

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号
特表2025-500815
(P2025-500815A)

(43)公表日 令和7年1月15日(2025.1.15)

(51)国際特許分類		F I			テーマコード (参考)
C 1 2 Q	1/6869(2018.01)	C 1 2 Q	1/6869	Z	4 B 0 2 9
C 1 2 M	1/00 (2006.01)	C 1 2 M	1/00	A	4 B 0 6 3
C 1 2 M	1/34 (2006.01)	C 1 2 M	1/34	Z	

審査請求 未請求 予備審査請求 未請求 (全151頁)			
(21)出願番号	特願2024-534541(P2024-534541)	(71)出願人	523473143
(86)(22)出願日	令和4年12月9日(2022.12.9)		エレメント バイオサイエンシズ, イン
(85)翻訳文提出日	令和6年8月2日(2024.8.2)		コーポレイティド
(86)国際出願番号	PCT/US2022/052448		アメリカ合衆国, カリフォルニア 9 2
(87)国際公開番号	WO2023/107719		1 2 1, サンディエゴ, バーンズ キャ
(87)国際公開日	令和5年6月15日(2023.6.15)		ニオン ロード 1 0 0 5 5
(31)優先権主張番号	17/547,602	(74)代理人	100145403
(32)優先日	令和3年12月10日(2021.12.10)		弁理士 山尾 憲人
(33)優先権主張国・地域又は機関	米国(US)	(74)代理人	100106518
(31)優先権主張番号	63/349,421		弁理士 松谷 道子
(32)優先日	令和4年6月6日(2022.6.6)	(74)代理人	100138911
(33)優先権主張国・地域又は機関	米国(US)		弁理士 櫻井 陽子
(31)優先権主張番号	17/854,042	(72)発明者	トンプソン, コナー
	最終頁に続く		アメリカ合衆国 9 2 1 2 1 カリフォルニ
			ア州サンディエゴ, バーンズ・キャニオ
			最終頁に続く

(54)【発明の名称】 次世代シーケンシングにおける一次分析

(57)【要約】

画像データ分析、及び特に、DNAシーケンシング中にフローセルのデジタル画像内でベースコーリングを実行するためのクラスタ又はポロニー位置を識別することが記載される。方法は、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することを含み得る。細胞試料は、その中に複数のコンカテマー分子を含み得る。第1の複数のフローセル画像について、ピクセル強度、及びピクセル強度の各々のそれぞれの色純度が判定され得る。ベースコーリングテンプレートは、ピクセル強度及びピクセル強度のそれぞれの色純度に基づいたベースコーリング位置を含み得る。ベースコーリングテンプレートは、1つ以上のサイクルの1つ以上の後続のサイクルにおける支持体の第2の複数のフローセル画像を位置合わせするためのものであってもよい。

【選択図】 図 5

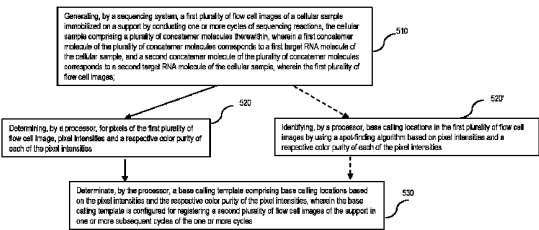


FIG. 5

【特許請求の範囲】

【請求項 1】

方法であって、

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標的RNA分子に対応する、生成することと、

プロセッサによって、前記第1の複数のフローセル画像のピクセルについて、ピクセル強度及び前記ピクセル強度の各々のそれぞれの色純度を判定することと、

前記プロセッサによって、前記ピクセル強度及び前記ピクセル強度の前記それぞれの色純度に基づいて、ベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリングテンプレートが、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を位置合わせするように構成される、判定することと、を含む、方法。

【請求項 2】

方法であって、

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標的RNA分子に対応する、生成することと、

プロセッサによって、前記第1の複数のフローセル画像のピクセルについて、ボクセル強度及びピクセル強度の各々のそれぞれの色純度を判定することと、

前記プロセッサによって、並びに前記ピクセル強度及び前記ピクセル強度の前記それぞれの色純度に基づいて、軸方向軸に沿った異なる軸方向位置を有するベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリングテンプレートが、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を位置合わせするように構成される、判定することと、を含む、方法。

【請求項 3】

方法であって、

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標的RNA分子に対応する、生成することと、

プロセッサによって、前記第1の複数のフローセル画像のピクセルについて、ボクセル強度及びピクセル強度の各々のそれぞれの色純度を判定することと、

前記プロセッサによって、並びに前記ピクセル強度及び前記ピクセル強度の前記それぞれの色純度に基づいて、ベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリングテンプレートが、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を位置合わせするように構成され、前記第1又は第2の複数のフローセル画像が、軸方向軸に沿った異なる軸方向位置で生成される、判定することと、を含む、方法。

【請求項 4】

方法であって、

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標的RNA分子に対応する、生成することと、

プロセッサによって、ピクセル強度及び前記ピクセル強度の各々のそれぞれの色純度に基づいたスポット探索アルゴリズムを使用することによって、前記第1の複数のフローセル画像内のベースコーリング位置を識別することと、

前記プロセッサによって、前記識別されたベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリングテンプレートが、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を位置合わせするように構成され、前記ベースコーリング位置が、軸方向軸に沿った異なる軸方向位置を含む、判定することと、を含む、方法。

10

【請求項5】

方法であって、

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標的RNA分子に対応する、生成することと、

20

プロセッサによって、ピクセル強度及び前記ピクセル強度の各々のそれぞれの色純度に基づいたスポット探索アルゴリズムを使用することによって、前記第1の複数のフローセル画像内のベースコーリング位置を識別することと、

前記プロセッサによって、前記識別されたベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリングテンプレートが、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を位置合わせするように構成され、前記ベースコーリング位置が、軸方向軸に沿った異なる軸方向位置を含む、判定することと、を含む、方法。

30

【請求項6】

方法であって、

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標的RNA分子に対応する、生成することと、

プロセッサによって、ピクセル強度及び前記ピクセル強度の各々のそれぞれの色純度に基づいたスポット探索アルゴリズムを使用することによって、前記第1の複数のフローセル画像内のベースコーリング位置を識別することと、

40

前記プロセッサによって、並びに前記ピクセル強度及び前記ピクセル強度の前記それぞれの色純度に基づいて、ベースコーリング位置を含むベースコーリングテンプレートを判定することと、

前記プロセッサによって、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を前記ベースコーリングテンプレートに対して位置合わせすることであって、前記第1又は第2の複数のフローセル画像が、軸方向軸に沿った異なる軸方向位置で生成される、位置合わせすることと、を含む、方法。

【請求項7】

方法であって、

50

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標的RNA分子に対応する、生成することと、

プロセッサによって、ピクセル強度及び前記ピクセル強度の各々のそれぞれの色純度に基づいたスポット探索アルゴリズムを使用することによって、前記第1の複数のフローセル画像内のベースコーリング位置を識別することと、

前記プロセッサによって、前記識別されたベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリング位置が、軸方向軸に沿った異なる軸方向位置を含む、判定することと、

前記プロセッサによって、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を前記ベースコーリングテンプレートに対して位置合わせすることと、

前記プロセッサによって、前記ベースコーリングテンプレートの前記ベースコーリング位置に基づいて、前記第2の複数のフローセル画像のベースコーリングを実行することとを含む、方法。

【請求項8】

前記ピクセルが、二次元である、先行請求項のいずれか一項に記載の方法。

【請求項9】

前記ピクセルが、画像平面内にある、先行請求項のいずれか一項に記載の方法。

【請求項10】

前記ピクセルが、三次元である、先行請求項のいずれか一項に記載の方法。

【請求項11】

前記ピクセルが、前記画像平面内の2つの次元と、前記軸方向軸と平行な第3の次元とを有する、先行請求項のいずれか一項に記載の方法。

【請求項12】

前記第1又は第2の複数のフローセル画像が、複数の所定の軸方向位置で取得される、先行請求項のいずれか一項に記載の方法。

【請求項13】

前記ベースコーリング位置の前記異なる軸方向位置のうちの少なくともいくつか、前記複数の所定の軸方向位置のうちのいずれとも異なる、先行請求項のいずれか一項に記載の方法。

【請求項14】

前記ベースコーリング位置の前記異なる軸方向位置のうちの少なくともいくつか、前記複数の所定の軸方向位置の2つの隣接する所定の軸方向位置の間にある、先行請求項のいずれか一項に記載の方法。

【請求項15】

前記支持体上に固定化された前記複数のコンカテマー分子の各々が、ポロニーに対応する、先行請求項のいずれか一項に記載の方法。

【請求項16】

前記支持体上に固定化された前記複数のコンカテマー分子の各々が、ベースコーリング位置に対応する、先行請求項のいずれか一項に記載の方法。

【請求項17】

前記複数の所定の軸方向位置が、3～500個の所定の軸方向位置を含む、先行請求項のいずれか一項に記載の方法。

【請求項18】

前記複数の所定の軸方向位置の各々が、その隣接する隣接軸方向位置から1、2、3、4、5、6、7、8、9、10、11、12、13、14、又は15um離隔している、

10

20

30

40

50

先行請求項のいずれか一項に記載の方法。

【請求項 19】

前記複数の所定の軸方向位置の各々が、その隣接する隣接軸方向位置から 0.1 ~ 400 μm 離隔している、先行請求項のいずれか一項に記載の方法。

【請求項 20】

前記 1 つ以上のサイクルが、1、2、3、4、5、6、7、8、9、又は 10 サイクルを含む、先行請求項のいずれか一項に記載の方法。

【請求項 21】

前記 1 つ以上のサイクルが、1 ~ 50 サイクルを含む、先行請求項のいずれか一項に記載の方法。

【請求項 22】

前記細胞試料が、1 つ以上のインサイチュ試料を含む、先行請求項のいずれか一項に記載の方法。

【請求項 23】

前記細胞試料が、1 つ以上の細胞又は組織を含む、先行請求項のいずれか一項に記載の方法。

【請求項 24】

前記第 1 又は第 2 の複数のフローセル画像の各々が、前記軸方向軸に直交する視野を含む、先行請求項のいずれか一項に記載の方法。

【請求項 25】

前記第 1 又は第 2 の複数のフローセル画像の各々の前記視野が、画像平面内で同一である、先行請求項のいずれか一項に記載の方法。

【請求項 26】

前記第 1 又は第 2 の複数のフローセル画像の各々の前記視野が、フローセルのタイルの少なくとも一部分を覆う、先行請求項のいずれか一項に記載の方法。

【請求項 27】

前記第 1 又は第 2 の複数のフローセル画像が、同一の画像解像度を含む、先行請求項のいずれか一項に記載の方法。

【請求項 28】

前記軸方向軸が、対物レンズから前記支持体まで延在する、先行請求項のいずれか一項に記載の方法。

【請求項 29】

前記軸方向軸が、画像平面に直交し、前記視野が、前記画像平面内にある、先行請求項のいずれか一項に記載の方法。

【請求項 30】

前記第 1 又は第 2 の複数のフローセル画像が、2、3、4、5、又は 6 つの異なる色チャンネルからのものである、先行請求項のいずれか一項に記載の方法。

【請求項 31】

前記第 1 又は第 2 の複数のフローセル画像が、各サイクルにおいて前記支持体上に固定化された複数の核酸テンプレート分子間のバランスのとれた多様性の A、G、C、及び T / U のヌクレオチド塩基を含む、先行請求項のいずれか一項に記載の方法。

【請求項 32】

前記支持体が、ガラス又はプラスチック基材を含む、先行請求項のいずれか一項に記載の方法。

【請求項 33】

前記支持体が、フローセルデバイスに含まれている、先行請求項のいずれか一項に記載の方法。

【請求項 34】

前記支持体上の前記複数のコンカテマー分子の密度が、 1 mm^2 当たり $10^1 \sim 10^{12}$ である、先行請求項のいずれか一項に記載の方法。

10

20

30

40

50

【請求項 35】

前記支持体上の前記支持体上の前記複数のコンカテマー分子の密度密度が、 1 mm^2 当たり $10^2 \sim 10^8$ である、先行請求項のいずれか一項に記載の方法。

【請求項 36】

前記複数のコンカテマー分子のうちの2つ以上の異なるコンカテマー分子が、異なる挿入配列を有する、先行請求項のいずれか一項に記載の方法。

【請求項 37】

異なる挿入配列が、異なる標的RNA分子又は標的cDNA分子に対応する、先行請求項のいずれか一項に記載の方法。

【請求項 38】

前記ベースコーリング位置の各々が、前記複数のコンカテマー分子の位置に対応する、先行請求項のいずれか一項に記載の方法。

【請求項 39】

前記第1の複数のフローセル画像が、前記支持体上に固定化された前記複数のコンカテマー分子間のバランスのとれた多様性のA、G、C、及びT/Uのヌクレオチド塩基に結合したヌクレオチド試薬から放出される光シグナルを含む、先行請求項のいずれか一項に記載の方法。

【請求項 40】

前記第2の複数のフローセル画像が、前記1つ以上の後続のサイクルにおいて前記支持体上に固定化された前記複数のコンカテマー分子間のアンバランスな多様性のA、G、C、及びT/Uのヌクレオチド塩基に結合したヌクレオチド試薬から放出される光シグナルを含む、先行請求項のいずれか一項に記載の方法。

【請求項 41】

前記複数のコンカテマー分子間の前記アンバランスな多様性のA、G、C、及びT/Uのヌクレオチド塩基が、(2)塩基の総数に対する(1)1つ以上の種類のヌクレオチド塩基の数のパーセンテージが、前記1つ以上のサイクルにおいて20%、15%、10%、又は5%未満であることを含む、先行請求項のいずれか一項に記載の方法。

【請求項 42】

前記複数のコンカテマー分子間の前記バランスのとれた多様性のA、G、C、及びT/Uのヌクレオチド塩基が、前記1つ以上のサイクルにおける(2)塩基の総数に対する(1)各種類のヌクレオチド塩基の数のパーセンテージが、10%、15%、又は20%を超えることを含む、先行請求項のいずれか一項に記載の方法。

【請求項 43】

前記シーケンシングシステムによって、1つ以上の後続のシーケンシング反応サイクルを行い、それによって、前記第2の複数のフローセル画像を生成することを更に含む、先行請求項のいずれか一項に記載の方法。

【請求項 44】

前記プロセッサによって、1つ以上の後続のシーケンシングサイクルからの前記第2の複数のフローセル画像を前記ベースコーリングテンプレートに対して位置合わせすること、

前記プロセッサによって、位置合わせされた前記第2の複数のフローセル画像からのシグナルを使用して、前記ベースコーリングテンプレート内の前記ベースコーリング位置において前記第2の複数のフローセル画像のベースコーリングを実行すること、を更に含む、先行請求項のいずれか一項に記載の方法。

【請求項 45】

前記プロセッサによって、前記第2の複数のフローセル画像からのシグナルを使用して、前記ベースコーリングテンプレート内の前記ベースコーリング位置においてのみベースコーリングを実行することを更に含む、先行請求項のいずれか一項に記載の方法。

【請求項 46】

前記ベースコーリング位置が、サブピクセル解像度である、先行請求項のいずれか一項

10

20

30

40

50

に記載の方法。

【請求項 47】

各ピクセル強度が、前記対応するピクセルの 1 つ以上のサブピクセル強度を含む、先行請求項のいずれか一項に記載の方法。

【請求項 48】

前記ピクセル強度の各々の前記それぞれの色純度が、前記対応するピクセルの前記 1 つ以上のサブピクセル強度に対応する 1 つ以上の色純度を含む、先行請求項のいずれか一項に記載の方法。

【請求項 49】

前記ピクセル強度の各々の前記それぞれの色純度が、1 つ以上の色チャンネルについての前記それぞれの色純度を含む、先行請求項のいずれか一項に記載の方法。

10

【請求項 50】

前記ピクセル強度の各々の前記それぞれの色純度を判定することが、(2) 他の種類のヌクレオチド塩基についてのシグナルの総量に対する (1) 特定の種類のヌクレオチド塩基に対応するシグナルの比率を判定することを含む、先行請求項のいずれか一項に記載の方法。

【請求項 51】

前記第 1 の複数のフローセル画像の前記ピクセルについて、前記ピクセル強度を判定することが、

前記対応するピクセル又はサブピクセル位置でのチャンネル強度のセットの比較に基づいて、それぞれの異なる蛍光波長に対応するチャンネル強度の前記セット内の各チャンネル強度を判定することを含む、先行請求項のいずれか一項に記載の方法。

20

【請求項 52】

前記 1 つ以上のサブピクセル強度が、2、3、4、5、6、7、8、9、10、12、13、14、15、16、17、18、19、又は 20 の異なる軸方向位置を有し、前記異なる軸方向位置が、前記軸方向軸に沿っており、前記複数の所定の軸方向位置のうちの 2 つの隣接する軸方向位置の間にある、先行請求項のいずれか一項に記載の方法。

【請求項 53】

前記 1 つ以上のサブピクセル強度が、1～80 の異なる軸方向位置を有し、前記異なる軸方向位置が、前記軸方向軸に沿っており、前記複数の所定の軸方向位置のうちの 2 つの隣接する軸方向位置の間にある、先行請求項のいずれか一項に記載の方法。

30

【請求項 54】

前記 1 つ以上の後続のシーケンシングサイクルからの前記第 2 の複数のフローセル画像を前記ベースコーリングテンプレートに対して位置合わせすることが、

前記ベースコーリングテンプレートとして、共通座標系における前記第 2 の複数のフローセル画像におけるポロニーの座標を生成することを含む、先行請求項のいずれか一項に記載の方法。

【請求項 55】

前記第 2 の複数のフローセル画像における前記ポロニーの前記座標が、三次元である、先行請求項のいずれか一項に記載の方法。

40

【請求項 56】

前記共通座標系が、三次元である、先行請求項のいずれか一項に記載の方法。

【請求項 57】

前記ピクセル強度及び前記ピクセル強度の前記それぞれの色純度に基づいて、ベースコーリングテンプレートを判定することが、

前記ピクセル又はサブピクセルの各々について、前記それぞれの色純度が、閾値距離内の他のピクセル又はサブピクセルの色純度よりも大きいと判定することと、

前記それぞれの色純度が、前記閾値距離内の他のピクセル又はサブピクセルの前記色純度よりも大きいとの判定に回答して、前記対応するピクセル又はサブピクセルの位置をベースコーリングテンプレートに追加することと、

50

前記それぞれの色純度が、前記閾値距離内の前記他のピクセル又はサブピクセルの前記色純度よりも大きくないとの判定に应答して、前記ベースコーリングテンプレートに変更を加えないことと、を含む、先行請求項のいずれか一項に記載の方法。

【請求項 58】

前記閾値距離が、三次元である、先行請求項のいずれか一項に記載の方法。

【請求項 59】

前記閾値距離が、二次元である、先行請求項のいずれか一項に記載の方法。

【請求項 60】

前記閾値距離が、軸方向軸に沿った第1の閾値距離及び前記軸方向軸に直交する平面内の第2の閾値を含む、先行請求項のいずれか一項に記載の方法。

10

【請求項 61】

前記閾値距離が、軸方向軸に沿った第1の閾値距離、並びに前記軸方向軸に直交する平面内の第2の閾値及び第3の閾値を含む、先行請求項のいずれか一項に記載の方法。

【請求項 62】

前記第1の閾値距離が、前記第2及び／又は第3の閾値距離と異なる、先行請求項のいずれか一項に記載の方法。

【請求項 63】

前記第1の閾値距離が、前記第2及び／又は第3の閾値距離と同一である、先行請求項のいずれか一項に記載の方法。

【請求項 64】

前記方法が、前記第1の標的RNA分子及び前記第2の標的RNA分子を含む複数のRNAを保有する前記細胞試料を提供することを更に含む、先行請求項のいずれか一項に記載の方法。

20

【請求項 65】

前記方法が、

前記細胞試料内で、前記第1の標的RNA分子に対応する第1の標的cDNA分子及び前記第2の標的RNA分子に対応する第2の標的cDNA分子を含む複数のcDNA分子を生成することを更に含む、先行請求項のいずれか一項に記載の方法。

【請求項 66】

前記方法が、

前記細胞試料中の前記複数のcDNA分子を、少なくとも第1の複数の第1の標的特異的パドロックプローブ及び第2の複数の第2の標的特異的パドロックプローブを含む、複数の標的特異的パドロックプローブと接触させることを更に含む、先行請求項のいずれか一項に記載の方法。

30

【請求項 67】

前記方法が、

前記細胞試料中の前記複数のRNA分子を、少なくとも第1の複数の標的特異的パドロックプローブ及び第2の複数の標的特異的パドロックプローブを含む、複数の標的特異的パドロックプローブと接触させることを更に含む、先行請求項のいずれか一項に記載の方法。

40

【請求項 68】

前記第1の複数の第1の標的特異的パドロックプローブにおける個々のパドロックプローブが、

第1及び第2の末端領域であって、前記第1の末端領域が、前記第1の標的cDNA分子又は前記第1の標的RNA分子の第1の領域に選択的にハイブリダイズし、前記第2の末端領域が、前記第1の標的cDNA分子又は前記第1の標的RNA分子の第2の領域に選択的にハイブリダイズする、第1及び第2の末端領域を含む、先行請求項のいずれか一項に記載の方法。

【請求項 69】

前記細胞試料中の前記複数のRNA分子を、前記複数の標的特異的パドロックプローブ

50

と接触させることが、

前記第1の標的特異的パドロックプロープの前記第1及び第2の末端領域を、前記第1の標的cDNA分子又は前記第1の標的RNA分子上の近位位置にハイブリダイズして、ハイブリダイズした前記第1及び第2の末端領域の間にニック又はギャップを有する環状化された第1の標的特異的パドロックプロープを形成することを含む、先行請求項のいずれか一項に記載の方法。

【請求項70】

前記第1の標的特異的パドロックプロープが、前記第1の標的cDNA配列又は前記第1の標的RNA配列に対応し、かつそれを一意に識別する第1の標的バーコード配列を含む、先行請求項のいずれか一項に記載の方法。

10

【請求項71】

前記第1の標的特異的パドロックプロープが、前記第1の標的cDNA分子又は前記第1の標的RNA配列に選択的にハイブリダイズする前記第1の標的特異的パドロックプロープの前記領域のうちの1つに隣接して位置する第1の標的バーコード配列を含む、先行請求項のいずれか一項に記載の方法。

【請求項72】

前記第1の標的特異的パドロックプロープが、少なくとも1つのユニバーサルアダプター配列を含む、先行請求項のいずれか一項に記載の方法。

【請求項73】

前記第1の標的特異的パドロックプロープが、ローリングサークル増幅プライマー又はその相補的配列のためのユニバーサルプライマー結合部位を含む、先行請求項のいずれか一項に記載の方法。

20

【請求項74】

前記第1の標的特異的パドロックプロープが、ユニバーサルコンパクションオリゴヌクレオチド結合部位又はその相補的配列を含む、先行請求項のいずれか一項に記載の方法。

【請求項75】

前記方法が、

酵素反応を行うことによって、前記少なくとも第1及び第2の環状化標的特異的パドロックプロープ内の前記ニック又はギャップを閉鎖し、それによって、前記細胞試料内に少なくとも第1の共有結合性閉鎖環状パドロックプロープ及び第2の共有結合性閉鎖環状パドロックプロープを生成することを更に含む、先行請求項のいずれか一項に記載の方法。

30

【請求項76】

前記方法が、

前記第1及び第2の共有結合性閉鎖環状パドロックプロープをテンプレート分子として使用して、前記細胞試料内でローリングサークル増幅反応を行い、それによって、前記第1の標的RNA分子に対応する少なくとも前記第1のコンカテマー分子、及び前記第2の標的RNA分子に対応する前記第2のコンカテマー分子を含む複数のコンカテマー分子を生成することを更に含む、先行請求項のいずれか一項に記載の方法。

【請求項77】

前記第1のコンカテマーが、前記第1の標的RNA又は前記第1の標的cDNA配列を一意に識別する第1の標的バーコード配列と、前記第1の標的RNA又は前記第1の標的cDNAに対応する第1の挿入配列と、第1のシーケンシングプライマー結合部位又はその相補的配列と、のタンデムリピート単位を含む、先行請求項のいずれか一項に記載の方法。

40

【請求項78】

前記第1のコンカテマーが、増幅プライマーのためのユニバーサル結合部位又はその相補的配列と、コンパクションオリゴヌクレオチドのためのユニバーサル結合部位又はその相補的配列と、を更に含む、先行請求項のいずれか一項に記載の方法。

【請求項79】

前記第2のコンカテマーが、前記第2の標的RNA又は前記第2の標的cDNA配列を

50

一意に識別する第2の標的バーコード配列と、前記第2の標的RNA又は前記第2の標的cDNAに対応する第2の挿入配列と、第2のシーケンシングプライマー結合部位又はその相補的配列と、のタンデムリピート単位を含む、先行請求項のいずれか一項に記載の方法。

【請求項80】

前記第2のコンカテマーが、増幅プライマーのためのユニバーサル結合部位又はその相補的配列、及びコンパクションオリゴヌクレオチドのためのユニバーサル結合部位又はその相補的配列を更に含む、先行請求項のいずれか一項に記載の方法。

【請求項81】

前記1つ以上のシーケンシング反応サイクルを行うことが、

10

前記細胞試料内の前記複数のコンカテマー分子を、(i)複数のユニバーサルシーケンシングプライマー、(ii)複数のシーケンシングポリメラーゼ、及び(iii)複数のヌクレオチド試薬と、前記複数のユニバーサルシーケンシングプライマーを前記コンカテマー上のそれらのそれぞれのユニバーサルシーケンシングプライマー結合部位にハイブリダイズするのに好適な条件下で接触させることを含む、先行請求項のいずれか一項に記載の方法。

【請求項82】

前記複数のヌクレオチド試薬が、多価分子、ヌクレオチド、ヌクレオチド類似体、又はそれらの組み合わせを含む、先行請求項のいずれか一項に記載の方法。

【請求項83】

20

個々のヌクレオチド又はヌクレオチド類似体が、検出可能に標識されているか、又は標識されていない、先行請求項のいずれか一項に記載の方法。

【請求項84】

前記検出可能に標識された個々のヌクレオチド又はヌクレオチド類似体が、A、G、C、及びT/Uのヌクレオチド塩基の各々の異なる種類に対応する、異なる検出可能な色ラベルを含む、先行請求項のいずれか一項に記載の方法。

【請求項85】

個々の多価分子が、複数のヌクレオチドアームが付着したコアを含み、前記個々の多価分子の各アームが、同じ種類のヌクレオチド塩基を含む、先行請求項のいずれか一項に記載の方法。

30

【請求項86】

前記第1の複数のフローセル画像を生成することが、

各サイクルにおいて、光学システムによって、前記複数のコンカテマー分子に結合した前記ヌクレオチド試薬から放出された光学カラーシグナルをイメージングすることを含む、先行請求項のいずれか一項に記載の方法。

【請求項87】

前記第1の複数のフローセル画像が、前記複数のコンカテマー分子に結合した前記ヌクレオチド試薬から放出された光学カラーシグナルを含む、先行請求項のいずれか一項に記載の方法。

【請求項88】

40

前記1つ以上のシーケンシング反応サイクルを行うことが、

前記第1のコンカテマーの前記第1の標的バーコード配列領域のみをシーケンシングし、それによって前記第1のシーケンシングリード生成物を生成することを含む、先行請求項のいずれか一項に記載の方法。

【請求項89】

前記1つ以上のシーケンシング反応サイクルを行うことが、

前記第1の標的バーコード配列領域及び前記第1のコンカテマーの前記第1の挿入配列の少なくとも一部分をシーケンシングし、それによって前記第1のシーケンシングリード生成物を生成することを含む、先行請求項のいずれか一項に記載の方法。

【請求項90】

50

前記 1 つ以上のシーケンシング反応サイクルを行うことが、

前記第 2 のコンカテマーの前記第 2 の標的バーコード配列領域のみをシーケンシングし、それによって前記第 2 のシーケンシングリード生成物を生成することを含む、先行請求項のいずれか一項に記載の方法。

【請求項 9 1】

前記 1 つ以上のシーケンシング反応サイクルを行うことが、

前記第 2 の標的バーコード配列領域及び前記第 2 のコンカテマーの前記第 2 の挿入配列の少なくとも一部分をシーケンシングし、それによって前記第 2 のシーケンシングリード生成物を生成することを含む、先行請求項のいずれか一項に記載の方法。

【請求項 9 2】

前記方法が、

第 1 のシーケンシングリード生成物を前記第 1 のコンカテマー分子から除去し、前記第 1 のコンカテマー分子を前記細胞試料中に保持することと、

第 2 のシーケンシングリード生成物を前記第 2 のコンカテマー分子から除去し、前記第 2 のコンカテマー分子を前記細胞試料中に保持することと、を更に含む、先行請求項のいずれか一項に記載の方法。

【請求項 9 3】

前記方法が、少なくとも 1 回、以下の操作：

1 つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の前記第 1 の複数のフローセル画像を生成し、それによって、前記第 1 のシーケンシングリード生成物及び前記第 2 のシーケンシング生成物を生成する操作であって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第 1 のコンカテマー分子が、前記細胞試料の第 1 の標的 RNA 分子に対応し、前記複数のコンカテマー分子の第 2 のコンカテマー分子が、前記細胞試料の第 2 の標的 RNA 分子に対応し、前記第 1 の複数のフローセル画像、生成する操作と、

