

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6800749号
(P6800749)

(45) 発行日 令和2年12月23日 (2020. 12. 23)

(24) 登録日 令和2年11月27日 (2020. 11. 27)

(51) Int. Cl.	F I
C 1 2 Q 1/6869 (2018. 01)	C 1 2 Q 1/6869 Z N A Z
C 1 2 Q 1/34 (2006. 01)	C 1 2 Q 1/34
C 1 2 N 9/14 (2006. 01)	C 1 2 N 9/14
G O 1 N 27/00 (2006. 01)	G O 1 N 27/00 Z
G O 1 N 27/04 (2006. 01)	G O 1 N 27/04 Z
請求項の数 27 (全 112 頁) 最終頁に続く	

(21) 出願番号	特願2016-521718 (P2016-521718)	(73) 特許権者	500358711
(86) (22) 出願日	平成26年11月26日 (2014. 11. 26)		イルミナ インコーポレイテッド
(65) 公表番号	特表2017-503473 (P2017-503473A)		アメリカ合衆国 カリフォルニア州 92
(43) 公表日	平成29年2月2日 (2017. 2. 2)		122 サンディエゴ イルミナ ウェイ
(86) 国際出願番号	PCT/US2014/067582		5200
(87) 国際公開番号	W02015/081178	(74) 代理人	100097456
(87) 国際公開日	平成27年6月4日 (2015. 6. 4)		弁理士 石川 徹
審査請求日	平成29年11月22日 (2017. 11. 22)	(72) 発明者	エリク スタバ
(31) 優先権主張番号	61/909, 316		アメリカ合衆国 カリフォルニア州 92
(32) 優先日	平成25年11月26日 (2013. 11. 26)		122 サンディエゴ イルミナ ウェイ
(33) 優先権主張国・地域又は機関	米国 (US)		5200
		(72) 発明者	ジェンス エイチ. ガンドラク
			アメリカ合衆国 ワシントン州 9811
			7 シアトル ノースウッド ロード エ
			ヌダブリュー 13900
			最終頁に続く

(54) 【発明の名称】 ポリヌクレオチド配列決定のための組成物及び方法

(57) 【特許請求の範囲】

【請求項 1】

標的ポリヌクレオチドを特徴決定する方法であって：

(a) Hel308ヘリカーゼ及び該標的ポリヌクレオチドと接触している細孔を横切るポテンシャルの差を印加する工程；

(b) 該Hel308ヘリカーゼの完全転位サイクルの間に該細孔を通して移動する該標的ポリヌクレオチドの1ヌクレオチドにつき2つの異なる大きさのシグナルを測定する工程であって、該2つのシグナルは、2つのフラクショナル転位工程により生成され、各フラクショナル転位工程は、該細孔を通る該ポリヌクレオチドの部分的転位である、前記測定する工程；並びに

(c) b)において測定された、該1ヌクレオチドにつき2つの異なる大きさのシグナルを使用して該標的ポリヌクレオチドを特徴決定する工程：を含む、

該標的ポリヌクレオチドの特徴決定が、該標的ポリヌクレオチドの配列、該標的ポリヌクレオチドの長さ、該標的ポリヌクレオチドの同一性、及び該標的ポリヌクレオチドの給源の1以上を同定することを含む、前記方法。

【請求項 2】

前記標的ポリヌクレオチドの特徴決定が、該標的ポリヌクレオチドの配列を同定することを含む、請求項1記載の方法。

【請求項 3】

前記方法が、前記工程(a)－(c)を1回以上反復することを更に含む、請求項1記載の方法

。

【請求項 4】

細孔を通じた標的ポリヌクレオチドのフラクショナル転位工程を調節する方法であって：

(a) Hel308ヘリカーゼ及び該標的ポリヌクレオチドと接触している細孔を横切るポテンシャルの差を印加する工程；

(b) 該Hel308ヘリカーゼをHel308ヘリカーゼ基質の参照濃度とは異なる濃度の該基質と接触させる工程であって、該基質濃度が、該参照濃度と比べた該基質濃度の差異に比例してフラクショナル転位工程の持続期間に変化を生じる工程；並びに

(c) 該標的ポリヌクレオチドの1ヌクレオチドにつき2つの異なる大きさのシグナルを測定する工程であって、該2つのシグナルは、該細孔を通る該標的ポリヌクレオチドの各ヌクレオチドの該Hel308ヘリカーゼによる2つ以上のフラクショナル転位工程により生成され、フラクショナル転位工程は、該細孔を通るヌクレオチドの部分的転位である、前記測定する工程：を含む、前記方法。

10

【請求項 5】

前記標的ポリヌクレオチドの転位が、該細孔を通じて転位するポリヌクレオチド上のポテンシャルの差により印加された加力の反対方向である、請求項1又は4記載の方法。

【請求項 6】

前記標的ポリヌクレオチドの転位が、該細孔を通じて転位するポリヌクレオチド上のポテンシャルの差により印加された加力の方向である、請求項1又は4記載の方法。

20

【請求項 7】

前記シグナルが電氣的シグナルである、請求項1又は4記載の方法。

【請求項 8】

前記電氣的シグナルが、電流、電圧、トンネル電流、抵抗、ポテンシャル、電圧、コンダクタンス、及び横方向電氣的測定値から選択された測定値である、請求項7記載の方法。

【請求項 9】

前記標的ポリヌクレオチドの1以上のヌクレオチド残基が、完全転位サイクルから得られた単独の電氣的シグナルを使用する1以上のヌクレオチドの特徴決定と比べ50%より大きい精度で、完全転位サイクルの2つのフラクショナル工程から得られた電氣的シグナルを用いて特徴決定される、請求項8記載の方法。

30

【請求項 10】

前記細孔が、生物学的細孔である、請求項1又は4記載の方法。

【請求項 11】

前記細孔が、固体状態の細孔である、請求項1又は4記載の方法。

【請求項 12】

前記細孔が、生物学的及び固体状態のハイブリッド細孔である、請求項1又は4記載の方法。

【請求項 13】

前記Hel308ヘリカーゼが、以下のGenInfo識別子のいずれか1つによって規定される、請求項1又は4記載の方法：

40

GI:121689265、GI:18202135、GI:490144641、GI:506391664、GI:118576895、GI:18202627、GI:503633371、GI:333911571、GI:240102927、GI:315231273、GI:242398904、GI:121723325、GI:24418451、GI:294495240、GI:24418450、GI:116665562、GI:336477283、GI:298674154、GI:500195255、GI:490388033、GI:226740606、GI:497573451、GI:256810263、GI:18202572、GI:502864579、GI:88603707、GI:635552454、GI:502709689、GI:495257384、及びGI:15791161。

【請求項 14】

前記標的ポリヌクレオチドが、1本鎖ヌクレオチド、2本鎖ヌクレオチド、及び部分的2本鎖ポリヌクレオチドからなる群から選択される、請求項1又は4記載の方法。

50

【請求項15】

前記標的ポリヌクレオチドの特徴決定が、該標的ポリヌクレオチドの配列、該標的ポリヌクレオチドの長さ、該標的ポリヌクレオチドの同一性、及び該ポリヌクレオチドの給源の1以上を同定することを含む、請求項4記載の方法。

【請求項16】

前記ポテンシャルの差が、電気ポテンシャルの差を含む、請求項1又は4記載の方法。

【請求項17】

前記基質濃度が、該Hel308ヘリカーゼ基質のサブ飽和濃度である、請求項4記載の方法。

【請求項18】

前記参照濃度が、該Hel308ヘリカーゼ基質の飽和濃度である、請求項4記載の方法。

【請求項19】

前記基質濃度及び該参照濃度の両方が、該基質の飽和濃度ではない、請求項4記載の方法。

【請求項20】

前記標的ポリヌクレオチドの1以上のヌクレオチド残基が、該参照濃度と比べより低い基質濃度でより大きい精度で特徴決定される、請求項4記載の方法。

【請求項21】

前記特徴決定が、改変されたViterbiアルゴリズムを適用することを含む、請求項1記載の方法。

【請求項22】

前記方法が：

(d)工程(c)の後に、該細孔を通じた該標的ポリヌクレオチドの該Hel308ヘリカーゼによる1以上のフラクショナル転位工程の時機を変動するために、少なくとも1種のパラメータを変動する工程；並びに

(e)変動された少なくとも1種のパラメータを用い、工程(a)－(c)を反復する工程：を更に含む、請求項1記載の方法。

【請求項23】

工程(c)及び(e)の間に発生したシグナルを組合せること、並びに組合せたシグナルを基に該標的ポリヌクレオチドを特徴決定することを更に含む、請求項22記載の方法。

【請求項24】

前記変動された少なくとも1種のパラメータが、温度、塩濃度、補因子濃度、ATP生成物の濃度、pH、及び使用した特定のHel308ヘリカーゼからなる群から選択される、請求項23記載の方法。

【請求項25】

前記特徴決定が、検出され且つ同定されたシグナルの大きさを基に標的ポリヌクレオチドの配列を決定及び出力することを含む、請求項1記載の方法。

【請求項26】

前記1以上のシグナル中のレベルの検出及び同定が、完全転位、フラクショナル転位、全転位からの1以上のシグナルの大きさ、及び、シグナル持続期間の出力を含む、請求項25記載の方法。

【請求項27】

(a)前記検出及び同定されたレベルを基にした標的ポリヌクレオチドの配列の決定及び出力が、前記完全転位、フラクショナル転位、全転位からの1以上のシグナルの大きさ、及び、シグナル持続期間を入力として採用し、該入力を基に複数の配列を呼び出し、並びに呼び出された複数の配列に関する信頼情報を基に呼び出された配列の少なくとも1種を選択及び出力することを含むか、又は

(b)前記検出及び同定されたレベルを基にした標的ポリヌクレオチドの配列の決定及び出力が、前記完全転位、フラクショナル転位、全転位からの1以上のシグナルの大きさ、及び、シグナル持続期間を入力として採用し、該入力を基に複数の配列を呼び出し、並び

10

20

30

40

50

に呼び出された複数の配列の一部に関する信頼情報を基に複数の呼び出された配列の互いの部分を選択及び連鎖させることを含むか、又は

(c)前記検出及び同定されたレベルを基にした標的ポリヌクレオチドの配列の決定及び出力が、前記完全転位、フラクショナル転位、全転位からの1以上のシグナルの大きさ、及び、シグナル持続期間を入力として採用し、該入力に基づき複数の配列を呼び出し、呼び出された配列をモデル配列と比較し、並びに呼び出された配列のモデル配列との比較に関する信頼情報を基に呼び出された配列の少なくとも1種を選択及び出力することを含むか、又は

(d)前記検出及び同定されたレベルを基にした標的ポリヌクレオチドの配列の決定及び出力が、前記完全転位、フラクショナル転位、全転位からの1以上のシグナルの大きさ、及び、シグナル持続期間を入力として採用し、該入力に基づき複数の配列を呼び出し、呼び出された配列をモデル配列と比較し、並びに呼び出された複数の配列の一部のモデル配列との比較に関する信頼情報を基に複数の呼び出された配列の互いの部分を選択及び連鎖することを含む、請求項26記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

(関連出願の相互参照)

本出願は、2013年11月26日に出願され、「ポリヌクレオチド配列決定のための組成物及び方法」と題する、米国特許仮出願第61/909,316号の優先権を主張するものであり、この仮出願の全内容が引用により本明細書中に組み込まれている。

【0002】

(配列表)

本出願は、ASCIIフォーマットで電子提出され、且つその全体が引用により本明細書中に組み込まれている配列表を含む。2014年11月25日に作製された該ASCIIコピーは、12957-139-228_SL.txtと名前をつけ、且つサイズは19,778バイトである。

【背景技術】

【0003】

(背景)

本開示は概して、標的ポリヌクレオチドの配列を特徴決定することを含む、標的ポリヌクレオチドを特徴決定するための方法及び組成物に関する。

【0004】

ポリヌクレオチド(例えば、DNA又はRNA)にコードされた情報は、医薬品及び生命科学に最も重要であるので、迅速且つ安価にポリヌクレオチドを配列決定する必要性が存在している。現在、市販の配列決定技術は、試料及びライブラリーの調製を必要とし、これらは両方共骨が折れる。更に読み出しは、多くの適用に関して望ましいものよりも遅い。そのために、処理量は限定され且つ費用は相対的に高い。ナノポア配列決定は、標的ポリヌクレオチドを迅速且つ安価に配列決定するように開発されたひとつの新規方法を表している。

【0005】

ナノポア配列決定は、イオン電流のためのチャネルを提供することができるナノポアを利用する。ポリヌクレオチドは、ナノポアを通して電気泳動的に駆動され、且つポリヌクレオチドはこのナノポアを通過する際に、ナノポアを通る電流(electrical current)を減少する。各々の通過するヌクレオチド、又は一連のヌクレオチドは、特徴的電流を生じ、且つこの電流レベルの記録は、そのポリヌクレオチドの配列に対応している。一部の電流レベルは、複数のヌクレオチド(一般に3~4個)により支配されるので、精度を向上するために当該技術分野の状態を改善する必要性が依然残っている。形状及び持続期間などのナノポアを通じたポリヌクレオチド転位(translocate)として得られる電流レベルに関する追加情報は、利点を提供することができる。

【0006】

ナノポア配列決定の一般的挑戦は、ナノポアを通じたポリヌクレオチドの転位は、非常に迅速であるので、個々のヌクレオチドの電流レベルは、解明されるには余りにも短いことである。ナノポア配列決定の一つのアプローチは、電位に対する、ヘリカーゼ、トランスロカーゼ、又はポリメラーゼなどのポリヌクレオチド結合タンパク質の誘導下での、ナノポアを通じたポリヌクレオチドの制御された転位に関する。この制御された転位にもかかわらず、数多くの配列決定誤差モードが、依然存在し、且つ配列決定の精度不良の一因となっている。

【0007】

従ってナノポアを通じたポリヌクレオチドの更に制御された転位及びヌクレオチド識別におけるヌクレオチド転位のより良い解像を提供する方法及び組成物の必要性が存在する。本開示は、この必要性を満たし、且つ関連する利点を提供する。

10

【発明の概要】

【0008】

(実施態様の概要)

標的ポリヌクレオチドを特徴決定する方法が提供される。本方法は：(a)Hel308ヘリカーゼ及び標的ポリヌクレオチドと接触している細孔を横切るポテンシャルの差を印加する工程；(b)この細孔を通じた標的ポリヌクレオチドの1以上のフラクショナル転位(fractional translocation)工程により発生された1以上のシグナルを測定する工程；並びに、(c)このフラクショナル転位工程の電氣的シグナルからこの標的ポリヌクレオチドを特徴決定する工程：を含む。標的ポリヌクレオチドの特徴決定は、(1)標的ポリヌクレオチドの配列；(2)標的ポリヌクレオチドの修飾；(3)標的ポリヌクレオチドの長さ；(4)標的ポリヌクレオチドの同一性；(5)標的ポリヌクレオチドの給源、又は、(6)標的ポリヌクレオチドの二次構造：の1以上を同定することを含む。細孔を通じた標的ポリヌクレオチドのフラクショナル転位工程を調節する方法、並びに1mM未満のATP溶液又はヌクレオチド類縁体溶液中含有された細孔、Hel308ヘリカーゼ及び標的ポリヌクレオチドを含む、標的ポリヌクレオチドを特徴決定するための組成物も提供される。

20

【図面の簡単な説明】

【0009】

(図面の簡単な説明)

【図1A】図1Aは、ヘリカーゼによるポリヌクレオチドの転位に関する静電氣的コマ送り(inchworm)モデルを示している。

30

【0010】

【図1B】図1Bは、一部の実施態様に従う、Hel308ヘリカーゼと接触している細孔を含む第一の例証的組成物を概略的に図示している。

【0011】

【図1C】図1Cは、一部の実施態様に従う、標的ポリヌクレオチドを特徴決定する例証的方法の工程を概略的に図示している。

【0012】

【図2A】図2Aは、一部の実施態様に従う、Phi29ポリメラーゼ及びHel308 Tgaヘリカーゼの転位事象の比較を示す。Hel308 Tgaヘリカーゼにより認められたフラクショナル転位工程は、phi29 DNAポリメラーゼにより認められた転位工程と比較して示している。

40

【0013】

【図2B】図2Bは、一部の実施態様に従う、Phi29ポリメラーゼ及びHel308 Tgaヘリカーゼの転位事象の比較を示す。Hel308 Tgaヘリカーゼにより認められたフラクショナル転位工程は、分子モーターとしてPhi29ポリメラーゼを使用するMspA-M2ナノポアを通じて転位する1本鎖ポリヌクレオチド鑄型により生じた予測電流レベルを、分子モーターとしてHel308 Tgaヘリカーゼを使用し認められるものと比較して示している。

【0014】

【図2C】図2Cは、一部の実施態様に従う、Phi29ポリメラーゼ及びHel308 Tgaヘリカーゼの転位事象の比較を示す。Hel308 Tgaヘリカーゼにより認められたフラクショナル転位

50

工程は、単純な反復ヌクレオチド配列(配列番号:74)に関してphi29 DNAポリメラーゼによる認められた転位工程と比較して示している。

【0015】

【図3】図3は、一部の実施態様に従う、フラクショナル転位工程に関する提唱された「グリップ-ベース」の機序を示している。

【0016】

【図4】図4A及び4Bは、一部の実施態様に従う、フラクショナル転位工程の滞在時間に対するATP濃度の例証的作用を示している。

【0017】

【図5】図5は、一部の実施態様に従う、追加ノイズの様々なレベルによりインシリコで発生した電流トレース(以下に説明)の完全工程(菱形)及び1/2工程(四角形)に関する配列決定の再構築精度(隠れマルコフモデル(HMM))をプロットしている。

10

【0018】

【図6A】図6Aは、一部の実施態様に従う、そこでポリヌクレオチドがモーター酵素により移動されるナノポアにおいて配列を解読するためのHMMに関して必要な非-ゼロ確率を伴う状態移行を描いている。このモーターは、phi29 DNAP又は1ヌクレオチド工程でポリヌクレオチドを移動する類似の酵素である。

【0019】

【図6B】図6Bは、一部の実施態様に従う、ポリヌクレオチドがモーター酵素により移動されるナノポアにおいて配列を解読するためのHMMに関して必要な非-ゼロ確率を伴う状態移行を描いている。このモーターは、Hel308ヘリカーゼ又はポリマーのフラクショナル移動を可能にする類似の酵素である。

20

【0020】

【図7】図7は、一部の実施態様に従う、ガウスシフトの関数としての電流パターンを見つける予期される精度を描いている。菱形は、完全ヌクレオチド工程を有するモーターを描いている。円形は、フラクショナル転位工程を有するモーターを描き、並びに四角形は、持続期間値と組合せたフラクショナル転位工程を有するモーターを描いている。

【0021】

【図8】図8は、一部の実施態様に従う、ピロリン酸の変動する濃度による、Hel308ヘリカーゼ活性の例証的調節を示している。

30

【0022】

【図9】図9は、一部の実施態様に従う、ヌクレオチドインヒビターであるオルトバナジウム酸ナトリウムによる、及びヌクレオチド類縁体であるアデノシン5'-(β , γ -イミド)三リン酸リチウム塩水和物による、Hel308ヘリカーゼ活性の例証的調節を示している。

【0023】

【図10】図10は、一部の実施態様に従う、レベル及びレベル持続期間を使用し、2つの独立した配列読み値から得ることができる、追加的フラクショナル転位工程により提供される情報を使用する方法の例を描いている。

【0024】

【図11】図11は、一部の実施態様に従う、レベル及びレベル持続期間を使用し、2つの同時実行の配列読み値から得ることができる、追加的フラクショナル転位工程により提供される情報を使用する方法の例を描いている。

40

【0025】

【図12】図12は、一部の実施態様に従う、持続期間情報を伴う又は伴わない、電流トレースを使用する、追加的フラクショナル転位工程により提供される情報を使用する方法の例を描いている。

【0026】

【図13】図13A-13Eは、一部の実施態様に従う、Hel308ヘリカーゼ3'側オーバーハング結合部位及びコレステロール二重層アンカーを伴う、3つ組ポリヌクレオチド複合体を基にした、Hel308ヘリカーゼにより制御されたポリヌクレオチド転位を示している。黒丸(

50

●)は、5'側リン酸を表す。黒菱形(◆)は、3'側コレステロールを表す。切り欠けのある塗りつぶしの半透明の円形は、Hel308ヘリカーゼを表す。点線は、任意の長さを示す。大きい灰色矢印は、ポリヌクレオチドの細孔の内外へのポリヌクレオチド移動の方向(印加された電場により又はこれに対抗して)を表す。大きい黒色矢印は、3'側から5'側への、ポリヌクレオチドに沿ったヘリカーゼ転位の方向を示す。細孔(漏斗状の円錐物)は、膜(2本の水平線)内に位置する。

【0027】

【図14】図14A-14Dは、一部の実施態様に従う、Hel308ヘリカーゼ3'側オーバーハング結合部位及びコレステロール二重層アンカーを伴う、3つ組ポリヌクレオチド複合体を基にした、Hel308ヘリカーゼにより制御されたポリヌクレオチド転位を示している。黒丸(●)は、5'側リン酸を表す。黒菱形(◆)は、3'側コレステロールを表す。切り欠けのある塗りつぶしの半透明の円形は、Hel308ヘリカーゼを表す。点線は、任意の長さを示す。大きい灰色矢印は、ポリヌクレオチドの細孔の内外へのポリヌクレオチド移動の方向(印加された電場により又はこれに対抗して)を表す。大きい黒色矢印は、3'側から5'側への、ポリヌクレオチドに沿ったヘリカーゼ転位の方向を示す。細孔(漏斗状の円錐物)は、膜(2本の水平線)内に位置する。記号は、図13A-13Eと同じである。このスキームにおいて、任意のコレステロール部分に結合するか又はまたこれを含むために、Hel308ヘリカーゼに関してポリヌクレオチド“ii”上に3'側オーバーハングを作出する単独のハイブリダイゼーションポリヌクレオチド“i”が存在する。

【0028】

【図15】図15A-15Cは、一部の実施態様に従う、勾配力と同じ方向の制御された転位を示している。切り欠けのある塗りつぶしの半透明の円形は、Hel308ヘリカーゼを表す。点線は、任意の長さを示す。大きい灰色矢印は、細孔の内側への印加された電場によるポリヌクレオチドの移動の方向を表す。大きい黒色矢印は、3'側から5'側への、ポリヌクレオチドに沿ったヘリカーゼ転位の方向を示す。細孔(漏斗状の円錐物)は、膜(2本の水平線)内に位置する。

【0029】

【図16】図16は、SF2ファミリーにおいて同定されている様々なモチーフ(出現する順番に、各々、配列番号:75-81)、例えばHel308が一員であるDEAD-ボックス(配列番号:2)ヘリカーゼなどを概略的に図示している(Tutejaらの文献「DNAヘリカーゼの解明:モチーフ、構造、機序及び機能(Unraveling DNA Helicases: Motif, structure, mechanism and function)」、European Journal of Biochemistry 271(10): 1849-1863 (2004)から改作)。

【0030】

【図17】図17A-17Dは、一部の実施態様に従う、特定のパラメータを使用する、Hel308 Mbuヘリカーゼ、Hel308 Tgaヘリカーゼ、及びphi29ポリメラーゼの転位事象により発生した例証的シグナルを示している。

【0031】

【図18】図18は、一部の実施態様に従う、ポリヌクレオチドバーコードを特徴決定するためにフラクショナル転位を使用するアッセイを実行するための例証的方法の工程を概略的に図示している。

【0032】

【図19A】図19Aは、細孔を通じたポリヌクレオチドの単-工程転位からのシグナルを特徴決定するために使用される例証的隠れマルコフモデル(HMM)の態様を概略的に図示している。

【0033】

【図19B】図19Bは、一部の実施態様に従う、Hel308ヘリカーゼを使用する、細孔を通じたポリヌクレオチドのフラクショナル工程転位からのシグナルを特徴決定するために使用される例証的HMMの態様を概略的に図示している。

【0034】

10

20

30

40

50

【図20A】図20Aは、一部の実施態様に従う、フラクショナル工程を使用する、デノボ配列決定の例証的結果を図示している。

【0035】

【図20B-C】図20B-20Cは、一部の実施態様に従う、フラクショナル工程を使用する、パターンマッチングの例証的結果を図示している。

【0036】

【図21】図21A-21Cは、一部の実施態様に従う、細孔を通じた、ポリヌクレオチドの異なる転位に関する時間の関数としての発生し得るシグナルを概略的に図示している。

【0037】

【図22】図22A-22Dは、一部の実施態様に従う、細孔を通じたポリヌクレオチドのフラクショナル転位により提供される情報を使用する説明的方法の工程を図示している。

10

【0038】

【図23】図23は、一部の実施態様に従う、それぞれのバーコードとして使用するのに適した、第一の説明的ポリヌクレオチド配列(配列番号:89)及び第二の説明的ポリヌクレオチド配列(配列番号:90)に関する、時間の関数としての発生し得る例証的なシミュレーションされたシグナルを図示している。

【0039】

【図24】図24A-24Dは、一部の実施態様に従う、それぞれのバーコードとして使用するのに適した、第一及び第二の説明的ポリヌクレオチド配列に関する、時間の関数としての発生し得る例証的なシミュレーションされたシグナルを図示している。

20

【0040】

【図25】図25A及び25Bは、各々、一部の実施態様に従う、それぞれのバーコードとして使用するのに適した、第一及び第二の説明的ポリヌクレオチド配列に関する、時間の関数としての発生し得る例証的なシミュレーションされたシグナルを図示している。

【0041】

【図26】図26A-26Dは、各々、一部の実施態様に従う、それぞれのバーコードとして使用するのに適した、第一及び第二の説明的ポリヌクレオチド配列に関する、時間の関数としての発生し得る例証的な測定されたシグナルを図示している。

【発明を実施するための形態】

【0042】

30

(実施態様の詳細な説明)

本開示は、細孔を通じた標的ポリヌクレオチドの転位の1以上のフラクショナル転位工程を使用する標的ポリヌクレオチドの配列の特徴決定を含む、標的ポリヌクレオチドを特徴決定する方法及び組成物を提供する。

【0043】

ナノポア配列決定技術の開発において、ナノポアを通じたポリヌクレオチドの制御された転位の特定のレベルは、電気ポテンシャルの差に対する(例えば、それにより発生した力に抵抗するように)ヘリカーゼ、トランスロカーゼ、又はポリメラーゼなどの分子モーターの誘導下で達成され得る。分子モーターは、通常1工程につき1以上のヌクレオチドによる段階的様式でポリヌクレオチドを移動することができる。この制御されたラチェット方式は、天然の速度の μ 秒/ヌクレオチドからミリ秒/ヌクレオチドまで、ナノポアを通じたポリヌクレオチド転位を遅らせる。

40

【0044】

分子モーターは、ナノポアを通じたポリヌクレオチドの転位を駆動するために、ヌクレオチド加水分解のエネルギーを使用することができる。ヘリカーゼは、ATP加水分解が、ポリヌクレオチド転位のエネルギー源である一例である。図1の略図は、ヘリカーゼによるポリヌクレオチドの転位に関する静電的コマ送りモデルを図示している(Frickらの文献、Current Pharmaceutical Design, 12:1315-1338 (2006)参照)。このモデルにおいて、1本鎖ポリヌクレオチドは、第三のドメインからヘリカーゼの2つのRecAドメインを分離している負帯電されたクレフト内に保持される。ATPが存在しない場合、ブックエンド残

50

基(例えば、HCVヘリカーゼのTrp501)及びクランプ残基(例えば、HCVヘリカーゼのArg393)は、1本鎖ポリヌクレオチドがクレフトを通過して滑るのを防ぐ。ATP結合時には、RecAドメインは回転し、正帯電したArg-クランプを移動させる。Arg-クランプは、負帯電した1本鎖ポリヌクレオチドを引き付け、このことは次にブックエンドを外す(clear)。次に1本鎖ポリヌクレオチドは、負帯電されたクレフトと反発し、且つ1本鎖ポリヌクレオチドは、ATPが加水分解されるまでヘリカーゼにより転位する。従ってこの例証的モデルにおいて、ヘリカーゼによるポリヌクレオチド転位には、少なくとも2つの工程が関与している：第一の工程では、ヘリカーゼがATPに結合し、且つ高次構造の変化を受け、並びに第二工程では、ATPは加水分解され、且つポリヌクレオチドは、ヘリカーゼにより転位する。

【0045】

図1Bは、一部の実施態様に従う、Hel308ヘリカーゼと接触している細孔を含む第一の例証的組成物を概略的に図示している。図1Bにおいて、切り欠けのある塗りつぶしの半透明の円形は、本明細書に提供されるHel308ヘリカーゼを表す。直線は、ポリヌクレオチドを表し、点線は、任意の長さのポリヌクレオチドを示す。大きい灰色矢印は、ポリヌクレオチドの細孔の内外へのポリヌクレオチド移動の方向を表し、並びに大きい黒色矢印は、3'側から5'側への、ポリヌクレオチドに沿ったヘリカーゼ転位の方向を示す。図示された実施態様において、細孔(漏斗状の円錐物)は、膜(2本の水平線)内に位置するが、他の細孔の立体配置を好適に使用することができる。図1Bに図示された実施態様において、ポリヌクレオチド移動の方向は、細孔を横切るポテンシャルの差により生じた印加された電場によるものであることができる(180Vの電気ポテンシャルの差が図示されているが、他のポテンシャルの差を好適に使用することができる)。細孔を横切るポテンシャルの差により発生された印加された電場で再度ポリヌクレオチド移動の方向を生じるために、DNAの配向は、図15A-15Cを参照し以下により詳細に説明されるようにフリップされることができる。本明細書においてより詳細に提供するように、Hel308ヘリカーゼは、細孔を通じたポリヌクレオチドのフラクショナル転位を引き起こすことができ、これはヌクレオチドの特徴決定を促進することができる。例えば、かかるフラクショナル転位は、それを基にポリヌクレオチドが特徴決定され得る1以上のシグナルを生じることができる。この1以上のシグナルは、本明細書において別所記載のような電氣的シグナルを含むことができるか、又は本明細書において別所記載のような光学的シグナルを含むことができる。例証的電氣的シグナルは、電流、電圧、トンネル電流(tunneling)、抵抗、電位、電圧、コンダク

【0046】

図示されたように、Hel308ヘリカーゼは、細孔を通じてポリヌクレオチドをフラクショナルに転位するので、細孔内の異なるヌクレオチド塩基の通過は、細孔を通る電流の測定可能な変化を生じ得；かかる電流は、「ブロック(blockade)」電流と称されることができる。本明細書においてより詳細に説明するように、ポリヌクレオチドの配列、ポリヌクレオチドの修飾、ポリヌクレオチドの長さ、ポリヌクレオチドの同一性、ポリヌクレオチドの給源、若しくはポリヌクレオチドの二次構造、又はそれらの任意の好適な組合せなどのポリヌクレオチドの1以上の特徴は、シグナルの変化を基に、例えば細孔を通る電流の変化を基に、決定することができ、それらの変化は、細孔を通じたポリヌクレオチドのHel308ヘリカーゼによるフラクショナル転位工程を基にしている。細孔が非対称である、例えば細孔基底よりもより大きい直径の細孔口を含む(例えばMspAについてなど)ような実施態様において、Hel308ヘリカーゼは、図1Bに図示したように、細孔口と接触することができる。このような立体配置は、「フォワード」立体配置と称することができる。より一般的には、「フォワード立体配置」は、細孔が、細孔基底よりもより広い細孔口を含むかどうかにかかわらず、分子が事実上細孔を移行することができる方向をいうことができる。或いは、「フォワード方向」は、任意に規定することができる。

【0047】

図1Cは、一部の実施態様に従う、標的ポリヌクレオチドを特徴決定する例証的方法の工程を概略的に図示している。この方法は、Hel308ヘリカーゼ及び標的ポリヌクレオチドと

10

20

30

40

50

接触している細孔を横切るポテンシャルの差を印加する工程(工程110)を含むことができる。図13A-13E及び14A-14Dを参照し更に以下に説明されるものに類似した様式で、ポリヌクレオチドの転位は、細孔を通じて転位するポリヌクレオチドに対するポテンシャルの差により引き起こされた加力(applied force)の反対方向であるか、又はポリヌクレオチドの転位は、細孔を通じて転位する(translating)ポリヌクレオチドに対するポテンシャルの差により引き起こされた加力の方向であることができる。任意に、工程110-130は、1回以上反復することができる。フラクショナル転位工程(工程120)は、Hel308ヘリカーゼの完全転位サイクルの第一のフラクショナル転位工程を含むことができるか、又はHel308ヘリカーゼの完全転位サイクルの第二の転位工程を含むことができる。

【0048】

本明細書において使用される用語「ポリヌクレオチド」とは、デオキシリボ核酸(DNA)、リボ核酸(RNA)又はそれらの類縁体をいう。ポリヌクレオチドは、1本鎖、2本鎖であるか、又は1本鎖及び2本鎖の両方の配列を含むことができる。ポリヌクレオチド分子は、2本鎖DNA(dsDNA)の形状(例えば、ゲノムDNA、PCR産物及び増幅産物など)が起源であることができるか、又はDNA(ssDNA)若しくはRNAのように1本鎖の形状を起源とすることができるか、且つdsDNA形状に転換することができるか又はその逆であることができる。ポリヌクレオチド分子の正確な配列は、分かっているか又は不明であることができる。以下は、ポリヌクレオチドの例証的例である：遺伝子又は遺伝子断片(例えば、プローブ、プライマー、EST又はSAGEタグ)、ゲノムDNA、ゲノムDNA断片、エキソン、イントロン、メッセンジャーRNA(mRNA)、トランスファーRNA、リボソームRNA、リボザイム、cDNA、組換えポリヌクレオチド、合成ポリヌクレオチド、分岐型ポリヌクレオチド、プラスミド、ベクター、任意の配列の単離されたDNA、任意の配列の単離されたRNA、核酸プローブ、プライマー又は前述のいずれかの増幅されたコピー。

【0049】

ポリヌクレオチドは、ヌクレオチド又はヌクレオチド類縁体で構成されることができる。ヌクレオチドは典型的には、糖、核酸塩基、及び少なくとも1個のリン酸基を含む。ヌクレオチドは、脱塩基(すなわち、核酸塩基を欠く)であることができる。ヌクレオチドは、デオキシリボヌクレオチド、修飾されたデオキシリボヌクレオチド、リボヌクレオチド、修飾されたリボヌクレオチド、ペプチドヌクレオチド、修飾されたペプチドヌクレオチド、修飾されたリン酸糖骨格ヌクレオチド及びそれらの混合物を含む。ヌクレオチドの例は、例えば、アデノシン一リン酸(AMP)、アデノシン二リン酸(ADP)、アデノシン三リン酸(ATP)、チミジン一リン酸(TMP)、チミジン二リン酸(TDP)、チミジン三リン酸(TTP)、シチジン一リン酸(CMP)、シチジン二リン酸(CDP)、シチジン三リン酸(CTP)、グアノシン一リン酸(GMP)、グアノシン二リン酸(GDP)、グアノシン三リン酸(GTP)、ウリジン一リン酸(UMP)、ウリジン二リン酸(UDP)、ウリジン三リン酸(UTP)、デオキシアデノシン一リン酸(dAMP)、デオキシアデノシン二リン酸(dADP)、デオキシアデノシン三リン酸(dATP)、デオキシチミジン一リン酸(dTMP)、デオキシチミジン二リン酸(dTDP)、デオキシチミジン三リン酸(dTTP)、デオキシシチジン二リン酸(dCDP)、デオキシシチジン三リン酸(dCTP)、デオキシグアノシン一リン酸(dGMP)、デオキシグアノシン二リン酸(dGDP)、デオキシグアノシン三リン酸(dGTP)、デオキシウリジン一リン酸(dUMP)、デオキシウリジン二リン酸(dUDP)、及びデオキシウリジン三リン酸(dUTP)を含む。修飾された核酸塩基を含むヌクレオチド類縁体もまた、本明細書記載の方法において使用することができる。ポリヌクレオチドに含まれる例証的修飾された核酸塩基は、未変性の骨格又は類縁体構造を有するかどうかにかかわらず、例えば、イノシン、キサントシン(xanthine)、ハイポキサントシン(hypoxanthine)、イソシトシン、イソグアニン、2-アミノプリン、5-メチルシトシン、5-ヒドロキシメチルシトシン、2-アミノアデニン、6-メチルアデニン、6-メチルグアニン、2-プロピルグアニン、2-プロピルアデニン、2-チオウラシル(thioLiracil)、2-チオチミン、2-チオシトシン、15-ハロウラシル、15-ハロシトシン、5-プロピニルウラシル、5-プロピニルシトシン、6-アゾウラシル、6-アゾシトシン、6-アゾチミン、5-ウラシル、4-チオウラシル、8-ハロアデニン又はグアニン、8-アミノアデニン又はグアニン、8-チオールアデニン又は

10

20

30

40

50

グアニン、8-チオアルキルアデニン又はグアニン、8-ヒドロキシルアデニン又はグアニン、5-ハロ置換されたウラシル又はシトシン、7-メチルグアニン、7-メチルアデニン、8-アザグアニン、8-アザアデニン、7-デアザグアニン、7-デアザアデニン、3-デアザグアニン、3-デアザアデニンなどを含む。当該技術分野において公知であるように、特定のヌクレオチド類縁体は、例えば、アデノシン5'-ホスホ硫酸のようなヌクレオチド類縁体などは、ポリヌクレオチドに組み込まれることができない。

【0050】

本明細書において使用される用語「細孔」とは、イオン及び／又は水溶性分子が障壁の片側から障壁の他の側へと超えることを可能にする、膜のような障壁を横切って広がる構造を意味することが意図されている。細孔は、膜中に生じることができ、必ずしもではない。例えば、イオン又は水溶性分子の通過を通常阻害する障壁は、イオン又は水溶性分子が障壁の片側から障壁の他の側へと通過することを可能にする、障壁を横切って広がる細孔構造を含むことができる。細孔(例えば、膜貫通細孔)は、例えば、生物学的細孔、固体状態の細孔、並びに生物学的及び固体状態のハイブリッド細孔を含む。

10

【0051】

本明細書において使用される用語「生物学的細孔」とは、例えば、イオン及び／又は水溶性分子が障壁の片側から障壁の他の側へと超えることを可能にする膜を含む、障壁を横切って広がる、生物学的起源の物質から製造される細孔を意味することが意図される。生物学的起源とは、生物又は細胞などの生物学的環境に由来するか又はそれから単離された物質、或いは生物学的に利用可能な構造の合成により製造された型をいう。生物学的細孔は、例えば、ポリペプチド細孔及びポリヌクレオチド細孔を含む。

20

【0052】

本明細書において使用される用語「ポリペプチド細孔」とは、例えば膜のような障壁を横切って広がり、且つ、イオン及び／又は水溶性分子が障壁の片側から障壁の他の側へと流動することを可能にする、1以上のポリペプチドを意味することが意図される。ポリペプチド細孔は、モノマー、ホモポリマー又はヘテロポリマーであることができる。ポリペプチド細孔の構造は、例えば、 α -ヘリックスバンドル細孔及び β -バレル細孔、並びに当該技術分野において周知の他の全てを含む。例証的ポリペプチド細孔は、 α -溶血素、マイコバクテリウム・スメグマチスポリンA、グラミシジンA、マルトポリン、OmpF、OmpC、PhoE、Tsx、F-線毛、SP1(Wangらの文献、Chem. Commun., 49:1741-1743, 2013)及びミトコンドリアのポリン(VDAC)XX、Tom40(米国特許第6,015,714号及びDerringtonらの文献、Proc. Natl. Acad. Sci. USA, 107:16060 (2010))を含む。「マイコバクテリウム・スメグマチスポリンA(MspA)」は、マイコバクテリアにより生成される膜ポリンであり、親水性分子がこの細菌へ侵入することを可能にする。MspAは、密に相互接続された八量体で、杯に似ている中央にチャンネル／細孔を含む膜貫通型 β -バレルを形成する。

30

【0053】

本明細書において使用される用語「ポリヌクレオチド細孔」とは、例えば膜のような障壁を横切って広がり、且つ、イオン及び／又は水溶性分子が障壁の片側から障壁の他の側へと流動することを可能にする、1以上のポリペプチドを意味することが意図される。ポリヌクレオチド細孔は、例えばポリヌクレオチドオリガミ(origami)を含むことができる。

40

【0054】

本明細書において使用される用語「固体状態の細孔」とは、イオン及び／又は水溶性分子が障壁の片側から障壁の他の側へと超えることを可能にする、例えば膜のような障壁を横切って広がる、非-生物学的起源の物質から製造される、細孔を意味することが意図される。固体状態とは、生物学的起源ではない物質を意味することが意図される。固体状態の細孔は、無機物質又は有機物質であることができる。固体状態の細孔は、例えば、窒化ケイ素細孔、二酸化ケイ素細孔、及びグラフェン細孔を含む。

【0055】

本明細書において使用される用語「生物学的及び固体状態のハイブリッド細孔」とは、

50

水和イオン及び／又は水溶性分子が障壁の片側から障壁の他の側へと超えることを可能にする、例えば膜のような障壁を横切って広がる、生物学的起源の物質及び非-生物学的起源の物質の両方から製造される、ハイブリッド細孔を意味することが意図される。生物学的起源の物質は、先に規定されており、例えばポリペプチド及びポリヌクレオチドを含む。生物学的及び固体状態のハイブリッド細孔は、例えば、ポリペプチド-固体状態のハイブリッド細孔及びポリヌクレオチド-固体状態の細孔を含む。

【0056】

本明細書において使用される用語「ヘリカーゼ」とは、2本鎖のポリヌクレオチドをほどこくために、例えばヌクレオチド三リン酸(NTP)などの加水分解から誘導されるエネルギーを利用する活性を有するポリヌクレオチド結合タンパク質を意味することが意図される。2本鎖ポリヌクレオチドのほどきは、ポリヌクレオチドの活性部位に沿ったポリヌクレオチドの転位を生じる。この用語は、1本鎖ポリヌクレオチド、並びに部分的2本鎖ポリヌクレオチドを転位又は結合する活性を有するポリペプチドを含むことが意図される。「Hel308ヘリカーゼ」は、ATP-依存型DNAヘリカーゼ及びスーパーファミリー2ヘリカーゼである。その最初のメンバーであるキイロショウジョウバエ(*Drosophila melanogaster*)由来のMus308は、C-末端DNAポリメラーゼドメインに融合されたN-末端SF2ヘリカーゼドメインからなる。ホモ・サピエンスのHel308は、限定された処理能力を有する3'→5' DNAヘリカーゼであるSF2として機能する。Hel308ヘリカーゼは、ski2-様ヘリカーゼと互換的に使用される。有用なホモログは、ヘリカーゼドメインのみからなることができる(すなわち、ポリメラーゼドメインが存在しない)。ヘリカーゼのみのホモログは、後生動物及び古細菌に存在する。後生動物の例は、ヒトHel308及びMus301である。古細菌の例は、Tga及びMbuである。

【0057】

本明細書において別に明確に説明されない限りは、本明細書において使用される用語「Hel308ヘリカーゼ基質」は、ヘリカーゼにより加水分解されることが可能であり、且つ2本鎖若しくは部分的2本鎖のポリヌクレオチドをほどこく又は1本鎖ポリヌクレオチドを転位するエネルギーを提供する、ヌクレオチド又はヌクレオチド類縁体を意味することが意図される。Hel308ヘリカーゼに共通の基質は、ATPを含む。しかしこの用語の意味に収まる他のHel308ヘリカーゼ基質は、先に説明されたものなどのATP以外のヌクレオチド、及びHel308ヘリカーゼにより加水分解されることが可能であるヌクレオチド類縁体を含む。例証的類縁体は、例えば、リン酸類縁体、例えばγチオール類縁体、αチオール類縁体など、ATP γ S、ATP α S、AMP、PNP、ApCpp、AppCp、及びAppNHpなどを含む。

【0058】

本明細書において使用される用語「転位する」又は「転位」とは、ヘリカーゼ及び／又は細孔に沿った(又はそれ内の)標的ポリヌクレオチドの移動を意味することが意図される。

【0059】

本明細書において使用される用語「完全転位サイクル」とは、ヘリカーゼに関して使用される場合、ヘリカーゼ及び／又は細孔に沿った標的ポリヌクレオチドの1以上のヌクレオチドのユニットの移動のための完全なインターバルを意味することが意図される。完全なインターバルは、サイクルの任意の時点で始まることができ、且つ例えば、ATP結合の工程及び結合したATPの加水分解の工程を含む図3に図示したインターバルを含むことができる。従って、本明細書において使用される完全転位サイクルは、ヌクレオチド基質の結合で始まり、且つヌクレオチド基質の加水分解で終了することができる。完全転位サイクルは同様に、ヌクレオチド基質の加水分解で始まり、且つヌクレオチドの結合で終了することができる。同様に、完全転位サイクルは、それが開始点の直前の工程で終了する限りは、先に例示した2つの開始点の間の任意の時点で始まることができる。

【0060】

本明細書において使用される用語「フラクショナル転位工程」とは、ヘリカーゼに関して使用される場合、完全転位サイクルの一部を特徴づける検出可能な事象を意味すること

が意図される。例えばフラクショナル転位工程は、ヘリカーゼ及び／又は細孔に沿った標的ポリヌクレオチドの1以上のヌクレオチドのユニットの部分的転位であることができる。特定の実施態様において、フラクショナル工程は、高次構造の変化が生じる場合、ATP結合と加水分解の間で生じることができる。この高次構造の変化は、完全転位サイクルを少なくとも2つの部分的又はフラクショナル転位工程に効果的に分ける。フラクショナル工程は、ヘリカーゼに沿った核酸移動に付随してもよくしなくともよい。

【0061】

本明細書において使用される用語「シグナル」とは、情報を表すインジケータを意味することが意図される。シグナルは、例えば、電気的シグナル及び光学的シグナルを含む。

10

【0062】

本明細書において使用される用語「電気的シグナル」とは、情報を表す電気的品質のインジケータを意味することが意図される。このインジケータは、例えば、電流、電圧、トンネル電流、抵抗、電位、電圧、コンダクタンス；並びに、横方向電気的測定値であることができる。「電流」とは、電荷の流れをいう。電気ポテンシャルの差が、細孔を横切って印加される場合に、電荷は流れる。

【0063】

本明細書において使用される用語「光学的シグナル」とは、情報を表す光学的品質のインジケータを意味することが意図される。光学的シグナルは、例えば、蛍光シグナル及びラマンシグナルを含む。

20

【0064】

本明細書において使用される用語「相同性」とは、2つのポリヌクレオチド間の又は2つのポリペプチド間の配列類似性を意味することが意図される。類似性は、比較目的で整理され得る各配列の位置を比較することにより、決定することができる。配列間の類似性の程度は、これらの配列により共有されるマッチング又は相同な位置の数の関数である。それらの%配列類似性を決定するための2つの配列のアラインメントは、例えば、Ausubelらの文献(「分子生物学の最新プロトコル(Current Protocols in Molecular Biology)」、John Wiley and Sons社、ボルチモア、MD (1999))に記載されたものなど、当該技術分野において公知のソフトウェアプログラムを使用し行うことができる。好ましくは、デフォルトのパラメータを、アラインメントに使用するが、その例は、以下に示している。使用することができる当該技術分野において周知の1つのアラインメントプログラムは、デフォルトのパラメータに設定したBLASTである。特に、プログラムは、以下のデフォルトのパラメータを使用するBLASTN及びBLASTPである：遺伝子コード＝標準；フィルター＝なし；鎖＝両方；カットオフ値＝60；期待値＝10；マトリクス＝BLOSUM62；説明＝50配列；検索＝HIGH SCORE；データベース＝非-冗長、GenBank＋EMBL＋DDBJ＋PDB＋GenBank CDS translations＋SwissProtein＋SPupdate＋PIR。これらのプログラムの詳細は、米国国立バイオテクノロジー情報センター(NCBI)で認めることができる。

30

【0065】

本開示は、標的ポリヌクレオチドを特徴決定する方法を提供する。本方法は：(a)Hel308ヘリカーゼ及び標的ポリヌクレオチドと接触している細孔を横切るポテンシャルの差を印加する工程；(b)細孔を通じた標的ポリヌクレオチドの1以上のフラクショナル転位工程により発生した1以上のシグナルを測定する工程；並びに、(c)フラクショナル転位工程の電気的シグナルから標的ポリヌクレオチドを特徴決定する工程：を含む。

40

【0066】

本明細書に記載のように、ポリヌクレオチドは、デオキシリボ核酸(DNA)、リボ核酸(RNA)又はそれらの類縁体を含む。ポリヌクレオチドは一般に、ホスホジエステル結合を含むが、場合によっては、ポリヌクレオチドは、例えば、ホスホラミド(Beaucageらの文献、Tetrahedron, 49(10):1925 (1993)、及びその中の参考文献；Letsingerの文献、J. Org. Chem., 35:3800 (1970)；Sprinzlらの文献、Eur. J. Biochem., 81:579 (1977)；Letsingerらの文献、Nucl. Acids Res., 14:3487 (1986)；Sawaiらの文献、Chem. Lett., 805 (19

50

84)、Letsingerらの文献、J. Am. Chem. Soc., 110:4470 (1988)；並びに、Pauwelsらの文献、Chemica Scripta, 26:141 (1986))、ホスホロチオエート(Magらの文献、Nucleic Acids Res., 19:1437 (1991)；並びに、米国特許第5,644,048号)、ホスホロジチオアート(Briuらの文献、J. Am. Chem. Soc., 111:2321 (1989))、0-メチルホスホロアミダイト連結(Ecksteinの文献、「オリゴヌクレオチド及び類縁体：実践方法(Oligonucleotides and Analogues: A Practical Approach)」、Oxford University Press社参照)、並びにペプチド核酸骨格及び連結(Egholmの文献、J. Am. Chem. Soc., 114:1895 (1992)；Meierらの文献、Chem. Int. Ed. Engl., 31:1008 (1992)；Nielsenの文献、Nature, 365:566 (1993)；Carlssonらの文献、Nature, 380:207 (1996)参照)を含む、代替の骨格も有することができる。他のポリヌクレオチドは、陽性骨格をもつもの(Denpcyらの文献、Proc. Natl. Acad. Sci. USA, 92:6097 (1995))；非イオン性骨格をもつもの(米国特許第5,386,023号、第5,637,684号、第5,602,240号、第5,216,141号及び第4,469,863号；Kiedrowskiらの文献、Angew. Chem. Int. Ed. English, 30:423 (1991)；Letsingerらの文献、J. Am. Chem. Soc., 110:4470 (1988)；Letsingerらの文献、Nucleosides & Nucleotides, 13:1597 (1994)；第2及び3章、ASC Symposium Series 580、「アンチセンス研究における糖修飾(Carbohydrate Modifications in Antisense Research)」、Y. S. Sanghui及びP. Dan Cook編集；Mesmaekerらの文献、Bioorganic & Medicinal Chem. Lett., 4:395 (1994)；Jeffsらの文献、J. Biomolecular NMR, 34:17 (1994)；Tetrahedron Lett., 37:743 (1996))、並びに非-リボース骨格をもつもの(米国特許第5,235,033号及び第5,034,506号、並びにASC Symposium Series 580、「アンチセンス研究における糖修飾」、Y. S. Sanghui及びP. Dan Cook編集の第6及び7章に記載されたものを含む)を含む。1以上の炭素環式糖を含むポリヌクレオチド分子も、ポリヌクレオチドの定義内に含まれる(Jenkinsらの文献、Chem. Soc. Rev., (1995) 169-176頁参照)。いくつかのポリヌクレオチドは、Rawlsの文献、C & E News, Jun. 2, 1997, 35頁に説明されている。

