

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号

特表2024-518095

(P2024-518095A)

(43)公表日 令和6年4月24日(2024.4.24)

(51)国際特許分類		F I		テーマコード(参考)	
C 1 2 Q	1/6869(2018.01)	C 1 2 Q	1/6869	Z Z N A	4 B 0 6 3
C 4 0 B	40/06 (2006.01)	C 4 0 B	40/06		
C 1 2 Q	1/6883(2018.01)	C 1 2 Q	1/6883	Z	
C 1 2 Q	1/6886(2018.01)	C 1 2 Q	1/6886	Z	
C 1 2 Q	1/6888(2018.01)	C 1 2 Q	1/6888	Z	
審査請求 未請求 予備審査請求 未請求 (全87頁) 最終頁に続く					
(21)出願番号	特願2023-570212(P2023-570212)	(71)出願人	595117091		
(86)(22)出願日	令和4年5月12日(2022.5.12)		ベクトン・ディキンソン・アンド・カン		
(85)翻訳文提出日	令和6年1月10日(2024.1.10)		パニー		
(86)国際出願番号	PCT/US2022/029057		BECTON, DICKINSON A		
(87)国際公開番号	WO2022/241158		ND COMPANY		
(87)国際公開日	令和4年11月17日(2022.11.17)		アメリカ合衆国 ニュー・ジャージー 0		
(31)優先権主張番号	63/243,443		7 4 1 7 - 1 8 8 0 フランクリン・レ		
(32)優先日	令和3年9月13日(2021.9.13)		イクス ベクトン・ドライブ 1		
(33)優先権主張国・地域又は機関	米国(US)	(74)代理人	100094569		
			弁理士 田中 伸一郎		
(31)優先権主張番号	63/189,032	(74)代理人	100103610		
(32)優先日	令和3年5月14日(2021.5.14)		弁理士 吉 田 和彦		
(33)優先権主張国・地域又は機関	米国(US)	(74)代理人	100109070		
			弁理士 須田 洋之		
(81)指定国・地域	AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA	(74)代理人	100119013		
	最終頁に続く			最終頁に続く	

(54)【発明の名称】 核酸シーケンシングのためのライブラリーを製する方法

(57)【要約】

本明細書に開示されるものには、核酸シーケンシングのためのライブラリー生成において使用するのに適した方法、組成物、及びキットが含まれる。一部の実施形態では、複数のタンパク質複合体が提供される。各タンパク質複合体は、トランスポソームと、標的二本鎖DNA (dsDNA) 上のユーザーが選択した結合部位に特異的に結合することができるプログラム可能なDNA結合ユニットとを含むことができる。複数のタンパク質複合体の各々に対する結合部位は、互いに異なることができる。トランスポソームは、トランスポゼース、第1のアダプター、及び第2のアダプターを含むことができる。第1のアダプター、第2のアダプター、又はその両方は、シーケンシングアダプターであり得る。

【選択図】 図3

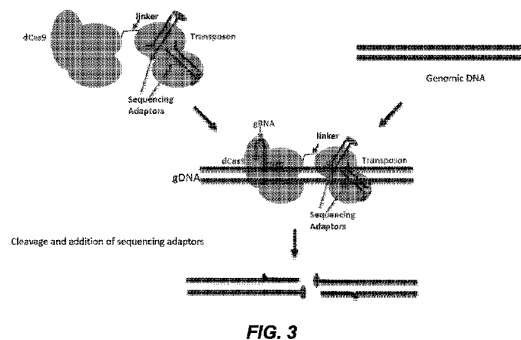


FIG. 3

【特許請求の範囲】

【請求項 1】

複数のタンパク質複合体を含む組成物であって、複数のタンパク質複合体の各々は、トランスポソームと、標的二本鎖 DNA (ds DNA) 上の結合部位に特異的に結合することができるプログラム可能な DNA 結合ユニットとを含み、トランスポソームは、トランスポゼースと、第 1 のアダプターと、第 2 のアダプターとを含み、複数のタンパク質複合体の各々に対する結合部位は互いに異なる、前記組成物。

【請求項 2】

複数のタンパク質複合体のうちの少なくとも 2 つが、同じトランスポソームを含む、請求項 1 に記載の組成物。

【請求項 3】

複数のタンパク質複合体のすべてが同じトランスポソームを含む、請求項 1 に記載の組成物。

【請求項 4】

複数のタンパク質複合体のすべてが同じトランスポゼースを含む、請求項 1 ~ 3 のいずれか 1 項に記載の組成物。

【請求項 5】

同じトランスポソーム内の第 1 のアダプターと第 2 のアダプターとが同じである、請求項 1 ~ 4 のいずれか 1 項に記載の組成物。

【請求項 6】

異なるトランスポソーム内の第 1 のアダプター、第 2 のアダプター、又はその両方が異なる、請求項 1 ~ 5 のいずれか 1 項に記載の組成物。

【請求項 7】

第 1 のアダプター、第 2 のアダプター、又はその両方が、dsDNA 又は RNA / DNA 二重鎖である、請求項 1 ~ 6 のいずれか 1 項に記載の組成物。

【請求項 8】

アダプターは、長さが約 3 ~ 200 塩基対である、請求項 1 ~ 7 のいずれか 1 項に記載の組成物。

【請求項 9】

第 1 のアダプター、第 2 のアダプター、又はその両方が、シーケンシングアダプターである、請求項 1 ~ 8 のいずれか 1 項に記載の組成物。

【請求項 10】

シーケンシングアダプターが P5 又は P7 プライマー配列を含む、請求項 9 に記載の組成物。

【請求項 11】

複数のタンパク質複合体のうちの少なくとも 2 つの結合部位が同じ標的 dsDNA 上にある、請求項 1 ~ 10 のいずれか 1 項に記載の組成物。

【請求項 12】

複数のタンパク質複合体のうちの少なくとも 2 つの結合部位が、同じ標的 dsDNA 上で約 1 ~ 50000 ヌクレオチド離れている、請求項 11 に記載の組成物。

【請求項 13】

複数のタンパク質複合体の一对の結合部位間の距離が、複数のタンパク質複合体の別の一对の結合部位間の距離と実質的に同じである、請求項 11 に記載の組成物。

【請求項 14】

複数のタンパク質複合体の一对の結合部位間の距離が、複数のタンパク質複合体の別の一对の結合部位間の距離と異なる、請求項 11 に記載の組成物。

【請求項 15】

複数のタンパク質複合体のうちの少なくとも 2 つの結合部位が、標的 dsDNA の異なる鎖上にある、請求項 1 ~ 11 のいずれか 1 項に記載の組成物。

【請求項 16】

10

20

30

40

50

複数のタンパク質複合体のうちの少なくとも2つが、異なる標的 dsDNA に特異的に結合することができる、請求項1～15のいずれか1項に記載の組成物。

【請求項17】

複数のタンパク質複合体が、約2～5000の標的 dsDNA に特異的に結合することができる、請求項1～15のいずれか1項に記載の組成物。

【請求項18】

トランスポゼースが、Tn5トランスポゼース、Tn7トランスポゼース、マリナーTc1様トランスポゼース、Himar1C9トランスポゼース、又はスリーピングビューティトランスポゼースである、請求項1～17のいずれか1項に記載の組成物。

【請求項19】

トランスポゼースが高活性トランスポゼースである、請求項1～18のいずれか1項に記載の組成物。

【請求項20】

プログラム可能なDNA結合ユニットが、ヌクレアーゼ欠損CRISPR関連タンパク質(dCASタンパク質)と、標的 dsDNA の結合部位に特異的に結合することができるガイドRNA(gRNA)とを含む、請求項1～19のいずれか1項に記載の組成物。

【請求項21】

トランスポソームが、トランスポゼースとdCASタンパク質とを接続するリンカーを介してプログラム可能なDNA結合ユニットと結び付けられている、請求項20に記載の組成物。

【請求項22】

リンカーが、ペプチドリinker、化学リンカー、又はその両方を含む、請求項21に記載の組成物。

【請求項23】

トランスポゼースが、dCASタンパク質を含む融合タンパク質中に存在する、請求項20に記載の組成物。

【請求項24】

dCASタンパク質がdCAS9、dCAS12、dCAS13、dCAS14、又はSpRY dCASである、請求項20～23のいずれか1項に記載の組成物。

【請求項25】

dCAS13タンパク質がdCAS13a、dCAS13b、dCAS13c、又はdCAS13dである、請求項24に記載の組成物。

【請求項26】

プログラム可能なDNA結合ユニットが、標的 dsDNA 上の結合部位に特異的に結合することができるタンパク質構成成分を含み、タンパク質構成成分が、エンドヌクレアーゼ欠損ジンクフィンガーヌクレアーゼ(ZFN)、エンドヌクレアーゼ欠損転写アクチベーター様エフェクターヌクレアーゼ(TALEN)、アルゴノートタンパク質、エンドヌクレアーゼ欠損メガヌクレアーゼ、リコンビナーゼ、又はそれらの組み合わせを含む、請求項1～19のいずれか1項に記載の組成物。

【請求項27】

トランスポソームが、トランスポゼースとタンパク質構成成分とを接続するリンカーを介してプログラム可能なDNA結合ユニットと結び付けられている、請求項26に記載の組成物。

【請求項28】

リンカーが、ペプチドリinker、化学リンカー、又はその両方を含む、請求項27に記載の組成物。

【請求項29】

ペプチドリinkerが、複数のグリシン、セリン、スレオニン、アラニン、リジン、グルタミン、又はそれらの組み合わせを含む、請求項28に記載の組成物。

【請求項30】

10

20

30

40

50

ペプチドリンカーがGSリンカーを含む、請求項29に記載の組成物。

【請求項31】

ペプチドリンカーがXTENリンカーである、請求項28に記載の組成物。

【請求項32】

タンパク質構成成分が、トランスポゼースを含む融合タンパク質中に存在する、請求項26に記載の組成物。

【請求項33】

請求項1～32のいずれか1項に記載の組成物と、

1種以上の標的dsDNAを含むことが疑われる試料核酸とを含む反応混合物。

10

【請求項34】

DNAポリメラーゼ、dNTP、又はそれらの組み合わせをさらに含む、請求項33に記載の反応混合物。

【請求項35】

アダプターが標的dsDNA又はその断片に共有結合的に付着している、請求項33～34のいずれか1項に記載の反応混合物。

【請求項36】

複数のdsDNA断片を含み、各々が、各末端にそれぞれ複数のタンパク質複合体のうちの1つの第1のアダプター及び第2のアダプターを含む、請求項33～35のいずれか1項に記載の反応混合物。

20

【請求項37】

試料核酸が、真核生物DNA、細菌DNA、ウイルスDNA、真菌DNA、原生動物DNA、又はそれらの組み合わせを含む、請求項33～36のいずれか1項に記載の反応混合物。

【請求項38】

標的dsDNAが、ゲノムDNA、ミトコンドリアDNA、プラスミドDNA、又はそれらの組み合わせである、請求項33～37のいずれか1項に記載の反応混合物。

【請求項39】

試料核酸が、生物学的試料、臨床試料、環境試料、又はそれらの組み合わせ由来である、請求項33～38のいずれか1項に記載の反応混合物。

30

【請求項40】

生物学的試料が、便、痰、末梢血、血漿、血清、リンパ節、呼吸組織、滲出液、体液、又はそれらの組み合わせを含む、請求項39に記載の反応混合物。

【請求項41】

核酸をタグ付けする方法であって、

請求項1～32のいずれか1項に記載の組成物を、複数の標的二本鎖DNA(dsDNA)を含むことが疑われる試料と接触させて、反応混合物を形成することと、

反応混合物をインキュベートして、複数のdsDNA断片であってその各々が、各末端にそれぞれ複数のタンパク質複合体のうちの1つの第1のアダプター及び第2のアダプターを含む前記複数のdsDNA断片を生成すること

40

を含む方法。

【請求項42】

シーケンシングライブラリーを生成する方法であって、

請求項1～32のいずれか1項に記載の組成物を、複数の標的二本鎖DNA(dsDNA)を含むことが疑われる試料と接触させて、反応混合物を形成することと、

反応混合物をインキュベートして、複数のdsDNA断片であってその各々が、各末端にそれぞれ複数のタンパク質複合体のうちの1つの第1のアダプター及び第2のアダプターを含む前記複数のdsDNA断片を生成することと、

dsDNA断片の末端でアダプターに結合することができるプライマーを用いて、複数のdsDNA断片を増幅して、シーケンシングライブラリーを生成することと

50

を含む方法。

【請求項 4 3】

プライマーの各々は、長さが約 5 ~ 8 0 ヌクレオチドである、請求項 4 2 に記載の方法。

【請求項 4 4】

プライマーを用いて複数の dsDNA 断片を増幅することが、ポリメラーゼ連鎖反応 (PCR) を用いて行われる、請求項 4 2 ~ 4 3 のいずれか 1 項に記載の方法。

【請求項 4 5】

PCR が、ループ媒介等温増幅 (LAMP)、ヘリカーゼ依存性増幅 (HDA)、組換えポリメラーゼ増幅 (RPA)、鎖置換増幅 (SDA)、核酸配列ベース増幅 (NASBA)、転写媒介増幅 (TMA)、ニックング酵素増幅反応 (NEAR)、ローリングサークル増幅 (RCA)、多置換増幅 (MDA)、分岐化 (RAM)、環状ヘリカーゼ依存性増幅 (cHDA)、単一プライマー等温増幅 (SPIA)、RNA 技術のシグナル媒介増幅 (SMART)、自己持続配列複製 (3SR)、ゲノム指数増幅反応 (GEAR)、又は等温多置換増幅 (IMDA) である、請求項 4 4 に記載の方法。

10

【請求項 4 6】

PCR が、リアルタイム PCR 又は定量的リアルタイム PCR (QRT-PCR) である、請求項 4 4 に記載の方法。

【請求項 4 7】

試料が、真核生物 DNA、細菌 DNA、ウイルス DNA、真菌 DNA、原生動物 DNA、又はそれらの組み合わせを含む、請求項 4 1 ~ 4 6 のいずれか 1 項に記載の方法。

20

【請求項 4 8】

複数の標的 dsDNA が、ゲノム DNA、ミトコンドリア DNA、プラスミド DNA、又はそれらの組み合わせを含む、請求項 4 1 ~ 4 7 のいずれか 1 項に記載の方法。

【請求項 4 9】

試料が、生物学的試料、臨床試料、環境試料、若しくはそれらの組み合わせであるか又はそれに由来する、請求項 4 1 ~ 4 8 のいずれか 1 項に記載の方法。

【請求項 5 0】

複数の標的 dsDNA が、少なくとも 2 つの異なる生物由来の DNA を含む、請求項 4 1 ~ 4 9 のいずれか 1 項に記載の方法。

30

【請求項 5 1】

複数の標的 dsDNA が、少なくとも 2 つの異なる遺伝子由来の DNA を含む、請求項 4 1 ~ 5 0 のいずれか 1 項に記載の方法。

【請求項 5 2】

逆転写酵素を用いて複数の標的 RNA から複数の標的 dsDNA を生成することをさらに含む、請求項 4 1 ~ 5 1 のいずれか 1 項に記載の方法。

【請求項 5 3】

複数の標的 dsDNA が、逆転写酵素を用いて標的 RNA から生成された標的 dsDNA を含む、請求項 4 1 ~ 5 1 のいずれか 1 項に記載の方法。

【請求項 5 4】

複数の標的 dsDNA が、目的とする遺伝子シグネチャを含む、請求項 4 1 ~ 5 3 のいずれか 1 項に記載の方法。

40

【請求項 5 5】

目的とする遺伝子シグネチャが、目的とする 1 種以上の突然変異を含む、請求項 5 4 に記載の方法。

【請求項 5 6】

目的とする 1 種以上の突然変異が、点突然変異、逆位、欠失、挿入、転座、複製、コピー数変動、又はそれらの組み合わせを含む、請求項 5 5 に記載の方法。

【請求項 5 7】

目的とする 1 種以上の突然変異が、ヌクレオチド置換、欠失、挿入、又はそれらの組み

50

合わせを含む、請求項 55 に記載の方法。

【請求項 58】

目的とする遺伝子シグネチャが、標的 dsDNA が由来する生物の病原体同定、抗生物質耐性又は抗生物質感受性を示す、請求項 54 ~ 57 のいずれか 1 項に記載の方法。

【請求項 59】

目的とする遺伝子シグネチャが、標的 dsDNA が由来する生物のがん状態を示す、請求項 54 ~ 57 のいずれか 1 項に記載の方法。

【請求項 60】

目的とする遺伝子シグネチャが、標的 dsDNA が由来する生物の遺伝的疾患の状態を示す、請求項 54 ~ 57 のいずれか 1 項に記載の方法。

10

【請求項 61】

遺伝的疾患が単一遺伝子疾患である、請求項 60 に記載の方法。

【請求項 62】

遺伝子疾患が、嚢胞性線維症、ハンチントン病、鎌状赤血球貧血、血友病、デュシェンヌ型筋ジストロフィー、サラセミア、脆弱 X 症候群、家族性高コレステロール血症、多嚢胞性腎疾患、神経線維腫症 I 型、遺伝性球状赤血球症、マルファン症候群、テイ-サックス病、フェニルケトン尿症、ムコ多糖症、リソソーム酸性リパーゼ欠損症、グリコーゲン貯蔵疾患、ガラクトース血症、又はヘモクロマトーシスである、請求項 60 に記載の方法。

【請求項 63】

20

複数の標的 dsDNA を複数のタンパク質複合体対と接触させることが、約 25 ~ 約 80 で行われる、請求項 41 ~ 62 のいずれか 1 項に記載の方法。

【請求項 64】

反応混合物をインキュベートすることが、反応混合物を約 37 ~ 約 55 でインキュベートすることを含む、請求項 41 ~ 63 のいずれか 1 項に記載の方法。

【請求項 65】

複数のタンパク質複合体対及び複数の標的 dsDNA が、約 2 : 1 ~ 約 2,000 : 1 の分子比で反応混合物中に存在する、請求項 41 ~ 64 のいずれか 1 項に記載の方法。

【請求項 66】

複数のタンパク質複合体対及び複数の標的 dsDNA が、約 2 : 1 ~ 約 200 : 1 の分子比で反応混合物中に存在する、請求項 41 ~ 64 のいずれか 1 項に記載の方法。

30

【請求項 67】

複数の dsDNA 断片のうちの 1 つ以上の一方又は両方の末端を標識することをさらに含む、請求項 41 ~ 66 のいずれか 1 項に記載の方法。

【請求項 68】

複数の dsDNA 断片のうちの 1 つ以上の 2 つの末端を異なるように標識することを含む、請求項 41 ~ 66 のいずれか 1 項に記載の方法。

【請求項 69】

標識が、アニオン標識、カチオン標識、中性標識、電気化学標識、タンパク質標識、蛍光標識、磁性標識、又はそれらの組み合わせによる標識を含む、請求項 67 ~ 68 のいずれか 1 項に記載の方法。

40

【請求項 70】

標識された dsDNA 断片を濃縮すること、標識された dsDNA 断片を捕捉すること、標識された dsDNA 断片を単離すること、及び / 又は標識された dsDNA 断片を視覚化することをさらに含む、請求項 67 ~ 69 のいずれか 1 項に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願

本出願は、2021年5月14日に提出された米国仮特許出願第 63 / 189,032

50

号、及び2021年9月13日に出願された米国仮特許出願第63/243,443号の35 U.S.C. § 119(e)に基づく利益を主張し、これらの関連出願の内容は、すべての目的のためにその全体が参照により本明細書に組み込まれる。

配列表への言及

本出願は、電子フォーマットによる配列表とともに提出される。配列表は、2022年5月12日に作製された68EB__317326__WO__Sequence__Listingという名称のファイルとして提供され、サイズは56.0キロバイトである。配列表の電子フォーマットにおける情報は、その全体が参照により本明細書に組み込まれる。

背景

分野

本開示は、一般的に、分子生物学の分野に関し、例えば、カスタマイズされた遺伝子座特異的シーケンシングライブラリーを生成するために核酸をタグ付けすることに関する。

【背景技術】

【0002】

関連技術の説明

核酸をシーケンシングするための従来のライブラリー調製方法は、作製するのに数時間を要し得、このプロセスは、ランダムに作製されるライブラリーを生成する。これらのライブラリーがランダムである理由は、核酸（物理的、酵素的、及び化学的断片化法を含む）を断片化するために使用される方法がランダムに核酸を断片化するからである。したがって、DNAシーケンシングのアウトプットは制御することができない。現在、標的化されたシーケンシングに使用されている2つの方法がある。第1のものは、アンプリコンシーケンシングである。この方法は、DNA増幅を介して目的とする領域を増幅するためにプライマーを使用することに依存する。この追加の増幅ステップは、標準的なライブラリー調製方法にさらなるコスト、時間、及び資源を追加する。第2の標的化されたシーケンシング方法は、標的捕捉である。この方法は、特定の核酸標的にハイブリダイズし得るようにプローブ又はプローブのプールを使用することに依存する。プローブのその標的へのハイブリダイゼーション及びこれらの標的の単離は、数日を要し得る時間のかかるプロセスである。さらに、この方法において使用されるプローブは、合成するのに費用がかかる。カスタマイズされた遺伝子座特異的ライブラリー調製のための組成物、方法、システム、及びキットが必要である。迅速な標的化されたシーケンシング（したがって、迅速なシーケンシングに基づく診断、例えば、2時間未満）を可能にする方法、組成物、キット、及びシステム、並びに適切な治療アプローチの同時診断及び決定を提供することができるセラノスティクス（theranostics）に対する必要性がある。

【発明の概要】

【0003】

概要

本明細書に開示されるものには、組成物が含まれる。一部の実施形態では、組成物は、複数のタンパク質複合体を含む。一部の実施形態では、複数のタンパク質複合体の各々は、トランスポソームと、標的二本鎖DNA（dsDNA）上の結合部位に特異的に結合することができるプログラム可能なDNA結合ユニットとを含む。一部の実施形態では、トランスポソームは、トランスポゼース、第1のアダプター及び第2のアダプターを含む。一部の実施形態では、複数のタンパク質複合体の各々に対する結合部位は、互いに異なる。

【0004】

一部の実施形態では、複数のタンパク質複合体のうちの少なくとも2つは、同じトランスポソームを含む。一部の実施形態では、複数のタンパク質複合体のすべては、同じトランスポソームを含む。一部の実施形態では、複数のタンパク質複合体のすべては、同じトランスポゼースを含む。一部の実施形態では、同じトランスポソーム内の第1のアダプターと第2のアダプターは同じである。一部の実施形態では、異なるトランスポソームにおける第1のアダプター、第2のアダプター、又はその両方は異なる。一部の実施形態では

10

20

30

40

50

、第1のアダプター、第2のアダプター、又はその両方は、dsDNA又はRNA/DNA二重鎖である。一部の実施形態では、アダプターは、長さが約3～200塩基対である。一部の実施形態では、第1のアダプター、第2のアダプター、又はその両方は、シーケンシングアダプターである。一部の実施形態では、シーケンシングアダプターは、P5又はP7プライマー配列を含む。

【0005】

一部の実施形態では、複数のタンパク質複合体のうち少なくとも2つの結合部位は、同じ標的dsDNA上にある。一部の実施形態では、複数のタンパク質複合体のうち少なくとも2つの結合部位は、同じ標的dsDNA上で約1～50000ヌクレオチド離れている。一部の実施形態では、複数のタンパク質複合体の一对の結合部位間の距離は、複数のタンパク質複合体の別の一对の結合部位間の距離と実質的に同じである。一部の実施形態では、複数のタンパク質複合体の一对の結合部位間の距離は、複数のタンパク質複合体の別の一对の結合部位間の距離と異なる。一部の実施形態では、複数のタンパク質複合体のうち少なくとも2つの結合部位は、標的dsDNAの異なる鎖上にある。一部の実施形態では、複数のタンパク質複合体のうち少なくとも2つは、異なる標的dsDNAに特異的に結合することができる。

10

【0006】

一部の実施形態では、複数のタンパク質複合体は、約2～5000の標的dsDNAに特異的に結合することができる。一部の実施形態では、トランスポゼースは、Tn5トランスポゼース、Tn7トランスポゼース、マリナーTc1様トランスポゼース、Himar1C9トランスポゼース、又はスリーピングビューティトランスポゼースである。一部の実施形態では、トランスポゼースは、高活性トランスポゼースである。一部の実施形態では、プログラム可能なDNA結合ユニットは、ヌクレアーゼ欠損CRISPR関連タンパク質(dCASタンパク質)と、標的dsDNAの結合部位に特異的に結合することができるガイドRNA(gRNA)とを含む。一部の実施形態では、トランスポソームは、トランスポゼース及びdCASタンパク質を連結するリンカーを介して、プログラム可能なDNA結合ユニットと結び付けられている。一部の実施形態では、リンカーは、ペプチドリンカー、化学リンカー、又はその両方を含む。一部の実施形態では、トランスポゼースは、dCASタンパク質を含む融合タンパク質中に存在する。一部の実施形態では、dCASタンパク質は、dCAS9、dCAS12、dCAS13、dCAS14、又はSpRY dCASである。一部の実施形態では、dCAS13タンパク質は、dCAS13a、dCAS13b、dCAS13c、又はdCAS13dである。

20

30

【0007】

一部の実施形態では、プログラム可能なDNA結合ユニットは、標的dsDNA上の結合部位に特異的に結合することができるタンパク質構成成分を含む。一部の実施形態では、タンパク質構成成分は、エンドヌクレアーゼ欠損ジンクフィンガーヌクレアーゼ(ZFN)、エンドヌクレアーゼ欠損転写アクチベーター様エフェクターヌクレアーゼ(TALEN)、アルゴノートタンパク質、エンドヌクレアーゼ欠損メガヌクレアーゼ、リコンビナーゼ、又はそれらの組み合わせを含む。一部の実施形態では、トランスポソームは、トランスポゼースとタンパク質構成成分を接続するリンカーを介して、プログラム可能なDNA結合ユニットと結び付けられている。一部の実施形態では、リンカーは、ペプチドリンカー、化学リンカー、又はその両方を含む。一部の実施形態では、ペプチドリンカーは、複数のグリシン、セリン、スレオニン、アラニン、リジン、グルタミン、又はそれらの組み合わせを含む。一部の実施形態では、ペプチドリンカーは、GSリンカーを含む。一部の実施形態では、ペプチドリンカーは、XTENリンカーである。一部の実施形態では、タンパク質構成成分は、トランスポゼースを含む融合タンパク質中に存在する。

40

【0008】

本明細書に開示されるものには、反応混合物が含まれる。一部の実施形態では、反応混合物は、本明細書に開示される組成物及び1種以上の標的dsDNAを含むことが疑われる試料核酸を含む。反応混合物は、DNAポリメラーゼ、dNTP、又はそれらの組み合

50

わせを含むことができる。一部の実施形態では、アダプターは、標的 dsDNA 又はその断片に共有結合的に付着される。反応混合物は、複数の dsDNA 断片を含むことができ、各々が、各末端にそれぞれ複数のタンパク質複合体のうちの 1 つの第 1 のアダプター及び第 2 のアダプターを含む。一部の実施形態では、試料核酸は、真核生物 DNA、細菌 DNA、ウイルス DNA、真菌 DNA、原生動物 DNA、又はそれらの組み合わせを含む。一部の実施形態では、標的 dsDNA は、ゲノム DNA、ミトコンドリア DNA、プラスミド DNA、又はそれらの組み合わせである。一部の実施形態では、試料核酸は、生物学的試料、臨床試料、環境試料、又はそれらの組み合わせ由来である。一部の実施形態では、生物学的試料は、便、痰、末梢血、血漿、血清、リンパ節、呼吸組織、滲出液、体液、又はそれらの組み合わせを含む。

10

本明細書に開示されるものには、核酸をタグ付けする方法が含まれる。一部の実施形態では、本方法は、本明細書に開示される組成物を、複数の標的 dsDNA を含むことが疑われる試料と接触させて反応混合物を形成することと、反応混合物をインキュベートして複数の dsDNA 断片を生成することとであって、各々が、各末端にそれぞれ複数のタンパク質複合体のうちの 1 つの第 1 のアダプター及び第 2 のアダプターを含むこととを含む。

【0009】

本明細書に開示されるものには、シーケンシングライブラリーを生成する方法が含まれる。一部の実施形態では、本方法は、本明細書に開示される組成物を、複数の標的 dsDNA を含むことが疑われる試料と接触させて、反応混合物を形成することを含む。本方法は、反応混合物をインキュベートして複数の dsDNA 断片を生成することとであって、各々が、各末端にそれぞれ複数のタンパク質複合体のうちの 1 つの第 1 のアダプター及び第 2 のアダプターを含むこととを含むことができる。本方法は、dsDNA 断片の末端でアダプターに結合することができるプライマーを用いて、複数の dsDNA 断片を増幅して、シーケンシングライブラリーを生成することとを含むことができる。

20

【0010】

一部の実施形態では、プライマーの各々は、約 5 ~ 80 ヌクレオチドの長さである。一部の実施形態では、プライマーを用いて複数の dsDNA 断片を増幅することは、ポリメラーゼ連鎖反応 (PCR) を用いて行われる。一部の実施形態では、PCR は、ループ媒介等温増幅 (LAMP)、ヘリカーゼ依存性増幅 (HDA)、組換えポリメラーゼ増幅 (RPA)、鎖置換増幅 (SDA)、核酸配列ベース増幅 (NASBA)、転写媒介増幅 (TMA)、ニックング酵素増幅反応 (NEAR)、ローリングサークル増幅 (RCA)、多置換増幅 (MDA)、分岐化 (RAM)、環状ヘリカーゼ依存性増幅 (cHDA)、単一プライマー等温増幅 (SPIA)、RNA 技術のシグナル媒介増幅 (SMART)、自己持続配列複製 (3SR)、ゲノム指数増幅反応 (GEAR)、又は等温多置換増幅 (IMDA) である。一部の実施形態では、PCR は、リアルタイム PCR 又は定量的リアルタイム PCR (QRT-PCR) である。一部の実施形態では、試料は、真核生物 DNA、細菌 DNA、ウイルス DNA、真菌 DNA、原生動物 DNA、又はそれらの組み合わせを含む。

30

【0011】

一部の実施形態では、複数の標的 dsDNA は、ゲノム DNA、ミトコンドリア DNA、プラスミド DNA、又はそれらの組み合わせを含む。一部の実施形態では、試料は、生物学的試料、臨床試料、環境試料、又はそれらの組み合わせであるか又はそれに由来する。一部の実施形態では、複数の標的 dsDNA は、少なくとも 2 つの異なる生物由来の DNA を含む。一部の実施形態では、複数の標的 dsDNA は、少なくとも 2 つの異なる遺伝子由来の DNA を含む。本方法は、逆転写酵素を用いて複数の標的 RNA から複数の標的 dsDNA を生成することとを含むことができる。一部の実施形態では、複数の標的 dsDNA は、逆転写酵素を用いて標的 RNA から生成された標的 dsDNA を含む。

40

【0012】

一部の実施形態では、複数の標的 dsDNA は、目的とする遺伝子シグネチャ (gene signature) を含む。一部の実施形態では、目的とする遺伝子シグネチャは、1 つ以上

50

の目的とする突然変異を含む。一部の実施形態では、目的とする1つ以上の突然変異は、点突然変異、逆位、欠失、挿入、転座、複製、コピー数変異、又はそれらの組み合わせを含む。一部の実施形態では、目的とする1つ以上の突然変異は、ヌクレオチド置換、欠失、挿入、又はそれらの組み合わせを含む。一部の実施形態では、目的とする遺伝子シグネチャは、標的 dsDNA が由来する生物の抗生物質耐性又は抗生物質感受性を示す (indicative)。一部の実施形態では、目的とする遺伝子シグネチャは、標的 dsDNA が由来する生物のがん状態を示す。一部の実施形態では、目的とする遺伝子シグネチャは、標的 dsDNA が由来する生物の遺伝的疾患の状態を示す。一部の実施形態では、遺伝子疾患は単一遺伝子疾患 (single-gene disorders) である。一部の実施形態では、遺伝子疾患は、嚢胞性線維症、ハンチントン病、鎌状赤血球貧血、血友病、デュシェンヌ型筋ジストロフィー、サラセミア、脆弱 X 症候群、家族性高コレステロール血症、多嚢胞性腎疾患、神経線維腫症 I 型、遺伝性球状赤血球症、マルファン症候群、テイ-サックス病、フェニルケトン尿症、ムコ多糖症、リソソーム (ライソゾーム) 酸性リパーゼ欠損症、グリコーゲン貯蔵疾患、ガラクトース血症、又はヘモクロマトーシスである。

10

【0013】

一部の実施形態では、複数の標的 dsDNA を複数のタンパク質複合体対と接触させることは、約 25 ~ 約 80 で行われる。一部の実施形態では、反応混合物をインキュベートすることは、反応混合物を約 37 ~ 約 55 でインキュベートすることを含む。一部の実施形態では、複数のタンパク質複合体対及び複数の標的 dsDNA は、約 2 : 1 ~ 約 2,000 : 1 の分子比で反応混合物中に存在する。一部の実施形態では、複数のタン

20

【0014】

本方法は、複数の dsDNA 断片のうちの一つ以上の一方又は両方の末端を標識することを含むことができる。本方法は、複数の dsDNA 断片のうちの一つ以上の二つの末端を異なるように標識することを含むことができる。一部の実施形態では、標識は、アニオン標識、カチオン標識、中性標識、電気化学標識、タンパク質標識、蛍光標識、磁性標識、又はそれらの組み合わせによる標識を含む。本方法は、標識された dsDNA 断片を濃縮すること、標識された dsDNA 断片を捕捉すること、標識された dsDNA 断片を単離すること、及び/又は標識された dsDNA 断片を視覚化することを含むことができる。

30

【図面の簡単な説明】

【0015】

【図1】次世代シーケンシングのための非限定的であり例示的な従来のライブラリー調製方法を示す図である。図示されたライゲーションに基づくライブラリーの調製は、www.idtdna.com/pages/technology/next-generation-sequencing/library-preparation/ligation-based-library-prep から再調製される。

【図2】IlluminaからのNextera XT Library Prep: Tips及びTroubleshooting (2015) から再調製される、タグメン

40

【図3】本明細書に開示されるカスタマイズされた遺伝子座特異的ライブラリー調製 (CLLP) の非限定的であり例示的な概略図を示す。

【図4】ゲノム編集ツール (Cas9及びガイドRNA) を使用する標的化されたシーケンシングの非限定的であり例示的な実施形態を示す図である。

【図5A - 5F】本明細書に開示されるカスタマイズされた遺伝子座特異的ライブラリー調製 (CLLP) の非限定的であり例示的な実施形態を示す図である。

【図6】ONT迅速シーケンシングキット、タグメンテーションベースを示す非限定的であり例示的な実施形態を示す図である。Nanoporeからの迅速シーケンシングキッ

50

トに記載されるワークフローから再調製される。

【図7A - 7H】は、Oxford Nanoporeからのものなどの、既存のシーケンシングプラットフォームにおいて使用するためのシーケンシングライブラリーを生成するためのゲノム編集タグメンテーション (GET) の非限定的であり例示的な実施形態を示す図である。

【図8】本明細書に提供されるタンパク質複合体の生成に使用するためのプラスミド構築物 (3XFlag-Cas9-Fl26-Tn5; 配列番号1) の非限定的であり例示的な概略図を示す。

【図9】本明細書に提供されるタンパク質複合体の生成に使用するためのプラスミド構築物 (3XFlag-Cas9-xTen-Tn5; 配列番号2) の非限定的であり例示的な概略図を示す。

10

【図10】本明細書に提供されるタンパク質複合体の生成に使用するためのプラスミド構築物 (pET-Tn5-xTen-dCas9; 配列番号3) の非限定的であり例示的な概略図を示す。

【図11】S.エンテリカ (S. enterica) InvA 遺伝子に対する例示的な sgRNA の相対的結合部位を示す図である。

【図12】S.エンテリカ FljC 遺伝子に対する例示的な sgRNA の相対的結合部位を示す図である。

【図13】ゲノムDNAの切断が予想されるサイズに特異的であることを示す例示的なバイオアナライザーデータのグラフを示し、サルモネラ・エンテリカ (Salmonella enterica) 用のガイドRNAが機能的であることを示す図である。また、表2も参照されたい。

20

【図14】プライマーとしてアダプターAを使用したTn5生成断片の増幅を示すテーブルステーション分析のグラフを示す。これは、アダプターが切断された分子の5'末端及び3'末端に付加されたことを示す。

【図15】プライマーとしてアダプターBを使用したTn5生成断片の増幅を示すテーブルステーション分析のグラフを示す。これは、アダプターが切断された分子の5'末端及び3'末端に付加されたことを示す。

【図16】組換え的に発現及び精製されたdCas9-Fl26-Tn5融合タンパク質の例示的なSDS-PAGEゲル分析を示す図である。矢印は融合タンパク質バンドを指す。

30

【図17】組換え的に発現及び精製されたdCas9-Fl26-Tn5融合タンパク質の例示的な電気泳動ゲルのバイオアナライザー分析を示す図である。

【図18】組換え的に発現及び精製されたdCas9-xTen-Tn5融合タンパク質の例示的なSDS-PAGEゲル分析を示す図である。矢印は融合タンパク質バンドを指す。

【図19】組換え的に発現及び精製されたdCas9-xTen-Tn5融合タンパク質の例示的な電気泳動分析からのバイオアナライザーデータを示す図である。

【図20】組換え的に発現及び精製されたTn5-Fl26-dCas9融合タンパク質の例示的なSDS-PAGEゲル分析を示す図である。矢印は融合タンパク質バンドを指す。

40

【図21】組換え的に発現及び精製されたTn5-xTen-dCas9融合タンパク質の例示的なSDS-PAGEゲル分析を示す図である。矢印は融合タンパク質バンドを指す。

【図22】触媒的に活性なCas9のみ (融合タンパク質なし) を使用する増幅反応のテーブルステーション分析を示す図である。増幅は観察されず、Cas9自身が消化された断片の5'末端及び3'末端にアダプターを付加することができないことを示唆する。可視シグナルは、Cas9とともにインキュベートされた試料由来であるが、PCRに供されない。下のピークは100bpサイズのマーカーであり、上のピークはゲノムDNAである。

50

【図23】 dCas9 - F126 - Tn5 とのタグメンテーション反応後の増幅反応のテーブステーション分析を示す図である。矢印は、Cas9 - Tn5 融合タンパク質とのインキュベーション後にPCR条件に供された反応からのシグナルを示す。この反応にはgRNAは含まれず、ランダムなタグメンテーションを示す広いピークをもたらした。下のピークは100bpサイズのマーカーであり、上のピークはゲノムDNAである。

【図24】 dCas9 - xTen - Tn5 とのタグメンテーション反応後の増幅反応の例示的なテーブステーション分析を示す図である。矢印は、Cas9 - Tn5 融合タンパク質とのインキュベーション後にPCR条件に供された反応からのシグナルを示す。この反応にはgRNAは含まれず、ランダムなタグメンテーションを示す広いピークをもたらした。下のピークは100bpサイズのマーカーであり、上のピークはゲノムDNAである

10

。【図25】 100nMのdCas9 - F126 - Tn5 融合タンパク質によるタグメンテーション後の増幅反応のテーブステーション分析を示す図である。矢印は、Cas9 - Tn5 融合タンパク質とのインキュベーション後にPCR条件に供された反応からのシグナルを示す。下のピークは100bpサイズのマーカーであり、上のピークはゲノムDNAである。

【図26】 1nMのdCas9 - F126 - Tn5 融合タンパク質とのタグメンテーション反応後の増幅反応のテーブステーション分析を示す図である。矢印は、Cas9 - Tn5 融合タンパク質とのインキュベーション後にPCR条件に供された反応からのシグナルを示す。下のピークは100bpサイズのマーカーであり、上のピークはゲノムDNAである。

