。Cpf1ゲノム編集システムに使用するためのgRNAは、一般に、ターゲティングドメインおよび相補性ドメイン(代わりに「ハンドル」とも呼ばれる)を含む。Cpf1と共に使用するためのgRNAでは、ターゲティングドメインが通常、Cas9 gRNAに関して上に記載した5'末端ではなく、3'末端にまたはその近くに存在する(ハンドルは、Cpf1 gRNAの5'末端にまたはその近くにある)ことにも留意すべきである。

【0157】

しかし、当業者は、異なる原核生物種由来のgRNAの間に、または、Cpf1 gRNAとCas9 gRNAの間に、構造の違いが存在する可能性はあるが、gRNAが動作する原理は一般に一致していることを認識するであろう。この動作の不変性から、gRNAは、そのターゲティングドメイン配列によって大まかに定義することができ、当業者は、所与のターゲティングドメイン配列を、単分子もしくはキメラgRNA、または1つ

10

20

30

40

50

または複数の化学的改変および／または配列的改変（置換、追加のヌクレオチド、切り詰めなど）を含む gRNA を含めた、あらゆる好適な gRNA に組み込むことができることを認識するであろう。したがって、この開示における提示の節約のために、gRNA は、そのターゲティングドメイン配列に関してのみ記載することができる。

【0158】

より一般的には、当業者は、本開示のいくつかの態様が、複数の RNA ガイド型ヌクレアーゼを使用して実行することができるシステム、方法、および組成物に関することを認識するであろう。この理由から、そうではないと指定されない限り、gRNA という用語は、Cas9 または Cpf1 のある特定の種と適合性のある gRNA だけでなく、あらゆる RNA ガイド型ヌクレアーゼと共に使用することができるあらゆる好適な gRNA を包含すると理解されるべきである。実例として、用語 gRNA には、ある種の実施態様では、II 型または V 型または CRISPR 系などのクラス 2 CRISPR 系に存在するあらゆる RNA ガイド型ヌクレアーゼ、またはそれに由来するまたはそこから適合される RNA ガイド型ヌクレアーゼと共に使用するための gRNA が含まれ得る。

10

【0159】

連結

本明細書によって開示された方法のいくつかの実施態様は、核酸配列、例えば、核酸鋳型の切断された末端を、オリゴヌクレオチド捕捉プローブまたはその混合物と連結するステップを含む。いくつかの実施態様では、連結のステップは、核酸、例えば、DNA および／または RNA リガーゼ上で作用するリガーゼ酵素を使用して達成される。様々なこうしたリガーゼは、当技術分野で公知であり、これらの多くは、市販品として入手可能である。

20

【0160】

本明細書によって開示された方法の種々の実施態様において使用することができるリガーゼの例としては、限定はされないが、T4 DNA リガーゼ、T3 DNA リガーゼ、T7 DNA リガーゼ、および E. coli リガーゼが挙げられる。

【0161】

選択されるリガーゼの種類は、切断組成物中に存在する切断される末端および／またはオリゴヌクレオチド捕捉プローブの捕捉末端の種類に依存し得る。

【0162】

例えば、切断組成物における切断された末端が、平滑末端を含む、または連結前に（例えば、本明細書に記載した通りの平滑化の追加のステップ中に）平滑化される切断された末端を含むならば、平滑末端を連結するのに適したリガーゼを選択することができる。

30

【0163】

例えば、切断組成物における切断された末端が、連結前に（例えば、本明細書に記載した通りの平滑化の追加のステップ中に）平滑化されない突出（「付着末端」）を含むならば、粘着末端を連結するのに適したリガーゼを選択することができる。

【0164】

いくつかのリガーゼは、平滑末端と突出を有する末端の両方に対して良く働き、これらのリガーゼのいずれかを、本開示の方法において使用することができる。さらに、2 つ以上のリガーゼのあらゆる組み合わせも、連結ステップ中に使用することができる。

40

【0165】

連結産物の分析

いくつかの実施態様では、本明細書に記載した複数の連結産物が分析される。いくつかの実施態様では、複数の連結産物は、増幅される。複数の連結産物が 1 つまたは複数の検出配列を含むいくつかの実施態様では、1 つまたは複数の検出配列を認識する増幅プライマーを使用することができる。

【0166】

いくつかのこうした実施態様では、増幅産物が分析される。例えば、いくつかの実施態様では、方法は、連結産物のレベルを決定するステップをさらに含む。決定されるレベル

50

は、絶対的および／または相対的であり得る。

【0167】

いくつかの実施態様では、方法は、連結産物の相対的存在量を算出するステップをさらに含む。この分析は、連結産物および／またはその増幅産物の核酸配列決定を含むことができる。非限定的な例として、次世代（ハイスループットシーケンシングとしても公知である）を実施することができる。

【0168】

いくつかの実施態様では、各ヌクレオチドが、例えば少なくとも7、少なくとも10、少なくとも15、少なくとも20よりも大きい、またはさらに大きい深度で、配列決定プロセス中に数回読み取りされることを意味する、ディープシーケンシングが実施される。ここでは、深度（D）は、

$$D = N \times L / G \text{ (式1)}$$

（式中、Nは、読み取りの数であり、Lは、実際のゲノムの長さであり、Gは、配列決定されるポリヌクレオチドの長さである）として算出される。

【0169】

いくつかの実施態様では、連結産物および／またはその増幅産物の少なくともいくつかを分析するために、サンガー配列決定が使用される。

【0170】

いくつかの実施態様では、配列決定データは、バイオインフォマティクス、例えば、図12に示された1つまたは複数のステップを使用して処理される。例えば、配列決定のリードは、最初に、適切な配列決定バーコードを使用して、その実行における異なる実験に逆多重化することができる。いくつかの場合、1、2、または3個までのミスマッチが、各バーコードにおいて許容される。次いで、さらなるダウンストリーム解析のために、各リードについてのUMIを抽出することができる。いくつかの方法では、次いで、（例えば、Martin, EMBnet. journal 17:10-12 (2011)に記載されている通りに）cutadaptなどの公知の方法を使用して、配列決定リードから、アダプター（例えば、3'アダプター）を取り除くことができる。いくつかの場合、1つまたは複数の参照アンプリコンが作製される。例えば、予想された切断部位を使用して、予想された染色体再構成（例えば、野生型、大きい欠失、逆位、転座など）を有する参照アンプリコンを構築することができる。参照アンプリコンは、あらゆる予測される遺伝子編集事象ならびに「オフターゲット」事象を含むことができる。いくつかの実施態様では、参照アンプリコンは、ベクターまたはプラスミド配列を含むことができる。ベクターまたはプラスミドからの配列を含む参照アンプリコンはまた、あらゆる予測される遺伝子編集事象および／またはあらゆる「オフターゲット」事象を含むことができる。

【0171】

いくつかの方法では、次いで、対形成させたリードを、例えば配列アライナー（aligner）を使用して、すべてまたはほとんどすべての参照アンプリコンに対して包括的に整列させることができる。ある例示的な配列アライナーは、bowtie2である（例えば、Langmead et al., Nat. Methods 9:357-9 (2012)に記載されている通り）。さらにまたはあるいは、SAMtoolsなどの好適なツールを使用して、ソートされたbamファイルを作成および／またはインデックス付けることができる（例えば、Li et al., Bioinformatics 25:2078-9 (2009)に記載されている通り）。いくつかの方法では、予測される接合部（例えば、10、11、12、13、14、15、16、17、18、19、または20bp）の周囲のウィンドウを完全にカバーする配列リードを抽出し、固有のUMIの総数を計数することができる。いくつかの方法では、参照アンプリコンに対してマッピングできない配列リードを抽出し、例えば、適切なバックグラウンド参照ゲノムに対してbowtie2を使用して、包括的にマッピングすることができる。

【0172】

ヌクレアーゼ切断を評価するための、本明細書に記載した実例プロセスの少なくとも一

部は、ハードウェアまたはハードウェアとソフトウェアの組み合わせを含む 1 つまたは複数のコンピュータシステムを使用して実行することができる。例えば、コンピュータシステムは、自動化された要素の動作を制御するための、様々な点に位置する様々な制御装置および/または処理装置を含むことができる。中央コンピュータは、様々な制御装置または処理装置間の動作を協調させることができる。中央コンピュータ、制御装置、および処理装置は、様々な自動化された要素の制御および協調を実行するために、様々なソフトウェアルーチンを遂行することができる。

【0173】

実例プロセスは、少なくともある程度、1 つまたは複数のコンピュータプログラム製品、例えば、1 つまたは複数の情報担体に有形的に組み込まれた 1 つまたは複数のコンピュータプログラム、例えば、1 つまたは複数のデータ処理装置による遂行のための、または当該装置の動作を制御するための、1 つまたは複数の非一過性の機械可読媒体など、（例えば、プログラム可能プロセッサ、コンピュータ、複数のコンピュータ、および/またはプログラム可能な論理構成要素）を使用して実行することができる。

10

【0174】

コンピュータプログラムは、コンパイルされるまたはインタプリタされる言語を含めた、あらゆる形式のプログラミング言語で書くことができ、スタンドアロンプログラムとしてまたはモジュールとして、コンポーネント、サブルーチン、またはコンピューティング環境における使用に適した他のユニットを含めた、あらゆる形式でデプロイすることができる。1 つのコンピュータ上でまたは 1 つのサイトで複数のコンピュータ上で遂行されることとなるコンピュータプログラムを、デプロイする、または複数のサイトに分散させ、ネットワークによって相互接続することができる。

20

【0175】

実例プロセスのすべてまたは一部を実行することに関連する動作は、本明細書に記載した機能を実施するために 1 つまたは複数のコンピュータプログラムを遂行する 1 つまたは複数のプログラム可能プロセッサによって実施することができる。実例プロセスのすべてまたは一部は、特殊目的の論理回路、例えば、FPGA（現場でプログラム可能なゲートアレイ（*field programmable gate array*））および/またはASIC（特定用途向け集積回路（*application-specific integrated circuit*））を使用して実行することができる。

30

【0176】

コンピュータプログラムの遂行に適したプロセッサとしては、例として、一般目的と特殊目的のマイクロプロセッサ、およびあらゆる種類のデジタルコンピュータのあらゆる 1 つまたは複数のプロセッサが挙げられる。一般に、プロセッサは、読み出し専用記憶領域またはランダムアクセス記憶領域またはこれらの両方からの命令およびデータを受け取ることとなる。コンピュータ（サーバーが含まれる）の要素には、命令を遂行するための 1 つまたは複数のプロセッサ、および命令およびデータを保管するための 1 つまたは複数の記憶領域装置が含まれる。一般に、コンピュータはまた、データを受け取る、データを移す、またはこれらの両方のために、1 つまたは複数の機械可読記憶媒体、例えばデータを保管するための大容量記憶装置、例えば、磁気、光磁気ディスク、または光ディスクを含む、またはそれと動作可能に接続されることとなる。コンピュータプログラム命令およびデータを組み込むのに適した機械可読記憶媒体には、例として、半導体記憶領域装置、例えば、EPROM、EEPROM、およびフラッシュ記憶領域装置；磁気ディスク、例えば、内蔵ハードディスクまたはリムーバブルディスク；光磁気ディスク；ならびにCD-ROMおよびDVD-ROMディスクを含めた、すべての形態の不揮発性の記憶領域が含まれる。

40

【0177】

本明細書に言及したすべての刊行物、特許出願、特許、および他の参考文献の全体を、参照によって組み込む。さらに、材料、方法、および実施例は、例示に過ぎず、限定的であることを意図しない。そうではないと定義されない限り、本明細書で使用されるすべて

50

の技術および科学用語は、こ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。本明細書に記載したものと類似のまたは等価な方法および材料を、本発明の実施または実験において使用することができるが、好適な方法および材料を、本明細書に記載する。

【0178】

この開示を、次の実施例によって、さらに例示する。これらの実施例は、例示目的に限って提供される。これらは、決してこの開示の範囲または内容を限定するものと解釈されるべきではない。

【実施例】

【0179】

実施例1：UDiTASを使用する標的遺伝子切断の分析

部位特異的ヌクレアーゼおよび種々のガイドRNAで処理したK562細胞由来のゲノムDNA試料を配列決定するための本開示の方法において、トランスポゾンを使用した。第1の実験では、(1)G1遺伝子的一部分に対して相補的な配列を含む第1の固定プライマーと、(2)トランスポゾンに相補的なヌクレオチド配列を含む選択的プライマーとを含むプライマー対を使用した。第2の実験では、(1)G2遺伝子的一部分に対して相補的な配列を含む第1の固定プライマーと、(2)トランスポゾンに相補的なヌクレオチド配列を含む選択的プライマーとを含むプライマー対を使用した。これらのプライマーを、図4に示す(固定プライマー「G1__F1」および「G2__F1」；選択的プライマー「Seq15__G1__F__UDITASdev」および「Seq16__G2__F__UDITASdev」)。その後、DNA断片を増幅および配列決定した。

【0180】

図5は、G1またはG2標的上であった配列リード(%)を示す。図5に示す通り、MiSeqリードの78~93%は、G1またはG2標的を正確に配列決定した。これは、標的遺伝子を、その標的遺伝子に特異的なたった1つの固定プライマーを使用して、成功裏に配列決定することができることを実証する。図6は、GenomeViewにおける配列リードを示す。

