

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号

特表2024-513187

(P2024-513187A)

(43)公表日 令和6年3月22日(2024.3.22)

(51)国際特許分類		F I		テーマコード(参考)	
C 1 2 Q	1/6876(2018.01)	C 1 2 Q	1/6876	Z	4 B 0 6 3
C 1 2 N	15/09 (2006.01)	C 1 2 N	15/09	Z	
C 1 2 Q	1/6869(2018.01)	C 1 2 Q	1/6869	Z	

審査請求 未請求 予備審査請求 未請求 (全78頁)

(21)出願番号	特願2023-558993(P2023-558993)	(71)出願人	500358711 イルミナ インコーポレイテッド 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
(86)(22)出願日	令和4年3月28日(2022.3.28)	(74)代理人	100094569 弁理士 田中 伸一郎
(85)翻訳文提出日	令和5年9月26日(2023.9.26)	(74)代理人	100103610 弁理士 吉 田 和彦
(86)国際出願番号	PCT/US2022/022184	(74)代理人	100109070 弁理士 須田 洋之
(87)国際公開番号	WO2022/212280	(74)代理人	100119013 弁理士 山崎 一夫
(87)国際公開日	令和4年10月6日(2022.10.6)	(74)代理人	100123777 弁理士 市川 さつき
(31)優先権主張番号	63/167,171		
(32)優先日	令和3年3月29日(2021.3.29)		
(33)優先権主張国・地域又は機関	米国(US)		
(31)優先権主張番号	63/227,550		
(32)優先日	令和3年7月30日(2021.7.30)		
(33)優先権主張国・地域又は機関	米国(US)		
(81)指定国・地域	AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA)		

最終頁に続く

最終頁に続く

(54)【発明の名称】 ライブラリー中のDNA損傷を評価し、アンプリコンサイズバイアスを正規化するための組成物及び方法

(57)【要約】

アンプリコンサイズバイアスを正規化する標準物質及び方法が本明細書に記載される。これらの標準物質は、固有分子識別子を含み得る。いくつかの実施形態では、標準物質及び方法は、次世代配列決定(NGS)アクセシと共に使用するためのものである。蛍光を使用してDNAを含む試料中のDNA損傷を定量化するため、又はライブラリー中のDNA損傷の存在を決定するための方法も本明細書に記載される。

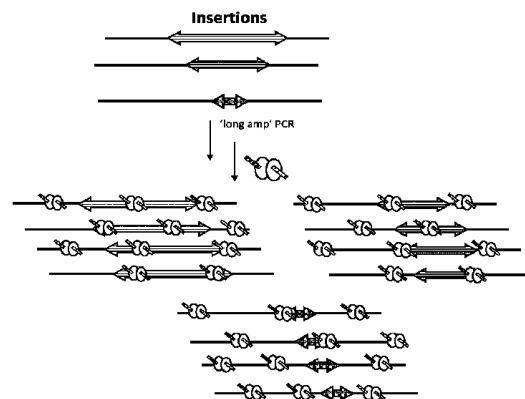


Fig. 1

【特許請求の範囲】

【請求項 1】

異なる長さの核酸標準物質のプールであって、前記核酸標準物質は、固有分子識別子 (UMI)、並びに、

- a. 5'ユニバーサルオリゴヌクレオチドであって、前記5'ユニバーサルオリゴヌクレオチドは全ての標準物質について同じである、5'ユニバーサルオリゴヌクレオチドと、
- b. 3'ユニバーサルオリゴヌクレオチドであって、前記3'ユニバーサルオリゴヌクレオチドは全ての標準物質について同じである、3'ユニバーサルオリゴヌクレオチドと、
- c. 前記UMIと前記5'ユニバーサルオリゴヌクレオチドとの間及び/又は前記UMIと前記3'ユニバーサルオリゴヌクレオチドとの間の少なくとも1つの領域と、を含み、

前記少なくとも1つの領域の長さは、前記標準物質の長さを決定する、標準物質のプール。

10

【請求項 2】

前記プールが、UMI、並びに、

- a. 5'ユニバーサルオリゴヌクレオチドであって、前記5'ユニバーサルオリゴヌクレオチドは全ての標準物質について同じである、5'ユニバーサルオリゴヌクレオチドと、
- b. 3'ユニバーサルオリゴヌクレオチドであって、前記3'ユニバーサルオリゴヌクレオチドは全ての標準物質について同じである、3'ユニバーサルオリゴヌクレオチドと、

を含む、更なる拡散標準物質を更に含み、

前記更なる核酸標準物質は、前記UMIと前記5'ユニバーサルオリゴヌクレオチドとの間又は前記UMIと前記3'ユニバーサルオリゴヌクレオチドとの間の少なくとも1つの領域を含まない、請求項1に記載の標準物質のプール。

20

【請求項 3】

前記UMIと前記5'ユニバーサルオリゴヌクレオチドとの間及び/又は前記UMIと前記3'ユニバーサルオリゴヌクレオチドとの間の前記少なくとも1つの領域が、0.2 kb ~ 1.0 kbを含む、請求項1に記載の標準物質のプール。

【請求項 4】

前記5'ユニバーサルオリゴヌクレオチド及び/又は前記3'ユニバーサルオリゴヌクレオチドがそれぞれ、目的の配列から増幅されたアンプリコンを含む、請求項1~3のいずれか一項に記載の標準物質のプール。

30

【請求項 5】

前記UMIと前記5'ユニバーサルオリゴヌクレオチドとの間及び/又は前記UMIと前記3'ユニバーサルオリゴヌクレオチドとの間の前記少なくとも1つの領域がそれぞれ、目的の配列から増幅されたアンプリコンを含む、請求項1又は3~4のいずれか一項に記載の標準物質のプール。

【請求項 6】

前記UMIと前記5'ユニバーサルオリゴヌクレオチドとの間及び/又は前記UMIと前記3'ユニバーサルオリゴヌクレオチドとの間の前記少なくとも1つの領域がそれぞれ、任意の配列を含む、請求項1又は3~5のいずれか一項に記載の標準物質のプール。

【請求項 7】

異なる長さの核酸標準物質のプールであって、前記核酸標準物質は、UMI、並びに、

- a. 5'部分的重複オリゴヌクレオチドであって、前記5'部分的重複オリゴヌクレオチドは、全ての前記標準物質についてその配列の少なくとも一部にわたって同一である、5'部分的重複オリゴヌクレオチド、及び/又は、
- b. 3'部分的重複オリゴヌクレオチドであって、前記3'部分的重複オリゴヌクレオチドは、全ての前記標準物質についてその配列の少なくとも一部にわたって同一である、3'部分的重複オリゴヌクレオチド、を含み、

前記5'部分的重複オリゴヌクレオチド及び/又は前記3'部分的重複オリゴヌクレオチドの長さは、前記標準物質の長さを決定する、標準物質のプール。

40

【請求項 8】

50

a. 前記 5' 部分的重複オリゴヌクレオチドが、目的の配列の少なくとも第 1 の部分を含み、

b. 前記 3' 部分的重複オリゴヌクレオチドが、目的の配列の少なくとも第 2 の部分を含む、請求項 7 に記載の標準物質のプール。

【請求項 9】

前記 5' 部分的重複オリゴヌクレオチド及び / 又は前記 3' 部分的重複オリゴヌクレオチドがそれぞれ、目的の配列よりも 20 bp ~ 1 kb 小さい配列を含む、請求項 7 ~ 8 のいずれか一項に記載の標準物質のプール。

【請求項 10】

前記 5' 部分的重複オリゴヌクレオチド及び / 又は前記 3' 部分的重複オリゴヌクレオチドがそれぞれ、目的の配列から増幅されたアンプリコンを含む、請求項 7 ~ 9 のいずれか一項に記載の標準物質のプール。

【請求項 11】

前記標準物質が、二本鎖である、請求項 1 ~ 10 のいずれか一項に記載の標準物質のプール。

【請求項 12】

前記標準物質が、二本鎖 DNA を含む、請求項 1 ~ 11 のいずれか一項に記載の標準物質のプール。

【請求項 13】

各標準物質が異なる UMI を含む、請求項 1 ~ 12 のいずれか一項に記載の標準物質のプール。

【請求項 14】

前記標準物質のプールに含まれる前記 UMI が、16 ~ 20 塩基対を含む配列のランダムセットである、請求項 1 ~ 13 のいずれか一項に記載の標準物質のプール。

【請求項 15】

前記標準物質のプールに含まれる前記 UMI が、18 塩基対を含む配列のランダムセットである、請求項 14 に記載の標準物質のプール。

【請求項 16】

前記標準物質のプールが、 1×10^{10} 以上、 10×10^{10} 以上、又は 100×10^{10} 以上の標準物質を含み、各標準物質が異なる UMI を含む、請求項 1 ~ 15 のいずれか一項に記載の標準物質のプール。

【請求項 17】

前記プール中の標準物質の数が、増幅反応によって生成されるアンプリコンの数より多い、請求項 1 ~ 16 のいずれか一項に記載の標準物質のプール。

【請求項 18】

標準物質のプールであって、前記標準物質の少なくとも第 1 の部分は、請求項 1 ~ 6 又は 11 ~ 17 のいずれか一項に由来し、前記標準物質の少なくとも第 2 の部分は、請求項 7 ~ 17 のいずれか一項に由来する、標準物質のプール。

【請求項 19】

核酸標準物質のプールを生成する方法であって、

a. 核酸を含む少なくとも 1 つの目的の配列の複数のコピーを提供することと、
 b. それぞれ UMI を含む一連のオリゴヌクレオチドを提供することと、
 c. 様々な長さの一連の挿入オリゴヌクレオチドを提供することと、
 d. (a) の少なくとも 1 つの目的の配列、(b) の UMI を含む少なくとも 1 つのオリゴヌクレオチド、及び (c) の少なくとも 1 つの挿入アンプリコンをライゲーションして、前記核酸標準物質のプールの複数の核酸標準物質を生成することと、を含む、方法。

【請求項 20】

前記少なくとも 1 つの目的の配列及び / 又は挿入オリゴヌクレオチドが、増幅によって調製される、請求項 19 に記載の方法。

【請求項 21】

40

50

前記目的の配列、それぞれUMIを含む前記オリゴヌクレオチド、及び/又は前記挿入オリゴヌクレオチドが、制限酵素切断部位を含む、請求項19又は20に記載の方法。

【請求項22】

前記制限酵素切断部位が、前記目的の配列、それぞれUMIを含む前記オリゴヌクレオチド、及び/又は前記挿入オリゴヌクレオチドの5'末端及び/又は3'末端の近位にある、請求項21に記載の方法。

【請求項23】

前記方法が、前記目的の配列、それぞれUMIを含む前記オリゴヌクレオチド、及び/又は前記挿入オリゴヌクレオチドを制限酵素で切断することを前記ライゲーションの前に更に含む、請求項21又は22に記載の方法。

10

【請求項24】

前記制限酵素により切断することが、前記ライゲーションのための付着末端を生成する、請求項23に記載の方法。

【請求項25】

核酸標準物質のプールを生成する方法であって、

- a. 核酸を含む少なくとも1つの目的の配列の複数のコピーを提供することと、
- b. それぞれUMIを含む一連のオリゴヌクレオチドを提供することと、
- c. (a)の少なくとも1つの目的の配列と(b)のUMIを含む少なくとも1つのオリゴヌクレオチドをライゲーションすることと、を含む、方法。

【請求項26】

前記少なくとも1つの目的の配列が、増幅によって調製される、請求項25に記載の方法。

20

【請求項27】

前記目的の配列及び/又はそれぞれUMIを含む前記オリゴヌクレオチドが、制限酵素切断部位を含む、請求項25又は26に記載の方法。

【請求項28】

前記制限酵素切断部位が、前記目的の配列及び/又はそれぞれUMIを含む前記オリゴヌクレオチドの5'末端及び/又は3'末端の近位にある、請求項27に記載の方法。

【請求項29】

前記方法が、前記目的の配列及び/又はそれぞれUMIを含む前記オリゴヌクレオチドを制限酵素で切断することを前記ライゲーションの前に更に含む、請求項27~28に記載の方法。

30

【請求項30】

前記制限酵素により切断することが、前記ライゲーションのための付着末端を生成する、請求項29に記載の方法。

【請求項31】

アンプリコンサイズバイアスを正規化する方法であって、

- a. 標的核酸を含む試料を、異なる長さの核酸標準物質のプールと組み合わせることと、
- b. 前記標準物質及び前記標的核酸に含まれる目的の配列のアンプリコンを増幅することと、
- c. 前記標準物質及び前記目的の配列のアンプリコンを配列決定して、配列決定データを生成することと、
- d. 前記標準物質由来の配列決定データを使用して、アンプリコンサイズに基づいてバイアスプロファイルを決定することと、
- e. 前記バイアスプロファイルを使用して、アンプリコンサイズバイアスを正規化することと、を含む、方法。

40

【請求項32】

前記核酸標準物質のプール中の前記標準物質が、0.2kb~20kb塩基対の範囲である、請求項31に記載の方法。

50

【請求項 33】

前記核酸標準物質のプール中に含まれる各標準物質が、異なるUMIを含む、請求項31又は32に記載の方法。

【請求項 34】

前記標準物質のプールに含まれる前記UMIが、16～20塩基対を含む配列のランダムセットである、請求項31～33に記載の方法。

【請求項 35】

前記標準物質のプールに含まれる前記UMIが、18塩基対を含む配列のランダムセットである、請求項31～34に記載の方法。

【請求項 36】

前記標準物質のプールが、 1×10^{10} 以上、 10×10^{10} 以上、又は 100×10^{10} 以上の標準物質を含み、各標準物質が異なるUMIを含む、請求項31～35のいずれか一項に記載の方法。

10

【請求項 37】

前記標準物質のプール中の標準物質の数が、前記増幅によって生成されるアンプリコンの数より多い、請求項31～36のいずれか一項に記載の方法。

【請求項 38】

前記核酸標準物質のプールが、請求項1～18のいずれか一項に記載の核酸標準物質のプールを含む、請求項31～37のいずれか一項に記載の方法。

【請求項 39】

前記核酸標準物質のプールが、請求項1～6又は11～17のいずれか一項に記載の核酸標準物質のプールを含む第1の部分と、請求項7～17のいずれか一項に記載の核酸標準物質のプールを含む第2の部分と、を含む、請求項31～37のいずれか一項に記載の方法。

20

【請求項 40】

前記目的の配列が、前記目的の配列の5'末端及び/若しくは3'末端にない、又は前記目的の配列の5'末端及び/若しくは3'末端に近接していない制限酵素切断部位を含む、請求項31～39のいずれか一項に記載の方法。

【請求項 41】

前記目的の配列が、挿入又は欠失変異を含み得る、請求項31～40のいずれか一項に記載の方法。

30

【請求項 42】

前記目的の配列が遺伝子編集に供されており、任意選択で、前記目的の配列が遺伝子編集によって導入された切断部位を含む、請求項31～41のいずれか一項に記載の方法。

【請求項 43】

前記目的の配列のアンプリコンを増幅することが、前記目的の配列の末端でプライマー結合配列に結合する一对のPCRプライマーを用いて前記標的核酸からアンプリコンを増幅することを含む、請求項31～42のいずれか一項に記載の方法。

【請求項 44】

前記標準物質が、前記目的の配列の末端にあるものと同じプライマー結合配列を含む、請求項31～43のいずれか一項に記載の方法。

40

【請求項 45】

前記増幅後かつ前記配列決定前に、断片のライブラリーを生成することを更に含む、請求項31～44のいずれか一項に記載の方法。

【請求項 46】

前記断片のライブラリーを生成することが、タグメンテーションによる、請求項31～45に記載の方法。

【請求項 47】

前記バイアスプロファイルを決定するために使用される前記標準物質由来の配列決定データが、前記標準物質に含まれるUMIの固有分子数である、請求項31～46のいずれ

50

か一項に記載の方法。

【請求項 48】

1つ以上のライブラリー分子を含むライブラリーにおけるDNA損傷の存在を決定する方法であって、各ライブラリー分子は、挿入部の各末端にヘアピンアダプターを有する二本鎖DNA挿入部を含み、前記方法は、

a. ライブラリー分子に含まれる前記二本鎖DNA挿入部の第1の鎖及び第2の鎖を変性することと、

b. フォワードプライマー及びリバースプライマーをライブラリー分子にアニーリングすることと、

c. ライブラリーアンプリコンを生成するために増幅することと、

d. 生成されたライブラリーアンプリコンの数に基づいてDNA損傷の存在を評価することと、を含む、方法。

10

【請求項 49】

前記フォワードプライマー及び/又は前記リバースプライマーが、一方又は両方のヘアピンアダプターに含まれる1つ以上の配列に結合する、請求項48に記載の方法。

【請求項 50】

前記フォワードプライマーが、前記二本鎖DNA挿入部の第1の末端に結合した前記ヘアピンアダプターに含まれる配列に結合し、前記リバースプライマーが、前記二本鎖DNA挿入部の第2の末端に結合した前記ヘアピンアダプターに含まれる配列に結合する、請求項48又は49に記載の方法。

20

【請求項 51】

前記生成されたライブラリーアンプリコンの数が、定量サイクル(Cq)値を測定することによって推定される、請求項48～50のいずれか一項に記載の方法。

【請求項 52】

ライブラリーアンプリコンの数が多いほど、Cq値が低くなる、請求項48～51のいずれか一項に記載の方法。

【請求項 53】

より低いCq値を有するライブラリーが、より少ないDNA損傷を有する、請求項48～52のいずれか一項に記載の方法。

【請求項 54】

前記Cq値に基づいて前記ライブラリーの分析のための条件を決定することを更に含む、請求項51～53のいずれか一項に記載の方法。

30

【請求項 55】

前記分析が配列決定である、請求項54に記載の方法。

【請求項 56】

前記増幅することが、5 kb以上、10 kb以上、15 kb以上、20 kb以上、25 kb以上、又は30 kb以上であるライブラリー分子を増幅するために最適化されている、請求項48～55のいずれか一項に記載の方法。

【請求項 57】

前記増幅することが、長いアンプリコンの増幅に最適化されているポリメラーゼを用いて行われる、請求項48～56のいずれか一項に記載の方法。

40

【請求項 58】

前記ポリメラーゼが、20 kb以上又は30 kb以上のアンプリコンの増幅に最適化されている、請求項57に記載の方法。

【請求項 59】

前記ポリメラーゼが、野生型Taqポリメラーゼと比較してより高い処理能力又は伸長速度を有する、請求項57又は58に記載の方法。

【請求項 60】

前記ポリメラーゼが、処理能力又は伸長速度を増加させる1つ以上の突然変異又は融合を含む、請求項59に記載の方法。

50

- 【請求項 6 1】
前記ポリメラーゼが、3 kb / 分以上の伸長速度を有する、請求項 5 9 又は 6 0 に記載の方法。
- 【請求項 6 2】
前記増幅することが指数関数的である、請求項 4 8 ~ 6 1 のいずれか一項に記載の方法。
- 【請求項 6 3】
3 0 サイクル以上又は 4 0 サイクル以上の増幅が行われる、請求項 4 8 ~ 6 2 のいずれか一項に記載の方法。
- 【請求項 6 4】
前記 DNA 損傷が、ライブラリー分子中の 1 つ以上のニックを含む、請求項 4 8 ~ 6 3 のいずれか一項に記載の方法。 10
- 【請求項 6 5】
前記 1 つ以上のニックが前記挿入部内にある、請求項 6 4 に記載の方法。
- 【請求項 6 6】
前記 C q 値が、前記ライブラリー中のより高い割合のライブラリー分子が 1 つ以上のニックを含む場合により高い、請求項 6 4 又は 6 5 に記載の方法。
- 【請求項 6 7】
前記 DNA 損傷が、ライブラリー分子中の 2 つ以上のニックを含み、前記ニックが、前記二本鎖 DNA 挿入部の同じ鎖中にある、請求項 6 4 ~ 6 6 のいずれか一項に記載の方法。 20
- 【請求項 6 8】
前記 DNA 損傷が、ライブラリー分子中の 2 つ以上のニックを含み、前記ニックが、前記二本鎖 DNA 挿入部の両方の鎖中にある、請求項 6 4 ~ 6 6 のいずれか一項に記載の方法。
- 【請求項 6 9】
前記ライブラリー分子が 1 つ以上のニックを含む場合、前記フォワードプライマー及び / 又は前記リバースプライマーが、前記ライブラリー分子の完全配列に対応するアンプリコンを生成することができない、請求項 4 8 ~ 6 8 のいずれか一項に記載の方法。
- 【請求項 7 0】
ニックを含むライブラリー分子から生成されたアンプリコンが、前記フォワードプライマー及び / 又は前記リバースプライマーに結合するための配列を欠いている、請求項 6 9 に記載の方法。 30
- 【請求項 7 1】
ニックを含むライブラリー分子が、ニックを含まないライブラリー分子と比較して、前記増幅中により少ないアンプリコンを生成する、請求項 6 4 ~ 7 0 のいずれか一項に記載の方法。
- 【請求項 7 2】
前記フォワードプライマー及び前記リバースプライマーをアニーリングする前に、ニックから二本鎖切断を生じさせることを更に含む、請求項 6 4 ~ 7 1 のいずれか一項に記載の方法。 40
- 【請求項 7 3】
前記二本鎖切断を生じさせることが、酵素反応を使用して行われる、請求項 7 2 に記載の方法。
- 【請求項 7 4】
前記酵素反応がエンドヌクレアーゼによって行われる、請求項 7 3 に記載の方法。
- 【請求項 7 5】
前記エンドヌクレアーゼが T 7 エンドヌクレアーゼである、請求項 7 4 に記載の方法。
- 【請求項 7 6】
二本鎖切断を含むライブラリー分子が、前記増幅中に前記ライブラリー分子の完全配列 50

に対応するアンプリコンを生成しない、請求項 72 ~ 75 のいずれか一項に記載の方法。

【請求項 77】

二本鎖切断を含むライブラリー分子から生成されたアンプリコンが、前記フォワードプライマー及び/又は前記リバースプライマーに結合するための配列を欠いている、請求項 72 ~ 76 に記載の方法。

【請求項 78】

蛍光を使用して DNA を含む試料中の DNA 損傷を定量化する方法であって、

a .

i . DNA を含む試料のアリコートと、

ii . 1 つ以上の DNA 修復酵素と、

iii . 1 つ以上の dNTP が蛍光標識されている、dNTP と、を組み合わせることと、

b . 修復された DNA を調製することと、

c . dNTP からリン酸を脱リン酸化することと、

d . 前記修復された DNA をカルボキシレートビーズ又はセルロースビーズに結合させることと、

e . 再懸濁緩衝液を用いて、前記結合した修復 DNA を前記カルボキシレートビーズ又はセルロースビーズから溶出することと、

f . 前記修復された DNA の蛍光を測定して、DNA 損傷の量を決定することと、を含む、方法。

【請求項 79】

前記修復された DNA のより高い蛍光が、より多い DNA 損傷を示す、請求項 78 に記載の方法。

【請求項 80】

前記修復された DNA の蛍光が、異なる量の DNA 損傷の範囲にわたって線形である、請求項 78 又は 79 に記載の方法。

【請求項 81】

前記アッセイが、前記試料の操作の前後に同じ試料のアリコートを評価することによって、前記操作によって誘発された DNA 損傷を評価することができる、請求項 78 ~ 80 のいずれか一項に記載の方法。

【請求項 82】

前記操作が、試料の配列決定である、請求項 81 に記載の方法。

【請求項 83】

前記修復された DNA の蛍光を測定することが、修復された DNA の希釈物の標準曲線を作成することと、前記修復された DNA の希釈物の蛍光を測定することと、を含む、請求項 81 又は 82 に記載の方法。

【請求項 84】

前記修復された DNA の蛍光を測定することが、前記修復された DNA の蛍光を、蛍光標識された前記 1 つ以上の dNTP のみの希釈物の別個の標準曲線に対して比較し、前記修復された DNA に含まれる蛍光色素分子の数を決定することを含む、請求項 78 ~ 83 のいずれか一項に記載の方法。

【請求項 85】

前記決定された蛍光色素分子の数を前記修復された DNA の質量で割ることによって、前記修復された DNA に含まれる蛍光色素分子の正規化された数を計算することを更に含む、請求項 84 の方法。

【請求項 86】

前記 DNA が、ゲノム DNA、cDNA、又は断片化二本鎖 DNA を含むライブラリーである、請求項 78 ~ 85 のいずれか一項に記載の方法。

【請求項 87】

前記 DNA が、ゲノム DNA 及び cDNA であり、前記方法が、前記 DNA 損傷の量を

10

20

30

40

50

決定した後にライブラリーを調製することを更に含む、請求項 86 の方法。

【請求項 88】

前記 DNA 損傷の量が全ヌクレオチドの 5% 以下、4% 以下、3% 以下、2% 以下、又は 1% 以下である場合にライブラリーが調製される、請求項 87 に記載の方法。

【請求項 89】

前記 DNA 損傷の量が全ヌクレオチドの 5% 以上、4% 以上、3% 以上、2% 以上、又は 1% 以上である場合にライブラリーが調製されない、請求項 78 ~ 88 のいずれか一項に記載の方法。

【請求項 90】

前記蛍光を測定する前に、前記修復された DNA をカルボキシレートビーズ又はセルロースビーズに結合させ、溶出することが 2 回以上行われる、請求項 78 ~ 89 のいずれか一項に記載の方法。 10

【請求項 91】

前記蛍光を測定する前に、前記修復された DNA をカルボキシレートビーズ又はセルロースビーズに結合させ、溶出することが 2 回行われる、請求項 90 に記載の方法。

【請求項 92】

前記カルボキシレートビーズ又はセルロースビーズが磁性である、請求項 78 ~ 91 のいずれか一項に記載の方法。

【請求項 93】

前記修復された DNA を調整することが、37 で行われる、請求項 78 ~ 92 のいずれか一項に記載の方法。 20

【請求項 94】

前記修復された DNA を調製することが、10 分間以上、20 分間以上、30 分間以上、45 分間以上、又は 60 分間以上行われる、請求項 78 ~ 93 のいずれか一項に記載の方法。

【請求項 95】

前記 dNTP からリン酸を脱リン酸化することが、酵素を用いて行われる、請求項 78 ~ 94 に記載の方法。

【請求項 96】

前記 dNTP からリン酸を脱リン酸化するための前記酵素が、エピアルカリホスファターゼ (SAP) 又は仔ウシ腸アルカリホスファターゼ (CIP) である、請求項 78 ~ 95 に記載の方法。 30

【請求項 97】

前記 1 つ以上の DNA 修復酵素が、DNA ポリメラーゼを含む、請求項 78 ~ 96 のいずれか一項に記載の方法。

【請求項 98】

前記 DNA ポリメラーゼが、5' 3' ポリメラーゼ活性を有するが、5' 3' エキソヌクレアーゼ活性を欠いている、請求項 97 に記載の方法。

【請求項 99】

前記 DNA ポリメラーゼが、Bst DNA ポリメラーゼ、ラージフラグメントである、請求項 97 に記載の方法。 40

【請求項 100】

前記 1 つ以上の DNA 修復酵素が、リガーゼを含む、請求項 78 ~ 99 のいずれかに記載の方法。

【請求項 101】

前記リガーゼが、Taq リガーゼである、請求項 100 に記載の方法。

【請求項 102】

前記 DNA 損傷が、二本鎖 DNA 中のニックを含む、請求項 78 ~ 101 のいずれか一項に記載の方法。

【請求項 103】

前記1つ以上のDNA修復酵素が、T4ピリミジンダイマーグリコシラーゼ(PDG)を含む、請求項78~102のいずれか一項に記載の方法。

【請求項104】

前記DNA損傷が、チミンダイマーを含む、請求項78~103のいずれか一項に記載の方法。

【請求項105】

前記チミンダイマーが、紫外線照射によって誘発された、請求項104に記載の方法。

【請求項106】

前記1つ以上のDNA修復酵素が、ウラシルDNAグリコシラーゼ(UDG)及び脱プリン又は脱ピリミジン部位リナーゼを含む、請求項78~105のいずれかに記載の方法 10

【請求項107】

前記DNA損傷が、ウラシルを含む、請求項78~106のいずれか一項に記載の方法。

【請求項108】

前記1つ以上のDNA修復酵素が、ホルムアミドピリミジンDNAグリコシラーゼ(FPG)及び脱プリン又は脱ピリミジン部位リナーゼを含む、請求項78~107のいずれかに記載の方法。

【請求項109】

前記DNA損傷が、酸化塩基を含む、請求項78~108に記載の方法。 20

【請求項110】

前記dNTPが、dATP、dGTP、dCTP、及びdTTP又はdUTPを含む、請求項78~109のいずれか一項に記載の方法。

【請求項111】

全ての前記dNTPが蛍光標識されている、請求項78~110のいずれか一項に記載の方法。

【請求項112】

dUTP及びdCTPが蛍光標識されている、請求項78~111に記載の方法。

【請求項113】

前記蛍光標識が、Alexa Fluor 488、Alexa Fluor 546、Alexa Fluor 555、Alexa Fluor 633、フルオレセインイソチオシアネート(FITC)、又はテトラメチルローダミン-5-(及び6)-イソチオシアネート(TRITC)である、請求項112に記載の方法。 30

【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、2021年3月29日に出願された米国特許仮出願第63/167,171号、及び2021年7月30日に出願された同第63/227,550号に対する優先権の利益を主張し、これらの各々は、任意の目的のためにその全体が参照により本明細書に組み込まれる。 40

【0002】

説明

本出願は、次世代配列決定(NGS)アッセイにおいてライブラリー損傷を評価し、アンプリコンサイズバイアスを正規化するための標準物質及び方法に関する。本出願はまた、蛍光を使用してDNAを含む試料中のDNA損傷を定量化することにも関する。

【背景技術】

【0003】

ゲノム編集又は腫瘍学用途における大きな挿入/欠失バリエーション(インデル)を検出及び定量化するための一般的な方法は、標的化された「長いアンプリコン」PCR(Lon 50

g A m p、1 k b超)と、それに続く長リード配列決定又は(短リード) N G Sのための短リードライブラリーへの変換を含む。しかしながら、「長い」P C R増幅におけるサイズに基づくバイアスは、大きなインデルバリエントの相対頻度を正確に定量化するプロセスを複雑にする。増幅前又は増幅中に固有分子識別子で標的D N A分子の末端にタグ付けする戦略は、同じN G Sリードにおいて同定されるバリエント及びU M Iを必要とする。したがって、長いアンプリコンライブラリーを用いるタグ付け方法は、長リード配列決定又は複雑な合成長リードライブラリー調製を必要とする。短リードN G Sは、バリエント配列及び元のアンプリコンU M Iを別々のリードに分断し得るため、短リードN G Sのための増幅後ライブラリー変換工程は、このU M I末端タグ付けを不適切なものにする。

【0004】

10

これらの既存の方法は、アンプリコンサイズバイアスを正規化するために、様々な長さのU M I含有合成D N A対照を有する短リードN G Sを組み込む。D N A対照は、標準物質及びU M Iの同一性が同じN G Sリードに含まれるように設計される。これらの標準物質を用いて対照アッセイを実行すること、又は各L o n g A m pアッセイに既知量のこれらの標準物質をスパイクインすることは、サイズに基づくP C Rバイアスのバイオインフォマティクス分析を可能にし、定量化されたP C Rサイズバイアスを考慮することによって、より良好に大きなインデルの頻度を推定することを容易にする。

【0005】

長リード配列決定のためのライブラリー(すなわち、長リードライブラリー)に関する別の問題は、損傷したライブラリー分子の存在である。長リードライブラリー調製の品質の評価は、後続のワークフロー工程及び配列決定の成功を予測するために使用され得る。長いライブラリー分子は、標準的なワークフローの間に容易にニック形成又は損傷され得、その結果アダプター配列と関連しないライブラリー分子を生じ、したがって、配列決定などのアダプターを必要とするワークフローにおいて使用できない。ライブラリー調製工程は、ピペティング、保存、又は他の取り扱い及び/若しくは手技的エラーのいずれかによって、D N Aを損傷し得る。ニックの入ったD N Aが、5'アダプター及び3'アダプターの両方を必要とするライブラリー調製を経る場合、ニックの入ったD N Aは、下流の工程において使用不可能である。したがって、考慮されないライブラリー損傷は、ライブラリー濃度の不正確な推定、不十分な配列決定カバレッジ、及び全体的に不十分な配列決定アッセイ基準を引き起こし得る。