【0067】

本標的ポリヌクレオチドは、本開示の方法に従い特徴決定することができる。例証的ポリヌクレオチドは、例えば、遺伝子又は遺伝子断片(例えば、プローブ、プライマー、EST又はSAGEタグ)、ゲノムDNA、ゲノムDNA断片、エキソン、イントロン、メッセンジャーRNA(mRNA)、トランスファーRNA、リボソームRNA、リボザイム、cDNA、組換えポリヌクレオチド、合成ポリヌクレオチド、分岐型ポリヌクレオチド、プラスミド、ベクター、任意の配列の単離されたDNA、任意の配列の単離されたRNA、核酸プローブ、プライマー又は前述のいずれかの増幅されたコピーを含む。

【0068】

本明細書の特定の実施態様において使用される標的ポリヌクレオチドは、様々な長さのいずれかであることができ、典型的には細孔を通じて広がり且つヘリカーゼにより細孔の片側上に結合されるのに十分な長さである。概して、このような長さは、少なくとも約10ヌクレオチド長である。しかし、この最小サイズよりもより長い多くの長さが、本開示の方法を使用する特徴決定に利用可能である。有用なポリヌクレオチドの例証的長さは、例えば少なくとも約10、20、30、40、50、60、70、80、90、100、150、200、300、400、500、1,000、5,000、又は10,000、100,000ヌクレオチド又はそれ以上を含む。或いは又は加えて、この長さは、1,000,000、100,000、10,000、1,000、100ヌクレオチド以下よりも短いことができる。従って本開示の方法を用い配列決定することができるポリヌクレオチドは、例えば、短いポリヌクレオチド、断片、cDNA、遺伝子及びゲノム断片からの範囲であることができる。

【0069】

本開示の方法において使用されるポリヌクレオチドは、1本鎖、2本鎖であるか、又は1本鎖と2本鎖の両方の配列を含むことができる。ポリヌクレオチド分子は、2本鎖ポリヌクレオチド(例えば、dsDNA)を起源とすることができ、並びに1本鎖ポリヌクレオチドへ転換されることができる。ポリヌクレオチド分子はまた、1本鎖ポリヌクレオチド(例えば、ssDNA、ssRNA)を起源とすることができ、並びにssDNAは、2本鎖ポリヌクレオチドへ転換さ

ることができる。本開示の一部の態様において、2本鎖又は部分的2本鎖ポリヌクレオチドは、ブロックするポリヌクレオチドを含む。かかるポリヌクレオチド種は、本明細書の図13A-13E、14A-14D、及び15A-15Cに結びつけて例示されるものを含むことができる。細孔を通じてポリヌクレオチドを転位する例証の様式は、WO 2013/057495に示されている。

【0070】

一部の態様において、本開示は、標的ポリヌクレオチドを特徴決定する方法を提供する。この方法は、以下を同定することを含む：(1)標的ポリヌクレオチドの配列；(2)標的ポリヌクレオチドの修飾；(3)標的ポリヌクレオチドの長さ；(4)標的ポリヌクレオチドの同一性；(5)標的ポリヌクレオチドの給源；又は、(6)標的ポリヌクレオチドの二次構造。

【0071】

ポリヌクレオチドの配列とは、ポリヌクレオチドの一次構造、又はポリヌクレオチド分子中のヌクレオチドの配列順番をいう。ポリヌクレオチドの配列は、細孔を通じた標的ポリヌクレオチドのフラクショナル転位工程により発生したシグナルを使用し、標的ポリヌクレオチド中のヌクレオチドを特徴決定することにより決定することができる。

【0072】

ポリヌクレオチドの修飾とは、ポリヌクレオチド中のヌクレオチドの何らかの共有又は非共有修飾をいい、例えば、ヌクレオチドのメチル化又はヒドロキシメチル化を含む。実際修飾は、例えば8-オキソグアノシン、5-ホルミルシトシン及び5-カルボキシシトシン並びに本明細書の別所に示した他のものを含む、ポリヌクレオチド鎖に組み込まれ得る任意の数のヌクレオチド類縁体を含むことができる。ヌクレオチドの修飾は、対応するシグナルの変化を提供する。従ってポリヌクレオチドの1又は複数の修飾は、細孔を通じた標的ポリヌクレオチドのフラクショナル転位工程により発生したシグナルを使用し、標的ポリヌクレオチド中の修飾されたヌクレオチドを特徴決定することにより決定することができる。

【0073】

ポリヌクレオチドの長さとは、ポリヌクレオチド中のヌクレオチドの数をいう。ポリヌクレオチドの長さは、例えば、ポリヌクレオチドの一次配列を決定することによるか、又は細孔中のその滞在時間を測定するか、又は細孔を通過するヌクレオチドの数をカウントすることにより、決定することができる。一部の実施態様において、滞在時間は、電流の一過性の変化の持続期間に対応している。一過性の変化は、ポリヌクレオチドの存在に起因した、細孔電流の何らかの逸脱であると考えられることができる。一部の実施態様において、この逸脱は、電流の大きさの減少を生じる。この減少は、一般に当初のブロックされない細孔電流の最大でも95%、90%、80%、60%、50%、40%、30%、20%若しくは10%又はそれ未満であることができる。或いは又は加えて、この減少は、少なくとも10%、20%、30%、40%、50%、60%、70%、80%、若しくは90%又はそれ以上であることができる。場合によっては、このポリヌクレオチドは、ブロックされない細孔に対し、電流の大きさの増加を生じることができる。持続期間とポリヌクレオチドの長さの関係は、使用した実験条件に依存する再現可能な数学的関数により説明することができる。この関数は、所定のポリヌクレオチド型(例えば、DNA又はRNA)について、線形又は非線形(例えば、シグモイド又は指数)関数であることができる。

【0074】

ポリヌクレオチドの同一性とは、ポリヌクレオチドの種類をいう。この同一性はまた、当該技術分野において公知であるようにポリヌクレオチドの名称をいうことができる。例えばポリヌクレオチドの同一性は、例えば、DNA、RNA、2本鎖ポリヌクレオチド、1本鎖ポリヌクレオチド及び／又は部分的2本鎖ポリヌクレオチドであることができる。ポリヌクレオチドの同一性はまた、ポリヌクレオチドの遺伝子産物又は構造機能の決定を含むことができる。例えばポリヌクレオチドは、ポリペプチドをコードすることができるか、又はこれは、リボソームRNAなどの構造ポリヌクレオチドであることができる。ポリヌクレオチドの同一性は、そのポリヌクレオチドの全て又は一部のヌクレオチド配列、そのポリヌクレオチドの全て又は一部と相補的である第二のポリヌクレオチドの配列、そのポリヌ

10

20

30

40

50

レオチドの全て又は一部によりコードされたRNAの配列、或いはそのポリヌクレオチドの全て又は一部によりコードされたタンパク質の配列から決定することができる。特定の例において、ポリヌクレオチドは、ポリヌクレオチドの一部を形成する「タグ」又は「バーコード」配列により同定することができる。かかる例において、ポリヌクレオチドの同一性は、タグ又はバーコードから予期されたシグナルパターンにより割り当てることができる。ポリヌクレオチドの給源とは、ポリヌクレオチドの起源の種又は合成起源をいうことができる。ポリヌクレオチドの同一性及び給源は、例えばBLASTNなど、当該技術分野において周知のプログラムを使用し、ポリヌクレオチド配列データベース中のポリヌクレオチドの配列を整列することにより決定することができる。

【0075】

ポリヌクレオチドの二次構造とは、ポリヌクレオチド分子内の自己相補性の領域の分子内の塩基対をいう。例証的二次構造は、例えば、二重らせん、ヘアピン、ループ、バルジ、二重鎖、ジャンクション、ステム、シュードノット、三重らせん、H-DNA、ハンマーヘッド、及び自己スプライシング型リボザイムを含む。ポリヌクレオチドの二次構造は、例えば、細孔中の滞在時間のその対応する変化を測定するか、又はフラクショナル転位工程により発生したシグナルの対応する変化を測定することにより、決定することができる。

【0076】

細孔は、障壁を横切って広がる構造であり、例えばイオン及び／又は水溶性分子が障壁の片側から障壁の他の側へと超えることを可能にする膜を含む。細孔は、膜中に生じることができるが、必ずしもではない。例えば、通常イオン又は水溶性分子の通過を阻害する障壁は、障壁の片側から障壁の他の側へとイオン又は水溶性分子の通過を可能にする障壁を横切って広がる細孔構造を含むことができる。本開示の膜は、例えば、同じ又は異なる組成物を有することができる2つの液体チャンバーを分離する、非透過性又は半透過性の障壁であることができる。膜が膜貫通細孔を含み、且つこの膜を横切るポテンシャルの差を維持するように構成され得る限りは、任意の膜を本開示に従い使用することができる。好適なポテンシャルの差は、以下に説明する。

【0077】

当該技術分野において周知の様々な膜を、本開示の組成物及び方法において使用することができる。当該技術分野において周知のかかる膜は、様々な異なる構造及び組成を含む。例えば、細孔がポリヌクレオチドの特徴決定に組み込まれ得る限りは、膜は単層又は多層構造であることができる。膜内の層とは、障壁を形成する非透過性又は半透過性の物質をいう。単層膜及び多層膜の例は、以下に更に説明する。

【0078】

膜-形成物質は、生物学的起源又は非生物学的起源であることができる。生物学的起源である物質とは、生体又は細胞などの生物学的環境から由来した物質又はこれから単離された物質、或いは生物学的に利用可能な構造の合成的に製造された型をいう。生物学的起源である物質から製造される例証的膜は、脂質二重層を含む。生物学的起源ではない物質はまた、固体状態の物質とも称され、且つ固体状態の膜を形成することができる。

【0079】

好適な脂質二重層及び脂質二重層を製造若しくは入手する方法は、当該技術分野において周知であり、且つ例えば、米国特許出願公開US 2010/0196203及びPCT出願特許公開公報WO 2006/100484に開示されている。好適な脂質二重層は、例えば細胞膜、細胞小器官膜、リポソーム、平面脂質二重層、及び支持された脂質二重層を含む。脂質二重層は、例えば、リン脂質の2つの反対側の層から形成され得、これはそれらの疎水性尾基は、互いに向かい合い、疎水性内部を形成するのに対し、脂質の親水性ヘッド基は、二重層の各側の水性環境へと外向きになるように配置される。脂質二重層はまた、例えば、Montal及びMullerの文献の方法(Proc. Natl. Acad. Sci. USA., 1972; 69: 3561-3566)により形成されることができる。ここでは脂質単層が、水溶液/空気境界面に垂直である開口部のいずれかの側を通過する水溶液/空気境界面を保持する。脂質は通常、最初にこれを有機溶媒中に溶解することにより、水性電解質溶液の表面に添加され、次にこの溶媒の液滴を開口部の

いずれかの側の水溶液の表面上で蒸発させる。一旦有機溶媒が蒸発されたならば、開口部のいずれかの側の液／気境界面は、二重層が形成されるまで、開口部を通して上下に物理的に移動される。二重層形成の他の一般的方法は、チップ-浸漬、ペインティング二重層、及びリボソーム二重層のパッチ-クランピングを含む。脂質二重層を入手する又は作製するための様々な他の方法も、当該技術分野において周知であり、且つ本開示の組成物及び方法における使用に関して同等に適用可能である。

【0080】

固体状態の膜は、当該技術分野において周知であり、且つ例えば、PCT出願特許公開公報WO 2000/079257において開示されている。先に説明されたように、固体状態の膜は、生物学的起源ではない物質の1以上の層から製造される。固体状態の膜は、支持基板上のコーティング又はフィルムなどの単層、又は自立型要素であることができる。固体状態の膜はまた、サンドイッチ型立体配置中の物質の多層の複合材料であることもできる。生じる固体状態の膜が、膜貫通細孔を含み且つ膜を横切るポテンシャルの差を設定するように配置され得る限りは、本開示に従い使用することができる物質に、特別な制限は存在しない。固体状態の膜は、例えば、マイクロエレクトロニクス材料、 Si_3N_4 、 Al_2O_3 、及び SiO_2 などの絶縁物質、ポリアミドなどの有機及び無機のポリマー、トリブロック共重合体(例えば、両親媒性PMOXA-PDMS-PMOXA ABAトリブロック共重合体)、テフロン(登録商標)などのプラスチック、又は2成分付加型硬化シリコーンゴムなどのエラストマー、並びにガラスを含む、有機物質及び無機物質の両方から製造することができる。加えて固体状態の膜は、二次元のハニカム格子へ密に充填された炭素原子の原子的に薄いシートであるグラフェンの単層、グラフェンの多層、又は他の固体状態の物質の1以上の層と混合されたグラフェンの1以上の層から製造することができる(PCT出願特許公開公報WO 2013/016486)。グラフェンを含む固体状態の膜は、標的ポリヌクレオチドを特徴決定するための電気センサーとして使用することができる、グラフェンナノリボン又はグラフェンナノギャップである少なくとも1つのグラフェン層を含むことができる(PCT出願特許公開公報WO 2013/016486参照)。固体状態の膜は、当該技術分野において周知の方法により製造することができる。例えば、グラフェン膜は、化学蒸着(CVD)又はグラファイトからの剥離(PCT出願特許公開公報WO 2013/016486)のいずれかにより調製することができる。

【0081】

本開示の組成物及び方法は、標的ポリヌクレオチドの特徴決定のために障壁中に位置する細孔を利用することができる。細孔は、生物学的起源又は非生物学的起源である物質から製造することができる。従って細孔は、例えば、生物学的細孔、固体状態の細孔、並びに生物学的及び固体状態のハイブリッド細孔を含む。

【0082】

細孔は、ポリヌクレオチド中のヌクレオチドの配列の検出を促進することに関連した機能性を有することができる。例えば細孔は、ポリヌクレオチドが細孔を通じて移行する速度を制御するために細孔に付着した、会合した、若しくは近傍に位置した、ヘリカーゼ又は他の機能性などの酵素を含むことができる。細孔は、例えば、パッチクランプ回路、トンネル電極回路、又は横方向コンダクタンス測定回路(グラフェンナノリボン、又はグラフェンナノギャップなど)を含む、これに関連した検出回路又はセンサーを有することができる。細孔はまた、細孔と断片の相互作用(例えば、細孔を通じた断片の通過)を基にヌクレオチド配列を決定する、ポリヌクレオチド上の、例えば、蛍光部分又はラマンシグナル発生部分を含む標識物を検出する光学的センサーを含むことができる。

【0083】

特定の実施態様において、ポリペプチド細孔及びポリヌクレオチド細孔を含む生物学的細孔は、細孔が障壁(例えば膜)を通じたポリヌクレオチドの通過を可能にする狭窄ゾーン(constriction zone)を有する限りは、本開示の組成物及び方法において使用することができる。狭窄ゾーンは、被検体(例えば、ポリヌクレオチド又はヌクレオチド)によるブロックが細孔により発生した検出可能なシグナルに影響を及ぼす、細孔の管腔中の位置である。例えば、少なくとも1、2、3、4、5、6、7、8、9、又は10ヌクレオチドの長さを含む

、様々な長さの狭窄ゾーンを有する細孔を、本開示の組成物及び方法において使用することができる。或いは又は加えて、最大で約10、9、8、7、6、5、4、3、2、又は1ヌクレオチドの長さを使用することができる。しかし狭窄ゾーンの長さは、シグナルの品質に影響を及ぼし得る。例えば比較的短い狭窄ゾーンは、ヌクレオチド転位のより良い解像又は再構築精度を生じることができる。一実施態様において、生物学的細孔は、約5ヌクレオチド以下の狭窄ゾーンを有し、狭窄ゾーンに位置する5又は5未満のヌクレオチドは、5よりも多いヌクレオチドから得られる電氣的シグナルよりもヌクレオチド転位のより良い解像を有するよう、電氣的シグナルを調節する。場合によっては、シグナル-対-ノイズ増強は、2ntよりもより小さい狭窄に関する配列決定精度の向上を生じない。これは、比較的小さい狭窄よりもより大きいホモポリマーが、最早検出できない場合に生じ、並びに酵素の確率論的移動のためにヌクレオチドがスキップされる場合の再読み込みの欠如は、精度を低下する。従って5ヌクレオチド以下の狭窄ゾーンを有する好適なポリペプチド細孔及びポリヌクレオチド細孔を、本開示に従い使用することができる。本明細書に提供される内容及び指針を考慮し、当業者は、どの長さの狭窄ゾーンが特定の必要性について適用可能であるかを理解するであろう。例えば当業者は、より高い品質の結果を必要とする適用において、より短い狭窄ゾーンを有する細孔を利用することができる。

10

【0084】

生物学的細孔は、イオン及び／又は水溶性分子が障壁の片側から障壁の他の側へと超えることを可能にする、障壁(例えば膜)を横切って広がる、生物学的起源の物質から製造された細孔である。本明細書に記されたように使用される膜と同様に、細孔に言及する場合、生物学的起源とは、生体又は細胞などの生物学的環境から由来した構造又はこれから単離された構造、或いは生物学的に利用可能な構造の合成的に製造された型をいう。生物学的起源の物質は、例えば、ポリペプチド及びポリヌクレオチドを含む。従って生物学的細孔は、例えば、ポリペプチド細孔及びポリヌクレオチド細孔を含む。

20

【0085】

脂質二重層などの障壁(例えば膜)に再構成されたポリペプチド細孔は、ナノポア配列決定のために使用することができる。ポリペプチドが障壁(例えば膜)を横切る標的ポリヌクレオチドの通過を可能にする狭窄ゾーンを形成することができる限りは、本開示に従い使用することができる様々なポリペプチド細孔が存在する。関与したポリペプチドに応じて、ポリペプチド細孔は、モノマー、ホモポリマー又はヘテロポリマーであることができる。ポリペプチド細孔は、7又は8サブユニットなどの、いくつかの反復サブユニットを含むことができる。従ってポリペプチド細孔は、例えば、六量体、七量体、又は八量体の細孔であることができる。

30

【0086】

ポリペプチド細孔は、例えば、 α -ヘリックスバンドル細孔及び β -バレル細孔、並びに他の当該技術分野において周知のもの全てを含む。 α -ヘリックスバンドル細孔は、 α -ヘリックスにより形成される細孔を含む。好適な α -ヘリックスバンドル細孔は、例えば、WZA及びClyA毒素などの、内膜タンパク質及び α 外膜タンパク質を含む。 β -バレル細孔は、 β -鎖により形成される細孔を含む。好適な β -バレル細孔は、例えば β -毒素、例えば α -溶血素、炭疽菌毒素及びロイコシジンなど、並びに細菌の外膜タンパク質／ポリン、例えばMspAを含むマイコバクテリウム・スメグマチスポリン(Msp)、外膜ポリンF(OmpF)、外膜ポリンG(OmpG)、外膜ホスホリパーゼA並びにナイセリアオートトランスポーターリポタンパク(NalP)を含む。他の細孔は、例えば、ライセニン(例えばWO 2013 153359参照)、又はノルカジア・ファルシニア由来のMspAホモログを含む。

40

【0087】

α -溶血素ポリペプチドは、本開示の方法及び組成物において使用することができる七量体ポリペプチド細孔である。これは、長さ $\sim 5\text{nm}$ の β -バレルに接続された 3.6nm の前庭で構成され、1本鎖ポリヌクレオチドの通過を可能にするが2本鎖ポリヌクレオチドは可能でない 1.4nm の狭窄を含む。 α -溶血素の長さ $\sim 5\text{nm}$ の円柱状 β -バレル細孔は、最大約10個のヌクレオチドを同時に収容することができる。この β -バレル中に位置したヌクレオチ

50

ドは、細孔電流を著しく調節し、引き続き最も狭い1.4nmの細孔狭窄において単独のヌクレオチドに特異的なイオン性シグネチャーを薄め、配列決定適用におけるヌクレオチド転位の全般的解像を低下する。

【0088】

MspAは、本開示の組成物及び方法において使用することができる八量体ポリペプチド細孔である。これは、狭窄長さ～0.5nmで、直径～1.2nmの単独の狭窄を含み；内側細孔は、α-溶血素の円柱状構造とは対照的に、先細の漏斗形状を形成する。Derringtonらの文献は、未変性のα-溶血素を上回るヌクレオチド分離効率において優に3.5倍の増強を伴うトリ-ヌクレオチドセット(AAA、GGG、TTT、CCC)間を識別する遺伝子操作されたMspAの能力を明らかにした(Derringtonらの文献、Proc. Natl. Acad. Sci. USA, 107:16060 (2010))。固定された1本鎖ポリヌクレオチドが関与する実験において、MspAの狭窄内又は近傍の3個と少ないヌクレオチドが、細孔電流に寄与し、未変性のα-溶血素におけるイオン電流を調節することが分かっている～10ヌクレオチドを上回る著しい改善が認められたことが報告された。著者らは、これは、恐らく位置指定変異誘発を通じて単独のヌクレオチドまで更に改善され、今後の目標はMspA変異体であろうと仮定している。

10

【0089】

一部の態様において、ポリペプチド細孔は、マイコバクテリウム・スメグマチスポリンA(MspA)である。一部の態様において、MspAは、配列番号:1のアミノ酸配列、又は少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、若しくは少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、若しくは少なくとも99%の配列番号:1との相同性を有するアミノ酸配列を有する。

20

【0090】

MspAは、好適なポリペプチド細孔である。加えて、MspA変異体は、細孔を通じたポリヌクレオチド転位を制御するために、本開示の組成物及び方法において使用することができる。本明細書の実施態様において使用されるMspA細孔は、

【化1】

GLDNELSLVDGQDRTLTVQQWDTFLNGVFPLDRNRLTREWFHSGRAKYIVAGPGAD
EFEGTLELGYQIGFPWSLGVGINFSYTTTPNILINNGNITAPPFGLNSVITPNLFPGVSISA
RLGNGPGIQEVATFSVRVSGAKGGVAVSNAHGTVTGAAGGVLLRPFARLIASGD
STGDS
VTTYGEPWNMN,

30

に相当する、配列番号:1のアミノ酸配列を有することができ、これは以下の変異を伴うMspA配列である：D90N、D91N、D93N、D118R、D134R及びE139K。配列番号:1のMspA細孔変異体は、「M2 NNN」と称される。少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、若しくは少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、若しくは少なくとも99%の配列番号:1との相同性を有する、他のMspA変異体は、本開示の組成物及び方法において使用することができる。ポリペプチド若しくはポリペプチド領域(又はポリヌクレオチド若しくはポリヌクレオチド領域)が、別の配列との特定の割合(例えば50%)の相同性を有するということは、整列した場合に、これら2つの配列を比較し、アミノ酸(又はヌクレオチド塩基)の割合が、同じであることを意味する。それらの%配列同一性を決定するための2つの配列のアラインメントは、本明細書記載のような、当該技術分野において公知のソフトウェアプログラムを使用し行うことができる。特定の領域又は特異的なアミノ酸残基の挿入、欠失、置換、又は他の選択された修飾を含む、未変性のMspAポリペプチドに対する変異は、MspAポリペプチドをコードしている核酸の位置指定変異誘発を含む、当該技術分野において周知の方法に従い作出することができる(Zoller, M.J.の文献、C

40

50

urr. Opin. Biotechnol., 3:348-354, (1992)). 有用なMspA変異体はまた、US 2012/0055792A1にも記されている。

【0091】

本開示の組成物及び方法において使用される未変性又は変異体MspAポリペプチドは、当該技術分野において周知の様々な方法、例えば、組み換え発現システム、沈降、ゲル濾過、イオン交換、逆相及びアフィニティクロマトグラフィーなどにより単離することができる。他の周知の方法は、Deutscherらの文献、「タンパク質精製の指針(Guide to Protein Purification)」、Methods in Enzymology, Vol. 182, (Academic Press社(1990))に説明されている。或いは、本開示の単離された未変性又は変異体MspAポリペプチドは、周知の組換え方法を使用し得ることができる。本開示の未変性又は変異体MspAポリペプチドの生化学的精製の方法及び条件は、当業者により選択されることができ、且つ精製は、例えば、機能アッセイによりモニタリングすることができる。

10

【0092】

未変性又は変異体MspAポリペプチドを調製する方法の一例は、当該技術分野において周知の方法を使用し、細菌細胞、酵母細胞、又は他の好適な細胞などの、好適な宿主細胞において、ポリペプチドをコードしているポリヌクレオチドを発現し、並びに本明細書記載の方法など、再度周知の精製方法を使用し、発現された未変性又は変異体MspAポリペプチドを回収することである。未変性又は変異体MspAポリペプチドは、本明細書に記載のように、発現ベクターで形質転換されている細胞から直接単離されることができる。組換えにより発現された未変性又は変異体MspAポリペプチドはまた、グルタチオンS転移酵素(GST)又はポリHisなどの適当なアフィニティタグとの、融合ポリペプチドとして発現され、並びにアフィニティ精製されることができる。未変性又は変異体MspAポリペプチドはまた、当業者に周知のポリペプチド合成の方法を使用し、化学合成により生成することもできる。

20

【0093】

脂質二重層などの障壁(例えば膜)へと再構成されたポリヌクレオチド細孔はまた、ナノポア配列決定に使用することができる。このポリヌクレオチド細孔は、障壁(例えば膜)を横切って広がり、且つイオン及び／又は水溶性分子が障壁の片側から障壁の他の側へと流動することを可能にする、1以上のポリヌクレオチドである。ポリヌクレオチドが、障壁(例えば膜)を横切って標的ポリペプチドを通過させる狭窄ゾーンを形成することができる限りは、任意のポリヌクレオチド細孔を、本開示の特定の実施態様に従い使用することができる。例証的ポリヌクレオチド細孔は、例えばポリヌクレオチドオリガミ細孔を含む。そのパターンが二次元又は三次元に広がるポリヌクレオチドオリガミ細孔は、Rothemundの文献(Nature, 440:297-302 (2006))に記載されたように、「オリガミ」を用いて作製することができる。オリガミは、拡大された構造を作出するために、ゲノムポリヌクレオチドの長い鎖及び多くのより短い合成「ステープル」ポリヌクレオチド鎖を使用する一般的技術である。当初のオリガミ構造は、実質的に二次元構造であった。このオリガミ技術は、三次元構造に拡大されている(Douglasらの文献、Nature 459:414-418 (2009); Keらの文献、Nano Letters, 6:2445-2447 (2009); Andersenらの文献、Nature 459:73-76 (2009))。

30

40

【0094】

固体状態の細孔もまた、本開示の組成物及び方法において使用することができる。固体状態の細孔は、イオン及び／又は水溶性分子が障壁の片側から障壁の他の側へと超えることを可能にする、障壁(例えば膜)を横切って広がる、非生物学的起源の物質から製造された細孔である。

【0095】

固体状態の細孔は、固体状態の障壁(例えば膜)中に細孔を作出することにより、形成することができる。従って固体状態の膜と同様に且つ本明細書記載のように、固体状態の細孔は、無機物質及び有機物質の両方を包含している、様々な物質により形成することができる。

50

【0096】

好適な固体状態の細孔は、例えば、酸化アルミニウム、酸化タンタル、酸化チタン、二酸化ケイ素、酸化ハフニウム、酸化ジルコニウム、窒化ホウ素、窒化ケイ素、グラフェン又はそれらのナノラミネート(例えばグラフェン- Al_2O_3)、又はそれらの任意の組合せを含む(PCT出願特許公開公報WO 2013016486A1)。固体状態の細孔は、特注のフィードバック制御されたイオンビーム彫刻ツールを使用することにより、又は膜中のナノポアを分解的にスパッターリングするための電界放出銃(FEG)TEMからの焦点を当てた収束性の電子線を使用し、又は当該技術分野において周知の任意の他の方法(PCT出願特許公開公報WO 2013016486A1)を使用することにより、作製することができる。例えばグラフェン- Al_2O_3 細孔などの、グラフェンナノラミネート細孔は、FEG TEMからの焦点を当てた収束性の電子線を使用する、グラフェン- Al_2O_3 膜を通る穿孔により製造することができる(Venkatesanらの文献、ACS Nano., 6:441-450 (2012))。

10

【0097】

生物学的及び固体状態のハイブリッド細孔は、本開示の組成物及び方法において使用することができる。生物学的及び固体状態のハイブリッド細孔は、イオン及び／又は水溶性分子が障壁の片側から障壁の他の側へと超えることを可能にする、障壁(例えば膜)を横切って広がる、生物学的起源及び非生物学的起源の両方の物質から製造される、ハイブリッド細孔である。生物学的起源の物質は、先に規定されており、且つ例えばポリペプチド及びポリヌクレオチドを含む。非生物学的起源の物質は、本明細書において説明されたように、固体状態の物質と称される。

20

【0098】

従って生物学的及び固体状態のハイブリッド細孔は、例えばポリペプチド-固体状態のハイブリッド細孔及びポリヌクレオチド-固体状態のハイブリッド細孔を含む。ポリペプチド-固体状態のハイブリッド細孔は、1以上のポリペプチド及び固体状態の物質を含む。ポリヌクレオチド-固体状態のハイブリッド細孔は、1以上のポリヌクレオチド及び固体状態の物質を含む。生物学的及び固体状態のハイブリッド細孔は、ポリペプチド又はポリヌクレオチド細孔を、固体状態の細孔により操作することにより、製造される(PCT出願特許公開公報WO 2013/016486参照)。好適なポリペプチド細孔、ポリヌクレオチド細孔、及び固体状態の細孔の例は、先に説明されている。

【0099】

ナノポア配列決定装置は、単独の又は複数の細孔を有することができる。複数の細孔は、同じ又は異なる組成物を有する、2以上の標的ヌクレオチドを特徴決定するためのナノポアアレイとして使用することができる。本明細書において使用される複数の細孔の数の例は、例えば少なくとも1、4、16、64、256、512、1028、4096、16384、32768、100000、100万、1000万個の細孔又はそれ以上を含む。好ましい実施態様において、複数の細孔の数は、4096よりも大きいであろう。ナノポアアレイは、当該技術分野において公知であり、且つ例えば、PCT出願特許公開公報WO 2013/016486において開示されている。例えば、直径~15nmの固体状態の細孔の高密度アレイは、 $\text{SiN}/\text{Al}_2\text{O}_3$ 膜において、電子線リソグラフィ工程及び反応性イオンエッチング工程を用い作り上げることができ、ポリヌクレオチド分子のハイスループット分析を促進する。

30

40

【0100】

本開示の方法は、障壁(例えば膜)を横切るポテンシャルの差を利用することができる。ポテンシャルの差は、電気ポテンシャルの差、化学ポテンシャルの差、又は電気化学ポテンシャルの差であることができる。電気ポテンシャルの差は、少なくとも1種の液体プールへ電流を投入若しくは加える電源を介して障壁(例えば膜)を横切るよう発生され得る。化学ポテンシャルは、2つのプールのイオン組成の差により、障壁を横切るよう発生され得る。電気化学ポテンシャルの差は、電気ポテンシャルと組合せた2つのプールのイオン組成の差により確立され得る。異なるイオン組成は、例えば、各プール中の異なるイオン、又は各プール中の同じイオンの異なる濃度であることができる。

【0101】

50

細孔を通じたポリヌクレオチドの転位を強制する細孔を横切る電気ポテンシャルの印加は、当該技術分野において周知であり、且つ本開示に従い使用することができる(Deamerらの文献、Trends Biotechnol., 18:147-151 (2000); Deamerらの文献、Ace Chem Res., 35:817-825 (2002); 及び、Liらの文献、Nat Mater., 2(9):611-615 (2003))。本開示の方法は、細孔を横切り印加された電圧により実行することができる。電圧の範囲は、40mVから1Vを超えるまで選択することができる。典型的には、本開示の方法は、100~200mVの範囲で試行されるであろう。具体例において、本方法は、140mV又は180mVで試行される。電圧は、モーターの移動の間は、静的である必要はない。電圧極性は、典型的には、負帯電したポリヌクレオチドが、細孔へと電気泳動的に駆動されるように加えられる。場合によっては、モーターの適切な機能を促進するために、電圧は、減少されるか、又は極性は逆転され得る。

10

【0102】

場合によっては、圧力格差の適用は、細孔を通じたポリヌクレオチドの転位を強制するために利用することができる。圧力格差は、本明細書に例証された方法において、電気ポテンシャル又は他のポテンシャルの差の代わりに使用することができる。

【0103】

本開示の方法は、細孔を通じた1以上のヌクレオチドの転位に対応する1以上のシグナルを生じる。従って、標的ポリヌクレオチドは細孔を移行するので、障壁を横切る電流は、例えば、狭窄の塩基-依存性ブロックのために変化する。その電流の変化からのシグナルは、本明細書記載の様々な方法又は当該技術分野において公知のそれ以外の方法のいずれかを使用し、測定することができる。各シグナルは、細孔中のヌクレオチド種に独自であり、結果的に生じるシグナルを使用し、先に説明されたようにポリヌクレオチドの特徴を決定することができる。例えば特徴的シグナルを生じるヌクレオチドの1以上の種の同一性を、決定することができる。本開示の方法において有用なシグナルは、例えば電氣的シグナル及び光学的シグナルを含み、これらは以下に更に説明される。一部の態様において、電氣的シグナルは、電流、電圧、トンネル電流、抵抗、電圧、コンダクタンスの測定値;又は、横方向電氣的測定値であることができる(PCT出願特許公開公報WO 2013/016486)。一部の態様において、電氣的シグナルは、細孔を通過する電氣的電流である。

20

【0104】

本明細書に示した方法において検出される電氣的シグナルは、細孔を通過する電荷の流れである電流であることができる(Deamerらの文献、Trends Biotechnol., 18:147-151 (2000); Deamerらの文献、Ace Chem Res., 35:817-825 (2002); 及び、Liらの文献、Nat Mater., 2(9):611-615 (2003))。本明細書記載のように、電氣的シグナルは、細孔に組合せた検出回路、例えば、パッチクランプ回路又はトンネル電極回路を用いて測定することができる。検出され得る電圧、トンネル電流、抵抗及びコンダクタンスシグナルの例、並びにそれらの検出のための装置の例は、例えば、Wanunuの文献、Phys Life Rev., 9(2):125-58 (2012); 及び、Venkatesanらの文献、Nat Nanotechnol., 6(10):615-24 (2011)に説明されたように、当該技術分野において公知である。

30

【0105】

本開示の方法において有用な光学的シグナルは、例えば蛍光及びラマンシグナルを含む。光学的シグナルは、標的ヌクレオチドを、光学的シグナルを発生する標識物、例えば、蛍光部分又はラマンシグナル発生部分と組合せることにより発生され得る。例えば、della Torreらの文献、Nanotechnology, 23(38):385308 (2012)において、TiO₂-コートされた膜の広範な領域を照射するために、全反射照明蛍光(TIRF)顕微鏡の光学的スキームを利用した。Soniらの文献、Rev Sci Instrum., 81(1):014301 (2010)において、方法は、2種の単独-分子測定モダリティ、すなわち全内部反射顕微鏡及びナノポアを使用する生体分子の電氣的検出を統合するために使用した。

40

【0106】

本明細書記載のように、細孔は、例えばパッチクランプ回路、トンネル電極回路、又は横方向コンダクタンス測定回路(グラフェンナノリボン、又はグラフェンナノギャップな

50

ど)を含む、検出回路と組合せて、本実施態様における電氣的シグナルを記録することができる。加えて細孔はまた、ポリヌクレオチド上の、例えば、蛍光部分又はラマンシグナル発生部分など、標識物を検出する光学的センサーと組合せることができる。

【0107】

ナノポア配列決定法は、細孔を通じた標的ポリヌクレオチドの転位を遅らせる機序を利用することができる。例えば、ヘリカーゼ、トランスロカーゼ、又はポリメラーゼなどのポリヌクレオチド結合タンパク質は、転位速度を制御するために、付着又は組み込まれることができる。この付着は、例えば一過性若しくは持続性であることができ、且つ細孔を通じて引き込まれる際に標的ポリヌクレオチドにより、又は当該技術分野において周知の様々なポリペプチド、化学リンカー若しくは捕獲部分により、媒介されることができる。例証的技術は、Manraoらの文献、Nat Biotechnol., 30(4):349-353 (2012)及びCherfらの文献、Nat Biotechnol., 30(4):344-348 (2102)に説明されている。特定の実施態様において、ヘリカーゼ又は他の分子モーターを使用し、細孔を通じた標的ポリヌクレオチドの転位を遅延若しくは停止することができる。例えば転位に作用するようヌクレオチドを加水分解するモーターを使用する場合、ヌクレオチドは、このモーターから省くことができ、並びに／又はこのモーターは、インヒビター(例えば加水分解不可能なヌクレオチド類縁体)に供され、結果的に標的ポリヌクレオチドは、モーターに結合され続け且つ細孔を通じて評価できるほどには転位しないことができる。一部の実施態様において、転位は引き続き、ヌクレオチドをモーターへ送達し及び／又はインヒビターを除去することにより生じ得る。本開示の方法は、細孔を標的ポリヌクレオチド及びHel308ヘリカーゼと接触させ、細孔を通じたポリヌクレオチドの転位速度を制御する工程を含むことができる。以下に更に説明されるように、Hel308ヘリカーゼは、ATP-依存型DNAヘリカーゼ及びスーパーファミリー2ヘリカーゼとして特徴決定され得る。本明細書に提供される内容及び指針を考慮し、当業者は、本実施態様に従い使用するために、任意のHel308ヘリカーゼを好適に選択するか又は適応させることができる。好適なHel308ヘリカーゼは以下に更に説明される。

【0108】

本明細書に示した方法の一部の態様において、標的ポリヌクレオチドの転位は、細孔を通る電流の方向の反対方向である。他の態様において、標的ポリヌクレオチドの転位は、細孔を通過する電流の方向と同じ方向である。

【0109】

従って本開示の方法は、少なくとも二つの様式で実行することができ、ここで標的ポリヌクレオチドの転位は、例えば細孔を通る電流又は他のポテンシャルの方向と反対又はこれと同じのいずれかである。この結果は、本開示のHel308ヘリカーゼを、標的ポリヌクレオチドの5'末端又は3'末端のいずれかに結合することにより、達成することができる。2本鎖ポリヌクレオチドに言及する場合、5'方向又は3'方向は、2本鎖ポリヌクレオチド内の1本鎖についていう。従ってHel308ヘリカーゼは、細孔の外又は内への、すなわち、電圧勾配により生じたポリヌクレオチドに対する力に対抗する方向で(図13A-13E及び14A-14D参照)、又は電圧勾配又は他のポテンシャルにより生じた力と同じ方向にポリヌクレオチドが移動するように、転位速度を制御するヘリカーゼを使用し(図15A-15C参照)、ポリヌクレオチドを引き出し又は供給のいずれかができる。

【0110】

図13A-13Eは、例えば一部の実施態様に従う、Hel308ヘリカーゼ3'側オーバーハング結合部位及びコレステロール二重層アンカーを伴う、3つ組ポリヌクレオチド複合体を基にした、電圧勾配などのポテンシャルにより生じた力に対する、Hel308ヘリカーゼにより制御されたポリヌクレオチド転位を示している。黒丸(●)は、5'側リン酸を表す。黒菱形(◆)は、3'側コレステロールを表す。切り欠けのある塗りつぶしの半透明の円形は、Hel308ヘリカーゼを表す。点線は、任意の長さを示す。大きい灰色矢印は、ポリヌクレオチドの細孔の内外へのポリヌクレオチド移動の方向(印加された電場により又はこれに対抗して)を表す。大きい黒色矢印は、3'側から5'側へのポリヌクレオチドに沿ったヘリカー

ゼ転位の方向を示す。細孔(漏斗状の円錐物)は、膜(2本の水平線)内に位置する。

【0111】

図13A-13Eは、ポリヌクレオチド配列決定のための、Hel308ヘリカーゼ3'側オーバーハング結合部位及びコレステロール二重層アンカーを伴う、3つ組ポリヌクレオチド複合体の使用を図示している。コレステロール-標識されたポリヌクレオチド“i”は、任意であり、且つ細孔を通じて転位し、全複合体の脂質二重層への動員を促進する標的ポリヌクレオチド“ii”へとハイブリダイズするために使用される(図13A)。5'側リン酸は、例えば、電圧勾配により、細孔を通じて引き寄せられ、標的ポリヌクレオチド“ii”の5'末端で最初の細孔への侵入を生じ、且つコレステロール-標識されたポリヌクレオチドが外れることを引き起こす(図13B)。リン酸-含有ポリヌクレオチドは、トランス側に細孔を通じて引き寄せられるので、第二のハイブリダイズされたポリヌクレオチド“iii”は、細孔が2本鎖ポリヌクレオチドの転位を可能にするには余りにも狭いために、外れる(図13C)。ポリヌクレオチドiiiの一つの目的は、Hel308ヘリカーゼが優先的に結合することができる一般に約8個のヌクレオチドの3'側1本鎖ポリヌクレオチドオーバーハングである、Hel308ヘリカーゼ結合部位を作製することである。更にHel308ヘリカーゼ分子を転位しているポリヌクレオチドの3'末端に強制的に結合させることにより、細孔を通じて転位しているポリヌクレオチドの長さは最大化される。この複合体のポリヌクレオチド“iii”は、任意の長さを含む、あらゆる長さであることができ、且つ3'末端は、ポリヌクレオチド“i”の5'末端に隣接する必要はない。細孔口に到達した時点で、Hel308ヘリカーゼは、電圧勾配に対してポリヌクレオチドをその3'から5'トランスロカーゼ活性を介して引き寄せ、シスチャンバーへ戻す(図13D及び図13E)。

【0112】

図14A-14Dもまた、一部の実施態様に従う、例えば、Hel308ヘリカーゼ3'側オーバーハング結合部位及びコレステロール二重層アンカーを伴う、3つ組ポリヌクレオチド複合体を基にした、電圧勾配などのポテンシャルにより生じた力に対する、Hel308ヘリカーゼにより制御されたポリヌクレオチド転位を図示する。黒丸(●)は、5'側リン酸を表す。黒菱形(◆)は、3'側コレステロールを表す。切り欠けのある塗りつぶしの半透明の円形は、Hel308ヘリカーゼを表す。点線は、任意の長さを示す。大きい灰色矢印は、ポリヌクレオチドの細孔の内外へのポリヌクレオチド移動の方向(印加された電場により又はこれに対抗して)を表す。大きい黒色矢印は、3'側から5'側へのポリヌクレオチドに沿ったヘリカーゼ転位の方向を示す。細孔(漏斗状の円錐物)は、膜(2本の水平線)内に位置する。しかしこのスキームは、任意のコレステロール部分に結合するか又はまたこれを含むためにHel308ヘリカーゼに関して標的ポリヌクレオチド“ii”上に3'側オーバーハングを作出する単独のハイブリダイゼーションポリヌクレオチド“i”の使用を例示している。Hel308ヘリカーゼは、ポリヌクレオチド“ii”の1本鎖領域上のどこかに結合することができる。複数のHel308ヘリカーゼ分子が示され、且つ“E1”、“E2”及び“E3”と示されている。最初に細孔口に達するヘリカーゼは、シス側に戻る制御された転位プロセスを開始するであろう。これが離れるとすると、次の結合したHel308ヘリカーゼ分子が細孔口に到達し、制御された転位を開始するまで、制御されない転位が続いて起こる。

【0113】

図15A-15Cは、電圧勾配などのポテンシャルにより生じた力と同じ方向のポリヌクレオチド移動としてのポリヌクレオチド転位速度を制御するためのHel308ヘリカーゼの使用を例示している。切り欠けのある塗りつぶしの半透明の円形は、Hel308ヘリカーゼを表す。点線は、任意の長さを示す。大きい灰色矢印は、細孔の内側への印加された電場によるポリヌクレオチド移動の方向を表す。大きい黒色矢印は、3'側から5'側へのポリヌクレオチドに沿ったヘリカーゼ転位の方向を示す。細孔(漏斗状の円錐物)は、膜(2本の水平線)内に位置する。この例証的スキームにおいて、標的ポリヌクレオチドは、最初に細孔の3'末端に侵入する。これは転位しているポリヌクレオチドに沿って3'側から5'側へ転位するので、Hel308ヘリカーゼは、細孔へのポリヌクレオチドの転位の速度を制御する。

【0114】

先に説明したように、Hel308ヘリカーゼの文脈でのフラクショナル転位工程とは、ヘリカーゼ及び／又は細孔に沿った標的ポリヌクレオチドの1以上のヌクレオチドの部分的転位をいうことができる。従ってフラクショナル転位工程とは、完全転位サイクル未満であるヌクレオチド工程の一部をいう。フラクショナル転位工程は、高次構造の変化が生じる場合に、ATP結合と加水分解の間で起こることができる。1以上のフラクショナル転位工程は、完全ヌクレオチド工程に必要とされ得る。高次構造変化は、完全転位サイクルを少なくとも2つの部分的又はフラクショナル転位工程に効果的に分ける。

【0115】

部分的又はフラクショナル転位工程は、細孔を移行する1以上のヌクレオチドを特徴決定するための独自のシグナルを発生するために同じ様式で利用することができる。従って本開示の方法は、細孔を通じた各々の1以上のヌクレオチド転位に関する各々のフラクショナル転位工程に対応する電流の変化に起因する、少なくとも2種の電氣的シグナルを生じることができる。従って一部の態様において、フラクショナル転位工程は、Hel308ヘリカーゼの完全転位サイクルの第一のフラクショナル転位工程を含む。別の態様において、フラクショナル転位工程は、Hel308ヘリカーゼの完全転位サイクルの第二のフラクショナル転位工程を含む。細孔の狭窄ゾーンを移行する1以上のヌクレオチドを特徴決定するために、各第一又は第二のフラクショナル転位工程は、単独で又はそのパートナーと一緒に、例えば各々第二若しくは第一のフラクショナル転位工程と一緒に、使用することができる。

【0116】

例えば、実施例Iに更に説明するように、Hel308ヘリカーゼは、ATPに結合し、高次構造変化を受け、第一のフラクショナル転位工程を提供することができ、並びにHel308ヘリカーゼは、標的ポリヌクレオチドの1以上のヌクレオチドをATP加水分解によりヘリカーゼ及び／又は細孔に沿って転位し、第二のフラクショナル転位工程を提供することができる。第一及び第二のフラクショナル転位工程のいずれか又は両方は、例えばシグナルを発生するヌクレオチド又は1以上のヌクレオチドのヌクレオチド配列を決定するために使用することができる。シグナルが2以上のヌクレオチドにより発生される場合は、このシグナルを発生しているポリヌクレオチドの一部が、ワードと称される。従ってかかるヌクレオチドワードは、長さが少なくとも4、5、6、7、8、9、10又はそれよりも多いヌクレオチドであり、且つ細孔の狭窄ゾーンの長さに対応することができる。或いは又は加えて、ヌクレオチドワードは、長さが最大でも10、9、8、7、6、5、若しくは4又はそれよりも少ないヌクレオチドであることができる。

【0117】

先に説明し且つ下記実施例IIIにおいて更に例証するように、ポリヌクレオチド中の1以上のヌクレオチド残基は、完全転位サイクルの2つのフラクショナル工程から得られた電氣的シグナルを用いて、同定することができる。両方のフラクショナル転位工程からのシグナルの利用は、同じ1以上のヌクレオチドに関して2つ組のシグナルを提供し、1回の決定のうちでのより良い精度を可能にする。従って両方のフラクショナル転位工程からのシグナルの利用は、増大した特徴決定精度を生じ、完全転位サイクルから得られる単独の電氣的又は他のシグナルを使用する1以上のヌクレオチドの同定と比べ、誤差率を25～50%だけ低下することができる。同様に両方のフラクショナル転位工程からのシグナルの利用は、少なくとも5%、10%、20%、30%、40%、45%、55%、60%、65%、70%、75%、80%、85%、90%又はそれよりも大きく減少した誤差率を生じることができる。本明細書に提供される内容及び指針を考慮し、当業者は、例えばヌクレオチド転位の解像を増大するために、先に説明されたように、狭窄ゾーンのサイズを減少することによるなど、所定の目的のために精度を調節する方法を知っているであろう。

【0118】

他の実施態様において、フラクショナル転位工程から得られた追加情報は、数多くの方式で、ナノポア配列決定を進めるために使用することができる。例えば同じヌクレオチド

ワードに関してフラクショナル転位工程から得られた測定値は、ナノポアベース-呼び出し精度を向上するためにアルゴリズムにおいて使用することができる。同じヌクレオチドワードは、1回の決定において2回読まれるので、同じヌクレオチドワードに関してフラクショナル転位工程から得られた測定値は、ホモポリマー読み値の誤差率を減少するために使用することができる。従って同じヌクレオチドワードに関するフラクショナル転位工程から得られた測定値は、未変性のポリヌクレオチド転位反応の分解できる解像を倍加し、配列-特異性パターンの増強された解像を生じる。後者の1利用は、反復配列又は単一ヌクレオチド多型(SNP)を検出するための、配列-特異的パターン認識アルゴリズムである。

【0119】

先に記したように、方法は、(a)Hel308ヘリカーゼ及び標的ポリヌクレオチドと接触している細孔を横切るポテンシャルの差を引き起こす工程；(b)細孔を通じた標的ポリヌクレオチドの1以上のフラクショナル転位工程により発生した1以上のシグナルを測定する工程；並びに、(c)フラクショナル転位工程の電氣的シグナルから標的ポリヌクレオチドを特徴決定する工程：を含むことができる。一部の態様において、本方法は更に、工程(a)－(c)の1回以上の繰り返しを含む。工程(a)－(c)を繰り返すことにより、隣接ヌクレオチド又は隣接ヌクレオチドワードは、特徴決定され得る。工程(a)－(c)の繰り返しは、標的ポリヌクレオチドの一部又は全てが特徴決定されるまで、望ましいなら繰り返すことができる。例えば標的ポリヌクレオチドの一部又は全ての配列は、工程(a)－(c)のいくつかの望ましい数の反復を通じて決定することができる。従って標的ポリヌクレオチドの全体又は部分に関する1以上の特徴決定を、決定することができる。

【0120】

本明細書に説明したように、任意のHel308ヘリカーゼ又はそれらの変種を、本実施態様に従い使用することができる。例証的Hel308ヘリカーゼは、下記表1及び表2に示している。