【0181】

実施例2：インデルを検出するためにUDiTASを使用すること

UDiTASがモデルゲノム標的についての小さいインデルの割合を特定する能力を検定した。モデルゲノム標的をトランスポゾンと接触させた後、(1)モデルゲノム標的的一部分に相補的な配列を含む第1の固定プライマーと、(2)トランスポゾンに相補的なヌクレオチド配列を含む選択的プライマーとを含むプライマー対を使用して、DNA断片を増幅した。UDiTAS方法を、2つの公知の方法 - アンブリコン配列決定およびT7E1配列決定と比較した。図7に示す通り、UDiTASがインデルを検出する能力は、標的配列決定およびT7E1アッセイと良く相関していた。

【0182】

実施例3：染色体内再構成を検出するためにUDiTASを使用すること

ゲノムDNAを部位特異的ヌクレアーゼおよびデュアルガイドで処理することは、複数の結果の可能性を得る可能性がある。図8に示す通り、デュアルガイドは、欠失、一方または両方の鎖におけるインデル、および逆位をもたらす可能性がある。UDiTAS方法が、こうした染色体内再構成を検出するために使用できるかどうかを決定するために、野生型HEK293T細胞由来のまたは安定なHEK293Tクローン由来のゲノムDNAを使用した。安定なクローンのゲノムDNAは、そのゲノムDNAの66%における欠失、およびそのゲノムDNAの33%における逆位を含む(図9、左側参照)。野生型由来とクローン由来のゲノムDNAを、様々な比率で混合し、本明細書に記載した通りのUDiTASによる配列決定にかけた。配列決定の結果を、ddPCR配列決定と比較した。

【0183】

図9(右側)に示す通り、混合物における野生型ゲノムDNAの%が、100%から0%まで低下するにつれて、UDiTASが検出できる野生型ゲノムDNAの頻度は低下し

10

20

30

40

50

、混合物における欠失および逆位を有するゲノムDNAの頻度は増大した。対照的に、ddPCRは、野生型および欠失を有するゲノムDNAを検出したのに対し、この方法は、逆位を有するゲノムDNAを検出できなかった。

【0184】

UDiTASが染色体内再構成を検出する能力はまた、2つの異なるプライマーを使用して、様々な割合の野生型ゲノムDNA、逆位を含むゲノムDNA、および大きい欠失を含むゲノムDNAを含むゲノムDNA試料を使用して評価した(図10参照)。図10に示す通り、UDiTASは、1)わずか0.1%の大きい欠失、または2)実質的に1%未満の逆位を有する試料における特定の種類の再構成の割合を正確に決定することができた。(図10の右側参照)。異なるプライマーを使用して、UDiTASは、より低いレベルの該当するアレルについてはより低い精度ではあったが、さらに低いレベルの逆位および欠失を検出することができた。(図10の左側を参照のこと)。

10

【0185】

実施例4：パイオインフォマティクス・パイプライン

ある例示的な方法では、配列決定データを、パイオインフォマティクス・パイプラインを使用して処理した。特殊化されたステップのための追加ソフトウェアと呼ばれるパイソン(python)コードを使用して、解析パイプラインを構築した。このパイオインフォマティクス・パイプラインを、図12に模式的に示し、また、これは、次のステップからなる。

【0186】

20

逆多重化

配列決定リードを、最初に、適切な配列決定バーコード(各バーコードにおいて1つまでのミスマッチが可能である)を使用して、その実行における異なる実験に逆多重化した。各リードについてのUMIを、さらなるダウストリーム解析のために抽出した。

【0187】

トリミング

3'アダプターを、cutadapt(Martin, EMBnet.journal 17:10-12(2011)), version 1.9.1を使用して取り除いた。

【0188】

参照アンプリコンの作製

30

予想された切断部位を使用して、予想された染色体再構成：野生型、大きい欠失、重複、逆位、転座、小さいインデルなどを有する参照アンプリコンを構築した。参照アンプリコンは、あらゆる予測される遺伝子編集事象、例えば「オフターゲット」事象を含むことができる。いくつかの実施態様では、参照アンプリコンは、ベクターまたはプラスミド配列を含むことができる。ベクターまたはプラスミドからの配列を含む参照アンプリコンはまた、1つまたは複数の予測される遺伝子編集事象、例えば「オフターゲット」事象を含むことができる。

【0189】

アラインメント

次いで、対形成させたリードを、bowtie2(Langmead et al., Nat. Methods 9:357-9(2012)), version 2.1.0を使用して、すべての参照アンプリコンに対して包括的に整列させた。最後に、SAMtools(Li et al., Bioinformatics 25:2078-9(2009)), version 1.3-5-g664cc5fを使用して、インデックスソートされたbamファイルを作成した。

40

【0190】

アラインメント解析

予測される接合部(15bp)の周囲のウィンドウを完全にカバーするリードを抽出し、固有のUMIの総数を計数した。

【0191】

50

最終のゲノムワイド解析

最後に、参照アンプリコンに対してマッピングできなかったリードを抽出し、適切なバックグラウンド参照ゲノムに対して *bowtie2* を使用して、包括的にマッピングした。

【0192】

実施例5：UDiTASを使用する標的遺伝子切断の分析

複雑な遺伝子編集事象の評価の、ある例示的な方法では、Gene 3 のイントロン内の領域を標的にするために、一对の黄色ブドウ球菌 (*S. aureus*) Cas9 (*SaCas9*) sgRNA を使用した。Gene 3 発現の低下は、いくつかのヒト疾患をもたらす。有害なイントロンスプライスドナー部位の除去が、Gene 3 の機能を回復させることが予測される。スプライスドナーの周囲の約 1000 塩基対を切断する単一のガイドRNA対、Gene 3 - ガイド1とGene 3 - ガイド2を、特徴付けのための内部スクリーニングから選択した。

10

【0193】

U-2 OS細胞に、SaCas9を発現するプラスミドと共に、sgRNA Gene 3 - ガイド1とGene 3 - ガイド2を発現する線状DNA断片を導入した。3日後、ゲノムDNAを単離し、UDiTAS配列決定ライブラリを作製した。このライブラリを、MiSeq上で配列決定し、バイオインフォマティクス・パイプラインを介してデータを処理した(図12)。

【0194】

Gene 3 - ガイド2切断部位に隣接する標的化プライマーを使用して(図13)、予想された切断部位の周囲の編集および再構成の範囲を観察し(図15A~15D)、自動的に分類し、集計した(図14A)。小さいインデル事象をもたらす編集が、予想通り、約17%の割合で観察された。このインデルは、再構成前の修復経路活性に起因して生じた可能性がある。さらに、所望される約1.1 kb欠失からの接合部も、約40%で存在していた。特に、2つの切断部位間の約1.1 kb断片の逆位も約18%で観察され、欠失に匹敵していた。同一の切断部位での相同染色体または姉妹染色体間の転座を含めた、他の、より低い頻度の接合部も、約0.75%で観察された(図15D)。

20

【0195】

線形性、精度、およびLLoD

この方法の線形性、精度、および検出下限(Lower Level of Detection)(LLoD)を特徴付けるために、所望される大きい編集事象に対して、Gene 3イントロン野生型遺伝子座、欠失、および逆位を含有する安定な細胞株およびプラスミドを構築した。HEK293細胞は、Gene 3遺伝子座の3つのコピーを有し、クローンは、2つの欠失と1つの逆位を有していた。編集されたHEK293細胞株由来のゲノムDNA(gDNA)を、親の編集されていないゲノムDNAと混合して、5桁の対数レンジにわたる約1.1 kb欠失および逆位のタイトレーションをもたらした。欠失および逆位についてのUDiTAS測定された編集率に対する予想された率のプロットは、およそ0.1%までの優れた相関を示した(図14B)。このUDiTASプロトコルは、50 ngのインプットDNA(これは、およそ14,300のヒトハプローム(haplome)と同等である)を使用した。標本抽出の二項分布および約20%工程歩留まりと仮定すると、UDiTASライブラリにおける0.1%での2~3例の観察が予想された(95%の信頼度)。これは、観察された割合と一致していた(図16)。感度は、インプットDNAを増大させることによって、また、配列決定読み取り深度を高めることによって、どちらも、解析における固有のUMIの数を増加させるために必要とされるので、増大した。ゲノムに対するオンターゲットマッピング率は、>95%であり、これは、このプロセスが頑強かつ生産的であることを示していた(図17)。

30

40

【0196】

UDiTASの線形性をさらに実証するために、また、これをAMP-Seqと比較するために、イントロン野生型遺伝子座、大きい欠失、および逆位を含有するように合成した参照プラスミドを分析した。

50

【 0 1 9 7 】

【 表 1 】

表1:この実施例で使したプラスミド

プラスミド名	説明
PLA380	Gene1およびGene2 SpyガイドのためのIVT PCR鋳型
PLA379	pUC57_Amp_Gene3_SNP _s 1
PLA370	pUC57_Amp_Gene3_大きい_逆位_SNP _s _1
PLA367	pUC57_Amp_Gene3_大きい_欠失_SNP _s _1
PLA371	pUC57_Amp_Gene3_大きい_逆位_SNP _s _2
PLA368	pUC57_Amp_Gene3_大きい_欠失_SNP _s _2
PLA372	pUC57_Amp_Gene3_大きい_逆位_SNP _s _3
PLA369	pUC57_Amp_Gene3_大きい_欠失_SNP _s _3
PLA377	pUC57_Amp_Gene2_SNP _s 1
PLA378	pUC57_Amp_Gene1_SNP _s 1
PLA361	pUC57_Amp_Gene2_Gene1_SNP _s _1
PLA362	pUC57_Amp_Gene2_Gene1_SNP _s _2
PLA363	pUC57_Amp_Gene2_Gene1_SNP _s _3
PLA364	pUC57_Amp_Gene2_Gene1_SNP _s _4
PLA365	pUC57_Amp_Gene2_Gene1_SNP _s _5
PLA366	pUC57_Amp_Gene2_Gene1_SNP _s _6
PLA13	pAF003 STITCHR骨格プラスミド

【 0 1 9 8 】

10

20

30

40

50

【表 2 - 1】

表2:この実施例で使用したオリゴ

オリゴ名	説明	配列
OLI7076	Gene2 IVTのためのフォワードプライマー	CACCGCTAGCTAATACGACTCACTATA GGCCACGGAGCGAGACATCTGTTTTAG AGCTAGAAATA (配列番号1)
OLI7077	Gene1 IVTのためのフォワードプライマー	CACCGCTAGCTAATACGACTCACTATA GCTGGTACACGGCAGGGTCAGTTTTAG AGCTAGAAATA (配列番号2)
OLI7078	IVT鋳型のための共通リバースプライマー	TTTTTTTTTTTTTTTTTTTTTGCACCGACT CGGTGCCACTTTTTCAAGTTGATA (配列 番号3)
OLI6062	Gene3ガイド2のためのUDiTaS およびAMP-seq遺伝子特異的 プライマー	GTGACTGGAGTTCAGACGTGTGCTCTTC CGATCTGGACCATGGATGCACTCTGTA AATTCTCAT (配列番号4)
OLI6256	Gene2のためのUDiTaSおよび AMP-seq遺伝子特異的プライ マー	GTGACTGGAGTTCAGACGTGTGCTCTTC CGATCTGCATGCCTTCTTAAACATCACG AGACTCTAA (配列番号5)
OLI6259	Gene1のためのUDiTaSおよび AMP-seq遺伝子特異的プライ マー	GTGACTGGAGTTCAGACGTGTGCTCTTC CGATCTGCACTGTTGCTCTTGAAGTCCA TAGACCTC (配列番号6)
OLI6380	UDiTaSアダプタートップオリゴ i5_N501_UMI_Tn5-A	AATGATACGGCGACCACCGAGATCTAC ACTAGATCGCNNNNNNNNNTCGTCGG CAGCGTCAGATGTGTATAAGAGACAG (配列番号7)

【 0 1 9 9 】

10

20

30

40

50

【表 2 - 2】

OLI6381	UDiTaSアダプタートップオリゴ i5_N502_UMI_Tn5-A	AATGATACGGCGACCACCGAGATCTAC ACCTCTCTATNNNNNNNNNNNTCGTCGG CAGCGTCAGATGTGTATAAGAGACAG (配列番号8)
OLI6382	UDiTaSアダプタートップオリゴ i5_N503_UMI_Tn5-A	AATGATACGGCGACCACCGAGATCTAC ACTATCCTCTNNNNNNNNNNNTCGTCGG CAGCGTCAGATGTGTATAAGAGACAG (配列番号9)
OLI6383	UDiTaSアダプタートップオリゴ i5_N504_UMI_Tn5-A	AATGATACGGCGACCACCGAGATCTAC ACAGAGTAGANNNNNNNNNNTCGTCG GCAGCGTCAGATGTGTATAAGAGACAG (配列番号10)
OLI6384	UDiTaSアダプタートップオリゴ i5_N505_UMI_Tn5-A	AATGATACGGCGACCACCGAGATCTAC ACGTAAGGAGNNNNNNNNNNNTCGTCG GCAGCGTCAGATGTGTATAAGAGACAG (配列番号11)
OLI6385	UDiTaSアダプタートップオリゴ i5_N506_UMI_Tn5-A	AATGATACGGCGACCACCGAGATCTAC ACACTGCATANNNNNNNNNNNNTCGTCGG CAGCGTCAGATGTGTATAAGAGACAG (配列番号12)
OLI6386	UDiTaSアダプタートップオリゴ i5_N507_UMI_Tn5-A	AATGATACGGCGACCACCGAGATCTAC ACAAGGAGTANNNNNNNNNNNNTCGTCG GCAGCGTCAGATGTGTATAAGAGACAG (配列番号13)
OLI6387	UDiTaSアダプタートップオリゴ	AATGATACGGCGACCACCGAGATCTAC ACCTAAGCCTNNNNNNNNNNNTCGTCGG CAGCGTCAGATGTGTATAAGAGACAG (配列番号14)