20

30

【0006】

ライブラリー調製中に損傷していないライブラリー分子を正確に定量するためのライブラリー品質管理(Q C)法は、この問題を解決するのに役立ち得る。本明細書に記載される定量P C R(q P C R) Q C法は、不正確な濃度のライブラリーによって後続のワークフロー工程を進めることを回避するために、ライブラリー調製品質を評価する。したがって、これらの方法は、ユーザーの時間、予算、並びに試薬及び他の消耗品の損失を回避することができる。

【0007】

更に、環境、試料の調製及び処理、又は保存条件由来のD N A損傷は、ライブラリー調製品質の一貫性に顕著に影響し得る。例えば、配列決定プロセスの間、配列決定サイクル間の低波長レーザー及び他の化学物質への曝露によるD N A損傷の蓄積は、配列決定の誤り率を増加させ得る。ユーザーは、この損傷の評価を望み得る。蛍光を使用してD N A損傷を定量化する方法が本明細書に記載される。蛍光を使用してD N A損傷を定量化するために開発された他のアッセイ(例えば、米国特許出願公開第2014/0030705号、国際公開第2010028388号、及び米国特許出願公開第20090042205号)は、取り込まれていない蛍光ヌクレオチドの非特異的結合におそらく部分的に起因する低いシグナル対ノイズ比によって妨げられてきた。D N A損傷を測定する本発明の方法は、d N T Pの脱リン酸化の工程と、カルボキシレート又はセルロースビーズから修復されたD N Aの結合/溶出の工程とを組み込み、シグナルを改善し、アッセイのより大きなダイナミックレンジを可能にする。

40

50

【発明の概要】

【0008】

異なる長さの核酸標準物質のプールが本明細書に記載され、核酸標準物質は、固有分子識別子（UMI）、並びに、

5'ユニバーサルオリゴヌクレオチドであって、5'ユニバーサルオリゴヌクレオチドは全ての標準物質について同じである、5'ユニバーサルオリゴヌクレオチドと、3'ユニバーサルオリゴヌクレオチドであって、3'ユニバーサルオリゴヌクレオチドは全ての標準物質について同じである、3'ユニバーサルオリゴヌクレオチドと、UMIと5'ユニバーサルオリゴヌクレオチドとの間、及び/又は、UMIと3'ユニバーサルオリゴヌクレオチドとの間の少なくとも1つの領域と、を含み、

10

少なくとも1つの領域の長さは、標準物質の長さを決定する。また、ライブラリーの品質管理の方法も本明細書に記載される。

【0009】

実施形態1．異なる長さの核酸標準物質のプールであって、核酸標準物質は、固有分子識別子（UMI）、並びに、

a．5'ユニバーサルオリゴヌクレオチドであって、5'ユニバーサルオリゴヌクレオチドは全ての標準物質について同じである、5'ユニバーサルオリゴヌクレオチドと、

b．3'ユニバーサルオリゴヌクレオチドであって、3'ユニバーサルオリゴヌクレオチドは全ての標準物質について同じである、5'ユニバーサルオリゴヌクレオチドと、

c．UMIと5'ユニバーサルオリゴヌクレオチドとの間及び/又はUMIと3'ユニバーサルオリゴヌクレオチドとの間の少なくとも1つの領域と、を含み、

20

少なくとも1つの領域の長さは、標準物質の長さを決定する、標準物質のプール。

【0010】

実施形態2．プールが、UMI、並びに、

a．5'ユニバーサルオリゴヌクレオチドであって、5'ユニバーサルオリゴヌクレオチドは全ての標準物質について同じである、5'ユニバーサルオリゴヌクレオチドと、

b．3'ユニバーサルオリゴヌクレオチドであって、3'ユニバーサルオリゴヌクレオチドは全ての標準物質について同じである、3'ユニバーサルオリゴヌクレオチドと、を含む、更なる核酸標準物質を更に含み、

更なる核酸標準物質は、UMIと5'ユニバーサルオリゴヌクレオチドとの間又はUMIと3'ユニバーサルオリゴヌクレオチドとの間の少なくとも1つの領域を含まない、実施形態1に記載の標準物質のプール。

30

【0011】

実施形態3．UMIと5'ユニバーサルオリゴヌクレオチドとの間及び/又はUMIと3'ユニバーサルオリゴヌクレオチドとの間の少なくとも1つの領域が、0.2kb~10kbを含む、実施形態1に記載の標準物質のプール。

【0012】

実施形態4．5'ユニバーサルオリゴヌクレオチド及び/又は3'ユニバーサルオリゴヌクレオチドがそれぞれ、目的の配列から増幅されたアンプリコンを含む、実施形態1~3のいずれか1つに記載の標準物質のプール。

40

【0013】

実施形態5．UMIと5'ユニバーサルオリゴヌクレオチドとの間及び/又はUMIと3'ユニバーサルオリゴヌクレオチドとの間の少なくとも1つの領域がそれぞれ、目的の配列から増幅されたアンプリコンを含む、実施形態1又は3~4のいずれか1つに記載の標準物質のプール。

【0014】

実施形態6．UMIと5'ユニバーサルオリゴヌクレオチドとの間及び/又はUMIと3'ユニバーサルオリゴヌクレオチドとの間の少なくとも1つの領域がそれぞれ、任意の配列を含む、実施形態1又は3~5のいずれか1つに記載の標準物質のプール。

【0015】

50

実施形態 7 . 異なる長さの核酸標準物質のプールであって、核酸標準物質は、UMI、並びに、

a . 5'部分的重複オリゴヌクレオチドであって、5'部分的重複オリゴヌクレオチドは、全ての標準物質についてその配列の少なくとも一部にわたって同一である、5'部分的重複オリゴヌクレオチド、及び/又は、

b . 3'部分的重複オリゴヌクレオチドであって、3'部分的重複オリゴヌクレオチドは、全ての標準物質についてその配列の少なくとも一部にわたって同一である、3'部分的重複オリゴヌクレオチド、を含み、

5'部分的重複オリゴヌクレオチド及び/又は3'部分的重複オリゴヌクレオチドの長さは、標準物質の長さを決定する、標準物質のプール。

10

【0016】

実施形態 8 .

a . 5'部分的重複オリゴヌクレオチドが、目的の配列の少なくとも第1の部分を含み、

b . 3'部分的重複オリゴヌクレオチドが、目的の配列の少なくとも第2の部分を含む、実施形態 7 に記載の標準物質のプール。

【0017】

実施形態 9 . 5'部分的重複オリゴヌクレオチド及び/又は3'部分的重複オリゴヌクレオチドがそれぞれ、目的の配列よりも20bp~1kb小さい配列を含む、実施形態 7 ~ 8 のいずれか1つに記載の標準物質のプール。

【0018】

20

実施形態 10 . 5'部分的重複オリゴヌクレオチド及び/又は3'部分的重複オリゴヌクレオチドがそれぞれ、目的の配列から増幅されたアンプリコンを含む、実施形態 7 ~ 9 のいずれか1つに記載の標準物質のプール。

【0019】

実施形態 11 . 標準物質が、二本鎖核酸を含む、実施形態 1 ~ 10 のいずれか1つに記載の標準物質のプール。

【0020】

実施形態 12 . 標準物質が、二本鎖DNAを含む、実施形態 1 ~ 11 のいずれか1つに記載の標準物質のプール。

【0021】

30

実施形態 13 . 各標準物質が異なるUMIを含む、実施形態 1 ~ 12 のいずれか1つに記載の標準物質のプール。

【0022】

実施形態 14 . 標準物質のプールに含まれるUMIが、16~20塩基対を含む配列のランダムセットである、実施形態 1 ~ 13 のいずれか1つに記載の標準物質のプール。

【0023】

実施形態 15 . 標準物質のプールに含まれるUMIが、18塩基対を含む配列のランダムセットである、実施形態 14 に記載の標準物質のプール。

【0024】

実施形態 16 . 標準物質のプールが、 1×10^{10} 以上、 10×10^{10} 以上、又は 100×10^{10} 以上の標準物質を含み、各標準物質が異なるUMIを含む、実施形態 1 ~ 15 のいずれか1つに記載の標準物質のプール。

40

【0025】

実施形態 17 . プール中の標準物質の数が、増幅反応によって生成されるアンプリコンの数より多い、実施形態 1 ~ 16 のいずれか1つに記載の標準物質のプール。

【0026】

実施形態 18 . 標準物質のプールであって、標準物質の少なくとも第1の部分は、実施形態 1 ~ 6 又は11~17のいずれか1つに由来し、標準物質の少なくとも第2の部分は、実施形態 7 ~ 17 のいずれか1つに由来する、標準物質のプール。

【0027】

50

実施形態 19 . 核酸標準物質のプールを生成する方法であって、

- a . 核酸を含む少なくとも 1 つの目的の配列の複数のコピーを提供することと、
- b . それぞれ U M I を含む一連のオリゴヌクレオチドを提供することと、
- c . 様々な長さの一連の挿入オリゴヌクレオチドを提供することと、
- d . (a) の少なくとも 1 つの目的の配列、(b) の U M I を含む少なくとも 1 つのオリゴヌクレオチド、及び (c) の少なくとも 1 つの挿入アンプリコンをライゲーションして、核酸標準物質のプールの複数の核酸標準物質を生成することと、を含む、方法。

【 0 0 2 8 】

実施形態 20 . 少なくとも 1 つの目的の配列及び / 又は挿入オリゴヌクレオチドが、増幅によって調製される、実施形態 19 に記載の方法。

10

【 0 0 2 9 】

実施形態 21 . 目的の配列、それぞれ U M I を含むオリゴヌクレオチド、及び / 又は挿入オリゴヌクレオチドが、制限酵素切断部位を含む、実施形態 19 又は実施形態 20 に記載の方法。

【 0 0 3 0 】

実施形態 22 . 制限酵素切断部位が、目的の配列、それぞれ U M I を含むオリゴヌクレオチド、及び / 又は挿入オリゴヌクレオチドの 5 ' 末端及び / 又は 3 ' 末端の近位にある、実施形態 21 に記載の方法。

【 0 0 3 1 】

実施形態 23 . 方法が、目的の配列、それぞれ U M I を含むオリゴヌクレオチド、及び / 又は挿入オリゴヌクレオチドを制限酵素で切断することをライゲーションの前に更に含む、実施形態 21 又は実施形態 22 に記載の方法。

20

【 0 0 3 2 】

実施形態 24 . 制限酵素により切断することが、ライゲーションのための付着末端を生成する、実施形態 23 に記載の方法。

【 0 0 3 3 】

実施形態 25 . 核酸標準物質のプールを生成する方法であって、

- a . 核酸を含む少なくとも 1 つの目的の配列の複数のコピーを提供することと、
- b . それぞれ U M I を含む一連のオリゴヌクレオチドを提供することと、
- c . (a) の少なくとも 1 つの目的の配列と (b) の U M I を含む少なくとも 1 つのオリゴヌクレオチドをライゲーションすることと、を含む、方法。

30

【 0 0 3 4 】

実施形態 26 . 少なくとも 1 つの目的の配列が、増幅によって調製される、実施形態 25 に記載の方法。

【 0 0 3 5 】

実施形態 27 . 目的の配列及び / 又はそれぞれ U M I を含むオリゴヌクレオチドが、制限酵素切断部位を含む、実施形態 25 又は実施形態 26 に記載の方法。

【 0 0 3 6 】

実施形態 28 . 制限酵素切断部位が、目的の配列及び / 又はそれぞれ U M I を含むオリゴヌクレオチドの 5 ' 末端及び / 又は 3 ' 末端の近位にある、実施形態 27 に記載の方法。

40

【 0 0 3 7 】

実施形態 29 . 方法が、目的の配列及び / 又はそれぞれ U M I を含むオリゴヌクレオチドを制限酵素で切断することをライゲーションの前に更に含む、実施形態 27 ~ 28 に記載の方法。

【 0 0 3 8 】

実施形態 30 . 制限酵素により切断することが、ライゲーションのための付着末端を生成する、実施形態 29 に記載の方法。

【 0 0 3 9 】

実施形態 31 . アンプリコンサイズバイアスを正規化する方法であって、

- a . 標的核酸を含む試料を、異なる長さの核酸標準物質のプールと組み合わせることで

50

あって、各標準物質はUMIを含む、ことと、

b．標準物質及び標的核酸に含まれる目的の配列のアンプリコンを増幅することと、

c．標準物質及び目的の配列のアンプリコンを配列決定して、配列決定データを生成することと、

d．標準物質由来の配列決定データを使用して、アンプリコンサイズに基づいてバイアスプロファイルを決定することと、

e．バイアスプロファイルを使用して、アンプリコンサイズバイアスを正規化することと、を含む、方法。

【0040】

実施形態32．核酸標準物質のプール中の標準物質が、0.2kb~20kb塩基対の範囲である、実施形態31に記載の方法。 10

【0041】

実施形態33．核酸標準物質のプール中に含まれる各標準物質が、異なるUMIを含む、実施形態31又は実施形態32に記載の方法。

【0042】

実施形態34．標準物質のプールに含まれるUMIが、16~20塩基対を含む配列のランダムセットである、実施形態31~33に記載の方法。

【0043】

実施形態35．標準物質のプールに含まれるUMIが、18塩基対を含む配列のランダムセットである、実施形態31~34に記載の方法。 20

【0044】

実施形態36．標準物質のプールが、 1×10^{10} 以上、 10×10^{10} 以上、又は 100×10^{10} 以上の標準物質を含み、各標準物質が異なるUMIを含む、実施形態31~35のいずれか1つに記載の方法。

【0045】

実施形態37．標準物質のプール中の標準物質の数が、増幅によって生成されるアンプリコンの数より多い、実施形態31~36のいずれか1つに記載の方法。

【0046】

実施形態38．核酸標準物質のプールが、実施形態1~18のいずれか1つに記載の核酸標準物質のプールを含む、実施形態31~37のいずれか1つに記載の方法。 30

【0047】

実施形態39．核酸標準物質のプールが、実施形態1~6又は11~17のいずれか1つに記載の核酸標準物質のプールを含む第1の部分と、実施形態7~17のいずれか1つに記載の核酸標準物質のプールを含む第2の部分と、を含む、実施形態31~37のいずれか1つに記載の方法。

【0048】

実施形態40．目的の配列が、目的の配列の5'末端及び/若しくは3'末端にない、又は目的の配列の5'末端及び/若しくは3'末端に近接していない制限酵素切断部位を含む、実施形態31~39のいずれか1つに記載の方法。

【0049】

実施形態41．目的の配列が、挿入又は欠失変異を含み得る、実施形態31~40のいずれか1つに記載の方法。 40

【0050】

実施形態42．目的の配列が遺伝子編集に供されており、任意選択で、目的の配列が遺伝子編集によって導入された切断部位を含む、実施形態31~41のいずれか1つに記載の方法。

【0051】

実施形態43．目的の配列のアンプリコンを増幅することが、目的の配列の末端でプライマー結合配列に結合する一対のPCRプライマーを用いて標的核酸からアンプリコンを増幅することを含む、実施形態31~42のいずれか1つに記載の方法。 50

【 0 0 5 2 】

実施形態 4 4 . 標準物質が、目的の配列の末端にあるものと同じプライマー結合配列を含む、実施形態 3 1 ~ 4 3 のいずれか 1 つに記載の方法。

【 0 0 5 3 】

実施形態 4 5 . 増幅後かつ配列決定前に、断片のライブラリーを生成することを更に含む、実施形態 3 1 ~ 4 4 のいずれか 1 つに記載の方法。

【 0 0 5 4 】

実施形態 4 6 . 断片のライブラリーを生成することが、タグメンテーションによる、実施形態 3 1 ~ 4 5 のいずれか 1 つに記載の方法。

【 0 0 5 5 】

実施形態 4 7 . バイアスプロファイルを決定するために使用される標準物質由来の配列決定データが、標準物質に含まれる U M I の固有分子数である、実施形態 3 1 ~ 4 6 のいずれか 1 つに記載の方法。

【 0 0 5 6 】

実施形態 4 8 . 1 つ以上のライブラリー分子を含むライブラリーにおける D N A 損傷の存在を決定する方法であって、各ライブラリー分子は、挿入部の各末端にヘアピンアダプターを有する二本鎖 D N A 挿入部を含み、方法は、

- a . ライブラリー分子に含まれる二本鎖 D N A 挿入部の第 1 の鎖及び第 2 の鎖を変性することと、
- b . フォワードプライマー及びリバースプライマーをライブラリー分子にアニーリングすることと、
- c . ライブラリーアンプリコンを生成するために増幅することと、
- d . 生成されたライブラリーアンプリコンの数に基づいて D N A 損傷の存在を評価することと、を含む、方法。

【 0 0 5 7 】

実施形態 4 9 . フォワードプライマー及び / 又はリバースプライマーが、一方又は両方のヘアピンアダプターに含まれる 1 つ以上の配列に結合する、実施形態 4 8 に記載の方法。

【 0 0 5 8 】

実施形態 5 0 . フォワードプライマーが、二本鎖 D N A 挿入部の第 1 の末端に結合したヘアピンアダプターに含まれる配列に結合し、リバースプライマーが、二本鎖 D N A 挿入部の第 2 の末端に結合したヘアピンアダプターに含まれる配列に結合する、実施形態 4 8 又は実施形態 4 9 に記載の方法。

【 0 0 5 9 】

実施形態 5 1 . 生成されたライブラリーアンプリコンの数が、定量サイクル (cycle o f quantification、C q) 値を測定することによって推定される、実施形態 4 8 ~ 5 0 のいずれか 1 つに記載の方法。

【 0 0 6 0 】

実施形態 5 2 . ライブラリーアンプリコンの数が多いほど、C q 値が低くなる、実施形態 4 8 ~ 5 1 のいずれか 1 つに記載の方法。

【 0 0 6 1 】

実施形態 5 3 . より低い C q 値を有するライブラリーが、より少ない D N A 損傷を有する、実施形態 4 8 ~ 5 2 のいずれか 1 つに記載の方法。

【 0 0 6 2 】

実施形態 5 4 . C q 値に基づいてライブラリーの分析のための条件を決定することを更に含む、実施形態 5 1 ~ 5 3 のいずれか 1 つに記載の方法。

【 0 0 6 3 】

実施形態 5 5 . 分析が配列決定である、実施形態 5 4 に記載の方法。

【 0 0 6 4 】

実施形態 5 6 . 増幅することが、5 k b 以上、1 0 k b 以上、1 5 k b 以上、2 0 k b

10

20

30

40

50

以上、25 kb 以上、又は30 kb 以上であるライブラリー分子を増幅するために最適化されている、実施形態48～55のいずれか1つに記載の方法。

【0065】

実施形態57．増幅することが、長いアンプリコンの増幅に最適化されているポリメラーゼを用いて行われる、実施形態48～56のいずれか1つに記載の方法。

【0066】

実施形態58．ポリメラーゼが、20 kb 以上又は30 kb 以上のアンプリコンの増幅に最適化されている、実施形態57に記載の方法。

【0067】

実施形態59．ポリメラーゼが、野生型Taqポリメラーゼと比較してより高い処理能力又は伸長速度を有する、実施形態57又は実施形態58に記載の方法。 10

【0068】

実施形態60．ポリメラーゼが、処理能力又は伸長速度を増加させる1つ以上の突然変異又は融合を含む、実施形態59に記載の方法。

【0069】

実施形態61．ポリメラーゼが、3 kb / 分より高い伸長速度を有する、実施形態59又は実施形態60に記載の方法。

【0070】

実施形態62．増幅することが指数関数的である、実施形態48～61のいずれか1つに記載の方法。 20

【0071】

実施形態63．30サイクル以上又は40サイクル以上の増幅が行われる、実施形態48～62のいずれか1つに記載の方法。

【0072】

実施形態64．DNA損傷が、ライブラリー分子中の1つ以上のニックを含む、実施形態48～63のいずれか1つに記載の方法。

【0073】

実施形態65．1つ以上のニックが挿入部内にある、実施形態64に記載の方法。

【0074】

実施形態66．Cq値が、ライブラリー中のより高い割合のライブラリー分子が1つ以上のニックを含む場合により高い、実施形態64又は実施形態65に記載の方法。 30

【0075】

実施形態67．DNA損傷が、ライブラリー分子中の2つ以上のニックを含み、ニックが、二本鎖DNA挿入部の同じ鎖中にある、実施形態64～66のいずれか1つに記載の方法。

【0076】

実施形態68．DNA損傷が、ライブラリー分子中の2つ以上のニックを含み、ニックが、二本鎖DNA挿入部の両方の鎖中にある、実施形態64～66のいずれか1つに記載の方法。

【0077】

実施形態69．ライブラリー分子が1つ以上のニックを含む場合、フォワードプライマー及び/又はリバースプライマーが、ライブラリー分子の完全配列に対応するアンプリコンを生成することができない、実施形態48～68のいずれか1つに記載の方法。 40

【0078】

実施形態70．ニックを含むライブラリー分子から生成されたアンプリコンが、フォワードプライマー及び/又はリバースプライマーに結合するための配列を欠いている、実施形態69に記載の方法。

【0079】

実施形態71．ニックを含むライブラリー分子が、ニックを含まないライブラリー分子と比較して、増幅中により少ないアンプリコンを生成する、実施形態64～70のいずれ 50

か 1 つに記載の方法。

【 0 0 8 0 】

実施形態 7 2 . フォワードプライマー及びリバースプライマーをアニーリングする前に、ニックから二本鎖切断を生じさせることを更に含む、実施形態 6 4 ~ 7 1 のいずれか 1 つに記載の方法。

実施形態 7 3 . 二本鎖切断を生じさせることが、酵素反応を使用して行われる、実施形態 7 2 に記載の方法。

【 0 0 8 1 】

実施形態 7 4 . 酵素反応がエンドヌクレアーゼによって行われる、実施形態 7 3 に記載の方法。

【 0 0 8 2 】

実施形態 7 5 . エンドヌクレアーゼが T 7 エンドヌクレアーゼである、実施形態 7 4 に記載の方法。

【 0 0 8 3 】

実施形態 7 6 . 二本鎖切断を含むライブラリー分子が、増幅中にライブラリー分子の完全配列に対応するアンプリコンを生成しない、実施形態 7 2 ~ 7 5 のいずれか 1 つに記載の方法。

【 0 0 8 4 】

実施形態 7 7 . 二本鎖切断を含むライブラリー分子から生成されたアンプリコンが、フォワードプライマー及び / 又はリバースプライマーに結合するための配列を欠いている、実施形態 7 2 ~ 7 6 に記載の方法。

【 0 0 8 5 】

実施形態 7 8 . 蛍光を使用して DNA を含む試料中の DNA 損傷を定量化する方法であって、

a .

i . DNA を含む試料のアリコートと、

i i . 1 つ以上の DNA 修復酵素と、

i i i . 1 つ以上の d N T P が蛍光標識されている、d N T P と、を組み合わせることと、

とと、

b . 修復された DNA を調製することと、

c . d N T P からリン酸を脱リン酸化することと、

d . 修復された DNA をカルボキシレートビーズ又はセルロースビーズに結合させることと、

e . 再懸濁緩衝液を用いて、結合した修復 DNA をカルボキシレートビーズ又はセルロースビーズから溶出することと、

f . 修復された DNA の蛍光を測定して、DNA 損傷の量を決定することと、を含む、方法。

【 0 0 8 6 】

実施形態 7 9 . 修復された DNA のより高い蛍光が、より多い DNA 損傷を示す、実施形態 7 8 に記載の方法。

【 0 0 8 7 】

実施形態 8 0 . 修復された DNA の蛍光が、異なる量の DNA 損傷の範囲にわたって線形である、実施形態 7 8 又は実施形態 7 9 に記載の方法。

【 0 0 8 8 】

実施形態 8 1 . アッセイが、試料の操作の前後に同じ試料のアリコートを評価することによって、操作によって誘発された DNA 損傷を評価することができる、実施形態 7 8 ~ 8 0 のいずれか 1 つに記載の方法。

実施形態 8 2 . 操作が、試料の配列決定である、実施形態 8 1 に記載の方法。

【 0 0 8 9 】

実施形態 8 3 . 修復された DNA の蛍光を測定することが、修復された DNA の希釈物

10

20

30

40

50

の標準曲線を作成することと、修復されたDNAの希釈物の蛍光を測定することと、を含む、実施形態81又は実施形態82に記載の方法。

【0090】

実施形態84．修復されたDNAの蛍光を測定することが、修復されたDNAの蛍光を、蛍光標識された1つ以上のdNTPのみの希釈物の別個の標準曲線に対して比較し、修復されたDNAに含まれる蛍光色素分子の数を決定することを含む、実施形態78～83のいずれか1つに記載の方法。

【0091】

実施形態85．決定された蛍光色素分子の数を修復されたDNAの質量で割ることによって、修復されたDNAに含まれる蛍光色素分子の正規化された数を計算することを含む、実施形態84の方法。

【0092】

実施形態86．DNAが、ゲノムDNA、cDNA、又は断片化二本鎖DNAを含むライブラリーである、実施形態78～85のいずれか1つに記載の方法。

【0093】

実施形態87．DNAが、ゲノムDNA及びcDNAであり、方法が、DNA損傷の量を決定した後にライブラリーを調製することを更に含む、実施形態86の方法。

【0094】

実施形態88．DNA損傷の量が全ヌクレオチドの5%以下、4%以下、3%以下、2%以下、又は1%以下である場合にライブラリーが調製される、実施形態87に記載の方法。

【0095】

実施形態89．DNA損傷の量が全ヌクレオチドの5%以上、4%以上、3%以上、2%以上、又は1%以上である場合にライブラリーが調製されない、実施形態78～88のいずれか1つに記載の方法。

【0096】

実施形態90．蛍光を測定する前に、修復されたDNAをカルボキシレートビーズ又はセルロースビーズに結合させ、溶出することが2回以上行われる、実施形態78～89のいずれか1つに記載の方法。

【0097】

実施形態91．蛍光を測定する前に、修復されたDNAをカルボキシレートビーズ又はセルロースビーズに結合させ、溶出することが2回行われる、実施形態90に記載の方法。

【0098】

実施形態92．カルボキシレートビーズ又はセルロースビーズが磁性である、実施形態78～91のいずれか1つに記載の方法。

【0099】

実施形態93．修復されたDNAを調整することが、37℃で行われる、実施形態78～92のいずれか1つに記載の方法。

【0100】

実施形態94．修復されたDNAを調製することが、10分間以上、20分間以上、30分間以上、45分間以上、又は60分間以上行われる、実施形態78～93のいずれか1つに記載の方法。

【0101】

実施形態95．dNTPからリン酸を脱リン酸化することが、酵素を用いて行われる、実施形態78～94に記載の方法。

【0102】

実施形態96．dNTPからリン酸を脱リン酸化するための酵素が、エピアルカリホスファターゼ(SAP)又は仔ウシ腸アルカリホスファターゼ(CIP)である、実施形態78～95に記載の方法。

10

20

30

40

50

【0103】

実施形態97．1つ以上のDNA修復酵素が、DNAポリメラーゼを含む、実施形態78～96のいずれか1つに記載の方法。

【0104】

実施形態98．DNAポリメラーゼが、5' - 3'ポリメラーゼ活性を有するが、5' - 3'エキソヌクレアーゼ活性を欠いている、実施形態97に記載の方法。

【0105】

実施形態99．DNAポリメラーゼが、Bst DNAポリメラーゼ、ラージフラグメントである、実施形態97に記載の方法。

【0106】

実施形態100．1つ以上のDNA修復酵素が、リガーゼを含む、実施形態78～99のいずれかに記載の方法。

【0107】

実施形態101．リガーゼが、Taqリガーゼである、実施形態100に記載の方法。

【0108】

実施形態102．DNA損傷が、二本鎖DNA中のニックを含む、実施形態78～101のいずれか1つに記載の方法。

【0109】

実施形態103．1つ以上のDNA修復酵素が、T4ピリミジンダイマーグリコシラーゼ(PDG)を含む、実施形態78～102のいずれか1つに記載の方法。

【0110】

実施形態104．DNA損傷が、チミンダイマーを含む、実施形態78～103のいずれか1つに記載の方法。

【0111】

実施形態105．チミンダイマーが、紫外線照射によって誘発された、実施形態104に記載の方法。

【0112】

実施形態106．1つ以上のDNA修復酵素が、ウラシルDNAグリコシラーゼ(UDG)及び脱プリン又は脱ピリミジン部位リアーゼを含む、実施形態78～105のいずれかに記載の方法。

【0113】

実施形態107．DNA損傷が、ウラシルを含む、実施形態78～106のいずれか1つに記載の方法。

【0114】

実施形態108．1つ以上のDNA修復酵素が、ホルムアミドピリミジンDNAグリコシラーゼ(FPG)及び脱プリン又は脱ピリミジン部位リアーゼを含む、実施形態78～107のいずれかに記載の方法。

【0115】

実施形態109．DNA損傷が、酸化塩基を含む、実施形態78～108に記載の方法。

【0116】

実施形態110．dNTPが、dATP、dGTP、dCTP、及びdTTP又はdUTPを含む、実施形態78～109のいずれか1つに記載の方法。

【0117】

実施形態111．全てのdNTPが蛍光標識されている、実施形態78～110のいずれか1つに記載の方法。

【0118】

実施形態112．dUTP及びdCTPが蛍光標識されている、実施形態78～111に記載の方法。

【0119】

10

20

30

40

50

実施形態 113 . 蛍光標識が、Alexa Fluor 488、Alexa Fluor 546、Alexa Fluor 555、Alexa Fluor 633、フルオレseinイソチオシアネート (FITC)、又はテトラメチルローダミン - 5 - (及び 6) - イソチオシアネート (TRITC) である、実施形態 112 に記載の方法。

【0120】

追加の目的及び利点は、以下の説明において部分的に記載され、一部は説明から明白であるか、又は実践によって習得され得る。目的及び利点は、添付の特許請求の範囲において特に指摘される要素及び組み合わせによって実現及び達成されるであろう。

【0121】

前述の一般的な説明及び以下の詳細な説明はいずれも例示的かつ説明的のみであり、特許請求の範囲を限定するものではないことを理解されたい。 10

【0122】

本明細書に組み込まれ、本明細書の一部を構成する添付図面は、1つの(いくつかの)実施形態を図解し、明細書とともに本明細書に記載される原理を説明する役割を果たす。

【図面の簡単な説明】

【0123】

【図1】大きなインデル検出のための代表的な標準的方法を示す。このような方法は、長いアンプリコン(約10kb)に最適化されたPCR条件を用いた、切断部位の周りの低サイクルPCR(低サイクル、約1kb野生型アンプリコン)を伴う。増幅後、Nexteraライブラリー調製(LP)をPCRアンプリコンに対して行う。アンプリコン分析は、「デノボ」アンプリコンアセンブリ及び固有の遺伝子編集事象(すなわち、固有のアンプリコンを生成する事象)の定量化を含む。 20

【図2A】ユニバーサルUMI二本鎖(ds)DNAオリゴヌクレオチドを使用して調製できる長い増幅(Long Amp)挿入対照を要約する。UMI dsDNAオリゴヌクレオチドは、商業的に供給され得る(Integrated DNA TechnologiesのgBlock遺伝子断片など)(A)。このオリゴヌクレオチドを用いて、Long Amp挿入対照を調製することができる(B)。RS(RS1などの)は、制限部位を指す。N18は、18個のランダムヌクレオチドを含むUMI配列を指す。LA-fwd及びLA-revは、それぞれ、Long Amp反応のためのフォワードプライマー及びリバースプライマーを指す。対照1、2、3、及びnは、それぞれ0.2kb、1kb、2kb、及び10kbの挿入部を含む。10kb標準物質の明るい領域は、この標準物質が縮尺通りに描かれていないことを示す。 30

【図2B】ユニバーサルUMI二本鎖(ds)DNAオリゴヌクレオチドを使用して調製できる長い増幅(Long Amp)挿入対照を要約する。UMI dsDNAオリゴヌクレオチドは、商業的に供給され得る(Integrated DNA TechnologiesのgBlock遺伝子断片など)(A)。このオリゴヌクレオチドを用いて、Long Amp挿入対照を調製することができる(B)。RS(RS1などの)は、制限部位を指す。N18は、18個のランダムヌクレオチドを含むUMI配列を指す。LA-fwd及びLA-revは、それぞれ、Long Amp反応のためのフォワードプライマー及びリバースプライマーを指す。対照1、2、3、及びnは、それぞれ0.2kb、1kb、2kb、及び10kbの挿入部を含む。10kb標準物質の明るい領域は、この標準物質が縮尺通りに描かれていないことを示す。 40