第 1 のシーケンシングリード生成物を前記第 1 のコンカテマー分子から除去し、前記第 1 のコンカテマー分子を前記細胞試料中に保持し、かつ第 2 のシーケンシングリード生成物を前記第 2 のコンカテマー分子から除去し、前記第 2 のコンカテマー分子を前記細胞試料中に保持する操作と、を繰り返すことによって、前記複数のコンカテマーを反復してシーケンシングすることを更に含む、先行請求項のいずれか一項に記載の方法。

【請求項 9 4】

前記第 1 のシーケンシングリード生成物が、前記第 1 のコンカテマー分子の 1 つ以上のタンデム単位における第 1 の標的バーコード配列、前記第 1 のコンカテマー分子の 1 つ以上のタンデム単位における第 1 の挿入配列、又はそれらの組み合わせのうちのいくつか又は全てを含む、先行請求項のいずれか一項に記載の方法。

【請求項 9 5】

前記方法が、前記第 1 の標的 RNA 分子、前記第 2 の標的 RNA 分子、又は両方の分子が、前記ベースコーリングテンプレート内の前記ベースコーリング位置で前記第 2 の複数のフローセル画像の前記実行されたベースコーリングに基づいて、前記細胞試料中に存在することを確認することを更に含む、先行請求項のいずれか一項に記載の方法。

【請求項 9 6】

前記方法が、前記シーケンシングシステムによって、前記 1 つ以上のサイクルの後に、後続のシーケンシング反応サイクルを行うことによって、前記支持体上に固定化された前記細胞試料の前記第 2 の複数のフローセル画像を生成することを更に含む、先行請求項のいずれか一項に記載の方法。

【請求項 9 7】

前記支持体上に固定化された前記細胞試料の前記第 1 の複数のフローセル画像を生成することが、

前記第 2 のコンカテマーのシーケンシングを阻害する条件下で、前記細胞試料内の少なくとも前記第 1 のコンカテマーをシーケンシングすることを含む、先行請求項のいずれか

10

20

30

40

50

一項に記載の方法。

【請求項 98】

前記細胞試料内の少なくとも前記第1のコンカテマーをシーケンシングすることが、複数の第1のシーケンシングリード生成物を生成することであって、前記第1のシーケンシングリード生成物の前記配列が、前記細胞試料中の前記第1の標的RNAの存在を確認するために、第1の標的参照配列とアラインメントされる、生成することを含む、先行請求項のいずれか一項に記載の方法。

【請求項 99】

前記支持体上に固定化された前記細胞試料の前記第1の複数のフローセル画像を生成することが、

前記第1のコンカテマーのシーケンシングを阻害する条件下で、前記細胞試料内の少なくとも前記第2のコンカテマーをシーケンシングすることを含む、先行請求項のいずれか一項に記載の方法。

【請求項 100】

前記細胞試料内の少なくとも前記第2のコンカテマーをシーケンシングすることが、複数の第2のシーケンシングリード生成物を生成することであって、前記第2のシーケンシングリード生成物の配列が、前記細胞試料中の前記第2の標的RNAの存在を確認するために、第2の標的参照配列とアラインメントされる、生成することを含む、先行請求項のいずれか一項に記載の方法。

【請求項 101】

システムであって、

1つ以上のハードウェアプロセッサと、

前記1つ以上のハードウェアプロセッサに操作を実行させるために、前記1つ以上のハードウェアプロセッサによって実行可能な命令を格納する1つ以上のデータストレージデバイスであって、前記操作が、

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標的RNA分子に対応する、生成することと、

プロセッサによって、前記第1の複数のフローセル画像のピクセルについて、ピクセル強度及び前記ピクセル強度の各々のそれぞれの色純度を判定することと、

前記プロセッサによって、前記ピクセル強度及び前記ピクセル強度の前記それぞれの色純度に基づいて、ベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリングテンプレートが、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を位置合わせするように構成される、判定することと、を含む、1つ以上のデータストレージデバイスと、を含む、システム。

【請求項 102】

システムであって、

1つ以上のハードウェアプロセッサと、

前記1つ以上のハードウェアプロセッサに操作を実行させるために、前記1つ以上のハードウェアプロセッサによって実行可能な命令を格納する1つ以上のデータストレージデバイスであって、前記操作が、

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標

10

20

30

40

50

的RNA分子に対応する、生成することと、

プロセッサによって、前記第1の複数のフローセル画像のピクセルについて、ボクセル強度及びピクセル強度の各々のそれぞれの色純度を判定することと、

前記プロセッサによって、並びに前記ピクセル強度及び前記ピクセル強度の前記それぞれの色純度に基づいて、軸方向軸に沿った異なる軸方向位置を有するベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリングテンプレートが、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を位置合わせするように構成される、判定することと、を含む、1つ以上のデータストレージデバイスと、を含む、システム。

【請求項103】

システムであって、

1つ以上のハードウェアプロセッサと、

前記1つ以上のハードウェアプロセッサに操作を実行させるために、前記1つ以上のハードウェアプロセッサによって実行可能な命令を格納する1つ以上のデータストレージデバイスであって、前記操作が、

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標

的RNA分子に対応する、生成することと、

プロセッサによって、前記第1の複数のフローセル画像のピクセルについて、ボクセル強度及びピクセル強度の各々のそれぞれの色純度を判定することと、

前記プロセッサによって、並びに前記ピクセル強度及び前記ピクセル強度の前記それぞれの色純度に基づいて、ベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリングテンプレートが、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を位置合わせするように構成され、前記第1又は第2の複数のフローセル画像が、軸方向軸に沿った異なる軸方向位置で生成される、判定することと、を含む、1つ以上のデータストレージデバイスと、を含む、システム。

【請求項104】

システムであって、

1つ以上のハードウェアプロセッサと、

前記1つ以上のハードウェアプロセッサに操作を実行させるために、前記1つ以上のハードウェアプロセッサによって実行可能な命令を格納する1つ以上のデータストレージデバイスであって、前記操作が、

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標

的RNA分子に対応する、生成することと、

プロセッサによって、ピクセル強度及び前記ピクセル強度の各々のそれぞれの色純度に基づいたスポット探索アルゴリズムを使用することによって、前記第1の複数のフローセル画像内のベースコーリング位置を識別することと、

前記プロセッサによって、前記識別されたベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリングテンプレートが、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を位置合わせするように構成され、前記ベースコーリング位置が、軸方向軸に沿った異なる軸方向位置を含む、判定することと、を含む、1つ以上のデータストレージデ

10

20

30

40

50

バイスと、を含む、システム。

【請求項 105】

システムであって、

1つ以上のハードウェアプロセッサと、

前記1つ以上のハードウェアプロセッサに操作を実行させるために、前記1つ以上のハードウェアプロセッサによって実行可能な命令を格納する1つ以上のデータストレージデバイスであって、前記操作が、

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標的RNA分子に対応する、生成することと、

プロセッサによって、ピクセル強度及び前記ピクセル強度の各々のそれぞれの色純度に基づいたスポット探索アルゴリズムを使用することによって、前記第1の複数のフローセル画像内のベースコーリング位置を識別することと、

前記プロセッサによって、前記識別されたベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリングテンプレートが、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を位置合わせするように構成され、前記ベースコーリング位置が、軸方向軸に沿った異なる軸方向位置を含む、判定することと、を含む、1つ以上のデータストレージデバイスと、を含む、システム。

【請求項 106】

システムであって、

1つ以上のハードウェアプロセッサと、

前記1つ以上のハードウェアプロセッサに操作を実行させるために、前記1つ以上のハードウェアプロセッサによって実行可能な命令を格納する1つ以上のデータストレージデバイスであって、前記操作が、

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標的RNA分子に対応する、生成することと、

プロセッサによって、ピクセル強度及び前記ピクセル強度の各々のそれぞれの色純度に基づいたスポット探索アルゴリズムを使用することによって、前記第1の複数のフローセル画像内のベースコーリング位置を識別することと、

前記プロセッサによって、並びに前記ピクセル強度及び前記ピクセル強度の前記それぞれの色純度に基づいて、ベースコーリング位置を含むベースコーリングテンプレートを判定することと、

前記プロセッサによって、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を前記ベースコーリングテンプレートに対して位置合わせすることであって、前記第1又は第2の複数のフローセル画像が、軸方向軸に沿った異なる軸方向位置で生成される、位置合わせすることと、を含む、1つ以上のデータストレージデバイスと、を含む、システム。

【請求項 107】

システムであって、

1つ以上のハードウェアプロセッサと、

前記1つ以上のハードウェアプロセッサに操作を実行させるために、前記1つ以上のハードウェアプロセッサによって実行可能な命令を格納する1つ以上のデータストレージ

バイスであって、前記操作が、

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標的RNA分子に対応する、生成することと、

プロセッサによって、ピクセル強度及び前記ピクセル強度の各々のそれぞれの色純度に基づいたスポット探索アルゴリズムを使用することによって、前記第1の複数のフローセル画像内のベースコーリング位置を識別することと、

前記プロセッサによって、前記識別されたベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリング位置が、軸方向軸に沿った異なる軸方向位置を含む、判定することと、

前記プロセッサによって、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を前記ベースコーリングテンプレートに対して位置合わせすることと、

前記プロセッサによって、前記ベースコーリングテンプレートの前記ベースコーリング位置に基づいて、前記第2の複数のフローセル画像のベースコーリングを実行することと、を含む、1つ以上のデータストレージデバイスと、を含む、システム。

【請求項108】

システムであって、

1つ以上のハードウェアプロセッサ、

前記1つ以上のハードウェアプロセッサに操作を実行させるために、前記1つ以上のハードウェアプロセッサによって実行可能な命令を格納する1つ以上のデータストレージデバイスであって、前記操作が、請求項8～100のいずれか一項を含む、1つ以上のデータストレージデバイス、を含む、システム。

【請求項109】

非一時的コンピュータ可読デバイスであって、それに格納された命令を有し、前記命令が、少なくとも1つの計算デバイスによって実行されると、前記少なくとも1つの計算デバイスに、

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標的RNA分子に対応する、生成することと、

プロセッサによって、前記第1の複数のフローセル画像のピクセルについて、ピクセル強度及び前記ピクセル強度の各々のそれぞれの色純度を判定することと、

前記プロセッサによって、前記ピクセル強度及び前記ピクセル強度の前記それぞれの色純度に基づいて、ベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリングテンプレートが、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を位置合わせするように構成される、判定することと、を含む、操作を実行させる、非一時的コンピュータ可読デバイス。

【請求項110】

非一時的コンピュータ可読デバイスであって、それに格納された命令を有し、前記命令が、少なくとも1つの計算デバイスによって実行されると、前記少なくとも1つの計算デバイスに、

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成するこ

10

20

30

40

50

とであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標的RNA分子に対応する、生成することと、

プロセッサによって、前記第1の複数のフローセル画像のピクセルについて、ボクセル強度及びピクセル強度の各々のそれぞれの色純度を判定することと、

前記プロセッサによって、並びに前記ピクセル強度及び前記ピクセル強度の前記それぞれの色純度に基づいて、軸方向軸に沿った異なる軸方向位置を有するベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリングテンプレートが、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を位置合わせするように構成される、判定することと、を含む、操作を実行させる、非一時的コンピュータ可読デバイス。

10

【請求項111】

非一時的コンピュータ可読デバイスであって、それに格納された命令を有し、前記命令が、少なくとも1つの計算デバイスによって実行されると、前記少なくとも1つの計算デバイスに、

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標的RNA分子に対応する、生成することと、

20

プロセッサによって、前記第1の複数のフローセル画像のピクセルについて、ボクセル強度及びピクセル強度の各々のそれぞれの色純度を判定することと、

前記プロセッサによって、並びに前記ピクセル強度及び前記ピクセル強度の前記それぞれの色純度に基づいて、ベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリングテンプレートが、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を位置合わせするように構成され、前記第1又は第2の複数のフローセル画像が、軸方向軸に沿った異なる軸方向位置で生成される、判定することと、を含む、操作を実行させる、非一時的コンピュータ可読デバイス。

30

【請求項112】

非一時的コンピュータ可読デバイスであって、それに格納された命令を有し、前記命令が、少なくとも1つの計算デバイスによって実行されると、前記少なくとも1つの計算デバイスに、

シーケンシングシステムによって、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第1のコンカテマー分子が、前記細胞試料の第1の標的RNA分子に対応し、前記複数のコンカテマー分子の第2のコンカテマー分子が、前記細胞試料の第2の標的RNA分子に対応する、生成することと、

40

プロセッサによって、ピクセル強度及び前記ピクセル強度の各々のそれぞれの色純度に基づいたスポット探索アルゴリズムを使用することによって、前記第1の複数のフローセル画像内のベースコーリング位置を識別することと、

前記プロセッサによって、前記識別されたベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリングテンプレートが、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を位置合わせするように構成され、前記ベースコーリング位置が、軸方向軸に沿った異なる軸方向位置を含む、判定することと、を含む、操作を実行させる、非一時的コンピュータ可読デバイス。

50

【請求項 1 1 3】

非一時的コンピュータ可読デバイスであって、それに格納された命令を有し、前記命令が、少なくとも 1 つの計算デバイスによって実行されると、前記少なくとも 1 つの計算デバイスに、

シーケンシングシステムによって、1 つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第 1 の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第 1 のコンカテマー分子が、前記細胞試料の第 1 の標的 RNA 分子に対応し、前記複数のコンカテマー分子の第 2 のコンカテマー分子が、前記細胞試料の第 2 の標的 RNA 分子に対応する、生成することと、

10

プロセッサによって、ピクセル強度及び前記ピクセル強度の各々のそれぞれの色純度に基づいたスポット探索アルゴリズムを使用することによって、前記第 1 の複数のフローセル画像内のベースコーリング位置を識別することと、

前記プロセッサによって、前記識別されたベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリングテンプレートが、前記 1 つ以上のサイクルの 1 つ以上の後続のサイクルにおける前記支持体の第 2 の複数のフローセル画像を位置合わせするように構成され、前記ベースコーリング位置が、軸方向軸に沿った異なる軸方向位置を含む、判定することと、を含む、操作を実行させる、非一時的コンピュータ可読デバイス。

20

【請求項 1 1 4】

非一時的コンピュータ可読デバイスであって、それに格納された命令を有し、前記命令が、少なくとも 1 つの計算デバイスによって実行されると、前記少なくとも 1 つの計算デバイスに、

シーケンシングシステムによって、1 つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第 1 の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第 1 のコンカテマー分子が、前記細胞試料の第 1 の標的 RNA 分子に対応し、前記複数のコンカテマー分子の第 2 のコンカテマー分子が、前記細胞試料の第 2 の標的 RNA 分子に対応する、生成することと、

30

プロセッサによって、ピクセル強度及び前記ピクセル強度の各々のそれぞれの色純度に基づいたスポット探索アルゴリズムを使用することによって、前記第 1 の複数のフローセル画像内のベースコーリング位置を識別することと、

前記プロセッサによって、並びに前記ピクセル強度及び前記ピクセル強度の前記それぞれの色純度に基づいて、ベースコーリング位置を含むベースコーリングテンプレートを判定することと、

前記プロセッサによって、前記 1 つ以上のサイクルの 1 つ以上の後続のサイクルにおける前記支持体の第 2 の複数のフローセル画像を前記ベースコーリングテンプレートに対して位置合わせすることであって、前記第 1 又は第 2 の複数のフローセル画像が、軸方向軸に沿った異なる軸方向位置で生成される、位置合わせすることと、を含む、操作を実行させる、非一時的コンピュータ可読デバイス。

40

【請求項 1 1 5】

非一時的コンピュータ可読デバイスであって、それに格納された命令を有し、前記命令が、少なくとも 1 つの計算デバイスによって実行されると、前記少なくとも 1 つの計算デバイスに、

シーケンシングシステムによって、1 つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の第 1 の複数のフローセル画像を生成することであって、前記細胞試料が、その中に複数のコンカテマー分子を含み、前記複数のコンカテマー分子の第 1 のコンカテマー分子が、前記細胞試料の第 1 の標的 RNA 分子に対応し、前記複数のコンカテマー分子の第 2 のコンカテマー分子が、前記細胞試料の第 2 の標的 RNA 分子に対応する、生成することと、

50

プロセッサによって、ピクセル強度及び前記ピクセル強度の各々のそれぞれの色純度に基づいたスポット探索アルゴリズムを使用することによって、前記第1の複数のフローセル画像内のベースコーリング位置を識別することと、

前記プロセッサによって、前記識別されたベースコーリング位置を含むベースコーリングテンプレートを判定することであって、前記ベースコーリング位置が、軸方向軸に沿った異なる軸方向位置を含む、判定することと、

前記プロセッサによって、前記1つ以上のサイクルの1つ以上の後続のサイクルにおける前記支持体の第2の複数のフローセル画像を前記ベースコーリングテンプレートに対して位置合わせすることと、

前記プロセッサによって、前記ベースコーリングテンプレートの前記ベースコーリング位置に基づいて、前記第2の複数のフローセル画像のベースコーリングを実行することと、を含む、操作を実行させる、非一時的コンピュータ可読デバイス。

10

【請求項116】

非一時的コンピュータ可読デバイスであって、それに格納された命令を有し、前記命令が、少なくとも1つの計算デバイスによって実行されると、前記少なくとも1つの計算デバイスに、請求項8～100のいずれか一項を含む、操作を実行させる、非一時的コンピュータ可読デバイス。

【請求項117】

データファイルであって、

1つ以上の細胞特徴のデータを含み、前記1つ以上の細胞特徴の各々の前記データが、細胞特徴インジケータと、

20

前記支持体上の前記細胞特徴の空間座標と、

前記細胞特徴に対応する1つ以上のシーケンシングパラメータ及びその値と、を含む

、
前記データファイルが、次世代シーケンシング（NGS）システムによって生成され、ベースコールを含むデータファイルと同一又は互換性のあるデータ形式を含む、データファイル。

【請求項118】

データファイルであって、

ベースコールと、

30

前記ベースコールに対応する品質インジケータと、

前記支持体上の前記ベースコールの空間座標と、

1つ以上の細胞特徴のデータであって、前記1つ以上の細胞特徴の各々の前記データが

、
細胞特徴インジケータと、

前記支持体上の前記細胞特徴の空間座標と、

前記細胞特徴に対応する1つ以上のシーケンシングパラメータ及びその値と、を含む

、1つ以上の細胞特徴のデータと、を含む、

前記データファイルが、次世代シーケンシング（NGS）システムによって生成され、ベースコールを含むデータファイルと同一又は互換性のあるデータ形式を含む、データファイル。

40

【請求項119】

データファイルを生成するための方法であって、

光学システムによって、中に1つ以上の細胞特徴を含む1つ以上の画像を生成することと、

前記1つ以上の画像内の前記1つ以上の細胞特徴の各々の空間座標を判定することと、

前記判定された空間座標に基づいて1つ以上の細胞特徴のデータを生成することであって、

前記1つ以上の細胞特徴の各々について、

細胞特徴インジケータを生成すること、

50

前記細胞特徴の前記空間座標を挿入すること、を含む、生成することと、を含み、
前記データファイルが、次世代シーケンシング（NGS）システムによって生成され、
ベースコールを含むデータファイルと同一又は互換性のあるデータ形式を含む、方法。

【請求項 1 2 0】

データファイルを生成するための方法であって、
光学システムによって、中に 1 つ以上の細胞特徴を含む 1 つ以上の画像を生成することと、

前記 1 つ以上の画像内の前記 1 つ以上の細胞特徴の各々の空間座標を判定することと、
前記判定された空間座標に基づいて、ベースコールを含む次世代シーケンシング（NGS）システムによって生成されるデータ形式と同一又は互換性のあるデータ形式で 1 つ以上の細胞特徴のデータを生成することであって、

前記 1 つ以上の細胞特徴の各々について、所定の回数の操作の反復を実行することであって、

前記細胞特徴の細胞特徴インジケータを挿入すること、

前記細胞特徴の前記空間座標の空間座標のセットを挿入すること、並びに

前記細胞特徴に対応する 1 つ以上のシーケンシングパラメータ及びその値を挿入すること、を含む、実行することを含む、生成することと、を含み、

前記データファイルが、前記 1 つ以上の細胞特徴及びベースコーリング結果を含む画像を再構築するように構成される、方法。

【請求項 1 2 1】

前記データファイルが、1 つ以上の区切り符号を更に含む、先行請求項のいずれか一項に記載の方法。

【請求項 1 2 2】

前記 1 つ以上の区切り符号の各々が、前記 1 つ以上の細胞特徴を区切るように構成される、先行請求項のいずれか一項に記載の方法。

【請求項 1 2 3】

前記 1 つ以上の細胞特徴の各々について、前記所定の回数の操作の反復を実行することであって、

前記 1 つ以上の区切り符号のうちの少なくとも 1 つを挿入することを更に含む、先行請求項のいずれか一項に記載の方法。

【請求項 1 2 4】

前記 1 つ以上の区切り符号の各々が、前記 1 つ以上の細胞特徴の前記空間座標を区切るように構成される、先行請求項のいずれか一項に記載の方法。

【請求項 1 2 5】

前記細胞特徴の各々の前記空間座標が、空間座標の複数のセットを含む、先行請求項のいずれか一項に記載の方法。

【請求項 1 2 6】

空間座標の各セットが、3Dでの一意の位置を示すように構成される、先行請求項のいずれか一項に記載の方法。

【請求項 1 2 7】

前記空間座標が、3Dである、先行請求項のいずれか一項に記載の方法。

【請求項 1 2 8】

前記方法が、

シーケンシングシステムによって、前記支持体上に固定化された前記細胞試料の第 1 の複数のフローセル画像を生成することを更に含む、先行請求項のいずれか一項に記載の方法。

【請求項 1 2 9】

前記方法が、プロセッサによって、前記細胞試料の前記第 1 の複数のフローセル画像に基づいて、ベースコールを生成することを更に含む、先行請求項のいずれか一項に記載の方法。

10

20

30

40

50

【請求項 130】

前記方法が、
ベースコールを挿入することと、
前記ベースコールに対応する品質インジケータを挿入することと、前記ベースコールの空間座標を挿入することと、を更に含む、先行請求項のいずれか一項に記載の方法。

【請求項 131】

前記ベースコールの前記空間座標と、前記細胞特徴の前記空間座標とが、同じ座標系内にある、先行請求項のいずれか一項に記載の方法。

【請求項 132】

前記細胞特徴が、前記支持体上に固定化された前記細胞試料のものである、先行請求項のいずれか一項に記載の方法。

【請求項 133】

前記ベースコールが、前記支持体上に固定化された前記細胞試料のシーケンシングから生成される、先行請求項のいずれか一項に記載の方法。

【請求項 134】

前記ベースコールが、2Dシーケンシング分析から生成される、先行請求項のいずれか一項に記載の方法。

【請求項 135】

前記ベースコールが、3Dシーケンシング分析から得られる、先行請求項のいずれか一項に記載の方法。

【請求項 136】

前記ベースコールが、前記細胞試料のシーケンシングから生成される、先行請求項のいずれか一項に記載の方法。

【請求項 137】

前記品質インジケータが、前記ベースコールの品質スコアを含む、先行請求項のいずれか一項に記載の方法。

【請求項 138】

前記細胞特徴インジケータが、核、膜、細胞質、及びミトコンドリアから選択される細胞要素を示すように構成される、先行請求項のいずれか一項に記載の方法。

【請求項 139】

前記細胞特徴インジケータが、核境界及び膜境界から選択される細胞要素を示すように構成される、先行請求項のいずれか一項に記載の方法。

【請求項 140】

前記1つ以上のシーケンシングパラメータが、標的配列を含む、先行請求項のいずれか一項に記載の方法。

【請求項 141】

前記ベースコールが、バーコード配列の1つ以上のベースコールを含む、先行請求項のいずれか一項に記載の方法。

【請求項 142】

前記ベースコールが、1つ以上のシーケンシングサイクルにおけるポロニーに由来する、先行請求項のいずれか一項に記載の方法。

【請求項 143】

前記1つ以上のシーケンシングパラメータが、シーケンシングシステムの名称を含む、先行請求項のいずれか一項に記載の方法。

【請求項 144】

1つ以上のシーケンシングパラメータが、シーケンシングランの識別を含む、先行請求項のいずれか一項に記載の方法。

【請求項 145】

1つ以上のシーケンシングパラメータが、フローセルの識別を含む、先行請求項のいずれか一項に記載の方法。

10

20

30

40

50

【請求項 1 4 6】

1つ以上のシーケンシングパラメータが、フローセルのレーン番号を含む、先行請求項のいずれか一項に記載の方法。

【請求項 1 4 7】

1つ以上のシーケンシングパラメータが、フローセルレーンのタイル数を含む、先行請求項のいずれか一項に記載の方法。

【請求項 1 4 8】

前記ベースコールの前記空間座標と、前記1つ以上の細胞特徴の前記空間座標が、x座標、y座標、z座標、又はそれらの組み合わせを含む、先行請求項のいずれか一項に記載の方法。

10

【請求項 1 4 9】

前記ベースコール、品質インジケータ、1つ以上の細胞特徴、及び空間座標が、1つ以上の文字、1つ以上の数字、1つ以上の記号、又はそれらの組み合わせを含む、先行請求項のいずれか一項に記載の方法。

【請求項 1 5 0】

前記データファイルが、テキストファイルである、先行請求項のいずれか一項に記載の方法。

【請求項 1 5 1】

前記データファイルが、FastQファイルである、先行請求項のいずれか一項に記載の方法。

20

【請求項 1 5 2】

前記データファイルが、ASCII符号化される、先行請求項のいずれか一項に記載の方法。

【請求項 1 5 3】

前記データファイルが、8ビット、12ビット、16ビット、18ビット、32ビット、又は36ビットの符号化形式である、先行請求項のいずれか一項に記載の方法。

【請求項 1 5 4】

前記1つ以上の区切り符号の各々が、文字、数字、記号、又はそれらの組み合わせを含む、先行請求項のいずれか一項に記載の方法。

【請求項 1 5 5】

前記1つ以上の区切り符号の各々が、文字、数字、記号、又はそれらの組み合わせを含む、先行請求項のいずれか一項に記載の方法。

30

【請求項 1 5 6】

1つ以上の細胞特徴の前記データが、前記細胞特徴の前記空間座標に基づいて、前記1つ以上の区切り符号によって区切られる、先行請求項のいずれか一項に記載の方法。

【請求項 1 5 7】

1つ以上の細胞特徴の前記データが、複数のエントリに分けられ、前記複数のエントリの各エントリが、前記支持体上の空間位置に対応する、先行請求項のいずれか一項に記載の方法。

【請求項 1 5 8】

前記空間位置が、3Dである、先行請求項のいずれか一項に記載の方法。

40

【請求項 1 5 9】

1つ以上の細胞特徴の前記データが、複数のエントリに分けられ、前記複数のエントリの各エントリが、空間座標のセットに対応する、先行請求項のいずれか一項に記載の方法。

【請求項 1 6 0】

空間座標の前記セットが、x座標、y座標、z座標、又はそれらの組み合わせを含む、先行請求項のいずれか一項に記載の方法。

【発明の詳細な説明】

【技術分野】

50

【0001】

関連出願の相互参照

本出願は、2020年8月31日に提出された米国仮特許出願第63/072,649号に対する優先権を主張する2021年3月31日に提出された米国特許出願第17/219,556号の継続である2021年12月10日に提出された米国特許出願第17/547,602号の継続である2022年6月30日に提出された米国特許出願第17/854,042号の一部継続である。本出願はまた、2022年7月11日に提出された米国仮特許出願第63/388,183号に対する優先権を主張する。本出願は、2022年6月6日に提出された米国仮特許出願第63/349,421号に対する優先権を主張する。上記の特許出願は、それらの全体が参照により本明細書に組み込まれる。

10

【0002】

本開示は、概して、画像データ分析に関し、特に、DNAシーケンシング中にフローセルのデジタル画像内でベースコーリングを実行するためのクラスタ又はポロニー位置を識別することに関する。

【背景技術】

【0003】

フローセルを使用した合成による次世代シーケンシングは、DNAの配列を識別するために使用され得る。シーケンシングライブラリからの一本鎖DNA断片がフローセルにわたってフラッディングされると、断片は、典型的には、フローセルの表面に結合した相補オリゴマー又はその上に存在するビーズに起因して、フローセルの表面にランダムに付着する。次いで、DNA断片に対して増幅プロセスを実行し、その結果、所与の断片のコピーが、変性クローン化ヌクレオチド鎖のクラスタ又はポロニーを形成する。いくつかの態様において、単一のビーズは、クラスタを含み得、ビーズは、ランダムな位置でフローセルに付着し得る。

20

【0004】

鎖の配列を識別するために、一度に1つのヌクレオチド塩基ずつ鎖対を再構築する。各塩基構築サイクル中に、各々が蛍光標識（又はタグ）及びブロッカーに付着した単一のヌクレオチドの混合物がフローセルにわたってフラッディングされる。ヌクレオチドは、鎖上の相補的な位置に付着する。ブロッカーは、単一のサイクル中に1つの塩基のみが任意の所与の鎖に付着するように含まれる。フローセルは励起光に曝露され、ラベルを励起して蛍光を発生させる。クローン化鎖は一緒にクラスタ化されるため、いずれか1つの断片の蛍光シグナルは、そのクローン化対応物からのシグナルによって増幅され、その結果、クラスタの蛍光が、イメージャによって記録され得る。フローセルがイメージングされた後、ブロッカーが切断され、流れたヌクレオチドから洗浄され、より多くのヌクレオチドがフローセル上にフラッディングされ、サイクルが繰り返される。各シーケンシングサイクルで、1つ以上の画像が記録される。

