

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6920275号
(P6920275)

(45) 発行日 令和3年8月18日(2021.8.18)

(24) 登録日 令和3年7月28日(2021.7.28)

(51) Int. Cl.			F I		
C 1 2 N	15/11	(2006.01)	C 1 2 N	15/11	Z N A Z
C 1 2 P	19/34	(2006.01)	C 1 2 P	19/34	Z
H O 1 L	27/10	(2006.01)	H O 1 L	27/10	4 4 9
C 1 2 Q	1/68	(2018.01)	C 1 2 Q	1/68	
C 1 2 M	1/00	(2006.01)	C 1 2 M	1/00	A

請求項の数 20 (全 44 頁)

(21) 出願番号 特願2018-501913 (P2018-501913)
 (86) (22) 出願日 平成28年7月13日 (2016.7.13)
 (65) 公表番号 特表2018-527900 (P2018-527900A)
 (43) 公表日 平成30年9月27日 (2018.9.27)
 (86) 国際出願番号 PCT/US2016/041981
 (87) 国際公開番号 W02017/011492
 (87) 国際公開日 平成29年1月19日 (2017.1.19)
 審査請求日 令和1年7月16日 (2019.7.16)
 (31) 優先権主張番号 62/191,982
 (32) 優先日 平成27年7月13日 (2015.7.13)
 (33) 優先権主張国・地域又は機関 米国 (US)

(73) 特許権者 507044516
 プレジデント アンド フェローズ オブ
 ハーバード カレッジ
 アメリカ合衆国 マサチューセッツ O 2
 1 3 8, ケンブリッジ, クインシー
 ストリート 1 7
 (74) 代理人 100079049
 弁理士 中島 淳
 (74) 代理人 100084995
 弁理士 加藤 和詳
 (72) 発明者 チャーチ、 ジョージ エム.
 アメリカ合衆国 O 2 4 4 6 マサチュー
 セッツ州 ブルックライン ケント スト
 リート 2 1 8

最終頁に続く

(54) 【発明の名称】 核酸を用いた回収可能な情報記憶のための方法

(57) 【特許請求の範囲】

【請求項 1】

バイナリエンコードされたポリマーを作製する方法であって、
 (i) 第一のモノマーペアにおける第一のモノマー又は第二のモノマーのいずれかのモノマー 1 個又は複数個、又は
 (i i) 第二のモノマーペアにおける第一のモノマー又は第二のモノマーのいずれかのモノマー 1 個又は複数個
 の伸長産物によって、成長ポリマー鎖を繰り返し伸長させること、
 ここで、前記伸長産物が、文字列、画像、ビデオ、又は音声のフォーマットから翻訳されたビットストリームに対応するバイナリ情報ビットを表し、
 前記第一のモノマーペアの前記第一のモノマー及び第二のモノマーがそれぞれ第一のバイナリ情報ビットを表し、
 前記第二のモノマーペアの前記第一のモノマー及び第二のモノマーがそれぞれ第二のバイナリ情報ビットを表す、及び
 前記伸長産物が同じバイナリ情報ビットを表し、且つ直接連続して存在する場合に、所与のモノマーペアの前記第一のモノマーと前記第二のモノマーとを交互に並べる (alternate) こと
 を含み、
 前記バイナリエンコードされたポリマーは、前記文字列、画像、ビデオ、又は音声のフォーマットをコードし、

前記バイナリエンコードされたポリマーは核酸であり、前記第一のモノマーペア及び前記第二のモノマーペアの各々における前記第一のモノマー及び前記第二のモノマーは互いに相補的である、

バイナリエンコードされたポリマーを作製する方法。

【請求項 2】

前記第一のモノマーペアにおける前記第一のモノマー又は前記第二のモノマーが、ヌクレオチドである、請求項 1 に記載の方法。

【請求項 3】

前記第二のモノマーペアにおける前記第一のモノマー又は前記第二のモノマーが、ヌクレオチドである、請求項 1 に記載の方法。

10

【請求項 4】

前記第一のモノマーペアが、アデニン (A)、及びチミン (T) 又はウラシル (U) を含む、請求項 1 に記載の方法。

【請求項 5】

前記第二のモノマーペアが、シトシン (C) 及びグアニン (G) を含む、請求項 1 に記載の方法。

【請求項 6】

前記伸長産物が、酵素及び選択されたモノマーを用いて、前記選択されたモノマーの付加を触媒する条件下で形成される、請求項 1 に記載の方法。

【請求項 7】

20

前記伸長産物が、ポリメラーゼ及び選択されたモノマーを用いて、前記選択されたモノマーの付加を触媒する条件下で形成される、請求項 1 に記載の方法。

【請求項 8】

前記伸長産物が、鋳型非依存的ポリメラーゼ及び選択されたモノマーを用いて、前記選択されたモノマーの付加を触媒する条件下で形成される、請求項 1 に記載の方法。

【請求項 9】

前記成長ポリマー鎖が、基材に結合されている、請求項 1 に記載の方法。

【請求項 10】

工程 (i) 及び工程 (i i) から形成された複数の成長ポリマー鎖を含む、請求項 1 に記載の方法。

30

【請求項 11】

工程 (i) 及び工程 (i i) から形成された複数の成長ポリマー鎖を含み、前記複数の成長ポリマー鎖は、基材に結合されている、請求項 1 に記載の方法。

【請求項 12】

前記第一のモノマーペアにおける前記第一のモノマー又は前記第二のモノマーが、天然ヌクレオチドである、請求項 1 に記載の方法。

【請求項 13】

前記第二のモノマーペアにおける前記第一のモノマー又は前記第二のモノマーが、天然ヌクレオチドである、請求項 1 に記載の方法。

【請求項 14】

40

前記第一のモノマーペア及び第二のモノマーペアが天然ヌクレオチドを含み、前記伸長産物が単一のヌクレオチド又は複数のヌクレオチドを付加するのに十分な条件下で天然ヌクレオチドの付加を触媒することによって作製される、請求項 1 に記載の方法。

【請求項 15】

前記第一のモノマーペア及び第二のモノマーペアが天然ヌクレオチドを含み、前記伸長産物が、基材上の 1 又は複数の位置で、ポリメラーゼ及び選択されたヌクレオチドの投与とヌクレオチド欠乏緩衝液の投与とを交互に行って、前記ヌクレオチドを付加するのに十分な条件下で天然ヌクレオチドの付加を触媒することによって作製される、請求項 1 に記載の方法。

【請求項 16】

50

バイナリエンコードされたポリマーを作製する方法であって、

(i) 第一のモノマーペアにおける第一のモノマー又は第二のモノマーのいずれかのモノマー 1 個又は複数個、又は

(i i) 第二のモノマーペアにおける第一のモノマー又は第二のモノマーのいずれかのモノマー 1 個又は複数個

の伸長産物によって、成長ポリマー鎖を繰り返し伸長させること、

ここで、前記伸長産物が、文字列、画像、ビデオ、又は音声のフォーマットから翻訳されたビットストリームに対応するバイナリ情報ビットを表し、

前記第一のモノマーペアの前記第一のモノマー及び第二のモノマーがそれぞれ第一のバイナリ情報ビットを表し、

前記第二のモノマーペアの前記第一のモノマー及び第二のモノマーがそれぞれ第二のバイナリ情報ビットを表す、及び

前記伸長産物が同じバイナリ情報ビットを表し、且つ直接連続して存在する場合に、所与のモノマーペアの前記第一のモノマーと前記第二のモノマーとを交互に並べることを含み、

前記伸長産物は、前記第一のモノマーペアにおける前記第一のモノマー又は前記第二のモノマーのいずれかの少なくとも 1 つのホモポリマー、又は前記第二のモノマーペアにおける前記第一のモノマー又は前記第二のモノマーのいずれかの少なくとも 1 つのホモポリマーを含み、且つ

前記バイナリエンコードされたポリマーは、前記文字列、画像、ビデオ、又は音声のフォーマットをコードし、

前記バイナリエンコードされたポリマーは核酸であり、前記第一のモノマーペア及び前記第二のモノマーペアの各々における前記第一のモノマー及び前記第二のモノマーは互いに相補的である、

バイナリエンコードされたポリマーを作製する方法。

【請求項 17】

バイナリエンコードされた核酸を、核酸配列から、文字列、画像、ビデオ、又は音声のフォーマットを表すバイナリ情報ビットの列へと翻訳する方法であって、アデニン、及びチミン又はウラシルが、第一のバイナリ情報ビットを表し、シトシン及びグアニンが、第二のバイナリ情報ビットを表し、

前記方法は、

前記核酸配列を読み取ること、及び

各アデニン、又は連続する場合は複数のアデニンに、前記第一のバイナリ情報ビットを割り当てること、

各チミン、又は連続する場合は複数のチミンに、前記第一のバイナリ情報ビットを割り当てること、

各ウラシル、又は連続する場合は複数のウラシルに、前記第一のバイナリ情報ビットを割り当てること、

各シトシン、又は連続する場合は複数のシトシンに、前記第二のバイナリ情報ビットを割り当てること、及び

各グアニン、又は連続する場合は複数のグアニンに、前記第二のバイナリ情報ビットを割り当てること

を含み、

前記核酸配列は、連続する 2 つ以上のアデニン、連続する 2 つ以上のチミン、連続する 2 つ以上のウラシル、連続する 2 つ以上のシトシン、又は連続する 2 つ以上のグアニンのうちの少なくとも 1 つを含む、

方法。

【請求項 18】

文字列、画像、ビデオ、又は音声のフォーマットをエンコード及びデコードする方法であって、

10

20

30

40

50

前記文字列、画像、ビデオ、又は音声のフォーマットをビットストリームの複数のビット列へと変換すること、

前記ビットストリームの前記複数のビット列に対応する核酸配列を、アデニン又はチミンを第一のバイナリ情報ビットに、シトシン又はグアニンを第二のバイナリ情報ビットに割り当てることで設計すること、

ここで、アデニン又はチミンの前記割り当ては、同じバイナリ情報ビットが直接連続して存在する場合は、交互に並べ、

シトシン又はグアニンの前記割り当ては、同じバイナリ情報ビットが直接連続して存在する場合は、交互に並べる、

前記核酸配列を合成すること、

前記合成された核酸配列を保存すること、

前記合成された核酸配列を読み取ること、及び

前記合成された核酸配列を、前記第一のバイナリ情報ビットをアデニン又はチミンに割り当て、前記第二のバイナリ情報ビットをシトシン又はグアニンに割り当てることによって、前記ビットストリームの前記複数のビット列にデコードすることを含む方法。

【請求項 19】

前記合成された核酸配列が、アデニン、チミン、シトシン、又はグアニンの少なくとも1つのホモポリマーを含み、前記合成された核酸配列の前記デコードが、前記第一のバイナリ情報ビットをアデニン又はチミンのホモポリマーに割り当てること、及び前記第二のバイナリ情報ビットをシトシン又はグアニンのホモポリマーに割り当てることを含む、請求項 18 に記載の方法。

【請求項 20】

バイナリエンコードされた核酸を、核酸配列からバイナリ情報ビットの列へと翻訳する方法であって、アデニン、及びチミン又はウラシルが、第一のバイナリ情報ビットを表し、シトシン及びグアニンが、第二のバイナリ情報ビットを表し、

前記方法は、

前記核酸配列を読み取ること、並びに

各アデニン、又は連続する場合は複数のアデニンに、前記第一のバイナリ情報ビットを割り当てること、

各チミン、又は連続する場合は複数のチミンに、前記第一のバイナリ情報ビットを割り当てること、

各ウラシル、又は連続する場合は複数のウラシルに、前記第一のバイナリ情報ビットを割り当てること、

各シトシン、又は連続する場合は複数のシトシンに、前記第二のバイナリ情報ビットを割り当てること、及び

各グアニン、又は連続する場合は複数のグアニンに、前記第二のバイナリ情報ビットを割り当てること

を含む方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、バイナリビット情報としてヌクレオチドなどのモノマーを用い、ヌクレオチドなどのモノマーの配列を用いてポリマーを形成して情報をエンコードする方法全般に関する。この方法により、ヌクレオチドなどのモノマーの配列を用いて、文字列又は画像又は音声などの情報を記憶することができる。

【背景技術】

【0002】

DNA は、情報記憶のための媒体として考えられてきた。Bancroft et al., Science 293, 1763-1765 (2001) を参照されたい。また、Davis, Art Journal 55, 70-74 (1996); G

10

20

30

40

50

ustafsson, Nature 458, 703 (2009)及びGibson, Science 329, 52-56 (2010) ; 米国特許出願第2003/0228611号及び国際公開第2014/014991号も参照されたい。さらに、米国特許出願第2010/0099080号及び国際公開第2014/014991号も参照されたい。

【発明の概要】

【0003】

本開示の実施形態は、モノマーを含むポリマー配列又はポリマー配列群を情報記憶のための媒体として用いる方法に関する。本開示のある実施形態は、ヌクレオチドを含む核酸配列又は核酸配列群を情報記憶のための媒体として用いる方法に関する。情報は、本質的に最も小さく最も正確に複製されるビットである塩基のペア自体にコードされる。一般的なヌクレオチドとしては、アデニン(「A」)、シトシン(「C」)、グアニン(「G」)、及びチミン(「T」)が挙げられる。ある態様では、チミンの代わりに、又はチミンに加えて、ウラシル(「U」)が用いられてもよい。当業者に公知の別の塩基のペアも意図され、6塩基の場合の3塩基のペア、及び12塩基の場合の6塩基のペアなどである。アミノ酸も、本明細書に記載のヌクレオチドと同様に、情報をコードするポリペプチドの作製に用いることができる。

10

【0004】

本開示の態様は、次世代シーケンシング技術及び合成技術を用いた、デジタル情報の強固で大規模な読み取り及び書き込みの方法に関する。ある態様では、文字列及び/又は画像及び/又は音声は、メガビットなどのビットの列へと変換される。ある態様では、文字列及び/又は画像及び/又は音声は、ビットストリームを含むメガビットへと変換される。次に、メガビットは、オリゴヌクレオチドなどのオリゴマーとしてエンコードされる。オリゴヌクレオチド配列などのオリゴマー配列は、設計され、続いて合成される。例えば、オリゴヌクレオチド配列は、設計され、続いて酵素オリゴヌクレオチド合成反応を用いて合成され、この場合、酵素及びヌクレオチドが、適切な反応条件下で基材上の所望の部位に配置され、このヌクレオチドが、支持体に結合された既存のヌクレオチドと共有結合される。オリゴヌクレオチド配列は、エラーブローンポリメラーゼなどのポリメラーゼを、単一のヌクレオチドが付加される可能性が最大化される時間及び条件下で基材上の位置に試薬が局在化される条件下で用いて合成されてもよい。ヌクレオチド付加の反応速度を考慮して2つ以上のヌクレオチドの付加が最小限に抑えられるように、所望の時点で適切な洗浄液が用いられて前記位置から試薬が除去されてもよい。この態様では、試薬は、適切な反応条件下、及び例えばポリメラーゼの存在下で、ヌクレオチドが付加に利用可能である時点を定めて、液体パルスとして基材上のある位置に添加されてもよい。同様に、洗浄液も、液体パルスとしてある位置に添加され、試薬がその位置から除去されてもよい。

20

30

【0005】

ある態様では、オリゴヌクレオチドなどのオリゴマーは、データブロック配列を含む。ある態様では、オリゴヌクレオチドなどのオリゴマーは、ビットストリーム中におけるデータブロックの位置を指定するアドレス配列(バーコード配列など)を含む。ある態様では、オリゴヌクレオチドは、オリゴヌクレオチドの両端に、増幅及びシーケンシングのための隣接共通配列を含む。ある態様では、オリゴヌクレオチドは、データブロック配列、ビットストリーム中におけるデータブロックの位置を指定するアドレス配列(バーコード配列など)、並びにオリゴヌクレオチドの両端に位置する増幅及びシーケンシングのための隣接共通配列のうちの1つ又は複数又は全てを含む。

40

【0006】

本開示のある態様では、1塩基あたり1ビットがコードされる。この態様では、単一のメッセージが複数の様式で、すなわち、0はA又はCで、数字1はG又はTでコードされてもよい。0はA若しくはG、数字1はC若しくはT、又は0はA若しくはT、数字1はG若しくはCなどの他の組み合わせも想定される。本明細書で考察されるように、他の組み合わせも意図される。ある態様では、ビットストリームは、アドレスされたデータプロ

50

ックに分けられる。この態様では、記録された情報を表すデータブロックのライブラリが作り出される。この方法は、記録された情報を全体として表す単一の長い核酸配列、又は比較的長い核酸配列を必要としない。

【 0 0 0 7 】

ある態様では、個々のオリゴヌクレオチドの多数のコピーが、ハイスループット次世代技術を用いて合成され、記録され、シークエンシングされる。合成及びシークエンシングの際のエラーは、同時に発生することがほとんどないことから、各分子コピーが他のコピーにおけるエラーを修正する。

【 0 0 0 8 】

ある態様では、オリゴヌクレオチドは、当業者に公知の方法を用いてシークエンシングされる。ヌクレオチド配列をバイナリ情報ビットに翻訳する目的で、エラープローンポリメラーゼの使用によって生じ得る特定のヌクレオチドのホモポリマーの並び（すなわち、連続する同じヌクレオチド又は他のモノマーの配列）は、バイナリ情報ビット、すなわち 0 又は 1 を割り当てるために単一ヌクレオチドとして処理される。ある他の態様では、ビットストリーム中の隣り合った 0 又は隣り合った 1 を区別するために、ヌクレオチド A 及び T などの 0 を表す 2 つのモノマーが、例えば、オリゴヌクレオチド配列の設計において交互に並べられる。このことによって、オリゴヌクレオチド合成の過程でホモポリマーの並びが生じ得る場合に、隣り合った 0 又は 1 を、別個のバイナリ情報ビットとして区別することが可能となる。例えば、ビットストリーム中で 2 つの 0 が互いに隣り合っている、すなわち、- 0 0 - の場合、対応するヌクレオチド配列は、- A T - 又は - T A - が選択される。この方法により、設計されたオリゴヌクレオチド配列の合成の過程で、- A A A T T T - などのホモポリマーの並びが得られる場合、ホモポリマーの並びは、単一のヌクレオチドとして解釈され、- 0 0 - に対応する - A T - として読み取られる。したがって、バイナリビットストリームを核酸配列にエンコードし、この核酸配列をデコードしてバイナリビットストリームに戻すことに関する本開示の方法により、不定な長さのヌクレオチドホモポリマーの並びが許容されると同時に、核酸をバイナリビットストリームに正確にデコードすることができる。

【 0 0 0 9 】

ある態様では、特定のフォーマットの情報を、各々が対応するビットバーコードを有するビットストリームの複数のビット列へと変換すること、前記複数のビット列を 1 塩基あたり 1 ビットのエンコードを用いて複数の対応するオリゴヌクレオチド配列へと変換すること、前記複数の対応するオリゴヌクレオチド配列を、成長オリゴヌクレオチド鎖にヌクレオチドが付加されるように酵素試薬及び洗浄液をパルスし、同期させることによって合成すること、並びに前記合成された複数の対応するオリゴヌクレオチド配列を保存することを含む、ビットを表すものとしてヌクレオチドを用いて情報を記憶する方法が提供される。ある態様では、オリゴヌクレオチド配列は、データブロック配列、ビットストリーム中におけるデータブロックの位置を指定するアドレス配列、又はオリゴヌクレオチドの両端に位置する増幅及びシークエンシングのための隣接共通配列のうちの 1 つ又は複数又は全てを含む。ある態様では、エラープローンポリメラーゼを用いて、複数の対応するオリゴヌクレオチド配列が合成されてもよい。

【 0 0 1 0 】

ある態様では、特定のフォーマットの情報のビット列をコードする複数のオリゴヌクレオチド配列を増幅すること、前記増幅されたオリゴヌクレオチド配列をシークエンシングすること、ホモポリマーの並びを単一のヌクレオチドとして解釈することによって前記オリゴヌクレオチド配列をビット列へと変換すること、及び前記ビット列を特定のフォーマットの情報へと変換することを含む、特定のフォーマットの情報を、前記特定のフォーマットの情報のビット列をコードする複数の合成されたオリゴヌクレオチド配列から回収する方法が提供される。ある態様では、オリゴヌクレオチド配列は、データブロック配列、ビットストリーム中におけるデータブロックの位置を指定するアドレス配列、又はオリゴヌクレオチドの両端に位置する増幅及びシークエンシングのための隣接共通配列のうちの

10

20

30

40

50

1つ又は複数又は全てを含む。ヌクレオチド配列をバイナリ情報ビットに翻訳する目的で、エラーブローンポリメラーゼの使用によって生じ得る特定のヌクレオチドのホモポリマーの並びは、バイナリ情報ビット、すなわち0又は1を割り当てるために単一ヌクレオチドとして処理される。