【図3】上流ユニバーサルPCRアダプターアンプリコン及び下流ユニバーサルPCRアダプターアンプリコンを生成する方法を示す。これらのアンプリコンは、それぞれ、5'ユニバーサルオリゴヌクレオチド及び3'ユニバーサルオリゴヌクレオチドとして使用され得る。RS1及びRS2を含み、目的の標的配列の5'領域又は3'領域中の相補鎖に結合するプライマーを使用して、それぞれLA-ampフォワードプライマー及びリバースプライマーを用い(例えば、上流アンプリコン用のLA-fwd/RS1プライマー及び下流アンプリコン用のLA-rev/RS2を用い)、上流ユニバーサルPCRアダプターアンプリコン(5'領域)及び下流ユニバーサルPCRアダプターアンプリコン(3'領 50

域)を生成できる。示される「切断部位」は、遺伝子編集(例えば、CRISPR Cas系を用いる)を介して代表的な目的の配列に導入された切断部位を指し、挿入及び欠失は、遺伝子編集に使用されるそのような切断部位の周囲でしばしば起こり得る。他の目的の配列(例えば、挿入/欠失変異について評価されているがん患者由来の試料に含まれる配列)は、導入された切断部位を持たない。

【図4】テール付加PCRプライマーを用いて異なるサイズの挿入アンプリコンを調製する方法を示す。この方法は、制限酵素切断部位(RS)の配列を含み、目的の配列内のプライマー結合配列に結合する2つのプライマーのセット(すなわち、示される、2つのプライマー、例えばRS1/RS3配列を含むもの、又は2つのプライマー、例えばRS2/RS4を含むもの)を使用する。挿入アンプリコンのサイズ及び挿入アンプリコンは、目的の配列とのプライマー結合部位に基づくプライマーの選択によって制御することができる。この図において、上流は、目的の配列の5'部分における配列を指し、下流は、目的の配列の3'部分における配列を指す。挿入アンプリコン対は、上流挿入アンプリコン及び下流挿入アンプリコンを指し得る。10kb標準物質の明るい領域は、この標準物質が縮尺通りに描かれていないことを示す。

10

【図5】欠失標準物質を作製する方法を示す。目的の配列の相補鎖上のRS3及びRS4に結合するプライマーは、LA-ampフォワードプライマー及びLA-ampリバースプライマー(例えば、LA-fwd/RS3プライマー又はLA-rev/RS4)を使用する欠失アンプリコンの生成に使用され得る。欠失アンプリコン対は、上流欠失アンプリコン及び下流欠失アンプリコンを指し得る。次いで、RS3及びRS4に対応する制限部位を使用して、切断されたアンプリコンをユニバーサルUMI dsDNAオリゴヌクレオチド(図6Aに示す)にライゲーションするための適切な末端を生成し、図6Bに示すLong Amp欠失標準物質を生成することができる。

20

【図6A】ユニバーサルUMI二本鎖(ds)DNAオリゴヌクレオチドを使用して調製できる長い増幅(Long Amp)欠失対照を要約する。UMI dsDNAオリゴヌクレオチドは、商業的に供給され得る(Integrated DNA TechnologiesのgBlock遺伝子断片など)(A)。このオリゴヌクレオチドを用いて、Long Amp欠失標準物質を調製することができる(B)。対照1、2、3、及びnは、それぞれ、-20塩基対(bp)、-50bp、又は約-1kbの欠失部を含む。

【図6B】ユニバーサルUMI二本鎖(ds)DNAオリゴヌクレオチドを使用して調製できる長い増幅(Long Amp)欠失対照を要約する。UMI dsDNAオリゴヌクレオチドは、商業的に供給され得る(Integrated DNA TechnologiesのgBlock遺伝子断片など)(A)。このオリゴヌクレオチドを用いて、Long Amp欠失標準物質を調製することができる(B)。対照1、2、3、及びnは、それぞれ、-20塩基対(bp)、-50bp、又は約-1kbの欠失部を含む。

30

【図7】UMI配列の重複を回避するためにLong Amp反応において存在し得る対照入力の高質量を示す。

【図8A】異なる長さの核酸標準物質のプールに含まれ得る代表的な個々の標準物質を示す。これらの標準物質は全て、UMI、並びにLA-rev及びLA-fwdプライマー結合配列を含み得る。以下の表1は、標準物質に含まれる標識領域及びオリゴヌクレオチドについての説明を提供する。完全長標準物質は、5'ユニバーサルオリゴヌクレオチド及び3'ユニバーサルオリゴヌクレオチドを含み得る(100及び101)(A)。挿入標準物質は、5'ユニバーサルオリゴヌクレオチド、3'ユニバーサルオリゴヌクレオチド、並びにUMIと5'ユニバーサルオリゴヌクレオチドとの間の領域及びUMIと3'ユニバーサルオリゴヌクレオチドとの間の領域を含み得る(100、101、並びに102及び103)(B)。挿入標準物質はまた、UMIと5'ユニバーサルオリゴヌクレオチドとの間の領域又はUMIと3'ユニバーサルオリゴヌクレオチドとの間の領域のいずれかを含み得るが、両方の領域を含まない(100、101、及び103を含むが102を含まない8Bの下部標準物質に示される)。欠失標準物質は、5'部分的重複オリゴヌクレオチド及び3'部分的重複オリゴヌクレオチドを含み得る(104及び105)(C)。

40

50

欠失標準物質は、5'部分的重複オリゴヌクレオチド又は3'部分的重複オリゴヌクレオチドのいずれかを含み得るが、両方を含まない(104を含むが105を含まない8Cの下部標準物質に示される)。本明細書に記載されるように、核酸標準物質のプールは、本明細書に示される任意の又は全ての異なる種類の標準物質を含み得る。

【図8B】異なる長さの核酸標準物質のプールに含まれ得る代表的な個々の標準物質を示す。これらの標準物質は全て、UMI、並びにLA-rev及びLA-fwプライマー結合配列を含み得る。以下の表1は、標準物質に含まれる標識領域及びオリゴヌクレオチドについての説明を提供する。完全長標準物質は、5'ユニバーサルオリゴヌクレオチド及び3'ユニバーサルオリゴヌクレオチドを含み得る(100及び101)(A)。挿入標準物質は、5'ユニバーサルオリゴヌクレオチド、3'ユニバーサルオリゴヌクレオチド、並びにUMIと5'ユニバーサルオリゴヌクレオチドとの間の領域及びUMIと3'ユニバーサルオリゴヌクレオチドとの間の領域を含み得る(100、101、並びに102及び103)(B)。挿入標準物質はまた、UMIと5'ユニバーサルオリゴヌクレオチドとの間の領域又はUMIと3'ユニバーサルオリゴヌクレオチドとの間の領域のいずれかを含み得るが、両方の領域を含まない(100、101、及び103を含むが102を含まない8Bの下部標準物質に示される)。欠失標準物質は、5'部分的重複オリゴヌクレオチド及び3'部分的重複オリゴヌクレオチドを含み得る(104及び105)(C)。欠失標準物質は、5'部分的重複オリゴヌクレオチド又は3'部分的重複オリゴヌクレオチドのいずれかを含み得るが、両方を含まない(104を含むが105を含まない8Cの下部標準物質に示される)。本明細書に記載されるように、核酸標準物質のプールは、本明細書に示される任意の又は全ての異なる種類の標準物質を含み得る。

【図8C】異なる長さの核酸標準物質のプールに含まれ得る代表的な個々の標準物質を示す。これらの標準物質は全て、UMI、並びにLA-rev及びLA-fwプライマー結合配列を含み得る。以下の表1は、標準物質に含まれる標識領域及びオリゴヌクレオチドについての説明を提供する。完全長標準物質は、5'ユニバーサルオリゴヌクレオチド及び3'ユニバーサルオリゴヌクレオチドを含み得る(100及び101)(A)。挿入標準物質は、5'ユニバーサルオリゴヌクレオチド、3'ユニバーサルオリゴヌクレオチド、並びにUMIと5'ユニバーサルオリゴヌクレオチドとの間の領域及びUMIと3'ユニバーサルオリゴヌクレオチドとの間の領域を含み得る(100、101、並びに102及び103)(B)。挿入標準物質はまた、UMIと5'ユニバーサルオリゴヌクレオチドとの間の領域又はUMIと3'ユニバーサルオリゴヌクレオチドとの間の領域のいずれかを含み得るが、両方の領域を含まない(100、101、及び103を含むが102を含まない8Bの下部標準物質に示される)。欠失標準物質は、5'部分的重複オリゴヌクレオチド及び3'部分的重複オリゴヌクレオチドを含み得る(104及び105)(C)。欠失標準物質は、5'部分的重複オリゴヌクレオチド又は3'部分的重複オリゴヌクレオチドのいずれかを含み得るが、両方を含まない(104を含むが105を含まない8Cの下部標準物質に示される)。本明細書に記載されるように、核酸標準物質のプールは、本明細書に示される任意の又は全ての異なる種類の標準物質を含み得る。

【0124】

【表1】

表1：標識の説明	
標識	説明
100	5'ユニバーサルオリゴヌクレオチド
101	3'ユニバーサルオリゴヌクレオチド
102	UMIと5'ユニバーサルオリゴヌクレオチドとの間の領域
103	UMIと3'ユニバーサルオリゴヌクレオチドとの間の領域
104	5'部分的重複オリゴヌクレオチド
105	3'部分的重複オリゴヌクレオチド

40

【図9】長いライブラリーにおいてDNA損傷を評価するための定量的PCR (qPCR) アッセイを要約する。このアッセイは、ライブラリー分子に含まれるヘアピンアダプター内の配列に結合するフォワードプライマー及びリバースプライマーを使用する。DNA損傷 (例えば、ニック) を含まないライブラリーは、より多くのシグナルを生成する (すなわち、より多くの完全長アンプリコンを産生する)。図に示すように、例示的なアッセイは、LongAmp PCRに最適化されたポリメラーゼ (例えば、PrimeStar GXL DNAポリメラーゼ、Takara) を用いた指数関数的増幅を含み得る。

【図10A】異なる濃度のニックで処理されたライブラリーに対するQCアッセイによる平均定量サイクル (Cq) 及び損傷%の結果を示す。10ngのライブラリーについてのCq (A) 及び損傷% (B) の結果、並びに20ngのライブラリーについてのCq (C) 及び損傷% (D) の結果を示す。

10

【図10B】異なる濃度のニックで処理されたライブラリーに対するQCアッセイによる平均定量サイクル (Cq) 及び損傷%の結果を示す。10ngのライブラリーについてのCq (A) 及び損傷% (B) の結果、並びに20ngのライブラリーについてのCq (C) 及び損傷% (D) の結果を示す。

【図10C】異なる濃度のニックで処理されたライブラリーに対するQCアッセイによる平均定量サイクル (Cq) 及び損傷%の結果を示す。10ngのライブラリーについてのCq (A) 及び損傷% (B) の結果、並びに20ngのライブラリーについてのCq (C) 及び損傷% (D) の結果を示す。

【図10D】異なる濃度のニックで処理されたライブラリーに対するQCアッセイによる平均定量サイクル (Cq) 及び損傷%の結果を示す。10ngのライブラリーについてのCq (A) 及び損傷% (B) の結果、並びに20ngのライブラリーについてのCq (C) 及び損傷% (D) の結果を示す。

20

【図11】例えばVibrio vulnificusヌクレアーゼ (VVN) とT7エンドヌクレアーゼ変異体との組み合わせを用いるなどの、ライブラリー分子中のニックを二本鎖切断に変換する方法の結果を示す。Endo = エンドヌクレアーゼ。

【図12A】ライブラリーに対するエンドヌクレアーゼ変異体での処理の有無によるCq値の差を要約する。(A) Cq値の概要。(B) TapeStation (登録商標)、Agilentを使用した自動電気泳動結果の概要。

【図12B】ライブラリーに対するエンドヌクレアーゼ変異体での処理の有無によるCq値の差を要約する。(A) Cq値の概要。(B) TapeStation (登録商標)、Agilentを使用した自動電気泳動結果の概要。

30

【図13A】SMRTbell 鋳型を定量的PCR (qPCR) で評価し、次いでPacBio Sequel 2システムで配列決定して、qPCR Cqが配列決定基準と相関するかどうかを決定した際の結果を示す。試料を最も低いCqから最も高いCqの順に並べる。(A) 平均Cq。(B) 総出力。(C) 変動 (P1%)。相関は、qPCR Cq及び総出力 (ギガベース、GB) について観察され、より低いCqは、より高い出力を示す (最も低いCqのライブラリー8の1つの外れ値を除く)。一般に、ライブラリーの平均Cq値は2~3であった。qPCRの結果は、ライブラリー13が低品質であることを予測し、これは比較的不十分な配列決定結果によって確認される。

40

【図13B】SMRTbell 鋳型を定量的PCR (qPCR) で評価し、次いでPacBio Sequel 2システムで配列決定して、qPCR Cqが配列決定基準と相関するかどうかを決定した際の結果を示す。試料を最も低いCqから最も高いCqの順に並べる。(A) 平均Cq。(B) 総出力。(C) 変動 (P1%)。相関は、qPCR Cq及び総出力 (ギガベース、GB) について観察され、より低いCqは、より高い出力を示す (最も低いCqのライブラリー8の1つの外れ値を除く)。一般に、ライブラリーの平均Cq値は2~3であった。qPCRの結果は、ライブラリー13が低品質であることを予測し、これは比較的不十分な配列決定結果によって確認される。

【図13C】SMRTbell 鋳型を定量的PCR (qPCR) で評価し、次いでPacBio Sequel 2システムで配列決定して、qPCR Cqが配列決定基準と相

50

関するかどうかを決定した際の結果を示す。試料を最も低いCqから最も高いCqの順に並べる。(A)平均Cq。(B)総出力。(C)変動(P1%)。相関は、qPCR Cq及び総出力(ギガベース、Gb)について観察され、より低いCqは、より高い出力を示す(最も低いCqのライブラリー8の1つの外れ値を除く)。一般に、ライブラリーの平均Cq値は2~3であった。qPCRの結果は、ライブラリー13が低品質であることを予測し、これは比較的不十分な配列決定結果によって確認される。

【図14A】qPCRで評価され、次いでPacBio Sequel 2システムで配列決定されたSMRTbell鑄型の別のセットを用いたデータを示す。(A)平均Cq値、試料は最も低いCqから最も高いCqへの順。(B)総出力(Gb)。(C)P1パーセント。相関は、qPCR Cq及び総出力について観察され、より低いCqは、より高い出力を示す(最も低いCqのライブラリー14の1つの外れ値を除く)。ほとんどのライブラリーの平均Cq値は3~4であった。qPCRは、ライブラリー10が低品質であることを予測し、これは配列決定によって確認される。

10

【図14B】qPCRで評価され、次いでPacBio Sequel 2システムで配列決定されたSMRTbell鑄型の別のセットを用いたデータを示す。(A)平均Cq値、試料は最も低いCqから最も高いCqへの順。(B)総出力(Gb)。(C)P1パーセント。相関は、qPCR Cq及び総出力について観察され、より低いCqは、より高い出力を示す(最も低いCqのライブラリー14の1つの外れ値を除く)。ほとんどのライブラリーの平均Cq値は3~4であった。qPCRは、ライブラリー10が低品質であることを予測し、これは配列決定によって確認される。

20

【図14C】qPCRで評価され、次いでPacBio Sequel 2システムで配列決定されたSMRTbell鑄型の別のセットを用いたデータを示す。(A)平均Cq値、試料は最も低いCqから最も高いCqへの順。(B)総出力(Gb)。(C)P1パーセント。相関は、qPCR Cq及び総出力について観察され、より低いCqは、より高い出力を示す(最も低いCqのライブラリー14の1つの外れ値を除く)。ほとんどのライブラリーの平均Cq値は3~4であった。qPCRは、ライブラリー10が低品質であることを予測し、これは配列決定によって確認される。

【図15A】配列決定前のいくつかのPacBio SMRTbellライブラリーについてのqPCR QCアッセイ結果、及び総Gb出力と相関したデータを示す。総出力はより低いCq値で増加し、このQCアッセイが配列決定性能を予測するための有用なツールとして役立つことを示唆している。ライブラリー20(A)、ライブラリー21(B)、及びライブラリー22(C)由来のライブラリー画分(F#)についてのCq値及びGb測定値。

30

【図15B】配列決定前のいくつかのPacBio SMRTbellライブラリーについてのqPCR QCアッセイ結果、及び総Gb出力と相関したデータを示す。総出力はより低いCq値で増加し、このQCアッセイが配列決定性能を予測するための有用なツールとして役立つことを示唆している。ライブラリー20(A)、ライブラリー21(B)、及びライブラリー22(C)由来のライブラリー画分(F#)についてのCq値及びGb測定値。

【図15C】配列決定前のいくつかのPacBio SMRTbellライブラリーについてのqPCR QCアッセイ結果、及び総Gb出力と相関したデータを示す。総出力はより低いCq値で増加し、このQCアッセイが配列決定性能を予測するための有用なツールとして役立つことを示唆している。ライブラリー20(A)、ライブラリー21(B)、及びライブラリー22(C)由来のライブラリー画分(F#)についてのCq値及びGb測定値。

40

【図16】DNA損傷検出ワークフローを示す。このアッセイのシグナル対ノイズ比は、組み込まれていない蛍光ヌクレオチドの非特異的結合を大きく減少させるために、エピアルカリホスファターゼ(SAP)消化及びストリンジェントな二重SPRIビーズベースの精製工程(すなわち、カルボキシレートビーズを用いた2回の精製)の両方を用いることによって増加した。

50

【図17】SAP消化及び単一のSPRIビーズベースの精製工程の結果を示す。単一のSPRIで精製された切断ゲノムDNAは、SAP処理なし(-SAP)とは対照的に、精製前にSAPで処理した場合(+SAP)、蛍光ヌクレオチドの非特異的結合の減少を示した。

【図18】2回のビーズベースの精製工程が蛍光ヌクレオチドの非特異的結合を実質的に減少させたことを示す。

【図19A】市販の修復ミックス(PreCR Repair mix(NEB)、パネル(A)に示す)と、Taqリガーゼ(40U)、Bstポリメラーゼラージフラグメント(8U)、及びT4 PDG(1U)を含むDNA修復酵素ミックスを用いる本発明の方法(パネル(B)に示す)との有効性の比較を示す。

10

【図19B】市販の修復ミックス(PreCR Repair mix(NEB)、パネル(A)に示す)と、Taqリガーゼ(40U)、Bstポリメラーゼラージフラグメント(8U)、及びT4 PDG(1U)を含むDNA修復酵素ミックスを用いる本発明の方法(パネル(B)に示す)との有効性の比較を示す。

【図20】ゲノムDNA試料に対する紫外線(UV)損傷の測定を示す。光のエネルギーが増加し、曝露時間が増加するにつれて、Taqリガーゼ、Bstポリメラーゼ、及びUV損傷特異的修復酵素であるT4ピリミジンダイマーグリコシラーゼ(T4 PDG)を含むカスタムDNA修復酵素ミックスで修復された試料において、蛍光の量もまた増加する。

【図21】ゲノムDNA試料に対するニックング損傷の測定を示す。ニックング酵素(Nt.BspQI)の量が増加するにつれて、本発明のアッセイを使用してTaqリガーゼ及びBstポリメラーゼで修復された試料中の蛍光シグナルも一般に増加する。

20

【発明を実施するための形態】

【0125】

長い増幅PCRは、標的核酸由来の目的の配列における標的化された長いインデル検出のために使用することができる。しかしながら、PCRは、小さな挿入及び欠失変異を有するものなどのより小さなアンプリコンに偏っており、長い挿入などのより長いアンプリコンに偏っている。長いアンプリコンは、短いアンプリコンと比較して、PCRサイクルにおいて長いアンプリコンが生成される可能性がより低い核酸の新鎖合成に時間がかかるので、このバイアスはPCR法に固有である。更に、長いアンプリコンは、事象が複製を停止し得る前に、完全なアンプリコンを産生する成功率がより低い。言い換えれば、長いアンプリコンの増幅は、より短いアンプリコンよりも高い割合で失敗する可能性がある。例えば、アンプリコンを産生するためにポリメラーゼを作用させるべき時間が長ければ長いほど、ランダムな脱落、DNA損傷への遭遇、又は処理速度を考慮する場合に時間不足のために、ポリメラーゼがアンプリコンの末端に到達しない可能性が高くなる。

30

【0126】

長いアンプリコンに対する既知のバイアスのために、長い増幅(Long Amp)PCRは、異なる事象の相対頻度を正確に決定するために使用することができない。したがって、異なる突然変異に関連するアンプリコンのサイズが異なって増幅されるため、Long Amp増幅の結果は、元の標的核酸試料中の特定の突然変異の相対数を定量化することができない。

40

【0127】

本明細書に記載される標準物質及び方法は、このアンプリコンサイズバイアスを正規化するのに役立つ。

【0128】

更に、本開示はまた、ライブラリー品質を評価するための品質管理(QC)方法も記載する。いくつかの実施形態では、ライブラリー、例えば長リード配列決定のためのものは、配列決定の前に評価される。いくつかの実施形態では、ライブラリーは、挿入部の両末端にヘアピンアダプターを有する二本鎖DNA挿入部を含むライブラリー分子を含む。いくつかの実施形態では、ライブラリーは、標的DNAを断片化し、タグメンテーション又

50

はライゲーションなどにより、断片の両端にヘアピンアダプターを組み込むことによって作製される。

【0129】

I. アンプリコンサイズバイアスを正規化するための標準物質

いくつかの実施形態では、アンプリコンサイズバイアスを正規化する方法において、異なる長さの核酸標準物質のプールを使用することができる。いくつかの実施形態では、これらの拡散標準物質は、固有分子識別子 (UMI) を含む。

【0130】

いくつかの実施形態では、核酸のプールは、目的の配列に含まれる様々な異なる配列を含み得る。

【0131】

いくつかの実施形態では、プール中の標準物質の数は、増幅反応によって生成されるアンプリコンの数より多い。いくつかの実施形態では、増幅反応は、目的の配列の増幅である。

【0132】

いくつかの実施形態では、標準物質の少なくとも第1の部分は、標準物質の1つのプール由来であり、標準物質の少なくとも第2の部分は、標準物質の別のプール由来である。

【0133】

いくつかの実施形態では、標準物質は、二本鎖である。いくつかの実施形態では、標準物質は、二本鎖DNAを含む。いくつかの実施形態では、各標準物質は、異なるUMIを含む。

【0134】

いくつかの実施形態では、増幅プライマー結合配列は、各標準物質の一方若しくは両方の末端に、又は各標準物質の一方若しくは両方の末端に近接して含まれる。本明細書を通して、「一方又は両方の末端に近接して」とは、末端から10ヌクレオチド以内を意味する。いくつかの実施形態では、増幅プライマー結合配列は、各標準物質の一方又は両方の末端に含まれる。いくつかの実施形態では、増幅プライマー結合配列は、各標準物質の一方又は両方の末端の1、2、3、4、5、6、7、8、又は9ヌクレオチドを含む。いくつかの実施形態では、標準物質は、その3'末端及びその5'末端の両方に増幅プライマー結合配列を含む。いくつかの実施形態では、標準物質は、その3'末端に対して3'末端に異なる増幅プライマー結合配列を含む。いくつかの実施形態では、標準物質は、UMIの5'側に1つ以上のオリゴヌクレオチドを含む。いくつかの実施形態では、標準物質は、UMIの3'側に1つ以上のオリゴヌクレオチドを含む。いくつかの実施形態では、標準物質は、UMIの5'側の1つ以上のオリゴヌクレオチド及びUMIの3'側の1つ以上のオリゴヌクレオチドを含む。

【0135】

A. UMI

いくつかの実施形態では、標準物質のプール中の標準物質はそれぞれ、UMIを含む。

【0136】

いくつかの実施形態では、UMIは、標準物質の5'末端及び/若しくは3'末端になく、又は標準物質の5'末端及び/若しくは3'末端に近接していない。いくつかの実施形態では、標準物質内の中心に位置するUMIは、標準物質の断片化(例えばタグメンテーションによる)が、UMI及び標準物質の残り由来の配列の全部又は一部(UMIの5'及び/又は3'のいずれか)を含む断片を生じる確率を増加させる。本明細書中で使用される場合、「中心に」位置する特徴は、標準物質の中心から10ヌクレオチド以内の位置にある特徴の中心をいう。いくつかの実施形態では、標準物質内の中心に位置するUMIは、標準物質の中心の1、2、3、4、5、6、7、8、又は9ヌクレオチド以内にUMIの中心を有する。

【0137】

対照的に、目的の配列の5'末端及び/又は3'末端の近位にUMIを配置すると、UMI

10

20

30

40

50

Iのみを含み、標準物質の残りの部分由来の追加の配列を含まない断片の割合が高くなる可能性がある。

【0138】

いくつかの実施形態では、UMIは、同じLong Amp標準物質から生成されるアンプリコンを同定するために使用される。言い換えれば、UMI及び上流/下流挿入ジャンクション塩基を含む標準物質の配列決定は、それぞれ、標準物質の固有の分子数及び対照同一性を提供することができる。これは、同じ標準物質から生成された各アンプリコンが同じ固有のUMIを有し、Long Amp標準物質から生成された他のアンプリコンが異なるUMIを有するためである。

【0139】

いくつかの実施形態では、UMIはランダム塩基対を含み、その結果、各固有のUMIは、プール内の他のUMIとは異なる配列を含む。いくつかの実施形態では、UMIは、10(N10)以上、12(N12)以上、14(N14)以上、16(N16)以上、18(N18)以上、20(N20)以上、又は22(N22)以上のランダム塩基対を含む。いくつかの実施形態では、UMIは、18塩基対(N18)を含む。いくつかの実施形態では、標準物質のプールに含まれるUMIは、16~20塩基対を含む配列のランダムセットである。

【0140】

多数のUMIを有するUMIプールの使用は、(UMI衝突を回避するのに役立つ)。より長いUMI(すなわち、N10の代わりにN18)を有することも、UMI衝突の機会を減少させる。

【0141】

本明細書で使用される場合、「UMI衝突」は、同じ配列及び同じUMIバーコードを有するが、2つの異なるゲノム分子に由来する2つのリードを観察する事象を指す。アンプリコン配列決定では、ゲノム中の特定の位置が何度も配列決定され、ゲノム全体の配列決定よりもはるかに大きいシーケンス深度をもたらす(Clement et al., Bioinformatics, 34, 2018, i202-i210参照)。このシーケンス深度に基づいて、異なるゲノム分子由来の多くの対立遺伝子は、同じ配列を共有し得、UMI衝突の可能性は、全ゲノム配列決定と比較してアンプリコン配列決定についてはるかに高い。

【0142】

いくつかの実施形態では、標準物質のプールは、 1×10^{10} 以上、 10×10^{10} 以上、又は 100×10^{10} 以上の標準物質を含み、各標準物質は異なるUMIを含む。図7は、必要とされるUMIを含む合成二本鎖DNAの量を含む、 6.87×10^{10} 個のUMIを含む実験を準備するための計算を示す。

【0143】

いくつかの実施形態では、本明細書に記載されるように、標準物質中のUMIは、比較的安価な市販の試薬に由来し得る。いくつかの実施形態では、UMIを含む二本鎖オリゴヌクレオチドはまた、標準物質の調製に使用するための1つ以上の制限酵素切断部位を含む。

【0144】

例えば、以下に記載されるように、代表的な合成dsDNAオリゴヌクレオチドは、挿入標準物質を調製するため(図2A)及び欠失標準物質を調製するため(図6A)に示される。いくつかの実施形態では、合成dsDNAオリゴヌクレオチドは、UMI及び制限酵素切断部位(又は図2A及び図6Aに示されるようなRS3及びRS4などの制限部位)を含む。いくつかの実施形態では、制限酵素切断部位を使用してオリゴヌクレオチドを切断し、次いで他のオリゴヌクレオチドにライゲーションして最終標準物質を調製することができる。UMI dsDNAオリゴヌクレオチドの供給源として、gBlock遺伝子断片(Integrated DNA Technologies)が挙げられる。

【0145】

10

20

30

40

50

B. 目的の配列

本明細書で使用される場合、「目的の配列」は、ユーザーが調査したい任意の配列であり得る。いくつかの実施形態では、目的の配列は、遺伝子編集に供されている。例えば、ユーザーは、遺伝子編集又は他の突然変異誘発（化学的突然変異誘発など）の方法を実施している可能性があり、目的の配列中の異なる突然変異（野生型配列と共に）を評価したいと望む。

【0146】

いくつかの実施形態では、遺伝子編集は、CRISPR Cas法を用いて行われる。いくつかの実施形態では、CRISPR Cas切断部位は、目的の配列中に存在する。いくつかの実施形態では、挿入変異又は欠失変異は、目的の配列内の切断部位の近くで生じる可能性が高い。例えば、図5は、CRISPR Casなどの遺伝子編集の方法を使用して導入された目的の配列内に存在する切断部位を示す。いくつかの目的の配列、例えばインデル突然変異について評価されている患者由来の腫瘍学的試料の配列は、遺伝子編集法によって導入された切断部位を有さない。

【0147】

いくつかの実施形態では、目的の配列は、目的の配列の5'末端及び/若しくは3'末端にない、又は目的の配列の5'末端及び/若しくは3'末端に近接していない制限酵素切断部位を含む。いくつかの実施形態では、このような切断部位は、標準物質を生成する際に有用であり得るか、又は目的の配列を評価するために使用され得る。

【0148】

いくつかの実施形態では、目的の配列は、長い増幅プライマー（すなわち、LA-fwd及びLA-revプライマー）に結合できるプライマー結合配列を含む。いくつかの実施形態では、ユーザーは、適切なLA-fwd及びLA-revプライマーを調製するために、目的の配列を評価し得る。

【0149】

いくつかの実施形態では、目的の配列は、挿入変異又は欠失変異を含み得る。例えば、目的の配列は、挿入変異を含み得るか、又は欠失変異であり得る（すなわち、目的の配列の全長配列を含まない）。

【0150】

本明細書で使用される場合、目的の「野生型」配列は、インデル突然変異を含まない目的の配列を指す。言い換えれば、野生型配列は、挿入変異を含まず、欠失変異も含まない配列を指す。本明細書で使用される場合、「野生型アンプリコン」は、目的の野生型配列を含むアンプリコンである。

【0151】

目的の配列は、任意のタイプの核酸配列であり得る。いくつかの実施形態では、目的の配列は、遺伝子編集法（CRISPRなど）に供されており、ユーザーは、固有の遺伝子編集事象の分析を望む。いくつかの実施形態では、遺伝子編集に供された目的の配列は、図3、図5、及び図6Bの代表例に示されるような「切断部位」を含み得る。そのような遺伝子編集法は、ユーザーが特徴付けることを望み得る様々な異なる種類のインデル突然変異をもたらし得る。

【0152】

いくつかの実施形態では、がん及び生殖系インデル変異を含む目的の配列は、転移性エレメント由来の挿入と同様に、この方法によって評価することができる。そのような実施形態では、目的の配列は、遺伝子編集による切断部位を含まなくてもよい。

【0153】

いくつかの実施形態では、目的の配列は、目的の遺伝子、例えば、がんに関連することが知られている遺伝子の全て又は一部であり得る。当業者は、患者が目的の配列を含む遺伝子中に有し得るインデルを特徴付けること、及び/又は異なる変異の相対量を特徴付けることを望む場合がある。例えば、当業者は、患者の試料由来の目的の配列中に存在する大きな挿入変異の数を特徴付けたい場合がある。

10

20

30

40

50

【 0 1 5 4 】

C. ユニバーサルオリゴヌクレオチドを含む標準物質

いくつかの実施形態では、核酸標準物質のプール内の全て又はいくつかの標準物質は、5'ユニバーサルオリゴヌクレオチド及び3'ユニバーサルオリゴヌクレオチドを含む。本明細書中で使用される場合、「ユニバーサルオリゴヌクレオチド」とは、このプール中の全ての標準物質に含まれるオリゴヌクレオチドをいう。本明細書で使用される場合、「5'ユニバーサルオリゴヌクレオチド」は、標準物質に含まれるUMIの5'側にあるオリゴヌクレオチドである(図8の100として表される)。本明細書で使用される場合、「3'ユニバーサルオリゴヌクレオチド」は、標準物質に含まれるUMIの3'側にあるオリゴヌクレオチドである(図8の101として表される)。