30

【0005】

ベースコーリングアルゴリズムを記録された画像に適用して、各クラスタからの連続したシグナルを「読み取り」、光シグナルを各断片に付加されたヌクレオチド塩基配列の識別に変換する。正確なベースコーリングは、連続したシグナルが正しい断片に帰属することを確実にするために、クラスタ中心の正確な識別を必要とする。

40

【発明の概要】

【0006】

システム、装置、製品、方法及び／若しくはコンピュータプログラム製品の態様、並びに／又はそれらの組み合わせ及びサブ組み合わせが本明細書に提供され、これらは、イメージャの解像度をその物理的解像度限界を超えて計算上改善し、かつ／あるいは画像内のより精度の高いソース位置を提供する。

【0007】

そのようなものの特定の用途として、フローセル内のベースコーリング位置のセットを識別するための方法及びシステムの態様が説明される。これらは、各シーケンシングサイ

50

クルの後にフローセル画像をキャプチャすることと、フローセル画像のうちの少なくとも1つにおいて候補クラスタ中心を識別することと、を含む。強度は、各候補クラスタ中心について判定される。純度は、強度に基づいて各候補クラスタ中心について判定される。いくつかの態様において、強度及び／又は純度は、サブピクセルレベルで判定される。距離閾値内の周囲の候補クラスタ中心の純度よりも高い純度を有する各候補クラスタ中心は、ベースコーリング位置のセットに追加される。ベースコーリング位置のセットは、本明細書においてテンプレートと称する場合がある。

【0008】

いくつかの態様において、候補クラスタ中心を識別することは、フローセル画像の各ピクセルを候補クラスタ中心として標識することを含む。

【0009】

いくつかの態様において、候補クラスタ中心を識別することは、スポット探索アルゴリズムを使用して潜在的なクラスタ中心位置のセットを検出することと、次いで、各潜在的なクラスタ中心位置の周りの追加のクラスタ位置を識別することと、を含む。

【0010】

本開示の更なる態様、特徴、及び利点、並びに本開示の様々な態様の構造及び操作は、添付の図面を参照して以下に詳細に記載される。

【0011】

本明細書に組み込まれ、本明細書の一部を形成する添付の図面は、本開示の態様を例示し、説明とともに本開示の原理を説明すること、及び技術分野（複数可）の当業者が態様を作製及び使用することを可能にすることに更に役立つ。

【図面の簡単な説明】

【0012】

【図1】いくつかの態様による、フローセル上のクラスタ位置を識別するためのシステムのブロック図を示す。

【図2】いくつかの態様による、候補クラスタ中心中心を有する例示的なフローセル画像を示す。

【図3】いくつかの態様による、ベースコーリングを実行する位置を識別するための方法を示すフローチャートである。

【図4】いくつかの態様による、本開示の様々な態様を実行するために使用され得るコンピュータのブロック図を示す。

【図5】いくつかの態様による、ベースコーリングを実行する位置を識別するための方法を示すフローチャートである。

【図6】パドロックプローブの例示的な態様を示す概略図である。

【図7】第1及び第2の標的RNA分子から（それぞれ）第1及び第2のcDNAを生成すること、第1及び第2のパドロックプローブを（それぞれ）第1及び第2のcDNA分子にハイブリダイズして、（それぞれ）第1及び第2の環状化パドロックプローブを生成すること、を含む、細胞内で環状化パドロックプローブを生成するためのワークフローを示す概略図である。

【図8】第1及び第2の共有結合性閉鎖環状分子を（それぞれ）使用してローリングサークル増幅を行うことによって第1及び第2のコンカテマーを生成することを含む、細胞内でのローリングサークル及びシーケンシングワークフローを示す概略図である。

【図9】細胞内で生成されるコンカテマーをシーケンシングするための例示的なワークフローを示す概略図である。

【図10】細胞内で生成されるコンカテマーをシーケンシングするための例示的なワークフローを示す概略図である。

【図11】細胞内で生成されるコンカテマーをシーケンシングするための例示的なワークフローを示す概略図である。

【図12】細胞内で生成されるコンカテマーをシーケンシングするための例示的なワークフローを示す概略図である。

10

20

30

40

50

【図 1 3】第 1 及び第 2 の標的 RNA 分子から（それぞれ）第 1 及び第 2 の cDNA を生成すること、第 1 及び第 2 のパドロックプローブを（それぞれ）第 1 及び第 2 の cDNA 分子にハイブリダイズして、（それぞれ）第 1 及び第 2 の環状化パドロックプローブを生成すること、を含む、環状化パドロックプローブを生成するためのワークフローを示す概略図である。

【図 1 4】第 1 及び第 2 の共有結合性閉鎖環状分子を（それぞれ）使用してローリングサークル増幅を行うことによって第 1 及び第 2 のコンカテマーを生成することを含む、ローリングサークル及びシーケンシングワークフローを示す概略図である。

【図 1 5】ガラス基材及び親水性コーティングの交互の層を含む例示的な低結合支持体の概略図であり、親水性コーティングの交互の層は、ガラスに共有結合的に又は非共有結合的に接着しており、化学反応性官能基を更に含み、化学反応性官能基は、オリゴヌクレオチドプライマー（例えば、キャプチャオリゴヌクレオチド）のための付着部位として機能する。

10

【図 1 6】多価分子の様々な例示的な構成の概略図である。

【図 1 7】複数のヌクレオチドアームに付着した一般的なコアを含む例示的な多価分子の概略図である。

【図 1 8】複数のヌクレオチドアームに付着したデンドリマーコアを含む例示的な多価分子の概略図である。

【図 1 9】複数のヌクレオチドアームに付着したコアを含む例示的な多価分子の概略図を示し、ヌクレオチドアームは、ビオチン、スパーサー、リンカー、及びヌクレオチド単位を含む。

20

【図 2 0】コア付着部分、スパーサー、リンカー、及びヌクレオチド単位を含む例示的なヌクレオチドアームの概略図である。

【図 2 1】例示的なスパーサーの化学構造（上）、並びに 11 原子リンカー、16 原子リンカー、23 原子リンカー、及び N3 リンカーを含む様々な例示的なリンカーの化学構造（下）を示す。

【図 2 2】リンカー 1～9 を含む様々な例示的なリンカーの化学構造を示す。

【図 2 3 A】ヌクレオチド単位に連結された／付着した様々な例示的なリンカーの化学構造を示す。

【図 2 3 B】ヌクレオチド単位に連結された／付着した様々な例示的なリンカーの化学構造を示す。

30

【図 2 3 C】ヌクレオチド単位に連結された／付着した様々な例示的なリンカーの化学構造を示す。

【図 2 3 D】ヌクレオチド単位に連結された／付着した様々な例示的なリンカーの化学構造を示す。

【図 2 4】例示的なヌクレオチドアームの化学構造を示す。

【図 2 5】グアニン四重鎖（例えば、G 四重鎖）の概略図である。

【図 2 6】例示的な分子内 G 四重鎖構造の概略図である。

【図 2 7 A】いくつかの態様による、細胞試料を固定化するための複数のタイルを有する例示的な支持体を示す。

40

【図 2 7 B】いくつかの態様による、軸方向位置のうちのいくつかにおけるベースコーリング位置及び 2 つの所定の軸方向位置間のベースコーリング位置を有する複数の軸方向位置におけるフローセル画像を示す概略図である。

【図 2 8】いくつかの態様による、細胞特徴及びインサイチュシーケンシング結果のデータを含む例示的なデータファイルと、インサイチュシーケンシング結果と細胞特徴とを組み合わせた再構築された画像とを示す。

【図 2 9 A】主な構造要素を有する細胞（上）の概略図である。各構造要素の対応する ASCI II マーカーが示されている。そのような ASCI II マーカーは、いくつかの態様に従って、例えば、インサイチュシーケンシングからの 3D 細胞又は組織形態がデータファイルから再構築され得るように、本明細書に開示されるデータファイルで使用さ

50

れ得る。図 29 A はまた、いくつかの態様による、データファイルで使用され得るデコーダ又はインジケータの例示的なリスト（下）を示す。

【図 29 B】いくつかの態様による、図 29 A に列挙されるインジケータを利用する例示的なデータファイルを示す。

【0013】

図面において、同様の参照番号は、概して、同じ又は同様の要素を示す。加えて、概して、参照番号の左端の数字（複数可）は、参照番号が最初に表示される図面を識別する。

【発明を実施するための形態】

【0014】

システム、装置、製品、方法及び／若しくはコンピュータプログラム製品の態様、並びに／又はそれらの組み合わせ及びサブ組み合わせが本明細書に提供され、これらは、イメージの解像度をその物理的解像度限界を超えて計算上改善し、かつ／あるいは画像内のより精度の高いソース位置を提供する。本明細書に記載される画像処理技法は、次世代シーケンシングにおけるベースコーリングに特に有用であり、ベースコーリングは、これらの技法の適用を説明するための本明細書の主要な例として使用される。しかしながら、そのような画像分析技術はまた、スポット探索及び／又は CCD イメージングが使用される他の用途において特に有用であり得る。例えば、知覚された光シグナルの実際の中心（例えば、ソース位置）を識別することは、位置検出及び追跡、天文撮影、ヒートマッピングなどの多くの他の分野において有用性を有する。加えて、本明細書に記載されるような技術は、イメージの物理的解像度限界に達すると、解像度を計算上上昇させることから利益を得る任意の他の用途において有用であり得る。

【0015】

DNA シーケンシングにおいて、クラスタ又はポロニー（しばしばビーズ上に形成される）の中心を識別することは、一次分析と呼ばれることがある。一次分析は、フローセルのためのテンプレートの形成を伴う。テンプレートは、共通座標系における全ての検出されたクラスタの推定位置を含む。テンプレートは、シーケンシングプロセスの最初のいくつかのフローにおける全ての画像内のクラスタ位置を識別することによって生成される。画像は、共通座標系を提供するために全ての画像にわたって位置合わせされ得る。異なる画像からのクラスタ位置は、座標系における近接度に基づいてマージされ得る。テンプレートが生成されると、全ての更なる画像がそれに対して位置合わせされ、シーケンシングが、テンプレート内のクラスタ位置に基づいて実行される。

【0016】

画像内のクラスタ中心を識別するための様々なアルゴリズムが存在する。これらの既存のアルゴリズムには、いくつかの欠点がある。上で考察したように、クラスタ中心は、それらが互いに近い場合、マージされたように見えることがある。近接は、精度の問題又は位置合わせの問題に起因する場合がある。したがって、異なるクラスタが、単一のクラスタとして扱われ、配列の識別が不適切になるか、あるいは配列が欠落する場合がある。

【0017】

加えて、アルゴリズムでは、テンプレートのクラスタ位置を識別するために、いくつかの画像にわたってクラスタを見つけることが必要となる場合がある。これにより、過度の処理時間が必要となる場合がある。

【0018】

図 1 は、一態様による、フローセル上のクラスタ位置を識別するためのシステム 100 のブロック図を示す。システム 100 は、フローセル 112、シーケンサ 114、イメージャ 116、データストレージ 122、及びユーザインターフェース 124 を含み得るシーケンシングシステム 110 を有する。シーケンシングシステム 110 は、クラウド 130 に接続され得る。シーケンシングシステム 110 は、専用プロセッサ 118、フィールドプログラマブルゲートアレイ（複数可）（FPGA（複数可））120、及びコンピュータシステム 126 のうちの 1 つ以上を含み得る。

【0019】

10

20

30

40

50

いくつかの態様において、フローセル 112 は、DNA 断片をキャプチャし、フローセル上でベースコーリングのための DNA 配列を形成するように構成される。シーケンサ 114 は、ヌクレオチド混合物をフローセル 112 上に流し、流すステップの間にヌクレオチドからブロッカーを切断し、フローセル 112 上で DNA 配列を形成するための他のステップを実行するように構成され得る。ヌクレオチドは、ヌクレオチドの種類を示す波長で光又はエネルギーを放出する蛍光要素を有し得る。各タイプの蛍光要素は、特定のヌクレオチド塩基（例えば、A、G、C、T）に対応し得る。蛍光要素は、可視波長の光を放出し得る。

【0020】

例えば、各ヌクレオチド塩基には、色が割り当てられてもよい。アデニンは、例えば、赤色であってもよく、シトシンは、青色であってもよく、グアニンは、緑色であってもよく、チミンは、黄色であってもよい。各ヌクレオチドの蛍光要素の色又は波長は、蛍光要素によって放出される光の波長に基づいて、ヌクレオチドが互いに区別可能であるように選択され得る。

【0021】

イメージャ 116 は、各流すステップの後にフローセル 112 の画像をキャプチャするように構成され得る。一態様では、イメージャ 116 は、CMOS 又は CCD カメラなどのデジタル画像をキャプチャするように構成されたカメラである。カメラは、ヌクレオチドに結合した蛍光要素の波長で画像をキャプチャするように構成され得る。

【0022】

イメージャ 116 の解像度は、ピクセルサイズを含むフローセル画像における詳細のレベルを制御する。既存のシステムでは、スポット探索アルゴリズムがクラスタ中心を識別する精度を制御するため、この解像度は非常に重要である。スポット探索の精度を高める 1 つの方法は、イメージャ 116 の解像度を改善すること、又はイメージャ 116 によって撮影された画像上で実行される処理を改善することである。本明細書に記載される方法は、スポット探索アルゴリズムによって検出されたピクセル以外のピクセル内のクラスタ中心を検出し得る。これらの方法は、イメージャ 116 の解像度を上昇させることなく、クラスタ中心の検出精度の改善を可能にする。イメージャの解像度は、同等の性能を有する既存のシステムよりも小さくてもよく、これは、シーケンシングシステム 110 のコストを削減し得る。

【0023】

一態様では、フローセルの画像は、グループでキャプチャされ得、グループ内の各画像は、蛍光要素のうちの 1 つだけに一致するか、又はそれを含む波長又はスペクトルで撮影される。別の態様では、画像は、蛍光要素の全ての波長をキャプチャする単一の画像としてキャプチャされ得る。

【0024】

シーケンシングシステム 100 は、フローセル画像に基づいてフローセル 112 上のクラスタ位置を識別するように構成され得る。クラスタを識別するための処理は、専用プロセッサ 118、FPGA（複数可）120、コンピューティングシステム 126、又はそれらの組み合わせによって実行され得る。クラスタ位置を識別又は判定することは、本明細書により具体的に記載されるクラスタ検出方法と組み合わせて、従来のクラスタ検出を実行することを伴い得る。

【0025】

汎用プロセッサは、Windows（商標）又は Linux（商標）などのオペレーティングシステム内で様々なプログラムを実行するためのインターフェースを提供する。そのようなオペレーティングシステムは、典型的には、ユーザに大きな柔軟性を提供する。

【0026】

いくつかの態様において、専用プロセッサ 118 は、本明細書に記載されるクラスタ検出方法のステップを実行するように構成され得る。これらは、汎用プロセッサではなく、それらのステップを実行するための特定のハードウェア又は命令を有するカスタムプロセ

10

20

30

40

50

ッサであり得る。専用プロセッサは、オペレーティングシステムなしで特定のソフトウェアを直接実行する。オペレーティングシステムの欠如は、プロセッサが実行し得るものの柔軟性を犠牲にして、オーバーヘッドを低減する。専用プロセッサは、汎用プロセッサ上で実行されるソフトウェアよりも効率的に動作するように設計され得るカスタムプログラミング言語を利用し得る。これは、ステップが実行される速度を増加させ、リアルタイム処理を可能にし得る。

【0027】

いくつかの態様において、FPGA（複数可）120は、本明細書に記載されるクラスタ検出方法のステップを実行するように構成され得る。FPGAは、特定のタスクのみを実行するハードウェアとしてプログラムされる。特別なプログラミング言語を使用して、ソフトウェアステップをハードウェア構成要素に変換してもよい。FPGAがプログラムされると、ハードウェアはソフトウェアを実行せずに提供されたデジタルデータを直接処理する。代わりに、FPGAは、デジタルデータを処理するために論理ゲート及びレジスタを使用する。オペレーティングシステムに必要なオーバーヘッドがないため、FPGAは、一般に、汎用プロセッサよりも速くデータを処理する。専用プロセッサと同様に、これは柔軟性を犠牲にしている。

【0028】

また、ソフトウェアオーバーヘッドがないと、FPGAは専用プロセッサよりも速く動作することができるが、これは、実行される正確な処理、並びに特定のFPGA及び専用プロセッサに依存する。

【0029】

FPGA（複数可）120のグループは、ステップを並列して実行するように構成され得る。例えば、いくつかのFPGA（複数可）120は、画像、画像のセット、又は1つ以上の画像内のクラスタ位置について処理ステップを実行するように構成され得る。各FPGA（複数可）120は、処理ステップの独自の部分を同時に実行し得、データを処理するために必要な時間を低減する。これにより、処理ステップをリアルタイムで完了することができる。FPGAの使用の更なる考察を以下に提供する。

【0030】

データは、受信されると処理され得るので、リアルタイムで処理ステップを実行することにより、システムが使用するメモリがより少なくなり得る。これは、従来のシステムよりも改善され、処理され得る前にデータを格納する必要があるが、これは、より多くのメモリを必要とし得るか、又はクラウド130内に位置するコンピュータシステムにアクセスし得る。

【0031】

いくつかの態様において、データストレージ122は、クラスタ位置の識別において使用される情報を格納するために使用される。この情報は、画像自体、又はイメージャ116によってキャプチャされた画像から導出される情報を含み得る。ベースコーリングから判定されたDNA配列は、データストレージ122内に格納され得る。クラスタ位置を識別するパラメータもまた、データストレージ122内に格納され得る。

【0032】

ユーザインターフェース124は、ユーザによって、シーケンシングシステムを操作するために、又はデータストレージ122若しくはコンピュータシステム126に格納されたデータにアクセスするために使用され得る。

【0033】

コンピュータシステム126は、シーケンシングシステムの一般的な動作を制御し得、ユーザインターフェース124にカップリングされ得る。また、クラスタ位置の識別及びベースコーリングのステップを実行し得る。いくつかの態様において、コンピュータシステム126は、図4でより詳細に説明されるように、コンピュータシステム400である。コンピュータシステム126は、構成情報、シーケンシングシステム110を動作させるための命令、又はユーザ情報など、シーケンシングシステム110の動作に関する情報

を格納し得る。コンピュータシステム 126 は、シーケンシングシステム 110 とクラウド 130 との間で情報を渡すように構成され得る。

【0034】

上で考察したように、シーケンシングシステム 110 は、専用プロセッサ 118、FPGA（複数可）120、又はコンピュータシステム 126 を有し得る。シーケンシングシステムは、上記の必要な処理を達成するために、これらの要素のうちの 1 つ、2 つ、又は全てを使用し得る。いくつかの態様において、これらの要素が一緒に存在するとき、処理タスクはそれらの間で分割される。例えば、FPGA（複数可）120 を使用して、本明細書に記載されるクラスタ中心検出方法を実行してもよく、一方、コンピュータシステム 126 は、シーケンシングシステム 110 のための他の処理機能を実行してもよい。当業者は、これらの要素の様々な組み合わせが、処理の効率及び速度と処理要素のコストとのバランスをとる様々なシステム態様を可能にすることを理解するであろう。

10

【0035】

クラウド 130 は、シーケンシングシステム 110 から分離されたネットワーク、リモートストレージ、又はいくつかの他のリモートコンピューティングシステムであり得る。クラウド 130 への接続は、シーケンシングシステム 110 の外部に格納されたデータへのアクセスを可能にし得るか、又はシーケンシングシステム 110 内のソフトウェアの更新を可能にし得る。

【0036】

図 3 は、ベースコーリングを実行する実際のクラスタ中心位置を識別するための方法 300 を示すフローチャートである。フローセル画像内のクラスタ中心は、物理フローセル上のクローン化クラスタの位置に対応する画像内の位置である。クラスタ中心で検出された光シグナルの波長は、その位置でフローセル上の断片に付加されたヌクレオチド塩基に相関する。DNA 配列が正しく判定されるためには、連続して検出される光シグナルは、一貫して正しい DNA 断片に帰属するものでなければならない。したがって、クラスタ中心の位置を正確に識別することは、その断片のベースコーリング精度を改善する。いくつかの態様において、実際のクラスタ中心が識別されると、そのような位置は、同じフローセルを使用して後続のベースコールサイクルで使用するためにテンプレート上にマッピングされ得る。方法 300 は、専用プロセッサ 118、FPGA（複数可）120、又はコンピュータシステム 126 によって実行され得る。

20

30

【0037】

ステップ 310 では、フローセル画像がキャプチャされる。上で考察したように、フローセル画像は、イメージャ 116 によってキャプチャされ得る。ステップ 310 は、以下のステップによって処理される画像を一度に 1 つキャプチャすることを伴い得るか、又は同時に処理するために画像のセットをキャプチャすることを伴い得る。画像のセットがキャプチャされる例では、画像のセット内の各画像は、異なる検出された波長に対応し得る。例えば、ヌクレオチドに結びついた上記色の表記を考慮すると、画像のセットは、各々が赤、青、緑、及び黄色の波長のそれぞれの 1 つでキャプチャされたシグナルに対応する 4 つの画像を含み得る。単一の画像がキャプチャされる例では、その画像は、対象となる全ての検出された波長を含み得る。各画像又は画像のセットは、フローセル上の単一の流すステップのためにキャプチャされ得る。いくつかの態様において、フローセル画像は、座標系を参照してキャプチャされる。

40

【0038】

支持体（例えば、フローセル）上に固定化された試料（複数可）に応じて、フローセル画像は、フローセル画像の画像平面に直交する軸方向軸に沿った単一又は複数の軸方向位置を含み得る。特に、三次元試料、例えば、細胞、組織、又は他のインサイチュ試料について、フローセル画像は、試料（複数可）全体を 3D でカバーするために、複数の z レベル（すなわち、軸方向位置）を含み得る。軸方向軸は、本明細書に開示される光学システムの対物レンズから支持体、例えば、フローセルまで延在することができる。軸方向軸は、フローセル画像の画像平面に直交し得る。フローセル画像の各軸方向位置（複数可）は

50

、隣接する軸方向位置から所定の距離、例えば、約 $0.1 \mu\text{m}$ ～ 約 $15 \mu\text{m}$ だけ隔てられ得る。フローセル画像の各軸方向位置は、隣接するレベル（複数可）から $0.5 \mu\text{m}$ ～ $10 \mu\text{m}$ だけ隔てられ得る。各軸方向位置において、フローセル画像は、1つ以上のシーケンシングサイクル及び／又は1つ以上のチャンネルから取得することができる。各フローセル画像は、その視野内に、フローセルの1つ以上のタイル又はサブタイルの少なくとも一部を含み得る。図27Aは、複数のタイル2710を有するフローセル2712の一部分を示す。画像平面は、x軸及びy軸によって画定される。そして、軸方向軸（すなわち、z軸）は、x-y平面に直交する。フローセル画像、試料、及び軸方向軸は、図27A～27Bに示されるように、直交座標系で説明されるが、任意の他の座標系を、本明細書で空間的位置及び関係を定義するために使用してもよい。他の座標系は、極座標系、円筒座標系、又は球面座標系を含むことができるが、これらに限定されない。

10

【0039】

図2は、クラスタからのシグナルが存在するフローセル画像200の概略図を示す。フローセル画像200は、ピクセル210A、210B、及び210Cなどのピクセル210で構成される。ステップ310の間に、イメージャは、例えば、フローセル上の断片に結合された蛍光要素の励起後にフローセルから受信された光シグナルを記録し、そのような断片は、断片のクローン化クラスタ内に位置する。

【0040】

ステップ320において、フローセル画像内の潜在的なクラスタ中心の位置が識別される。例えば、いくつかの態様において、ステップ310でイメージングされた光シグナルは、スポット探索アルゴリズムが潜在的なクラスタ中心のセットを出力するように、スポット探索アルゴリズムに入力され得る。いくつかの態様において、潜在的なクラスタ中心は、単一のフローセル画像（例えば、目的の全ての波長を含む1つの画像）のみを使用して識別され得る。いくつかの他の態様では、潜在的なクラスタ中心は、フローセル上の単一の流すサイクル（例えば、目的の各波長における1つの画像）からの画像のセットから識別され得る。スポット探索アルゴリズムは、将来の流すサイクルからの追加の画像が取得されるのを待つ必要がないため、単一の流すサイクルからのフローセル画像又はフローセル画像のセットのみを使用することは、処理時間の量を有利に低減する。

20

【0041】

更に他の態様では、スポット探索アルゴリズムは、2つ以上のシーケンシングサイクルからの画像に適用され得、潜在的なクラスタ中心は、それらの画像のいくつかの組み合わせを使用して探索され得る。例えば、潜在的なクラスタ中心は、2つ以上のシーケンシングサイクルからの画像内の同じ位置を占有するスポットの存在によって識別され得る。

30

【0042】

スポット探索アルゴリズムによって識別された潜在的なクラスタ中心位置は、潜在的なクラスタ中心位置220A、220B、及び220Cなど、図2に「X」で示されている。フローセルへの断片付着がランダムであるという性質により、クローン化クラスタのうちのいくつかは互いに近くてもよく、他のクラスタは更に離隔していてもよく、又は単独であってもよい。その結果、図2のいくつかの「X」は、他のものよりもより互いに近くに配置される。加えて、いくつかのピクセルは、潜在的なクラスタ中心を含むものとして識別され得るが、他のピクセルはそのようには識別されない。例えば、ピクセル210Aは、潜在的なクラスタ中心位置225Aを含むものとして識別され得るが、ピクセル210Bは、潜在的なクラスタ中心位置220を含むものとして最初に識別されない。

40

【0043】

いくつかの態様において、スポット探索アルゴリズムは、ピクセル210にわたって補間することによって、サブピクセル解像度で潜在的なクラスタ中心位置220を識別し得る。例えば、潜在的なクラスタ中心位置225Aは、ピクセル210Aの中心ではなく、ピクセル210Aの右下側に位置する。他の潜在的なクラスタ中心位置は、それらのそれぞれのピクセル210の異なる領域に位置し得る。例えば、潜在的なクラスタ中心位置225Bは、ピクセルの右上に位置し、潜在的なクラスタ中心位置225Cは、ピクセルの

50

左上に位置する。補間は、補間関数によって実行され得る。

【0044】

いくつかの態様において、補間関数は、当業者に既知のガウス補間関数である。サブピクセル解像度は、潜在的なクラスタ位置が、例えば、10分の1のピクセル解像度で判定されることを可能にし得るが、他の解像度も考慮される。態様では、例えば、解像度は、4分の1のピクセル解像度、2分の1のピクセル解像度などであり得る。補間機能は、この解像度を判定するように構成され得る。

【0045】

補間関数は、1つ以上のピクセル210における光の強度に適合するように使用され得る。この補間は、サブピクセル位置の識別を可能にする。補間関数は、潜在的なクラスタ中心位置220を含むピクセル210のセットにわたって適用され得る。一態様では、補間関数は、中に潜在的なクラスタ中心位置220を有するピクセル210、及びそのピクセル210の縁に接触する周囲のピクセル210に適合され得る。

【0046】

いくつかの態様において、補間関数は、画像内のいくつかの点で判定され得る。解像度は、各ピクセル210に位置する点の数を判定する。例えば、解像度がピクセルの10分の1である場合、ピクセルエッジに垂直な線に沿って、ピクセル210全体で9つの点及び各エッジで1つの点が計算され、ピクセル210を10個の部分に分割する。いくつかの態様において、補間関数は、各点で計算され、各点での補間関数とピクセル強度との差が判定される。補間関数の中心は、補間関数と各ピクセル210の強度との差を最小にするようにシフトされる。このサブピクセル補間は、システムがより低解像度のイメージャでより高い解像度を達成することを可能にし、システムのコスト及び／又は複雑さを低減する。

【0047】

いくつかの態様において、補間は、5×5グリッド上で実行され得る。グリッドは、潜在的なクラスタ中心位置を有するピクセル内のピクセル中心210を中心にしてもよい。

【0048】

ステップ320のいくつかの態様は、潜在的なクラスタ中心位置を識別するためにスポット探索アルゴリズムを使用するが、ステップ320のいくつかの他の態様は、最初に、キャプチャされたフローセル画像内の各ピクセルを潜在的なクラスタ中心位置として識別する。例えば、図2において、全てのピクセル210は、潜在的なクラスタ中心位置220として識別され得る。このアプローチは、方法300を実施するために必要とされる処理の種類を単純化し得る、スポット探索アルゴリズムの必要性を排除する。このアプローチは、各潜在的なクラスタ中心位置が並列に処理され得るため、大規模な並列処理が利用可能である場合に有利であり得る。これは、増加した専用プロセッサ118又はFPGA（複数可）120などの追加のハードウェアのコストが潜在的にかかるが、処理時間を低減し得る。いくつかの態様において、次いで、補間関数は、フローセル画像の全体にわたってサブピクセル解像度で強度を識別するために上記に記載されるように使用され得る。