【0011】

ある態様では、特定のフォーマットの情報のビット列をコードする複数のオリゴヌクレオチド配列を増幅すること、前記増幅されたオリゴヌクレオチド配列をシーケンシングすること、ホモポリマーの並びを単一のヌクレオチドとして解釈することによって前記オリゴヌクレオチド配列をビット列へと変換すること、前記ビット列を特定のフォーマットの情報へと変換すること、及び前記特定のフォーマットの情報を可視化するか又は前記特定のフォーマットの情報を音声化することを含む、特定のフォーマットの情報に、前記特定のフォーマットの情報のビット列をコードする複数の合成されたオリゴヌクレオチド配列からアクセスする方法が提供される。ある態様では、オリゴヌクレオチド配列は、データブロック配列、ビットストリーム中におけるデータブロックの位置を指定するアドレス配列、又はオリゴヌクレオチドの両端に位置する増幅及びシーケンシングのための隣接共通配列のうちの1つ又は複数又は全てを含む。ヌクレオチド配列をバイナリ情報ビットに翻訳する目的で、エラーブローンポリメラーゼの使用によって生じ得る特定のヌクレオチドのホモポリマーの並びは、バイナリ情報ビット、すなわち0又は1を割り当てるために単一ヌクレオチドとして処理される。

10

【0012】

ある態様では、特定のフォーマットの情報をビットストリームへと変換すること、ビット列を対応するオリゴヌクレオチド配列へとエンコードすること、前記オリゴヌクレオチド配列を、例えば、エラーブローンポリメラーゼなどの酵素を用い、2つ以上のヌクレオチドの結合が最小限に抑えられるように試薬及び洗浄液をパルスし、同期させることによって合成すること、前記オリゴヌクレオチド配列をシーケンシングすること、前記オリゴヌクレオチド配列を、ホモポリマーの並びを単一のヌクレオチドとして解釈することによってビット列へとデコードすること、前記ビット列をビットストリームへとアセンブルすること、及び前記ビットストリームを前記特定のフォーマットの情報へと変換することを含む、ヌクレオチドを用いて情報を記憶する方法が提供される。ある態様では、オリゴヌクレオチド配列は、データブロック配列、ビットストリーム中におけるデータブロックの位置を指定するアドレス配列、又はオリゴヌクレオチドの両端に位置する増幅及びシーケンシングのための隣接共通配列のうちの1つ又は複数又は全てを含む。ヌクレオチド配列をバイナリ情報ビットに翻訳する目的で、エラーブローンポリメラーゼの使用によって生じ得る特定のヌクレオチドのホモポリマーの並びは、バイナリ情報ビット、すなわち0又は1を割り当てるために単一ヌクレオチドとして処理される。

20

30

【0013】

第一のフォーマットの情報を第一のビットストリームへと変換すること、第一のビット列を対応するオリゴヌクレオチド配列へとエンコードすること、前記オリゴヌクレオチド配列を、例えば、エラーブローンポリメラーゼを用い、2つ以上のヌクレオチドの結合が最小限に抑えられるように試薬及び洗浄液をパルスし、同期させることによって合成すること、前記オリゴヌクレオチド配列をシーケンシングすること、前記オリゴヌクレオチド配列を、ホモポリマーの並びを単一のヌクレオチドとして解釈することによって第二のビット列へとデコードすること、前記第二のビット列を第二のビットストリームへとアセンブルすること、及び前記第二のビットストリームを第二のフォーマットの情報へと変換することを含む、ヌクレオチドを用いて情報を記憶する方法が提供される。ある態様では、オリゴヌクレオチド配列は、データブロック配列、ビットストリーム中におけるデータブロックの位置を指定するアドレス配列、又はオリゴヌクレオチドの両端に位置する増幅及びシーケンシングのための隣接共通配列のうちの1つ又は複数又は全てを含む。ヌクレオチド配列をバイナリ情報ビットに翻訳する目的で、エラーブローンポリメラーゼの使用によって生じ得る特定のヌクレオチドのホモポリマーの並びは、バイナリ情報ビット、

40

50

すなわち0又は1を割り当てるために単一ヌクレオチドとして処理される。

【0014】

本開示の実施形態は、ヌクレオチドなどの分子を情報のバイナリビットとして用いることに関する。ヌクレオチドは、0又は1などのバイナリ状態を表してよく、0又は1などのバイナリ状態の列を表すヌクレオチドの配列は、文字列フォーマット、画像フォーマット、ビデオフォーマット、又は音声フォーマットを表してもよい。この方法により、書面資料、画像、音声コンポーネント若しくは音声記録を含むビデオ、又は他のいかなる表現媒体も、ビットを表すものとして核酸を用いることによって記憶することができる。ある態様では、記憶される情報は、例えばコンピュータ及び適切なソフトウェアを用いて、情報を表す0及び1の列であるASCIIコードによるものなどのバイナリビットへと変換される。本技術分野で公知であるように、記憶される情報が、情報の他のコードされたビットに変換されてもよいことを理解されたい。次に、コンピュータ及び適切なソフトウェアを用いることなどによって、0及び1などの情報のコードされたビットの列を表すヌクレオチドの列が特定される。次に、ヌクレオチドの列は、合成され、記憶媒体に記憶される。情報がアクセスされる場合、前クレオチドの列が特定され、次に、コンピュータ及び適切なソフトウェアを用いることなどによって0及び1の列へと翻訳されて、続いてこれが、例えばコンピュータ及び適切なソフトウェアを用いて、情報に翻訳される。このように、本開示は、完全に若しくは部分的に一本鎖、二本鎖、又は多重鎖であるかどうかに関わらず、情報記憶媒体として核酸を使用することに関する。ある態様では、核酸は、規則的であっても、又はランダムであってもよい状態で、支持基材に含まれる。

10

20

【0015】

ある態様では、修飾の有無にかかわらず、エラーブローン鑄型依存的ポリメラーゼを含むがこれらに限定されないポリメラーゼを用いて、記憶される情報を表すものであるバイナリビットを表す所望のヌクレオチド配列を有するヌクレオチドポリマーを作製することができる。修飾の有無にかかわらず、鑄型非依存的ポリメラーゼを用いて核酸をdenovo作製することもできる。ある態様では、A、T/U、C、又はGなどの通常のヌクレオチドが用いられる。ある態様では、鎖停止(chain terminating)部分を有しないヌクレオチドが用いられる。ある態様では、鎖停止ヌクレオチドは、ヌクレオチドポリマーを作製する方法では用いられない。この態様では、鑄型非依存的ポリメラーゼを用いて、核酸配列を作製してもよい。そのような鑄型非依存的ポリメラーゼは、ホモポリマーをもたらず2つ以上のヌクレオチドの付加を引き起こし得るエラーブローンであってもよい。本開示のある態様では、ホモポリマーの並びは、このホモポリマーの並びが情報のどのバイナリビットを表すかを特定する目的で、単一のヌクレオチドとして解釈される。光活性化センサーなどのセンサー、リガンドによって活性化される代謝産物又は化学物質が、そのようなポリメラーゼと共に用いられてもよい。

30

【0016】

核酸ポリマーは、当業者に公知の方法を用いてシーケンシングすることができる。核酸配列が決定されると、核酸配列を翻訳して、バイナリビット、すなわち0及び1の列とすることができ、次にこれを、バイナリビットの列によって表される情報に翻訳することができる。

40

【0017】

本開示の特定の実施形態のさらなる特徴及び利点は、実施形態の以下の記述及びその図面において、並びに請求項からより十分に明らかとなるであろう。

【0018】

前述した本発明の特徴及びその他の特徴、並びに利点は、添付の図面と共に、例示的实施形態についての以下の詳細な説明から、より十分に理解されるであろう。

【図面の簡単な説明】

【0019】

【図1】オリゴヌクレオチドが合成される所定の領域を有する基材を横切るように流動するヌクレオチドのパルスを示す模式図である。

50

【図2】オリゴヌクレオチドが合成される所定の領域を有する基材を横切るように流動するヌクレオチドのパルスを示す模式図である。

【図3】オリゴヌクレオチドが合成される所定の領域を有する基材を横切るように流動するヌクレオチドのパルスを示す模式図である。

【発明を実施するための形態】

【0020】

本開示は、オリゴマーを用いて情報を記憶する方法に関する。そのようなオリゴマーは、例えばバイナリビットを表し得るモノマーから形成することができる。例示的なモノマーとしては、ヌクレオチドが挙げられる。例示的なオリゴマーとしては、オリゴヌクレオチドが挙げられる。ある態様では、ビット列がヌクレオチドなどのモノマーの配列へと変換される情報をエンコードする方法が提供され、前記モノマーの配列は、オリゴヌクレオチドなどのポリマーである。ある態様では、本明細書に記載される方法は、オリゴヌクレオチドの合成に用いられてよく、又は市販の又は公知のポリマー若しくは核酸の合成方法が用いられてもよい。ある態様では、市販の又は公知の核酸増幅方法が用いられる。ある態様では、市販の又は公知の核酸シーケンシング方法が用いられる。ある態様では、市販の又は公知のポリマー中のモノマーを識別する方法が用いられる。

10

【0021】

ある態様では、文字列及び/又は画像を有するhtml文書などのhtmlフォーマットの情報などの文字列、画像、ビデオ、又は音声のフォーマットなど、特定のフォーマットの情報の一部又は部分が、例えばコンピュータ及び適切なソフトウェアを用いて、ビット、すなわち0及び1へと変換され、ビットバーコードが追加されて、ビット列、すなわち一般的に理解されるように0及び1の列が形成される。ビットへと変換することができる他のフォーマットの情報も、当業者に公知である。ある態様では、ビットへと変換されるhtmlフォーマットの情報の部分は、バイト部分と呼ばれることがある。ビットバーコードは、htmlフォーマットの情報全体の中でのエンコードされたビットの位置を決定することができる。次に、ビット列が、コンピュータ及び適切なソフトウェアなどにより、1塩基あたり1ビットのエンコーディング(A又はC=0; T/U又はG=1)を用いて、ヌクレオチドの設計された配列、すなわち、オリゴヌクレオチド又はDNA又はRNAへと変換され(エンコードされ)、対応するエンコードされたオリゴヌクレオチド配列が形成される、すなわち、オリゴヌクレオチド配列は、ビット列に対応するか、又はビット列をコードする。

20

30

【0022】

htmlフォーマットの情報の一部又は全体に対応する複数のビット列が作製される。したがって、複数の対応するエンコードされたオリゴヌクレオチド配列が作製され、これらは併せてライブラリと呼ばれることがある。エンコードされたオリゴヌクレオチド配列のライブラリは、htmlフォーマットの情報を表す。ある態様では、オリゴヌクレオチドは、ビットデータブロック部分、ビットストリーム中におけるデータブロックの位置を指定するビットアドレス部分、並びに増幅及びシーケンシングのための隣接共通配列を含む。例えば、159ヌクレオチドのオリゴヌクレオチドは、96ビットのデータブロック(96nt)、19ビットのアドレス(19nt)、及びオリゴヌクレオチドに隣接する22ビットの共通配列(22nt)を含み得る。

40

【0023】

ある例示的な態様では、エンコードされたオリゴヌクレオチド配列は、次に、鋳型非依存のエラープロンポリメラーゼなどのエラープロンポリメラーゼ、及び未修飾であってもよい通常の又は天然の核酸を用いて合成される。この態様では、イニシエータ配列又はプライマーが、アドレス可能アレイにおける場合などの既知であってもよいか、又はランダムであってもよい様々な位置で、二酸化ケイ素基材などの基材と結合される。選択されたヌクレオチド、鋳型非依存のポリメラーゼ、及びポリメラーゼの酵素活性に必要な他の試薬を少なくとも含む試薬が、イニシエータ配列が配置されている基材の1又は複数の位置に、ポリメラーゼが1個又は2個以上又は複数個のヌクレオチドをイニシエータ配列

50

に付加してイニシエータ配列を伸長させる条件下で適用される。ある態様では、ヌクレオチド（「dNTP」）は、ポリメラーゼの重合速度（又はスイッチング速度）を考慮して、既知の時間的及び空間的な様式、若しくは幅、若しくは条件の周期的な適用又は波（wave）によって適用されるか、又は流される。この例示的方法では、dNTPの酵素的付加の確率を1つのdNTPに限定又は低減するのに十分であるように、すなわち、反応速度を考慮に入れた選択された反応条件を用いて1つのdNTPが付加されるように、反応条件が選択されるため、保護基又は可逆的ターミネータは、dNTPと共に用いられない。しかし、特定の実施形態において保護基又は可逆的ターミネータを有するヌクレオチドが用いられてもよいことは理解されたい。保護基又は可逆的ターミネータを有するヌクレオチドは、当業者に公知である。他の実施形態では、反応条件が合う場合、通常の又は天然のヌクレオチドが鑄型非依存的エラープロードポリメラーゼと共に用いられた場合に、2つ以上のdNTPが付加されてホモポリマーの並びが形成され得る。しかし、本明細書に記載される方法のシーケンシング工程の過程で、各ホモポリマーの並びは、単一のdNTPを表すとして解釈される。このように、本明細書に記載される記録及び読み取りの方法によりホモポリマーの並びがもたらされ、この合成方法では、エラープロードであってもよい鑄型非依存的ポリメラーゼを用いた場合に必要であり得るような、単一のdNTPのみを添加するという必要がない。

【0024】

さらに、本開示は、同じバイナリビットが並んでいる核酸を作製する際に、単一のバイナリビットを代表するものである2つの異なるモノマーを交互に並べる。例えば、2つの0が並んでいる「00」を意図する場合、対応する核酸は、1つ目の0を表す「A」及び2つ目の0を表す「C」を有することになり、それによって、エラープロードポリメラーゼに起因して複数のA又は複数のCのホモポリマーの並びが存在する場合であっても、1つ目の「0」及び2つ目の「0」を、互いに別個のビットとして区別することができる。4つのヌクレオチドの場合、1つのペアのヌクレオチドが第一のバイナリビットを表し、残りのヌクレオチドペアが第二のバイナリビットを表すことは理解されたい。したがって、A及びCが第一のバイナリビットを表してよく、T及びGが第二のバイナリビットを表す。あるいは、A及びGが第一のバイナリビットを表してよく、T及びCが第二のバイナリビットを表す。あるいは、A及びTが第一のバイナリビットを表してよく、C及びCが第二のバイナリビットを表す。あるいは、4つのヌクレオチドが、「トリナリ（ternary）」データシステムに用いられてもよく、3つのヌクレオチドが、0、1、及び2を表し、残りのヌクレオチドが、ホモポリマーの並びを区別するために、連続する3つのヌクレオチドの次に用いられる。例えば、4つのヌクレオチドでトリナリデータシステムを用いる場合、A、C、Tがそれぞれ0、1、及び2を表してよく、ホモポリマーの並びが存在し得る場合、Gが、例えばA及び連続するAを区別するために、連続するA、C、又はTの次に用いられてもよい。あるいは、3つのヌクレオチドがバイナリシステムを表すために用いられてもよく、この場合、2つのヌクレオチドがそれぞれ0及び1を表してよく、ホモポリマーの並びを区別するために、第三のヌクレオチドが連続するヌクレオチドの次に用いられてもよい。

【0025】

ポリメラーゼ活性は、単一のdNTPを超えるdNTPの付加が最小限に抑えられるように、反応条件として光化学的調整又は電気化学的調整を用いて調整されてもよい。次に、洗浄液が1又は複数の位置に適用されて、試薬が除去される。試薬及び洗浄液を適用する工程は、所望の核酸が作製されるまで繰り返される。ある態様では、試薬は、基材上の1つ若しくは2つ以上若しくは複数の位置に、順に若しくは並行して添加されてもよく、又は試薬は、試薬を支持体の表面を横切るように流動させることなどにより、支持体の表面全体と接触させてもよい。ある態様では、反応条件は、ポリメラーゼがイニシエータ配列又は成長オリゴヌクレオチドの末端に2つ以上のヌクレオチドを結合させる能力が制限されるように、例えば、反応速度又はポリメラーゼの活性に基づいて決定される。

【0026】

さらに、ある実施形態では、ポリメラーゼは、光に基づく方法のために光感受性となるように修飾されてもよい。この態様では、単一のヌクレオチドのみを付加するようにポリメラーゼを調節するために、光が調整される。光は、ポリメラーゼ、ヌクレオチド、及び適切な試薬、並びに反応条件が存在する基材の個々の位置又はピクセルに対して照射される。この方法により、ヌクレオチドは、ポリメラーゼが光によって活性化された場合に、イニシエータ配列又は既存のヌクレオチドに付加される。

【0027】

ある態様では、エラーブローンポリメラーゼを用いて、情報を表すオリゴヌクレオチドを作製することができる。そのようなオリゴヌクレオチドは、エラーブローンプロセスを用いて作製される。エラーブローンプロセスは、ホモポリマーの並びを含む場合があり、すなわち、単一のヌクレオチドを結合する代わりに、2つ以上のヌクレオチドを連続して結合する場合がある。本開示によると、ホモポリマーの並びは、バイナリビット情報に翻訳するために核酸をシーケンシングする際に、ホモポリマーの単一のヌクレオチドであるとして処理される。

【0028】

ある実施形態では、鑄型依存的エラーブローンポリメラーゼが用いられてもよい。ある実施形態では、エラーブローンとなり得る鑄型依存的ポリメラーゼが用いられてもよい。ある実施形態では、鑄型非依存的RNAポリメラーゼが用いられてもよい。

【0029】

さらに、核酸配列の有用な作製方法は、その内容が全体として参照により本明細書に引用される"Large-scale de novo DNA synthesis: technologies and applications," by Sriram Kosuri and George M. Church, Nature Methods, May, 2014, Vol. 11, No. 5, pp. 499-507に開示されている。

【0030】

ある態様では、基材上で局所的にpHに影響を与える、又はpHを変化させる、又はpHを発生させることによって、記憶される情報をコードする核酸配列を作製するために用いることができる例示的システムは、CustomArray社から市販されているCustomArrayシステムである。基材上の特定の位置でpHに影響を与える、又はpHを変化させる、又はpHを発生させるために、他の方法が用いられてもよいことは理解されたい。CustomArrayシステムは、pH勾配を用い、半導体に基づいた電気化学的合成プロセスを用いて所望のオリゴヌクレオチドマイクロアレイを合成する。各オリゴヌクレオチドプローブは、合成装置のコンピュータによって独立制御される白金電極を介して合成される。本明細書に記載の方法によると、基材上の特定の所望の位置のpH感受性ポリメラーゼを活性化させて、特定の所望の位置の水性媒体中に存在するヌクレオチドを付加するpH勾配が作り出される。この態様では、ポリメラーゼに単一のヌクレオチドのみを付加させるようにpHが調整される。本明細書に記載の態様では、オリゴヌクレオチドの形成方法において、CustomArrayシステムなどのシステムを用いて、pH依存的ポリメラーゼ、ヌクレオチド、及び他の適切な試薬が水性媒体中に存在する基材上で局所的にpHに影響を与え、又はpHを変化させ、又はpHを発生させて、イニシエータ配列又は既存のヌクレオチド又はオリゴヌクレオチドにヌクレオチドを付加させることができる。本明細書に記載の例示的な方法は、オリゴヌクレオチドの形成方法において、水性溶媒及びpHを用いてTdTなどの鑄型非依存的ポリメラーゼなどのポリメラーゼの活性を調整し、基材上の所望の位置で、既存のイニシエータ配列、既存のヌクレオチド、又は既存のオリゴヌクレオチドにヌクレオチドを付加する。

【0031】

ある態様では、入り口部及び出口部を有するマイクロ流体チャンネル又は複数のマイクロ流体チャンネルなどのフローセル又は他のチャンネルを用いて、ポリメラーゼ、ヌクレオチド、及び他の適切な試薬などの試薬、並びに洗浄液を含む流体が、反応チャンバー内などのフローセル内の基材上の特定の位置に送達される。ある態様では、反応条件は、基材上の位置を選択的に活性化及び不活性化するように選択される。この方法により、基材又はア

10

20

30

40

50

レイ上のグリッドポイントなどの所望の位置に反応条件を与えて、イニシエータ配列、既存のヌクレオチド、既存のオリゴヌクレオチドへのヌクレオチドの共有結合を促進することができ、且つ、反応条件を与えることにより、同じ位置に他のヌクレオチドがさらに結合されることを防止することができる。次に、その所望の位置でオリゴヌクレオチドを作製する方法では、既存のヌクレオチドへのヌクレオチドの共有結合を促進する反応条件を同じ位置に与えることができる。

【 0 0 3 2 】

ある態様では、支持体はその表面上に所望の成長核酸を有した時点で、第二の基材が、この核酸を表面上に有する基材の表面に追加されてよく、第二の基材上に核酸の第二の層が作製されてもよい。また、このプロセスは、他の基材に対して繰り返されて、核酸を有する多くの又は複数の基材層を有する層状の基材が作製されてもよい。この態様では、記録媒体の各層のオリゴヌクレオチドを用いて情報を記憶する記録媒体を作製することができる。