【 0 2 0 0 】

【表 2 - 3】

Tn5-A ボトム	UDiTaSアダプターボトムオリゴ	[Phos]CTGTCTCTTATACA[ddC] (配列番号 15)
OLI5589	UDiTaSおよびAMP-seqラウン ド1および2 PCRプライマー P5/i5	AATGATACGGCGACCACCGAGATCTAC AC (配列番号16)
OLI5639	UDiTaSおよびAMP-seqラウン ド2 PCRプライマー i7_N701_SBS12	CAAGCAGAAGACGGCATAACGAGATAGC GGAATGTGACTGGAGTTCAGACGTGT (配列番号17)
OLI5640	UDiTaSおよびAMP-seqラウン ド2 PCRプライマー i7_N702_SBS12	CAAGCAGAAGACGGCATAACGAGATGAT CATGCGTGACTGGAGTTCAGACGTGT (配列番号18)
OLI5641	UDiTaSおよびAMP-seqラウン ド2 PCRプライマー i7_N703_SBS12	CAAGCAGAAGACGGCATAACGAGATAA GACGGAGTGACTGGAGTTCAGACGTGT (配列番号19)
OLI5642	UDiTaSおよびAMP-seqラウン ド2 PCRプライマー i7_N704_SBS12	CAAGCAGAAGACGGCATAACGAGATCGA GTCCTGTGACTGGAGTTCAGACGTGT (配列番号20)
OLI5643	UDiTaSおよびAMP-seqラウン ド2 PCRプライマー i7_N705_SBS12	CAAGCAGAAGACGGCATAACGAGATTCC TCAGGGTGACTGGAGTTCAGACGTGT (配列番号21)
OLI5644	UDiTaSおよびAMP-seqラウン ド2 PCRプライマー i7_N706_SBS12	CAAGCAGAAGACGGCATAACGAGATGTA CGGATGTGACTGGAGTTCAGACGTGT (配列番号22)
OLI5645	UDiTaSおよびAMP-seqラウン ド2 PCRプライマー i7_N707_SBS12	CAAGCAGAAGACGGCATAACGAGATCAT CTCTCGTGACTGGAGTTCAGACGTGT (配列番号23)

【 0 2 0 1 】

10

20

30

40

50

【表 2 - 4】

OLI5646	UDiTaSおよびAMP-seqラウンド2 PCRプライマー i7_N710_SBS12	CAAGCAGAAGACGGCATACGAGATGTC GGAGCGTGACTGGAGTTCAGACGTGT (配列番号24)
OLI5647	UDiTaSおよびAMP-seqラウンド2 PCRプライマー i7_N711_SBS12	CAAGCAGAAGACGGCATACGAGATACG GAGAAGTGACTGGAGTTCAGACGTGT (配列番号25)
OLI5648	UDiTaSおよびAMP-seqラウンド2 PCRプライマー i7_N712_SBS12	CAAGCAGAAGACGGCATACGAGATAG GAGATGGTGACTGGAGTTCAGACGTGT (配列番号26)
OLI5649	UDiTaSおよびAMP-seqラウンド2 PCRプライマー i7_N714_SBS12	CAAGCAGAAGACGGCATACGAGATAGT ACTCGGTGACTGGAGTTCAGACGTGT (配列番号27)
OLI5650	UDiTaSおよびAMP-seqラウンド2 PCRプライマー i7_N715_SBS12	CAAGCAGAAGACGGCATACGAGATGG ACTCTAGTGACTGGAGTTCAGACGTGT (配列番号28)
OLI2909	インデックス1-AMP-seqトップアダプター	AATGATACGGCGACCACCGAGATCTAC ACACTGCATANNWNNWNNACACTCTTT CCCTACACGACGCTCTTCCGATC*T (配列番号29)
OLI2910	インデックス1-AMP-seqトップアダプター	AATGATACGGCGACCACCGAGATCTAC ACAAGGAGTANNWNNWNNACACTCTT TCCCTACACGACGCTCTTCCGATC*T (配列番号30)
AMP-seqボトムアダプター	AMP-seqボトムアダプター	[Phos]GATCGGAAGAGC*C*A (配列番号31)

【0202】

Gene 1 / Gene 2 均衡型転座または Gene 3 構造多型を含有するプラスミドを合成し、これは、配列決定後にプラスミドを特定するために、インサート内に、操作された固有の SNP を含有していた。約 2,200 ~ 約 714,000 ゲノム当量の範囲のプラスミドを、マウスゲノム DNA に加え、UDiTaS および AMP - Seq 方法を介して処理した。UDiTaS と AMP - Seq の両方について、検出されたプラスミドの数に対するインプットプラスミドの数をプロットした。UDiTaS は、最も低い希釈までの優れた線形性を示した (図 18 A および 18 B)。UDiTaS は、AMP - Seq と比較した場合に、より高い精度と線形性を示した (図 18 A および 18 B)。

【0203】

UDiTaS ライブラリは、同じ DNA インプット材料を与えた AMP - Seq と比較すると、より線形性が高く、より高いダイナミックレンジを有していた。これは、AMP - Seq の切断およびアダプター連結と比較して、より効率的なタグメンテーションプロセスに起因していた。また、UDiTaS が必要としているのは、AMP - Seq と比較して、全収量を高める、より少なく、合理化された処理ステップであった。

【0204】

担体マウスゲノムDNAがUDiTAS反応に影響を与えていないことを確実にするために、いかなる担体DNAも存在しないプラスミドを用いて、UDiTAS反応の追加のセットを実行した。図19Aは、野生型、大きい欠失、または大きい挿入配列を含有するGene3プラスミド(PLA379、PLA367、およびPLA370)を示す。図19Bは、均衡型転座を含有するGene1/Gene2プラスミド(PLA377、PLA378、PLA365、およびPLA366)を示す。プラスミドは希釈し、プラスミド混合物を、UDiTASおよび解析パイプラインを介して処理した。ある構造多型についての測定された頻度に対する、所与の構造多型についての予想された頻度をプロットした(プロットされた線によって $x = y$ が示される場合)。4つすべてのプライマーを用いて、どちらの遺伝子座についても、精度と線形性は優れており、約0.01%~0.1%のLLODは、このプロセスの頑強性をさらに示した。

10

【0205】

染色体間転座率

UDiTASはまた、染色体間転座率を測定した。これを示すために、Gene1およびGene2遺伝子を標的にするsgRNAとの、2つの化膿レンサ球菌(*S. pyogenes*) Cas9(SpCas9)リボ核タンパク質(RNP)複合体を、活性化されたヒトCD4⁺T細胞にヌクレオフェクト(nucleofect)させた。これらの細胞からのUDiTASライブラリを、各ガイド部位に隣接するプライマーを使用して調製した。10例の末端連結結果の可能性のうち、2つのターゲティングプライマーを使用して、7例を測定することが可能であった。7つのすべての結果が検出され、定量化された。これらには、切断部位でのNHEJ(非相同末端連結)編集、ならびに相同染色体または姉妹染色体間の、また別個の無関係の染色体間の、均衡型の、無動原体性の、および二動原体性の融合が含まれていた(図14C)。いくつかの研究が、遺伝子編集の状況での転座の特徴付けを公表しているが、どれも、すべての事象を測定していない(Huet al., Nat. Protoc. 2016; 11: 853-71; Jiang et al., Sci. Rep. 2016; 6: 21918; Ghezraoui et al. Mol. Cell. 2014; 55: 829-42; Frock et al., Nat. Biotechnol. 2015; 33: 179-186)。代わりに、これらの研究は、転座をつなぎ合わせることを必要とする転座分析に頼っており、より小さなインデルの割合という状況に沿うことはできなかった。この実施例は、インデルを伴う二重の遺伝子編集転座率の定量化を示した。この実施例はまた、各ガイドRNAに対してそれぞれ約82%および約91%というオンターゲットインデル編集と共に、約2.5%の染色体間転座率を示した(図14C)。Gene1およびGene2転座位置でのアッセイの線形性および約0.01%というLLODが、Gene3遺伝子座と同様の様式で、プラスミドを用いて特徴付けられ、転座事象を検出するためのUDiTASの高い感度の実証された(図18Aおよび図19B)。

20

30

【0206】

この実施例は、UDiTASが、小さいインデルと、大きい欠失、逆位、および転座などの、より大きな構造的再構成の同時測定を可能にする配列決定および解析方法論であることを実証した。UDiTAS方法は、頑強であり、拡大縮小可能であり、低い(50ng)インプットDNAを使用する場合にアンカープライマーを多重化するのに適していた。これは、試料が限られる場合に、候補となるオフターゲット編集部位のパネルのモニタリングを可能にした。さらに、ここに記載したカスタムなトランスポゾン、GUIDE-Seqなどのアンカープライマーと共にDNA剪断を利用する方法を改良する可能性を有する。UDiTAS方法はまた、正確かつ線形であった。UDiTASは、遺伝子編集技術の有効性の評価のために重要な、意図されるまたは意図されない大きい構造変化の正確な定量化を可能にする、重要な配列決定方法論である。

40

【0207】

実施例の方法

U-2 OSバルク編集トランスフェクション実験

50

U-2 OS細胞(ATCC)を、DMEM(高グルコース、Glutamaxおよびピルビン酸ナトリウム(ThermoFisher)、10%ウシ胎児血清を含有、1%ペニシリン/ストレプトマイシンを添加)中で維持した。細胞は、4D nucleofectorシステムを使用して、Lonzaヌクレオフェクションによって遺伝子導入した。簡単に言うと、250,000個の細胞に、CMVプロモーターによって駆動されるSaCas9を発現する1.5μgプラスミドpAF003と、U6プロモーターによって駆動されるgRNAを発現する500ngの線状DNA断片(250ngの各ガイド)を導入した。細胞を、SEキットおよびパルス符号DN-100を使用してヌクレオフェクトし、6ウェルプレートにプレATINGした。細胞を、ヌクレオフェクション後、3日間培養し、トランスフェクション技術の3連の反復物を共にブールした。ゲノムDNAは、製造業者の説明書に従ってAgencourt DNAdvanceキット(Beckman Coulter)を使用して単離した。

10

【0208】

Gene3遺伝子座でのHEK293細胞株作製

HEK293細胞(ATCC)を、DMEM(高グルコース、Glutamaxおよびピルビン酸ナトリウム(ThermoFisher)、10%ウシ胎児血清を含有、1%ペニシリン/ストレプトマイシンを添加)中で維持した。細胞は、製造業者の説明書に従ってMirus TransIT 293キットを使用して遺伝子導入した。簡単に言うと、120,000個の細胞を、トランスフェクションの24時間前に、24ウェルプレートのウェルに播種した。細胞に、CMVプロモーターによって駆動されるSaCas9を発現する750ngプラスミドpAF003と、U6プロモーターによって駆動されるgRNAを発現する250ngの線状DNA断片(125ngの各ガイド)を導入した。増殖後、細胞を、トリプシン処理し、希釈し、3つのウェルあたりおよそ1つの細胞の希釈度で96ウェルプレートに再プレATINGした。細胞を、視覚によって観察して、単一細胞コロニーを確認し、24ウェルプレートで増殖させた。編集を決定するために、ゲノムDNAを、製造業者の説明書に従ってAgencourt DNAdvanceキット(Beckman Coulter)を使用して、クローンから単離した。クローンを、ddPCRによってスクリーニングし、サンガー配列決定によって確認した。

20

【0209】

T細胞 - Gene1 / Gene2

Gene1およびGene2遺伝子座を標的にする化膿レンサ球菌(*S. pyogenes*)ガイドRNAを、製造業者のプロトコルに従ってT7-Scribe™ Standard RNA IVT Kit(CELLSCRIPT)を使用して、PCR産物のインビトロ転写によって産生した。インビトロ転写反応のためのPCR産物は、鋳型としてのプラスミドPLA380、および通共通のリバースプライマー(OLI7078)と共に示されたフォワードプライマー(Gene2のためのOLI7076、およびGene1のためのOLI7077)を使用して産生した。

30

【0210】

インビトロで転写されたガイドRNAを、野生型Cas9タンパク質と、2:1のモル比で複合体形成させて、リボ核タンパク質(RNP)を産生した。複合体形成の完全性を、示差走査蛍光定量法(DSF)によって評価した。簡単に言うと、5μLの複合体形成されたRNPを、5μL 2X Dye Mixで希釈した。2X Dye Mixは、5000XストックSYPRO Orange Protein Gel Stain染料(Life Technologies, S6651)を、MgCl₂を含む10xHEPES-生理食塩水(Boston Bio Products, C-6767)に入れたもの(ヌクレアーゼフリー水で1xに希釈)から産生した。複合体形成された試料と、複合体形成されていないタンパク質対照を、384ウェルプレートに入れ、次のプロトコルを使用するBioRadサーモサイクラーに入れた: 20 で1分、1 ずつ増える20 ~ 95 までの融解曲線(Melt Curve)、4 で1分。成功裏の複合体形成は、複合体形成されていない対照試料と複合体形成されたRNPとの間の明らかな温度変化