表1. Hel308ヘリカーゼの例

【表 1】

アクセッション	説明
NP_578406.1	ski2-様ヘリカーゼ[パイロコッカス・フリオサスDSM 3638] >sp O73946.1 HELS_PYRFU RecName: Full=推定ski2-型 ヘリカーゼ>pdb 2ZJ2 A 鎖A, Archaeal DNAヘリカーゼHjm Apo State 1型において >pdb 2ZJ5 A 鎖A, ArchaealDNAヘリカーゼHjm Complexed With Adp 1型において>pdb 2ZJ8 A 鎖A, ArchaealDNA ヘリカーゼHjm Apo State 2型において >pdb 2ZJA A鎖A, Archaeal DNAヘリカーゼHjm Complexed With Amppep 2型において >dbj BAA32016.1 ヘリカーゼ[パイロコッカス・フリオサス]>gb AAL80801.1 ヘリカーゼ[パイロコッカス・フリオサスDSM 3638]
NP_126564.1	ski2-様ヘリカーゼ[パイロコッカス・アビシGE5] >sp Q9V0A9.1 HELS_PYRAB RecName: Full=推定ski2-型 ヘリカーゼ >emb CAB49795.1 DNAヘリカーゼ[パイロコッカス・アビシGE5]
NP_143168.1	ski2-様ヘリカーゼ[パイロコッカス・ホリコシOT3] >sp O59025.1 HELS _ PYRHO RecName: Full=推定ski2-型ヘリカーゼ >dbj BAA30383.1 715aaの長い仮定のタンパク質[パイロコッカス・ ホリコシOT3]
YP_004424773.1	ski2-様ヘリカーゼ[パイロコッカス種NA2] >gb AEC52769.1 ski2-様 ヘリカーゼ[パイロコッカス種NA2]
YP_004623750.1	ski2-様ヘリカーゼ[パイロコッカス・ヤヤノシCHI] >gb AEH24478.1 ski2- 様ヘリカーゼ[パイロコッカス・ヤヤノシCHI]
YP_002307730.1	ski2-様ヘリカーゼ[サーモコッカス・オウレウスNA1] >gb ACJ16833.1 DNAヘリカーゼ[サーモコッカス・オウレウスNA1]
YP_004763427.1	ski2-様ヘリカーゼ[サーモコッカス種4557] >gb AEK73750.1 ski2- 様ヘリカーゼ [サーモコッカス種4557]
YP_002959236.1	ski2-様ヘリカーゼ[サーモコッカス・ガンマトレランスEJ3] >gb ACS33372.1 ski2-型ヘリカーゼ、推定[サーモコッカス・ ガンマトレランスEJ3]
YP_004071709.1	ski2-型ヘリカーゼ[サーモコッカス・パロフィラスMP] >gb ADT84486.1 推定ski2-型ヘリカーゼ[サーモコッカス・パロフィラスMP]
YP_002994328.1	推定ski2-型ヘリカーゼ[サーモコッカス・シビリカスMM 739] >gb ACS89979.1 推定ski2-型ヘリカーゼ[サーモコッカス・シビリカス MM 739]

10

20

30

ZP_04875329.1	III型制限酵素, resサブユニットファミリー[アシデュリプロフンダム・ボーネイT469] >gb EDY35111.1 III型制限酵素, resサブユニットファミリー[アシデュリプロフンダム・ボーネイT469]	
YP_003436565.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ [フェログロバス・プラシダスDSM 10642] >gb ADC66290.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質 [フェログロバス・プラシダスDSM 10642]	
YP_004485304.1	ski2-型ヘリカーゼ[メタノトリス・イグネウスKol5] >gb AEF97239.1 ski2-型ヘリカーゼ[メタノトリス・イグネウスKol 5]	10
YP_004616424.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン-含有タンパク質[メタノサルサム・ズィリナーエDSM 4017] >gb AEH61205.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[メタノサルサム・ズィリナーエDSM 4017]	
ZP_04873370.1	III型制限酵素, resサブユニットファミリー[アシデュリプロフンダム・ボーネイT469] >ref YP_003482774.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[アシデュリプロフンダム・ボーネイT469] >gb EDY36687.1 III型制限酵素, resサブユニットファミリー[アシデュリプロフンダム・ボーネイT469] >gb ADD08212.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[アシデュリプロフンダム・ボーネイT469]	20
YP_004342552.1	ski2-型ヘリカーゼ[アルカエグロブス・ベネフィカスSNP6] >gb AEA47837.1 ski2-型ヘリカーゼ[アルカエグロブス・ベネフィカスSNP6]	
NP_071282.1	SKI2-ファミリーヘリカーゼ[アルカエグロブス・フルギダスDSM 4304]	
2P6R_A	A鎖,ほどこれたDNAと複合体のスーパーファミリー2ヘリカーゼHel308の結晶構造 >pdb 2P6U A鎖, Hel308スーパーファミリー2ヘリカーゼのアポ構造	30
YP_685308.1	ski2-様ヘリカーゼ[非培養のメタン生成性古細菌RC-1] >sp Q0W6L1.1 HELS_UNCMA RecName:Full=推定ski2-型ヘリカーゼ>emb CAJ35982.1 推定ski2-型ヘリカーゼ[非培養のメタン生成性古細菌RC-1]	
YP_001048404.1	ski2-様ヘリカーゼ[メタノクレウス・マリスニグリJR1] >gb ABN58422.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[メタノクレウス・マリスニグリJR1]	
YP_919908.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン-含有タンパク質[テルモフィルム・ペンデンスHrk 5] >gb ABL77905.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[テルモフィルム・ペンデンスHrk 5]	40

YP_843229.1	ski2-様ヘリカーゼ[メタノサエタ・テルモフィリアPT] >gb ABK14589.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[メタノサエタ・テルモフィリアPT]	
ZP_08045937.1	ski2-様ヘリカーゼ[ハラダブタツス・パウチハロフィラスDX253] >gb EFW90585.1 ski2-様ヘリカーゼ[ハラダブタツス・パウチハロフィラスDX253]	
NP_280985.1	ski2-様ヘリカーゼ[ハロバクテリウム種NRC-1] >ref YP_001690117.1 ski2-様ヘリカーゼ[ハロバクテリウム・サリナラムR1] .sp Q9HNV6.1 HELS_HALSA RecName: Full=推定ski2-型ヘリカーゼ >sp B0R7Q2.1 HELS_HALS3 RecName: Full=推定ski2-型ヘリカーゼ >gb AAG20465.1 DNA修復タンパク質[ハロバクテリウム種NRC-1] >emb CAP14771.1 推定DNAヘリカーゼ [ハロバクテリウム・サリナラムR1]	10
YP_003357840.1	ホリデージャンクション移動ヘリカーゼ[メタノセラ・パルディコラSANAЕ] dbj BAI62857.1 ホリデージャンクション移動ヘリカーゼ [メタノセラ・パルディコラSANAЕ]	
YP_003457479.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[メタノカルドコッカス種FS406-22] >gb ADC68743.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質 [メタノカルドコッカス種FS406-22]	20
YP_003127632.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[メタノカルドコッカス・フェルベンスAG86] >gb ACV24132.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[メタノカルドコッカス・フェルベンスAG86]	
YP_003735335.1	ski2-様ヘリカーゼ[ハラルカリコッカス・ジェオタガリB3] >gb ADJ13543.1 ski2-様ヘリカーゼ[ハラルカリコッカス・ジェオタガリB3]	
YP_503885.1	ski2-様ヘリカーゼ[メタノスピリラム・フンガテイJF-1] >gb ABD42166.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ-様タンパク質[メタノスピリラム・フンガテイJF-1]	30
BAJ48115.1	ヘリカーゼ[カンジダタス・カルディアカエウム・サブテラネウム] >dbj BAJ48144.1 ヘリカーゼ[カンジダタス・カルディアカエウム・サブテラネウム] >dbj BAJ50919.1 ヘリカーゼ[カンジダタス・カルディアカエウム・サブテラネウム]	
YP_001405615.1	ski2-様ヘリカーゼ[カンジダタス・メタノレギュラ・ブーネイ6A8] >sp A71B61.1 HELS_METB6 RecName: Full=推定 ski2-型ヘリカーゼ >gb ABS56972.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[メタノレギュラ・ブーネイ6A8]	40
YP_306959.1	ski2-様ヘリカーゼ[メタノサルキナ・バーケリ株フサロ] >sp Q465R3.1 HELS_METBF RecName: Full=推定ski2-型ヘリカーゼ >gb AAZ72379.1 ヘリカーゼ[メタノサルキナ・バーケリ株フサロ]	

YP_001031179.1	ski2-様ヘリカーゼ[メタノコルプスクラム・ラブレアヌムZ] >gb ABN07912.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメインタンパク質[メタノコルプスクラム・ラブレアヌムZ]	
YP_003541733.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ [メタノハロフィラス・マヒイDSM 5219] >gb ADE36088.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン タンパク質[メタノハロフィラス・マヒイDSM 5219]	
YP_004384692.1	推定Ski2-型ヘリカーゼ[メタノサエタ・コンシリGP6] >gb AEB68874.1 推定Ski2-型ヘリカーゼ[メタノサエタ・コンシリ GP6]	10
YP_003725904.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[メタノハロビウム・エベスチガツムZ-7303] >gb AD173108.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメインタンパク質[メタノハロビウム・エベスチガツムZ-7303]	
YP_003405271.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ [ハロテリゲナ・ツルカメニカDSM 5511] >gb ADB62598.1 DEAD (配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質 [ハロテリゲナ・ツルカメニカDSM 5511]	20
YP_004244914.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ [ブルカニサエタ・モウトノブスキア768-28] >gb ADY01412.1 DEAD(配列 番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質 [ブルカニサエタ・モウトノブスキア768-28]	
YP_001540156.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[カルディビルガ・マキリンゲンシスIC-167] >sp A8MB76.1 HELS_CALMQ RecName: Full=推定ski2-型 ヘリカーゼ >gb ABW01166.1 DEAD(配列番号:2)/DEAH(配列 番号:3)ボックスヘリカーゼドメインタンパク質[カルディビルガ・マキリンゲンシス IC-167]	30
NP_618094.1	ski2-様ヘリカーゼ[メタノサルキナ・アセチボランスC2A] >sp Q8TL39.1 HELS_METAC RecName: Full=推定ski2-型 ヘリカーゼ >gb AAM06574.1 ヘリカーゼ[メタノサルキナ・アセチボランス C2A]	
YP_003900980.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ ドメイン-含有タンパク質[ブルカニサエタ・ディストリブタDSM 14429] >gb ADN49929.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメインタンパク質[ブルカニサエタ・ディストリブタDSM 14429]	
YP_003896003.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[メタノプラナス・ペトロレアリウスDSM 11571] >gb ADN37565.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメインタンパク質[メタノプラナス・ペトロレアリウスDSM 11571]	40

YP_003615773.1	DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメイン タンパク質[メタノカルドコッカス・インフェルナスME] >gb ADG12809.1 DEAD (配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質 [メタノカルドコッカス・インフェルナスME]
YP_183745.1	RNAヘリカーゼSki2-様タンパク質[サーモコッカス・コダカレンシスKOD1] >sp Q5JGV6.1 HELS_PYRKO RecName: Full=推定ski2-型 ヘリカーゼ; 含有: RecName: Full=エンドヌクレアーゼP1-PkoHel; AltName: Full=Pko Hel intein >dbj BAD85521.1 RNAヘリカーゼSki2 ホモログ[サーモコッカス・コダカレンシスKOD1]
YP_001322557.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[メタノコッカス・バニエリSB] >sp A6UN73.1 HELS_METVS RecName: Full=推定ski2-型 ヘリカーゼ >gb ABR53945.1 DEAD(配列番号:2)/DEAH(配列 番号:3)ボックスヘリカーゼドメインタンパク質[メタノコッカス・バニエリSB]
YP_002467772.1	ski2-様ヘリカーゼ[メタノスファエルラ・パルストリスE1-9c] >gb ACL18049.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメインタンパク質[メタノスファエルラ・パルストリスE1-9c]
YP_003480097.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ [ナトリアルバ・マガディ ATCC 43099] >gb ADD05535.1 DEAD(配列 番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質 [ナトリアルバ・マガディ ATCC 43099]
YP_004577043.1	ski2-型ヘリカーゼ[メタノテルモコッカス・オキナウエンシスIH1] >gb AEH07265.1 ski2-型ヘリカーゼ[メタノテルモコッカス・ オキナウエンシスIH1]
YP_004742641.1	スーパーファミリーIIヘリカーゼ[メタノコッカス・マリパルディスXI] >gb AEK19898.1 スーパーファミリーIIヘリカーゼ[メタノコッカス・ マリパルディスXI]
NP_632449.1	ski2-様ヘリカーゼ[メタノサルキナ・マゼイGo1]>sp Q8PZR7.1 HELS_METMA RecName: Full=推定ski2-型ヘリカーゼ >gb AAM30121.1 ヘリカーゼ[メタノサルキナ・マゼイGo1]
YP_001097223.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[メタノコッカス・マリパルディスC5 >gb ABO35008.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックス ヘリカーゼドメインタンパク質[メタノコッカス・マリパルディスC5]
YP_004742247.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[メタノコッカス・マリパルディスXI] >gb AEK19504.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[メタノコッカス・マリパルディスXI]

10

20

30

40

YP_004794766.1	ski2-様ヘリカーゼ[ハロアーキュラ・ヒスパニカATCC 33960] >gb AEM55778.1 ski2-様ヘリカーゼ[ハロアーキュラ・ヒスパニカATCC 33960]	
NP_988010.1	スーパーファミリーIIヘリカーゼ[メタノコッカス・マリパルディスS2] >emb CAF30446.1 スーパーファミリーIIヘリカーゼ[メタノコッカス・マリパルディスS2]	
YP_565780.1	ski2-様ヘリカーゼ[メタノコッコイデス・ブルトニーDSM 6242] >sp Q12WZ6.1 HELS_METBU RecName: Full=推定ski2-型 ヘリカーゼ >gb ABE52030.1 DEAD(配列番号:2)/DEAH(配列 番号:3)ボックスヘリカーゼ-様タンパク質[メタノコッコイデス・ブルトニーDSM 6242]	10
YP_001549808.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[メタノコッカス・マリパルディスC6] >gb ABX02576.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメインタンパク質[メタノコッカス・マリパルディスC6]	
YP_001548609.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[メタノコッカス・マリパルディスC6] >gb ABX01377.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメインタンパク質[メタノコッカス・マリパルディスC6]	20
YP_001329359.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[メタノコッカス・マリパルディスC7] >gb ABR65208.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメインタンパク質[メタノコッカス・マリパルディスC7]	
YP_004595982.1	ski2-型ヘリカーゼ[ハロピゲル・キサナデュエンシスSH-6] >gb AEH36103.1 ski2-型ヘリカーゼ[ハロピゲル・キサナデュエンシスSH-6]	
YP_656795.1	ski2-様ヘリカーゼ[ハロクアドラツム・ワルスビDSM 16790] >emb CAJ51138.1 ATP-依存型DNAヘリカーゼ[ハロクアドラツム・ ワルスビDSM 16790]	30
CCC38992.1	ATP-依存型DNAヘリカーゼHel308[ハロクアドラツム・ワルスビC23]	
YP_004035272.1	スーパーファミリーiiヘリカーゼ[ハロゲオメトリカム・ポリンクエンシスDSM 11551]	
YP_137330.1	ski2-様ヘリカーゼ[ハロアーキュラ・マリスモルツイATCC 43049] >sp Q5UYM9, 1 HELS_HALMA RecName: Full=推定ski2- 型ヘリカーゼ >gb AAV47624.1 推定ski2-型ヘリカーゼ[ハロアーキュラ・ マリスモルツイATCC 43049]	40
YP_001581577.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[ニトロソブミルス・マリチムSCM] >gb ABX12139.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックス ヘリカーゼドメインタンパク質[ニトロソブミルス・マリチムSCM1]	

EET90255.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン タンパク質[カンジダツス・ミクラケアウム・アシディフィルムARMAN-2]	
NP_376477.1	ヘリカーゼ[スルフォロブス・トコダイ株7 >sp Q974S1.1 HELS_SULTO RecName: Full=推定ski2-型ヘリカーゼ>dbj BAK54341.1 ホリディジャンクション移動ヘリカーゼ[スルフォロブス・ トコダイ株7]	
YP_001097792.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ(helices)ドメイン- 含有タンパク質[メタノコッカス・マリパルディスC5] >gb ABO35578.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックス ヘリカーゼドメインタンパク質[メタノコッカス・マリパルディスC5]	10
ZP_08667240.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン タンパク質[ニトロソブミルス種MYI] >gb EGP92972.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメインタンパク質[ニトロソブミルス種MYI]	
YP_254972.1	DNAヘリカーゼ[スルホロブス・アシドカルダリウスDMS 639] >sp Q4JC00.1 HELS_SULAC RecName: Full=推定ski2-型 ヘリカーゼ >gb AAY79679.1 DNAヘリカーゼ[スルホロブス・アシドカルダリウスDSM 639]	
EFD92533.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン タンパク質[カンジダツス・パルバルカエウム・ アシドフィラスARMAN-5]	20
YP_003176257.1	ski2-様ヘリカーゼ[ハロミクロビウム・ムコハタエイDSM 12286] >gb ACV46820.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン タンパク質[ハロミクロビウム・ムコハタエイDSM 12286]	
EGD71904.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン タンパク質[カンジダツス・パルバルカエウム・アシドフィラスARMAN-5_ '5-way FS']	
YP_001040230.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[スタフィロテルムス・マリナスF1] >gb ABN69322.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ ドメインタンパク質[スタフィロテルムス・マリナスF1]	30
ABZ07376.1	推定DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ [非培養の海洋性古細菌 HF400 AN1W133M9]	
YP_001097458.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ ドメイン-含有タンパク質[メタノコッカス・マリパルディスC5] >gb AB035243.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックス ヘリカーゼドメインタンパク質[メタノコッカス・マリパルディスC5]	
ABZ08606.1	推定DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ [非培養の海洋性古細菌HF4000_APKG3H9]	40

YP_325906.1	ski2-様ヘリカーゼ[ナトロノモナス・ファラオニスDSM 2160] >sp Q31U46.1 HELS_NATPD RecName: Full=推定ski2- 型ヘリカーゼ >emb CA148337.1 ATP-依存型DNAヘリカーゼ [ナトロノモナス・ファラオニスDSM 2160]	
YP_930665.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[ピロバキュラム・イスランディカスDSM 4184] >gb ABL88322.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックス ヘリカーゼドメインタンパク質[ピロバキュラム・イスランディカスDSM 4184]	
YP_001435870.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ [イグニコッカス・ホスピタリスKIN4/1] >gb ABU82463.1 DEAD(配列 番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質 [イグニコッカス・ホスピタリスKIN4/1]	10
YP_003668634.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[スタフィロテルムス・ヘレニカスDSM 12710] >gb AD131735. DEAD(配列番号:2)/DEAH(配列番号:3)ボックス ヘリカーゼドメインタンパク質[スタフィロテルムス・ヘレニカスDSM 12710]	
ZP_08558598.1	ski2-様ヘリカーゼ[ハロラブダス・ティアマテアSARL4B] >g EGM36528.1 ski2-様ヘリカーゼ[ハロラブダス・ティアマテアSARL4B]	
YP_002428409.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[デスルフロコックス・カムチャケンシス1221n] >gb ACL11042.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメインタンパク質[デスルフロコックス・カムチャケンシス1221n]	20
YP_004336918.1	ATP-依存型, DNA結合ヘリカーゼ[テルモプロテウス・ウゾニエンシス 768-20] >gb AEA11606.1 ATP-依存型, DNA結合ヘリカーゼ [テルモプロテウス・ウゾニエンシス768-20]	
ZP_08257442.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[カンジダツス・ニトロソアルカエウム・リミナSFB1] >gb EGG41989.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメイン-含有タンパク質[カンジダツス・ニトロソアルカエウム・ リミナSFB1]	30
YP_004459284.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[アシジアヌス・ホスピタリスW1] >gb AEE94986.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメインタンパク質[アシジアヌス・ホスピタリスW1]	
NP_558924.1	ATP-依存型, DNA結合ヘリカーゼ[ピロバキュラム・アエロフィラム株 IM2] >gb AAL63106.1 ATP-依存型, DNA結合ヘリカーゼ [ピロバキュラム・アエロフィラム株IM2]	
YP_004409449.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[メタロスファエラ・クブリナAr-4] >gb AEB94965.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメイン-含有タンパク質[メタロスファエラ・クブリナAr- 4]	40

YP_003649556.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン-含有タンパク質[テルモスファエラ・アグレガンスDSM 11486] >gb ADG90604.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[テルモスファエラ・アグレガンスDSM 11486]	
ZP_06387115.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[スルフォロブス・ソルファタリカス98/2] >gb ACX90562.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[スルフォロブス・ソルファタリカス98/2]	
2VA8_A	鎖A, Dna修復ヘリカーゼHel308>pdb 2VA8 B 鎖B, Dna修復ヘリカーゼHel308>emb CA085626.1 DNAヘリカーゼ[スルフォロブス・ソルファタリカス]	10
YP_004809267.1	ski2-型ヘリカーゼ[好塩性古細菌DL31]>gb AEN06894.1 ski2-型ヘリカーゼ[好塩性古細菌DL31]	
ADX84345.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[スルフォロブス・イスランディカスREY15A]>gb ADX81629.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[スルフォロブス・イスランディカスHVE10/4]	20
YP_002828439.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ[スルフォロブス・イスランディカスM.14.25] >ref YP_002842325.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[スルフォロブス・イスランディカスM.16.27]>gb ACP37141.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[スルフォロブス・イスランディカスM.14.25]>gb ACP54280.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[スルフォロブス・イスランディカスM.16.27]	
YP_002913571.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[スルフォロブス・イスランディカスM.16.4]>gb ACR40903.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[スルフォロブス・イスランディカスM.16.4]	30
Q97VY9.1	RecName: Full=推定ski2-型ヘリカーゼ	
YP_002841682.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[スルフォロブス・イスランディカスY.N.15.51]>gb ACP49760.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[スルフォロブス・イスランディカスY.N.15.51]	

YP_002831080.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[スルフォロブス・イスランディカスL.S.2.15]>ref YP_003418425.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[スルフォロブス・イスランディカスL.D.8.5]>gb ACP34435.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[スルフォロブス・イスランディカスL.S.2.15]>gb ADB86055.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[スルフォロブス・イスランディカスL.D.8.5]	
YP_001054984.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン-含有タンパク質[ピロバキュラム・カリディフォンティスJCM 11548]>sp A3MSA.1.1 HELS_PYRCJ RecName: Full=推定ski2-型ヘリカーゼ>gb ABO07518.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[ピロバキュラム・カリディフォンティスJCM 11548]	10
NP_343811.1	DNAヘリカーゼ関連タンパク質[スルフォロブス・ソルファタリカスP2]>ref YP_002836469.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ[スルフォロブス・イスランディカスY.G.57.14]>gb AAK42601.1 DNAヘリカーゼ関連タンパク質[スルフォロブス・ソルファタリカスP2]>gb ACP44547.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[スルフォロブス・イスランディカスY.G.57.14]	20
YP_001152379.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン-含有タンパク質[ピロバキュラム・アルセナティカムDSM 13514] >gb ABP49727.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[ピロバキュラム・アルセナティカムDSM 13514]	
YP_001191456.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン-含有タンパク質[メタロスファエラ・セデュラDSM 5348] >gb ABP95532.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[メタロスファエラ・セデュラDSM 5348]	30
NP_147034.2	ホリデージャンクション遊走ヘリカーゼ[アエロピラム・ペルニックスK1] >sp Q9YFQ8.2 HELS_AERPE RecName: Full=推定ski2-型ヘリカーゼ >dbj BAA79103.2 ホリデージャンクション遊走ヘリカーゼ[アエロピラム・ペルニックスK1]	
YP_024158.1	ski2-様ヘリカーゼ[ピコロフィラス・トリダスDSM 9790] >gb AAT43965.1 UV-防御関与ヘリカーゼ[ピコロフィラス・トリダスDSM 9790]	
YP_003816358.1	推定ski2-型ヘリカーゼ[アシディオロブス・サッカロボランス345-15] >gb ADL19327.1 推定ski2-型ヘリカーゼ[アシディオロブス・サッカロボランス345-15]	40

YP_003860265.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン タンパク質[イグニスファエラ・アグレガンスDSM 17230] >gb ADM28385.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン タンパク質[イグニスファエラ・アグレガンスDSM 17230]	
NP_394295.1	ski2-様ヘリカーゼ[サーモプラズマ・アシドフィラムDSM 1728] >sp Q9HJX7.1 HELS_THEAC RecName: Full=推定ski2-型 ヘリカーゼ >emb CAC11964.1 DNAヘリカーゼ関連タンパク質 [サーモプラズマ・アシドフィラム]	
YP_876638.1	スーパーファミリーIIヘリカーゼ[セナルカエウム・シンビオサムA] >gb ABK78334.1 スーパーファミリーIIヘリカーゼ[セナルカエウム・シンビオサム A]	10
ZP_05571398.1	ski2-様ヘリカーゼ[フェロプラズマ・アシダルマヌsfer1]	
YP_004176252.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[デスフロコックス・ムコサDSM 2162] >gb ADV64770.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメインタンパク質[デスフロコックス・ムコサDSM 2162]	
YP_001737782.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[カンジダツス・コラルカエム・クリプトフィルムOPF8] >gb ACB08099.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメインタンパク質[カンジダツス・コラルカエム・クリプトフィルム OPF8]	20
EGQ40435.1	スーパーファミリーIIヘリカーゼ[カンジダツス・ナノサリナルム種J07AB56]	
YP_002567343.1	ski2-様ヘリカーゼ[ハロラブラム・ラキウスプロフンディ ATCC 49239] >gb ACM58273.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメインタンパク質[ハロラブラム・ラキウスプロフンディ ATCC 49239]	30
YP_001793507.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[テルモプロテウス属好中球V24Sta] >gb ACB39061.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメインタンパク質[テルモプロテウス属好中球V24Sta]	
YP_003534088.1	ATP-依存型DNAヘリカーゼHel308a[ハロフェラックス・ボルカニイDS2] >gb ADE04048.1 ATP-依存型DNAヘリカーゼHel308a[ハロフェラックス・ ボルカニイDS2]	
YP_004037165.1	スーパーファミリーiiヘリカーゼ[ハログオメトリカム・ボリンクエンズDSM 11551] >gb ADQ67720.1 スーパーファミリーIIヘリカーゼ[ハログオメトリカム・ ボリンクエンズDSM 11551]	40

NP_111333.1	ski2-様ヘリカーゼ[テルモプラズマ・ボルカニウムGSS1] >sp Q97AI2.1 HELS_THEVO RecName: Full=推定ski2-型 ヘリカーゼ>dbj BAB59970.1 DNAヘリカーゼ[テルモプラズマ・ボルカニウム GSS1]	
YP_002565871.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ [ハロラブラム・ラキユスプロフンディATCC 49239] >gb ACM56801.1 DEAD (配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質 [ハロラブラム・ラキユスプロフンディATCC 49239]	
CCC39675.1	ATP-依存型DNAヘリカーゼHel308[ハロクアドラツム・ワルスビC23]	10
YP_657401.1	ATP-依存型DNAヘリカーゼ[ハロクアドラツム・ワルスビDSM 16790] >emb CAJ51759.1 ATP-依存型DNAヘリカーゼ[ハロクアドラツム・ ワルスビDSM 16790]	
YP_003535028.1	ATP依存型DNAヘリカーゼHel308b[ハロフェラックス・ボルカニイ DS2]>gb ADE02398.1 ATP-依存型DNAヘリカーゼHEL308b [ハロフェラックス・ボルカニイDS2]	
YP_003706863.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[メタノコッカス・ボルタック A3]>gb ADI35890.1 DEAD(配列番号:2)/DEAH(配列番号: 3)ボックスヘリカーゼドメインタンパク質[メタノコッカス・ボルタックA3]	20
ABD17736.1	ヘリカーゼ[メタノコッカス・ボルタエPS]	
NP_613398.1	スーパーファミリーIIヘリカーゼ[メタノピラス・カンドレリ AV19]>gb AAM01328.1 推定スーパーファミリーIIヘリカーゼ [メタノピラス・カンドレリAV19]	
CBH38575.1	推定ski2-型ヘリカーゼ[非培養の古細菌]	
EEZ93258.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン タンパク質[カンジダツス・パルバルカエウム・アシドフィラムARMAN-4]	
YP_004004246.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン タンパク質[メタノテルムス・フェルビダスDSM 2088]>gb ADP77484.1 DEAD(配列番号:2)/DEAH(配列番号: 3)ボックスヘリカーゼドメインタンパク質[メタノテルムス・フェルビダスDSM 2088]	30
YP_003850109.1	ヘリカーゼ[メタノテルモバクテル・マルブルゲネシス種 Marburg]>gb ADL58796.1 推定ヘリカーゼ [メタノテルモバクテル・マルブルゲネシス種Marburg]	
YP_003424423.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[メタノプレビバクテル・ルミナンティウム M1]>gb ADC47531.1 DEAD(配列番号:2)/DEAH(配列番号: 3)ボックスヘリカーゼドメイン-含有タンパク質[メタノプレビバクテル・ ルミナンティウムM1]	
YP_004291107.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン- 含有タンパク質[メタノバクテリウム種AL- 21]>gb ADZ10135.1 DEAD(配列番号:2)/DEAH(配列番号: 3)ボックスヘリカーゼドメインタンパク質[メタノバクテリウム種AL-21]	40
YP_447162.1	ヘリカーゼ[メタノスファエラ・スタッドマナエDSM 3091] >gb ABC56519.1 推定ヘリカーゼ[メタノスファエラ・スタッドマナエ DSM 3091]	

YP_004519549.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン-含有タンパク質[メタノバクテリウム種SWAN-1]>gb AEG17748.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[メタノバクテリウム種SWAN-1]	
NP_275949.1	DNAヘリカーゼ関連タンパク質[メタノテルモバクテル・テルマウトトロフィカス種δH]>sp O26901.1 HELS...METTH RecName: Full=推定ski2-型ヘリカーゼ>gb AAB85310.1 DNAヘリカーゼ関連タンパク質[メタノテルモバクテル・テルマウトトロフィカス種δH]	
ZP_05975717.2	推定Ski2-型ヘリカーゼ[メタノブレビバクテル・スミチDSM 2374]>gb EFC93382.1 推定Ski2-型ヘリカーゼ [メタノブレビバクテル・スミチDSM 2374]	10
ZP_03607647.1	仮想タンパク質METSMIALI...00751 [メタノブレビバクテル・スミチDSM 2375]>gb EEE41862.1 仮想タンパク質 METSMIALI_00751[メタノブレビバクテル・スミチDSM 2375]	
YP_001273412.1	ATP-依存型ヘリカーゼ[メタノブレビバクテル・スミチATCC 35061]>gb ABQ87044.1 ATP-依存型ヘリカーゼ [メタノブレビバクテル・スミチATCC 35061]	
YP_003247505.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[メタノカルドコッカス・ブルカニウスM7]>gb ACX73023.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[メタノカルドコッカス・ブルカニウスM7]	20
NP_248116.1	SKI2ファミリーヘリカーゼ[メタノカルドコッカス・ジャナシイDSM 2661]>sp Q58524.1 HELS...METJA RecName: Full=推定ski2-型ヘリカーゼ; 含有: RecName: Full=エンドヌクレアーゼPI-MjaHel; AltName: Full=Mja Helインテイン; AltName: Full=Mja Pep3インテイン>gb AAB99126.1 推定SKI2-ファミリーヘリカーゼ [メタノカルドコッカス・ジャナシイDSM 26621]	
YP_001324295.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン-含有タンパク質[メタノコッカス・アエオリカス・ナカイ-3]>gb ABR 55683.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[メタノコッカス・アエオリカス・ナカイ-3]	30
YP_003536960.1	プレ-mRNAスプライシングヘリカーゼ[ハロフェラックス・ボルカニイDS2]>gb ADE02332.1 プレ-mRNAスプライシングヘリカーゼ[ハロフェラックス・ボルカニイDS2]	
YP_003131029.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[ハロラブダス・ウタヘンシスDSM 12940]>gb ACV12296.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[ハロラブダス・ウタヘンシスDSM 12940]	
YP_002567151.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ[ハロルブラム・ラキウスプロファンディATCC 49239]>gb ACM58081.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[ハロルブラム・ラキウスプロファンディATCC 49239]	40
YP_004035351.1	スーパーファミリーiiヘリカーゼ[ハロゲオメトリカム・ボリンクエンズDSM 1151]>gb ADQ65912.1 スーパーファミリーIIヘリカーゼ[ハロゲオメトリカム・ボリンクエンズDSM 11551]	

YP_004808851.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン-含有タンパク質[好塩性古細菌DL31]>gb AEN06478.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[好塩性古細菌DL31]
XP_002716686.1	推定: DNAポリメラーゼ θ アイソフォームI[オリクトラグス・クニキュラス]
YP_656834.1	ATP依存型DNAヘリカーゼ[ハロクアドラツム・ワルスビDSM 16790]>emb CAJ51176.1 ATP-依存型DNAヘリカーゼ[ハロクアドラツム・ワルスビDSM 16790]
XP_003248103.1	推定: DNAポリメラーゼ θ -様アイソフォームI[アシルトシフォン・ピサム]
ABC72356.1	ATP-依存型DNAヘリカーゼ[ハロクアドラツム・ワルスビ]
CCC39031.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ[ハロクアドラツム・ワルスビC23]
XP_001165150.2	推定: DNAポリメラーゼ θ アイソフォームI[チンパンジー(Pan troglodytes)]
XP_003225852.1	推定: DNAポリメラーゼ θ -様[グリーンアノール(Anolis carolinensis)]
XP_615375.3	推定: DNAポリメラーゼ θ [ウシ(Bos taurus)]>ref[XP_002684835.1] 推定: ポリメラーゼ(DNA指向), θ -様[ウシ]>gb DAA33456.1 ポリメラーゼ(DNA指向), θ -様[ウシ]
XP_002813286.1	推定: 低品質タンパク質: DNAポリメラーゼ θ -様[オランウータン(Pongo abelii)]
AAR08421.2	DNAポリメラーゼ θ [ヒト(Homo sapiens)]
EAW79510.1	ポリメラーゼ(DNA指向) θ , アイソフォームCRA_a [ヒト]
NP_955452.3	DNAポリメラーゼ θ [ヒト] >sp 075417.2 DPOLQ...ヒトRecName: Full=DNAポリメラーゼ θ ; AltName: Full=DNAポリメラーゼeta >gb aa172289.1 ポリメラーゼ(DNA指向), θ [合成ポリヌクレオチド]
NP_001099348.1	DNAポリメラーゼ θ [ラット(Rattus norvegicus)]>gb EDM11249.41 ポリメラーゼ(DNA指向), θ (推定), アイソフォームCRA_a [ラット]
XP_003341262.1	推定: 低品質タンパク質: DNAポリメラーゼ θ -様[ハイイロジネズミオポッサム(Monodelphis domestica)]
XP_001502374.3	推定: DNAポリメラーゼ θ [ウマ(Equus caballus)]
XP_545125.3	推定: 低品質タンパク質: DNAポリメラーゼ θ [イヌ(Canis lupus familiaris)]
XP_002928855.1	推定: 低品質タンパク質: DNAポリメラーゼ θ -様[ジャイアントパンダ(Ailuropoda melanoleuca)]
NP_084253.1	DNAポリメラーゼ θ アイソフォームI[マウス(Mus musculus)]>gb AAL77225.1 DNAポリメラーゼ θ [マウス] >gb EDK 97951.1 ポリメラーゼ(DNA指向), θ , アイソフォームCRA_a [マウス]>gb AA138361.1 ポリメラーゼ(DNA指向), θ [マウス]>gbAA157901.1 ポリメラーゼ(DNA指向), θ [マウス]
AAK39635.1	DNAポリメラーゼ θ [ヒト]
AAN393838.1	DNAポリメラーゼQ[マウス]
XP_003412882.1	推定: DNAポリメラーゼ θ [アフリカゾウ(Loxodonta africana)]

10

20

30

40

YP_003735206.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン-含有タンパク質[ハラルカリコッカス・ジェオタガリ B3]>gb ADJ13414.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[ハラルカリコッカス・ジェオタガリ B3]	
YP_004794841.1	プレ-mRNAスプライシングヘリカーゼ[ハロアーキュラ・ヒスパニカ ATCC 33960]>gb AEM55853.1 プレ-mRNAスプライシングヘリカーゼ[ハロアーキュラ・ヒスパニカ ATCC 33960]	
XP_416549.2	推定: DNAポリメラーゼ θ に類似[セキショクヤケイ (<i>Gallus gallus</i>)]	
XP_003427319.1	推定: ヘリカーゼ POLQ-様アイソフォーム2[キョウソヤドリコバチ (<i>Nasonia vitripennis</i>)]	
XP_003202748.1	推定: DNAポリメラーゼ θ -様[シチメンチョウ (<i>Meleagris gallopavo</i>)]	
XP_969311.1	推定: DNAポリメラーゼ θ に類似[コクヌストモドキ (<i>Tribolium castaneum</i>)]>gb EEZ97532.1 仮想タンパク質 TcasGA2..TC011380 [コクヌストモドキ]	10
ZP_08046037.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[ハラダブタツス・パウチハロフィラス DX253]>gb EFW90685.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[ハラダブタツス・パウチハロフィラス DX253]	
YP_461714.1	ヘリカーゼ[シントロファス・アシディトロフィカス SB]>gb ABC77546.1 ヘリカーゼ[シントロファス・アシディトロフィカス SB]	
YP_003176510.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ[ハロミクロビウム・ムコハタエイ DSM 12286]>gb AAV47694.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[ハロミクロビウム・ムコハタエイ DSM 12286]	20
YP_137400.1	プレ-mRNAスプライシングヘリカーゼ[ハロアーキュラ・マリスモルツイ ATCC43049]>gb AAV47694.1 プレ-mRNAスプライシングヘリカーゼ[ハロアーキュラ・マリスモルツイ ATCC 43049]	
NP_001184156.1	ポリメラーゼ(DNA指向), θ [アフリカツメガエル (<i>Xenopus tropicalis</i>) (Silurana)]	
NP_280861.1	プレ-mRNAスプライシングヘリカーゼ[ハロバクテリウム種 NRC-1]>ref YP...001689987.1 ATP-依存型DNAヘリカーゼ[ハロバクテリウム・サリナラム RI]>gb AAG20341.1 プレ-mRNAスプライシングヘリカーゼ[ハロバクテリウム種 NRC-1]>emb CAP14641.1 ATP-依存型DNAヘリカーゼ[ハロバクテリウム・サリナラム RI]	30
YP_004595640.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン-含有タンパク質[ハロピゲル・キサナデュエンシス SH 6]>gb AEH35761.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[ハロピゲル・キサナデュエンシス SH-6]	
XP_001521144.2	推定: DNAポリメラーゼ θ , 部分[カモノハシ (<i>Ornithorhynchus anatinus</i>)]	
XP_003261953.1	推定: DNAポリメラーゼ θ , 部分[テナガザル (<i>Nomascus leucogenys</i>)]	
XP_001358456.2	GA19301 [ウスグロショウジョウバエ (<i>Drosophila pseudoobscura pseudoobscura</i>)]>gb EAI...27595.2 GA 19301 [ウスグロショウジョウバエ]	
ZP_0856003.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[ハロラブダス・ティアマテア SARL4B]>gb EGM34502.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[ハロラブダス・ティアマテア SARL4B]	40
XP_002187783.1	推定: ポリメラーゼに類似(DNA指向), θ [キンカチョウ (<i>Taeniopygia guttata</i>)]	

XP_002112587.1	仮想タンパク質TRIADDRAFT...25163[センモウヒラムシ(<i>Trichoplax adhaerens</i>)]>gb EDV24697.1 仮想タンパク質 TRIADDRAFT...25163 [センモウヒラムシ]	
YP_003405139.1	DEATH/DEAHボックスヘリカーゼ[ハロテリゲナ・ツルクメニカDSM 5511]>gb ADB62466.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[ハロテリゲナ・ツルクメニカDSM 5511]	
EGV92665.1	DNAポリメラーゼ θ [チャイニーズハムスター(<i>Cricetulus griseus</i>)]	
CBY24305.1	無名のタンパク質産物[ワカレオタマボヤ(<i>Oikopleura dioica</i>)]	
YP_003130565.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[ハロラブダス・ウタヘンシスDSM 12940]>gb ACV11832.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[ハロラブダス・ウタヘンシスDSM 12940]	10
YP_003479811.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ[ナトリアルバ・マガディATCC 43099]>gb ADD05249.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[ナトリアルバ・マガディATCC 43099]	
EFB22383.1	仮想タンパク質PANDA_000253[ジャイアントパンダ]	
YP_003357334.1	推定ATP-依存型ヘリカーゼ[メタノセラ・パルディコラSANAЕ]>dbj BA162351.1 推定ATP-依存型ヘリカーゼ[メタノセラ・パルディコラSANAЕ]	
YP_325942.1	ATP-依存型DNAヘリカーゼ2[ナトロノモナス・ファラオニスDSM 2160]>emb CA148373.2 ATP-依存型DNAヘリカーゼ2[ナトロノモナス・ファラオニスDSM 2160]	20
XP_002912509.1	推定：低品質タンパク質：ヘリカーゼPOLQ様[パンダ]	
XP_002704678.1	推定：ヘリカーゼ, POLQ-様[ウシ]	
CAE47762.2	ヒトDNA-指向ポリメラーゼ θ (POLQ)に類似した新規タンパク質[DANIO rerio]	
XP_003205636.1	推定：ヘリカーゼPOLQ-様[シチメンチョウ]	
XP_544959.2	推定：ヘリカーゼPOLQ-様[イヌ]	
EFX86757.1	仮想タンパク質DAPPUDRAFT_312857[ミジンコ(<i>Daphnia pulex</i>)]	
YP_003389641.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ[スピロソマ・リングアルDSM 74]>gb ADB40842.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[スピロソマ・リングアルDSM 74]	30
XP_002602932.1	仮想タンパク質BRAFLDRAFT_251779[ブランキオストマ・フロリダエ]>gb EEN58944.1 仮想タンパク質 BRAFLDRAFT_251779[ブランキオストマ・フロリダエ]	
YP_004144962.1	ペプチダーゼC14カスパーゼ触媒サブユニットp20 [メゾリゾビウム・シセリ・ピオバル・ピセルラエWSM1271]>ref YP_004614892.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメイン-含有タンパク質[メゾリゾビウム・オポチュニスツムWSM2075]>gb ADV14912.1 ペプチダーゼC14カスパーゼ触媒サブユニットp20 [メゾリゾビウム・シセリ・ピオバル・ピセルラエWSM1271]>gb AEH90798.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[メゾリゾビウム・オポチュニスツムWSM2075]	40
XP_002124758.1	推定：DNAポリメラーゼ θ に類似[カタユウレイボヤ(<i>Ciona intestinalis</i>)]	

XP_694437.5	推定：DNAポリメラーゼθ[ゼブラフィッシュ(Danio rerio)]
XP_420565.1	推定：DNAヘリカーゼHEL308に類似[セキショクヤケイ]
XP_003129397.1	推定：ヘリカーゼPOLQ-様[イノシシ(Sus scrofa)]
EDL20278.1	mCG128467, アイソフォームCRA_b[マウス]
XP_001517710.2	推定：ヘリカーゼPOLQ, 部分[カモノハシ(Ornithorhynchus anatinus)]
AAH82601.1	ヘリカーゼ, mus308-様(ショウジョウバエ)[マウス]
XP_003384429.1	推定：DNAポリメラーゼθ-様[アムフィメドン・クイーンズランディカ(Amphimedon queenslandica)]
XP_003221282.1	推定：ヘリカーゼPOLQ-様[グリーンアノール]
NP_524333.1	変異原-感受性308[キイロショウジョウバエ] >gb AAB67306.1 Mus308[キイロショウジョウバエ] >gb ACH92234.1 FI03732p[キイロショウジョウバエ]
AAX33507.1	LP14642p[キイロショウジョウバエ]
NP_001074576.1	ヘリカーゼPOLQ-様[マウス] >sp Q2VPA6.2 HELQ_マウスRecName : Full=ヘリカーゼPOLQ-様 ; AltName : Full=Mus308-様ヘリカーゼ : AltName : Full-POLQ-様ヘリカーゼ >gb AAI09171.2 ヘリカーゼ, mus308-様(ショウジョウバエ)[マウス]
YP_003523727.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[シデルオキシダンス・リソトロフィカスES-1] >gb ADE11340.1 DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼドメインタンパク質[シデルオキシダンス・リソトロフィカスES-1]
XP_002120889.1	推定：DNAヘリカーゼHEL308に類似[カタユウレイボヤ]
XP_001892566.1	III型制限酵素, resサブユニットファミリータンパク質[マレー糸状虫(Brugia malayi)] >gb EDP38603.1 III型制限酵素, resサブユニットファミリータンパク質[マレー糸状虫]
ABZ09232.1	推定ヘリカーゼ保存されたC-末端ドメインタンパク質[非培養の海洋性古細菌HF4000_APKG7F11]
XP_002814981.1	推定：低品質タンパク質：ヘリカーゼPOLQ-様[オランウータン]
XP_002717082.1	推定：DNAヘリカーゼHEL308 [アナウサギ(Oryctolagus cuniculus)]
XP_001104832.1	推定：ヘリカーゼ, POLQ-様[アカゲザル]
AAL85274.1	DNAヘリカーゼHEL308[ヒト]
NP_598375.2	ヘリカーゼPOLQ-様[ヒト] >gb EAX05934.1 DNAヘリカーゼHEL308, アイソフォームCRA_a[ヒト] >gb AAI41525.1 ヘリカーゼ, POLQ-様[合成ポリヌクレオチド]
Q8TDG4.2	RecName : Full=ヘリカーゼPOLQ-様 ; AltName : Full=Mus308-様ヘリカーゼ ; AltName : Full=POLQ-様ヘリカーゼ
XP_003265889.1	推定：ヘリカーゼPOLQ [テナガザル]
XP_002745688.1	推定：ヘリカーゼPOLQ-様 [マーモセット(Callithrix jacchus)]
XP_003310356.1	推定：低品質タンパク質：ヘリカーゼPOLQ-様[チンパンジー]

10

20

30

40

NP_001014156.2	ヘリカーゼ, POLQ-様[ラット] >ref XP_001060858.1 推定: ヘリカーゼ, POLQ-様[ラット] >gb EDL99554.1 rCG37823, アイソフォームCRA_c [ラット]
XP_001850567.1	ATP-依存型DNAヘリカーゼMER3 [熱帯イエカ(Culex quinquefasciatus)] >gb EDS32308.1 ATP-依存型DNAヘリカーゼMER3[熱帯 イエカ]
XP_003427318.1	推定: ヘリカーゼPOLQ-様アイソフォーム1[キョウソヤドリコバチ]
XP_003143912.1	仮想タンパク質LOAG_08332 [ロア・ロア(Loa loa)] >gb EFO20157.1 仮想タンパク質LOAG_08332 [ロア・ロア]
CAG11187.1	無名のタンパク質生成物[テトラオドン・ニグロビリディス]
XP_001111254.2	推定: DNAポリメラーゼ θ アイソフォーム2[アカゲザル]
XP_003414242.1	推定: ヘリカーゼPOLQ [アフリカゾウ]
XP_002681870.1	推定タンパク質[ネグレリア・グルベリ] >gb EFC49126.1 推定 タンパク質[ネグレリア・グルベリ]
EAX05935.1	DNAヘリカーゼHEL308, アイソフォームCRA_b [ヒト]
AAH59917.1	Asce3タンパク質[マウス]
ZP_07082808.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ ドメインタンパク質[スフィンゴバクテリウム・スピリチボラム ATCC 33861] >gb EFK55937.1 DEAD(配列番号:2)/DEAH(配列番号:3) ボックスヘリカーゼドメインタンパク質[スフィンゴバクテリウム・ スピリチボラム ATCC 33861]
XP_001494572.3	推定: 低品質タンパク質: ヘリカーゼPOLQ-様 [ウマ]
XP_002714920.1	推定: 活性化シグナルコインテグレーター1複合体サブユニット3 [アナウサギ]
XP_002598278.1	仮想タンパク質BRAFLDRAFT_204526[ブランキオストマ・ フロリダエ] >gb EEN54290.1 仮想タンパク質 BRAFLDRAFT_204526 [ブランキオストマ・フロリダエ]
XP_001943294.1	推定: ヘリカーゼPOLQ-様アイソフォーム1[エンドウヒゲナガアブラムシ (Acyrtosiphon pisum)] >ref XP_003240510.1 推定: ヘリカーゼPOLQ-様 アイソフォーム2[エンドウヒゲナガアブラムシ]
XP_002803889.1	推定: 活性化シグナルコインテグレーター1複合体サブユニット3- 様[アカゲザル]
XP_001651546.1	DNAポリメラーゼ θ [ネッタイシマカ] >gb EAT42599.1 DNA ポリメラーゼ θ [ネッタイシマカ]

10

20

30

40

CAA11679.1	RNAヘリカーゼ[ヒト]
XP_002837795.1	仮想タンパク質[トリュフ(Tuber melanosporum)Mel28] >emb CAZ1986.1 無名のタンパク質生成物 [トリュフ]
EGT47882.1	仮想タンパク質CAEBREN_02542 [カエノルハブディティス・ブレネリ]
EDL99655.1	活性化シグナルコインテグレーター1複合体サブユニット3(推定), アイソフォームCRA_b[ラット]
NP_932124.2	活性化シグナルコインテグレーター1複合体サブユニット3[マウス]
EDL05054.1	mCG119534 [マウス]
gi 352115865 ZP_08963952.1	DEAD(配列番号:2)/DEAH(配列番号:3)ボックスヘリカーゼ ドメインタンパク質[ナトリアルバ・ベリルブラムDSM 15624]

10

【0121】

Hel308ヘリカーゼ、並びにHel308モチーフ、及び拡大されたHel308モチーフの更なる実施態様は、下記表2に示している。

表2：Hel308ヘリカーゼ、Hel308モチーフ、及び拡大されたHel308モチーフの例

20

【表2】

GI番号	ヘリ カーゼ	名称	% 同一性	% 同一性	Hel308モチーフ	拡張された Hel308モチーフ
			Hel308 Pfu	Hel308 Mbu		
GI:12168 9265	Hel308 Mbu	メタノコッコイデス・ ブルトニー	37%		QMAGRAGR (配列番号:4)	QMAGRGRP (配列番号:5)
GI:18202 135	Hel308 Pfu	パイロコッカス・ フリオサスDSM 3638		37%	QMLGRAGR (配列番号:6)	QMLGRGRP (配列番号:7)
GI: 49014464 1	Hel308 Hvo	ハロフェラックス・ ボルカニ	34%	41%	QMMGRAGR (配列番号:8)	QMMGRGRP (配列番号:9)
GI: 50639166 4	Hel308 Hla	ハロルブラム・ ラキウスプロファンディ	35%	42%	QMCGRAGR (配列番号:10)	QMGRGRP (配列番号:11)
GI:11857 6895	Hel308 Csy	セナルカエウム・ シンピオサム	34%	34%	QLCGRAGR (配列番号:12)	QLCGRGRP (配列番号:13)
GI:18202 627	Hel308 Sso	スルフォロプス・ ソルファタリカス	35%	33%	QMSGRAGR (配列番号:14)	QMSGRGRP (配列番号:15)