【0049】

上で考察したように、クラスタ中心は、物理フローセル上のクローン化クラスタの位置に対応する、ピクセル210などの画像内の位置を識別する。潜在的なクラスタ中心位置220は、1つ以上の波長の光がイメージャによって検出される画像内の位置である。いくつかの場合では、クラスタの物理的位置がピクセル210の1つのセットに対応するが、例えば、カメラ内の対応するセンサの飽和（「ブルーミング」とも称される）に起因して、そのクラスタからの光シグナルがその1つのセットに隣接する追加のピクセル210にオーバーフローすることが可能である。加えて、クラスタ同士が近接して配置されるとき、クラスタ自体がそうでなくても、それらのクラスタからの光シグナルが重複し得る。真のクラスタ中心を識別することは、検出されたシグナルが正しいDNA断片に帰属することを可能にし、したがって、ベースコーリングアルゴリズムの精度を向上させる。

【0050】

したがって、ステップ 3 2 5 において、追加のクラスタシグナル位置 2 2 5 は、各潜在的なクラスタ中心位置 2 2 0 の周りで識別される。これらは、図 2 の「+」、例えば、2 2 5 A、2 2 5 B、及び 2 2 5 C で示される。これらの追加のクラスタシグナル位置 2 2 5 は、潜在的なクラスタ中心位置 2 2 0 として既に識別されている位置の代わりに、又はそれに加えて、クラスタ中心を構成し得るフローセル内の他の位置に対応する。

【0051】

いくつかの態様において、追加のクラスタ位置 2 2 5 は、潜在的なクラスタ中心位置 2 2 0 の周りに配置される。いくつかの態様において、これらの追加のクラスタ位置 2 2 5 は、スポット探索アルゴリズムによって最初に識別されるのではなく、スポット探索アルゴリズムによって識別される各潜在的なクラスタ中心位置 2 2 0 の周りにパターンで配置される。追加のクラスタ位置 2 2 5 は、実際の検出されたクラスタ中心を表すのではなく、そうでなければ検出されない可能性のあるクラスタ中心をチェックするための潜在的な位置を表す。そのようなクラスタ中心は、近接クラスタ中心からのシグナル間の混合、スポット探索アルゴリズムのエラー、又は他の影響のために検出されない場合がある。

【0052】

一例として、追加のクラスタ位置 2 2 5 A は、潜在的なクラスタ中心位置 2 2 0 A の位置に基づいて、ピクセル 2 1 0 A に配置され得る。この文脈において、これは、追加のクラスタ位置 2 2 5 A が潜在的なクラスタ中心位置 2 2 0 A から離隔したピクセルの幅に配置されることを意味し得る。他の追加のクラスタ位置 2 2 5 も、潜在的なクラスタ中心位置 2 2 0 A の周りに配置される。潜在的なクラスタ中心位置 2 2 0 がフローセル画像 2 0 0 のエッジの近くにある場合など、追加のクラスタ位置 2 2 5 は、ピクセル 2 1 0 が存在しない場所に配置されないことを理解されたい。

【0053】

いくつかの態様において、追加のクラスタシグナル位置 2 2 5 は、潜在的なクラスタ中心位置 2 2 0 を中心とするグリッドに配置される。追加のクラスタ位置 2 2 5 は、ピクセル幅だけ互いから及び潜在的なクラスタ中心位置 2 2 0 から離隔していてもよい。いくつかの態様において、追加のクラスタ位置 2 2 5 及び潜在的なクラスタ中心位置 2 2 0 は、正方形のグリッドを形成する。グリッドは、5 ピクセル×5 ピクセル、9×9、15×15、又は他の寸法の面積を有し得る。

【0054】

いくつかの態様において、潜在的なクラスタ中心位置 2 2 0 は、追加のクラスタ位置 2 2 5 の対応するグリッドを重複させるのに十分に互いに近いものであり得る。これにより、同じピクセル 2 1 0 が、2 つ以上の潜在的なクラスタ中心位置 2 2 0 からの追加のクラスタ位置 2 2 5 を含み、2 つ以上の追加のクラスタ位置 2 2 5 が同じピクセル 2 1 0 に帰属するようになり得る。例えば、ピクセル 2 1 0 C は、（潜在的なクラスタ中心位置 2 2 0 B に基づいて識別された）追加のクラスタ位置 2 2 5 B、並びに（潜在的なクラスタ中心位置 2 2 0 C に基づいて識別された）追加のクラスタ位置 2 2 5 C を含む。

【0055】

いくつかの態様において、同じピクセル 2 1 0 内の追加のクラスタ位置 2 2 5 が十分に互いに近い場合、追加のクラスタ位置 2 2 5 のうちの一方が破棄され、他方が両方を表すのに使用される。2 つの追加のクラスタ位置 2 2 5 は、それらが、例えば、限定されないが、ピクセルの 10 分の 2、ピクセルの 10 分の 1、又はいくつかの他のサブピクセル距離内にある場合、そのような処理にとって十分に互いに近いと考えられ得る。

【0056】

当業者は、ステップ 3 2 0 で、全てのピクセル位置が潜在的なクラスタ中心位置 2 2 0 として識別された場合、ステップ 3 2 5 は、識別された潜在的なクラスタ中心位置に加えて考慮すべき他のピクセルがフローセル画像内に残っていないためにスキップされ得ることを理解するであろう。

【0057】

いくつかの態様において、潜在的なクラスタ中心位置 2 2 0（及び識別された場合、周

10

20

30

40

50

囲の追加のクラスタ位置 2 2 5) は、一緒に全ての候補クラスタ中心のセットを構成する。これらの候補クラスタ中心は、キャプチャされた各フローセル画像についての実際のクラスタ中心を識別するために処理され得る。

【0058】

したがって、潜在的なクラスタ中心位置 2 2 0 及び追加のクラスタ位置 2 2 5 のそれらの周囲のグリッド（すなわち、候補クラスタ中心）が識別されると、それらは、クラスタ中心の実際の位置を判定するための出発点として使用され得、これは、最初に識別された潜在的なクラスタ中心位置 2 2 0 と同じであっても、そうでなくてもよい。

【0059】

図 3 に戻り、ステップ 3 3 0 において、各キャプチャされたフローセル画像上の各候補クラスタ中心の純度値が判定される。純度値は、ヌクレオチドに結合された蛍光要素の波長及びキャプチャされたフローセル画像内のピクセルの強度に基づいて判定され得る。

【0060】

各候補クラスタ中心において、ピクセルの強度は、イメージャのスペクトル帯域幅にわたって放出されるエネルギー又は光の組み合わせである。いくつかの態様において、各ヌクレオチド塩基の蛍光スペクトル帯域幅に対応するエネルギー又は光の量が探索され得る。特定のヌクレオチド塩基に対応する各シグナルの純度は、1つのヌクレオチド塩基シグナルについてのエネルギーの量と、各他のヌクレオチド塩基シグナルについてのエネルギーの総量との比率として探索され得る（例えば、「赤」シグナルの純度は、検出された青、緑、及び黄色の波長の各々と比較して、そのピクセル又はサブピクセルについて検出された赤い波長の相対強度に基づいて判定され得る）。ピクセルの全体的な純度は、最大の比率、最小の比率、比率の平均、又は比率の中央値であり得る。次いで、計算された純度は、そのピクセル又はサブピクセルに割り当てられ得る。

【0061】

上述のように、フローセル画像のセットは、単一のシーケンシングサイクルのためにキャプチャされ得る。セット内の各画像は、異なる波長でキャプチャされ、各波長は、ヌクレオチドに結合した蛍光要素のうちの 1 つに対応する。画像のセットにわたる所与のクラスタ中心の純度は、画像のセットについての純度のセットからの最高、最低、中央値、又は平均純度であり得る。

【0062】

いくつかの態様において、候補クラスタ中心の純度は、ピクセルの波長からの 2 番目に高い強度又はエネルギーと、そのピクセルの波長からの最高の強度又はエネルギーとの比率を差し引いたものとして判定され得る。閾値は、高純度又は低純度を構成するものについて設定され得る。例えば、最高純度は、1 であってもよく、低純度のピクセルは、ゼロに近い純度値を有してもよい。閾値は、その間に設定され得る。

【0063】

いくつかの態様において、2つの強度の比率は、2番目に高い強度及び最高強度の両方にオフセットを加えることによって修正され得る。オフセットは、品質スコアの改善された精度を提供し得る。例えば、いくつかの場合において、2つの比率強度は、わずかな量の絶対最大強度（高パーセンテージの最高強度でもある）だけ異なることがある。具体的な非限定的な例として、最高強度は 1 0 であり得、最低強度は、最大の可能な強度が 1 0 0 0 であり得る。この場合の比率は 0 . 1 であり、純度は 0 . 9 となる。それ以上がなければ、これは高品質スコアとして読み取られる可能性がある。これは、例えば、最高強度の場合は 5 0 0、及び 2 番目に高い強度の場合は 4 9 0 である強度とは対照的である。この例は、ほぼ同じ絶対差を有しているが、比率は 1 に近く、純度はゼロに近い。第 1 の場合、全体的な強度が低いことはポロニーが存在しないことを示唆するので、純度は誤解を招く。第 2 の場合、純度はより正確であり、ピクセルが近くに位置する 2 つの異なるクラスタ中心からの強度又はエネルギーを表示していることを示す。

【0064】

オフセットは、そのような問題を解決するための比率の強度に加えられた値であり得る

。例えば、オフセットが最大振幅の10%である場合、上記の例では、オフセットは100であり、第1の比率は110に対して101になり、これは1にはるかに近く、2つの波長強度間の小さなデルタを正確に反映するゼロに近い純度をもたらす。第2の比率では、比率は590に対して600であり、これは依然として1に近く、やはりゼロに近い純度をもたらす。

【0065】

オフセットを組み込む別の例として、最高強度が800であり、最低強度が1である場合、比率がほぼゼロであるため、オフセットのない純度は1に近い。オフセットが再び100であると、比率は900に対して101になる。これにより、純度は1から約0.89にわずかに低下する。これは純度を低下させる可能性があるが、計算された純度は依然として高い。オフセット値は、この影響を低減するように設定してもよい。例えば、別の場合のオフセットは、10であり得る。最高強度が10であり、最低強度が1である前の例を使用すると、純度は、 $1 - 10 / 11$ 、すなわち約0.09となり、これは強度間の小さな差を正確に反映している。最高強度が600であり、最低強度が590である例では、純度は、 $1 - 600 / 610$ 、すなわち約0.016であり、これもやはり強度間の小さな差を反映している。最高強度が800であり、最低強度が1である例では、純度は、 $1 - 11 / 810$ 、すなわち約0.99であり、これは純度のはるかに小さい低下であり、強度間の大きな差を反映している。

【0066】

ステップ340では、実際のクラスタ中心は、フローセルの候補クラスタ中心について計算された純度値に基づいて識別される。実際のクラスタ中心は、ステップ320及び325で識別された候補クラスタ位置のサブセットであってもよい。

【0067】

いくつかの態様において、実際のクラスタ中心は、フローセル画像全体の各候補クラスタ中心の純度を、同じ画像内の近くの候補クラスタ中心と比較することによって識別される。いくつかの態様において、比較されている2つの候補クラスタ中心を考慮すると、より高い純度を有する候補クラスタ中心が維持される。いくつかの態様において、候補クラスタ中心は、ある特定の距離内の他の候補クラスタ中心とのみ比較される。例えば、この距離閾値は、ピクセルサイズ及び所与のヌクレオチドのクラスタのサイズに基づき得る。

【0068】

例えば、クラスタの平均サイズが4つのピクセル幅／高さである場合、距離閾値は、互いの2つのピクセル幅／高さ内の任意の候補クラスタ中心が同じクラスタに属するか、又はより高い強度を有する可能性が高い（候補クラスタ中心が実際に2つの別個のクラスタのエッジ上にあることを示す）ため、2つのピクセル幅／高さであり得る。

【0069】

純度が複数のフローセルサイクルにわたって計算されるいくつかの態様において、候補クラスタ中心が、複数のフローセル画像にわたって周囲の候補クラスタ中心よりも高い純度を一貫して有することを判定することは、位置が実際のクラスタ中心である可能性を更に高め得る。より低い純度は、候補クラスタ中心で検出されたシグナルが実際のクラスタ中心ではなく、ノイズ、他のシグナルの混合、又はいくつかの他の現象であることを示し得る。

【0070】

ステップ350では、実際のクラスタ中心は、フローセル画像上でベースコーリングを実行するために使用される。例えば、実際のクラスタ中心から検出された波長を判定してもよく、これは特定のヌクレオチド塩基（例えば、A、G、C、T）に相関する。次いで、そのヌクレオチド塩基は、実際のクラスタ中心に対応する配列に付加されたものとして記録される。

【0071】

連続したフローセル画像における実際のクラスタ中心でのシーケンシングサイクル及び蛍光波長識別の連続した反復により、フローセル上の各実際のクラスタ中心に対応するD

10

20

30

40

50

N A 断片の配列が構築され得る。

【0072】

いくつかの態様において、テンプレートは、単一のシーケンシングサイクルについて識別された実際のクラスタ中心から形成される。次に、テンプレート内の実際のクラスタ位置を使用して、後続のシーケンシングサイクルからの画像内でベースコーリングを実行する場所を識別し得る。

【0073】

異なるシーケンシングサイクルでキャプチャされたフローセル画像は、シーケンシングサイクル間のフローセル又はイメージャの位置のシフトに起因して、位置合わせの問題を有し得る。したがって、いくつかの態様において、ステップ350は、連続した画像を適切に整列させるための位置合わせステップを含み得る。これにより、実際のクラスタがテンプレートの中心を各フローセル画像上の同じ位置に正確にマッピングすることが保証され、したがって、ベースコールの精度が向上する。

【0074】

テンプレートが後続の画像内の実際のクラスタ中心を識別するために使用されるいくつかの態様において、それらの後続の画像内の関連する位置に対応するデータのみが維持及び／又は処理される必要がある。処理されるデータの量のこの減少は、レガシーシステムよりも正確な結果がより迅速に得られ得るように、処理の速度及び／又は効率を増加させる。加えて、格納されるデータの量の減少は、シーケンサに必要なストレージの量を減少させ、したがって、必要なリソースの量及び／又はコストを減少させる。

【0075】

加えて、いくつかのレガシーシステムは、クラスタ位置を識別するために、異なる画像を互いに比較することを必要とする。この比較は、複数のフローセルサイクルからの画像にスポット探索アルゴリズムを適用し、次いで、画像にわたってスポット探索結果を比較することを含み得る。これは、複数のシーケンシングサイクルの各々の画像又はスポット探索結果を格納することを必要とし得る。画像を直接比較する必要がないため、方法300は、クラスタ探索の処理及び格納効率を改善し得る。代わりに、画像は、リアルタイムで処理され得、純度情報及び／又は最終的なテンプレート位置のみが格納される必要がある。

【0076】

シーケンシングサイクル及び画像作成プロセスは、多くの場合、画像を分析するスポット探索及びベースコーリングプログラムよりも高速に実行される。実行時間のこの差は、各シーケンシングサイクルの後にフローセル画像を格納すること、又は画像分析プロセスのうちのいくつか又は全てが完了するのを待っている間のシーケンシングサイクルの遅延を必要とし得る。

【0077】

FPGAを使用することで、精度を犠牲にすることなく処理速度を上昇させることができる。FPGA上で本明細書に記載されるプロセスの一部分又は全てを実施することは、プロセッサオーバーヘッドを低減し、FPGA上での並列処理を可能にし得る。例えば、各可能なクラスタ位置は、異なるFPGAによって処理されてもよく、又は可能なクラスタ位置を同時に並列に処理するように構成された単一のFPGAによって処理されてもよい。適切に実施される場合、これは、リアルタイム処理を可能にし得る。リアルタイム処理は、画像生成時に画像が処理され得るという利点を有する。FPGAは、シーケンシングシステムがフローセルを準備した時間までに、次の画像を処理する準備ができています。シーケンシングシステムは、後処理を待つ必要はなく、一次分析の全体のプロセスは、ほんのわずかの時間で完了し得る。加えて、画像全体が受信時に処理されているため、格納する必要がある情報は、ベースコーリングを実行するためのデータのみである。全ての画像を格納する代わりに、特定のピクセルの純度又は強度のみを格納する必要がある。これは、シーケンシングシステムにおけるデータ格納又は画像のリモート格納の必要性を大幅に低減する。

【0078】

いくつかの態様において、画像位置合わせ、強度抽出、純度計算、ベースコーリング、及び他のステップを含むプロセス全体が、FPGAによって実行される。これは、最もコンパクトな実施を提供し得、リアルタイム処理に必要な速度及びスループットを提供し得る。

【0079】

いくつかの態様において、処理責任は、FPGAと関連付けられたコンピュータシステムとの間などで共有される。例えば、いくつかの態様において、FPGAは、画像位置合わせ、強度抽出、及び純度計算を処理し得る。次いで、FPGAは、ベースコーリングのために情報をコンピュータシステムに渡す。このアプローチは、FPGAとコンピュータシステムとの間の通信をスケールダウンすることを含めて、FPGAとコンピュータシステムリソースとの間の負荷のバランスをとる。また、迅速なアルゴリズムチューンアップ機能を備えたベースコーリングを処理するコンピュータシステム上のソフトウェアに柔軟性を提供する。そのようなアプローチは、リアルタイム処理を提供し得る。

10

【0080】

FPGA、専用プロセッサ、及びコンピュータシステムの異なる構成が、様々なステップを実行するために使用され得ることが、当業者には認識されよう。所与の構成の選択は、フローセル画像サイズ、イメージャ解像度、処理する画像の数、所望の精度、及び必要な速度に基づいてもよい。FPGA、専用プロセッサ、及びコンピュータシステムの実装コスト及びハードウェアコストもまた、構成の選択に影響を及ぼし得る。

20

【0081】

既存の方法と本明細書記載の方法の態様との間の性能を比較する非限定的な例として、クラスタの密度が低いフローセルとクラスタの密度が高いフローセルの2つの例示的なフローセル上で試験を実行した。比較のために、識別されたクラスタ上の偽陽性の特定の平均エラー率を標的とするために各方法を使用して試験を実行した。

【0082】

低密度フローセルの場合、平均誤差率は、0.3%であった。既存の方法は、約78,000のクラスタ中心を識別したが、本明細書に記載される方法は、約98,000のクラスタ中心を識別した。高密度フローセルの場合、平均誤差率は、1.1%であった。既存の方法は、約63,000のクラスタ中心を識別したが、本明細書に記載される方法は、約170,000のクラスタを識別した。

30

【0083】

結果は、本明細書に記載される方法が既存の方法よりも多くのクラスタを効果的に識別することを示唆している。更に、同じ誤差率であっても、フローセル上のクラスタの密度が増加するとき、本明細書に開示される方法は、更に優れた性能を発揮し、ほぼ3倍の数のクラスタを識別する。いくつかの態様において、これは、既存の方法で典型的に経験される性能損失なしに、フローセルがより高い密度で流れることを可能にし得る。

【0084】

インサイチュシーケンシング

いくつかの態様において、ベースコーリング位置を判定するための方法、システム、及び媒体は、例えば、3Dボルメトリック試料のインサイチュシーケンシングにおいて利用され得る。本明細書に記載されるインサイチュシーケンシングでは、RNAは細胞試料から抽出されず、シーケンシング情報を追跡し、細胞試料の画像にマッピングし戻す必要はない。むしろ、RNAは、細胞内の標的RNAの空間位置の直接イメージングを可能にするために、細胞試料中に保持される。加えて、細胞試料内のRNAは断片化されず、標的RNAの濃縮は必要ない。標的特異的及び／又はランダム配列逆転写プライマーの使用は、ユニプレックスモード又はマルチプレックスモードのいずれかで、ポリA RNA及び非ポリA RNAの両方の検出を可能にする。

40

【0085】

いくつかの態様において、本明細書におけるインサイチュシーケンシングは、テンプレ

50

ート分子（例えば、コンカテマー分子）の同じ領域の複数のシーケンシングサイクルを繰り返し行うことを含む。反復シーケンシングサイクル（例えば、100サイクルを超えるリードと比較して30又は40サイクル未満の少数）を行うことによって、細胞試料のRNA含有量を発見することができる。長いリードシーケンシングワークフローと比較して、本明細書に記載される短い反復シーケンシングサイクルは、コストを低減し、時間を節約する低減された量のシーケンシング試薬を使用する。短い（例えば、2～40サイクル）反復シーケンシングサイクルを行うための方法は、目的の特定のRNA、変異体RNA配列、スプライスバリエント、及びそれらの存在量レベルの検出を含むが、これらに限定されない多くの用途を有する。

【0086】

本明細書におけるコンカテマーは、目的のcDNA、ユニバーサルシーケンシングプライマー結合部位、及び標的バーコード配列のタンデムリピート単位を保有する。コンカテマーは、細胞試料内でシーケンシングされ、ここで、複数のシーケンシングサイクル（例えば、2～40サイクル）が各ラウンドに対して行われ、複数ラウンドの短いリードシーケンシングが行われる。標的バーコード及びcDNA領域の全長は、シーケンシングされない。代わりに、標的バーコード領域の少なくとも一部分が反復してシーケンシングされる。いくつかの態様において、cDNA領域をシーケンシングする必要はない。いくつかの態様において、標的バーコード及びcDNA領域の一部分は、反復してシーケンシングされる。いくつかの態様において、cDNA領域の全長をシーケンシングする必要はない。いくつかの態様において、シーケンシングリードをアセンブルする必要も、目的のcDNAの全長配列を得る必要もない。本明細書の短いシーケンシングリードから得られる冗長なシーケンシング情報は、コンカテマーの相補鎖をシーケンシングする必要性を有利に排除し得る。したがって、ペアワイズシーケンシングは必要ない。

【0087】

加えて、いくつかの態様において、コンカテマーにおけるcDNA領域の短い部分は、同じ開始位置から少なくとも1回（例えば、反復シーケンシング）再シーケンシングされて、参照配列にアラインメントされ得る重複シーケンシングリードを生成する。例えば、コンカテマー分子の同じ部分は、少なくとも2回、3回、4回、5回、又は最大50回シーケンシングすることができる。開始シーケンシング部位は、コンカテマーの任意の位置であり得、コンカテマー内の選択された位置にアニーリングするように設計されたシーケンシングプライマーによって決定される。短い反復シーケンシングリードは、cDNA領域における個々の塩基についてのシーケンシング情報の冗長性を増加させる。コンカテマーテンプレート分子の1つの鎖を反復してシーケンシングすることにより、細胞試料中の標的RNAの存在を明らかにするのに十分な塩基カバレッジが得られ、その結果、相補鎖のペアワイズシーケンシングが不要となる。

【0088】

コンカテマーテンプレート分子は、増加したシーケンシング深度のために複数の使用可能なシーケンシングリードを生成するために使用することができる、同じコンカテマー分子に沿った複数のシーケンシングプライマー結合部位を含み得る。一緒に、コンカテマーテンプレートの一方の鎖を反復してシーケンシングすることにより、ワンコピーテンプレート分子のシーケンシングと比較して、シーケンシング塩基カバレッジ及びシーケンシング深度が増加し得る。

【0089】

本明細書に記載されるインサイチュシーケンシングは、ユニプレックスモード又はマルチプレックスモードで行われ得る。異なる逆転写プライマー、異なる標的特異的パドロックプローブ、及びユニバーサルシーケンシングプライマーを使用して、2つ以上の異なる標的RNAを細胞試料内で同時に検出及びイメージングすることができる。例えば、細胞試料中のハウスキーピングRNA及び少なくとも1つの標的RNAの存在は、同時に検出され、本明細書に記載される短い反復リードシーケンシング方法のうちのいずれかを使用してイメージングされ得る。

10

20

30

40

50

【0090】

3Dベースコーリングテンプレート

いくつかの態様において、インサイチュシーケンシングで細胞試料（複数可）のベースコーリングテンプレートを生成するための方法が本明細書に開示される。ベースコーリングテンプレートは、単一の2D平面、複数の2D平面、又は3D空間からのベースコーリング位置を含み得る。ベースコーリングテンプレートは、単一の2D平面、複数の2D平面、又は3D空間内のフローセル画像を使用して生成され得る。ベースコーリングテンプレートは、単一の2D平面、複数の2D平面、又は3D空間内で取得されたフローセル画像を共通座標系に対して位置合わせするために使用され得る。

【0091】

従来の2Dフローセル画像は、平坦化された試料からのクラスタ又はポロニーを示すことができ、ベースコーリングは、それらの対応する画像強度を使用して実行することができる。光学システムは、クラスタ又はポロニーが合焦している間にそれらをイメージングするように調整することができる（本明細書における「合焦」は、対応するフローセル画像（複数可）を取得するために、実質的に合焦及び／又は被写界深度内を含み得る）。しかしながら、細胞又は組織などのインサイチュ試料は、単一の2D画像内で合焦することができない軸方向又はz方向に沿った厚さを有することができる。異なる軸方向位置でクラスタ又はポロニーをカバーするために、複数の2Dフローセル画像のスタックが必要とされ得る。フローセル画像のスタックには、焦点外ポロニー及び細胞構成要素からのバックグラウンドシグナルからなどの干渉が発生し得る。例えば、第1の軸方向位置に位置するポロニーは、第1のフローセル画像に現れることができ、また、ポロニーが焦点外である第1のフローセル画像の隣接する軸方向位置で撮影された第2のフローセル画像にシグナルのプロブを生成することができる。シグナルのプロブは、第2のフローセル画像内の同じx-y位置において、又はその近傍においてポロニーの強度と干渉し得、したがって、インサイチュ試料のベースコーリングの精度及び信頼性を低下させる。更に、同じポロニーは、2つの隣接する軸方向位置に対応する両方のポロニーマップに含まれ得、それによって不正確なベースコーリング結果を引き起こす。したがって、3Dベースコーリングに使用することができるように、3Dボルメトリック試料からポロニー又はクラスタのための正確で信頼性の高いベースコーリングテンプレートを生成する必要がある。

【0092】

本明細書に開示される技術は、有利には、インサイチュシーケンシング分析のための3Dベースコーリングテンプレート（すなわち、ポロニーマップ）を生成する。フローセル画像におけるポロニー又はクラスタの特性に基づいて、本明細書の技術は、フローセル画像における、及び最終的には3Dポロニーマップにおける、合焦ポロニー又はクラスタに影響を及ぼすことなく、焦点外ポロニー及びバックグラウンドオブジェクトを効率的に除去する。3Dポロニーマップは、ベースコーリングのポロニー強度を抽出するために使用することができる。単一の平坦化された2Dポロニーマップと比較して、3Dポロニーマップは、2Dポロニーマップ内で減衰され得るか、又は完全に除去され得るポロニー及びクラスタの情報を有利に保持する。いくつかの態様において、本明細書の3Dポロニーマップは、フローセル画像を生成するための所定の軸方向位置でのみベースコーリング位置を生成することに限定されない。3Dポロニーマップは、ベースコーリング位置がフローセル画像のための2つの隣接する軸方向位置の間にあることができるように、軸方向軸に沿ったサブピクセル解像度を含み得る。3Dポロニーマップは、いくつかの初期のシーケンシングサイクルで生成され、その後のシーケンシングサイクルで使用されて、新しいポロニーマップを再計算する追加の計算負荷及び新しいポロニーマップを格納するための格納スペースを最小限に抑えることができる。いくつかの態様において、本明細書に開示される技術は、1つ以上の細胞特徴を示す画像へのポロニー又はクラスタの正確かつ効率的な位置合わせのためにバックグラウンド情報を保持する画像を有利に利用する。

【0093】

図5は、ベースコーリングテンプレートに含まれるベースコーリング位置を識別するこ

とによって、インサイチュシーケンシングのためのベースコーリングテンプレートを生成するための方法500の例示的な態様のフローチャートを示す。方法500は、本明細書に開示される操作のうちのいくつか又は全てを含むことができる。操作は、本明細書に記載される順序で実行され得るが、これに限定されない。

【0094】

方法500は、本明細書に開示される1つ以上のプロセッサ（例えば、図4の404）によって実行することができる。いくつかの態様において、プロセッサは、プロセッシングユニット、集積回路、又はそれらの組み合わせのうちの1つ以上を含むことができる。例えば、プロセッシングユニットは、セントラルプロセッシングユニット（CPU）及び／又はグラフィックスプロセッシングユニット（GPU）を含むことができる。集積回路は、フィールドプログラマブルゲートアレイ（FPGA）などのチップを含むことができる。いくつかの態様において、プロセッサは、コンピューティングシステム400を含み得る。

10

【0095】

いくつかの態様において、方法500におけるいくつか又は全ての操作は、FPGA（複数可）によって実行され得る。いくつかの操作がFPGA（複数可）によって実行される態様では、FPGA（複数可）によって実行される操作の後のデータは、FPGA（複数可）によってCPU（複数可）に通信されることが可能であり、その結果、CPU（複数可）は、そのようなデータを使用して方法500における後続の操作（複数可）を実行することができる。同様に、データはまた、FPGA（複数可）による処理のために、CPU（複数可）からFPGA（複数可）に伝達され得る。いくつかの態様において、方法500における全ての操作は、CPU（複数可）によって実行され得る。あるいは、CPU（複数可）によって実行される操作は、専用プロセッサ、又はGPU（複数可）などの他のプロセッサによって実行され得る。いくつかの態様において、方法500における全ての操作は、FPGA（複数可）によって実行され得る。