40

50

と定義される。

【0211】

ヒトT細胞を、製造業者のプロトコルに従ってMiltenyi CD4マイクロビーズを使用して、バフィーコートから単離した。第0日に、T細胞を、T細胞増殖および活性化のためのDynabeads（登録商標）ヒトT-Activator CD3/CD28（ThermoFisher Scientific）を使用して活性化した。第2日に、ビーズを除去した。第4日に、製造業者のプロトコルに従ってトリパンブルー（Trypan Blue）（ThermoFisher Scientific）およびTC20（商標）自動セルカウンター（Automated Cell Counter）（Bio-Rad）を使用して、細胞を計数した。各条件について、500,000個のT細胞を、2μMの総RNPを含有する22μLのPrimary Cell Nucleofector Solution P2（Lonza）に再懸濁した。試料を、16ウェルNucleocuvette™ Strips（Lonza）に移し、4D-Nucleofector（商標）システムのプログラムDS130（Lonza）を使用して電気穿孔した。その後、細胞を、未処理96ウェル丸底プレートに移し、5% ヒトAB血清（Gemini BioProduct）、1.6mg/mL N-アセチルシステイン（Sigma）、2mM L-アラニル-L-グルタミン（Thermo Scientific）、50IU/mL IL-2（PeproTech）、5ng/mL IL-7（PeproTech）、および0.5ng/mL IL-15（PeproTech）を含有する、200μLのX-Vivo15培地（Lonza）中で培養した。

10

20

【0212】

ヌクレオフェクションの4日後、試料をベレット化し、製造業者のプロトコルに従ってAgencourt DNAdvanceキット（Beckman Coulter）を使用して、ゲノムDNAを精製した。UDiTASを、次の改変を伴って実施した。タグメンテーション後、1μLの0.2% SDSの添加、ピペット混合、および5分の室温インキュベーションを用いて、酵素を失活させた。タグメンテーションされたDNAは、プライマーOLI6259およびOLI6256を使用するラウンド1 PCRに直接的に添加した（図14C）。

【0213】

プラスミド感度実験

Gene3については、プラスミドに基づく感度実験PLA370、PLA367、およびPLA379を使用した（図19A）。Gene1/Gene2転座については、プラスミドに基づく感度実験PLA365、PLA366、PLA377、およびPLA378を使用した（図19B）。プラスミド濃度は、NanoDrop2000分光光度計を使用して決定し、すべてのプラスミドについて、10ng/μLの作業用希釈液を作製した。簡単に言うと、Gene2感度実験については、第1の試料は、PLA370（逆位）とPLA367（大きい欠失）の50%混合物からなり、対照プラスミド（PLA379）を含有しない。その後、PLA370/PLA367混合物を、様々な希釈全体を通して10ng/μLの全プラスミド濃度を維持しながら、対照プラスミド（PLA379）で連続的に希釈する（sqrt10希釈係数）ことによって、10種の希釈物を作製した。最後の試料は、対照プラスミド（PLA379）のみから構成されていた。Gene1/Gene2感度実験については、第1の試料は、PLA365とPLA366の50%混合物からなり、対照プラスミド（PLA377/PLA378）を含有しない。その後、PLA366を、様々な希釈全体を通して10ng/μLの全プラスミド濃度を維持しながら、対照プラスミドPLA377/PLA378の等量混合物で連続的に希釈する（sqrt10希釈係数）ことによって、10種の希釈物を作製した。最後の試料は、対照プラスミド（PLA377/PLA378の等量混合物）のみから構成されていた。その後、すべての試料を、UDiTASにかけた：Gene1/Gene2転座試料は、OLI6256およびOLI6259を用いて増幅させ、一方で、Gene3プラスミドは、OLI6062を用いて増幅させた。UDiTASプロトコルを、次の改変：第1お

30

40

50

よび第2ラウンドのPCRについて6サイクルを伴って適用した。

【0214】

マウスDNAに加えられるプラスミドについては、異なる量の固有のプラスミドが、マウスgDNAに混合された。Gene3プラスミド添加実験については、表3に示す推定コピー数を有するプラスミドが、10ng/μLマウスgDNAに加えられた。

【0215】

【表3】

表3.Gene3添加実験に使用されるプラスミドおよび量。

プラスミド名	予想されたプラスミドコピー数 (50ngの反応物あたり)
PLA370	1,032,035
PLA367	326,363
PLA371	103,204
PLA368	32,636
PLA372	10,320
PLA369	3,264
PLA379	4,684,365

【0216】

Gene1/Gene2プラスミド添加実験については、プラスミドは、表4に記載した通りに、10ng/μLマウスgDNAに加えられた。

【0217】

10

20

30

40

50

【表 4】

表4.Gene1/Gene2添加実験に使用されるプラスミドおよび量。

プラスミド名	予想されたプラスミドコピー数(50ng の反応物あたり)
PLA361 PLA362 PLA363 PLA364 PLA365 PLA366 PLA377/PLA378(等量混合物) PLA361	1,031,968
PLA362	326,339
PLA363	103,197
PLA364	32,634
PLA365	10,320
PLA366	3,263
PLA377/PLA378(等量混合物)	4,695,992

【 0 2 1 8 】

10

20

30

40

50

【 表 5 】

表5:アニーリングステップ、およびTn5:オリゴとの複合体形成のために使用されるアダプターオリゴ配列。モザイク末端配列には下線を引いてある。

オリゴ名	配列 5'-3'
Tn5-A ボトム	[Phos]CTGTCTCTTATACA[ddC] (配列番号 32)
i5_N501_UMI_Tn5-A	AATGATACGGCGACCACCGAGATCTACACTAGA TCGCNNNNNNNNNNNTCGTCGGCAGCGTCAGATG TGTATAAGAGACAG (配列番号 33)
i5_N502_UMI_Tn5-A	AATGATACGGCGACCACCGAGATCTACACCTCT CTATNNNNNNNNNNNTCGTCGGCAGCGTCAGATG TGTATAAGAGACAG (配列番号 34)
i5_N503_UMI_Tn5-A	AATGATACGGCGACCACCGAGATCTACACTATCC TCTNNNNNNNNNNNTCGTCGGCAGCGTCAGATGT GTATAAGAGACAG (配列番号 35)
i5_N504_UMI_Tn5-A	AATGATACGGCGACCACCGAGATCTACACAGAG TAGANNNNNNNNNNTCGTCGGCAGCGTCAGATG TGTATAAGAGACAG (配列番号 36)
i5_N505_UMI_Tn5-A	AATGATACGGCGACCACCGAGATCTACACGTAA GGAGNNNNNNNNNNNTCGTCGGCAGCGTCAGATG TGTATAAGAGACAG (配列番号 37)
i5_N506_UMI_Tn5-A	AATGATACGGCGACCACCGAGATCTACACACTG CATANNNNNNNNNNTCGTCGGCAGCGTCAGATG TGTATAAGAGACAG (配列番号 38)
i5_N507_UMI_Tn5-A	AATGATACGGCGACCACCGAGATCTACACAAGG AGTANNNNNNNNNNTCGTCGGCAGCGTCAGATG TGTATAAGAGACAG (配列番号 39)
i5_N508_UMI_Tn5-A	AATGATACGGCGACCACCGAGATCTACACCTAA GCCTNNNNNNNNNNNTCGTCGGCAGCGTCAGATG TGTATAAGAGACAG (配列番号 40)

【 0 2 1 9 】

10

20

30

40

50

【表 6】

表6.

オリゴ名	i5バーコード 配列	i5バーコードリード(UMIについてはT×10)
i5_N501_UMI_Tn5-A	TAGATCGC	TAGATCGCTTTTTTTTTTT (配列番号 41)
i5_N502_UMI_Tn5-A	CTCTCTAT	CTCTCTATTTTTTTTTTTT (配列番号 42)
i5_N503_UMI_Tn5-A	TATCCTCT	TATCCTCTTTTTTTTTTTT (配列番号 43)
i5_N504_UMI_Tn5-A	AGAGTAGA	AGAGTAGATTTTTTTTTTT (配列番号 44)
i5_N505_UMI_Tn5-A	GTAAGGAG	GTAAGGAGTTTTTTTTTTT (配列番号 45)
i5_N506_UMI_Tn5-A	ACTGCATA	ACTGCATATTTTTTTTTTT (配列番号 46)
i5_N507_UMI_Tn5-A	AAGGAGTA	AAGGAGTATTTTTTTTTTT (配列番号 47)
i5_N508_UMI_Tn5-A	CTAAGCCT	CTAAGCCTTTTTTTTTTTT (配列番号 48)

【 0 2 2 0 】

【表 7】

表7.ラウンド1および2 PCRのためのP5/i5オリゴ配列:

オリゴ名	配列 5'-3'
i5	AATGATACGGCGACCACCGAGATCTACAC (配列番号 49)

【 0 2 2 1 】

10

20

30

40

50

【表 8 - 1】

表8.ラウンド2 PCRのためのi7バーコードオリゴ配列:

オリゴ名	配列 (5'-3')	i7 BC 配列	i7 BC リード
i7_N701_SBS12	CAAGCAGAAGACGGCATAACG AGATAGCGGAATGTGACTGGA G TTCAGACGTGT (配列番号 50)	<u>AGCGGAAT</u> (配列番号 62)	ATTCCGCT (配列番号 74)
i7_N702_SBS12	CAAGCAGAAGACGGCATAACG AGATGATCATGCGTGACTGGA G TTCAGACGTGT (配列番号 51)	<u>GATCATGC</u> (配列番号 63)	GCATGATG (配列番号 75)
i7_N703_SBS12	CAAGCAGAAGACGGCATAACG AGATAAGACGGAGTGACTGGA G TTCAGACGTGT (配列番号 52)	<u>AAGACGGA</u> (配列番号 64)	TCCGTCTT (配列番号 76)
i7_N704_SBS12	CAAGCAGAAGACGGCATAACG AGATCGAGTCCTGTGACTGGA G TTCAGACGTGT (配列番号 53)	<u>CGAGTCCT</u> (配列番号 65)	AGGACTCG (配列番号 77)
i7_N705_SBS12	CAAGCAGAAGACGGCATAACG AGATTCCTCAGGGTGACTGGA G TTCAGACGTGT (配列番号 54)	<u>TCCTCAGG</u> (配列番号 66)	CCTGAGGA (配列番号 78)
i7_N706_SBS12	CAAGCAGAAGACGGCATAACG AGATGTACGGATGTGACTGGA G TTCAGACGTGT (配列番号 55)	<u>GTACGGAT</u> (配列番号 67)	ATCCGTAC (配列番号 79)

【 0 2 2 2 】

10

20

30

40

50

【表 8 - 2】

i7_N707_SBS12	CAAGCAGAAGACGGCATAACG AGATCATCTCTCGTGACTGGA GTTTCAGACGTGT (配列番号 56)	CATCTCTC (配列番号 68)	GAGAGATG (配列番号 80)
i7_N710_SBS12	CAAGCAGAAGACGGCATAACG AGATGTCGGAGCGTGACTGGA GTTTCAGACGTGT (配列番号 57)	GTCGGAGC (配列番号 69)	GCTCCGAC (配列番号 81)
i7_N711_SBS12	CAAGCAGAAGACGGCATAACG AGATACGGAGAAAGTGACTGGA GTTTCAGACGTGT (配列番号 58)	ACGGAGAA (配列番号 70)	TTCTCCGT (配列番号 82)
i7_N712_SBS12	CAAGCAGAAGACGGCATAACG AGATAGGAGATGGTGACTGGA GTTTCAGACGTGT (配列番号 59)	AGGAGATG (配列番号 71)	CATCTCCT (配列番号 83)
i7_N714_SBS12	CAAGCAGAAGACGGCATAACG AGATAGTACTCGGTGACTGGA GTTTCAGACGTGT (配列番号 60)	AGTACTCG (配列番号 72)	CGAGTACT (配列番号 84)
i7_N715_SBS12	CAAGCAGAAGACGGCATAACG AGATGGACTCTAGTGACTGGA GTTTCAGACGTGT (配列番号 61)	GGACTCTA (配列番号 73)	TAGAGTCC (配列番号 85)

【0223】

等価物

本発明を、その詳細な説明と共に記載してきたが、前述の説明は、添付の特許請求の範囲の範囲によって定義される本発明の範囲を例示することが意図され、本発明の範囲を限定しないことが理解されよう。他の態様、利点、および改変は、以下の特許請求の範囲の範囲内である。

本発明は、例えば、以下の項目を提供する。

(項目 1)

方法であって、

(a) 部位特異的ヌクレアーゼと接触させた細胞または組織から得られたゲノム DNA 試料を、トランスポゾンの 5' 末端に第 1 の検出配列を含む前記トランスポゾンと、前記トランスポゾンが前記ゲノム DNA 試料に挿入され、かつ前記ゲノム DNA 試料が断片化されて、複数のタグメンテーションされた二本鎖核酸断片であって、前記核酸断片の 5' 末端に付着された前記トランスポゾンを含む二本鎖核酸断片になる条件下で接触させることと；

(b)