30

40

GI番号	ヘリ カーゼ	名称	% 同一性	% 同一性	Hel308モチーフ	拡張された Hel308モチーフ
**	Hel308 Mfr	メタノゲニウム・ フリギダム	37%	44%	QMAGRAGR (配列番号:16)	QMAGRAGRP (配列番号:17)
GI: 50363337 1	Hel308 Mok	メタノテルモコッカ ス・オキナウエンシス	37%	34%	QCIGRAGR (配列番号:18)	QCIGRAGRP (配列番号:19)
GI:33391 1571	Hel308 Mig	メタノトリス・ イグネウスKol 5	40%	35%	QCIGRAGR (配列番号:20)	QCIGRAGRP (配列番号:21)
GI:24010 2927	Hel308 Tga	サーモコッカス・ ガンマトレランス EJ3	60%	38%	QMMGRAGR (配列番号:22)	QMMGRAGRP (配列番号:23)
GI:31523 1273	Hel308 Tba	サーモコッカス・ バロフィラスMP	57%	35%	QMIGRAGR (配列番号:24)	QMIGRAGRP (配列番号:25)
GI:24239 8904	Hel308 Tsi	サーモコッカス・ シビリカスMM 739	56%	35%	QMMGRAGR (配列番号:26)	QMMGRAGRP (配列番号:27)
GI: 12172332 5	Hel308 Mba	メタノサルキナ・ バーケリ株Fusaro	39%	60%	QMAGRAGR (配列番号:28)	QMAGRAGRP (配列番号:29)
GI:24418 451	Hel308 Mac	メタノサルキナ・ アセチボランス	38%	60%	QMAGRAGR (配列番号:30)	QMAGRAGRP (配列番号:31)
GI:29449 5240	Hel308 Mmah	メタノハロフィラス・ マヒDSM 5219	38%	60%	QMAGRAGR (配列番号:32)	QMAGRAGRP (配列番号:33)
GI:24418 450	Hel308 Mmaz	メタノサルキナ・ マゼイ	38%	60%	QMAGRAGR (配列番号:34)	QMAGRAGRP (配列番号:35)
GI:11666 5562	Hel308 Mth	メタノサエタ・ テルモフィラPT	39%	46%	QMAGRAGR (配列番号:36)	QMAGRAGRP (配列番号:37)
GI:33647 7283	Hel308 Mzh	メタノサルサム・ ジリナエDSM 4017	39%	57%	QMAGRAGR (配列番号:38)	QMAGRAGRP (配列番号:39)
GI:29867 4154	Hel308 Mev	メタノハロピウム・ エベスチガツムZ- 7303	38%	61%	QMAGRAGR (配列番号:40)	QMAGRAGRP (配列番号:41)
GI:50019 5255	Hel308 Mma	メタノコッカス・ マリパルディス	36%	32%	QCIGRAGR (配列番号:42)	QCIGRAGRP (配列番号:43)
GI:49038 8033	Hel308 Nma	ナトリアルバ・ マガディ	37%	43%	QMMGRAGR (配列番号:44)	QMMGRAGRP (配列番号:45)

10

20

30

40

GI番号	ヘリカーゼ	名称	% 同一性	% 同一性	Hel308モチーフ	拡張された Hel308モチーフ
GI: 22674060 6	Hel308 Mbo	メタノレグラ・ ブーンキ6A8	38%	45%	QMAGRAGR (配列番号:46)	QMAGRGRP (配列番号:47)
GI:49757 3451	Hel308 Fac	フェロプラズマ・ アシダルマヌス	34%	32%	QMIGRAGR (配列番号:48)	QMIGRGRP (配列番号:49)
GI:25681 0263	Hel308 Mfe	メタノカルドコッカス・ フェルベンスAG86	40%	35%	QCIGRAGR (配列番号:50)	QCIGRGRP (配列番号:51)
GI: 18202572	Hel308 Mja	メタノカルドコッカ ス・ジャンナキ	24%	22%	QCIGRAGR (配列番号:52)	QCIGRGRP (配列番号:53)
GI:50286 4579	Hel308 Min	メタノカルドコッカ ス・インフェルナス	41%	33%	QCIGRAGR (配列番号:54)	QCIGRGRP (配列番号:55)
GI:88603 707	Hel308 Mhu	メタノスピリラム・ フンガテイJF-1	36%	40%	QMAGRAGR (配列番号:56)	QMAGRGRP (配列番号:57)
GI: 63555245 4	Hel308 Afu	アルカエグロブス・ フルギダスDSM 4304	40%	40%	QMAGRAGR (配列番号:58)	QMAGRGRP (配列番号:59)
GI:50270 9689	Hel308 Htu	ハロテリゲナ・ ツルクメニカ	35%	43%	QMAGRAGR (配列番号:60)	QMAGRGRP (配列番号:61)
GI:49525 7384	Hel308 Hpa	ハルアダプツタス・ パウシハロフィラス DX253	38%	45%	QMFGRAGR (配列番号:62)	QMFGRGRP (配列番号:63)
GI: 15791161	Ski2- like helicase	ハロバクテリウム種 NRC-1	36.8%	42.0%	QMFGRAGR (配列番号:64)	QMFGRGRP (配列番号:65)

**：このHel308ヘリカーゼの配列に関する更なる詳細については、国際公開公報WO 2013/057495を参照。

【0122】

ポリヌクレオチド結合活性及びヘリカーゼ酵素活性を維持しているHel308ヘリカーゼの変種又は変異体もまた、本実施態様において使用することができる。かかる変種又は変異体は、未変性のHel308ヘリカーゼをコードしている核酸の位置指定変異誘発を含む、当該技術分野において周知である方法に従い入手することができる(Zoller, M.J.の文献、Cur. Opin. Biotechnol., 3:348-354, (1992))。

【0123】

加えて、先に注記し且つ当該技術分野において公知であるように、Hel308ヘリカーゼは、SF2ファミリーであり、且つ3'→5'ヘリカーゼである(これはまたA型ヘリカーゼとも称され得る)。様々なヘリカーゼのコアドメインは、ヌクレオチド結合及び加水分解に関与したWalker Aモチーフ(これはまたモチーフIとも称され得る)及びWalker Bモチーフ(これはまたモチーフIIとも称される)、並びにモチーフVIを含む、RecA結合フォールドなどの、互いに共通のモチーフを含むことができる。更なる詳細については、Flechsigsらの文献、「スーパーファミリー2ヘリカーゼタンパク質における高次構造移動のインシリコ研究(In Silico Investigation of Conformational Motions in Superfamily 2 Helicase Proteins)」、PLoS One: 6(7): e 21809 (2011)を参照されたい。加えてファミリーSF2のヘリカーゼは、9つの保存モチーフを共有することができ、これはQ、I、Ia、Ib、II、III、IV、V、及びVIと称される。モチーフII(DEAD(配列番号:2)又はDEAH(配列番号:3)又はDEXH)の配列のために、SF2ヘリカーゼファミリーはまた、DEAD-ボックス(配列番号:2)タン

10

20

30

40

50

パク質又はDEAH-ボックス(配列番号:3)ヘリカーゼとも称される。SF2ファミリーに含まれるヘリカーゼは、RecQ-様ファミリー及びSnf2-様酵素を含む。XPDファミリーなどのいくつかの例外はあるが、多くのSF2ヘリカーゼは、A型である。SF2ファミリーのX線結晶解析研究は、保存されたヘリカーゼモチーフは、タンパク質の三次構造において密接に関連していること、及びこれらは大型の機能ドメインを形成し得ることを示唆している。更なる詳細に関しては、Tutejaらの文献、「DNAヘリカーゼの解明：モチーフ、構造、機序及び機能(Unraveling DNA Helicases: Motif, structure, mechanism and function)」、European Journal of Biochemistry 271(10): 1849-1863 (2004)、及びHallらの文献、「ヘリカーゼモチーフ：DNAをほどく力をもたらしエンジン(Helicase motifs: the engine that powers DNA unwinding)」、Molecular Microbiology 34: 867-877 (1999)を参照されたい。図16は、Tutejaらの文献の改作であるが、これはHel308がその一員であるSF2ファミリーにおいて同定された様々なモチーフ、例えばDEAD-ボックス(配列番号:2)ヘリカーゼを概略的に描いている。Tutejaの文献で説明されたように、オープンボックスは、保存されたモチーフを表している。各ヘリカーゼモチーフのコンセンサス配列は、1文字コードで表され、例えば、図16において“C”は、D、E、H、K、又はRであり；図16において“O”は、S又はTであり；並びに、図16において“X”は、任意のアミノ酸であることができる。モチーフに割り当てられた名称、例えばQ、I、Ia、Ib、II、III、IV、V、及びVIもまた、図16に示されている。更に先に記したように、モチーフIは、Walker Aモチーフと称され、且つTutejaの文献において、ATPaseA Walker Iと称され、並びにモチーフIIは、Walker Bモチーフと称され、且つTutejaの文献において、ATPaseB Walker IIと称される。矢印が示しているモチーフ間の番号は、それらのモチーフ間に配置されたアミノ酸残基の代表的範囲である。

【0124】

加えて、WO 2013/057495に記載されたように、Hel308ヘリカーゼは、アミノ酸モチーフ【化2】

Q-X1-X2-G-R-A-G-R (配列番号:66),

を含むことができ、ここでX1は、C、M、又はLであることができ；X1は、Cであることができ；X2は、A、F、M、C、V、L、I、S、T、P、又はRなど、疎水性残基又は中性残基を含む、任意の残基であることができる。任意に上記モチーフ内の末端Rは、Pに結合することができる。

【0125】

本明細書に提供される内容及び指針を考慮し、当業者は、先の例証的Hel308ヘリカーゼの1以上との配列同一性又はアラインメントを決定することにより、参照ヘリカーゼは、Hel308ヘリカーゼであるかどうかを決定することができる。

【0126】

加えて本明細書に提供される内容及び指針を考慮し、当業者は、例えば加水分解を遅くすることができる様式で別のタンパク質の相同モチーフに類似したHel308のモチーフを変異することにより、Hel308ヘリカーゼが行う加水分解工程を遅くすることにより、細孔を通じたポリヌクレオチドのフラクショナル転位を遅くするようにHel308ヘリカーゼを適切に変異することができる。一例として、Tanakaらの文献、「組換え依存型DNA複製に必須の大腸菌PriAタンパク質のATPase/ヘリカーゼモチーフ変異体(ATPase/helicase motif mutants of Escherichia coli PriA protein essential for recombination-dependent DNA replication)」、Genes to Cells 8: 251-261 (2003)は、その保存されたATPase/DNAヘリカーゼモチーフ、すなわちWalker A、B、及びQXXGRXGRモチーフ内にアミノ酸置換を保持するPriAタンパク質(DEXH-型ヘリカーゼ)の変異体を説明している。Tanakaの文献によると、特定の変異体は、特定条件でのATPの加水分解で高度に損なわれ、且つWalker A及びWalker B変異体タンパク質は全て、特定の条件で、高度に減弱されたDNAヘリカーゼ活性を示した。従って、Tanakaらの文献に開示されたものに類似しているHel308ヘリカーゼ

のWalker A及びWalker Bモチーフに対する変異は、DNAヘリカーゼ活性を減弱するか又はATP加水分解を遅くすることを予想することができ、このことは、細孔を通じたポリヌクレオチドのフラクショナル転位を遅くし、結果的にそのポリヌクレオチドの特徴決定を増強することが予想され得ることを、想定することができる。別の例として、Hishidaらの文献、「ホリデージャンクションの分岐移動を促進するRuvBタンパク質のWalkerモチーフAの役割：RuvBのATP結合活性、ATP加水分解活性、及びDNA結合活性に影響を及ぼすWalkerモチーフA変異(Role of Walker Motif A of RuvB Protein in Promoting Branch Migration of Holliday Junctions: Walker motif A mutations affect ATP binding, ATP hydrolyzing, and DNA binding activities of RuvB)」、Journal of Biological Chemistry 274(36): 25335-25342 (1999)は、ATP-依存性六量体DNAヘリカーゼである大腸菌RuvBタンパク質の変異体を説明している。Hishidaの文献によると、WalkerモチーフAへの特定の点変異は、ATP加水分解及びATP結合のRuvB活性、並びにDNA結合の活性、六量体形成、及び分岐移動の促進に影響を及ぼした。従って、Hishidaの文献に開示されたものに類似しているHel308ヘリカーゼのWalker Aモチーフへの変異は、ATP加水分解及びATP結合に影響を及ぼすことを予想することができ、このことは、いくつかの実施態様において、細孔を通じたポリヌクレオチドのフラクショナル転位を遅くし、結果的にそのポリヌクレオチドの特徴決定を増強することが予想され得ることを、想定することができる。

10

【0127】

従って本開示は、標的ポリヌクレオチドを特徴決定する方法を提供する。この方法は、(a)Hel308ヘリカーゼ及び標的ポリヌクレオチドと接触している細孔を横切るポテンシャルの差を印加する工程；(b)該細孔を通じた該標的ポリヌクレオチドの該Hel308ヘリカーゼの1以上のフラクショナル転位工程により発生した1以上のシグナルを測定する工程；並びに、(c)該フラクショナル転位工程により発生した該1以上のシグナルから該標的ポリヌクレオチドを特徴決定する工程：を含むことができる。

20

【0128】

本開示は、ポテンシャルの差が電気ポテンシャルの差を含む、標的ポリヌクレオチドを特徴決定する方法を更に提供する。同じくシグナルが、電氣的シグナル又は光学的シグナルを含む、標的ポリヌクレオチドを特徴決定する方法も提供する。電氣的シグナルは、電流、電圧、トンネル電流、抵抗、ポテンシャル、電圧、コンダクタンス、及び横方向電氣的測定値から選択された測定値であることができる。電氣的シグナルは、細孔を通過する電氣的電流を含む。

30

【0129】

別の態様において、本開示は、フラクショナル転位工程が、Hel308ヘリカーゼの完全転位サイクルの第一のフラクショナル転位工程を含む、標的ポリヌクレオチドを特徴決定する方法を提供する。フラクショナル転位工程はまた、Hel308ヘリカーゼの完全転位サイクルの第二のフラクショナル転位工程を含むことができる。標的ポリヌクレオチドの転位は、細孔を通じて転位するポリヌクレオチド上の加力の反対の方向であるか、又は細孔を通じて転位するポリヌクレオチド上の加力の方向である。

【0130】

追加的に、標的ポリヌクレオチドの1以上のヌクレオチド残基が、完全転位サイクルから得られた単独の電氣的シグナルを使用する1以上のヌクレオチドの特徴決定と比べ50%を上回る精度で、完全転位サイクルの2つのフラクショナル工程から得られた電氣的シグナルを用いて特徴決定される、標的ポリヌクレオチドを特徴決定する方法が提供される。

40

【0131】

更に、細孔が生物学的細孔である、標的ポリヌクレオチドを特徴決定する方法が提供される。生物学的細孔は、ポリペプチド細孔又はポリヌクレオチド細孔であることができる。一部の態様において、ポリペプチド細孔は、5ヌクレオチド以下の狭窄ゾーンを有する。他の態様において、ポリペプチド細孔は、マイコバクテリウム・スメグマチスポリンA(MspA)を含む。MspAは、配列番号:1のアミノ酸配列、又は配列番号:1と少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、

50

少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、又は少なくとも70%の相同性を有するアミノ酸配列を有することができる。

【0132】

同じく、細孔が、固体状態の細孔又は生物学的及び固体状態のハイブリッド細孔である、標的ポリヌクレオチドを特徴決定する方法も提供する。生物学的及び固体状態のハイブリッド細孔は、ポリペプチドー固体状態のハイブリッド細孔又はポリヌクレオチドー固体状態のハイブリッド細孔を含む。

【0133】

本開示は加えて、Hel308ヘリカーゼが、表1及び表2に示されたヘリカーゼ又はそれらの変種である、標的ポリヌクレオチドを特徴決定する方法を提供する。更に標的ポリヌクレオチドが、1本鎖、2本鎖、及び部分的2本鎖ポリヌクレオチドからなる群から選択される、標的ポリヌクレオチドを特徴決定する方法を提供する。

10

【0134】

一部の実施態様において、該フラクショナル転位工程により発生した該1以上のシグナルからのポリヌクレオチドの特徴決定は、改変されたViterbiアルゴリズムを適用することを含む。

【0135】

一部の実施態様において、本方法は更に、(d)工程(c)の後に、該細孔を通じた該標的ポリヌクレオチドの該Hel308ヘリカーゼによる1以上のフラクショナル転位工程の時機を変更するために、少なくとも1つのパラメータを変更する工程；並びに、(e)変更された少なくとも1つのパラメータを使用し、工程(a)－(c)を繰り返す工程を含む。本方法は、工程(c)及び(e)の間に発生したシグナルを組合せること、並びに組合せたシグナルを基に該標的ポリヌクレオチドを特徴決定することを更に含むことができる。一部の実施態様において、変更された少なくとも1種のパラメータは、温度、塩濃度、補因子濃度、ATP生成物(例えば無機リン酸)の濃度、ADPの濃度、pH、及び使用した特定のHel308ヘリカーゼからなる群から選択される。

20

【0136】

一部の実施態様において、該特徴決定は、1以上のシグナル中のレベルを検出及び同定し、並びに検出され且つ同定されたレベルを基に標的ポリヌクレオチドの配列を決定及び出力することを含む。

30

【0137】

図示されたように、該1以上のシグナルのレベルの検出及び同定は、完全レベル、フラクショナルレベル、全レベル、及びレベル識別子の1以上の出力を含む。

【0138】

前述の検出及び同定されたレベルを基にした標的ポリヌクレオチドの配列の決定及び出力は、完全レベル、フラクショナルレベル、全レベル、及びレベル識別子の1以上を入力として採用し、該入力を基に複数の配列を呼び出し、並びに呼び出された複数の配列に関する信頼情報を基に呼び出された配列の少なくとも1種を選択及び出力することを含む。

【0139】

前述の検出及び同定されたレベルを基にした標的ポリヌクレオチドの配列の決定及び出力は、完全レベル、フラクショナルレベル、全レベル、及びレベル識別子の1以上を入力として採用し、該入力を基に複数の配列を呼び出し、並びに呼び出された複数の配列の一部に関する信頼情報を基に複数の呼び出された配列のお互いの部分を選択及び連鎖させることを含む。

40

【0140】

前述の検出及び同定されたレベルを基にした標的ポリヌクレオチドの配列の決定及び出力は、完全レベル、フラクショナルレベル、全レベル、及びレベル識別子の1以上を入力として採用し、該入力を基に複数の配列を呼び出し、呼び出された配列をモデル配列と比較し、並びに呼び出された配列のモデル配列との比較に関する信頼情報を基に呼び出された配列の少なくとも1種を選択及び出力することを含む。

50

【0141】

前述の検出及び同定されたレベルを基にした標的ポリヌクレオチドの配列の決定及び出力は、完全レベル、フラクショナルレベル、全レベル、及びレベル識別子の1以上を入力として採用し、該入力を基に複数の配列を呼び出し、呼び出された配列をモデル配列と比較し、並びに呼び出された複数の配列の一部のモデル配列との比較に関する信頼情報を基に複数の呼び出された配列のお互いの部分を選択及び連鎖することを含む。

【0142】

本開示はまた、細孔を通じた標的ポリヌクレオチドのフラクショナル転位工程を調節する方法を提供する。この方法は：(a)Hel308ヘリカーゼ及び標的ポリヌクレオチドと接触している細孔を横切るポテンシャルの差を印加する工程；(b)Hel308ヘリカーゼをHel308ヘリカーゼ基質の参照濃度とは異なる濃度の基質と接触させる工程であって、この基質濃度が、参照濃度と比べた基質濃度の差異に比例してフラクショナル転位工程の持続期間の変化を生じる工程；並びに、(c)細孔を通じた標的ポリヌクレオチドのフラクショナル転位工程により発生した1以上のシグナルを測定する工程：を含むことができる。工程(b)は同様に、フラクショナル転位工程の持続期間の変化を達成するために基質類縁体又はインヒビターを使用することを含むことができる。従って、本明細書記載又は当該技術分野において公知の任意の基質類縁体又はヌクレオチドインヒビターを、Hel308ヘリカーゼ基質のいずれか、参照濃度として使用されるHel308基質、又はHel308ヘリカーゼ基質のいずれかと参照濃度として使用されるHel308基質の両方として、フラクショナル転位工程の調節に関する開示の方法において使用することができる。

【0143】

ポリヌクレオチドのフラクショナル転位工程の調節が可能であるHel308ヘリカーゼ基質は、ヘリカーゼにより加水分解されることが可能である、ヌクレオチド又はヌクレオチド類縁体であることができる。ヌクレオチド基質は、2本鎖又は部分的に2本鎖のポリヌクレオチドをほどこか、又は細孔を通じて1本鎖ポリヌクレオチドを転位するためのエネルギーを提供する。Hel308ヘリカーゼに共通の基質は、例えばATPを含む。Hel308ヘリカーゼ基質はまた、ヘリカーゼにより加水分解されることが可能であるヌクレオチド及びヌクレオチド類縁体を含む。

【0144】

本明細書記載のように、ヌクレオチド基質結合に関係している1以上のフラクショナル転位工程に関する滞在時間は、Hel308ヘリカーゼ基質の濃度に反比例することができる。例えば、試験される一部の条件下で、1つのヌクレオチド転位につき認められる2つのフラクショナル転位工程の一方のみの滞在時間は、Hel308ヘリカーゼ基質の濃度に反比例する。従って1つの転位工程は、別の転位工程がないならば、基質濃度に感受性があり得る。

【0145】

フラクショナル転位工程の異なる長さを得るためのフラクショナル転位工程の調節は、Hel308ヘリカーゼ基質の濃度を変更することにより達成することができる。調節の程度又は大きさは、当業者が、所望の標的ポリヌクレオチド特徴決定に適した特定の長さのフラクショナル転位工程を選択することができるように、決定することができる。調節の程度は、基質の参照濃度とは異なるHel308ヘリカーゼ基質の濃度中にHel308ヘリカーゼを配置することにより決定することができる。参照濃度と比較した基質濃度の変化は、参照濃度と比較した基質濃度の差に比例するフラクショナル転位工程の異なる滞在時間を生じる。

【0146】

従って細孔を通じた標的ポリヌクレオチドのフラクショナル転位工程は、基質の参照濃度とは異なるHel308ヘリカーゼ基質の濃度を使用することにより、調節することができる。ヘリカーゼ溶液内の他の成分又は反応条件もまた、フラクショナル転位工程の滞在時間、従って1回の転位サイクルに関するフラクショナル転位工程の長さを変更するために使用することができる。異なるフラクショナル転位工程は同様に、標的ポリヌクレオチド特徴決定の精度を上昇するために、追加のシグナル情報を獲得するために使用することができる。

。

10

20

30

40

50

【0147】

例えばHel308ヘリカーゼへの基質結合及びヘリカーゼによる基質加水分解の速度論に影響を及ぼす反応の成分及び反応条件は、フラクショナル転位工程の滞在時間を変更するために使用することができる。このような他の要因は、例えば、反応条件の温度、2価の金属濃度を含む金属濃度、イオン濃度、溶媒粘度を含む。加水分解工程は、例えば前記要因及び条件により、並びにリン酸及び／又はピロリン酸の濃度により、影響され得る。加えて細孔を横切る電圧は、例えば基質結合及び／又はHel308ヘリカーゼの滞在時間を構成するヘリカーゼの休止期(pause)に影響を及ぼすことができる。他の要因は、例えば、pH、陽イオンの種類又は2価陽イオンの濃度及び種類、ヘリカーゼ変異などを含み、全て滞在時間に影響を及ぼし得る。これに関して、例えばピロリン酸濃度の上昇は、Hel308ヘリカーゼの触媒速度を遅くするため、従って滞在時間を増大するために使用することができる。更に例えば、オルトバナジウム酸ナトリウム及びアデノシン5'-(β , γ -イミド)三リン酸リチウム塩水和物も、ヘリカーゼ活性を遅くするために使用することができる。ヘリカーゼ活性を調節するためのピロリン酸及びヌクレオチド類縁体の使用は、実施例Vにおいて以下に例示されている。

10

【0148】

逐次工程間の電流差は増加するので、データ解析のためのフラクショナル状態を使用する恩恵も増大する。第一の近似で、フラクショナル転位工程は、隣接の完全転位工程の間にある値をとるであろう。フラクショナル転位工程が、1/2ヌクレオチド(0.3オングストローム)よりもはるかに少ないならば、場合によって、又は多くの場合においてさえも、小数値は認めることが困難であるか不可能でさえある。フラクショナル転位工程が正確にヌクレオチドの1/2の長さであるならば、生じる電流は、平均して、完全ヌクレオチド工程に対応する先行する電流値及び後続の電流値とは最大に異なることができる。この酵素の修飾は、ナノメーターのごく一部(fractions)によりポリマーサブユニットの再配置を可能にすることができる。これは、ナノポアの制限のある狭窄に対し、酵素の活性加水分解部位の相対高さを増加又は減少する酵素修飾を通じて生じ得る。一部の実施態様において、これは、ヘリカーゼのアミノ酸の付加若しくは除去、又はより大きい流体力学半径を持つアミノ酸の置換を通じて、達成することができる。他の実施態様において、これは、ナノポアの周縁の静電的反発力又は引力を変更することができるアミノ酸電荷の変更を通じて、達成することができる。いずれかの理論に結びつけられることを欲するものではないが、「グリップ-ベース」仮説が正しいならば(図3を参照しより詳細に説明される)、特定の変異は、ヘリカーゼがヘリカーゼ-ポリヌクレオチド複合体を上向きに押す程度に影響を及ぼすことは可能であり、これはヌクレオチドのz-軸転位率の変化に翻訳することができる。

20

30

【0149】

フラクショナル転位工程の持続期間を調和することが意図され：ヘリカーゼATPaseドメインに対する特定の変異は、ATPが加水分解される速度に影響を及ぼすと予想することは妥当である。これは次に、フラクショナル転位工程の1つの滞在時間に影響を及ぼすと予想されるであろう。例えば加水分解速度が遅くされたならば、フラクショナル転位工程の1つの滞在時間は、増加すると予想される。他の変異は、ATPがヘリカーゼに結合する速度(k_{on})に影響を及ぼすであろう。この場合ATPが結合するのに要する時間が増加するので、対応するフラクショナル転位工程の滞在時間は増加するであろう。

40

【0150】

Hel308ヘリカーゼの参照濃度は、例えば、標的ポリヌクレオチド特徴決定において一般に使用される基質の量であることができるか、又は異なることができる。例えば一般に使用されるHel308ヘリカーゼ基質の濃度が1.0mMである場合、1mMは、参照濃度に相当する。この参照濃度は、経験的に誘導されるか、又は当該技術分野において周知の報告から入手することができる。この具体例において、1mM以外の基質の濃度は、参照濃度とは異なるHel308ヘリカーゼ基質であろう。更に以下に説明するように、様々な濃度のHel308ヘリカーゼ基質及び参照基質を利用し、フラクショナル転位工程の変更の量を調節又は決定する

50

ことができる。

【0151】

Hel308ヘリカーゼ基質濃度及び参照基質濃度の濃度は、両方の濃度が飽和濃度でない限りは、変更することができる。図示されたように、Hel308ヘリカーゼ基質の飽和濃度は、約1mMのヌクレオチド基質である。従って参照濃度が1mMである場合、変更されるべきHel308ヘリカーゼ基質濃度は、濃度1mM未満の任意の濃度、例えば0.1 μ M、1.0 μ M、10 μ M、100 μ M、200 μ M、300 μ M、400 μ M、500 μ M、600 μ M、700 μ M、800 μ M、900 μ Mを含むことができる。Hel308ヘリカーゼ基質濃度及び／又は参照に応じて、他の例証的濃度は、例えば1.0mM、2.0mM、3.0mM、4.0mM及び4.9mM又はそれよりも少ないことができる。同様に、Hel308ヘリカーゼ基質及び参照基質濃度の両方の濃度は、それらが異なる限りは、飽和でない濃度であることができる。従ってHel308ヘリカーゼ基質濃度及び参照濃度は、先に列挙した例証的濃度のいずれか、並びに例えば、0.01 μ M～5mMの範囲の任意の濃度及びこの範囲内の全ての濃度など、任意の濃度であることができる。

10

【0152】

フラクショナル転位工程を調節するための本開示の方法は、標的ポリヌクレオチドを特徴決定する方法に関して先に説明したように行うことができる。一旦特定の必要性に適したHel308ヘリカーゼ基質濃度が決定されたならば、その基質濃度を、標的ポリヌクレオチドを特徴決定するために本明細書に記載した方法において利用することができる。類似の様式で、同様の決定を、例えば特定の必要性に適している成分濃度又は反応条件を決定するための基質結合及び加水分解の速度論に影響を及ぼす反応の成分及び条件により、実行することができる。この好適な濃度又は条件は次に、標的ポリヌクレオチドを特徴決定するための本開示の方法において利用することができる。新たな基質濃度、反応成分濃度及び／又は反応条件は、実施例IXを参照し以下に説明されるような様式で決定の精度を増強するための追加のシグナル情報を提供することができる異なる滞在時間を生じるであろう。

20

【0153】

従って本開示は、1以上のフラクショナル転位工程の1以上のシグナルからの標的ポリヌクレオチドの特徴決定を更に含む、標的ポリヌクレオチドのフラクショナル転位工程を調節する方法を提供する。この特徴決定は、以下の1以上の同定を含むことができる：(1)標的ポリヌクレオチドの配列；(2)標的ポリヌクレオチドの修飾；(3)標的ポリヌクレオチドの長さ；(4)標的ポリヌクレオチドの同一性；(5)標的ポリヌクレオチドの給源；又は、(6)標的ポリヌクレオチドの二次構造。

30

【0154】

本開示はまた、方法が、電気ポテンシャルの差を含むポテンシャルの差を利用する、標的ポリヌクレオチドのフラクショナル転位工程を調節する方法を提供する。フラクショナル転位工程により発生したシグナルが、電氣的シグナル又は光学的シグナルを含む、標的ポリヌクレオチドのフラクショナル転位工程を調節する方法を更に提供する。加えて、電氣的シグナルが、電流、電圧、トンネル電流、抵抗、ポテンシャル、電圧、コンダクタンス、及び横方向電氣的測定値から選択された測定値である、標的ポリヌクレオチドのフラクショナル転位工程を調節する方法を更に提供する。電氣的シグナルはまた、細孔を通過する電氣的電流であることができる。

40

【0155】

基質濃度が、Hel308ヘリカーゼ基質のサブ飽和濃度である、標的ポリヌクレオチドのフラクショナル転位工程を調節する方法をまた更に提供する。一部の実施態様において、参照濃度は、Hel308ヘリカーゼ基質の飽和濃度である。別の態様において、基質濃度及び参照濃度の両方は、Hel308ヘリカーゼ基質のサブ飽和濃度である。更にHel308ヘリカーゼ基質が、アデノシン三リン酸(ATP)である、標的ポリヌクレオチドのフラクショナル転位工程を調節する方法をまた更に提供する。

【0156】

フラクショナル転位工程が、Hel308ヘリカーゼの完全転位サイクルの第一のフラクシヨ

50

ナル転位工程又はHel308ヘリカーゼの完全転位サイクルの第二のフラクショナル転位工程を含む、標的ポリヌクレオチドのフラクショナル転位工程を調節する方法をまた更に提供する。標的ポリヌクレオチドの転位は、細孔を通じて転位するポリヌクレオチド上の加力の反対の方向であるか、又は細孔を通じて転位するポリヌクレオチド上の加力の方向であることができる。

【0157】

また本開示により、標的ポリヌクレオチドの1以上のヌクレオチド残基が、完全転位サイクルから得られた単独の電氣的シグナルを使用する1以上のヌクレオチドの特徴決定と比べ50%より大きい精度で、完全転位サイクルの2つのフラクショナル工程から得られた電氣的シグナルを用いて特徴決定される、標的ポリヌクレオチドのフラクショナル転位工程を調節する方法を更に提供する。本開示の方法の一部の態様において、標的ポリヌクレオチド中の1以上のヌクレオチド残基は、参照濃度と比べより低い基質濃度でより大きい精度で特徴決定される。

10

【0158】

加えて細孔が生物学的細孔である、標的ポリヌクレオチドのフラクショナル転位工程を調節する方法を提供する。生物学的細孔は、ポリペプチド細孔又はポリヌクレオチド細孔であることができる。一部の態様において、ポリペプチド細孔は、5ヌクレオチド以下の狭窄ゾーンを有する。他の態様において、ポリペプチド細孔は、マイコバクテリウム・スメグマチスポリンA(MspA)を含む。MspAは、配列番号:1のアミノ酸配列、又は配列番号:1との少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、又は少なくとも99%の相同性を有するアミノ酸配列を有することができる。

20

【0159】

細孔が、固体状態の細孔又は生物学的及び固体状態のハイブリッド細孔である、標的ポリヌクレオチドのフラクショナル転位工程を調節する方法をまた更に提供する。生物学的及び固体状態のハイブリッド細孔は、ポリペプチドー固体状態のハイブリッド細孔又はポリヌクレオチドー固体状態のハイブリッド細孔であることができる。

30

【0160】

また本方法のHel308ヘリカーゼは、表1及び表2に示されたヘリカーゼ又はそれらの変種である、標的ポリヌクレオチドのフラクショナル転位工程を調節する方法を提供する。標的ポリヌクレオチドは、1本鎖、2本鎖、及び部分的2本鎖ポリヌクレオチドからなる群から選択される。

【0161】

本開示は更に、標的ポリヌクレオチドを特徴決定するための組成物を提供する。この組成物は、1mM未満のATPの溶液又はヌクレオチド類縁体の溶液に含まれた、細孔、Hel308ヘリカーゼ及び標的ポリヌクレオチドを含む。本組成物の一部の態様において、1mM未満のATPの溶液は、0.1 μ M、1.0 μ M、10 μ M、100 μ M、0.5mM又は0.9mMのATPである。

【0162】

40

本開示の組成物は、ポリヌクレオチドの特徴決定又は標的ポリヌクレオチド転位のフラクショナル転位工程の調節のための本開示の方法において使用される、先に又は以下に説明された成分のいずれかを含むことができる。例えば、組成物は、先に説明された細孔を含むことができる。本明細書に提供される内容及び指針に従い、細孔は、例えばポリペプチド細孔又はポリヌクレオチド細孔などの生物学的細孔であることができる。或いは、細孔は、先に説明された固体状態の細孔又はハイブリッド細孔であることができる。

【0163】

加えて本組成物は、特徴決定のための標的ポリヌクレオチド、Hel308ヘリカーゼ及びHel308ヘリカーゼ基質を含むであろう。細孔のように、標的ポリヌクレオチド、Hel308ヘリカーゼ及びHel308ヘリカーゼ基質は、例証的ポリヌクレオチド、Hel308ヘリカーゼ、基質

50

及び本明細書に記載された変種及び類縁体、並びに当該技術分野において周知のもののいずれかであることができる。

【0164】

従って本開示は、細孔が生物学的細孔である、標的ポリヌクレオチドを特徴決定するための組成物を提供する。生物学的細孔は、ポリペプチド細孔又はポリヌクレオチド細孔であることができる。ポリペプチド細孔は、5ヌクレオチド以下の狭窄ゾーンを有し、且つマイコバクテリウム・スメグマチスポリンA(MspA)であることができる。MspAは、配列番号:1のアミノ酸配列、又は配列番号:1と少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、又は少なくとも99%の相同性を有するアミノ酸配列を有する。

10

【0165】

また細孔が固体状態の細孔である、標的ポリヌクレオチドを特徴決定するための組成物を提供する。加えて、細孔が生物学的及び固体状態のハイブリッド細孔である、標的ポリヌクレオチドを特徴決定するための組成物を提供する。生物学的及び固体状態のハイブリッド細孔は、ポリペプチド-固体状態のハイブリッド細孔又はポリヌクレオチド-固体状態のハイブリッド細孔であることができる。

【0166】

Hel308ヘリカーゼが表1及び表2に示されたヘリカーゼ又はそれらの変種である、標的ポリヌクレオチドを特徴決定するための組成物を更に提供する。加えて、標的ポリヌクレオチドが、1本鎖、2本鎖、及び部分的2本鎖ポリヌクレオチドからなる群から選択される、標的ポリヌクレオチドを特徴決定するための組成物を提供する。

20

【0167】

本開示の様々な実施態様の活性に実質的に影響を及ぼさない修飾もまた、本明細書に提供される開示の規定内に含まれることが理解される。従って、以下の実施例は、例証することを意図し、本開示を限定するものではない。

【実施例】

【0168】

(実施例I: Hel308ヘリカーゼによる フラクショナル転位工程)

30

実施例Iは、例証的Hel308ヘリカーゼにより観察された フラクショナル転位工程 を説明している。

【0169】

脂質二重層は、1,2-ジフィタノイル-sn-グリセロ-3-ホスホコリン(Avanti Polar Lipids社)から形成した。この二重層は、テフロン中に水平方向に直径~20ミクロンの開口部が広がっている。M2-NNN-MspAを、この二重層の基底側(grounded side)に、濃度~2.5ng/mlで添加した。一旦単独の細孔が挿入されたならば、更なる挿入を避けるために、このコンパートメントを実験用緩衝液でフラッシュした。Axopatch-200Bパッチクランプ増幅器(Axon Instruments社)で、180mVの二重層を横切る電圧を印加し、且つイオン電流を測定した。類縁体シグナルは、4-極Besselフィルターにより50kHzで低域フィルターを通過させ、その後低域フィルター周波数を5回デジタル化した。データ獲得は、LabWindows/CVI (National Instruments社)により書かれた特注のソフトウェアにより制御した。二重層の両側の~60 µlコンパートメントは、0.3M KCl、1mM EDTA、1mM DTT、10mM MgCl₂、及び10mM HEPES/KOHで緩衝された実験用緩衝液(pH8.0)を含んだ。野生型Hel308 Tga又は野生型Phi29ポリメラーゼのいずれかを、モーターとして使用した。Hel308 Tgaの存在下で、緩衝液に、1mM ATPを補充した。Phi29の存在下で、緩衝液に、dCTP、dATP、dTTP及びdGTPの各々10 µMを補充した。

40

【0170】

図2A-2Cは、一部の実施態様に従う、Phi29ポリメラーゼとHel308 Tgaヘリカーゼの転位事象の比較を示している。図2Aは、phi29 DNAポリメラーゼ(DNAP)で観察された転位工程

50

と比較した、Hel308 Tgaヘリカーゼで観察されたフラクショナル転位工程を示している。
転位するポリヌクレオチド

【化 3】

(配列番号: 67:

/5Phos/AAACCTTCCXCCCGTACCGTGCCGTACCGTTCCGTTCGGTACCGTA
TTTTTTTT TCTCACTATCGCATTCTCATGCAGGTCGTAGCC、式中X=脱塩基)

は、コレステロール-含有ポリヌクレオチド

【化 4】

(配列番号: 68:

CCTGCATGAGAATGCGATAGTGAGA TTTTTTTTTTTTTTTTTTTT-CholTEG)

へハイブリダイズさせた。MspA-M2ナノポアを使用した。Hel308 Tgaヘリカーゼポリヌクレオチド転位に関して認められたレベルの数は、phi29 DNAPについて認められたレベルの数のほぼ2倍であった。トレース間に引かれた線は、対応するレベルを示す。phi29トレース(上側)はコンセンサスであるのに対し、Hel308ヘリカーゼトレース(下側)は、測定された単独の転位事象である。コンセンサスとは、同じ配列由来の複数の読み値からの信頼できる検出されたレベルの組合せをいうことができる。かかる組合せは潜在的に、単回の読み値よりもより信頼することができ、その理由はこれは、単-分子転位、例えば当該技術分野において公知であるようなヌクレオチド「スキッピング」又はヌクレオチド「トグリ

【0 1 7 1】

図2Bは、分子モーターとしてHel308 Tgaヘリカーゼを使用し認められたものと、分子モーターとしてPhi29ポリメラーゼを用いMspA-M2ナノポアを通じて転位する1本鎖ポリヌクレオチド鋳型により生じる推定電流レベルとを比較した、Hel308 Tgaヘリカーゼにより認められたフラクショナル転位工程を示す。転位するポリヌクレオチド

【化 5】

(配列番号: 69: /5Phos/

CTCACCTATCCTTCCACTXXCCCCCTTTGGGTTTAAATTTTTTCAGATCTCACTAT
C TTTTTAAAGTTT TCTCACTATCGCATTCTCATGCAGGTCGTAGCC、式中X=脱塩基)

は、コレステロール-含有ポリヌクレオチド

【化 6】

(配列番号: 68:

CCTGCATGAGAATGCGATAGTGAGATTTTTTTTTTTTTTTTTTTT-CholTEG)

へハイブリダイズさせた。phi29トレース(上側)は推定であるのに対し、Hel308ヘリカーゼトレース(下側)は、測定された転位事象である。コンセンサスのような推定は、単-分子転位、例えば当該技術分野において公知であるようなヌクレオチド「スキッピング」又はヌクレオチド「トグリ」などにより生じ得る誤差を、必ずしも含まない。推定とは、先に収集された、k-mer表を基にシミュレートされたデータをいうことができる。Phi29推定されたパターンは、1塩基につき完全工程を基にし、且つ予測することができるパターンの型が、完全工程分子モーターであることを図示している。比較すると、Hel308 Tgaヘリカーゼはフラクショナル工程を有することが、明確に認められる。

【0 1 7 2】

図2Cは、phi29 DNAPにより認められた転位工程と比較した、Hel308 Tgaヘリカーゼによ

10

20

30

40

50

り認められたフラクショナル転位工程を示している。転位するポリヌクレオチド配列は、
【化 7】

配列番号: 70:

/5Phos/CATCATCATCATCATCATXXCCCCCTAAACAAGAATACCACGACTAGCAT
TTTTCAGATCTCACTATCGCATTCTCATGCAGGTCGTAGCC

である。この転位するポリヌクレオチドは、コレステロール-含有ポリヌクレオチド
【化 8】

(配列番号: 68:

CCTGCATGAGAATGCGATAGTGAGA TTTTTTTTTTTTTTTTTTTT-CholTEG)

にハイブリダイズさせ、且つMspA-M2ナノポアを通過させた。単純な反復配列5'-CAT-3'を使用し、反復パターンを示した。Hel308 Tgaヘリカーゼポリヌクレオチド転位に関して認められたレベルの数は、phi29 DNAPについて認められたレベルの数の2倍であった。phi29及びHel308ヘリカーゼの両トレースは、コンセンサストレースであった。コンセンサストレースの使用は、異なる分子モーター間の転位工程サイズの比較を促進することができ、且つそうでなければ潜在的に解釈を複雑にするスキップ及びトグルなどのアーチファクトを減少若しくは除去することができる。

【0173】

いずれかの理論に結びつけられることを欲するものではないが、フラクショナル転位工程の更なる説明において、「グリップ-ベースの」機序が提唱される。図3は、一部の実施態様に従う、提唱されたフラクショナル転位工程に関する「グリップ-ベースの」機序を示している。ポリヌクレオチド(黒実線)は、ヘリカーゼ(塗りつぶし横線による形状)により結合されている。ATP結合時(工程1)に、ヘリカーゼは、高次構造の変化を受ける(工程2)。ポリヌクレオチドはヘリカーゼによりグリップされているので、ヘリカーゼに対するこのポリヌクレオチドの位置は、必ずしも変化しない。ヘリカーゼ上の参照点(灰色三角)は、ヘリカーゼによりグリップされたポリヌクレオチドに対して移動しない(グリップされたポリヌクレオチド上の参照点、灰色四角を参照)。ヘリカーゼの高次構造の変化は、ヘリカーゼ-ポリヌクレオチド複合体を、ナノポアの上面に押し出し、それと共に細孔狭窄中のポリヌクレオチドを引き寄せる(矢頭を持つ黒色線で指摘した黒色線)。第二のポリヌクレオチド参照点(白色円)は、高次構造変化(工程2)時の細孔狭窄に対するポリヌクレオチド移動を示し、これはフラクショナル工程に関する測定された電流変化を生じる。最後に、ATPは、加水分解され、且つヘリカーゼは、ポリヌクレオチドに沿って、転位する(工程3)。これは、ポリヌクレオチドが、ヘリカーゼ及び細孔に対し、完全ヌクレオチドを移動することを引き起こす。まとめると、第一のフラクショナル転位工程において、Hel308ヘリカーゼはATPに結合し、ヘリカーゼによりグリップされたポリヌクレオチドを引き寄せる高次構造変化を受け、且つフラクショナル的な1ヌクレオチドだけポリヌクレオチドをシフトし、これは次に測定可能な電流変化を生じる。第二のフラクショナル転位工程において、ATPは加水分解され、且つHel308ヘリカーゼは、ナノポアを通る1ヌクレオチドの転位を完了する。他の機序は、このフラクショナル転位工程の知見を説明するために適切に使用することができる。

【0174】

(実施例II：ATP濃度とフラクショナル転位工程の関係)

実施例IIは、フラクショナル転位工程の滞在時間に対するATP濃度の作用を説明する。

【0175】

フラクショナル転位工程の生化学的機序を更に解明するために、フラクショナル転位工程の滞在時間を、ATPの変動する濃度下で試験した。シスウェル及びトランスウェルを、最初に400mM KCl、10mM HEPESからなる緩衝液(pH8)で満たした。DPhPCからなる脂質二重

10

20

30

40

50

層を、直径～25 μmのテフロン細孔の上をヘキサデカン及び脂質の混合物で塗装することにより形成し、脂質二重層とテフロン細孔の間のギガオーム密封を確実にするために、コンダクタンス測定を行った。全ての電氣的測定は、シスウェル及びトランスウェルに接続したAg/AgCl電極対へ接続した、Axopatch 200Bパッチクランプ増幅器を用いて行った。膜形成後、MspAナノポアを、シスウェルへ注入し、そこで脂質二重層へのナノポア組込みを、コンダクタンス測定によりモニタリングした。単独のナノポアの二重層への組込み時には、シスチャンバーは、複数の細孔の挿入を防止するために、灌流した。次に1本鎖ポリヌクレオチドを、このシスチャンバーへ最終濃度10nMで注入し、電圧をこの膜を横切って印加し、且つ細孔を通じたポリヌクレオチド転位を、過渡電流反応により検出した。ポリヌクレオチド転位検出時に、電圧は、0Vに設定し、且つ1mM MgCl₂、115nM Hel308ヘリカーゼ、及び様々な濃度のATP(10 μM、30 μM、100 μM、及び1mM)を、シスウェルへ注入した。その後電圧を、保持電位(0.01、0.1、及び1mM ATPに関して140mV；0.03mM ATPに関して180mV)に設定し、電流を記録した。図14A-14Dに示したように、シスウェルへの注入前に、転位するポリヌクレオチド

10

【化 9】

(配列番号: 71:

/5Phos/CATCATCATCATCATCATXXCCCCCTAAACAAGAATACCACGACTAGCAT
TTTTCAGATCTCACTATCGCATTCTCATGCAGGTCGTAGCC、式中X=脱塩基)

20

は、コレステロール-含有ポリヌクレオチド

【化 10】

(配列番号: 68:

CCTGCATGAGAATGCGATAGTGAGA TTTTTTTTTTTTTTTTTTTTTT-CholTEG)

へハイブリダイズさせた(本明細書別書により詳細に記載した)。この様式で、ポリヌクレオチドの5'末端は、ナノポアを通じて最初に転位され、次にHel308ヘリカーゼの処理により、ナノポアを通じて引き戻された。Axopatch増幅器は、50kHzのサンプリング速度で及び低域フィルター10kHzにより、このシステムの電流反応を記録した。フラクショナル転位工程を含む、ナノポアを通じたポリヌクレオチドのHel308ヘリカーゼ処理に起因した工程移行は、この周波数範囲内で明確に同定可能であった。実験後、コンピュータアルゴリズムを使用し、ポリヌクレオチド転位事象を同定した。隣接値間の統計学的有意性を決定するために当該技術分野において一般に公知であるスチューデントt-検定を用い、統計学的に有意な電流レベルが、これらの転位事象内で同定された(更なる詳細は、Carterらの文献、本明細書別所に引用、又はJohn E. Freundの文献、「数学的統計(Mathematical Statistics)」、第5版、Prentice Hallを参照されたい)。この特定の配列から観察された電流に関して、処理されたヌクレオチドが存在するとして同定されたものよりも、ほぼ2倍多い統計学的に有意な電流レベルが存在し、トポロジー(電流レベルのピーク及びトラフ)は、直接観察により測定された場合、単-工程分子モーターに関するよりも、各ピーク間及び各トラフ間で、ほぼ2倍多いレベルを有した。

30

40

【0176】

実験誤差を減少するために、ナノポアを通じたポリヌクレオチド転位の持続期間のデータ解析を、ヌクレオチド転位の大きい解像度の領域において行った。鎖配列決定、及び特にポリヌクレオチドのナノポア配列決定において脱塩基領域は、隣接ポリヌクレオチド配列のシグナル-対-ノイズ比と比べ、ブロックされるイオン流れの有意差のために、比較的高いシグナル-対-ノイズ比を生じ得る。この理由のために、脱塩基領域の近傍内の統計学的に有意なレベルは、一部の隠された「ノイジー」効果によるよりも、ナノポアを通るヌクレオチド処理による可能性が潜在的により大きい。この理由のために、電流レベル持続期間-ベースのデータ解析のために脱塩基電流ピークを取り巻き且つこれを含む27電流レ

50

ベルの持続期間を選択した。

【0177】

図4A及び図4Bは、一部の実施態様に従う、ATP濃度のフラクショナル転位工程の滞在時間に対する例証的作用を示している。図4Aにおいて、フラクショナル転位工程によりナノポアを通じて転位するポリヌクレオチド配列からの脱塩基電流ピークの電流レベルを取り巻く電流レベルの包含は、連続して1から27としてラベルされ、且つ持続期間の中央値をプロットした。