10

【 0 1 5 5 】

いくつかの実施形態では、標準物質の少なくとも第1の部分は、標準物質の1つのプール由来であり、標準物質の少なくとも第2の部分は、標準物質の別のプール由来である。言い換えれば、各標準物質が5'ユニバーサルオリゴヌクレオチド及び3'ユニバーサルオリゴヌクレオチドを含む標準物質のプールを、5'ユニバーサルオリゴヌクレオチド及び/又は3'ユニバーサルオリゴヌクレオチドを含まない標準物質の異なるプールと組み合わせることができる。

【 0 1 5 6 】

いくつかの実施形態では、核酸標準物質のプールは、異なる長さの標準物質を含み、核酸標準物質は、固有分子識別子(UMI)、並びに、5'ユニバーサルオリゴヌクレオチドであって、5'ユニバーサルオリゴヌクレオチドは全ての標準物質について同じである、5'ユニバーサルオリゴヌクレオチドと、3'ユニバーサルオリゴヌクレオチドであって、3'ユニバーサルオリゴヌクレオチドは全ての標準物質について同じである、3'ユニバーサルオリゴヌクレオチドと、UMIと5'ユニバーサルオリゴヌクレオチドとの間、及び/又は、UMIと3'ユニバーサルオリゴヌクレオチドとの間の少なくとも1つの領域と、を含み、少なくとも1つの領域の長さは、標準物質の長さを決定する。UMIと5'ユニバーサルオリゴヌクレオチドとの間の領域は図8Bに102として示され、UMIと3'ユニバーサルオリゴヌクレオチドとの間の領域は図8Bに103として示される。

20

【 0 1 5 7 】

いくつかの実施形態では、5'ユニバーサルオリゴヌクレオチド及び3'ユニバーサルオリゴヌクレオチドを含み、追加の配列(例えば、UMIと5'ユニバーサルオリゴヌクレオチドとの間の領域及び/又はUMIと3'ユニバーサルオリゴヌクレオチドとの間の領域)も含む標準物質は、「挿入標準物質」と称され得る。これは、挿入標準物質が目的の野生型配列よりも長さが長い可能性があるためである。これらの挿入突然変異は対象の野生型配列よりも大きいため、このようにして、挿入標準物質は、対象の野生型配列における挿入突然変異のアンプリコンサイズバイアスを正規化するために制御することができる。

30

【 0 1 5 8 】

いくつかの実施形態では、プールは、UMI、並びに、5'ユニバーサルオリゴヌクレオチドであって、5'ユニバーサルオリゴヌクレオチドは全ての標準物質について同じである、5'ユニバーサルオリゴヌクレオチドと、3'ユニバーサルオリゴヌクレオチドであって、3'ユニバーサルオリゴヌクレオチドが全ての標準物質について同じである、3'ユニバーサルオリゴヌクレオチドと、を含む、核酸標準物質を更に含み、更なる核酸標準物質は、UMIと5'ユニバーサルオリゴヌクレオチドとの間又はUMIと3'ユニバーサルオリゴヌクレオチドとの間の少なくとも1つの領域を含まない。5'ユニバーサルオリゴヌクレオチド(100)及び3'ユニバーサルオリゴヌクレオチド(101)を含む標準物質は、図8Aに示されるように、完全長標準物質と称され得る。完全長標準物質は、挿入又は欠失変異のいずれも有さない目的の野生型配列(すなわち、インデルを有さない野生型配列)と同様の長さを有し得る。

40

【 0 1 5 9 】

50

いくつかの実施形態では、UMIと5'ユニバーサルオリゴヌクレオチドとの間及び/又はUMIと3'ユニバーサルオリゴヌクレオチドとの間の少なくとも1つの領域は、挿入標準物質の長さを決定する。いくつかの実施形態では、UMIと5'ユニバーサルオリゴヌクレオチドとの間及び/又はUMIと3'ユニバーサルオリゴヌクレオチドとの間の少なくとも1つの領域は、目的の挿入変異の潜在的な長さに対応するキロベース(kb)数を含む。いくつかの実施形態では、UMIと5'ユニバーサルオリゴヌクレオチドとの間及び/又はUMIと3'ユニバーサルオリゴヌクレオチドとの間の少なくとも1つの領域は、0.2kb~10kbを含む。

【0160】

5'ユニバーサルオリゴヌクレオチド及び/又は3'ユニバーサルオリゴヌクレオチドは10、目的の配列に含まれる配列を含み得る。いくつかの実施形態では、5'ユニバーサルオリゴヌクレオチド及び/又は3'ユニバーサルオリゴヌクレオチドはそれぞれ、目的の配列から増幅されたアンプリコンを含む。言い換えれば、5'ユニバーサルオリゴヌクレオチド及び/又は3'ユニバーサルオリゴヌクレオチドは、図3に示されるように、増幅によって調製され得る。

【0161】

5'ユニバーサルオリゴヌクレオチドが増幅によって調製される場合、それは「5'ユニバーサルPCRアダプターアンプリコン」又は「上流ユニバーサルPCRアダプターアンプリコン」と称され得る。図3は、代表的な上流ユニバーサルPCRアダプターアンプリコンが、長い増幅フォワードプライマー(LA-fwd)及び、目的の配列に結合し、かつ制限酵素切断部位(RS1)を含むプライマーを使用して生成され得る方法を示す。20

【0162】

3'ユニバーサルオリゴヌクレオチドが増幅によって調製される場合、それは「3'ユニバーサルPCRアダプターアンプリコン」又は「下流ユニバーサルPCRアダプターアンプリコン」と称され得る。図3は、代表的な下流ユニバーサルPCRアダプターアンプリコンが、長い増幅リバースプライマー(LA-rev)及び、目的の配列に結合し、かつ制限酵素切断部位(RS2)を含むプライマーを使用して生成され得る方法を示す。

【0163】

いくつかの実施形態では、上流ユニバーサルPCRアダプターアンプリコン及び下流ユニバーサルPCRアダプターアンプリコンを適切な制限酵素で切断して(図3に示す例ではRS1及びRS2で切断できる)、UMI、並びに、5'ユニバーサルオリゴヌクレオチドであって、5'ユニバーサルオリゴヌクレオチドは全ての標準物質について同じである、5'ユニバーサルオリゴヌクレオチドと、3'ユニバーサルオリゴヌクレオチドであって、3'ユニバーサルオリゴヌクレオチドは全ての標準物質について同じである、3'ユニバーサルオリゴヌクレオチドと、を含む、標準物質を調製することができる。この切断は、標準物質を作製する方法の記述において以下で考察されるように、これらのアンプリコンを標準物質の他の部分(例えば、UMIと5'ユニバーサルオリゴヌクレオチドとの間及び/又はUMIと3'ユニバーサルオリゴヌクレオチドとの間の領域)にライゲーションするのに適合する末端を生成し得る。30

【0164】

いくつかの実施形態では、UMIと5'ユニバーサルオリゴヌクレオチドとの間及び/又はUMIと3'ユニバーサルオリゴヌクレオチドとの間の少なくとも1つの領域は、それぞれ任意の配列を含む。本明細書中で使用される場合、「任意の配列」とは、ヌクレオチドを含む任意の配列をいい、特定の核酸配列がこの任意の配列に含まれることを何ら必要としない。例えば、当業者は、任意の配列がランダムであり、目的の配列に関連しない挿入標準物質を調製することを所望し得る。別の実施形態では、任意の配列は、ランダムではない既知の配列であってよいが、それはまた、目的の配列に関連しない(例えば、関連しない遺伝子配列)。任意の配列を含む標準物質を使用して、挿入変異のアンプリコンサイズバイアスについて正規化することができ、これは、このバイアスの多くがアンプリコンサイズに関連し、挿入配列に含まれる正確な配列に関連しないためである。いくつか40

の実施形態では、任意の配列は、二本鎖である。

【0165】

いくつかの実施形態では、UMIと5'ユニバーサルオリゴヌクレオチドとの間及び/又はUMIと3'ユニバーサルオリゴヌクレオチドとの間の少なくとも1つの領域は、それぞれ目的の配列から増幅されたアンプリコンを含む。言い換えれば、UMIと5'ユニバーサルオリゴヌクレオチドとの間及び/又はUMIと3'ユニバーサルオリゴヌクレオチドとの間の領域を、増幅によって調製することができる。いくつかの実施形態では、この増幅は、図4に示されるように、目的の配列由来である。

【0166】

1. 挿入アンプリコン

本明細書で使用される場合、UMIと5'ユニバーサルオリゴヌクレオチドとの間の領域は、増幅によって調製されるとき、「5'挿入アンプリコン」又は「上流挿入アンプリコン」と称され得る。図4は、目的の配列に結合し、かつ制限酵素切断部位(RS1及びRS3)を含むプライマーを使用する、代表的な上流挿入アンプリコンが生成され得る方法を示す。

10

【0167】

本明細書で使用される場合、UMIと3'ユニバーサルオリゴヌクレオチドとの間の領域は、増幅によって調製されるとき、「3'挿入アンプリコン」又は「下流挿入アンプリコン」と称され得る。図4は、制限酵素切断部位(RS2及びRS4)を使用する、代表的な上流挿入アンプリコンが生成され得る方法を示す。

20

【0168】

いくつかの実施形態では、挿入アンプリコンを調製するために使用されるリバースプライマー及びフォワードプライマーは、挿入アンプリコンのサイズを決定する。いくつかの実施形態では、単一のプライマー対は、所望のサイズの挿入アンプリコンを生成する。

【0169】

本明細書で使用される場合、「挿入アンプリコン」は、5'挿入アンプリコン又は3'挿入アンプリコンのいずれかであるアンプリコンを指し得る。一般に、「挿入アンプリコン」は、標準物質中の配置によって限定されない。

【0170】

いくつかの実施形態では、標準物質は、上流挿入アンプリコン及び下流挿入アンプリコンの両方を含む(図4に示される)。これらは、「挿入アンプリコン対」と称され得る。しかし、標準物質はまた、上流挿入アンプリコン又は下流挿入アンプリコンのいずれかのみを含み得る。

30

【0171】

図2Bは、5'ユニバーサルオリゴヌクレオチド及び3'ユニバーサルオリゴヌクレオチドを含む核酸標準物質のプールを含む標準物質の代表的なプールを示す。図2Bに示されるように、標準物質のプールは、図4に示されるように調製された上流挿入アンプリコン及び下流挿入アンプリコンを含み得る。

【0172】

D. 部分的重複オリゴヌクレオチドを含む標準物質

いくつかの実施形態では、異なる長さの核酸標準物質のプールは、UMI、並びに、5'部分的重複オリゴヌクレオチドであって、5'部分的重複オリゴヌクレオチドは、全ての標準物質についてその配列の少なくとも一部にわたって同一である、5'部分的重複オリゴヌクレオチド、及び/又は、3'部分的重複オリゴヌクレオチドであって、3'部分的重複オリゴヌクレオチドは、全ての標準物質についてその配列の少なくとも一部にわたって同一である、3'部分的重複オリゴヌクレオチド、を含む、核酸標準物質を含み、5'部分的重複オリゴヌクレオチド及び/又は3'部分的重複オリゴヌクレオチドの長さは、標準物質の長さを決定する。

40

【0173】

本明細書中で使用される場合、「部分的重複オリゴヌクレオチド」とは、全ての標準物

50

質についてその配列の少なくとも一部にわたって同一であるオリゴヌクレオチドを指す。いくつかの実施形態では、標準物質は、5'部分的重複オリゴヌクレオチド及び3'部分的重複オリゴヌクレオチドの両方を含む。

【0174】

本明細書で使用される場合、「5'部分的重複オリゴヌクレオチド」は、図8Cの104によって表されるような、標準物質に含まれるUMIの5'側にあるオリゴヌクレオチドである。本明細書で使用される場合、「3'部分的重複オリゴヌクレオチド」は、図8Cの105によって表されるような、標準物質に含まれるUMIの3'側にあるオリゴヌクレオチドである。いくつかの実施形態では、5'部分的重複オリゴヌクレオチド及び3'部分的重複オリゴヌクレオチドは異なる。いくつかの実施形態では、5'部分的重複オリゴヌクレオチド及び3'部分的重複オリゴヌクレオチドは、異なる数のヌクレオチドを含む。

【0175】

いくつかの実施形態では、5'部分的重複オリゴヌクレオチドは、目的の配列の少なくとも第1の部分を含み、3'部分的重複オリゴヌクレオチドは、目的の配列の少なくとも第2の部分を含む。言い換えれば、5'部分的重複オリゴヌクレオチドは、目的の配列の少なくとも第1の部分を含み、3'部分的重複オリゴヌクレオチドは、目的の配列の異なる部分に対応し得る。

【0176】

いくつかの実施形態では、標準物質は、5'部分的重複オリゴヌクレオチドのみを含む(3'部分的重複オリゴヌクレオチドは含まない)。いくつかの実施形態では、標準物質は、3'部分的重複オリゴヌクレオチドのみを含む(5'部分的重複オリゴヌクレオチドは含まない)。5'部分的重複オリゴヌクレオチド又は3'部分的重複オリゴヌクレオチドのみを含む標準物質は、目的の配列中の大きな領域の損失を生じる欠失変異について制御するために有用であり得る。

【0177】

いくつかの実施形態では、5'部分的重複オリゴヌクレオチド及び/又は3'部分的重複オリゴヌクレオチドはそれぞれ、図5に示されるように、目的の配列から増幅されたアンプリコンを含む。

【0178】

1. 欠失アンプリコン

5'部分的重複オリゴヌクレオチドは、目的の配列からの増幅によって生成される場合、5'欠失アンプリコン又は上流欠失アンプリコンと称され得る。3'部分的重複オリゴヌクレオチドは、目的の配列からの増幅によって生成される場合、3'欠失アンプリコン又は下流欠失アンプリコンと称され得る。例えば、図5に示されるように、上流欠失アンプリコンの各々は、目的の配列の一部(黒色で示される)を含み、下流欠失アンプリコンの各々もまた、目的の配列の一部(黒色で示される)を含む。いくつかの実施形態では、上流欠失アンプリコン及び下流欠失アンプリコンに含まれる目的の配列の部分は異なってもよい。図5は、制限酵素切断部位(例えば、RS3及びRS4)を含み、かつLA-fwd及びLA-revプライマー結合配列及び目的の配列に含まれる他の配列に結合するプライマーを使用して、代表的な上流欠失アンプリコン及び下流欠失アンプリコンが生成され得る方法を示す。

【0179】

本明細書で使用される場合、「欠失アンプリコン」は、5'欠失アンプリコン又は3'欠失アンプリコンのいずれかであるアンプリコンを指し得る。一般に、「欠失アンプリコン」は、標準物質におけるその配置によって限定されない。

【0180】

いくつかの実施形態では、欠失アンプリコンを調製するために使用されるリバースプライマー及びフォワードプライマーは、欠失アンプリコンのサイズを決定する。いくつかの実施形態では、単一のプライマー対は、所望のサイズの欠失アンプリコンを生成する。

【0181】

いくつかの実施形態では、標準物質は、上流欠失アンプリコン及び下流欠失アンプリコンの両方を含む（図5に示される）。これらは、「欠失アンプリコン対」と称され得る。しかし、標準物質はまた、上流欠失アンプリコン又は下流欠失アンプリコンのいずれかのみを含み得る。

【0182】

いくつかの実施形態では、5'部分的重複オリゴヌクレオチド及び/又は3'部分的重複オリゴヌクレオチドはそれぞれ、目的の配列よりも20bp~1kb小さい配列を含む。言い換えれば、5'部分的重複オリゴヌクレオチド及び/又は3'部分的重複オリゴヌクレオチドは、目的の配列の欠失変異において見出される配列に対応し得る。

10

【0183】

図6Bは、図5に示されるように調製された、上流欠失アンプリコン及び下流欠失アンプリコンを含む核酸標準物質のプールを含む標準物質の代表的なプールを示す。

【0184】

II. 標準物質を作製する方法

本発明の標準物質及び使用方法は、標準物質を生成する手段によって限定されない。いくつかの実施形態では、標準物質は、標準物質を調製するためにオリゴヌクレオチドと一緒にライゲーションすることによって生成される。

【0185】

核酸標準物質のプールを生成する方法が本明細書に記載され、この方法は、核酸を含む少なくとも1つの目的の配列の複数のコピーを提供することと、それぞれUMIを含む一連のオリゴヌクレオチドを提供することと、様々な長さの一連の挿入オリゴヌクレオチドを提供することと、少なくとも1つの目的の配列、UMIを含む少なくとも1つのオリゴヌクレオチド、及び少なくとも1つの挿入アンプリコンをライゲーションして、核酸標準物質のプールの複数の核酸標準物質を生成することと、を含む。

20

【0186】

いくつかの実施形態では、少なくとも1つの目的の配列及び/又は挿入オリゴヌクレオチドは、増幅によって調製される。

【0187】

いくつかの実施形態では、目的の配列、それぞれUMIを含むオリゴヌクレオチド、及び/又は挿入オリゴヌクレオチドは、制限酵素切断部位を含む。いくつかの実施形態では、制限酵素切断部位は、目的の配列、それぞれUMIを含むオリゴヌクレオチド、及び/又は挿入オリゴヌクレオチドの5'末端及び/又は3'末端の近位にある。

30

【0188】

いくつかの実施形態では、方法は、目的の配列、それぞれUMIを含むオリゴヌクレオチド、及び/又は挿入オリゴヌクレオチドを制限酵素で切断することをライゲーションの前に更に含む。いくつかの実施形態では、制限酵素により切断することは、ライゲーションのための付着末端を生成する。いくつかの実施形態では、UMIを含むオリゴヌクレオチドは、目的の配列にも含まれる所望の制限酵素切断部位を含むように設計される。

【0189】

核酸標準物質のプールを生成する方法が本明細書に更に記載され、この方法は、核酸を含む少なくとも1つの目的の配列の複数のコピーを提供することと、それぞれUMIを含む一連のオリゴヌクレオチドを提供することと、少なくとも1つの目的の配列と、UMIを含む少なくとも1つのオリゴヌクレオチドとをライゲーションすることと、を含む。

40

【0190】

いくつかの実施形態では、少なくとも1つの目的の配列は、増幅によって調製される。いくつかの実施形態では、目的の配列及び/又はそれぞれUMIを含むオリゴヌクレオチドは、制限酵素切断部位を含む。いくつかの実施形態では、制限酵素切断部位は、目的の配列及び/又はそれぞれUMIを含むオリゴヌクレオチドの5'末端及び/又は3'末端の近位にある。

50

【0191】

いくつかの実施形態では、方法は、目的の配列及び/又はそれぞれUMIを含むオリゴヌクレオチドを制限酵素で切断することをライゲーションの前に更に含む。

【0192】

いくつかの実施形態では、制限酵素により切断することは、ライゲーションのための付着末端を生成する。

【0193】

いくつかの実施形態では、使用されているLong Amp標準物質の数と比較して、より多数のUMIが利用可能である。このようにして、UMIの数は、作成される標準物質の数より多くなり、UMIの重複が最小限に抑えられる。

10

【0194】

III. アンプリコンサイズバイアスを正規化する方法

本明細書に記載される標準物質のプールは、アンプリコンサイズバイアスを正規化するための方法において使用され得る。

【0195】

アンプリコンサイズバイアスを正規化する方法が本明細書に記載され、この方法は、標的核酸を含む試料を、異なる長さの核酸標準物質のプールと組み合わせることによって、各標準物質はUMIを含む、ことと、標準物質及び標的核酸に含まれる目的の配列のアンプリコンを増幅することと、標準物質及び目的の配列のアンプリコンを配列決定して、配列決定データを生成することと、標準物質由来の配列決定データを使用して、アンプリコンサイズに基づいてバイアスプロファイルを決定することと、バイアスプロファイルを使用して、アンプリコンサイズバイアスを正規化することと、を含む。

20

【0196】

本明細書で使用される場合、「アンプリコンサイズバイアス」は、異なるサイズのアンプリコンが異なって増幅されるという事実を指す。いくつかの実施形態では、所与の増幅反応において、より短いアンプリコンと比較して、より少ない大きなアンプリコンが生成される。いくつかの実施形態では、増幅は、PCR増幅を含む。いくつかの実施形態では、増幅は、Long Amp PCRである。

【0197】

Long Amp PCRは、通常のPCR法又は試薬を使用して典型的に増幅することができないDNA長の増幅を含む。Long Amp PCRに最適化された酵素は、長距離ポリメラーゼと称され得る。Long Amp PCRの結果は、完全なアンプリコンが生成される場合に改善されるため、サイクルにおける不完全なアンプリコンの生成は、後のPCRサイクルにおける不完全なアンプリコンの更なる生成につながる。いくつかの実施形態では、長距離ポリメラーゼは、高い処理能力(すなわち、DNAポリメラーゼによる単一の結合事象の間に比較的多数のヌクレオチドを組み込む)及び/又は高い伸長速度を有する。

30

【0198】

高い処理能力及び高い伸長速度を有する長距離ポリメラーゼは、長い鋳型の効率的なDNA合成を確実にし、サイクル時間を短縮するのに役立つ。Long Amp PCRで使用するための多種多様なプロトコル及び長距離ポリメラーゼ、例えばLong Amp Taq DNAポリメラーゼ及びPhusion DNAポリメラーゼ(New England Biolabs)が知られている。いくつかの実施形態では、長距離ポリメラーゼは、PrimeSTAR GXL DNAポリメラーゼ(Takara)である。

40

【0199】

いくつかの実施形態では、Long Amp PCRにおけるアンプリコンサイズバイアスは、本明細書に記載の核酸標準物質を使用する方法で正規化することができる。いくつかの実施形態では、標準物質は、バイアスプロファイルを生成するために使用され、このバイアスプロファイルは、目的の配列から生成されたアンプリコンについてのデータを正規化するために使用され得る。いくつかの実施形態では、目的の配列からのアンプリコン

50

の増幅へのアンプリコンサイズの影響は、本明細書に記載される標準物質を用いて生成されたデータを使用して正規化され得る。

【0200】

いくつかの実施形態では、目的の配列のアンプリコンを増幅することは、目的の配列の末端でプライマー結合配列に結合する一対のPCRプライマーを用いて標的核酸からアンプリコンを増幅することを含む。いくつかの実施形態では、標準物質は、目的の配列の末端にあるものと同じプライマー結合配列を含む。

【0201】

いくつかの実施形態では、方法は、増幅後かつ配列決定前に、断片のライブラリーを生成することを更に含む。

【0202】

いくつかの実施形態では、断片のライブラリーを生成することは、タグメンテーションによる。このような方法を図1に示し、ここで、断片は、Nextera断片化プロトコルによって生成される。そのような方法は、例えば、異なる挿入突然変異(図1において矢印で示される)を含む断片を生成する。この「long amp」PCR及び断片化工程では、PCR中のアンプリコンサイズバイアスを正規化するために、本明細書に記載される標準物質のプールを添加することができる。このようにして、標準物質のプールは、目的の配列と同じ増幅及び断片化条件に供される。

【0203】

いくつかの実施形態では、バイアスプロファイルを決定するために使用される標準物質由来の配列決定データは、標準物質に含まれるUMIの固有分子数である。言い換えれば、当業者は、配列決定データの標準的な分析を使用して、異なる標準物質から複製されたUMIの数を決定することができる。これらのUMIは異なる長さの標準物質に由来するため、異なるUMIの数は、バイアスプロファイルを生成するための異なるサイズのアンプリコンの増幅効率の尺度を提供することができる。このようにして、目的の配列とは異なる配列について生成されたアンプリコン(目的の野生型配列及びインデルを含む目的の配列からも生成されたアンプリコンを含む)の数を、バイアスプロファイルと比較することができる。言い換えれば、標準物質と比較する目的の配列から生成されたデータの比較を使用して、アンプリコンサイズバイアスについて配列決定データを正規化することができる。例えば、目的の配列の大きな挿入変異と同様のサイズの挿入標準物質が、目的の野生型配列と同様のサイズの標準物質よりも3倍低い速度で増幅された場合、ユーザーは、これらの大きな挿入変異のコピー数を、野生型配列と比較して正規化することができる。同様に、当業者は、欠失標準物質を使用して野生型配列と比較して、多数の大きな欠失変異(すなわち、大量の配列が失われている場合)について正規化することができる。

【0204】

A. 長い増幅PCR及び配列決定

長い増幅PCR(Long Amp)は、長いアンプリコンに最適化されたPCR反応を指す。そのようなLong Amp反応を図1に示す(「long amp」PCR)。そのような最適化されたLong Amp PCRの方法は、当技術分野において周知である。

【0205】

いくつかの実施形態では、長いアンプリコンは、5,000キロベース超、10,000キロベース超、又は20,000キロベース超であり得る。

【0206】

いくつかの実施形態では、長いアンプリコンは、大きな挿入変異を含み得る目的の配列から生成される。例えば、長いアンプリコンは、約10,000キロベースであり得るが、この目的の配列由来の野生型アンプリコンは、約1,000キロベースである。

【0207】

いくつかの実施形態では、Long Ampは、目的の配列における長い挿入突然変異の同定を最適化するために使用される。

【0208】

10

20

30

40

50

Long Amp PCRの後、ライブラリー断片の配列決定の前にライブラリー調製を行ってもよい。例えば、タグメンテーションを使用して（例えば、IlluminaのNexteraシステムを用いて）、配列決定のためのライブラリーを調製できる。

【0209】

いくつかの実施形態では、標準物質は、対照アッセイを実行するために使用される。いくつかの実施形態では、これらの対照アッセイは、Long Amp PCR反応とは別個である。いくつかの実施形態では、標準物質は、各Long Amp PCR反応への既知の量のスパイクインである。「スパイクイン」とは、標準物質がLong Amp PCR反応と同じ反応溶液中で増幅されることを意味する。

【0210】

IV. ライブラリー中のDNA損傷を決定する方法

ライブラリーを品質管理（QC）するための定量的PCR（qPCR）方法が本明細書に記載される。このような方法は、ユーザーが、配列決定などのライブラリーの更なる分析を実施する前に、ライブラリー中に存在するDNA損傷の量を決定することを可能にし得る。いくつかの実施形態では、QCアッセイは、異なるレベルの損傷を有するライブラリーを区別する。

【0211】

いくつかの実施形態では、これらのライブラリーは、配列決定のために使用され得る。いくつかの実施形態では、ライブラリーは、長リード配列決定のために意図される。いくつかの実施形態では、ライブラリーは、タグメンテーション及び/又はビーズ連結トランスポソームを使用して調製される。ライブラリー中のDNA損傷を決定する本発明の方法は、任意の方法によって生成されたライブラリーで使用することができる。

【0212】

本明細書中で使用される場合、「ライブラリー分子」は、ライブラリー内に含まれる単一分子をいう。いくつかの実施形態では、各ライブラリー分子は、標的核酸とは異なる挿入部を含み得る。ライブラリー分子は、当該分野で周知である標準的なタグメンテーション又はライゲーションプロトコルを用いて生成され得る。

【0213】

多くの配列決定用途は、ライブラリー分子中に1つ以上のアダプターの存在を必要とする。多くの場合、これらのアダプター配列は挿入部の両端にある。いくつかの実施形態では、アダプターに含まれる配列は、例えば、フローセルへのライブラリー分子の結合又はライブラリー分子への配列決定プライマーの結合を可能にするために、配列決定用途において使用される。いくつかの実施形態では、アダプター配列は、2つの異なる配列決定プライマー配列への結合のためなど、配列決定用途のために挿入部の両端に必要とされる。このようなシナリオにおいて、1つのアダプター配列を欠くライブラリー分子（例えば、ニックの入ったライブラリー又はそのアンプリコン）は、配列決定を成功できない。

【0214】

いくつかの実施形態では、ライブラリーは、長リードヘアピンアダプターを含むライブラリー分子を含む。長リードライブラリー分子中の挿入部サイズは、5 kb以上、10 kb以上、15 kb以上、20 kb以上、25 kb以上、又は30 kb以上であり得る。いくつかの実施形態では、ヘアピンアダプターを、ライブラリー分子内の挿入部に含まれるDNAの長い領域に付加することができる。いくつかの実施形態では、ヘアピンアダプターは、ライゲーション又はタグメンテーションプロトコルを使用して挿入部に付加され得る。例えば、NEBのNEBNext Multiplex Oligos for Illumina（登録商標）は、アダプター二量体形成を最小限にする独特のヘアピンループ構造を有するアダプターライゲーションを使用する。

【0215】

いくつかの実施形態では、ヘアピンアダプターは、タグメンテーション反応の間に挿入部に付加され得る。「タグメンテーション」は、断片及びタグ核酸に対するトランスポザーゼの使用を指す。タグメンテーションは、トランスポゾン末端配列（本明細書でトラン

10

20

30

40

50

スポゾンと称される)を含む1つ以上のタグ(アダプター配列など)で複合体化されたトランスポザーゼ酵素を含むトランスポソーム複合体によるDNAの修飾を含む。タグメンテーションは、DNAの断片化及び二重鎖断片の両方の鎖の5'末端へのアダプターのライゲーションを同時にもたらし得る。しかし、タグメンテーションは、ライブラリーを生成する1つの方法にすぎず、他の方法(ライゲーションなど)もまた、本発明のQCアッセイで使用するためのライブラリーを生成するために使用することができる。

【0216】

いくつかの実施形態では、1つ以上のライブラリー分子を含むライブラリーにおけるDNA損傷の存在を決定する方法であって、各ライブラリー分子は、挿入部の各末端にヘアピンアダプターを有する二本鎖DNA挿入部を含み、この方法は、ライブラリー分子に含まれる二本鎖DNA挿入部の第1の鎖及び第2の鎖を変性することと、フォワードプライマー及びリバースプライマーをライブラリー分子にアニーリングすることと、ライブラリーアンプリコンを生成するために増幅することと、生成されたライブラリーアンプリコンの数に基づいてDNA損傷の存在を評価することと、を含む。例示的な方法を図9に示し、これは、ニックを有するライブラリー分子が完全長アンプリコンを生成しないことを示す。

10

【0217】

本明細書に記載の方法は、長距離ポリメラーゼを使用して、QCのためのライブラリー分子を増幅することができる。いくつかの実施形態では、QCアッセイは、異なるレベルの損傷を有するライブラリーを区別し、ライブラリー調製物における損傷率と相関するCq値をもたらし得る。本明細書に記載の方法は、1つ以上のヘアピンアダプターを含む任意のライブラリーに適用することができる。特に、長リード配列決定のための長い挿入部のライブラリー調製に使用することができる。いくつかの実施形態では、本発明のQCアッセイの使用は、損傷したライブラリーの使用を回避し、時間、費用、及び消耗品の節約をもたらし得る。

20

【0218】

A. ライブラリーにおけるDNA損傷

ライブラリー調製の全ての方法は、調製プロセスの間に核酸に損傷を導入し得る。例えば、任意のピペッティング工程は、核酸の剪断をもたらし得る。ユーザーは潜在的な損傷を低減するための措置を講じることができるが、この損傷を完全に回避又は予測することはできない。

30

【0219】

ライブラリー分子内の挿入部は、1つ以上のより大きな核酸由来の断片として得られる二本鎖核酸を含み得る。断片化は、例えば、噴霧化、超音波処理、化学的開裂、酵素的開裂、又は物理的剪断を含む、当技術分野で公知の様々な技術のいずれかを用いて行うことができる。しかし、これらの断片化方法のいずれも、DNA損傷(例えば、DNAのニックング)を導入する可能性を有する。

【0220】

したがって、ライブラリー中のDNA損傷を評価できることが重要である。例えば、配列決定の質が低いため、ユーザーは、広範なDNA損傷を有するライブラリーに対して更なる配列決定を行うことを望まない。同様に、ライブラリーの多くが損傷されている場合、ユーザーは、配列決定するためのライブラリー産物の適切な量を決定することが困難であり得る。多くの配列決定プラットフォームについて、ライブラリー分子は、フローセルへの結合又は配列決定プライマーへの結合などの使用のために、断片の両端にアダプター配列を必要とする。適切なアダプターが存在しない場合、例えば、ライブラリー分子がDNA損傷を有する場合、ライブラリー分子(及びそのアンプリコン)は、分析可能な配列決定データを生成しない。