10

【 0 0 3 3 】

合成されたオリゴヌクレオチドは、次に、当業者に公知の方法を用いて増幅され、オリゴヌクレオチドのライブラリが形成される。オリゴヌクレオチドのライブラリは、次に、次世代シーケンシング法などの当業者に公知の方法を用いてシーケンシングされる。シーケンシングされたオリゴヌクレオチドは、次に、例えばhtmlフォーマットの情報に対応するビット列へと変換される。ビット列は、当業者に公知の方法を用いて、前記フォーマットの情報へと変換することができる。前記フォーマットの情報は、当業者に公知の方法及びデバイスを用いて、可視化、又は表示、又は音声フォーマットの場合は再生することができる。

20

【 0 0 3 4 】

本明細書で用いられる核酸化学、生化学、遺伝学、及び分子生物学の用語並びに記号は、この分野の標準的な取り決め及び教科書、例えば、Komberg and Baker, DNA Replication, Second Edition (W.H. Freeman, New York, 1992); Lehninger, Biochemistry, Second Edition (Worth Publishers, New York, 1975); Strachan and Read, Human Molecular Genetics, Second Edition (Wiley-Liss, New York, 1999); Eckstein, editor, Oligonucleotides and Analogs: A Practical Approach (Oxford University Press, New York, 1991); Gait, editor, Oligonucleotide Synthesis: A Practical Approach (IRL Press, Oxford, 1984)などの用語及び記号に従う。

30

【 0 0 3 5 】

ビット

本明細書で用いられる場合、「ビット」の用語は、その当業者にとって一般的な意味に従って理解されるべきである。「ビット (b i t)」の用語は、「バイナリ数値 (b i n a r y d i g i t)」の縮約形であり得、電子計算及び通信における情報の基本容量を意味し得る。「ビット」は、1又は0 (1又はゼロ) などの第一の状態又は第二の状態のいずれかのみを表す。様々なシステムにおいて、2状態デバイスを用いて表現を実行することができる。

【 0 0 3 6 】

核酸及びヌクレオチド

本明細書で用いられる場合、「核酸分子」、「核酸配列」、「核酸断片」、及び「オリゴマー」の用語は、交換可能に用いられ、これらに限定されないが、様々な長さであってもよいポリマー形態のヌクレオチドを含むことを意図しており、デオキシリボヌクレオチド若しくはリボヌクレオチド、又はこれらの類似体を含む。

40

【 0 0 3 7 】

一般的に、「核酸分子」、「核酸配列」、「核酸断片」、「オリゴヌクレオチド」、及び「ポリヌクレオチド」の用語は、交換可能に用いられ、これらに限定されないが、様々な長さであってもよいポリマー形態のヌクレオチドを含むことを意図しており、デオキシリボヌクレオチド (D N A) 若しくはリボヌクレオチド (R N A)、又はこれらの類似体

50

を含む。オリゴヌクレオチドは、典型的には、4つのヌクレオチド塩基、である、アデニン(A)；シトシン(C)；グアニン(G)；及びチミン(T)（ポリヌクレオチドがRNAである場合は、チミン(T)に代わってウラシル(U)）の特異的配列から構成される。ある態様では、デオキシヌクレオチド(dATP、dCTP、dGTP、dTTPなどのdNTP)が用いられてもよい。ある態様では、リボヌクレオチド三リン酸(rNTP)が用いられてもよい。ある態様では、リボヌクレオチド二リン酸(rNDP)が用いられてもよい。

【0038】

用語「オリゴヌクレオチド配列」は、ポリヌクレオチド分子のアルファベット表示であるか、または、この用語は、ポリヌクレオチド分子自体にも適用可能である。このアルファベット表示は、中央処理装置を有するコンピュータのデータベースに入力可能であり、機能ゲノム科学及び相同性検索などの生物情報学用途に用いることができる。オリゴヌクレオチドは、必要に応じて、非標準的ヌクレオチド、ヌクレオチド類似体、及び/又は修飾ヌクレオチドのうちの1つ又は複数を含んでいてもよい。本開示は、いずれのデオキシリボヌクレオチド又はリボヌクレオチド、及びこれらの塩基のメチル化形態、ヒドロキシメチル化形態、又はグルコシル化形態などの化学的変異体も意図する。ある態様では、天然ヌクレオチドが、核酸の作製方法に用いられる。天然ヌクレオチドは、鎖停止部分を有しない。別の態様では、本明細書に記載の核酸の作製方法では、当業者に公知である可逆的ターミネータなどの停止核酸を用いないか、又は停止核酸が存在しない。この方法は、鎖停止核酸の非存在下で行われるか、又は核酸が鎖停止核酸以外である。

【0039】

修飾されたヌクレオチドの例としては、これらに限定されないが、ジアミノプリン、5²T、5^{フルオロ}ウラシル、5^{プロモ}ウラシル、5^{クロロ}ウラシル、5^{ヨード}ウラシル、ヒポキサンチン、キサンチン、4^{アセチル}シトシン、5^(カルボキシヒドロキシメチル)ウラシル、5^{カルボキシメチルアミノメチル} 2^{チオウリジン}、5^{カルボキシメチルアミノメチルウラシル}、ジヒドロウラシル、ベータ^Dガラクトシルキユーオシン、イノシン、N⁶イソペンテニルアデニン、1^{メチル}グアニン、1^{メチル}イノシン、2²ジメチルグアニン、2^{メチル}アデニン、2^{メチル}グアニン、3^{メチル}シトシン、5^{メチル}シトシン、N⁶アデニン、7^{メチル}グアニン、5^{メチルアミノメチルウラシル}、5^{メトキシアミノメチル} 2^{チオウラシル}、ベータ^Dマンノシルキユーオシン、5^{メトキシカルボキシメチルウラシル}、5^{メトキシウラシル}、2^{メチルチオ} D⁴⁶イソペンテニルアデニン、ウラシル 5^{オキシ酢酸(v)}、ワイブトキシソシン(wybutoxosine)、シュードウラシル、キユーオシン、2^{チオ}シトシン、5^{メチル} 2^{チオウラシル}、2^{チオウラシル}、4^{チオウラシル}、5^{メチルウラシル}、ウラシル 5^{オキシ酢酸メチルエステル}、ウラシル 5^{オキシ酢酸(v)}、5^{メチル} 2^{チオウラシル}、3^(3アミノ3N2カルボキシプロピル)ウラシル、(acp3)w、及び2⁶ジアミノプリンなどが挙げられる。核酸分子はまた、塩基部分において(例えば、相補的ヌクレオチドとの水素結合の形成に通常は利用可能である1つ又は複数の原子において、及び/又は相補的ヌクレオチドとの水素結合を通常は形成することができない1つ又は複数の原子において)、糖部分において、又はリン酸骨格において修飾されていてもよい。核酸分子はまた、N^{ヒドロキシ}スクシンイミドエステル(NHS)などのアミン反応性部分の共有結合を可能にする、アミノアリル dUTP(aa dUTP)及びアミノヘキシルアクリルアミド dCTP(aha dCTP)などのアミン修飾された基を含んでいてもよい。

【0040】

本開示のオリゴヌクレオチドにおける標準的なDNA塩基のペア又はRNA塩基のペアの代替物により、1mm³あたりのピット密度がより高くなり、安全性(天然毒の偶発的又は意図的な合成に対する耐性)が高くなり、光プログラム化ポリメラーゼによる識別がより簡便になり、又は二次構造がより低くなる。de novo合成及び/又は増幅合成において天然ポリメラーゼ及び変異型ポリメラーゼと適合するこのような代替塩基のペア

10

20

30

40

50

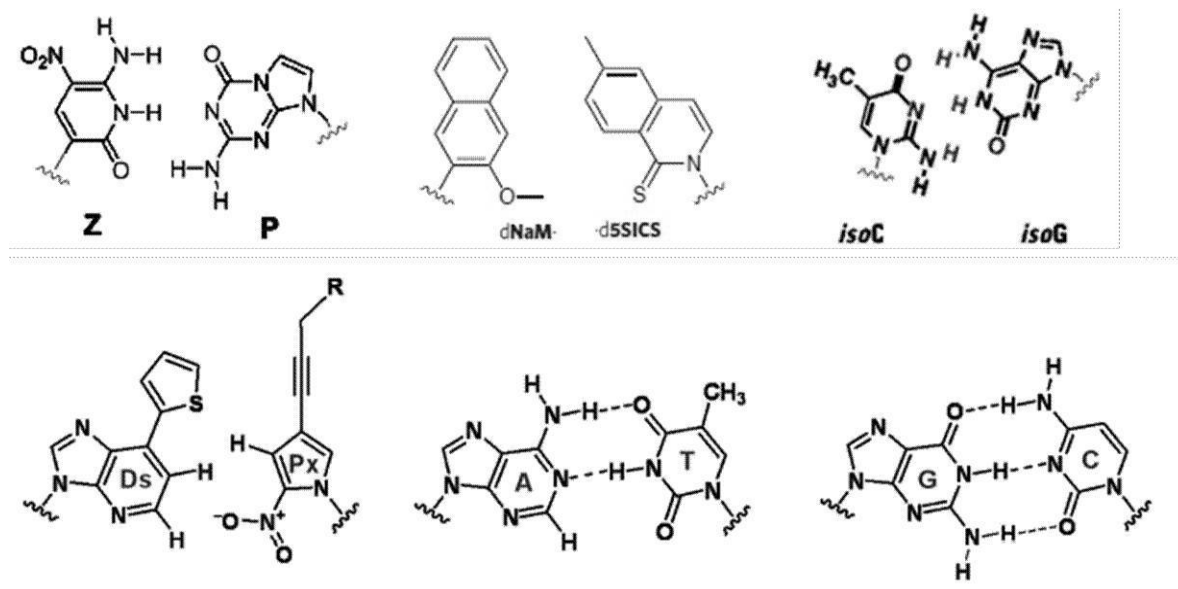
は、Betz K, Malyshev DA, Lavergne T, Welte W, Diederichs K, Dwyer TJ, Ordoukhani an P, Romesberg FE, Marx A (2012) KlenTaq polymerase replicates unnatural base pairs by inducing a Watson-Crick geometry, *Nature Chem. Biol.* 8:612-614 ; See YJ, Malyshev DA, Lavergne T, Ordoukhanian P, Romesberg FE. *J Am Chem Soc.* 2011 Dec 14;133(49):19878-88, Site-specific labeling of DNA and RNA using an efficiently replicated and transcribed class of unnatural base pairs ; Switzer CY, Moroney SE , Benner SA. (1993) *Biochemistry.* 32(39):10489-96. Enzymatic recognition of the base pair between isocytidine and isoguanosine ; Yamashige R, Kimoto M, Takezawa Y, Sato A, Mitsui T, Yokoyama S, Hirao I. *Nucleic Acids Res.* 2012 Mar;40(6):2793-806. Highly specific unnatural base pair systems as a third base pair for PCR amplification ; 及び Yang Z, Chen F, Alvarado JB, Benner SA. *J Am Chem Soc.* 2011 Sep 28;133(38):15105-12, Amplification, mutation, and sequencing of a six-letter synthetic genetic systemに記載されている。その全体が参照により本明細書に援用される Malyshev, D.A., et al., *Nature*, vol. 509, pp. 385-388 (15 May 2014) に記載されるような他の非標準的ヌクレオチドが用いられてもよい。

【 0 0 4 1 】

以下の6つのペア (A T、G C、Z P、Ds Px、NAM SSICS、イソ C イソ G) は、ポリメラーゼと適合することが示されており、互いに直交している (すなわち、交差対形成 (cross pairing) のレベルが低い)。

【 0 0 4 2 】

【 化 1 】



【 0 0 4 3 】

したがって、本開示の態様は、4塩基システムでの2つの異なる塩基のペアとは対照的に、12塩基システムで6つの異なる塩基対を用いることを意図するものである。したがって、本開示の態様は、4塩基システムでの2つの異なる塩基のペアとは対照的に、6塩基システムで3つの異なる塩基のペアを用いることを意図するものである。

【 0 0 4 4 】

ある態様では、mRNA非依存的リボソームは、例えば20種の標準的アミノ酸を用いた本明細書に記載のヌクレオチドパルスに類似したtRNAのパルスと共に用いられて、情報をコードするポリマーが合成される。この態様により、より多くの種類のモノマー及びよりコンパクトなコードが得られ、すなわち、平均サイズが小さくなることによって3Xビット/gとなり、4塩基(2ビット)に対して20の標準的AA及び12の非標準的AA(5ビット)といった多様性の増加により5/2Xビット/gとなる。この実施形態では、一つの位置あたり、より多くのパルスが必要となる(4パルスに対して20パルス

10

20

30

40

50

又は32パルス)。

【0045】

非ヌクレオチドモノマー

本開示の実施形態は、ヌクレオチドに関して本明細書に記載されるように、ビットを表すことができ、且つポリマーを形成して情報を記録することができる他のモノマー分子を含む。そのようなポリマー及びそのモノマーとしては、ペプチド及びポリペプチド(コラーゲン及びバンコマイシンなど)、ケチド及びポリケチド(脂肪及びテトラサイクリンなど)、脂肪酸及び脂質、脂肪酸及び糖脂質、単糖及びリポ多糖、リン脂質、ホルモン、多糖(セルロース及びデンプンなど)、テルペン及びポリテルペン(コレステロール及びゴムなど)、アミノ酸及びポリアミノ酸(リグニン及びポリアルカロイドなど)、ピロール及びポリピロール(ヘム及びビタミンB12など)、並びにエステル及びポリエステル(PHA、PHVなど)などのモノマー及びバイオポリマーが挙げられる。他のポリマーとしては、シロキサン及びポリシロキサン、アクリルアミド及びポリアクリルアミドなどを含む直鎖状ポリマーなどの非生物学的ポリマーが挙げられる。そのようなオリゴマーは、十分な熱安定性を有し、又はナノ細孔若しくは他のポリマーシーケンシング装置中での検出が容易であり得る。非ヌクレオチドモノマーを用いてポリマーを作製する場合、そのようなモノマーの検出には当業者に公知の方法が用いられる。

10

【0046】

ある態様では、本明細書で特定される非ヌクレオチド系ポリマーを含むポリマーは、ナノ細孔又はナノギャップ又はナノチャンネルにポリマーを通過させることによってシーケンシングされて、ポリマー中の個々のモノマーが特定されてもよい。簡潔に述べると、ポリマーは、電気伝導性媒体中に存在し、電圧差の影響下でナノ細孔中を通過させられる。イオン電流の界面依存的変化を用いて、個々のモノマーの区別が行われる。

20

【0047】

「ナノ細孔」とは、ナノメートルスケールの幅を有する孔又は通路を意味する。例示的なナノ細孔としては、多量体タンパク質環によって形成された、膜を通した孔又は通路が挙げられる。通常、通路は、0.2~25nmの幅である。本明細書で用いられる場合、ナノ細孔は、膜を通した分子の移動を可能とする膜貫通構造を含み得る。ナノ細孔の例としては、ヘモリジン(スタフィロコッカス・アウレウス(*Staphylococcus aureus*))及びMspA(マイコバクテリウム・スメグマチス(*Mycobacterium smegmatis*))が挙げられる。ナノ細孔の他の例は、ナノ細孔シーケンシングについて述べている先行技術に見出され得るか、又はPFTのPanton ValentineロイコシジンS、アエロリジン、及びクロストリジウムイブシロン毒素、PFTの細胞溶解素A、二成分PFT(binary PFT)の炭疽毒素、又は他のニューモリシン若しくはグラミシジンなど、細孔形成毒素として本技術分野において記載され得る。ナノ細孔は、ナノ細孔シーケンシング技術の登場により、技術的にも経済的にも重要となってきた。ナノ細孔シーケンシングの方法は、例えば参照により援用される米国特許第5,795,782号に記載のように、本技術分野にて公知である。簡潔に述べると、ナノ細孔検出には、例えば、KCl、NaCl、NiCl、LiCl、又は当業者に公知の他のイオン形成無機化合物を含むイオン溶液などの電圧伝導性流体(voltage-conducting fluid)中に浸漬したナノ細孔が穿孔された膜が関与する。この膜を通して電圧が印加され、ナノ細孔を通過するイオンの伝導によって電流が流れる。ナノ細孔が、DNA又は他の非DNAポリマーなどのポリマーと相互作用すると、ナノ細孔を通る流れがモノマー特異的に変化して電流の変化を引き起こし、モノマーの検出が可能となる。本開示の範囲内のナノ細孔は、当業者に公知の固体非タンパク質ナノ細孔、及び当業者に公知のDNA折り紙ナノ細孔を含む。そのようなナノ細孔では、公知のタンパク質ナノ細孔よりも大きいナノ細孔幅が得られ、それによって、複合体がナノ細孔を通過する際のイオン電流の変化を検出するのに十分な感度を有しつつ、検出のためのより大きい分子の通過が可能となる。

30

40

【0048】

「ナノ細孔シーケンシング」とは、ポリマーのナノ細孔との相互作用に基づいて、ポ

50

リマーの構成成分を決定する方法を意味する。ナノ細孔シーケンシングは、ポリマーとの相互作用によって開口部の大きさが変化する際に発生する、ナノ細孔を通じたイオンのコンダクタンスの変化を測定することによって実現することができる。

【0049】

ナノ細孔に加えて、本開示は、2つの電極間のギャップであるとして本技術分野で公知であるナノギャップの使用も想定しており、このギャップは、約0.2 nm～約25 nm又は約2～約5 nmなどのおよそ数ナノメートルの幅である。ギャップは、ナノ細孔の開口部を模倣しており、ギャップ及び電極間をポリマーが通過可能である。本開示の態様はまた、ナノチャンネルの使用も想定しており、電極が、ポリマーが通過するナノチャンネルに隣接して配置される。当業者であれば、分子又は部分の電場中の移動、電場を通過する構造を表す電場のひずみの発生に基づいて、分子又は部分を同定し、シーケンシングする別の実施形態を容易に想定するであろうことは理解されたい。

【0050】

核酸合成

オリゴヌクレオチドは、末端デオキシヌクレオチジルトランスフェラーゼ (TdT) 若しくはエラープロンポリメラーゼを用いて、及び/又は本明細書に記載のパルス/同期法を用いて、本明細書に記載の方法で作製されてもよい。ある態様では、所望の位置での単一のヌクレオチドの結合を促進するパルス及び同期パラメータは、基材の寸法、試薬、濃度、反応温度、並びに試薬及び洗浄液のパルスを引き起こし送達するために用いられる構造に基づいて決定されてもよい。同期とは、単一のヌクレオチドのみの付加を最適化するために、酵素及び他の反応物質の存在下で、所望の位置での反応物質の希釈若しくは除去によって、又は所望の位置での試薬の不活性化によって反応速度にも影響を与え得る洗浄の後に、ある位置にヌクレオチドがとどまる時間を意味する。ある態様では、pH及び他の反応体及び反応条件が、TdTを用いて既存のヌクレオチド又はオリゴヌクレオチドに鑄型非依存的にdNTPが付加されるように最適化されてもよい。例えば、その内容が全体として参照により本明細書に援用されるAshley et al., Virology 77, 367-375 (1977)では、イニシエータサイズ、二価カチオン、及びpHなどのdNTP付加のための特定の試薬及び反応が同定されている。TdTは、広いpH範囲にわたって活性であり、最適pHは6.85であることが報告されている。

【0051】

ある例示的实施形態では、オリゴヌクレオチド配列は、1つ又は複数のホスホラミダイトリンカーの使用、及び/又は当業者に公知の連結法によるシーケンシングによって作製されてもよい。オリゴヌクレオチド配列は、例えば、本明細書において以下に記載の方法、及びBeaucage and Carruthers ((1981) Tetrahedron Lett. 22: 1859)に記載の方法などの標準的なホスホラミダイト法、又はMatteucci et al. (1981) J. Am. Chem. Soc. 103:3185)によるトリエステル法、又は市販の自動オリゴヌクレオチド合成装置若しくは当業者に公知のハイスループット高密度アレイ法を用いた他の化学的方法(その内容が全体として参照により本明細書に援用される、米国特許第5,602,244号、同第5,574,146号、同第5,554,744号、同第5,428,148号、同第5,264,566号、同第5,141,813号、同第5,959,463号、同第4,861,571号、及び同第4,659,774号参照)などの任意の適切な方法によって作製されてもよい。予め合成されたオリゴヌクレオチドも、様々な供給業者から市販のものを入手可能である。

【0052】

ある例示的实施形態では、オリゴヌクレオチド配列を、本技術分野で公知の様々なマイクロアレイ技術を用いて作製してもよい。予め合成されたオリゴヌクレオチド配列及び/又はポリヌクレオチド配列が、支持体に結合されてもよく、又は以下の文献: McGall et al. (1996) Proc. Natl. Acad. Sci. U.S.A. 93:13555; Synthetic DNA Arrays In Genetic Engineering, Vol. 20:111, Plenum Press (1998); Duggan et al. (1999) Nat. Genet. S21:10; Microarrays: Making Them and Using Them In Microarray Bioinformatics,