【0178】

電流レベルは、隣接する電流値間の統計学的有意性を決定するために、スチューデント t-検定を使用するアルゴリズムにより検出した(Carterらの文献、本明細書別所の引用を参照されたい)。速度の閾値及びカイ二乗最小化を含む他の技術は、これを可能にし、これらも全て、ヌクレオチド処理に関連した電流変化に関して、並びに画像処理における検出工程に関して、当該技術分野において公知である。これらのレベルの各々に関連した持続期間が存在し、且つ細孔を横切る同じ配列の複数のポリヌクレオチドにわたる同じレベルの比較において、各レベルに関する持続期間の中央値を計算した。従ってこれらの持続期間の中央値は、各レベルに関連した典型的持続期間の代表であった。しかし滞在時間の指数分布のために、これらの滞在時間の時間定数は、それらのATP依存性のより良い指標である。この理由のために、図4Bにおいて、偶数及び奇数レベルの持続期間ヒストグラム(ここで「偶数」及び「奇数」は図4Aのレベルインデックスに関連している)は、指数関数型減衰曲線($a \times e^{-t/\tau}$)にフィットし、且つ各々の時間定数を、プロットした。このレベルの持続期間のヒストグラムは、毎ポリヌクレオチド-転位事象の各レベルの持続期間は、同等のサイズのビン(equivalently-sized bins)へ組込むことにより、構築した。次にこれらのヒストグラムを、市販の曲線-フィッティングアルゴリズム(Matlab Curve Fitting Toolbox)を使用しフィットさせ、このアルゴリズムは、指数関数型減衰モデルへデータをフィットするために最小二乗法を使用する。この方法は、残差の平方和を最小化し、ここで残差は、データポイントとそのポイントに対しフィットした応答の間の差として規定される。これは、データをパラメトリックモデルにフィットさせる上での標準技術である。図4Bのエラーバーは、各フィットの95%信頼限界に対応している。

【0179】

図4Bに示すように、偶数レベルに関する滞在時間は、ATP濃度の減少により増大するのに対し、奇数レベルに関する滞在時間は、一定であり続ける。従って第一のフラクショナル転位工程に対応している偶数レベルに関する滞在時間は、ATP結合に表面上は関連し、並びに指数分布によりATP濃度に反比例するのに対し、第二のフラクショナル転位工程に対応している奇数レベルに関する滞在時間は、ATP加水分解及びATP依存性に表面上は関連していた。

【0180】

(実施例III：ポリヌクレオチド配列決定におけるフラクショナル転位工程の有用性)

実施例IIIは、完全転位サイクルの2つのフラクショナル転位工程から得られた電氣的シグナルを使用することによる、完全転位サイクルから得られた単独の電氣的シグナルを使用するものと比べ、増大した配列決定精度を説明している。

【0181】

MspA「リードヘッド」は、狭窄ゾーン内の4つのヌクレオチドの部分配列(4-mer)に対し感受性があるので、MspAナノポアにおいて認められた全ての4-mer組合せに対応する電流を測定する四量体マップから、電流トレースを生じた。4-mer組合せに対応する電流の測定に関する更なる詳細については、Laszloらの文献、「天然DNAの長いナノポア配列決定読み値の解釈(Decoding long nanopore sequencing reads of natural DNA)」、Nature Biotechnology 32: 829-833 (2014)を参照されたい。しかし異なる細孔は、狭窄ゾーン内の異なる数のヌクレオチドに対し感受性があり得ることは、理解されるべきである。本実施例において、配列決定精度は、以下に説明するように隠れマルコフモデル(HMM)結果を、オリジナルのde Bruijn配列と比較することにより、決定した。図5に図示したような典

型的実験ノイズレベル(すなわち、 $\sim 0.5\text{-}2\text{pA}$ 、又はおよそ $0.5\sim 1.5\text{pA}$)に関して、完全工程を使用する再構築精度(菱形)は、フラクショナル工程(四角形)と比べ低下した。

【0182】

簡単に述べると、細孔を、先に説明された方法により確立した(Butlerらの文献、Proc. Natl. Acad. Sci. USA, 105:20647-20652 (2008); Manraoらの文献、PLoS ONE, 6:e25723 (2011))。簡単に述べると、脂質二重層を、1,2-ジフィタノイル-sn-グリセロール-3-ホスホコリン(Avanti Polar Lipids社)からのテフロン中に水平方向に直径 ~ 20 ミクロンの開口部を横切って形成した。この二重層の両側のコンパートメントは、 10mM Hepes、 $\text{pH}8.0$ 、 400mM KCl、 1mM DTT、及び 10mM MgCl_2 の実験用緩衝液を含んだ。Axopatch-200B(Axon Instruments社)を使用し、二重層を横切る電圧(140mV 又は 180mV)を印加し、且つイオン電流を測定した。MspAを、基底側のシスコンパートメントへ、濃度 $\sim 2.5\text{ng/ml}$ で添加した。一旦単独のMspAタンパク質が、テフロン開口部へ挿入されたならば、更なる挿入を阻害又は回避するために、シスコンパートメントは、実験用緩衝液でフラッシュした。全ての実験は、 23°C で行った。類縁体イオン電流シグナルは、4-極Besselフィルターにより、 20kHz で低域フィルターを通過させ、且つNational Instruments 6363デジタイザーを使用し、 100kHz でデジタル化した。データ獲得は、LabWindows/CVI(National Instruments社)で書かれた特注のソフトウェアにより制御した。データは、Matlab(The Mathworks社)で書かれた特注のソフトウェアにより解析した。その場合ATP濃度が $10\mu\text{M}\sim 1\text{mM}$ の範囲である、ATP力価決定実験を除いて、ATPは典型的には、 1mM で使用した。コレステロール-含有ポリヌクレオチドにハイブリダイズされた転位するポリヌクレオチドは、 10nM で使用した。Hel308 Tgaヘリカーゼは、最終濃度 115nM で使用した。ポリヌクレオチド及びATPを、シスチャンバーへ添加し、引き続き最後にHel308 Tgaヘリカーゼを添加した。或いは当該技術分野において周知のATP再生システムを、利用することができる。一つの実験システムは、 2mM ATP、 10mM クレアチンリン酸二ナトリウム塩、 3.5U/mL クレアチンキナーゼ及び 0.6U/mL 無機ピロホスファターゼを含む。

【0183】

図5は、一部の実施態様に従う、様々なレベルの追加されたノイズによりインシリコで生じた電流トレース(以下に説明)での、完全工程(菱形)及び $1/2$ 工程(四角形)に関する配列決定再構築精度(隠れマルコフモデル(HMM))をプロットしている。図5は、de Bruijn配列(256-mer)に関するモデル電流ブロックトレースのHMM/Viterbiアルゴリズム解析由来の配列再構築精度を示している。一般的HMMアルゴリズムは、いくつかの観点において、Timprらの文献(Biophys J. 2012 May 16;102(10):L37-9. doi: 10.1016/j.bpj.2012.04.009)に記載されたものに類似している。このアルゴリズムは、一連の観察された測定値から、M「状態」の基礎となるセットを回復することができる。このアルゴリズムの基本形は、確率の2つの実験で決定されたセットに頼っている：状態-状態「推移」確率、及び状態-観察「出力」確率。測定は、工程 $i=1,2,3\cdots N$ において N 個の測定値を生じる。1つの確率セットは、所定の i 回、及び状態 S_i (ここで S は、 M 状態のセット中の状態である)、後続の状態 S_{i+1} (ここで、 S_{i+1} は、必ずしも S_i ではない)について、確率を説明する推移マトリクスである。ナノポアシステムに関して、 4nt に感受性があり、且つ4種の正準のヌクレオチド(A,C,G,T)を試験しているナノポアは、 $4^4=256$ 状態を生じ、これは 4nt の各組合せに対応している。これらの状態の各々は、4つの隣接状態の一つへのみ推移することができる。

【0184】

図6Aは、一部の実施態様に従う、ポリヌクレオチドがモーター酵素により移動されるナノポアにおいて配列を解読するためのHMMに必要な非-ゼロ確率を伴う状態推移を描いている。このモーターは、phi29 DNAP又は1ヌクレオチド工程においてポリヌクレオチドを移動する類似の酵素である。図6Bは、一部の実施態様に従う、ポリヌクレオチドがモーター酵素により運動されるナノポアにおいて配列を解読するためのHMMに必要な非-ゼロ確率を伴う状態推移を描いている。このモーターは、Hel308ヘリカーゼ又はポリマーのフラクショナル移動を可能にする類似の酵素である。

【0185】

このシステムの推移マトリクスに関する非-ゼロ推移確率を、単ヌクレオチド工程において移動する酵素について、図6Aに図示している。この種の酵素を使用し、各ポリヌクレオチド状態又はn-merは、4隣接n-mer状態の1つへと進行しなければならない。1フラクショナル転位工程を採用する酵素に関して、より多くの状態が存在するであろう。これに関して、所定の完全-工程状態は、別の完全-工程状態を認めることができる前に、半-工程(又はフラクショナル-工程)状態へと進行しなければならない。従って、より多くの識別できる経路により利用可能なより多くの状態が存在し、結果的にポリヌクレオチド特徴決定の精度を補助する。

【0186】

状態の数は、 $q \times 4^{n+1}$ によりもたらされ、ここでnは、ナノポアのリードサイズであり、並びにqは、完全転位サイクルを完了するために必要とされる工程の数である。Hel308ヘリカーゼ及びM2-NNN MspAにより認められるような、 $q=2$ 及び $n=4$ に関して、2048状態が存在する。推移確率マトリクスは、フラクショナルヌクレオチド工程において移動する酵素に関して、図6Bに図解している。完全状態に対応する各状態は、1024の「半状態」又は「フラクショナル状態」のわずかに1つに推移することができるが、半(又はフラクショナル)状態の各々は、ナノポアのリードヘッド内の新たなものに対応している、4つの異なる状態に推移することができる。HMM解読アルゴリズムに関して、別の確率セットが使用される：時点tでの電流測定値 C_t が、状態 S_t に属する確率。この確率のセットは、実験により決定されるか、又は先の実験知見から推定される。かかる推定は、Laszloらの文献(2014(本明細書別所に引用))に開示されたようなアラインメントアルゴリズムの反復適用によるか、又はHMMの最大の予測により、達成することができる。フラクショナル転位工程の有用性を評価するために、フラクショナルヌクレオチド工程による酵素に関する配列決定精度を、単ヌクレオチド工程による酵素のそれと比較した。HMM Viterbi解読アルゴリズムを、MATLABにおいて実行される特注のソフトウェアにより実行し、且つインシリコ実験における10モンテカルロシミュレーションを、各条件について生じた。配列決定の平均及び標準偏差は、これらの100モンテカルロシミュレーションの平均値及び標準偏差から得た。電流レベルは、Manraoの文献(2012(本明細書別所に引用))からの結果を基に作製した。ガウスノイズを付加し、ガウス幅は、図5のX-軸上に示された値により生じ、配列再構築に使用したインシリコで観察された電流値をシフトさせ、且つ典型的ナノポア配列決定実験は、平均レベルにおいて、約 1pA_{rms} の変動を有する。幅 0.5pA を持つ追加されたガウスシフトに関して、フラクショナル及び完全の両工程の再構築は、100%に相応する配列決定精度を生じた。 0.5pA を上回る幅を持つ追加されたガウスシフトを上回ると、フラクショナル転位工程の配列決定精度は、非-フラクショナル転位工程に関する配列決定精度よりも大きかった。従って、フラクショナル転位工程の余分な情報は、 0.5pA よりも大きいガウスノイズが平均電流レベルに付加される場合に、増強された配列再構築精度を提供又は付与する。

【0187】

ノイズ変動に加え、酵素の確率論的移动により引き起こされたスキップされたレベルは、配列決定精度を低下するか又は低下すると予想することができる。この精度の減少は、隣接四量体のヌクレオチドパターンの再読み取りにより、部分的に相殺されるか又は相殺することができる。追加されたフラクショナル転位工程により、ヌクレオチドパターンの追加の再読み取りが存在する。例えば所定のk-merに関する情報は、隣接するフラクショナル工程中に含まれ、そのためk-merは、これらの隣接するフラクショナル工程の間に、「再読み取り」される。例えば4-ヌクレオチド-感受性リードヘッドを持つナノポアを通じて、配列ATCGTCを有するポリヌクレオチドがフラクショナルに転位されると仮定する。いずれかの理論に結びつけられることを欲するものではないが、完全-工程を辿る(steping)モーターに関して、CとGの間の領域がリードヘッドの中心である場合、4-mer TCGTは、唯一の読みである(すなわち、先行する「工程」は、リードヘッドの中心に「TC」を有し、且つATCGのみが読み取られ；後続の工程は、リードヘッドの中心に「GT」を有し、且

10

20

30

40

50

つCGTCのみが読み取られる)。従って、TCGT読み取り工程が、このモーターによりスキップされる場合、その特定の4-merに関する情報は、これまで測定されていない。しかし、いずれかの理論に結びつけられることを欲するものではないが、Hel308ヘリカーゼなどのフラクショナル-工程を辿るモーターにより、完全工程時には、2つの隣接するヌクレオチド間の領域は、リードヘッドの中心であることができるが、他方フラクショナル工程時には、単独ヌクレオチドが、リードヘッドの中心であることができる。そのため前述のポリヌクレオチドの「CG」がリードヘッドの中心である場合、完全-工程を辿る場合と同じように、TCGTが読み取られる。先行するフラクショナル工程は、リードヘッドの中心にCのみを有することができ、且つATCGTに関する情報を読み取ることができ；後続のフラクショナル工程は、リードヘッドの中心にGのみを有することができ、且つTCGTCに関する情報を読み取ることができ。「TCGT」に関する情報は、フラクショナル工程を辿る場合に3回及び完全-工程を辿る場合にわずかに1回読み取ることができるので、モーターが完全-工程を辿る場合には恐らく真ではないことに関する工程をスキップしたとしても、4-merのこの追加の「再読み取り」は、TCGTに関して情報が得られることを可能にし得る。除かれたレベルの全ての画分に関して、2~8%の配列決定精度の向上が存在する。これは、電流レベルのランダム除去を実行する、インシリコモンテカルロシミュレーションにおいて追加的に示された。結論として、ナノポア配列決定実験において認められる誤差モードに関して、配列決定精度の堅固な増加が存在した。図10、図11及び図12は、それにより追加的フラクショナル転位工程情報を、配列精度を改善するために使用することができるスキームを描いている。これらのスキームは更に、実施例VIにおいて以下に例示される。説明されたスキームは、例証的使用であり、限定されることを意図するものではない。

【0188】

(実施例IV：パターンマッチングにおけるフラクショナル転位工程の有用性)

実施例IVは、公知のアルゴリズムを使用するレベルを同定するためのフラクショナル転位工程の例証的使用について説明する。Needleman-Wunschアラインメントなどの動的プログラミングアルゴリズムを使用し、追加のレベルが、多くのレベル内パターンを正確に見出す助けを提供した。Needleman Wunschアラインメントアルゴリズムに関する更なる詳細については、Durbinらの文献、「生物学的配列解析(Biological Sequence Analysis)」、11版(Cambridge University Press社、ケンブリッジ、英国、2006)を参照されたい。加えて又は代わりに、レベル持続期間を使用する、レベル電流平均値、レベル電流標準偏差、又はレベル分布は、パターンマッチング精度を更に増強することができる。本実施例において、Needleman Wunschアラインメントアルゴリズムを使用し、1000塩基配列に対応するレベル内に包埋された15-塩基配列に対応するレベルを同定した。以下の使用を、比較した：(1)完全ヌクレオチド移動に対応するレベル；(2)2つの半(又はフラクショナル)工程移動に対応するレベル；(3)2つの半-工程(又はフラクショナル-工程)移動に対応するレベル及び持続期間。知見は、変動する幅のガウス分布から作製されたランダム値により生じた値だけシフトされたレベルによる、10モンテカルロシミュレーションにより、インシリコで作製した。結果は、図7に示し、ここでは一部の実施態様に従う、ガウスシフトの関数として認められる電流パターンの予想された精度を描いている。アラインメント精度の平均及び標準偏差は、10モンテカルロシミュレーションの平均及び標準偏差から出した。図7において、菱形は、完全ヌクレオチド工程を有するモーターを描き、並びに円形は、フラクショナル転位工程を有するモーターを描き、並びに四角形は、持続期間値と組合せたフラクショナル転位工程を有するモーターを描いている。簡単に述べると、15-ヌクレオチドに対応するレベルパターンを、ランダム1000-ヌクレオチド配列に対応するレベルパターン内に包埋した。レベルは、phi29 DNAPなどの完全ヌクレオチド工程(菱形)を伴うモーター(完全転位工程のみ)、又はHel308ヘリカーゼなどのフラクショナル転位工程を伴うモーター(円形)のいずれかに対応した。マッチを更に改善するために、持続期間を、電流値に加えて使用した(四角形)。図7の結果から、ノイズの増加に関して、フラクショナル転位工程移動を使用するアルゴリズムに関して、マッチングの品質は、かなり大きいことを理解することができる。マッチング品質は、持続期間値も使用される場合に、更に

改善された。Needleman Wunschアルゴリズムによりレベルをマッチするために、入力レベル類似性測定、又はスコア化を、レベルを比較するために使用した。本試験において、スチューデントt-検定を利用し、電流レベルを比較した。2つの持続期間の類似性を比較(スコア化)するために、持続期間の自然対数の差異を決定し、且つスチューデントt-検定によりもたらされたスコアに加算した。用語「スコア」は、Needleman Wunschアルゴリズムの命名において規定することができる。これらのスコア化関数は、シグナルレベル(例えば電流値)及び持続期間を比較するために使用することができる方法の非限定的例を表している。

【0189】

(実施例V：ヘリカーゼフラクショナル工程の調節)

10

実施例Vは、Hel308ヘリカーゼ滞在時間を変動するために、変動する反応成分の使用を例示している。

【0190】

図8は、一部の実施態様に従う、ピロリン酸の変動する濃度による、Hel308ヘリカーゼ活性の例証的調節を示している。図9は、一部の実施態様に従う、ヌクレオチドインヒビターのオルトバナジウム酸ナトリウムによる、及びヌクレオチド類縁体のアデノシン5'-(β , γ -イミド)三リン酸リチウム塩水和物による、Hel308ヘリカーゼ活性の例証的調節を示している。

【0191】

Hel308ヘリカーゼ活性は、ピロリン酸濃度を増加することにより、調節された。簡単に述べると、反応条件は、0~50mMの範囲の様々な濃度のピロリン酸、例えば0mM(対照)、5mM、10mM、20mM、30mM、40mM、及び50mMを包含する、実施例IIIに説明されたものであった。結果は、図8に示し、これはピロリン酸の存在しないヘリカーゼ活性(対照)と比較したヘリカーゼ活性の割合(%)を示している。5mM及び10mMのピロリン酸濃度は、対照の75%よりも多いヘリカーゼ活性の低下を生じた。10mMよりも大きいピロリン酸濃度は、ヘリカーゼ活性の、結果的にヘリカーゼ滞在時間の更なる減少を生じた。蛍光アッセイを使用し、ヘリカーゼの二重鎖DNAをほどく能力をモニタリングした。49-nt FRETポリヌクレオチド(最終濃度50nM)は、5'フルオレセイン基(/FAM/)を含んだ。40-nt消光剤-含有ポリヌクレオチド(最終50nM)は、蛍光消光剤Black Hole消光剤(/BHQ1/)を含んだ。これら2種のポリヌクレオチドは、当該技術分野において周知の方法を用い、ポリヌクレオチドをそれらの融解温度75℃まで加熱し、室温までゆっくり冷却することにより、一緒にハイブリダイズさせた。この二重鎖は、3'→5'ヘリカーゼがそれに結合し得る、9-塩基の3'側オーバーハングを含んだ。40-nt消光剤-含有ポリヌクレオチドと100%相補的である相補的FRET 40-ntポリヌクレオチドは、最低でも10倍モル過剰に存在した。消光剤及び発蛍光団は最初に密接しているので、蛍光は消光された。ヘリカーゼが二重鎖DNAをほどくことを基に、40-nt消光剤-含有ポリヌクレオチドは、49-nt FRETポリヌクレオチドへ再結合するよりも、相補的FRET 40-ntポリヌクレオチドへ結合する可能性が高まり始めた。従って新たな1本鎖49-nt FRETポリヌクレオチドは、好適な励起光源の存在下で、発蛍光された。このアッセイ緩衝液は、10mM HEPES、pH8.0、400mM KCl、1mM MgCl₂、1mM DTT、1mM ATPを含んだ。この反応は、蛍光を読み取る前に室温で20分間継続させた。

20

30

40

【0192】

Hel308ヘリカーゼ活性、従って滞在時間はまた、それぞれヌクレオチドインヒビター又は類縁体であるオルトバナジウム酸ナトリウム及びアデノシン5'-(β , γ -イミド)三リン酸リチウム塩水和物のいずれかの存在下で減少されることが示された。簡単に述べると、反応条件は、5mMのインヒビター又は類縁体濃度で、オルトバナジウム酸ナトリウム(図9の“オルトバナジウム酸Na”)又はアデノシン5'-(β , γ -イミド)三リン酸リチウム塩水和物(図9の“AMP-PNP”)いずれかを含有する、実施例IIIに説明したものであった。結果は、図9に示し、これは、ヌクレオチドインヒビター又は類縁体の非存在でのヘリカーゼ活性(対照)と比較したヘリカーゼ活性の割合(%)を示している。濃度5mMのインヒビター又は類縁体は、対照の85%よりも多いヘリカーゼ活性の低下を生じ、従って、ヘリカーゼ

50

滞在時間、又はヘリカーゼがDNAに沿って移動するのに要する時間を増加すると予想することができる。例えば、滞在時間の増加は、フラクショナル工程の時間を延長することができる、結果的にシグナルの獲得時間をより長くすることができる。

【0193】

(実施例VI：配列決定精度を向上するためのフラクショナル工程情報の処理法)

実施例VIは、配列決定精度を向上するためのフラクショナル転位工程から得られる追加情報の3つの処理方法を例証している。

【0194】

図10は、電流レベル及び持続期間の情報を使用する、フラクショナル転位工程から得られた追加情報の処理方法を例示している。この方法は、2つの独立した配列読み値に適用することができる。このスキームを使用し、電流トレースは、工程検出アルゴリズムに供され、そこでは電流レベル及びこれらのレベルの持続期間が認められる。少なくとも部分的にこれらのレベルに関する持続期間を基にし、2つの状態のHMMは、完全工程(長い)として又は半(若しくはフラクショナル)工程(短い)としてのレベル、又はこれらの知見内のスキップの可能性を同定する。これらの同定された長い及び短い工程並びにスキップ情報は次に、HMM、Viterbi、又はパターンマッチングアルゴリズム、若しくはそれらの好適な組合せにより使用され、2つの状態型(各々、完全工程及び半(又はフラクショナル)工程レベルに対応する、長い及び短い)について個別に、ポリヌクレオチド配列を再構築する。呼び出された配列は次に、比較され、並びに例えば、HMM、Viterbi、又はパターンマッチングアルゴリズムを調節することにより、ポリヌクレオチド配列決定精度を向上するために使用される。アラインメントを使用し、2つの独立した配列読み値のマッチングの悪い位置を同定することができる。

【0195】

図11は、電流レベル及び持続期間の情報を使用する、フラクショナル転位工程から得られた追加情報を処理する例証的方法を図示している。この方法は、2つの同時の配列読み値に適用することができる。この方法において、電流トレースは最初に工程検出アルゴリズムに供され、それらのレベルを認める。その後平均(又は中央値)レベル電流値及び各レベルの持続期間を、二次元HMM、Viterbi、又はパターンマッチングアルゴリズム、若しくはそれらの好適な組合せに、対として入力し、これらのアルゴリズムは、持続期間及び電流値を試験し、且つ半(又はフラクショナル)状態及び完全状態の最適配列を推定するか若しくは呼び出す。この技術において、HMM出力確率は、二次元である： $P_i(\text{出力}_t) = P_i(\text{cur}_t, \text{dur}_t) = P_i(\text{cur}_t) \times P_i(\text{dur}_t)$ 、式中*i*は、ポリヌクレオチドの長い工程又は短い(フラクショナル)転位工程に対応する「状態」であり、並びに cur_t 及び dur_t は、各々、*t*番目のレベル電流及びレベル持続期間である。この二次元HMMは、長いレベル(完全状態)に関するコンセンサスマップ及び確率分布、並びに短いレベル(半(若しくはフラクショナル)状態)に関するコンセンサスマップ及び確率分布の入力として採用することができる。二次元HMMは、ヌクレオチド配列の呼び出しの出力として提供することができる。

【0196】

図12は、電流トレースを直接使用する、フラクショナル転位工程から得られた追加情報を処理する例証的方法を図示している。この方法は、持続期間情報の使用を伴う又は伴わずに適用することができる。持続期間情報の使用を参照し、この方法において、電流トレースは、持続期間-依存型HMMにより直接分析される。この型のHMMにおいて、レベルの持続期間は、最も可能性のある配列及び完全又は半(若しくはフラクショナル)工程状態と同時に決定される。状態が、2つの時間反復の間で変化されずにある場合は、所定の状態に関する持続期間は増加するであろう。その後この持続期間は、その状態が完全状態又はフラクショナル状態のいずれであるかの評価を向上するために使用される。

【0197】

(実施例VII：配列決定精度を向上するためのフラクショナル工程情報を処理する追加方法)

実施例VIIは、配列決定精度を向上するためのフラクショナル工程情報を処理する追加

の例証的方法を説明する。

【0198】

隠れマルコフモデル(HMM)及びViterbiアルゴリズムは、単-工程分子モーターを使用する、ナノポアを通じて転位するポリヌクレオチドからのシグナルを基にした塩基-呼び出しのために、これまで使用されてきた。更なる詳細については、Timpらの文献、「Viterbiアルゴリズムを使用するナノポアからのDNA塩基-呼び出し(DNA Base-Calling from a Nanopore Using a Viterbi Algorithm)」、Biophysical Journal 102: L37-L39 (May 2012)を参照されたい。図19Aは、例えば所定のシグナルレベルが、例えばポリメラーゼ又はヘリカーゼによる細孔を通じた1ヌクレオチドの転位に対応しているような、細孔を通じたポリヌクレオチドの単-工程転位からのシグナルを特徴決定するために使用した、例証的隠れマルコフモデル(HMM)の態様を概略的に図示している。本明細書別所記載のように、シグナルレベルは、細孔の狭窄内の単ヌクレオチドの存在に必ずしも対応していないが、代わりに、複数のヌクレオチド、例えば2、3、4、5、6、7、8、9、10又は10よりも多いヌクレオチドを含む、「ワード」の存在に対応することができる。このような「ワード」はまた、「k-mer」と称することができる。図19Aに例示された実施態様において、「ワード」又は「k-mer」は、細孔の狭窄中の4個のヌクレオチドの存在を基にしたシグナルレベルに対応している、4ヌクレオチド長であるか、又は「四量体」若しくは「4-mer」である。

【0199】

図19Aにおいて、細孔を通じて転位するポリヌクレオチドの所定の位置*i*について、細孔の狭窄内の所定の四量体は、4個のヌクレオチドの可能な組合せ、例えばAAAA、AAAC、AAAG、AAAT、・・・TTTTを含むことができることが示されている。その四量体を独自に同定することは、かかる四量体に対応するシグナルレベルを基に必ずしも可能ではない。例えば、2つの異なる四量体、例えば予想された配列内に互いに隣接する2つの異なる四量体は、他方に対し同じシグナルレベルを有する可能性がある。Timpは、特定のトリプレットが他方と同じシグナルレベルを有することが認められることを基にし、従ってそのトリプレットに対応する電流レベルのみを基にした、トリプレット中の塩基のヌクレオチド塩基呼び出しを阻害する、DNAトリプレット(3-mer)に関する例証的電流値を明らかにしている。いくつかの四量体(及びより一般的にはいくつかのk-mer)は、他方から識別不可能なシグナルレベルを有し得、従って4-mer又はk-merのそれに対応する電流レベルのみを基にしたその四量体又はk-merにおける塩基のヌクレオチド塩基呼び出しを阻害することは、理解されるべきである。従って、HMMの専門用語を使用し、シグナルレベルの知見を基に互いに識別不可能であるかかる四量体又はk-merの塩基は、「隠れ状態」としてモデル化することができる。

【0200】

細孔狭窄内のポリヌクレオチドの他の単-工程位置の知見を基にした追加情報を、その四量体又はk-mer中の塩基を精確に同定する尤度、従って「隠れ状態」を精確に同定する尤度を増大するために、使用することができる。例えば図19Aにおいて、細孔を通じて転位しているポリヌクレオチドの次の位置*i+1*について、位置*i*の最後の3個のヌクレオチドは、位置*i+1*の最初の3個のヌクレオチドに対応するので、細孔の狭窄内の所定の四量体は、4つのヌクレオチドの特定の可能な組合せのみを有することができることも認めることができる。従って、*i*状態及び*i+1*状態に関するシグナルの測定値は、*i*位置及び*i+1*位置の(又は同等に、*i-1*位置及び*i*位置の)一方又は両方に存在する四量体を正確に同定する尤度を増大するために使用することができる。例えば、ポリヌクレオチドの位置*i*に対応する配列AAAAを基に、4つの配列AAAA、AAAC、AAAG、及びAAATのみが、位置*i+1*において利用可能である。位置*i+1*での利用可能な4つの配列は、位置*i*での各可能性のある配列に関して容易に同定することができる。同様に、ポリヌクレオチドの位置*i+1*での配列を基に、ポリヌクレオチドの位置*i+2*で利用可能な4つの配列は、容易に同定することができる。単工程モーターに関するViterbiアルゴリズムでは、シグナルレベルと位置*i*、*i+1*、*i+2*、・・・*i+n*の間に1-対-1の対応が存在する(ここで*n*は、ポリヌクレオチドにおけるヌクレオチドの数である)が、これは、 $L = \{l_1, l_2, \dots, l_n\}$ の順序付けられたレベルのセットからの

シグナルを表すことができる。ポリヌクレオチドの位置*i*に対応している、各レベル l_i は、そのシグナルレベルの平均(mean_i)、そのシグナルレベルの標準偏差(std_i)、又はそのシグナルレベルの持続期間(dur_i)の1以上として表すことができる。可能な四量体のセットは、 $\text{prev}(q) = \{q_1, q_2, \dots, q_4\}$ として表すことができ、これはポリヌクレオチドのこの位置(*i*位置)に対応する四量体が q である場合の、ポリヌクレオチドの先の位置(*i*-1位置)に対応する四量体の可能性のある値を規定している。例えば、 $\text{prev}(\text{AACC}) = \{\text{AAAC}, \text{CAAC}, \text{GAAC}, \text{TAAC}\}$ である。*j*番目に観察されたレベルに対応する観察されたシグナルレベル 0_j を基に、所定の四量体 q が位置*i*に存在する尤度スコアは、下記式で表すことができ：
【数 1】

$$\text{score}(\text{level}, i, q) = \max \left\{ \begin{array}{l} \text{score}(\text{level}, i-1, q) + \text{InsPen} \\ \max_{p \in \text{prev}(q)} \text{score}(\text{level}, i-1, p) + s(l_i | q) \\ \max_{p \in \text{prev}(q)} \text{score}(\text{level}, i, p) + \text{DelPen} \end{array} \right\} \quad (1)$$

式中、 $s(l_i | q)$ は、所定の四量体 q で、レベル l_i を観察する尤度を表す判定(award)に対応し、InsPenは、挿入ペナルティ(観察されたシグナルレベルに対応するが、ポリヌクレオチドの四量体には対応しない、ペナルティ)であり、並びにDelPenは、欠失ペナルティ(ポリヌクレオチドにおける四量体に対応するが、対応するシグナルレベルを有さない、ペナルティ)である。

【0201】

図19Bは、一部の実施態様に従う、Hel308ヘリカーゼを使用する、細孔を通じたポリヌクレオチドのフラクショナル工程転位からのシグナルを特徴決定するために使用した、例証的HMMの態様を概略的に図示している。図19Bにおいて、細孔を通じて転位するポリヌクレオチドの所定の位置*i*について、細孔の狭窄内の所定の四量体は、4個のヌクレオチドの可能な組合せのいずれかを、例えばAAAA、AAAC、AAAG、AAAT、・・・TTTTを含むことができることを再度認めることができる。細孔狭窄内のポリヌクレオチドのフラクショナル工程位置、並びに他の単工程位置の観察を基にした追加情報を、四量体又は*k*-mer中の塩基を精確に同定する尤度、従って単工程位置単独のみの使用に対し改善された精度を伴う「隠れ状態」を精確に同定する尤度を増大するために、使用することができる。

【0202】

例えば、図19Bにおいて、フラクショナル工程モーターに関して、細孔を通じて転位しているポリヌクレオチドの次の位置は、「iフラクショナル」であり、且つそこで位置*i*の最後の3個のヌクレオチドは、位置「iフラクショナル」の最初の3個のヌクレオチドに対応するので、細孔の狭窄内の所定の四量体は、4つのヌクレオチドの特定の可能な組合せのみを有することができることも認めることができる。従って*i*状態及び*iフラクショナル*状態に関するシグナルの測定値は、存在する四量体を正確に同定する尤度を増大するために使用することができる。例えば、ポリヌクレオチドの位置*i*に対応する配列AAAAを基に、4つの配列AAAA、AAAC、AAAG、及びAAATのみが、位置*iフラクショナル*について利用可能である。位置*iフラクショナル*での利用可能な4つの配列は、位置*i*での各可能性のある配列に関して容易に同定することができる。

【0203】

加えて図19Bにおいて、iフラクショナルの直後に続く位置である細孔を通じて転位しているポリヌクレオチドの次の位置*i*+1完全について、位置*i*+1完全の4個のヌクレオチドは、位置*iフラクショナル*と同じヌクレオチドに対応するので、細孔の狭窄内の所定の四量体は、一つの可能な配列のみを有することができることも認めることができる。従って、*i*位置、iフラクショナル位置及び*i*+1完全位置に対応するシグナルの測定値は、*i*位置、iフラクショナル位置及び*i*+1完全位置の一部又は全て(又は同等に、*i*-1位置及び*i*位置)に存在する四量体を正確に同定する尤度を増大するために使用することができる。例えば、ポリヌクレオチドの位置*i*に対応する配列AAAAを基に、4つの配列AAAA、AAAC、AAAG、及び

AAATのみが、位置*i*フラクショナル及び*i*+1完全について利用可能である。フラクショナル工程モーターに関する改変されたViterbiアルゴリズムでは、シグナルレベルとフラクショナル工程及び完全工程の両方の位置*i*、iフラクショナル、i+1完全、i+1フラクショナル、i+2完全、i+2フラクショナル、 $\dots$ i+*n*フラクショナル、i+*n*完全の間に対応が存在する(ここで*n*は、ポリヌクレオチドにおけるヌクレオチドの数である)が、これは、レベルのセットとしてシグナルレベル*l*を表すことができる。図19Aに関する先の考察と同様に、i完全又はiフラクショナル位置に対応している、各シグナルレベル*l_i*は、そのシグナルレベルの平均(mean_i)、そのシグナルレベルの標準偏差(std_i)、又はそのシグナルレベルの持続期間(dur_i)の1以上として表すことができる。電流フラクショナル転位工程での四量体*q*を考慮し、先の完全転位性工程に対応する可能性のある四量体のセットは $\text{prev}(q)=\{q_1, q_2, \dots, q_4\}$ として規定することができる。例えば、 $\text{prev}(\text{AACC}) = \{\text{AAAC}, \text{CAAC}, \text{GAAC}, \text{TAAC}\}$ 。

10

【0204】

位置*i*に対応する観察されたシグナルレベル*l_i*を基に、完全転位状態に対応する配置に関して位置*i*に存在する所定の四量体*q*に関する尤度スコア score_f 、及び半(又はフラクショナル)転位状態に対応する配置に関して位置*i*に存在する所定の四量体*q*に関する尤度スコア score_h は、下記式で表すことができ：

【数2】

$$\text{score}_f(\text{level}_i, i, q) = \max \left\{ \begin{array}{l} \text{score}_f(\text{level}, i-1, q) + \text{InsPen} \\ \max_{p \in \text{prev}(q)} \text{score}_h(\text{level}, i-1, p) + s_f(l_i | q) \\ \max_{p \in \text{prev}(q)} \text{score}_h(\text{level}, i, p) + \text{DelPen} \end{array} \right\} \quad (2)$$

20

$$\text{score}_h(\text{level}_i, i, q) = \max \left\{ \begin{array}{l} \text{score}_h(\text{level}, i-1, q) + \text{InsPen} \\ \max_{p \in \text{prev}(q)} \text{score}_f(\text{level}, i-1, p) + s_h(l_i | p, q) \\ \max_{p \in \text{prev}(q)} \text{score}_f(\text{level}, i, p) + \text{DelPen} \end{array} \right\} \quad (3)$$

式中、 $s_f(l_i | q)$ は、完全転位状態の所定の四量体*q*で、レベル*l_i*を観察する尤度を表す判定に対応し、 $s_h(l_i | p, q)$ は、フラクショナル転位状態の所定の四量体*q*及び先の四量体*p*で、レベル*l_i*を観察する尤度を表す判定に対応し、InsPenは、挿入ペナルティ(観察されたシグナルレベルに対応するが、ポリヌクレオチドの四量体には対応しない、ペナルティ)であり、並びにDelPenは、欠失ペナルティ(ポリヌクレオチドにおける四量体に対応するが、対応するシグナルレベルを有さない、ペナルティ)である。

30

【0205】

加えて、動的プログラミングを、フラクショナル工程分子モーター(Hel308ヘリカーゼなど)に関するパターンマッチングのために、使用することができる。動的パターンマッチングは、単-工程分子モーターについて、Laszloらの文献、「天然DNAの長いナノポア配列決定読み値の解読(Decoding long nanopore sequencing reads of natural DNA)」、Nature Biotechnology 32: 829-833 (2014)に説明されている。例えば、単-工程分子モーターについて、シグナルレベル*l*は、レベルのセット $L = \{l_1, l_2, \dots, l_n\}$ として表すことができ、ここでポリヌクレオチドの完全転位工程位置に対応する各シグナルレベル*l_i*は、そのシグナルレベルの平均(mean_i)、そのシグナルレベルの標準偏差(std_i)、又はそのシグナルレベルの持続期間(dur_i)の1以上として表すことができる。観察されたシグナルレベル*l_i*を基に、測定された所定の四量体*q_j*に関する尤度スコアは、下記式で表すことができ：

40

【数 3】

$$score(i, j) = \max \left\{ \begin{array}{l} score(i-1, j) + InsPen \\ score(i-1, j-1) + s(l_i | q_j) \\ score(i, j-1) + DelPen \end{array} \right\} \quad (4)$$

式中、 i は、レベル配列の位置を表し； j は、DNA配列中の位置を表し、四量体 q_j の最後の塩基は、位置 j での塩基であり； $score(i, j)$ は、レベル $l_1 \cdots l_i$ と、四量体 $q_1 \cdots q_j$ の間のマッチの程度であり； $s(l_i | q_j)$ は、所定の四量体 q_j で、レベル l_i を観察する尤度を表す判定に対応し； $InsPen$ は、挿入ペナルティ（観察されたシグナルレベルに対応するが、ポリヌクレオチドの四量体には対応しない、ペナルティ）であり； $DelPen$ は、欠失ペナルティ（ポリヌクレオチドにおける四量体に対応するが、対応するシグナルレベルを有さない、ペナルティ）である。

【0206】

Hel308などのフラクショナル-工程分子モーターに関して、シグナルレベル l は、レベルのセット $L = \{l_1, l_2, \cdots, l_n\}$ として表すことができ、ここでポリヌクレオチドの i 完全位置又は i フラクショナル位置に対応する各シグナルレベル l_i は、そのシグナルレベルの平均($mean_i$)、そのシグナルレベルの標準偏差(std_i)、又はそのシグナルレベルの持続期間(dur_i)の1以上として表すことができる。観察されたシグナルレベル l_i を基に、完全転位状態に対応して測定された所定の四量体 q_j に関する尤度スコア $score_f$ 、及び半(又はフラクショナル)転位状態に対応して測定された所定の四量体 q_j に関する尤度スコア $score_h$ は、下記式で表すことができ：

【数 4】

$$score_f(i, j) = \max \left\{ \begin{array}{l} score_h(i-1, j) + InsPen \\ score_h(i-1, j-1) + s_f(l_i | q_j) \\ score_f(i, j-1) + DelPen \end{array} \right\} \quad (5)$$

$$score_h(i, j) = \max \left\{ \begin{array}{l} score_f(i, j) + InsPen \\ score_f(i, j-1) + s_h(l_i | q_{j-1}q_j) \\ score_h(i, j-1) + DelPen \end{array} \right\} \quad (6)$$

式中、 i は、レベル配列の位置を表し； j は、DNA配列中の位置を表し、四量体 q_j の最後の塩基は、位置 j での塩基であり； $score_f(i, j)$ 及び $score_h(i, j)$ は、レベル $l_1 \cdots l_i$ と、四量体 $q_1 \cdots q_j$ の間のマッチの程度であり、各々、完全状態又はフラクショナル状態を仮定し； $s_f(l_i | q_j)$ 及び $s_h(l_i | q_j)$ は、各々、完全状態及びフラクショナル状態における所定の四量体 q_j で、レベル l_i を観察する尤度を表す判定に対応し； $InsPen$ は、挿入ペナルティ（観察されたシグナルレベルに対応するが、ポリヌクレオチドの四量体には対応しない、ペナルティ）であり； $DelPen$ は、欠失ペナルティ（ポリヌクレオチドにおける四量体に対応するが、対応するシグナルレベルを有さない、ペナルティ）である。

【0207】

ここでフラクショナル工程を使用するデノボ配列決定結果のいくつかの例を、図20Aを参照し説明する。75種の500-merポリヌクレオチドのライブラリーを、ヒトDNAを基に作製し、ナノポアデータを、実施例II及びIIIを参照し、本明細書別所記載と同様に収集した。このデータを基にしたヌクレオチド塩基-呼び出しを、先に式(2)及び式(3)を用いて説明した改変Viterbiアルゴリズムを用いて解析した。次に塩基-呼び出し配列を、その75種は真の500-merであり、且つその75種は「デコイ」又は「ダミー」500-mer配列である、150種の500-merのセットと整列した。図20Aは、アラインメントの精度の関数として読み値

の長さを図示する(Kielbasaらの文献、「適応シーズに従うゲノム配列比較(Adaptive seeds tame genomic sequence comparison)」、Genome Research21: 487-493 (2011)に説明されたようなLASTALアライナーを使用)が、この図において、白菱形は、塩基-呼び出し配列が、正確な(“標的”)配列に整列されている結果に対応し、並びに黒菱形は、塩基-呼び出し配列が、「デコイ」又は「ダミー」配列に整列されている結果に対応している。図20Aから、約200塩基対よりも大きい読み値の長さについて、約60%よりも大きい精度を得ることができることを理解することができる。この精度は、DNAの両鎖の読み取りなど、公知の技術を使用し、更に増大することができる可能性がある。

【0208】

ここでフラクショナル工程を使用するパターンマッチング結果のいくつかの例を、図20B-20Cを参照し説明する。図20Aを参照し先に説明したものと同一、75種の500-merヌクレオチドのライブラリー及び同じ実験プロトコールを使用した。このデータを基にしたヌクレオチド塩基-呼び出しを、先に式(5)及び(6)を用いて説明したパターンマッチングに関する動的プログラミングを用いて解析した。次に塩基-呼び出し配列を、その75種は真の500-merであり、且つその75種は「デコイ」又は「ダミー」500-mer配列である、150種の500-merのセットと整列した。図20Bは、アラインメントスコアの関数としてアラインメントサイズを図示するが、この図において、白菱形は、塩基-呼び出し配列が、正確な(“標的”)配列に整列されている結果に対応し、並びに黒菱形は、塩基-呼び出し配列が、「デコイ」又は「ダミー」配列に整列されている結果に対応している。図20Bから、約40よりも大きいアラインメントスコアを、約200塩基対よりも大きいアラインメントサイズについて得ることができることを理解することができる。図20Cも、アラインメントスコアの関数としてアラインメントサイズを図示するが、この図において、白菱形は、塩基-呼び出し配列が、正確な(“標的”)配列に整列されている結果に対応し、並びに黒菱形は、塩基-呼び出し配列が、「デコイ」又は「ダミー」配列に整列されている結果に対応している。図20Cから、約20よりも大きいアラインメントスコアを、約50塩基対よりも大きいアラインメントサイズについて得ることができることを理解することができる。フラクショナル工程モデルは、単工程モデルよりも、より多くの事象を精確に同定することができることを、認めることができる。

【0209】

加えて、1332レベルを伴う転位事象に関して、80kbのデータセットに対するパターンマッチング(式5及び式6)は、1本鎖について約145秒を要するのに対し、1332レベルを伴う同じ事象に関して、そのデータセットに対するデノボ配列決定(式2及び式3)は、1本鎖について約69秒を要することが認められた。パターンマッチング複雑度は、ヌクレオチドデータセットにより線形に成長するのに対し、デノボ配列決定の複雑度は、データセットとは無関係であることが認められた。パターンマッチングは、デノボ配列決定は同定に失敗するような比較的短い事象を精確に同定することが認められた。加えてパターンマッチングに関するフラクショナル工程モデルは、単工程モデルよりもより真の陽性を生じることが認められ、このことは、フラクショナル工程モデルは、ヘリカーゼデータを説明するためのより良いモデルであり得ることを示している。

【0210】

(実施例VIII：追加のHel308ヘリカーゼによるフラクショナル転位工程)

実施例VIIIは、分子モーターとして使用された例証的Hel308ヘリカーゼにより観察されたフラクショナル転位工程を説明している。

【0211】

実施例VIIIの実験は、DphPC脂質二重層中の単独の2NNN MspAナノポアを使用し、並びに表3に列挙されたパラメータを使用し、実施例Iを参照し先に説明されたものと同様に行い、ここで「Hel308 Mbu (A)」は、Hel308 Mbuを使用する第一の実験に使用されるパラメータのセットを指し、並びに「Hel308 Mbu (B)」は、Hel308 Mbuを使用する第二の実験について使用されるパラメータのセットを指す。脂質二重層は、1,2-ジフィタノイル-sn-グリセロ-3-ホスホコリン(Avanti Polar Lipids社)から形成した。この二重層は、テフロン中

に水平方向に直径～20ミクロンの開口部を広げた。M2-NNN-MspAを、濃度～2.5ng/mlで二重層の基底側に添加した。一旦単独の細孔が挿入されたならば、更なる挿入を避けるために、コンパートメントを、実験用緩衝液でフラッシュした。Axopatch-200Bパッチクランプ増幅器(Axon Instruments社)を使用し、二重層を横切る電圧180mVを印加し、且つイオン電流を測定した。類縁体シグナルは、4-極Besselフィルターにより、50kHzで低域フィルターを通過させ、その後低域フィルター周波数を5回デジタル化した。データ獲得は、LabWindows/CVI(National Instruments社)で書かれた特注のソフトウェアにより制御した。二重層の両側の～60 μ lコンパートメントは、適した濃度のKCl、1mM EDTA、1mM DTT、1mM ATP、5mM MgCl₂、及び10mM HEPES/KOH緩衝液でpH8.0の実験用緩衝液を含んだ。野生型Mbu Hel 308ヘリカーゼを、分子モーターとして、指定された濃度で使用した。

10

【0212】

Hel308 Mbu実験及びHel308 Tga実験の両方において、DNAは、3'側から5'側方向に読み取ったが、phi29ポリメラーゼ実験においては、DNAは、5'側から3'側方向に読み取った。

表3

【表3】

パラメータ	Hel308 Mbu (A)	Hel308 Mbu (B)	Hel308 Tga	phi29
[KCl]	300 mM	600 mM	400 mM	300 Mm
pH	8.0	8.0	8.0	8.0
[MgCl ₂]	5 mM	5 mM	5 mM	5 mM
[ATP]	1 mM	1 mM	1 mM	-
[DTT]	1 mM	1 mM	1 mM	1 mM
[EDTA]	1 mM	1 mM	1 mM	1 mM
[DNA]	10 nM	10 nM	10 nM	10 nM
[分子モーター]	150 nM	150 nM	150 nM	1.5 μ M
電圧	180 mV	180 mV	180 mV	180 mV

20

30

【0213】

図17A-17Dは、一部の実施態様に従う、特定のパラメータを使用する、Hel308 Mbuヘリカーゼ、Hel308 Tgaヘリカーゼ、及びphi29ポリメラーゼの転位事象の比較を示している。図17Aは、表3に示された「Hel308 Mbu (A)」パラメータを使用する、Hel308 Mbuヘリカーゼにより観察された転位工程を示す。転位するポリヌクレオチド

40

【化11】

(配列番号: 72:

/5Phos/AAACCTTCCXCCCGTACCGTGCCGTACCGTTCCGTTCGGTACCGTATTTTT
TTTTCTCACTATCGCATTCTCATGCAGGTCGTAGCC、ここでX=脱塩基)

は、コレステロール-含有ポリヌクレオチド

【化 1 2】

(配列番号: 73:

AAAAAAAATACGGTACGGAACGGAACGGTACGGCACGGTACGGG

TTTTTTTTTTTTTTTT/3CholTEG)

へハイブリダイズした。図17Bは、表3に示された「Hel308 Mbu (B)」パラメータを使用し、且つ図17Aと同じポリヌクレオチド配列を使用する、Hel308 Mbuヘリカーゼにより観察された転位工程を示す。図17Cは、表3に示された「Hel308 Tga」パラメータを使用し、且つ図17Aと同じポリヌクレオチド配列を使用する、Hel308 Tgaヘリカーゼにより観察された転位工程を示す。図17Dは、表3に示された「phi29」パラメータを使用し、且つ図17Aと同じポリヌクレオチド配列を使用する、phi29ポリメラーゼにより観察された転位工程を示し；図17Dのphi29プロットは、図17A、17B、17C、及び17Dの間の比較を促進するために、ほぼ垂直軸を反映している。

10

【0 2 1 4】

図17A-17Dにおいて、各ヘリカーゼの配列決定に関して、ナノポアは、一般に“a”(シグナルの谷に対応)及び“b”(シグナルのピークに対応)と称される特徴を検出したことを認めることができる。Hel308 Tgaヘリカーゼによる配列決定に関して(図17C)、phi29ヘリカーゼについて(図17D)のほぼ2倍の数のレベルが観察されたことも認められる。同じく、“Hel308 Mbu (B)”条件下でのHel308 Mbuヘリカーゼの配列決定(図17B)について、“Hel308 Mbu (A)”条件下でのHel308 Mbuヘリカーゼについて(図17A)よりもより大きい数のレベルが観察されたことも認められる。同じく、“Hel308 Mbu (B)”条件下でのHel308 Mbuヘリカーゼによる配列決定に関して(図17B)、Hel308 Tgaヘリカーゼについて(図17C)より少ないレベルが観察されるが、phi29ヘリカーゼについて(図17D)より多いレベルが観察されたことも認められる。図17A-17Dは、(1)Hel308ヘリカーゼの複数の変種(例えば、Tga及びMbuの両方)は、フラクショナル工程を表すのに対し、ポリメラーゼPhi29についてフラクショナル工程は観察されなかったこと；並びに、(2)フラクショナル工程は、環境の変数又はパラメータ、例えば、KCl濃度などを変更することにより、解明することができること：を意味するものとして解釈することができる。加えて、他のデータは、Mbuを利用する場合、レベルの持続期間は、ATP濃度の減少により増大すること、例えばMbuにおけるフラクショナル工程の持続期間、従って物理的機序も、ATP-依存型であり得ることを示している。

20

30

【0 2 1 5】

(実施例IX：任意にマルチ-モダリティと組合せたストレッサーの使用)