20

【0096】

いくつかの態様において、方法500のうちのいくつか又は全ての操作は、シーケンシングランにおけるシーケンシングサイクルNの間又はその前に実行される。ベースコーリングテンプレート、例えば、ポロニーマップは、サイクル1～Nのうちのいくつか又は全てから得られた情報を使用して生成され得る。そのようなサイクル内の1つ以上のチャンネルからのポロニー又はクラスタは、参照座標系内のテンプレートに含まれ得るが、サイクルN及び／又はその後続のサイクルのフローセル画像は、まだキャプチャされていないか、又は現在キャプチャされていない。いくつかの態様において、サイクルNは、現在のサイクルである。Nは、任意の非ゼロ整数、例えば、3、4、5、6、又は10であり得る。

30

【0097】

本明細書に開示されるように、「ベースコーリングテンプレート」及び「ポロニーマップ」は等価であり、交換可能に使用することができる。ベースコーリングテンプレート又はポロニーマップは、本明細書に開示されるような1つ以上のベースコーリング位置を含み得る。本明細書に開示されるベースコーリングテンプレート又はポロニーマップは、2D又は3Dのいずれかであり得る。ベースコーリングテンプレート又はポロニーマップ内の全てのベースコーリング位置が同じ2D平面内にある場合、ベースコーリングテンプレート又はポロニーマップは2Dである。ベースコーリング位置のうちのいくつかが同じ2D平面ではなく3D空間内にある場合、ベースコーリングテンプレート又はポロニーマップは3Dである。

40

【0098】

本明細書における空間位置及び関係を定義するために様々な座標系を使用することができる。本明細書において空間位置を定義するために使用されるが、空間位置及び関係を定義するために様々な代替的な座標系を使用することができる。直交系以外のいくつかの例示的な座標系は、極座標系、円筒座標系、又は球面座標系を含むことができるが、これらに

50

限定されない。他の座標系は、均質な座標系又は非均質な座標系を含むことができる。

【0099】

いくつかの態様において、方法500は、第1の標的RNA分子及び第2の標的RNA分子を含む複数のRNAを保有する細胞試料を提供する操作を含むことができる。

【0100】

いくつかの態様において、細胞試料は、固定され、透過化される。いくつかの態様において、細胞試料は、2～25個の異なる標的RNA分子を保有するか、又は25～50個の異なる標的RNA分子を保有するか、又は50～75個の異なる標的RNA分子を保有するか、又は75～100個の異なる標的RNA分子を保有する。いくつかの態様において、細胞試料は、100個を超える異なる標的RNA分子、又は250個を超える異なる標的RNA分子、又は500個を超える異なる標的分子、又は1000個を超える異なる標的RNA分子、又はそれ以上を保有する。いくつかの態様において、細胞試料は、10,000個を超える異なる標的RNA分子を有する。いくつかの態様において、細胞試料は、全細胞、複数の全細胞、無傷の組織、又は無傷の腫瘍を含む。いくつかの態様において、細胞試料は、新鮮な細胞試料、新鮮凍結細胞試料、切片化細胞試料、FFPE細胞試料、又は切片化FFPE細胞試料を含む。いくつかの態様において、細胞試料は、固体支持体上に堆積される。いくつかの態様において、細胞試料は、細胞接着を促進するコーティングで不動態化された固体支持体上に堆積される。いくつかの態様において、細胞試料は、固定化されたキャプチャオリゴヌクレオチドを欠く支持体上に堆積される。いくつかの態様において、細胞試料は、細胞試料を固体支持体に堆積させる前又は後に培養される。いくつかの態様において、細胞試料は、以下に記載されるステップ(b)を行う前に培養される。いくつかの態様において、細胞試料は、単純又は複合細胞培養培地中で培養された拡張細胞試料を含む。いくつかの態様において、細胞試料は、細胞試料内に複数のcDNA分子を生成する前に培養又は拡張されない。

【0101】

いくつかの態様において、方法500は、細胞試料内で、第1の標的RNA分子に対応する第1の標的cDNA分子及び第2の標的RNA分子に対応する第2の標的cDNA分子を含む複数のcDNA分子を生成する操作を含むことができる。

【0102】

いくつかの態様において、複数のcDNA分子を生成する操作は、2～10,000個の異なる標的RNA分子に対応する少なくとも2～10,000個の異なる標的cDNA分子を生成することを含む。いくつかの態様において、cDNA分子を生成する操作は、細胞試料内の複数のRNAを、(i)複数の逆転写プライマー、(ii)複数の逆転写酵素、及び(iii)複数のヌクレオチドと、細胞試料中で複数のcDNA分子(例えば、複数の第1の鎖cDNA分子)を生成するための逆転写反応を行うのに好適な条件下で接触させることを含む(例えば、図7)。

【0103】

いくつかの態様において、複数の逆転写プライマーは、第1の標的RNAに選択的にハイブリダイズする標的特異的逆転写プライマーの第1のサブ集団を含み、第2の標的RNAに選択的にハイブリダイズする標的特異的逆転写プライマーの第2のサブ集団を含む。いくつかの態様において、標的特異的逆転写プライマーの第1及び第2のサブ集団は、同じ配列又は異なる配列を有する。

【0104】

いくつかの態様において、標的特異的逆転写プライマーの第1のサブ集団の全長は、第1の標的RNA分子にハイブリダイズする。いくつかの態様において、標的特異的逆転写プライマーの第1のサブ集団は、第1の標的RNA分子にハイブリダイズする一部分と、第1の標的RNA分子にハイブリダイズしない一部分とを有するテール付きプライマーを含む。いくつかの態様において、標的特異的逆転写プライマーの第1のサブ集団は、ポリT配列を有する少なくとも一部分を含む。いくつかの態様において、標的特異的逆転写プライマーの第1のサブ集団は、ランダム配列を有する少なくとも一部分及び／又は標的特

異的配列を有する少なくとも一部分を含む。

【0105】

いくつかの態様において、標的特異的逆転写プライマーの第2のサブ集団の全長は、第2の標的RNA分子にハイブリダイズする。いくつかの態様において、標的特異的逆転写プライマーの第2のサブ集団は、第2の標的RNA分子にハイブリダイズする一部分と、第2の標的RNA分子にハイブリダイズしない一部分とを有するテール付きプライマーを含む。いくつかの態様において、標的特異的逆転写プライマーの第2のサブ集団は、ポリT配列を有する少なくとも一部分を含む。いくつかの態様において、標的特異的逆転写プライマーの第2のサブ集団は、ランダム配列を有する少なくとも一部分及び／又は標的特異的配列を有する少なくとも一部分を含む。

10

【0106】

いくつかの態様において、cDNA分子にハイブリダイズされる標的RNA分子は、RNA/DNA二本鎖中のRNAを分解するのに好適な条件下で、リボヌクレアーゼを使用して酵素分解に供され得る。いくつかの態様において、cDNA分子にハイブリダイズされる標的RNA分子は、酵素分解に供されない。

【0107】

いくつかの態様において、方法500は、細胞試料中の複数のcDNA分子を、少なくとも第1の複数の第1の標的特異的パドロックプローブ及び第2の複数の第2の標的特異的パドロックプローブを含む、複数の標的特異的パドロックプローブと接触させることを更に含む。いくつかの態様において、本方法は、細胞試料中の複数のcDNA分子を、少なくとも2〜10,000個の異なる標的特異的パドロックプローブと接触させることを含む。

20

【0108】

代替の態様において、cDNAは、細胞試料内のRNAから生成されない。いくつかの態様において、方法500は、細胞内のRNAを複数の標的特異的パドロックプローブと接触させ、環状化パドロックプローブを生成する操作を含む。いくつかの態様において、方法は、細胞試料中の複数のRNA分子を、少なくとも第1の複数の標的特異的パドロックプローブ及び第2の複数の標的特異的パドロックプローブを含む、複数の標的特異的パドロックプローブと接触させる操作を含む。いくつかの態様において、本方法は、細胞試料中の複数のcDNA分子を、少なくとも2〜10,000個の異なる標的特異的パドロックプローブと接触させることを含む。いくつかの態様において、標的RNA分子は、リボヌクレアーゼを使用して酵素分解に供され得る。いくつかの態様において、標的RNA分子は、酵素分解に供されない。

30

【0109】

いくつかの態様において、複数の第1の標的特異的パドロックプローブにおける個々のパドロックプローブは、第1及び第2の末端領域（例えば、第1及び第2のパドロック結合アーム）を含み、第1の末端領域は、第1の標的cDNA分子（又は第1の標的RNA分子）の第1の領域に選択的にハイブリダイズし、第2の末端領域は、第1の標的cDNA分子（又は第1の標的RNA分子）の第2の領域に選択的にハイブリダイズする。いくつかの態様において、複数のcDNA分子を接触させることは、第1の標的特異的パドロックプローブの第1及び第2の末端領域を第1の標的cDNA分子（又は第1の標的RNA分子）上の近位位置にハイブリダイズして、ハイブリダイズした第1及び第2の末端領域の間にニック又はギャップを有する環状化された第1の標的特異的パドロックプローブを形成することを含む（例えば、図7、左）。いくつかの態様において、第1の標的特異的パドロックプローブは、第1の標的cDNA配列（又は第1の標的RNA配列）に対応し、かつそれを一意に識別する第1の標的バーコード配列（標的BC-1）を含む。いくつかの態様において、第1の標的特異的パドロックプローブは、第1の標的cDNA分子（又は第1の標的RNA配列）に選択的にハイブリダイズする第1の標的特異的パドロックプローブの領域のうちの1つに隣接して位置する第1の標的バーコード配列を含む。いくつかの態様において、第1の標的特異的パドロックプローブは、少なくとも1つのユニ

40

50

バーサルアダプター配列、例えば、ユニバーサルシーケンシングプライマー結合部位（又はその相補的配列）を含む。いくつかの態様において、第1の標的特異的パドロックプローブは、ローリングサークル増幅プライマー（又はその相補的配列）のためのユニバーサルプライマー結合部位を含む。いくつかの態様において、第1の標的特異的パドロックプローブは、ユニバーサルコンパクションオリゴヌクレオチド結合部位（又はその相補的配列）を含む。

【0110】

いくつかの態様において、複数の第2の標的特異的パドロックプローブにおける個々のパドロックプローブは、第1及び第2の末端領域（例えば、第1及び第2のパドロック結合アーム）を含み、第1の末端領域は、第2の標的cDNA分子（又は第2の標的RNA分子）の第1の領域に選択的にハイブリダイズし、第2の末端領域は、第2の標的cDNA分子（又は第2の標的RNA分子）の第2の領域に選択的にハイブリダイズする。いくつかの態様において、ステップ（c）の接触は、第2の標的特異的パドロックプローブの第1及び第2の末端領域を第2の標的cDNA分子（又は第2の標的RNA分子）上の近位位置にハイブリダイズして、ハイブリダイズした第1及び第2の末端領域の間にニック又はギャップを有する環状化された第2の標的特異的パドロックプローブを形成することを含む（例えば、図7、右）。いくつかの態様において、第2の標的特異的パドロックプローブは、第2の標的cDNA配列（又は第2の標的RNA配列）に対応し、かつそれを一意に識別する第2の標的バーコード配列（標的BC-2）を含む。いくつかの態様において、第2の標的特異的パドロックプローブは、第2の標的cDNA分子（又は第2の標的RNA配列）に選択的にハイブリダイズする第2の標的特異的パドロックプローブの領域のうちの1つに隣接して位置する第2の標的バーコード配列を含む。いくつかの態様において、第2の標的特異的パドロックプローブは、少なくとも1つのユニバーサルアダプター配列、例えば、ユニバーサルシーケンシングプライマー結合部位（又はその相補的配列）を含む。いくつかの態様において、第2の標的特異的パドロックプローブは、ローリングサークル増幅プライマー（又はその相補的配列）のためのユニバーサルプライマー結合部位を含む。いくつかの態様において、第2の標的特異的パドロックプローブは、ユニバーサルコンパクションオリゴヌクレオチド結合部位（又はその相補的配列）を含む。

【0111】

いくつかの態様において、第1の標的バーコード配列（標的BC-1）及び第2の標的バーコード配列（標的BC-2）は、異なる配列を有し、多重RNA検出及びシーケンシングを行うために使用され得る。いくつかの態様において、第1の標的バーコード配列（標的BC-1）及び第2の標的バーコード配列（標的BC-2）は、同じ配列を有し、ユニプレックスRNAの検出及びシーケンシングを行うために使用され得る。

【0112】

いくつかの態様において、第1及び第2の標的特異的パドロックプローブは、コンカテマーの標的バーコード領域が最初にシーケンシングされるように、互いに隣接するユニバーサルシーケンシングプライマー結合部位及び標的バーコード配列を含む。標的バーコード配列は、いずれかの長さ、例えば、3～15塩基、又は15～25塩基、又は25～40塩基、又はより長い長さであることができる。

【0113】

いくつかの態様において、細胞試料中の複数のRNA分子を複数の標的特異的パドロックプローブと接触させる操作は、第1の標的特異的パドロックプローブの第1及び第2の末端領域を第1の標的cDNA分子又は第1の標的RNA分子上の近位位置にハイブリダイズして、ハイブリダイズした第1及び第2の末端領域の間にニック又はギャップを有する環状化された第1の標的特異的パドロックプローブを形成することを含む。

【0114】

図6は、パドロックプローブの例示的な態様を示す概略図である。いくつかの態様において、パドロックプローブは、2つの末端領域（例えば、第1及び第2の結合アーム）及び内部領域を有する一本鎖核酸分子を含む。いくつかの態様において、個々のパドロック

10

20

30

40

50

プローブの第1の末端領域は、標的RNA又は標的cDNA分子の第1の領域に選択的にハイブリダイズする第1の標的特異的配列を有し、個々のパドロックプローブの第2の末端領域は、同じ標的RNA又は標的cDNA分子の第2の領域に選択的にハイブリダイズする第2の標的特異的配列を有する。いくつかの態様において、パドロックの内部領域は、所与の標的RNA又は標的cDNAに対応する標的バーコード配列（例えば、それぞれ、標的BC-1又は標的BC-2、左及び右の概略図）を含む。いくつかの態様において、標的バーコード配列は、標的RNA又は標的cDNAを一意に識別する。いくつかの態様において、パドロックの内部領域は、シーケンシングプライマー（又はその相補的配列）のためのユニバーサルプライマー結合部位を含む。いくつかの態様において、パドロックの内部領域は、ローリングサークル増幅プライマー（又はその相補的配列）のためのユニバーサルプライマー結合部位を含む。いくつかの態様において、パドロックの内部領域は、コンパクションオリゴヌクレオチド結合（又はその相補的配列）のためのユニバーサル結合部位を含む。いくつかの態様において、パドロックプローブの内部領域は、任意の配置及び方向で、標的バーコード配列及び少なくとも1つのユニバーサルプライマー結合部位（例えば、シーケンシングプライマーを結合するため、ローリングサークル増幅プライマーを結合するため、及び／又はコンパクションオリゴヌクレオチドを結合するため）を含む（図6、上及び下）。

10

【0115】

図7は、第1及び第2の標的RNA分子から（それぞれ）第1及び第2のcDNAを生成すること、第1及び第2のパドロックプローブを（それぞれ）第1及び第2のcDNA分子にハイブリダイズして、（それぞれ）第1及び第2の環状化パドロックプローブを生成すること、を含む、細胞内で環状化パドロックプローブを生成するためのワークフローを示す概略図である。第1のパドロックプローブは、(i)第1の標的RNA又は第1の標的cDNAを一意に識別する第1の標的バーコード配列（標的BC-1）、(ii)第1のシーケンシングプライマー結合部位（又はその相補的配列）、(iii)増幅プライマー（ユニバーサルRCA）（又はその相補的配列）のためのユニバーサル結合部位、及び(iv)コンパクションオリゴヌクレオチド（又はその相補的配列）のためのユニバーサル結合部位を含む。第2のパドロックプローブは、(i)第2の標的RNA又は第2の標的cDNAを一意に識別する第2の標的バーコード配列（標的BC-2）、(ii)第2のシーケンシングプライマー結合部位（又はその相補的配列）、(iii)増幅プライマーのためのユニバーサル結合部位（ユニバーサルRCA）（又はその相補的配列）、及び(iv)コンパクションオリゴヌクレオチドのためのユニバーサル結合部位（又はその相補的配列）を含む。

20

30

【0116】

いくつかの態様において、方法500は、酵素反応を行うことによって、少なくとも第1及び第2の環状化標的特異的パドロックプローブ内のニック又はギャップを閉鎖し、それによって、細胞試料内に少なくとも第1の共有結合性閉鎖環状パドロックプローブ及び第2の共有結合性閉鎖環状パドロックプローブを生成する操作を含む。いくつかの態様において、第1及び第2の環状化パドロックプローブにおけるニックを閉じることは、酵素ライゲーション反応を行うことを含む。いくつかの態様において、第1及び第2の環状化パドロックプローブ内のギャップを閉じることは、第1若しくは第2の標的cDNA分子（又は第1若しくは第2のRNA分子）をテンプレートとして使用してポリメラーゼ触媒充填反応を行うことと、酵素ライゲーション反応を行うことと、を含む。いくつかの態様において、方法は、1つ以上の酵素反応を行うことによって、少なくとも2〜10,000個の環状化標的特異的パドロックプローブ内のニック又はギャップを閉鎖し、それによって、細胞試料内に少なくとも2〜10,000個の共有結合性閉鎖環状パドロックプローブを生成することを含む。図8は、例示的な共有結合性閉鎖環状パドロックプローブを示す。

40

【0117】

いくつかの態様において、方法500は、第1及び第2の共有結合性閉鎖環状パドロッ

50

クプローブをテンプレート分子として使用して、細胞試料内でローリングサークル増幅反応を行い、それによって、第1の標的RNA分子に対応する少なくとも第1のコンカテマー分子を含む複数のコンカテマー分子を生成する操作を含み、複数のコンカテマー分子は、第2の標的RNA分子に対応する少なくとも第2のコンカテマー分子を含む。いくつかの態様において、第1のコンカテマー分子は、タンデムリピート単位を含み、単位は、第1の標的cDNA（又は第1の標的RNA）、第1の標的バーコード配列、及びユニバーサルシーケンシングプライマー結合部位（又はその相補的配列）に対応する配列を含む。いくつかの態様において、第2のコンカテマー分子は、タンデムリピート単位を含み、単位は、第2の標的cDNA（又は第2の標的RNA）、第2の標的バーコード配列、及びユニバーサルシーケンシングプライマー結合部位（又はその相補的配列）に対応する配列を含む。

10

【0118】

いくつかの態様において、本明細書におけるローリングサークル増幅反応は、共有結合性閉鎖環状化パドロックプローブを、増幅プライマー（例えば、ユニバーサルローリングサークル増幅プライマー）、鎖置換DNAポリメラーゼ、及び複数のヌクレオチドと、個々の増幅プライマーを共有結合性閉鎖パドロックプローブにハイブリダイズするのに好適な条件下で、並びに共有結合性閉鎖パドロックプローブをテンプレート分子として使用してプライマー伸長を行うのに好適な条件下で接触させて、核酸コンカテマーを生成することを含む。いくつかの態様において、方法は、少なくとも2～10,000個の共有結合性閉鎖環状パドロックプローブをテンプレート分子として使用して、細胞試料内でローリングサークル増幅反応を行い、それによって、少なくとも2～10,000個の標的RNA分子に対応する少なくとも2～10,000個のコンカテマー分子を生成することを含む。いくつかの態様において、細胞試料内で生成された複数のコンカテマーは、非崩壊コンカテマーと比較してよりコンパクトな形状及びサイズを有するDNAナノボールに崩壊する。

20

【0119】

いくつかの態様において、方法500は、1つ以上のシーケンシング反応サイクルを行うことによって、支持体上に固定化された細胞試料の複数のフローセル画像を生成する操作510を含む。

【0120】

複数のフローセル画像は、シーケンシング反応の1つ以上のサイクルにおいて取得された第1の複数のフローセル画像であり得る。1つ以上のサイクルは、シーケンシングラン中の任意のサイクルであり得る。1つ以上のサイクルは、シーケンシングラン中のより早期のサイクル、例えば、サイクル1～サイクル10又はサイクル1～サイクル5であり得る。

30

【0121】

いくつかの態様において、細胞試料は、その中に複数のコンカテマー分子を含む。複数のコンカテマー分子の第1のコンカテマー分子（複数可）は、細胞試料の第1の標的RNA分子に対応し得る。複数のコンカテマー分子の第2のコンカテマー分子（複数可）は、細胞試料の第2の標的RNA分子に対応し得る。

40

【0122】

いくつかの態様において、細胞試料は、支持体上で培養され得る。いくつかの態様において、方法500は、細胞試料を2～10世代又はそれ以上拡張するのに好適な条件下で、支持体上で細胞試料を培養する操作を含む。培養した細胞試料は、細胞のコロニーを生成することができる。いくつかの態様において、細胞試料は、1つ以上のインサイチュ試料を含む。いくつかの態様において、細胞試料は、1つ以上の細胞又は組織を含む。いくつかの態様において、この方法は、細胞試料をコンフルエンス又は非コンフルエンスに培養することを含み得る。いくつかの態様において、この方法は、単純又は複合細胞培養培地中で、支持体上の細胞試料を培養することを含む。例えば、細胞培地は、D-MEM高グルコース（例えば、Thermo Fisher Scientificから、カタロ

50

グ番号11965118)、ウシ胎児血清(例えば、10%FBS;例えば、Thermo Fisher Scientificから、カタログ番号A3160402)、MEM非必須アミノ酸(例えば、0.1mMのMEM、例えば、Thermo Fisher Scientificから、カタログ番号11140050)、L-グルタミン(例えば、6mMのL-グルタミン、例えば、Thermo Fisher Scientificから、カタログ番号A2916801)、MEMピルビン酸ナトリウム(例えば、1mMのピルビン酸ナトリウム、例えば、Thermo Fisher Scientificから、カタログ番号11360070)、及び抗生物質(例えば、1%ペニシリンーストレプトマイシン-グルタミン、例えば、Thermo Fisher Scientificから、カタログ番号10378016)を含む。いくつかの態様において、方法は、支持体上で細胞(複数可)を培養するのに好適な湿度及び温度で細胞試料を培養することを含む。例示的な好適な条件は、空気中におよそ5~10%の二酸化炭素の加湿雰囲気を含む。細胞試料は、酸素及び/又は窒素による好適な通気で培養することができる。

10

【0123】

本明細書に記載される方法のうちのいずれにおいても、細胞試料は、標的RNA及び/又は非標的RNAを含む複数のRNAを保有する。いくつかの態様において、細胞は、典型的には、DNA(例えば、ゲノムDNA)のRNA分子への転写を含む遺伝子発現によってRNAを産生する。転写されたRNAは、スプライシングを受けることができるか、又はスプライシングされ得ない。転写されたRNAは、ポリペプチド(例えば、コーディングRNA)に翻訳され得るか、又は翻訳を受けないが、tRNA又はrRNA(例えば、ノンコーディングRNA)にプロセッシングされ得る。

20

【0124】

いくつかの態様において、細胞試料は、その中に複数のRNAを含むことができる。RNAは、細胞試料の個々の細胞内にあり得る。複数のRNAは、細胞試料によって保有され得る。RNAは、標的及び非標的RNAを含むことができる。いくつかの態様において、細胞試料によって保有される複数のRNAは、野生型RNA、変異体RNA、又はスプライスバリエントRNAを含む。いくつかの態様において、細胞試料によって保有される複数のRNAは、事前にスプライシングされたRNA、部分的にスプライシングされたRNA、又は完全にスプライシングされたRNAを含む。いくつかの態様において、細胞試料によって保有される複数のRNAは、コーディングRNA、ノンコーディングRNA、mRNA、tRNA、rRNA、マイクロRNA(miRNA)、成熟マイクロRNA、又は未成熟マイクロRNAを含む。いくつかの態様において、細胞試料によって保有される複数のRNAは、ハウスキーピングRNA、細胞特異的RNA、組織特異的RNA、又は疾患特異的RNAを含む。いくつかの態様において、細胞試料によって保有される複数のRNAは、熱、光、化学物質、又は薬物などの刺激に応答して1つ以上の細胞によって発現されるRNAを含む。いくつかの態様において、細胞試料によって保有される複数のRNAは、健康な細胞又は罹患した細胞において見出されるRNAを含む。いくつかの態様において、細胞試料によって保有される複数のRNAは、組換えDNA手順を使用して細胞試料に導入されるトランスジェニックDNA配列から転写されたRNAを含む。例えば、RNAは、誘導性又は構成的プロモーター配列によって制御されるトランスジェニックDNA配列から転写され得る。いくつかの態様において、細胞試料によって保有される複数のRNAは、トランスジェニックではないDNA配列から転写されるRNAを含む。

30

40

【0125】

いくつかの態様において、操作510は、1つ以上のサイクルの各々において、複数のコンカテマー分子に結合したヌクレオチド試薬から放出された光学カラーシグナルをイメージングすることを含む。イメージング操作は、本明細書のシーケンシングシステム100の光学システム又はイメージャ116によって実行され得る。第1の複数のフローセル画像は、複数のコンカテマー分子に結合したヌクレオチド試薬から放出された光学カラーシグナルを含み得る。

50

【0126】

いくつかの態様において、操作510は、細胞試料内の複数のコンカテマー分子をシーケンシングする操作を含み、これは、複数の第1のシーケンシングリード生成物を生成するために、所定の数以下のシーケンシングサイクル（例えば、2〜30サイクル）を行うことによって第1のコンカテマー分子をシーケンシングすることと、複数の第2のシーケンシングリード生成物を生成するために、所定の数以下のシーケンシングサイクルを行うことによって第2のコンカテマー分子をシーケンシングすることと、を含む（図8）。いくつかの態様において、シーケンシング操作は、複数の第1のシーケンシングリード生成物を生成するために、第1のコンカテマー分子の所定の数以下の塩基（例えば、2〜30塩基）をシーケンシングすることを含み、これは、複数の第2のシーケンシングリード生成物を生成するために、第2のコンカテマー分子の所定の数以下のシーケンシングサイクル塩基をシーケンシングすることを含む。いくつかの態様において、本方法は、細胞試料内の少なくとも2〜10,000個のコンカテマー分子をシーケンシングすることを含み、これは、複数のシーケンシングリード生成物を生成するために、2〜10,000個のコンカテマー分子上で所定数以下のシーケンシングサイクルを行うことを含む。

10

【0127】

いくつかの態様において、第1のコンカテマー分子の第1の標的バーコード領域のみがシーケンシングされる（例えば、図8、上）。いくつかの態様において、第1のコンカテマー分子の第1の標的バーコードの少なくとも一部分又は全長がシーケンシングされる（例えば、図8、上）。いくつかの態様において、第1の標的バーコードがシーケンシングされ、第1のコンカテマー分子の第1のcDNA領域（又は第1のRNA領域）の一部分がシーケンシングされる。いくつかの態様において、第1のコンカテマー分子の第1のcDNA領域（又は第1のRNA領域）の少なくとも一部分がシーケンシングされる。

20

【0128】

いくつかの態様において、第2のコンカテマー分子の第2の標的バーコード領域のみがシーケンシングされる（例えば、図8、下）。いくつかの態様において、第2のコンカテマー分子の第2の標的バーコードの少なくとも一部分又は全長がシーケンシングされる（例えば、図8、下）。いくつかの態様において、第2の標的バーコードがシーケンシングされ、第2のコンカテマー分子の第2のcDNA領域（又は第2のRNA領域）の一部分がシーケンシングされる。いくつかの態様において、第2のコンカテマー分子の第2のcDNA領域（又は第2のRNA領域）の少なくとも一部分がシーケンシングされる。

30

【0129】

いくつかの態様において、シーケンシング操作は、細胞試料内の複数のコンカテマー分子を、(i) 複数のユニバーサルシーケンシングプライマー、(ii) 複数のシーケンシングポリメラーゼ、及び(iii) 複数のヌクレオチド試薬と、複数のユニバーサルシーケンシングプライマーをコンカテマー上のそれらのそれぞれのユニバーサルシーケンシングプライマー結合部位にハイブリダイズするのに好適な条件下で接触させることを含む。いくつかの態様において、シーケンシング操作は、所定の数以下のシーケンシングサイクル（例えば、2〜30サイクル）を行って、少なくとも第1の標的バーコード領域（標的BC-1）をシーケンシングすることによって少なくとも第1の複数のシーケンシングリード生成物を生成し、任意選択で、所定の数以下のシーケンシングサイクルを行って、少なくとも第2の標的バーコード領域（標的BC-2）をシーケンシングすることによって少なくとも第2の複数のシーケンシングリード生成物を生成することを更に含む。いくつかの態様において、ヌクレオチド試薬は、多価分子、ヌクレオチド及び／又はヌクレオチド類似体を含む。

40

【0130】

いくつかの態様において、シーケンシング操作は、 1.0 mm^2 を超える視野（FOV）を含む光学イメージングシステムを使用して、第1及び第2の核酸コンカテマーの少なくとも一部分をシーケンシングすることを含む。

【0131】

50

いくつかの態様において、複数の第1及び第2のシーケンシングリード生成物は、イメージングによって検出可能であり、シーケンシングすることは、所定の数以下のシーケンシングサイクルの間に得られた画像から複数の第1及び第2のシーケンシングリード生成物を復号することを含む。

【0132】

いくつかの態様において、複数の第1及び第2のシーケンシングリード生成物は、イメージングによって検出可能であり、シーケンシングは、細胞試料中の複数の第1及び第2の検出可能なシーケンシングリード生成物を同時にイメージングすることを含む（第1及び第2のシーケンシングリード生成物の共局在化）。