20

【図27】 100pMのdCas9 - F126 - Tn5 融合タンパク質とのタグメンテーション反応後の増幅反応のテーブステーション分析を示す図である。矢印は、Cas9 - Tn5 融合タンパク質とのインキュベーション後にPCR条件に供された反応からのシグナルを示す。下のピークは100bpサイズのマーカーであり、上のピークはゲノムDNAである。

【図28】 拡大された図27からの100pMのdCas9 - F126 - Tn5 融合タンパク質とのタグメンテーション反応後の増幅反応のテーブステーション分析を示す図である。

【図29】 100pMのdCas9 - xTen - Tn5 融合タンパク質とのタグメンテーション反応後の増幅反応のテーブステーション分析を示す図である。より低い、100bpより低いマーカー。

30

【図30】 10pMのdCas9 - xTen - Tn5 融合タンパク質とのタグメンテーション反応後の増幅反応のテーブステーション分析を示す図である。より低い、100bpより低いマーカー。

【図31】 1pMのdCas9 - xTen - Tn5 融合タンパク質とのタグメンテーション反応後の増幅反応のテーブステーション分析を示す図である。

【図32】 1つのアダプター（アダプターB）のみをロードした、Tn5のみのタグメンテーションによって調製されたライブラリーからの増幅のバイオアナライザー分析を示す図である。

40

【図33】 1つのアダプター（アダプターB）のみをロードした、dCas9 - F126 - Tn5 誘導タグメンテーションによって調製されたライブラリーからの増幅のバイオアナライザー分析を示す図である。この実験では、より短いインキュベーションプロトコルを使用した。

【図34】 1つのアダプター（アダプターB）のみをロードした、dCas9 - F126 - Tn5 誘導タグメンテーションによって調製されたライブラリーからの増幅のバイオアナライザー分析を示す図である。この実験では、より長いインキュベーションプロトコルを使用した。

【図35】 アダプターAとBの両方をロードした、dCas9 - F126 - Tn5 誘導タグメンテーションによって調製されたライブラリーからの増幅の例示的なバイオアナライ

50

ザー分析を示す図である。この実験では、より長いインキュベーションプロトコルを使用した。

【図36】アダプターAとBの両方をロードした、dCas9-Fl26-Tn5誘導タグメンテーションによって調製されたライブラリーからの増幅の例示的なバイオアナライザー分析を示す図である。この実験では、より短いインキュベーションプロトコルを使用した。

【図37】本明細書に開示されるCasTn-NEBNextライゲーションベースのライブラリー法を用いて、NGS配列アダプターで標識されたDNA断片の例示的な実施形態を示す図である。

【図38】S.エンテリカsgRNAをロードした、dCas9-xTen-Tn5とともにインキュベートしたS.エンテリカゲノムDNA試料からのPCR増幅の例示的なテーブレーション分析を示す図である。より低い、100bpより低いマーカー。 10

【図39】sgRNAなしでdCas9-xTen-Tn5とともにインキュベートしたS.エンテリカ試料からのPCR増幅の例示的なテーブレーション分析を示す図である。より低い、100bpより低いマーカー。

【図40】単一のアダプター（例えば、アダプターB）を使用するdCas9-Tn5生成断片の例を示す図である。

【図41】Tn5が2つの異なるアダプター（例えば、アダプターA及びアダプターB）がロードされた反応から生成されたdCas9-Tn5断片を示す図である。

【図42A-42B】次世代シーケンシングのためのNEBNextライゲーションベースのライブラリー調製の例を示す図である。アダプター及びプライマー配列の主要標識部分に示される記号。NEBNextライブラリー調製によるdCas9-Tn5タグメンテーションによって生成された断片を図37に示す。 20

【図43】次世代シーケンシングのためのタグメンテーションベースのNexteraライブラリー調製の例を示す図である。

【図44】dCas9-Tn5誘導タグメンテーションを使用したタグメンテーションベースのライブラリー調製を示す図である。

【発明を実施するための形態】

【0016】

詳細な説明

以下の詳細な説明において、その一部を形成する添付図面に言及する。図面においては、文脈上別段の指示がない限り、同様の記号は、典型的には、同様の構成成分を識別する。詳細な説明、図面、及び特許請求の範囲に記載される例示的な実施形態は、限定されるものではない。本明細書に提示される主題の精神又は範囲から逸脱することなく、他の実施形態を利用することができ、他の変更を行うことができる。本開示の態様は、本明細書に一般的に記載され、図面に例示されるように、多種多様な異なる形態で配置、置換、結合、分離、及び設計することができ、これらのすべては、本明細書中で明示的に企図され、本明細書中の開示の一部を構成することができることは容易に理解される。 30

すべての特許、公開された特許出願、他の刊行物、及びGenBank、並びに本明細書に言及される他のデータベースからの配列は、関連する技術に関して、それらの全体が参照により本明細書に組み込まれる。 40

本明細書に開示されるものには、組成物が含まれる。一部の実施形態では、組成物は、複数のタンパク質複合体を含む。一部の実施形態では、複数のタンパク質複合体の各々は、トランスポソームと、標的二本鎖DNA(dsDNA)上の結合部位に特異的に結合することができるプログラム可能なDNA結合ユニットとを含む。一部の実施形態では、トランスポソームは、トランスポゼース、第1のアダプター及び第2のアダプターを含む。一部の実施形態では、複数のタンパク質複合体の各々に対する結合部位は、互いに異なる。

【0017】

本明細書に開示されるものには、反応混合物が含まれる。一部の実施形態では、反応混 50

合物は、本明細書に開示される組成物及び1種以上の標的 dsDNA を含むことが疑われる試料核酸を含む。反応混合物は、DNAポリメラーゼ、dNTP、又はそれらの組み合わせを含むことができる。一部の実施形態では、アダプターは、標的 dsDNA 又はその断片に共有結合的に付着される。反応混合物は、複数の dsDNA 断片を含むことができ、各々が、各末端にそれぞれ複数のタンパク質複合体のうちの1つの第1のアダプター及び第2のアダプターを含む。一部の実施形態では、試料核酸は、真核生物DNA、細菌DNA、ウイルスDNA、真菌DNA、原生動物DNA、又はそれらの組み合わせを含む。一部の実施形態では、標的 dsDNA は、ゲノムDNA、ミトコンドリアDNA、プラスミドDNA、又はそれらの組み合わせである。一部の実施形態では、試料核酸は、生物学的試料、臨床試料、環境試料、又はそれらの組み合わせからのものである。一部の実施形態では、生物学的試料は、便、痰、末梢血、血漿、血清、リンパ節、呼吸組織、滲出液、体液、又はそれらの組み合わせを含む。

本明細書に開示されるものには、核酸をタグ付けする方法が含まれる。一部の実施形態では、本方法は、本明細書に開示される組成物を、複数の標的 dsDNA を含むことが疑われる試料と接触させて反応混合物を形成することと、反応混合物をインキュベートして複数の dsDNA 断片を生成することと、各々が、各末端にそれぞれ複数のタンパク質複合体のうちの1つの第1のアダプター及び第2のアダプターを含むこととを含む。

【0018】

本明細書に開示されるものには、シーケンシングライブラリーを生成する方法が含まれる。一部の実施形態では、本方法は、本明細書に開示される組成物を、複数の標的二本鎖DNA (dsDNA) を含むことが疑われる試料と接触させて、反応混合物を形成することを含む。本方法は、反応混合物をインキュベートして複数の dsDNA 断片を生成することと、各々が、各末端にそれぞれ複数のタンパク質複合体のうちの1つの第1のアダプター及び第2のアダプターを含むこととを含むことができる。本方法は、dsDNA 断片の末端でアダプターに結合することができるプライマーを用いて、複数の dsDNA 断片を増幅して、シーケンシングライブラリーを生成することとを含むことができる。

【0019】

定義

別段の定義がない限り、本明細書において使用される技術用語及び科学用語は、本開示が属する当業者によって一般的に理解されるものと同じ意味を有する。例えば、Singleton et al., Dictionary of Microbiology and Molecular Biology 2nd ed., J. Wiley & Sons (New York, NY 1994); Sambrook et al., Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press (Cold Spring Harbor, NY 1989) を参照されたい。本開示の目的のために、以下の用語を以下に定義する。

【0020】

本明細書で使用される場合、用語「アダプター」とは、結び付けられている核酸の増幅又はシーケンシングを容易にすることができる配列を意味することができる。結び付けられている核酸は、標的核酸を含むことができる。結び付けられている核酸は、空間標識、標的標識、試料標識、インデックス化標識、又はバーコード配列（例えば、分子標識）のうちの1種以上を含み得る。アダプターは、直鎖状であり得る。アダプターは、プレアデニル化アダプターであり得る。アダプターは、二本鎖でも一本鎖であり得る。1つ以上のアダプターは、核酸の5'末端又は3'末端に位置することができる。アダプターが5'末端及び3'末端に公知の配列を含む場合、公知の配列は同一又は異なる配列であり得る。ポリヌクレオチドの5'末端及び/又は3'末端に位置するアダプターは、表面上に固定化された1つ以上のオリゴヌクレオチドにハイブリダイズすることができる。アダプターは、一部の実施形態では、ユニバーサル配列を含むことができる。ユニバーサル配列は、2つ以上の核酸分子に共通のヌクレオチド配列の領域であり得る。2つ以上の核酸分子はまた、異なる配列の領域を有することができる。したがって、例えば、5'アダプターは、同一及び/又はユニバーサル核酸配列を含み得、3'アダプターは、同一及び/又はユニバーサル配列を含み得る。複数の核酸分子の異なるメンバーに存在し得るユニバーサル配

列は、ユニバーサル配列に相補的な単一のユニバーサルプライマーを使用して、複数の異なる配列の複製又は増幅を可能にすることができる。同様に、核酸分子群の異なるメンバーに存在し得る少なくとも1つ、2つ（例えば、1対）又は複数のユニバーサル配列は、ユニバーサル配列に相補的である少なくとも1つ、2つ（例えば、1対）又は複数の単一ユニバーサルプライマーを使用して、複数の異なる配列の複製又は増幅を可能にすることができる。したがって、ユニバーサルプライマーは、このようなユニバーサル配列にハイブリダイズし得る配列を含む。標的核酸配列を有する分子は、異なる標的核酸配列の一方又は両方の末端にアダプター（例えば、非標的核酸配列）を付着させるように修飾され得る。標的核酸に付着した1つ以上のユニバーサルプライマーは、ユニバーサルプライマーのハイブリダイゼーションのための部位を提供することができる。標的核酸に付着した1つ以上のユニバーサルプライマーは、互いに同一であるか又は異なるものであり得る。

10

【0021】

本明細書中使用される場合、用語「結び付けられている」又は「と結び付けられている」とは、2つ以上の種がある時点で共存しているものとして同定可能であることを意味することができる。結び付きとは、2つ以上の種が同様の容器内にあるか又はその内にあったことを意味し得る。結び付きは、インフォマティクス上の結び付きであり得る。例えば、2つ以上の種に関するデジタル情報を保存することができ、1つ以上の種がある時点で共存していたことを決定するために使用することができる。結び付きはまた、物理的結び付きであり得る。一部の実施形態では、2つ以上の結び付けられている種は、互いに又は共通の固体若しくは半固体表面に「係留」、「付着」、又は「固定」されている。結び付きとは、標識をビーズなどの固体又は半固体支持体に付着させるための共有結合又は非共有結合手段を指すことができる。結び付きは、標的と標識の間の共有結合であり得る。結び付きは、2つの分子（標的分子及び標識など）間のハイブリダイゼーションを含み得る。

20

【0022】

本明細書で使用される場合、用語「相補的」とは、2つのヌクレオチド間の正確な対形成の能力を指すことができる。例えば、核酸の所定の位置にあるヌクレオチドが、他の核酸のヌクレオチドと水素結合し得る場合、2つの核酸は、その位置で互いに相補的であるとみなされる。2つの一本鎖核酸分子間の相補性は、ヌクレオチドの一部のみが結合する「部分的」であり得るか、又は一本鎖分子間に全体の相補性が存在する場合に完全であり得る。第1のヌクレオチド配列は、第1のヌクレオチド配列が第2のヌクレオチド配列に相補的である場合、第2の配列の「相補体」と言うことができる。第1のヌクレオチド配列が第2の配列の逆（すなわち、ヌクレオチドの順序が逆である）の配列と相補的である場合、第1のヌクレオチド配列は、第2の配列の「逆相補体」と言うことができる。本明細書で使用される場合、「相補的」配列は、配列の「相補体」又は「逆相補体」を指すことができる。分子が別の分子とハイブリダイズすることができる場合、それはハイブリダイズしている分子に対して相補的であり得るか、又は部分的に相補的であり得ることが本開示から理解される。

30

【0023】

本明細書で使用される場合、用語「標識」又は「標識（複数）」とは、試料内で標的と結び付けられている核酸コードを指すことができる。標識は、例えば、核酸標識であり得る。標識は、完全に又は部分的に増幅可能な標識であり得る。標識は、完全に又は部分的に連続した標識であり得る。標識は、明らかのように同定可能である天然核酸の一部であり得る。標識は、公知の配列であり得る。標識は、核酸配列の接合部、例えば、天然配列及び非天然配列の接合部を含むことができる。本明細書で使用される場合、用語「標識」は、用語「インデックス」、「タグ」、又は「標識タグ」と互換的に使用することができる。標識は情報を伝達することができる。例えば、様々な実施形態では、標識を使用して、試料の同一性、試料の供給源、細胞の同一性、及び/又は標的を決定することができる。

40

【0024】

50

本明細書で使用される場合、用語「核酸」は、ポリヌクレオチド配列、又はその断片を指す。核酸は、ヌクレオチドを含み得る。核酸は、細胞に対して外因性又は内因性であり得る。核酸は、無細胞環境中に存在し得る。核酸は、遺伝子又はその断片であり得る。核酸は、DNAであり得る。核酸は、RNAであり得る。核酸は、1種以上の類似体（例えば、改変された骨格、糖、又は核酸塩基）を含み得る。類似体の一部の非限定的な例としては、5-プロモウラシル、ペプチド核酸、キセノ核酸、モルホリノース、ロックド核酸、グリコール核酸、トレオース核酸、ジデオキシヌクレオチド、コルジセピン、7-デアザ-GTP、フルオロフォア（例えば、糖に連結されたローダミン又はフルオレセイン）、チオール含有ヌクレオチド、ピオチン連結ヌクレオチド、蛍光塩基類似体、CpGアイランド、メチル-7-グアノシン、メチル化ヌクレオチド、イノシン、チオウリジン、プソイドウリジン、ジヒドロウリジン、クエオシン、及びウヨシンが挙げられる。「核酸」、「ポリヌクレオチド」、「標的ポリヌクレオチド」、及び「標的核酸」は、互換的に使用することができる。

【0025】

核酸は、1種以上の修飾（例えば、塩基修飾、骨格修飾）を含み、新規又は増強された特徴（例えば、改良された安定性）を核酸に提供することができる。核酸は、核酸親和性タグを含むことができる。ヌクレオシドは塩基と糖の組み合わせであり得る。ヌクレオシドの塩基部分は複素環塩基であり得る。このような複素環塩基の最も一般的な2種類はプリンとピリミジンである。ヌクレオチドは、ヌクレオシドの糖部分に共有結合的に連結したリン酸基をさらに含むヌクレオシドであり得る。ペントフラノシル糖を含むそれらのヌクレオチドでは、リン酸基は糖の2'、3'、又は5'ヒドロキシル部分に連結することができる。核酸を形成する際に、リン酸基は、隣接するヌクレオチドを互いに共有結合的に連結させて、直鎖状高分子化合物を形成することができる。次に、この直鎖状高分子化合物のそれぞれの末端をさらに接続して環状化合物を形成することができるが、一般的に直鎖状化合物が適している。さらに、直鎖状化合物は、内部ヌクレオチド塩基相補性を有し得、したがって、完全に又は部分的に二本鎖の化合物を生成するように折り畳まれ得る。核酸内では、リン酸基は、一般的に、核酸のヌクレオチド間骨格の形成と称され得る。連結又は骨格は、3'から5'のホスホジエステル連結であり得る。

【0026】

核酸は、修飾された骨格及び/又は修飾されたヌクレオチド間連結を含むことができる。修飾された骨格には、骨格中にリン原子を保持するもの、及び骨格中にリン原子を有しないものが含まれ得る。そこにリン原子を含有する適切な修飾された核酸骨格には、例えば、ホスホロチオエート、キラルホスホロチオエート、ホスホロジチオエート、ホスホトリエステル、アミノアルキルホスホトリエステル、メチル及び他のアルキルホスホネート、例えば、3'-アルキレンホスホネート、5'-アルキレンホスホネート、キラルホスホネート、ホスフィネート、ホスホロアミデート、例えば、3'-アミノホスホロアミデート及びアミノアルキルホスホロアミデート、ホスホロジアミデート、チオノホスホロアミデート、チオノアルキルホスホネート、チオノアルキルホスホトリエステル、セレノホスフェート、及び正常な3'-5'連結を有するボラノホスフェート、2'-5'連結した類似体、及び1種以上のヌクレオチド間連結が3'から3'、5'から5'、又は2'から2'連結である逆極性を有するものが含まれ得る。

【0027】

核酸は、短鎖アルキル又はシクロアルキルヌクレオチド間連結、混合ヘテロ原子及びアルキル又はシクロアルキルヌクレオチド間連結、又は1つ以上の短鎖ヘテロ原子若しくは複素環ヌクレオチド間連結によって形成されるポリヌクレオチド骨格を含むことができる。これらは、モルホリノ連結（ヌクレオチドの糖部分から部分的に形成される）；シロキサン骨格；スルフィド、スルホキシド及びスルホン骨格；ホルムアセチル及びチオホルムアセチル骨格；メチレンホルムアセチル及びチオホルムアセチル骨格；リボアセチル骨格；アルキレン含有骨格；スルファメート骨格；メチレンイミノ及びメチレンヒドラジノ骨格；スルホネート及びスルホンアミド骨格；アミド骨格；並びに混合N、O、S及びCH

2構成成分部分を有する他のものを有するものを含むことができる。

【0028】

核酸は、核酸模倣物 (mimetic) を含むことができる。用語「模倣物」は、フラノース環のみ、又はフラノース環とヌクレオチド間連結の両方が非フラノース基で置換されているポリヌクレオチドを含むことを意図することができる。フラノース環のみの置換はまた糖代替物と称することができる。複素環塩基部分又は修飾された複素環塩基部分は、適切な標的核酸とのハイブリダイゼーションのために維持することができる。1つのこのような核酸はペプチド核酸 (PNA) であり得る。PNAでは、ポリヌクレオチドの糖骨格は、アミド含有骨格、特にアミノエチルグリシン骨格で置換することができる。ヌクレオチドは保持され得、直接的又は間接的に骨格のアミド部分のアザ窒素原子に結合される。PNA化合物の骨格は、PNAにアミド含有骨格を与える2種以上の連結アミノエチルグリシンユニットを含むことができる。複素環塩基部分は、骨格のアミド部分のアザ窒素原子に直接的又は間接的に結合することができる。

核酸は、モルホリノ骨格構造を含むことができる。例えば、核酸は、リボース環の代わりに6員モルホリノ環を含むことができる。これらの実施形態の一部では、ホスホロジアミデート又は他の非ホスホジエステルヌクレオシド間連結は、ホスホジエステル連結を置換することができる。

【0029】

核酸は、モルホリノ環に付着した複素環塩基を有する連結モルホリノユニット (例えば、モルホリノ核酸) を含むことができる。連結基は、モルホリノ核酸中のモルホリノ単量体ユニットを連結することができる。非イオン性モルホリノベースのオリゴマー化合物は、細胞タンパク質との望ましくない相互作用をより少なくすることができる。モルホリノベースのポリヌクレオチドは、核酸の非イオン性模倣物であり得る。モルホリノクラス内の様々な化合物は、異なる連結基を用いて接続することができる。ポリヌクレオチド模倣物のさらなるクラスは、シクロヘキセニル核酸 (CeNA) と称され得る。核酸分子中に通常存在するフラノース環は、シクロヘキセニル環で置換することができる。CeNA DMT保護されたホスホロアミダイト単量体を調製し、ホスホロアミダイト化学を用いたオリゴマー化合物合成に使用することができる。核酸鎖へのCeNA単量体の取り込みは、DNA/RNAハイブリッドの安定性を増加させることができる。CeNAオリゴアデニル酸は、天然複合体と同様の安定性を有する核酸相補体と複合体を形成することができる。さらなる修飾には、2'-ヒドロキシル基が糖環の4'炭素原子に連結し、それによって2'-C, 4'-C-オキシメチレン連結を形成し、それによって二環糖部分を形成するロックド核酸 (LNA) が含まれ得る。連結は、2'酸素原子及び4'炭素原子を架橋するメチレン (-CH₂) 基であり得、nは1又は2である。LNA及びLNA類似体は、相補的核酸との非常に高い二重鎖熱安定性 (T_m = +3 ~ +10)、3'-エキソヌクレアーゼ分解への安定性、及び良好な溶解性を示すことができる。

【0030】

核酸はまた、核酸塩基 (しばしば単に「塩基」と呼ばれる) 修飾又は置換を含み得る。本明細書で使用される場合、「未修飾」又は「天然」核酸塩基は、プリン塩基 (例えば、アデニン (A) 及びグアニン (G))、並びにピリミジン塩基 (例えば、チミン (T)、シトシン (C) 及びウラシル (U)) を含むことができる。修飾された核酸塩基は、他の合成及び天然の核酸塩基、例えば、5-メチルシトシン (5-me-C)、5-ヒドロキシメチルシトシン、キサントシン、ヒポキサントシン、2-アミノアデニン、アデニン及びグアニンの6-メチル及び他のアルキル誘導体、アデニン及びグアニンの2-プロピル及び他のアルキル誘導体、2-チオウラシル、2-チオチミン及び2-チオシトシン、5-ハロウラシル及びシトシン、5-プロピニル (-C=C-CH₃) ウラシル及びシトシン及びピリミジン塩基の他のアルキニル誘導体、6-アゾウラシル、シトシン及びチミン、5-ウラシル (プソイドウラシル)、4-チオウラシル、8-ハロ、8-アミノ、8-チオール、8-チオアルキル、8-ヒドロキシル及び他の8-置換アデニン及びグアニン、5-ハロ、特に5-プロモ、5-トリフルオロメチル及び他の5-置換ウラシル及びシトシ

ン、7 - メチルグアニン及び7 - メチルアデニン、2 - F - アデニン、2 - アミノアデニン、8 - アザグアニン及び8 - アザアデニン、7 - デアザグアニン及び7 - デアザアデニン及び3 - デアザグアニン及び3 - デアザアデニンを含み得る。修飾された核酸塩基は、三環ピリミジン、例えば、フェノキサジンシチジン(1H - ピリミド(5, 4 - b)(1, 4)ベンゾキサジン - 2(3H) - オン)、フェノチアジンシチジン(1H - ピリミド(5, 4 - b)(1, 4)ベンゾチアジン - 2(3H) - オン)、G - クランプ、例えば、置換フェノキサジンシチジン(例えば、9 - (2 - アミノエトキシ) - H - ピリミド(5, 4 - (b)(1, 4)ベンゾキサジン - 2(3H) - オン)、フェノチアジンシチジン(1H - ピリミド(5, 4 - b)(1, 4)ベンゾチアジン - 2(3H) - オン)、G - クランプ、例えば、置換フェノキサジンシチジン(例えば、9 - (2 - アミノエトキシ) - H - ピリミド(5, 4 - (b)(1, 4)ベンゾキサジン - 2(3H) - オン)、カルバゾールシチジン(2H - ピリミド(4, 5 - b)インドール - 2 - オン)、ピリドインドールシチジン(H - ピリド(3', 2' : 4, 5)ピロロ[2, 3 - d]ピリミジン - 2 - オン)を含み得る。

10

20

30

40

50

【0031】

本明細書で使用される場合、用語「標的」とは、目的とする核酸(例えば、標的 ds DNA)を指すことができる。一部の実施形態では、標的は、アダプター及び/又はバーコードと結び付けられ得る。開示された方法、デバイス、及びシステムによる分析に適した例示的な標的としては、オリゴヌクレオチド、DNA、RNA、mRNA、マイクロRNA、tRNAなどが挙げられる。標的は一本鎖又は二本鎖であり得る。一部の実施形態では、標的は、タンパク質、ペプチド、又はポリペプチドであり得る。一部の実施形態では、標的は脂質である。本明細書で使用される場合、「標的」は、「種」と互換的に使用することができる。

【0032】

本明細書で使用される場合、用語「逆転写酵素」とは、逆転写酵素活性(すなわち、RNA鋳型からのDNAの合成を触媒する)を有する一群の酵素を指すことができる。一般的に、このような酵素は、限定されないが、レトロウイルス逆転写酵素、レトロトランスポゾン逆転写酵素、レトロプラスミド逆転写酵素、レトロ逆転写酵素、細菌逆転写酵素、グループIIイントロン由来逆転写酵素、及びそれらの突然変異体、バリエーション又は誘導体を含む。非レトロウイルス逆転写酵素には、非LTRレトロトランスポゾン逆転写酵素、レトロプラスミド逆転写酵素、レトロ逆転写酵素、及びグループIIイントロン逆転写酵素が含まれる。グループIIイントロン逆転写酵素の例としては、ラクトコッカス・ラクチス(Lactococcus lactis)LI.LtrBイントロン逆転写酵素、サーモシネコッカス・エロンガ투스(Thermosynechococcus elongatus)TeI4cイントロン逆転写酵素、又はゲオバチルス・ステアロサーモフィルス(Geobacillus stearothermophilus)GsI-IIcイントロン逆転写酵素が挙げられる。他のクラスの逆転写酵素は、多くのクラスの非レトロウイルス逆転写酵素(すなわち、とりわけレトロ、グループIIイントロン、及び多様性生成レトロエレメント)を含むことができる。

【0033】

本明細書で使用される場合、用語「核酸を単離する」とは、1種以上の細胞構成成分からの核酸の精製を指すことができる。当業者は、そこから「核酸を単離する」ために処理された試料が、核酸以外の構成成分及び不純物を含み得ることを理解する。単離された核酸を含む試料は、当該技術分野において公知である任意の許容される方法を用いて検体から調製することができる。例えば、細胞は、公知の溶解剤を用いて溶解することができ、核酸は、他の細胞構成成分から精製することができ、又は部分的に精製することができる。DNA及びRNA抽出のための適切な試薬及びプロトコールは、例えば、それぞれ米国特許出願公開第2010/0009351号、及び米国特許出願公開第2009/0131650号に見出すことができる(各々は、その全体が参照により本明細書に組み込まれる)。

本明細書で使用される場合、「鋳型」とは、少なくとも1つの標的ヌクレオチド配列を含むポリヌクレオチドの全部又は一部を指すことができる。

【0034】

本明細書で使用される場合、「プライマー」とは、核酸鎖伸長反応を開始するのに役立つことができるポリヌクレオチドを指すことができる。プライマーの長さは、例えば、約5～約100ヌクレオチド、約10～約50ヌクレオチド、約15～約40ヌクレオチド、又は約20～約30ヌクレオチドの範囲で変動し得る。プライマーの長さは、約10ヌクレオチド、約20ヌクレオチド、約25ヌクレオチド、約30ヌクレオチド、約35ヌクレオチド、約40ヌクレオチド、約50ヌクレオチド、約75ヌクレオチド、約100ヌクレオチド、又はこれらの値の任意の2つの間の範囲であり得る。一部の実施形態では、プライマーは、10～約50ヌクレオチド、すなわち、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50又はそれ以上のヌクレオチドの長さを有する。一部の実施形態では、プライマーは18～32ヌクレオチドの長さを有する。

10

【0035】

本明細書で使用される場合、「プローブ」とは、ハイブリダイゼーションを可能にする条件下で、核酸中の標的配列に（例えば、特異的に）ハイブリダイズすることができ、それによって標的配列又は増幅された核酸の検出を可能にするポリヌクレオチドを指すことができる。プローブの「標的」とは、一般的に、標準的な水素結合（すなわち、塩基対形成）によってプローブオリゴマーの少なくとも一部に特異的にハイブリダイズする、増幅された核酸配列内の配列又はそのサブセットを指す。プローブは、標的特異的配列、及びプローブの三次元立体構造に寄与する他の配列を含み得る。配列は、プローブオリゴマーの適切なハイブリダイゼーション条件において、プローブの標的特異的配列と完全には相補的でない標的配列への安定なハイブリダイゼーションを可能にする場合、「十分に相補的」である。プローブの長さは、例えば、約5～約100ヌクレオチド、約10～約50ヌクレオチド、約15～約40ヌクレオチド、又は約20～約30ヌクレオチドの範囲で変動し得る。プローブの長さは、約10ヌクレオチド、約20ヌクレオチド、約25ヌクレオチド、約30ヌクレオチド、約35ヌクレオチド、約40ヌクレオチド、約50ヌクレオチド、約100ヌクレオチド、又はこれらの値の任意の2つの間の範囲であり得る。一部の実施形態では、プローブは、10～約50ヌクレオチドの長さを有する。例えば、プライマー及びプローブは、少なくとも10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50又はそれ以上のヌクレオチドであり得る。一部の実施形態では、プローブは、非配列特異的であり得る。

20

30

【0036】

好ましくは、プライマー及び/又はプローブは、8～45ヌクレオチドの長さであり得る。例えば、プライマー及びプローブは、長さが少なくとも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、又はそれ以上のヌクレオチドであり得る。プライマー及びプローブは、5'末端又は3'末端、又はその両方にさらなるヌクレオチドを含有するように修飾することができる。当業者は、増幅プライマーの3'末端に対するさらなる塩基（必ずしもプローブではない）が、一般的に、鋳型配列に対して相補的であることを理解する。プライマー及びプローブ配列はまた、5'末端又は3'末端のヌクレオチドを除去するために修飾することができる。当業者は、増幅用に機能するために、プライマー又はプローブは、本明細書に開示されるように、最小の長さ及びアニーリング温度であることを理解する。

40

50

【0037】

プライマー及びプローブは、融解温度（ T_m ）未満の温度であるアニーリング温度でそれらの標的に結合することができる。本明細書で使用される場合、「 T_m 」及び「融解温度」は、二本鎖ポリヌクレオチド分子の集団の50%が一本鎖に解離するようになる温度を指す交換可能な用語である。ポリヌクレオチドの T_m を計算するための式は、当該技術分野において周知である。例えば、 T_m は、以下の式： $T_m = 69.3 + 0.41 \times (G + C) \% - 6.50 / L$ （式中、 L はヌクレオチド中のプローブの長さである）によって計算することができる。ハイブリッドポリヌクレオチドの T_m はまた、1M塩中のハイブリダイゼーションアッセイから採用された式を用いて推定することができ、PCRプライマーについての T_m の計算に一般的に使用される： $[(A + T \text{の数}) \times 2 + (G + C \text{の数}) \times 4]$ 。例えば、C. R. Newton et al. PCR, 2nd ed., Springer-Verlag (New York: 1997), p.24（その全体が参照により本明細書に組み込まれる）を参照されたい。他のより洗練されたコンピュータ計算は、 T_m の計算のために構造的及び配列特性を考慮に入れた当該技術分野に存在する。オリゴヌクレオチドの融解温度は、オリゴヌクレオチドプライマー又はプローブと結合配列との間の相補性、及び塩の条件に依存し得る。一部の実施形態では、本明細書に提供されるオリゴヌクレオチドプライマー又はプローブは、50mMのKCl、10mMのTris-HCl緩衝液中では約90未満の T_m を有し、例えば、約89、88、87、86、85、84、83、82、81、80、79、78、77、76、75、74、73、72、71、70、69、68、67、66、65、64、63、62、61、60、59、58、57、56、55、54、53、52、50、49、48、47、46、45、44、43、42、41、40、39、又はそれ未満であり、列挙された値の任意の2つの間の範囲が含まれる。

10

20

【0038】

一部の実施形態では、本明細書に開示されるプライマー、例えば増幅プライマーは、例えば、フォワードプライマー及びリバースプライマー（第1の増幅プライマー及び第2の増幅プライマー）を含む増幅プライマー対として提供することができる。好ましくは、フォワードプライマー及びリバースプライマーは、10を超えて異ならず、例えば、10未満、9未満、8未満、7未満、6未満、5未満、4未満、3未満、2未満、又は1未満で異なる T_m を有する。

【0039】

プライマー配列及びプローブ配列は、オリゴヌクレオチドが標的核酸配列に特異的にハイブリダイズするのに十分な相補性を含有することを条件として、オリゴヌクレオチド配列内に（標的配列に対して）ヌクレオチド置換を有することによって修飾することができる。このようにして、少なくとも1、2、3、4、又は最大約5ヌクレオチドを置換することができる。本明細書で使用される場合、用語「相補的」とは、2つのポリヌクレオチド鎖の領域間、又は同じポリヌクレオチド鎖の2つの領域間の配列相補性を指すことができる。ポリヌクレオチドの第1の領域は、2つの領域が反平行に配置される場合、第1の領域の少なくとも1つのヌクレオチドが第2の領域の塩基と塩基対形成することができる場合に、同一又は異なるポリヌクレオチドの第2の領域に相補的である。したがって、2つの相補的なポリヌクレオチドがすべてのヌクレオチド位置で塩基対形成する必要はない。「十分に相補的」とは、第2のポリヌクレオチドと100%又は「十分に」相補的であり、したがって、すべてのヌクレオチド位置で塩基対形成する第1のポリヌクレオチドを指すことができる。また、「部分的に相補的」とは、100%相補的ではなく（例えば、90%、又は80%、又は70%相補的である）、1つ以上のヌクレオチド位置にミスマッチしたヌクレオチドを含有する第1のポリヌクレオチドを指すことができる。一部の実施形態では、オリゴヌクレオチドは、ユニバーサル塩基を含む。

30

40

【0040】

本明細書で使用される場合、用語「十分に相補的」とは、一連の相補的塩基間の水素結合によって別の塩基配列にハイブリダイズすることができる連続する核酸塩基配列を指すことができる。相補的塩基配列は、標準的な塩基対形成（例えば、G:C、A:T又はA

50

：U)を用いることによってオリゴマー配列の各位置で相補的であり得るか、又は相補的ではない(非塩基性の位置を含む)が、全体の相補的塩基配列が適切なハイブリダイゼーション条件下で別の塩基配列と特異的にハイブリダイズすることができる1つ以上の残基を含有し得る。隣接塩基は、オリゴマーがハイブリダイズすることを意図する配列と少なくとも約80%、少なくとも約85%、少なくとも約90%、少なくとも約95%、少なくとも約99%、又は100%相補的であり得る。実質的に相補的な配列は、参照配列と比較して、100、99、98、97、96、95、94、93、92、91、90、89、88、87、86、85、84、83、82、81、80、75、70若しくはそれ以下、又はその間の任意の数の同一性のパーセンテージの範囲の配列を指すことができる。当業者は、塩基配列組成に基づいて予測することができる適切なハイブリダイゼーション条件を容易に選択することができるか、又は習慣的な試験(例えば、Green and Sambrook, Molecular Cloning, A Laboratory Manual, 4th ed. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2012)を参照されたい)を用いて決定することができる。

本明細書で使用される場合、用語「多重PCR」とは、1種の単一の標的、又は2種以上の異なる標的を単一の反応容器(例えば、チューブ)内で増幅することを可能にする反応に1を超えるプライマーセットが含まれるPCRのタイプを指す。多重PCRは、例えば、リアルタイムPCRであり得る。

【0041】

本明細書に開示されるものには、迅速な標的化されたシーケンシング(したがって、迅速なシーケンシングに基づく診断、例えば、2時間未満)を可能にする方法、組成物、キット、及びシステム、並びに適切な治療アプローチの同時診断及び決定を必要とするセラノスティックスが含まれる。一部の実施形態では、迅速な標的化されたシーケンシングアプローチの適用は、迅速な病原体診断、迅速ながん診断、稀な疾患診断(例えば、嚢胞性線維症)を含むことができる。

【0042】

本明細書に開示されるものには、ゲノム編集ツール(例えば、Casタンパク質、亜鉛フィンガーヌクレアーゼ(ZFN)、転写アクチベーター様エフェクターヌクレアーゼ(TALEN)、及びアルゴノートタンパク質)を用いてDNA及びRNAシーケンシングのためのカスタマイズされた遺伝子座特異的ライブラリーを作製し、ユーザー定義された遺伝子座で核酸を切断するための酵素(例えば、トランスポゼース)を誘導する方法が含まれる。この酵素(例えば、トランスポゼース)は、シーケンシング、例えば、次世代又は第3世代のシーケンシング技術(限定されないが、Illumina、PacBio、Roche、ThermoFisher、及びOxford Nanoporeによるシーケンシング技術を含む)によるシーケンシングのためのアダプターをこれらの部位に付加することができる。

核酸をシーケンシングするための従来のライブラリー調製方法は、作製するのに数時間を要し得、このプロセスは、ランダムに作製されるライブラリーを生成する(図1)。これらのライブラリーがランダムである理由は、核酸(物理的、酵素的、及び化学的断片化法を含む)を断片化するために使用される方法がランダムに核酸を断片化するからであり、その結果、DNAシーケンシングのアウトプットを制御することができない。

【0043】

ゲノム中の特定の遺伝子座を研究することに関心がある場合、これらの遺伝子座について十分な配列情報が得られることを期待して、何百万もの塩基のシーケンシングが必要となる。これらのデータがすべて獲得されたら、目的とする遺伝子座に関する情報を抽出するために、バイオインフォマティクス手法を用いなければならない。このプロセスは、調製され、シーケンシングされたDNAの大部分が、目的とするこれらの領域に関連性がないため、バイオインフォマティクスの及び計算学的に集約的なものになり得る。さらに、ライブラリー調製プロセスがランダムであるため、これらの領域に十分な情報(カバレッジ)が存在しないリスクがある。この場合、これらの領域が十分にカバーされることを期

待して、時間と資源の両方を浪費して、別のライブラリーを作製し、再度シーケンシングしなければならない。

【0044】

カスタマイズされた遺伝子座特異的ライブラリー調製 (CLLP) 方法を用いてシーケンシングするための本明細書に開示される迅速な標的化されたライブラリー調製方法は、従来のライブラリー調製方法を用いてライブラリーを作製するのに必要な数時間の代わりに、作製するのにわずか数分かかる迅速なプロセスである。さらに、本明細書に開示される CLLP 法によって作製されたライブラリーはランダムではない。一部の実施形態では、選択された遺伝子座のみがシーケンシングされ、他のすべては無視することができ、費用対効果、時間及び資源の節約、及び精度を提供する。さらに、目的とする領域をシーケンシングすることによってのみ、標準的な方法と比較して、必要なバイオインフォマティクス資源及び分析は最小限に留まる。本明細書に開示されるカスタマイズされた遺伝子座特異的ライブラリー調製 (CLLP) 方法は、診断及び/又はセラノスティックスのための迅速で安価な方法としての DNA シーケンシングの使用を可能にする。

10

【0045】

CLLP の一部の実施形態では、ゲノム編集ツール及びトランスポゼース (例えば、高活性トランスポゼース) は、標的化された断片化を達成するために使用される。ユーザーが定義した DNA の二本鎖切断又は RNA の一本鎖切断を可能にする任意のゲノム編集ツールを用いることができる。これらのツールには、限定されないが、CAS タンパク質、ZFN、TALEN、アルゴノートタンパク質、又はそれらの任意の組み合わせが含まれる。一部の実施形態では、ゲノム編集ツールは、正確に選択することができるゲノムの特定の領域に核酸の断片化を制御及び指示するために使用される。ゲノム編集ツールによってなされる切断は、シーケンシングアダプターのためのプライミング部位として使用することができる。これは、順に、シーケンシングされるゲノムの領域に大きな偏りをもたらす。本明細書に開示されるプログラム可能な断片化プロセスは、標的化されたシーケンシングに導くことができる。さらに、この方法は、任意のシーケンシング技術 (限定されないが、Illumina、PacBio、Oxford Nanopore、Roche 及び Thermo Fisher のシーケンシング技術を含む) で使用することができる。図 5A ~ 図 5F は、本明細書に開示されるカスタマイズされた遺伝子座特異的ライブラリー調製 (CLLP) の非限定的であり例示的な実施形態を示す。