本明細書に提供される開示を基に明らかにされるように、多くの環境変数又はパラメータは、特定のポリヌクレオチド配列を基に、ナノポアシステムを読み取る様式、又はシグナルを発生する様式に影響を及ぼすことができる。かかる作用を提供し得る変数又はパラメータの例は、温度、塩濃度(例えばMg、Cl)、補因子(例えばATP)濃度、ピロリン酸などのATP生成物の濃度、pH、使用される特定の分子モーター(例えば使用される特定のHel308ヘリカーゼ)、圧力などを含むことができる。

40

【0 2 1 6】

例えば、実施例II並びに図4A及び4Bを参照し先に説明したように、ATPの濃度は、いくつかの転位工程に対応するレベルの滞在時間に影響を及ぼすことができる。例えば、第一のフラクショナル転位工程に関する滞在時間は、ATP濃度が減少するにつれ、増大され、表面上はATP結合に関連し、且つATP濃度に反比例することが観察された。実施例V及び図8を参照し先に説明したような別の例として、ピロリン酸の濃度は、Hel308ヘリカーゼの活性に影響を及ぼすことができる。例えば、Hel308ヘリカーゼの活性は、ピロリン酸濃度が増加するにつれ、減少され、結果的にヘリカーゼ滞在時間を増大することが観察された。実施例V及び図9を参照し先に説明したような別の例として、ヌクレオチドインヒビター又は類縁体の濃度は、Hel308ヘリカーゼの活性に影響を及ぼすことができる。例えば、オル

50

トバナジウム酸ナトリウム及びアデノシン5'-(β , γ -イミド)三リン酸リチウム塩水和物(AMP-PNP)の存在を基に減少したHe1308ヘリカーゼの活性は、ヘリカーゼ活性を減少し、結果的にヘリカーゼ滞在時間を増加することが観察された。更に実施例VIII並びに図17A及び17Bを参照し先に説明したような別の例として、塩の濃度は、観察されたレベルの数に影響を及ぼすことができる。例えば、塩(例えばKCl)濃度の増加は、He1308 Mbuヘリカーゼによる配列決定時に観察されたレベルの数を増大したことが観察された。当業者は容易に、シグナルがポリヌクレオチド配列を基に発生される様式を調節するために、いずれか好適なパラメータを調節することを想起するであろう。

【0217】

加えてかかるパラメータの異なる組合せは、配列決定の精度、並びに配列決定の処理量に影響を及ぼし得ることは、理解されるべきである。例えばヘリカーゼの滞在時間の増加は、精度を増加することができ、例えば観察されたレベルの数を増加することができるが、潜在的に配列決定の処理量は減少し得る。フラクショナル工程の観察を基にした配列決定に関して、一部の工程は、別の工程のセットよりも特定の変数により、潜在的により影響を受け得る。変数-非依存性工程を使用し、精度のベースラインを設定する一方で、他の工程は、特定の配列決定の要求に合致するように調節することができる(例えば、より低い処理量により増大した精度、又は減少した精度により増加した処理量)。一部の実施態様において、マルチ-モーダル装置は、シーケンサーの要求を基に精度及び処理量を調和することにより、例えば配列決定時に1以上のパラメータを調節することにより、これを利用することができる。一つの非限定的な例証的例として、先に記載したように、He1308 TgaによるATP濃度の減少は、フラクショナル状態の持続期間を増加することができることが観察される。フラクショナル状態の持続期間の増加は、例えばフラクショナル状態の読み値のシグナル-対-ノイズ比(SNR)を向上するか、又はより低い-周波数フィルターが適用されることを可能にすることにより、配列決定精度を増大することができるが、処理量を減少することができる。マルチ-モーダル装置は、配列の粗い「スカフォールド」を比較的迅速に決定するために、配列決定試行を高濃度のATPで開始することにより、これを利用することができ、且つ次に品質は高いが遅い読み値によりスカフォールドの「ギャップを充填」するために、ATP濃度を減少することができる。

【0218】

加えて細孔を通じた標的ポリヌクレオチドのHe1308ヘリカーゼによる転位により生じた1以上のシグナルの解像を増大するために、いずれか好適な数の異なるパラメータを、逐次又は互いに並行して使用することができることに留意。図21A-21Cは、一部の実施態様に従う、細孔を通じたポリヌクレオチドの様々な転位に関する時間の関数として生じ得るシグナルを概略的に図示している。図21A-21Cは各々、分子モーターによるよりもむしろ、ポテンシャルの差のみ印加された加力下で細孔を通じてポリヌクレオチドが転位する条件下で生じた理想的なシグナルに対応し、且つ無限のシグナル解像を伴う、破曲線を描いている。このような条件下で、シグナルは、ヌクレオチドが細孔を通過する際に、ヌクレオチドの位置及び配列の連続して変化する関数である。

【0219】

図21Aはまた、垂直の点線により示された時点で生じる完全転位工程のみ使用して発生された例証的シグナル(太線)を図示している。このシグナルは、本明細書において別所記載のように、電氣的又は光学的シグナルであることができる。加えて、このシグナルは、平均シグナルレベル、シグナル持続期間、又は標準偏差(例えば、広帯域ノイズ又は帯域制限ノイズ)などの、かかる電氣的又は光学的シグナルのいずれか好適な特徴を含むことができる。シグナルは、単独の工程により比較的低いレベルから比較的高いレベルまで、次に再度単独の工程により比較的低いレベルまで、変化することが、図21Aにおいて認めことができ、これはポリヌクレオチドが細孔を通じて転位する際に、完全転位工程間に生じる移行に対応している。シグナルは、異なる時点で(a)点、(b)点、及び(c)点で理想的なシグナルと交差し、結果的にこれらの点で理想的なシグナルを「サンプリング」すると考えられ得ることも図21Aにおいて認めることができる。しかし、有効サンプリング速

度は比較的遅いので、シグナルは、理想的シグナルを余り良くはサンプリングしない。例えば、(a)点及び(b)点での値は他方と同じであり、これは異なる転位工程に関するシグナルレベルを縮重する(degenerate)ことに対応している。シグナルは、(a)点と(b)点の間を線で結ぶ理想的曲線の部分を適切にサンプリングしないので、(a)点及び(b)点に対応する物理的転位工程は、他方から識別することができず、これはこのポリヌクレオチド配列に関する情報の喪失を生じる。加えて、シグナルは、(b)点と(c)点の間を線で結ぶ理想的曲線の部分を適切にサンプリングしないので、(b)点と(c)点の間の理想的曲線の下向き勾配に対応する物理的転位工程は、このような工程の間に細孔を通じて転位されたポリヌクレオチド部分を潜在的に部分的にのみ特徴づけることができる。

【0220】

図21Bはまた、先に説明された様に破曲線で表された理想的シグナルに加え、垂直の点線により示された時点で生じた時間分離型完全転位工程の組合せ、又は完全及びフラクショナル転位工程の組合せを使用し発生された、例証的シグナル(太線)を図示している。時間-分離型完全転位工程は、各々がポリヌクレオチドを転位する2つの分子モーターにより発生されるが、その時点で、他方に対してシフトされる、例えば、完全転位サイクルの持続期間時間のおよそ50%だけ他方に対してシフトされるシグナルに対応することができる。完全及びフラクショナル転位工程の組合せは、フラクショナル転位工程が、完全転位サイクルの持続期間時間のおよそ50%で生じるような、部分的及び完全転位工程を通じてポリヌクレオチドをフラクショナルに転位する単独の分子モーター(例えばHel308ヘリカーゼ)により発生されるシグナルに対応することができる。このシグナルは、図21Aを参照し先に説明したものであることができる。シグナルは、一連の工程により比較的低いレベルから比較的高いレベルまで、次に再度別の一連の工程により比較的低いレベルまで、変化することが、図21Bにおいて認めることができ、これはポリヌクレオチドが細孔を通じて転位する際に、時間-分離型完全転位工程間に、又は完全及びフラクショナル転位工程の組合せにより、生じる移行に対応している。シグナルは、図21Aよりも有意に多い数の点(及び回数)で理想的なシグナルと交差し、結果的にこれらの点で理想的なシグナルを「サンプリング」すると考えられ得ることも図21Bにおいて認めることができる。有効サンプリング速度は、図21Aよりも比較的速いので、このシグナルは、理想的シグナルを図21Aよりも比較的良くサンプリングする。例えば、(a)点及び(b)点での値は、他方と同じであり、これは異なる転位工程のためにシグナルレベルを縮重することに対応している。図21Aのシグナルはまた、(a)点と(b)点の間を線で結ぶ理想的曲線の部分をサンプリングするので、(a)点及び(b)点に対応する物理的転位工程は、他方から識別することができ、これはポリヌクレオチド配列に関する追加情報を生じる。加えて、図21Bのシグナルは、(b)点と(c)点の間を線で結ぶ理想的曲線の部分をより完全にサンプリングするので、(b)点と(c)点の間の理想的曲線の下向き勾配に対応する物理的転位工程は、図21Aのシグナルを使用し達成され得る可能性のあるものよりも、このような工程の間に細孔を通じて転位されたポリヌクレオチド部分をより良く特徴づけることができる。

【0221】

図21Cはまた、先に説明された様に破曲線で表された理想的シグナルに加え、垂直の点線により示された時点で生じた時間分離型完全転位工程の組合せ、又は完全及びフラクショナル転位工程の組合せを使用し発生された、別の例証的シグナル(太線)を図示している。時間-分離型完全転位工程は、各々がポリヌクレオチドを転位する複数の分子モーターにより発生されるが、その時点で、他方に対してシフトされる、例えば、完全転位サイクルの持続期間時間のおよそ25%、50%、及び75%だけ他方に対してシフトされるシグナルに対応することができる。完全及びフラクショナル転位工程の組合せは、フラクショナル転位工程が、完全転位サイクルの持続期間時間のおよそ25%、50%、及び75%で生じるような、部分的及び完全転位工程を通じてポリヌクレオチドをフラクショナルに転位する単独の分子モーター(例えばHel308ヘリカーゼ)により発生されるシグナルに対応することができる。このシグナルは、図21Aを参照し先に説明したものであることができる。シグナルは、図21Bよりもより大きい数の一連の工程により比較的低いレベルから比較的高いレ

10

20

30

40

50

ベルまで、次に再度図21Bよりもより大きい数の別の一連の工程により比較的低いレベルまで、変化することが、図21Cにおいて認めることができ、これはポリヌクレオチドが細孔を通じて転位する際に、時間-分離型完全転位工程間に、又は完全及びフラクショナル転位工程の組合せにより、生じる移行に対応している。シグナルは、図21Bよりも有意に多い数の点(及び回数)で理想的なシグナルと交差し、結果的にこれらの点で理想的なシグナルを「サンプリング」すると考えられ得ることも図21Cにおいて認めることができる。有効サンプリング速度は、図21Bよりも比較的速いので、このシグナルは、理想的シグナルを図21Bよりも比較的良くサンプリングし、且つ結果的に図21A又は21Bのシグナルを使用し達成され得る可能性のあるものよりも、このような工程の間に細孔を通じて転位されたポリヌクレオチドをより良く特徴づけることができる。

10

【0222】

理想的な試料曲線の任意の選択された部分のサンプリングを増大するために、いずれか好適なパラメータの選択を使用することができることは理解されるべきである。例えば、先に言及したように、異なる分子モーターからの時間-シフト型(相-シフト型)完全転位工程の組合せを、使用することができる。これに関して、図21Bは、完全転位工程の時間の50%だけ他方から分子モーターを時間-シフトすることを説明し、並びに図21Bは、完全転位工程の時間の25%、50%、及び75%だけ他方から分子モーターを時間-シフトすることを説明するが、かかる値は、純粹に例証的であり、且つ代わりの分子モーターは、他方からいずれか好適な時間量だけ時間-シフトされることができ、例えば他方から完全転位工程の時間の5%~95%シフトされることができ、例えば他方から完全転位工程の時間の10%~90%シフトされることができ、例えば他方から完全転位工程の時間の25%~75%シフトされることができ、例えば他方から完全転位工程の時間の40%~60%シフトされることができ。別の例として、完全及びフラクショナル転位工程の組合せは、部分的及び完全転位工程を通じてポリヌクレオチドをフラクショナルに転位する単独の分子モーター(例えばHel308ヘリカーゼ)により発生されるシグナルに対応することができる。図21Bは、他方から完全転位工程の時間の50%で生じるフラクショナル転位工程を説明するが、並びに図21Cは、他方から完全転位工程の時間の25%、50%、及び75%で生じるフラクショナル転位工程を説明するが、かかる値は、純粹に例証的であり、且つ代わりのフラクショナル転位は、例えば完全転位工程の時間の5%~95%で、例えば完全転位工程の時間の10%~90%で、例えば完全転位工程の時間の25%~75%で、例えば完全転位工程の時間の40%~60%で、完全転位工程に対するいずれか好適な時間で生じることができる。

20

30

【0223】

加えて、完全又はフラクショナル工程が生じる相対時間、従ってシグナルが、理想的シグナルをサンプリングする時間は、いずれか好適なパラメータを変動することにより、好適なことに調節することができることは理解されるべきである。例えば先に注記したように、シグナル発生に作用し得る例証的変数又はパラメータは、温度、塩濃度(例えばMg、Cl)、補因子(例えばATP)濃度、ピロリン酸などのATP生成物の濃度、pH、使用される特定の分子モーターなどを含むことができる。一部の実施態様において、第一のシグナルは、時間の第一の個別化されたセットで、理想的シグナルをサンプリングするために、パラメータの第一のセットを基に生じることができ、並びに第二のシグナルは、時間の第二の個別化されたセットで、理想的シグナルをサンプリングするために、パラメータの第二のセット(少なくとも1つに関して第一のパラメータのセットとは異なる)を基に生じることができる。第一又は第二シグナルのいずれか単独よりも、より良い解像を伴う理想的シグナルをサンプリングするシグナル曲線を提供するために、第一及び第二のシグナルは組合せることができる。これらのシグナルの個々のシグナルよりも、より良い解像を伴う理想的シグナルをサンプリングするシグナル曲線を提供するために、いずれか好適な数のシグナルを、同様の様式で組合せることができることは理解されるべきである。

40

【0224】

(実施例X：配列同定のための追加アプローチ)

配列同定のためのいくつかの追加アプローチを、実施例Xを参照し説明する。

50

【0225】

一部の実施態様において、配列-特異的情報を得るために、特定の型の情報を、単独で、又は他と組合せて、使用することができる：(A)完全工程反応情報のみ、(B)フラクショナル工程反応情報のみ、(C)識別子を伴わずに完全工程及びフラクショナル工程の反応情報を一緒に、並びに(D)識別子を伴い完全工程及びフラクショナル工程の反応情報を一緒に。

【0226】

「反応情報」とは、k-mer又はk-merサブセット(関心対象のk-merを含む)に対し独自である所定のポリヌクレオチド配列(k-mer)に対するシステムの反応から得られたデータを意味する。反応情報の例は、平均レベル電流、中央値レベル電流、広帯域レベル電流ノイズ、帯域制限レベル電流ノイズ、レベル持続期間などを含む。

10

【0227】

「識別子」とは、ポリヌクレオチド(k-mer)が、「理想的反応」に沿って、特定のレベルが他のレベルと結ばれる(lie)場所を同定するナノポア環境と相互作用する間に、得られたデータを意味する。例えば、比較的高い又は比較的低いレベルのATP濃度の存在下で、分子モーターとしてHel308 Tgaヘリカーゼを利用するシステムは、あらゆる他のレベルに関して、各々、比較的短い又は比較的長い持続期間を示すことができ、ここであらゆる他のレベルは、隣接するレベルから理想的反応に沿っておよそ50%である。この例において、レベル持続期間は、識別子として使用することができ、その理由は、これは(隣接するレベルに関して)理想的反応に沿って配列位置を同定するために使用することができるからである。

20

【0228】

「理想的反応」とは、ポリヌクレオチドの十分に小さい移動を解像することができるよう、十分に高い解像度を伴う、ナノポアを通じて転位する特定のポリヌクレオチドのシステムの反応を意味する。例えば理想的反応は、ナノポアを通じて転位するDNAの無限に-高い解像度の連続電流トレースである。

【0229】

本実施例において先に更に言及された項目(A)-(D)を再度参照し、項目(A)-(D)の各々は、ポリヌクレオチド配列を同定するために、独立して、又は1以上の他の項目(A)-(D)と一緒に、使用することができる。例えば1以上の項目(A)-(D)は、例えば、コンピュータリソースの制限、時間の制限、最適アプローチの先験的知識などのために、いずれか他の項目(A)-(D)から独立して計算することができる。計算される2以上の項目(A)-(D)を基に、ただ1つの項目(A)-(D)の結果を使用することができる。そのような計算のどれを使用するかは、結果の信頼性を基に決定することができる。例えば結果の信頼性は、以下の1以上を基にすることができる：(a)反応情報それ自身(例えば高レベルのATPは、Hel308 Tgaのフラクショナル工程サイズを短くすることができ、これは項目(A)に対し項目(B)の信頼性を低下し得る)；(b)配列決定アルゴリズムそれ自身(例えば、Viterbiアルゴリズムは、それが目的とする最適配列に関する尤度スコアを生じることができ、これは提唱された配列の信頼性レベルを決定するために使用することができる)；(c)配列決定アルゴリズムにより作製された配列(例えば、信頼性は、アルゴリズムにより提唱された配列と、配列のルックアップテーブル及び／又は配列決定されるポリヌクレオチドの先験的知識のいずれかの間の比較を基に割り当てることができる)；又は、(d)項目(a)-(c)のいずれか好適な組合せ。

30

40

【0230】

一部の状況において、項目(A)-(D)の2種以上から提唱された配列を利用することにより、実際の配列を決定することは有益であることに留意。例えばかかる提唱された配列の一部又は全てを基に、コンセンサス配列を決定することができる。コンセンサス配列は、提唱された配列の全て又は一部を基に決定することができる。コンセンサス配列は、ポリヌクレオチド配列全体に全般的に、又はその配列の一部に局所的に、適用することができる。コンセンサス配列は、項目(A)-(D)の一部又は全てからの信頼値を基に決定することが

50

できる。信頼値は、本実施例において先に更に説明されたものであることができる。信頼値は、配列の一部へ局所的に、又は配列全体に全般的に、適用することができる。最終コンセンサス配列は、前述のアプローチの複数ラウンドにより決定することができ、ここで各ラウンドで得られたコンセンサスは、提唱された配列として使用することができ、且つ各ラウンドに関する信頼性-決定方法は、ラウンド間で異なることができる。

【0231】

一例として、Viterbiアルゴリズムを使用し、ナノポアを通して転位するDNAの完全工程のみ及びフラクショナル工程のみを配列決定することによる(本実施例において先に示した項目(A)及び(B))、2つの異なる提唱された配列を決定することができる。DNAの各ベースに関するこのアルゴリズムの尤度スコアを使用し、提唱された配列の各領域に関する信頼性を決定し、且つ各領域の信頼性の集合(aggregation)は、第一ラウンドの提唱されたコンセンサス配列を生じることができる。このコンセンサス配列は次に、既知の配列のルックアップテーブルに関して、2つの最初に提唱された配列と比較することができる。ルックアップテーブルとこれら3種の提唱された各配列の間の類似性は、3種の提唱された各配列の各領域に関する信頼値を生じることができる。3種の提唱された配列間の信頼性-ベースの比較の第二ラウンドは、最終の提唱されたコンセンサス配列を生じることができる。

【0232】

一部の実施態様において、図22A-22Dは、一部の実施態様に従う、細孔を通じたポリヌクレオチドのフラクショナル転位により提供される情報を使用する説明的方法の工程を図示している。図22Aは、一部の実施態様に従う、細孔を通じたポリヌクレオチドのフラクショナル転位により提供される情報を使用する方法の高レベルの概要を図示している。図22Aに図示された方法は、本明細書の別所により詳細に説明しているような、細孔を通じた標的ポリヌクレオチドのHel308ヘリカーゼによる1以上のフラクショナル転位工程により発生した1以上のシグナルなどの、シグナル獲得(工程2210)を含む。図22Aに図示された方法はまた、例えばシグナル中の異なるシグナルレベルを検出及び同定する、例えば細孔を通じたポリヌクレオチドのフラクショナル転位工程に対応したレベルを検出及び同定する、並びに細孔を通じたポリヌクレオチドの完全転位工程に対応したレベルを同じく検出及び同定する、レベル検出及び同定(工程2220)を含む。図22Aに図示された方法はまた、例えば、シグナル中の検出及び同定された異なるシグナルレベルを基にポリヌクレオチドの配列を特徴決定する、配列決定(工程2230)を含む。図22Aに図示された方法はまた、例えば、配列呼び出しの結果を基に実際のヌクレオチドの可能性のあるヌクレオチド配列を出力する、配列出力(工程2240)を含む。

【0233】

図22B-22Dは、図22Aに図示された工程の1以上の任意の部分工程を図示している。例えば、図22Bは、図22Aに図示された工程2210及び2220の1つの可能性のある実行の追加の詳細を図示している。図22Bに図示された方法はまた、本明細書の別所により詳細に説明しているような、細孔を通じた標的ポリヌクレオチドのHel308ヘリカーゼによる1以上のフラクショナル転位工程により発生した1以上のシグナルなどの、シグナル獲得(工程2210)を含む。図22Bに図示された方法はまた任意に、入力パラメータの獲得(2211)を含むことができる。かかる入力パラメータは、特徴的シグナルの特色がシグナルに対応するよう検出及び決定されるべきであることを規定するパラメータを含むが、これらに限定されるものではない。例えば、入力パラメータは、シグナル値の閾値の大きさの変化を規定することができる。或いは例えば入力パラメータは、完全転位工程に対応するシグナルレベルのみ、又はフラクショナル転位工程に対応するシグナルレベルのみ、又は完全及びフラクショナル転位工程の両方に対応するシグナルレベルが、検出されるべきであることを規定することができる。入力パラメータはまた、レベルを決定する際に考慮することができる、特定のエラーモードの性向及び/又は程度を恐らく含む、エラーモード(例えば、ヌクレオチドスキッピング又はヌクレオチドトグリング)に関連した情報を含むことができる。

入力パラメータはまた、所定のシグナルに関するレベルを決定するために使用することができる、ナノポア、分子モーター及びポリヌクレオチドがその間で相互作用する特定の環境(例えば、温度、塩分濃度、pH、補因子濃度など)に関連した情報を含むことができる。図22Bに図示された方法はまた、細孔を通じたポリヌクレオチドのフラクショナル転位工程に対応するシグナルにおいて、レベルの検出、例えば、異なるシグナルレベルの検出(工程2221)を含む。例えば、工程2210で得られたシグナル及び工程2211で得られた入力パラメータを基に、このようなレベル検出は、そのレベルに対応するシグナルの他の領域とは十分に統計学的に有意に異なるシグナルの領域を検出することができる。レベル検出(これは縁検出又は工程検出とも称される)の例証的方法は、当該技術分野において公知であり、且つスチューデントt-検定及びカイ二乗最大化を含む。工程2221でレベル検出に使用するために適宜改変され得る工程-検出アルゴリズムの一部の例に関して、Carterらの文献、「工程検出法の比較：どの程度成功するか？(A Comparison of Step-Detection Methods: How Well Can You Do?)」、Biophysical Journal 94: 306-308 (January 2008)を参照されたい。

【0234】

図22Bに図示された方法はまた、工程2221のレベル検出を基にしたレベル情報の出力(工程2222)を含む。レベル情報は、平均値、中央値、モード、分布、持続期間、所定のレベルについて検出される最大及び／又は最小電流、若しくはこれらの値の任意の組合せ、又は所定のレベルについての電流値のサブセットに關与するこれらの値を含むことができる(例えば、エラーモードに関連した電流情報を最初に除いた後の平均値電流を利用することができる)。レベル情報はまた、電流の標準偏差、又は電流の周波数帯域制限サブセット(例えば、低域通過、高域通過、帯域通過、若しくは帯域消去フィルター、又はこれらのフィルターの任意の組合せの適用後得られた電流)を含むことができる。レベル情報はまた、レベルの持続期間に関連した情報、並びにこのレベルに関連したエラーモード情報を含むことができる。図22Bに図示された方法はまた、レベル同定(工程2223)、例えば、それについてのレベル情報が工程2222で出力される工程2221で検出されたどのレベルが、標的ポリヌクレオチドの完全又はフラクショナル転位工程に対応するかを決定することを含む。例えば、工程2223は、それについてのレベル情報が工程2222で出力される工程2221で検出された異なるレベルの持続期間を分析し、並びにかかる持続期間を基に、完全転位工程に対応するものとして特定のレベルを同定し、及びフラクショナル転位工程に対応するものとして他の特定のレベルを同定することを含むことができる。一例として、第一の閾値よりも短い持続期間を有するシグナルレベルは、ノイズに対応すると仮定され、従って廃棄されることができるのに対し、第一の閾値よりも長く且つ第二の閾値よりも短い持続期間を有するシグナルレベルは、フラクショナル転位工程に対応すると推定され、従ってそのように同定されるのに対し、第二の閾値よりも長く且つ第三の閾値よりも短い持続期間を有するシグナルレベルは、完全転位工程に対応すると推定され、従ってそのように同定されるのに対し、第三の閾値よりも長い持続期間を有するシグナルレベルは、エラー、又はポリヌクレオチドの不在に対応すると仮定され、従って廃棄されることができる。

【0235】

図22Bに図示された方法はまた、以下の出力の1以上を出力することを含む：完全レベル、フラクショナルレベル、全レベル、及びレベル識別子。例えば、先に注記したように、工程2211で得られた入力パラメータは、完全転位工程に対応するシグナルレベルのみ、又はフラクショナル転位工程に対応するシグナルレベルのみ、又は完全及びフラクショナル転位工程の両方に対応するシグナルレベル(例えば、「全レベル」)が検出されるべきであることを規定することができる。一部の実施態様において、入力パラメータによる「全レベル」の選択は、レベル同定工程を迂回することに対応することができ、その結果このレベル検出工程2221は、全てのレベルを直接出力することに留意。或いは、レベル検出2223及び入力パラメータ2211の結果を基に、所望のシグナルの同定されたレベルは例えば、図22C及び22Dを参照し以下に説明されるような更なる処理のために出力することができる。レベル識別子は、例えば、同定されたレベルに対応する移行の型を示すために、工程2223

時に使用された完全又はフラクショナル工程に関する持続期間を示す指標など、レベルの更なる分析を促進するいずれか好適な情報を含むことができる。

【0236】

再度図22Aを参照し、図22Bに図示された方法を使用するか又は別の好適な方法を使用し作成することができる、完全レベル、フラクショナルレベル、全レベル、及びレベル識別子の1以上を、配列の決定を実行するための入力として使用することができる(図22Aの工程2230)。例えば、図22Cは、例えば、完全レベル、フラクショナルレベル、全レベル、及びレベル識別子の1以上を入力として採用する、かかる完全レベル、フラクショナルレベル、全レベル、及びレベル識別子の1以上を基に配列の決定を実行するための、第一の例証的方法を図示している。図22Cに図示された方法は、完全レベル、フラクショナルレベル、全レベル、及びレベル識別子の1以上の入力を基にした、配列呼び出し工程(工程2231)を含む。配列呼び出しは、標的ポリヌクレオチドのヌクレオチド塩基が、入力シグナルレベルを基に呼び出されるいずれか好適な方法を含むことができる。配列呼び出しの例証的方法は、図19Aを参照し実施例VIIに説明されたものなどのViterbiアルゴリズム、又は図19Bを参照し実施例VIIに説明されたものなどの改変Viterbiアルゴリズム、又は実施例XIに説明されたものに類似のパターンマッチングを含むが、これらに限定されるものではない。配列呼び出しの他の方法を、適切に使用することができる。配列呼び出しの出力(工程2231)は、複数の呼び出された配列、例えば配列A、配列B、・・・配列N、並びにこのような呼び出された各配列に関する信頼情報を含むことができる。様々な呼び出された配列は、別のものよりも、工程2231への様々な入力を基にすることができる。例えば第一の呼び出された配列(例えば配列A)は、工程2210で得られた所定のシグナルを基に完全転位レベルのみが同定される工程2231への入力を基にすることができ、第二の呼び出された配列(例えば配列B)は、フラクショナル転位レベルのみが同定される工程2231への入力を基にすることができ、並びに第三の呼び出された配列(例えば配列N)は、全ての転位レベル(例えば完全及びフラクショナル転位レベルの両方)が同定される工程2231への入力を基にすることができる。或いは又は追加して、他の呼び出された配列は、これを上回るシグナルの大きさの変化はレベルに対応するとして検出することができる、シグナル値の異なる閾値大きさ変化など、どの特徴的シグナルの特色がシグナルに対応するように検出及び決定されるべきかを規定するパラメータの異なる値など、工程2211で得られた代替入力パラメータを基に同定された他のレベルを基にすることができる。各異なる呼び出された配列は、関連した信頼情報、例えば、呼び出された配列が標的ヌクレオチドの実際の配列に対応する尤度を表す値を有することができる。

【0237】

図22Cに図示された実施態様において、配列選択の工程(工程2232)は、呼び出された配列の1以上を選択し、且つ出力として選択された配列を提供することができる(工程2240)。一例として、配列選択の工程(工程2232)は、様々な呼び出された配列に関する信頼情報を比較することを含むことができ、且つ工程2240で最高の信頼性、例えば実際の配列に対応する最高の尤度を有する呼び出された配列を選択及び出力することができる。別の例において、所定の呼び出された配列の信頼情報は、各々、呼び出された配列の対応する部分がその部分について標的ポリヌクレオチドの実際の配列に対応する尤度を表している、複数の信頼値を含むことができる。呼び出された配列の様々な部分について(例えば、10塩基対長、又は50塩基対長、又は100塩基対長、又は10~100塩基対長、又は10~50塩基対長、又は50~100塩基対長である部分)、配列選択の工程(工程2232)は、その部分で様々な呼び出された配列に関して信頼値を比較し、且つその部分について最高値を有する呼び出された配列の一部を選択することを含むことができる。その選択された部分は、そのような部分について最高値を有する他の呼び出された配列の選択された部分と、連鎖されるか又は整列されることができる。

【0238】

図22Dは、配列決定(2230)に使用することができる代替方法を図示している。図22Dに図示された方法は、複数の呼び出された配列、例えば、配列A、配列B、・・・配列Nの入力

、並びにそのような呼び出された各配列に関する信頼情報として得ることを含むことができ、これは図22Cを参照し先に説明されたものに類似することができる。これに関して、具体的に図示されないが、図22Dに図示された方法は、図22Cを参照し先に説明されるものに類似している入力を受け取り、図22Cを参照し先に説明されるものに類似している出力を提供し、並びに工程2231と同様に操作する、配列呼び出しの工程2231を含むことができる。或いは、図22Dに図示された方法は、いずれか他の好適な給源から複数の呼び出された配列を得ることができる。

【0239】

図22Dに図示された方法はまた、モデル配列の獲得(工程2234)を含むことができる。例えば、かかる配列は、1以上の異なる病原体など、1以上の異なる種に関する、先験的に公知の配列を含むことができる。図示されたように、このモデル配列は、非一時的なコンピュータで読み取り可能な媒体に保存された、ルックアップテーブル、データベース、又は他の好適なデータ構造に保存することができる。図22Dに図示された方法はまた、配列選択の工程(工程2233)を含むことができる。図22Dに図示された実施態様において、配列選択の工程は、工程2234で得られたモデル配列の1以上を基に入力として受け取られた呼び出された配列の1以上を選択し、且つ提唱された配列及び新規信頼情報を出力として提供することができる。一例として配列選択の工程(図22Dの工程2233)は、様々な呼び出された配列の1以上を、工程2234で得られたモデル配列の1以上と比較することを含むことができ、且つ最高の新規信頼情報、例えばモデル配列に対応する最高尤度を有する呼び出された配列に対応することができる提唱された配列を選択及び出力することができる。入力信頼情報は、モデル配列(又はモデル配列内の領域)に合致するように、配列(又はその配列内の領域)の尤度により重みをつけ、最も可能性のある配列を決定することができ、これは提唱された配列として出力されることができる。例えば、モデル配列Zと最良のアラインメントである入力配列A、及びモデル配列Yと最良のアラインメントである入力配列Bに関して、提唱された配列は、BとYの間よりも、AとZの間のより良いアラインメントを基にモデル配列Zであることができる。しかし、BとYがより良いアラインメントを有する場合において、BよりもAがより高い信頼値を有する場合は、Zを提唱された配列とすることができる。同じく別のシナリオにおいて、配列の領域を比較し、出力提唱された配列に、A、B、Z、及びYからの配列情報を含ませることができる。或いは、所定の呼び出された配列は、新規信頼情報、例えば各々呼び出された配列の対応する部分が、その部分に関するモデル配列の1以上の部分に対応する尤度を表している、複数の新規信頼値を含むことができる。呼び出された配列の様々な部分(例えば、10塩基対長、又は50塩基対長、又は100塩基対長、又は10~100塩基対長、又は10~50塩基対長、又は50~100塩基対長である部分)について、配列選択の工程(工程2233)は、異なる呼び出された配列について新規信頼値をモデル配列のその部分と比較し、且つその部分に関して最高の新規信頼値を有する呼び出された配列の部分を選択することを含むことができる。その選択された部分は、そのような部分について最高の新規信頼値を有する他の呼び出された配列の選択された部分と、連鎖されるか又は整列されることができる。

【0240】

図22Dに図示された方法は、工程2233による新規信頼情報出力を基に、同じく工程2233により出力される提唱された配列に関する新規信頼情報が、必要要件と合致するかどうかを決定すること(工程2235)を更に含むことができる。一例として、工程2235は、新規信頼値であることができる新規信頼情報を、それ以上では提唱された配列がモデルと十分に合致すると決定され得、並びにそれ未満では提唱された配列がモデルと不十分に合致すると決定され得る、閾値信頼値と、比較することができる。新規信頼情報は、入力信頼性情報の結果、提唱された配列と入力配列の間の関係、提唱された配列とモデル配列の間の関係、及び／又は入力配列とモデル配列の間の関係を含むことができる。例えば提唱された配列が、単純に入力配列のものである場合において、新規信頼情報は、入力配列の入力信頼値とモデル配列と最良のアラインメントのそのアラインメントスコアの間の重み付き平均であることができる。別の場合において、提唱された配列が、入力配列の領域の組合せで

ある場合のように、新規信頼情報は、提唱された配列内の領域間での入力信頼値の重み付き平均とアラインメントスコア(モデル配列に対する)の重み付き平均を含むことができる。工程2235での新規信頼情報は必要要件に合致する(“イエス”)という決定を基に、工程2235は、提唱された配列を出力として提供する(工程2240)。工程2235での新規信頼情報は必要要件に合致しないという決定を基に、工程2235は、工程2233へ戻り、そこで例えば呼び出された配列のモデル配列との更なる比較を行うことにより、配列選択が継続される。配列選択アルゴリズム又はモデル配列のセットは、パラメータによって左右され得、これは、提唱された配列、新規信頼情報、配列選択アルゴリズムが試行される回数、及び既に利用されたモデル配列の1以上を含むことができる。例えば、配列選択アルゴリズムを通る初回通過は、比較的少ないモデル配列を利用することができる(例えば、処理量のために)。しかし、入力配列とモデル配列の間のアラインメントがあまり良くない場合は、新規信頼情報は、必要要件に合致せず、従ってモデル配列の新規の又はより精緻化されたセットとの比較が、工程2233へ戻る際に実行されることができる。

10

【0241】

(実施例XI：任意のSNP同定のためのパターン再認識)

一部の実施態様において、本明細書に開示された方法及び組成物を、マルチプレックス核酸検出、遺伝子型判定及び増幅の方法と組合せて使用することができる。マルチプレックス核酸検出、遺伝子型判定及び増幅の方法は、当該技術分野において周知であり、且つ当業者により容易に選択及び適用することができる。例えば一実施態様において、本明細書に開示された方法及び組成物は、米国特許第6,890,741号、第6,913,884号、第7,955,794号、第7,582,420号、及び第8,288,103号、並びに米国特許出願公開2013-0244882に開示されたマルチプレックス核酸検出、遺伝子型判定及び増幅の方法と組合せて使用することができ、これらは引用により本明細書中に組み込まれている。

20

【0242】

一部の実施態様において、本明細書に開示された方法及び組成物と組合せることができるマルチプレックス核酸検出、遺伝子型判定及び増幅の方法は、アレイ(ランダム及び規則的の両方)又はビーズなどの固体支持体において又はこれと組合せて実行される方法を含む。例えば一部の態様において、ゲノムDNAなどのアッセイされるべき標的ポリヌクレオチドは、固体支持体に固定されることができる。かかる固定された標的ポリヌクレオチドは、当該技術分野において周知のマルチプレックス核酸検出及び遺伝子型判定の方法に供することができる。得られる標的ポリヌクレオチドは、本明細書に開示された方法を使用し特徴づけることができる。

30

【0243】

一部の実施態様において、標的ポリヌクレオチドを特徴決定する方法は、アッセイされるべき標的ポリヌクレオチドを生じるのに必要な工程を更に含むことができる。従って一部の実施態様において、本方法は、(a)各々が、3'側から5'側へ、第一、第二及び第三の標的ドメインを含む複数の標的核酸配列を提供する工程であって、第一標的ドメインが、検出位置を含み、第二標的ドメインは、少なくとも1個のヌクレオチドである工程；(b)標的核酸配列を、各標的配列のためのプローブのセットと接触させ、第一のハイブリダイゼーション複合体のセットを形成する工程であって、プローブの各セットが：5'側から3'側へ、ユニバーサルプライミング配列、及び標的配列の第一の標的ドメインに実質的に相補性である配列、及び検出位置と塩基対形成するのに適した照合位置(例えば3'側の4個の末端塩基内)を含む第一のプローブ、並びに5'側から3'側へ、標的配列の第三の標的ドメインに実質的に相補性である配列、及びユニバーサルプライミング配列を含む第二のプローブであって、ここで任意に少なくとも1種のプローブが、遺伝子座同定配列(例えば、タグ又はバーコード)を含む、工程；(c)ハイブリダイゼーション複合体を、伸長酵素及びdNTPと、照合位置の塩基が検出位置の塩基と完全に相補性である場合、第二の標的ドメインを介した第一のプローブの伸長が生じるような条件下で接触させ、第二のハイブリダイゼーション複合体を形成する工程；並びに、(d)伸長された第一のプローブを、第二のプローブヘライゲーションし、増幅鋳型を形成する工程：を含むことができる。この方法の一

40

50

部の態様において、このプローブセットの第一又は第二のプローブは、対立遺伝子同定配列(例えばタグ又はバーコード)を含むことができる。

【0244】

一部の実施態様において、標的ポリヌクレオチドを特徴決定する方法は、(a)各々が、3'側から5'側へ、第一、第二及び第三の標的ドメインを含む複数の標的核酸配列を提供する工程であって、第一標的ドメインが、検出位置を含み、第二標的ドメインは、少なくとも1個のヌクレオチドである工程；(b)標的核酸配列を、プローブと接触させる工程であって、各プローブが：5'側から3'側へ、ユニバーサルプライミング配列、及び標的配列の第一の標的ドメインに実質的に相補性である配列、及び検出位置と塩基対形成するのに適した照合位置(例えば3'側の4個の末端塩基内)を含み、ここで任意にプローブが、遺伝子座同定配列(例えば、タグ又はバーコード)を含む、工程；(c)ハイブリダイゼーション複合体を、伸長酵素及びdNTPと、照合位置の塩基が検出位置の塩基と完全に相補性である場合、第二及び第三の標的ドメインを介したプローブの伸長が生じるような条件下で接触させ、増幅鋳型として作用することができる伸長されたプローブを形成する工程：を更に含むことができる。

10

【0245】

本明細書に開示された方法においてアッセイするための標的ポリヌクレオチドを生成する方法は、増幅鋳型を増幅し、アンプリコンを生成する工程を更に含むことができる。一部の態様において、第一又は第二のプローブのためにユニバーサルプライミング配列を含むプライマーはまた、どの同定配列が増幅鋳型に既に組込まれているかに応じて、対立遺伝子同定配列又は遺伝子座同定配列(例えばタグ又はバーコード)も含む。これらのアンプリコンは、遺伝子座同定配列及び対立遺伝子同定配列の両方を含むことができるが、これらは、本明細書に開示された方法を用いて特徴づけることができる。標的配列の特徴決定は、遺伝子座及び対立遺伝子の同定配列の存在を基に試料の遺伝子型を示すことができる。

20

【0246】

一部の実施態様において、アンプリコンを作製するために使用されるプライマーは、増幅時に使用される伸長酵素が残基を横断することができないように1以上の修飾された残基を含む。例えば一部の態様において、一つのプライマーは、伸長酵素がプライマー伸長を継続することを防ぐために、脱塩基部位(脱プリン/脱ピリミジン部位)、C3スパーサーホスホロアミダイト(Int C3 Spacer)、トリエチレングリコールスパーサー(Int Spacer 9)又は18-原子ヘキサ-エチレングリコールスパーサー(Int Spacer 18)を含む。当業者は、この同じ機能を実行することができる他の修飾された残基を選択することができるが、理解される。十分な長さの5'側オーバーハングが、本明細書記載の方法を使用し標的ポリヌクレオチドの特徴決定のために作製される限りは、1以上の修飾された残基は、対立遺伝子同定配列内又は対立遺伝子同定配列のいずれかの側に位置することができる。例えば、5'側オーバーハングは、アンプリコンの不動化を可能にするのに十分な長さである。

30

【0247】

一部の実施態様において、前記方法により作製されたアンプリコンは、第二のプローブ配列中又はその近傍に3'側オーバーハングを作製するために、ニッキングエンドヌクレアーゼと更に接触される。このようなニッキング酵素は、2本鎖産物の1本鎖のみを切断するよう、特異性のある配列であることができる。様々なニッキングエンドヌクレアーゼは、当該技術分野において周知であり、且つ当業者は、プローブ及びプライミング配列を基に好適なエンドヌクレアーゼを容易に選択することができることが認められる。ニッキングエンドヌクレアーゼによる切断後に3'側オーバーハングを作製するために、例えばニック鎖のより小さい部分がアンプリコンから放出されるのに対し、アンプリコンの残りは一緒にハイブリダイズされるように、アンプリコンを部分的に変性することを含む、当該技術分野において公知のいくつかの方法を使用することができる。アンプリコンのより小さい部分が除去される便宜を図るために、このより小さい部分のリバース補完物(reverse

40

50

complement)を、望ましくない鎖にハイブリダイズさせるために添加することができる。

【0248】

一部の実施態様において、3'側オーバーハングは、先の方法で説明された第二のプロープ配列中に1以上のウラシル残基を含み、且つウラシルの位置に単ヌクレオチドギャップを特異的に作製するウラシル-特異的酵素とこのアンプリコンを接触させることにより、作製することができる。このようなウラシル-特異的酵素の非限定的例は、Uracil-Specific Excision Reagent (USER (商標)) 酵素 (New England Biolabs 社) である。従って、この酵素により作製されたより小さい分散型断片は、周知の方法を用いアンプリコンから容易に変性され得る。

【0249】

特定の態様において、作製される3'側オーバーハングは、本明細書記載のヘリカーゼの結合を促進するのに十分な長さである。従って一部の態様において、3'側オーバーハングは、長さが少なくとも4のヌクレオチドを含む。別の態様において、3'側オーバーハングは、長さが4~20のヌクレオチドを、又は特定の態様においては8~16の、又は別の態様においては長さが10及び16のヌクレオチドを含む。

【0250】

語句「遺伝子座同定配列」とは、標的ポリヌクレオチドに割り当てられるか、又は標的ポリヌクレオチド上の特定の位置に接続されることが分かっている、核酸残基の配列 (例えばタグ又はバーコード) をいう。標的ポリヌクレオチドの位置は、アッセイされる対立遺伝子に近接しているゲノム上の、例えば、遺伝子、遺伝子の一部分 (例えばエキソン若しくはイントロン) 又は非コード領域 (例えばプロモーター若しくはエンハンサー) であることができる。遺伝子座同定配列は、関心対象の標的配列の位置に特異的である天然の配列及び／又は関心対象の標的配列に対し未変性ではない合成配列であることができる。遺伝子座同定配列は、タグ又はバーコードから予想されるシグナルパターンにより割り当てることができる。

【0251】

語句「対立遺伝子同定配列」とは、標的ポリヌクレオチドの検出位置内である特異的核酸残基に割り当てられた、核酸残基の配列 (例えばタグ又はバーコード) をいう。対立遺伝子同定配列は、検出位置内の核酸残基 (例えば、A、T、C、又はG) の存在を示すことができる。対立遺伝子同定配列もまた、タグ又はバーコードから予想されるシグナルパターンにより割り当てることができる。

【0252】

別の実施態様において、標的ポリヌクレオチドを特徴決定する方法は、図18に説明された工程を更に含むことができる。かかる方法は、(a) 関心対象の異なる標的核酸配列を有する試料を提供する工程であって、ここで異なる標的核酸配列は、固体支持体上に任意に固定されている工程；(b) 試料を、関心対象の異なる標的核酸配列の各々に関するプローブのセットと接触させ、ハイブリダイゼーション複合体を形成する工程であって、各セットが：5'側から3'側へ、第一のユニバーサルプライミング配列、及び第一の標的ドメインに実質的に相補性であり且つ検出位置による塩基対形成に適した照合位置を有する配列を含む、第一のプローブ；並びに、5'側から3'側へ、第三の標的ドメインと実質的に相補性である配列、及び第二のユニバーサルプライミング配列を含む、第二のプローブ：を含み、ここで少なくとも1つのプローブが、関心対象の標的配列に対し未変性ではない遺伝子座同定配列 (例えば、タグ又はバーコード) を含む工程；(c) ハイブリダイゼーション複合体を、伸長酵素及びdNTPと接触させる工程であって、ここで各ハイブリダイゼーション複合体は、照合位置の塩基が検出位置の塩基と完全に相補性である場合、第一のプローブが、第二の標的ドメインに沿って伸長される工程；(d) 伸長された第一のプローブを、第二のプローブヘライゲーションし、増幅鋳型を形成する工程；(e) 増幅鋳型を、第一及び第二のユニバーサルプライマーにより増幅し、アンプリコンを作製する工程であって、ここで少なくとも1つのプライマーが、対立遺伝子同定配列 (例えばタグ又はバーコード) を含み、ここで対立遺伝子同定配列が、脱塩基部位を含む、工程；(f) アンプリコンをニッキ

10

20

30

40

50

ングエンドヌクレアーゼと接触させ、第二のプライマー配列に3'側オーバーハングを製作する工程；並びに、(g)本明細書記載の標的ポリヌクレオチド特徴決定の方法を用い、異なるアンプリコンの遺伝子座同定配列及び対立遺伝子同定配列の両方の存在を検出し、これにより試料中の関心対象の異なる標的配列の存在を示す工程：を含む。

【0253】

本明細書において使用される語句「マルチプレックス」又は文法上の同等物は、2よりも多い関心対象の標的配列の検出、分析又は増幅をいう。一実施態様において、マルチプレックスは、少なくとも100又は200の異なる標的配列をいう一方で、少なくとも500の異なる標的配列が、好ましい。より好ましくは、少なくとも1000で、5000又は10,000より多くが特に好ましく、並びに50,000又は100,000より多くが最も好ましい。検出は、本明細書記載の様々なプラットフォーム上で行うことができる。

10

【0254】

一部の態様において、本明細書の開示は、試料中の核酸標的配列の検出の方法を提供する。当業者が理解するように、試料溶液は、体液(非限定的に、事実上あらゆる生物の、血液、尿、血清、リンパ液、唾液、肛門及び膣の分泌物、汗及び精液を含み、哺乳動物の試料が好ましく、且つヒト試料が特に好ましい)；環境的試料(非限定的に、空気、農業試料、水及び土壌試料を含む)；生物学的軍事物質試料；研究試料；精製されたゲノムDNA、RNA、タンパク質などの、精製された試料；生の試料(細菌、ウイルス、ゲノムDNAなど)を含むが、これらに限定されるものではないいくつかのものを含有し得る。当業者が理解するように、事実上任意の実験操作は、この試料について行われ得る。

20

【0255】

必要ならば、標的ポリヌクレオチドは、公知の技術を用いて、調製される。例えば、当業者に理解されるように、試料は、既知の溶解緩衝液、音波処理、電気穿孔などを用い、細胞を溶解するように処理され、必要に応じて起こる以下に概説する精製及び増幅を伴う。加えて、本明細書に概説した反応は、当業者に理解されるように、様々な様式で達成される。以下に概説する好ましい実施態様により、反応成分は、同時に、逐次に、いずれかの順番で添加される。加えて反応は、アッセイに含まれる様々な他の試薬を含むことができる。これらは、塩、緩衝液、中性タンパク質、例えばアルブミン、界面活性剤などの試薬を含み、これらは、最適なハイブリダイゼーション及び検出を促進し、並びに／又は非特異的若しくはバックグラウンド相互作用を減少するために使用することができる。また、

30

プロテアーゼインヒビター、ヌクレアーゼインヒビター、抗微生物剤などの、そうでなければアッセイの効率を向上する試薬も、試料調製方法及び標的の純度に応じて使用することができる。

【0256】

加えてほとんどの実施態様において、本明細書記載のプライマー及び他のプローブのハイブリダイゼーションを可能にするために、2本鎖標的ポリヌクレオチドは変性され、それらは1本鎖とされる。一実施態様は、pH変化及び他の技術も使用するが、一般に反応の温度を約95℃まで上昇することによる、温熱工程を利用する。

【0257】

本明細書に概説するように、標的ポリヌクレオチドは、反応からの検出配列などの反応生成物、ライゲーションされたプローブ、PCR反応からの伸長されたプローブ、又はPCR増幅産物(“アンプリコン”)などであることができる。

40

【0258】

一部の実施態様において、標的ポリヌクレオチドは、それについて配列情報が望ましい位置を含み、これは一般に本明細書において「検出位置」と称される。特定の実施態様において、検出位置は、単独のヌクレオチドであるが、一部の実施態様において、これは、互いに連続しているか又は1以上のヌクレオチドにより隔てられているかのいずれかの、複数のヌクレオチドを含むことができる。本明細書において使用される「複数」は、少なくとも2を意味する。本明細書において使用されるように、ハイブリッドにおいて検出位置の塩基と塩基対形成する塩基は、「読み出し位置」又は「照合位置」と称され；従って

50

、本発明の第一又は第二工程のプロープの多くは、照合位置を含む。

【0259】

本明細書に開示された方法は、図面に示され且つ本明細書により詳細に説明された、多種多様な立体配置をとることができる。一般にこれらのコンポーネントは、複雑度低下コンポーネント、特異性コンポーネント及び増幅コンポーネントを含む。これらのコンポーネントは、以下に開示された様々な様式で構成され得る。すなわち、一実施態様においては、複雑度低下工程が、最初に行われる。これには、増幅工程又は特異性工程のいずれかが続く。或いは、特異性工程が最初に行われる。これには、複雑度低下工程又は増幅工程が続くことができる。或いは、増幅が最初に行われる。これには、複雑度工程及び特異性工程が続く。