10

20

30

40

50

Cambridge University Press, 2003 ; 米国特許出願公開第 2 0 0 3 / 0 0 6 8 6 3 3 号及び同第 2 0 0 2 / 0 0 8 1 5 8 2 号 ; 米国特許第 6 , 8 3 3 , 4 5 0 号、同第 6 , 8 3 0 , 8 9 0 号、同第 6 , 8 2 4 , 8 6 6 号、同第 6 , 8 0 0 , 4 3 9 号、同第 6 , 3 7 5 , 9 0 3 号、及び同第 5 , 7 0 0 , 6 3 7 号 ; 並びに P C T 出願番号国際公開第 0 4 / 0 3 1 3 9 9 号、国際公開第 0 4 / 0 3 1 3 5 1 号、国際公開第 0 4 / 0 2 9 5 8 6 号、国際公開第 0 3 / 1 0 0 0 1 2 号、国際公開第 0 3 / 0 6 6 2 1 2 号、国際公開第 0 3 / 0 6 5 0 3 8 号、国際公開第 0 3 / 0 6 4 6 9 9 号、国際公開第 0 3 / 0 6 4 0 2 7 号、国際公開第 0 3 / 0 6 4 0 2 6 号、国際公開第 0 3 / 0 4 6 2 2 3 号、国際公開第 0 3 / 0 4 0 4 1 0 号、及び国際公開第 0 2 / 2 4 5 9 7 号に記載の、光による制御を用いた方法 (light-directed methods)、フローチャネル及びスポットティング法、インクジェット法、ピンに基づいた方法、並びにビーズに基づいた方法を用いて *in situ* で合成されてもよい。

10

【 0 0 5 3 】

ある態様では、オリゴヌクレオチド配列は、当業者に公知のインクジェット技術、当業者に公知の電気化学的技術、当業者に公知のマイクロ流体技術、当業者に公知の光発生酸、又は当業者に公知の光脱保護モノマーを用いて作製されてもよい。そのような技術は、高速でのオリゴヌクレオチド作製、低コスト、毒性の少ない化学物質、高い携帯性、及び *de novo* (デジタル又はアナログ) 合成で DNA 生化学 (例 : 修飾、ポリメラーゼ、ハイブリダイゼーションなど) をインタリーブする (interleave) 能力という利点を有する。例えば、カメラ光学系から直接の、又はデジタルマイクロミラー表示装置 (DMD) からの空間的にパターン化された光を、水溶性化学と共に用いてもよい。米国特許出願公開第 2 0 0 3 / 0 2 2 8 6 1 1 号を参照されたい。例えば、末端デオキシヌクレオチジルトランスフェラーゼ (TdT) 又はポリ (A) ポリメラーゼなどの鋳型非依存的ポリメラーゼ、あるいは、Taq 又は Phi 29 誘導体などの鋳型依存的ポリメラーゼは、ポリメラーゼ又は 5 から 3 エキソヌクレアーゼドメイン (存在する場合) の活性部位へのアゾベンゼンアミノ酸の取り込みにより、基本的なポリメラーゼ機能、塩基特異性又は忠実性を、光によりプログラム可能なものとすることができる (Hoppmann C, Schmieder P, Heinrich N, Beyermann M. (2011) *Chembiochem*.12(17):2555-9. doi: 10.1002/cbic.201100578. Epub 2011 Oct 13, Photoswitchable click amino acids: light control of conformation and bioactivity 参照)。

20

30

【 0 0 5 4 】

光感受性ニューロン (光遺伝学) は、イオン感受性ポリメラーゼを誘発可能であるか (Zamft B, Marblestone A, Kording K, Schmidt D, Martin-Alarcon D, Tyo K, Boyden E, Church GM(2012) *Measuring Cation Dependent DNA Polymerase Fidelity Landscapes by Deep Sequencing*. *PLoS One*, in press 参照)、又はある用途では、イオン束パターン自体が記憶されたデータセットを構築し得る。

【 0 0 5 5 】

ある態様では、核酸は、電極アレイ、従来のカメラ光学系、顕微鏡光学系、平面光学系 (フレネルマイクロレンズ又はビーズマイクロレンズ (bead microlens))、湾曲イメージング平面 (curved imaging planes) を用いて基材上に製造されてもよい。(デジタル写真の場合のように) 正方形、三角形、六角形、若しくは他の繰り返しモチーフのアレイ、又は (従来ハロゲン化銀写真の場合のように) アナログイメージングである。光が用いられる場合、空間的パターン化は、DMD (デジタルマイクロミラー表示装置)、他のデジタル投影法、又は自然の (アナログ) 光照射野を介するものであってもよい。

40

【 0 0 5 6 】

ある態様では、核酸は、その内容が全体として参照により本明細書に援用される米国特許第 6 , 0 9 3 , 3 0 2 号に開示されるように、電気化学的固相合成によって作製されてもよい。この態様では、多様な配列の個別のポリマー又は核酸配列が、電気化学的に生成された試薬の拡散に起因する電極間の化学的クロストークを防止するために緩衝溶液又は捕捉溶液と選択的に接触している、少なくとも 1 つの電極を含む基材上の特定の位置での

50

モノマー又はヌクレオチドの電気化学的配置を用いて作製される。

【0057】

ある態様では、その内容が全体として参照により本明細書に援用されるChurch et al., Nature, Vol. 432, 23/30 December 2004に記載のように、光発生酸を用いて核酸が合成されてもよい。

【0058】

ある態様では、dNTP、rNTP、若しくはrNDPを提供する又は送達する方法が、核酸の作製に有用である。dNTPが充填されたりポソーム内部のpH感受性ウィルス粒子からのリパーゼ又は他の膜溶解性酵素の放出が、J Clin Microbiol. May 1988; 26(5): 804-807に記載されている。NTPを放出し得る光ケージ化(Photo-caged) rNTP 又はdNTP、通常は、350nm光に対して感受性があるニトロベンジル誘導体は、Life technologies社から市販されている。ヌクレオチドの小胞分泌又は他の分泌をもたらすロドプシン又はバクテリオオプシンで誘発されるシグナル伝達は、本技術分野で公知である。dNTPを送達するためのこれらの方法では、ヌクレオチドは、プライマーとポリメラーゼの最初の遭遇と任意の下流との間で除去又は捕捉されるべきである。

【0059】

ある態様では、pH又は光を用いてポリメラーゼ活性を調整する方法が、核酸の作製に有用である。ヌクレオチド組み込みのための最適pH範囲及び可逆的活性が発生するpH範囲を有するポリメラーゼは、本技術分野で公知である。ACS Nano. 2014 May 27;8(5):4157-65に記載されるように、アゾベンゼンアミノ酸を、合成ペプチドを介して、又は改変されたtRNAによる独特の遺伝コードを介して、DNA又はRNAポリメラーゼに組み込むことができる。他の有用な方法は、Nature, 500(7463) August 22, 2013に記載されている。

【0060】

ポリメラーゼ

本発明の別の実施形態では、ポリメラーゼを用いて、情報を表す核酸分子が構築され、このことは、本明細書において、核酸配列に記録されるといわれるか、又は本明細書において、核酸が記憶媒体であるといわれる。ポリメラーゼは、例えば、DNA又はRNAを鋳型として用いて核酸配列を生成する酵素である。RNAポリマーを生成するポリメラーゼは、RNAポリメラーゼとして知られ、一方、DNAポリマーを生成するポリメラーゼは、DNAポリメラーゼとして知られる。エラーを組込むポリメラーゼが、本技術分野において知られており、本明細書において、「エラープロンポリメラーゼ」と呼ばれる。鋳型非依存的ポリメラーゼは、エラープロンポリメラーゼであってもよい。エラープロンポリメラーゼを用いることにより、特定の塩基をDNA分子の正確な位置に組み込むことが可能となる。エラープロンポリメラーゼは、可逆的鎖停止塩基などの非標準的塩基を許容することになるか、又はエラープロンポリメラーゼが鋳型を複製しようとする際に選択的にエラープロンポリメラーゼに与えられる天然若しくは未修飾のヌクレオチドを組込むことになる。DNAヌクレオチジルエキソトランスフェラーゼ(DNTT)又は末端トランスフェラーゼとしても知られる末端デオキシヌクレオチジルトランスフェラーゼ(TdT)などの鋳型非依存的ポリメラーゼは、鋳型なしでDNA分子の3'末端へのヌクレオチド付加を触媒することによって核酸鎖を作り出す。TdTの好ましい基質は3'オーバーハングであるが、平滑3'末端又は陥凹3'末端にもヌクレオチドを付加することができる。コバルトは補助因子であるが、*in vitro*でMg及びMnを投与すると、酵素は反応を触媒する。核酸イニシエータは、4ヌクレオチド若しくは5ヌクレオチド、又はそれより長くてもよく、一本鎖又は二本鎖であってもよい。二本鎖イニシエータは、3'オーバーハングを有してよく、平滑末端を有してよく、又は3'陥凹末端を有していてもよい。

【0061】

すべてのDNAポリメラーゼと同様に、TdTにも、触媒のために二価金属イオンが必

10

20

30

40

50

要である。しかし、TdTは、 Co^{2+} 、 Mn^{2+} 、 Zn^{2+} 、及び Mg^{2+} などの様々な二価カチオンを用いることができるという点で独特である。一般的に、二価金属イオンの存在下でのdATPによるプライマーの伸長速度 $p(dA)_n$ (n は、4~50の鎖長である)は、 $Mg^{2+} > Zn^{2+} > Co^{2+} > Mn^{2+}$ の順番である。加えて、各金属イオンは、ヌクレオチド組み込みの反応速度に対して異なる効果を有する。例えば、 Mg^{2+} は、dGTP及びdATPの優先的な使用を促進するが、 Co^{2+} は、ピリミジン類、dCTP、及びdTTPの触媒重合効率を高める。 Zn^{2+} は、 Zn^{2+} のマイクロモル量の添加によって Mg^{2+} との反応速度が増すことから、TdTに対する独特な正のエフェクターとして働く。この向上は、より高い触媒効率をもたらすTdTの立体構造変化を誘導する Zn^{2+} の能力を反映し得るものである。重合速度は、 Mg^{2+} と比較して Mn^{2+} の存在下ではより低く、 Mn^{2+} が Mg^{2+} ほど効率的には反応を支持しないことが示唆される。TdTについては、その内容が全体として参照により本明細書に援用されるBiochim Biophys Acta., May 2010; 1804(5): 1151-1166にさらに記載されている。さらに、ヌクレオチドパルス中の Mg^{2+} 、 Zn^{2+} 、 Co^{2+} 、又は Mn^{2+} を、ヌクレオチドの結合を調整するように設計された他のカチオンで置き換えてもよい。例えば、ヌクレオチドパルスの Mg^{++} を、 Na^+ 、 K^+ 、 Rb^+ 、 Be^{++} 、 Ca^{++} 、又は Sr^{++} などの他のカチオンで置き換えた場合、ヌクレオチドは結合可能であるが、組み込まれず、それによって、ヌクレオチドが組み込まれるかどうかは制御される。次に、(必要に応じて)ヌクレオチド又は Mg^{++} を含まない予備洗浄液のパルスが提供されてよく、又は次に、ヌクレオチドを含まない Mg^{++} 緩衝液が提供されてもよい。

10

20

【0062】

ポリメラーゼが利用可能であるヌクレオチドを制限することにより、特定の核酸のポリマーへの組み込みを制御することができる。よって、これらのポリメラーゼは、鋳型配列に対して非依存的にヌクレオチドを組込むことができ、したがって、核酸配列のde novo作製にとって有益である。エラープローンポリメラーゼとプライマー配列との組み合わせが、核酸配列に情報を付与するための書き込み機構として働く。

【0063】

鋳型非依存的ポリメラーゼが利用可能であるヌクレオチドを制限することにより、イニシエータ配列又は既存のヌクレオチド若しくはオリゴヌクレオチドへのヌクレオチドの付加を制御して、伸長によってオリゴヌクレオチドを作製することができる。したがって、これらのポリメラーゼは、鋳型配列なしにヌクレオチドを組込むことができ、したがって、核酸配列のde novo作製にとって有益である。

30

【0064】

イータポリメラーゼ(Matsuda et al. (2000) Nature 404(6781):1011-1013)は、4種すべてのdNTPの存在下で高い変異率(約10%)及び3 ミスマッチに対する高い許容度を有するポリメラーゼの例であり、1種又は2種のdNTPに制限されると、恐らくはより高くなる。したがって、イータポリメラーゼは、末端デオキシヌクレオチジルトランスフェラーゼ(TdT)同様に核酸情報のde novo記録体(de novo recorder)であるが、このポリメラーゼによって生成された生成物は、連続的な二本鎖であるという利点を有する。二本鎖DNAは、一本鎖DNAよりも付着性の低い二次構造を有し、より予測可能な二次構造を有する。さらに、二本鎖DNAは、ポリメラーゼの良好な支持体及び/又はDNA結合タンパク質テザーとしても働く。

40

【0065】

ある態様では、鋳型依存的又は鋳型半依存的エラープローンポリメラーゼが用いられてもよい。ある実施形態では、鋳型依存的ポリメラーゼが用いられてもよく、これはエラープローンとなり得る。ある実施形態では、鋳型非依存的RNAポリメラーゼが用いられてもよい。鋳型依存的又は鋳型半依存的ポリメラーゼが用いられる場合、ユニバーサル塩基(universal bases)を有する鋳型の任意の組み合わせが用いられてもよく、多くの種類のヌクレオチドの受容が促進される。また、 Mn^+ などのエラー許容性カチオンが用いられてもよい。さらに、本開示は、エラー許容性ポリメラーゼ変異体の使用も意図する。参

50

照により本明細書に援用されるBerger et al., Universal Bases for Hybridization, Replication and Chain Termination, Nucleic Acids Research 2000, August 1, 28(15) p. 2911-2914を参照されたい。

【 0 0 6 6 】

ある態様では、タンパク質、核酸、又は他のポリマーが、ポリメラーゼ、核酸、又は他のポリマーと共有結合又は非共有結合されて、ポリメラーゼがポリマーにモノマーを付加する能力を変化させるように、ポリメラーゼ及びプライマーの関連性が改変されてもよい。

【 0 0 6 7 】

ある態様では、DNA分解酵素又はRNA分解酵素は、逆行させることができ、核酸の作製に有用である。一例としては、リボNTPを作製するためのポリヌクレオチドホスホリラーゼがある。

【 0 0 6 8 】

ある態様では、リガーゼは、核酸の作製に有用である。そのようなリガーゼには、当業者に公知のDNAリガーゼ及び当業者に公知のRNAリガーゼが含まれる。DNAリガーゼには、細菌DNAリガーゼ及び哺乳類DNAリガーゼが含まれる。例示的なりガーゼとしては、T3リガーゼ、T4リガーゼ、T7リガーゼ、大腸菌DNAリガーゼ、Taq DNAリガーゼ、及びcircリガーゼなどが挙げられる。

【 0 0 6 9 】

ある態様では、支持体表面上で合成された核酸は、当業者に公知の切断可能リンカーなどによって除去することができる。核酸は、製造密度よりも高い密度などで異なる基材上に配置されてよく、又は記憶媒体として働くことになる異なる基材上に配置されてもよい。また、さらなる核酸合成のための新たな基材として働く基材の別の層が追加されてもよい。したがって、第一の基材上で複数のオリゴヌクレオチドを作製すること、複数のオリゴヌクレオチドを第一の基材から除去すること、及びそれらをランダムに又は規則的に、所望の密度で第二の基材に結合させることによる、高密度核酸記憶デバイスの作製方法が提供される。

【 0 0 7 0 】

支持体及び結合

ある例示的实施形態では、本明細書に記載の1又は複数のオリゴヌクレオチド配列は、支持体（例えば、固体支持体及び/又は半固体支持体）上に固定される。ある態様では、オリゴヌクレオチド配列は、本明細書に記載のホスホロアミダイトリンカーのうちの1つ又は複数を用いて支持体に結合されてもよい。適切な支持体としては、これらに限定されないが、スライド、ビーズ、チップ、粒子、ストランド、ゲル、シート、チューブ、球体、容器、キャピラリー、パッド、スライス、フィルム、及びプレートなどが挙げられる。様々な実施形態では、固体支持体は、生物学的支持体、非生物学的支持体、有機支持体、無機支持体、又はこれらのいずれの組み合わせであってもよい。本発明の支持体は、必要に応じて、いかなる形状、寸法、又は表面形状であってもよい。例えば、支持体は、正方形、長方形、円形、平坦、平面、環状、管状、及び球状などであってもよい。実質的に平面の支持体を用いる場合、支持体を、例えば、トレンチ、溝、ウエル、又は化学的障壁（例えば、疎水性コーティングなど）で物理的に複数の領域に分けてもよい。支持体は、ガラス（二酸化ケイ素）、金属、セラミック、ポリマー、又は当業者に公知の他の材料から作られていてもよい。支持体は、固体、半固体、エラストマー、又はゲルであってもよい。ある例示的实施形態では、支持体はマイクロアレイである。本明細書で用いられる場合、「マイクロアレイ」の用語は、ある実施形態において、固定化ハイブリダイゼーションプローブをそれぞれが含む空間的に定められた非重複領域又は部位のアレイが存在する、実質的に平面である表面を有する固相支持体を含むアレイの種類を意味する。「実質的に平面」とは、プローブ部位などの表面上に存在する対象となる特性（feature）又は物体が、表面上又は表面下に広がる体積を占めていてよく、この体積の寸法は表面の寸法に対して小さいことを意味する。例えば、光ファイバー束の面上に配置されたビーズは、プロ

10

20

30

40

50

ープ部位の実質的に平面である表面を作り出し、又は多孔性平面基材上に配置又は合成されたオリゴヌクレオチドは、実質的に平面である表面を作り出す。空間的に定められた部位は、その位置及びその位置における固定化プローブの同一性が既知又は特定可能であるという点において、さらに、「アドレス可能」であってもよい。

【0071】

固体支持体は、固体成分及び液体成分の両方を有する圧縮可能マトリックスなどの半固体支持体も含んでいてもよく、液体は、細孔、空間、又は固体マトリックス要素間の他の空隙を占めている。好ましくは、半固体支持体材料としては、ポリアクリルアミド、セルロース、ポリジメチルシロキサン、ポリアミド（ナイロン）、並びに架橋アガロース、架橋デキストラン、及び架橋ポリエチレングリコールが挙げられる。固体支持体及び半固体支持体は、一緒に用いられても、又は互いに別々に用いられてもよい。

10

【0072】

支持体はまた、固定化媒体を含んでいてもよい。本発明において有用であるそのような固定化媒体は、核酸分子の堆積及び増幅に必要とされる条件下で物理的に安定であり、且つ化学的に不活性である。有用な支持体マトリックスは、PCRに必要とされる温度の急激な変化及び極端な温度に耐える。支持体材料は、核酸の酵素的合成を可能とするものである。ある基材がそのようであるかどうかは未知である場合、本発明による一連のアレイを製造する任意の試みを行う前に、基材が実験的に試験される。本発明のある実施形態では、支持体の構造は、半固体（すなわち、ゼラチン質）格子又はマトリックスを含み、格子又はマトリックス要素間の空隙又は細孔は、水性媒体又は他の液体媒体で満たされ、典型的な細孔（又は「ふるい」）の寸法は、 $100\ \mu\text{m} \sim 5\ \text{nm}$ の範囲内である。マトリックス要素間のより大きい空間も許容範囲内であるが、増幅産物が固定される前に拡散する可能性が増大する。半固体支持体は、圧縮可能である。支持体は、印刷の目的のために、平面又は実質的に平面であるように作製される。例えば、実質的に平面である支持体は、円筒形であってもよく、その場合、アレイの核酸は、平面又は円筒形である他の支持体と、一方が他方の上を転がることによって接触するために、その外側表面全体に分布されている。次に、本発明において有用である支持体材料は、当業者に公知の手段によって、支持体材料へのアレイの核酸特性の固定化（共有結合）を可能とするものである。このような必要条件を満たす材料は、有機物質及び無機物質の両方を含み、これらに限定されないが、ポリアクリルアミド、セルロース、及びポリアミド（ナイロン）、並びに架橋アガロース、架橋デキストラン、又は架橋ポリエチレングリコールが挙げられる。

20

30

【0073】

ある実施形態は、プレート、スライド、又はチップなどのガラス支持体上の薄いポリアクリルアミドゲルに関する。この種のポリアクリルアミドシートは、以下のようにして合成される。ゲルに用いられる混合物の全体としての割合がポリアクリルアミドシートに必要とされる引張特性を与えるように調節された場合に所望の細孔寸法が得られる個々のポリマー鎖間の架橋度が得られるように設計された比率で、アクリルアミド及びビスアクリルアミドが混合される（例えば、シークエンシングゲルの場合、38：2の比率が典型的である）。ポリアクリルアミドゲルのキャスト法は、本技術分野において公知であり（その内容が全体として参照により本明細書に援用されるSambrook et al., 1989, Molecular Cloning. A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY参照）、当業者であれば、そのような調節を行うのに何の困難もない。