(i) 前記ゲノム DNA 内のあらかじめ決定された位置に相補的なヌクレオチド配列を含み、かつその 5' 末端に第 2 の検出配列を含む第 1 の固定のプライマーと、

(ii) 前記第 1 の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第 2 の選択的プライマーと

を使用して、前記タグメンテーションされた核酸断片を増幅させて、前記第 1 の検出配列、前記核酸断片の前記 5' 末端に付着された前記トランスポゾン、および前記第 2 の検出配列を含む、増幅された核酸断片を形成させることと；

(c) 前記増幅された核酸断片を配列決定することとを含む方法。

(項目 2)

前記第 1 の検出配列が、第 1 の配列決定タグを含む、項目 1 に記載の方法。

(項目 3)

ステップ (c) が、前記増幅された核酸断片を、前記第 1 の配列決定タグとハイブリッド形成する第 1 の配列決定プライマーと接触させることを含む、項目 2 に記載の方法。

(項目 4)

前記第 2 の検出配列が、第 2 の配列決定タグを含む、項目 1 ~ 3 のいずれか一項に記載の方法。

(項目 5)

ステップ (c) が、前記増幅された核酸断片を、前記第 2 の配列決定タグとハイブリッド形成する第 2 の配列決定プライマーと接触させることを含む、項目 4 に記載の方法。

(項目 6)

前記ゲノム DNA 試料が、単一の細胞から得られる、項目 1 ~ 5 のいずれか一項に記載の方法。

(項目 7)

ステップ (b) が、それぞれが前記ゲノム DNA 内の異なるあらかじめ決定された位置に相補的なヌクレオチド配列を含む、複数の第 1 の固定プライマーを使用することを含む、項目 1 ~ 6 のいずれか一項に記載の方法。

(項目 8)

ステップ (b) が、それぞれの反応が異なる第 1 の固定プライマーを使用する、複数の増幅反応を実施することを含む、項目 7 に記載の方法。

(項目 9)

前記複数の第 1 の固定プライマーのそれぞれが、(i) 同じ配列決定タグと (ii) 固有のバーコードとを含む検出配列を含む、項目 7 または 8 に記載の方法。

(項目 10)

前記二本鎖核酸断片が、約 500 ~ 約 5000 bp を含む、項目 1 ~ 9 のいずれか一項に記載の方法。

(項目 11)

前記部位特異的ヌクレアーゼが、Cas9 である、項目 1 ~ 10 のいずれか一項に記載の方法。

(項目 12)

方法であって、

(a) 細胞または組織から得られたゲノム DNA 試料を、RNA ガイド型ヌクレアーゼと、前記ヌクレアーゼによる前記ゲノム DNA の切断に好都合な条件下で接触させることと；

(b) 前記切断されたゲノム DNA を、トランスポゾンの 5' 末端に第 1 の検出配列を含む前記トランスポゾンと、前記切断されたゲノム DNA に前記トランスポゾンが挿入され、かつ前記切断されたゲノム DNA が断片化されて、複数のタグメンテーションされた二本鎖核酸断片であって、前記核酸断片の 5' 末端に付着された前記トランスポゾンを含む二本鎖核酸断片になる条件下で接触させることと；

(c)

(i) 前記ゲノム DNA 内のあらかじめ決定された位置に相補的なヌクレオチド配列を含み、かつその 5' 末端に第 2 の検出配列を含む第 1 の固定プライマーと、

(ii) 前記第 1 の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第 2 の選択的プライマーと

を使用して、前記タグメンテーションされた核酸断片を増幅させて、前記第 1 の検出配列、前記核酸断片の前記 5' 末端に付着された前記トランスポゾン、および前記第 2 の検出配列を含む、増幅された核酸断片を形成させることと；

(d) 増幅された核酸断片を配列決定することとを含む方法。

(項目 13)

前記第 1 の検出配列が、第 1 の配列決定タグを含む、項目 12 に記載の方法。

10

20

30

40

50

(項目 1 4)

ステップ (d) が、前記増幅された核酸断片を、前記第 1 の配列決定タグとハイブリッド形成する第 1 の配列決定プライマーと接触させることを含む、項目 1 3 に記載の方法。

(項目 1 5)

前記第 2 の検出配列が、第 2 の配列決定タグを含む、項目 1 2 ~ 1 4 のいずれか一項に記載の方法。

(項目 1 6)

ステップ (d) が、前記増幅された核酸断片を、前記第 2 の配列決定タグとハイブリッド形成する第 2 の配列決定プライマーと接触させることを含む、項目 1 5 に記載の方法。

(項目 1 7)

前記ゲノム DNA 試料が、単一の細胞から得られる、項目 1 2 ~ 1 6 のいずれか一項に記載の方法。

(項目 1 8)

ステップ (c) が、それぞれが前記ゲノム DNA 内の異なるあらかじめ決定された位置に相補的なヌクレオチド配列を含む、複数の第 1 の固定プライマーを使用することを含む、項目 1 2 ~ 1 7 のいずれか一項に記載の方法。

(項目 1 9)

ステップ (c) が、それぞれの反応が異なる第 1 の固定プライマーを使用する、複数の増幅反応を実施することを含む、項目 1 8 に記載の方法。

(項目 2 0)

前記複数の第 1 の固定プライマーのそれぞれが、(i) 同じ配列決定タグと (i i) 固有のバーコードとを含む検出配列を含む、項目 1 8 または 1 9 に記載の方法。

(項目 2 1)

前記二本鎖核酸断片が、約 5 0 0 ~ 約 5 0 0 0 b p を含む、項目 1 2 ~ 2 0 のいずれか一項に記載の方法。

(項目 2 2)

前記 RNA ガイド型ヌクレアーゼが、C a s 9 である、項目 1 2 ~ 2 1 のいずれか一項に記載の方法。

(項目 2 3)

方法であって、

(a) 細胞または組織から得られたゲノム DNA 試料を、トランスポソンの 5 ' 末端に第 1 の検出配列を含む前記トランスポソンと、前記トランスポソンが前記ゲノム DNA に挿入され、かつ前記ゲノム DNA が断片化されて、複数のタグメンテーションされた二本鎖核酸断片であって、前記核酸断片の 5 ' 末端に付着された前記トランスポソンを含む二本鎖核酸断片になる条件下で接触させることであって、前記ゲノム DNA は、部位特異的ヌクレアーゼによって切断された少なくとも 1 つの部位に組み込まれた二本鎖オリゴヌクレオチドを含む、ことと；

(b)

(i) 前記二本鎖オリゴヌクレオチドの少なくとも一部分に相補的なヌクレオチド配列を含み、かつその 5 ' 末端に第 2 の検出配列を含む第 1 の固定プライマーと、

(i i) 前記第 1 の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第 2 の選択的プライマーと

を使用して、前記タグメンテーションされた核酸断片を増幅させて、前記第 1 の検出配列、前記核酸断片の前記 5 ' 末端に付着された前記トランスポソン、および前記第 2 の検出配列を含む、増幅された核酸断片を形成させることと；

(c) 前記増幅された核酸断片を配列決定することとを含む方法。

(項目 2 4)

前記部位特異的ヌクレアーゼが、C a s 9 である、項目 1 9 に記載の方法。

(項目 2 5)

10

20

30

40

50

(a) 細胞または組織から得られたゲノムDNA試料を、ヌクレアーゼによるゲノムDNAの切断、および二本鎖オリゴヌクレオチドの少なくとも1つの切断部位への組み込みに好都合な条件下で、前記RNAガイド型ヌクレアーゼおよび前記二本鎖オリゴヌクレオチドと接触させることであって、それによって、前記組み込まれた二本鎖オリゴヌクレオチドを含むタグ付けされたゲノムDNAが産生される、ことと；

(b) 前記タグ付けされたゲノムDNAを、トランスポソンの5'末端に第1の検出配列を含む前記トランスポソンと、前記タグ付けされたゲノムDNAに前記トランスポソンが挿入され、かつ前記タグ付けされたゲノムDNAが断片化されて、複数のタグメンテーションされた二本鎖核酸断片であって、前記核酸断片の5'末端に付着された前記トランスポソンを含む二本鎖核酸断片になる条件下で接触させることと；

10

(c)

(i) 前記二本鎖オリゴヌクレオチドの少なくとも一部分に相補的なヌクレオチド配列を含み、かつその5'末端に第2の検出配列を含む第1の固定プライマーと、

(ii) 前記第1の検出配列の少なくとも一部分に相補的なヌクレオチド配列を含む第2の選択的プライマーと

を使用して、前記タグメンテーションされた核酸断片を増幅させて、前記第1の検出配列、前記核酸断片の前記5'末端に付着された前記トランスポソン、および前記第2の検出配列を含む、増幅された核酸断片を形成させることと；

(d) 前記増幅された核酸断片を配列決定することとを含む方法。

20

(項目26)

前記RNAガイド型ヌクレアーゼが、Cas9である、項目25に記載の方法。

(項目27)

前記二本鎖オリゴヌクレオチドのどちらの鎖も、前記ゲノムDNAに対してオルソログスである、項目23～26のいずれか一項に記載の方法。

(項目28)

前記二本鎖オリゴヌクレオチドの各5'末端が、リン酸基を含む、項目23～27のいずれか一項に記載の方法。

(項目29)

前記二本鎖オリゴヌクレオチドの各3'末端が、ホスホロチオエート結合を含む、項目23～28のいずれか一項に記載の方法。

30

(項目30)

前記二本鎖オリゴヌクレオチドの各5'末端が、ホスホロチオエート結合を含む、項目23～29のいずれか一項に記載の方法。

(項目31)

前記二本鎖オリゴヌクレオチドが、30～35個のヌクレオチドまたは60～65個のヌクレオチドを含む、項目23～30のいずれか一項に記載の方法。

(項目32)

前記二本鎖オリゴヌクレオチドが、平滑末端化される、項目23～31のいずれか一項に記載の方法。

40

(項目33)

前記二本鎖オリゴヌクレオチドが、1、2、3、または4塩基の突出ヌクレオチドを有する5'末端を含む、項目23～31のいずれか一項に記載の方法。

(項目34)

前記第1の検出配列が、第1の配列決定タグを含む、項目23～33のいずれか一項に記載の方法。

(項目35)

配列決定することが、前記増幅された核酸断片を、前記第1の配列決定タグとハイブリッド形成する第1の配列決定プライマーと接触させることを含む、項目34に記載の方法。

(項目36)

50

前記第 2 の検出配列が、第 2 の配列決定タグを含む、項目 23 ~ 35 のいずれか一項に記載の方法。

(項目 37)

配列決定することが、前記増幅された核酸断片を、前記第 2 の配列決定タグとハイブリッド形成する第 2 の配列決定プライマーと接触させることを含む、項目 36 に記載の方法。

(項目 38)

前記二本鎖核酸断片が、約 500 ~ 約 5000 bp を含む、項目 19 ~ 37 のいずれか一項に記載の方法。

(項目 39)

前記配列決定された核酸断片が、参照ゲノム DNA 試料に対する転座を含む、項目 1 ~ 38 のいずれか一項に記載の方法。

10

(項目 40)

前記ゲノム DNA 試料が、ウイルスによって導入された配列を含む、項目 1 ~ 39 のいずれか一項に記載の方法。

(項目 41)

前記ゲノム DNA、前記切断されたゲノム DNA、または前記タグ付けされたゲノム DNA を前記トランスポゾンと接触させる前に、ウイルスベクターを使用して、配列を前記ゲノム DNA に導入するステップをさらに含む、項目 1 ~ 40 のいずれか一項に記載の方法。

(項目 42)

20

前記配列決定された核酸断片を前記参照ゲノム DNA 試料と比較することをさらに含む、項目 1 ~ 41 のいずれか一項に記載の方法。

(項目 43)

少なくとも 1 つの配列決定された核酸が、前記参照ゲノム DNA 試料に対する転座を含む、項目 42 に記載の方法。

(項目 44)

前記配列決定された核酸の 1 % 未満が、前記参照ゲノム DNA 試料に対する転座を含む、項目 43 に記載の方法。

(項目 45)

前記ゲノム DNA 試料が、生きている細胞または組織を含有しない試料から得られる、項目 1 ~ 44 のいずれか一項に記載の方法。

30

(項目 46)

前記ゲノム DNA 試料が、痕跡量の細胞および / または組織を含有する試料から得られる、項目 1 ~ 44 のいずれか一項に記載の方法。

(項目 47)

前記ゲノム DNA 試料が、真核細胞から得られる、項目 1 ~ 44 または 46 のいずれか一項に記載の方法。

(項目 48)

前記ゲノム DNA 試料が、原核細胞から得られる、項目 1 ~ 44 または 46 のいずれか一項に記載の方法。

40

(項目 49)

前記ゲノム DNA 試料が、1 つまたは複数の微生物叢から得られる、項目 1 ~ 44、46、または 48 のいずれか一項に記載の方法。

(項目 50)

1 つまたは複数の処理装置によって実施される方法であって、複数の実体についてのヌクレオチド配列を含むデータセットであって、次世代配列決定データを含むデータセットを受け取ることと；

前記データセットから第 1 のデータを得るために前記データセットを逆多重化することであって、前記第 1 のデータは、前記複数の実体の中のある実体から得られるヌクレオチド配列を含み、前記ヌクレオチド配列は、前記実体を特定するバーコードを含み、前記ヌク