【0133】

例示的なシーケンシングリード生成物は、図8～10に破線矢印として示される。第1のシーケンシングリード生成物は、第1の標的バーコード配列、及び任意選択で、第1の挿入配列の一部を含んでもよい。第1のシーケンシングリード生成物は、標的バーコード配列の相補鎖又は逆相補鎖、及び任意選択で挿入配列の一部を含み得る。

【0134】

図8は、第1及び第2の共有結合性閉鎖環状分子を使用してローリングサークル増幅を行うことによって第1及び第2のコンカテマーを生成することを含む、細胞内のローリングサークル及びシーケンシングワークフローを示す概略図である。第1及び第2のコンカテマーは、ユニバーサルシーケンシングプライマー、シーケンシングポリメラーゼ、及び複数のヌクレオチド試薬を使用したシーケンシングワークフローに供される。

【0135】

図9は、細胞内で生成されるコンカテマーをシーケンシングするための例示的なワークフローを示す概略図である。コンカテマーは、タンデムリピート単位を含み、各単位は、(i) ユニバーサルシーケンシングプライマー結合部位 (Seq) と、(ii) ユニバーサルコンパクションオリゴヌクレオチド結合部位 (CO) と、(iii) 所与の標的 cDNA に対応する挿入配列と、(iv) 所与の標的 cDNA (BC) に対応する標的バーコード配列と、を含む。いくつかの態様において、ユニバーサルシーケンシングプライマー（実線矢印）は、ユニバーサルシーケンシングプライマー結合部位にハイブリダイズし、30回以下のシーケンシングサイクルを行って、複数の第1のシーケンシングリード生成物（破線矢印）を生成し、第1のシーケンシングリード生成物は、標的バーコード配列のみを含む。複数の第1のシーケンシングリード生成物をコンカテマーから除去し、シーケンシングを繰り返し、30回以下のシーケンシングサイクルを行って、別の複数の第1のシーケンシングリード生成物（破線矢印）を生成し、第1のシーケンシングリード生成物は、標的バーコード配列のみを含む。複数の第1のシーケンシングリード生成物をコンカテマーから除去し、シーケンシングを再び繰り返し、30回以下のシーケンシングサイクルを行って、別の複数の第1のシーケンシングリード生成物（破線矢印）を生成し、第1のシーケンシングリード生成物は、標的バーコード配列のみを含む。いくつかの態様において、反復シーケンシングは、最大50回行うことができる。第1のシーケンシングリード生成物の全ての配列を判定し、第1の参照配列（例えば、参照バーコード配列）とアラインメントして、細胞試料内の第1の標的 RNA 分子の存在を確認することができる。

【0136】

いくつかの態様において、方法500は、複数の第1のシーケンシングリード生成物を第1のコンカテマー分子から除去し、第1のコンカテマー分子を細胞試料中に保持し、複数の第2のシーケンシングリード生成物を第2のコンカテマー分子から除去し、第2のコンカテマー分子を細胞試料中に保持する操作を含む。いくつかの態様において、第1及び第2のコンカテマー分子は、細胞試料中に残るが、シーケンシングリード生成物としての相補鎖又は逆相補鎖は除去される。

【0137】

いくつかの態様において、方法500は、510の操作と、複数の第1のシーケンシングリード生成物を第1のコンカテマー分子から除去し、第1のコンカテマー分子を細胞試

10

20

30

40

50

料中に保持し、複数の第2のシーケンシングリード生成物を第2のコンカテマー分子から除去し、第2のコンカテマー分子を細胞試料中に保持する操作と、をn回繰り返すことによって複数のコンカテマーを反復してシーケンシングする操作を含む。

【0138】

いくつかの態様において、方法500は、所定の数以下のシーケンシングサイクルを行うことによって第1のコンカテマー分子（複数可）及び第2のコンカテマー分子（複数可）をシーケンシングして、複数の第1及び第2のシーケンシングリード生成物を生成する操作と、複数の第1のシーケンシングリード生成物を第1のコンカテマー分子から除去し、第1のコンカテマー分子を細胞試料中に保持し、複数の第2のシーケンシングリード生成物を第2のコンカテマー分子から除去し、第2のコンカテマー分子を細胞試料中に保持する操作と、をn回繰り返すことによって、複数のコンカテマーを反復してシーケンシングする操作を含む。

10

【0139】

いくつかの態様において、複数の第1のシーケンシングリード生成物の配列は、細胞試料中の第1の標的RNA分子の存在を確認し、複数の第2のシーケンシングリード生成物の配列は、細胞試料中の第2の標的RNA分子の存在を確認する。

【0140】

いくつかの態様において、コンカテマーの少なくとも1つの領域を反復してシーケンシングすることは、シーケンシング操作及び除去操作をn回繰り返すことを含み、nは、1～20、1～30、1～100、又は更に大きな整数範囲の範囲であり得る。

20

【0141】

反復シーケンシングの例を図9～12の概略図に示す。反復シーケンシングは、シーケンシングエラーを低減し、シーケンシング精度及び信頼性を改善するのに役立つことができる。

【0142】

いくつかの態様において、少なくとも1つのコンカテマーは、非反復シーケンシングによってシーケンシングされる。いくつかの態様において、少なくとも1つのコンカテマーは、シーケンシング操作及び除去操作を1回又は2回繰り返すことによってシーケンシングされる。

【0143】

図10は、細胞内で生成されるコンカテマーをシーケンシングするための例示的なワークフローを示す概略図である。コンカテマーは、タンデムリピート単位を含み、各単位は、(i)ユニバーサルシーケンシングプライマー結合部位(Seq)と、(ii)ユニバーサルコンパクションオリゴヌクレオチド結合部位(CO)と、(iii)所与の標的cDNAに対応する挿入配列と、(iv)所与の標的cDNA(BC)に対応する標的バーコード配列と、を含む。いくつかの態様において、ユニバーサルシーケンシングプライマー（実線矢印）は、ユニバーサルシーケンシングプライマー結合部位にハイブリダイズし、30回以下のシーケンシングサイクルを行って、複数の第1のシーケンシングリード生成物（破線矢印）を生成し、第1のシーケンシングリード生成物は、標的バーコード配列及び挿入配列の一部を含む。複数の第1のシーケンシングリード生成物をコンカテマーから除去し、シーケンシングを繰り返し、30回以下のシーケンシングサイクルを行って、別の複数の第1のシーケンシングリード生成物（破線矢印）を生成し、第1のシーケンシングリード生成物は、標的バーコード配列及び挿入配列の一部を含む。複数の第1のシーケンシングリード生成物をコンカテマーから除去し、シーケンシングを再び繰り返し、30回以下のシーケンシングサイクルを行って、別の複数の第1のシーケンシングリード生成物（破線矢印）を生成し、第1のシーケンシングリード生成物は、標的バーコード配列及び挿入配列の一部を含む。いくつかの態様において、反復シーケンシングは、最大50回行うことができる。第1のシーケンシングリード生成物の全ての配列を判定し、第1の参照配列（例えば、参照バーコード配列及び標的RNAに対応する挿入配列）とアラインメントして、細胞試料内の第1の標的RNA分子の存在を確認することができる。

30

40

50

【0144】

図11は、細胞内で生成されるコンカテマーをシーケンシングするための例示的なワークフローを示す概略図である。コンカテマーは、タンデムリピート単位を含み、各単位は、(i)ユニバーサルシーケンシングプライマー結合部位(Seq)と、(ii)ユニバーサルコンパクションオリゴヌクレオチド結合部位(CO)と、(iii)所与の標的cDNAに対応する挿入配列と、を含む。いくつかの態様において、ユニバーサルシーケンシングプライマー(実線矢印)は、ユニバーサルシーケンシングプライマー結合部位にハイブリダイズし、30回以下のシーケンシングサイクルを行って、複数の第1のシーケンシングリード生成物(破線矢印)を生成し、第1のシーケンシングリード生成物は、挿入配列の一部分を含む。複数の第1のシーケンシングリード生成物をコンカテマーから除去し、シーケンシングを繰り返し、30回以下のシーケンシングサイクルを行って、別の複数の第1のシーケンシングリード生成物(破線矢印)を生成し、第1のシーケンシングリード生成物は、挿入配列の一部分を含む。複数の第1のシーケンシングリード生成物をコンカテマーから除去し、シーケンシングを再び繰り返し、30回以下のシーケンシングサイクルを行って、別の複数の第1のシーケンシングリード生成物(破線矢印)を生成し、第1のシーケンシングリード生成物は、挿入配列の一部分を含む。いくつかの態様において、反復シーケンシングは、最大50回行うことができる。第1のシーケンシングリード生成物の全ての配列を判定し、第1の参照配列(例えば、標的RNAに対応する挿入配列)とアラインメントして、細胞試料内の第1の標的RNA分子の存在を確認することができる。

10

20

【0145】

図12は、細胞内で生成されるコンカテマーをシーケンシングするための例示的なワークフローを示す概略図である。コンカテマーは、タンデムリピート単位を含み、各単位は、(i)ユニバーサルシーケンシングプライマー結合部位(Seq)と、(ii)所与の標的cDNAに対応する挿入配列と、を含む。いくつかの態様において、ユニバーサルシーケンシングプライマー(実線矢印)は、ユニバーサルシーケンシングプライマー結合部位にハイブリダイズし、30回以下のシーケンシングサイクルを行って、複数の第1のシーケンシングリード生成物(破線矢印)を生成し、第1のシーケンシングリード生成物は、挿入配列の一部分を含む。複数の第1のシーケンシングリード生成物をコンカテマーから除去し、シーケンシングを繰り返し、30回以下のシーケンシングサイクルを行って、別の複数の第1のシーケンシングリード生成物(破線矢印)を生成し、第1のシーケンシングリード生成物は、挿入配列の一部分を含む。複数の第1のシーケンシングリード生成物をコンカテマーから除去し、シーケンシングを再び繰り返し、30回以下のシーケンシングサイクルを行って、別の複数の第1のシーケンシングリード生成物(破線矢印)を生成し、第1のシーケンシングリード生成物は、挿入配列の一部分を含む。いくつかの態様において、反復シーケンシングは、最大50回行うことができる。第1のシーケンシングリード生成物の全ての配列を判定し、第1の参照配列(例えば、標的RNAに対応する挿入配列)とアラインメントして、細胞試料内の第1の標的RNA分子の存在を確認することができる。

30

【0146】

いくつかの態様において、操作510は、1つ以上のサイクルの各々において、第1のコンカテマー分子の第1の標的バーコード配列領域のみをシーケンシングし、それによって、第1のシーケンシングリード生成物を生成することを含む。いくつかの態様において、操作510は、1つ以上のサイクルの各々において、第1の標的バーコード配列領域のうちのいくつか又は全て及び第1のコンカテマーの第1の挿入配列の少なくとも一部分をシーケンシングし、それによって、第1のシーケンシングリード生成物を生成することを含むことができる。

40

【0147】

いくつかの態様において、操作510は、1つ以上のサイクルの各々において、第2のコンカテマーの第2の標的バーコード配列領域のみをシーケンシングし、それによって、

50

第2のシーケンシングリード生成物を生成することを含むことができる。いくつかの態様において、操作510は、1つ以上のサイクルの各々において、第2の標的バーコード配列領域及び第2のコンカテマーの第2の挿入配列の少なくとも一部分をシーケンシングし、それによって、第2のシーケンシングリード生成物を生成することを含むことができる。

【0148】

いくつかの態様において、操作510は、1つ以上のサイクルの各々において、第1のコンカテマー分子の第1の標的バーコード配列領域及び第2のコンカテマーの第2の標的バーコード配列領域のみをシーケンシングするが、第1のコンカテマーの第1の挿入配列のいかなる部分も、第2のコンカテマーの第2の挿入配列のいかなる部分をシーケンシングせず、それによって、第1のシーケンシングリード生成物及び第2のシーケンシングリード生成物を生成することを含む。いくつかの態様において、操作510は、1つ以上のサイクルの各々において、第1の標的バーコード配列領域のうちのいくつか又は全て及び第1のコンカテマーの第1の挿入配列の少なくとも一部分をシーケンシングし、それによって、第1のシーケンシングリード生成物を生成することを含むことができる。

10

【0149】

いくつかの態様において、支持体上に固定化された複数のコンカテマー分子の各々は、ポロニーに対応する。いくつかの態様において、支持体上に固定化された複数のコンカテマー分子の各々は、ベースコーリングテンプレートのベースコーリング位置に対応する。複数のコンカテマー分子のうちの2つ以上の異なるコンカテマー分子は、異なる挿入配列を有し得る。例示的な挿入配列を図8に示す。異なる挿入配列は、異なる標的RNA分子又は標的cDNA分子に対応する。

20

【0150】

ヌクレオチド多様性の固定化されたコンカテマー分子の集団は、シーケンシングサイクルに存在するヌクレオチドA、G、C、及びT/Uの相対的な割合を指すことができる。高又はバランスのとれた多様性データは、概して、シーケンシングランの各サイクルにおいて表される4つのヌクレオチド全てがほぼ等しい割合であるバーコード及び／又は挿入領域を含むことができる。低い又はアンバランスな多様性データは、概して、ある特定のヌクレオチドの割合が高く、かつ他のヌクレオチドの割合が低いバーコード及び／又は目的の配列（挿入）領域を含むことができる。低多様性ライブラリの問題を克服するために、Phixバクテリオファージから調製された少量の高多様性ライブラリは、典型的には、目的のライブラリ（例えば、Phixスパイクインライブラリ）と混合され、同じフローセル上で一緒にシーケンシングされる。Phixライブラリスパイクインライブラリは、ヌクレオチド多様性を提供することができるが、それはまた、フローセル上の空間を占有し、それによって、目的の配列を保有する標的ライブラリを置き換え、標的ライブラリから得られるシーケンシングデータの量を低減し（例えば、シーケンシングスループットを低減する）。低多様性ライブラリの問題を克服する別の方法は、カラーバランスがとれているように設計され、したがって高い多様性を有するいくつかのインデックス及び／若しくはバーコード配列を有する標的ライブラリ分子を調製することであり得る。しかしながら、多数のインデックス又はバーコードセット、例えば、16-plex、24-plex、96-plex、又はより大きなplexレベルのためのコンカテマーのセットを設計することが望ましい場合がある。全ての試料インデックス又はバーコード配列がカラーシグナルバランスがとれた大規模な試料インデックスセットのインデックス又はバーコード配列を設計することは困難であり得る。

30

40

【0151】

本明細書におけるフローセル画像は、支持体上に固定化された複数のコンカテマー分子間のバランスのとれた多様性のA、G、C、及びT/Uのヌクレオチド塩基に結合したヌクレオチド試薬から放出される光シグナルを含み得る。本明細書における複数のフローセル画像は、支持体上に固定化された複数のコンカテマー分子間のアンバランスな多様性のA、G、C及びT/Uのヌクレオチド塩基に結合したヌクレオチド試薬から放出される光

50

シグナルを含み得る。バランスのとれた多様性又はアンバランスな多様性の塩基は、特定のシーケンシングサイクル又は複数のシーケンシングサイクル、例えば本明細書のバーコード配列を包含するサイクルで判定され得る。

【0152】

バランスのとれた多様性の塩基を有することは、1つ以上のシーケンシングサイクルにおいてアンバランスな多様性の塩基を使用することよりも有利に、ベースコーリングテンプレートをより正確かつ確実に生成することを容易にし得る。いくつかの態様において、本明細書におけるバーコード配列は、1つ以上のシーケンシングサイクルにおいてバランスのとれた多様性の塩基を提供し得るように判定されてもよい。

【0153】

いくつかの態様において、第2の複数のフローセル画像は、1つ以上の後続のサイクルにおいて支持体上に固定化された複数のコンカテマー分子間のアンバランスな多様性のA、G、C、及びT/Uのヌクレオチド塩基に結合したヌクレオチド試薬から放出される光シグナルを含む。

【0154】

複数のコンカテマー分子間のアンバランスな多様性のA、G、C、及びT/Uのヌクレオチド塩基は、(2)視野内の全てのポロニーの塩基の総数に対する(1)1つ以上の種類のヌクレオチド塩基の数のパーセンテージを含み得、このパーセンテージは、1つ以上のサイクルにおいて20%、15%、10%、5%、3%、2%、又は1%未満である。

【0155】

複数のコンカテマー分子間のバランスのとれた多様性のA、G、C、及びT/Uのヌクレオチド塩基は、1つ以上のサイクルにおける(2)視野内の全てのポロニーの塩基の総数に対する(1)各種類のヌクレオチド塩基の数のパーセンテージを含み、そのパーセンテージは、5%、8%、10%、12%、15%、18%、又は20%を超える。例えば、アンバランスな多様性のヌクレオチド塩基は、サイクルのコンカテマー中の全てのヌクレオチド塩基の総数の約5%であるヌクレオチド塩基Aを含む。別の例として、アンバランスな多様性のヌクレオチド塩基は、サイクルのコンカテマー分子間の全てのヌクレオチド塩基の総数の約8%であるヌクレオチド塩基C及び約1%であるヌクレオチド塩基Tを含む。

【0156】

単一のシーケンシングラン中に、ベースコーリング位置、すなわち、ポロニー又はクラスタは、異なるサイクル及び/又はクロスチャネルから得られたフローセル画像内でシフト、回転、又はそうでなければ空間的に変換することができる。結果として、シーケンシングランにおけるベースコーリング位置が空間的にアラインメントされ、ベースコールが対応するポリニー又はクラスタ、テンプレート分子、及び試料に正確に割り当てられることを確実にするために、テンプレートが必要とされる。シーケンシングランの早期にテンプレートを生成し、テンプレートが生成された後の全てのシーケンシングサイクルにおけるベースコーリング位置の位置合わせを可能にすることが有利であり得る。テンプレートの早期生成はまた、有利には、後続のサイクルにおけるフローセル画像の一次分析の遅延を低減し、シーケンシングランがまだ進行中である間にリアルタイム分析を可能にすることができる。例えば、テンプレートが最初の3~5サイクルから生成される場合、一次分析は、サイクル6ではフローセル画像が取得された後にリアルタイムで実行され得、サイクル7ではシーケンシング及びイメージング操作と並行して実行され得る。サイクル6後の各サイクルの同様の分析は、その後続のシーケンシングサイクルがまだ進行中である間に実行することができる。したがって、一次分析は、シーケンシングランが完了したときではないにしても、完全にその直後であり得る。

【0157】

インサイチュバッチ特異的シーケンシングを行うための方法

いくつかの態様において、方法500は、本明細書に開示されるように、符号化されたパドロックプローブを使用して、インサイチュ多重化及びマルチオミクス検出及び識別を

10

20

30

40

50

含むインサイチュシーケンシング及びその分析を行うように構成される。パドロックプローブは、標的RNAを選択的に検出するように設計することができる。

【0158】

RNA特異的パドロックプローブは、標的RNAに対応するcDNAに選択的にハイブリダイズし得る。RNA特異的プローブは、cDNAを一意に識別するバーコードを保有し得る。いくつかの態様において、RNA特異的パドロックプローブはまた、バッチ特異的シーケンシングプライマー結合部位を保有し得る。

【0159】

異なる種類（例えば、2つ以上）のパドロックプローブを使用して、バッチ特異的シーケンシング結合部位及びバーコードの複数のコピーを有するコンカテマーを生成することができる。コンカテマーは、シーケンシング中に増加したシグナル強度及び色分化をもたらすコンパクトな形状及びサイズを有するDNAナノボールに崩壊することができる。

【0160】

インサイチュシーケンシングの場合、光学解像度の限界は、高度に多重化されたシーケンシングを実行する能力を妨げる可能性がある。パドロックプローブ上のバッチ特異的シーケンシングプライマー結合部位は、過密シグナル及び画像を低減するために、選択されたバッチ特異的シーケンシングプライマーを使用して、コンカテマーの所望のサブセット（例えば、バッチ）をシーケンシングすることを可能にしてもよく、その結果、（既存のインサイチュシーケンシング方法が使用するものよりも）単位イメージング空間当たりより高い細胞密度で細胞試料を調製することができ、シーケンシングスループットは、既存のインサイチュシーケンシング方法と比較して、バッチシーケンシングを使用して改善されてもよい。バッチ特異的シーケンシングプライマーを使用することにより、鮮明かつ解像可能な光学画像が生成される。異なるバッチ特異的シーケンシングプライマーを使用して、同じ細胞試料上で複数ラウンドのシーケンシングを行うことによって、多数の標的RNAを明らかにするための多重シーケンシングを可能にし得る。

【0161】

本明細書に記載されるバッチ特異的シーケンシングは、多くの用途を有し得る。例えば、イメージングされ、シーケンシングに関連するスポットの数をカウントすることができる。カウントされたスポットは、細胞試料中のRNAレベルの尺度として使用され得る。

【0162】

いくつかの態様において、方法500は、インサイチュバッチ特異的シーケンシング用途におけるベースコーリングテンプレートを判定するように構成され得る。いくつかの態様において、インサイチュバッチ特異的シーケンシングの文脈における方法は、固体支持体上に堆積された細胞試料を準備する操作を含み、細胞試料は、（i）第1の標的cDNA又はRNA分子に対応する第1の複数のDNAアンプリコン（例えば、第1のコンカテマー）、及び（ii）第2の標的cDNA又はRNA分子に対応する第2の複数のDNAアンプリコン（例えば、第2のコンカテマー）を保有する。

【0163】

いくつかの態様において、インサイチュバッチ特異的シーケンシングにおける操作510は、第2の複数のDNAアンプリコンのシーケンシングを阻害する条件下で、細胞試料内の第1の複数のDNAアンプリコンをシーケンシングする操作を含み、細胞試料内の第1の複数のDNAアンプリコンをシーケンシングすることは、複数の第1のシーケンシングリード生成物を生成することを含み、第1のシーケンシングリード生成物の配列は、第1の標的参照配列とアラインメントして、細胞試料中の第1の標的RNAの存在を確認する。いくつかの態様において、第1のアンプリコンは、2～30個以下のシーケンシングサイクルを行うことによって反復してシーケンシングされ得るか、又は1～250個のシーケンシングサイクルを行うことによって反復してシーケンシングされ得る。

【0164】

いくつかの態様において、インサイチュバッチ特異的シーケンシングにおける操作510は、第1の複数のDNAアンプリコンのシーケンシングを阻害する条件下で、細胞試料

内の第2の複数のDNAアンプリコンをシーケンシングする操作を更に含んでもよく、細胞試料内の第2の複数のDNAアンプリコンをシーケンシングすることは、複数の第2のシーケンシングリード生成物を生成することを含み、第2のシーケンシングリード生成物の配列は、第2の標的参照配列とアラインメントして、細胞試料中の第2の標的RNAの存在を確認する。いくつかの態様において、第2のアンプリコンは、2～30個以下のシーケンシングサイクルを行うことによって反復してシーケンシングされ得るか、又は1～250個のシーケンシングサイクルを行うことによって反復してシーケンシングされ得る。

【0165】

いくつかの態様において、インサイチュバッチ特異的シーケンシングにおける方法500は、固体支持体上に堆積された細胞試料を準備する操作を含み、細胞試料は、第1の複数の標的RNA及び第2の複数の標的RNAを保有する。いくつかの態様において、第1の複数の標的RNAは、第1のポリペプチドをコードする。いくつかの態様において、第2の複数の標的RNAは、第2のポリペプチドをコードする。いくつかの態様において、細胞試料は、固定され、透過化される。

【0166】

いくつかの態様において、細胞試料は、2～25個の異なる標的RNA分子を保有するか、又は25～50個の異なる標的RNA分子を保有するか、又は50～75個の異なる標的RNA分子を保有するか、又は75～100個の異なる標的RNA分子を保有する。いくつかの態様において、細胞試料は、100個を超える異なる標的RNA分子、又は250個を超える異なる標的RNA分子、又は500個を超える異なる標的分子、又は1000個を超える異なる標的RNA分子、又はそれ以上を保有する。いくつかの態様において、細胞試料は、10,000個を超える異なる標的RNA分子を有する。いくつかの態様において、細胞試料は、全細胞、複数の全細胞、無傷の組織、又は無傷の腫瘍を含む。いくつかの態様において、細胞試料は、新鮮な細胞試料、新鮮凍結細胞試料、切片化細胞試料、又はFFPE細胞試料を含む。いくつかの態様において、細胞試料は、固体支持体上に堆積される。いくつかの態様において、細胞試料は、細胞接着を促進するコーティングで不動態化された固体支持体上に堆積される。いくつかの態様において、細胞試料は、固定化されたキャプチャオリゴヌクレオチドを欠く支持体上に堆積される。いくつかの態様において、細胞試料は、以下に記載されるステップ(b)を行う前に培養される。

【0167】

いくつかの態様において、細胞試料は、2～25個の異なる標的ポリペプチド分子を保有するか、又は25～50個の異なる標的ポリペプチド分子を保有するか、又は50～75個の異なる標的ポリペプチド分子を保有するか、又は75～100個の異なる標的ポリペプチド分子を保有する。いくつかの態様において、細胞試料は、100個を超える異なる標的ポリペプチド分子、又は250個を超える異なる標的ポリペプチド分子、又は500個を超える異なる標的分子、又は1000個を超える異なる標的ポリペプチド分子、又はそれ以上を保有する。いくつかの態様において、細胞試料は、10,000個を超える異なる標的ポリペプチド分子を有する。標的ポリペプチド分子は、標的RNA分子によってコードされる。

【0168】

いくつかの態様において、インサイチュバッチ特異的シーケンシングにおける方法500は、(i)第1の複数の標的RNAから少なくとも第1の複数の標的cDNAを生成することと、(ii)第2の複数の標的RNAから少なくとも第2の複数の標的cDNAを生成することと、によって、細胞試料内に複数のcDNAを生成する操作を含む(例えば、図13)。いくつかの態様において、第1の標的cDNAは、第1の標的RNA分子に対応する。いくつかの態様において、第2の標的cDNAは、第2の標的RNA分子に対応する。いくつかの態様において、方法は、2～10,000個の異なる標的RNA分子に対応する少なくとも2～10,000個の異なる標的cDNA分子を生成することを含む。いくつかの態様において、細胞試料内で複数のcDNAを生成する操作は、細胞試料

内の複数のRNAを、(i) 複数の逆転写プライマー、(ii) 複数の逆転写酵素、及び (iii) 複数のヌクレオチドと、細胞試料中で複数のcDNA分子（例えば、複数の第1の鎖cDNA分子）を生成するための逆転写反応を行うのに好適な条件下で接触させる操作を含む。いくつかの態様において、複数の逆転写プライマーは、第1の標的RNAに選択的にハイブリダイズする標的特異的逆転写プライマーの第1のサブ集団を含み、かつ／又は第2の標的RNAに選択的にハイブリダイズする標的特異的逆転写プライマーの第2のサブ集団を含む。いくつかの態様において、複数の逆転写プライマーは、第1の標的RNAにハイブリダイズするランダム配列逆転写プライマーの第1のサブ集団を含み、第2の標的RNAにハイブリダイズするランダム配列逆転写プライマーの第2のサブ集団を含む。

10

【0169】

いくつかの態様において、インサイチュバッチ特異的シーケンシングにおける方法500は、細胞試料内で、第1及び第2の複数の標的RNA分子に対応する複数のDNAコンカテマーを生成する操作を含み、この操作は、(1) 第1の複数の標的cDNAを第1の複数のパドロックプローブと接触させることによって、第1の複数の共有結合性閉鎖環状パドロックプローブを生成することであって、接触させることが、第1のパドロックプローブの第1の結合アーム及び第2の結合アームを、それらそれぞれの第1の標的cDNA分子上の近位位置にハイブリダイズして、ハイブリダイズした第1の結合アームと第2の結合アームとの間に各々ニック又はギャップを有する第1の複数の環状パドロックプローブを形成するのに好適である条件下で行われ、第1のパドロックプローブが、(i) 第1の標的RNA又はcDNAを一意に識別する第1の標的バーコード配列（標的BC-1）、(ii) 第1のバッチ特異的シーケンシングプライマー結合部位（バッチSeq-1）（又はその相補的配列）、及び(iii) 増幅プライマーのためのユニバーサル結合部位（ユニバーサルRCA）（又はその相補的配列）（例えば、図13、左）を含む、生成することと、(2) 第1の複数の共有結合性閉鎖環状パドロックプローブ内のニック又はギャップを酵素的に閉鎖して、第1の複数の共有結合性閉鎖パドロックプローブを形成することと、(3) 第1の共有結合性閉鎖環状パドロックプローブをテンプレート分子として使用して、細胞試料内でローリングサークル増幅を行い、それによって、第1の複数の標的RNA又はcDNA分子に対応する第1の複数のコンカテマー分子を生成することと、を含む。いくつかの態様において、ローリングサークル増幅反応は、複数のコンパクションオリゴヌクレオチドの存在下又は不在下で行われ得る。いくつかの態様において、本方法は、細胞試料中の複数のcDNA分子を、少なくとも2～10,000個の異なる標的特異的パドロックプローブと接触させることを含む。いくつかの態様において、第1のパドロックプローブは、ユニバーサルコンパクションオリゴヌクレオチド結合部位（又はその相補的配列）を更に含む。いくつかの態様において、第1の環状化パドロックプローブにおけるニックを閉じることは、酵素ライゲーション反応を行うことを含む。いくつかの態様において、第1の環状化パドロックプローブ内のギャップを閉じることは、第1の標的cDNA分子をテンプレートとして使用してポリメラーゼ触媒充填反応を行うことと、酵素ライゲーション反応を行うことと、を含む。いくつかの態様において、方法は、酵素反応を行うことによって、少なくとも2～10,000個の環状化標的特異的パドロック