40

【0074】

ゲルシートは、2つの剛性面間にキャストされ、そのうちの少なくとも一方は、他方を取り外した後もゲルシートが付着したままとなるガラスである。重合完了後に取り外されることになるキャスト面は、ゲルの重合を阻害しない潤滑剤でコーティングされており、この目的のために、一般的にはシランが用いられる。シランの層は、ドラフトチャンバー下で面上に広げられ、ほぼ乾燥するまで静置される。次に、過剰のシランがエタノールで除去される（拭き取られるか、又は小さい物体の場合は、十分にリンスされる）。ゲルシ

50

ートが付着したままとなるガラス面は、キャストする前は「架橋シラン」と呼ばれることの多いメタクリロキシプロピルトリメトキシシラン（カタログ番号 M6514、Sigma社；ミズーリ州、セントルイス）で処理される。ゲルと接触することになるガラス面は、この薬剤で3回コーティングされる。1200 cm²の面積の1回の処理ごとに、25 mlのエタノール中の125 μlの架橋シランが必要である。この溶液は、ガラス面全体に広げられる直前に、750 μlの水及び75 μlの氷酢酸の混合物と混合され、激しく振とうされる。エタノール溶媒は、コーティングの合間（ドラフトチャンパー下で約5分間）に蒸発させ、最後のコーティングが乾燥した後、他方の支持プレートがゲル上へ「サンドイッチすること」を防止する目的で、過剰の架橋シランが、十分なエタノール洗浄を介してできる限り完全に除去される。次に、プレートが組み立てられ、必要に応じてゲルがキャストされる。

10

【0075】

本発明において有用であるゲルの大きさを決定する唯一の操作上の制約は、当業者がそのようなゲルをキャストする身体的能力である。1メートルまでの長さのゲルのキャストは、楽ではないが、核酸シーケンシング技術の当業者に公知の手順である。より大きいゲルも、作製された場合、本発明において有用である。極めて小さいゲルは、より大きいゲル全体から、重合完了後に切り出される。

【0076】

ポリアクリルアミドゲルを、そのマトリックス中にトラップされる酵素などの生物活性物質と共にキャストするための少なくとも1つの手順が、本技術分野において公知であることに留意されたい（その内容が全体として参照により本明細書に援用されるO' Driscoll, 1976, *Methods Enzymol.*, 44: 169-183）。光架橋性ポリエチレングリコール樹脂を用いて、ゲルマトリックス中に生細胞をトラップさせる類似のプロトコルも報告されている（その内容が全体として参照により本明細書に援用されるNojima and Yamada, 1987, *Methods Enzymol.*, 136: 380-394）。そのような方法は、本発明において有用である。以下に記載されるように、パルスフィールドゲル電気泳動に適するマトリックス中に無処置の染色体DNAを送達する目的で、又はBenton and Davisの方法（その内容が全体として参照により本明細書に援用されるManiatis et al., 1982, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY参照）よりプラークを取り出す前に、バクテリオファージの増殖を支持する宿主細胞の「菌叢（lawn）」として利用するために、通常、全細胞がアガロース中にキャストされる。簡潔に述べると、電気泳動グレードのアガロース（例えば、Ultrapure；Life Technologies / Gibco BRL）を生理学的（等張性）緩衝液中に溶解し、管、瓶、又はフラスコ中、50 ~ 52 の温度に平衡化させる。次に、細胞をアガロースに添加し、（瓶又は管の場合は、蓋をして反転させることによって、フラスコの場合は、回旋することによって）十分にしかし素早く混合し、その後、混合物をゲルトレーにデカントするか、又はピペットで移す。低融点アガロースが用いられる場合、細胞の添加の前に、かなり低い温度とされてもよい（アガロースの濃度に応じて、およそ室温まで低下）。これは、ある細胞型の場合は望ましいが、電気泳動が、得られた核酸プールの分子を支持体に共有結合させる前の細胞溶解に続いて行われる場合は、4 ~ 10 の「低温」室などの冷蔵下で行われる。

20

30

40

【0077】

マイクロアレイ上に固定されたオリゴヌクレオチドは、アッセイ反応において作製された、又はアッセイ反応から作製された核酸を含む。典型的には、マイクロアレイ上のオリゴヌクレオチド又はポリヌクレオチドは、一本鎖であり、通常は5'末端又は3'末端によって固相支持体に共有結合している。ある例示的实施形態では、プローブは、1つ又は複数の切断可能なリンカーを介して固定されている。マイクロアレイにおける核酸を含む非重複領域の密度は、典型的には、100 / cm²より大きく、より典型的には、1000 / cm²より大きい。核酸プローブに関するマイクロアレイ技術は、以下の例示的文獻に概説されている：Schena, Editor, *Microarrays: A Practical Approach* (IRL Press,

50

Oxford, 2000) ; Southern, *Current Opin. Chem. Biol.*, 2: 404-410 (1998) ; *Nature Genetics Supplement*, 21:1-60 (1999) ; 並びにFodor et al, 米国特許第 5 , 4 2 4 , 1 8 6 号 ; 同第 5 , 4 4 5 , 9 3 4 号 ; 及び同第 5 , 7 4 4 , 3 0 5 号。

【 0 0 7 8 】

支持体にオリゴヌクレオチドを固定する方法は、本技術分野において公知である (ビーズ : Dressman et al. (2003) *Proc. Natl. Acad. Sci. USA* 100:8817、Brenner et al. (2000) *Nat. Biotech.* 18:630、Albretsen et al. (1990) *Anal. Biochem.* 189:40、及びLang et al. *Nucleic Acids Res.* (1988) 16:10861 ; ニトロセルロース : Ranki et al. (1983) *Gene* 21:77 ; セルロース : Goldkorn (1986) *Nucleic Acids Res.* 14:9171 ; ポリスチレン : Ruth et al. (1987) *Conference of Therapeutic and Diagnostic Applications of Synthetic Nucleic Acids*, Cambridge U.K. ; テフロン アクリルアミド : Duncan et al. (1988) *Anal. Biochem.* 169:104 ; ポリプロピレン : Polsky-Cynkin et al. (1985) *Clin. Chem.* 31:1438 ; ナイロン : Van Ness et al. (1991) *Nucleic Acids Res.* 19:3345 ; アガロース : Polsky-Cynkin et al., *Clin. Chem.* (1985) 31:1438 ; 並びにセファクリル : Langdale et al. (1985) *Gene* 36:201 ; ラテックス : Wolf et al. (1987) *Nucleic Acids Res.* 15:2911) 。 支持体は、アミノ シラン、NH₂ エステル、クリックケミストリー、ポリリジンなどの結合化学物質又はポリマーでコーティングされてもよい。

【 0 0 7 9 】

本明細書で用いられる場合、用語「結合」は、共有結合的相互作用及び非共有結合的相互作用の両方を意味する。共有結合的相互作用は、1対の電子(すなわち単結合)、2対の電子(すなわち二重結合)、若しくは3対の電子(すなわち三重結合)の共有により形成される2個の原子又はラジカル間の化学結合である。共有結合的相互作用は、本技術分野において、電子対相互作用又は電子対結合としても知られる。非共有結合的相互作用としては、これらに限定されないが、ファンデルワールス相互作用、水素結合、弱化学結合(すなわち、短距離非共有結合力を介する)、疎水性相互作用、及びイオン結合などが挙げられる。非共有結合的相互作用の概説は、Alberts et al., in *Molecular Biology of the Cell*, 3d edition, Garland Publishing, 1994を参照されたい。

【 0 0 8 0 】

ある態様では、核酸分子の基材への付着又は固定は、酸化3-メチルウリジン、アクリリル基、及びヘキサエチレングリコールを含む群から選択される共有結合リンカーを用いて行われる。支持体への指向性結合で用いるプールの分子ヘリンカー配列を結合することに加えて、制限部位又は制御エレメント(プロモーターエレメント、キャップ部位、又は翻訳終止シグナルなど)が、必要に応じてプールのメンバーと接合される。リンカーはまた、酵素、光、熱、pH緩衝液、及びレドックス試薬などの薬剤で必要に応じて切断可能である化学反応性セグメントを有するように設計されてもよい。そのようなリンカーを用いて、別個の特性のそれぞれに対する異なる溶液相プライマーの*in situ*固相不活性リザーバ(*in situ* solid-phase inactive reservoir)が予め作製されてもよい。リンカーが切断されると、恐らくは熱サイクルプロセスからの熱をきっかけとして用いることによって、プライマーがPCRの溶液中に放出されることになる。

【 0 0 8 1 】

基材への核酸分子の付着は、基材と共有結合されている核酸分子に、プールのメンバーをハイブリダイズすることによって行われることも意図される。

【 0 0 8 2 】

本発明による支持体マトリックスへの核酸分子の固定は、いくつかの手順のうちのいずれによって行われてもよい。支持体との共有結合に適した化学基を有する3'-末端タグの使用による直接固定、基材と既に結合されているオリゴヌクレオチドプライマーへの核酸分子プール的一本鎖分子のハイブリダイゼーション、又はプレートの前若しくは後に添加される、基材と共有結合し得るプライマーの導入に伴う核酸分子の基材上への分散が実施され得る。予め固定されたプライマーが用いられる場合、プライマーは、広範囲の配列モチーフ(例えば、ヘキサマーなど、所与の鎖長の考え得るすべての多量体)、特定

10

20

30

40

50

の配列に対する相同性を有する核酸、又は特定の配列モチーフ上に変異を含む核酸を捕捉するように設計される。あるいは、プライマーは、リンカー配列など、核酸分子のプールのすべてのメンバーに共通の合成分子特性を包含する。

【0083】

核酸分子をポリアクリルアミドゲルシートに架橋させる2つの手段について詳細に考察する。第一の手段は (Khrapko et al., 1996、米国特許第5,552,270号に記載)、3-メチルウリジンによる核酸分子の3'-キャッピングを含む。この方法を用いることにより、本発明のライブラリの核酸分子は、3'-末端にこの修飾塩基を含むように作製される。引用されたプロトコルでは、厚さ30 μ mの8%ポリアクリルアミドゲル(30:1、アクリルアミド:ビス-アクリルアミド)シートがキャストされ、次に50%ヒドロ

10

20

30

40

50

60

70

80

90

100

110

120

130

140

150

160

170

180

190

200

210

220

230

240

250

260

270

280

290

300

310

320

330

340

350

360

370

380

390

400

410

420

430

440

450

460

470

480

490

500

510

520

530

540

550

560

570

580

590

600

610

620

630

640

650

660

670

680

690

700

710

720

730

740

750

760

770

780

790

800

810

820

830

840

850

860

870

880

890

900

910

920

930

940

950

960

970

980

990

1000

1010

1020

1030

1040

1050

1060

1070

1080

1090

1100

1110

1120

1130

1140

1150

1160

1170

1180

1190

1200

1210

1220

1230

1240

1250

1260

1270

1280

1290

1300

1310

1320

1330

1340

1350

1360

1370

1380

1390

1400

1410

1420

1430

1440

1450

1460

1470

1480

1490

1500

1510

1520

1530

1540

1550

1560

1570

1580

1590

1600

1610

1620

1630

1640

1650

1660

1670

1680

1690

1700

1710

1720

1730

1740

1750

1760

1770

1780

1790

1800

1810

1820

1830

1840

1850

1860

1870

1880

1890

1900

1910

1920

1930

1940

1950

1960

1970

1980

1990

2000

2010

2020

2030

2040

2050

2060

2070

2080

2090

2100

2110

2120

2130

2140

2150

2160

2170

2180

2190

2200

2210

2220

2230

2240

2250

2260

2270

2280

2290

2300

2310

2320

2330

2340

2350

2360

2370

2380

2390

2400

2410

2420

2430

2440

2450

2460

2470

2480

2490

2500

2510

2520

2530

2540

2550

2560

2570

2580

2590

2600

2610

2620

2630

2640

2650

2660

2670

2680

2690

2700

2710

2720

2730

2740

2750

2760

2770

2780

2790

2800

2810

2820

2830

2840

2850

2860

2870

2880

2890

2900

2910

2920

2930

2940

2950

2960

2970

2980

2990

3000

3010

3020

3030

3040

3050

3060

3070

3080

3090

3100

3110

3120

3130

3140

3150

3160

3170

3180

3190

3200

3210

3220

3230

3240

3250

3260

3270

3280

3290

3300

3310

3320

3330

3340

3350

3360

3370

3380

3390

3400

3410

3420

3430

3440

3450

3460

3470

3480

3490

3500

3510

3520

3530

3540

3550

3560

3570

3580

3590

3600

3610

3620

3630

3640

3650

3660

3670

3680

3690

3700

3710

3720

3730

3740

3750

3760

3770

3780

3790

3800

3810

3820

3830

3840

3850

3860

3870

3880

3890

3900

3910

3920

3930

3940

3950

3960

3970

3980

3990

4000

4010

4020

4030

4040

4050

4060

4070

4080

4090

4100

4110

4120

4130

4140

4150

4160

4170

4180

4190

4200

4210

4220

4230

4240

4250

4260

4270

4280

4290

4300

4310

4320

4330

4340

4350

4360

4370

4380

4390

4400

4410

4420

4430

4440

4450

4460

4470

4480

4490

4500

4510

4520

4530

4540

4550

4560

4570

4580

4590

4600

4610

4620

4630

4640

4650

4660

4670

4680

4690

4700

4710

4720

4730

4740

4750

4760

4770

4780

4790

4800

4810

4820

4830

4840

4850

4860

4870

4880

4890

4900

4910

4920

4930

4940

4950

4960

4970

4980

4990

5000

5010

5020

5030

5040

5050

5060

5070

5080

5090

5100

5110

5120

5130

5140

5150

5160

5170

5180

5190

5200

5210

5220

5230

5240

5250

5260

5270

5280

5290

5300

5310

5320

5330

5340

5350

5360

5370

5380

5390

5400

5410

5420

5430

5440

5450

5460

5470

5480

5490

5500

5510

5520

5530

5540

5550

5560

5570

5580

5590

5600

5610

5620

5630

5640

5650

5660

5670

5680

5690

5700

5710

5720

5730

5740

5750

5760

5770

5780

5790

5800

5810

5820

5830

5840

5850

5860

5870

5880

5890

5900

5910

5920

5930

5940

5950

5960

5970

5980

5990

6000

6010

6020

6030

6040

6050

6060

6070

6080

6090

6100

6110

6120

6130

6140

6150

6160

6170

6180

6190

6200

6210

6220

6230

6240

6250

6260

6270

6280

6290

6300

6310

6320

6330

6340

6350

6360

6370

6380

6390

6400

6410

6420

6430

6440

6450

6460

6470

6480

6490

6500

6510

6520

6530

6540

6550

6560

6570

6580

6590

6600

6610

6620

6630

6640

6650

6660

6670

6680

6690

6700

6710

6720

6730

6740

6750

6760

6770

6780

6790

6800

6810

6820

6830

6840

6850

6860

6870

6880

6890

6900

6910

6920

6930

6940

6950

6960

6970

6980

6990

7000

7010

7020

7030

7040

7050

7060

7070

7080

7090

7100

7110

7120

7130

7140

7150

7160

7170

7180

7190

7200

7210

7220

7230

7240

7250

7260

7270

7280

7290

7300

7310

7320

7330

7340

7350

7360

7370

7380

7390

7400

7410

7420

7430

7440

7450

7460

7470

7480

7490

7500

7510

7520

7530

7540

7550

7560

7570

7580

7590

7600

7610

7620

7630

7640

7650

7660

7670

7680

7690

7700

7710

7720

7730

7740

7750

7760

7770

7780

7790

7800

7810

7820

7830

7840

7850

7860

7870

7880

7890

7900

7910

7920

7930

7940

7950

7960

7970

7980

7990

8000

8010

8020

8030

8040

8050

8060

8070

8080

8090

8100

8110

8120

8130

8140

8150

8160

8170

8180

8190

8200

8210

8220

8230

8240

8250

8260

8270

8280

8290

8300

8310

8320

8330

8340

8350

8360

8370

8380

8390

8400

8410

8420

8430

8440

8450

8460

8470

8480

8490

8500

8510

8520

8530

8540

8550

8560

8570

8580

8590

8600

8610

8620

8630

8640

8650

8660

8670

8680

8690

8700

8710

8720

8730

8740

8750

8760

8770

8780

8790

8800

8810

8820

8830

8840

8850

8860

8870

8880

8890

8900

8910

8920

8930

8940

8950

8960

8970

8980

8990

9000

9010

9020

9030

9040

9050

9060

9070

9080

9090

9100

9110

9120

9130

9140

9150

9160

9170

9180

9190

9200

9210

9220

9230

9240

9250

9260

9270

9280

9290

9300

9310

9320

9330

9340

9350

9360

9370

9380

9390

9400

9410

9420

9430

9440

9450

9460

9470

9480

9490

9500

9510

9520

9530

9540

9550

9560

9570

9580

9590

9600

9610

9620

9630

9640

9650

9660

9670

9680

9690

9700

9710

9720

9730

9740

9750

9760

9770

9780

9790

9800

9810

9820

9830

9840

9850

9860

9870

9880

9890

9900

9910

9920

9930

9940

9950

9960

9970

9980

9990

10000

【0084】

ポリアクリルアミドシートに核酸分子を共有結合させるのに有用である第二の架橋部分は、プライマーに結合された5'-アクリリル基である。そのような修飾された塩基を5'-末端に有するオリゴヌクレオチドプライマーを、本発明に従って用いてもよい。特に、そのようなオリゴヌクレオチドは、アクリリル基が重合マトリックスの不可欠な共有結合部分となるように、ゲル中に直接キャストされる。プライマーの3'-末端は、自由にプールの核酸分子と相互作用及びハイブリダイズし、且つ酵素による第二鎖の合成を刺激するよう

30

40

50

60

70

80

90

100

110

120

130

140

150

160

170

180

190

200

210

220

230

240

250

260

270

280

290

300

310

320

330

340

350

360

370

380

390

400

410

420

430

440

450

460

470

480

490

500

510

520

530

540

550

560

570

580

590

600

610

620

630

640

650

660

670

680

690

700

710

720

730

740

750

760

770

780

790

800

810

820

830

840

850

860

870

880

890

900

910

920

930

940

950

960

970

980

990

1000

1010

1020

1030

1040

1050

1060

1070

1080

1090

1100

1110

1120

1130

1140

1150

1160

1170

1180

1190

1200

1210

1220

1230

1240

1250

1260

1270

1280

1290

1300

1310

1320

1330

1340

1350

1360

1370

1380

1390

1400

1410

1420

1430

1440

1450

1460

1470

1480

1490

1500

1510

1520

1530

1540

1550

1560

1570

1580

1590

1600

1610

1620

1630

1640

1650

1660

1670

1680

1690

1700

1710

1720

1730

1740

1750

1760

1770

1780

1790

1800

1810

1820

1830

1840

1850

1860

1870

1880

1890

1900

1910

1920

1930

1940

1950

1960

1970

1980

1990

2000

2010

2020

2030

2040

2050

2060

2070

2080

2090

2100

2110

2120

2130

2140

2150

2160

2170

2180

2190

2200

2210

2220

2230

2240

2250

2260

2270

2280

2290

2300

2310

2320

2330

2340

2350

2360

2370

2380

2390

2400

2410

2420

2430

2440

2450

2460

2470

2480

2490

2500

2510

2520

2530

2540

2550

2560

2570

2580

2590

2600

2610

2620

2630

2640

2650

2660

2670

2680

2690

2700

2710

2720

2730

2740

2750

2760

2770

2780

2790

2800

2810

2820

2830

2840

2850

2860

2870

2880

2890

2900

2910

2920

2930

2940

2950

2960

2970

2980

2990

3000

3010

3020

3030

3040

3050

3060

3070

3080

3090

3100

3110

3120

3130

3140

3150

3160

3170

3180

3190

3200

3210

3220

3230

3240

3250

3260

3270

3280

3290

3300

3

T Pを共有結合するように、パルスとして送達される。選択されたヌクレオチド試薬液は、パルスされるか、又はアレイ上を流動され、続いて、ヌクレオチドを含まない緩衝液又は洗浄液がパルスされる。パルスの継続時間は、例えば、酵素反応であっても又はそれ以外であっても、反応速度によって決定される。パルスの継続時間は、ヌクレオチドと洗浄液とで異なっているもよく、又は同じであってもよい。例えば、試薬の送達及び洗浄液の送達のいずれにおいても、0.2秒パルスが有効である。ある態様では、試薬送達及び洗浄液送達のタイミングは、反応のためのヌクレオチドのアクセス及び反応速度に関して同期される。