50

レオチド配列はまた、1つまたは複数の識別子を含み、前記1つまたは複数の識別子は、前記ヌクレオチド配列上で実施される遺伝子編集事象に基づいている、ことと；
前記第1のデータに基づいて、第2のデータによって表される参照アンプリコンに対して、前記ヌクレオチド配列の少なくとも一部を整列させることであって、前記参照アンプリコンは、変異、欠失、挿入、インデル、転座、逆位、および/または重複を含む1つまたは複数の予測される遺伝子編集事象の提示を含む参照ヌクレオチド配列を含む、ことと；
前記参照アンプリコンに対して整列される前記ヌクレオチド配列の前記少なくとも一部に含有される、前記識別子の数を決定することと；
識別子の前記数に基づいて、前記遺伝子編集事象の結果を評価することとを含む方法。

10

(項目51)

前記1つまたは複数の識別子が、前記ヌクレオチド配列上で実施される複数の遺伝子編集事象に基づいており；かつ
評価することが、識別子の前記数に基づく前記複数の遺伝子編集事象の結果を評価することを含む、項目50に記載の方法。

(項目52)

前記ヌクレオチド配列の前記少なくとも一部を、前記参照アンプリコンに対して整列させることが、前記ヌクレオチド配列のヌクレオチドを、前記参照アンプリコンのヌクレオチドと適合させることを含む、項目50に記載の方法。

(項目53)

バーコードを含む前記ヌクレオチド配列が、前記バーコード中の前記ヌクレオチドの1つを除いてすべてが適合するならば、前記実体を特定する、項目50に記載の方法。

20

(項目54)

整列の前に、前記ヌクレオチド配列から1つまたは複数のアダプターを取り除くことをさらに含む、項目50に記載の方法。

(項目55)

1つまたは複数の表示装置を与えるために、前記遺伝子編集事象の精度を代表する画像データを作成することと；

前記1つまたは複数の表示装置上に、前記画像データに基づいて、1つまたは複数の画像を与えることと

30

をさらに含む、項目50に記載の方法。

(項目56)

前記参照アンプリコンが、1つまたは複数のベクター配列または1つまたは複数のプラスミド配列を含む、項目50に記載の方法。

(項目57)

前記1つまたは複数の識別子が、固有分子インデックス(UMI)を含む、項目50に記載の方法。

(項目58)

前記1つまたは複数の識別子が、1つまたは複数の予測される遺伝子編集事象を示すものである、項目50に記載の方法。

40

(項目59)

前記参照アンプリコンに対して整列される前記ヌクレオチド配列の前記少なくとも一部における前記識別子の前記数を決定することが、予測される遺伝子編集事象に対して整列される識別子の数を計数することを含む、項目50に記載の方法。

(項目60)

前記データセットが、UDiTaSライブラリに基づいている、項目50に記載の方法。

(項目61)

前記バーコードが、前記ヌクレオチド配列に対して固有であり、その結果、前記ヌクレオチド配列に由来する配列決定ライブラリのすべてのメンバーが、同じバーコードを含むようになる、項目50に記載の方法。

50

(項目 6 2)

前記遺伝子編集事象の精度に基づいて、遺伝子編集プロセスを選択すること
をさらに含む、項目 5 0 に記載の方法。

(項目 6 3)

1 つまたは複数の非一過性の機械可読記憶媒体であって、

複数の実体についてのヌクレオチド配列を含むデータセットであって、次世代配列決定
データを含むデータセットを受け取ることと；

前記データセットから第 1 のデータを得るために前記データセットを逆多重化すること
であって、前記第 1 のデータは、前記複数の実体の中のある実体から得られるヌクレオチド
配列を含み、前記ヌクレオチド配列は、前記実体を特定するバーコードを含み、前記ヌク
レオチド配列はまた、1 つまたは複数の識別子を含み、前記 1 つまたは複数の識別子は、
前記ヌクレオチド配列上で実施される遺伝子編集事象に基づいている、ことと；

前記第 1 のデータに基づいて、第 2 のデータによって表される参照アンプリコンに対して
、前記ヌクレオチド配列の少なくとも一部を整列させることであって、前記参照アンプリ
コンは、1 つまたは複数のオンターゲット遺伝子編集事象、1 つまたは複数のオフターゲ
ット遺伝子編集事象、または 1 つまたは複数のオンターゲット遺伝子編集事象と 1 つまた
は複数のオフターゲット遺伝子編集事象との両方の提示を含む参照ヌクレオチド配列を含
む、ことと；

前記参照アンプリコンに対して整列される前記ヌクレオチド配列の前記少なくとも一部に
含有される、前記識別子の数を決定することと；

識別子の前記数に基づいて、前記遺伝子編集事象の結果を評価することと

を含む、動作を実施するための 1 つまたは複数の処理装置によって実行可能である、命令
を保存する 1 つまたは複数の非一過性の機械可読記憶媒体。

(項目 6 4)

前記 1 つまたは複数の識別子が、前記ヌクレオチド配列上で実施される複数の遺伝子編
集事象に基づいており；かつ

評価することが、識別子の前記数に基づく前記複数の遺伝子編集事象の結果を評価するこ
とを含む、項目 6 3 に記載の 1 つまたは複数の非一過性の機械可読記憶媒体。

(項目 6 5)

前記ヌクレオチド配列の少なくとも一部を、前記参照アンプリコンに対して整列させる
ことが、前記ヌクレオチド配列のヌクレオチドを、前記参照アンプリコンのヌクレオチド
と適合させることを含む、項目 6 3 に記載の 1 つまたは複数の非一過性の機械可読記憶媒
体。

(項目 6 6)

前記バーコードを含む前記ヌクレオチド配列が、前記バーコード中の前記ヌクレオチド
の 1 つを除いてすべてが適合するならば、前記実体を特定する、項目 6 3 に記載の 1 つま
たは複数の非一過性の機械可読記憶媒体。

(項目 6 7)

前記動作が、

整列の前に、前記ヌクレオチド配列から 1 つまたは複数のアダプターを取り除くことを
含む、項目 6 3 に記載の 1 つまたは複数の非一過性の機械可読記憶媒体。

(項目 6 8)

前記動作が、

1 つまたは複数の表示装置を与えるために、前記遺伝子編集事象の精度を代表する画像
データを作成することと；

前記 1 つまたは複数の表示装置上に、前記画像データに基づいて、1 つまたは複数の画像
を与えることと

を含む、項目 6 3 に記載の 1 つまたは複数の非一過性の機械可読記憶媒体。

(項目 6 9)

前記参照アンプリコンが、1 つまたは複数のベクター配列または 1 つまたは複数のプラ

10

20

30

40

50

スミド配列を含む、項目 6 3 に記載の 1 つまたは複数の非一過性の機械可読記憶媒体。
(項目 7 0)

前記 1 つまたは複数の識別子が、固有分子インデックス (U M I) を含む、項目 6 3 に記載の 1 つまたは複数の非一過性の機械可読記憶媒体。

(項目 7 1)

前記 1 つまたは複数の識別子が、1 つまたは複数の予測される遺伝子編集事象を示すものである、項目 6 3 に記載の 1 つまたは複数の非一過性の機械可読記憶媒体。

(項目 7 2)

前記参照アンプリコンに対して整列される前記ヌクレオチド配列の前記少なくとも一部における前記識別子の前記数を決定することが、予測される遺伝子編集事象に対して整列される識別子の数を計数することを含む、項目 6 3 に記載の 1 つまたは複数の非一過性の機械可読記憶媒体。

10

(項目 7 3)

前記データセットが、U D i T a S ライブラリに基づいている、項目 6 3 に記載の 1 つまたは複数の非一過性の機械可読記憶媒体。

(項目 7 4)

前記バーコードが、前記ヌクレオチド配列に対して固有であり、その結果、前記ヌクレオチド配列に由来する配列決定ライブラリのすべてのメンバーが、同じバーコードを含むようになる、項目 6 3 に記載の 1 つまたは複数の非一過性の機械可読記憶媒体。

(項目 7 5)

20

前記遺伝子編集事象の精度に基づいて、遺伝子編集プロセスを選択することをさらに含む、項目 6 3 に記載の 1 つまたは複数の非一過性の機械可読記憶媒体。

(項目 7 6)

システムであって、

実行可能である命令を保存する非一過性の機械可読ストレージと；

複数の実体についてのヌクレオチド配列を含むデータセットであって、次世代配列決定データを含むデータセットを受け取ることと；

前記データセットから第 1 のデータを得るために前記データセットを逆多重化することであって、前記第 1 のデータは、前記複数の実体の中のある実体から得られるヌクレオチド配列を含み、前記ヌクレオチド配列は、前記実体を特定するバーコードを含み、前記ヌクレオチド配列はまた、1 つまたは複数の識別子を含み、前記 1 つまたは複数の識別子は、前記ヌクレオチド配列上で実施される遺伝子編集事象に基づいている、ことと；

30

前記第 1 のデータに基づいて、第 2 のデータによって表される参照アンプリコンに対して、前記ヌクレオチド配列の少なくとも一部を整列させることであって、前記参照アンプリコンは、1 つまたは複数のオンターゲット遺伝子編集事象、1 つまたは複数のオフターゲット遺伝子編集事象、または 1 つまたは複数のオンターゲット遺伝子編集事象と 1 つまたは複数のオフターゲット遺伝子編集事象との両方の提示を含む参照ヌクレオチド配列を含む、ことと；

前記参照アンプリコンに対して整列される前記ヌクレオチド配列の前記少なくとも一部に含有される、前記識別子の数を決定することと；

40

識別子の前記数に基づいて、前記遺伝子編集事象の結果を評価することと

を含む動作を実施するための前記命令を遂行するための 1 つまたは複数の処理装置とを含むシステム。

(項目 7 7)

前記 1 つまたは複数の識別子が、前記ヌクレオチド配列上で実施される複数の遺伝子編集事象に基づいており；かつ

評価することが、識別子の前記数に基づく前記複数の遺伝子編集事象の結果を評価することを含む、項目 7 6 に記載のシステム。

(項目 7 8)

前記ヌクレオチド配列の前記少なくとも一部を、前記参照アンプリコンに対して整列さ

50

せることが、前記ヌクレオチド配列のヌクレオチドを、前記参照アンプリコンのヌクレオチドと適合させることを含む、項目 7 6 に記載のシステム。

(項目 7 9)

バーコードを含む前記ヌクレオチド配列が、前記バーコード中の前記ヌクレオチドの 1 つを除いてすべてが適合するならば、前記実体を特定する、項目 7 6 に記載のシステム。

(項目 8 0)

前記動作が、
整列の前に、前記ヌクレオチド配列から 1 つまたは複数のアダプターを取り除くことを含む、項目 7 6 に記載のシステム。

(項目 8 1)

前記動作が、
1 つまたは複数の表示装置を与えるために、前記遺伝子編集事象の精度を代表する画像データを作成することと；
前記 1 つまたは複数の表示装置上に、前記画像データに基づいて、1 つまたは複数の画像を与えることと
を含む、項目 7 6 に記載のシステム。

(項目 8 2)

前記参照アンプリコンが、1 つまたは複数のベクター配列または 1 つまたは複数のプラスミド配列を含む、項目 7 6 に記載のシステム。

(項目 8 3)

前記 1 つまたは複数の識別子が、固有分子インデックス (UMI) を含む、項目 7 6 に記載のシステム。

(項目 8 4)

前記 1 つまたは複数の識別子が、1 つまたは複数の予測される遺伝子編集事象を示すものである、項目 7 6 に記載のシステム。

(項目 8 5)

前記参照アンプリコンに対して整列される前記ヌクレオチド配列の前記少なくとも一部における前記識別子の前記数を決定することが、予測される遺伝子編集事象に対して整列される識別子の前記数を計数することを含む、項目 7 6 に記載のシステム。

(項目 8 6)

前記データセットが、UDiTaSライブラリに基づいている、項目 7 6 に記載のシステム。

(項目 8 7)

前記バーコードが、前記ヌクレオチド配列に対して固有であり、その結果、前記ヌクレオチド配列に由来する配列決定ライブラリのすべてのメンバーが、同じバーコードを含むようになる、項目 7 6 に記載のシステム。

(項目 8 8)