20

30

40

【0170】

50

いくつかの態様において、インサイチュバッチ特異的シーケンシングにおける方法500は、細胞試料内で、第2の複数の標的RNA分子に対応する複数のDNAコンカテマーを生成する操作を含み、この操作は、(1)第2の複数の標的cDNAを第1の複数のパドロックプローブと接触させることによって、第2の複数の共有結合性閉鎖環状パドロックプローブを生成することであって、接触させることが、第2のパドロックプローブの第1の結合アーム及び第2の結合アームを、それらそれぞれの第2の標的cDNA分子上の近位位置にハイブリダイズして、ハイブリダイズした第1の結合アームと第2の結合アームとの間に各々ニック又はギャップを有する第2の複数の環状パドロックプローブを形成するのに好適である条件下で行われ、第2のパドロックプローブが、(i)第2の標的cDNA又はRNAを一意に識別する第2の標的バーコード配列(標的BC-2)、(ii)第2のバッチ特異的シーケンシングプライマー結合部位(バッチSeq-2)(又はその相補的配列)であって、第2のバッチ特異的シーケンシングプライマー結合部位の配列が、第1のバッチ特異的シーケンシングプライマー結合部位の配列と異なる、第2のバッチ特異的シーケンシングプライマー結合部位、及び(iii)増幅プライマーのためのユニバーサル結合部位(ユニバーサルRCA)(又はその相補的配列)(例えば、図13、右)を含む、生成することと、(2)第2の複数の共有結合性閉鎖環状パドロックプローブ内のニック又はギャップを酵素的に閉鎖して、第2の複数の共有結合性閉鎖パドロックプローブを形成することと、(3)第2の共有結合性閉鎖環状パドロックプローブをテンプレート分子として使用して、細胞試料内でローリングサークル増幅を行い、それによって、第2の複数の標的RNAに対応する第2の複数のコンカテマー分子を生成することと、を含む。いくつかの態様において、ローリングサークル増幅反応は、複数のコンパクションオリゴヌクレオチドの存在下又は不在下で行われ得る。いくつかの態様において、本方法は、細胞試料中の複数のcDNA分子を、少なくとも2~10,000個の異なる標的特異的パドロックプローブと接触させることを含む。いくつかの態様において、第2のパドロックプローブは、ユニバーサルコンパクションオリゴヌクレオチド結合部位(又はその相補的配列)を更に含む。いくつかの態様において、第2の環状化パドロックプローブにおけるニックを閉じることは、酵素ライゲーション反応を行うことを含む。いくつかの態様において、第2の環状化パドロックプローブ内のギャップを閉じることは、第2の標的cDNA分子をテンプレートとして使用してポリメラーゼ触媒充填反応を行うことと、酵素ライゲーション反応を行うことと、を含む。いくつかの態様において、方法は、酵素反応を行うことによって、少なくとも2~10,000個の環状化標的特異的パドロックプローブ内のニック又はギャップを閉鎖し、それによって、細胞試料内に少なくとも2~10,000個の共有結合性閉鎖環状パドロックプローブを生成することを含む。いくつかの態様において、第2の複数のうちの各コンカテマー分子は、タンデムリピート単位を含み、単位は、第2の標的cDNAの配列と、(i)第2の標的cDNA又はRNAを一意に特定する第2の標的バーコード配列(標的BC-2)、(ii)第2のバッチ特異的シーケンシングプライマー結合部位(バッチSeq-2)(又はその相補的配列)、及び(iii)増幅プライマーのためのユニバーサル結合部位(ユニバーサルRCA)(又はその相補的配列)と、を含む。いくつかの態様において、単位は、ユニバーサルコンパクションオリゴヌクレオチド結合部位(又はその相補的配列)を更に含む。

10

20

30

40

【0171】

いくつかの態様において、インサイチュバッチ特異的シーケンシングにおいて1つ以上のシーケンシング反応サイクルを行うことは、第2の複数のコンカテマーのシーケンシングを阻害する条件下で、細胞試料内の第1の複数のコンカテマー分子をシーケンシングすることを含む(例えば、図14)。いくつかの態様において、細胞試料内の第1の複数のコンカテマーをシーケンシングすることは、2~30個以下のシーケンシングサイクルを行って、複数の第1のシーケンシングリード生成物を生成することを含み、第1のシーケンシングリード生成物の配列を、第1の標的参照配列とアラインメントして、細胞試料中の第1の標的RNAの存在を確認する。いくつかの態様において、細胞試料内の第1の複数のコンカテマーをシーケンシングすることは、1~250個のシーケンシングサイクル

50

を行って、複数の第1のシーケンシングリード生成物を生成することを含み、第1のシーケンシングリード生成物の配列を、第1の標的参照配列とアラインメントして、細胞試料中の第1の標的RNAの存在を確認する。

【0172】

いくつかの態様において、第1のコンカテマー分子では、第1の標的バーコード領域（標的BC-1）のみがシーケンシングされる。いくつかの態様において、第1のコンカテマー分子において、第1の標的バーコード（標的BC-1）の少なくとも一部分又は全長がシーケンシングされる。いくつかの態様において、第1のコンカテマー分子において、第1の標的バーコード（標的BC-1）がシーケンシングされ、第1のcDNA領域の一部分がシーケンシングされる。

10

【0173】

いくつかの態様において、インサイチュバッチ特異的シーケンシングにおいて第1のコンカテマーをシーケンシングすることは、細胞試料内の第1の複数のコンカテマー分子を、(i) 複数の第1のバッチ特異的シーケンシングプライマー、(ii) 複数のシーケンシングポリメラーゼ、及び(iii) 複数のヌクレオチド試薬と、複数の第1のバッチ特異的シーケンシングプライマーを第1のコンカテマー上のそれらのそれぞれの第1のバッチ特異的シーケンシングプライマー結合部位にハイブリダイズするのに好適な条件下で接触させるステップ(1)を含む。いくつかの態様において、シーケンシングすることは、テンプレート分子として第1のコンカテマーを使用して、第1の複数のシーケンシングリード生成物を生成するために、2〜30回以下のシーケンシングサイクルを行うステップ(2)を更に含む。

20

【0174】

いくつかの態様において、本明細書におけるインサイチュバッチ特異的シーケンシングにおけるシーケンシング操作は、1.0mm²を超える視野(FOV)を含む光学イメージングシステムを使用して、第1の核酸コンカテマーの少なくとも一部分をシーケンシングすることを含む。

【0175】

いくつかの態様において、複数の第1のシーケンシングリード生成物は、イメージングによって検出可能であり、シーケンシングすることは、2〜30個以下のシーケンシングサイクル中に取得された画像から、又は1〜250個のシーケンシングサイクル中に取得された画像から、複数の第1のシーケンシングリード生成物を復号することを含む。

30

【0176】

いくつかの態様において、インサイチュバッチ特異的シーケンシングにおける方法500は、複数の第1のシーケンシングリード生成物を第1のコンカテマー分子から除去することと、第1のコンカテマー分子を細胞試料中に保持することと、を含む。いくつかの態様において、更なるシーケンシング反応を阻害するために、3'ブロック部分第1のシーケンシングリード生成物に付加することができる。例えば、ヌクレオチド類似体が後続のヌクレオチドの組み込みを阻害する場合、ヌクレオチド類似体を組み込むことができる。例示的なブロックヌクレオチド類似体としては、ジデオキシヌクレオチド又は2'若しくは3'鎖終結部分を有するヌクレオチドが挙げられる。

40

【0177】

いくつかの態様において、インサイチュバッチ特異的シーケンシングにおける方法は、ステップ(d)及び(e)を少なくとも1回繰り返すことによって、複数の第1のコンカテマーを反復してシーケンシングすることを更に含む。いくつかの態様において、反復シーケンシングは、任意選択である。

【0178】

いくつかの態様において、インサイチュバッチ特異的シーケンシングにおいて第1のコンカテマーをシーケンシングすることは、細胞試料内の第1の複数のコンカテマー分子を、(i) 複数の第1のバッチ特異的シーケンシングプライマー、(ii) 複数のシーケンシングポリメラーゼ、及び(iii) 複数のヌクレオチド試薬と、複数の第1のバッチ特

50

異的シーケンシングプライマーを第1のコンカテマー上のそれらのそれぞれの第1のバッチ特異的シーケンシングプライマー結合部位にハイブリダイズするのに好適な条件下で接触させるステップ(1)を含む。いくつかの態様において、シーケンシングすることは、テンプレート分子として第1のコンカテマーを使用して、第1の複数のシーケンシングリード生成物を生成するために、2~30回以下のシーケンシングサイクルを行うステップ(2)を更に含む。いくつかの態様において、シーケンシングすることは、第1の複数のシーケンシングリード生成物を第1のコンカテマーから除去し、複数の第1のコンカテマーを細胞試料中に保持するステップ(3)を更に含む。いくつかの態様において、シーケンシングすることは、ステップ(1)~(3)を少なくとも1回繰り返すステップ(4)を更に含む(例えば、図14)。いくつかの態様において、ステップ(4)は、ステップ(1)~(3)を少なくとも2回、少なくとも3回、少なくとも4回、少なくとも5回、少なくとも6回、少なくとも7回、少なくとも8回、少なくとも9回、又は少なくとも10回繰り返すことを含む。いくつかの態様において、ステップ(4)は、ステップ(1)~(3)を最大10回、最大20回、最大30回、最大40回、又は最大50回繰り返すことを含む。

10

【0179】

いくつかの態様において、第1のコンカテマーの反復シーケンシングは、結合によるシーケンシング手順、標識及び／若しくは非標識鎖終結ヌクレオチド、又は多価分子を使用して行うことができる。これらの3つのシーケンシング方法の説明は、本明細書に記載されている。

20

【0180】

いくつかの態様において、複数のユニバーサルシーケンシングプライマーをホルムアミド(例えば、10~20%ホルムアミド)を含むSSC緩衝液(例えば、2倍の生理食塩水-クエン酸ナトリウム)緩衝液を含むハイブリダイゼーション試薬と、コンカテマーテンプレート分子にハイブリダイズさせることができる。ハイブリダイゼーション条件は、約10~60分間、約20~30℃の温度を含む。

【0181】

いくつかの態様において、複数のシーケンシングリード生成物をコンカテマーから除去することができ、複数のコンカテマーは、例えば30~90℃などの核酸変性を促進する温度で、ホルムアミドの有無にかかわらず、SSC緩衝液(例えば、生理食塩水-クエン酸ナトリウム)緩衝液を含む脱ハイブリダイゼーション試薬を使用して細胞試料中に保持することができる。

30

【0182】

いくつかの態様において、インサイチュバッチ特異的シーケンシングにおける方法500は、第1の複数のコンカテマーのシーケンシングを阻害する条件下で、細胞試料内の第2の複数のコンカテマー分子をシーケンシングすることを更に含む(例えば、図14)。いくつかの態様において、細胞試料内の第2の複数のコンカテマーをシーケンシングすることは、2~30個以下のシーケンシングサイクルを行って、複数の第2のシーケンシングリード生成物を生成することを含み、第2のシーケンシングリード生成物の配列を、第2の標的参照配列とアラインメントして、細胞試料中の第2の標的RNAの存在を確認する。いくつかの態様において、細胞試料内の第2の複数のコンカテマーをシーケンシングすることは、1~250個のシーケンシングサイクルを行って、複数の第2のシーケンシングリード生成物を生成することを含み、第2のシーケンシングリード生成物の配列を、第2の標的参照配列とアラインメントして、細胞試料中の第2の標的RNAの存在を確認する。

40

【0183】

いくつかの態様において、第2のコンカテマー分子では、第2の標的バーコード領域(標的BC-2)のみがシーケンシングされる。いくつかの態様において、第2のコンカテマー分子において、第2の標的バーコード(標的BC-2)の少なくとも一部分又は全長がシーケンシングされる。いくつかの態様において、第2のコンカテマー分子において、

50

第2の標的バーコード（標的BC-2）がシーケンシングされ、第2のcDNA領域の一部がシーケンシングされる。

【0184】

いくつかの態様において、第2のコンカテマーをシーケンシングすることは、細胞試料内の第2の複数のコンカテマー分子を、（i）複数の第2のバッチ特異的シーケンシングプライマー、（ii）複数のシーケンシングポリメラーゼ、及び（iii）複数のヌクレオチド試薬と、複数の第2のバッチ特異的シーケンシングプライマーを第2のコンカテマー上のそれらのそれぞれの第2のバッチ特異的シーケンシングプライマー結合部位にハイブリダイズするのに好適な条件下で接触させるステップ（1）を含む。いくつかの態様において、シーケンシングすることは、テンプレート分子として第2のコンカテマーを使用して、第2の複数のシーケンシングリード生成物を生成するために、2〜30回以下のシーケンシングサイクルを行うステップ（2）を更に含む。

10

【0185】

いくつかの態様において、第2の複数のコンカテマー分子をシーケンシングすることは、1.0mm²を超える視野（FOV）を含む光学イメージングシステムを使用して、第2の核酸コンカテマーの少なくとも一部分をシーケンシングすることを含む。

【0186】

いくつかの態様において、第2の複数のコンカテマー分子のシーケンシングにおいて、複数の第2のシーケンシングリード生成物は、イメージングによって検出可能であり、シーケンシングは、2〜30個以下のシーケンシングサイクル中に得られた画像から、又は1〜250個のシーケンシングサイクル中に得られた画像から、複数の第2のシーケンシングリード生成物を復号することを含む。

20

【0187】

いくつかの態様において、バッチ特異的インサイチュシーケンシングにおける方法500は、複数の第2のシーケンシングリード生成物を第2のコンカテマー分子から除去することと、第2のコンカテマー分子を細胞試料中に保持することと、を更に含む。いくつかの態様において、更なるシーケンシング反応を阻害するために、3'ブロック部分を第2のシーケンシングリード生成物に付加することができる。例えば、ヌクレオチド類似体が後続のヌクレオチドの組み込みを阻害する場合、ヌクレオチド類似体を組み込むことができる。例示的なブロックヌクレオチド類似体としては、ジデオキシヌクレオチド又は2'若しくは3'鎖終結部分を有するヌクレオチドが挙げられる。

30

【0188】

いくつかの態様において、バッチ特異的インサイチュシーケンシングにおける方法500は、（1）第1の複数のコンカテマーのシーケンシングを阻害する条件下で、細胞試料内の第2の複数のコンカテマー分子をシーケンシングすることと、（2）第2のコンカテマー分子から複数の第2のシーケンシングリード生成物を除去し、第2のコンカテマー分子を細胞試料内に保持することと、を少なくとも1回繰り返すことによって、複数の第2のコンカテマーを反復してシーケンシングすることを更に含む。いくつかの態様において、反復シーケンシングは、任意選択である。

【0189】

いくつかの態様において、第2のコンカテマーをシーケンシングすることは、細胞試料内の第2の複数のコンカテマー分子を、（i）複数の第2のバッチ特異的シーケンシングプライマー、（ii）複数のシーケンシングポリメラーゼ、及び（iii）複数のヌクレオチド試薬と、複数の第2のバッチ特異的シーケンシングプライマーを第2のコンカテマー上のそれらのそれぞれの第2のバッチ特異的シーケンシングプライマー結合部位にハイブリダイズするのに好適な条件下で接触させるステップ（1）を含む。いくつかの態様において、シーケンシングすることは、テンプレート分子として第2のコンカテマーを使用して、第1の複数のシーケンシングリード生成物を生成するために、2〜30回以下のシーケンシングサイクルを行うステップ（2）を更に含む。いくつかの態様において、シーケンシングすることは、第2の複数のシーケンシングリード生成物を第1のコンカテマー

40

50

から除去し、複数の第2のコンカテマーを細胞試料中に保持するステップ(3)を更に含む。いくつかの態様において、シーケンシングすることは、ステップ(1)～(3)を少なくとも1回繰り返すステップ(4)を更に含む(例えば、図14)。いくつかの態様において、ステップ(4)は、ステップ(1)～(3)を少なくとも2回、少なくとも3回、少なくとも4回、少なくとも5回、少なくとも6回、少なくとも7回、少なくとも8回、少なくとも9回、又は少なくとも10回繰り返すことを含む。いくつかの態様において、ステップ(4)は、ステップ(1)～(3)を最大10回、最大20回、最大30回、最大40回、又は最大50回繰り返すことを含む。

【0190】

いくつかの態様において、ステップ(i)の第2のコンカテマーの反復シーケンシングは、結合によるシーケンシング手順、標識及び／若しくは非標識鎖終結ヌクレオチド、又は多価分子を使用して行うことができる。これらの3つのシーケンシング方法の説明は、以下に記載されている。

【0191】

いくつかの態様において、操作において使用される複数のヌクレオチド試薬は、検出可能に標識されているか、又は標識されていない複数のヌクレオチドを含む。いくつかの態様において、個々のヌクレオチドは、検出可能なレポーター部分に連結されている。いくつかの態様において、検出可能なレポーター部分は、フルオロフォアを含む。いくつかの態様において、複数の検出可能に標識されたヌクレオチド類似体は、複数の鎖終結ヌクレオチドを含み、鎖終結部分は、3'ヌクレオチド糖位に連結されて、3'ブロックされたヌクレオチド類似体を形成する。いくつかの態様において、鎖終結部分を除去して、3'ブロックされたヌクレオチド類似体を、糖上に3'OH基を有する伸長可能なヌクレオチドに変換することができる。いくつかの態様において、標識されたヌクレオチド類似体は、ヌクレオ塩基のアデニン、シトシン、グアニン、チミン、又はウラシルに対応する異なるフルオロフォアに連結され、異なるフルオロフォアは、蛍光シグナルを発する。いくつかの態様において、シーケンシングサイクルは、(1)シーケンシングポリメラーゼ及び検出可能に標識された鎖終結ヌクレオチドと、シーケンシングプライマーの末端への検出可能に標識された鎖終結ヌクレオチドのポリメラーゼ触媒組み込みに適した条件下で、コンカテマー／シーケンシングプライマー二重鎖を接触させることと、(2)組み込まれた鎖終結ヌクレオチドによって放出される蛍光シグナル及び色を検出及びイメージングすることと、(3)組み込まれたヌクレオチドから鎖終結部分(例えば、ブロック解除)及びフルオロフォアを除去し、コンカテマー／シーケンシングプライマー二重鎖を保持することと、を含む。いくつかの態様において、細胞試料内の複数のコンカテマーに対して2～30回以下のシーケンシングサイクルを行って、複数のシーケンシングリード生成物を生成する。いくつかの態様において、第1のシーケンシングリード生成物の配列を判定し、第1の参照配列とアラインメントして、細胞試料内の第1の標的RNA分子の存在を確認することができる。いくつかの態様において、第2のシーケンシングリード生成物の配列を判定し、第2の参照配列とアラインメントして、細胞試料内の第2の標的RNA分子の存在を確認することができる。

【0192】

いくつかの態様において、第1及び第2のシーケンシングリード生成物の配列は、30塩基以下の長さの第1及び第2のシーケンシングリード生成物を生成する各ラウンドの後、又は第1及び第2のシーケンシングリード生成物が30塩基以下の長さの生成物を読み取る反復シーケンシングリード生成物のセットを生成した後にアラインメントすることができる。いくつかの態様において、シーケンシング反応は、細胞試料内でのシーケンシング反応から蛍光シグナルをキャプチャする検出器を有するシーケンシング装置上で行われる。シーケンシング装置は、検出器によってキャプチャされた蛍光シグナルデータを、細胞試料中で共存する異なる蛍光スポットの画像を表示するようにプログラムされたコンピュータシステムに中継するように構成され得、個々の蛍光スポットは、異なる標的RNA分子に対応する。いくつかの態様において、シーケンシングが、異なるヌクレオ塩基(例

えば、A、G、C、T/U) に対応する異なる蛍光標識ヌクレオチド試薬を使用して行われる場合、画像は、異なるシーケンシングサイクルで同じ細胞試料に共存する異なる色の蛍光スポットを有することができる。

【0193】

いくつかの態様において、非同期フェーズ及び／又はプレフェーズ事象は、クローン的に増幅されたテンプレートアンプリコン上の同期シーケンシング反応中に生じ得、シーケンシング反応は、検出可能に標識された鎖ターミネーターヌクレオチドを用いたポリメラーゼ触媒シーケンシング反応を含む。いくつかの態様において、クローン的に増幅されたテンプレート分子内の1つのテンプレート分子上のシーケンシング反応は、クローン的に増幅されたテンプレート分子内の他のテンプレート分子のシーケンシングの前進（例えば、事前フェージング）又は後退（例えば、フェージング）する。シーケンシング中、典型的には、標識された鎖ターミネーターヌクレオチドの組み込みに対応する蛍光シグナルが検出される。したがって、フェージング及び事前フェージング事象は、標識されたチェーンターミネーターヌクレオチドの組み込みを使用して検出及び監視することができる。

10

【0194】

いくつかの態様において、ステップ(d)及び(g)の複数のヌクレオチド試薬は、複数のヌクレオチドアームに付着するコアを各々含む複数の多価分子を含み、ヌクレオチドアームは、ヌクレオチド単位付着合する。いくつかの態様において、個々の多価分子は、検出可能なレポーター部分に連結されている。いくつかの態様において、検出可能なレポーター部分は、フルオロフォアを含む。いくつかの態様において、多価分子のコアは、フルオロフォアで標識され、多価分子の所与のコアに結合するフルオロフォアは、ヌクレオチドアームのヌクレオチド塩基（例えば、アデニン、グアニン、シトシン、チミン、又はウラシル）に対応する。いくつかの態様において、多価分子のヌクレオチドアームの少なくとも1つは、フルオロフォアに結合するリンカー及び／又はヌクレオチド塩基を含み、所与のヌクレオチド塩基に結合するフルオロフォアは、ヌクレオチドアームのヌクレオチド塩基（例えば、アデニン、グアニン、シトシン、チミン、又はウラシル）に対応する。いくつかの態様において、シーケンシングサイクルは、(1)コンカテマー／シーケンシングプライマー二重鎖を第1のシーケンシングポリメラーゼと接触させて、複合型ポリメラーゼを形成することと、(2)複合型ポリメラーゼを検出可能に標識された多価分子と、多価分子の相補的ヌクレオチド単位を複合型ポリメラーゼに結合させるのに好適な条件で接触させ、それによって多価結合複合体を形成することであって、この条件は、シーケンシングプライマーの末端への相補的ヌクレオチド単位の組み込みを阻害するのに好適である、接触させることと、(3)結合した、検出可能に標識された多価分子によって放出される蛍光シグナル及び色をイメージングすることと、(4)第1のシーケンシングポリメラーゼ及び結合した、検出可能に標識された多価分子を除去し、コンカテマー／シーケンシングプライマー二重鎖を保持することと、(5)保持されたコンカテマー／シーケンシングプライマー二重鎖を、シーケンシングプライマーの末端への非標識鎖終結ヌクレオチドの組み込みに好適な条件下で、第2のシーケンシングポリメラーゼ及び非標識鎖終結ヌクレオチドと接触させることと、(6)鎖終結部分を除去し（例えば、ブロック解除し）、コンカテマー／シーケンシングプライマー二重鎖を保持することと、を含む。いくつかの態様において、細胞試料内の複数のコンカテマーに対して2〜30回以下のシーケンシングサイクルを行って、複数のシーケンシングリード生成物を生成する。いくつかの態様において、第1のシーケンシングリード生成物の配列を判定し、第1の参照配列とアラインメントして、細胞試料内の第1の標的RNA分子の存在を確認することができる。いくつかの態様において、第2のシーケンシングリード生成物の配列を判定し、第2の参照配列とアラインメントして、細胞試料内の第2の標的RNA分子の存在を確認することができる。いくつかの態様において、第1及び第2のシーケンシングリード生成物の配列は、30塩基以下の長さの第1及び第2のシーケンシングリード生成物を生成する各ラウンドの後、又は第1及び第2のシーケンシングリード生成物が30塩基以下の長さの生成物を読み取る反復シーケンシングリード生成物のセットを生成した後にアラインメントする

20

30

40

50

ことができる。いくつかの態様において、シーケンシング反応は、細胞試料内でのシーケンシング反応から蛍光シグナルをキャプチャする検出器を有するシーケンシング装置上で行われる。シーケンシング装置は、検出器によってキャプチャされた蛍光シグナルデータを、細胞試料中で共存する異なる蛍光スポットの画像を表示するようにプログラムされたコンピュータシステムに中継するように構成され得、個々の蛍光スポットは、異なる標的RNA分子に対応する。いくつかの態様において、個々のサイクル時間は、30分未満で達成され得る。いくつかの態様において、視野（FOV）は、 1 mm^2 を超えることができ、大面積（ $>10\text{ mm}^2$ ）を走査するためのサイクル時間は、5分未満であり得る。

【0195】

図13は、第1及び第2の標的RNA分子から（それぞれ）第1及び第2のcDNAを生成すること、第1及び第2のパドロックプローブを（それぞれ）第1及び第2のcDNA分子にハイブリダイズして、（それぞれ）第1及び第2の環状化パドロックプローブを生成すること、を含む、環状化パドロックプローブを生成するためのワークフローを示す概略図である。第1のパドロックプローブは、（i）第1の標的RNAを一意に識別する第1の標的バーコード配列（標的BC-1）、（ii）第1のバッチ特異的シーケンシングプライマー結合部位（バッチSeq-1）（又はその相補的配列）、（iii）増幅プライマー（ユニバーサルRCA）（又はその相補的配列）のためのユニバーサル結合部位、及び（iv）コンパクションオリゴヌクレオチド（又はその相補的配列）のためのユニバーサル結合部位を含む。第2のパドロックプローブは、（i）第2の標的RNAを一意に識別する第2の標的バーコード配列（標的BC-2）、（ii）第2のバッチ特異的シーケンシングプライマー結合部位（バッチSeq-2）（又はその相補的配列）、（iii）増幅プライマーのためのユニバーサル結合部位（ユニバーサルRCA）（又はその相補的配列）、及び（iv）コンパクションオリゴヌクレオチドのためのユニバーサル結合部位（又はその相補的配列）を含む。

【0196】

図14は、第1及び第2の共有結合性閉鎖環状分子を（それぞれ）使用してローリングサークル増幅を行うことによって第1及び第2のコンカテマーを生成することを含む、ローリングサークル及びシーケンシングワークフローを示す概略図である。第1及び第2のコンカテマーは、第1のバッチ特異的シーケンシングプライマー、シーケンシングポリメラーゼ、及び複数のヌクレオチド試薬を使用した第1のシーケンシングワークフローに供される。第1のコンカテマーは、反復シーケンシングを受けるが、第2のコンカテマーは、反復シーケンシングを受けない。第1及び第2のコンカテマーは、第2のバッチ特異的シーケンシングプライマー、シーケンシングポリメラーゼ、及び複数のヌクレオチド試薬を使用した第2のシーケンシングワークフローに供される。第2のコンカテマーは、反復シーケンシングを受けるが、第1のコンカテマーは、反復シーケンシングを受けない。

【0197】

方法500は、第1の複数のフローセル画像のピクセルについてのピクセル強度及びピクセル強度の各々のそれぞれの色純度を判定する操作520を含むことができる。操作520は、本明細書に開示されるプロセッサ404によって実行され得る。ピクセル強度の各々は、ピクセルの強度及び／又はピクセルに対応するサブピクセルの1つ以上のサブピクセル強度を含むことができる。例えば、サブピクセルは、ピクセル内にあってもよい。ピクセル強度の各々のそれぞれの色純度は、ピクセルの1つの色純度及び／又は対応するピクセルの1つ以上のサブピクセル強度に対応する1つ以上の色純度を含むことができる。ピクセル強度の個々のピクセル強度のそれぞれの色純度は、1つ以上の色チャンネルのそれぞれの色純度を含むことができる。

【0198】

インサイチュバッチ特異的シーケンシングにおいて、操作520を、個々のバッチの対応する複数のフローセル画像に対して繰り返してもよい。例えば、4つのバッチがある場合、操作520は、同じバッチからのフローセル画像に対してのみ、4回及び毎回実行することができる。

10

20

30

40

50

【0199】

いくつかの態様において、ピクセル強度及びそれぞれの色純度を判定する操作520は、第1の軸方向位置でそのようなピクセルの各々についてのピクセル強度及びそれぞれの色純度を判定することと、第2の軸方向位置でそのようなピクセルの各々についてのピクセル強度及びそれぞれの色純度を判定することと、を含むことができる。第1及び第2の軸方向位置は、フローセル画像を生成するための異なる所定の軸方向位置にあり得る。例えば、第1の軸方向位置は、 $z = 0$ においてであり得、第2の軸方向位置は、その隣接する z レベルにおいてであり得、 $z = 0.1$ 、 0.2 、又は $0.3 \mu\text{m}$ である。