【0087】

例えば、図1、図2、及び図3を参照すると、基材は、円で示される例示的な反応位置を備えている。反応位置は、ランダムであってもよく、又は規則的アレイの場合のように、予め定められた領域であってもよい。基材の表面は囲まれて反応領域若しくは反応体の中を流動し得るフローセルが作られていてもよく、又は反応体の中を流動し得るフローセル内に基材が配置されていてもよい。反応領域は、反応体が添加され、基材の表面に接触され、除去され得るように、入口部及び出口部を有する。別個の試薬、又は試薬若しくは洗浄液の混合物を、ポンプ若しくは電極、又は流体をチャンネル若しくはマイクロ流体チャンネルを通して移動させる技術分野の当業者に公知の他の方法を用いて、ヌクレオチドが添加される所望の位置に試薬が接触可能であるように1つ又は複数のチャンネルを通して基材表面に位置する反応領域又は反応容器へ送達することができる、様々なフローセル形態、フローチャンネル形態、又はマイクロ流体チャンネル形態が想定される。所望の位置は、特定のpHを作り出すための、又は酵素が所望の位置でのヌクレオチドの付加について活性化、不活性化、又は制御されることになる特定のpHで特定の体積の流体を送達するための、電極又は他のデバイスも含んでいてもよい。

【0088】

図1に示されるように、ヌクレオチドAを含む試薬の第一のパルスが、基材の表面を横切るように流動される。ヌクレオチドAは、ヌクレオチドAがイニシエータ配列又は既存のヌクレオチド若しくはオリゴヌクレオチドに酵素的に付加されるのに十分な反応条件である基材上の1つ又は複数の所望の位置に添加される。ヌクレオチドAの添加後、洗浄液が基材の表面を横切るように流動されて、ヌクレオチドAが除去されてもよい。次に、ヌクレオチドTを含む試薬の第二のパルスが、基材の表面を横切るように流動される。ヌクレオチドTは、ヌクレオチドTがイニシエータ配列又は既存のヌクレオチド若しくはオリゴヌクレオチドに酵素的に付加されるのに十分な反応条件である基材上の1つ又は複数の所望の位置に添加される。ヌクレオチドTの添加後、洗浄液が基材の表面を横切るように流動されて、ヌクレオチドTが除去されてもよい。次に、ヌクレオチドGを含む試薬の第三のパルスが、基材の表面を横切るように流動される。ヌクレオチドGは、ヌクレオチドGがイニシエータ配列又は既存のヌクレオチド若しくはオリゴヌクレオチドに酵素的に付加されるのに十分な反応条件である基材上の1つ又は複数の所望の位置に添加される。ヌクレオチドGの添加後、洗浄液が基材の表面を横切るように流動されて、ヌクレオチドGが除去されてもよい。図2は、ヌクレオチドCを含む試薬の第四のパルス、及びヌクレオチドAを含む試薬の第五のパルスを示す。図3は、パルスが基材の表面を横切るように移動し、反応領域から出ていくことを示している。ある例示的な態様では、所望の位置でのpHが、所望の位置でのpH感受性酵素の活性を制御してこの所望の位置で試薬流体中のヌクレオチドを酵素的に付加するように制御される。ある例示的な態様では、所望の位置での光は、所望の位置での光感受性酵素の活性を制御してこの所望の位置で試薬流体中のヌクレオチドを酵素的に付加するように制御される。活性化刺激によって酵素を活性化する方法は、当業者に公知であり、本発明において、所望の位置で試薬流体中のヌクレオチドを酵素的に付加するのに有用である。

【0089】

ある態様では、体積100ナノリットル未満の液体の多数の小さなプールを多重操作することを可能とするデバイスが提供される。2つの平行な平面、及び液体サンプルが表面

10

20

30

40

50

張力によって各貫通孔に維持される寸法とされた貫通孔を有するプラテンから成る、複数の液体サンプルを分析するためのシステムが本技術分野において知られている（その内容が全体として参照により本明細書に援用される欧州特許第1051259(A1)号、2000年11月15日）。サンプルは、毛管作用を用いて平面から吸い出すことができ、希釈及び混合が可能である。各貫通孔は、光学的放射線によって調べることができる。このデバイス、さらにはGene Logic社のFlow Thru Chip（商標）（その内容が全体として参照により本明細書に援用されるTorres et al., 国際公開第01/45843(A2)号、2001年6月28日）などの類似のデバイスは、本明細書に記載の方法において有用である。各チャンパーの内壁は、5'結合鋳型核酸配列で官能化されてもよく、他の必要な試薬はすべて（部位特異的リコンビナーゼ又はエラープロ- 10
ーンポリメラーゼ、及びヌクレオチドなど）、別個のチャンパー（又は「ハニカム」セル）のそれぞれに液相で送達される。

【0090】

ある実施形態では、顕微鏡スライドなどの基材は、1) 同じプールの非分析核酸分子が、スライド上のすべての特性の全体に均一に分散されることを意図する場合、湿潤性表面境界領域によって分割されてよく、又は2) 各特性が、異なるプールの非分析核酸分子及び/若しくはプライマーを用いてスポットされる場合、非湿潤性表面境界領域によって分割されてもよい。非湿潤性境界でそれぞれ囲まれた連続的な湿潤性領域のより大きい群にさらに分割されたスライドなど、上記の組み合わせも可能であり、この場合、各湿潤可能領域は、それぞれ異なるプライマーがスポットされたより小さい特性にさらに分割される 20

【0091】

別の実施形態では、連続的な疎水性境界で分割された小さい非連続的な親水性の特性を有するスライドを、希釈DNA鋳型分子の水溶液に浸漬することによって、単一のDNA分子を区画分けすることも可能である。スライドを取り出し、横にして緩やかに拭き取ることにより、親水性の特性上に液体の小さなビーズが形成されることになり、それによって、0、1、又は2のDNA鋳型を有する小さい非連続的な液体のプールが作り出される（関連技術の記載に関し、その内容が全体として参照により本明細書に援用されるBrennan、米国特許第6,210,894(B1)号、2001年4月3日参照）。

【0092】

別の実施形態では、マイクロ流体デバイスは、1又は複数の試薬を含む1又は複数のリザーバを備え、1又は複数の試薬は、試薬が混合されて反応が行われる反応ゾーンへマイクロチャンネルを介して移送される。そのようなマイクロ流体デバイス及び流体試薬をそのようなマイクロ流体デバイスを通して移動させる方法は、当業者に公知である。 30

【0093】

固定された核酸分子は、必要に応じて、試薬含有溶液の集中したバースト(burst)を支持体上へ噴霧するデバイス（例えば、実質的に改変されない形態で用いられてもよい、任意の市販のインクジェットプリンター）を用いて作製されてもよい（その内容が全体として参照により本明細書に援用されるCastellino (1997) Genome Res. 7:943-976参照）。そのような方法は、現在、Incyte Pharmaceuticals社及びRosetta Biosystems社で実用化されており、後者は、「最小限に改変したEpsonインクジェットカートリッジ」を用いている（Epson America社；カリフォルニア州、トランス）。インクジェット堆積の方法は、圧電効果に依存しており、対象となる液体（この場合、オリゴヌクレオチド合成試薬）を含有する細管をアダプターが取り囲んでいる。アダプターを通して送られた電荷により、管とは異なる率でアダプターが膨張し、液体試薬の小滴が管からコーティングされたスライド又は他の支持体へ押出される。 40

【0094】

試薬は、各領域がアレイの特性を形成するように、支持体の別個の領域へ堆積されてもよい。所望の核酸配列は、本技術分野で公知の他の方法がそうであるように、全体として 50

堆積されるか、又は各位置で1滴ずつ合成されてもよい。試薬の分散角度が狭い場合、多くの特性を含むアレイを作製することが可能である。あるいは、より広い角度で核酸合成試薬を分散させるように、噴霧デバイスがより広範囲に焦点を合わせている場合、1回ごとに支持体のほぼ全体が被覆され、各メンバーが同じ配列を有するアレイが作製される(すなわち、アレイは単一の特性のみを有する)。

【0095】

ある態様によると、ヌクレオチド前駆体(dNTP/rNTP/rNDP)の結合のための時間、及び成長プライマー3'末端と共有結合を形成するのに費やされる時間に対して、異なる分布が考えられる。前駆体のパルスが短く維持されると、1つが結合することができ、共有結合形成反応の待機中に、試薬の継続的な流れが、ヌクレオチド除去緩衝液中に流入し得る。第一の結合反応は、ヌクレオチド濃度依存的である。Xヌクレオチドの平均長さを目的とする場合、分布は、ポアソン分布になると考えられ、 $X = 1$ ヌクレオチドの平均に対する理論的最大値は、収率37%となり、それ以外は、0、2個、3個などのヌクレオチドが組み込まれる。対照的に、本明細書に記載のパルスによる手法では、 $X = 1$ に対して37%よりも高い収率を得ることができる。

10

【0096】

ある態様では、標準的な合成及びシーケンシング手順に類似した、アレイに基づくフローセル技術が用いられる。出発TdTPライマーが、平坦な二酸化ケイ素(又は10ミクロン厚のポリマー層)と既知の位置で結合される。dNTPは、TdTP重合(又はスイッチング)速度に対して調節された既知の時間的及び空間的な幅の周期的な波で流動することになる。TdTP活性は、所望の位置でのpHを変化させることなど、光化学的又は電気化学的に制御される。オリゴヌクレオチドを作製するための位置の数は、90,000~5,000,000の範囲内であってもよい。

20

【0097】

高pH又は低pHなど、pHが制御されるように酸又は塩基を既知の位置で生成させるための例示的方法としては少なくとも1つの電極が挙げられ、電気化学的に生成された試薬の拡散に起因する電極間の化学的クロストークを防止するために緩衝溶液又は捕捉溶液と接触していることが好ましい。参照により本明細書に援用される米国特許第6,093,302号を参照されたい。あるいは、所望の位置でのpHの制御に、光で生成された酸が用いられてもよい。参照により本明細書に援用されるTian et al., Nature, Vol. 432, 23/30 December 2004, pp. 1050-1054を参照されたい。

30

【0098】

pH又は光を用いてポリメラーゼ活性又はdNTP/rNTP/rNDPのアクセスを調節する方法は、(1)(dNTPが充填された)リポソーム内部のpH感受性ウイルス粒子からのリパーゼ(又は他の膜溶解酵素)の放出(J. Clinical Microbiology 1988 May; 26(5) 804-807)及びJ. Control Release 2013 Nov 28; 172(1): 341-50参照;(2)350nm光で機能する、ニトロベンジル誘導体などの光ケージ化rNTP又はdNTP;(3)ヌクレオチドの小胞分泌又は他の分泌をもたらすロドプシン又はバクテリオオプシンで誘発されるシグナル伝達;(4)組み込みのための最適pH範囲及び可逆的活性が生じるpH範囲を有するポリメラーゼ;(5)(合成ペプチド又は改変tRNAによる新規な遺伝コードを介して)DNA又はRNAポリメラーゼに組み込まれたアゾベンゼンアミノ酸(ACS Nano 2014 May 27; 8(5): 4157-65参照)、並びに(6)Konerman et al., Nature (2013), Optical Control of mammalian Endogenous Transcriptionに記載の方法によって実現することができる。ある態様では、記載される(1)~(3)の方法を用いてヌクレオチドを放出するには、放出されたヌクレオチドを、最初のプライマーとポリメラーゼの最初の遭遇と任意の下流との間で除去又は捕捉するための方法が必要である。ポリメラーゼ調節法(4)~(6)には、そのような作業は必要ない。

40

【0099】

ある態様では、RNAポリメラーゼ又はDNA分解酵素若しくはRNA分解酵素の逆行、例えば、リボNTPのためのポリヌクレオチドホスホリラーゼ(PNP)、RNAリガ

50

ーゼ、circリガーゼなどのリガーゼも意図される。鋳型非依存的ポリメラーゼに加えて、(エラープローン又はバイパスポリメラーゼとしても知られる)半依存的ポリメラーゼも意図される。タンパク質、核酸、又は他のポリマー環が、ポリメラーゼと共有結合又は非共有結合で結合することができ、ポリメラーゼとプライマーとの会合が進行しやすくなるように、核酸又は他のポリマートラック (polymer track) を取り囲む。

【0100】

増幅

一般的に、「増幅」とは、プライマーを用いた酵素的合成の繰り返しによって、核酸分子のコピーを作製することを含む。「In situ」増幅は、溶液中ではなく、支持体又はビーズ上に配置された鋳型核酸分子を用いて増幅が行われることを示す。In situ増幅法は、米国特許第6,432,360号に記載されている。温度、鎖置換、及びブルーフリーディングなど、異なった特性によって、多様なポリメラーゼの選択が存在する。増幅は、等温であってもよく、Dean et al., Comprehensive human genome amplification using multiple displacement amplification, Proc. Natl. Acad. Sci. U.S.A., vol. 99, p. 5261-5266. 2002;さらにはDean et al., Rapid amplification of plasmid and phage DNA using phi29 DNA polymerase and multiply-primed rolling circle amplification, Genome Res., vol. 11, p. 1095-1099. 2001;さらにはAviel-Ronen et al., Large fragment Bst DNA polymerase for whole genome amplification of DNA formalin-fixed paraffin-embedded tissues, BMC Genomics, vol. 7, p. 312. 2006に記載の多置換増幅 (MDA) などの類似の適用であってもよい。増幅はまた、異なる温度計画 (temperature regiments) によるサイクルであってもよく、Mullis et al., Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. Cold Spring Harbor Symp. Quant. Biol., vol. 51, p. 263-273. 1986によって普及した従来のポリメラーゼ連鎖反応 (PCR) などである。ゲノム増幅により適用可能である変法は、Zhang et al., Whole genome amplification from a single cell: implications for genetic analysis, Proc. Natl. Acad. Sci. U.S.A., vol. 89, p. 5847-5851. 1992;及びTelenius et al., Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer, Genomics, vol. 13, p. 718-725. 1992に記載されている。他の方法としては、Mittra and Church, In situ localized amplification and contact replication of many individual DNA molecules, Nuc. Acid. Res., vol. 27, pages e34. 1999に記載されている Polony PCR; Shendure et al., Accurate multiplex polony sequencing of an evolved bacterial genome, Science, vol. 309, p. 1728-32. 2005;及びWilliams et al., Amplification of complex gene libraries by emulsion PCR, Nat. Methods, vol. 3, p. 545-550. 2006に記載されているエマルジョンPCR (ePCR) が挙げられる。いずれの増幅法も、事前に、逆転写工程と組み合わせ、RNAを増幅させてもよい。ある態様では、十分な感度を有するプローブ、レポーター、及び検出システムを用いて、記載されている鋳型とハイブリダイズしない核酸構造を用いた単一分子の検出を可能とすることができることから、増幅は、必要とは限らない。システムにおける感度を適合させる方法としては、励起源 (例えば、照明) 及び検出 (例えば、光検出器、光電子増幅器) の選択が挙げられる。シグナルレベルを適合させる方法としては、レポーターのスタッキングを可能にするプローブが挙げられ、高強度レポーター (例えば、量子ドット) を用いることも可能である。

【0101】

本開示において有用である増幅法は、核酸を、ハイブリダイゼーション及び鎖伸長を促進する条件下で核酸に特異的にハイブリダイズする1又は複数のプライマーと接触させることを含んでいてもよい。核酸を増幅するための例示的な方法としては、ポリメラーゼ連鎖反応法 (PCR) (例えばMullis et al. (1986) Cold Spring Harb. Symp. Quant. Biol. 51 Pt 1:263及びCleary et al. (2004) Nature Methods 1:241;並びに米国特許第4,683,195号及び同第4,683,202号参照)、アンカーPCR、RACE PCR、ライゲーション連鎖反応 (LCR) (例えばLandegran et al. (1988) Science

10

20

30

40

50

241:1077-1080 ; 及び Nakazawa et al. (1994) Proc. Natl. Acad. Sci. U.S.A. 91:360-364 参照)、自家持続配列複製法 (Guatelli et al. (1990) Proc. Natl. Acad. Sci. U.S.A. 87:1874)、転写増幅系 (Kwoh et al. (1989) Proc. Natl. Acad. Sci. U.S.A. 86:1173)、Q ベータレプリカーゼ (Lizardi et al. (1988) BioTechnology 6:1197)、再帰的 PCR (recursive PCR) (Jaffe et al. (2000) J. Biol. Chem. 275:2619 ; 及び Williams et al. (2002) J. Biol. Chem. 277:7790)、米国特許第 6,391,544 号、同第 6,365,375 号、同第 6,294,323 号、同第 6,261,797 号、同第 6,124,090 号、及び同第 5,612,199 号に記載の増幅方法、等温増幅 (例えば、ローリングサークル増幅 (RCA)、超分岐ローリングサークル増幅 (HRCA)、鎖置換増幅 (SDA)、ヘリカーゼ依存的増幅 (HDA)、PWGA)、又は当業者に公知の技術を用いた他の任意の核酸増幅法も挙げられる。

10

【0102】

例示の実施形態では、本明細書に開示される方法は、PCR 増幅を利用する。「ポリメラーゼ連鎖反応」又は「PCR」とは、DNA 相補鎖の同時プライマー伸長により、特定の DNA 配列を *in vitro* 増幅する反応を意味する。すなわち、PCR は、プライマー結合部位に挟まれた標的核酸の複数のコピー又は複製物を作製する反応であり、そのような反応は、以下の工程：(i) 標的核酸を変性する工程、(ii) プライマー結合部位にプライマーをアニーリングする工程、及び (iii) ヌクレオシド三リン酸の存在下で核酸ポリメラーゼによりプライマーを伸長する工程の 1 回又は複数回の繰り返しを含む。通常、反応は、サーマルサイクラー装置により、各工程に対して最適化された異なる温度のサイクルで行われる。特定の温度、各工程での継続時間、及び工程間の変化率は、当業者に公知の多くの因子に応じて異なり、例えば、文献 McPherson et al., editors, PCR: A Practical Approach 及び PCR2: A Practical Approach (それぞれ、IRL Press, Oxford, 1991 及び 1995) に例示されている。例えば、Taq DNA ポリメラーゼを用いる従来の PCR では、二本鎖標的核酸の変性は、90 超の温度で行われてよく、プライマーのアニーリングは、50 ~ 75 の範囲内の温度で行われてよく、プライマーの伸長は、68 ~ 78 の範囲内の温度で行われてもよい。

20

【0103】

「PCR」の用語には、これらに限定されないが、RT-PCR、リアルタイム PCR、ネステッド PCR、定量的 PCR、マルチプレックス PCR、アセンブリ PCR など、反応の派生型が包含される。反応容量は、数百ナノリットル、例えば 200 nL、から数百マイクロリットル、例えば 200 µL の範囲である。「逆転写 PCR」又は「RT-PCR」は、標的 RNA を相補的一本鎖 DNA へと変換する逆転写反応が先に行われ、次に DNA が増幅される PCR を意味し、例えば、Tecott et al.、米国特許第 5,168,038 号である。「リアルタイム PCR」は、反応生成物、すなわち増幅産物の量が、反応の進行に従ってモニタリングされる PCR を意味する。リアルタイム PCR には、反応生成物のモニタリングに用いられる検出化学が主に異なっている多くの形態が存在し、例えば、Gelfand et al.、米国特許第 5,210,015 号 (「Taqman」) ; Wittwer et al.、米国特許第 6,174,670 号及び同第 6,569,627 号 (インターカレート色素) ; Tyagi et al.、米国特許第 5,925,517 号 (モレキュラービーコン) である。リアルタイム PCR の検出化学は、Mackay et al., Nucleic Acids Research, 30: 1292-1305 (2002) に概説されている。「ネステッド PCR」は、二段階 PCR を意味し、第一の PCR の増幅産物が、新しいプライマーセットを用いる第二の PCR のための試料となり、このプライマーの少なくとも一つが第一の増幅産物の内側に結合する。本明細書で用いられる場合、ネステッド増幅反応に関する「一次プライマー」は、第一の増幅産物の生成に用いられるプライマーを意味し、「二次プライマー」は、第二の増幅産物又はネステッド増幅産物の生成に用いられる 1 又は複数のプライマーを意味する。「マルチプレックス PCR」は、複数の標的配列 (又は単一標的配列、及び 1 若しくは複数の参照配列) が同一の反応混合物中で同時に行われる PCR を意味し、例えば、Bernard et al. (1999) Anal. Biochem., 273:221-228 (2 色のリアルタイム PCR) である。通常、

30

40

50

増幅される配列のそれぞれに対して異なるプライマーセットが用いられる。「定量的PCR」は、サンプル又は試料中の1又は複数の特異的標的配列の存在量を測定するように設計されたPCRを意味する。以下の文献Freeman et al., *Biotechniques*, 26: 112-126 (1999); Becker-Andre et al., *Nucleic Acids Research*, 17:9437-9447 (1989); Zimmerman et al., *Biotechniques*, 21:268-279 (1996); Diviacco et al., *Gene*, 122:3013-3020 (1992); 及びBecker-Andre et al., *Nucleic Acids Research*, 17:9437-9446 (1989)などに例示されるように、定量的PCRのための技術は当業者に公知である。