前記遺伝子編集事象の精度に基づいて、遺伝子編集プロセスを選択すること
をさらに含む、項目 7 6 に記載のシステム。

10

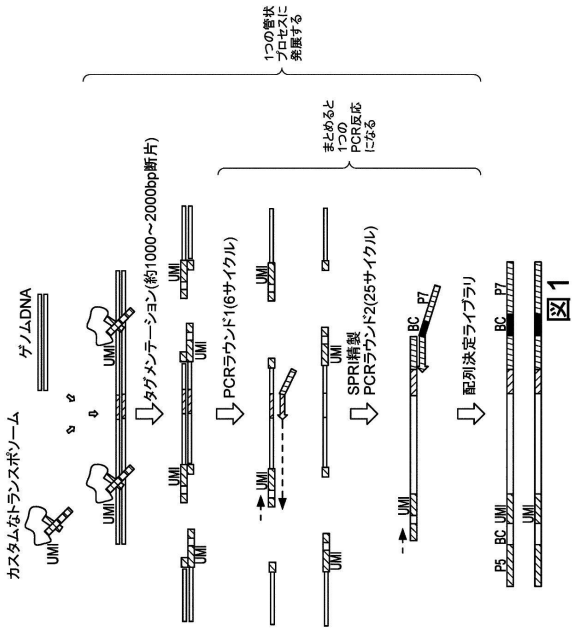
20

30

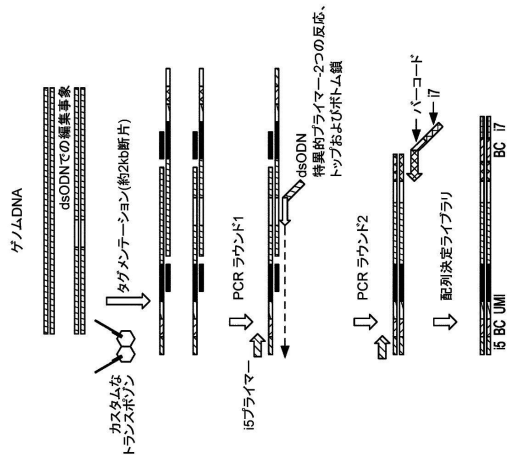
40

50

【図面】
【図 1】



【図 3】



【図 2】

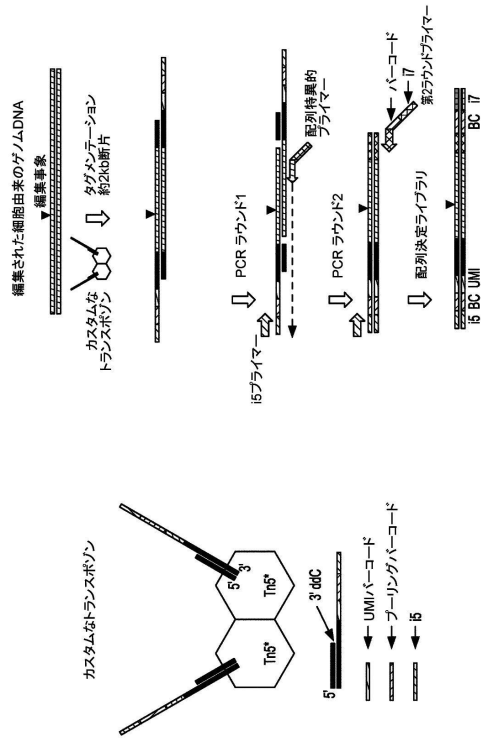


図 2B

図 2A

【図 4 - 1】

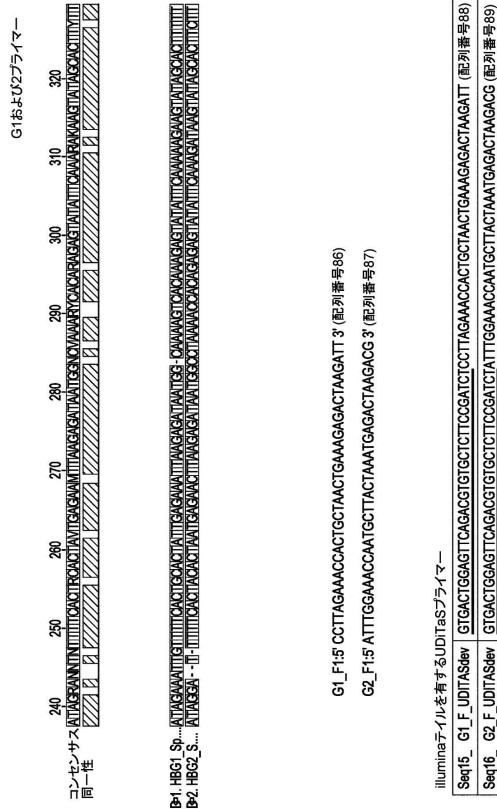
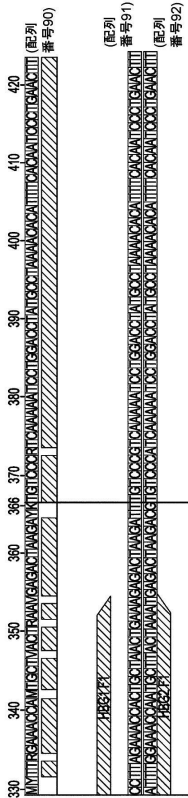


図 4 続き

【図 4 - 2】



【図 6】

試料 JHEEE46_3 (K582細胞)
G1およびG2プライマー-UDITaSアライメントは、標的的特異的である
(リード拡張プロット)

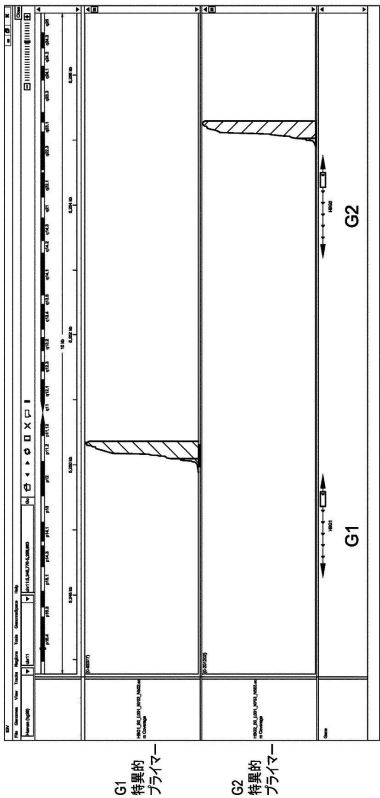


図 4 続き

【図 5】

正しい標的に対するMiSeqリードの78%~93%アラインメント

試料ID	ゲノムDNA名	ガイドRNA名	合計_R1/R2リード	合計_標的リード	標的リード/合計リード
NGS req-097_ A.1 amp_01_	JHEEE46_1	.9	101512	83565	82
NGS req-097_ A.2 amp_02_	JHEEE46_2	.34	91628	79697	87
NGS req-097_ A.3 amp_03_	JHEEE46_3	.35	116184	94552	80
NGS req-097_ A.4 amp_04_	JHEEE46_4	.36	82818	69476	84
NGS req-097_ A.5 amp_05_	JHEEE46_5	.37	76414	63341	83
NGS req-097_ A.6 amp_06_	JHEEE46_6	.38	104338	81211	78
NGS req-097_ A.7 amp_07_	JHEEE46_7	.41	19076	99933	84
NGS req-097_ A.8 amp_08_	JHEEE46_8	.43	118350	90221	76
NGS req-097_ A.9 amp_09_	JHEEE46_9	neg ctrl	19618	17567	88
NGS req-097_ A.1 amp_10_	JHEEE46_10	.9	17060	14782	86
NGS req-097_ A.2 amp_11_	JHEEE46_2	.34	136770	125338	92
NGS req-097_ A.3 amp_12_	JHEEE46_3	.35	220634	203468	92
NGS req-097_ A.4 amp_13_	JHEEE46_4	.36	171376	157285	92
NGS req-097_ A.5 amp_14_	JHEEE46_5	.37	141086	128307	91
NGS req-097_ A.6 amp_15_	JHEEE46_6	.38	133158	116597	88
NGS req-097_ A.7 amp_16_	JHEEE46_7	.41	187056	173777	93
NGS req-097_ A.8 amp_17_	JHEEE46_8	.43	152182	130977	86
NGS req-097_ A.9 amp_18_	JHEEE46_9	neg ctrl	211478	198876	93

【図 7】

小さいインデル率についての方法比較
UDITaSは、標的配列決定およびT7E1アッセイと良く関連している

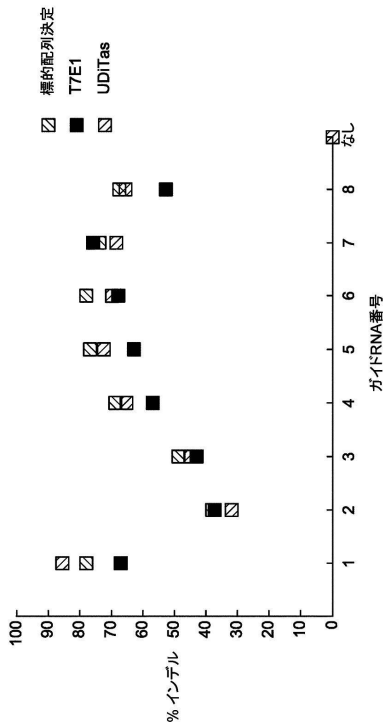


図 5

図 7

10

20

30

40

50

【図 8】

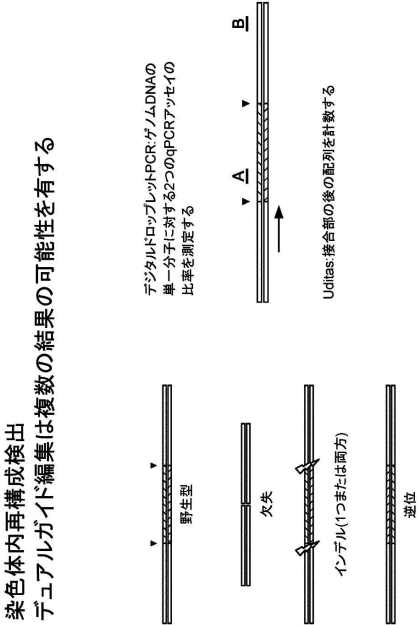


図 8

【図 9】

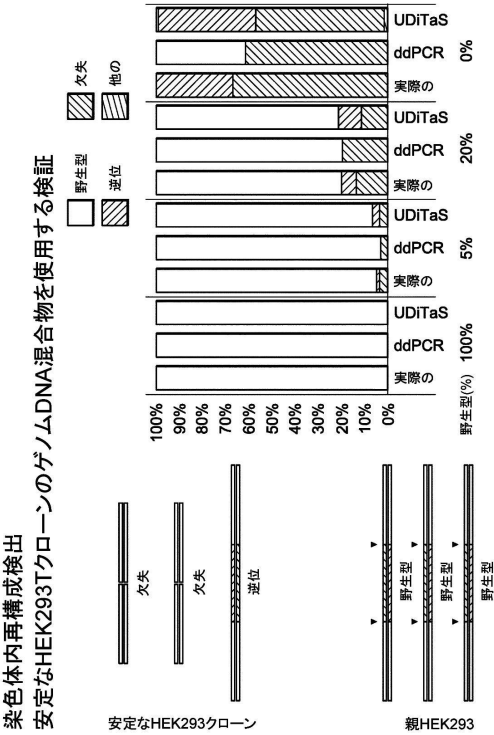


図 9

【図 10】

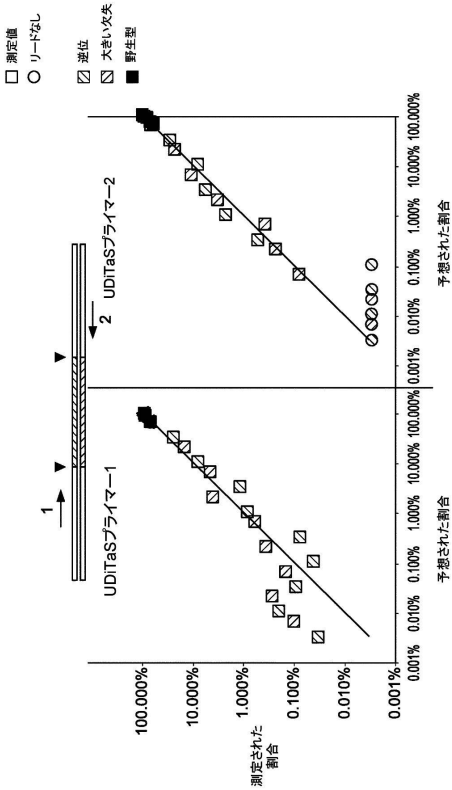


図 10

【図 11】

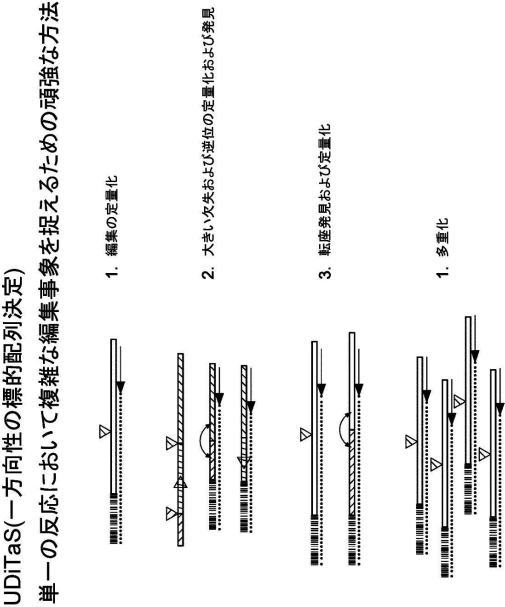


図 11

【 図 1 4 A B 】

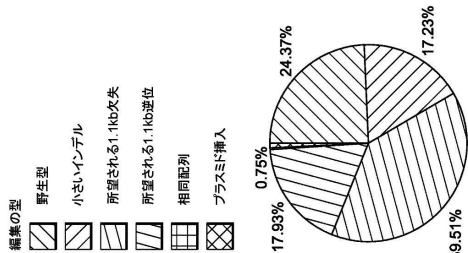


図 14A

図 14B

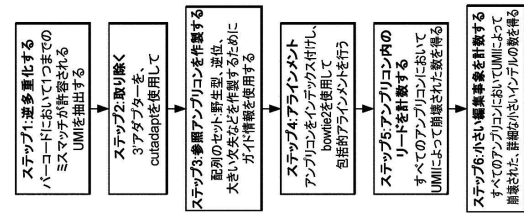


図 12

ステップ	タスク	命令
1	逆多量化する	demultiplex (sample_dir)
2	取り除く	cutadapt -m 10 -e 0.33 -a reverse complement (SBS12) -A reverse complement (NV27) -o file_cutadapt_R1 -p file_cutadapt_R2 file_R1 file_R2
3	参照 アンプリコンを 作成する	bowtie2-build amplicons.fa
4	アライメント	bowtie2 --very-sensitive -p ncpu -X 5000 -k 2 -x amplicons -1 file_R1 -2 file_R2 samtools view -Sb alignments.sam samtools sort alignments.bam -o sorted_alignments.bam
5	アンプリコン 内のリードを 計数する	for amplicon in amplicons.fa: do for read in sorted_alignments.fetch(amplicon) if read in window: UMI_list = UMI_list + [UMI[read]] counts['amplicon'] = len(set(UMI_list)) done
6	小さい編集 のリードを 計数する	for amplicon in amplicons.fa: do for read in sorted_alignments.fetch(amplicon) if read in window: if parse cigar for indel (read): UMI_indel_list = UMI_indel_list + [UMI [read]] count_indels['amplicon'] = len (set(UMI_indel_list)) done