20

30

【0046】

一部の実施形態では、タグメンテーションは、高活性トランスポゼースを利用する DNA シーケンシングのためのライブラリーを作製することを含む。タグメンテーションは、二本鎖 DNA を切断し、切断部位に DNA アダプターを貼り付けるトランスポゾンを使用する (図 2)。タグメンテーションは、ライブラリーを標準的なライブラリー調製方法よりもかなり少ない時間で作製する非常に速いプロセスである。しかしながら、トランスポゾンはゲノムをランダムに偏りなく切断する。

【0047】

一部の実施形態では、ライブラリーの調製速度を向上させるために、本明細書に開示される方法は、ゲノム編集ツールに連結されたトランスポゾンを使用する。例えば、dCAS9 タンパク質は、ゲノム編集ツールとして使用することができる。dCAS9 タンパク質は、ゲノム中の特定の領域にプログラム可能なガイド RNA に結合することができる。dCAS9 は CAS9 タンパク質の一種であり、CAS9 タンパク質のヌクレアーゼ活性は失われるが、標的特異性は保持されるように突然変異されている。dCAS9 がその標的に結合すると、CAS9 タンパク質に付着しているトランスポゼースが DNA を切断し、その切断部位アダプターに付着して配列を決定する。最終結果は、シーケンシングの準備が整った標的化 DNA 断片であり、標的化ライブラリー調製プロセスを数時間ではなく数分に短縮する (図 3)。

40

本開示の方法、組成物、キット及びシステムが可能にする非限定的な利点には、結果に至るまでの迅速な時間；先行技術よりも少ない実験室、バイオインフォマティクス及び計

50

算資源の使用；希少及び低頻度のバリエーションの迅速な検出及び定量化を可能にすること；全ゲノムシーケンシングよりも多くの試料を分析し得ることができること；複数のカスタマイズ可能な標的数を同時に検出することができる迅速な診断ツールとして使用できること；より簡単であり容易なデータ分析；並びにそれらの任意の組み合わせが含まれる。

【0048】

現在、標的化されたシーケンシングに使用されている2つの方法がある。第1のものは、アンプリコンシーケンシングである。この方法は、DNA増幅を介して目的とする領域を増幅するためにプライマーを使用することに依存する。この追加の増幅ステップは、標準的なライブラリー調製方法にさらなるコスト、時間、及び資源を追加する。第2の標的化されたシーケンシング方法は、標的捕捉である。この方法は、特定の核酸標的にハイブリダイズすることができるようにプローブ、又はプローブのプールを使用することに依存する。プローブのその標的へのハイブリダイゼーション及びこれらの標的の単離は、数日を要し得る時間のかかるプロセスである。さらに、この方法において使用されるプローブは、合成するのに高価である。

10

【0049】

一部の実施形態では、dCAS9タンパク質に連結された高活性トランスポゼースTn5を使用することができる。dCAS9は、CAS9タンパク質の触媒死滅型であり、CAS9タンパク質のヌクレアーゼ活性は失われるが、プログラム可能なDNA結合活性は保持するように突然変異される。dCAS9タンパク質のN末端は、リンカー（例えば、X-TEN）、SNAPタグ又はCLIPタグを介してTn5トランスポゼースのC末端に付着される。両方のタンパク質を付着させるには、多くの異なる方法を用いることができるが、Tn5トランスポゼースは、シーケンシング技術プラットフォームに特異的であるシーケンシングアダプターがロードされる。dCas9タンパク質には、ユーザーが定義した遺伝子座に特異的であるガイドRNA（sgRNA）が付着する。各々dCAS9タンパク質に結合した複数のsgRNAと、各々異なる遺伝子座を標的とする複数のsgRNAを用いて、複数の遺伝子座を選択する。

20

sgRNAに付着するdCAS9がsgRNA配列に相補的である分子を見つけると、付着したTn5トランスポゼースは特定の部位でDNAを切断し、その切断部位アダプターに付着して配列を決定することができる。最終結果は、シーケンシングの準備が整った標的化DNA断片であり、標的化されたライブラリー調製プロセスを数時間ではなく数分に短縮する（図3）。

30

【0050】

一部の実施形態では、Tn5トランスポゼースでない高活性トランスポゼースは、マリナーTc1様トランスポゾン、Himar1C9トランスポゼース、スリーピングビューティトランスポゼース、Tn7トランスポゾン、又はそれらの組み合わせであり得る。一部の実施形態では、dCas9タンパク質の代替物を、プログラム可能なDNA結合活性のために使用することができる。例えば、FOK1ヌクレアーゼに結合していない亜鉛フィンガーの使用を用いることができる。同様に、FOK1ヌクレアーゼをもたないTALEN分子を用いることもできる。一部の実施形態では、配列特異的プライマーと組み合わせたリコンビナーゼの使用は、プログラム可能なDNA結合分子として使用することができる。一部の実施形態では、遺伝子座特異的ライブラリーを作製するための代替の方法は、トランスポゼースの支援なしに、ゲノム編集ツール（例えば、Casタンパク質、亜鉛フィンガーヌクレアーゼ（ZFN）、転写アクチベーター様エフェクターヌクレアーゼ（TALEN）、アルゴノートタンパク質）のみを使用することによって行うことができる。これは、遺伝子座特異的なシーケンシングライブラリーを作製するためにさらに使用することができる核酸（図4）のプログラム可能な断片化法をもたらす。

40

ゲノムの特定の領域を標的とするプログラム可能なツールとしてのゲノム編集ツールの使用、及びシーケンシングライブラリーを作製するのに必要なアダプターを切断及び貼り付けるためのトランスポゼースの使用が、本明細書に開示される。

【0051】

50

一部の実施形態は、病原体/疾患の原因(遺伝子突然変異)を同定し、同時に抗生物質に対する感受性を同定するように構成された疾患パネル(例えば、敗血症パネル)を提供する。一部の実施形態では、がんパネルは、がん細胞における複数の突然変異の同定を含むことができる。一部の実施形態では、稀な疾患パネルは、遺伝子疾患(例えば、嚢胞性線維症)に導くことができる突然変異と関連している特定の遺伝子座のシーケンシングを含むことができる。

以下の特許出願公開の各々は、その全体が参照により本明細書に組み込まれる: 国際公開第2016028843A2号及び国際公開第2018175872A1号、米国特許出願公開第20190144920号、及びカナダ特許出願公開第3026206号。

【0052】

本明細書に開示されるものには、組成物が含まれる。一部の実施形態では、組成物は、複数のタンパク質複合体を含む。一部の実施形態では、複数のタンパク質複合体の各々は、トランスポソームと、標的 dsDNA 上の結合部位に特異的に結合することができるプログラム可能な DNA 結合ユニットとを含む。一部の実施形態では、トランスポソームは、トランスポゼース、第1のアダプター、及び第2のアダプターを含む。一部の実施形態では、複数のタンパク質複合体の各々に対する結合部位は、互いに異なる。

【0053】

本明細書に開示されるものには、反応混合物が含まれる。一部の実施形態では、反応混合物は、本明細書に開示される組成物及び1種以上の標的 dsDNA を含むことが疑われる試料核酸を含む。反応混合物は、DNA ポリメラーゼ、dNTP、又はそれらの組み合わせを含むことができる。アダプターは、標的 dsDNA 又はその断片に共有結合的に付着され得る。反応混合物は、各々が、各末端にそれぞれ複数のタンパク質複合体のうちの1つの第1のアダプター及び第2のアダプターを含む複数の dsDNA 断片を含むことができる。

本明細書に開示されるものには、核酸をタグ付けする方法が含まれる。一部の実施形態では、本方法は、本明細書に開示された組成物を、複数の標的 dsDNA を含むことが疑われる試料と接触させて、反応混合物を形成することと、反応混合物をインキュベートして、各々が、各末端にそれぞれ複数のタンパク質複合体のうちの1つの第1のアダプター及び第2のアダプターを含む複数の dsDNA 断片を生成することとを含む。

【0054】

本明細書に開示されるものには、シーケンシングライブラリーを生成する方法が含まれる。一部の実施形態では、本方法は、本明細書に開示される組成物を、複数の標的 dsDNA を含むことが疑われる試料と接触させて、反応混合物を形成することを含む。本方法は、反応混合物をインキュベートして、各々が、各末端にそれぞれ複数のタンパク質複合体のうちの1つの第1のアダプター及び第2のアダプターを含む複数の dsDNA 断片を生成することを含むことができる。複数の標的 dsDNA を複数のタンパク質複合体対と接触させることは、約25~約85 (例えば、約25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、45、50、55、60、65、70、75、80、85、又はこれらの値の任意の2つの間の数値若しくは範囲)で行うことができる。反応混合物をインキュベートすることは、約37~約55 (例えば、約37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、又はこれらの値の任意の2つの間の数値若しくは範囲)で反応混合物をインキュベートすることができる。

【0055】

複数のタンパク質複合体対及び複数の標的 dsDNA は、約2:1~約2000:1 (例えば、2:1、2.5:1、3:1、4:1、5:1、6:1、7:1、8:1、9:1、10:1、11:1、12:1、13:1、14:1、15:1、16:1、17:1、18:1、19:1、20:1、21:1、22:1、23:1、24:1、25:

10

20

30

40

50

1、26：1、27：1、28：1、29：1、30：1、31：1、32：1、33：1、34：1、35：1、36：1、37：1、38：1、39：1、40：1、41：1、42：1、43：1、44：1、45：1、46：1、47：1、48：1、49：1、50：1、51：1、52：1、53：1、54：1、55：1、56：1、57：1、58：1、59：1、60：1、61：1、62：1、63：1、64：1、65：1、66：1、67：1、68：1、69：1、70：1、71：1、72：1、73：1、74：1、75：1、76：1、77：1、78：1、79：1、80：1、81：1、82：1、83：1、84：1、85：1、86：1、87：1、88：1、89：1、90：1、91：1、92：1、93：1、94：1、95：1、96：1、97：1、98：1、99：1、100：1、200：1、300：1、400：1、500：1、600：1、700：1、800：1、900：1、1000：1、2000：1、又はこれらの値の任意の2つの間の数値若しくは範囲)の分子比で反応混合物中に存在することができる。一部の実施形態では、複数のタンパク質複合体対及び複数の標的 dsDNA は、約 2：1～約 200：1 (例えば、2：1、2.5：1、3：1、4：1、5：1、6：1、7：1、8：1、9：1、10：1、11：1、12：1、13：1、14：1、15：1、16：1、17：1、18：1、19：1、20：1、21：1、22：1、23：1、24：1、25：1、26：1、27：1、28：1、29：1、30：1、31：1、32：1、33：1、34：1、35：1、36：1、37：1、38：1、39：1、40：1、41：1、42：1、43：1、44：1、45：1、46：1、47：1、48：1、49：1、50：1、51：1、52：1、53：1、54：1、55：1、56：1、57：1、58：1、59：1、60：1、61：1、62：1、63：1、64：1、65：1、66：1、67：1、68：1、69：1、70：1、71：1、72：1、73：1、74：1、75：1、76：1、77：1、78：1、79：1、80：1、81：1、82：1、83：1、84：1、85：1、86：1、87：1、88：1、89：1、90：1、91：1、92：1、93：1、94：1、95：1、96：1、97：1、98：1、99：1、100：1、200：1、又はこれらの値の任意の2つの間の数値若しくは範囲)の分子比で反応混合物中に存在する。

【0056】

複数のタンパク質複合体のうちの少なくとも2つの結合部位は、同じ標的 dsDNA 上にあり得る。複数のタンパク質複合体のうちの少なくとも2つの結合部位は、同じ標的 dsDNA 上で約 1～50000ヌクレオチド離れていることができる。一部の実施形態では、複数のタンパク質複合体のうちの少なくとも2つの結合部位は、同じ標的 dsDNA 上に1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、20000、30000、40000、50000、60000、70000、80000、90000、100000ヌクレオチド若しくはほぼこれらの数値、又はこれらの値の任意の2つの間の数値若しくは範囲で離れていることができる。一部の実施形態では、複数のタンパク質複合体のうちの少なくとも2つの結合部位は、同じ標的 dsDNA 上に少なくとも若しくは多くとも1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、4

7、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、20000、30000、40000、50000、60000、70000、80000、90000、又は100000ヌクレオチド離れていることができる。複数のタンパク質複合体の一对の結合部位間の距離は、複数のタンパク質複合体の別の一对の結合部位間の距離と実質的に同じであり得る。複数のタンパク質複合体の一对の結合部位間の距離は、複数のタンパク質複合体の別の一对の結合部位間の距離と異なることができる。複数のタンパク質複合体のうち少なくとも2つの結合部位は、標的 dsDNA の異なる鎖上にあり得る。複数のタンパク質複合体のうち少なくとも2つは、異なる標的 dsDNA に特異的に結合することができる。複数のタンパク質複合体は、約2~5000の標的 dsDNA に特異的に結合することができる。一部の実施形態では、複数のタンパク質複合体は、約1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、60、70、80、90、100、110、120、128、130、140、150、160、170、180、190、200、210、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400、410、420、430、440、450、460、470、480、490、500、510、520、530、540、550、560、570、580、590、600、610、620、630、640、650、660、670、680、690、700、710、720、730、740、750、760、770、780、790、800、810、820、830、840、850、860、870、880、890、900、910、920、930、940、950、960、970、980、990、1000、1100、1200、1300、1400、1500、1600、1700、1800、1900、2000、2100、2200、2300、2400、2500、2600、2700、2800、2900、3000、3250、3500、3750、4000、4250、4500、4750、5000、5250、5500、6000、6500、7000、7500、8000、8500、9000、9500、10000、又はこれらの値の任意の2つの間の数値若しくは範囲の標的 dsDNA に特異的に結合することができる。

10

20

30

【0057】

トランスポソーム

一部の実施形態では、トランスポソームは、トランスポゼース、第1のアダプター、及び第2のアダプターを含む。複数のタンパク質複合体のうち少なくとも2つは、同じトランスポソームを含むことができる。複数のタンパク質複合体はすべて、同じトランスポソームを含むことができる。複数のタンパク質複合体はすべて、同じトランスポゼースを含むことができる。トランスポゼースは、Tn5トランスポゼース、Tn7トランスポゼース、マリナーTc1様トランスポゼース、Himar1C9トランスポゼース、又はスリーピングビューティトランスポゼースであり得る。トランスポゼースは、高活性トランスポゼースであり得る。

40

【0058】

トランスポゼースは、Tn5、Tn7、MuA、又はビブリオ・ハーベリ(Vibrio harveyi)トランスポゼース、又はそれらの活性突然変異体であり得る。一部の実施形態では、トランスポゼースは、Tn5トランスポゼース又はその突然変異体である。一部の実施形態では、Tn5トランスポゼースは、高活性Tn5トランスポゼース、又はその活性突然変異体である。一部の実施形態では、Tn5トランスポゼースは、参照により本明細書に組み込まれる国際公開第2015/160895号に記載されるような

50

Tn5トランスポゼースである。一部の実施形態では、Tn5トランスポゼースは、野生型Tn5トランスポゼースに対して54、56、372、212、214、251、及び338位に突然変異を有する高活性Tn5である。一部の実施形態では、Tn5トランスポゼースは、野生型Tn5トランスポゼースに対して以下の突然変異：E54K、M56A、L372P、K212R、P214R、G251R、及びA338Vを有する高活性Tn5である。一部の実施形態では、Tn5トランスポゼースは融合タンパク質である。一部の実施形態では、Tn5トランスポゼース融合タンパク質は、融合伸長因子Ts (Tsf) タグを含む。一部の実施形態では、Tn5トランスポゼースは、野生型配列に対してアミノ酸54、56、及び372に突然変異を含む高活性Tn5トランスポゼースである。一部の実施形態では、高活性Tn5トランスポゼースは融合タンパク質である。一部の実施形態では、認識部位は、Tn5型トランスポゼース認識部位である。

10

【0059】

トランスポゼースは、単一のタンパク質を含み得るか、又は複数のタンパク質サブユニットを含み得る。トランスポゼースは、トランスポゾン末端又はトランスポゾン末端配列と機能的複合体を形成することができる酵素であり得る。一部の実施形態では、トランスポゼース複合体は、第1の単量体及び第2の単量体を含むトランスポゼース (例えば、Tn5トランスポゼース) 二量体を含む。一部の実施形態では、トランスポソーム複合体は、2分子のトランスポゼースの二量体を含む。

【0060】

トランスポゼース及び/又はトランスポソームは、実施形態に依存して変動し得る。トランスポゼースは、Tn5トランスポゼースを含むことができる。トランスポゼースは、Tnトランスポゼース (例えば、Tn3、Tn5、Tn7、Tn10、Tn552、Tn903)、MuAトランスポゼース、Vibhartトランスポゼース (例えば、ビブリオ・ハーベリ由来)、Ac-Ds、Ascot-1、Bs1、Cin4、Copia、En/Spm、Fエレメント、hobo、Hsmar1、Hsmar2、IN(HIV)、IS1、IS2、IS3、IS4、IS5、IS6、IS10、IS21、IS30、IS50、IS51、IS150、IS256、IS407、IS427、IS630、IS903、IS911、IS982、IS1031、ISL2、L1、マリナー、Pエレメント、Tam3、Tc1、Tc3、Tel、THE-1、Tn/O、TnA、Tn3、Tn5、Tn7、Tn10、Tn552、Tn903、Tol1、Tol2、Tn10、Ty1、任意の原核生物トランスポゼース、又は上記に列挙したものに関連する及び/若しくはそれらに由来する任意のトランスポゼースであり得る。一部の実施形態では、親トランスポゼースに関連する及び/又はそれに由来するトランスポゼースは、親トランスポゼースの対応するペプチド断片と少なくとも約50%、約55%、約60%、約65%、約70%、約75%、約80%、約85%、約90%、約91%、約92%、約93%、約94%、約95%、約96%、約97%、約98%、又は約99%のアミノ酸配列相同性を有するペプチド断片を含むことができる。ペプチド断片は、長さが少なくとも約10、約15、約20、約25、約30、約35、約40、約45、約50、約60、約70、約80、約90、約100、約150、約200、約250、約300、約400、又は約500のアミノ酸であり得る。例えば、Tn5由来のトランスポゼースは、長さが50

20

30

40

【0061】

アダプター

同じトランスポソーム内の第1のアダプター及び第2のアダプターは同じであり得る。第1のアダプター、第2のアダプター、又はその両方は、異なるトランスポソーム中で異なることができる。第1のアダプター、第2のアダプター、又はその両方は、dsDNA又はRNA/DNA二重鎖であり得る。アダプターは、長さが約3~200塩基対 (例え

50

ば、長さが約 3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、60、70、80、90、100、150、200、又はこれらの値の任意の2つの間の数値若しくは範囲のヌクレオチド)であり得る。一部の実施形態では、アダプターは、長さが3~500塩基対(例えば、長さが約3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、60、70、80、90、100、150、200、300、500、又はこれらの値の任意の2つの間の数値若しくは範囲のヌクレオチド)であり得る。第1のアダプター、第2のアダプター、又はその両方は、シーケンシングアダプターであり得る。シーケンシングアダプターは、所定のシーケンシングプロトコールにおいて採用される1種以上の構成成分、例えば、シーケンシングプラットフォームアダプター構築物、インデックス化ドメイン、クラスター化ドメインなどを含むことができる。シーケンシングアダプターは、P5又はP7プライマー配列を含むことができる。一部の実施形態では、第1のアダプター及び/又は第2のアダプターは、バーコード(例えば、確率バーコード)を含む。一部の実施形態では、第1のアダプター及び/又は第2のアダプターは、ユニバーサル配列を含む。一部の実施形態では、第1のアダプター及び/又は第2のアダプターは、一本鎖部分及び/又は二本鎖部分を含む。一部の実施形態では、アダプターは、トランスポゼースに結合するトランスポゾン末端配列を含む。トランスポゾン末端配列は二本鎖であり得る。一部の実施形態では、トランスポゾン末端配列はモザイク末端(ME)配列である。特定の実施形態では、トランスポゾン末端はモザイク末端、又はトランスポゾン末端の高活性型である。アダプター配列は、2つのトランスポゾン末端配列のうちの1つに付着することができる。したがって、一部の実施形態では、第1のアダプタートランスポゾン末端配列はME配列であり、第2のアダプター末端配列はME'配列である。

10

20

【0062】

第1のアダプター及び/又は第2のアダプターは、修飾されるか又はそうでなければ天然に存在しない1種以上のヌクレオチド(又はその類似体)を含むことができる。例えば、第1のアダプター及び/又は第2のアダプターは、1種以上のヌクレオチド類似体(例えば、LNA、FANA、2'-O-Me RNA、2'-フルオロRNAなど)、連結修飾(例えば、ホスホロチオエート、3'-3'及び5'-5'逆連結)、5'及び/又は3'末端修飾(例えば、5'及び/又は3'アミノ、ビオチン、DIG、リン酸塩、チオール、色素、消光剤など)、1種以上の蛍光標識されたヌクレオチド、又は所望の機能性を提供する任意の他の特徴を含むことができる。

30

【0063】

第1のアダプター及び/又は第2のアダプターは、シーケンシングプラットフォームアダプター構築物の全部又は構成成分を含むことができる。「シーケンシングプラットフォームアダプター構築物」とは、Illumina(登録商標)(例えば、HiSeq(商標)、MiSeq(商標)及び/又はGenome Analyzer(商標)シーケンシングシステム); Ion Torrent(商標)(例えば、Ion PGM(商標)及び/又はIon Proton(商標)シーケンシングシステム); Pacific Biosciences(例えば、PACBIO RS IIシーケンシングシステム); Life Technologies(商標)(例えば、SOLIDシーケンシングシステム); Roche(例えば、454GS FLX+及び/又はGS Juniorシーケンシングシステム); 又は目的とする任意の他のシーケンシングプラットフォームによって提供されるシーケンシングプラットフォームなどの、目的とするシーケンシングプラットフォームによって利用される核酸ドメイン(例えば、シーケンシングプラットフォームアダプター核酸配列)の少なくとも一部を含む核酸構築物を意味する。第1のアダプター及び/又は第2のアダプターは、表面に付着したシーケンシングプラットフォームオリゴヌクレオチド(例えば、Illumina(登録商標)シーケンシングシステムにおいてフローセルの表面に付着したP5又はP7オリゴヌクレオチド)に特異的に結合するドメイン(例えば、「捕捉部位」又は「捕捉配列」); シーケンシングプライマー結合ド

40

50

メイン（例えば、Illumina（登録商標）プラットフォームのリード1又はリード2プライマーが結合し得るドメイン）；バーコードドメイン（例えば、特定のバーコード又は「タグ」で所与の試料からのすべての分子をマーキングすることによって試料の多重化を可能にするためにシーケンシングされる核酸の試料源を固有に同定するドメイン）；バーコードシーケンシングプライマー結合ドメイン（バーコードをシーケンシングするために使用されるプライマーが結合するドメイン）；分子同定ドメイン（例えば、4、6、又は他の数のヌクレオチドのランダム化されたタグなどの分子インデックスタグ）であって、固有のタグがシーケンシングされる事例の数に基づいて発現レベルを決定するために、目的とする分子を固有にマーキングするためのもの；又はこのようなドメインの任意の組み合わせから選択される1種以上の核酸ドメインを含むことができる。一部の実施形態では、バーコードドメイン（例えば、試料インデックスタグ）及び分子同定ドメイン（例えば、分子インデックスタグ）は、同じ核酸に含まれ得る。

10

【0064】

シーケンシングプラットフォームアダプタードメインは、第1のアダプター及び/又は第2のアダプターに存在する場合、目的とするシーケンシングプラットフォームに適した任意の長さ及び配列の1種以上の核酸ドメインを含むことができる。核酸ドメインは、目的とするシーケンシングプラットフォームによって採用されるポリヌクレオチド（例えば、オリゴヌクレオチド）が、例えば、核酸ドメインに隣接するcDNA挿入物の合成による固相増幅及び/又はシーケンシングのために、核酸ドメインに特異的に結合することを可能にする長さ及び配列を有し得る。例示的な核酸ドメインは、Illumina（登録商標）ベースのシーケンシングプラットフォーム上で採用されるP5、P7、リード1プライマー及びリード2プライマードメインを含む。他の例示的な核酸ドメインは、Ion Torrent（商標）ベースのシーケンシングプラットフォーム上で採用されるAアダプター及びP1アダプタードメインを含む。

20

【0065】

目的とするシーケンシングプラットフォーム上でのシーケンシングに有用な核酸ドメインのヌクレオチド配列は、経時的に変動及び/又は変化し得る。アダプター配列は、典型的には、シーケンシングプラットフォームの製造者によって提供される（例えば、シーケンシングシステムとともに提供される技術文書及び/又は製造者のウェブサイトで利用可能である）。このような情報に基づいて、本明細書に提供されるアダプターの配列は、目的とするプラットフォーム上の標的dsDNAのシーケンシングを可能にする構成において、1種以上の核酸ドメインのすべて又は一部を含むように設計することができる。

30

【0066】

第1のアダプター及び/又は第2のアダプターは、Ion Torrent（商標）シーケンシングプラットフォーム（例えば、Ion PGM（商標）及び/又はIon Proton（商標）シーケンシングシステム）の構成成分を含むことができる。第1のアダプター及び/又は第2のアダプターは、P1アダプター、Aアダプター、Ion Xpress（商標）バーコードアダプター、Ion P1アダプター、及び/又はIon Xpress（商標）バーコードXアダプターを含むことができる。

【0067】

第1のアダプター及び/又は第2のアダプターは、ヘアピンを含むことができる。第1のアダプター及び/又は第2のアダプターは、SMRTbell（商標）技術ライブラリーを生成するように構成することができる。本明細書に提供される方法は、二本鎖断片の末端にライゲートされたヘアピンアダプターをもたらし、末端に中心の二本鎖部分及び一本鎖ヘアピループを有する環状鋳型分子を生成することができる（Pacific Biosciences（登録商標）のSMRTbell（商標）を参照されたい）。SMRTBELL（登録商標）鋳型などの環状鋳型を調製及び使用するための方法は、例えば、「クローナルシーケンシングのためのDNAのエラーフリー増幅」と題する米国特許第8,003,330号、及び「核酸試料調製のための方法及び組成物」と題する米国特許出願公開第2009/0280538号に記載され、全開示は、すべての目的のために参

40

50

照により本明細書に組み込まれる。

【0068】

第1のアダプター及び/又は第2のアダプターは、ONT機器(例えば、SmidgION、MinION、GridION、PromethION)上のタグ付き核酸の下流での使用のために構成することができる。図6は、ONT迅速シーケンシングキット、タグメンテーションベースの迅速シーケンシングキットを示す非限定的であり例示的な実施形態を示す。第1のアダプター及び/又は第2のアダプターは、(i)スパーサー；(ii)モータータンパク質の活性部位がスパーサーによって占められる、スパーサー上でストールしたモータータンパク質；及び/又は(iii)ブロッキング部分が、モータータンパク質がスパーサーから移動するのを妨げる、アダプターに結合したブロッキング部分を含むことができる。第1のアダプター及び/又は第2のアダプターは、ヘアピンループアダプターを含むことができる。ヘアピンループアダプターは、一本のポリヌクレオチド鎖を含むアダプターであり得、ポリヌクレオチド鎖の末端は互いにハイブリダイズすることができるか、又は互いにハイブリダイズされ、ポリヌクレオチドの中央部分はループを形成する。適切なヘアピンループアダプターは、当該技術分野において公知である方法を使用して設計することができる。第1のアダプター及び/又は第2のアダプターは、直鎖状アダプターを含むことができる。第1のアダプター及び/又は第2のアダプターは、Yアダプターであり得る。Yアダプターは、典型的には、ポリヌクレオチドアダプターである。Yアダプターは、典型的には、二本鎖であり、(a)一端において、2本の鎖が互いにハイブリダイズする領域、及び(b)他端において、2本の鎖が相補的でない領域を含む。鎖の非相補的部分は、典型的には突出部を形成する。Yアダプター内に非相補的領域が存在すると、2本の鎖は、典型的には、二本鎖部分とは異なり互いにハイブリダイズしないため、アダプターのY形状が得られる。Yアダプターの2つの一本鎖部分は、同じ長さであり得るか又は異なる長さであり得る。モータータンパク質は、Yアダプターなどのアダプターの突出部に結合することができる。一部の実施形態では、モータータンパク質は、二本鎖領域に結合し得る。一部の実施形態では、モータータンパク質は、アダプターの本鎖領域及び/又は二本鎖領域に結合し得る。一部の実施形態では、第1のモータータンパク質は、このようなアダプターの本鎖領域に結合することができ、第2のモータータンパク質は、アダプターの本鎖領域に結合することができる。第1のアダプター及び/又は第2のアダプターは、ナノ細孔シーケンシング反応を促進するさらなる結合した構成成分、例えば、結合酵素(例えば、ヘリカーゼ、ポリメラーゼ、又は他のモータータンパク質)、膜結合部分(例えば、コレステロール)などを含むことができる。典型的には、モータータンパク質は、ヘリカーゼ、ポリメラーゼ、エキソヌクレアーゼ、トポイソメラーゼ、又はそれらのバリエーションである。一部の実施形態では、ポリヌクレオチドアダプターのスパーサー上のモータータンパク質は、(スパーサーの末端を通過すること以外の)モータータンパク質がスパーサーから離れることを防止するように修飾される。モータータンパク質は、任意の適切な方法で適合させることができる。図7A~図7Hは、Oxford Nanoporeからのものなどの既存のシーケンシングプラットフォームにおいて使用するためのシーケンシングライブラリーを生成するためのゲノム編集タグメンテーション(GET)の非限定的であり例示的な実施形態を示す。

【0069】

本明細書に提供されるアダプター(例えば、第1のアダプター及び/又は第2のアダプター)は、バーコード、例えば、確率バーコードを含むことができ、1種以上の標識を含むことができる。例えば、Fu et al., Proc Natl Acad Sci U.S.A., 2011 May 31, 108(22):9026-31; 米国特許出願公開第2011/0160078号; Fan et al., Science, 2015, 347(6222):1258367; 米国特許出願公開第2015/0299784号; 国際公開第2015/031691号には、確率バーコーディングなどのバーコーディングが記載されており、これらの各々の内容は、任意の支持若しくは補足情報又は材料を含み、その全体が参照により本明細書に組み込まれる。一部の実施形態では、本明細書に開示されるバーコードは、標的を確率的に標識(例えば、バーコード、タグ)する

ために使用され得るポリヌクレオチド配列であり得る確率バーコードであり得る。確率バーコードの異なるバーコード配列の数と、標識されるべき標的のうちのいずれかの出現の数との比が、1 : 1、2 : 1、3 : 1、4 : 1、5 : 1、6 : 1、7 : 1、8 : 1、9 : 1、10 : 1、11 : 1、12 : 1、13 : 1、14 : 1、15 : 1、16 : 1、17 : 1、18 : 1、19 : 1、20 : 1、30 : 1、40 : 1、50 : 1、60 : 1、70 : 1、80 : 1、90 : 1、100 : 1若しくはほぼこれらの数値、又はこれらの値の任意の2つの間の数値若しくは範囲であり得る場合、バーコードは確率バーコードと称することができる。標的は、同一又はほぼ同一の配列を有するmRNA分子を含むmRNA種であり得る。確率バーコードの異なるバーコード配列の数と、標識されるべき標的のうちのいずれかの出現の数との比が少なくとも若しくは多くとも1 : 1、2 : 1、3 : 1、4 : 1、5 : 1、6 : 1、7 : 1、8 : 1、9 : 1、10 : 1、11 : 1、12 : 1、13 : 1、14 : 1、15 : 1、16 : 1、17 : 1、18 : 1、19 : 1、20 : 1、30 : 1、40 : 1、50 : 1、60 : 1、70 : 1、80 : 1、90 : 1、100 : 1である場合、バーコードは確率バーコードと称することができる。確率バーコードのバーコード配列を分子標識と称することができる。

10

【0070】

アダプター及び/又はバーコードは、1種以上のユニバーサル標識を含むことができる。一部の実施形態では、1種以上のユニバーサル標識は、すべてのバーコード及び/又はアダプターに対して同じであり得る。一部の実施形態では、ユニバーサル標識は、シーケンシングプライマーにハイブリダイズすることができる核酸配列を含むことができる。シーケンシングプライマーは、ユニバーサル標識を含むバーコードをシーケンシングするために使用することができる。シーケンシングプライマー（例えば、ユニバーサルシーケンシングプライマー）は、ハイスループットシーケンシングプラットフォームと関連するシーケンシングプライマーを含むことができる。一部の実施形態では、ユニバーサル標識は、PCRプライマーにハイブリダイズすることができる核酸配列を含むことができる。一部の実施形態では、ユニバーサル標識は、シーケンシングプライマー及びPCRプライマーにハイブリダイズすることができる核酸配列を含むことができる。シーケンシング又はPCRプライマーにハイブリダイズすることができるユニバーサル標識の核酸配列は、プライマー結合部位と称することができる。ユニバーサル標識は、バーコードの転写を開始するために使用することができる配列を含むことができる。ユニバーサル標識は、バーコード又はバーコード内の領域の伸長に使用できる配列を含むことができる。ユニバーサル標識は、長さが、1、2、3、4、5、10、15、20、25、30、35、40、45、50若しくはほぼこれらの数値、又はこれらの値の任意の2つの間の数値若しくは範囲のヌクレオチドであり得る。例えば、ユニバーサル標識は、少なくとも約10ヌクレオチドを含むことができる。ユニバーサル標識は、長さが少なくとも若しくは多くとも1、2、3、4、5、10、15、20、25、30、35、40、45、50、100、200、又は300ヌクレオチドであり得る。

20

30

【0071】

バーコード、例えば、確率バーコードは、1種以上の標識を含むことができる。例示的な標識には、ユニバーサル標識、細胞標識、バーコード配列（例えば、分子標識）、試料標識、プレート標識、空間標識、及び/又はプレ空間標識を含むことができる。バーコードは、ユニバーサル標識、寸法標識、空間標識、細胞標識、及び/又は分子標識を含むことができる。バーコード中の異なる標識（限定されないが、ユニバーサル標識、寸法標識、空間標識、細胞標識、及び分子標識を含む）の順序は、変動し得る。例えば、ユニバーサル標識は5'最大標識であり得、分子標識は3'最大標識であり得る。空間標識、寸法標識、及び細胞標識は、任意の順序であり得る。一部の実施形態では、ユニバーサル標識、空間標識、寸法標識、細胞標識、及び分子標識は、任意の順序である。一部の実施形態では、バーコードの標識（例えば、ユニバーサル標識、寸法標識、空間標識、細胞標識、及びバーコード配列）は、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、若しくは20又はそれ以上のヌクレオチドによ

40

50

って分離され得る。

【0072】

標識、例えば、細胞標識は、定義された長さ、例えば、各々7つのヌクレオチド（いくつかのハミングエラー訂正コードにおいて使用されるビットの数に等しい）の固有の核酸サブ配列のセットを含み得、エラー訂正能力を提供するように設計され得る。7つのヌクレオチド配列を含むエラー訂正サブ配列のセットは、セット中の配列の任意の対になった組み合わせが、定義された「遺伝子距離」（又はミスマッチ塩基の数）を示すように設計することができ、例えば、エラー訂正サブ配列のセットは、3つのヌクレオチドの遺伝子距離を示すように設計することができる。この場合、標識された標的核酸分子（以下により十分に記載される）についての配列データのセットにおけるエラー訂正配列の総説は、増幅又はシーケンシングエラーを検出又は訂正することを可能にすることができる。一部の実施形態では、エラー訂正コードを作製するために使用される核酸サブ配列の長さは変動することができ、例えば、1、2、3、4、5、6、7、8、9、10、15、20、30、31、40、50若しくはほぼこれらの数値、又はこれらの値の任意の2つの間の数値若しくは範囲のヌクレオチドの長さであり得る。一部の実施形態では、他の長さの核酸サブ配列は、エラー訂正コードを作製するために使用することができる。

10

【0073】

CRISPR関連タンパク質

プログラム可能なDNA結合ユニットは、ヌクレアーゼ欠損CRISPR関連タンパク質（dCASタンパク質）と、標的dsDNAの結合部位に特異的に結合することができるガイドRNA（gRNA）とを含むことができる。dCASタンパク質は、dCAS9、dCAS12、dCAS13、dCAS14、又はSpRY dCASであり得る。dCAS13タンパク質は、dCAS13a、dCAS13b、dCAS13c、又はdCAS13dであり得る。

20

【0074】

一部の実施形態では、Cas9タンパク質は、不活性な（例えば、不活性化された）DNA切断ドメインを有する。ヌクレアーゼで不活性化されたCas9タンパク質は、互換的に「dCas9」タンパク質（ヌクレアーゼ死滅型Cas9の場合）と称することができる。不活性なDNA切断ドメインを有するCas9タンパク質（又はその断片）を生成する方法は公知である（例えば、各々の全内容が参照により本明細書に組み込まれるJinek et al., Science.337:816-821(2012); Qi et al., (2013) Cell.28; 152(5): 1173-83を参照されたい）。例えば、Cas9のDNA切断ドメインは、2つのサブドメインであるHNHヌクレアーゼサブドメイン及びRuvC1サブドメインを含むことが公知である。HNHサブドメインは、gRNAに相補的である鎖を切断するが、RuvC1サブドメインは非相補的な鎖を切断する。これらのサブドメイン内の突然変異は、Cas9のヌクレアーゼ活性を沈黙させることができる。例えば、突然変異D10A及びH840Aは、S.ピオゲネス（S. pyogenes）Cas9のヌクレアーゼ活性を完全に不活性化する（Jinek et al., and Qi et al.）。

30

【0075】

プログラム可能なDNA結合ユニットは、ガイドRNAになおも結合することができる適切なヌクレアーゼ欠損Casタンパク質を含むことができる。プログラム可能なDNA結合ユニットは、クラス2タイプII Casタンパク質を含むことができる。クラス2タイプII Casタンパク質は、野生型の対応物と比較して突然変異したCasタンパク質であり得る。突然変異したCasタンパク質は、ヌクレアーゼ欠損であり得る。突然変異したCasタンパク質は、突然変異したCas9であり得る。突然変異したCas9は、Cas9D10Aであり得る。Cas9における突然変異の他の例としては、H820A、D839A、H840A、N863A、又はそれらの任意の組み合わせ、例えば、D10A/H820A、D10A、D10A/D839A/H840A、及びD10A/D839A/H840A/N863Aが挙げられる。本明細書に記載される突然変異は、SpCas9を参照しており、SpCas9以外のCRISPRタンパク質における類似