40

【0221】

DNA損傷の評価は、損傷したライブラリーのユーザーによる更なる使用の回避を可能にし得る。このようにして、ユーザーは、低いライブラリー品質が高品質のデータの生成

50

を妨げる場合、配列決定などの用途のための時間及び試薬コストを節約することができる。いくつかの実施形態では、低品質のライブラリーは、配列決定から除外される。

【0222】

いくつかの実施形態では、DNA損傷は、1つ以上のニックである。いくつかの実施形態では、QCアッセイを実施する前に、1つ以上のニックを二本鎖切断に変換することができる。

【0223】

1. ニック

いくつかの実施形態では、DNA損傷は、ライブラリー分子中の1つ以上のニックを含む。本明細書中で使用される場合、1つ以上のニックは、単一のニック又は複数の別個のニックであってよい。

10

【0224】

いくつかの実施形態では、1つ以上のニックは、ライブラリー分子に含まれる挿入部内にある。挿入部は二本鎖挿入部であり得るので、ニックは挿入部の一方の鎖における切断を指し、切断はその位置において他方の鎖に存在しない。したがって、本明細書で使用される場合、ニックは、二本鎖DNA挿入部における不連続性を指すことができ、そこでは一方の鎖の隣接するヌクレオチド間にホスホジエステル結合が存在しない。いくつかの実施形態では、1つ以上のニックは、ライブラリー調製中のDNA損傷によって生成された。例えば、ピペティング中の剪断は、ライブラリー分子中にニックをもたらし得る。

【0225】

いくつかの実施形態では、以下で述べられるように、ライブラリー中に1つ以上のニックを含むライブラリー分子の割合がより大きい場合、QCアッセイで生成されるCq値がより大きい。

20

【0226】

いくつかの実施形態では、DNA損傷は、ライブラリー分子中の2つ以上のニックを含み、ニックは、二本鎖DNA挿入部の同じ鎖中にある。

【0227】

いくつかの実施形態では、DNA損傷は、ライブラリー分子中の2つ以上のニックを含み、ニックは、二本鎖DNA挿入部の両方の鎖中にある。2つ以上のニックが異なる鎖にある場合、これらのニックは、以下に記載される二本鎖DNA切断と区別するために、異なる位置にあってもよい。

30

【0228】

増幅中にニックに遭遇した場合、DNAポリメラーゼは、ニックを越えてアンプリコンを伸長することができない可能性がある。したがって、1つ以上のニックは、ライブラリー分子の完全な配列を有さない不完全なアンプリコンの生成をもたらし得る。いくつかの実施形態では、ライブラリー分子が1つ以上のニックを含む場合、フォワードプライマー及び/又はリバースプライマーは、ライブラリー分子の完全配列に対応するアンプリコンを生成することができない。ライブラリーの完全配列を含まないこのようなアンプリコンは、配列決定不可能であり得る（挿入部の一方又は両方の末端にあるべきアダプター配列の欠如に起因する）。

40

【0229】

いくつかの実施形態では、ニックを含むライブラリー分子から生成されたアンプリコンは、フォワードプライマー及び/又はリバースプライマーに結合するための配列を欠いている。

【0230】

いくつかの実施形態では、ニックを含むライブラリー分子は、ニックを含まないライブラリー分子と比較して、増幅中により少ないアンプリコンを生成する。以下に記載されるように、本発明のQC方法は、ニックを含むライブラリー分子のCq値を推定ことができ、したがって、ライブラリーが比較的low品質（高いCq値を有する）又は比較的高品質（低いCq値を有する）であることをユーザーに示すことができる。このようにして、

50

Cq値は、配列決定などによってライブラリーを更に評価するかどうかを評価するために所与のライブラリーの質を推定するため、及び不十分なデータを生じるライブラリーの配列決定に関連する時間及び費用を回避するために使用され得る。

【0231】

2. ニックから生成される二本鎖DNA切断

いくつかの実施形態では、方法は、ニックから二本鎖切断を生成することを更に含む。いくつかの実施形態では、二本鎖切断は、QC法においてフォワードプライマー及びリバースプライマーをアニーリングする前にニックから生成される。

【0232】

いくつかの実施形態では、酵素を用いて、ニックから二本鎖切断を調製する。言い換えれば、二本鎖切断を生成することは、酵素反応を使用して行われ得る。いくつかの実施形態では、酵素反応はエンドヌクレアーゼによって行われる。いくつかの実施形態では、エンドヌクレアーゼは、T7エンドヌクレアーゼである。

10

【0233】

いくつかの実施形態では、二本鎖切断を含むライブラリー分子は、増幅中にライブラリー分子の完全配列に対応するアンプリコンを生成しない。いくつかの実施形態では、二本鎖切断は、挿入部内のライブラリー分子を切断し、ライブラリー分子の完全長アンプリコンは、切断後に生成され得ない。

【0234】

いくつかの実施形態では、二本鎖切断を含むライブラリー分子から生成されたアンプリコンが、フォワードプライマー及び/又はリバースプライマーに結合するための配列を欠いている。いくつかの実施形態では、二本鎖切断は、挿入部内のライブラリー分子を切断し、2つの異なるヘアピンアダプター（ライブラリー挿入部の2つの末端に関連する）に含まれるプライマー結合配列が分離される。いくつかの実施形態では、切断後、フォワードプライマーもリバースプライマーも、ライブラリー分子への結合後に完全長アンプリコンを生成することができない。

20

【0235】

B. ヘアピンアダプター

本明細書で使用される場合、「ヘアピン」とは、互いに少なくとも部分的に相補的である一対の核酸配列を含む核酸を指す。少なくとも部分的に相補的であるこれらの2つの核酸配列は、互いに結合し、核酸の折り畳みを媒介することができる。いくつかの実施形態では、少なくとも部分的に相補的である2つの核酸配列は、ヘアピン二次構造を有する核酸を生成する。

30

【0236】

本明細書で使用される「ヘアピンアダプター」は、互いに少なくとも部分的に相補的である少なくとも1対の核酸配列を含むアダプターを指す。いくつかの実施形態では、ヘアピンアダプターは、折り畳まれた二次構造を有する。

【0237】

いくつかの実施形態では、ヘアピンアダプターは、1つ以上のアダプター配列を含む。いくつかの実施形態では、アダプター配列は、プライマー配列、インデックスタグ配列、捕捉配列、バーコード配列、切断配列、若しくは配列決定関連配列、又はそれらの組み合わせを含む。本明細書で使用される場合、配列決定関連配列は、後の配列決定工程に関連する任意の配列であり得る。配列決定関連配列は、下流の配列決定工程を単純化するように作用し得る。例えば、配列決定関連配列は、他の点でアダプターを核酸断片にライゲーションする工程を介して組み込まれる配列であり得る。いくつかの実施形態では、アダプター配列は、特定の配列決定方法においてフローセルへの結合を容易にするために、P5又はP7配列（又はそれらの相補体）を含む。

40

【0238】

いくつかの実施形態では、ヘアピンアダプターは、増幅プライマー配列（すなわち、増幅プライマー配列に結合する配列）を含む。いくつかの実施形態では、ヘアピンアダプタ

50

ーは、増幅プライマー配列と、アダプター配列に少なくとも部分的に相補的な配列の全部又は一部と、を含む。いくつかの実施形態では、ヘアピンに含まれた増幅プライマー配列は、ユニバーサルプライマー配列である。ユニバーサル配列は、2つ以上の核酸断片に共通の(すなわち、共有される)ヌクレオチド配列の領域である。

【0239】

いくつかの実施形態では、フォワードプライマー又はリバースプライマーのいずれかは、一方又は両方のヘアピンアダプターに含まれる1つ以上の配列に結合する。いくつかの実施形態では、フォワードプライマー及びリバースプライマーの両方は、一方又は両方のヘアピンアダプターに含まれる1つ以上の配列に結合する。いくつかの実施形態では、フォワードプライマーは、二本鎖DNA挿入部の第1の末端に結合したヘアピンアダプター

10

【0240】

いくつかの実施形態では、ライブラリー分子は、二本鎖核酸を含む挿入部及び挿入部の両末端にヘアピンアダプターを含む。いくつかの実施形態では、挿入部は、標的核酸由来の断片を含む。ヘアピンアダプターを組み込む方法は、ライゲーション又はタグメンテーションによるものなど、当技術分野で周知である。

【0241】

例えば、NEBNext(登録商標)Multiplex Oligos for Illumina(登録商標)(New England BioLabs)は、ライブラリー産物の収量を増加させるためのヘアピンアダプター及びプライマーを提供する。いくつかの実施形態では、ヘアピンアダプターは、アダプター二量体形成を最小化するヘアピングループ構造を含む。いくつかの実施形態では、ヘアピンアダプターは、末端修復されたdAテール付加DNAにライゲーションされる。いくつかの実施形態では、ヘアピンアダプターは、USER試薬による処理によって除去されるウラシルを含有するループを含む。いくつかの実施形態では、USER酵素は、ウラシルDNAグリコシラーゼ(UDG)及びDNAグリコシラーゼ-リアーゼ(エンドヌクレアーゼVIIなど)の混合物である。いくつかの実施形態では、USER処理は、ヘアピンアダプターのループを開き、インデックスプライマーを組み込むための増幅及びその後の配列決定のための基質として利用可能にすることができる。

20

30

【0242】

いくつかの実施形態では、ヘアピンアダプターは、遺伝子座特異的プライマー及びUSER試薬を使用して組み込まれ、ヘアピンアダプターをライゲーションするためのオーバーハングを生成する。例示的な方法は、SMRTbellライブラリー調製である(Pacific Biosciences, SMRTbell Library Preparation & SMRT Sequencing Workflow Updates, 2017参照)。

【0243】

いくつかの実施形態では、ヘアピンアダプターは、比較的大きな挿入部を有するライブラリー分子に含まれ、ライブラリー分子は、長リード配列決定のために設計される。

40

【0244】

いくつかの実施形態では、各ヘアピンアダプターは、増幅プライマー結合部位を含む。いくつかの実施形態では、挿入部の第1の末端におけるヘアピンアダプターは、挿入部の第2の末端におけるヘアピンアダプターとは異なる増幅プライマー結合部位を含む。いくつかの実施形態では、挿入部の第1の末端におけるヘアピンアダプターは、第1の増幅プライマー結合部位を含み、挿入部の第2の末端におけるヘアピンアダプターは、第2の増幅プライマー結合部位を含む。いくつかの実施形態では、第1の増幅プライマー結合部位及び第2の増幅プライマー結合部位は、反対方向の増幅を媒介する。

【0245】

図9に示されるようないくつかの実施形態では、挿入部の第1の末端のヘアピンアダプ

50

ターは、フォワード増幅プライマー結合部位を含んでもよく、挿入部の第2の末端のヘアピンアダプターは、リバース増幅プライマー結合部位を含んでもよい。

【0246】

C. 増幅

いくつかの実施形態では、方法は、増幅プライマー配列に結合する増幅プライマーを使用してライブラリー分子を増幅することを更に含む。いくつかの実施形態では、ライブラリー分子に含まれる一方又は両方のヘアピンアダプターは、増幅プライマーを含む。

【0247】

いくつかの実施形態では、増幅することは、5 kb以上、10 kb以上、15 kb以上、20 kb以上、25 kb以上、又は30 kb以上であるライブラリー分子を増幅するために最適化されている。

10

【0248】

いくつかの実施形態では、増幅することは、長いアンプリコンの増幅に最適化されているポリメラーゼを用いて行われる。いくつかの実施形態では、ポリメラーゼは、20 kb以上又は30 kb以上のアンプリコンの増幅に最適化されている。

【0249】

長いアンプリコンの増幅に最適化されている多くの例示的なポリメラーゼが、当該分野で既知である。1つの例示的なポリメラーゼは、PrimeSTAR GXL DNAポリメラーゼ(Takara)である。

【0250】

いくつかの実施形態では、ポリメラーゼは、野生型Taqポリメラーゼと比較してより高い処理能力及び/又は伸長速度を有する。いくつかの実施形態では、ポリメラーゼは、処理能力又は伸長速度を増加させる1つ以上の突然変異又は融合を含む。

20

【0251】

本明細書中で使用される場合、ポリメラーゼの「処理能力」とは、ポリメラーゼが、DNA鋳型から解離する前に、単一の鋳型結合事象の間にDNAに組み込まれ得るヌクレオチドの数を指す。したがって、比較的高い処理能力を有するポリメラーゼは、単一の鋳型結合事象の間に多数のヌクレオチドを組み込むことができる。より高い処理能力は、完全なアンプリコンがPCRサイクルの間に生成される可能性を増加させ得る。

【0252】

本明細書中で使用される場合、ポリメラーゼの「伸長速度」は、ある期間にわたってDNAに組み込まれ得るヌクレオチドの数である。いくつかの実施形態、比較的高い伸長速度を有するポリメラーゼは、PCRサイクル中にライブラリー分子の完全なアンプリコンを生成することができる。いくつかの実施形態では、ポリメラーゼは、2 kb/分以上、3 kb/分以上、又は4 kb/分以上の伸長速度を有する。

30

【0253】

いくつかの実施形態では、ポリメラーゼは、3 kb/分以上の伸長速度を有する。

【0254】

いくつかの実施形態では、増幅は指数関数的である。

【0255】

いくつかの実施形態では、30サイクル以上又は40サイクル以上の増幅が行われる。

40

【0256】

いくつかの実施形態では、増幅プライマーは、インデックス配列を含み得る。これらのインデックス配列は、試料及びアレイ中の位置を同定するために使用され得る。いくつかの実施形態では、インデックス配列は、固有分子識別子(unique molecular identifier、UMI)を含む。UMIは、国際公開第2016/176091号、同第2018/197950号、同第2018/197945号、同第2018/200380号、及び同第2018/204423号に記載されている(これらの各々は、参照によりその全体が本明細書に組み込まれる)。

【0257】

50

いくつかの実施形態では、試料は、固体支持体上で増幅される。

【0258】

例えば、いくつかの実施形態では、試料は、米国特許第7,985,565号及び同第7,115,400号の開示によって例示されるように、クラスター増幅方法論を使用して増幅される（これらの各々の内容は、参照によりその全体が本明細書に組み込まれる）。米国特許第7,985,565号及び同第7,115,400号の組み込まれている材料は、固定化された核酸分子のクラスター又は「コロニー」からなるアレイを形成するために、増幅産物を固体支持体上に固定化することを可能にする固相核酸増幅の方法を記載する。そのようなアレイ上の各クラスター又はコロニーは、複数の同一の固定化されたポリヌクレオチド鎖及び複数の同一の固定化された相補的ポリヌクレオチド鎖から形成される。そのように形成されたアレイは、本明細書では一般に「クラスター化アレイ」と称される。米国特許第7,985,565号及び同第7,115,400号に記載されているような固相増幅反応の産物は、固定化されたポリヌクレオチド鎖と固定化された相補鎖の対をアニーリングすることによって形成されるいわゆる「ブリッジ (bridged)」構造体であり、両方の鎖は、いくつかの実施形態では、共有結合を介して、5'末端で固体支持体上に固定化されている。クラスター増幅方法論は、固定化された核酸鋳型を使用して固定化されたアンプリコンを産生する方法の例である。他の好適な方法論を使用して、本明細書で提供される方法に従って産生された固定化されたDNA断片から固定化されたアンプリコンを産生することもできる。例えば、1つ以上のクラスター又はコロニーは、増幅プライマーの各対の一方又は両方のプライマーが固定化されているかどうかにかかわらず、固相PCRによって形成することができる。

10

20

【0259】

他の実施形態では、試料は、溶液中で増幅される。例えば、いくつかの実施形態では、試料は、切断されるか、又は他の方法で固体支持体から遊離され、次いで、増幅プライマーは、溶液中で遊離分子にハイブリダイズする。他の実施形態では、増幅プライマーを、1つ以上の初期の増幅工程のために所望の試料にハイブリダイズさせ、続いて、溶液中で後続の増幅工程を行う。いくつかの実施形態では、固定化された核酸鋳型を使用して、液相アンプリコンを生成することができる。

【0260】

本明細書に記載されているか又は当該技術分野で一般に知られている増幅方法論のいずれかを、ユニバーサルプライマー又は標的特異的プライマーとともに利用して、所望の試料を増幅できることが理解されるであろう。増幅のための好適な方法には、参照によりその全体が本明細書に組み込まれる、米国特許第8,003,354号に記載されるように、ポリメラーゼ連鎖反応 (PCR)、鎖置換増幅 (SDA)、転写媒介増幅 (TMA) 及び核酸配列ベースの増幅 (NASBA) が含まれるが、これらに限定されない。上記の増幅方法論を使用して、目的とする1つ以上の核酸を増幅することができる。例えば、多重PCR、SDA、TMA、NASBAなどを含むPCRを利用して、固定化DNA断片を増幅することができる。いくつかの実施形態では、目的とする核酸に特異的に向けられるプライマーは、増幅反応に含まれる。

30

【0261】

ポリヌクレオチドの増幅に好適な他の方法としては、オリゴヌクレオチド伸長及びライゲーション、ローリングサークル増幅 (rolling circle amplification、RCA) (Lizardi et al., Nat. Genet. 19: 225-232 (1998))、参照により本明細書に組み込まれる)、及びオリゴヌクレオチドライゲーションアッセイ (oligonucleotide ligation assay、OLA) (一般に米国特許第7,582,420号、同第5,185,243号、同第5,679,524号、及び同第5,573,907号、欧州特許第0320308 (B1)号、同第0336731 (B1)号、同第0439182 (B1)号、国際公開第90/01069号、同第89/12696号、及び同第89/09835号、その全てが参照により組み込まれる) 技術、を挙げることができる。これらの増幅方法論は、固定化DNA断片を増幅するように設計され得るこ

40

50

とを理解されたい。例えば、いくつかの実施形態では、増幅方法は、目的とする核酸に特異的に向けられるプライマーを含むライゲーションプロブ増幅又はオリゴヌクレオチドライゲーションアッセイ (OLA) 反応を含み得る。いくつかの実施形態では、増幅方法は、目的の核酸に特異的に指向されるプライマーを含有するプライマー伸長ライゲーション反応を含んでもよい。目的の核酸を増幅するように特別に設計することができるプライマー伸長及びライゲーションプライマーの非限定的な例として、増幅は、米国特許第7,582,420号及び同第7,611,869号(これらの各々は、参照によりその全体が本明細書に組み込まれる)に例示されるようなGolden Gateアッセイ(Illumina, Inc., San Diego, CA)に使用されるプライマーを含み得る。
【0262】

本開示の方法で使用され得る例示的な等温増幅方法としては、例えばDean et al., Proc. Natl. Acad. Sci. USA 99:5261-66(2002)により例示される多置換増幅(Multiple Displacement Amplification, MDA)、又は例えば米国特許第6,214,587号(これらの各々は、参照によりその全体が本明細書に組み込まれる)により例示される等温鎖置換核酸増幅が挙げられるが、これらに限定されない。本開示で使用され得る他の非PCRベースの方法としては、例えば、Walker et al., Molecular Methods for Virus Detection, Academic Press, Inc., 1995、米国特許第5,455,166号、及び同第5,130,238号、並びにWalker et al., Nucl. Acids Res. 20:1691-96(1992)に記載されている鎖置換増幅(SDA)又は例えば、Lage et al., Genome Research 13:294-307(2003)に記載されている超分枝鎖置換増幅が挙げられ、これらの各々は、参照によりその全体が本明細書に組み込まれる。等温増幅法は、ゲノムDNAのランダムプライマー増幅のために、鎖置換Phi 29ポリメラーゼ又はBst DNAポリメラーゼの大型断片、5³ エキソ-とともに使用することができる。これらのポリメラーゼの使用は、それらの高い加工性及び鎖置換活性を利用する。高い加工性により、ポリメラーゼは、10~20kbの長さの断片を産生することが可能になる。上記のように、クレノウポリメラーゼなどの低いプロセスビティと鎖置換活性を有するポリメラーゼを使用して、等温条件下でより小さな断片を産生することができる。増幅反応、条件、及び成分の更なる説明は、米国特許第7,670,810号の開示に詳細に記載されている(参照によりその全体が本明細書に組み込まれる)。
【0263】

D. 配列決定

いくつかの実施形態では、方法は、ライブラリー産物及び増幅されたライブラリー産物(すなわち、アンプリコン)を配列決定することを更に含む。いくつかの実施形態では、QCアッセイ後のライブラリーの分析は配列決定である。

【0264】

いくつかの実施形態では、方法は、Cq値に基づいてライブラリーの分析のための条件を決定することを含む。いくつかの実施形態では、QCアッセイは、ライブラリーを配列決定する条件を決定するために使用される。いくつかの実施形態では、QCアッセイは、所与のライブラリーを配列決定すべきでないことを決定するために使用される。例えば、QCアッセイは、所与のライブラリー中に十分なライブラリー分子が存在せず、その結果、ライブラリーから生成された配列決定データが低品質であることを推定し得る。

【0265】

いくつかの実施形態では、方法は、インサートの全配列の配列決定を可能にする。

【0266】

1つの例示的な配列決定方法論は、合成による配列決定(sequencing-by-synthesis, SBS)である。SBSにおいて、核酸テンプレートに沿った核酸プライマーの伸長を監視して、テンプレート中のヌクレオチド配列を判定する。基礎となる化学プロセスは、重合(例えば、ポリメラーゼ酵素によって触媒される)であり得る。特定のポリマーベ

10

20

30

40

50

ーアの SBS の実施形態では、プライマーに付加されるヌクレオチドの順序及び種類の検出を使用して鋳型の配列を決定することができるように、蛍光標識ヌクレオチドを鋳型依存的にプライマーに付加する（それによってプライマーを伸長させる）。

【0267】

フローセルは、配列決定のための便利な固体支持体を提供する。例えば、第1の SBS サイクルを開始するために、1つ以上の標識されたヌクレオチド、DNAポリマーゼなどを、1つ以上の増幅された核酸分子を収容するフローセルに流入/通過させることができる。プライマー伸長によって標識ヌクレオチドを組み込む部位は、検出することができる。任意選択的に、ヌクレオチドは、ヌクレオチドがプライマーに付加されると、更なるプライマー伸長を終結する可逆的終結特性を更に含むことができる。例えば、脱ブロック作用因子が送達されてその部分を除去するまで、その後の伸長が起こらないように、可逆的ターミネータ部分を有するヌクレオチド類似体をプライマーに付加することができる。したがって、可逆的終端を使用する実施形態では、脱ブロック試薬をフローセルに送達することができる（検出発生の前又は後）。洗浄は、様々な送達工程の間を実施することができる。次に、サイクルをn回繰り返してプライマーをn個のヌクレオチドで伸長し、それによって長さnの配列を検出することができる。本開示の方法によって生成されるアンプリコンとともに使用するために容易に適合させることができる例示的な SBS 手順、流体系及び検出プラットフォームは、例えば、Bentley et al., Nature 456: 53-59 (2008)、国際公開第04/018497号、米国特許第7,057,026号、国際公開第91/06678号、同第07/123744号、米国特許第7,329,492号、同第7,211,414号、同第7,315,019号、同第7,405,281号、及び米国特許出願公開第2008/0108082号に記載されており、これらの各々は、参照により本明細書に組み込まれる。

10

20

【0268】

パイロシーケンシングなどの、サイクル反応を使用する他の配列決定手順を使用することができる。パイロシーケンシングは、特定のヌクレオチドが新生核酸鎖に取り込まれる際の、無機ピロリン酸 (PPi) の放出を検出する (Ronaghi et al., Analytical Biochemistry 242 (1), 84-9 (1996)、Ronaghi, Genome Res. 11 (1), 3-11 (2001)、Ronaghi et al. Science 281 (5375), 363 (1998)、米国特許第6,210,891号、同第6,258,568号、及び同第6,274,320号、これらの各々は参照により本明細書に組み込まれる)。パイロシーケンシングでは、放出された PPi は、ATP スルフィラーゼによってアデノシン三リン酸 (adenosine triphosphate, ATP) に即座に変換されることで検出することができ、生成された ATP のレベルは、ルシフェラーゼ産生光子を介して検出することができる。したがって、配列決定反応は、発光検出システムを介して監視することができる。蛍光ベースの検出システムに使用される励起放射線源は、パイロシーケンシング手順には必要ない。本開示に従って生成されたアンプリコンへのパイロシーケンシングの適用に適合させることができる有用な流体システム、検出器、及び手順は、例えば国際公開第2012058096号、米国特許出願公開第2005/0191698 (A1) 号、米国特許第7,595,883号、及び同第7,244,559号に記載されており、これらの各々は、参照により本明細書に組み込まれる。

30

40

【0269】

いくつかの実施形態は、DNAポリマーゼ活性のリアルタイムモニタリングを伴う方法を利用することができる。例えば、ヌクレオチドの取り込みは、フルオロフォア担持ポリマーゼと -リン酸標識ヌクレオチドとの間の蛍光共鳴エネルギー移動 (fluorescence resonance energy transfer, FRET) 相互作用を介して、又はゼロモード導波路 (zeromode waveguide, ZMW) を用いて検出することができる。FRET ベースの配列決定のための技術及び試薬は、例えば、Levene et al. Science 299, 682-686 (2003)、Lundquist et al. O

50

pt. Lett. 33, 1026-1028 (2008)、Korlach et al. Proc. Natl. Acad. Sci. USA 105, 1176-1181 (2008)に記載され、それらの開示は参照により本明細書に組み込まれる。

【0270】

いくつかのSBS実施形態は、伸長産物へのヌクレオチドの組み込み時に放出されるプロトンの検出を含む。例えば、放出されたプロトンの検出に基づく配列決定は、Ion Torrent (Guilford, CT, Life Technologiesの子会社)から市販されている電気検出器及び関連技術、又は、米国特許出願公開第2009/0026082(A1)号、同第2009/0127589(A1)号、同第2010/0137143(A1)号、若しくは同第2010/0282617(A1)号に記載されている配列決定方法及びシステムであり、これらの各々は、参照により本明細書に組み込まれる。速度論的排除(kinetic exclusion)を使用して核酸を増幅するための本明細書に記載の方法は、プロトンを検出するために使用される基質に容易に適用することができる。より具体的には、本明細書に記載の方法を使用して、プロトンを検出するために使用されるアンプリコンのクローン集団を生成することができる。

10

【0271】

別の有用なシーケンシング技術は、ナノポアシーケンシングである(例えば、Deamer et al. Trends Biotechnol. 18, 147-151 (2000); Deamer et al. Acc. Chem. Res. 35: 817-825 (2002)、Li et al. Nat. Mater. 2: 611-615 (2003)を参照されたく、これらの開示は、参照により本明細書に組み込まれる)。いくつかのナノポアの実施形態では、核酸又は核酸から除去された個々のヌクレオチド個々のヌクレオチドは、ナノポアを通過する。核酸又はヌクレオチドがナノポアを通過するとき、各ヌクレオチドの種類は、ポアの電気コンダクタンスの変動を測定することによって識別され得る。(米国特許第7,001,792号、Soni et al. Clin. Chem. 53, 1996-2001 (2007)、Healy, Nanomed. 2, 459-481 (2007)、Cockroft et al. J. Am. Chem. Soc. 130, 818-820 (2008)、これらの開示は、参照により本明細書に組み込まれる)。

20

【0272】

本開示による検出に適用することができるアレイベースの発現及び遺伝子型解析のための例示的な方法は、米国特許第7,582,420号、同第6,890,741号、同第6,913,884号、若しくは同第6,355,431号、又は米国特許公開第2005/0053980(A1)号、同第2009/0186349(A1)号、若しくは同第2005/0181440(A1)号に記載されており、これらの各々は参照により本明細書に組み込まれる。

30

【0273】

本明細書に記載の方法の利点は、複数の核酸を並行して迅速かつ効率的に検出することを提供することである。したがって、本開示は、上記で例示されるものなどの当該技術分野において既知の技術を使用して核酸を調製及び検出することができる統合システムを提供する。したがって、本開示の統合システムは、増幅試薬及び/又は配列決定試薬を1つ以上の固定化されたDNA断片に送達することができる流体構成要素を含むことができ、システムは、ポンプ、弁、リザーバ、流体ラインなどの構成要素を含む。フローセルは、核酸を検出するための統合システムで構成及び/又は使用することができる。例示的なフローセルは、例えば、米国特許出願公開第2010/0111768(A1)号及び同第2012/0270305(A1)号に記載されており、これらの各々は、参照により本明細書に組み込まれる。フローセルについて例示されるように、統合システムの流体構成要素の1つ以上を増幅方法及び検出方法に使用することができる。核酸配列決定の実施形態を一例として取ると、統合システムの流体構成要素の1つ以上を、本明細書に記載の増幅方法、及び上記に例示したような配列決定方法における配列決定試薬の送達に使用する

40

50

ことができる。代替的に、統合システムは、増幅方法を実施し、検出方法を実施するための別々の流体システムを含み得る。増幅された核酸を作成し、又核酸の配列を決定することができる統合シーケンシングシステムの例としては、MiSeq（商標）プラットフォーム（Illumina, Inc., San Diego, CA）、及び参照により本明細書に組み込まれる、米国特許出願公開第2012/0270305号（参照により本明細書に組み込まれる）に記載の装置が挙げられるが、これらに限定されない。

【0274】**E. Cq値**

いくつかの実施形態では、産生されたライブラリーアンプリコンの数は、定量的PCR（qPCR）によって推定される。いくつかの実施形態では、産生されたライブラリーアンプリコンの数は、定量のサイクル（Cq、定量サイクルとしても知られる）値を測定することによって推定される。

10

【0275】

本明細書で使用される場合、Cq値は、試料の反応曲線が閾値線と交差するPCRサイクル数である。したがって、Cq値は、所与の試料についてノイズを超えるシグナルを検出するのに必要なPCRのサイクル数を示す。

【0276】

これは、蛍光色素及びプローブを用いて決定され得、この方法は、蛍光を検出するために必要とされる増幅サイクルの数を測定する。この方法を使用すると、Cq値は、PCR産物の蛍光がバックグラウンドシグナルを超えて検出され得るサイクル数である。したがって、より高いCq値は、より少ない核酸が試料中に存在することを示す。

20

【0277】

Bustin et al., *Clinical Chemistry* 55(4): 611-622 (2009)に記載されているように、閾値サイクル（Ct）、交差点（Cp）、及び出発点（TOP）という用語は全て、Cq値と同じ測定値を指し、名称の違いは単に異なる機器に基づく。これらの用語（Ct、Cp、及びTOP）は全て、試料の反応曲線が閾値線と交差するPCRサイクル数を決定する方法を指し、したがって、これらの値は全て、Cq値の同義語である。

【0278】

いくつかの実施形態では、ライブラリーアンプリコンの数が多ほど、Cq値が低くなる。いくつかの実施形態では、より低いCq値を有するライブラリーは、より少ないDNA損傷を有する。いくつかの実施形態では、より少ないDNA損傷を有するライブラリーは、より良好な配列決定結果を生じる。

30

【0279】

いくつかの実施形態では、ニックを含むこれらのライブラリー産物は、ライブラリー分子の完全配列に対応するアンプリコンを生成しない。いくつかの実施形態では、増幅サイクル中の伸長（すなわち、アンプリコンの生成）は、ライブラリー分子中のニックの部位で停止する。

【0280】

例えば、図9は、増幅がフォワード及びリバース増幅プライマー結合部位の両方を有するライブラリー分子の完全配列を産生しないために、ニックを有するライブラリー分子（すなわち、損傷ライブラリー）がより少ないシグナルを生成する様子を示す。

40

【0281】

いくつかの実施形態では、Cq値は、ライブラリー中の損傷の割合と相関する。いくつかの実施形態では、損傷はライブラリー調製中に導入された。

【0282】

いくつかの実施形態では、高いCq値は、ライブラリー分子のより多くのDNA損傷と相関する。いくつかの実施形態では、高いCq値を有するライブラリーは、より低い配列決定性能を示す。いくつかの実施形態では、より低い配列決定性能は、総出力（Gb）又はP1パーセントによって測定される。