10

【0260】

先に3つのコンポーネントの各々は任意の順番で行うことができることを指摘しているが、当業者は、増幅が最初に行われる場合には、恐らくある程度の複雑度低下又は特異性が関与するであろうことを理解するであろう。加えて、特異性コンポーネントが最初に行われる場合、ある程度の複雑度低下が存在するであろう。加えて一部の実施態様において、増幅が最初に行われる場合、ある程度の特異性及び複雑度低下が存在するであろう。しかし以下に説明するように、本方法は一般に、3コンポーネントを含む。

【0261】

(プローブ及びプライマー)

【0262】

当業者が理解するように、本明細書で開示された方法において使用することができるいくつかのプローブ又はプライマーが存在する。これらのプローブ／プライマーは、様々な立体配置をとることができ、且つ以下により詳細に説明される様々な構造上のコンポーネントを有することができる。第一の工程プローブは、対立遺伝子特異的プローブ又は遺伝子座特異的プローブのいずれかであることができる。「対立遺伝子特異的」プローブ又はプライマーは、標的配列へハイブリダイズし且つ対立遺伝子の間を区別するか、又は標的配列へハイブリダイズし且つ対立遺伝子特異的様式で修飾されるかのいずれかのプローブ又はプライマーを意味する。「遺伝子座特異的」プローブ又はプライマーは、遺伝子座特異的様式で標的配列へハイブリダイズするが、必ずしも対立遺伝子間を区別しないプローブ又はプライマーを意味する。遺伝子座特異的プライマーはまた、修飾され、すなわち以下に説明されるように伸長され、結果的にこれは特定の対立遺伝子に関する情報を含むが、遺伝子座特異的プライマーは、対立遺伝子間を区別しない。

20

30

【0263】

多くの実施態様において、プローブ又はプライマーは、1以上のユニバーサルプライミング部位及び／又は同定配列を含む。例えば一つの立体配置において、4つの対立遺伝子塩基の各々は、異なる配列に、すなわち対立遺伝子同定配列(例えばタグ又はバーコード)、類似の増幅効率を有する各配列に会合している。別の構成において、プローブの1種は、遺伝子座同定配列(例えばタグ又はバーコード)を含む。

【0264】

当業者が理解するように、プライマー及びプローブ核酸のサイズは、変動することができる。プローブの各部分及び変動するプローブの全体の長さは、概して長さ5～500ヌクレオチドを変動する。各部分は、用途及び増幅技術に応じて、長さ10～300、15～250、又は10～35ヌクレオチドであることができる。従って例えば、プローブのユニバーサルプライミング部位は、長さが15～20ヌクレオチドであることができ、いくつかの実施態様において18が使用される。プローブの遺伝子座及び／又は対立遺伝子同定配列は、長さ10～300ヌクレオチドであることができ、いくつかの実施態様においては20～100が使用される。プローブの標的特異的部分は、長さ15～50ヌクレオチドであることができる。加えてプライマーは、追加の増幅プライミング部位を含むことができる。

40

【0265】

一実施態様において、対立遺伝子又は遺伝子座特異的プローブ又はプローブ類は、標的

50

配列の第一のドメインに実質的に相補的である標的ドメインを含む。概してプローブは、標的配列(本明細書に記載のように、試料の標的配列又は他のプローブ配列のいずれかに対し)に対し相補的であるように設計することができ、結果的に標的と本明細書記載のプローブのハイブリダイゼーションが起こる。この相補性は、完全である必要性はなく；標的配列と本発明の1本鎖核酸の間のハイブリダイゼーションを妨害するいくつかの塩基対ミスマッチが存在することができる。しかしこの変異の数が非常に多いと、ハイブリダイゼーション条件の最低ストリンジェント下であっても、ハイブリダイゼーションは起こりえず、この配列は、相補的標的配列ではない。従って本明細書における「実質的に相補的」とは、プローブが、選択された反応条件下でハイブリダイズするのに、標的配列に対し十分に相補性であることを意味する。

10

【0266】

同じく、本明細書記載の方法において使用されるプローブは、後続の増幅スキームのための必須のプライミング部位又は部位類を含むように構築されることができる。いくつかの実施態様において、プライミング部位は、ユニバーサルプライミング部位である。本明細書における「ユニバーサルプライミング部位」又は「ユニバーサルプライミング配列」は、増幅のためにプライマーに結合するプローブの配列を意味する。

【0267】

当業者が理解するように、概して高度に多重化された反応が行われ、全てのユニバーサルプライミング部位は、全ての反応について同じである。或いは、ユニバーサルプライミング部位及び対応するプローブの「セット」を、同時又は連続して使用することができる。ユニバーサルプライミング部位は、本明細書に概略されたように、修飾されたプローブを増幅し、複数のアンプリコンを形成し、その後様々な様式で検出するために使用される。

20

【0268】

従って本明細書記載の方法は、第一の標的プローブセットを提供する。本明細書の「プローブセット」は、特定の多重化されたアッセイにおいて使用される複数の標的プローブを意味する。この文脈において、アッセイ、試料及び試験の目的に応じて、複数は、少なくとも2を意味し、10より多くが好ましい。一実施態様において、プローブセットは、100よりも多くを含み、500よりも多いプローブが好ましく、1000よりも多くが特に好ましい。特に好ましい実施態様において、各プローブは、少なくとも5000を含み、10,000よりも多いプローブが、最も好ましい。

30

【0269】

(複雑度低下コンポーネント)

【0270】

複雑度低下は、本明細書に記す多重スキームセットのコンポーネントであることができる。一般に複雑度低下は、特定の標的又は遺伝子座を濃厚化する方法である。すなわち複雑度低下は、試料からの非-標的核酸の除去又は標的核酸に正確に若しくは全くハイブリダイズしなかったプローブ/プライマーの除去を生じる方法と考えられる。加えて複雑度低下は、酵素工程時に修飾されなかったプローブの除去を含む。すなわち複雑度低下は、酵素工程前、すなわち増幅工程若しくは特異的工程のいずれか、又は両方の前の非-標的核酸の除去、すなわち標的核酸の濃厚化又はハイブリダイズしないプローブ若しくはプライマーの除去を含む。

40

【0271】

複雑度低下工程を含む様々な方法が存在する。これらは、標的核酸又は標的特異の様式で修飾されているプローブ/プライマーの選択的固定、非-標的核酸の選択的除去、並びに非-標的核酸の選択的破壊を含むが、これらに限定されるものではない。かかる破壊は、非-標的核酸の変性、分解又は切断を含むが、これらに限定されるものではない。加えて複雑度低下は、標的選択的増幅などのコンポーネントを含むことができるが、これはまた増幅及びコンポーネントも含む。

【0272】

50

いくつかの実施態様において、複雑度低下は、標的特異の様式で修飾されたプライマーの選択的固定により達成される。すなわち、遺伝子座特異的プライマー又は対立遺伝子特異的プライマーのいずれかは、標的とハイブリダイズされる。この標的は、固定されるか又は溶液中に存在することができる。ハイブリダイゼーション後、プライマーは、プライマー伸長反応において伸長される。一部の態様において、プライマー又はNTPのいずれかは、伸長産物の反応混合物からの除去又は精製を可能にする精製タグを含む。一旦伸長されると、一般に修飾されたプライマーは、固体支持体上に固定されることができ、修飾されたプライマーの固定後、支持体は、洗浄され、非-標的核酸及び修飾されない、すなわち伸長されないプライマーの両方を除去する。こうして固定されたプライマーは、特定の対立遺伝子情報を含む標的遺伝子座に関する情報を含む。これは、標的核酸の濃厚化又は非-標的核酸の除去を生じる。

10

【0273】

別の実施態様において、複雑度低下コンポーネントは、標的ポリヌクレオチドの選択的固定を含む。すなわち標的ポリヌクレオチドは、非-標的核酸よりも支持体上に優先的に固定される。

【0274】

一実施態様において、標的ポリヌクレオチド、プローブ又は修飾されたプライマーを含むプライマーは、固体支持体に付着される。「固体支持体」又は本明細書の他の文法上の同等物は、標的配列の付着に適しているか又はこれに適しているように修飾することができる任意の物質を意味する。当業者が理解するように、可能な基板の数は非常に多い。可能な基板は、ガラス及び修飾又は機能化されたガラス、プラスチック(アクリル系、ポリスチレン、及びスチレンと他の物質のコポリマー、ポリプロピレン、ポリエチレン、ポリブチレン、ポリウレタン、テフロン(商標)などを含む)、多糖、ナイロン又はニトロセルロース、セラミック、樹脂、ケイ素及び修飾ケイ素を含むシリカ又はシリカ-ベースの物質、炭素、金属、無機ガラス、プラスチック、光ファイバー束、及び様々な他のポリマーを含むが、これらに限定されるものではない。磁気ビーズ及びハイスループットマイクロタイタープレートが、特に好ましい。

20

【0275】

固体支持体の組成及び幾何は、その用途により変動する。いくつかの実施態様において、ミクロスフェア又はビーズを含有する支持体を、固体支持体に使用することができる。「ミクロスフェア」又は「ビーズ」又は本明細書の文法上の同等物は、小さい個別の粒子を意味する。ビーズの組成は、生体活性物質のクラス及び合成方法に応じて、変動するであろう。好適なビーズ組成物は、非限定的に、プラスチック、セラミック、ガラス、ポリスチレン、メチルスチレン、アクリル系ポリマー、常磁性金属、トリウム土壌(thoria soil)、カーボングラファイト、二酸化チタン、ラテックス又はセファロースなどの架橋されたデキストラン、セルロース、ナイロン、架橋されたミセル及びテフロンを含む、ペプチド、核酸及び有機部分の合成において使用されるものを含み、並びに固体支持体に関して本明細書において概説された他の物質を全て使用することができる。Bangs Laboratories社(フィッシャー、IN)の「ミクロスフェア検出ガイド(Microsphere Detection Guide)」は、有用な指針である。好ましくはこの実施態様において、複雑度低下が行われる場合、ミクロスフェアは、磁気ミクロスフェア又はビーズである。

30

40

【0276】

一旦固体支持体へ付着されると、標的配列、プローブ又はプライマーは、本明細書に記載のように分析されやすい。

【0277】

高、中及び低ストリンジェンシー条件を含む、様々なハイブリダイゼーション又は洗浄の条件を、本発明において使用することができ；例えば、Maniatisらの文献、「分子クローニング：実験マニュアル(Molecular Cloning: A Laboratory Manual)」、第2版、1989年、及びAusubelら編集の「分子生物学の簡単なプロトコール(Short Protocols in Molecular Biology)」を参照し、これは引用により本明細書中に組み込まれている。ストリン

50

ジェント条件は、配列-依存性であり、且つ異なる状況において異なるであろう。より長い配列は、より高い温度で特異的にハイブリダイズする。核酸ハイブリダイゼーションの広範な指針は、Tijssenの文献、「生化学及び分子生物学の技術－核酸プローブによるハイブリダイゼーション(Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes)」、「ハイブリダイゼーションの原理の概要及び核酸アッセイの戦略(Overview of principles of hybridization and the strategy of nucleic acid assays)」(1993)に認められる。一般に、ストリンジェント条件は、規定されたイオン強度及びpHで、特定の配列について融解温度(T_m)よりも約5~10℃低いように選択される。 T_m は、標的に相補的なプローブの50%が(規定されたイオン強度、pH及び核酸濃度の下で)、平衡状態で、標的配列にハイブリダイズする温度である(標的配列は過剰に存在するので、 T_m で、平衡状態で、プローブの50%は占拠される)。ストリンジェント条件は、塩濃度が、pH7.0~8.3で、約1.0Mナトリウムイオン未満、典型的には約0.01~1.0Mナトリウムイオン濃度(又は他の塩)であり、並びに温度は、短いプローブ(例えば10~50ヌクレオチド)に関しては少なくとも約30℃、及び長いプローブ(例えば50ヌクレオチド以上)に関しては少なくとも約60℃であるものであろう。ストリンジェント条件はまた、ホルムアミドなどのヘリックス不安定化剤の添加により、達成されることができる。

【0278】

本明細書の「伸長酵素」は、NTPの添加により配列が伸長する酵素であることを意味する。当該技術分野において周知であるように、多種多様な好適な伸長酵素が存在し、その中でポリメラーゼ(標的配列及びプレサークル(precircle)プローブの組成に応じて、RNA及びDNAの両方)が、好ましい。好ましいポリメラーゼは、鎖置換活性を欠くものであり、標的化ドメインに対し相補的であるヌクレオチドを含むようにプローブを更に伸長することなく、従って環化を防止するように、プローブの末端に必要な塩基のみ付加することが可能である。好適なポリメラーゼは、DNAポリメラーゼIのクレノウ断片、SEQUENASE 1.0及びSEQUENASE 2.0(U.S. Biochemical社)、T5 DNAポリメラーゼ、Phi29 DNAポリメラーゼ及び例えばサーマス(Thermus)種由来の様々なRNAポリメラーゼ、又はバクテリオファージ由来のQ β レプリカーゼを含む、DNA及びRNAポリメラーゼの両方を含むが、これらに限定されるものではなく、とりわけ、SP6、T3、T4及びT7 RNAポリメラーゼも使用することができる。

【0279】

プローブが、プローブの5'末端を超えて伸長されないことを確実にするために、ポリメラーゼはまた、5'→3'エキソヌクレアーゼ活性が本質的に欠けているものを含むことができる。5'→3'エキソヌクレアーゼ活性を欠いている例証的酵素は、DNAポリメラーゼのクレノウ断片及びDNAPtaqポリメラーゼのStoffel断片を含む。例えばTaq DNAポリメラーゼのStoffel断片は、遺伝子操作のために、5'→3'エキソヌクレアーゼ活性を欠き、これはN-末端の289個のアミノ酸を欠いている切断型タンパク質の生成を生じる(例えば、Lawyerらの文献、J. Biol. Chem., 264:6427-6437 (1989); 及び、Lawyerらの文献、PCR Meth. Appl., 2:275-287 (1993)参照)。類似の変異体ポリメラーゼは、T. maritima、Tsp17、TZ05、Tth及びTaf由来のポリメラーゼについて生成される。

【0280】

追加のポリメラーゼは、3'→5'エキソヌクレアーゼ活性を欠くものであり、これは通常、プルーフリーディング活性と称され、且つプライマー-鋳型二重鎖の3'末端でミスマッチした塩基を除去する。3'→5'エキソヌクレアーゼ活性の存在は、合成された鎖において増大した忠実度を提供するが、Tma(5'→3'エキソヌクレアーゼ活性を欠いているTmaの変異体型を含む)などの熱安定性のDNAポリメラーゼにおいて認められた3'→5'エキソヌクレアーゼ活性もまた、PCRにおいて使用されるプライマー、1本鎖鋳型及び1本鎖PCR産物などの1本鎖DNAを分解する。プライマー伸長プロセスにおいて使用されるオリゴヌクレオチドライマーの3'末端の完全性は、これは新生鎖の伸長が始まるのがこの末端からであるので、重要である。3'末端の分解は、短いオリゴヌクレオチドに繋がり、これは次にプライミング反応の特異性の喪失を生じる(すなわち、より短いプライマーは、偽の又は非特異的

プライミングが生じ始める可能性がより大きい)。

【0281】

更に追加のポリメラーゼは、熱安定性ポリメラーゼである。熱抵抗性酵素は、最適条件下で、40℃で1時間後に、その活性のほとんどを維持している任意の酵素を含むことができる。5'→3'エキソヌクレアーゼ及び3'→5'エキソヌクレアーゼの両方を欠いている熱安定性ポリメラーゼの例は、Taq DNAポリメラーゼのStoffel断片を含む。このポリメラーゼは、遺伝子操作のために、5'→3'エキソヌクレアーゼ活性を欠いており、且つTaqポリメラーゼは、3'→5'エキソヌクレアーゼ活性を天然に欠いているので、3'→5'活性は存在しない。Tth DNAポリメラーゼは、サーマス・サーモフィルス(*Thermus thermophilus*)に由来し、且つEpicentre Technologies社、Molecular Biology Resource社、又はPerkin-Elmer社から入手可能である。3'エキソヌクレアーゼ活性を欠いている他の有用なDNAポリメラーゼは、New England Biolabs社から入手可能なVent[R](exo-)(古細菌サーモコッカス・リトラリス(*Thermococcus litoralis*)由来のDNAポリメラーゼ遺伝子を保持する大腸菌株から精製された)、及びサーマス・フラバス(*Thermus flavus*)由来で、Amersham社から入手できるHot Tub DNAポリメラーゼを含む。熱安定性であり且つ5'→3'エキソヌクレアーゼ活性及び3'→5'エキソヌクレアーゼ活性を欠いた他の好ましい酵素は、AmpliTaq Goldを含む。少なくとも実質的に同等である他のDNAポリメラーゼは、他のN-末端切断型サーマス・アクアティクス(*Thermus aquaticus*)(Taq)DNAポリメラーゼIのように使用することができる。KlenTaq I及びKlenTaq LAと称されるポリメラーゼは、この目的に極めて適している。当然、これらの特徴を有する任意の他のポリメラーゼもまた、本発明に従い使用することができる。

【0282】

プローブの3'末端で1以上のヌクレオチドの付加を実行する条件は、使用される特定の酵素によって左右され、且つ一般に使用される酵素の製造業者により推奨される条件に従う。

【0283】

(特異性コンポーネント)

【0284】

一般に複雑度低下工程後に、特異性工程が、本明細書記載の方法に含まれる。「特異性コンポーネント」は、好ましくは対立遺伝子のレベルで、標的核酸間を区別する工程を意味する。すなわち、特異性コンポーネントは、対立遺伝子特異的工程(例えば遺伝子型判定又はSNP解析)である。特異性の一部のレベルは、対立遺伝子特異的プローブの鋳型(すなわち先の複雑度低下工程の生成物)への単純なハイブリダイズにより、達成され得るが、好ましい実施態様において、特異性工程は、酵素的工程を含む。すなわち、酵素的工程の忠実度は、対立遺伝子区別に関する特異性を改善する。好ましい酵素は、本明細書においてより詳細に説明される、DNAポリメラーゼ、RNAポリメラーゼ及びリガーゼを含む。

【0285】

先に説明されたポリメラーゼはまた、特異性工程に適することができる。

【0286】

多くのリガーゼは、公知であり、且つ本明細書記載の方法における使用に適する。例証的リガーゼは、Lehmanの文献、Science, 186: 790-797 (1974); Englerらの文献、「DNAリガーゼ(DNA Ligases)」、3-30頁、編集Boyer, The Enzymes, Vol. 15B(Academic Press社、ニューヨーク、1982); 及び同類のものに説明されている。好ましいリガーゼは、T4 DNAリガーゼ、T7 DNAリガーゼ、大腸菌DNAリガーゼ、Taqリガーゼ、Pfuリガーゼ、及びTthリガーゼを含む。それらの使用に関するプロトコールは、周知であり、例えばSambrookらの文献(先に言及); Baranyの文献、PCR Methods an Applications, 1: 5-16 (1991); Marshらの文献、Strategies, 5: 73-76 (1992); 及び同類のものにおいて周知である。一般に、リガーゼは、隣接している鎖の3'ヒドロキシルへのライゲーシヨンのために、5'リン酸基が存在することを必要とする。好ましいリガーゼは、熱安定性又は(好熱性)リガーゼ、例えばpfuリガーゼ、Tthリガーゼ、Taqリガーゼ及びAmpligase(商標)DNAリガーゼ(Ep

icentre Technologies社、マジソン、Wis.)を含む。Ampligaseは、低い鈍的末端ライゲーション活性を有する。

【0287】

特定の実施態様では、リガーゼは、最小のミスマッチライゲーションを有するものである。リガーゼの特異性は、より少ない特異性のT4 DNAリガーゼを、大腸菌リガーゼ及び(熱安定性)Taqリガーゼなどのより特異的NAD⁺-依存性リガーゼで置換することにより、増加することができる。ライゲーション反応におけるNAD類縁体の使用は、ライゲーション反応の特異性を更に増加する。Wallace らの米国特許第5,508,179号を参照されたい。

【0288】

一実施態様において、特異性コンポーネントは、固定された標的により実行される。すなわち、複雑度低下工程の生成物は、本明細書に概説したように、固体支持体上に固定される。本明細書において考察したように、特異性反応の標的は、「特異性標的」と称される。すなわち、複雑度低下工程の生成物は、特異性標的である。

10

【0289】

一実施態様において、支持体は、最初の複雑度低下工程におけるものと同じ支持体である。この実施態様において、標的核酸は、特異性アッセイの前に、固体支持体から除去される。標的核酸は、標的核酸の放出を生じるハイブリダイゼーション複合体を変性する任意の方法により除去することができる。当業者が理解するように、この実施態様において、標的核酸は、固体支持体に共有結合されない。すなわちこれは、支持体へ安定して付着される標的プローブである。すなわち、プローブの付着は必ずしも共有的ではないが、これはハイブリダイゼーション複合体の変性及び付着していない標的核酸の除去に耐えるのに十分に安定している。

20

【0290】

代替の実施態様において、特異性標的は、溶液中にある。すなわち、複雑度低下工程後、固定された標的核酸と標的プローブの間のハイブリダイゼーション複合体は、変性され、且つ修飾された標的プローブは、ハイブリダイゼーション複合体から溶離される。特定の実施態様において、特異性標的は、溶液中で分析される。代替の実施態様において、液相の特異性標的は、後続の固体支持体上に固定される。

【0291】

これらの特異性アッセイ、すなわち遺伝子型判定技術は、以下の5つの一般的範疇に収まる：(1)検出位置でヌクレオチドを識別するために、ストリンジェンシー条件(温度、緩衝液条件など)の変動を利用する、従来のハイブリダイゼーション法に頼る技術；(2)検出位置でヌクレオチドと塩基対形成するようある塩基(「この塩基」)を付加する、伸長技術；(3)検出位置に完全な相補性が存在する場合に、ライゲーション反応が優先的に起こるよう、リガーゼ酵素の特異性(又は場合によっては、化学技術の特異性)に頼る、ライゲーション技術；(4)完全な相補性が存在する場合に、優先的に切断が起こるよう、同じく酵素的又は化学的特異性に頼る、切断技術；並びに、(5)これらの方法を組合せた技術。一般に米国特許第6,890,741号、第6,913,884号、第7,955,794号、第7582,420号、及び第8,288,103号、並びに米国特許出願公開2013-0244882を参照し、これらは引用により本明細書中に組み込まれている。

30

40

【0292】

いくつかの実施態様において、伸長遺伝子型判定が行われる。この実施態様において、検出位置に隣接する標的配列にハイブリダイズされたプローブの読み出し位置へヌクレオチドを付加するために、いくつかの技術を使用することができる。酵素的特異性に頼ることにより、優先的に完全に相補性の塩基が付加される。本明細書記載の方法の一部は、検出位置でのヌクレオチドの酵素的組込みに頼っている。これは、単塩基伸長又は多-塩基伸長などの、いくつかの当該技術分野において周知の方法を使用し、行うことができる。いくつかの実施態様において、遺伝子型判定は、鎖終結ヌクレオチドを使用しないプライマー伸長により達成される。従ってこの遺伝子型判定は、多-塩基伸長と考えられる。この方法は、所定のSNPの1つの対立遺伝子を検出するように設計された照合子オリゴヌクレ

50

オチドを提供することを含む。オリゴヌクレオチドの数は、プロービングされる個別のSNP対立遺伝子の数によって決まる。例えば、各々が2つの対立遺伝子を伴う1000 SNPをプロービングする場合、2000オリゴヌクレオチドが必要であろう。照合子は、SNP位置に対応している各照合子の末端塩基によるか、又は照合子の最後の1、2、3又は4つのヌクレオチド内のSNP-特異的位置により、SNP含有DNAの部分配列と相補的である。一部の実施態様において、照合子は、プライマーの末端位置ではないが、プライマーの3'末端から1、2、3、4、5又は6ヌクレオチドの位置の残基である。例えば、SNPがA及びC対立遺伝子を有する場合、T及びGに末端のある照合子が提供され、且つ一部の実施態様において、これら2つを検出するために、個別の要素(ビーズ)上に固定される。マッチ及びミスマッチの両方が所定の対立遺伝子にハイブリダイズするが、マッチのみが、DNAポリメラーゼ伸長反応のプライマーとして作用することができる。従って、標的DNAによるプローブのハイブリダイゼーション後、ポリメラーゼ反応が実行される。これは、dNTPの存在下でのDNAポリメラーゼとのハイブリッドの伸長を生じる。

【0293】

いくつかの実施態様において、伸長しないプローブ又はプライマーは、捕獲プローブへの結合において伸長したプライマーと競合することができるので、アッセイ混合物から、特に固体支持体から、伸長しない又は反応しないプローブ又はプライマーを除去することが、望ましい。十分なプライマーアニーリングを生じるために大きく過剰のプライマーが通常必要とされるので、伸長したプライマーに対する伸長しないプライマーの濃度は、比較的高い。従って、数多くの異なる技術を使用し、伸長されないプローブ又はプライマーの除去を促進することができる。これらは一般に、固体支持体への結合、反応したプライマーの保護及び伸長していないものの分解、並びに反応していないプライマーと及び反応したプライマーの分離による、反応していないプライマーの除去を基にした方法を含む。

【0294】

(増幅コンポーネント)

【0295】

この実施態様において、ポリヌクレオチドの増幅を含む方法が、本明細書に提供され、且つ核酸増幅反応の産物、すなわちアンプリコンを、ポリヌクレオチドを特徴決定する方法において使用することができる。好適な増幅方法は、標的増幅及びシグナル増幅の両方を含む。標的増幅は、検出されるべき標的配列の増幅(すなわち複製)に関与し、標的分子の数の著しい増加を生じる。標的増幅戦略は、ポリメラーゼ連鎖反応(PCR)、鎖置換増幅(SDA)、核酸配列ベースの増幅(NASBA)、及びローリング-サークル増幅(RCA)を含むが、これらに限定されるものではない。かかる増幅戦略は、当業者に周知であり、且つ本記載された方法における使用のために容易に選択することができる。

【0296】

或いは代替技術は、標的の増幅よりもむしろ、シグナリングプローブを複製するための鋳型として標的を使用し、少数の標的分子が、その後検出され得る多数のシグナリングプローブを生じることを可能にする。シグナル増幅戦略は、リガーゼ連鎖反応(LCR)、サイクリングプローブ技術(CPT)、Invader(商標)技術などの侵襲的切断技術、Q-βレプリカーゼ(QR)技術、及び単独の標的配列に結合する複数の標識物プローブを生じる「分岐したDNA」などの「増幅プローブ」の使用を含む。

【0297】

これらの方法は全て、標的配列へハイブリダイズし、ハイブリダイゼーション複合体を形成するプライマー核酸(核酸類縁体を含む)を含むことができ、且つ一部の様式ではプライマーを修飾する酵素が、それに添加され、修飾されたプライマーを形成する。例えばPCRは一般に、2つのプライマー、dNTP及びDNAポリメラーゼを必要とし；LCRは、標的配列に隣接してハイブリダイズする2つのプライマー及びリガーゼを必要とし；CPTは、1つの切断可能なプライマー及び切断酵素を必要とし；侵襲的切断は、2つのプライマー及び切断酵素を必要とするなどである。従って、一般に、標的核酸は、必要な増幅コンポーネントを含有する反応混合物へ添加され、並びに修飾されたプライマーが形成される。

【0298】

概して修飾されたプライマーは、二次反応のための標的配列として役立ち、これは次に数多くの増幅された鎖を生じ、これは本明細書に概説されたように検出されることができ。当業者に理解され且つ本明細書に概説されるように、様々な様式で、必要に応じ、未反応のプライマーは除去される。従ってこの反応は、プライマー核酸の、標的配列への添加により始まり、これはハイブリダイゼーション複合体を形成する。一旦プライマーと標的配列の間にハイブリダイゼーション複合体が形成されたならば、時に「増幅酵素」と称される酵素を使用し、プライマーを修飾する。本明細書に概説される全ての方法に関して、酵素は、プライマーの添加前、添加時又は添加後のいずれかの、アッセイ時の任意の時点で添加されてよい。酵素の同一性は、使用される増幅技術によって決まる。同様に、修飾は、増幅技術によって決まるであろう。

10

【0299】

いくつかの実施態様において、標的増幅技術は、ポリメラーゼ連鎖反応(PCR)である。PCRは、広範に使用され且つ説明され、これは標的配列を増幅するためのサーマルサイクリングと組合せたプライマー伸長の使用に関与している；米国特許第4,683,195号及び第4,683,202号、並びに「PCR必須データ(PCR Essential Data)」、J. W. Wiley & sons、編集C. R. Newton、1995年を参照し、これらは全て引用により本明細書中に組み込まれている。加えて、同じく本発明における使用が認められる、数多くのPCRの変形が存在し、これはとりわけ、「定量競合的PCR」又は「QC-PCR」、「任意配列プライマー使用するPCR」又は「AP-PCR」、「イムノ-PCR」、「Alu-PCR」、「PCR単鎖高次構造多型」又は「PCR-SSCP」、「逆転写酵素PCR」又は「RT-PCR」、「ビオチン捕獲PCR」、「ベクトレッテPCR」、「パンハンドルPCR」、並びに「PCRセレクトcDNAサブトラクション」、「対立遺伝子-特異的PCR」を含む。当業者は、本明細書記載の方法において使用することができる好適なPCRの変形を容易に選択することができることが、理解される。

20

【0300】

いくつかの実施態様において、増幅反応は、本明細書記載のように多重増幅反応である。一実施態様において、増幅反応は、複数の標的配列を増幅するために、複数のPCRプライマーを使用する。この実施態様において、複数の標的配列は、複数の増幅プライマー対により、同時に増幅される。

【0301】

30

代替の実施態様で、多重PCR反応は、本明細書記載のユニバーサルプライマーを使用する。すなわちユニバーサルPCRプライマーは、標的配列上のユニバーサルプライミング部位にハイブリダイズされ、これにより複数の標的配列を増幅する。この実施態様は、限定された数のPCRプライマーのみを必要とするので、潜在的に好ましい。すなわち、わずかに1つのプライマー対が、複数の標的配列を増幅することができる。

【0302】

Golden Gateアンプリコンは、先に説明したように鋳型としてヒトDNAを用いて作製される(Cold Spring Harb Symp Quant Biol. 2003;68:69-78、「高度に並行したSNP遺伝子型判定(Highly parallel SNP genotyping)」、Fan JBらの文献)。得られるアンプリコンは、対立遺伝子に応じて、P1及びP2と称される2つのプライマーの一方を有する。更に、ユニバーサルリバースプライマー(“リバースP3”)が、全てのアンプリコン上に存在する。

40

【0303】

P1:

【化13】

TCTCGTCGCTCATCAACT (配列番号: 82)

【0304】

P2:

【化 1 4】

GAGTCGAGGTCATATCGT (配列番号: 83)

【0 3 0 5】

リバーソP3:

【化 1 5】

GTCTGCCTATAGTGAGTC (配列番号: 84)

10

【0 3 0 6】

16サイクルを使用するPCRの第二ラウンドを使用し、“P1_バーコード_A”及び“P1_バーコード_B”と称される、対立遺伝子をバーコードしているプライマーを追加した。複数のデオキシウラシル残基を含む伸長されたユニバーサルリバーソプライマー(“ユニバーサルdUリバーソ”)を使用した。

【0 3 0 7】

P1_バーコード_A:

【化 1 6】

/5phos/TTTTTTTTTTTTTTTCCTTCCXXTTTTCTTCTTCTCAAGAAGAAGATCTCGT
CGCTCATCAACT (配列番号: 85)

20

【0 3 0 8】

P2_バーコード_B:

【化 1 7】

/5phos/TTTTTTTTTTTTTTTCCTTCCXXTTTTTAATTAATTTTGTGTTGTGAGTCG
AGGTCATATCGT (配列番号: 86)

【0 3 0 9】

ユニバーサルdUリバーソ:

【化 1 8】

ATACGGCG/dU/CCACCGACC/dU/CAGCGTC/dU/GCCTATAG/dU/GAGTC
(配列番号: 87)

30

【0 3 1 0】

ここで/5phos/は5'リン酸を意味し、/dU/はデオキシウラシル塩基であり、並びにXは脱塩基部分である。

【0 3 1 1】

PCR後、この試料は、USER酵素(New England Biolabs社、イプスウィッチ、MA)と共に、2.5時間、37℃でインキュベーションし、dU残基が配置された1本鎖のギャップを作出した。この試料を、65℃で10分間加熱し、断片化されたDNAを除去し、3'側オーバーハングを作出した。PCR Cleanup Kit(Qiagen社)を使用し、試料を精製した。

40

【0 3 1 2】

試料を、コレステロール-含有オリゴ“P3_Chol”へ、分子比1:1で、65℃で加熱し、且つゆっくり冷却することにより、アニーリングした。

【0 3 1 3】

P3_Chol:

【0 3 1 4】

50

【化 19】

ACCGACACTGCGTCTGCCTATAGTGAGTC/iSp9//3CholTEG/ (配列番号:88)

【0315】

ここで、/iSp9/は9-原子トリエチレングリコールスパーサーを意味し、並びに/3CholTEG/は3' コレステロールTEG(トリエチレングリコール)部分を意味する。

【0316】

脂質二重層は、1,2-ジフィタノイル-sn-グリセロ-3-ホスホコリン(Avanti Polar Lipids社)から形成した。この二重層は、テフロン中に水平方向に直径~20ミクロンの開口部を広げた。M2-NNN-MspAを、濃度~2.5ng/mlで二重層の基底側に添加した。一旦単独の細孔が挿入されたならば、更なる挿入を避けるために、コンパートメントを、実験用緩衝液でフラッシュした。Axopatch-200Bパッチクランプ増幅器(Axon Instruments社)で、二重層を横切る電圧180mVを印加し、且つイオン電流を測定した。類縁体シグナルは、4-極Besselフィルターにより、50kHzで低域フィルターを通過させ、その後低域フィルター周波数を5回デジタル化した。データ獲得は、LabWindows/CVI(National Instruments社)により書かれた特注のソフトウェアにより制御した。

【0317】

二重層の両側の~60 µlコンパートメントは、0.4M KCl、1mM EDTA、1mM DTT、1mM ATP、10mM MgCl₂、及び10mM HEPES/KOHで緩衝されたpH8.0の実験用緩衝液を含んだ。野生型Hel308 Tgaを、分子モーターとして、150nmで使用した。

【0318】

図23は、一部の実施態様に従う、それぞれのバーコードとして使用するのに適した、第一の説明的ポリヌクレオチド配列(配列番号:89)及び第二の説明的ポリヌクレオチド配列(配列番号:90)に関する、時間の関数としての、発生し得る例証的シミュレーションされたシグナルを図示している。図24において、Hel308ヘリカーゼによる細孔を通じた第一の説明的ポリヌクレオチド配列のフラクショナル転位に対応するシミュレーションされたシグナル(1)は、比較的高シグナルレベルでの2つの“ピーク”、それに続く下落、それに続く比較的低シグナルレベルでの2つ以上の“ピーク”を含む特徴的パターンを有するのに対し、Hel308ヘリカーゼによる細孔を通じた第二の説明的ポリヌクレオチド配列のフラクショナル転位に対応するシミュレーションされたシグナル(2)は、比較的低シグナルレベルでの2つの“ピーク”、それに続く増加、それに続く比較的高シグナルレベルでの2つ以上の“ピーク”を含むことを認めることができる。従って、シミュレーションされたシグナル(1)及び(2)のような個別の特徴を含む実際のシグナルは、例えばパターンマッチングを使用し、他方から容易に識別することができ、従ってアッセイ結果の他方からの識別を促進することができることを予想することができる。例えば、図24A-24Dは、一部の実施態様に従う、それぞれのバーコードとして使用するのに適した、第一及び第二の説明的ポリヌクレオチド配列に関する、時間の関数としての、発生し得る例証的シミュレーションされたシグナルを図示している。これは、図24A及び24Bにおいて一般に“バーコード”と指定されるシミュレーションされたシグナルのセクションは、比較的高シグナルレベルでの2つの“ピーク”、それに続く下落、それに続く比較的低シグナルレベルでの2つ以上の“ピーク”を含むことを認めることができ、従って第一の説明的ポリヌクレオチド配列に対応すると理解することができる。同じく図24C及び24Dにおいて一般に“バーコード”と指定されるシミュレーションされたシグナルのセクションは、比較的低シグナルレベルでの2つの“ピーク”、それに続く増加、それに続く比較的高シグナルレベルでの2つ以上の“ピーク”を含むことを認めることができ、従って第二の説明的ポリヌクレオチド配列に対応すると理解することができる。

【0319】

別の例において、本実施例において先に説明されたものと同様の様式で、2NNN MspA細孔が、DPhPC脂質二重層に挿入された。この緩衝液は、400mM KCl、10mM HEPES、pH8、5mM

MgCl₂、及び1mM EDTAを含んだ。これらの試薬は、1mM DTT及び1mM ATPを含んだ。酵素は、およそ150mM Hel308 Tgaを含んだ。DNAは、およそ10nMであり、且つ配列決定された1本鎖(RS1801131 SNP1及びSNP2と称される)は、コレステロール-含有ポリヌクレオチドへハイブリダイズした。このような鎖の配列決定時に得られたシグナルは、レベル-発見及び本明細書において別所記載のようなアルゴリズムを使用する予測された配列へのアラインメントを含む、ポスト-プロセッシングを用いて解読した。

【0320】

図25A及び25Bは、各々、一部の実施態様に従う、それぞれのバーコードとして使用するのに適した、第一及び第二の説明的ポリヌクレオチド配列に関する、時間の関数として発生し得る例証的シミュレーションされたシグナルを図示している。図25A及び25Bにおいて点線のボックス内のシミュレーションされたシグナルのセクションは、各々、rs1801131 SNP1及びrs1801131 SNP2と各々指定された配列に関する独特のパターンを含み、従ってそれぞれのバーコードとして使用することができる。バーコードとして使用される配列は、図23に図示したものと同一であった。

【0321】

図26A-26Dは、各々、一部の実施態様に従う、それぞれのバーコードとして使用するのに適した、第一及び第二の説明的ポリヌクレオチド配列に関する、時間の関数として発生し得る例証的測定されたシグナルを図示している。図26A及び26Bにおいて点線のボックス内の測定されたシグナルのセクションは、各々、rs1801131 SNP1と指定された配列のバーコードに対応することが認められる独特のパターンを含むのに対し、図26C及び26Dにおいて点線のボックス内の測定されたシグナルのセクションは、各々、rs1801131 SNP2と指定された配列のバーコードに対応することが認められる独特のパターンを含み、同じくrs1801131 SNP1と指定される配列のバーコードから容易に識別される。

【0322】

(参考文献の組み込み)

本出願を通じて、括弧内に又は括弧なしで、様々な刊行物が言及されている。これらの刊行物の開示は、それらの全体が、この開示が関与する技術分野の状況をより完全に説明するという目的を含むが、これに限定されるものではない全ての目的のために、本出願中に引用により組み込まれている。

【0323】

(他の代替実施態様)

本明細書に提供されるシステム及び方法は、本明細書に開示されたオペレーションを実施するために、指示(例えば、ソフトウェア指示)を実行する様々な種類のデータプロセッサ環境を用いて(例えば、1以上のデータプロセッサ上で)実行され得ることは留意されるべきである。非限定的例は、単独の汎用コンピュータ若しくはワークステーション上で、又はネットワークシステム上での、或いはクライアント-サーバー構成においての、又はアプリケーションサービスプロバイダーの構成においての、実行を含む。例えば本明細書記載の方法及びシステムは、装置の処理サブシステムにより実行可能であるプログラム指示を含むプログラムコードにより、多くの異なる型の処理装置において、実行することができる。ソフトウェアプログラム指示は、ソースコード、オブジェクトコード、マシンコード、又は処理システムに本明細書記載の方法及びオペレーションの実行を起こすよう操作可能であるいずれか他の保存されたデータを含むことができる。しかしファームウェアなど、又は本明細書記載の方法及びシステムを実行するために構築された適宜設計されたハードウェアであっても、また他の実装も、使用することができる。

【0324】

本システム及び方法は、1以上のデータ処理装置との連絡のために、ネットワーク(例えば、ローカルエリアネットワーク、ワイドエリアネットワーク、インターネット、それらの組合せなど)、光ファイバー媒体、搬送波、ワイヤレスネットワークなどを介して搬送されるデータシグナルを含むことができることも、更に注目される。データシグナルは、装置に又は装置から提供される本明細書に明らかにされたデータのいずれか又は全てを搬

送することができる。

【0325】

本システム及び方法のデータ(例えば、関連付け、データ入力、データ出力、中間データ結果、最終データ結果など)は、異なる種類の記憶装置及びプログラミング構成体(例えば、RAM、ROM、フラッシュメモリー、単層ファイル、データベース、プログラミングデータ構造、プログラミング変数、IF-THEN(又は類似型)ステートメント構成体など)などの、1以上の異なる種類のコンピュータ-実装されたデータ格納庫に、保存し且つ実装することができる。データ構造は、コンピュータプログラムにより使用するための、データベース、プログラム、メモリー、又は他のコンピュータ-読み取り可能な媒体における、データの組織化及び保存において使用するためのフォーマットを説明することに留意されたい。

10

【0326】

本システム及び方法は、本方法のオペレーションを行い且つ本明細書記載のシステムを実行するためにプロセッサによる実行において使用するための指示(例えばソフトウェア)を含む、コンピュータ記憶機構(例えば、CD-ROMなど持続型媒体、ディスク、RAM、フラッシュメモリー、コンピュータのハードドライブなど)を含む、多くの異なる種類のコンピュータ-読み取り可能な記憶媒体上で、更に提供することができる。

【0327】

更に、本明細書に提供されるコンピュータ構成要素、ソフトウェアモジュール、関数、データ格納及びデータ構造は、それらのオペレーションに必要とされるデータフローを可能にするために、互いに直接又は間接に接続され得る。モジュール又はプロセッサは、ソフトウェアオペレーションを実施するコードのユニットを含むが、これらに限定されるものではなく、並びに例えば、コードのサブルーチンユニットとして、若しくはコードのソフトウェア関数ユニットとして、又はオブジェクトとして(オブジェクト指向パラダイムにおいてとして)、又はアプレットとして、又はコンピュータスクリプト言語において、又は別の種類のコンピュータコードとして、実行されることができるとも、留意されたい。ソフトウェア構成要素及び／又は機能性は、単独のコンピュータ上に配置するか、又は当面の状況に応じて複数のコンピュータにわたって分配することができる。

20

【0328】

本開示は、明らかにされた実施態様を参照し説明されているが、当業者は、先に詳述された具体的な実施例及び試験は、本開示の単なる例証であることを容易に理解するであろう。本開示の精神から逸脱しない、様々な変更を行うことができることは理解されるべきである。従って、本開示は、以下の請求項によってのみ限定される。

30

本件出願は、以下の構成の発明を提供する。

(構成1)

標的ポリヌクレオチドを特徴決定する方法であって：

(a)Hel308ヘリカーゼ及び標的ポリヌクレオチドと接触している細孔を横切るポテンシャルの差を印加する工程；

(b)該細孔を通じた該標的ポリヌクレオチドの該Hel308ヘリカーゼによる1以上のフラクショナル転位工程により発生した1以上のシグナルを測定する工程；並びに

(c)該フラクショナル転位工程により発生した該1以上のシグナルから該標的ポリヌクレオチドを特徴決定する工程：を含む、前記方法。

40

(構成2)

前記標的ポリヌクレオチドの特徴決定が、該標的ポリヌクレオチドの配列、該標的ポリヌクレオチドの修飾、該標的ポリヌクレオチドの長さ、該標的ポリヌクレオチドの同一性、該標的ポリヌクレオチドの給源、及び該標的ポリヌクレオチドの二次構造：の1以上を同定することを含む、構成1記載の方法。

(構成3)

前記ポテンシャルの差が、電気ポテンシャルの差を含む、構成1記載の方法。

(構成4)

前記1以上のシグナルが、電氣的シグナルを含む、構成1記載の方法。

50

(構成 5)

前記1以上のシグナルが、光学的シグナルを含む、構成1記載の方法。

(構成 6)

前記工程(a)－(c)を1回以上反復することを更に含む、構成2～5のいずれか一項記載の方法。

(構成 7)

前記フラクショナル転位工程が、該Hel308ヘリカーゼの完全転位サイクルの第一のフラクショナル転位工程を含む、構成1記載の方法。

(構成 8)

前記フラクショナル転位工程が、該Hel308ヘリカーゼの完全転位サイクルの第二のフラクショナル転位工程を含む、構成1記載の方法。

10

(構成 9)

前記標的ポリヌクレオチドの転位が、該細孔を通じて転位するポリヌクレオチド上のポテンシャルの差により印加された加力の反対方向である、構成1記載の方法。

(構成 10)

前記標的ポリヌクレオチドの転位が、該細孔を通じて転位するポリヌクレオチド上のポテンシャルの差により印加された加力の方向である、構成1記載の方法。

(構成 11)

前記電氣的シグナルが、電流、電圧、トンネル電流、抵抗、ポテンシャル、電圧、コンダクタンス、及び横方向電氣的測定値から選択された測定値である、構成4記載の方法。

20

(構成 12)

前記電氣的シグナルが、該細孔を通過する電流を含む、構成11記載の方法。

(構成 13)

前記標的ポリヌクレオチドの1以上のヌクレオチド残基が、完全転位サイクルから得られた単独の電氣的シグナルを使用する1以上のヌクレオチドの特徴決定と比べ50%より大きい精度で、完全転位サイクルの2つのフラクショナル工程から得られた電氣的シグナルを用いて特徴決定される、構成4記載の方法。

(構成 14)

前記細孔が、生物学的細孔である、構成1記載の方法。

(構成 15)

前記生物学的細孔が、ポリペプチド細孔である、構成14記載の方法。

30

(構成 16)

前記生物学的細孔が、ポリヌクレオチド細孔である、構成14記載の方法。

(構成 17)

前記ポリペプチド細孔が、5ヌクレオチド以下の狭窄ゾーンを有する、構成15記載の方法。

(構成 18)

前記ポリペプチド細孔が、マイコバクテリウム・スメグマチスポリリンA(MspA)を含む、構成15記載の方法。

(構成 19)

前記MspAが、配列番号:1のアミノ酸配列、又は配列番号:1との少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、又は少なくとも70%の相同性を有するアミノ酸配列を有する、構成18記載の方法。

40

(構成 20)

前記細孔が、固体状態の細孔である、構成1記載の方法。

(構成 21)

前記細孔が、生物学的及び固体状態のハイブリッド細孔である、構成1記載の方法。

(構成 22)

前記生物学的及び固体状態のハイブリッド細孔が、ポリペプチドー固体状態のハイブリ

50

ッド細孔である、構成21記載の方法。

(構成 2 3)

前記生物学的及び固体状態のハイブリッド細孔が、ポリヌクレオチドー固体状態のハイブリッド細孔である、構成21記載の方法。

(構成 2 4)

前記Hel308ヘリカーゼが、表1及び表2に示されたヘリカーゼ又はそれらの変種である、構成1記載の方法。

(構成 2 5)

前記標的ポリヌクレオチドが、1本鎖ヌクレオチド、2本鎖ヌクレオチド、及び部分的2本鎖ポリヌクレオチドからなる群から選択される、構成1記載の方法。

(構成 2 6)

細孔を通じた標的ポリヌクレオチドのフラクショナル転位工程を調節する方法であって：

(a)Hel308ヘリカーゼ及び標的ポリヌクレオチドと接触している細孔を横切るポテンシャルの差を印加する工程；

(b)該Hel308ヘリカーゼをHel308ヘリカーゼ基質の参照濃度とは異なる濃度の該基質と接触させる工程であって、該基質濃度が、該参照濃度と比べた該基質濃度の差異に比例してフラクショナル転位工程の持続期間に変化を生じる工程；並びに

(c)該細孔を通じた該標的ポリヌクレオチドの該Hel308ヘリカーゼによる1以上のフラクショナル転位工程により発生した1以上のシグナルを測定する工程：を含む、前記方法。

(構成 2 7)

前記1以上のフラクショナル転位工程により発生した該1以上のシグナルから該標的ポリヌクレオチドを特徴決定する工程を更に含む、構成26記載の方法。

(構成 2 8)

前記標的ポリヌクレオチドの特徴決定が、該標的ポリヌクレオチドの配列、該標的ポリヌクレオチドの修飾、該標的ポリヌクレオチドの長さ、該標的ポリヌクレオチドの同一性、該標的ポリヌクレオチドの給源、及び該標的ポリヌクレオチドの二次構造：の1以上を同定することを含む、構成27記載の方法。

(構成 2 9)

前記ポテンシャルの差が、電気ポテンシャルの差を含む、構成26記載の方法。

(構成 3 0)

前記1以上のシグナルが、電氣的シグナルを含む、構成26記載の方法。

(構成 3 1)

前記1以上のシグナルが、光学的シグナルを含む、構成26記載の方法。

(構成 3 2)

前記基質濃度が、該Hel308ヘリカーゼ基質のサブ飽和濃度である、構成26記載の方法。

(構成 3 3)

前記参照濃度が、該Hel308ヘリカーゼ基質の飽和濃度である、構成26記載の方法。

(構成 3 4)

前記基質濃度及び該参照濃度の両方が、該基質の飽和濃度ではない、構成26記載の方法。

(構成 3 5)

前記基質濃度及び該参照濃度が、該Hel308ヘリカーゼ基質のサブ飽和濃度である、構成26記載の方法。

(構成 3 6)

前記Hel308ヘリカーゼ基質が、アデノシン三リン酸(ATP)である、構成26記載の方法。

(構成 3 7)

前記フラクショナル転位工程が、該Hel308ヘリカーゼの完全転位サイクルの第一のフラクショナル転位工程を含む、構成26記載の方法。

(構成 3 8)

10

20

30

40

50

前記フラクショナル転位工程が、該Hel308ヘリカーゼの完全転位サイクルの第二のフラクショナル転位工程を含む、構成26記載の方法。

(構成39)

前記標的ポリヌクレオチドの転位が、該細孔を通じて転位するポリヌクレオチド上のポテンシャルの差により印加された加力の反対方向である、構成26記載の方法。

(構成40)

前記標的ポリヌクレオチドの転位が、該細孔を通じて転位するポリヌクレオチド上のポテンシャルの差により印加された加力の方向である、構成26記載の方法。

(構成41)

前記電気的シグナルが、電流、電圧、トンネル電流、抵抗、ポテンシャル、電圧、コンダクタンス、及び横方向電気的測定値から選択された測定値である、構成30記載の方法。

(構成42)

前記電気的シグナルが、該細孔を通過する電流を含む、構成41記載の方法。

(構成43)

前記標的ポリヌクレオチドの1以上のヌクレオチド残基が、完全転位サイクルから得られた単独の電気的シグナルを使用する1以上のヌクレオチドの特徴決定と比べ50%より大きい精度で、完全転位サイクルの2つのフラクショナル工程から得られた電気的シグナルを用いて特徴決定される、構成30記載の方法。