図 12

【 図 1 4 C 】

Gene 1 (Chr 14)	Gene 2 (Chr 15)
型	型
編集なし	編集なし
インデル	インデル
二動原体性	二動原体性
無動原体性	無動原体性
編集なし	編集なし
インデル	インデル
二動原体性	二動原体性
無動原体性	無動原体性
均衡型	均衡型
均衡型	均衡型
無動原体性	無動原体性
二動原体性	二動原体性

図 14C

【 図 1 3 】

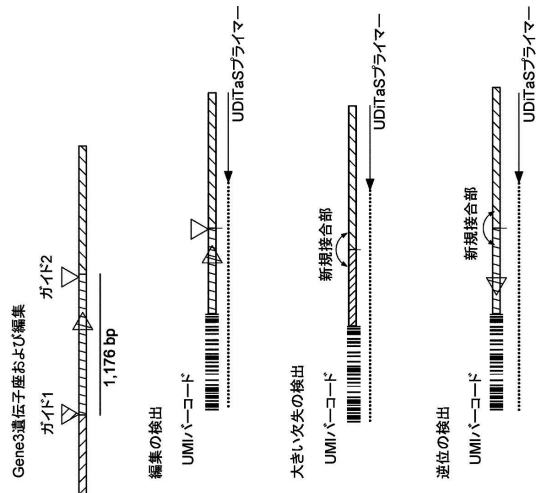
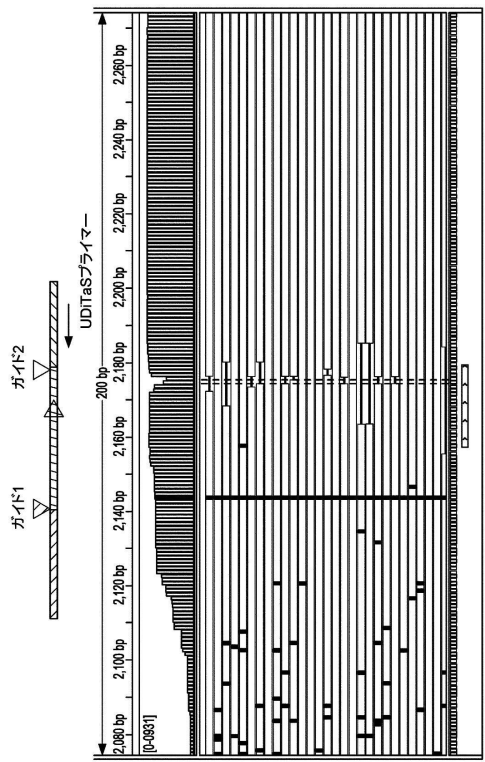


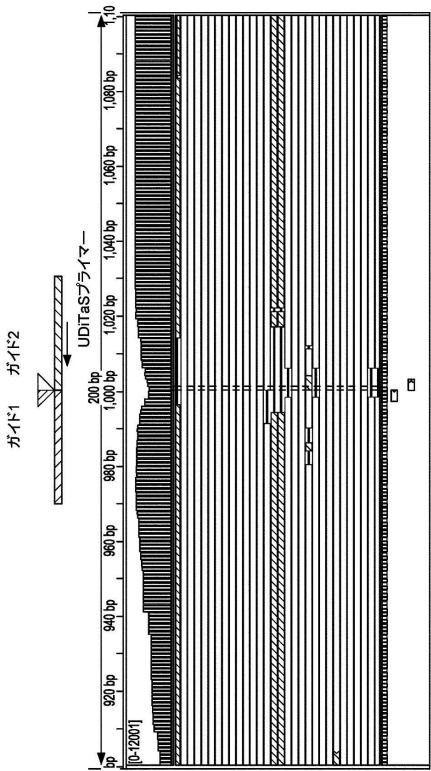
図 13

【 図 1 2 】

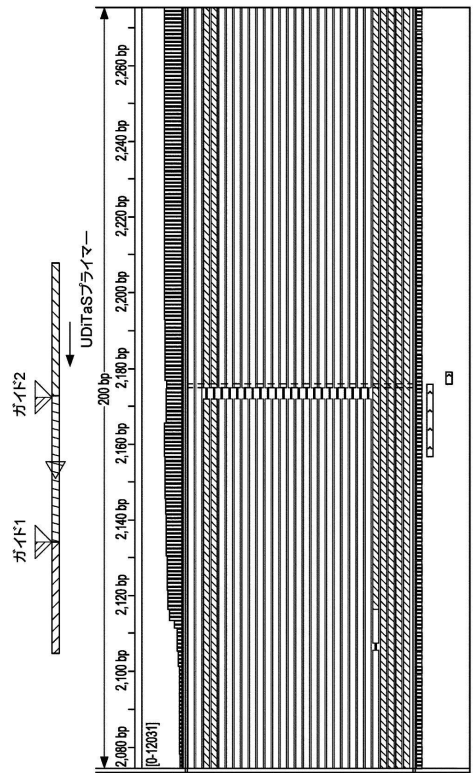
【図 15 A】



【図 15 B】



【図 15 C】



【図 15 D】

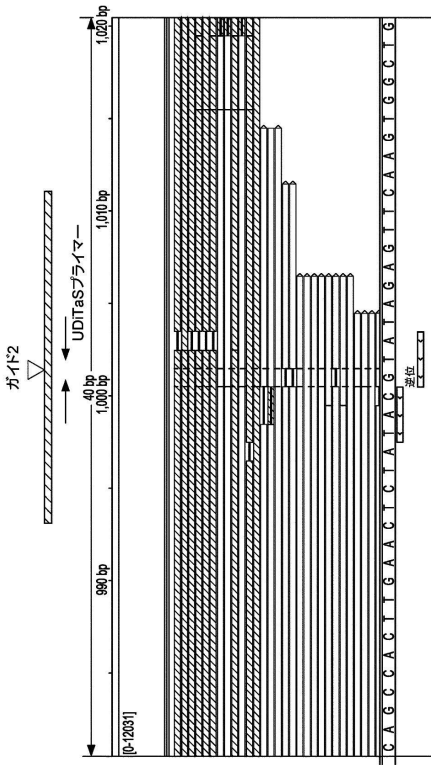


図15B

図15D

10

20

30

40

50

【 図 1 6 】

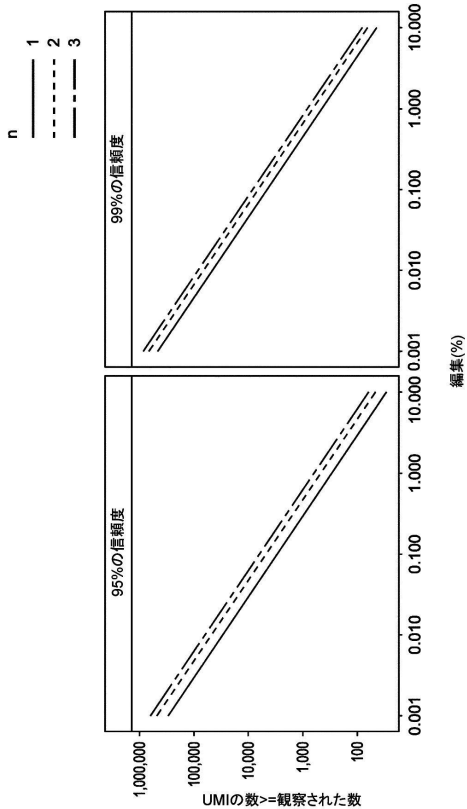


図16

【 図 1 7 】

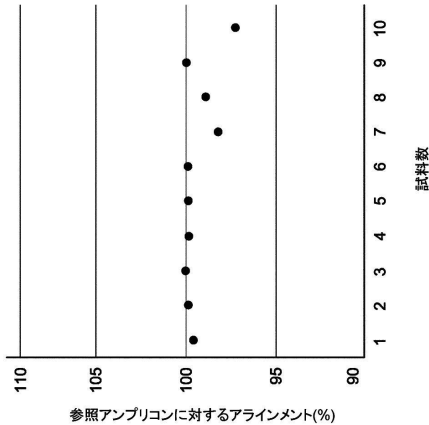


図17

【 図 1 8 】

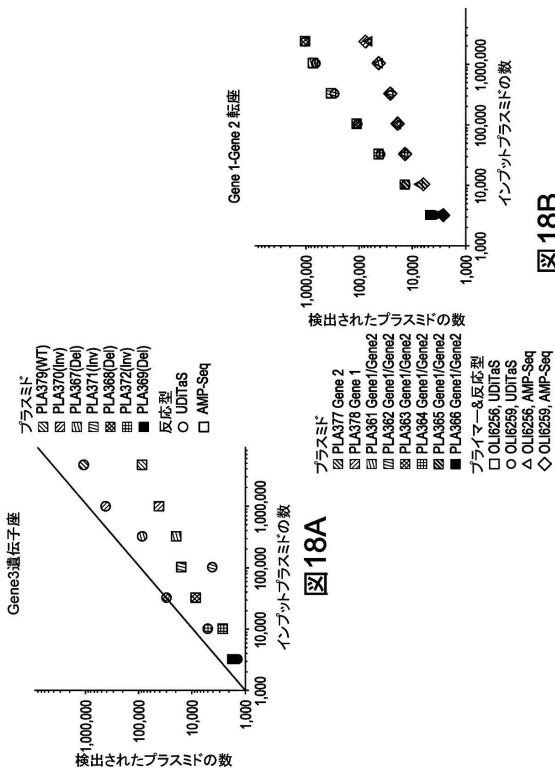


図18B

【 図 1 9 】

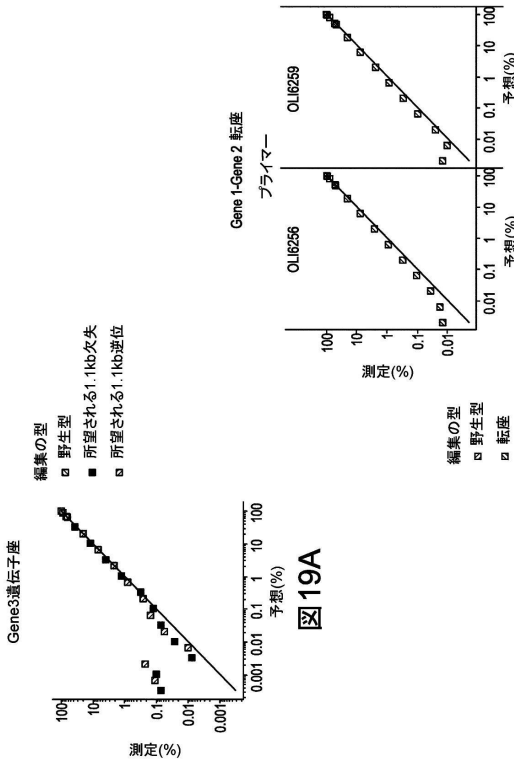


図 19A

図19B

【配列表】

0007229923000001.app

10

20

30

40

50

フロントページの続き

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 62/570,300

(32)優先日 平成29年10月10日(2017.10.10)

(33)優先権主張国・地域又は機関

米国(US)

弁護士 山本 健策

(72)発明者 ジアンヌコス, ジョージア

アメリカ合衆国 マサチューセッツ 02141, ケンブリッジ, ハーレイ ストリート 11,
エディタス メディシン, インコーポレイテッド 気付

(72)発明者 ウィルソン, クリストファー

アメリカ合衆国 マサチューセッツ 02141, ケンブリッジ, ハーレイ ストリート 11,
エディタス メディシン, インコーポレイテッド 気付

(72)発明者 チューラ, ドーン

アメリカ合衆国 マサチューセッツ 02141, ケンブリッジ, ハーレイ ストリート 11,
エディタス メディシン, インコーポレイテッド 気付

審査官 山内 達人

(56)参考文献 米国特許出願公開第2016/0017396(US, A1)

特表2001-510055(JP, A)

(58)調査した分野 (Int.Cl., DB名)

C12N

C12M

C12Q

JSTPlus/JMEDPlus/JST7580(JDreamIII)

CAPLUS/MEDLINE/EMBASE/BIOSIS(STN)