50

【0283】

いくつかの実施形態では、変則的に低い（例えば、2.58未満）Cq値はまた、より低い配列決定性能を有し得る。

【0284】

いくつかの実施形態では、次のライブラリーの使用に応じて適切なデータ品質を有する配列決定ランを生じる所望のCq範囲を決定することができる。いくつかの実施形態では、所望のCq範囲は、2.58～5であり得る。Cq範囲は、使用されるライブラリーの特定の種に基づいて変動し得る。したがって、ユーザーは、十分な品質の配列決定データをもたらす所望のCq範囲を決定するために最初の実験を行い、次いで、この範囲内のCq値を有する配列ライブラリーのみを選択することができる。所望のCq範囲を決定する

10

【0285】

F. 長リード配列決定

標準的な短リード配列決定は、短距離情報を提供するために正確な塩基レベルの配列を提供するが、短リード配列決定は、長距離ゲノム情報を提供しない場合がある。更に、ハプロタイプ情報は、配列決定されたゲノム又は短リードデータを有する参照について保持されないため、長距離ハプロタイプの再構築は、標準的な方法では困難である。したがって、標準的な配列決定及び分析アプローチは、一般に、単一ヌクレオチドバリエーション（single nucleotide variant、SNV）をコールすることができるが、これらの方法は、個々のゲノムにおいてみられる構造変異の全スペクトルを同定しない場合がある。ゲノムの「構造変異」は、本明細書で使用される場合、50塩基対以上の事象を含む、SNVより大きな事象を指す。代表的な構造バリエーションには、コピー数変異、逆位、欠失、及び重複が含まれる。

20

【0286】

「連結長リード配列決定」又は「連結リード配列決定」は、ゲノム配列に関する長距離情報を提供する配列決定方法を指す。

【0287】

いくつかの実施形態では、連結リード配列決定は、分子バーコードを使用して、同じ長いDNA断片に由来するリードをタグ付けする。固有のバーコードが、個々のDNA分子から生成された全てのリードに付加される場合、リードは、そのDNA分子と一緒に連結することができる。言い換えれば、バーコードを共有するリードは、単一の長いインプット分子に由来するものとして分類することができ、長距離情報を短リードからアセンブリ

30

【0288】

いくつかの実施形態では、連結リード配列決定は、ハプロタイプ再構築のために使用され得る。いくつかの実施形態では、連結リード配列決定は、構造バリエーションのコーリングを改善する。いくつかの実施形態では、連結リード配列決定は、アクセス可能性が制限されたゲノムの領域へのアクセスを改善する。いくつかの実施形態では、連結リード配列決定は、デノボ二倍体アセンブリのために使用される。いくつかの実施形態では、連結リード配列決定は、デノボアセンブリを必要とする高度に多型の配列（ヒト白血球抗原遺伝子など）の配列決定を改善する。

40

【0289】

いくつかの実施形態では、配列決定は、5kb以上、10kb以上、15kb以上、20kb以上、25kb以上、又は30kb以上であるライブラリー分子の長リード配列決定である。

【0290】

G. 二本鎖DNA切断の調製を含む方法

いくつかの実施形態では、ニックは、二本鎖DNA切断に変換される。ニックから二本鎖DNA切断を生成する利点は、二本鎖切断がライブラリー産物において生成された後に

50

、完全なライブラリー分子に対応するアンプリコンが生成され得ないことである。このようにして、ニックを含むライブラリー分子は、ライブラリー産物の完全配列に対応するアンプリコンを全く生成しない。対照的に、二本鎖挿入部の一本鎖にニックを含むニックの入ったライブラリー分子は、より少ないアンプリコンを生成するが、ライブラリー産物の完全配列に対応するいくつかのアンプリコンを生成することができる（図9に示される）。これは、リバースプライマー又はフォワードプライマーのいずれかが、ライブラリー分子の完全配列に対応するアンプリコンを生成し得るためである。

【0291】

ニックから二本鎖切断を生成する利点は、二本鎖切断を有するライブラリー分子が、フォワードプライマー及びリバースプライマーとの結合部位の両方を有する完全長アンプリコンを生成することができないことである。

10

【0292】

いくつかの実施形態では、ニックは、エンドヌクレアーゼを使用して二本鎖切断に変換される。いくつかの実施形態では、エンドヌクレアーゼは、変異型T7エンドヌクレアーゼである。いくつかの実施形態では、変異型エンドヌクレアーゼは、マルトース結合タンパク質(MBP)-T7 Endo Iである。いくつかの実施形態では、T7エンドヌクレアーゼは、それまでニックが一本鎖中に位置していたDNAにおいて二本鎖切断を生成するために、カウンターニックを生成する。そのようなニックからの二本鎖切断の生成は、ニックを横切る切断と称され得る。

【0293】

H. SMRTbell 鑄型を用いる方法

いくつかの実施形態では、ライブラリー分子は、二本鎖DNA断片の末端にライゲーションされる2つのヘアピンアダプターを含む。いくつかの実施形態では、そのようなアダプターは、閉ループを形成する。

20

【0294】

本発明はこの調製方法に限定されないが、いくつかの実施形態では、ライブラリー分子はSMRTbell 鑄型である。SMRTbell 鑄型は、1分子リアルタイム(SMRT)配列決定と共に使用するために当該分野で周知である。いくつかの実施形態では、SMRT配列決定は、Pacific Biosciences (PacBio)の方法を使用する(例えば、Rhoads and Au, Genomics Proteomics Bioinformatics 13:278-289(2015)参照)。本明細書で使用される場合、SMRT配列決定及びPacBio配列決定は、互換的に使用され得る。

30

【0295】

SMRT配列決定手法は、環状コンセンサス配列決定(CCS)を利用して、>99%の精度かつ3回のパスを伴う高精度で長い高忠実度のリードを生成する。配列決定ランごとにHiFiリードの最高出力を生成するために、一定のローリングサークル増幅(RCA)を可能できる高品質のSMRTbell 鑄型が生成されなくてはならない。例えば、PacBio Sequelシステムは、オンプラットフォームRCAを使用して、ヘアピンアダプターがライゲーションされたライブラリー分子を配列決定できる。したがって、CCSリードを生成するために、ポリメラーゼは、挿入部の長さの3倍の長いポリメラーゼリード長を生成するように、反復パスで配列決定すべきである。

40

【0296】

SMRTシステムにおけるポリメラーゼが効率的に配列決定するためには、入力ライブラリーは高品質でなければならない。ライブラリー調製プロセスの間に、ピペティング、保存又は他の取り扱い及び/又は手技的エラーのいずれかによって、損傷がDNAに導入され得る。ニックの入ったSMRTbell 鑄型が配列決定のためにBio Sequelシステムにロードされる場合、ポリメラーゼはニック部位で脱落し、RCAを終結させ、結果として、P1パーセントは、その配列決定ランからのCCS出力と共に減少する。

50

【0297】

S M R T 配列決定の利点は、特定の他の配列決定方法よりも長いリード長と速いランである。例えば、P a c B i o システムは、60キロベースを超えるリード長を生成できることが知られている。これらのより長いリード長は、単一のリード内の反復領域の正確な位置及び配列を可能で、これは、他の配列決定プラットフォームでは利用できない場合がある。

【0298】

要約すると、S M R T 配列決定は、いくつかの他の方法よりも低いスループット、高いエラー率、及び高い塩基あたりコストを有することが知られており、ユーザーはこれらの欠点を最小限にすることを望むであろう。いくつかの実施形態では、本発明のライブラリーの品質管理の方法は、ユーザーが、S M R T 配列決定などの方法を用いて十分な品質の配列決定ランを生成する可能性が高い配列決定のためのライブラリーを選択することを可能にする。このようにして、ユーザーは、質の高い配列決定データを生成する能力を制限する D N A 損傷を有する配列決定ランに費やされる費用及び時間を回避することができる。

10

【0299】

いくつかの実施形態では、本明細書に記載される Q C 方法は、S M R T 配列決定ランからの P 1 パーセント及び総出力を最大化する。いくつかの実施形態では、本明細書に記載される q P C R Q C 方法は、顧客による S M R T 配列決定プラットフォーム上への損傷したライブラリーのロードの回避を可能にし、したがって時間、費用、試薬、及び消耗品の節約を可能にする。図 1 3 A ~ 図 1 5 C は、S M R T 配列決定を用いた Q C アッセイのいくつかの代表的なデータを示す。

20

【0300】

V . 蛍光を使用して D N A 損傷を決定する方法

D N A を含む試料中の D N A 損傷の量はまた、本明細書中に記載される方法によって蛍光を使用して測定され得る。いくつかの実施形態では、D N A 損傷は、ライブラリーが調製される前に、蛍光を使用して試料 D N A において定量化され得る。そのようなワークフローは、配列決定などの下流のアッセイに弊害をもたらす、試料中の D N A 損傷が多すぎることをユーザーが決定することを可能にするため、非常に魅力的であり得る。例えば、ユーザーは、試料中の D N A 損傷を定量化し、次いで、低レベル（例えば、5%以下）の D N A 損傷が存在する場合にのみ、試料からライブラリーを調製し得る。このようにして、ユーザーは、中程度（例えば、5%超）レベルの D N A 損傷を有する試料からライブラリーを調製しないことによって、時間及び資源を節約することができる。

30

【0301】

いくつかの実施形態では、蛍光を使用して D N A を含む試料中の D N A 損傷を定量化する方法は、

a .

i . D N A を含む試料のアリコートと、

i i . 1 つ以上の D N A 修復酵素と、

i i i . 1 つ以上の d N T P が蛍光標識されている、d N T P と、を組み合わせるこ

40

とと、

b . 修復された D N A を調製することと、

c . d N T P からリン酸を脱リン酸化することと、

d . 修復された D N A をカルボキシレートビーズ又はセルロースビーズに結合させることと、

e . 再懸濁緩衝液を用いて、結合した修復 D N A をカルボキシレートビーズ又はセルロースビーズから溶出することと、

f . 修復された D N A の蛍光を測定して、D N A 損傷の量を決定することと、を含む。

【0302】

D N A 損傷を定量化する方法の概要を図 1 6 に示し、この方法を用いた代表的な実験の

50

結果を図 17 ~ 図 21 に示す。

【0303】

いくつかの実施形態では、修復された DNA のより高い蛍光は、より多い DNA 損傷を示す。言い換えれば、より高いレベルの DNA 損傷が存在する場合、より多くの蛍光標識された dNTP が組み込まれる。

【0304】

いくつかの実施形態では、修復された DNA の蛍光は、異なる量の DNA 損傷の範囲にわたって線形である。このようにして、アッセイのダイナミックレンジ（すなわち、正確に測定できる DNA 損傷の全範囲）が改善され、したがって、ユーザーは、様々なライブラリーに対する損傷の相対的な差を評価することができる。いくつかの実施形態では、比較的少量の DNA 損傷が結果に悪影響を及ぼし得る高感度の下流アッセイのためにユーザーが試料を評価している場合、広い線形範囲は、その量の DNA 損傷を正確に決定するのに役立つ。

10

【0305】

いくつかの実施形態では、方法は、試料のアリコート中の DNA 損傷を評価することができる。言い換えれば、ユーザーは、少量の試料を採取し、DNA 損傷を定量化し、次いで、DNA 損傷の定量化の結果に基づいて、潜在的により多くのアッセイ（ライブラリー調製又は配列決定等）を行うことができる。

【0306】

いくつかの実施形態では、方法は、操作の前後に同じ試料のアリコートを評価することによって、試料の操作によって誘発された DNA 損傷を評価することができる。このようにして、ユーザーは、操作によって誘発された任意の DNA 損傷を直接測定することができる。

20

【0307】

いくつかの実施形態では、操作は試料の配列決定である。例えば、ユーザーは、特定の試薬が DNA 損傷を誘発するかどうかを決定するために、DNA を含む試料に対する異なる配列決定試薬の影響の評価を望む場合がある。

【0308】

いくつかの実施形態では、修復された DNA の蛍光を測定することは、修復された DNA の希釈物の標準曲線を作成することと、修復された DNA の希釈物の蛍光を測定することと、を含む。いくつかの実施形態では、標準曲線の使用は、アッセイのダイナミックレンジを増加させて、少量の DNA 損傷の定量化を可能にできる。少量の DNA 損傷を定量化するためのこのような方法は、少量の DNA 損傷であっても下流のアッセイ（配列決定など）の結果に弊害をもたらし得る場合に有用であり得る。

30

【0309】

いくつかの実施形態では、修復された DNA の蛍光を測定することは、修復された DNA の蛍光を、蛍光標識された 1 つ以上の dNTP のみの希釈物の別個の標準曲線に対して比較し、修復された DNA に含まれる蛍光色素分子の数を決定することを含む。

【0310】

いくつかの実施形態では、方法は、決定された蛍光色素分子の数を修復された DNA の質量で割ることによって、修復された DNA に含まれる蛍光色素分子の正規化された数を計算することを更に含む。そのような尺度は、DNA の損傷されている割合を推定することができる。

40

【0311】

いくつかの実施形態では、DNA は、ゲノム DNA、cDNA、又は断片化二本鎖 DNA を含むライブラリーである。DNA がゲノム DNA 又は cDNA である場合、この方法は、ライブラリー調製の前に実施され得る。

【0312】

いくつかの実施形態では、DNA は、ゲノム DNA 及び cDNA であり、方法は、DNA 損傷の量を決定した後ライブラリーを調製することを更に含む。

50

【0313】

いくつかの実施形態では、DNA損傷の量が全ヌクレオチドの5%以下、4%以下、3%以下、2%以下、又は1%以下である場合にライブラリーが調製される。言い換えれば、ライブラリーは、DNA損傷が低いと決定される場合に調製され得る。ライブラリー調製又は他の下流アッセイに受容可能なDNA損傷の量は、下流アッセイの感度及びDNA損傷の種類に依存する。例えば、短リード配列決定は、中程度のレベルのDNA損傷（例えば、5%以下）であっても、許容可能な配列決定結果を与え得る。対照的に、長リード配列決定は、許容可能な結果のためにより低いレベルのDNA損傷（例えば、2%以下）を必要とする場合があり、また、ニックングによって誘発される損傷に対してより感受性であり得る。

10

【0314】

いくつかの実施形態では、本発明のアッセイが特定の種類の損傷（ニックングなど）の存在を決定する場合、この損傷は、ライブラリー調製又は配列決定などの更なる工程の前に修復され得る。

【0315】

いくつかの実施形態では、DNA損傷の量が全ヌクレオチドの5%以上、4%以上、3%以上、2%以上、又は1%以上である場合、ライブラリーは調製されない。このようにして、ユーザーは、下流アッセイの結果に悪影響を及ぼすレベルのDNA損傷が存在する場合に、ライブラリーを調製する（及び配列決定などの更なる下流アッセイを実行する）際に時間及び資源を浪費することを回避する。

20

【0316】

いくつかの実施形態では、蛍光を測定する前に、修復されたDNAをカルボキシレートビーズ又はセルロースビーズに結合させ、溶出することが2回以上行われる。いくつかの実施形態では、複数回のビーズベースの精製は、方法の結果を改善する。いくつかの実施形態では、複数回のビーズベースの精製は、非特異的シグナルを減少させる。いくつかの実施形態では、蛍光を測定する前に、ビーズベースの精製が複数回、修復されたDNAをカルボキシレートビーズ又はセルロースビーズに結合させ、溶出することが2回行われる。

【0317】

カルボキシレートビーズ（例えば、SPRIビーズ）及びセルロースビーズは、DNA精製及びサイズ選択用途のために市販されており、このようなビーズは、本発明の方法において使用され得る。

30

【0318】

いくつかの実施形態では、カルボキシレート又はセルロースビーズは、磁性である。この特性は、修復されたDNAの結合後のビーズの洗浄に役立ち得る。

【0319】

いくつかの実施形態では、修復されたDNAの調製は、37℃で行われる。いくつかの実施形態では、修復されたDNAを調製することは、10分間以上、20分間以上、30分間以上、45分間以上、又は60分間以上行われる。

【0320】

いくつかの実施形態では、dNTPからリン酸を脱リン酸化することは、dNTPの非特異的結合を減少させ、アッセイ結果を改善する。

40

【0321】

いくつかの実施形態では、dNTPからリン酸を脱リン酸化することは、酵素を用いて行われる。いくつかの実施形態では、dNTPからリン酸を脱リン酸化するための酵素は、エピアルカリホスファターゼ（SAP）又は仔ウシ腸アルカリホスファターゼ（CIP）である。

【0322】

様々な異なるDNA修復酵素をこの方法で使用することができ、本明細書で使用される場合、「DNA損傷」は、単一試料中に含まれるDNA中に存在し得る複数の異なる種類

50

のDNA修飾（例えば、ニック及びチミンダイマー）を指し得る。

【0323】

いくつかの実施形態では、1つ以上のDNA修復酵素は、DNAポリメラーゼを含む。いくつかの実施形態では、DNAポリメラーゼは、5' - 3'ポリメラーゼ活性を有するが、5' - 3'エキソヌクレアーゼ活性を欠いている。いくつかの実施形態では、DNAポリメラーゼは、Bst DNAポリメラーゼ、ラージフラグメントである。いくつかの実施形態では、1つ以上のDNA修復酵素は、リガーゼを含む。いくつかの実施形態では、リガーゼは、Taqリガーゼである。いくつかの実施形態では、DNA損傷は、二本鎖DNA中のニックを含む。

【0324】

いくつかの実施形態では、1つ以上のDNA修復酵素は、T4ピリミジンダイマーグリコシラーゼ（PDG）を含む。いくつかの実施形態では、DNA損傷は、チミンダイマーを含む。いくつかの実施形態では、チミンダイマーは、紫外線照射によって誘発された。

【0325】

いくつかの実施形態では、1つ以上のDNA修復酵素は、ウラシルDNAグリコシラーゼ（UDG）及び脱プリン又は脱ピリミジン部位リアーゼを含む。いくつかの実施形態では、DNA損傷は、ウラシルを含む。

【0326】

いくつかの実施形態では、1つ以上のDNA修復酵素は、ホルムアミドピリミジンDNAグリコシラーゼ（FPG）及び脱プリン又は脱ピリミジン部位リアーゼを含む。いくつかの実施形態では、DNA損傷は、酸化塩基を含む。

【0327】

いくつかの実施形態では、2つ以上のDNA修復酵素が使用される。いくつかの実施形態では、1つ以上のDNA修復酵素は、複数のDNA修復酵素の混合物である。そのようなアプローチは、DNA損傷がDNAへの2つ以上の種類の損傷修飾（すなわち、チミンダイマー及びニック又は任意の他の修飾の組み合わせ）を含み得るとユーザーが疑う場合に使用され得る。

【0328】

いくつかの実施形態では、dNTPは、dATP、dGTP、dCTP、及びdTTP又はdUTPを含む。いずれか又は全てのdNTPを蛍光標識してもよい。いくつかの実施形態では、全てのdNTPが蛍光標識される。いくつかの実施形態では、dUTP及びdCTPが蛍光標識される。

【0329】

任意の適切な蛍光標識が、dNTPに含まれ得る。いくつかの実施形態では、蛍光標識は、Alexa Fluor 488、Alexa Fluor 546、Alexa Fluor 555、Alexa Fluor 633、フルオレセインイソチオシアネート（FITC）、又はテトラメチルローダミン - 5 - （及び6） - イソチオシアネート（TRITC）であるが、励起スペクトルにおける広範な他の蛍光標識が使用されてもよい。いくつかの実施形態では、蛍光標識は、DNAを損傷しない励起波長を有する。

【0330】

実施例

実施例1．標準物質を使用したLongAmp PCR反応のアンプリコンサイズバイアスの正規化

図1は、代表的なLongAmp PCR反応と、それに続く、例えばNextera製品（Illumina）を用いた断片化を示す。本明細書に記載されているように、異なる長さの核酸標準物質のプールを使用して、この実験におけるアンプリコンサイズバイアスを正規化することができる。

【0331】

長い増幅PCRを行って、試料内の標的核酸断片に含まれる目的の配列からアンプリコンを生成することができる（図1に示される）。試料は、遺伝子編集に供された核酸から

10

20

30

40

50

構成される試料であってもよく、ユーザーは、いくつかの異なる種類のインデル変異が存在し得ると予想する。

【0332】

このPCR反応の間、本明細書に記載されるように、異なる長さの核酸標準物質のプールを反応に含めることができる。このプールは、完全長標準物質（例えば、図8Aに示されるもの）、挿入標準物質（例えば、図8Bに示されるもの）、及び欠失標準物質（例えば、図8Cに示されるもの）を含み得る。このようにして、標準物質は、目的の配列と同じ条件下で増幅される。

【0333】

挿入標準物質を作成する代表的な方法は以下の通りである。

10

工程1) N18 UMIを含む図2Aに示されるオリゴヌクレオチドを、制限部位3 (RS3) 及び制限部位4 (RS4) で切断する制限酵素を用いて消化する。

工程2) 図3のPCR産物を、制限部位1 (RS1) 及び制限部位2 (RS2) で切断する制限酵素によって消化する。

工程3) 図4のPCR産物をRS1及びRS2によって消化する。

工程4) 工程2及び3の産物をライゲーションする。

工程5) 図4のPCR産物をRS3及びRS4によって消化する。

工程6) 工程5の産物を工程1の産物とライゲーションする。

【0334】

挿入標準物質を調製するためのこれらの工程は、図2Bに示される産物を生成することが期待される。RS消化の順序は固定されていない。更に、RSで消化する全ての制限酵素が緩衝液相溶性である場合、全ての消化工程を組み合わせてもよい。あるいは、消化工程は別々の工程で行われてもよい。ライゲーション工程（工程4及び6）はまた、挿入標準物質を調製する方法における最終工程として組み合わせられ得る。

20

【0335】

欠失標準物質を作成する代表的な方法は以下の通りである。

工程1) 図6Aに示すオリゴヌクレオチド（図2Aに示すオリゴヌクレオチドと同一であり、N18 UMIを含む）を、RS3及びRS4によって消化する。

工程2) 図5のPCR産物をRS3及びRS4によって消化する。

工程3) 工程2の産物を工程1の産物とライゲーションする。

30

【0336】

欠失標準物質を調製するためのこれらの工程は、図6Bに示される産物を生成することが期待される。

【0337】

目的の配列を標準物質と共に増幅した後、次いで、アンプリコン（標準物質由来及び目的の配列由来）を、配列決定ライブラリーを調製するための方法に供し得る。図1Aは、これが、トランスポザーゼが断片の両端にアダプター配列を組み込むNextera断片化（すなわち、タグメンテーション）であり得ることを示す。次いで、断片は、これらのアダプター配列に含まれる配列（例えば、配列決定プライマー結合部位）を使用して配列決定され得る。

40

【0338】

次いで、ライブラリー（目的の配列及び標準物質から生成された断片から構成される）を、配列決定できる。個々の標準物質に含まれるUMIを使用して、バイアスプロファイルを生成することができる。このバイアスプロファイルは、所与の標準物質の複製物が標準物質のUMIを使用して同定され得るため、より大きな標準物質がより少ない固有の複製物を有するという事実を説明する。これらのデータを使用して、アンプリコンサイズバイアスを正規化することができる。このようにして、ユーザーは、目的の配列中の所与のインデル変異を有した元のコピー数を概算することができる。言い換えれば、この方法は、目的の配列の大きな挿入変異（得られる目的の配列のアンプリコンが、有意に大きくなる）が、目的の野生型配列又は目的の配列の欠失変異よりも少ないアンプリコンを生成す

50

るという事実を制御できる。

【0339】

実施例2．ライブラリーの品質管理評価

ライブラリーの品質管理（QC）のために定量的PCR（qPCR）アッセイを実施した。QC qPCRアッセイは、ニックのついていない鋳型鎖を増幅するために、PrimeStar GXL DNAポリメラーゼ（Takara）（長い標的（例えば、30 kb超）を高い忠実度で増幅することができることが知られている長距離ポリメラーゼ）を使用した。増幅の間、ライブラリー分子に含まれるヘアピンアダプターに特異的なフォワードプライマーは、鋳型がニックによって破壊されない場合にのみ、反対のアダプターに伸長し、リバースプライマーのための新しい鋳型鎖を生成する。対照的に、新しい鋳型鎖由来のシグナルは、ポリメラーゼがニックに遭遇した場合には生成されない（図9に示される）。

10

【0340】

コントロール実験を行って、ニックがCq値にどのように影響するかを決定した。qPCRマスターミックスは、0.5 U長距離ポリメラーゼ（PrimeStar GXLポリメラーゼ）、ヘアピンアダプター内の特定の配列に結合するようにそれぞれ設計されたフォワードプライマー及びリバースプライマー、1×EvaGreen、200 μMの各dNTP、1×PrimeStarバッファー、及び約200 pg/μLのDNA入力（入力は、必要に応じてfg範囲に減少させ得る）から構成された。

20

【0341】

20×EvaGreenを水中で5×に希釈し、次いで、効率的な増幅を確認するために試料を用いて作製した標準曲線（Nexteraアダプター及びP5/P7増幅プライマーを有するライブラリー）と共に反応プレート上に含めた。以下のサイクルパラメータを実施した。95 で2分間の初期変性、続いて95 で30秒間、50 で30秒間、及び68 で15秒間を30サイクル。反応を2回行い、Cq値を平均した。

【0342】

表2は、qPCRマスターミックスの概要を提供する。

【0343】

【表2】

表2：qPCRマスターミックス			
試薬	μL	最終	ベンダー／カタログ#
1 ng/μLの約10 kbヘアピンライブラリー	2	200 pg/μL	
100 μM フォワードプライマー	0.05	500 nM	IDT
100 μM リバースプライマー	0.05	500 nM	IDT
5×EvaGreen	1.6	0.8X	Biotium、 #31000
2.5 mM 各dNTP	0.8	200 μMずつ	Takara R 050A
5×PrimeStar緩衝液	2	1X	Takara R 050A
PrimeStar GXLポリメラーゼ (1.25 U/μL)	0.2	0.025 U/μL	Takara R 050A
H ₂ O	3.3		
総容量	10		

30

40

【0344】

50

EvaGreen (登録商標) Dye及びEvaGreen (登録商標) Plus Dyeは、それ自体は本質的に非蛍光性であるが、dsDNAに結合すると高度蛍光性になる緑色蛍光核酸色素である。したがって、EvaGreenは、デジタルPCR及び等温増幅用途に使用することができる。

【0345】

ニック処理により、10ngライブラリー(図10A及び図10B)及び20ngライブラリー(図10C及び図10D)の両方について、DNA損傷及び平均Cqの用量依存的増加が引き起こされた。これらの結果は、このQCアッセイのqPCR結果が、より高品質のライブラリーに対してより低いCqをもたらす、損傷ライブラリー(例えば、ニックを含有するライブラリー分子から構成されるもの)に対してより高いCqをもたらすことを示す。

10

【0346】

Vibrio vulnificusヌクレアーゼ(VVN、非特異的ヌクレアーゼ)とT7エンドヌクレアーゼ変異体との組み合わせを使用してニックから二本鎖切断を調製するためのエンドヌクレアーゼ処理(図11並びに図12A及び図12B)後に、同様の結果が見られた。したがって、ニックの入った鋳型から二本鎖切断を調製し、増幅に必要なプライマー配列の分離をもたらすことは、QCアッセイが不十分な品質のライブラリーを同定できることを更に実証する。

【0347】

実施例3. SMRTbellライブラリーの品質管理

20

図13A~図15Cは、実施例2に記載の方法を用いた、二本鎖断片の両端にヘアピンアダプターを含有するSMRTbellライブラリーを用いた更なる実験を示す。異なるライブラリーについてのこれらの分析は、総配列決定出力が、より低いCq値を有するライブラリーについて一貫して増加することを確認する。言い換えれば、QC工程におけるqPCR結果と測定された総配列決定出力(すなわち、配列決定されたギガベース)との間に強い相関があった。一般に、QCアッセイにおいてより低いCq値を有するライブラリーは、より高い総配列決定出力を有した。例えば、QCアッセイにおいて約3のCq値を有するライブラリーでは、Cq値が9を超えたときの17%と比較して、39%~67%のP1パーセントの変動が見られた(図13A~図13C)。ライブラリー8は、この関係に対する外れ値として注目される。

30

【0348】

更に、図14A~図14Cのデータは、3~4の範囲のCq値が平均で約366ギガベースを生成したことを示す。対照的に、ライブラリー10は、6を超えるそのQC値に基づいて実施不十分であると予測され(図14A)、配列決定結果は、比較的低い総出力とP1パーセントを示した(図14B及び図14C)。したがって、QCアッセイは、不十分な配列決定性能を有するライブラリーを予測することができた。一般的に、ライブラリーの平均Cqが低いほど、P1パーセントが高いという関係が見られたが、これは、ライブラリー14(図13A~図13Cにおけるライブラリー8に対応する)については当てはまらなかった。

【0349】

40

図15A~図15Cは、同様に、高いCq値を有するライブラリー画分と比較して、QCアッセイにおいてより低いCq値を有するライブラリー画分(すなわち、F4、F5、及びF6などの同じライブラリーから調製された異なる画分)について、最も良い総配列決定出力(ギガベース)が見られたことを示す。

【0350】

したがって、本発明のQC法は、個々のライブラリーを配列決定すること(又は配列決定しないこと)を決定するための有用なツールである。そのようなQC方法は、既存のQC方法のみに基づいてユーザーが予測できないようにライブラリーの品質が変化し得るため、特に有益である。例えば、1つの試料で使用されるピペッティング力は、同じユーザーによって生成された他のライブラリーでは見られない分解を引き起こし得る。既に作製

50

されたライブラリーの質を評価できる方法のみが、配列決定データの質に影響を与えるランダムな変化を制御することができる。したがって、当業者は、最初の実験を使用して、使用される特定のライブラリーに基づいて、所望のCq値の範囲を生成することができ、これは、QC法を使用する配列決定のためのライブラリーを選択するために使用され得る。

【0351】

実施例4．蛍光を用いたDNA損傷の測定

ユーザーはまた、蛍光を使用してDNA損傷を測定することを望み得る。例えば、ユーザーは、試料中のDNA損傷のレベルが許容可能であることを保証するために、ライブラリーを調製する前にDNA損傷を測定することを望む場合がある。例えば、ユーザーは、ライブラリー調製前のゲノムDNA又はcDNAに対して、又は既に調製されたライブラリーに対して使用するのに柔軟なDNA損傷を定量化する方法を使用することを望む場合がある。しかしながら、蛍光標識されたヌクレオチド及びタンパク質の両方を含有する現在のアッセイは、組み込まれていない蛍光ヌクレオチドの高い非特異的結合に悩まされることが多い。

10

【0352】

本発明のアッセイは、蛍光定量のシグナル対ノイズ比を改善するために開発された。この方法は、非特異的結合を顕著に減少させるために、エピアルカリホスファターゼ(SAP)消化及びSPRI(カルボキシレートビーズ)結合/溶出工程の両方を用いる。ユーザーの好みに依存して、セルロースビーズを、カルボキシレートビーズの代わりに使用でき、仔ウシ腸アルカリホスフェートを、記載される方法のいずれかにおいてSAPの代わりに使用できる。

20

【0353】

図16は、蛍光標識されたdNTPの存在下でのDNA修復工程(この例では、Bstポリメラーゼ及びTaqリガーゼを用いる)と、それに続くSAPによる処理及びSPRIビーズベースの精製の2工程を組み込む、本発明の方法を概説する。次いで、修復されたDNAを含む処理された試料を測定して、蛍光の量を決定する。

【0354】

最初の実験は、dNTPの非特異的結合を減少させるための異なる条件を試験した。図17は、単一のSPRIビーズベースの精製により、切断されたゲノムDNA(gDNA)のSAP処理が、SAP処理を伴わないアッセイと比較して、蛍光ヌクレオチドの非特異的結合を実質的に低減したことを示す。言い換えれば、SAP処理を伴うビーズベースの精製工程は、非特異的蛍光を減少させた。

30

【0355】

更に、図18は、2回目のSPRIビーズベースの精製工程が、蛍光ヌクレオチドの非特異的結合を緩衝液に匹敵するレベルまで低下させたことを示す。このような低いバックグラウンドは、少量のDNA損傷(すなわち、DNA中の低い割合のヌクレオチドが損傷されている場合)を正確に測定するために重要である。