(構成44)

前記標的ポリヌクレオチドの1以上のヌクレオチド残基が、該参照濃度と比べより低い基質濃度でより大きい精度で特徴決定される、構成26記載の方法。

(構成45)

前記細孔が、生物学的細孔である、構成26記載の方法。

(構成46)

前記生物学的細孔が、ポリペプチド細孔である、構成45記載の方法。

(構成47)

前記生物学的細孔が、ポリヌクレオチド細孔である、構成45記載の方法。

(構成48)

前記ポリペプチド細孔が、5ヌクレオチド以下の狭窄ゾーンを有する、構成46記載の方法。

(構成49)

前記ポリペプチド細孔が、マイコバクテリウム・スメグマチスポリリンA(MspA)を含む、構成46記載の方法。

(構成50)

前記MspAが、配列番号:1のアミノ酸配列、又は配列番号:1との少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、又は少なくとも99%の相同性を有するアミノ酸配列を有する、構成49記載の方法。

(構成51)

前記細孔が、固体状態の細孔である、構成26記載の方法。

(構成52)

前記細孔が、生物学的及び固体状態のハイブリッド細孔である、構成26記載の方法。

(構成53)

前記生物学的及び固体状態のハイブリッド細孔が、ポリペプチドー固体状態のハイブリッド細孔である、構成52記載の方法。

(構成54)

前記生物学的及び固体状態のハイブリッド細孔が、ポリヌクレオチドー固体状態のハイブリッド細孔である、構成52記載の方法。

10

20

30

40

50

(構成 5 5)

前記Hel308ヘリカーゼが、表1及び表2に示されたヘリカーゼ又はそれらの変種である、構成26記載の方法。

(構成 5 6)

前記標的ポリヌクレオチドが、1本鎖ポリヌクレオチド、2本鎖ポリヌクレオチド、及び部分的2本鎖ポリヌクレオチドからなる群から選択される、構成26記載の方法。

(構成 5 7)

1mM未満のATPを含むか又はヌクレオチド類縁体を含む溶液中に含有された、細孔、Hel308ヘリカーゼ及び標的ポリヌクレオチドを含む、標的ポリヌクレオチドを特徴決定するための組成物。

(構成 5 8)

前記1mM未満のATPを含む溶液が、0.1 μ M、1.0 μ M、10 μ M、100 μ M、0.5mM及び0.9mMのATPからなる群から選択される濃度を含む、構成57記載の組成物。

(構成 5 9)

前記細孔が、生物学的細孔である、構成57記載の組成物。

(構成 6 0)

前記生物学的細孔が、ポリペプチド細孔である、構成59記載の組成物。

(構成 6 1)

前記生物学的細孔が、ポリヌクレオチド細孔である、構成59記載の組成物。

(構成 6 2)

前記ポリペプチド細孔が、5ヌクレオチド以下の狭窄ゾーンを有する、構成60記載の組成物。

(構成 6 3)

前記ポリペプチド細孔が、マイコバクテリウム・スメグマチスポリンA(MspA)を含む、構成58記載の組成物。

(構成 6 4)

前記MspAが、配列番号:1のアミノ酸配列、又は配列番号:1との少なくとも15%、少なくとも20%、少なくとも25%、少なくとも30%、少なくとも35%、少なくとも40%、少なくとも45%、少なくとも50%、少なくとも55%、少なくとも60%、少なくとも65%、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、又は少なくとも99%の相同性を有するアミノ酸配列を有する、構成63記載の組成物。

(構成 6 5)

前記細孔が、固体状態の細孔である、構成57記載の組成物。

(構成 6 6)

前記細孔が、生物学的及び固体状態のハイブリッド細孔である、構成57記載の組成物。

(構成 6 7)

前記生物学的及び固体状態のハイブリッド細孔が、ポリペプチドー固体状態のハイブリッド細孔である、構成66記載の組成物。

(構成 6 8)

前記生物学的及び固体状態のハイブリッド細孔が、ポリヌクレオチドー固体状態のハイブリッド細孔である、構成66記載の組成物。

(構成 6 9)

前記Hel308ヘリカーゼが、表1及び表2に示されたヘリカーゼ又はそれらの変種である、構成57記載の組成物。

(構成 7 0)

前記標的ポリヌクレオチドが、1本鎖ポリヌクレオチド、2本鎖ポリヌクレオチド、及び部分的2本鎖ポリヌクレオチドからなる群から選択される、構成57記載の組成物。

(構成 7 1)

前記特徴決定が、改変されたViterbiアルゴリズムを適用することを含む、構成1記載の

10

20

30

40

50

方法。

(構成 7 2)

前記方法が：

(d) 工程(c)の後に、該細孔を通じた該標的ポリヌクレオチドの該Hel308ヘリカーゼによる1以上のフラクショナル転位工程の時機を変動するために、少なくとも1種のパラメータを変動する工程；並びに

(e) 変動された少なくとも1種のパラメータを用い、工程(a)－(c)を反復する工程：を更に含む、構成1記載の方法。

(構成 7 3)

工程(c)及び(e)の間に発生したシグナルを組合せること、並びに組合せたシグナルを基に該標的ポリヌクレオチドを特徴決定することを更に含む、構成72記載の方法。

10

(構成 7 4)

前記変動された少なくとも1種のパラメータが、温度、塩濃度、補因子濃度、ATP生成物の濃度、pH、及び使用した特定のHel308ヘリカーゼからなる群から選択される、構成73記載の方法。

(構成 7 5)

前記特徴決定が、1以上のシグナル中のレベルを検出及び同定し、並びに検出され且つ同定されたレベルを基に標的ポリヌクレオチドの配列を決定及び出力することを含む、構成1記載の方法。

(構成 7 6)

前記1以上のシグナル中のレベルの検出及び同定が、完全レベル、フラクショナルレベル、全レベル、及びレベル識別子の1以上の出力を含む、構成75記載の方法。

20

(構成 7 7)

前記検出及び同定されたレベルを基にした標的ポリヌクレオチドの配列の決定及び出力が、完全レベル、フラクショナルレベル、全レベル、及びレベル識別子の1以上を入力として採用し、該入力を基に複数の配列を呼び出し、並びに呼び出された複数の配列に関する信頼情報を基に呼び出された配列の少なくとも1種を選択及び出力することを含む、構成76記載の方法。

(構成 7 8)

前記検出及び同定されたレベルを基にした標的ポリヌクレオチドの配列の決定及び出力が、完全レベル、フラクショナルレベル、全レベル、及びレベル識別子の1以上を入力として採用し、該入力を基に複数の配列を呼び出し、並びに呼び出された複数の配列の一部に関する信頼情報を基に複数の呼び出された配列の互いの部分を選択及び連鎖させることを含む、構成76記載の方法。

30

(構成 7 9)

前記検出及び同定されたレベルを基にした標的ポリヌクレオチドの配列の決定及び出力が、完全レベル、フラクショナルレベル、全レベル、及びレベル識別子の1以上を入力として採用し、該入力を基に複数の配列を呼び出し、呼び出された配列をモデル配列と比較し、並びに呼び出された配列のモデル配列との比較に関する信頼情報を基に呼び出された配列の少なくとも1種を選択及び出力することを含む、構成76記載の方法。

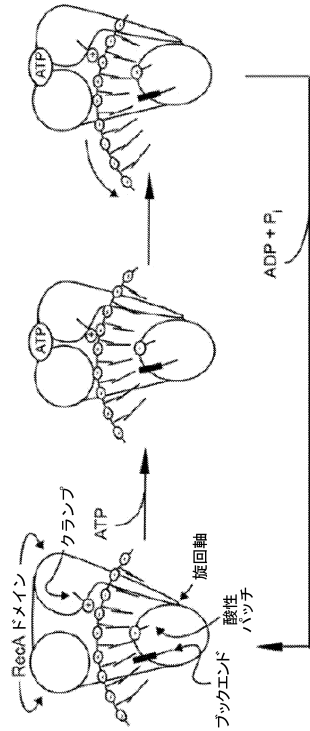
40

(構成 8 0)

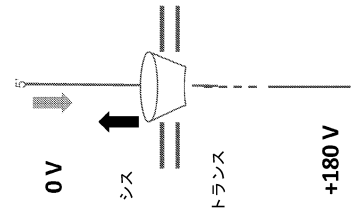
前記検出及び同定されたレベルを基にした標的ポリヌクレオチドの配列の決定及び出力が、完全レベル、フラクショナルレベル、全レベル、及びレベル識別子の1以上を入力として採用し、該入力を基に複数の配列を呼び出し、呼び出された配列をモデル配列と比較し、並びに呼び出された複数の配列の一部のモデル配列との比較に関する信頼情報を基に複数の呼び出された配列の互いの部分を選択及び連鎖することを含む、構成76記載の方法。

。

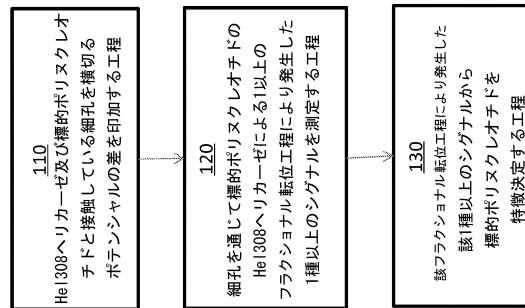
【図 1 A】



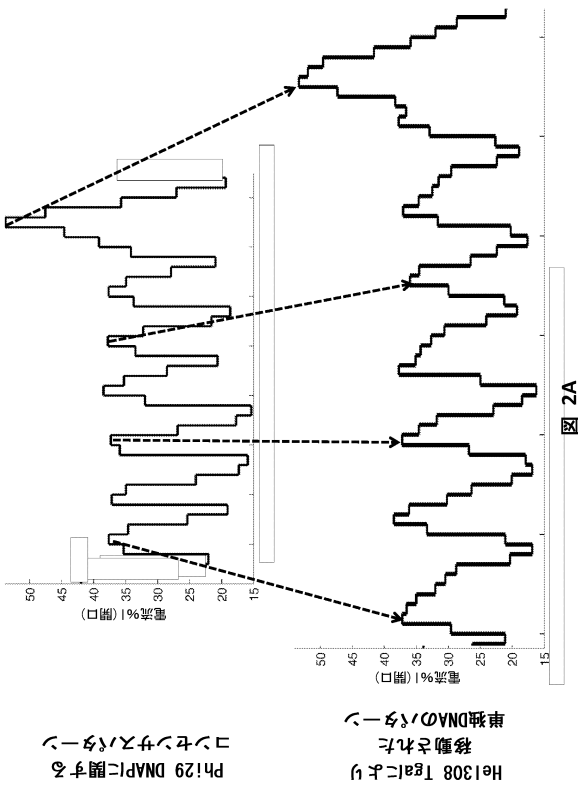
【図 1 B】



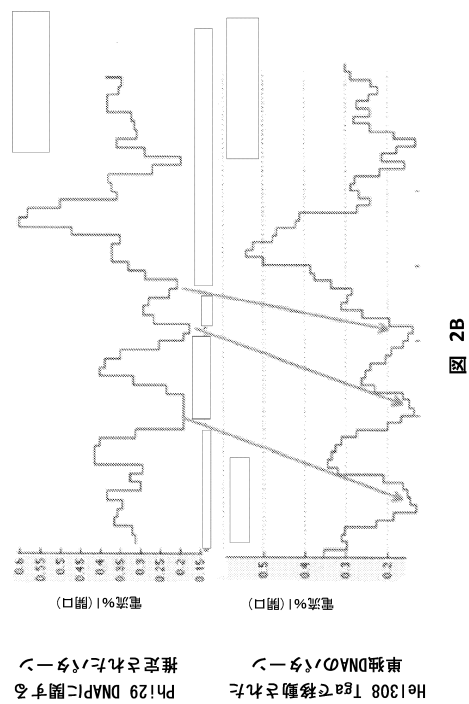
【図 1 C】



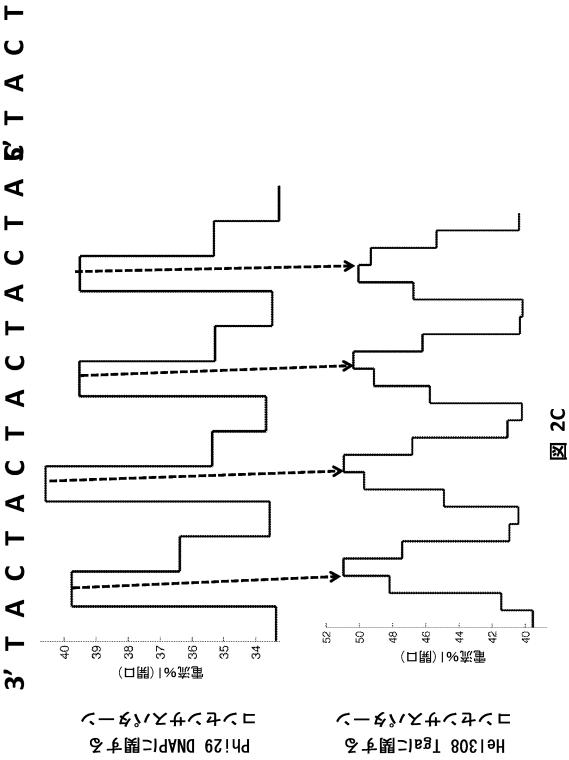
【図 2 A】



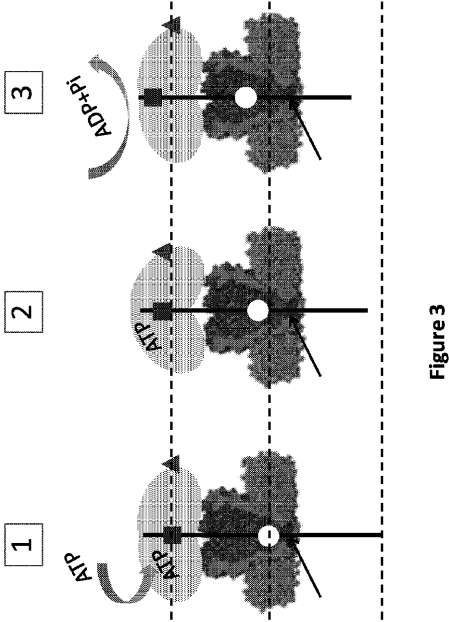
【図 2 B】



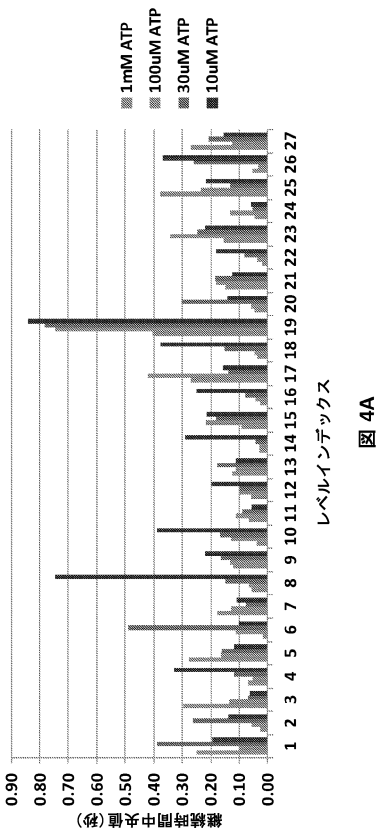
【図 2 C】



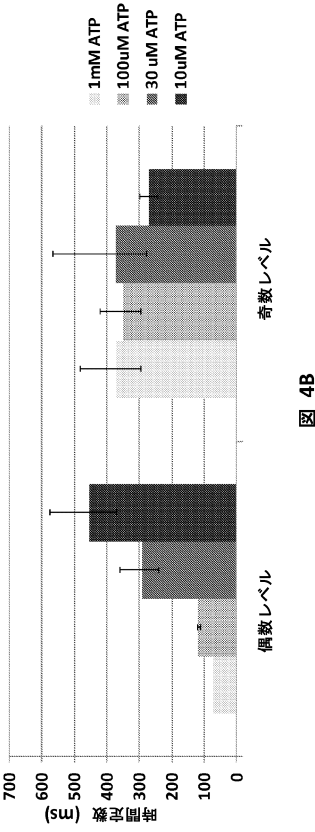
【図 3】



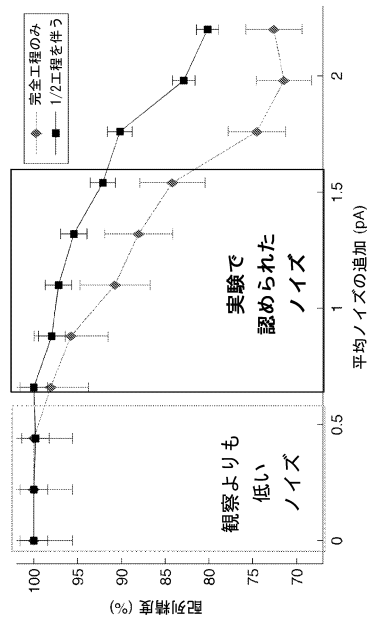
【図 4 A】



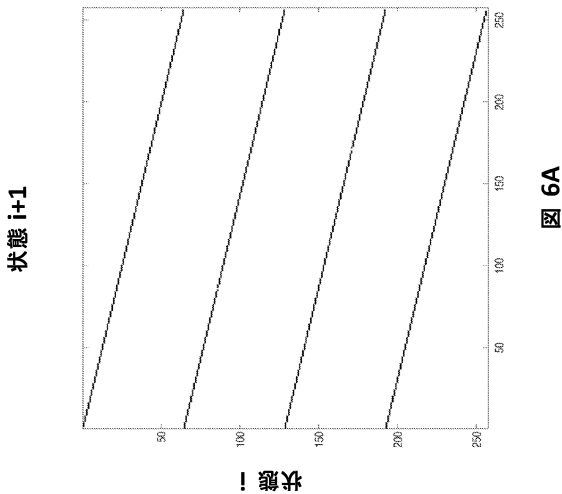
【図 4 B】



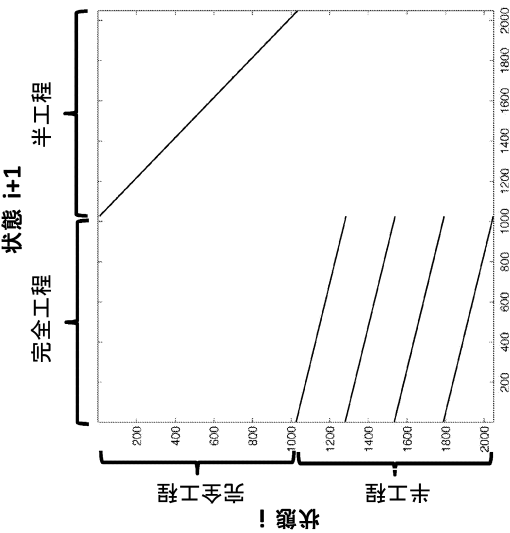
【図 5】



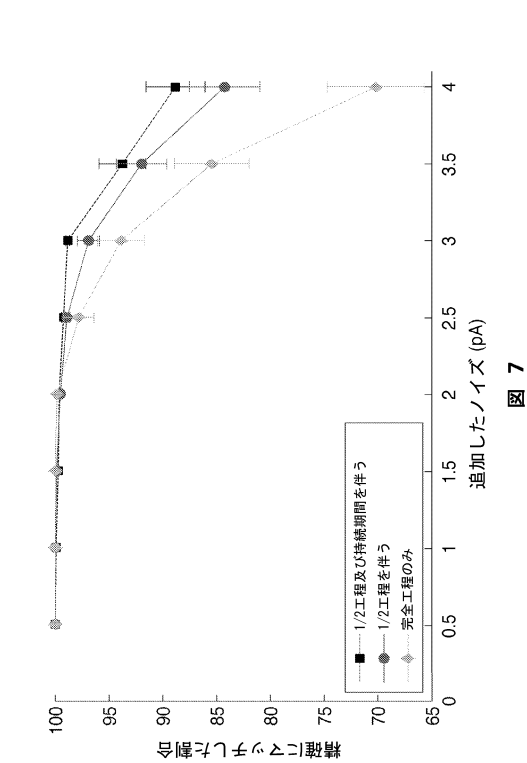
【図 6 A】



【図 6 B】



【図 7】



【図 8】

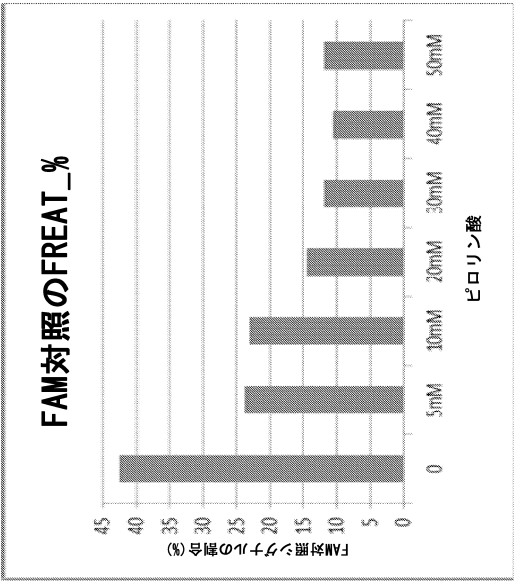


図 8

【図 9】

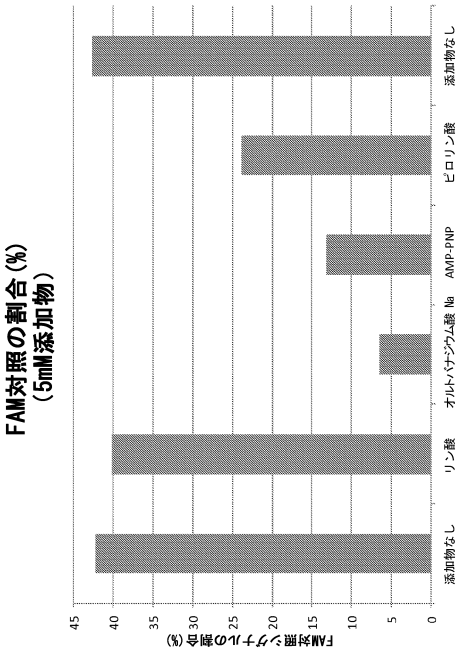


図 9

【図 10】

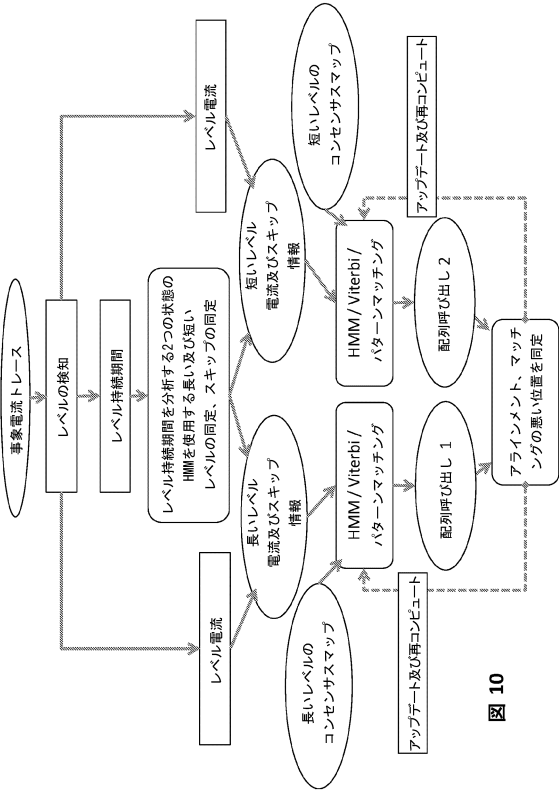


図 10

【図 11】

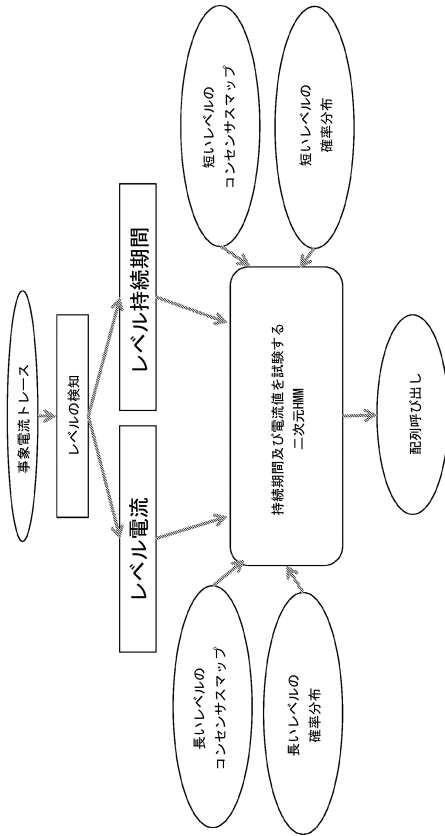


図 11

【図 1 2】

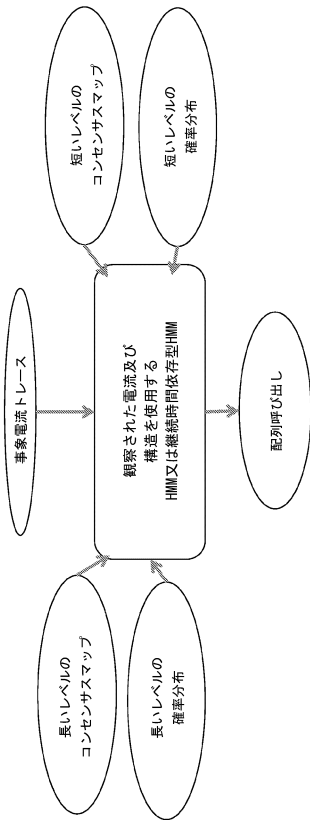
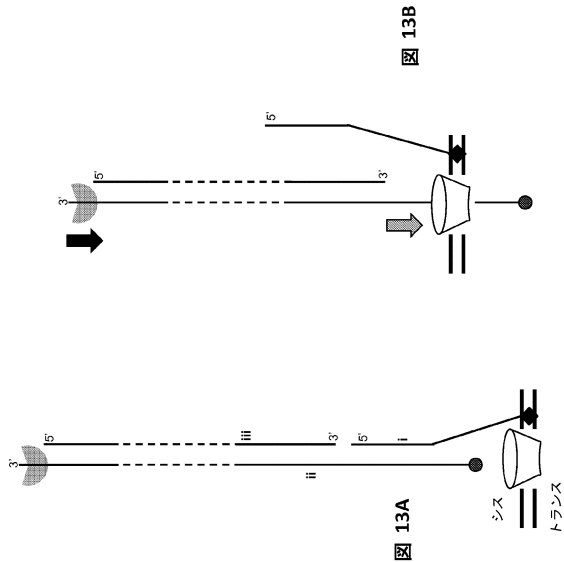
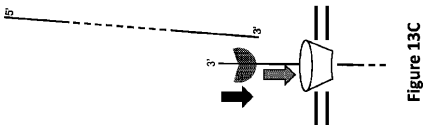


図 12

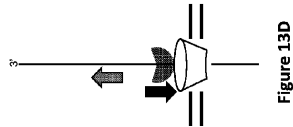
【図 1 3 A – B】



【図 1 3 C】



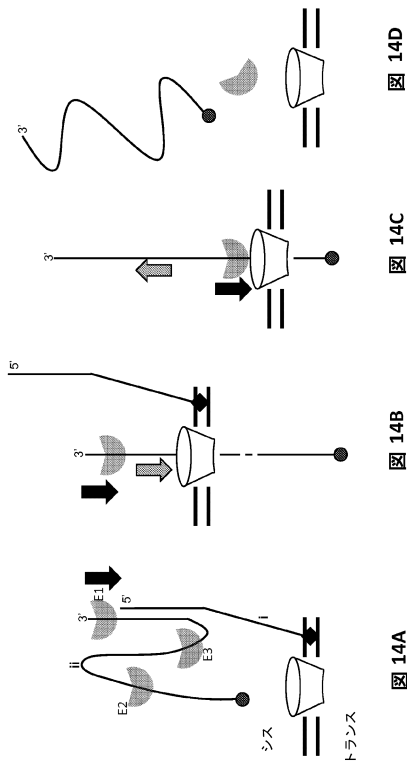
【図 1 3 D】



【図 1 3 E】



【図 1 4】



【図 1 5】

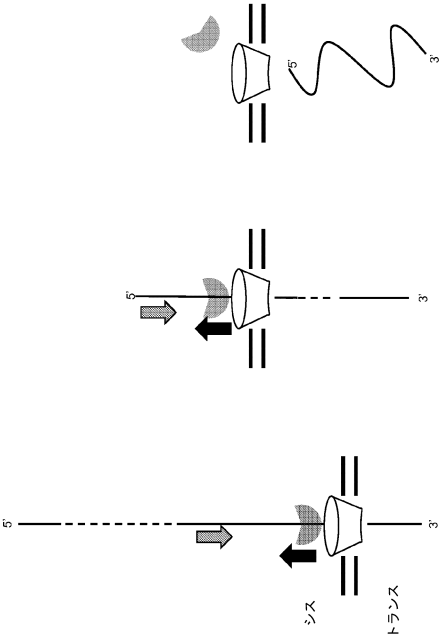


図 15A

図 15B

図 15C

【図 1 6】

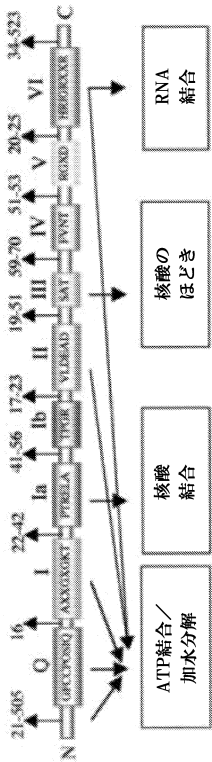
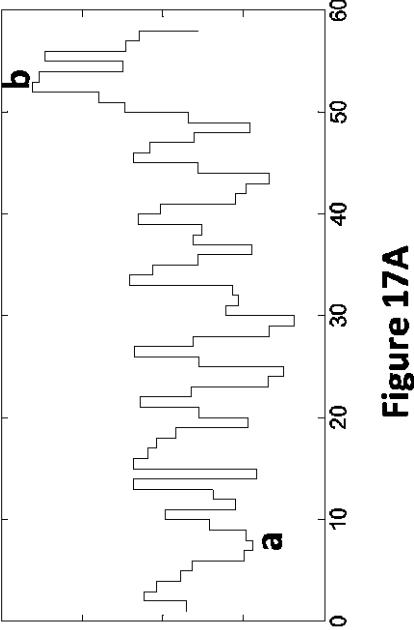
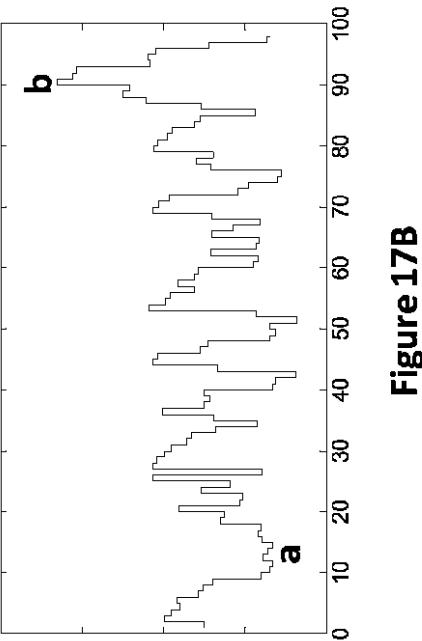


図 16. Tutejaらの文献から改作

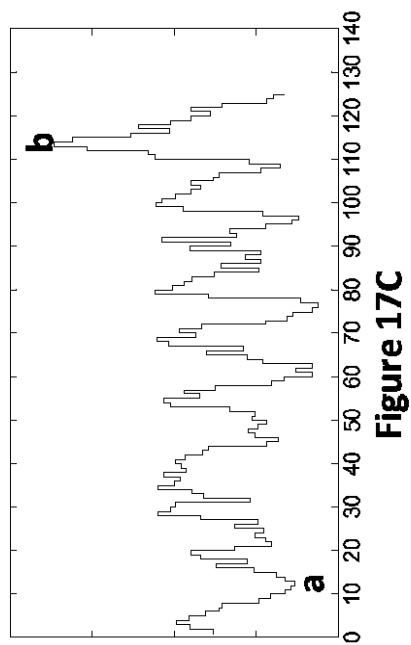
【図 1 7 A】



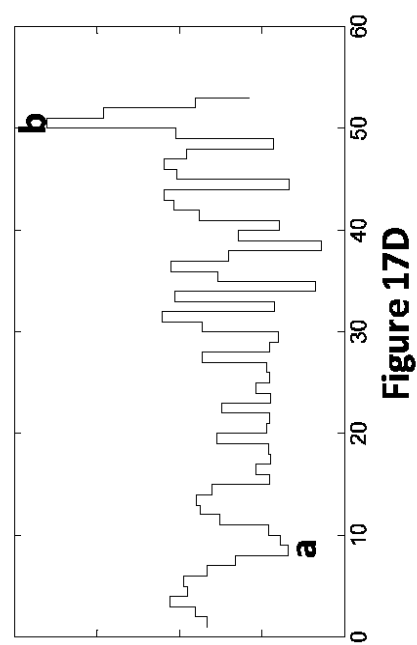
【図 1 7 B】



【図 17 C】



【図 17 D】



【図 18】

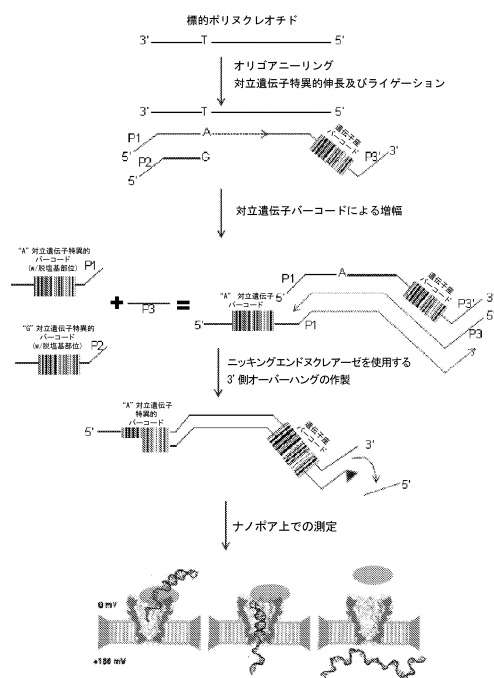


図 18

【図 19 A】

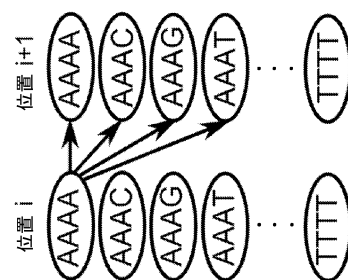


図 19A

【図 19 B】

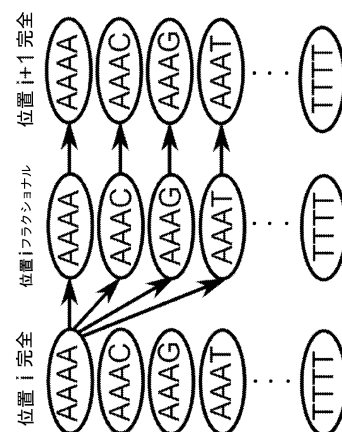
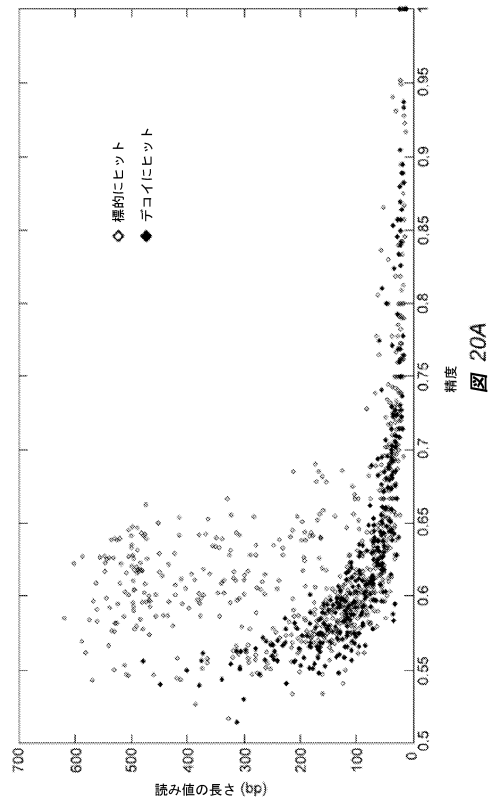
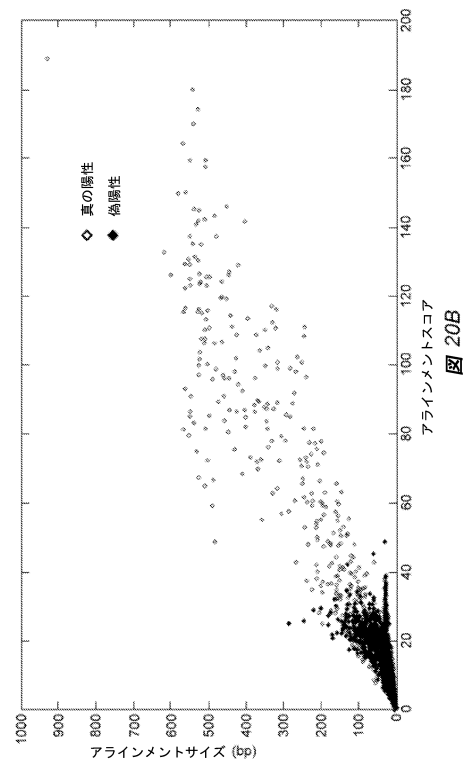


図 19B

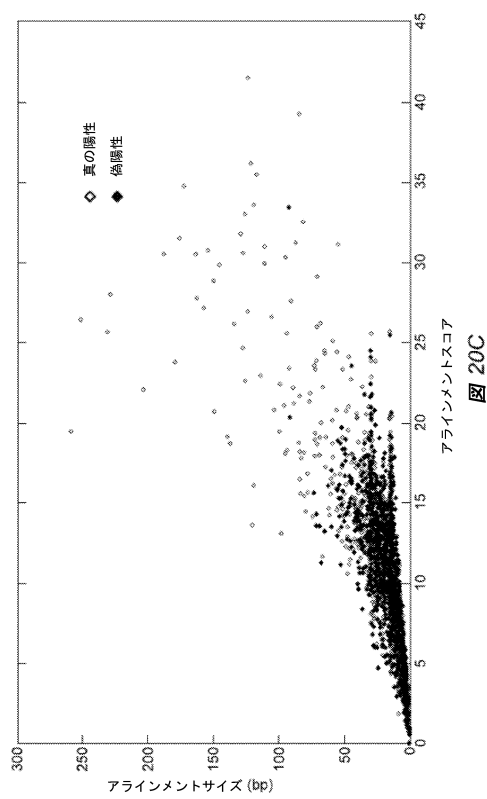
【図 20 A】



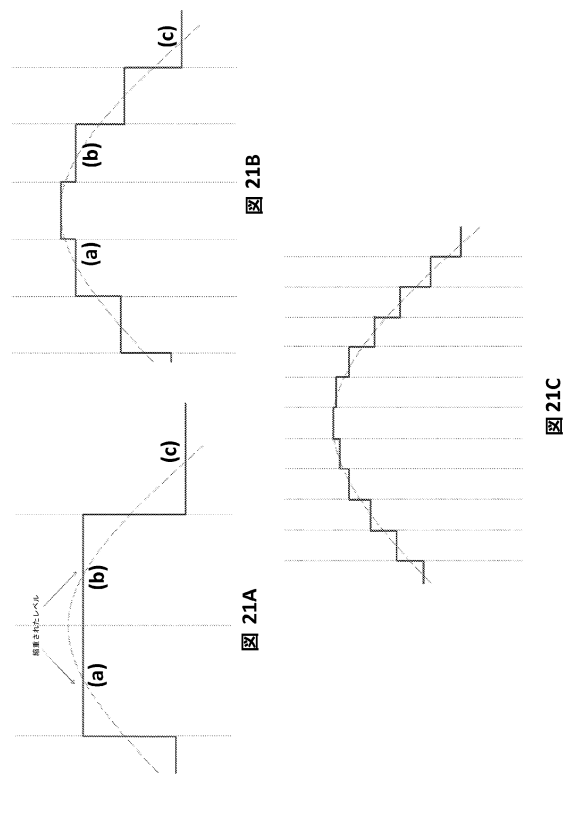
【図 20 B】



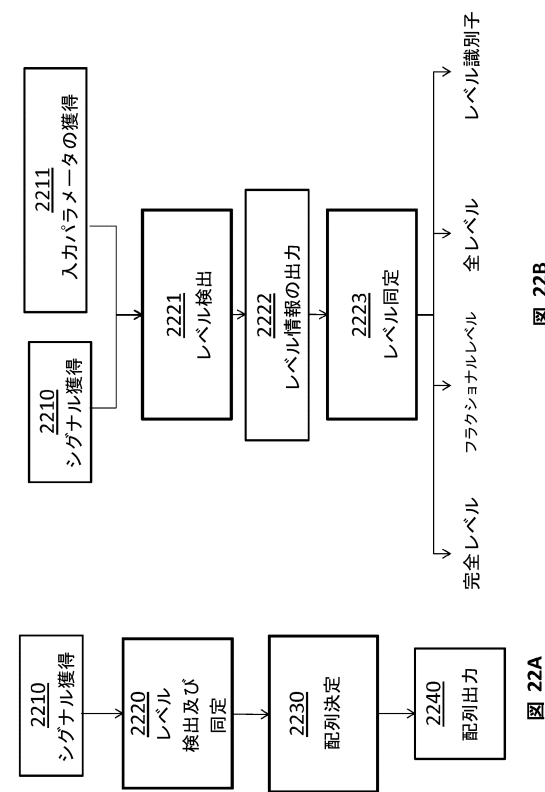
【図 20 C】



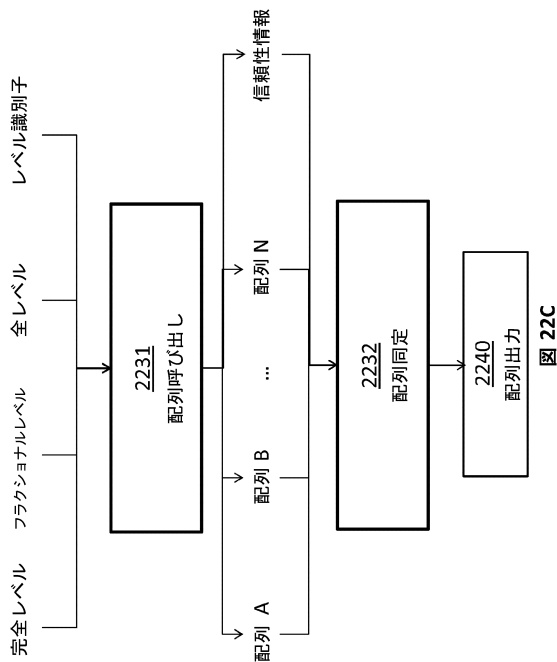
【図 21】



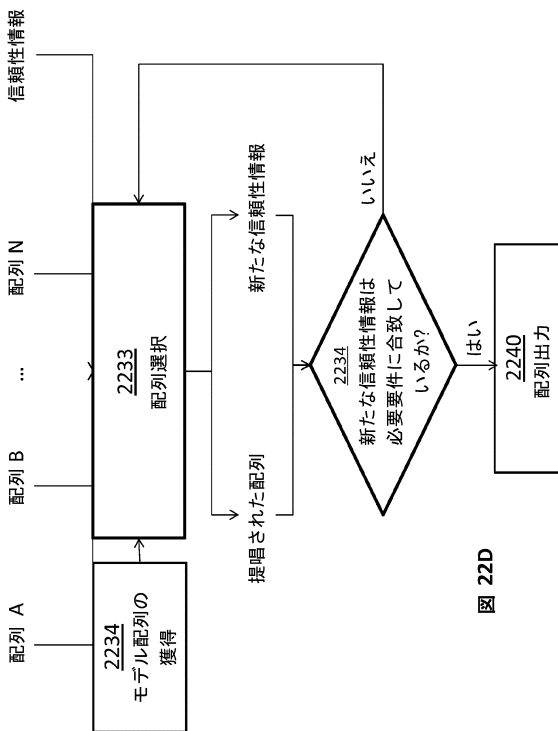
【図 2 2 A - B】



【図 2 2 C】



【図 2 2 D】



【図 2 3】

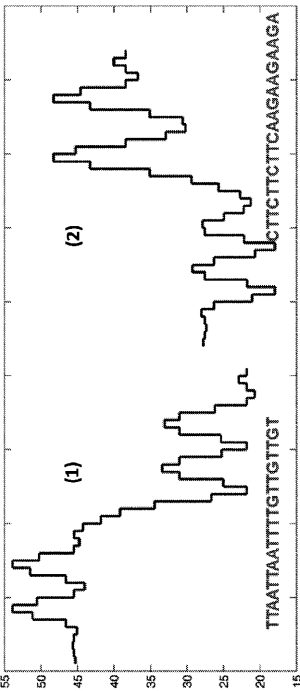
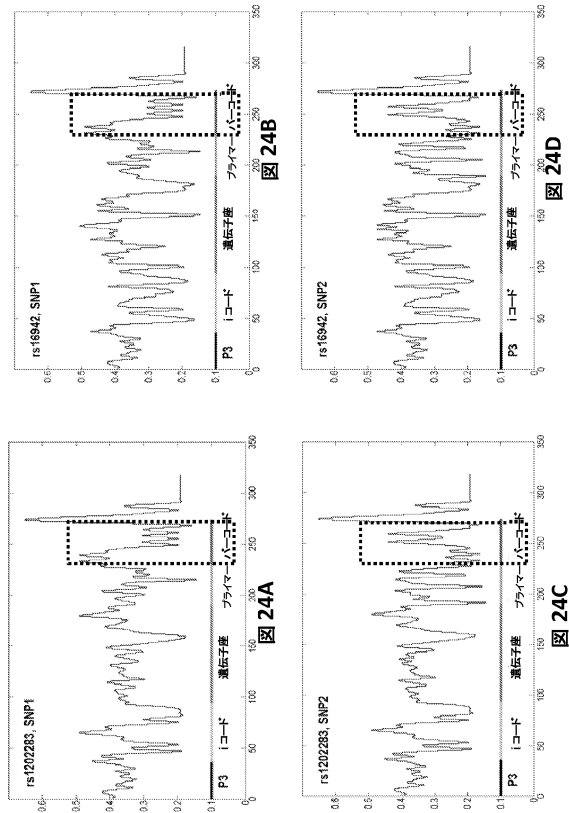


Figure 23

【図 2 4】



【図 2 5 A】

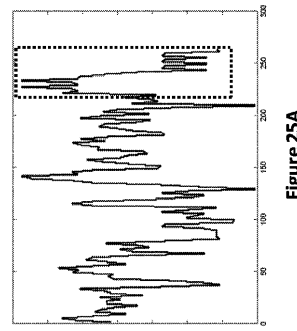


Figure 25A

【図 2 5 B】

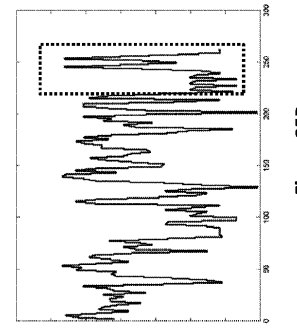


Figure 25B

【図 2 6 A - B】

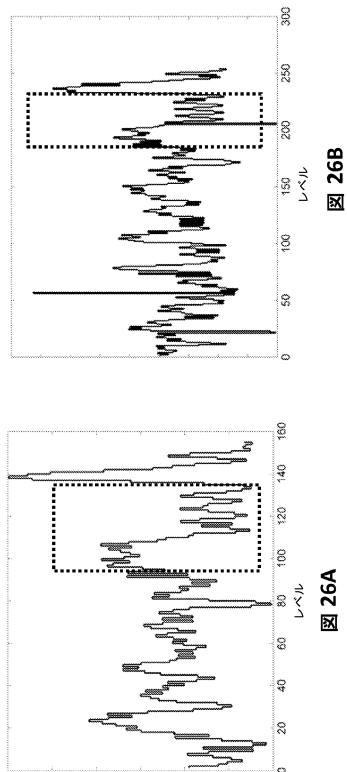


図 26A

図 26B

【図 2 6 C - D】

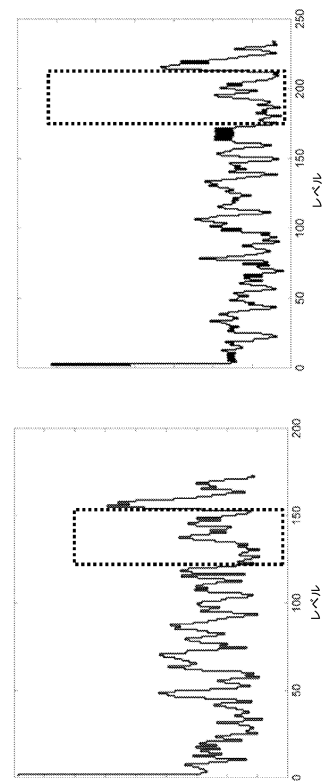


図 26C

図 26D

【配列表】

0006800749000001.app

フロントページの続き

(51)Int.Cl.		F I		
G 0 1 N	21/65	(2006.01)	G 0 1 N	21/65
G 0 1 N	21/64	(2006.01)	G 0 1 N	21/64
C 1 2 N	15/09	(2006.01)	C 1 2 N	15/09

(72)発明者 ジェフリ ジー. マンデル
 アメリカ合衆国 カリフォルニア州 9 2 1 2 2 サンディエゴ イルミナ ウェイ 5 2 0 0

(72)発明者 ケビン エル. ガンダーソン
 アメリカ合衆国 カリフォルニア州 9 2 1 2 2 サンディエゴ イルミナ ウェイ 5 2 0 0

(72)発明者 イアン エム. デリングトン
 アメリカ合衆国 ワシントン州 9 8 1 0 3 シアトル デイトン アベニュー ノース 4 番
 3 6 3 6

(72)発明者 ホセイン モヒマニ
 アメリカ合衆国 カリフォルニア州 9 2 1 2 2 サンディエゴ イルミナ ウェイ 5 2 0 0

審査官 市島 洋介

(56)参考文献 国際公開第 2 0 1 3 / 0 5 7 4 9 5 (WO, A 1)
 Nat. Nanotechnol., 2 0 1 1 年, Vol.6, pp.615-624
 Biophys. J., 2 0 1 2 年, Vol.102, pp.L37-L39

(58)調査した分野(Int.Cl., DB名)
 C 1 2 Q 1 / 0 0 - 3 / 0 0
 C A p l u s / M E D L I N E / E M B A S E / B I O S I S / W P I D S (STN)