40

50

の突然変異も含む。プログラム可能なDNA結合ユニットは、ストレプトコッカス・ピオゲネス (*Streptococcus pyogenes*) Cas9 (SpCas9)、スタフィロコッカス・アウレウス (*Staphylococcus aureus*) Cas9 (SaCas9)、Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas6、Cas7、Cas8、Cas9、Cas100、Csy1、Csy2、Csy3、Cse1、Cse2、Csc1、Csc2、Csa5、Csn2、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Csb1、Csb2、Csb3、Csx17、Csx14、Csx10、Csx16、CsaX、Csx3、Csx1、Csx15、Csf1、Csf2、Csf3、Csf4、Cpf1、C2c1、C2c3、Cas12a、Cas12b、Cas12c、Cas12d、Cas12e、Cas13a、Cas13b、Cas13c、その誘導体、又はそれらの任意の組み合わせを含むことができる。様々な種のCas9分子を、本明細書に記載される方法及び組成物において使用することができる。S.ピオゲネス及びS.アウレウス (*S. aureus*) Cas9分子は、本明細書中の開示の大部分の主題であるが、本明細書に列挙される他の種のCas9タンパク質のCas9分子、それ由来のCas9分子、又はそれに基づくCas9分子を同様に使用することができる。これらには、例えば、アシドボラックス・アベナエ (*Acidovorax avenae*)、アクチノバチルス・ブルウロニューモニエ (*Actinobacillus pleuropneumoniae*)、アクチノバチルス・スクシノゲネス (*Actinobacillus succinogenes*)、アクチノバチルス・スイス (*Actinobacillus suis*)、アクチノミセス属種 (*Actinomyces sp.*)、シクリフィルス・デニトリフィカンス (*Cyclophilus denitrificans*)、アミノモナス・パウシボランス (*Aminomonas paucivorans*)、バチルス・セレウス (*Bacillus cereus*)、バチルス・スミス (*Bacillus smithii*)、バチルス・チューリングエンシス (*Bacillus thuringiensis*)、バクテロイデス属種 (*Bacteroides sp.*)、ブラストピレルラ・マリーナ (*Blastopirellula marina*)、ブラディリゾビウム属種 (*Bradyrhizobium sp.*)、プレビバチルス・ラテロスポラス (*Brevibacillus laterosporus*)、カンピロバクター・コリ (*Campylobacter coli*)、カンピロバクター・ジェジュニ (*Campylobacter jejuni*)、カンピロバクター・ラリ (*Campylobacter lari*)、カンジダトゥス・プニセイスピラム (*Candidatus Punicteispirillum*)、クロストリジウム・セルロリテイクム (*Clostridium cellulolyticum*)、クロストリジウム・パーフリングENS (*Clostridium perfringens*)、コリネバクテリウム・アコレNS (*Corynebacterium accolens*)、コリネバクテリウム・ジフテリア (*Corynebacterium diphtheria*)、コリネバクテリウム・マトウルコティ (*Corynebacterium matruchotii*)、ディノロセオバクター・シバエ (*Dinoroseobacter shibae*)、ユーバクテリウム・ドリクム (*Eubacterium dolichum*)、ガンマプロテオバクテリウム (*gamma proteobacterium*)、グルコンアセトバクター・ジアゾトロフィカス (*Gluconacetobacter diazotrophicus*)、ヘモフィルス・パラインフルエンザ (*Haemophilus parainfluenzae*)、ヘモフィルス・スプトルム (*Haemophilus sputorum*)、ヘリコバクター・カナデンシス (*Helicobacter canadensis*)、ヘリコバクター・シナエディ (*Helicobacter cinaedi*)、ヘリコバクター・ムステラエ (*Helicobacter mustelae*)、イリオバクター・ポリトロパス (*Ilyobacter polytropus*)、キングラ・キング (*Kingella kingae*)、ラクトバチルス・クリスパツス (*Lactobacillus crispatus*)、リステリア・イ

10

20

30

40

50

バノヴィ (*Listeria ivanovii*)、リステリア・モノサイトゲネス (*Listeria monocytogenes*)、リステリア・バクテリウム (*Listeria bacterium*)、メチロシステイス属種 (*Methylocystis sp.*)、メチロシヌス・トリコスポリウム (*Methylosinus trichosporium*)、モビルンカス・ムリエリス (*Mobiluncus mulieris*)、ナイセリア・バシリフォルミス (*Neisseria bacilliformis*)、ナイセリア・シネレア (*Neisseria cinerea*)、ナイセリア・フラベセンス (*Neisseria flavescens*)、ナイセリア・ラクタミカ (*Neisseria lactamica*)、ナイセリア・メニンギチジス (*Neisseria meningitidis*)、ナイセリア属種 (*Neisseria sp.*)、ナイセリア・ワズウォルチイ (*Neisseria wadsworthii*)、ニトロソモナス属種 (*Nitrosomonas sp.*)、バルビバクルム・ラヴァメンチボランス (*Parvibaculum lavamentivorans*)、パストレラ・ムルトシダ (*Pasteurella multocida*)、ファスコラルクトバクテリウム・サクシナツテンス (*Phascolarctobacterium succinatutens*)、ラルストニア・シジギイ (*Ralstonia solygyi*)、ロドプシュードモナス・パルストリス (*Rhodopseudomonas palustris*)、ロドブルム属種 (*Rhodovulum sp.*)、シモンシエラ・ムエレリ (*Simonsiella muelleri*)、スフィンゴモナス属種 (*Sphingomonas sp.*)、スポロラクトバチルス・ヴィネエ (*Sporolactobacillus vineae*)、スタフィロコッカス・ルグドゥネンシス (*Staphylococcus lugdunensis*)、ストレプトコッカス属種 (*Streptococcus sp.*)、スブドリグラヌラム属種 (*Subdoligranulum sp.*)、チストレラ・モビリス (*Tistrella mobilis*)、トレポネーマ属種 (*Treponema sp.*)、又はベルミネフロバクター・エイセニアエ (*Verminephrobacter eiseniae*) 由来の Cas 9 分子が含まれる。触媒的に不活性化する突然変異、及び上記突然変異体のヌクレアーゼ活性を評価する手段は、当業者に公知である。

【0076】

プログラム可能な DNA 結合ユニットは、ガイド分子を含むことができる。ガイド RNA 分子 (*sgRNA*) は、標的特異的な *crRNA* と、Cas 分子に結合する *tracrRNA* との 2 つの別々の分子で構成され得る。一部の実施形態では、*crRNA* 及び *tracrRNA* は、別個の分子として提供され、1 つは機能的な *sgRNA* にするためにアニールしなければならない。本明細書で使用される場合、用語「ガイド配列」及び「ガイド分子」は、CRISPR-Cas システムとの関連で、選択された結合部位とハイブリダイズするために選択された結合部位と十分な相補性を有する任意のポリヌクレオチド配列、及び選択された結合部位へのプログラム可能な DNA 結合ユニットの直接的な配列特異的結合を含む。*gRNA* 分子は、*gRNA* 分子 / Cas 9 分子複合体の標的結合部位への特異的標的化又はホーミングを促進する核酸を指すことができる。*gRNA* 分子は、単分子 (単一 RNA 分子を有する) (例えば、キメラ) 又はモジュラー (1 を超える、典型的には 2 つの別々の RNA 分子を含む) であり得る。本明細書に開示される方法を用いて作製されるガイド配列は、全長ガイド配列、切断型ガイド配列、全長 *sgRNA* 配列、切断型 *sgRNA* 配列、又は E + F *sgRNA* 配列であり得る。一部の実施形態では、適切なアラインメントアルゴリズムを用いて最適に整列された場合、所与の結合部位に対するガイド配列の相補性の程度は、約 50%、60%、75%、80%、85%、90%、95%、97.5%、99% 又はそれ以上である。ある特定の例示的な実施形態では、ガイド分子は、ガイド配列と結合部位の間に RNA 二重鎖が形成するように、結合部位と少なくとも 1 つのミスマッチを有するように設計され得るガイド配列を含む。したがって、相補性の程度は好ましくは 99% 未満である。例えば、ガイド配列が 24 ヌクレオチドからなる場合、相補性の程度は、より具体的には約 96% 又はそれ未満である。特定の実施

形態では、ガイド配列は、ガイド配列全体にわたる相補性の程度がさらに減少するように、2種以上の隣接するミスマッチヌクレオチドのストレッチを有するように設計される。例えば、ガイド配列が24ヌクレオチドからなる場合、相補性の程度は、2種以上のミスマッチヌクレオチドのストレッチが2、3、4、5、6又は7個のヌクレオチドなどを含むかどうかによって依存して、より具体的には約96%又はそれ未満、より具体的には約92%又はそれ未満、より具体的には約88%又はそれ未満、より具体的には約84%又はそれ未満、より具体的には約80%又はそれ未満、より具体的には約76%又はそれ未満、より具体的には約72%又はそれ未満である。一部の実施形態では、1種以上のミスマッチヌクレオチドのストレッチを除き、適切なアラインメントアルゴリズムを用いて最適に整列された場合の相補性の程度は、約50%、60%、75%、80%、85%、90%、95%、97.5%、99%又はそれ以上である。最適アラインメントは、配列を整列させるための任意の適切なアルゴリズムを使用して決定することができ、その非限定的な例には、Smith-Watermanアルゴリズム、Needleman-Wunschアルゴリズム、Burrows-Wheeler変換に基づくアルゴリズム(例えば、Burrows Wheeler Aligner)、Clustal W、Clustal X、Clustal Omega、BLAT、Novoalign(Novocraft Technologies; www.novocraft.comで入手可能)、ELAND(Illumina, San Diego, CA)、SOAP(soap.genomics.org.cnで入手可能)、及びMaq(maq.sourceforge.netで入手可能)が挙げられる。選択された結合部位へのプログラム可能なDNA結合ユニットの配列特異的結合を指示するガイド配列(核酸標的化ガイドRNA内)の能力は、任意の適切なアッセイによって評価され得る。一部の実施形態では、ガイド配列は、長さが10~50ntであるが、より具体的には約20~30nt、有利には約20nt、23~25nt又は24ntのRNA配列である。ガイド配列は、それが選択された結合部位にハイブリダイズすることを確実にするように選択することができる。

10

20

30

40

50

【0077】

死滅型ガイド配列

プログラム可能なDNA結合ユニットは、CRISPR関連タンパク質(CASタンパク質)と、標的dsDNAの結合部位に特異的に結合することができるガイドRNA(gRNA)とを含むことができる。一部の実施形態では、ガイド配列は、CRISPR Cas複合体の形成を可能にし、結合部位への結合に成功し、一方で同時に、成功するヌクレアーゼ活性を許容しない様式で修飾される。このような修飾されたガイド配列は、「死滅型ガイド」又は「死滅型ガイド配列」と称される。これらの死滅型ガイド又は死滅型ガイド配列は、ヌクレアーゼ活性に関して触媒的に不活性であるか又は立体構造的に不活性であると考えられることができる。プログラム可能なDNA結合ユニットは、機能的Casタンパク質及びガイドRNA(gRNA)又はcrRNAを含むことができ、gRNA又はcrRNAは、死滅型ガイド配列を含み、それにより、gRNAは、Casタンパク質が、非突然変異体Casタンパク質の検出可能な切断活性なしに選択された結合部位に向けられるように、選択された結合部位にハイブリダイズすることができる。CRISPR複合体の結合部位への配列特異的結合を指示する死滅型ガイド配列の能力は、任意の適切なアッセイによって評価され得る。死滅型ガイド配列は、典型的には、活性切断をもたらすそれぞれのガイド配列よりも短くすることができる。特定の実施形態では、死滅型ガイドは、同じ結合部位に向けられたそれぞれのガイドよりも5%、10%、20%、30%、40%、50%短い。

【0078】

タンパク質構成成分

プログラム可能なDNA結合ユニットは、標的dsDNA上の結合部位に特異的に結合することができるタンパク質構成成分を含むことができる。タンパク質構成成分は、エンドヌクレアーゼ欠損ジンクフィンガーヌクレアーゼ(ZFN)、エンドヌクレアーゼ欠損転写アクチベーター様エフェクターヌクレアーゼ(TALEN)、アルゴノートタンパク

質、エンドヌクレアーゼ欠損メガヌクレアーゼ、リコンビナーゼ、又はそれらの組み合わせを含むことができる。一部の実施形態では、プログラム可能なDNA結合ユニットは、ヌクレアーゼドメインを有さない。一部の実施形態では、プログラム可能なDNA結合ユニットは、1種以上の突然変異を介して触媒的に不活性化にされたヌクレアーゼドメインを有する。触媒的に不活性化する突然変異及び上記突然変異体のヌクレアーゼ活性を評価する手段は、当業者に公知である。

【0079】

転写アクチベーター様エフェクター (T A L E)

プログラム可能なDNA結合ユニットは、エンドヌクレアーゼ欠損転写アクチベーター様エフェクターヌクレアーゼ (T A L E N)、その機能的断片、又はそのバリエーションを含むことができる。転写アクチベーター様エフェクター (T A L E) は、実質的に任意の所望のDNA配列に結合するように操作することができる。T A L E N システムを使用する標的化の例示的な方法は、例えば、Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, et al. Efficient design and assembly of custom TAL EN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res.* 2011;39:e82; Zhang F, Cong L, Lodato S, Kosuri S, Church G M, Arlotta P Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol.* 2011;29:149-153、米国特許第8,450,471号、米国特許第8,440,431号及び米国特許第8,440,432号に見出され得、すべては参照により具体的に組み込まれる。

10

20

【0080】

プログラム可能なDNA結合ユニットは、T A L E ポリペプチドを含むことができる。T A L E は、植物病原体キサントモナス (*Xanthomonas*) 由来の転写因子であり、新しいDNA標的に結合するように容易に操作することができる。本明細書に提供される一部の実施形態では、T A L E は、エンドヌクレアーゼの触媒ドメイン (例えば、F o k I) と連結していない。本明細書に提供される一部の実施形態では、プログラム可能なDNA結合ユニットは、エンドヌクレアーゼドメインが触媒的に不活性であるT A L E Nを含むことができる。T A L E ポリペプチドは、高度に保存された単量体ポリペプチドのタンデム反復で構成される核酸結合ドメインを含み、長さは主に33、34又は35アミノ酸であり、主にアミノ酸位置12及び13において互いに異なる。本明細書で使用される場合、用語「ポリペプチド単量体」又は「T A L E 単量体」は、T A L E 核酸結合ドメイン内の高度に保存された反復ポリペプチド配列を指すために使用され、用語「反復可変二残基」又は「R V D」は、ポリペプチドモノマーの位置12及び13における高度に変化するアミノ酸を指すために使用される。T A L E 単量体は、そのR V D中のアミノ酸の同一性によって決定されるヌクレオチド結合親和性を有する。例えば、N IのR V Dを有するポリペプチド単量体は、アデニン (A) に優先的に結合し、N GのR V Dを有するポリペプチド単量体は、チミン (T) に優先的に結合し、H DのR V Dを有するポリペプチド単量体は、シトシン (C) に優先的に結合し、N NのR V Dを有するポリペプチド単量体は、アデニン (A) とグアニン (G) の両方に優先的に結合する。一部の実施形態では、I GのR V Dを有するポリペプチド単量体は、Tに優先的に結合する。したがって、T A L E の核酸結合ドメインにおけるポリペプチド単量体反復の数及び順序が、その核酸標的特異性を決定する。本明細書に提供されるなおさらなる実施形態では、N SのR V Dを有するポリペプチド単量体は、すべての4つの塩基対を認識し、A、T、G又はCに結合し得る。T A L E の構造及び機能は、例えば、Moscou et al., *Science* 326:1501 (2009); Boch et al., *Science* 326:1509-1512 (2009); Zhang et al., *Nature Biotechnology* 29: 149-153 (2011)にさらに記載され、各々は参照によりその全体が組み込まれる。プログラム可能なDNA結合ユニットは、特定の核酸配列を標的とするように設計されたポリペプチド単量体反復を含むことができる。

30

40

【0081】

Zhang et al., *Nature Biotechnology* 29:149-153 (2011)に記載されるよう

50

に、T A L E ポリペプチド結合効率は、天然に存在するT A L E のD N A 結合領域の直接的にN末端又はC末端である「キャッピング領域」からのアミノ酸配列を、操作されたT A L E D N A 結合領域のN末端又はC末端の位置で操作されたT A L E に含めることによって増大させることができる。したがって、ある特定の実施形態では、本明細書に記載されるT A L E ポリペプチドは、N末端キャッピング領域及び/又はC末端キャッピング領域をさらに含む。

本明細書で使用される場合、N末端キャッピング領域の所定の「N末端」から「C末端」配向への、反復T A L E 単量体及びC末端キャッピング領域を含むD N A 結合ドメインは、本明細書に提供されるT A L E 又はポリペプチドにおける異なるドメインの組織化のための構造的基礎を提供する。

N末端及び/又はC末端キャッピング領域全体は、D N A 結合領域の結合活性を増強するために必要ではない。したがって、ある特定の実施形態では、N末端及び/又はC末端キャッピング領域の断片は、本明細書に記載されるT A L E ポリペプチドに含まれる。

【0082】

ある特定の実施形態では、本明細書に記載されるT A L E ポリペプチドは、N末端キャッピング領域の少なくとも10、20、30、40、50、54、60、70、80、87、90、94、100、102、110、117、120、130、140、147、150、160、170、180、190、200、210、220、230、240、250、260又は270個のアミノ酸を含んだN末端キャッピング領域断片を含む。ある特定の実施形態では、N末端キャッピング領域断片アミノ酸は、N末端キャッピング領域のC末端(D N A 結合領域近位端)のものである。Zhang et al., *Nature Biotechnology* 29:149-153 (2011)に記載されるように、C末端240アミノ酸を含むN末端キャッピング領域断片は、全長キャッピング領域に等しい結合活性を増強し、一方、C末端147アミノ酸を含む断片は、全長キャッピング領域の有効性の80%を超えて保持し、C末端117アミノ酸を含む断片は、全長キャッピング領域の活性の50%を超えて保持する。

【0083】

一部の実施形態では、本明細書に記載されるT A L E ポリペプチドは、C末端キャッピング領域の少なくとも6、10、20、30、37、40、50、60、68、70、80、90、100、110、120、127、130、140、150、155、160、170、180アミノ酸を含んだC末端キャッピング領域断片を含む。ある特定の実施形態では、C末端キャッピング領域断片アミノ酸は、C末端キャッピング領域のN末端(D N A 結合領域近位端)のものである。Zhang et al., *Nature Biotechnology* 29:149-153 (2011)に記載されるように、C末端68アミノ酸を含むC末端キャッピング領域断片は、全長キャッピング領域に等しい結合活性を増強し、一方、C末端20アミノ酸を含む断片は、全長キャッピング領域の有効性の50%を超える保持する。

【0084】

ジンクフィンガー(ZF)タンパク質

プログラム可能なD N A 結合ユニットは、Znフィンガー(ZF)ヌクレアーゼ、その機能的断片、又はそのバリエーションを含むことができる。プログラム可能なD N A 結合ユニットは、エンドヌクレアーゼ欠損ZFヌクレアーゼ、その機能的断片、又はそのバリエーションを含み得、エンドヌクレアーゼのドメイン(例えば、FokI)は、触媒的に不活性であるか又は存在しない。プログラム可能なD N A 結合ユニットは、ZFタンパク質(ZFP)を含むことができる。ZFPは、選択される標的部位に結合するように操作することができる。例えば、Beerli et al. (2002) *Nature Biotechnol.* 20:135-141; Pabo et al. (2001) *Ann. Rev. Biochem.* 70:313-340; Isalan et al. (2001) *Nature Biotechnol.* 19:656-660; Segal et al. (2001) *Curr. Opin. Biotechnol.* 12:632-637; Choo et al. (2000) *Curr. Opin. Struct Biol.* 10:411-416; U. S. Pat. Nos. 6,453,242; 6,534,261; 6,599,692; 6,503,717; 6,689,558; 7,030,215; 6,794,136; 7,067,317; 7,262,054; 7,070,934; 7,361,635; 7,25

10

20

30

40

50

3,273 ; 米国特許出願公開第 2005 / 0064474 号 ; 米国特許出願公開第 2007 / 0218528 号 ; 米国特許出願公開第 2005 / 0267061 号を参照されたい。ZFP は、所望の DNA 結合部位を標的とする ZF モジュールのアレイを含むことができる。ZF アレイの各フィンガーモジュールは、3つの DNA 塩基を標的とすることができる。個々のジンクフィンガードメインのカスタマイズされたアレイを ZFP にアセンブルさせることができる。

【0085】

メガヌクレアーゼ

プログラム可能な DNA 結合ユニットは、エンドヌクレアーゼ欠損メガヌクレアーゼ、その機能的断片、又はそのバリエーションであり得る。メガヌクレアーゼの DNA 結合ドメインは、12 ~ 45 bp の二本鎖 DNA 標的配列を有し得る。一部の実施形態では、メガヌクレアーゼは、各メガヌクレアーゼドメインが単量体上にある二量体酵素、又は単一のポリペプチド上に2つのドメインを含む単量体酵素のいずれかである。野生型メガヌクレアーゼだけでなく、様々なメガヌクレアーゼバリエーションもまた、無数の固有の配列の組み合わせをカバーするためのタンパク質工学によって生成されている。一部の実施形態では、メガヌクレアーゼ A の半部位及びタンパク質 B の半部位で構成される認識部位を有するキメラメガヌクレアーゼを使用することもできる。このようなキメラメガヌクレアーゼの具体例は、I - Dmo I 及び I - Cre I のタンパク質ドメインを含む。メガヌクレアーゼの例には、LAGLIDADG ファミリー由来のホーミングエンドヌクレアーゼが含まれる。「LAGLIDADG メガヌクレアーゼ」とは、Stoddard et al (Stoddard, 2005) に定義される LAGLIDADG ファミリー由来のホーミングエンドヌクレアーゼ、又は上記天然ホーミングエンドヌクレアーゼと少なくとも 80%、85%、90%、95%、97.5%、99% 若しくはそれ以上の同一性又は類似性を共有するポリペプチドを含む操作されたバリエーションを指す。このような操作された LAGLIDADG メガヌクレアーゼは、単量体又は二量体メガヌクレアーゼに由来し得る。二量体メガヌクレアーゼに由来する場合、このような操作された LAGLIDADG メガヌクレアーゼは、一本鎖又は二量体エンドヌクレアーゼであり得る。メガヌクレアーゼは、当業者に周知である技術を用いてそれらの認識配列を修飾することによって、特定の配列を標的とすることができる。例えば、Epinat et al., 2003, Nuc. Acid Res., 31(11):2952-62 and Stoddard, 2005, Quarterly Review of Biophysics, pp. 1-47 を参照されたい。

【0086】

LAGLIDADG メガヌクレアーゼは、I - Sce I、I - Chu I、I - Cre I、I - Csm I、PI - Sce I、PI - Tli I、PI - Mtu I、I - Ceu I、I - Sce II、I - Sce III、HO、PI - Civ I、PI - Ctr I、PI - Aae I、PI - Bsu I、PI - Dha I、PI - Dra I、PI - Mav I、PI - Mch I、PI - Mfu I、PI - Mfl I、PI - Mga I、PI - Mgo I、PI - Min I、PI - Mka I、PI - Mle I、PI - Mma I、PI - Msh I、PI - Msm I、PI - Mth I、PI - Mtu I、PI - Mxe I、PI - Npu I、PI - Pfu I、PI - Rma I、PI - Spb I、PI - Ssp I、PI - Fac I、PI - Mja I、PI - Pho I、PI - Tag I、PI - Thy I、PI - Tko I、PI - Tsp I、若しくは I - Mso I であり得るか ; 又は、ホモ二量体、ヘテロ二量体若しくは単量体のいずれであっても、その機能的な突然変異体若しくはバリエーションであり得る。一部の実施形態では、LAGLIDADG メガヌクレアーゼは、I - Cre I 誘導体である。一部の実施形態では、LAGLIDADG メガヌクレアーゼは、天然 I - Cre I LAGLIDADG メガヌクレアーゼと少なくとも 80% の類似性を共有する。一部の実施形態では、LAGLIDADG メガヌクレアーゼは、天然 I - Cre I LAGLIDADG メガヌクレアーゼの残基 1 ~ 152 と少なくとも 80% の類似性を共有する。一部の実施形態では、LAGLIDADG メガヌクレアーゼは、リンカーペプチドの有無にかかわらず、一緒に連結した天然 I - Cre I LAGLIDADG メガヌクレアーゼの残基 1 ~ 152 と少なくとも 80% の類似性を共有する2つの単量体からなり得る。

【0087】

アルゴノートタンパク質

一部の実施形態では、プログラム可能なDNA結合ユニットは、ヌクレアーゼ不活性アルゴノートを含む。一部の実施形態では、プログラム可能なDNA結合ユニットは、ナトロノバクテリウム・グレゴリ (*Natronobacterium gregoryi*) (NgAgO) 由来のアルゴノートタンパク質、その機能的断片、又はそのバリエーションを含む。NgAgOは、ssDNA誘導エンドヌクレアーゼである。NgAgOはおよそ24ヌクレオチドの5'リン酸化されたssDNA (gDNA) に結合して標的部位に到達し、gDNA部位でDNA二本鎖切断を起こす。一部の実施形態では、プログラム可能なDNA結合ユニットは、ヌクレアーゼ不活性NgAgO (dNgAgO) を含む。NgAgOの特徴付け及び使用は、Gao et al, Nat Biotechnol. Epub 2016 May 2. PubMed PMID: 27136078; Swarts et al, Nature. 507(7491) (2014):258-61; Swarts et al, Nucleic Acids Res. 43(10) (2015):5120-9に記載され、各々が参照により本明細書に組み込まれる。NgAgOベースのプログラム可能なDNA結合ユニットは、少なくとも1つのガイドDNAエレメント、又はガイドDNAエレメントをコードする核酸配列を含む核酸を含むことができ、結合部位のDNAと直接塩基対を介して、結合部位の特異的な標的化又は認識を達成することができる。アルゴノートタンパク質の原核生物相同体は公知であり、例えば、Makarova K., et al., "Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements", Biol. Direct. 2009 Aug. 25; 4:29. doi: 10.1186/1745-6150-4-29に記載され、参照により本明細書に組み込まれる。一部の実施形態では、プログラム可能なDNA結合ユニットは、マリニトガ・ピエゾフィラ (*Marinitoga piezophila*) アルゴノート (MpAgO) タンパク質、その機能的断片、又はそのバリエーションである。

10

20

30

40

50

【0088】

リコンビナーゼ

一部の実施形態では、プログラム可能なDNA結合ユニットは、標的dsDNA上の結合部位に結合するように構成されたリコンビナーゼを含む。部位特異的リコンビナーゼは、当該技術分野において周知であり、一般的にインベルターゼ、リゾルベース、又はインテグラーゼと称され得る。部位特異的リコンビナーゼの非限定的な例としては、限定されないが、ラムダイнтеグラーゼ、Cre、Int、IHF、Xis、Flp、Fis、Hin、Gin、phiC31、Cin、Tn3レゾルベース、TndX、XerC、XerD、TnpX、Hjc、Gin、SpCCel、及びParAが挙げられる。

【0089】

リンカー

トランスポソームは、トランスポゼースとdCASタンパク質を接続するリンカーを介して、プログラム可能なDNA結合ユニットと結び付けられ得る。リンカーは、ペプチドリンカー、化学リンカー、又はその両方を含むことができる。トランスポゼースは、dCASタンパク質を含む融合タンパク質中に存在することができる。トランスポソームは、トランスポゼースとタンパク質構成成分を接続するリンカーを介して、プログラム可能なDNA結合ユニットと結び付けられ得る。ペプチドリンカーは、複数のグリシン、セリン、スレオニン、アラニン、リジン、グルタミン、又はそれらの組み合わせを含むことができる。ペプチドリンカーは、GSリンカーを含み得る。ペプチドリンカーは、XTENリンカーであり得る。タンパク質構成成分は、トランスポゼースを含む融合タンパク質中に存在することができる。本明細書で使用される用語「リンカー」とは、分子間又は分子の部分間の相互作用を促進する分子を指す。一部の実施形態では、リンカーは、ポリペプチドリンカーである。一部の実施形態では、リンカーは、化学リンカーである。「ペプチドリンカー」又は「ポリペプチドリンカー」という用語は、本明細書で使用される場合、ペプチド結合によって接続された2種以上のアミノ酸残基を含むペプチド又はポリペプチドを意味する。このようなペプチド又はポリペプチドリンカーは、当技術分野において周知

である。リンカーは、天然に存在し及び/若しくは天然に存在しないペプチド又はポリペプチドを含むことができる。リンカーは、トランスポゼース及び/又はプログラム可能なDNA結合ユニットのC末端及び/又はN末端と結び付けられ得る。

リンカーは、化学リンカー又はペプチドリッカーであり得る。したがって、実施形態は、ペプチド結合を介して他の分子にコンジュゲートされたポリペプチド、及び化学的コンジュゲーションを介して他の分子にコンジュゲートされたポリペプチドに関する。

【0090】

ある程度の柔軟性を有するペプチドリッカーを使用することができる。ペプチドリッカーは、適切なペプチドリッカーが、一般的に柔軟なペプチドをもたらす配列を有することを念頭に置いて、実質的に任意のアミノ酸配列を有することができる。グリシン及びアラニンなどの小さなアミノ酸の使用は、柔軟なペプチドを作製するのに使用されるものである。このような配列の作製は、当業者にとって習慣的である。

10

【0091】

適切なリンカーは、容易に選択され得、任意の適切な長さを有するものであり得て、例えば、1アミノ酸(例えば、Gly)~50アミノ酸、例えば、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、又はこれらの値の任意の2つの間の数値若しくは範囲のアミノ酸である。

20

【0092】

好ましいペプチドリッカー配列は、柔軟な伸長した立体構造を採用し、秩序だった二次構造を発生する傾向を示さない。ある特定の実施形態では、リンカーは、単量体、二量体、多量体又は高分子であり得る化学部分であり得る。好ましくは、リンカーはアミノ酸を含む。柔軟なリンカー中の典型的なアミノ酸には、Gly、Asn及びSerが含まれる。したがって、特定の実施形態では、リンカーは、Gly、Asn及びSerアミノ酸のうち1種以上の組み合わせを含む。リンカー配列においては、Thr及びAlaなどの他の中性に近いアミノ酸も使用することができる。柔軟なリンカーの例としては、グリシンポリマー(G)_n(配列番号32)、グリシン-セリンポリマー(例えば、(GS)_n(配列番号33)、(GSGGS)_n(配列番号34)、(G4S)_n(配列番号35)及び(GGGS)_n(配列番号36)が挙げられ、ここでnは少なくとも1の整数である。一部の実施形態では、nは、少なくとも、せいぜい、又は正確に1、2、3、4、5、6、7、8、9又は10(又はその中の任意の導出可能な範囲)である。グリシン-アラニンポリマー、アラニン-セリンポリマー、及び当技術分野において公知である他の柔軟なリンカー。グリシン及びグリシン-セリンポリマーを使用することができ、Gly及びSerはいずれも比較的構造化されておらず、したがって、構成成分間の中立的係留として働くことができる。グリシンポリマーを用いることができ、グリシンはアラニンよりもさらに著しく大きなファイ-プサイ空間にアクセスし、より長い側鎖を有する残基よりもはるかに低く制限される。例示的スペーサーは、限定されないが、GGS(配列番号37)、GGS(配列番号38)、GSGSG(配列番号39)、GSGGG(配列番号40)、GGGS(配列番号41)、GSSSG(配列番号42)などを含むアミノ酸配列を含むことができる。リンカー配列においては、Thr及びAlaなどの他の中性に近いアミノ酸も使用することができる。リンカー配列の長さは、融合タンパク質の機能又は活性に著しく影響することなく変動し得る(例えば、米国特許第6,087,329号を参照されたい)。一部の実施形態では、リンカーは、少なくとも、せいぜい、又は正確に4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、55、60、65、70、75、80、85、90、95、又は100アミノ酸残基(又はその中で導出可能な任意の範囲)であり得る。

30

40

50

【 0 0 9 3 】

一部の実施形態では、ポリペプチドリンカーは、X T E N リンカーである。一部の実施形態では、リンカーは、X T E N リンカー、又は X T E N リンカーのバリエーション、例えば、S G S E T P G T S E S A (配列番号 4 3)、S G S E T P G T S E S A T P E S (配列番号 4 4) 若しくは S G S E T P G T S E S A T P E G G S G G S (配列番号 4 5) である。X T E N リンカーは、例えば、Schellenberger et al. (2009), Nature Biotechnology 27: 1186-1190 に記載されており、その全内容は参照により本明細書に組み込まれる。

【 0 0 9 4 】

本明細書に提供される方法において使用するための適切なリンカーは、当業者に周知であり、限定されないが、直鎖若しくは分枝鎖炭素リンカー、複素環炭素リンカー、又はペプチドリンカーを含む。しかしながら、本明細書で使用される場合、リンカーはまた、共有結合(炭素-炭素結合又は炭素-ヘテロ原子結合)であり得る。特定の実施形態では、リンカーは、各タンパク質がその必要な機能的特性を保持することを確実にするのに十分な距離だけ、トランスポソーム及びプログラム可能な DNA 結合ユニットを分離するために使用される。

【 0 0 9 5 】

リンカーを用いて、2つのタンパク質パートナーを融合させて融合タンパク質を形成することができる。「リンカー」は、2つの分子又は部分、例えば融合タンパク質の2つのドメインを連結する化学基又は分子であり得る。典型的には、リンカーは、2つの基、分子、ドメイン、又は他の部分の間に配置され(それに隣接する)、共有結合を介して各々に接続され、したがって、2つを接続する。一部の実施形態では、リンカーは、アミノ酸又は複数のアミノ酸(例えば、ペプチド又はタンパク質)である。一部の実施形態では、リンカーは、有機分子、基、ポリマー(例えば、非天然ポリマー、非ペプチドポリマー)、又は化学的部分である。一部の実施形態では、リンカーは、直接結合又は原子、例えば、酸素(O)又は硫黄(S)、ユニット、例えば、- N R - (Rは、水素又はアルキル、- C (O) -、- C (O) O -、- C (O) N H -、S O、S O₂、- S O₂ N H - である)、又は原子の鎖、例えば、置換又は非置換アルキル、置換又は非置換アルケニル、置換又は非置換アルキニル、アリールアルキル、ヘテロアリールアルキルを含む。一部の実施形態では、原子の鎖中の1つ以上のメチレンは、O、S、S(O)、S O₂、- S O₂ N H -、- N R -、- N R₂、- C (O) -、- C (O) O -、- C (O) N H -、切断可能な連結基、置換又は非置換アリール、置換又は非置換ヘテロアリール、及び置換又は非置換ヘテロ環のうち1種以上で置換され得る。また、リンカーの例には、化学部分及びコンジュゲート剤、例えば、スルホ-スクシンイミジル誘導體(スルホ-S M C C、スルホ-S M P B)、ジスクシンイミジルスベレート(D S S)、ジスクシンイミジルグルタル酸(D S G)及びジスクシンイミジル酒石酸(D S T)が含まれ得る。リンカーの例は、C N (ここで、N = 1 ~ 100 炭素原子)などの直鎖状炭素鎖をさらに含む。一部の実施形態では、リンカーは、ジペプチドリンカー、例えば、バリン-シトルリン(v a l - c i t)、フェニルアラニン-リジン(p h e - l y s)リンカー、又はマレイミドカブロン-バリン-シトルリン-p-アミノベンジルオキシカルボニル(v c)リンカーであり得る。一部の実施形態では、リンカーは、スルホスクシンイミジル-4-[N-マレイミドメチル]シクロヘキサン-1-カルボキシレート(s m c c)である。スルホ-s m c c コンジュゲーションは、スルフヒドリル(チオール、- S H)と反応するマレイミド基を介して起こり、一方、そのスルホ-N H S エステルは、第1級アミン(リジン及びタンパク質又はペプチドN末端に見られる)に対して反応性である。さらに、リンカーはマレイミドカブロイル(m e)であり得る。一部の実施形態では、共有連結は、T r a u t 試薬の使用を介して達成され得る。

【 0 0 9 6 】

図 8 ~ 10 は、本明細書に提供されるタンパク質複合体の生成に使用するための、プラスミド構築物 3 X F l a g - C a s 9 - F l 2 6 - T n 5 (配列番号 1)、3 X F l a g

10

20

30

40

50

- C a s 9 - x T e n - T n 5 (配列番号 2)、及び p E T - T n 5 - x T e n - d C a s 9 (配列番号 3) のそれぞれの非限定的であり例示的な概略図を示す。本明細書に開示されるタンパク質複合体、リンカー、プログラム可能な DNA 結合ユニット、及び/又はトランスポゼースは、配列番号 1 ~ 3 にコードされるタンパク質複合体、リンカー、プログラム可能な DNA 結合ユニット、及び/又はトランスポゼースと、少なくとも約 50 %、51 %、52 %、53 %、54 %、55 %、56 %、57 %、58 %、59 %、60 %、61 %、62 %、63 %、64 %、65 %、66 %、67 %、68 %、69 %、70 %、71 %、72 %、73 %、74 %、75 %、76 %、77 %、78 %、79 %、80 %、81 %、82 %、83 %、84 %、85 %、86 %、87 %、88 %、89 %、90 %、91 %、92 %、93 %、94 %、95 %、96 %、97 %、98 %、99 %、100 %、又はこれらの値の任意の 2 つの間の数値若しくは範囲で同一であるヌクレオチド配列によってコードされ得る。

【 0 0 9 7 】

増幅

本方法は、dsDNA断片の末端でアダプターに結合することができるプライマーを用いて、複数のdsDNA断片を増幅することを含むことができる。増幅は、核酸増幅産物を生成することができる。核酸増幅産物は、ライブラリー（例えば、シーケンシングライブラリー）を含み得る。各プライマーは、長さが約 5 ~ 80 ヌクレオチド（例えば、長さが約 5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、60、70、80、又はこれらの値の任意の 2 つの間の数値若しくは範囲のヌクレオチド）であり得る。複数の dsDNA 断片をプライマーで増幅することは、ポリメラーゼ連鎖反応（PCR）を用いて行うことができる。PCRは、ループ媒介等温増幅（LAMP）、ヘリカーゼ依存性増幅（HDA）、リコンビナーゼポリメラーゼ増幅（RPA）、鎖置換増幅（SDA）、核酸配列ベース増幅（NASBA）、転写媒介増幅（TMA）、ニックング酵素増幅反応（NEAR）、ローリングサークル増幅（RCA）、多置換増幅（MDA）、分岐化（RAM）、環状ヘリカーゼ依存性増幅（cHDA）、単一プライマー等温増幅（SPIA）、RNA技術のシグナル媒介増幅（SMART）、自己持続配列複製（3SR）、ゲノム指数増幅反応（GEAR）、又は等温多置換増幅（IMDA）であり得る。PCRは、リアルタイムPCR又は定量的リアルタイムPCR（QRT-PCR）であり得る。

【 0 0 9 8 】

本明細書で使用される場合、核酸増幅は、配列特異的方法を使用して、標的核酸配列又はその相補体若しくはその断片の複数コピーを得るための任意の公知の手法を指すことができる。公知の増幅方法の例としては、限定されないが、ポリメラーゼ連鎖反応（PCR）、リガーゼ連鎖反応（LCR）、ループ媒介等温増幅（LAMP）、鎖置換増幅（SDA）（例えば、多置換増幅（MDA））、レプリカーゼ媒介増幅、免疫増幅、核酸配列ベース増幅（NASBA）、自己持続配列複製（3SR）、ローリングサークル増幅、及び転写媒介増幅（TMA）が挙げられる。一部の実施形態では、2種以上の上記核酸増幅方法を、例えば、連続的に行うことができる。

【 0 0 9 9 】

例えば、LCR増幅は、ハイブリダイゼーション、ライゲーション、及び変性の複数サイクルを用いることによって、標的及びその相補鎖を増幅するために、少なくとも4つの別々のオリゴヌクレオチドを使用する。SDAは、標的配列を含むヘミ修飾されたDNA二重鎖の一方の鎖にニックを入れ、続いて、一連のプライマー伸長及び鎖置換ステップにおける増幅を行う制限エンドヌクレアーゼの認識部位を含有するプライマーを用いることによって増幅する。

【 0 1 0 0 】

PCRは、核酸の増幅のための当該技術分野において周知である方法である。PCRは、標的配列に隣接する2種以上の伸長可能な配列特異的オリゴヌクレオチドプライマーを用いた標的配列の増幅を伴う。目的とする標的配列を含有する核酸は、プライマー、熱安