【0356】

最初の実験に基づいて、SPRIビーズに基づく精製の2回の工程を、更なる実験においてSAP処理後に行った。本発明の方法を用いて、市販の修復ミックス対我々のDNA修復酵素ミックスの有効性の比較を行った。PreCR Repair Mix(NEB)を、本プロトコルを用いて、Taqリガーゼ(40U)、Bstポリメラーゼラージフラグメント(8U)、及びT4PDG(1U)のカスタム修復ミックスと比較した。図19Aに示されるように、PreCRミックスは、試料の損傷が増加するにつれて予想される蛍光増加を示さなかったが、カスタム修復ミックスは、これらの予想される増加を示した。PreCR Mix試料はまた、より大きな標準偏差及び低いシグナルを有し、このような不一致はまた、DNA損傷修復処方を最適化するグループからの文献において見出され得る。対照的に、本発明の方法を使用したカスタム修復酵素ミックスは、低い標準偏差及びより高いシグナル対ノイズ比を有した(図19B)。

40

50

【0357】

ユーザーによって決定されたDNA修復酵素のカスタムミックスを用いる本発明の方法はまた、ユーザーがアッセイにおいてどの修復酵素を利用するかを選択することができるため、ワークフローに柔軟性を加える。例えば、本発明のアッセイは、異なるDNA損傷修復酵素を利用することによって、DNA中の異なる種類の損傷を検出するように設計することができる。DNA修復酵素混合物中にT4ピリミジンダイマーグリコシラーゼ(T4PDG)酵素を組み込むことにより、チミンダイマーなどのUV照射によって引き起こされる損傷の修復及びその後の検出を可能にする。図20に示すように、Taqリガーゼ、Bstポリメラーゼ、及びT4PDG(UV損傷特異的修復酵素)を含むDNA修復酵素混合物を使用する方法は、UV誘発性DNA損傷を評価することができた。UV光の量及び曝露時間が増加するにつれて、本発明のアッセイによって測定されるDNA損傷も増加し、本発明のアッセイが広い範囲にわたってDNA損傷を測定する能力を示す。

10

【0358】

図21は更に、DNA試料が異なる量のニッキング酵素(Nt.BspQI)に曝露されたとき、DNA損傷測定の蛍光シグナルが増加したことを示す。したがって、本発明のアッセイは、広い範囲にわたってニックの入ったDNAの量を感度よく測定することができる。

【0359】

ユーザーが望む場合、ウラシルDNAグリコシラーゼ(UDG)及び脱プリン若しくは脱ピリミジン部位リナーゼ、並びに/又は、ホルムアミドピリミジンDNAグリコシラーゼ(FPG)及び脱プリン若しくは脱ピリミジン部位リナーゼを酵素修復ミックスに組み込むことにより、それぞれウラシル又は酸化塩基の修復及びその後の検出を可能にする。

20

【0360】

このアッセイのモジュール方式により、使用される酵素の活性及び特異性に基づいて、二本鎖DNAにおける異なる種類の損傷を検出するための柔軟でカスタマイズ可能なツールにする。

【0361】

実施例5. 蛍光を用いたDNA損傷の測定

最初の実験に基づいて、例示的なアッセイプロトコールを、Taqリガーゼ、Bstポリメラーゼ、及びT4PDGを含むDNA修復酵素ミックスと共に使用するために開発した。表3は、このアッセイで使用するための試薬を提供し、表4は、dNTPマスターミックス含量を提供し、表5は、DNA損傷アッセイの含量を提供する。

30

【0362】

40

50

【表 3】

表 3 : DNA 損傷を測定するための試薬		
材料	サプライヤ	部品番号
M i l l i Q 水	--	--
200 プルーフェタノール	S i g m a A l d r i c h	E 7 0 2 3
0. 2mL ストリップチューブ	U S A S c i e n t i f i c	1 4 0 2 - 4 7 0 8
A M P u r e P B ビーズ	P a c i f i c B i o s c i e n c e s	1 0 0 - 2 6 5 - 9 0 0
10×ThermoPol 緩衝液	N e w E n g l a n d B i o l a b s	B 9 0 0 4 S
100×NAD+	N e w E n g l a n d B i o l a b s	B 9 0 0 7 S
A l e x a F l u o r 5 4 6 - 1 4 - d U T P、1mM	T h e r m o F i s h e r	C 1 1 4 0 1
A l e x a F l u o r 5 5 5 - a h a - d C T P、1mM	T h e r m o F i s h e r	A 3 2 7 7 0
dGTP、100mM	P r o m e g a	U 1 3 3 0
dATP、100mM	P r o m e g a	U 1 3 3 0
再懸濁緩衝液	I l l u m i n a	--
B s t ポリメラーゼ、ラージフラグメント (8U/uL)	N e w E n g l a n d B i o l a b s	M 0 2 7 5 S
T a q リガーゼ (40U/μL)	N e w E n g l a n d B i o l a b s	M 0 2 0 8 S
T 4 P D G (10U/uL)	N e w E n g l a n d B i o l a b s	M 0 3 0 8 S
エビアルカリホスファターゼ (1U/μL)	N e w E n g l a n d B i o l a b s	M 0 3 7 1 S
黒色 96 ウェルハーフエリアプレート	C o r n i n g	3 6 9 4
Q u b i t 1×dsDNA 高感度キット	T h e r m o F i s h e r	Q 3 3 2 3 1
0. 5mL Q u b i t チューブ	T h e r m o F i s h e r	Q 3 2 8 5 6

10

20

30

40

50

【 0 3 6 3 】

【表 4】

成分	体積
Alexa Fluor 546-14-dUTP、100 μ M	0.25 μ L
Alexa Fluor 555-ah-a-dCTP、100 μ M	0.25 μ L
dGTP、100 μ M	0.25 μ L
dATP、100 μ M	0.25 μ L
MilliQ水	9 μ L
総容量	10 μ L

10

【0364】

【表 5】

成分	体積
MilliQ水	2.4 μ L
10 \times ThermoPol緩衝液	1 μ L
dNTPマスターミックス、10 μ M	2.5 μ L
Bst、ラージフラグメント (8U/ μ L)	1 μ L
Taqリガーゼ (40U/ μ L)	1 μ L
T4 PDG (10U/ μ L)	0.1 μ L
gDNA (200ng)	2 μ L
総容量	10 μ L

20

【0365】

30

代表的なアッセイプロトコールを以下のように実施できる。

1. dNTP希釈物及びdNTPマスターミックスを表4及び5に記載のように調製する。氷上に置く。

2. Qubitを使用して試料及び対照gDNAを定量化する。gDNAを100ng/ μ Lに希釈し、氷上に置く。

3. アッセイ混合物を氷上のストリップチューブ中に、試料当たり2つ調製し、穏やかにピペティングして混合する。加熱した蓋を備えたサーモサイクラーにおいて37で30分間インキュベートする。

4. 30分後、サーモサイクラーから取り出し、1 μ Lのエピアルカリホスファターゼ(SAP)を各試料に添加する。穏やかにピペティングして混合し、加熱した蓋を備えたサーモサイクラーにおいて37で60分間インキュベートする。

40

5. インキュベーション後、再懸濁緩衝液(RSB)で100 μ Lに希釈する。AMPure PB(SPRI)ビーズをボルテックスして混合し、100 μ LのSPRIビーズを添加する。ピペティングして混合し、室温で15分間穏やかに振盪する。

6. 卓上磁気ラックを使用してビーズを磁化し、ビーズペレットを乱すことなく、試料を100 μ Lの80%エタノールで2回洗浄する。2回目の洗浄後、確実に遠沈させ、全てのエタノールを完全に吸引する。

7. ビーズを100 μ LのRSB中に再懸濁する。室温で15分間穏やかに振盪する。

8. 卓上磁気ラックを使用してビーズを磁化し、上清を新しいストリップチューブに吸引する。

50

9. 任意選択で、S P R I の洗浄を繰り返す（工程 5 ~ 8）。

10. 5 n M で開始し、濃度を半分に減少させる R S B 中の A F - 5 4 6 d U T P を使用して、100 μ L の標準曲線用試料を作製する。（5 n M、2.5 n M、1.25 n M、625 p M、312 p M、156 p M、78 p M、及び39 p M）

11. 45 μ L の各精製試料を96ウェルプレートに2箇所ずつピペットで入れる。45 μ L の標準曲線用試料を96ウェルプレートに2箇所ずつピペットで入れる。

12. プレートを C y t a t i o n 5 マルチモードリーダー（A g i l e n t）のプレートホルダーに入れる。A l e x a F l u o r 5 4 6 をフルオロフォアとして選択し、試料及び標準曲線用試料の蛍光を1回の読み取りで測定する。

13. 残りの試料及び対照を R S B 中で 1 : 10 に希釈し、回収した D N A を Q u b i t で定量する。標準曲線を用いて、D N A に取り込まれた色素分子を計算する。色素分子の数を回収された g D N A の質量で割り、色素分子の正規化された数を決定する。

【0366】

当業者は、試料中の D N A 損傷を評価するために、彼らの好みの D N A 修復酵素混合物を用いてこの代表的なプロトコールを使用することができる。

【0367】

均等物

前述の明細書は、当業者が実施形態を實踐することを可能にするのに十分であると考えられる。前述の説明及び実施例は、特定の実施形態を詳細に詳述し、本発明者らによって想到される最良の様式を説明する。しかしながら、前述の内容が本文にどれほど詳細にあらわれていても、実施形態は多くの方式で実施することができ、添付の特許請求の範囲及びその任意の均等物に従って解釈されるべきであることが理解されるであろう。

【0368】

本明細書で使用するとき、約という用語は、明示的に示されているか否かにかかわらず、例えば、整数、割合、及び百分率を含む数値を指す。用語は、一般に、当業者が列挙された値（例えば、同じ機能又は結果を有する）と同等であると考えられる数値の範囲（例えば、列挙された範囲の $\pm 5 \sim 10\%$ ）を指す。少なくとも及び約などの用語が数値又は範囲のリストの前にある場合、その用語はリスト内に提供される値又は範囲の全てを変更する。場合によっては、約という用語は、最も近い有効数字に丸められた数値を含んでもよい。

10

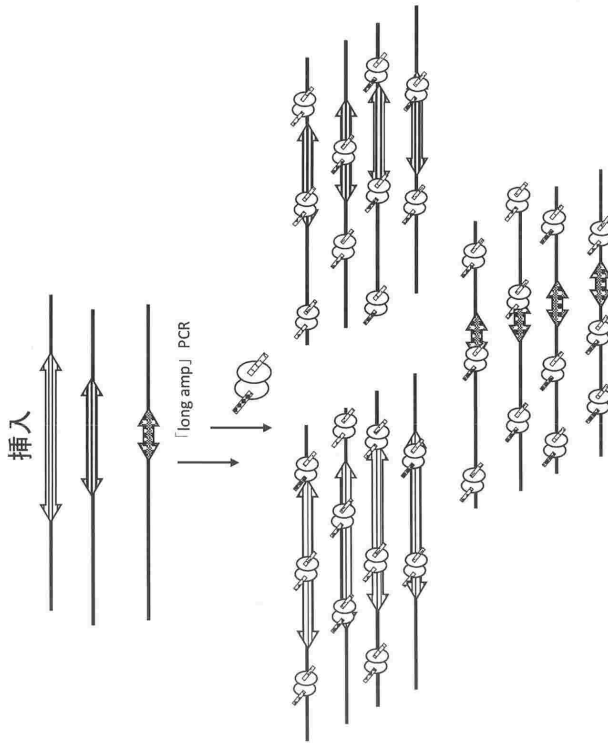
20

30

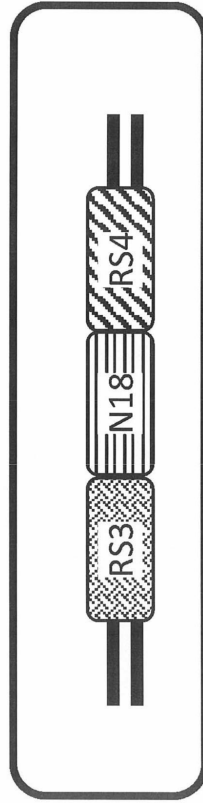
40

50

【 図 面 】
【 図 1 】



【 図 2 A 】



10

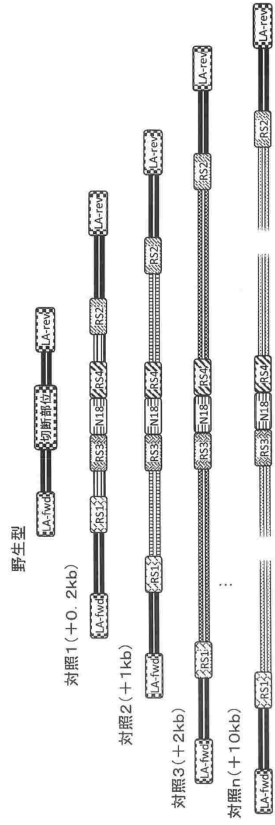
20

30

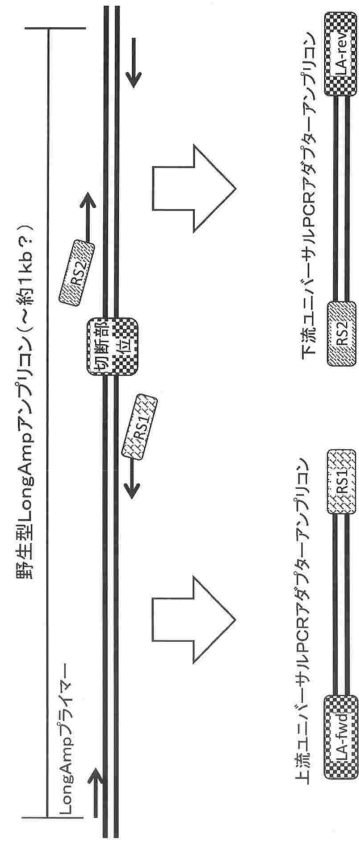
40

50

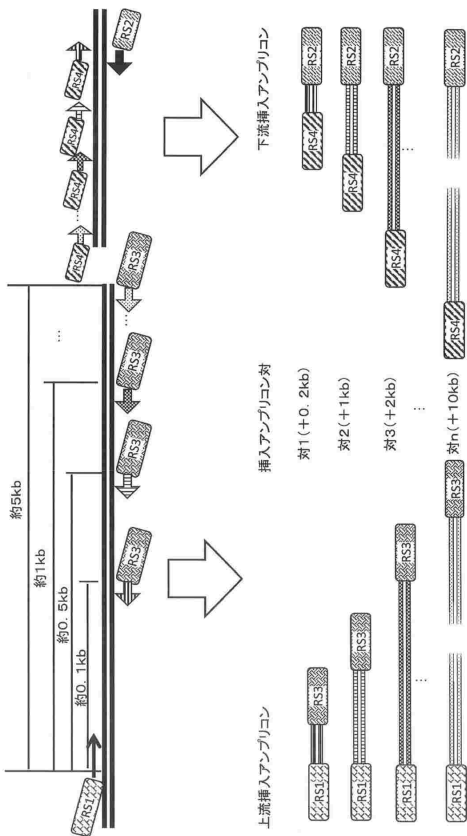
【 図 2 B 】



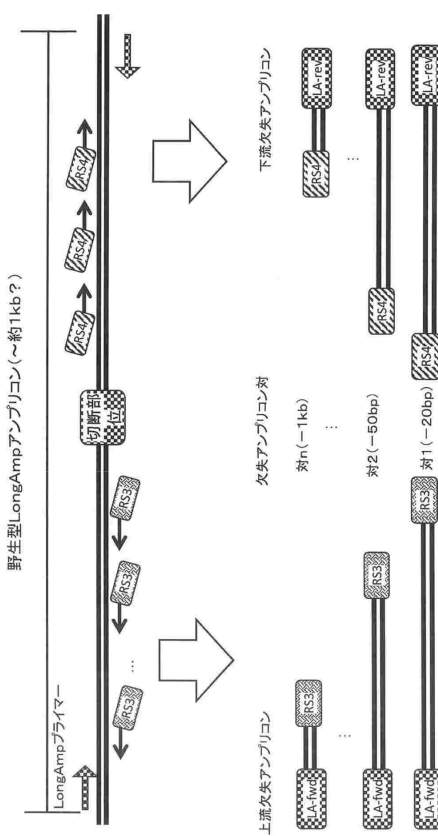
【 図 3 】



【 図 4 】



【 図 5 】



10

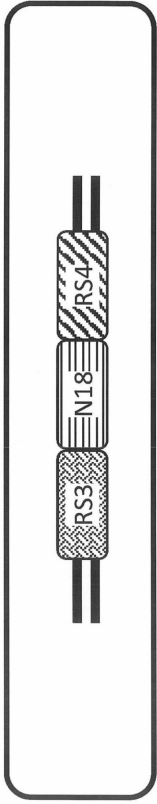
20

30

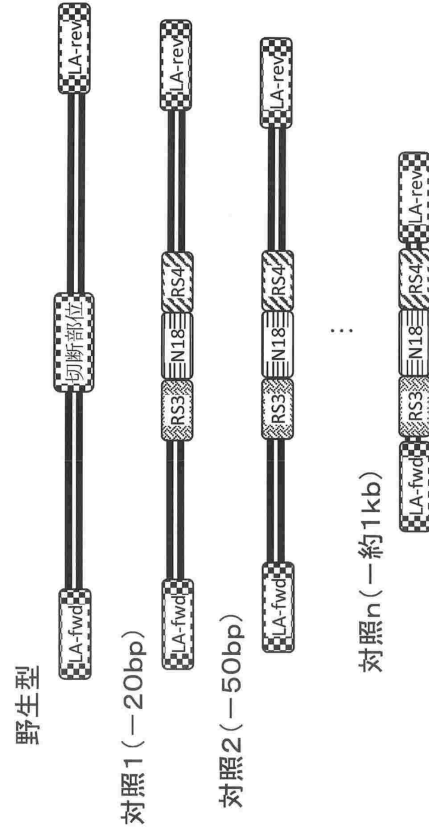
40

50

【 図 6 A 】



【 図 6 B 】



10

20

【 図 7 】

UMI重複を回避するためのLongAmp反応への対照入力の本数

隣接領域長(bp): 1050bp

プライマーについて1kb+50bpと仮定

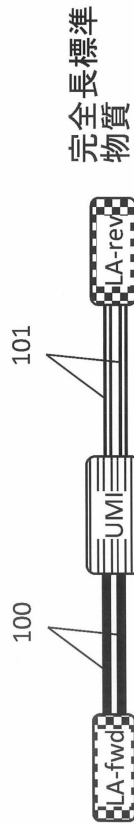
合成dsDNAオリゴ (IDT, gBlock; 200ngスケール)

		全てのUMIについての理論量					
UMIの長さ	UMIの#	コスト	dsDNAオリゴ長	MW(g/mol)	#mol	#グラム	#ナノグラム
18	6.87E+10	1439	268	1.77E+05	1.14E-13	2.02E-08	20.2

対照アンプリコン(UMI+dsオリゴを含有)

タイプ	対照アンプリコン#	挿入/欠失サイズ	全てのUMIについての理論量(UMI重複を回避するための最大量)				
			MW(g/mol)	最小分子(#UMI)	#mol	#グラム	
挿入	1	200	8.25E+05	6.87E+10	1.14E-13	9.42E-08	94.2
挿入	2	1000	1.35E+06	6.87E+10	1.14E-13	1.54E-07	154.4
挿入	3	2000	2.01E+06	6.87E+10	1.14E-13	2.30E-07	229.8
挿入	n	10000	7.29E+06	6.87E+10	1.14E-13	8.33E-07	832.5
欠失	1	-20	6.80E+05	6.87E+10	1.14E-13	7.76E-08	77.6
欠失	2	-50	6.60E+05	6.87E+10	1.14E-13	7.53E-08	75.3
欠失	n	-1000	3.30E+04	6.87E+10	1.14E-13	3.77E-09	3.8

【 図 8 A 】



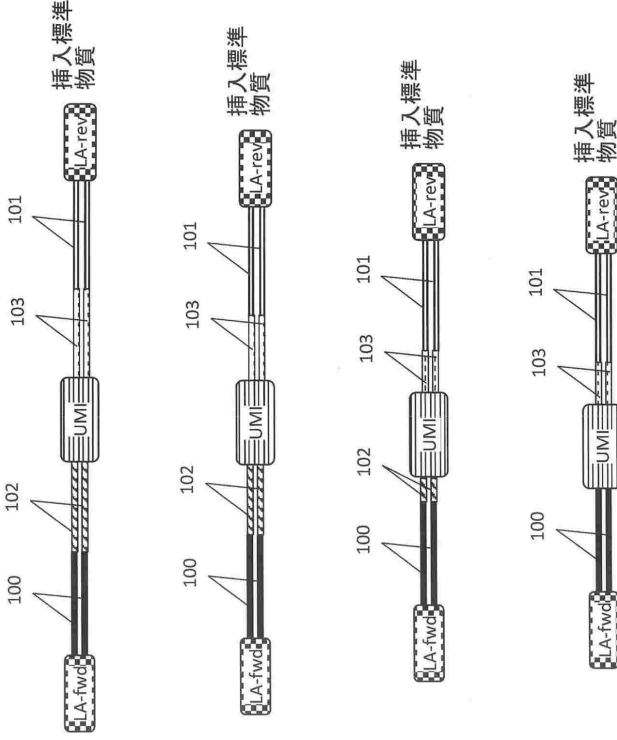
完全長標準物質

30

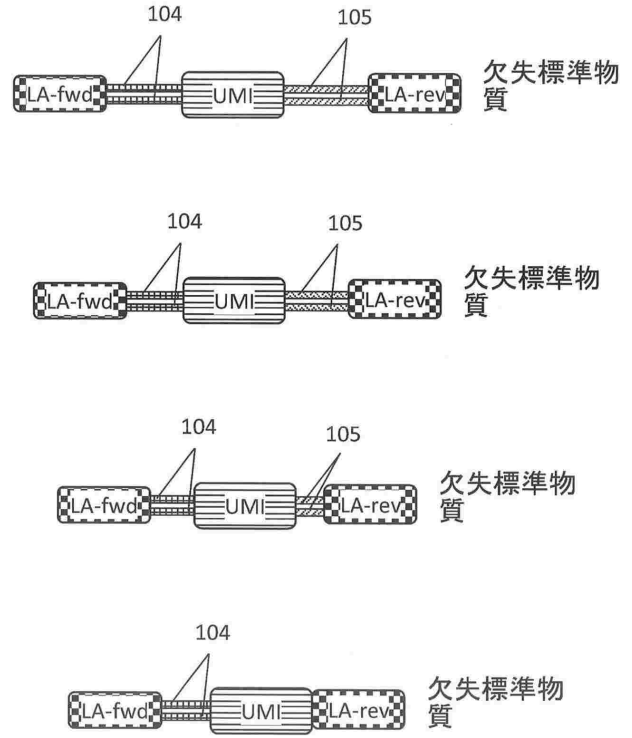
40

50

【 図 8 B 】



【 図 8 C 】

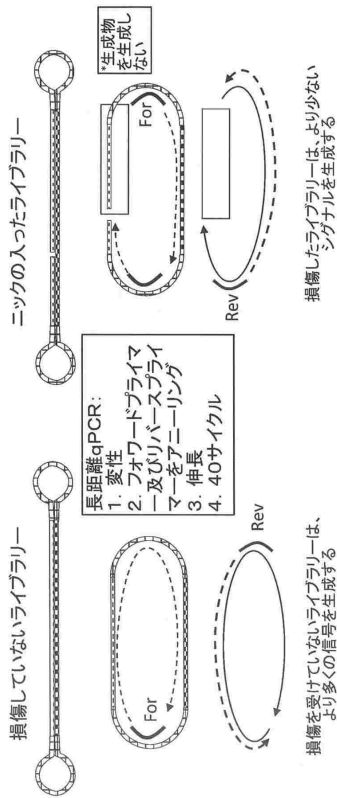


10

20

【 図 9 】

長いライブラリーにおいてDMS損傷を評価するためのqPCR QCアッセイ
 指数関数的増幅アプローチ(PrimeStar GXL DNA ポリメラーゼ; 30kb まで)

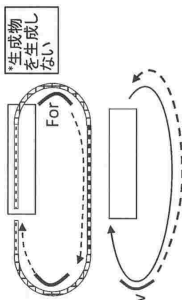


損傷を受けていないライブラリーは、より多くの倍増を生成する
 損傷したライブラリーは、より少ない倍増を生成する

損傷したライブラリーは、より少ないシグナルを生成する

長距離qPCR:

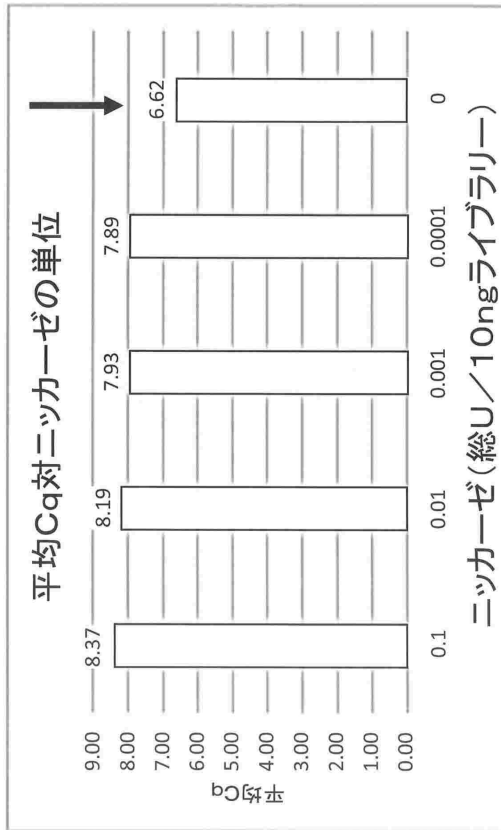
1. 変性
2. フォワードプライマー及びリバースプライマーをアニーリング
3. 伸長
4. 40サイクル