【0104】

ローリングサークル増幅(RCA)(Zhong (2001) *Proc. Natl. Acad. Sci. USA* 98(7):3940-3945)は、連続複製であり、熱サイクルを必要としないことから、*polony* 増幅に対する代替法の代表例である。この増幅法は、1種のプライマー(又はニック)のみによって、時間に比例した速度で、元の環から1つの長いテールを成長させる。環状又は直鎖状核酸鋳型の等温増幅はまた、Tabor and Richardson(国際公開第00/41524号)に従って、核酸分子の酵素合成がオリゴヌクレオチドプライマーの非存在下で行われる方法を用いて行うことができる。反対側の鎖からの第二のプライマーも含まれている場合、高分岐鎖状構造が生成され、初期は、質量が時間に関して指数関数的に増加する($m = k * \exp(t)$ 、又は少なくとも $m = k t^2$)。

【0105】

本明細書に記載のRCAプロセスのモデリングにより、時間、化学物質、及び光学パターンの関数として、3Dアレイに層を構築する方法が示される。複製がスライドガラスの平坦面(又は他の面)上の均一層で開始する場合、重合反応は、すぐ上のnm厚さの層でのみ発生することができる。RCAにおけるポリメラーゼ(イータPol又はイータ様BstPol)の鎖置換活性は、鎖置換DNA合成を開始するのにニック又はプライマーを必要とする。RCAプライマーの一部が固定されていると、超分岐DNA産物は、空間的及び時間的に非常に安定となる。粗い(ミクロンスケール、5Hz)パターンを、メガピクセルマイクロミラー光学系で設定することができるが、より微細な細部(nm、250Hz)は、自走式又はRNA Pol イータDNA Pol 融合ステッパー(RNAPol-etaDNA Pol-fusion stepper)によって得られる。ナノスケールの記録は、時間の成分を有することから、必ずしも「冗長」ではなく、ミクロンスケールの光パターンによって制限されるものでもない。厚さ、ひいては記録容量は、それぞれ、位置決定及び記録に用いられる特定のNTPパルス及び/又はdNTPパルスの時空間的精度によって達成される。

【0106】

好ましい実施形態では、堆積された層は、核酸に結合した(又は核酸によって配置された)様々な化学物質を含んでよい。ある態様では、4つのdNTPのそれぞれに対するレドックス感受性フルオロフォア「側鎖」が開発されている。別の態様では、4つのdNTPのそれぞれの光感受性型が、当業者に公知の方法を用いて(Rob Mitra, unpublished data (2000))開発可能である。さらに別の態様では、金属結合基及びワイヤ(Braun (1998) *Nature* 391(6669):775-778)、量子ドット(Michler (2000) *Nature* 406(6799):968-70)、量子ワイヤ(Emiliani (2001) *J. Microsc.* 202(Pt 1):229-240)、磁気ドット(Co wburn (2000) *Science* 287(5457):1466-1468)、又は屈折ドット(refractive dot)(Yguerabide (1998) *Anal. Biochem.* 262(2):157-76)を、この方法によって集合させることができる。本発明の3Dアレイにより、シグナル一致(signal coincidence)及び/又は(神経回路における学習及び計算に類似する)トラフィックレベルに基づく、高速電子光学経路が提供される。RCAで見られる自然に超分岐した構造は、この方向に向かう第一ステップである。

【0107】

denovoポリマーは、ポリメラーゼ増幅をして、又はポリメラーゼ増幅をせずに、記憶及び読み取りが可能である。増幅は、熱サイクルによるか、又は等温によるものであってもよい。増幅産物は、現行の化学合成において都合が良いように短くてもよく(100~200mer)、又はポリメラーゼを用いて実現可能であり得る1Mbp以下であ

10

20

30

40

50

ってもよい。

【0108】

シーケンシング

取り込まれたヌクレオチドの種類は、a) dNTP 溶液の周期的パターンにおけるその時点で存在する特定の dNTP (又は rNTP 又は他のモノマー種) と一致する光パルスの交差、b) 「ケージ化」(すなわち、光活性化可能又は光不活性化可能な) dNTP、rNTP、又はカチオン、c) 塩基特異的で光調節された立体選択性又は立体構造選択性 (Hoppmann C, Schmieder P, Heinrich N, Beyermann M. (2011) *Chembiochem.* 12(17):2555-9. doi: 10.1002/cbic.201100578. Epub 2011 Oct 13. Photoswitchable click amino acids: light control of conformation and bioactivity参照) により決定され得る。ポリ(A)ポリメラーゼは、他の rNTP と比較した ATP に対する特異性が、(架橋又は未架橋のアゾベンゼンのように) 光感受性アミノ酸結合により模倣可能な立体構造の変化に起因するため、特に有用である。

10

【0109】

本明細書に記載の方法は、大量のデータ(数十億ビット)を作り出すことができる。したがって、Mitra (1999) *Nucleic Acids Res.* 27(24):e34; pp.1-6に開示されるような、これらの核酸分子をシーケンシングするハイスループットな方法が有用である。好ましい実施形態では、ハイスループットな方法が、100bp未満の長さを有するPCR増幅産物又は他の核酸分子と共に用いられる。他の好ましい実施形態では、100bp、110bp、120bp、130bp、140bp、150bp、160bp、170bp、180bp、190bp、200bp、250bp、300bp、350bp、400bp、450bp、500bp、550bp、600bp、650bp、700bp、750bp、800bp、850bp、900bp、950bp、1000bp、又はそれ以上のPCR増幅産物が用いられてもよい。

20

【0110】

ローリングサークル増幅(RCA)(Zhong (2001) *Proc. Natl. Acad. Sci. USA* 98(7):3940-3945)は、連続複製であり、熱サイクルを必要としないことから、polony増幅に対する代替法の代表例である。この増幅法は、1種のプライマー(又はニック)のみによって、時間に比例した速度で、元の環から1つの長いテールを成長させる。環状又は直鎖状核酸鑄型の等温増幅はまた、Tabor and Richardson(国際公開第00/41524号)に従って、核酸分子の酵素合成がオリゴヌクレオチドプライマーの非存在下で行われる方法を用いて行うことができる。

30

【0111】

本開示において有用なシーケンシング法としては、Shendure et al., *Accurate multiplex polony sequencing of an evolved bacterial genome*, *Science*, vol. 309, p. 1728-32. 2005; Drmanac et al., *Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays*, *Science*, vol. 327, p. 78-81. 2009; McKernan et al., *Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two-base encoding*, *Genome Res.*, vol. 19, p. 1527-41. 2009; Rodrigue et al., *Unlocking short read sequencing for metagenomics*, *PLoS One*, vol. 28, e11840. 2010; Rothberg et al., *An integrated semiconductor device enabling non-optical genome sequencing*, *Nature*, vol. 475, p. 348-352. 2011; Margulies et al., *Genome sequencing in microfabricated high-density picolitre reactors*, *Nature*, vol. 437, p. 376-380. 2005; Rasko et al. *Origins of the E. coli strain causing an outbreak of hemolytic-uremic syndrome in Germany*, *N. Engl. J. Med.*, Epub. 2011; Hutter et al., *Labeled nucleoside triphosphates with reversibly terminating aminoalkoxyl groups*, *Nucleos. Nucleot. Nucl.*, vol. 92, p. 879-895. 2010; Seo et al., *Four-color DNA sequencing by synthesis on a chip using photocleavable fluorescent nucleotides*, *Proc. Natl. Acad. Sci. USA.*, Vol. 102, P. 5926-5931 (2005); Olejnik et al.; *Photocleavable biotin derivatives:*

40

50

a versatile approach for the isolation of biomolecules, Proc. Natl. Acad. Sci. U.S.A., vol. 92, p. 7590-7594. 1995; 米国特許第 5,750,34号、米国特許出願公開第 2009/0062129号、及び同第 2009/0191553号が挙げられる。

【0112】

本開示によるシーケンシングプライマーは、標的ポリヌクレオチドの既知の結合領域に結合可能であり、本開示のオリゴヌクレオチドプローブのライゲーションを促進可能であるものである。シーケンシングプライマーは、例えば、DNA Works、又は Gene 2 Oligo などのコンピュータプログラムの補助で設計されてもよい。結合領域の長さは様々であり得るが、シーケンシングプライマーとハイブリダイズするのに十分な長さが必要である。標的ポリヌクレオチドは、複数の異なった結合領域を有していてもよく、それによって、標的ポリヌクレオチドの様々な部分をシーケンシングすることが可能となる。シーケンシングプライマーは、連続的なライゲーションサイクルの間、ハイブリダイズした状態が維持されるよう、高度に安定な二重鎖を形成するように選択される。シーケンシングプライマーは、ライゲーションが 5 から 3 方向若しくは 3 から 5 方向のいずれか、又は両方向に進行可能であるように選択されてもよい。シーケンシングプライマーは、ハイブリダイゼーション効率を高めるために、又は安定性を向上するために、又はどちらか一方の端部からの伸長を阻害するために、修飾されたヌクレオチド又は修飾された結合を含んでいてもよい。

10

【0113】

ある態様では、一本鎖 DNA 鋳型 (ssDNA) が RCA によって作製されてシーケンシングプライマーと共に用いられる。あるいは、一本鎖鋳型は、エマルジョン中のビーズ又はナノ粒子と結合され、ePCR によって増幅される。この結果、単一の増幅 ssDNA 鋳型を有するクローンビーズが得られる。

20

【0114】

いくつかの鋳型ヌクレオチド配列を並行して同定する目的で、鋳型は、pH 7.4 の PBS 緩衝液で希釈され、ビオチン ストレプトアビジン、アジド アルキル (例えば、クリックケミストリー)、NHS エステル、又はシラン処理 (例えば、アルデヒドシラン、エポキシシラン、アミノシラン) などの様々な結合方法を用いて、パターン化基材又は非パターン化基材に結合される。ある態様では、コロニー (colony) が SiO₂ 固体表面などのパターン化表面に結合され、1% アミノシラン (体積 / 体積) で処理されて、特定の時間 (通常、5 分間 ~ 2 時間) 相互作用が起こされる。続いて、Wash 1 緩衝液を用いて、未結合の鋳型がすべて洗い流される。

30

【0115】

次に、シーケンシングプライマーが作製され、シーケンシングプライマーハイブリダイゼーション部位にハイブリダイズされる。ある態様では、鋳型の既知の配列にハイブリダイズすることができるシーケンシングプライマーが作製されてもよい。あるいは、鋳型作製の過程で、既知の核酸配列を有するアダプターが、当業者に公知であり本明細書に記載の方法によるライゲーション、増幅、転位、又は組み換えによって、未知の核酸配列に付加される。あるいは、特定のレベルの縮重を有するシーケンシングプライマーを用いて、鋳型に沿った特定の位置にハイブリダイズさせてもよい。ある態様では、プライマー縮重を用いて、プライマーを鋳型に沿って準ランダムにハイブリダイズさせることが可能となる。プライマー縮重は、プライマーが鋳型の長さに沿って特定の間隔でハイブリダイズすることを促進するように、当業者に公知の統計的方法に基づいて選択される。この態様では、100 塩基ごと、200 塩基ごと、2000 塩基ごと、100,000 塩基ごとなど、N 塩基ごとの結合を促進する特定の縮重を有するプライマーを設計することができる。鋳型の長さに沿ったプライマーの結合は、プライマーの設計、及びプライマー設計が鋳型の長さに沿っておよそ N 塩基ごとに結合することになる統計的尤度に基づいている。シーケンシングプライマー P1 はライゲーションによって伸長されるため、シーケンシングプライマーの末端基は、通常、DNA リガーゼによってオリゴヌクレオチドプローブと容易に共有結合されるように合成される。ライゲーションが、シーケンシング

40

50

プライマーの 5' 末端とオリゴヌクレオチドプローブの 3' 末端との間で発生する場合、シーケンシングプライマー上にはリン酸基 (5' P O 4') が存在し、オリゴヌクレオチドプローブ上にはヒドロキシル基 (3' O H) が存在しなければならず、逆もまた同様である。シーケンシングプライマーをシーケンシングプライマーハイブリダイゼーション部位にハイブリダイズさせるために、5 × S S P E 緩衝液で希釈した 1 μ M のシーケンシングプライマーが用いられる。次に、混合物を、室温よりも高い温度で数分間インキュベートして、適切なアニーリングを促進する (通常、25 ~ 55 の温度で 1 ~ 5 分間)。

【 0 1 1 6 】

HTML ファイルの DNA セグメントへのエンコード

10

ある態様では、文字列などの情報を、(埋め込み jpg 画像を有する) HTML フォーマットへと変換し、次にビット形式で読み込むことができる。個々のビットは、0 については A 又は C に、1 については T 又は G へと変換される。ビットストリームのアドレスは、19 ビット長であり、0000000000000000001 から始まるなど連続的に番号付与されている。Bits2DNA.pl として識別される以下のプログラムは、HTML ファイルを DNA セグメントへとエンコードするために用いられる。

【 0 1 1 7 】

【表 1】

```
# cd "\Perl\gmc\Bin_DNA"
```

```
# \Perl\bin\perl Bits2DNA.pl GMC Jul-2011 & 27-May-2012
```

20

```
# docstore.mik.ua/orelly/perl/cookbook/ch02_05.htm (bin) ch01_05.htm (char)
```

```
# http://perldoc.perl.org/functions/pack.html rand.html
```

```
# Each oligo is L(19)+8N(12)= 115 bp, long flanked by 22-mer amplification primers.
```

```
# DNA Encoded Artifacts Registry (DEAR) to coordinate global standards.
```

```
open IN,"in.html"; open OUT,">Bits2DNA.txt"; binmode IN;
```

30

```
$t{"0"}="a"; $t{"1"}="G"; # lowercase a,c = zero bit.
```

```
$t{"a"}="c"; $t{"G"}="T"; $t{"c"}="a"; $t{"T"}="G";
```

```
$u1=""; $u2=""; $u3=""; # Initialize; keep homopolymer runs < 4
```

```
$N=12; # Length of segment in bytes (not including segment number)
```

```
$L=19; # 2^19 = 524,288 = max number of oligos L=00010011
```

```
$seed=2; srand($seed); # remove this line to get a random seed
```

```
print int2bp(262144)," ",int2bp(262145);
```

40

```
$f="CTACACGACGCTCTCCGATCT"; # forward 'universal' sequencing &
amplification primer
```

```
$r="AGATCGGAAGAGCGGTTCAGCA"; # reverse 22-mer primer
```

【 0 1 1 8 】

【表 2】

```

$N=0; print OUT $f,int2bp(0),"";    ###
while (read (IN, $text, 65536)) {
    @ascii_num = unpack("C*", $text);
    foreach $val (@ascii_num) {
        print OUT byt2bp($val);      ###
        $n++;
        if($n%$N==0){
            print OUT $r, "\n", $f,int2bp($n/$N),""; ###
        } # N bases per output line
    } # each byte
} # 65 Kbytes
for ($k=$n%$N; $k<$N; $k++){
    print OUT byt2bp(int(rand(256))); ###
} # pad last data line to keep all oligos same size.
print OUT "$r\n";    ###

```

10

20

【 0 1 1 9 】

【表 3】

```
sub byt2bp { # convert rightmost 8 bits (MSB first byte) to 8 bp
```

```
  my $b = unpack("B32", pack("N", shift));
```

```
  $p="";
```

```
  for ($i=24; $i<=31; $i++){
```

```
    $x=substr($b,$i,1); # bits 24 to 31 inclusive
```

```
    $u=St{$x};
```

```
    if(rand(2)<1){$u=St{$u};} # pick synonym a=c; G=T
```

```
    if(($u eq $u1) && ($u eq $u2) && ($u eq $u3)){ $u=St{$u};}
```

```
    $u1=$u2; $u2=$u3; $u3=$u; # Shift previous base string
```

```
    $p = $p.$u;
```

```
  }
```

```
  return $p;
```

```
}
```

```
sub int2bp { # convert rightmost $L bits of 32 bit integers to $L bp
```

```
  my $b = unpack("B32", pack("N", shift));
```

```
  $p="";
```

```
  for ($i=31; $i>=32-$L; $i--){
```

```
    $x=substr($b,$i,1); # bits 31 to $L
```

```
    $u=St{$x};
```

```
    if(rand(2)<1){$u=St{$u};} # pick synonym a=c; G=T
```

```
    if(($u eq $u1) && ($u eq $u2) && ($u eq $u3)){ $u=St{$u};}
```

```
    $u1=$u2; $u2=$u3; $u3=$u; # Shift previous base string
```

```
    $p = $p.$u;
```

```
  }
```

```
  return $p;
```

```
}
```

```
【 0 1 2 0 】
```

10

20

30

40

【表 4】
buildConsensus.py

```
import sys
```

```
#builds consensus sequence from individual base counts
```

```
def getConsensus(finalbuckets):
```

```
    sequence = "
```

```
    for i in range(len(finalbuckets)):
```

```
        letterindex = finalbuckets[i].index(max(finalbuckets[i]))
```

```
        if letterindex == 0:
```

```
            sequence += 'A'
```

```
        elif letterindex == 1:
```

```
            sequence += 'C'
```

```
        elif letterindex == 2:
```

```
            sequence += 'G'
```

```
        elif letterindex == 3:
```

```
            sequence += 'T'
```

```
    return sequence
```

```
oligolength = 115
```

```
currentbarcode = "
```

```
#initialize vector to building consensus
```

```
buckets = [[0 for col in range(4)] for row in range(oligolength)]
```

```
【 0 1 2 1 】
```

10

20

30

【表 5】

```
for line in sys.stdin:
```

```
    splitline = line.split()
```

```
    count = int(splitline[0])
```

```
    barcode = splitline[1]
```

```
    sequence = splitline[2]
```

```
    if not barcode == currentbarcode:
```

```
        if not currentbarcode == ":
```

```
            print getConsensus(buckets)
```

```
        buckets = [[0 for col in range(4)] for row in range(oligolength)]
```

```
        currentbarcode = barcode
```

```
    for i in range(oligolength):
```

```
        if sequence[i] == 'A':
```

```
            buckets[i][0] += count
```

```
        elif sequence[i] == 'C':
```

```
            buckets[i][1] += count
```

```
        elif sequence[i] == 'G':
```

```
            buckets[i][2] += count
```

```
        elif sequence[i] == 'T':
```

```
            buckets[i][3] += count
```

```
#print final consensus
```

```
print getConsensus(buckets)
```

【 0 1 2 2 】

本明細書で開示される方法の実行には、従来の生物学的方法、ソフトウェア、コンピュータ、及びコンピュータシステムが用いられてもよい。したがって、本明細書に記載の方法は、全体が、又は部分的に、コンピュータで実行される方法であってもよい。本開示の方法で用いられるコンピュータソフトウェアは、本発明の方法の論理ステップを行うためのコンピュータ実行可能な命令を有するコンピュータ読み取り可能媒体を含む。適切なコンピュータ読み取り可能媒体としては、これらに限定されないが、フロッピーディスク、CD ROM / DVD / DVD ROM、ハードディスクドライブ、フラッシュメモリ、ROM / RAM、磁気テープ、及び開発され得る他の媒体が挙げられる。コンピュータ実行可能な命令は、適切なコンピュータ言語、又は複数のコンピュータ言語の組み合わせで書かれていてもよい。本明細書に記載の方法はまた、文字列又は画像のバイナリコードへの翻訳、バイナリコードを表す核酸配列の設計、核酸配列からのシーケンシングデータの解析、核酸配列データのバイナリコードへの翻訳、及びバイナリコードの文字列又は画像への翻訳を含む様々な目的のために、様々な市販のコンピュータ、及びコンピュータプログラム製品、及びソフトウェアを利用してよい。