定性DNAポリメラーゼ（例えば、Taqポリメラーゼ）及び種々のdNTPの存在下での熱サイクル（変性、アニーリング及び伸長）の複数ラウンドのプログラムに供され、標的配列の増幅がもたらされる。PCRは、DNA分子の定義された領域の相補鎖が熱安定性DNAポリメラーゼによって同時に合成される複数ラウンドのプライマー伸長反応を使用する。各サイクルの終わりに、新たに合成された各DNA分子は次のサイクルの鋳型として働く。これらの反応の反復ラウンド中に、新たに合成されたDNA鎖の数は指数関数的に増加し、そのため、20～30回の反応サイクル後、最初の鋳型DNAは数千倍又は百万倍複製される。

PCRは、増幅後処理に適した二本鎖増幅産物を生成することができる。所望であれば、増幅産物は、アガロースゲル電気泳動による可視化、プローブベースの比色検出を用いる酵素イムノアッセイフォーマット、蛍光発光技術、又は当業者に公知である他の検出手段によって検出することができる。

10

【0101】

PCR法の例としては、限定されないが、リアルタイムPCR、エンドポイントPCR、増幅断片長多型PCR（AFLP-PCR）、Alu-PCR、非対称PCR、コロニーPCR、DD-PCR、縮退PCR、ホットスタートPCR、インサイチュPCR、インバースPCR、ロング-PCR、多重PCR、ネステッドPCR、PCR-ELISA、PCR-RFLP、PCR-一本鎖立体構造多型（PCR-SSCP）、定量的競合PCR（QC-PCR）、cDNA末端の迅速増幅-PCR（RACE-PCR）、多型DNAのランダム増幅-PCR（RAPD-PCR）、リアルタイムPCR、反復遺伝子外
回文PCR（Rep-PCR）、逆転写酵素PCR（RT-PCR）、TAIL-PCR、タッチダウンPCR、及びVectorette PCRが挙げられる。

20

【0102】

定量的リアルタイムポリメラーゼ連鎖反応（QRT-PCR）とも呼ばれるリアルタイムPCRは、所与の核酸分子の特定の部分を同時に定量及び増幅するために使用することができる。これは、特定の配列が試料中に存在するかどうか、及び存在する場合には、存在する配列のコピー数を決定するために使用することができる。用語「リアルタイム」とは、PCR中の定期的なモニタリングを指すことができる。ABI 7700及び7900HT配列検出システム（Applied Biosystems、Foster City、CA）などのある特定のシステムは、予め決定されているか又はユーザーが定義した
点で、各熱サイクル中にモニタリングを行う。蛍光共鳴エネルギー移動（FRET）プローブを用いたPCRのリアルタイム分析は、サイクルからサイクルへの蛍光色素シグナル変化、好ましくは任意の内部制御シグナルを減じた蛍光色素シグナル変化を測定する。リアルタイム手法はPCRの一般的なパターンに従うが、核酸は増幅の各ラウンド後に定量される。定量化の方法の2つの例は、二本鎖DNAにインターカレートする蛍光色素（例えば、SYBR Green）、及び相補的DNAとハイブリダイズした場合に蛍光を発する修飾されたDNAオリゴヌクレオチドプローブの使用である。インターカレーティング剤は、結合していない場合、比較的低い蛍光であり、二本鎖核酸への結合時は比較的高い蛍光を有する。したがって、インターカレーティング剤を使用して、核酸増幅反応中の二本鎖核酸の蓄積をモニタリングすることができる。本明細書に開示される実施形態
において有用なこのような非特異的色素の例としては、SYBR Green I（Molecular Probes）、ヨウ化プロピジウム、臭化エチジウムなどのインターカレーティング剤が挙げられる。

30

40

【0103】

標識

本明細書に記載される方法は、（例えば、検出可能な標識を用いて）複数のdsDNA断片のうちの一つ以上の一方又は両方の末端を標識することを含むことができる。本方法は、複数のdsDNA断片のうちの一つ以上の二つの末端を異なるように標識することを含むことができる。標識は、検出可能な標識（例えば、アニオン標識、カチオン標識、中性標識、電気化学標識、タンパク質標識、蛍光標識、磁性標識、又はそれらの組み合わせ

50

)による標識を含むことができる。本方法は、標識された dsDNA断片を濃縮すること、標識された dsDNA断片を捕捉すること、標識された dsDNA断片を単離すること、及び/又は標識された dsDNA断片を可視化することを含むことができる。本方法は、検出可能な標識のモニタリング(例えば、化学的モニタリング)を含むことができる。

【0104】

一部の実施形態では、検出可能な部分(例えば、検出可能な標識)は、光学部分、発光部分、電気化学的に活性な部分、ナノ粒子、又はそれらの組み合わせを含む。一部の実施形態では、発光部分は、化学発光部分、エレクトロルミネッセンス部分、フォトルミネッセンス部分、又はそれらの組み合わせを含む。一部の実施形態では、フォトルミネッセンス部分は、蛍光部分、リン光性部分、又はそれらの組み合わせを含む。一部の実施形態では、蛍光部分は、蛍光色素を含む。一部の実施形態では、ナノ粒子は、量子ドットを含む。一部の実施形態では、本方法は、検出可能な部分前駆体を検出可能な部分に変換するための反応を行うことを含む。一部の実施形態では、検出可能な部分前駆体を検出可能な部分に変換するための反応を行うことは、検出可能な部分前駆体を基質と接触させることを含む。一部のこのような実施形態では、検出可能な部分前駆体を基質と接触させると、2つの分子間の反応の検出可能な副産物が生じる。

10

【0105】

増幅産物の検出及び定量化

本明細書に提供される方法のいくつかは、核酸増幅産物を生成するために、複数の dsDNA断片を増幅することを含む。本明細書に記載される方法は、核酸増幅産物、若しくはその産物を検出及び/又は定量することをさらに含み得る。増幅産物又はその産物は、例えば、本明細書に記載される任意の検出方法又は定量化方法を含む、任意の適切な検出及び/又は定量化方法によって検出及び/又は定量化することができる。検出及び/又は定量化方法の非限定的な例としては、分子ビーコン(例えば、リアルタイム、エンドポイント)、側方流動、蛍光共鳴エネルギー移動(FRET)、蛍光偏光(FP)、表面捕捉、5'から3'へのエキソヌクレアーゼ加水分解プローブ(例えば、TAQMAN)、インターカレーティング色素/結合色素、吸光度法(例えば、比色、濁度)、電気泳動(例えば、ゲル電気泳動、キャピラリー電気泳動)、質量分析、核酸シーケンシング、デジタル増幅、プライマー伸長方法(例えば、iPLEX(商標))、Affymetrixからの分子逆方向プローブ(MIP)技術、制限断片長多型(RFLP分析)、アレルト異的オリゴヌクレオチド(ASO)分析、メチル化特異的PCR(MSPCR)、パイロシーケンシング分析、アシクロプライム分析、リバースドットプロット、GeneChipマイクロアレイ、ダイナミックな対立遺伝子特異的ハイブリダイゼーション(DASH)、ペプチド核酸(PNA)及びロックド核酸(LNA)プローブ、アルファスクリーン、SNPストリーム、遺伝子ピット分析(GBA)、多重ミニシーケンシング、SNAPSHOT GOODアッセイ、マイクロアレイmini-seq、アレイされるプライマー伸長(APEX)、マイクロアレイプライマー伸長、Tagアレイ、コードされたマイクロスフェア、鑄型指向性組込み(TDI)、比色分析オリゴヌクレオチドライゲーションアッセイ(OLA)、配列コード化OLA、マイクロアレイライゲーション、リガーゼ連鎖反応、パドロックプローブ、インベーターアッセイ、少なくとも1つのプローブを用いるハイブリダイゼーション、少なくとも1つの蛍光標識されたプローブを用いるハイブリダイゼーション、クローニング及びシーケンシング、ハイブリダイゼーションプローブ及び定量的リアルタイムポリメラーゼ連鎖反応(QRT-PCR)の使用、ナノ細孔シーケンシング、チップ及びそれらの組み合わせが挙げられる。核酸増幅産物の検出は、リアルタイム検出方法(すなわち、生成物が増幅プロセス中に検出され及び/又は連続的にモニタリングされる)の使用、エンドポイント検出方法(すなわち、生成物が増幅プロセスを完了するか又は停止した後に検出される)の使用、又はその両方を含むことができる。核酸検出方法はまた、標的配列に直接取り込まれたか、又は標的に相補的な配列を含有するプローブに取り込まれた標識ヌクレオチドの使用を採用することができる。このような標識は、本質的に放射性及び/又は蛍光性であり得、本明細書において検討される任意の方

20

30

40

50

法で解像することができる。一部の実施形態では、核酸増幅産物の定量化は、以下に記載される1種以上の検出方法を用いて達成される。検出方法は、シグナル強度の測定、及び/又は核酸増幅産物の定量化のための標準曲線及び/又はルックアップ表の生成(又は参照)とともに使用することができる。

【0106】

核酸増幅産物の検出は、分子ビーコン技術の使用を含むことができる。分子ビーコンという用語は、一般的に、検出可能な分子を指し、分子の検出可能な特性は、ある特定の条件下で検出可能であり、それによって、分子が特異的であり情報的なシグナルとして機能することを可能にする。検出可能な特性の非限定的な例には、光学特性(例えば、蛍光)、電気特性、磁気特性、化学特性、及び公知のサイズの開口を通る時間又は速度が挙げられる。核酸分子を検出するための分子ビーコンは、例えば、一端にフルオロフォアを含み、反対端に消光色素を含有するヘアピン型オリゴヌクレオチドであり得る。ヘアピンのループは、標的配列に相補的なプローブ配列を含有し得、ステムは、プローブ配列のいずれか側に位置する相補的なアーム配列のアニリングによって形成される。フルオロフォア及び消光分子は、各アームの両端で共有結合的に連結され得る。オリゴヌクレオチドがその相補的な標的にハイブリダイズするのを妨げる条件下、又は分子ビーコンが溶液中で遊離である場合、蛍光分子及び消光分子は互いに近接しており、FRETを妨げる。分子ビーコンが標的分子(例えば、核酸増幅産物)に遭遇した場合、ハイブリダイゼーションが起こり得、ループ構造が安定なより硬い立体構造に変換され、蛍光につながるフルオロフォア及び消光剤分子の分離を引き起こす。プローブの特異性のために、蛍光の生成は、一般的に、意図された増幅された産物の合成に独占的に起因する。一部の実施形態では、分子ビーコンプローブ配列は、標的核酸中の配列と同一又は相補的である増幅産物中の配列とハイブリダイズする。一部の実施形態では、分子ビーコンプローブ配列は、標的核酸中の配列と同一ではないか又は相補的でない増幅産物中の配列とハイブリダイズする(例えば、尾部増幅プライマー又はライゲーションにより増幅産物に付加された配列とハイブリダイズする)。分子ビーコンは、異なる着色フルオロフォア及び異なる標的配列を用いて合成することができ、同じ反応(例えば、多重反応)におけるいくつかの生成物の同時検出を可能にする。定量的増幅プロセスでは、分子ビーコンは増幅の各サイクル後に増幅された標的に特異的に結合することができ、ハイブリダイズしていない分子ビーコンは暗いため、増幅された産物の量を定量的に決定するためにプローブ-標的ハイブリッドを単離する必要はない。得られたシグナルは、増幅された産物の量に比例する。分子ビーコンを用いた検出は、リアルタイムで、又は終点検出方法として行うことができる。

【0107】

核酸増幅産物の検出は、側方流動の使用を含むことができ、側方流動デバイスは、一般的に、毛細管力によって流体が流れる固相流体透過性流路を含むことができる。例示的なデバイスには、限定されないが、ディップスティックアッセイ及び種々の適切な被覆を有する薄層クロマトグラフィープレートが挙げられる。流路上に固定化されたものは、試料のための種々の結合試薬、試料及びシグナル生成システムのための結合パートナーを含む結合パートナー又はコンジュゲートである。検出は、例えば、酵素検出、ナノ粒子検出、比色検出、及び蛍光検出を含むいくつかの方法で達成することができる。

【0108】

核酸増幅産物の検出は、2つの発色団:ドナー及びアクセプタ分子の間のエネルギー移動メカニズムであるFRETの使用を含むことができる。簡単に説明すると、ドナーフルオロフォア分子は、特定の励起波長で励起される。その後、基底状態に戻るときにドナー分子から放出される励起エネルギーは、長距離双極子-双極子相互作用を介してアクセプタ分子に移動され得る。アクセプタ分子の発光強度をモニターすることができ、ドナーとアクセプタの間の距離、ドナー発光スペクトルとアクセプタ吸収スペクトルとの重なり、及びドナー発光双極子モーメントとアクセプタ吸収双極子モーメントの配向の関数である。FRETは、例えば、分子ビーコンについて記載されるようなDNA-DNA相互作用において、分子動力学を定量するために有用であり得る。特定の生成物の生成をモニター

10

20

30

40

50

するために、プローブは、一方の端にドナー分子、及び他方にアクセプタ分子で標識することができる。プローブ-標的ハイブリダイゼーションは、ドナー及びアクセプタの距離又は配向の変化をもたらし、FRET変化が観察される。

【0109】

核酸増幅産物の検出は、一般的に、直線偏光によって励起された場合に蛍光標識された化合物が、その回転速度と反比例する偏光度を有する蛍光を発するという原理に基づくFPの使用を含む。したがって、例えば、蛍光標識を有するトレーサー-核酸コンジュゲートなどの分子を直線偏光で励起した場合、発光された光は、フルオロフォアが、光が吸収され発光される時間の間に回転するのを制限されるため、高度に偏光されたままである。遊離トレーサー化合物（すなわち、核酸に結合していない）が直線偏光によって励起された場合、その回転は、対応するトレーサー-核酸コンジュゲートよりもはるかに速く、分子はよりランダムに配向され、したがって、放出された光は脱分極される。したがって、蛍光偏光は、増幅反応において生成されたトレーサー-核酸コンジュゲートの量を測定するための定量的手段を提供する。 10

核酸増幅産物の検出は、特異的オリゴヌクレオチドを、高感度及び選択的なバイオセンサーを生成する表面に固定化することによって達成することができる表面捕捉の使用を含む。

核酸増幅産物の検出は、5'から3'へのエキソヌクレアーゼ加水分解プローブ（例えば、TAQMAN）の使用を含むことができる。例えば、TAQMANプローブは、定量的増幅法（例えば、定量的PCR）の特異性を増加させることができる加水分解プローブである。TAQMANプローブ原理は、1)相補的な標的配列へのハイブリダイゼーション中に二重標識プローブを切断するためのTaqポリメラーゼの5'から3'へのエキソヌクレアーゼ活性、及び2)フルオロフォアベースの検出に依存する。得られた蛍光シグナルは、増幅の指数関数的段階の間の増幅産物の蓄積を定量的に測定することを可能にする。 20

【0110】

核酸増幅産物の検出は、インターカレーティング色素及び/又は結合色素、例えば、核酸を特異的に染色することができる色素の使用を含む。例えば、インターカレーティング色素は、DNA又はRNAに結合すると増強された蛍光を示す。色素の非限定的な例としては、SYTO（登録商標）82、アクリジンオレンジ、エチジウムブロマイド、ヘキスト色素、PicoGreen（登録商標）、ヨウ化プロピジウム、SYBR（登録商標）I（非対称シアニン色素）、SYBR（登録商標）II、TOTO（チアキソールオレンジ二量体）及びYOYO（オキサゾール黄色二量体）が挙げられる。 30

【0111】

核酸増幅産物の検出は、吸光度法（例えば、比色法、濁度）の使用を含む。核酸の検出及び/又は定量は、例えば、吸光度（例えば、260nmにおけるUV吸光度測定）を濃度に直接変換することによって達成することができる。核酸の直接測定は、測定の経路長及び吸光係数を使用して、吸光度を濃度に関連付けるBeer-Lambert法を使用して濃度に変換することができる。

核酸増幅産物の検出は、電気泳動（例えば、ゲル電気泳動、キャピラリー電気泳動）、質量分析、核酸シーケンシング、デジタル増幅（例えば、デジタルPCR）、又はそれらの任意の組み合わせの使用を含むことができる。 40

【0112】

目的とする遺伝子シグネチャ

複数の標的dsDNAは、目的とする遺伝子シグネチャ（例えば、バイオマーカーシグネチャ）を含み得る。目的とする遺伝子シグネチャは、目的とする1種以上の突然変異（例えば、バイオマーカー）を含み得る。目的とする1種以上の突然変異は、点突然変異、逆位、欠失、挿入、転座、重複、コピー数変動、又はそれらの組み合わせを含み得る。目的とする1種以上の突然変異は、ヌクレオチド置換、欠失、挿入、又はそれらの組み合わせを含み得る。目的とする遺伝子シグネチャは、標的dsDNAが由来する生物の抗生物質耐性又は抗生物質感受性を示すことができる。目的とする遺伝子シグネチャは、標的d 50

s DNA が由来する生物のがん状態を示すことができる。目的とする遺伝子シグネチャは、標的 ds DNA が由来する生物の遺伝子疾患の状態を示すことができる。遺伝子疾患は単一遺伝子障害であり得る。遺伝子疾患は、嚢胞性線維症、ハンチントン病、鎌状赤血球貧血、血友病、デュシェンヌ型筋ジストロフィー、サラセミア、脆弱 X 症候群、家族性高コレステロール血症、多嚢胞性腎疾患、神経線維腫症 I 型、遺伝性球状赤血球症、マルファン症候群、テイ-サックス病、フェニルケトン尿症、ムコ多糖症、リソソーム酸性リパーゼ欠損症、グリコーゲン貯蔵疾患、ガラクトース血症、又はヘモクロマトーシスであり得る。目的とする遺伝子シグネチャ（例えば、バイオマーカーシグネチャ）は、本明細書に提供される方法及び組成物を用いて検出することができる。診断的評価は、本明細書に提供される方法及び組成物を用いて行うことができる。

10

【0113】

診断評価は、本明細書に記載されるように、単独で又は他の評価若しくは因子と組み合わせ、バイオマーカーシグネチャ（例えば、目的とする遺伝子シグネチャ）に基づいて行われる。本明細書には、疾患若しくは状態を発症するリスクを評価し、上記疾患を予測し、上記疾患若しくは状態を診断し、上記疾患若しくは状態の進行又は退行をモニタリングし、或いは治療の有効性を評価し、又はバイオマーカーシグネチャ（例えば、目的とする遺伝子シグネチャ）に基づいて、上記疾患若しくは状態を改善又は治療することができる化合物を同定するための組成物及び方法が提供される。

【0114】

疾患及び状態

本明細書に提供される方法は、種々の疾患又は状態と関連するバイオマーカーシグネチャ（例えば、目的とする遺伝子シグネチャ）に基づいて、種々の疾患又は状態に適用することができる。開示された組成物及び方法の対象となる目的とする遺伝子シグネチャを有する例示的な疾患又は状態としては、心血管疾患若しくは状態、腎臓と関連する疾患若しくは状態、出生前若しくは妊娠に関連した疾患若しくは状態、神経学的若しくは神経精神疾患若しくは状態、自己免疫若しくは免疫に関連した疾患若しくは状態、がん、感染性疾患若しくは状態、小児疾患、障害若しくは状態、ミトコンドリア障害、呼吸-消化管疾患若しくは状態、生殖系疾患若しくは状態、眼疾患若しくは状態、筋骨格疾患若しくは状態、又は皮膚疾患若しくは状態が挙げられる。

20

【0115】

試料

試料は、真核生物 DNA、細菌 DNA、ウイルス DNA、真菌 DNA、原生動物 DNA、又はそれらの組み合わせを含むことができる。複数の標的 ds DNA は、ゲノム DNA、ミトコンドリア DNA、プラスミド DNA、又はそれらの組み合わせを含むことができる。試料は、生物学的試料、臨床試料、環境試料、若しくはそれらの組み合わせであり得るか又はそれに由来し得る。複数の標的 ds DNA は、少なくとも 2 つの異なる生物由来の DNA を含むことができる。複数の標的 ds DNA は、少なくとも 2 つの異なる遺伝子由来の DNA を含むことができる。本方法は、逆転写酵素を用いて複数の標的 RNA から複数の標的 ds DNA を生成することを含むことができる。複数の標的 ds DNA は、逆転写酵素を用いて標的 RNA から生成された標的 ds DNA を含むことができる。試料核酸は、真核生物 DNA、細菌 DNA、ウイルス DNA、真菌 DNA、原生動物 DNA、又はそれらの組み合わせを含むことができる。標的 ds DNA は、ゲノム DNA、ミトコンドリア DNA、プラスミド DNA、又はそれらの組み合わせであり得る。試料核酸は、生物学的試料、臨床試料、環境試料、又はそれらの組み合わせ由来であり得る。生物学的試料は、便、痰、末梢血、血漿、血清、リンパ節、呼吸組織、滲出液、体液、又はそれらの組み合わせを含むことができる。

30

40

【0116】

本明細書に記載される方法において利用される核酸は、任意の適切な生物学的試料から得ることができ、しばしば対象から得られる試料から単離される。対象は、限定されないが、ヒト、非ヒト動物、植物、細菌、真菌、ウイルス、又は原生生物を含む、任意の生き

50

ている生物又は生きていない生物であり得る。任意のヒト又は非ヒト動物を選択することができ、限定されないが、哺乳動物、爬虫類、鳥類、両生類、魚類、有蹄動物、反芻動物、ウシ (bovine) (例えば、ウシ (cattle))、ウマ (equine) (例えば、ウマ (horse))、ヤギ (caprine) 及びヒツジ (ovine) (例えば、ヒツジ (sheep))、ヤギ (goat)、ブタ (swine) (例えば、ブタ (pig))、ラクダ (camelid) (例えば、ラクダ (camel))、ラマ、アルパカ)、サル、類人猿 (例えば、ゴリラ、チンパンジー)、クマ (ursid) (例えば、クマ (bear))、家禽、イヌ、ネコ、マウス、ラット、魚、イルカ、クジラ及びサメが含まれる。対象は、男性又は女性であり得、対象は、任意の年齢 (例えば、胚、胎児、乳児、小児、成人) であり得る。

10

【0117】

試料又は試験試料は、対象若しくはその一部から単離されるか又は得られる任意の検体であり得る。検体の非限定的な例としては、対象由来の体液又は組織が挙げられ、限定されないが、血液又は血液生成物 (例えば、血清、血漿など)、臍帯血、骨髄、絨毛、羊水、脳脊髄液、脊髄液、洗浄液 (例えば、気管支肺胞、胃、腹腔、管、耳、関節鏡)、生検試料、セロセンチス試料、細胞 (例えば、血液細胞) 又はその一部 (例えば、ミトコンドリア、核、抽出物など)、女性生殖器の洗浄液、尿、便、痰、唾液、鼻粘膜、前立腺液、洗浄液、精液、リンパ液、胆汁、涙、汗、母乳、乳房液、硬組織 (例えば、肝臓、脾臓、腎臓、肺、又は卵巣) など、又はそれらの組み合わせが含まれる。血液という用語は、従来から定義されているように、全血、血液生成物又は血液の任意の画分、例えば、血清、血漿、パフィーコートなどを包含する。血液血漿とは、抗凝固剤で処理された血液の遠心分離から生じる全血の画分を指す。血液血清とは、血液試料が凝固した後に残った流体の水様部分を指す。体液又は組織の試料は、しばしば、病院又は診療所が一般的に守る標準的なプロトコールに従って採取される。血液については、適切な量の末梢血 (例えば、3 ~ 40 ミリリットル) がしばしば採取され、調製前又は調製後に標準的な手法に従って保存され得る。

20

【0118】

試料又は試験試料は、芽胞、ウイルス、細胞、原核生物若しくは真核生物由来の核酸、又は任意の遊離核酸を含有する試料を含み得る。例えば、本明細書に記載される方法は、(例えば、溶解の必要なしに) 芽胞の外側で核酸を検出するために使用され得る。試料は、上述の対象由来など、標的配列を含有することが疑われる任意の物質から単離することができる。一部の実施形態では、標的配列は、空気、植物、土壌、又は生物学的生物を含有することが疑われる他の物質中に存在する。

30

【0119】

核酸は、当該技術分野において公知である方法によって、1種以上の供給源から誘導 (例えば、単離、抽出、精製) することができる。任意の適切な方法は、生物学的試料から核酸を単離、抽出及び/又は精製するために使用することができ、その非限定的な例には、当該技術分野におけるDNA調製の方法、及び種々の市販される試薬又はキット、例えば、Qiagen's QIAamp 循環核酸キット、QIAamp DNA ミニキット又はQIAamp DNA 血液ミニキット (Qiagen、Hilden、Germany)、Genomic Prep (商標) 血液DNA単離キット (Promega、Madison、Wis)、GFX (商標) ゲノム血液DNA精製キット (Amersham、Piscataway、N.J.) など、或いはそれらの組み合わせが挙げられる。

40

【0120】

一部の実施形態では、細胞溶解手法が行われる。細胞溶解は、本明細書に提供される反応の開始前に行うことができる。細胞溶解手法及び試薬は当該技術分野において公知であり、一般的に、化学的溶解法 (例えば、界面活性剤、低張液、酵素的な手法など、又はそれらの組み合わせ)、物理的溶解法 (例えば、加圧型細胞破壊、超音波処理など)、又は電解的溶解法によって行うことができる。適切な溶解手法を利用することができる。例えば、化学的方法は、一般的に、細胞を破壊し、細胞から核酸を抽出し、続いてカオトロピッ

50

ク塩で処理するために溶解剤を採用する。一部の実施形態では、細胞溶解は、界面活性剤（例えば、イオン性、非イオン性、陰イオン性、双性イオン性）の使用を含む。一部の実施形態では、細胞溶解は、イオン性界面活性剤（例えば、ドデシル硫酸ナトリウム（SDS）、ラウリル硫酸ナトリウム（SLS）、デオキシコール酸塩、コール酸塩、サルコシル）の使用を含む。凍結/解凍、続く粉碎、細胞圧搾の使用などの物理的方法も有用であり得る。高塩分溶解手法もまた用いることができる。例えば、アルカリ溶解手法を利用することができる。後者の手法は、従来、フェノール-クロロホルム溶液の使用を組み込んでおり、3つの溶液を伴う代替的なフェノール-クロロホルム不含手法を利用することができる。後者の手法では、1つの溶液は、15 mM Tris、pH 8.0；10 mM EDTA及び100 µg/ml RNase Aを含有することができ；第2の溶液は、0.2 N NaOH及び1% SDSを含有することができ；第3の溶液は、例えば、3 M KOAc、pH 5.5を含有することができる。一部の実施形態では、細胞溶解緩衝液は、本明細書に記載される方法及び構成成分とともに使用される。

【0121】

核酸は、核酸を含有する試料を処理することなく、本明細書に記載される方法を行うために提供され得る。例えば、一部の実施形態では、核酸は、以前の核酸精製なしに、本明細書に記載される増幅方法を行うために提供される。一部の実施形態では、標的配列は、（例えば、任意の核酸抽出、単離、精製及び/又は部分的精製ステップを行うことなく）試料から直接増幅される。一部の実施形態では、核酸は、核酸を含有する試料の処理後に、本明細書に記載される方法を行うために提供される。例えば、核酸は、試料から抽出、単離、精製、又は部分的に精製することができる。用語「単離された」とは、一般的に、その元の環境（例えば、それが天然に存在する場合の自然環境、又は外因性に発現される場合の宿主細胞）から取り出された核酸を指し、したがって、その元の環境からのヒトの介入（例えば、「ヒトの手による」）によって変更される。用語「単離された核酸」とは、対象（例えば、ヒト対象）から取り出された核酸を指すことができる。単離された核酸は、供給源試料中に存在する分量よりも少ない非核酸成分（例えば、タンパク質、脂質、糖質）を提供することができる。単離された核酸を含む組成物は、非核酸成分を約50%～99%を超えて含まない場合がある。単離された核酸を含む組成物は、非核酸成分を約90%、91%、92%、93%、94%、95%、96%、97%、98%、99%又は99%を超えて含まない場合がある。用語「精製された」とは、一般的に、核酸を精製手法に供する前に存在する非核酸成分の量よりも少ない非核酸成分（例えば、タンパク質、脂質、糖質）を含有する核酸を指す。精製された核酸を含む組成物は、他の非核酸成分を約80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%又は99%を超えて含まない場合がある。

【0122】

核酸は、核酸を修飾することなく、本明細書に記載される方法を行うために提供され得る。修飾には、例えば、変性、消化、ニックング、巻き戻し、不均一配列の組込み及び/又はライゲーション、エピジェネティック修飾の付加、標識（例えば、³²P、³³P、¹²⁵I若しくは³⁵Sなどの放射性標識；アルカリホスファターゼなどの酵素標識；フルオレセインイソチオシアネート（FITC）などの蛍光標識；又はビオチン、アビジン、ジゴキシゲニン、抗原、ハプテン、蛍光色素などの他の標識）などの付加が含まれ得る。したがって、一部の実施形態では、未修飾核酸は増幅される。

【0123】

試料中の標的核酸配列（一本鎖又は二本鎖DNA及び/又はRNA）を検出するための本開示の方法は、高程度の感度で標的核酸配列（例えば、DNA又はRNA）を検出することができる。一部の実施形態では、本開示の方法を使用して、複数のRNA/DNA（標的RNA/DNA及び複数の非標的RNA/DNAを含む）を含む試料中に存在する標的RNA/DNAを検出することができ、標的RNA/DNAは、1コピー以上/10⁷個の非標的RNA/DNA（例えば、1コピー以上/10⁶個の非標的RNA/DNA、

1コピー以上/10⁵個の非標的RNA/DNA、1コピー以上/10⁴個の非標的RNA/DNA、1コピー以上/10³個の非標的RNA/DNA、1コピー以上/10²個の非標的RNA/DNA、1コピー以上/50個の非標的RNA/DNA、1コピー以上/20個の非標的RNA/DNA、1コピー以上/10個の非標的RNA/DNA、又は1コピー以上/5個の非標的RNA/DNA)で存在する。一部の実施形態では、本開示の方法を使用して、複数のRNA/DNA(標的RNA/DNA及び複数の非標的RNA/DNAを含む)を含む試料中に存在する標的RNA/DNAを検出することができ、標的RNA/DNAは、1コピー以上/10¹⁸個の非標的RNA/DNA(例えば、1コピー以上/10¹⁵個の非標的RNA/DNA、1コピー以上/10¹²個の非標的RNA/DNA、1コピー以上/10⁹個の非標的RNA/DNA、1コピー以上/10⁶個の非標的RNA/DNA、1コピー以上/10⁵個の非標的RNA/DNA、1コピー以上/10⁴個の非標的RNA/DNA、1コピー以上/10³個の非標的RNA/DNA、1コピー以上/10²個の非標的RNA/DNA、1コピー以上/50個の非標的RNA/DNA、1コピー以上/20個の非標的RNA/DNA、1コピー以上/10個の非標的RNA/DNA、又は1コピー以上/5個の非標的RNA/DNA)で存在する。本明細書で使用される場合、用語「RNA/DNA」及び「RNAs/DNAs」は、それらの通常の意味を与えられるものとし、DNA、若しくはRNA、又はDNAとRNAの組み合わせをも指すものとする。

10

【0124】

一部の実施形態では、本開示の方法は、試料中に存在する標的RNA/DNAを検出することができ、標的RNA/DNAは、1コピー/10⁷個の非標的RNA/DNA~1コピー/10個の非標的RNA/DNA(例えば、1コピー/10⁷個の非標的RNA/DNA~1コピー/10²個の非標的RNA/DNA、1コピー/10⁷個の非標的RNA/DNA~1コピー/10³個の非標的RNA/DNA、1コピー/10⁷個の非標的RNA/DNA~1コピー/10⁴個の非標的RNA/DNA、1コピー/10⁷個の非標的RNA/DNA~1コピー/10⁵個の非標的RNA/DNA、1コピー/10⁷個の非標的RNA/DNA~1コピー/10⁶個の非標的RNA/DNA、1コピー/10⁷個の非標的RNA/DNA~1コピー/10個の非標的RNA/DNA、1コピー/10⁶個の非標的RNA/DNA~1コピー/10²個の非標的RNA/DNA、1コピー/10⁶個の非標的RNA/DNA~1コピー/10³個の非標的RNA/DNA、1コピー/10⁶個の非標的RNA/DNA~1コピー/10⁴個の非標的RNA/DNA、1コピー/10⁶個の非標的RNA/DNA~1コピー/10⁵個の非標的RNA/DNA、1コピー/10⁵個の非標的RNA/DNA~1コピー/10個の非標的RNA/DNA、1コピー/10⁵個の非標的RNA/DNA~1コピー/10²個の非標的RNA/DNA、1コピー/10⁵個の非標的RNA/DNA~1コピー/10³個の非標的RNA/DNA、又は1コピー/10⁵個の非標的RNA/DNA~1コピー/10⁴個の非標的RNA/DNA)で存在する。

20

30

【0125】

一部の実施形態では、本開示の方法は、試料中に存在する標的RNA/DNAを検出することができ、標的RNA/DNAは、1コピー/10¹⁸個の非標的RNA/DNA~1コピー/10個の非標的RNA/DNA(例えば、1コピー/10¹⁸個の非標的RNA/DNA~1コピー/10²個の非標的RNA/DNA、1コピー/10¹⁵個の非標的RNA/DNA~1コピー/10²個の非標的RNA/DNA、1コピー/10¹²個の非標的RNA/DNA~1コピー/10²個の非標的RNA/DNA、1コピー/10⁹個の非標的RNA/DNA~1コピー/10²個の非標的RNA/DNA、1コピー/10⁷個の非標的RNA/DNA~1コピー/10²個の非標的RNA/DNA、1コピー/10⁷個の非標的RNA/DNA~1コピー/10³個の非標的RNA/DNA、1コピー/10⁷個の非標的RNA/DNA~1コピー/10⁴個の非標的RNA/DNA、1コピー/10⁷個の非標的RNA/DNA~1コピー/10⁵個の非標的RNA/DNA、1コピー/10⁷個の非標的RNA/DNA~1コピー/10⁶個の非標的RNA/DNA

40

50

RNA、1コピー/10⁶個の非標的RNA/DNA～1コピー/10個の非標的RNA/DNA、1コピー/10⁶個の非標的RNA/DNA～1コピー/10²個の非標的RNA/DNA、1コピー/10⁶個の非標的RNA/DNA～1コピー/10³個の非標的RNA/DNA、1コピー/10⁶個の非標的RNA/DNA～1コピー/10⁴個の非標的RNA/DNA、1コピー/10⁶個の非標的RNA/DNA～1コピー/10⁵個の非標的RNA/DNA、1コピー/10⁵個の非標的RNA/DNA～1コピー/10個の非標的RNA/DNA、1コピー/10⁵個の非標的RNA/DNA～1コピー/10²個の非標的RNA/DNA、1コピー/10⁵個の非標的RNA/DNA～1コピー/10³個の非標的RNA/DNA、又は1コピー/10⁵個の非標的RNA/DNA～1コピー/10⁴個の非標的RNA/DNA)で存在する。

10

【0126】

一部の実施形態では、本開示の方法は、試料中に存在する標的RNA/DNAを検出することができ、標的RNA/DNAは、1コピー/10⁷個の非標的RNA/DNA～1コピー/100個の非標的RNA/DNA(例えば、1コピー/10⁷個の非標的RNA/DNA～1コピー/10²個の非標的RNA/DNA、1コピー/10⁷個の非標的RNA/DNA～1コピー/10³個の非標的RNA/DNA、1コピー/10⁷個の非標的RNA/DNA～1コピー/10⁴個の非標的RNA/DNA、1コピー/10⁷個の非標的RNA/DNA～1コピー/10⁵個の非標的RNA/DNA、1コピー/10⁷個の非標的RNA/DNA～1コピー/10⁶個の非標的RNA/DNA、1コピー/10⁶個の非標的RNA/DNA～1コピー/100個の非標的RNA/DNA、1コピー/10⁶個の非標的RNA/DNA～1コピー/10²個の非標的RNA/DNA、1コピー/10⁶個の非標的RNA/DNA～1コピー/10³個の非標的RNA/DNA、1コピー/10⁶個の非標的RNA/DNA～1コピー/10⁴個の非標的RNA/DNA、1コピー/10⁶個の非標的RNA/DNA～1コピー/10⁵個の非標的RNA/DNA、1コピー/10⁵個の非標的RNA/DNA～1コピー/100個の非標的RNA/DNA、1コピー/10⁵個の非標的RNA/DNA～1コピー/10²個の非標的RNA/DNA、1コピー/10⁵個の非標的RNA/DNA～1コピー/10³個の非標的RNA/DNA、又は1コピー/10⁵個の非標的RNA/DNA～1コピー/10⁴個の非標的RNA/DNA)で存在する。

20

【0127】

一部の実施形態では、試料中の標的RNA/DNAを検出する主題の方法についての検出閾値は、10nM以下である。用語「検出閾値」は、検出が起こるために試料中に存在しなければならない最小量の標的RNA/DNAを記載するために本明細書で使用される。したがって、例示的な例として、検出閾値が10nMである場合、標的RNA/DNAが10nM以上の濃度で試料中に存在する場合には、シグナルを検出することができる。一部の実施形態では、本開示の方法は、5nM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、1nM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、0.5nM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、0.1nM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、0.05nM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、0.01nM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、0.005nM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、0.001nM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、0.0005nM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、0.0001nM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、0.00005nM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、0.00001nM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、10pM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、1pM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、500fM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、250fM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、100fM以下の検

30

40

50

出閾値を有する。一部の実施形態では、本開示の方法は、50 fM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、500 aM（アトモル濃度）以下の検出閾値を有する。一部の実施形態では、本開示の方法は、250 aM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、100 aM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、50 aM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、10 aM以下の検出閾値を有する。一部の実施形態では、本開示の方法は、1 aM以下の検出閾値を有する。

【0128】

一部の実施形態では、検出閾値（主題の方法において標的RNA/DNAを検出するため）は、500 fM～1 nM（例えば、500 fM～500 pM、500 fM～200 pM、500 fM～100 pM、500 fM～10 pM、500 fM～1 pM、800 fM～1 nM、800 fM～500 pM、800 fM～200 pM、800 fM～100 pM、800 fM～10 pM、800 fM～1 pM、1 pM～1 nM、1 pM～500 pM、1 pM～200 pM、1 pM～100 pM、又は1 pM～10 pM）の範囲である（濃度は標的RNA/DNAが検出され得る標的RNA/DNAの閾値濃度を指す）。一部の実施形態では、本開示の方法は、800 fM～100 pMの範囲の検出閾値を有する。一部の実施形態では、本開示の方法は、1 pM～10 pMの範囲の検出閾値を有する。一部の実施形態では、本開示の方法は、10 fM～500 fM、例えば、10 fM～50 fM、50 fM～100 fM、100 fM～250 fM、又は250 fM～500 fMの範囲の検出閾値を有する。

【0129】

一部の実施形態では、標的RNA/DNAが試料中で検出され得る最低濃度は、500 fM～1 nM（例えば、500 fM～500 pM、500 fM～200 pM、500 fM～100 pM、500 fM～10 pM、500 fM～1 pM、800 fM～1 nM、800 fM～500 pM、800 fM～200 pM、800 fM～100 pM、800 fM～10 pM、800 fM～1 pM、1 pM～1 nM、1 pM～500 pM、1 pM～200 pM、1 pM～100 pM、又は1 pM～10 pM）の範囲である。一部の実施形態では、標的RNA/DNAが試料中で検出され得る最低濃度は、800 fM～100 pMの範囲である。一部の実施形態では、標的RNA/DNAが試料中で検出され得る最低濃度は、1 pM～10 pMの範囲である。