ニックの入ったライブラリー

損傷していないライブラリー

【 図 10 A 】

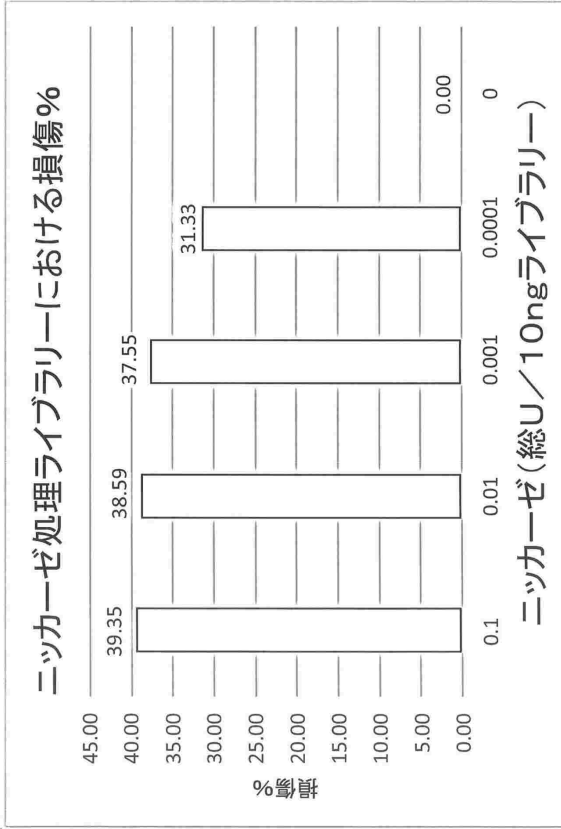


30

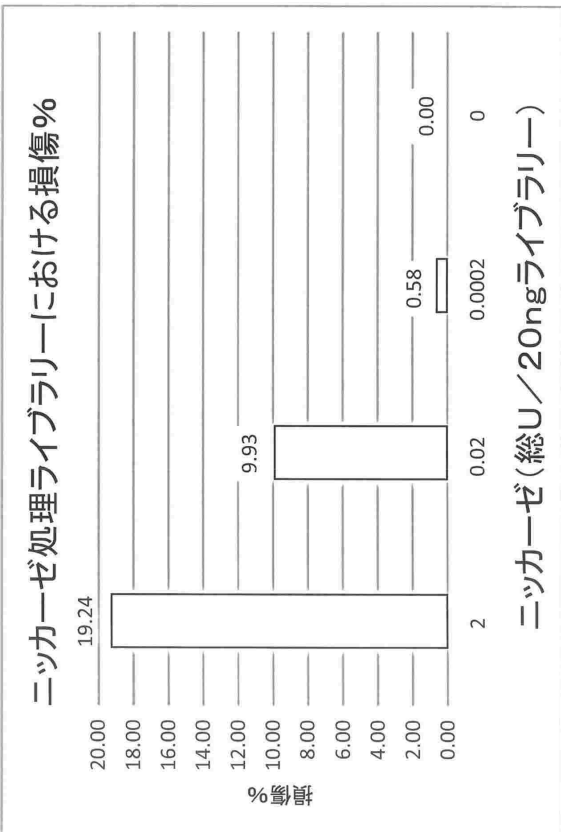
40

50

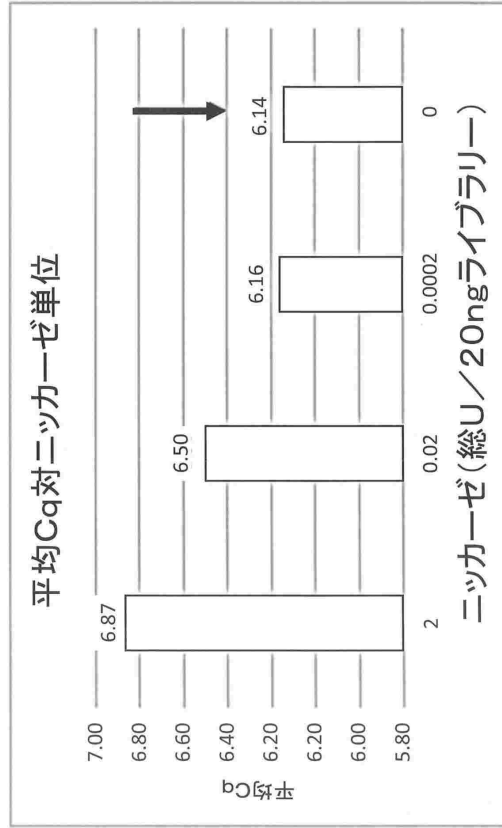
【 図 1 0 B 】



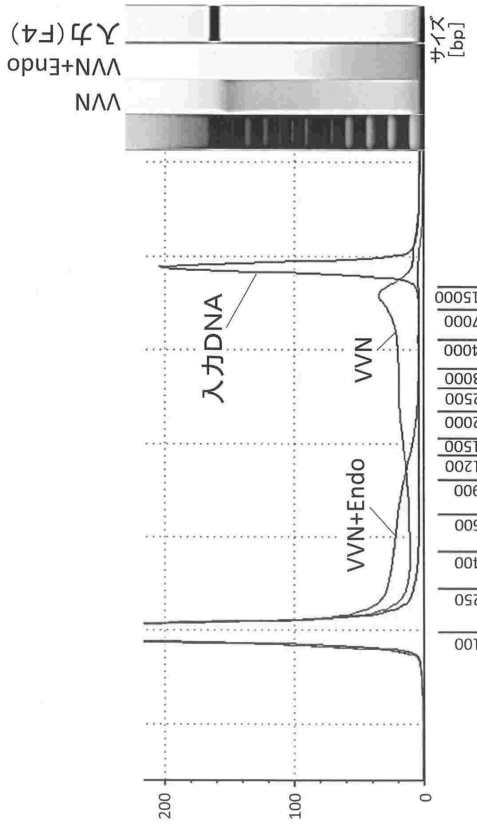
【 図 1 0 D 】



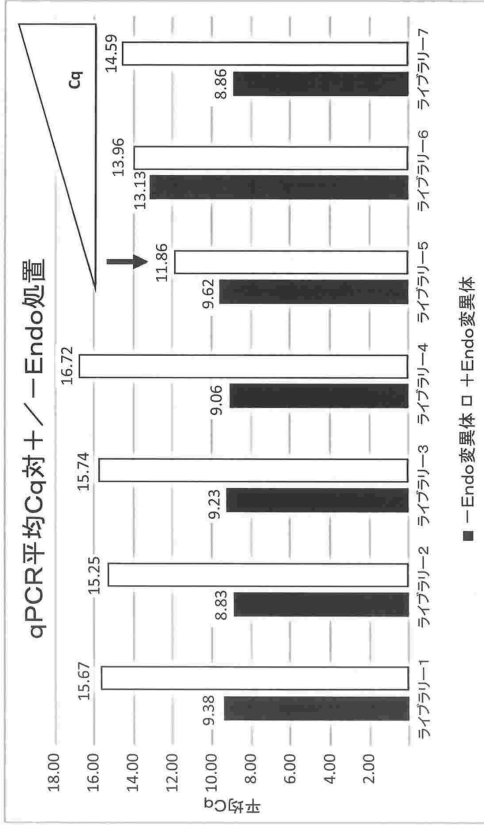
【 図 1 0 C 】



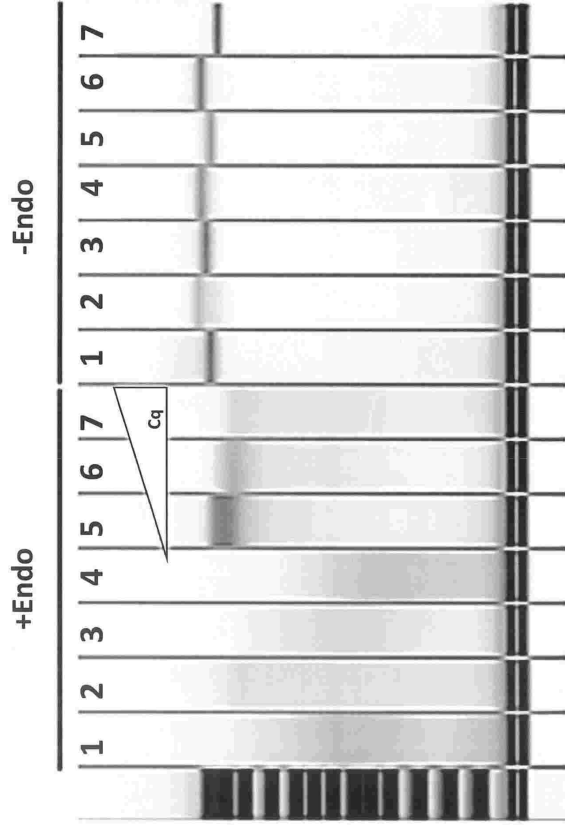
【 図 1 1 】



【 図 1 2 A 】



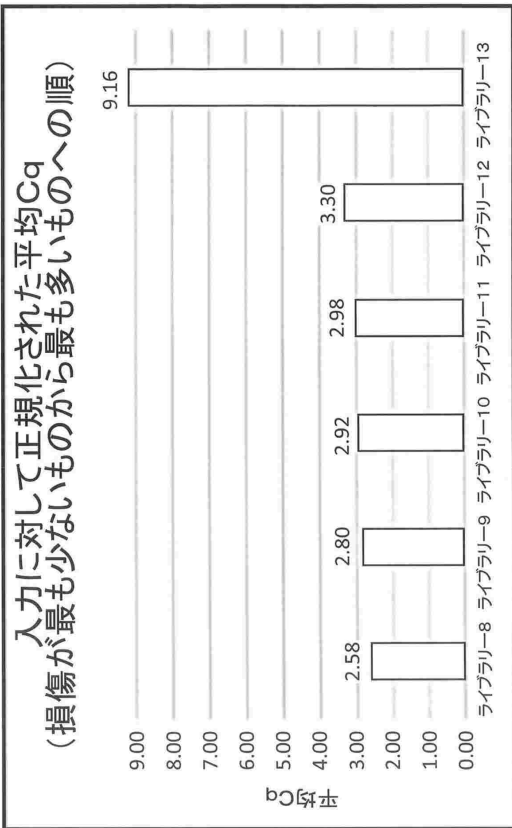
【 図 1 2 B 】



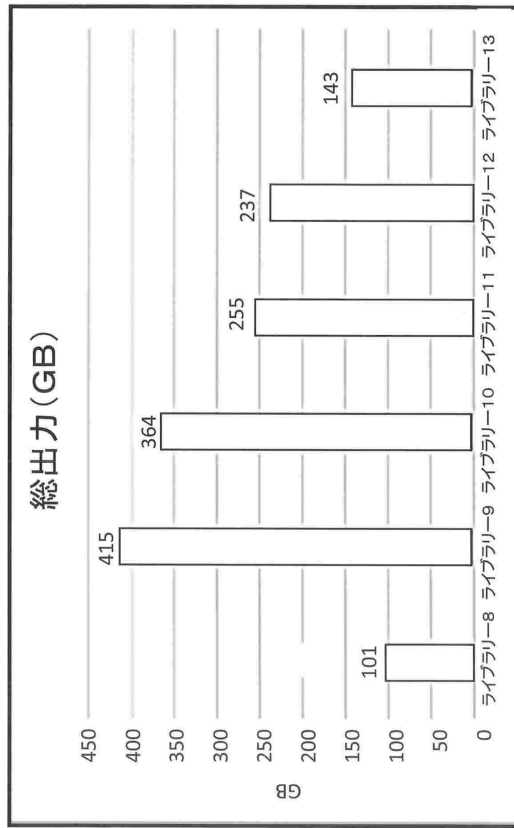
10

20

【 図 1 3 A 】



【 図 1 3 B 】

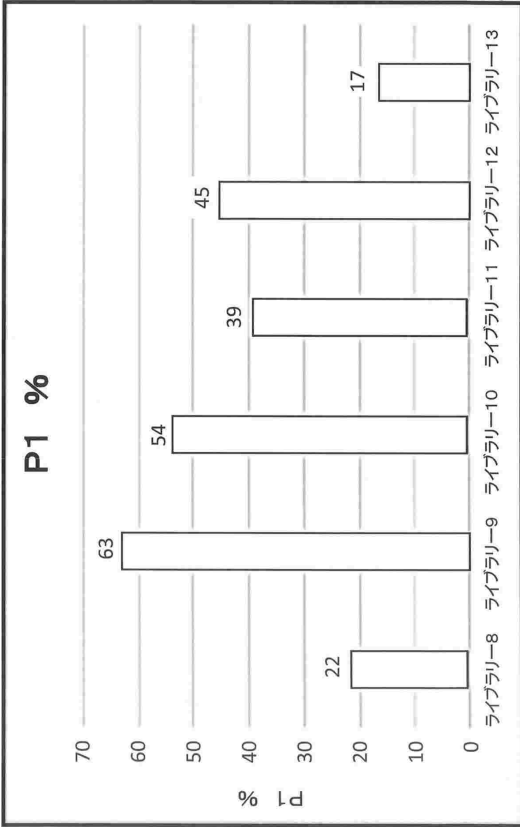


30

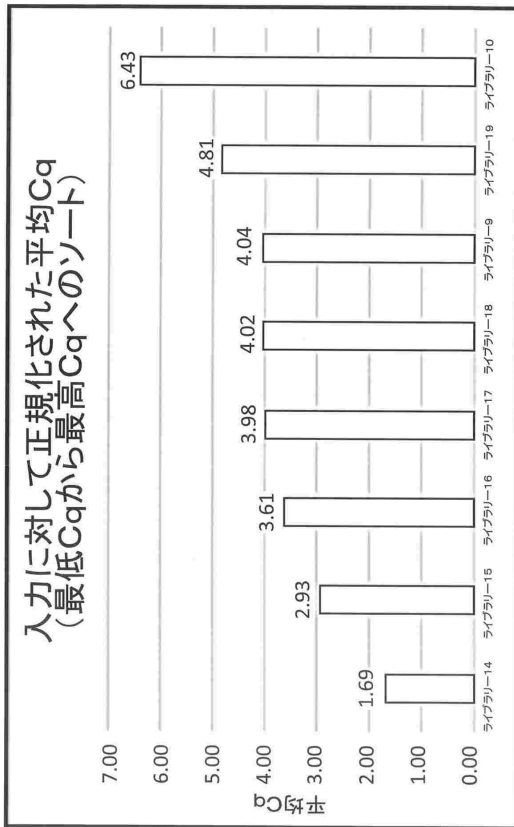
40

50

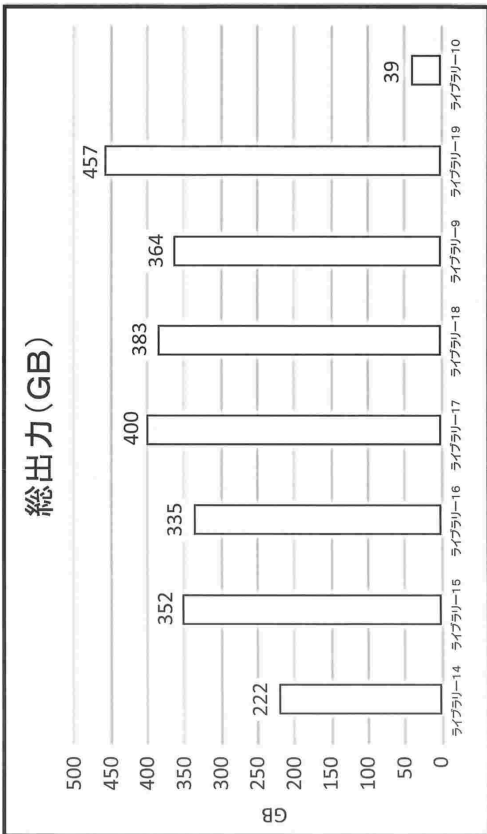
【 図 1 3 C 】



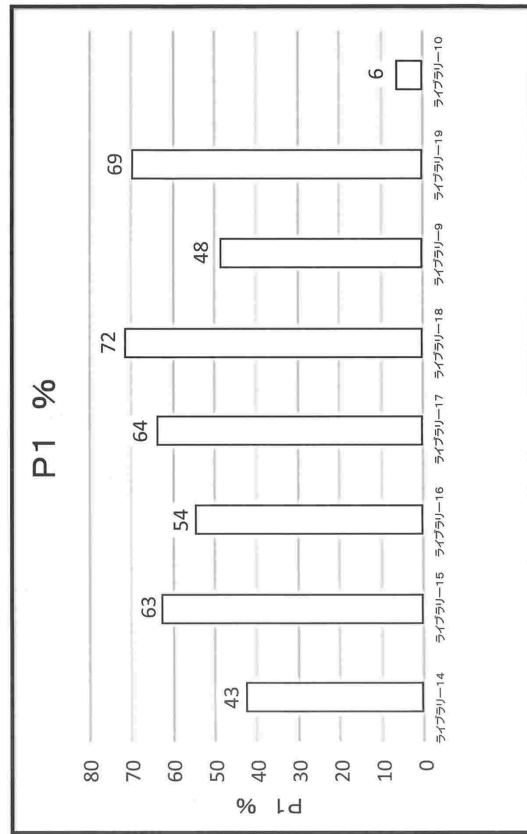
【 図 1 4 A 】



【 図 1 4 B 】



【 図 1 4 C 】



10

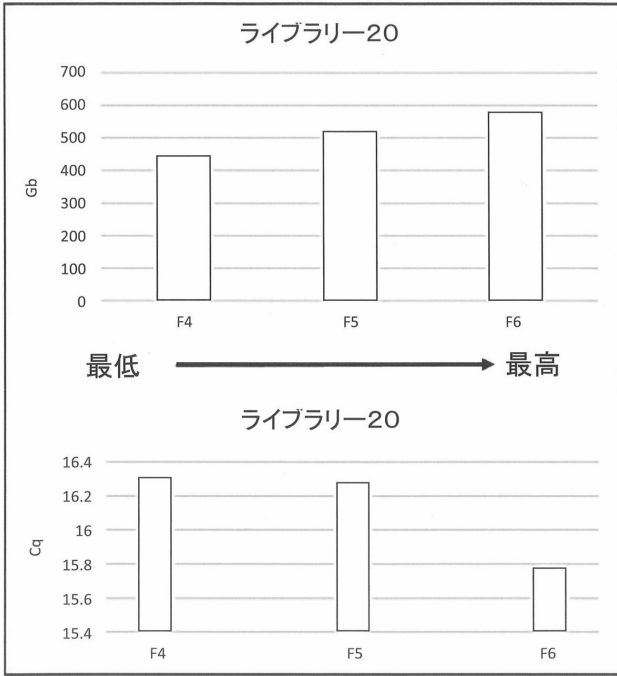
20

30

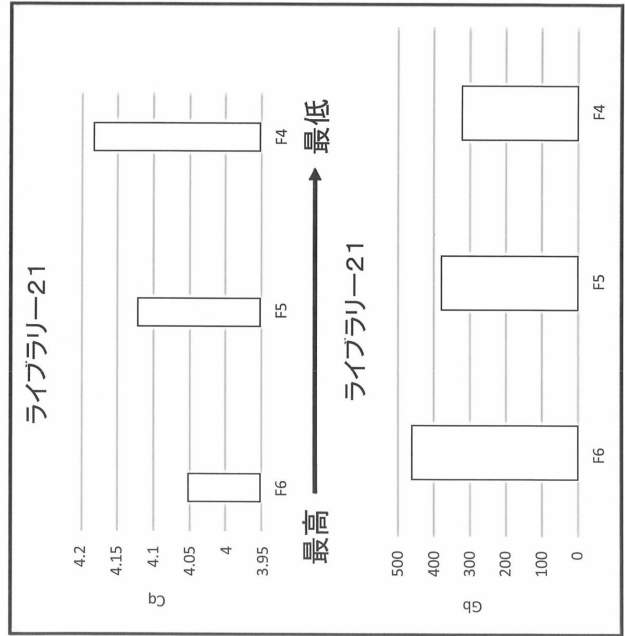
40

50

【 図 1 5 A 】



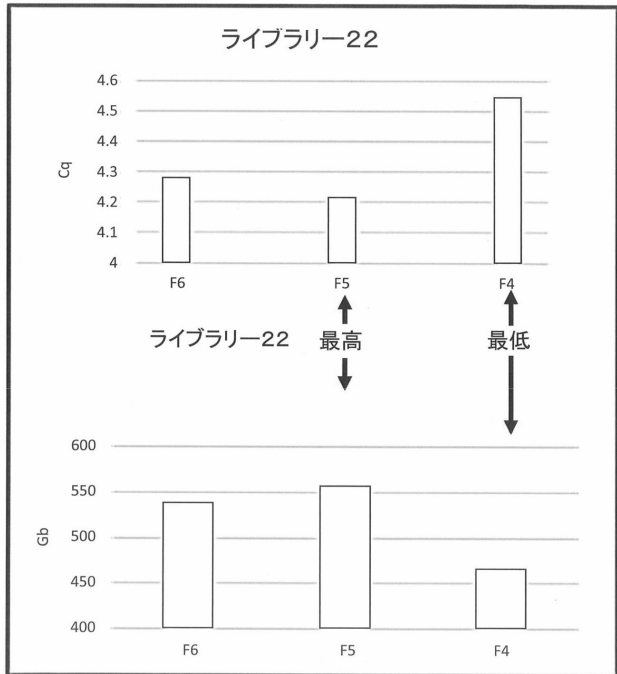
【 図 1 5 B 】



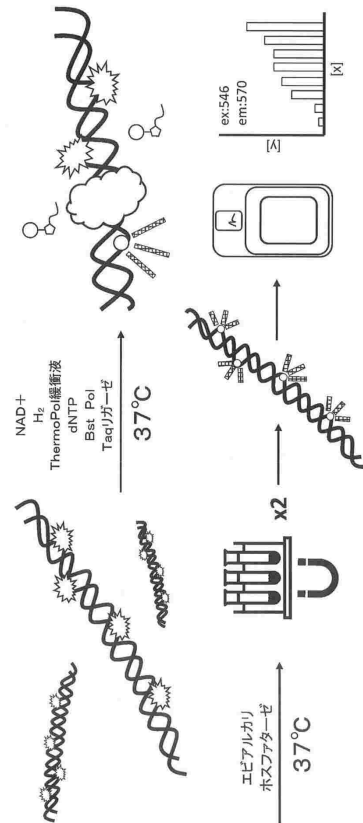
10

20

【 図 1 5 C 】



【 図 1 6 】

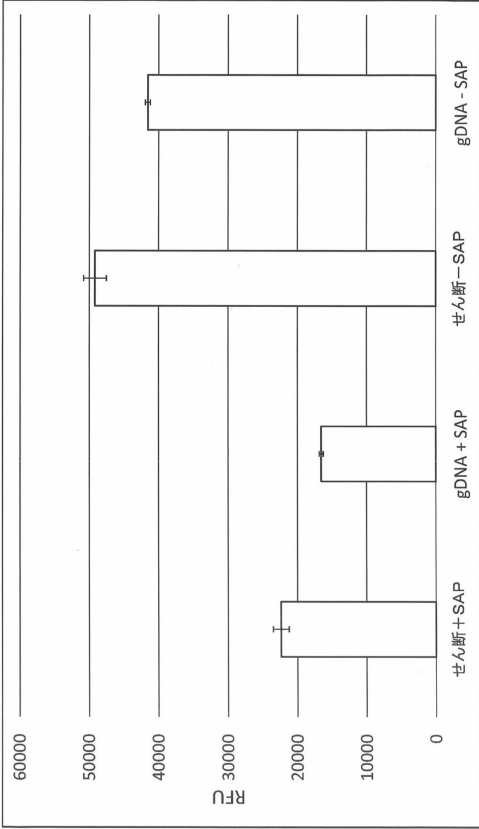


30

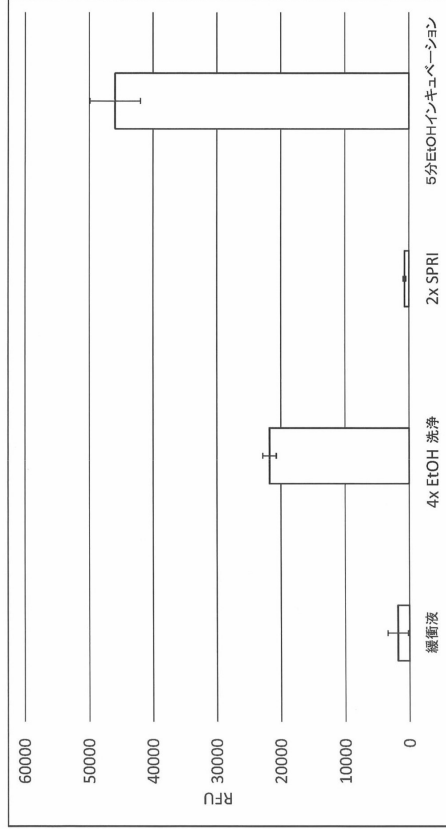
40

50

【 図 1 7 】



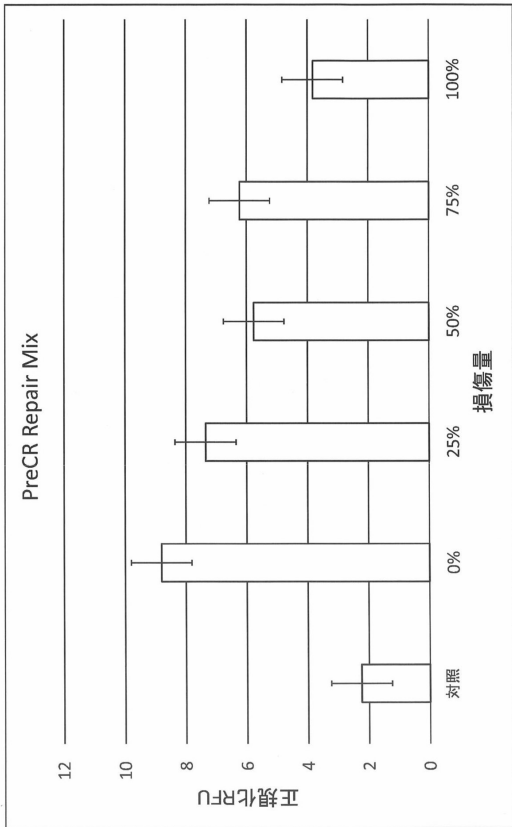
【 図 1 8 】



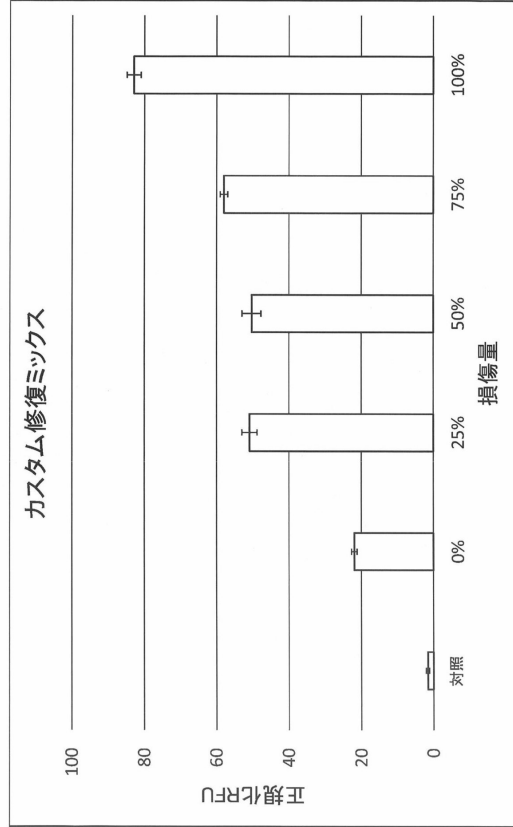
10

20

【 図 1 9 A 】



【 図 1 9 B 】

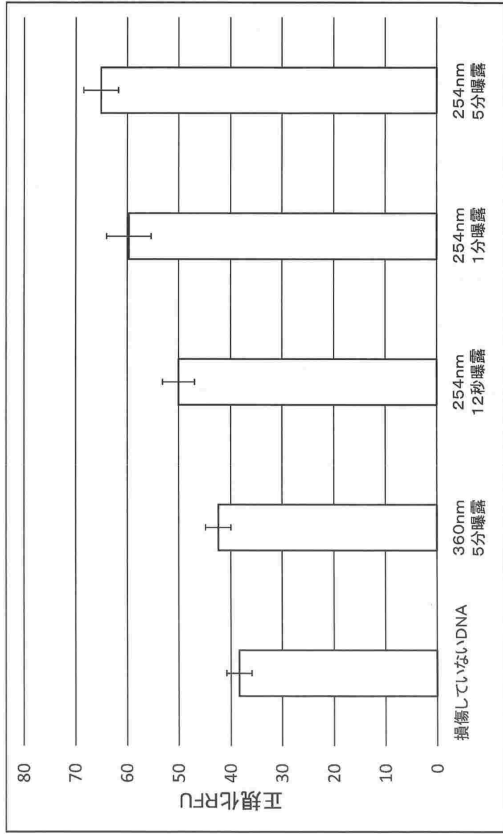


30

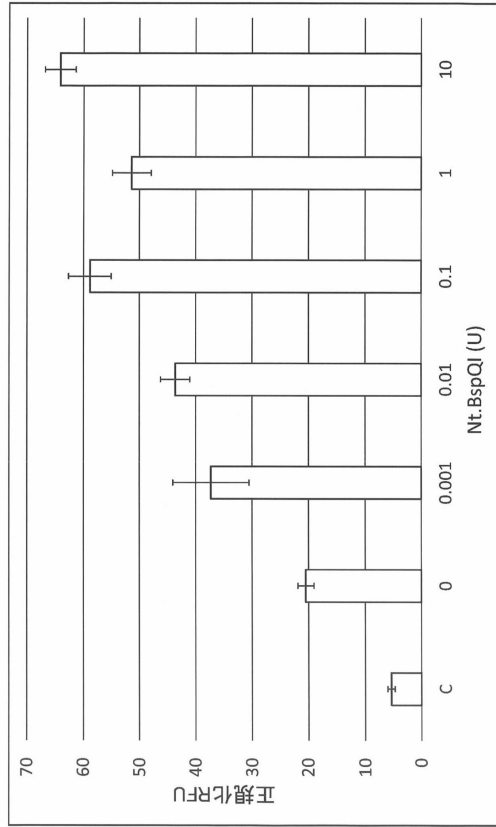
40

50

【 図 2 0 】



【 図 2 1 】



10

20

30

40

50

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2022/022184

A. CLASSIFICATION OF SUBJECT MATTER INV. C12Q1/6806 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) C12Q		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, WPI Data, BIOSIS, EMBASE		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2020/132628 A1 (GUARDANT HEALTH INC [US]) 25 June 2020 (2020-06-25) paragraphs [0138], [0117], [0188], [0210]; claims 1, 5; figure 1A -----	1-3, 6, 11-18
X	WO 2020/072990 A1 (ARC BIO LLC [US]) 9 April 2020 (2020-04-09) paragraph [0496]; claims 1, 76 -----	1-3, 6, 11-18
X	WO 2017/165864 A1 (KARIUS INC [US]) 28 September 2017 (2017-09-28) paragraph [0269]; claim 1; figure 4; example 7 -----	1-3, 6, 11-18
X	WO 2019/002366 A1 (HOFFMANN LA ROCHE [CH]; ROCHE DIAGNOSTICS GMBH [DE] ET AL.) 3 January 2019 (2019-01-03) claims 1, 15; figures 1, 2B -----	1-3, 6, 11-18
	-/-	
<input checked="" type="checkbox"/>	Further documents are listed in the continuation of Box C.	<input checked="" type="checkbox"/> See patent family annex.
* Special categories of cited documents :		
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	
"E" earlier application or patent but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family	
"P" document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search 16 June 2022	Date of mailing of the international search report 16/08/2022	
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer Reuter, Uwe	

1

Form PCT/ISA/210 (second sheet) (April 2005)

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2022/022184

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2018/231818 A1 (LIFE TECHNOLOGIES CORP [US]) 20 December 2018 (2018-12-20) claim 1; examples 2, 7 -----	1-6, 11-18

10

20

30

40

1

50

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2022/022184

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

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

10

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

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

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

20

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

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

see additional sheet

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

30

2. As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims;; it is covered by claims Nos.:

40

1-6 (completely) ; 11-18 (partially)

Remark on Protest

The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

No protest accompanied the payment of additional search fees.

International Application No. PCT/US2022/022184

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-6(completely); 11-18(partially)

A pool of nucleic acid standards of different lengths, wherein the nucleic acid standards comprise a unique molecular identifier (UMI) and: a. a 5' universal oligonucleotide, wherein the 5' universal oligonucleotide is the same for all standards;b. a 3' universal oligonucleotide, wherein the 3' universal oligonucleotide is the same for all standards; and c. at least one region between the UMI and the 5' universal oligonucleotide and/or between the UMI and the 3' universal oligonucleotide;wherein the length of the at least one region determines the length of the standard.

2. claims: 7-10(completely); 11-18(partially)

A pool of nucleic acid standards of different lengths, wherein the nucleic acid standards comprise a UMI and: a 5' partially overlapping oligonucleotide, wherein the 5' partially overlapping oligonucleotide is identical over at least a portion of its sequence for all the standards; and/orb. a 3' partially overlapping oligonucleotide, wherein the 3' partially overlapping oligonucleotide is identical over at least a portion of its sequence for all the standards;wherein the lengths of the 5' partially overlapping oligonucleotide and/or the 3' partially overlapping oligonucleotide determines the length of the standard.

3. claims: 19-30

A method of generating a pool of nucleic acid standards comprising: a. providing multiple copies of at least one sequence of interest comprising nucleic acids;b. providing a collection of oligonucleotides each comprising a UMI; and c. ligating at least one sequence of interest of (a) and at least one oligonucleotide comprising a UMI of (b) optionally the method including insertion oligonucleotides of varying lengths.

4. claims: 31-47

A method of normalizing amplicon size bias comprising:a. combining a sample comprising a target nucleic acid with a pool of nucleic acid standards of different lengths, wherein each standard comprises a UMI;b. amplifying the standards and amplicons of a sequence of interest comprised in the target nucleic acid;c. sequencing the standards and the amplicons of the sequence of interest to generate sequencing

10

20

30

40

50

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

data;d. determining a bias profile based on amplicon size using sequencing data from the standards; and e. normalizing amplicon size bias using the bias profile.

5. claims: 48-77

A method of determining the presence of DNA damage in a library comprising one or more library molecule, wherein each library molecule comprises a double-stranded DNA insert with a hairpin adapter at each end of the insert, comprising: a. denaturing the first stand and second strand of the double-stranded DNA inserts comprised in library molecules;b. annealing a forward primer and a reverse primer to library molecules;c. amplifying to produce library amplicons; and d. assessing the presence of DNA damage based on the number of library amplicons produced.

6. claims: 78-113

A method of quantifying DNA damage in a sample comprising DNA using fluorescence comprising: a. combining: i. an aliquot of a sample comprising DNA,ii. one or more DNA repair enzyme; and iii. dNTPs, wherein one or more dNTP is fluorescently labeled;b. preparing repaired DNA;c. dephosphorylating the phosphates from dNTPs; d. binding the repaired DNA to carboxylate or cellulose beads;e. eluting the bound repaired DNA from the carboxylate or cellulose beads with a resuspension buffer; and f. measuring fluorescence of the repaired DNA to determine the amount of DNA damage.

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2022/022184

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2020132628 A1	25-06-2020	CA 3119980 A1	25-06-2020
		CN 113454218 A	28-09-2021
		EP 3898969 A1	27-10-2021
		JP 2022514010 A	09-02-2022
		SG 11202104701X A	29-06-2021
		US 2020232010 A1	23-07-2020
		WO 2020132628 A1	25-06-2020
WO 2020072990 A1	09-04-2020	AU 2019356016 A1	13-05-2021
		CA 3115281 A1	09-04-2020
		CN 113207299 A	03-08-2021
		EP 3861135 A1	11-08-2021
		IL 281982 A	31-05-2021
		JP 2022513343 A	07-02-2022
		US 2022002781 A1	06-01-2022
		WO 2020072990 A1	09-04-2020
WO 2017165864 A1	28-09-2017	AU 2017237199 A1	16-08-2018
		AU 2021200637 A1	04-03-2021
		AU 2021215106 A1	26-08-2021
		BR 112018069557 A2	29-01-2019
		CA 3014070 A1	28-09-2017
		CN 109312400 A	05-02-2019
		EP 3433382 A1	30-01-2019
		EP 3978627 A1	06-04-2022
		IL 278000 A	30-11-2020
		JP 6743268 B2	19-08-2020
		JP 7091397 B2	27-06-2022
		JP 2019509068 A	04-04-2019
		JP 2020182491 A	12-11-2020
		KR 20180129844 A	05-12-2018
		KR 20200128194 A	11-11-2020
		KR 20210138154 A	18-11-2021
		US 2017275691 A1	28-09-2017
		US 2018237851 A1	23-08-2018
		US 2021324467 A1	21-10-2021
		WO 2017165864 A1	28-09-2017
WO 2019002366 A1	03-01-2019	CN 110785493 A	11-02-2020
		DK 3645717 T3	08-11-2021
		EP 3645717 A1	06-05-2020
		ES 2898644 T3	08-03-2022
		JP 7030857 B2	07-03-2022
		JP 2020529833 A	15-10-2020
		LT 3645717 T	10-12-2021
		US 2020109397 A1	09-04-2020
		WO 2019002366 A1	03-01-2019
		WO 2018231818 A1	20-12-2018
US 2020131563 A1	30-04-2020		
WO 2018231818 A1	20-12-2018		

Form PCT/ISA/210 (patent family annex) (April 2005)

10

20

30

40

50

フロントページの続き

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

(74)代理人 100111796

弁理士 服部 博信

(74)代理人 100123766

弁理士 松田 七重

(74)代理人 100111501

弁理士 滝澤 敏雄

(72)発明者 ケネディ アンドリュー ビー

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 イルミ
ナ インコーポレイテッド内

(72)発明者 ストームス レナ

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 イルミ
ナ インコーポレイテッド内

(72)発明者 シェン フェイ

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 イルミ
ナ インコーポレイテッド内

(72)発明者 ベニス オリヴィア

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 イルミ
ナ インコーポレイテッド内

(72)発明者 ムルトフェルト エリック

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 イルミ
ナ インコーポレイテッド内

(72)発明者 ブグリエーセ ケイトリン

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 イルミ
ナ インコーポレイテッド内

(72)発明者 ハワード マイケル

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 イルミ
ナ インコーポレイテッド内

F ターム (参考) 4B063 QA13 QQ28 QQ42 QQ52 QR08 QR32 QR35 QR62 QS25 QS36

QX02