10

20

30

40

50

【 0 1 2 3 】

特定の方法の実施形態

本開示の実施形態には、バイナリエンコードされたポリマーを作製する方法であって、(i) 第一のモノマーペアにおける第一のモノマー又は第二のモノマーのいずれかのモノマー 1 個又は複数個、又は (i i) 第二のモノマーペアにおける第一のモノマー又は第二のモノマーのいずれかのモノマー 1 個又は複数個の伸長産物によって、成長ポリマー鎖を繰り返し伸長させる工程であって、ここで、前記伸長産物が、文字列、画像、ビデオ、又は音声のフォーマットから翻訳されたビットストリームに対応するバイナリ情報ビットを表し、前記第一のモノマーペアの前記第一のモノマー及び第二のモノマーがそれぞれ第一のバイナリ情報ビットを表し、前記第二のモノマーペアの前記第一のモノマー及び第二のモノマーがそれぞれ第二のバイナリ情報ビットを表す工程；及び前記伸長産物が同じバイナリ情報ビットを表し、且つ直接連続して存在する場合に、所与のモノマーペアの前記第一のモノマーと前記第二のモノマーとを交互に並べる工程を含み、前記バイナリエンコードされたポリマーは、前記文字列、画像、ビデオ、又は音声のフォーマットをコードする、バイナリエンコードされたポリマーを作製する方法を提供することが含まれる。ある態様では、前記ポリマーは、核酸である。ある態様では、前記第一のモノマーペアにおける前記第一のモノマー又は前記第二のモノマーは、ヌクレオチドである。ある態様では、前記第二のモノマーペアにおける前記第一のモノマー又は前記第二のモノマーは、ヌクレオチドである。ある態様では、前記第一のモノマーペアは、アデニン (A)、及びチミン (T) 又はウラシル (U) を含む。ある態様では、前記第二のモノマーペアは、シトシン (C) 及びグアニン (G) を含む。ある態様では、前記伸長産物は、酵素及び選択されたモノマーを用い、前記選択されたモノマーの付加を触媒する条件下で形成される。ある態様では、前記伸長産物は、ポリメラーゼ及び選択されたモノマーを用いて、前記選択されたモノマーの付加を触媒する条件下で形成される。ある態様では、前記伸長産物は、鑄型非依存的ポリメラーゼ及び選択されたモノマーを用いて、前記選択されたモノマーの付加を触媒する条件下で形成される。ある態様では、前記成長ポリマー鎖は、基材に結合されている。ある態様では、工程 (i) 及び工程 (i i) から形成された複数の成長ポリマー鎖が提供される。ある態様では、工程 (i) 及び工程 (i i) から形成された複数の成長ポリマー鎖が提供され、前記複数の成長ポリマー鎖は、基材に結合されている。ある態様では、前記第一のモノマーペアにおける前記第一のモノマー又は前記第二のモノマーは、天然ヌクレオチドである。ある態様では、前記第二のモノマーペアにおける前記第一のモノマー又は前記第二のモノマーは、天然ヌクレオチドである。ある態様では、前記第一及び第二のモノマーペアは天然ヌクレオチドを含み、前記伸長産物は、単一のヌクレオチド又は複数のヌクレオチドを付加するのに十分な条件下で天然ヌクレオチドの付加を触媒することによって作製される。ある態様では、前記第一のモノマーペア及び前記第二のモノマーペアは天然ヌクレオチドを含み、前記伸長産物は、前記基材上の 1 又は複数の位置で、ポリメラーゼ及び選択されたヌクレオチドの投与とヌクレオチド欠乏緩衝液の投与とを交互に行って、前記ヌクレオチドを付加するのに十分な条件下で天然ヌクレオチドの付加を触媒することによって作製される。

【 0 1 2 4 】

ある態様では、バイナリエンコードされたポリマーを作製する方法であって、(i) 第一のモノマーペアにおける第一のモノマー又は第二のモノマーのいずれかのモノマー 1 個又複数個、又は (i i) 第二のモノマーペアにおける第一のモノマー又は第二のモノマーのいずれかモノマー 1 個又複数個の伸長産物によって成長ポリマー鎖を繰り返し伸長させる工程であって、ここで、前記伸長産物が、文字列、画像、ビデオ、又は音声のフォーマットから翻訳されたビットストリームに対応するバイナリ情報ビットを表し、前記第一のモノマーペアの前記第一のモノマー及び第二のモノマーがそれぞれ第一のバイナリ情報ビットを表し、前記第二のモノマーペアの前記第一のモノマー及び第二のモノマーがそれぞれ第二のバイナリ情報ビットを表す工程；及び前記伸長産物が同じバイナリ情報ビットを表し、且つ直接連続して存在する場合に、所与のモノマーペアの前記第一のモノマーと前

10

20

30

40

50

記第二のモノマーとを交互に並べる工程を含み、前記伸長産物は、前記第一のモノマーペアにおける前記第一のモノマー又は前記第二のモノマーのいずれかの少なくとも1つのホモポリマー、又は前記第二のモノマーペアにおける前記第一のモノマー又は前記第二のモノマーのいずれかの少なくとも1つのホモポリマーを含み、且つ、前記バイナリエンコードされたポリマーは、前記文字列、画像、ビデオ、又は音声のフォーマットをコードする、バイナリエンコードされたポリマーを作製する方法が提供される。

【0125】

ある態様では、バイナリエンコードされた核酸を、核酸配列から文字列、画像、ビデオ、又は音声のフォーマットを表すバイナリ情報ビットの列へと翻訳する方法であって、アデニン、及びチミン又はウラシルが、第一のバイナリ情報ビットを表し、シトシン及びグアニンが、第二のバイナリ情報ビットを表し、前記方法は、前記核酸配列を読み取る工程、及び各アデニン、又は連続する場合は複数のアデニンに、前記第一のバイナリ情報ビットを割り当てる工程、各チミン、又は連続する場合は複数のチミンに、前記第一のバイナリ情報ビットを割り当てる工程、各ウラシル、又は連続する場合は複数のウラシルに、前記第一のバイナリ情報ビットを割り当てる工程、各シトシン、又は連続する場合は複数のシトシンに、前記第二のバイナリ情報ビットを割り当てる工程、及び各グアニン、又は連続する場合は複数のグアニンに、前記第二のバイナリ情報ビットを割り当てる工程を含み、前記核酸配列は、連続する2つ以上のアデニン、連続する2つ以上のチミン、連続する2つ以上のウラシル、連続する2つ以上のシトシン、又は連続する2つ以上のグアニンのうちの少なくとも1つを含む、方法が提供される。

【0126】

ある態様では、文字列、画像、ビデオ、又は音声のフォーマットをエンコード及びデコードする方法であって、前記文字列、画像、ビデオ、又は音声のフォーマットをビットストリームの複数のビット列へと変換する工程、前記ビットストリームの前記複数のビット列に対応する核酸配列を、アデニン又はチミンを第一のバイナリ情報ビットに、及びシトシン又はグアニンを第二のバイナリ情報ビットに割り当てることで設計する工程であって、ここで、アデニン又はチミンの割り当ては、同じバイナリ情報ビットが直接連続して存在する場合は、交互に並べ、シトシン又はグアニンの割り当ては、同じバイナリ情報ビットが直接連続して存在する場合は、交互に並べる工程；前記核酸配列を合成する工程；前記合成された核酸配列を保存する工程；前記合成された核酸配列を読み取る工程；及び前記合成された核酸配列を、前記第一のバイナリ情報ビットをアデニン又はチミンに割り当て、前記第二のバイナリ情報ビットをシトシン又はグアニンに割り当てることによって、前記ビットストリームの前記複数のビット列にデコードする工程を含む方法が提供される。ある態様では、前記合成された核酸配列は、アデニン、チミン、シトシン、又はグアニンの少なくとも1つのホモポリマーを含み、前記合成された核酸配列の前記デコードは、前記第一のバイナリ情報ビットをアデニン又はチミンのホモポリマーに割り当てること、及び前記第二のバイナリ情報ビットをシトシン又はグアニンのホモポリマーに割り当てることを含む。

【0127】

ある態様では、文字列、画像、ビデオ、又は音声のフォーマットをコードするビットストリームの複数のビット列を表す核酸配列群を用いて基材上に情報を記憶する方法であって、一本鎖核酸イニシエータ配列が結合した基材をアレイ上の領域に提供する工程；1又は複数の位置を、 Co^{2+} 、 Mn^{2+} 、 Zn^{2+} 、及び Mg^{2+} のうちの1つ又は複数、鑄型非依存的ポリメラーゼ、並びに選択された天然ヌクレオチドと接触させる工程；前記1又は複数の位置で、標的一本鎖核酸イニシエータ配列の3'ヒドロキシル末端への前記選択された天然ヌクレオチドの付加を触媒する工程；及び選択された天然ヌクレオチドの付加を触媒する工程を繰り返して、前記基材の既知の位置に複数の所定の配列を生成する工程であって、ここで、前記複数の所定の配列は、前記文字列、画像、ビデオ、又は音声のフォーマットをコードする前記ビットストリームの前記複数のビット列を表す工程を含む方法が提供される。ある態様では、前記触媒する工程及び前記繰り返す工程は、前記基材上

の複数の位置で順に行われて、前記基材の既知の位置で複数の所定の配列が生成される。ある態様では、前記触媒する工程及び前記繰り返す工程は、前記基材上の複数の位置で同時に行われて、複数の所定の配列のそれぞれが前記基材の対応する既知の位置で生成され、対応する既知の位置で所定の配列のアレイが生成される。ある態様では、前記複数の所定の配列の1又は複数の、シーケンシングされ、前記配列は、バイナリビット情報に翻訳され、前記バイナリビット情報は次に、前記文字列、画像、ビデオ、又は音声のフォーマットに翻訳される。ある態様では、前記基材は、少なくとも 10^2 個、 10^3 個、 10^4 個、 10^5 個、 10^6 個、 10^7 個、 10^8 個、 10^9 個、 10^{10} 個の所定の配列を、各々対応する既知領域に含む。ある態様では、前記所定の配列の長さは、100ヌクレオチド超、500ヌクレオチド超、又は1000ヌクレオチド超である。ある態様では、前記鑄型非依存的ポリメラーゼは、末端デオキシヌクレオチジルトランスフェラーゼである。ある態様では、前記天然ヌクレオチドの前記伸長産物は、前記天然ヌクレオチドの反応時間を限定することによって生成され、ヌクレオチド除去緩衝液が添加されて前記天然ヌクレオチドが除去され、それによって前記天然ヌクレオチドの反応時間が限定されるか、又は前記基材の表面を横切る天然ヌクレオチドのパルス流速が、特定の所望の位置での前記天然ヌクレオチドの反応時間を限定する。

10

【0128】

ある態様では、バイナリエンコードされた核酸を、核酸配列からバイナリ情報ビットの列へと翻訳する方法であって、アデニン、及びチミン又はウラシルが、第一のバイナリ情報ビットを表し、シトシン及びグアニンが、第二のバイナリ情報ビットを表し、前記方法は、前記核酸配列を読み取る工程；並びに各アデニン、又は連続する場合は複数のアデニンに、前記第一のバイナリ情報ビットを割り当てる工程、各チミン、又は連続する場合は複数のチミンに、前記第一のバイナリ情報ビットを割り当てる工程、各ウラシル、又は連続する場合は複数のウラシルに、前記第一のバイナリ情報ビットを割り当てる工程、各シトシン、又は連続する場合は複数のシトシンに、前記第二のバイナリ情報ビットを割り当てる工程、及び各グアニンに、又は連続する場合は2つ以上のグアニンに前記第二のバイナリ情報ビットを割り当てる工程を含む方法が提供される。

20

【0129】

ある態様では、記憶された情報を有し、基材、及び前記基材上に配置された複数の核酸配列を含む情報記憶デバイスであって、前記複数の核酸配列は、前記記憶された情報に対応するバイナリ情報ビットの列をコードし、アデニン又はアデニンの列、チミン又はチミンの列、及びウラシル又はウラシルの列は、第一のバイナリ情報ビットを表し、且つシトシン又はシトシンの列、及びグアニン又はグアニンの列は、第二のバイナリ情報ビットを表す、情報記憶デバイスが提供される。

30

【0130】

ある態様では、特定のフォーマットの情報を、各々が対応するビットバーコードを有するビットストリームの複数のビット列へと変換する工程；前記複数のビット列を1塩基あたり1ビットのエンコードを用いて複数の対応するオリゴヌクレオチド配列へと変換する工程；前記複数の対応するオリゴヌクレオチド配列を、複数の反応位置を有する基材の表面を横切るように試薬及び洗浄液をパルスし、同期させることによって合成する工程；並びに前記合成された複数の対応するオリゴヌクレオチド配列を保存する工程を含む、ヌクレオチドを用いて情報を記憶する方法が提供される。ある態様では、前記オリゴヌクレオチド配列は、データブロック配列、前記ビットストリーム中における前記データブロックの位置を指定するアドレス配列、又は前記オリゴヌクレオチドの両端に位置する増幅及びシーケンシングのための隣接共通配列のうちの1つ又は複数又は全てを含む。

40

【0131】

ある態様では、特定のフォーマットの情報のビット列をコードする複数のオリゴヌクレオチド配列を増幅する工程；前記増幅されたオリゴヌクレオチド配列をシーケンシングする工程；ホモポリマーの並びを単一のヌクレオチドとして解釈することによって前記オリゴヌクレオチド配列をビット列へと変換する工程；及び前記ビット列を前記特定のフォ

50

フォーマットの情報へと変換する工程を含む、特定のフォーマットの情報を、前記特定のフォーマットの情報のビット列をコードする複数の合成されたオリゴヌクレオチド配列から回収する方法が提供される。ある態様では、前記オリゴヌクレオチド配列は、データブロック配列、前記ビットストリーム中における前記データブロックの位置を指定するアドレス配列、又は前記オリゴヌクレオチドの両端に位置する増幅及びシーケンシングのための隣接共通配列のうちの1つ又は複数又は全てを含む。

【0132】

ある態様では、特定のフォーマットの情報のビット列をコードする複数のオリゴヌクレオチド配列を増幅する工程；前記増幅されたオリゴヌクレオチド配列をシーケンシングする工程；ホモポリマーの並びを単一のヌクレオチドとして解釈することによって前記オリゴヌクレオチド配列をビット列へと変換する工程；前記ビット列を特定のフォーマットの情報へと変換する工程；及び前記特定のフォーマットの情報を出力する工程を含む、特定のフォーマットの情報に、前記特定のフォーマットの情報のビット列をコードする複数の合成されたオリゴヌクレオチド配列からアクセスする方法が提供される。ある態様では、前記オリゴヌクレオチド配列は、データブロック配列、前記ビットストリーム中における前記データブロックの位置を指定するアドレス配列、又は前記オリゴヌクレオチドの両端に位置する増幅及びシーケンシングのための隣接共通配列のうちの1つ又は複数又は全てを含む。

10

【0133】

ある態様では、特定のフォーマットの情報をビットストリームへと変換する工程；ビット列を対応するオリゴヌクレオチド配列へとエンコードする工程；前記オリゴヌクレオチド配列を、複数の反応位置を有する基材の表面を横切るように試薬及び洗浄液をパルスし、同期させることによって合成する工程；前記オリゴヌクレオチド配列をシーケンシングする工程；前記オリゴヌクレオチド配列を、ホモポリマーの並びを単一のヌクレオチドとして解釈することによってビット列へとデコードする工程；前記ビット列をビットストリームへとアSEMBルする工程；及び前記ビットストリームを前記特定のフォーマットの情報へと変換する工程を含む、ヌクレオチドを用いて情報を記憶する方法が提供される。ある態様では、前記オリゴヌクレオチド配列は、データブロック配列、前記ビットストリーム中における前記データブロックの位置を指定するアドレス配列、又は前記オリゴヌクレオチドの両端に位置する増幅及びシーケンシングのための隣接共通配列のうちの1つ又は複数又は全てを含む。

20

30

【0134】

ある態様では、第一のフォーマットの情報を第一のビットストリームへと変換する工程；第一のビット列を対応するオリゴヌクレオチド配列へとエンコードする工程；前記オリゴヌクレオチド配列を、複数の反応位置を有する基材の表面を横切るように試薬及び洗浄液をパルスし、同期させることによって合成する工程；前記オリゴヌクレオチド配列をシーケンシングする工程；前記オリゴヌクレオチド配列を、ホモポリマーの並びを単一のヌクレオチドとして解釈することによって第二のビット列へとデコードする工程；前記第二のビット列を第二のビットストリームへとアSEMBルする工程；及び前記第二のビットストリームを第二のフォーマットの情報へと変換する工程を含む、ヌクレオチドを用いて情報を記憶する方法が提供される。ある態様では、前記オリゴヌクレオチド配列は、データブロック配列、前記ビットストリーム中における前記データブロックの位置を指定するアドレス配列、又は前記オリゴヌクレオチドの両端に位置する増幅及びシーケンシングのための隣接共通配列のうちの1つ又は複数又は全てを含む。

40

【0135】

本出願全体を通して引用されるすべての参考文献、特許、及び公開特許出願の内容は、その全体があらゆる目的のために参照により本明細書に援用される。

【0136】

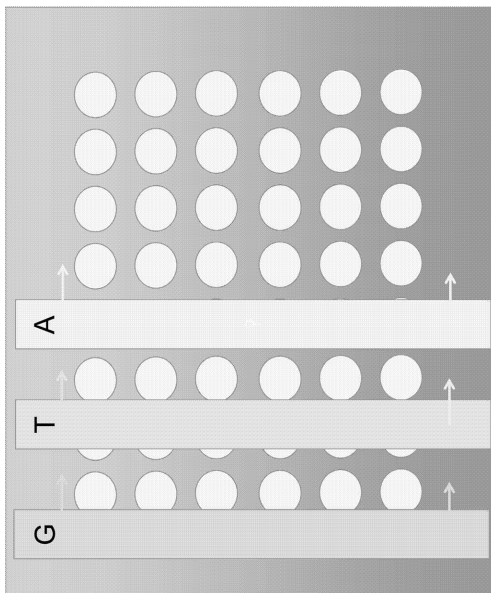
他の実施形態

他の実施形態は、当業者であれば明らかであろう。上記の記載は、明確化の目的で提供

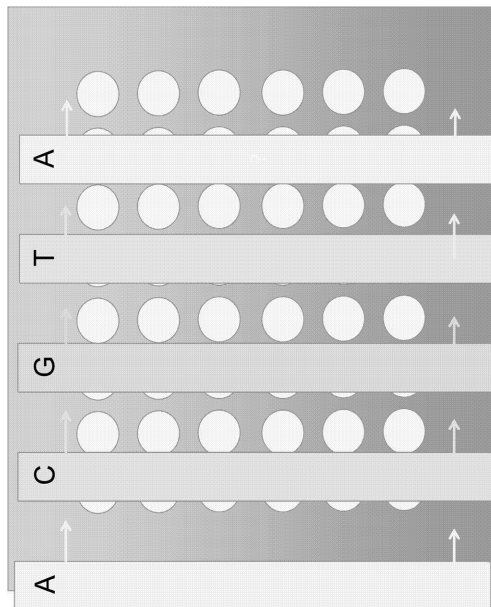
50

されるものであり、単なる例示であると理解されるべきである。本発明の趣旨及び範囲は、上記の例に限定されるものではなく、以下の請求項に包含される。上記で引用されたすべての文献及び特許出願は、個々の文献又は特許出願が参照により取り込まれることが具体的かつ個々に記された場合と同程度に、その内容が全体として参照により本明細書中に取り込まれる。

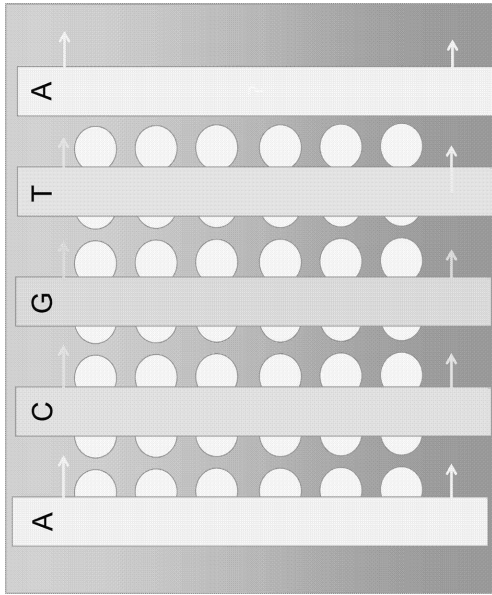
【図1】



【図2】



【図 3】



【配列表】

0006920275000001.app

フロントページの続き

審査官 野村 英雄

- (56)参考文献 国際公開第2014/014991(WO, A1)
特表2006-522356(JP, A)
国際公開第2003/025123(WO, A1)
米国特許第06312911(US, B1)
Jonathan P.L. COX, Long-term data storage in DNA, TRENDS in Biotechnology, 2001年
, Vol.19, No.7, Pages 247-250
C. BANCROFT, LONG-TERM STORAGE OF INFORMATION IN DNA, SCIENCE, 2001年 9月 7日
, V293 N5536, P1763-1765
COX JONATHAN P L, LONG-TERM DATA STORAGE IN DNA, TRENDS IN BIOTECHNOLOGY, 英国, ELSEVIER PUBLICATIONS, 2001年 7月 1日, V19 N7, P247-250
GEORGE M. CHURCH, NEXT-GENERATION DIGITAL INFORMATION STORAGE IN DNA, SCIENCE, 2012年 9月 28日, V337 N6102, P1628
NICK GOLDMAN, TOWARDS PRACTICAL, HIGH-CAPACITY, LOW-MAINTENANCE INFORMATION STORAGE IN SYNTHESIZED DNA, NATURE, 2013年 1月 1日, V494, P77-80

(58)調査した分野(Int.Cl., DB名)

C12N 15/00 - 15/90

C12Q 1/00 - 3/00

C12M 1/00 - 3/10

JSTPlus/JMEDPlus/JST7580(JDreamIII)

PubMed