【0130】

一部の実施形態では、検出閾値（主題の方法において標的RNA/DNAを検出するため）は、1 aM～1 nM（例えば、1 aM～500 pM、1 aM～200 pM、1 aM～100 pM、1 aM～10 pM、1 aM～1 pM、100 aM～1 nM、100 aM～500 pM、100 aM～200 pM、100 aM～100 pM、100 aM～10 pM、100 aM～1 pM、250 aM～1 nM、250 aM～500 pM、250 aM～200 pM、250 aM～100 pM、250 aM～10 pM、250 aM～1 pM、500 aM～1 nM、500 aM～500 pM、500 aM～200 pM、500 aM～100 pM、500 aM～10 pM、500 aM～1 pM、750 aM～1 nM、750 aM～500 pM、750 aM～200 pM、750 aM～100 pM、750 aM～10 pM、750 aM～1 pM、1 fM～1 nM、1 fM～500 pM、1 fM～200 pM、1 fM～100 pM、1 fM～10 pM、1 fM～1 pM、500 fM～500 pM、500 fM～200 pM、500 fM～100 pM、500 fM～10 pM、500 fM～1 pM、800 fM～1 nM、800 fM～500 pM、800 fM～200 pM、800 fM～100 pM、800 fM～10 pM、800 fM～1 pM、1 pM～1 nM、1 pM～500 pM、1 pM～200 pM、1 pM～100 pM、又は1 pM～10 pM）の範囲である（濃度は標的RNA/DNAが検出され得る標的RNA/DNAの閾値濃度を指す）。一部の実施形態では、本開示の方法は、1 aM～800 aMの範囲の検出閾値を有する。一部の実施形態では、本開示の方法は、50 aM～1 pMの範囲の検出閾値を有する。一部の実施形態では、本開示の方法は、50 aM～500 fMの範囲の検出閾値を

有する。

【0131】

一部の実施形態では、標的RNA/DNAが試料中で検出され得る最低濃度は、1 aM ~ 1 nM (例えば、1 aM ~ 500 pM、1 aM ~ 200 pM、1 aM ~ 100 pM、1 aM ~ 10 pM、1 aM ~ 1 pM、100 aM ~ 1 nM、100 aM ~ 500 pM、100 aM ~ 200 pM、100 aM ~ 100 pM、100 aM ~ 10 pM、100 aM ~ 1 pM、250 aM ~ 1 nM、250 aM ~ 500 pM、250 aM ~ 200 pM、250 aM ~ 100 pM、250 aM ~ 10 pM、250 aM ~ 1 pM、500 aM ~ 1 nM、500 aM ~ 500 pM、500 aM ~ 200 pM、500 aM ~ 100 pM、500 aM ~ 10 pM、500 aM ~ 1 pM、750 aM ~ 1 nM、750 aM ~ 500 pM、750 aM ~ 200 pM、750 aM ~ 100 pM、750 aM ~ 10 pM、750 aM ~ 1 pM、1 fM ~ 1 nM、1 fM ~ 500 pM、1 fM ~ 200 pM、1 fM ~ 100 pM、1 fM ~ 10 pM、1 fM ~ 1 pM、500 fM ~ 500 pM、500 fM ~ 200 pM、500 fM ~ 100 pM、500 fM ~ 10 pM、500 fM ~ 1 pM、800 fM ~ 1 nM、800 fM ~ 500 pM、800 fM ~ 200 pM、800 fM ~ 100 pM、800 fM ~ 10 pM、800 fM ~ 1 pM、1 pM ~ 1 nM、1 pM ~ 500 pM、1 pM ~ 200 pM、1 pM ~ 100 pM、又は1 pM ~ 10 pM)の範囲である。一部の実施形態では、標的RNA/DNAが試料中で検出され得る最低濃度は、1 aM ~ 500 pMの範囲である。一部の実施形態では、標的RNA/DNAが試料中で検出され得る最低濃度は、100 aM ~ 500 pMの範囲である。

一部の実施形態では、開示される組成物又は方法は、アトモル濃度(aM)の検出感度を示す。一部の実施形態では、開示される組成物又は方法は、フェムトモル濃度(fM)の検出感度を示す。一部の実施形態では、開示される組成物又は方法は、ピコモル濃度(pM)の検出感度を示す。一部の実施形態では、開示される組成物又は方法は、ナノモル濃度(nM)の検出感度を示す。

【0132】

開示される試料は、試料核酸(例えば、複数の試料核酸)を含む。用語「複数」とは、本明細書において、2種以上を意味するために使用される。したがって、一部の実施形態では、試料は、2種以上(例えば、3種以上、5種以上、10種以上、20種以上、50種以上、100種以上、500種以上、1,000種以上、又は5,000種以上)の試料核酸(例えば、RNA)を含む。開示される方法は、試料中(例えば、RNAなどの核酸の複合混合物中)に存在する標的核酸を検出するための非常に感度の高い方法として用いることができる。一部の実施形態では、試料は、配列が互いに異なる5種以上のDNA(例えば、10種以上、20種以上、50種以上、100種以上、500種以上、1,000種以上、又は5,000種以上のRNA)を含む。一部の実施形態では、試料は、10種以上、20種以上、50種以上、100種以上、500種以上、 10^3 種以上、 5×10^3 種以上、 10^4 種以上、 5×10^4 種以上、 10^5 種以上、 5×10^5 種以上、 10^6 種以上、 5×10^6 種以上、又は 10^7 種以上のDNAを含む。一部の実施形態では、試料は、10 ~ 20種、20 ~ 50種、50 ~ 100種、100 ~ 500種、500 ~ 10^3 種、 10^3 ~ 5×10^3 種、 5×10^3 ~ 10^4 種、 10^4 ~ 5×10^4 種、 5×10^4 ~ 10^5 種、 10^5 ~ 5×10^5 種、 5×10^5 ~ 10^6 種、 10^6 ~ 5×10^6 種、若しくは 5×10^6 ~ 10^7 種、又は 10^7 種以上のDNAを含む。一部の実施形態では、試料は、5 ~ 10^7 種のRNA(例えば、配列が互いに異なる)(例えば、5 ~ 10^6 種、5 ~ 10^5 種、5 ~ 50,000種、5 ~ 30,000種、10 ~ 10^6 種、10 ~ 10^5 種、10 ~ 50,000種、10 ~ 30,000種、20 ~ 10^6 種、20 ~ 10^5 種、20 ~ 50,000種、又は20 ~ 30,000種のDNA)を含む。一部の実施形態では、試料は、配列が互いに異なる20種以上のRNAを含む。一部の実施形態では、試料は、細胞溶解物(例えば、真核細胞溶解物、哺乳動物細胞溶解物、ヒト細胞溶解物、原核生物細胞溶解物、植物細胞溶解物など)由来のRNAを含む。例えば、一部の実施形態では、試料は、真核細胞などの細胞、例えば、ヒト細胞などの哺乳動物細胞由来のDNAを含む。

【0133】

用語「試料」は、本明細書で使用される場合、その通常の意味を与えられ、（例えば、標的DNA及び／又は標的RNAがRNA及び／又はDNAの集団中に存在するかどうかを決定するために）RNA及び／又はDNAを含む任意の試料を含まなければならない。試料は、任意の供給源に由来することができ、例えば、試料は、精製されたDNA及び／又はRNAの合成の組み合わせであり得；試料は、細胞溶解物、DNA/RNA濃縮細胞溶解物、又は細胞溶解物から単離及び／若しくは精製されたDNA/RNAであり得る。試料は、（例えば、診断の目的のために）患者由来であり得る。試料は、透過処理された細胞由来であり得る。試料は、架橋された細胞由来であり得る。試料は、組織切片中であり得る。試料は、架橋し、続いて、脱脂及び調整して均一な屈折率を作製することによって調製された組織由来であり得る。

10

【0134】

適切な試料には、限定されないが、唾液、血液、血清、血漿、尿、吸引物、及び生検試料が含まれる。試料は、患者由来であり得、生物学的起源の血液及び他の液体試料、固体組織試料、例えば、生検検体又はそれらに由来する組織培養物若しくは細胞、及びそれらの子孫を包含する。この定義はまた、例えば、試薬による処理、洗浄、又はがん細胞などのある特定の細胞集団についての濃縮などによって、それらの調達後に任意の方法で操作された試料を含む。この定義はまた、特定タイプの分子、例えば、RNAについて濃縮された試料を含む。用語「試料」は、生物学的試料、例えば、臨床試料、例えば、血液、血漿、血清、吸引物、脳脊髄液（CSF）を包含し、また、外科的切除によって得られた組織、生検によって得られた組織、培養細胞、細胞上清、細胞溶解物、組織試料、臓器、骨髓なども含む。「生物学的試料」は、それに由来する生物学的流体（例えば、がん性細胞、感染細胞など）、例えば、このような細胞から得られるRNAを含む試料（例えば、RNAを含む細胞溶解物又は他の細胞抽出物）を含む。

20

【0135】

一部の実施形態では、試料の供給源は、疾患細胞、体液、組織、又は臓器である（又はその疑いがある）。一部の実施形態では、試料の供給源は、正常な（非疾患）細胞、体液、組織、又は臓器である。一部の実施形態では、試料の供給源は、病原体に感染した細胞、組織、又は臓器である（又はその疑いがある）。例えば、試料の供給源は、感染され得るか又は感染され得ない個体であり得、試料は、個体から採取される任意の生物学的試料（例えば、血液、唾液、生検、血漿、血清、気管支肺胞洗浄、痰、糞便試料、脳脊髄液、細針吸引物、スワブ試料（例えば、口腔スワブ、頸部スワブ、鼻スワブ）、間質液、滑液、鼻汁、涙、パフィーコート、粘膜試料、上皮細胞試料（例えば、上皮細胞剥離）など）であり得る。一部の実施形態では、試料は、無細胞液体試料である。一部の実施形態では、試料は、細胞を含むことができる液体試料である。病原体には、ウイルス、真菌、蠕虫、原虫、マラリア寄生虫、プラズモジウム（*Plasmodium*）寄生虫、トキソプラズマ（*Toxoplasma*）寄生虫、住血吸虫（*Schistosoma*）寄生虫などが含まれる。「蠕虫」には、回虫、心臓線虫、及び植物性線虫（ネマトーダ）、吸虫（テマトーダ）、アカントセファラ、並びに糸虫（セストーダ）が含まれる。原虫感染症には、ジアルジア属種（*Giardia* spp.）、トリコモナス属種（*Trichomonas* spp.）、アフリカトリパノソーマ症、アメーバ赤痢、バベシア症、バランチジア赤痢、シャーガス病、コクシジウム症、マラリア及びトキソプラズマ症由来の感染が含まれる。寄生/原虫病原体などの病原体の例には、限定されないが、熱帯熱マラリア原虫（*Plasmodium falciparum*）、三日熱マラリア原虫（*Plasmodium vivax*）、トリパノソーマ・クルーズ（*Trypanosoma cruzi*）及びトキソプラズマ原虫（*Toxoplasma gondii*）が挙げられる。真菌病原体には、限定されないが、クリプトコッカス・ネオフォルマンズ（*Cryptococcus neoformans*）、ヒストプラズマ・カプスラツム（*Histoplasma capsulatum*）、コクシジオイデス・イミティス（*Coccidioides immitis*）、プラストミセス・デルマチチジス（*Blastom*

30

40

50

yces dermatitidis）、クラミジア・トラコマチス (*Chlamydia trachomatis*)、及びカンジダ・アルビカンス (*Candida albicans*) が含まれる。病原性ウイルスには、例えば、免疫不全ウイルス (例えば、HIV)；インフルエンザウイルス；デング；西ナイルウイルス；ヘルペスウイルス；黄熱ウイルス；C型肝炎ウイルス；A型肝炎ウイルス；B型肝炎ウイルス；パピローマウイルスなどが含まれる。病原性ウイルスには、DNAウイルス、例えば、パポバウイルス (例えば、ヒトパピローマウイルス (HPV)、ポリオーマウイルス)；ヘパドナウイルス (例えば、B型肝炎ウイルス (HBV))；ヘルペスウイルス (例えば、単純ヘルペスウイルス (HSV))、水痘帯状疱疹ウイルス (VZV)、エプスタイン・バーウイルス (EBV)、サイトメガロウイルス (CMV)、ヘルペスリンパ球向性ウイルス、バラ色靴糠疹 (*Pityriasis Rosea*)、カポジ肉腫関連ヘルペスウイルス)；アデノウイルス (例えば、アタデノウイルス、アピアデノウイルス、イクタデノウイルス、マストアデノウイルス、シアデノウイルス)；ポックスウイルス (例えば、天然痘ウイルス、ワクシニアウイルス、牛痘ウイルス、サル痘ウイルス、オルフウイルス、仮性痘瘡ウイルス、ウシ丘疹性口内炎ウイルス)；タナポックスウイルス、ヤバザル腫瘍ウイルス；伝染性軟属腫ウイルス (MCV)；パルボウイルス (例えば、アデノ随伴ウイルス (AAV)、パルボウイルス B19、ヒトボカウイルス、ブファウイルス、ヒトパルブ4 G1)；ジェミニウイルス科 (*Geminiviridae*)；ナノウイルス科 (*Nanoviridae*)；フィコドナウイルス科 (*Phycodnaviridae*) などが含まれ得る。病原体には、例えば、DNAウイルス [例えば、パポバウイルス (例えば、ヒトパピローマウイルス (HPV)、ポリオーマウイルス)；ヘパドナウイルス (例えば、B型肝炎ウイルス (HBV))；ヘルペスウイルス (例えば、単純ヘルペスウイルス (HSV))、水痘帯状疱疹ウイルス (VZV)、エプスタイン・バーウイルス (EBV)、サイトメガロウイルス (CMV)、ヘルペスリンパ球向性ウイルス、バラ色靴糠疹、カポジ肉腫関連ヘルペスウイルス)；アデノウイルス (例えば、アタデノウイルス、アピアデノウイルス、イクタデノウイルス、マストアデノウイルス、シアデノウイルス)；ポックスウイルス (例えば、天然痘ウイルス、ワクシニアウイルス、牛痘ウイルス、サル痘ウイルス、オルフウイルス、仮性痘瘡ウイルス、ウシ丘疹性口内炎ウイルス)；タナポックスウイルス、ヤバザル腫瘍ウイルス；伝染性軟属腫ウイルス (MCV)；パルボウイルス (例えば、アデノ随伴ウイルス (AAV)、パルボウイルス B19、ヒトボカウイルス、ブファウイルス、ヒトパルブ4 G1)；ジェミニウイルス科；ナノウイルス科；フィコドナウイルス科など]、ヒト型結核菌 (*Mycobacterium tuberculosis*)、ストレプトコッカス・アガラクティエ (*Streptococcus agalactiae*)、メチシリン耐性黄色ブドウ球菌、レジオネラ・ニューモフィラ (*Legionella pneumophila*)、化膿連鎖球菌、大腸菌 (*Escherichia coli*)、淋菌 (*Neisseria gonorrhoeae*)、髄膜炎菌、肺炎球菌 (*Pneumococcus*)、クリプトコッカス・ネオフォルマンズ、ヒストプラズマ・カプスラタム、B型インフルエンザ菌 (*Hemophilus influenzae B*)、梅毒トレポネーマ (*Treponema pallidum*)、ライム病スピロヘータ、緑膿菌 (*Pseudomonas aeruginosa*)、らい菌 (*Mycobacterium leprae*)、ブルセラ・アボルツス (*Bruceella abortus*)、狂犬病ウイルス、インフルエンザウイルス、サイトメガロウイルス、単純ヘルペスウイルス I、単純ヘルペスウイルス II、ヒト血清パルボ様ウイルス、呼吸器合胞体ウイルス、水痘・帯状疱疹ウイルス、B型肝炎ウイルス、C型肝炎ウイルス、麻疹ウイルス、アデノウイルス、ヒトT細胞白血病ウイルス、エプスタイン・バーウイルス、マウス白血病ウイルス、ムンプスウイルス、水疱性口内炎ウイルス、シンドビスウイルス、リンパ球性脈絡髄膜炎ウイルス、イボウイルス、青舌ウイルス、センダイウイルス、ネコ白血病ウイルス、レオウイルス、ポリオウイルス、シミアンウイルス 40、マウス乳がんウイルス、デング熱ウイルス、風疹ウイルス、西ナイルウイルス、熱帯熱マラリア原虫、三日熱マラリア原虫、トキソプラズマ原虫、トリパノソーマ・ランゲリ (*Trypan*

osoma rangeli)、トリパノソーマ・クルーズ、トリパノソーマ・ローデシ
 エンス(Trypanosoma rhodesiense)、トリパノソーマ・ブルー
 セイ(Trypanosoma brucei)、マンソン住血吸虫(Schistosoma mansoni)、日本住血吸虫(Schistosoma japonicum)、バベシア・ボビス(Babesia bovis)、アメリカ・テネラ(Eimeria tenella)、回旋糸状虫(Onchocerca volvulus)、熱帯リーシュマニア(Leishmania tropica)、ヒト型結核菌、旋毛虫(Trichinella spiralis)、東沿岸熱タイレリア(Theileria parva)、胞状条虫(Taenia hydatigena)、タエニア・オビス(Taenia ovis)、無鉤条虫(Taenia saginata)、単包条虫(Echinococcus granulosus)、メソセストイデス・コルチ(Mesocestoides corti)、マイコプラズマ・アルスリティディス(Mycoplasma arthritidis)、M.ヒオルイニス(M. hyorhinis)、M.オラル(M. orale)、M.アルギニニ(M. arginini)、アコレプラズマ・レイドロウイ(Acholeplasma laidlawii)、M.サリバリウム(M. salivarium)、及びM.ニューモニエ(M. pneumoniae)が含まれ得る。病原性ウイルスは、SARS-CoV-2、A型インフルエンザ、B型インフルエンザ、及びノ又はC型インフルエンザのうちの一つ以上を含むことができる。

10

【0136】

20

試料は、生物学的試料、例えば、臨床試料であり得る。一部の実施形態では、試料は、生物学的供給源、例えば、腔、尿道、陰茎、肛門、咽喉、子宮頸部、発酵ブロス、細胞培養物などから採取される。試料は、例えば、糞便試料由来の流体及び細胞を含むことができる。生物学的試料は、(i)対象又は供給源から直接得られるように、又は(ii)前処理の後に、試料の特性を修飾するために使用することができる。したがって、試験試料は、使用前に、例えば、細胞又はウイルス粒子を破壊し、固体材料から液体を調製し、粘性流体を希釈し、液体をろ過し、液体を濃縮し、妨害成分を不活化し、試薬を添加し、核酸を精製することなどによって前処理することができる。したがって、本明細書で使用される「生物学的試料」には、臨床検体又は生物学的検体から抽出された核酸(DNA、RNA又は全核酸)が含まれる。試料調製はまた、分析用に試料を調製するために使用される緩衝液、塩、界面活性剤などを含有する溶液を使用することを含むことができる。一部の実施形態では、試料は、分子試験前に処理される。一部の実施形態では、試料は、直接分析され、試験前に前処理されない。試料は、例えば、糞便試料であり得る。一部の実施形態では、試料は、急性胃腸炎の臨床症状を有する患者由来の糞便試料である。

30

【0137】

一部の実施形態では、試験されるべき試料は、本明細書に開示される方法を行う前に処理される。例えば、一部の実施形態では、試料は、本明細書に開示される方法を行う前に、単離、濃縮、又は種々の他の処理ステップに供することができる。例えば、一部の実施形態では、試料は、本明細書に開示されるように、試料をオリゴヌクレオチドと接触させる前に、試料から核酸を単離するために処理することができる。一部の実施形態では、本明細書に開示される方法は、試料をインビトロで培養することなく、試料に対して行われる。一部の実施形態では、本明細書に開示される方法は、試料を本明細書に開示されるオリゴヌクレオチドと接触させる前に、試料から核酸を単離することなく試料上で行われる。

40

【0138】

試料は、1種以上の核酸(例えば、複数の核酸)を含むことができる。本明細書で使用される用語「複数」とは、2種以上を指すことができる。したがって、一部の実施形態では、試料は、2種以上(例えば、3種以上、5種以上、10種以上、20種以上、50種以上、100種以上、500種以上、1,000種以上、又は5,000種以上)の核酸(例えば、gDNA、mRNA)を含む。開示される方法は、試料中(例えば、gDNA

50

などの核酸の複合混合物中)に存在する標的核酸を検出するための非常に感度の高い方法として用いることができる。一部の実施形態では、試料は、配列が互いに異なる5種以上の核酸(例えば、10種以上、20種以上、50種以上、100種以上、500種以上、1,000種以上、又は5,000種以上のRNA)を含む。一部の実施形態では、試料は、10種以上、20種以上、50種以上、100種以上、500種以上、 10^3 種以上、 5×10^3 種以上、 10^4 種以上、 5×10^4 種以上、 10^5 種以上、 5×10^5 種以上、 10^6 種以上、 5×10^6 種以上、又は 10^7 種以上の核酸を含む。

【0139】

一部の実施形態では、試料は、10~20種、20~50種、50~100種、100~500種、500~ 10^3 種、 10^3 ~ 5×10^3 種、 5×10^3 ~ 10^4 種、 10^4 ~ 5×10^4 種、 5×10^4 ~ 10^5 種、 10^5 ~ 5×10^5 種、 5×10^5 ~ 10^6 種、 10^6 ~ 5×10^6 種、若しくは 5×10^6 ~ 10^7 種、又は 10^7 種以上の核酸を含む。一部の実施形態では、試料は、5~ 10^7 種の核酸(例えば、配列が互いに異なる)(例えば、5~ 10^6 種、5~ 10^5 種、5~50,000種、5~30,000種、10~ 10^6 種、10~ 10^5 種、10~50,000種、10~30,000種、20~ 10^6 種、20~ 10^5 種、20~50,000種、若しくは20~30,000種の核酸、又はこれらの値の任意の2つの間の数値若しくは範囲)を含む。一部の実施形態では、試料は、配列が互いに異なる20種以上の核酸を含む。

【0140】

試料は、(例えば、標的核酸が核酸集団中に存在するかどうかを決定するために)核酸を含む任意の試料であり得る。試料は、任意の供給源から誘導することができ、例えば、試料は、精製された核酸の合成の組み合わせであり得る；試料は、細胞溶解物、DNA濃縮細胞溶解物、又は細胞溶解物から単離及び/若しくは精製された核酸であり得る。試料は、(例えば、診断の目的のために)患者由来であり得る。試料は、透過処理された細胞由来であり得る。試料は、架橋された細胞由来であり得る。試料は、組織切片中にあり得る。試料は、架橋し、続いて、脱脂及び調整して均一な屈折率を作製することによって調製された組織由来であり得る。

【0141】

試料は、標的核酸及び複数の非標的核酸を含むことができる。一部の実施形態では、標的核酸は、1コピー/10個の非標的核酸、1コピー/20個の非標的核酸、1コピー/25個の非標的核酸、1コピー/50個の非標的核酸、1コピー/100個の非標的核酸、1コピー/500個の非標的核酸、1コピー/ 10^3 個の非標的核酸、1コピー/ 5×10^3 個の非標的核酸、1コピー/ 10^4 個の非標的核酸、1コピー/ 5×10^4 個の非標的核酸、1コピー/ 10^5 個の非標的核酸、1コピー/ 5×10^5 個の非標的核酸、1コピー/ 10^6 個の非標的核酸、1コピー未満/ 10^6 個の非標的核酸、又はこれらの値の任意の2つの間の数値若しくは範囲で試料中に存在する。一部の実施形態では、標的核酸は、1コピー/10個の非標的核酸~1コピー/20個の非標的核酸、1コピー/20個の非標的核酸~1コピー/50個の非標的核酸、1コピー/50個の非標的核酸~1コピー/100個の非標的核酸、1コピー/100個の非標的核酸~1コピー/500個の非標的核酸、1コピー/500個の非標的核酸~1コピー/ 10^3 個の非標的核酸、1コピー/ 10^3 個の非標的核酸~1コピー/ 5×10^3 個の非標的核酸、1コピー/ 5×10^3 個の非標的核酸~1コピー/ 10^4 個の非標的核酸、1コピー/ 10^4 個の非標的核酸~1コピー/ 10^5 個の非標的核酸、1コピー/ 10^5 個の非標的核酸~1コピー/ 10^6 個の非標的核酸、若しくは1コピー/ 10^6 個の非標的核酸~1コピー/ 10^7 個の非標的核酸、又はこれらの値の任意の2つの間の数値若しくは範囲で試料中に存在する。

【0142】

適切な試料には、限定されないが、唾液、血液、血清、血漿、尿、吸引物、及び生検試料が含まれる。したがって、患者に関する用語「試料」は、生物学的起源の血液及び他の液体試料、固体組織試料、例えば、生検検体又はそれらに由来する組織培養物若しくは細胞、及びそれらの子孫を包含する。この定義はまた、例えば、試薬による処理、洗浄、又

10

20

30

40

50

はがん細胞などのある特定の細胞集団についての濃縮などによって、それらの調達後に任意の方法で操作された試料を含む。この定義はまた、特定のタイプの分子、例えば、核酸について濃縮された試料を含む。用語「試料」は、生物学的試料、例えば、臨床試料、例えば、血液、血漿、血清、吸引物、脳脊髄液（CSF）を包含し、また、外科的切除によって得られた組織、生検によって得られた組織、培養細胞、細胞上清、細胞溶解物、組織試料、臓器、骨髄なども含む。「生物学的試料」は、それに由来する生物学的流体（例えば、がん性細胞、感染細胞など）、例えば、このような細胞から得られる核酸を含む試料（例えば、核酸を含む細胞溶解物又は他の細胞抽出物）を含む。

【0143】

本明細書に開示される方法において使用するための適切な試料には、植物、動物、細菌などの生物又はその一部から得られる任意の慣用的な生物学的試料が含まれる。特定の実施形態では、生物学的試料は、ヒト対象などの動物対象から得られる。生物学的試料は、限定されないが、とりわけ細菌、酵母、原生動物、及びアメーバなどの単細胞生物、多細胞生物を含む任意の生物から得られるか、排泄されるか、又は分泌される任意の固体又は流体の試料（植物又は動物など、健康な若しくは見かけ上健康なヒト対象、又は病原性細菌若しくはウイルスなどの病原性微生物による感染など、診断又は調査されるべき対象の状態若しくは疾患に罹患しているヒト患者由来の試料を含む）である。例えば、生物学的試料は、例えば、血液、血漿、血清、尿、糞便、痰、粘液、リンパ液、滑液、胆汁、腹水、胸水、漿液腫、唾液、脳脊髄液、房水若しくは硝子体液、又は任意の分泌液、浸出液、滲出液（例えば、膿瘍、又は感染若しくは炎症の任意の他の部位から得られる流体）、又は関節（例えば、関節リウマチ、変形性関節症、痛風又は敗血症性関節炎）から得られる流体、又は皮膚若しくは粘膜表面のスワブから得られる生物学的流体であり得る。

【0144】

試料はまた、任意の臓器又は組織から得られた試料（生検又は剖検検体、例えば腫瘍生検を含む）であり得るか、又は細胞（初代細胞又は培養細胞のいずれか）又は任意の細胞、組織若しくは臓器によって馴化された培地を含むことができる。例示的な試料としては、限定されないが、細胞、細胞溶解物、血液スメア、細胞遠心分離調製物、細胞診スメア、体液（例えば、血液、血漿、血清、唾液、痰、尿、気管支肺胞洗浄、精液など）、組織生検（例えば、腫瘍生検）、細針吸引物、及び/又は組織切片（例えば、クライオスタット組織切片及び/又はパラフィン包埋組織切片）が挙げられる。他の例では、試料は循環腫瘍細胞（細胞表面マーカーによって同定することができる）を含む。特別な例では、試料は、直接的に使用され（例えば、新鮮又は凍結）、又は使用前に、例えば、固定（例えば、ホルマリンを使用）及び/又はワックス（例えば、ホルマリン固定パラフィン包埋（FFPE）組織試料など）に包埋することによって操作され得る。対象から組織を得る任意の方法を利用することができるが、使用される方法の選択は、組織のタイプ、対象の年齢、又は開業医が利用できる手法などの様々な要因に依存することが理解される。このような試料を獲得するための標準的な技術は、当該技術分野において利用可能である。

試料は、水、土壌などの環境試料、又は工業表面若しくは医療表面などの表面であり得る。

本明細書に開示される実施形態の感度が増大するため、ある特定の例示的な実施形態では、アッセイ及び方法は、粗試料、又は検出されるべき標的分子が試料からさらに分画又は精製されていない試料上で行うことができる。

【0145】

細胞を溶解して標的分子（例えば、標的 dsDNA）を遊離させることができる。細胞溶解は、例えば、化学的若しくは生化学的手段、浸透ショック、又は熱的溶解、機械的溶解、若しくは光学的溶解のいずれかの種々の手段によって達成することができる。細胞は、界面活性剤（例えば、SDS、ドデシル硫酸 Li、Triton X-100、Tween-20、又は NP-40）、有機溶媒（例えば、メタノール又はアセトン）、若しくは消化酵素（例えば、プロテイナーゼ K、ペプシン、又はトリプシン）、又はそれらの任意の組み合わせを含む細胞溶解緩衝液の添加によって溶解され得る。標的とバーコードと

の結び付きを増加させるために、標的分子の拡散速度は、例えば、温度を低下させ、及び/又は溶解物の粘度を増加させることによって変化させることができる。

一部の実施形態では、試料は、ろ紙を用いて溶解され得る。ろ紙は、ろ紙の上部に溶解緩衝液を浸すことができる。ろ紙は、試料の溶解及び試料の標的の基質へのハイブリダイゼーションを容易にすることができる圧力で試料に適用することができる。

【0146】

一部の実施形態では、溶解は、機械的溶解、熱溶解、光学的溶解、及び/又は化学的溶解によって行うことができる。化学的溶解には、プロテイナーゼK、ペプシン、及びトリプシンなどの消化酵素の使用が含まれ得る。溶解は、基質に溶解緩衝液を添加することによって行うことができる。溶解緩衝液はTris HClを含むことができる。溶解緩衝液は、少なくとも約0.01、0.05、0.1、0.5、若しくは1M又はそれ以上のTris HClを含むことができる。溶解緩衝液は、多くとも約0.01、0.05、0.1、0.5、若しくは1M又はそれ以上のTris HClを含むことができる。溶解緩衝液は、約0.1MのTris HClを含むことができる。溶解緩衝液のpHは、低くとも約1、2、3、4、5、6、7、8、9、10、又はそれ以上であり得る。溶解緩衝液のpHは、高くとも約1、2、3、4、5、6、7、8、9、10、又はそれ以上であり得る。一部の実施形態では、溶解緩衝液のpHは約7.5である。溶解緩衝液は、塩(例えば、LiCl)を含むことができる。溶解緩衝液中の塩の濃度は、低くとも約0.1、0.5、若しくは1M又はそれ以上であり得る。溶解緩衝液中の塩の濃度は、高くとも約0.1、0.5、若しくは1M又はそれ以上であり得る。一部の実施形態では、溶解緩衝液中の塩の濃度は約0.5Mである。溶解緩衝液は、界面活性剤(例えば、SDS、ドデシル硫酸Li、triton X、tween、NP-40)を含むことができる。溶解緩衝液中の界面活性剤の濃度は、低くとも約0.0001%、0.0005%、0.001%、0.005%、0.01%、0.05%、0.1%、0.5%、1%、2%、3%、4%、5%、6%、若しくは7%又はそれ以上であり得る。溶解緩衝液中の界面活性剤の濃度は、高くとも約0.0001%、0.0005%、0.001%、0.005%、0.01%、0.05%、0.1%、0.5%、1%、2%、3%、4%、5%、6%、若しくは7%又はそれ以上であり得る。一部の実施形態では、溶解緩衝液中の界面活性剤の濃度は、約1%ドデシル硫酸Liである。溶解方法において使用する時間は、使用する界面活性剤の量によって異なることがある。一部の実施形態では、使用される界面活性剤が多いほど、溶解に必要な時間はより少なくなる。溶解緩衝液は、キレート化剤(例えば、EDTA、EGTA)を含むことができる。溶解緩衝液中のキレート剤の濃度は、低くとも約1、5、10、15、20、25、若しくは30mM又はそれ以上であり得る。溶解緩衝液中のキレート剤の濃度は、高くとも約1、5、10、15、20、25、若しくは30mM又はそれ以上であり得る。一部の実施形態では、溶解緩衝液中のキレート化剤の濃度は約10mMである。溶解緩衝液は還元試薬(例えば、ベータ-メルカプトエタノール、DTT)を含むことができる。溶解緩衝液中の還元試薬の濃度は、低くとも約1、5、10、15、若しくは20mM又はそれ以上であり得る。溶解緩衝液中の還元試薬の濃度は、高くとも約1、5、10、15、若しくは20mM又はそれ以上であり得る。一部の実施形態では、溶解緩衝液中の還元試薬の濃度は約5mMである。一部の実施形態では、溶解緩衝液は、約0.1MのTris HCl、約pH7.5、約0.5M LiCl、約1%ドデシル硫酸リチウム、約10mM EDTA、及び約5mM DTTを含むことができる。

【0147】

溶解は、約4、10、15、20、25、又は30の温度で行うことができる。溶解は、約1、5、10、15、若しくは20分間又はそれ以上で行うことができる。溶解細胞は、少なくとも約100000個、200000個、300000個、400000個、500000個、600000個、若しくは700000個又はそれ以上の標的核酸分子を含むことができる。溶解細胞は、多くとも約100000個、200000個、300000個、400000個、500000個、600000個、若しくは700000

10

20

30

40

50

個又はそれ以上の標的核酸分子を含むことができる。

【0148】

キット

本明細書に記載されるキットは、複数のタンパク質複合体を含むことができる。一部の実施形態では、複数のタンパク質複合体の各々は、トランスポソームと、標的二本鎖DNA (dsDNA) 上の結合部位に特異的に結合することができるプログラム可能なDNA結合ユニットとを含む。一部の実施形態では、トランスポソームは、トランスポゼース、第1のアダプター、及び第2のアダプターを含む。一部の実施形態では、複数のタンパク質複合体の各々に対する結合部位は互いに異なる。一部の実施形態では、キットは、核酸増幅産物についてリアルタイム検出活性を提供する少なくとも1つの成分を含む。リアルタイム検出活性は、分子ビーコンによって提供することができる。乾燥組成物は、逆転写酵素及び/又は逆転写プライマーを含むことができる。

10

キットは、例えば、本明細書に記載されるように、1種以上のポリメラーゼ及び1種以上のプライマー、並びに場合により1種以上の逆転写酵素及び/又は逆転写プライマーを含むことができる。1つの標的が増幅される場合、1対のプライマー(フォワード及びリバース)をキットに含めることができる。複数の標的配列が増幅される場合、複数のプライマー対をキットに含めることができる。キットは、対照ポリヌクレオチドを含むことができ、複数の標的配列が増幅される場合、複数の対照ポリヌクレオチドをキットに含めることができる。

【0149】

20

キットはまた、任意の数の別々の容器(vessel)、チャンバー、容器(container)、バケツ、チューブ、バイアル、マイクロタイタープレートなどに1つ以上の成分を含むことができ、又は成分は、このような容器中で種々の組み合わせで組み合わせることができる。キットの成分は、例えば、1つ以上の容器中に存在することができる。一部の実施形態では、すべての構成要素は、1つの容器内に提供される。一部の実施形態では、酵素(例えば、ポリメラーゼ及び/又は逆転写酵素)は、プライマーから別々の容器内に提供することができる。成分は、例えば、凍結乾燥(lyophilized)、加熱乾燥、凍結乾燥(freeze dried)、又は安定な緩衝液中であり得る。一部の実施形態では、ポリメラーゼ及び/又は逆転写酵素は、凍結乾燥形態であるか、又は単一容器中で加熱乾燥形態であり、プライマーは、異なる容器中で凍結乾燥され、加熱乾燥され、凍結乾燥され、又は緩衝液中である。一部の実施形態では、ポリメラーゼ及び/又は逆転写酵素、及びプライマーは、凍結乾燥形態又は加熱乾燥形態で単一の容器内にある。

30

キットは、例えば、反応に使用されるdNTP、又は反応に使用される修飾ヌクレオチド、容器、キュベット若しくは他の容器、又は凍結乾燥若しくは加熱乾燥成分を再水和するための水又は緩衝液のバイアルをさらに含むことができる。使用される緩衝液は、例えば、ポリメラーゼとプライマーアニーリング活性の両方に適していることができる。

【0150】

キットはまた、本明細書に記載される1つ以上の方法を行うための説明書、及び/又は本明細書に記載される1つ以上の構成要素の説明を含むことができる。指示及び/又は説明は、印刷形態であり得、キットインサートに含めることができる。キットはまた、このような指示又は説明を提供するインターネットロケーションの書面による説明を含むことができる。

40

キットは、検出方法に使用される試薬、例えば、FRETに使用される試薬、側方流動デバイス、ディップスティック、蛍光色素、コロイド金粒子、ラテックス粒子、分子ビーコン、又はポリスチレンビーズをさらに含むことができる。

本開示の図1、3、4、5A-5F及び7A-7Hは、BioRender.comを用いて作製された。

【実施例】

【0151】

50

上記で検討した実施形態の一部の態様は、以下の実施例においてさらに詳細に開示され、本開示の範囲を限定することを意図したものではない。

(実施例1)

融合タンパク質及びガイドRNA (sgRNA) の設計及び検証

融合タンパク質を生成するための4つの構築物を設計した: dCAS9 - F126 - Tn5、dCAS9 - xTen - Tn5、Tn5 - F126 - dCas9、Tn5 - xTen - dCas9 (例えば、図8~図10を参照されたい)。これらの構築物は、F126リンカー又はxTenリンカーによって分離された融合タンパク質のN末端にdCas9又はTn5配列のいずれかを有する。プラスミド設計は、一部の実施形態では、"Chen, S.P. & Wang, H.H. (2019). An Engineered Cas-Transposon System for Programmable and Site-Directed DNA Transposition. The CRISPR Journal. Vol 2, Number 6. DOI: 10.1089/crispr.2019.0030 及びPicelli S., Bjorklund, A.K., Reinius, B., Sgasser, S., Wingerb, G., & Sandbert, R. (2014)"; 及び"Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. Genome Research. 24:2033-2040. ISSN 1088-9051/14"に基づく。

10

【0152】

sgRNA設計

サルモネラ・エンテリカのInvA及びFliC遺伝子を標的とするsgRNAを設計した。サルモネラ・エンテリカ株ATCC 13311の配列を用いた。sgRNAは、統合DNA技術(IDT)からのツールを用いて設計された(表1)。InvA及びFliC遺伝子のsgRNAの相対的位置をそれぞれ図11及び図12に示す。

20

【0153】

【表1】

表1: S.エンテリカsgRNAs

試料名	チューブラベル	[μM]	種	遺伝子	配列	配列番号
CD.Cas9.RNXS0617.A A	Sg1	10	サルモネラ・エンテリカ	InvA	GCTATTTTGACCAT TTCAAT	10
CD.Cas9.RNXS0617.A B	Sg2	10	サルモネラ・エンテリカ	InvA	CGGAGGACAAATC CATACCA	11
CD.Cas9.RNXS0617.A C	Sg3	10	サルモネラ・エンテリカ	InvA	CAGTTTATCGTTAT TACCAA	12
CD.Cas9.RNXS0617.A D	Sg4	10	サルモネラ・エンテリカ	InvA	ACTTATAACCATGCT GACCAT	13
CD.Cas9.RZZJ8230.A A	Sg5	10	サルモネラ・エンテリカ	FliC	GGACAACACCCTGA CCATCC	14
CD.Cas9.RZZJ8230.A C	Sg6	10	サルモネラ・エンテリカ	FliC	GTCTGACCTCGACT CCATCC	15
CD.Cas9.RZZJ8230.A D	Sg7	10	サルモネラ・エンテリカ	FliC	GAACATCAAAGGTC TGACTC	16

ACからAA PAMから5' sgRNAまで79bp。ADからAC PAMから5' sgRNAまで127 bp。

30

40

InvAでは264bp、8bp、148bp、292bp、458bp及び195bp断片が予想された。FliCでは約130bp、82bp、232bpの断片が予想された。

【0154】

サルモネラ・エンテリカsgRNAの検証

sgRNAの特異性を検証するために、ゲノム試料をCas9により切断した。アダブ

50

ターを Cas 9 により切断された DNA にライゲートし、PCR 増幅断片をバイオアナライザーにより可視化した。

【 0 1 5 5 】

【 表 2 】

表2: sgRNA活性のバイオアナライザー分析

Cas 9 予想断片 [bp]	ユニバーサルアダプターサイズ [bp]	バーコードアダプターサイズ [bp]	バイオアナライザー予想サイズ [bp]	バイオアナライザー予想サイズ 低範囲 [bp]	バイオアナライザー予想サイズ 低範囲 [bp] ²	実測 [bp]
8	57	63	128	115.2	140.8	125
82	57	63	202	181.8	222.2	225
130	57	63	250	225	275	266
148	57	63	268	241.2	294.8	272
195	57	63	315	283.5	346.5	296
232	57	63	352	316.8	387.2	323
264	57	63	384	345.6	422.4	382
292	57	63	412	370.8	453.2	425
458	57	63	578	520.2	635.8	

10

図 1 3 及び表 2 は、gDNA の切断が予想されるサイズに特異的であったことを示しており（表 2 の「バイオアナライザー予測サイズ [bp] カラム」を「実測 [bp]」カラムと比較する）、したがって、サルモネラ・エンテリカのためのガイド RNA が機能的であることを示す。

次に、ヒト遺伝子 EXT 1、BCL 9、HOXA 13、HOXD 11、及び OLIG 2 を標的とする sgRNA を 10 の総 sgRNA について設計した（表 3 A ~ 表 3 C）。sgRNA は GenScript のツールを用いて設計された。

【 0 1 5 6 】

【 表 3 】

表3A: ヒトsgRNA標的

遺伝子	切断サイズ [bp]	アダプターサイズ [bp]	断片全体のサイズ [bp]	gRNA 設計	腫瘍型
EXT1	255	34	323	1,4	外骨腫症、骨肉腫
BCL9	66	34	134	2, 4	B細胞性急性リンパ性白血病
HOXA13	445	34	513	1,4	急性骨髄性白血病
HOXD11	225	34	293	1,3	急性骨髄性白血病
OLIG2	114	34	182	1,2	T細胞性急性リンパ芽球性白血病

30

40

【 0 1 5 7 】

50

【表 4】

表3B: ヒトsgRNA標的

遺伝子	オンター ゲットス コア1	オンター ゲットス コア2	総合ス コア1	総合ス コア2	オンター ゲットA VG	総合ス コアAV G	オンター ゲットス コア1
EXT1	85.69	84.4	82	58	85.045	70	85.69
BCL9	88.35	80.6	76	46	84.475	61	88.35
HOXA13	91.27	73.1	77	60	82.185	68.5	91.27
HOXD11	88.1	74.58	75	66	81.34	70.5	88.1
OLIG2	64.69	96.08	65	57	80.385	61	64.69

10

【 0 1 5 8 】

【表 5】

表3C: ヒトsgRNA標的

名称	インプット配列	配列番号
EXT1-1	ATATCACGTCCATAACGGGG	17
EXT1-4	CACTTGGCCTGACTACACCG	18
BCL9-2	GGGTTGGCATCGGAACCACG	19
BCL9-4	GATGCCCTCTCCAAATGCCG	20
HOXA13-1	GTAGCCATAGGGCAGCGCCG	21
HOXA13-4	TTTCTCTACGACAACGGCGG	22
HOXD11-1	GGGCTTCGACCAGTTCTACG	23
HOXD11-3	GGGCTACGCTCCCTACTACG	24
OLIG2-1	ACTGGTGAGCGAGATCTACG	25
OLIG2-2	GCACGCCGCACATCACCCCG	26

20

gRNAはまた、クラミジア・トラコマチス遺伝子多型膜タンパク質A (pmpA) を標的とするように設計された(表4)。全部で5つのsgRNAをIDTからのツールを用いて設計した。

30

【 0 1 5 9 】

40

50

【表 6】

表4: C.トラコマチスsgRNA標的

設計ID	遺伝子記号	位置	鎖	配列	配列番号	PAM	オンターゲットスコア	断片サイズ [bp]
CD.Cas9.CVSJ0588.AF	pmpA	45	+	GAAATTAATGGTT TAAGCTT	4	TGG	76	161
CD.Cas9.CVSJ0588.AA	pmpA	206	-	AGGTGAGCAAGAT TTCCATT	5	TGG	100	169
CD.Cas9.CVSJ0588.AC	pmpA	375	-	TCAAGGACATATT CTCCTGT	6	TGG	85	48
CD.Cas9.CVSJ0588.AJ	pmpA	423	-	AATGTGCTCCATA AGGAATT	7	AGG	71	140
CD.Cas9.CVSJ0588.AG	pmpA	563	-	AACTTTCCTTCT GAGGAG	8	TGG	75	260
CD.Cas9.CVSJ0588.AB	pmpA	823	+	AAGATCACACCTA TGGGAAA	9	TGG	100	

10

【0160】

20

トランスポゼース Tn5 の検証

図14～図15は、Tn5が、PCR増幅のためのDNA断片に、設計されたアダプターA（5'-TCGTCGGCAGCGTCAGATGTTGTTATAAGAGACAG-3'、配列番号27）及びアダプターB（5'-GTCTCGTGGGCTCGGAGATGTTGTTATAAGAGACAG-3'、配列番号28）をライゲートすることができることを示し、機能性を示す。最初に、Tn5を用いてS.エンテリカ由来のgDNAをカスタムアダプターで切断し、ペーストした。次に、標識断片をPCRにより増幅した。図14～図15のデータは、Tn5トランスポゼースがカスタムアダプターとともにロードされたことを示す。

【0161】

30

融合タンパク質の検証

dCAS9-F126-Tn5、dCAS9-xTen-Tn5、Tn5-F126-dCas9、Tn5-xTen-dCas9を組換え発現させ、次に精製した。一部の実施形態では、組換えタンパク質は、キチンカラム上の自己切断部分（インテイン）を用いて単離された。精製された融合タンパク質を、SDS-PAGEゲル上で予測されるサイズ及び純度について分析した（図16～図21）。

dCAS9-F126-Tn5のSDS-PAGE分析を図16に示す。試料は純度：>80%であることが観察された。一部の実施形態では、融合タンパク質はまた、インテインドメインを含み得る。図17のバイオアナライザー分析は、作製されたタンパク質の一部（44.91のピーク）が正しいサイズ（インテインを含まない）であることを示す。

40

dCAS9-xTen-Tn5のSDS-PAGE分析を図18に示す。試料は純度：>70%であることが観察された。一部の実施形態では、融合タンパク質はまた、インテインドメインを含み得、その結果、予想よりも大きいサイズとなる。図19のバイオアナライザー分析は、作製されたタンパク質の一部（44.62のピーク）が正しいサイズ（インテインを含まない）であることを示す。

図20は、組換え発現及び精製されたTn5-F126-dCas9のSDS-PAGE分析を示す。図21は、組換え発現及び精製されたTn5-xTen-dCas9のSDS-PAGE分析を示す。試料は純度：>65%であることが観察された。

【0162】

50

機能性についての融合タンパク質の試験

C a s 9 - F l 2 6 - T n 5 及び d C a s 9 - x T e n - T n 5 を機能性について試験した。プロトコールは、以下：(1) s g R N A 及びアダプターを融合タンパク質(特に断りのない限り、使用したヒト s g R N A) にロードする、(2) 誘導タグメンテーション、(3) クリーンアップ、(4) P C R 増幅、(5) 品質管理(Q C)、及び(6) 結果分析であった。

s g R N A とアダプターを融合タンパク質にロードする

融合タンパク質を 1 : 1 : 2 (1 分子の d C a s 9 - T n 5 対 1 分子の s g R N A 対 2 分子のアダプター) の比でロードした。混合物を 2 4 で 3 0 分間インキュベートした。

【 0 1 6 3 】

誘導タグメンテーション

1 0 0 m M の d C a s 9 - T n 5 (6 . 0 2 e 1 0 分子) 及び 5 0 0 n g のヒト g D N A (1 . 5 2 e 5 分子) を、g D N A 対 d C a s 9 - T n 5 の 1 ~ 3 . 9 5 e 5 の比で組み合わせた。混合物を 3 7 で 6 0 分間、及び 5 5 で 6 0 分間インキュベートして、タグ付けされた断片を生成した。いくつかのインキュベーション方法が試みられ、一部の実施形態では、d C a s 9 は、2 5 ~ 4 2 の範囲で機能することができる。T n 5 は、3 7 ~ 6 0 の範囲で機能することができる。P C R 増幅プログラムを表 5 に示す。

【 0 1 6 4 】

【表 7】

表5: PCR増幅

温度	時間	
95	30秒	
72	3分	
95	10秒	X 35サイクル
64	30秒	
72	30秒	
72	5分	
4	長期	

図 2 2 は、C a s 9 のみの対照反応に関するデータを示す。可視線は、C a s 9 消化後の D N A のテープステーション分析を示す。P C R 増幅反応後の試料分析はシグナルを示さない。このデータは、C a s 9 自体が D N A 断片の 5 ' 末端又は 3 ' 末端にアダプターを付加することができないことを示した。

【 0 1 6 5 】

図 2 3 ~ 図 2 4 は、それぞれ d C a s 9 - F l 2 6 - T n 5 又は d C a s 9 - x T e n - T n 5 によるアダプターの消化及びライゲーション後の P C R 増幅の結果を示す。グラフの矢印は、P C R 後の試料からのシグナルを指示する。P C R 増幅が検出されたが、これは、両方の融合タンパク質(d C a s 9 - F l 2 6 - T n 5 及び d C a s 9 - x T e n - T n 5) が、転位(例えば、D N A 分子の 5 ' 末端及び 3 ' 末端にアダプター(アダプター 4 0 - B) を付加すること)を行うことのできる場合にのみ可能である。

結果

結果は、T n 5 がヒト g D N A にカスタムアダプターを付加することができることを示す。C a s 9 のみの対照は、このプロセスが増幅するために T n 5 を必要とすることを示した。これらの結果は、d C a s 9 に融合した T n 5 の機能性を示した。

【 0 1 6 6 】

融合タンパク質対 D N A 比試験

次に、g D N A 対 C a s 9 - T n 5 比に低下させる効果を試験した。D N A 濃度は、C a s - T n 融合タンパク質濃度：1 0 0 n M (d C a s 9 - T n 5 の 1 9 4 , 0 7 1 分子対 D N A の 1 ゲノムコピー)、1 n M (1 , 9 4 0 : 1)、1 0 0 p M (1 9 4 : 1)、

10

20

30

50

10 pM (19.4 : 1)、1 pM (1.94 : 1) を低下させながら一定に保った。結果を図 25 ~ 図 31 に示す。図 25 は、dCas9 - Tn5 の 194,071 : 1 の比を用いた誘導タグメンテーション反応後の PCR 増幅の結果を示し、PCR 後の広いピークを示し、非特異的タグメンテーションを示す。dCas9 - Tn5 量の減少 (図 26 ~ 図 31) は、PCR 反応から検出可能なピークの生成をもたらす、これは、融合タンパク質対 DNA の比を低下させることが、タグメンテーションに特異性を付加することを示す。

結果

結果は、Tn5 がヒト gDNA にカスタムアダプターを付加することができたことを示す。Cas9 のみの対照は、このプロセスが DNA を増幅するために Tn5 を必要とすることを示した。Tn5 は、機能的であることが示され、誘導転位の証拠があった。したがって、dCas9 と Tn5 活性の両方を含む融合タンパク質の証拠がある。

10

S . エンテリカ上の融合タンパク質及び sgRNA

図 38 ~ 図 39 は、dCas9 - xTen - Tn5 上の S . エンテリカ sgRNA を使用する誘導タグメンテーションを示す。このデータは、sgRNA の付加が特異性を付加することを示す。図 39 は、sgRNA を伴わない誘導タグメンテーションがランダムであることを示す。図 38 は、sgRNA の付加が特異性を付与することを示す。

【0167】

(実施例 2)

試料ライブラリーの調製

誘導タグメンテーションライブラリー

20

本明細書には、Illumina NextSeq 上でシーケンシングするためのライブラリーを生成する方法及び組成物が記載される。

3 つのライブラリーをライゲーションベースの方法 (図 37、図 40、図 42A ~ 図 42B) を用いて作製し、NEBNext シーケンシングアダプターを単一アダプター (例えば、Tn5 単独又は dCas9 - Tn5 融合のいずれかを用いるアダプター B) を用いてタグメンテーションするステップ後に添加し、2 つのライブラリーを誘導タグメンテーションベースの方法 (図 41、図 43 ~ 図 44) を用いて作製し、NGS に必要な配列は、アダプター A 及び B 上の誘導タグメンテーションステップに含まれた。すべてのライブラリーは、ヒト sgRNA を用いて調製された。誘導タグメンテーションには、dCas9 - F126 - Tn5 融合タンパク質を用いた。これらの実験において、DNA は、長い又は短いインキュベーションプロトコル下で、dCas9 - Tn5 とともにインキュベートされた。短いプロトコルでは、反応液を 30 で 30 分間インキュベートし、次に 37 で 30 分間インキュベートした。長いプロトコルでは、反応液を 30 で 30 分間インキュベートし、続いて 38 で 60 分間インキュベートし、次に 55 で 60 分間インキュベートした。

30

図 32 は、Tn5 のみを用いた高度に多重化された単一プライマー DNA 増幅を示す。バイオアナライザー分析は、PCR による非特異的 DNA 増幅を示し、1 つのプライマー (アダプター B) のみを用いて DNA を増幅し得ることを示す。

【0168】

dCas9 - Tn5 融合タンパク質を用いる高度に多重化された単一プライマー DNA 増幅を支持する証拠は、図 33 (短いインキュベーションプロトコル) 及び図 34 (より長いインキュベーションプロトコル) に示される。PCR 増幅のバイオアナライザー分析は、1 プライマー (アダプター B) のみを用いた数種の DNA 断片の同時特異的増幅を示した。

40

カスタマイズされた遺伝子座特異的配列ライブラリー調製を支持する証拠は、図 35 (より長いインキュベーションプロトコル) 及び図 36 (より短いインキュベーションプロトコル) に示される。バイオアナライザー分析は、シーケンシングライブラリーが作製されることを示す。Illumina プラットフォームでのシーケンシングに必要なアダプター A 及び B の付加は、誘導タグメンテーションを使用してシーケンシングライブラリーを作製し得ることを示す。

50

【 0 1 6 9 】

前述の実施形態の少なくとも一部において、実施形態で使用される1種以上のエレメントは、そのような置換が技術的に実施可能でない場合を除き、別の実施形態で互換的に使用することができる。当業者であれば、請求された主題事項の範囲から逸脱することなく、上述の方法及び構造に対して、種々の他の省略、追加及び修飾を行うことができることが理解される。このような修飾及び変更はすべて、添付の特許請求の範囲によって定義されるように、主題の範囲内に入ることが意図される。

【 0 1 7 0 】

本明細書における実質的に任意の複数の用語及び/又は単数の用語の使用に関して、当業者は、文脈及び/又は用途に適切のように、複数から単数及び/又は単数から複数に翻訳することができる。種々の単数/複数の並べ替えは、明瞭化のために、本明細書に明示的に記載することができる。本明細書及び添付の特許請求の範囲で使用されるように、文脈が明確に別段の指示をしない限り、単数形「1つの(a)」、「1つの(an)」、及び「その(the)」は複数の指示対象を含む。本明細書における「又は」への任意の言及は、特に断らない限り、「及び/又は」を包含することを意図する。

【 0 1 7 1 】

一般的に、本明細書で使用される用語、特に添付の特許請求の範囲(例えば、添付の特許請求の範囲の本体)で使用される用語は、一般的に「開かれた」用語として意図されていることが、当業者によって理解される(例えば、用語「含んでいる」は、「限定されないが、含んでいる」として解釈されるべきであり、用語「有している」は、「少なくとも有する」として解釈されるべきであり、用語「含む」は、「限定されないが、含む」として解釈されるべきである)。さらに、導入された請求項の記載の特定の数が意図されている場合には、このような意図が請求項に明示的に記載され、このような記載がない限り、このような意図は存在しないことが、当業者によって理解される。例えば、理解を助けるために、以下の添付の特許請求の範囲は、請求項の記載を導入するための「少なくとも1つの」及び「1つ以上の」という導入用語の使用を含有し得る。しかしながら、このような語句の使用は、たとえ同じ請求項が「1つ以上の」又は「少なくとも1つの」の導入文、及び「1つの(a)」又は「1つの(an)」などの不定冠詞を含む場合であっても、不定冠詞「1つの(a)」又は「1つの(an)」による請求項の記載の導入が、このような導入された請求項の記載を含有する任意の特定の請求項を、このような記載を1つのみを含む実施態様に限定することを意味すると解釈されるべきではなく(例えば、「1つの(a)」又は「1つの(an)」は、「少なくとも1つの」又は「1つ以上の」を意味すると解釈されるべきである);請求項の記載を導入するために使用される定冠詞の使用についても同様である。さらに、導入された請求項の特定の数の記載が明示的に記載されている場合であっても、当業者は、このような記載が、少なくとも記載された数を意味する(例えば、他の修飾語なしでの「2つの記載」のそのままの記載は、少なくとも2つの記載、又は2つ以上の記載を意味する)と解釈されるべきであることを認識する。さらに、「A、B、及びCなどのうちの少なくとも1つ」に類似する慣例が用いられる場合には、一般的に、このような構造は、当業者が該慣例を理解することが意図されるという意味で意図される(例えば、「A、B、及びCのうちの少なくとも1つを有するシステム」には、限定されないが、Aのみ、Bのみ、Cのみ、AとBの組み合わせ、AとCの組み合わせ、BとCの組み合わせ、及び/又はAとBとCの組み合わせなどを有するシステムが含まれる)。「A、B、又はCなどのうちの少なくとも1つ」に類似する慣例が使用される場合には、一般的に、このような構造は、当業者が該慣例を理解するという意味で意図される(例えば、「A、B、又はCのうちの少なくとも1つを有するシステム」には、限定されないが、Aのみ、Bのみ、Cのみ、AとBの組み合わせ、AとCの組み合わせ、BとCの組み合わせ、及び/又はAとBとCの組み合わせなどを有するシステムが含まれる)。明細書、特許請求の範囲又は図面のいずれにおいても、2つ以上の代替用語を提示する実質的に任意の分離語及び/又は語句は、用語の1つ、用語のいずれか又は両方の用語を含む可能性を考慮するように理解されるべきであることが、当業者によってさらに理

10

20

30

40

50

解される。

さらに、本開示の特徴又は態様がマーカッシュ群という観点から記載されている場合、当業者は、本開示がまた、マーカッシュ群のメンバーの任意の個々のメンバー又はサブグループという観点からも記載されていることを認識する。

【0172】

当業者に理解されるように、例えば、書面による説明を提供するという観点から、任意の及びすべての目的のために、本明細書に開示されるすべての範囲はまた、その任意の及びすべての可能な部分的範囲、並びにその部分的範囲の組み合わせを包含する。任意の列挙された範囲は、同じ範囲が少なくとも等しい半分、3分の1、4分の1、5分の1、10分の1などに分けられることを十分に記載し、可能にするものとして容易に認識することができる。非限定的な例として、本明細書において検討される各範囲は、下部3分の1、中央部3分の1及び上部3分の1などに容易に分けることができる。また、当業者に理解されるように、「最大」、「少なくとも」、「より大きい」、「より小さい」などのすべての言語は、引用された数を含み、上記で検討した部分的範囲内に実質的に分けることができる範囲を指す。最後に、当業者に理解されるように、範囲は、各個々のメンバーを含む。したがって、例えば、1～3個の物品を有する群は、1、2、又は3個の物品を有する群を指す。同様に、1～5個の物品を有する群は、1、2、3、4、又は5個の物品を有する群を指すなどである。

【0173】

本明細書には様々な態様及び実施形態が開示されているが、他の態様及び実施形態は当業者には明らかである。本明細書に開示される様々な態様及び実施形態は、例示の目的であり、限定することを意図しない。真の範囲及び精神は、続く特許請求の範囲によって示される。

【図面】

【図1】

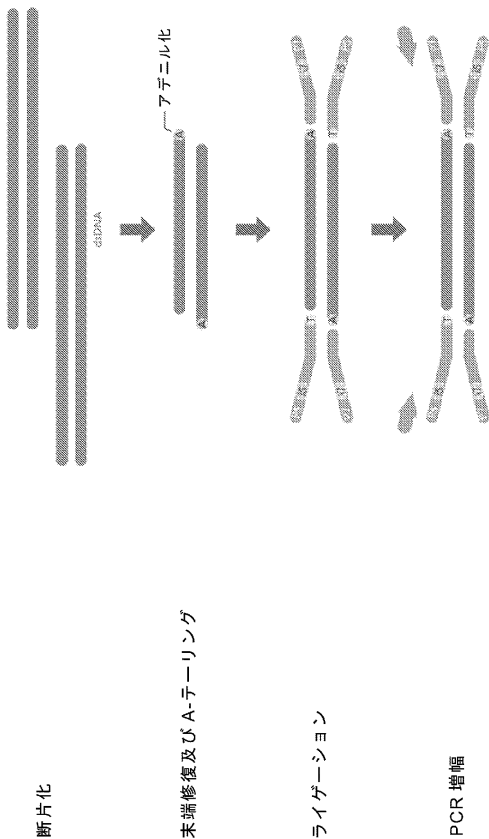


FIG. 1

【図2】

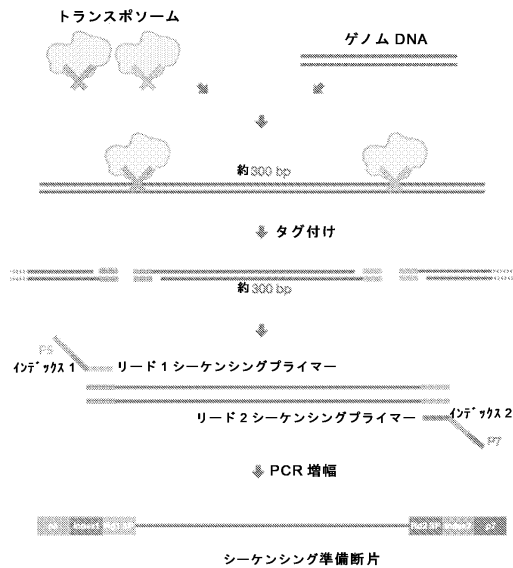


FIG. 2

10

20

30

40

50

【 図 3 】

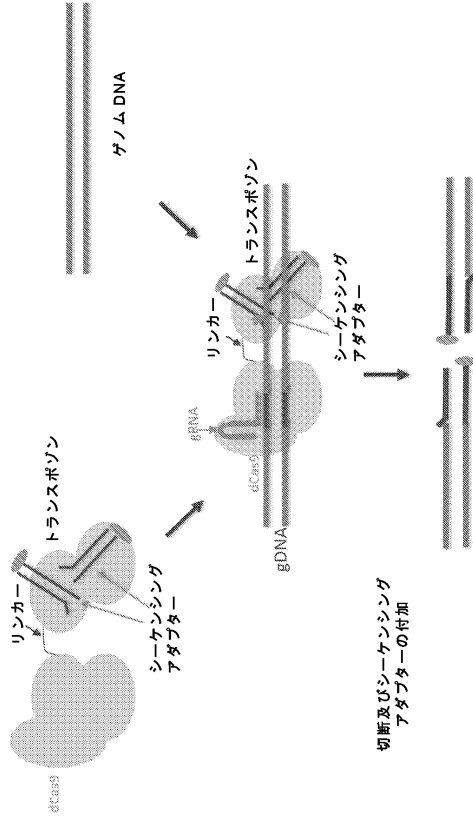


FIG. 3

【 図 4 】

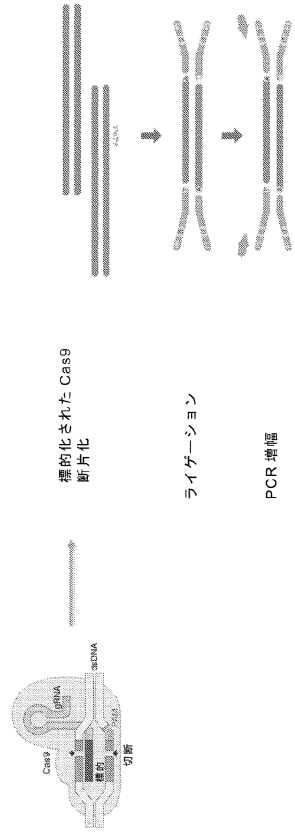


FIG. 4

【 図 5 A 】

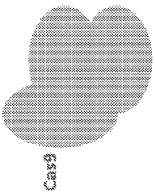


FIG. 5A

【 図 5 B 】

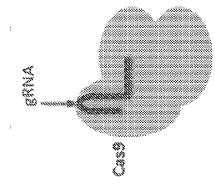


FIG. 5B

10

20

30

40

50

【 図 5 C 】

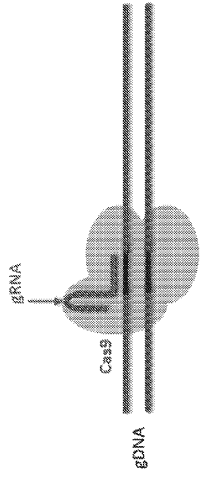


FIG. 5C

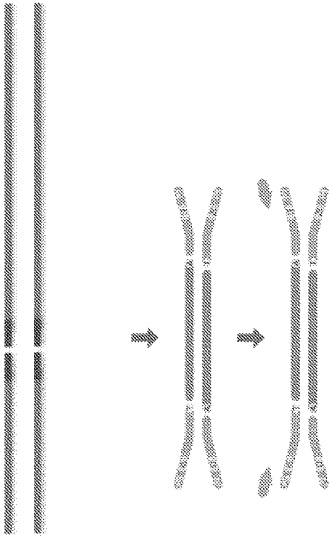
【 図 5 D 】



FIG. 5D

10

【 図 5 E 】



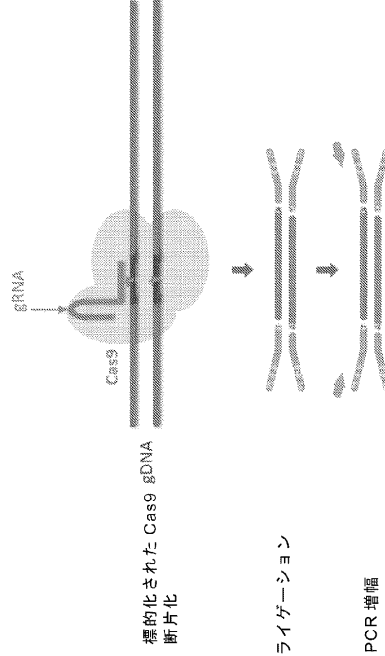
標的化された Cas9 断片化

ライゲーション

PCR 増幅

FIG. 5E

【 図 5 F 】



標的化された Cas9 gDNA 断片化

ライゲーション

PCR 増幅

FIG. 5F

20

30

40

50

【 図 6 】

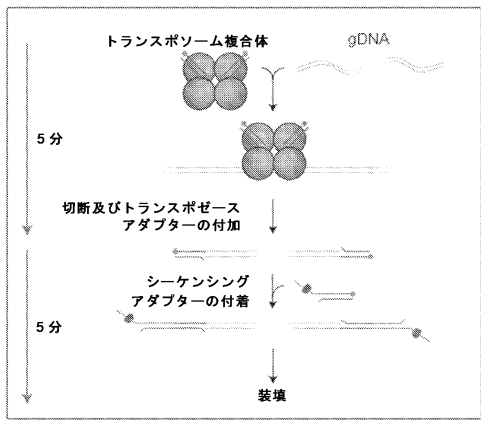


FIG. 6

【 図 7 A 】

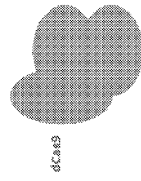
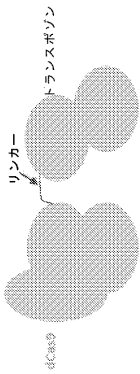


FIG. 7A

10

20

【 図 7 B 】



【 図 7 C 】

FIG. 7B

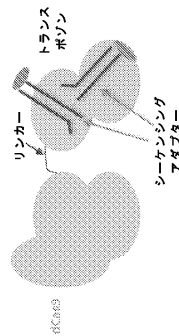


FIG. 7C

30

40

50

【 図 7 D 】

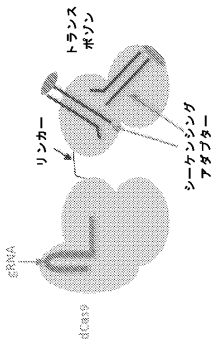


FIG. 7D

【 図 7 E 】

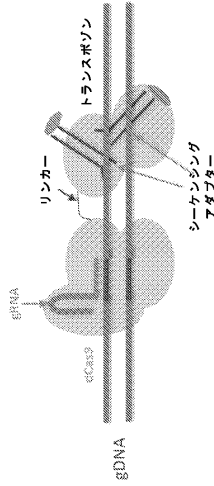


FIG. 7E

10

【 図 7 F 】



FIG. 7F

【 図 7 G 】

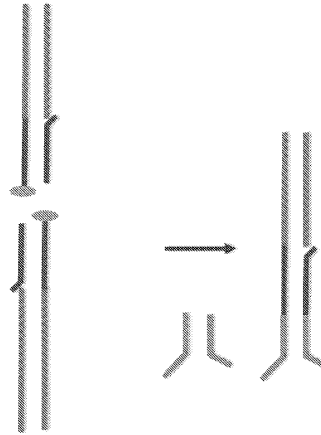


FIG. 7G

20

30

40

50

【 図 7 H 】

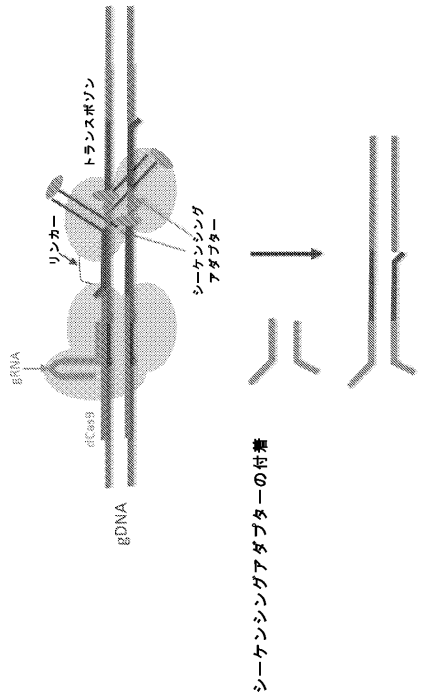


FIG. 7H

【 図 8 】

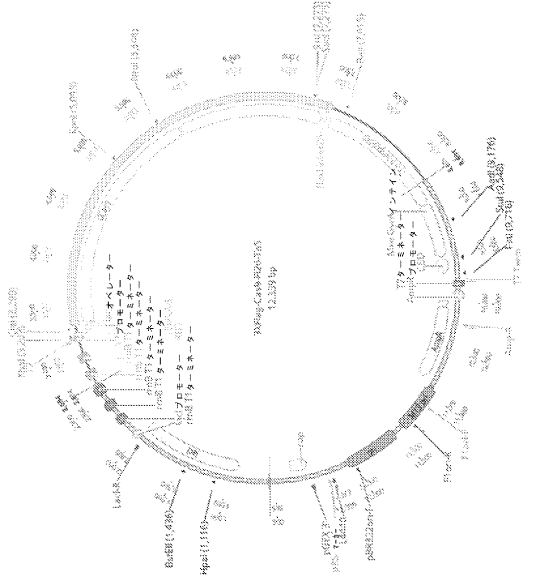


FIG. 8

【 図 9 】

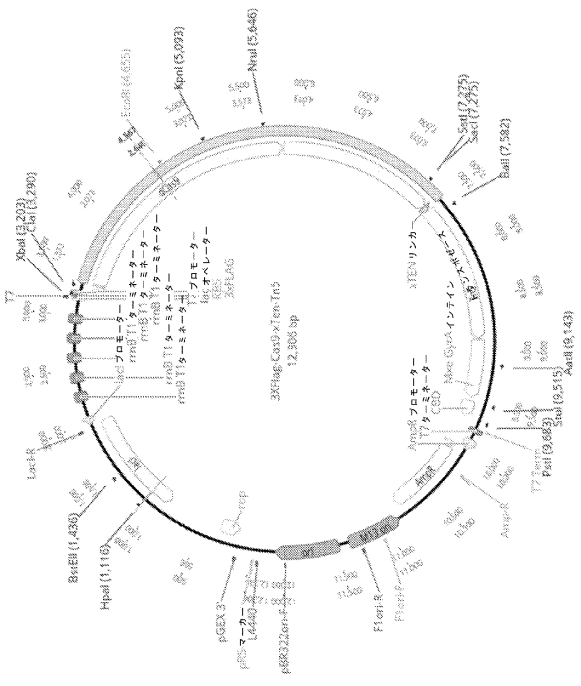


FIG. 9

【 図 10 】

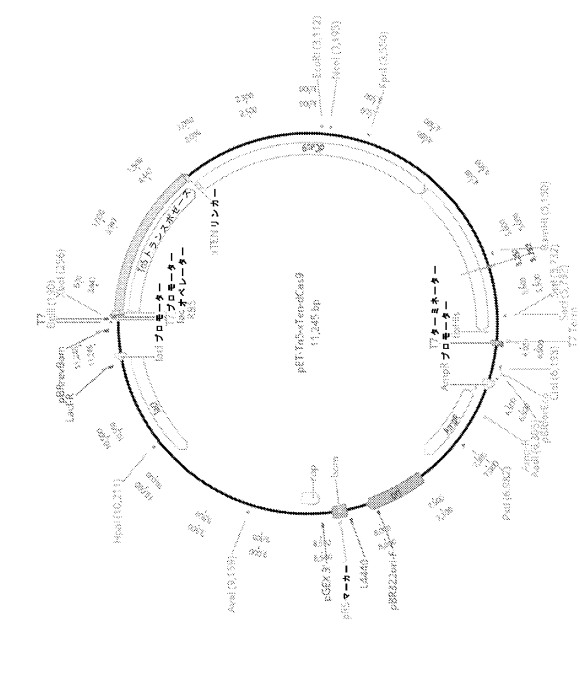


FIG. 10

10

20

30

40

50

【 図 1 1 】



FIG. 11

【 図 1 2 】



FIG. 12

【 図 1 3 】

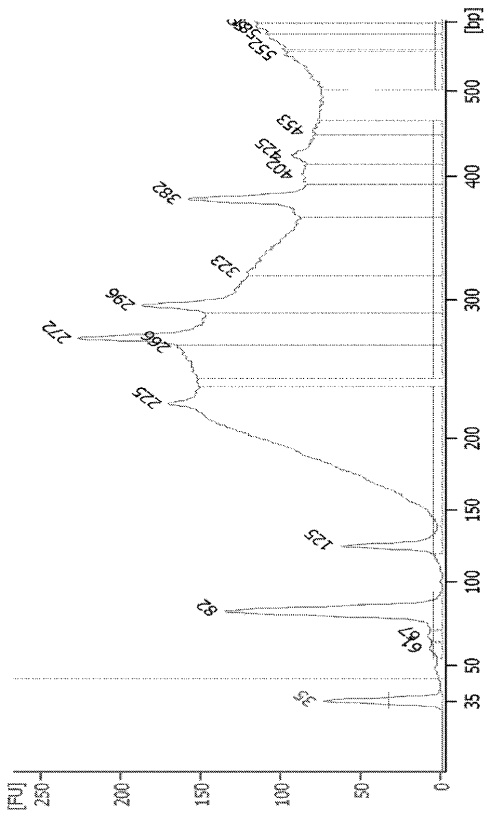


FIG. 13

【 図 1 4 】

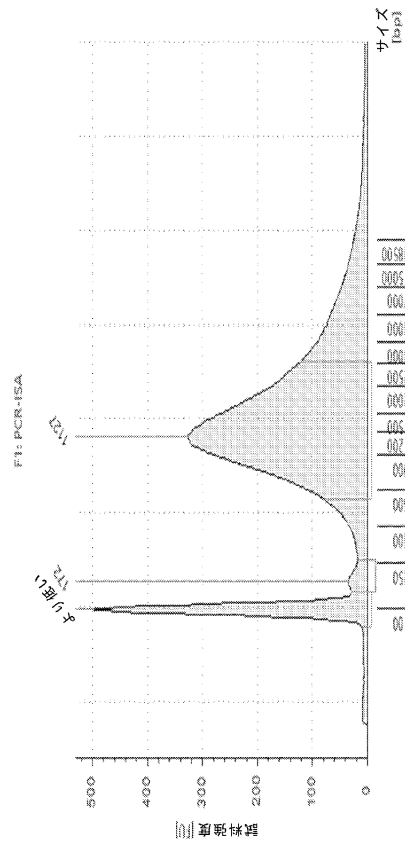


FIG. 14

10

20

30

40

50

【 図 1 5 】

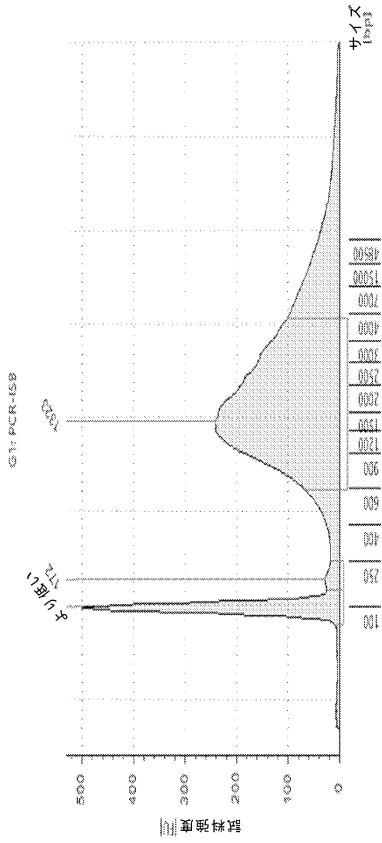


FIG. 15

【 図 1 6 】

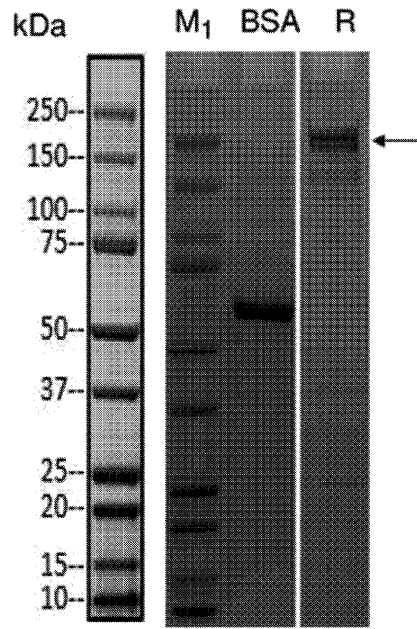


FIG. 16

10

20

【 図 1 7 】

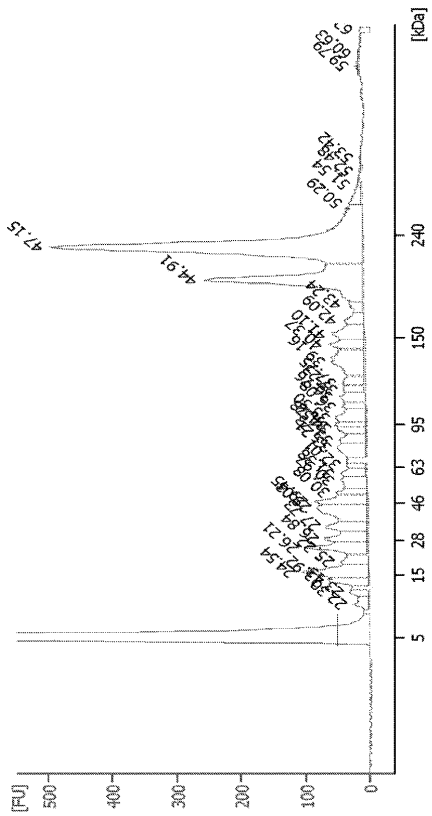


FIG. 17

【 図 1 8 】

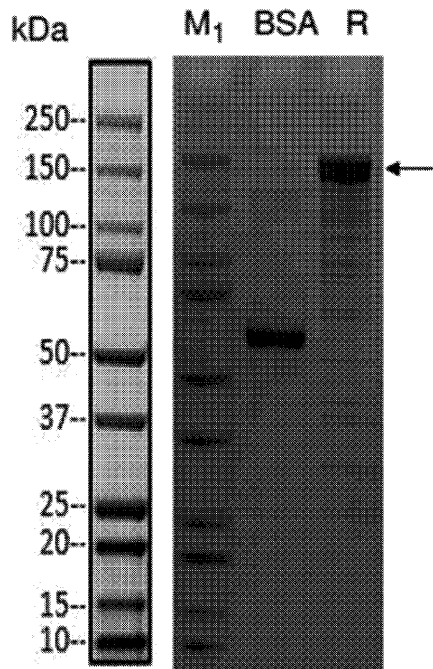


FIG. 18

30

40

50

【 図 19 】

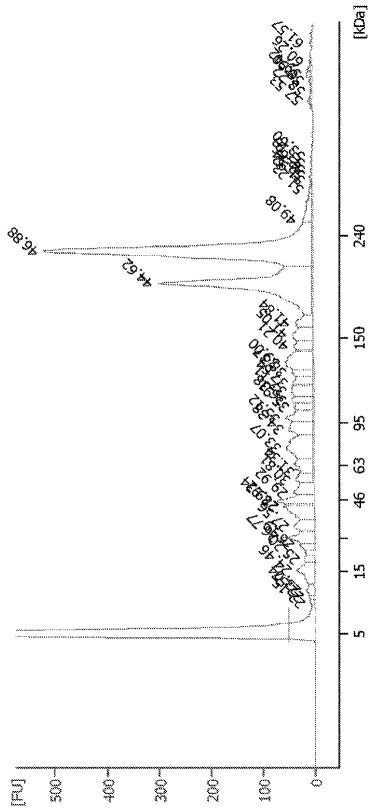


FIG. 19

【 図 20 】

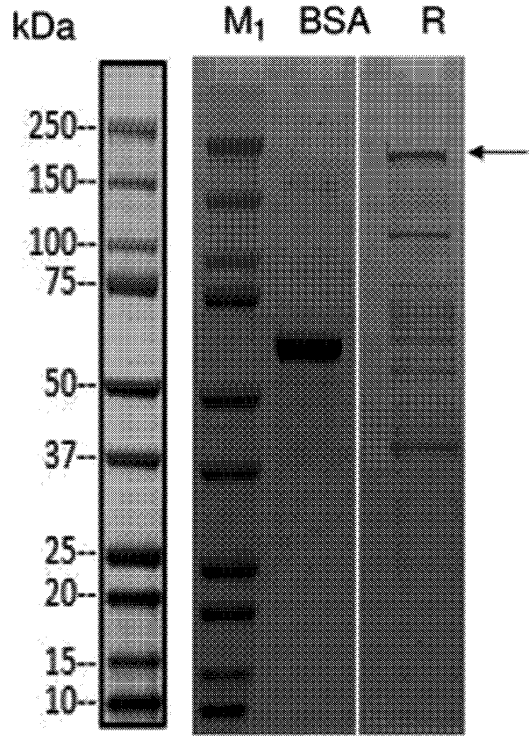


FIG. 20

10

20

【 図 21 】

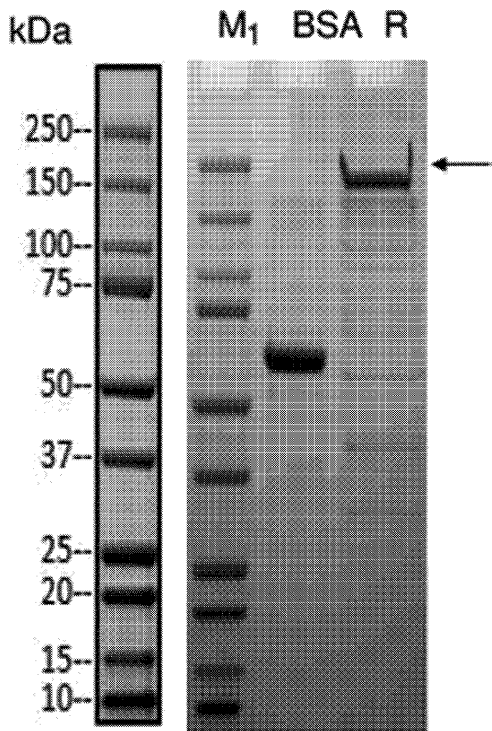


FIG. 21

【 図 22 】

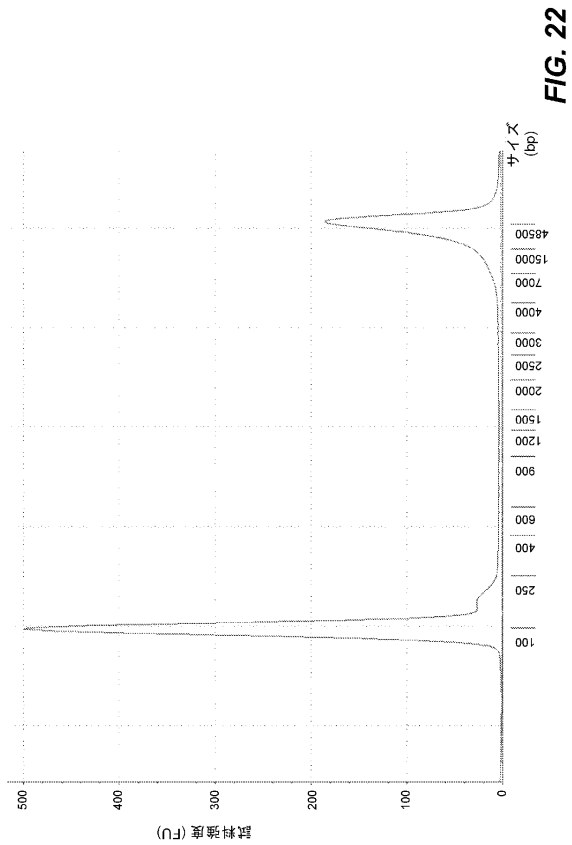


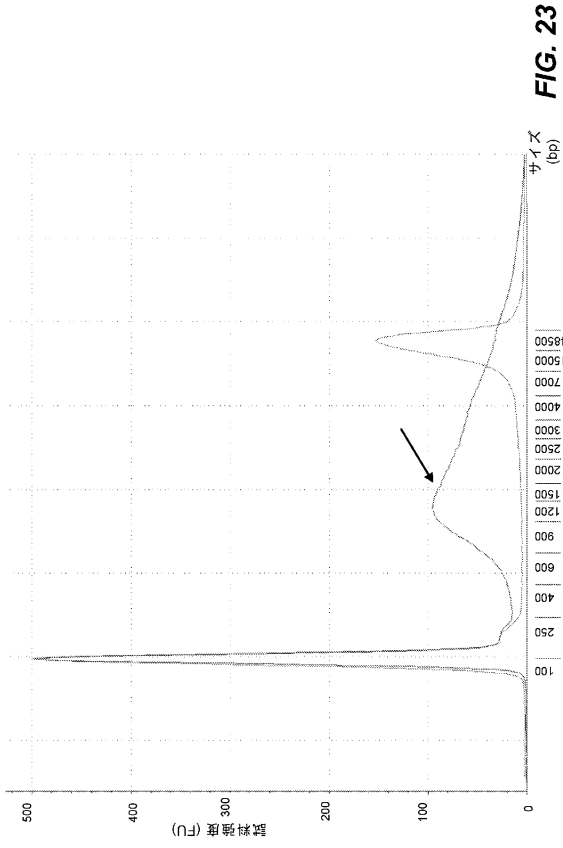
FIG. 22

30

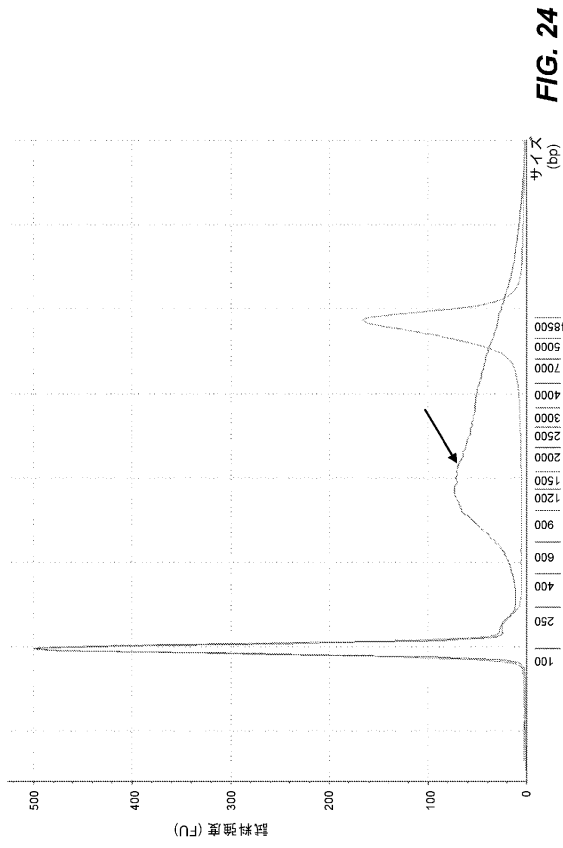
40

50

【 図 2 3 】



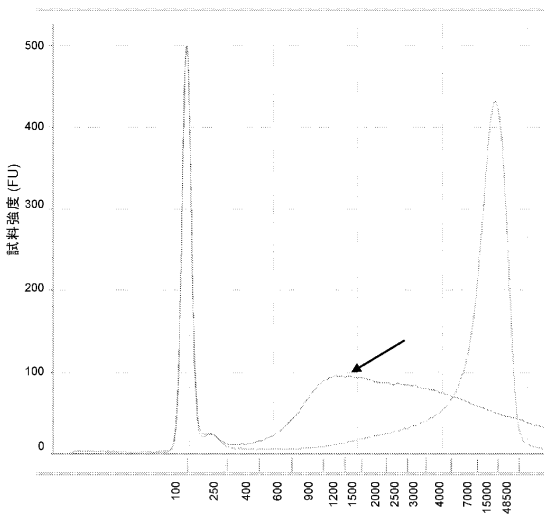
【 図 2 4 】



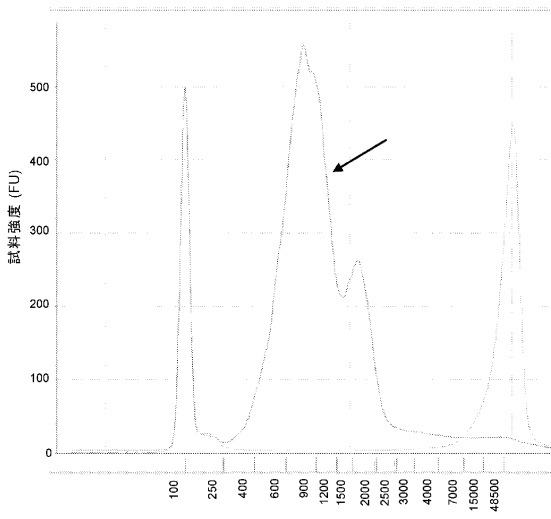
10

20

【 図 2 5 】



【 図 2 6 】



30

40

FIG. 25

FIG. 26

【 図 2 7 】

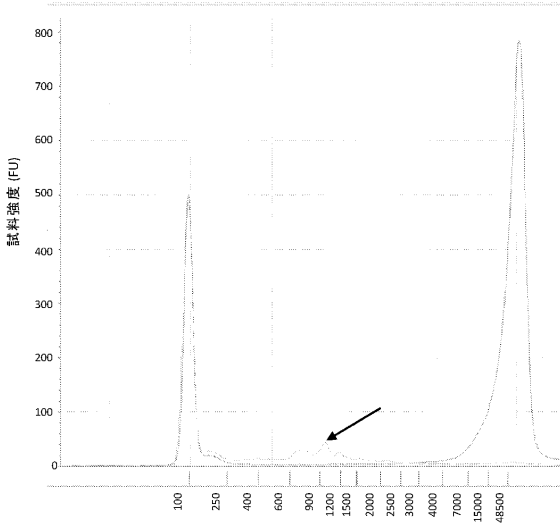


FIG. 27

【 図 2 8 】

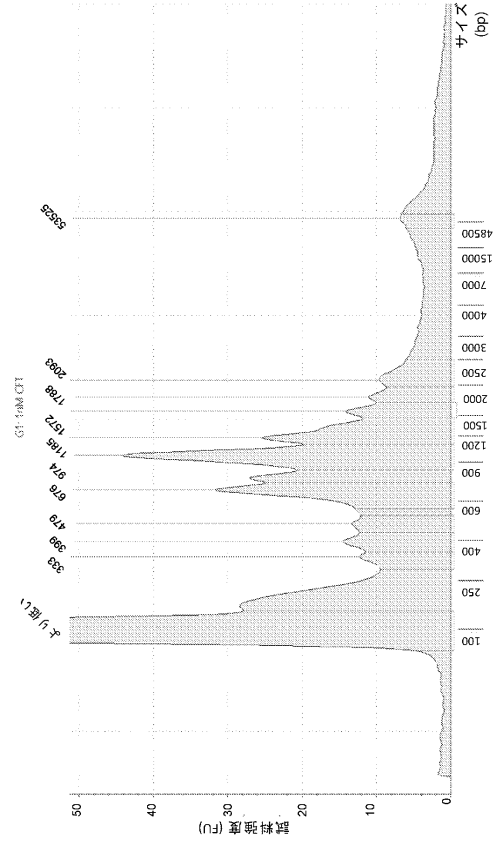


FIG. 28

10

20

【 図 2 9 】

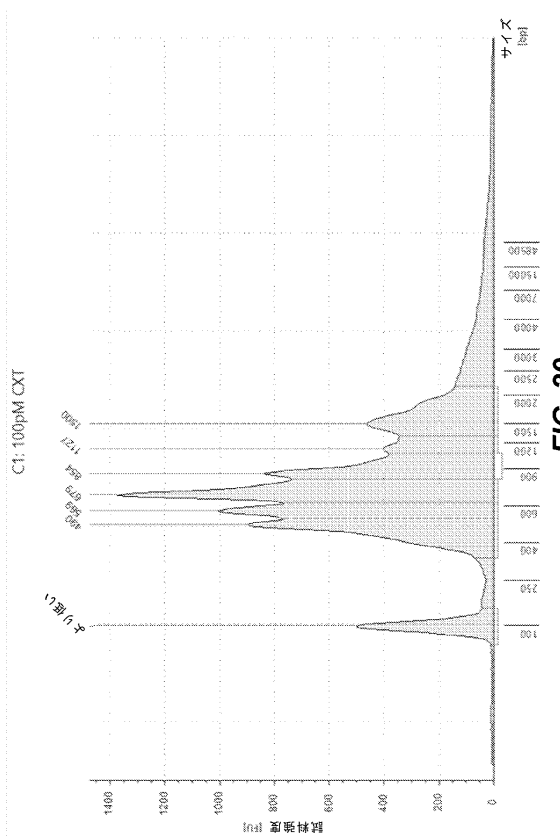


FIG. 29

30

40

【 図 3 0 】

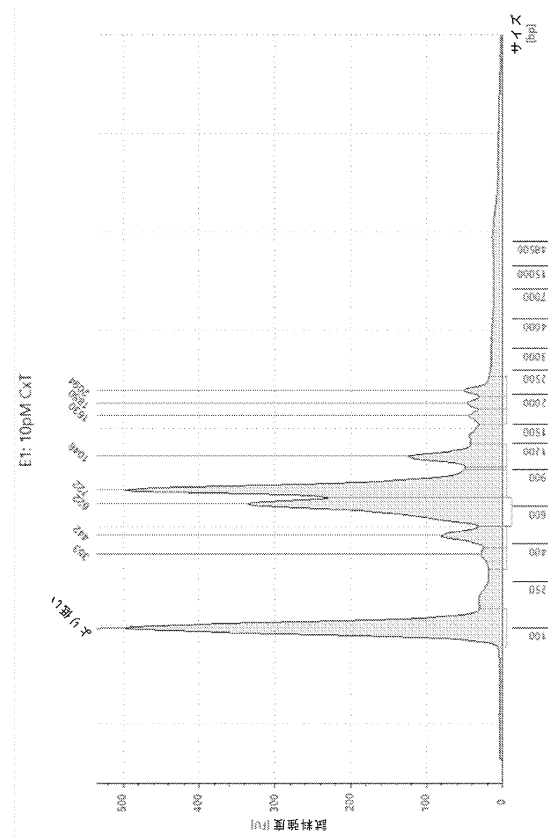


FIG. 30

50

【 図 3 1 】

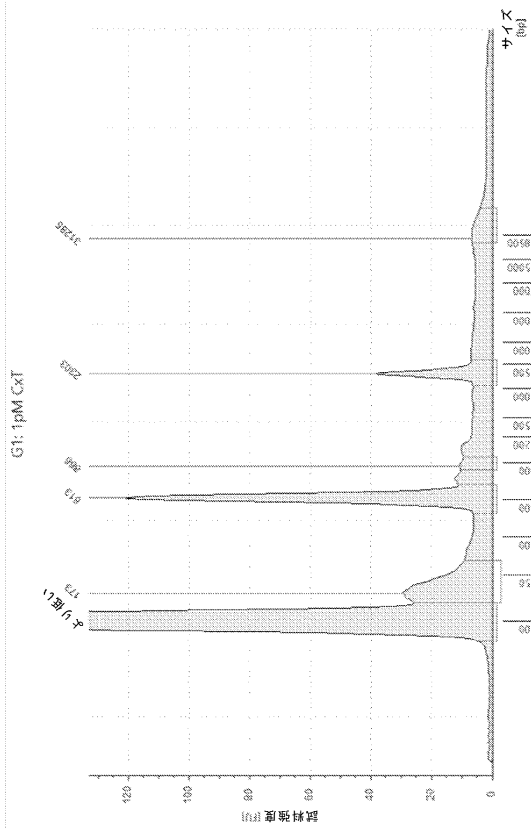


FIG. 31

【 図 3 2 】

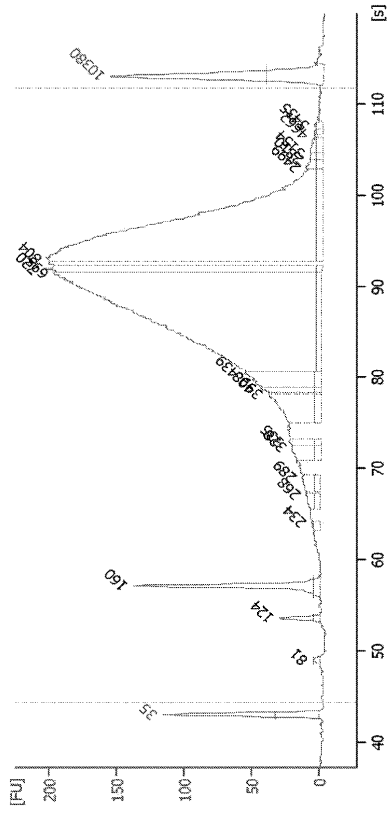


FIG. 32

【 図 3 3 】

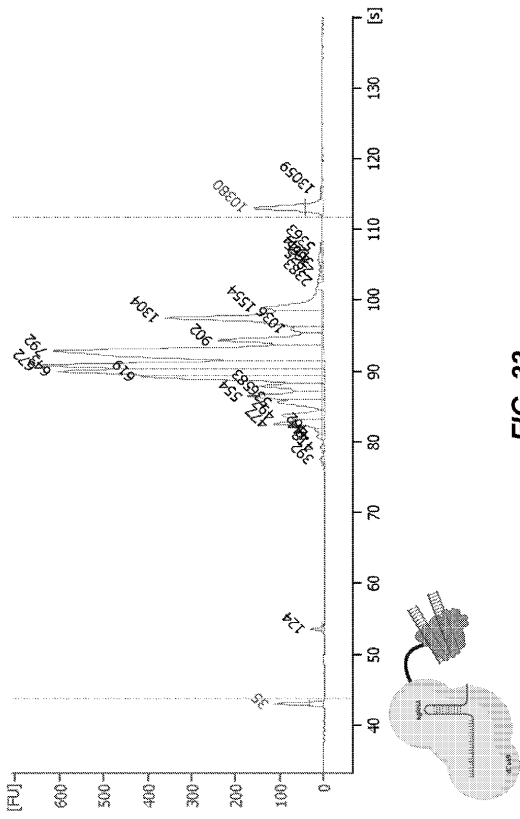


FIG. 33

【 図 3 4 】

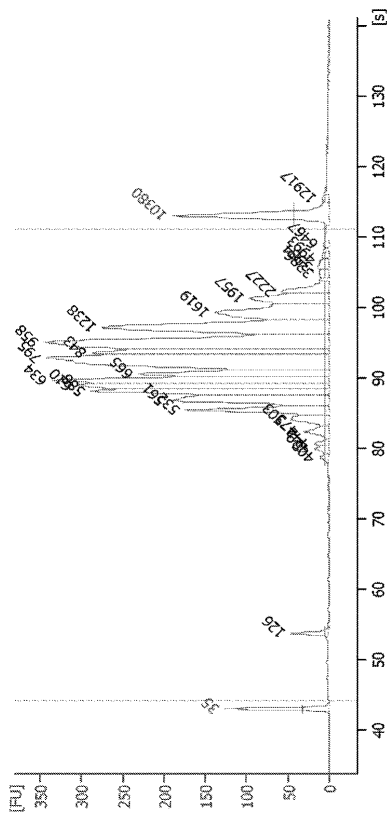


FIG. 34

【 図 3 5 】

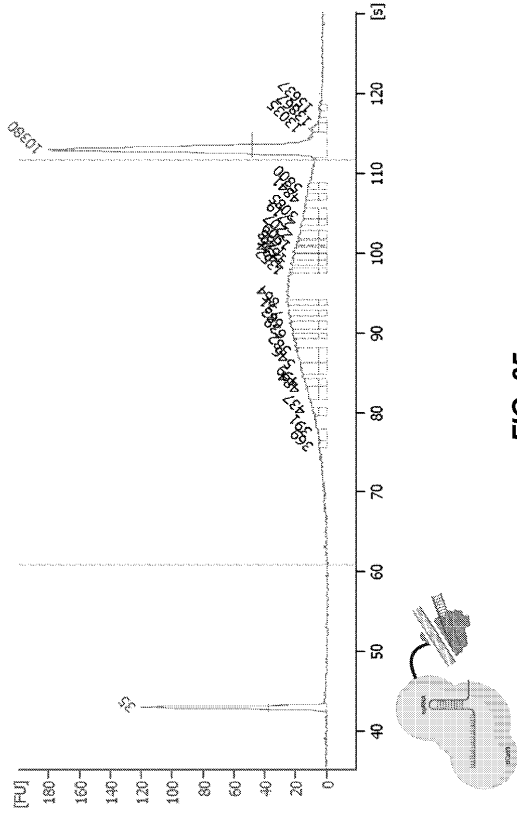


FIG. 35

【 図 3 6 】

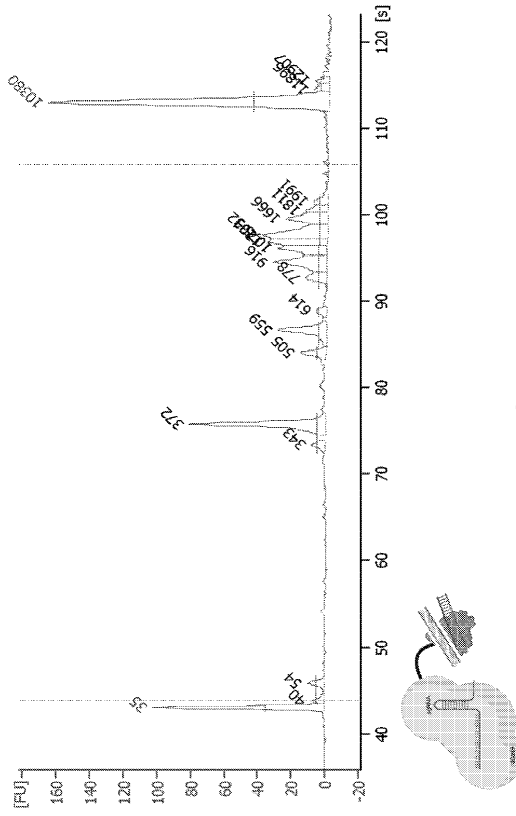


FIG. 36

【 図 3 7 】

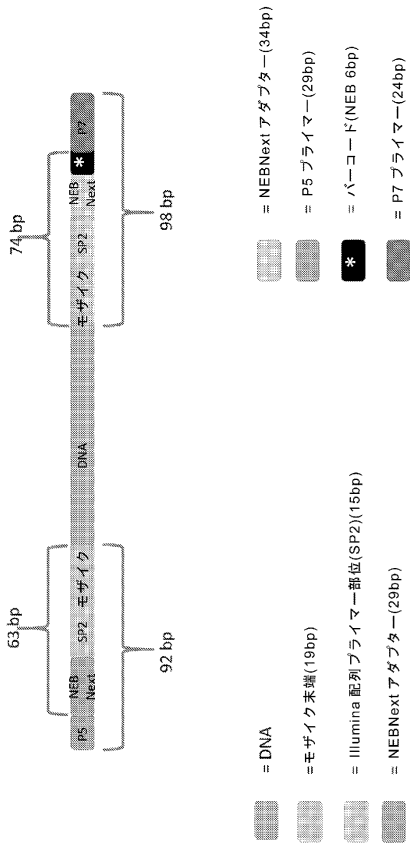


FIG. 37

【 図 3 8 】

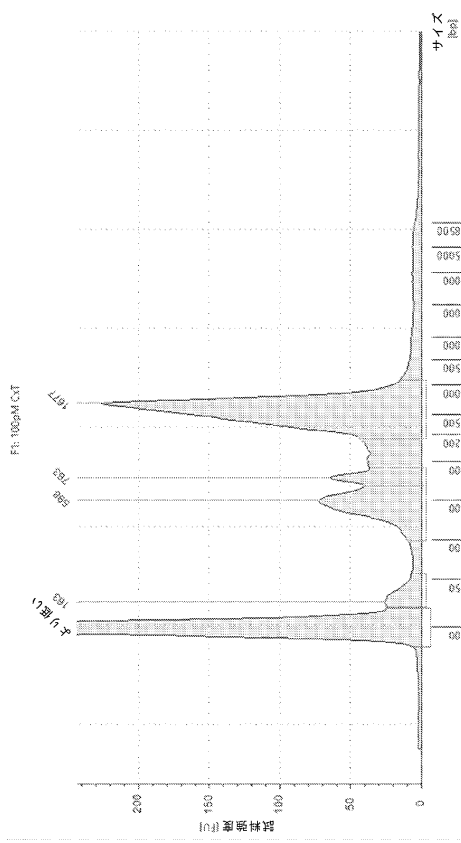


FIG. 38

10

20

30

40

50

【 図 39 】

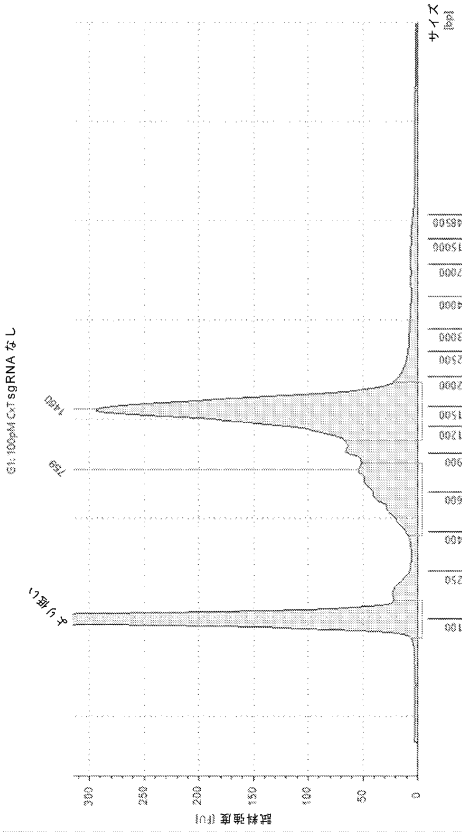


FIG. 39

【 図 40 】

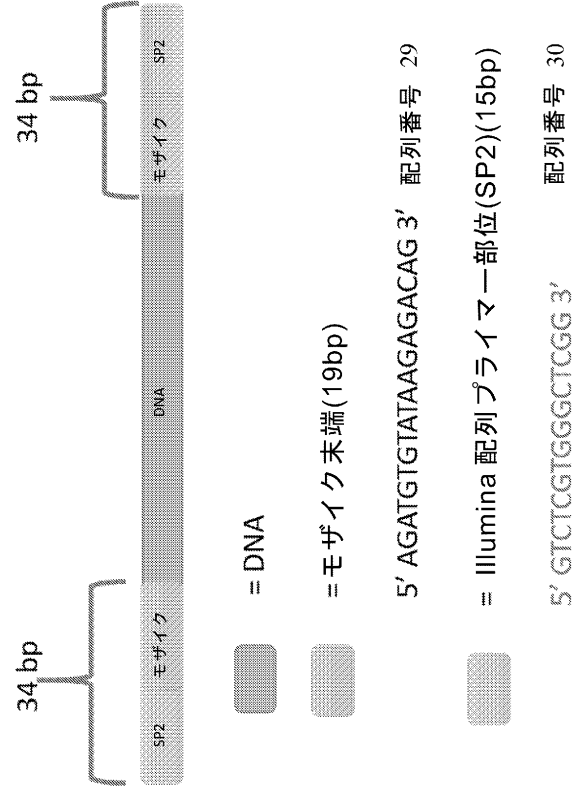


FIG. 40

10

20

【 図 41 】

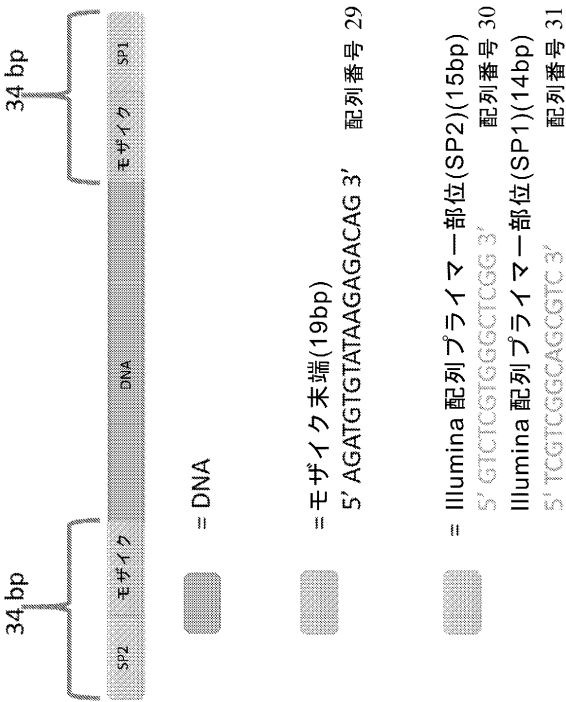
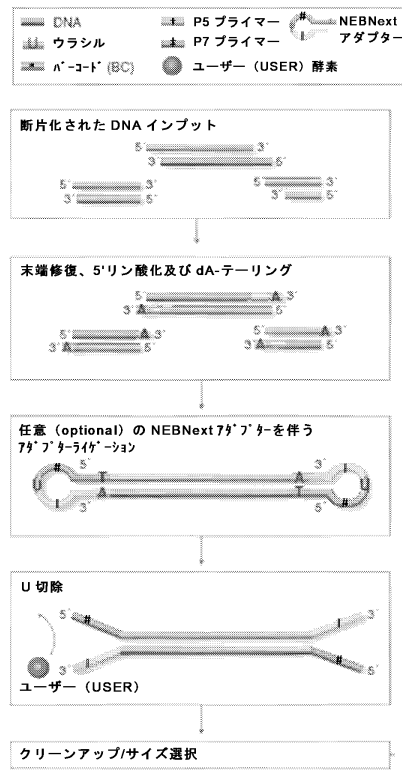


FIG. 41

【 図 42 A 】



30

40

50

【 図 4 2 B 】

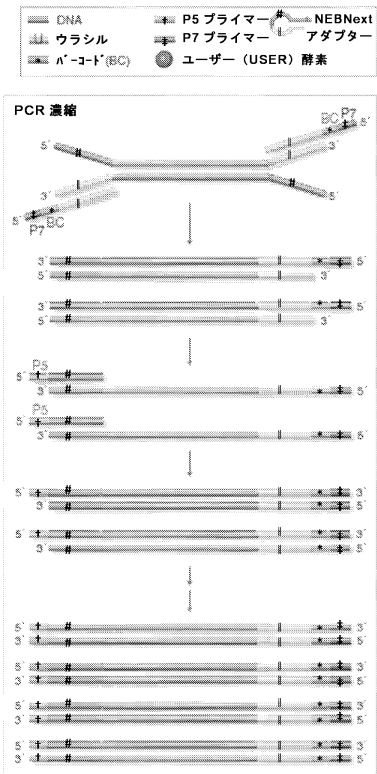


FIG. 42B

【 図 4 3 】

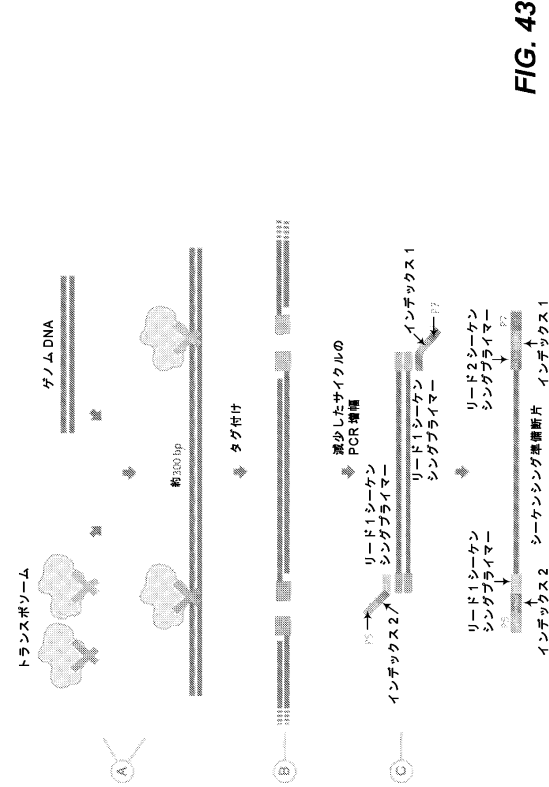


FIG. 43

10

20

【 図 4 4 】

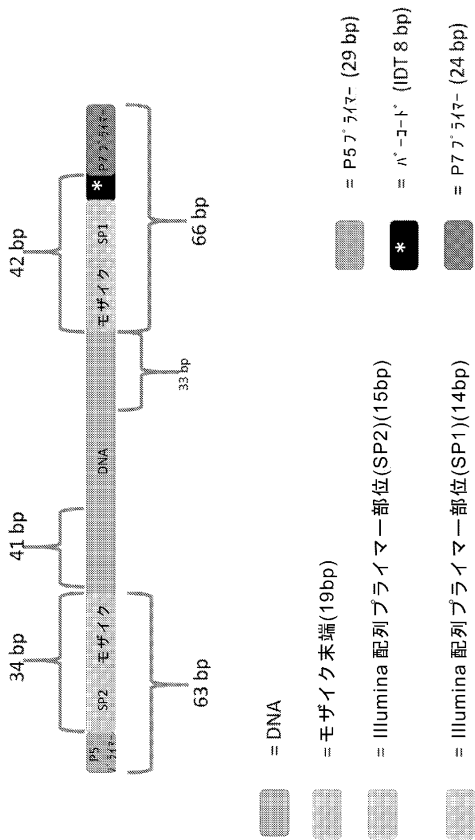


FIG. 44

30

40

50

【配列表】

2024518095000001.app

10

20

30

40

50

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2022/029057

A. CLASSIFICATION OF SUBJECT MATTER
 IPC(8) - C12N 9/12; C12N 9/22; C12N 15/10; C12N 15/11 (2022.01)
 CPC - C12N 15/1065; C12N 9/22; C12N 9/1241; C12N 15/111; C07K 2319/80; C12N 2310/20 (2022.05)

According to International Patent Classification (IPC) or to both national classification and IPC

10

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
 See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
 See Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
 See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

20

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2021/041922 A1 (THE BROAD INSTITUTE INC. et al) 04 March 2021 (04.03.2021) entire document	1-4
A	US 2020/0180487 A1 (THE BROAD INSTITUTE INC. et al) 18 June 2020 (18.06.2020) entire document	1-4
A	WO 2020/243085 A1 (THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK) 03 December 2020 (03.12.2020) entire document	1-4
P, X	WO 2022/040176 A1 (ILLUMINA INC. et al) 24 February 2022 (24.02.2022) entire document	1-4
P, X	WO 2021/257997 A3 (THE BROAD INSTITUTE INC. et al) 10 February 2022 (10.02.2022) entire document	1-4

30

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"D" document cited by the applicant in the international application

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

40

Date of the actual completion of the international search 12 July 2022	Date of mailing of the international search report JUL 29 2022
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-8300	Authorized officer Taina Matos Telephone No. PCT Helpdesk: 571-272-4300

50

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2022/029057

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item I.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

- a. forming part of the international application as filed:
 - in the form of an Annex C/ST.25 text file.
 - on paper or in the form of an image file.
- b. furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
- c. furnished subsequent to the international filing date for the purposes of international search only:
 - in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).
 - on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).

2. In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2022/029057

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

- 1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely: 10

- 2. Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

- 3. Claims Nos.: 5-70
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a). 20

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

- 1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims. 30
- 2. As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
- 3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
- 4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

- Remark on Protest
- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
 - The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
 - No protest accompanied the payment of additional search fees.

10

20

30

40

50

フロントページの続き

(51)国際特許分類

C 1 2 Q 1/689(2018.01)

F I

C 1 2 Q

1/689

Z

テーマコード(参考)

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

(特許庁注：以下のものは登録商標)

1 . T R I T O N

2 . T W E E N

弁理士 山崎 一夫

(74)代理人 100123777

弁理士 市川 さつき

(74)代理人 100111796

弁理士 服部 博信

(74)代理人 100123766

弁理士 松田 七重

(74)代理人 100168631

弁理士 佐々木 康匡

(72)発明者 ゴディネス アルヴァロ

アメリカ合衆国 カリフォルニア州 9 5 1 3 1 サンノゼ クム ドライヴ 2 3 5 0

F ターム(参考) 4B063 QA13 QA18 QA19 QQ03 QQ08 QQ18 QQ42 QR08 QR42 QR55

QR62 QS24 QS34 QX01