【0200】

いくつかの態様において、「ピクセル」は、当業者に既知のデジタル画像（例えば、フローセル画像）の2D空間要素であり得る。インサイチュシーケンシングの文脈におけるいくつかの態様において、「ピクセル」は、当業者に既知の「ボクセル」に相当する3D空間要素である。「ピクセル」が3D空間要素である態様では、フローセル画像は依然として2Dであり得るが、 $0.3 \mu\text{m} \sim 2 \mu\text{m}$ を超える被写界深度を有する。フローセル画像が依然として2Dであり得るが、 $0.5 \mu\text{m}$ 、 $0.6 \mu\text{m}$ 、 $0.7 \mu\text{m}$ 、 $0.8 \mu\text{m}$ 、 $0.9 \mu\text{m}$ 、 $1 \mu\text{m}$ 、 $1.1 \mu\text{m}$ 、 $1.2 \mu\text{m}$ 、 $1.3 \mu\text{m}$ 、 $1.5 \mu\text{m}$ 、 $1.8 \mu\text{m}$ 、 $2 \mu\text{m}$ 、 $2.2 \mu\text{m}$ 、 $2.5 \mu\text{m}$ 、 $2.8 \mu\text{m}$ 、 $3 \mu\text{m}$ 、 $3.5 \mu\text{m}$ 、 $3.8 \mu\text{m}$ 、 $4 \mu\text{m}$ 、 $4.5 \mu\text{m}$ 、 $4.8 \mu\text{m}$ 、 $5 \mu\text{m}$ 、 $5.5 \mu\text{m}$ 、 $6 \mu\text{m}$ 、 $6.5 \mu\text{m}$ 、 $7 \mu\text{m}$ 、 $8 \mu\text{m}$ 、 $9 \mu\text{m}$ 又はそれ以上である被写界深さを有する場合、「ピクセル」が3D空間要素である態様において、2Dフローセル画像は、3Dボリュームを覆うために、画像平面に直交する第3の次元に積み重ねられ得る。いくつかの態様において、フローセル画像は、単一の2D平面又は複数の2D平面内にあり、各フローセル画像は、第3の次元における焦点深度を有する。いくつかの態様において、ピクセルは二次元である。ピクセルは、画像平面内にあってもよい。いくつかの態様において、ピクセルは、三次元である。ピクセルは、画像平面内の2つの次元及び軸方向軸（すなわち、 z 軸）と平行な第3の次元を有し得る。

【0201】

ピクセル強度を判定することは、チャンネル強度のセット内の各チャンネル強度を判定することを含み得、各チャンネルは、それぞれ異なる蛍光波長に対応する。そのような判定は、対応するピクセル又は1つ以上のサブピクセル位置におけるチャンネル強度のセットの比較に基づくことができる。

【0202】

ピクセル強度のそれぞれの色純度を判定することは、（1）特定の種類のヌクレオチド塩基に対応するシグナル及び任意選択の第1の定数と、（2）他の種類のヌクレオチド塩基についてのシグナルの総量及び任意選択の第2の定数との定数と、を判定することを含むことができる。第1及び第2の定数は、異なる又は同一であり得る。いくつかの態様において、ピクセル強度の各々のそれぞれの色純度を判定することは、本明細書に開示される操作330を少なくとも部分的に含むことができる。

【0203】

方法500は、操作520で判定されたピクセル強度のピクセル強度及びそれぞれの色純度に基づいて、ベースコーリング位置を含むベースコーリングテンプレートを判定する操作530を含むことができる。

【0204】

いくつかの態様において、本明細書におけるフローセル画像は、単一の軸方向位置において取得される。いくつかの態様において、フローセル画像は、複数の所定の軸方向位置において取得される。複数の所定の軸方向位置は、 $2 \sim 500$ の所定の軸方向位置を含み得る。図27Bは、3つの例示的な所定の軸方向位置を示す。いくつかの態様において、複数の所定の軸方向位置は、 $2 \sim 30$ の所定の軸方向位置を含む。いくつかの態様において、複数の所定の軸方向位置は、 $2 \sim 20$ の所定の軸方向位置を含む。複数の所定の軸方向位置の各々は、その隣接する隣接軸方向位置から $0.1 \sim 500 \mu\text{m}$ 離隔していてもよ

10

20

30

40

50

い。所定の軸方向位置は、 $0.1 \sim 500 \mu\text{m}$ の範囲内にある同一の距離だけ離隔していてもよい。所定の軸方向位置は、 $0.1 \sim 500 \mu\text{m}$ の範囲内の異なる距離によって離隔していてもよい。複数の所定の軸方向位置の各々は、その隣接する隣接軸方向位置から1、2、3、4、5、6、7、8、9、10、11、12、13、14、又は $15 \mu\text{m}$ 離隔していてもよい。

【0205】

フローセル画像は、第1及び／又は第2の複数のフローセル画像を含み得る。いくつかの態様において、ベースコーリング位置は、同じ軸方向位置にある。言い換えれば、ベースコーリング位置は、2Dであり得る。いくつかの態様において、ベースコーリング位置のうちの少なくともいくつかは、異なる軸方向位置又はzレベルにある。言い換えれば、ベースコーリング位置は、3Dである。ベースコーリング位置のうちの異なる軸方向位置は、フローセル画像のための所定の軸方向位置のうちのいくつか又は全てにあり得る。いくつかの態様において、ベースコーリング位置の軸方向位置のうちの少なくともいくつかは、フローセル画像の複数の所定の軸方向位置のうちのいずれとも異なる。いくつかの態様において、ベースコーリング位置の異なる軸方向位置のうちの少なくともいくつかは、複数の所定の軸方向位置の2つの隣接する所定の軸方向位置の間にある。第1又は第2の複数のフローセル画像は、2、3、4、5、又は6つの異なる色チャンネルからのものであり得る。

【0206】

フローセル画像がベースコーリングテンプレートを生成するために取得される1つ以上のサイクルは、1、2、3、4、5、6、7、8、9、10、又はそれ以上のサイクルを含み得る。1つ以上のサイクルは、 $1 \sim 50$ サイクルを含む。1つ以上のサイクルは、 $1 \sim 6$ サイクル、 $2 \sim 5$ サイクル、又は $2 \sim 6$ サイクルを含む。

【0207】

第1又は第2の複数のフローセル画像の各々は、軸方向軸（すなわち、z軸）に直交する視野を含む。いくつかの態様において、第1又は第2の複数のフローセル画像の各々の視野は、画像平面内で同一である。第1又は第2の複数のフローセル画像の各々の視野は、フローセルのタイルの少なくとも一部分を覆い得る。いくつかの態様において、第1又は第2の複数のフローセル画像は、同一の画像解像度を含む。軸方向軸は、対物レンズから支持体まで延在してもよい。軸方向軸は、画像平面に直交し、視野は、画像平面内にある。

【0208】

支持体上の複数のコンカテマー分子の密度が、 1 mm^2 当たり $10^1 \sim 10^{12}$ である、先行請求項のいずれか一項に記載の方法。

【0209】

支持体上の支持体上の複数のコンカテマー分子の密度密度が、 1 mm^2 当たり $10^2 \sim 10^8$ である、先行請求項のいずれか一項に記載の方法。

【0210】

本明細書に開示されるように、ピクセル強度及びピクセル強度のそれぞれの色純度に基づいてベースコーリングテンプレートを判定する操作530は、ピクセル又はサブピクセルの各々について、それぞれの色純度が距離閾値内の他のピクセル又はサブピクセルの色純度よりも大きいかどうかを判定することを含み得る。それぞれの色純度が、距離閾値内の他のピクセル又はサブピクセルの色純度よりも大きいとの判定に応答して、ベースコーリングテンプレートにおいて対応するピクセル又はサブピクセルの位置を追加又は確認する。それぞれの色純度が、距離閾値内の他のピクセル又はサブピクセルの色純度よりも大きくないとの判定に応答して、ベースコーリングテンプレートに変更を加えず、次のピクセルに進んで、全てのピクセル又はピクセルの少なくともサブセットが同様の判定操作を通過するまで、判定操作を繰り返す。

【0211】

いくつかの態様において、操作530は、重複したポロニー、焦点外ポロニー、及び／

10

20

30

40

50

又は細胞構成要素からのバックグラウンドシグナルなどの合焦ポロニーの強度に干渉する他の光シグナルを除去するように構成される。

【0212】

距離閾値は、干渉シグナル（例えば、焦点外の及び／又は重複したポロニー）を排除する効果を最適化しながら、重複してもおらず、焦点外でもない、より弱い又はより大きなポロニーを維持するようにカスタマイズすることができる。距離閾値は、2つのポロニーの中心の間として判定することができる。いくつかの態様において、距離閾値は、2つのピクセル又はサブピクセルの中心の間で判定され得る。2つのポロニーが単一の2D平面にあるとき、距離閾値は2Dである。2つのピクセル又はサブピクセルが複数の2D平面にあるか、又は3つの3D空間内にある（例えば、インサイチューレーション）態様では、距離閾値は3Dである。例えば、図27Bに示されるように、ピクセル（ x_1 、 y_1 、 z_2 ）におけるポロニー p_2 は、3D距離閾値内にあるピクセル（ x_1 、 y_1 、 z_1 ）における第2のポロニー p_1 及び（ x_1 、 y_1 、 z_3 ）における第3のポロニー p_3 を有し得る。3D距離閾値は、ポロニー p_2 の周りの円筒を判定し得る。 z_1 、 z_2 、及び z_3 は、フローセル画像を取得するための所定の軸方向位置であり、約2 μm 離隔している。ポロニー p_1 は、3D距離閾値の外にある。距離閾値内の3つのポロニーについて、 p_1 は最も低い純度を有しており、その結果、 p_1 は、重複したポロニー又は焦点外ポロニーのいずれかとして除去され得る。ベースコーリング位置は、 z_3 及び z_2 の所定の軸方向位置の間と $z_3 - 1$ にあると判定される。軸方向位置 $z_3 - 1$ は、複数のポロニー、例えば、 p_2 及び p_3 のそれぞれの純度による線形補間又は重み付けによって判定され得る。 p_2 が p_1 よりも高い色純度を有する場合、軸方向位置 $z_3 - 1$ は p_2 に近接物であり得る。この態様では、各ピクセルは3Dであり、軸方向軸に厚さを含み得る。いくつかの態様において、2つ以上のポロニーが軸方向位置に沿った距離閾値内にある場合、それらのうちの1つのみが選択される。唯一のポロニーは、重み付け、補間、平均化、又は様々な統計的若しくは数学的関数に基づいて選択され得る。

10

20

【0213】

3D距離閾値は、1、2、3、又はそれ以上の距離要素を含んでよく、各距離要素は、 x 、 y 、又は z 方向の距離に対応してよい。例えば、3D距離閾値は、3D距離閾値が球面領域を判定し、球面内のポロニーが距離閾値内にあるように、 x 、 y 、及び z 方向に3つの同一の距離要素を含み得る。別の例として、 z 方向の距離要素は、 x 方向又は y 方向の距離要素とは異なり得るため、円筒、楕円、又は球体内のポロニーは、距離閾値内にある。距離閾値は、図27A～Bに示されるように、直交座標系において x 、 y 、又は z 方向の3つの距離要素に変換され得る様々な数の要素（例えば、異なる非直交座標系において）を含み得る。

30

【0214】

いくつかの態様において、距離閾値は、 x 、 y 、及び／又は z 方向における画像解像度に基づいてカスタマイズされ得る。いくつかの態様において、距離閾値は、 x 、 y 、及び／又は z 方向の画像解像度、並びにポロニー又はクラスタのサイズに基づいてカスタマイズすることができる。 z 方向の画像解像度は、2つの隣接する z レベルでのフローセル画像間の距離であり得る。例えば、2つの隣接する z レベルのフローセル画像は、互いに1 μm ～10 μm 離隔していてもよく、 z 解像度は、そのギャップとして判定されてもよい。

40

【0215】

いくつかの態様において、3D距離閾値は、軸方向軸（すなわち、 z 軸）に沿った第1の要素距離及び軸方向軸に直交する平面内の第2の要素距離を含む。いくつかの態様において、3D距離閾値は、軸方向軸（すなわち、 z 軸）に沿った第1の要素距離、並びに軸方向軸に直交する平面内の第2の要素距離及び第3の要素距離を含む。いくつかの態様において、第1の要素の距離は、第2の要素及び／又は第3の要素の距離と異なる。いくつかの態様において、第1の要素の距離は、第2の要素の距離及び／又は第3の要素の距離と同一である。

50

【0216】

ベースコーリングテンプレートを判定する操作は、少なくとも部分的に、340の操作を含むことができる。

【0217】

ベースコーリングテンプレートを判定した後、フローセル画像、例えば、1つ以上のサイクルに続く1つ以上のサイクルにおけるフローセルデバイスの第2の複数のフローセル画像を登録するように構成することができる。いくつかの態様において、第2の複数のフローセル画像は、ベースコーリングテンプレートを生成するためのサイクルに続く1つ以上のサイクルにおいて、インサイチュ試料のものである。第2の複数のフローセル画像は、複数の軸方向位置で取得され得る。

10

【0218】

いくつかの態様において、方法500は、第2の複数のフローセル画像を生成する操作を更に含み得る。第2の複数のフローセル画像は、ベースコーリングテンプレートの生成に対応する1つ以上のサイクルに続く1つ以上のサイクルで生成され得る。

【0219】

いくつかの態様において、方法500は、1つ以上の後続のシーケンシングサイクルからの第2の複数のフローセル画像を、ベースコーリングテンプレートに対して位置合わせ又は整列させる操作を更に備えることができる。いくつかの態様において、第2の複数のフローセル画像を位置合わせする操作は、共通座標系における第2の複数のフローセル画像におけるポロニーの座標を生成することを含むことができる。ベースコーリングテンプレートは、共通座標系にもある。いくつかの態様において、第2の複数のフローセル画像を位置合わせする操作は、フローセル画像のサブタイトルに対応する複数の変換（例えば、各個々の軸方向位置又は2D平面で）を利用して、フローセル画像全体の画像変換を推定することができる。フローセル画像を位置合わせする操作は、2Dにおけるフローセル画像の画像変換の推定を提供するサブタイトルの変換を生成することを含むことができる。隣接するサブタイトル内の情報は、サブタイトルの各個々の変換を判定するときに使用され得る。いくつかの態様において、軸方向軸に沿った変換が無視できる（例えば、0.05 μm 、0.02 μm 、0.001 μm 、0.0001 μm 、又はそれ以上）場合、2D変換は、フローセル画像をベースコーリングテンプレートに確実に位置合わせ又は整列させることができる。いくつかの態様において、軸方向軸に沿った変換を含める必要がある場合（例えば、0.05 μm 、0.02 μm 、0.001 μm 、0.0001 μm 、又はそれ以上）、x-y平面内の2D変換は、軸方向軸に沿ったシフト（例えば、z軸）と組み合わせて、第2のフローセル画像をベースコーリングテンプレートに対して位置合わせ又は整列させることができる。軸方向軸に沿った変換は、光学システムの物理的特性に基づいたシフトによって確実に表され得る。

20

30

【0220】

いくつかの態様において、第2の複数のフローセル画像を位置合わせする操作は、フローセル画像の画像変換を推定するために、ボルメトリック試料のサブボリュームに対応する（例えば、複数の個々の軸方向位置又は2D平面での）複数の3D変換を利用することができる。フローセル画像を位置合わせする操作は、フローセル画像の画像変換の推定を3Dで提供する変換を生成することを含むことができる。隣接するボリューム又はサブボリューム内の情報は、個々の変換を判定するときに使用できる。

40

【0221】

いくつかの態様において、ポロニーの座標は、一次元ベクトル又はリストに保存され得る。ベクトル又はリストの各エントリは、ポロニー、その座標、及びサイクルの1つ以上のチャンネルにおけるピクセル強度などの他の関連情報の一意の識別を含むことができる。

【0222】

いくつかの態様において、方法500は、位置合わせされた第2の複数のフローセル画像からのシグナルを使用して、ベースコーリングテンプレート内のベースコーリング位置において第2の複数のフローセル画像のベースコーリングを実行する操作を更に含むこと

50

ができる。いくつかの態様において、方法500は、第2の複数のフローセル画像のベースコーリングを、ベースコーリングテンプレート内のベースコーリング位置のうちのいくつか又は全てにおいてのみ実行する操作を更に含むことができる。言い換えれば、ベースコーリングテンプレートは、フローセル上の全ての正確なポロニー位置を識別し、重複したポロニー又は他の干渉シグナルからのベースコールを回避できるように、ポロニーである可能性が低い他のシグナル位置を無視するポロニーマップとして機能する。本明細書における一次分析におけるベースコールは、共通座標系における位置合わせされた画像強度に基づくことができる。本明細書に開示されるように、ベースコーリング位置は、 x 、 y 、及び／又は z 方向のサブピクセル位置にあり得る。

【0223】

第2の複数のフローセル画像のベースコーリングを実行する操作は、図3の操作350を少なくとも部分的に含むことができる。

【0224】

いくつかの態様において、フローセル画像の所定の軸方向位置のいずれにもないベースコーリングテンプレート内のベースコーリング位置は、 $x-y$ 平面内のサブピクセルベースコーリング位置と同様に取得され得る。いくつかの態様において、軸方向軸に沿った補間は、サブピクセル位置を判定するために使用され得る。軸方向軸に沿ったサブピクセル位置は、フローセル画像の2つの隣接する z レベルの間にある任意の軸方向位置であり得る。軸方向軸に沿った補間は、 $x-y$ 平面内の2D補間のために開示された任意の補間関数を使用し得る。いくつかの態様において、他の様々な補間機能が使用され得る。いくつかの態様において、軸方向軸に沿った線形補間は、顕微鏡の光学特性に基づいて使用され得る。いくつかの態様において、補間は使用されず、ベースコーリング位置は、フローセル画像の所定の軸方向位置にあり得る。いくつかの態様において、所定の軸方向位置は、補間が必要とされないように、細胞試料のサイズに対して調整され得る。

【0225】

本明細書のサブピクセル強度は、軸方向軸に沿った2、3、4、5、6、7、8、9、10、12、13、14、15、16、17、18、19、又は20の異なる軸方向位置を有し得る。本明細書のサブピクセル強度は、複数の所定の軸方向位置のうちの2つの隣接する軸方向位置間にある2、3、4、5、6、7、8、9、10、12、13、14、15、16、17、18、19、又は20の異なる軸方向位置を有し得る。本明細書におけるサブピクセル強度は、軸方向軸に沿った2～15の異なる軸方向位置を有し得る。本明細書におけるサブピクセル強度は、複数の所定の軸方向位置のうちの2つの隣接する軸方向位置間にある1～15の異なる軸方向位置を有し得る。サブピクセル強度の各々は、軸方向軸に沿っており、複数の所定の軸方向位置のうちの2つの隣接する軸方向位置間にある1～80の異なる軸方向位置を有し得る。いくつかの態様において、軸方向に沿ったサブピクセル解像度は、細胞試料のサイズ及び軸方向軸に沿ったコンカテマーの密度に基づき得る。

【0226】

支持体上のコンカテマー分子の密度は、 1 mm^2 当たり約 $10^1 \sim 10^{12}$ であり得る。支持体上のコンカテマー分子の密度は、 1 mm^2 当たり約 $10^2 \sim 10^8$ であり得る。支持体上のコンカテマー分子の密度は、 1 mm^2 当たり約 $10^4 \sim 10^8$ であり得る。支持体上のコンカテマー分子の密度は、 1 mm^2 当たり約 $10^2 \sim 10^5$ であり得る。支持体上のコンカテマー分子の密度は、 1 mm^2 当たり $10^1 \sim 10^{12}$ であり得る。支持体上のコンカテマー分子の密度は、 1 mm^2 当たり $10^2 \sim 10^8$ であり得る。支持体上のコンカテマー分子の密度は、 1 mm^2 当たり $10^4 \sim 10^8$ であり得る。支持体上のコンカテマー分子の密度は、 1 mm^2 当たり $10^2 \sim 10^5$ であり得る。

【0227】

いくつかの態様において、方法500は、操作520の代わりに操作520'を含み得る。操作520'は、ピクセル強度及びピクセル強度のそれぞれの色純度に基づいて、スポット探索アルゴリズムを使用することによって、第1の複数のフローセル画像内のベー

10

20

30

40

50

スコ어링位置を識別することを含み得る。操作520と520'との間の違いは、方法300に関連して記載される。いくつかの態様において、操作520は、フローセル画像のうちのいくつか又は全てのピクセルを候補ベーススコ어링位置として扱い、判定された各ベーススコ어링位置をベーススコ어링テンプレートに追加することによって、候補ベーススコ어링位置の中のベーススコ어링位置を判定し得る。いくつかの態様において、操作520'は、スポット探索アルゴリズムが潜在的な候補ベーススコ어링位置のセットを出力するように、フローセル画像の光シグナルをスポット探索アルゴリズムに入力することを含む。スポット探索アルゴリズムは、複数のzレベルからフローセル画像を取得し、潜在的な候補ベーススコ어링位置のセットを出力できるように、3Dで機能し得る。潜在的な候補ベーススコ어링位置のセットは、少なくとも2つの異なるz位置であり得る。

10

【0228】

シーケンシング結果及び細胞特徴のデータファイル

F a s t Qファイルなどのシーケンシング結果の従来のデータファイルは、一般に、ベースコール及びそれらに対応する品質スコアを含む。例えば、インサイチュ細胞又は組織のシーケンシングは、従来の2Dシーケンシングよりも、シーケンシングシステム及びシーケンシング分析が更に困難かつ複雑になっている。インサイチュ試料のシーケンシングは、シーケンシング結果の空間情報を判定するために従来の空間生物学に頼っている。現在、シーケンシングデータは、1つの機器、例えば、シーケンシングシステムで収集され、空間生物学データは、異なる機器、例えば、様々な細胞染色を有する従来の顕微鏡で収集され得る。したがって、空間情報及びシーケンシング結果は、異なるデータファイル形式で格納され、それによって、空間情報に関連してシーケンシング結果を分析するための複数のソフトウェア及び／又はプロセスを必要とする。更に、シーケンシング及び細胞特徴を含む画像は、複雑なデータ形式で保存され得、格納空間についての課題をもたらし得る。インサイチュシーケンシング分析を容易にするために、シーケンシングデータとシームレスに統合する空間生物学データが必要とされている。

20

【0229】

本開示の態様は、データファイルを提供する。本明細書のデータファイルは、シーケンシング結果を含むことができる。本明細書のデータファイルは、従来のデータファイル、例えば、従来のF a s t Qファイルと同様に、複数のベースコールを含むことができる。ベースコールは、単一の軸方向位置（すなわち、z位置）でのフローセル画像の一次分析から生成され得る。図28は、小さな明るい点（例えば、2つのピクセルの大きさ）としてのポロニーを有するインサイチュ細胞のシミュレートされたフローセル画像（右上）を示す。いくつかの態様において、ベースコールは、複数の軸方向位置又は3D空間におけるフローセル画像のシーケンシング分析から得られ得る。

30

【0230】

本開示の態様では、本明細書のデータファイル及び対応する形式は、有利には、空間生物学的情報コンテンツを含み、複雑な画像データセットを単一のデータ形式に圧縮することができる。本明細書に開示されるデータファイルは、有利には、3Dシーケンシングに特有の情報を含むことができるため、そのようなデータファイルは、3Dシーケンシング分析のために便利かつより効率的に利用することができる。本明細書に開示されるシステム及び方法の他の利点は、細胞又は組織形態などの3D細胞特徴をデータファイルから効率的に再構築することを可能にすること、データファイル内のバイオマーカー情報及びそのデータ処理を簡素化すること、単一のファイルタイプからより簡単でより便利なデータ学習を可能にして、様々なアッセイタイプからの情報をシームレスに統合することができるようにすること、並びに異なるアッセイタイプにわたって重要な情報コンテンツを保持しながらデータ圧縮を可能にすることを含むが、これらに限定されない。

40

【0231】

いくつかの態様において、ベースコールは、1つ以上のシーケンシングサイクルにおいてポロニーから生成される。いくつかの態様において、ベースコールは、1～50サイク

50

ル、2～30サイクル、2～20サイクル、又は2～60サイクルのみを含む短いシーケンシングリードにおいてポロニーから生成される。

【0232】

いくつかの態様において、ベースコールは、バーコード配列（複数可）の1つ以上のベースコールを含み得る。いくつかの態様において、ベースコールは、インデックス配列（複数可）の1つ以上のベースコールを含み得る。いくつかの態様において、ベースコールは、目的の挿入配列（複数可）の1つ以上のベースコールを含み得る。いくつかの態様において、ベースコールは、バーコード配列（複数可）からのベースコールのみを含み得る。いくつかの態様において、ベースコールは、インデックス配列（複数可）からのベースコールのみを含み得る。いくつかの態様において、ベースコールは、目的の挿入配列（複数可）の塩基を欠く。

10

【0233】

本明細書のデータファイルは、ベースコールのうちのいくつか又は全てに対応する品質インジケータを含み得る。品質インジケータは、ベースコールの品質スコアを含み得る。

【0234】

データファイルは、支持体上のベースコールの空間座標を含み得る。加えて、データファイルは、1つ以上の細胞特徴のデータを含み得る。細胞特徴は、支持体上に固定化された細胞試料のものであってもよい。ベースコールは、支持体上に固定化された細胞試料のシーケンシングから生成され得る。細胞特徴のデータは、データファイル内のベースコールを生成するようにシーケンシングされた同じインサイチュ試料から取得され得る。1つ以上の細胞特徴の各々のデータは、細胞特徴インジケータを含み得る。細胞特徴インジケータは、核、膜、細胞質、及びミトコンドリアから選択される細胞要素を示すように構成され得る。細胞特徴インジケータは、核境界及び膜境界から選択される細胞要素を示すように構成され得る。例えば、「m」は、ミトコンドリア又は膜を示し得、「c」は、サイトゾルを示し得、「n」は、核を示し得る。データファイルは、支持体上の細胞特徴の空間座標を含み得る。ベースコールの空間座標及び細胞特徴の空間座標は、同じ座標系内にあり得る。細胞特徴の各々の空間座標は、空間座標の複数のセットを含み得る。1つ以上の細胞特徴のデータは、複数のエントリに分けられていてもよく、複数のエントリの各エントリは、支持体上の空間位置に対応する。空間位置は、2D又は3Dであり得る。1つ以上の細胞特徴のデータは、複数のエントリに分けられていてもよく、複数のエントリの各エントリは、空間座標のセットに対応し得る。空間座標のセットは、x座標、y座標、z座標、又はそれらの組み合わせを含み得る。空間座標の各セットは、2D又は3Dで一意の位置を示すように構成される。例えば、細胞特徴は、1つ以上の細胞の細胞膜（複数可）であってもよい。細胞膜上にある画像内の各ピクセルは、エントリであることができ、全ての複数のエントリが一緒になって、細胞膜（複数可）の3D空間位置を判定する。2D及び／又は3D直交座標系が、本明細書における2D又は3D空間における空間位置を定義するために使用されるが、本明細書における空間位置及び関係を定義するために様々な座標系を使用することができる。直交系以外のいくつかの例示的な座標系は、極座標系、円筒座標系、又は球面座標系を含むことができるが、これらに限定されない。他の座標系は、均質な座標系又は非均質な座標系を含むことができる。

20

30

40

【0235】

ベースコール又は細胞特徴の空間座標は、3Dであり得る。ベースコールの空間座標及び1つ以上の細胞特徴の空間座標は、x座標、y座標、z座標、又はそれらの組み合わせを含み得る。

【0236】

いくつかの態様において、ベースコール、品質インジケータ、1つ以上の細胞特徴、ベースコールの空間座標、又は細胞特徴の空間座標は、1つ以上の文字、1つ以上の数字、1つ以上の記号、又はそれらの組み合わせを含む。

【0237】

データファイルは、テキストファイルであり得る。データファイルは、FastQファ

50

イルであり得る。データファイルは、テキストファイル用の様々な符号化形式を使用して符号化され得る。データファイルは、ASCII符号化され得る。データファイルは、8ビット、12ビット、16ビット、18ビット、32ビット、又は36ビットの符号化形式であり得る。

【0238】

いくつかの態様において、データファイルは、細胞特徴に対応する1つ以上のシーケンシングパラメータ及びその値を含む。1つ以上のシーケンシングパラメータは、標的配列を含み得る。1つ以上のシーケンシングパラメータは、シーケンシングシステムの名称を含み得る。1つ以上のシーケンシングパラメータは、シーケンシングランの識別を含み得る。1つ以上のシーケンシングパラメータは、フローセルの識別を含み得る。1つ以上のシーケンシングパラメータは、フローセルのレーン番号又はレーン識別を含み得る。1つ以上のシーケンシングパラメータは、タイル番号又はサブタイル番号を含み得る。

10

【0239】

いくつかの態様において、データファイルは、1つ以上の区切り符号を含む。1つ以上の区切り符号の各々は、1つ以上の細胞特徴を区切る、及び／又はベースコールを1つ以上の細胞特徴から区切るように構成され得る。1つ以上の区切り符号の各々は、文字、数字、記号、又はそれらの組み合わせを含み得る。いくつかの態様において、1つ以上の細胞特徴のデータは、細胞特徴の空間座標に基づいて、1つ以上の区切り符号によって区切られる。いくつかの態様において、1つ以上の細胞特徴の複数のエントリの各エントリは、1つ以上の区切り符号によって区切られる。

20

【0240】

いくつかの態様において、本明細書のデータファイルは、次世代シーケンシング（NGS）システムによって生成され、ベースコールを含むデータファイル、例えば、従来のFastQファイル、と同一又は互換性のあるデータ形式を含む。いくつかの態様において、本明細書のデータファイルの少なくとも一部は、従来のシーケンシングデータファイル、例えば、FastQファイル、を処理するように構成されたソフトウェア又はコンピュータプログラムによって分析又は処理され得る。例えば、少なくともベースコール、及び任意選択でそれらの対応する品質スコアを含む本明細書のデータファイルのシーケンシング結果部分は、従来のシーケンシングデータファイル、例えば、FastQファイル、を処理するように構成されたソフトウェア又はコンピュータプログラムによって分析又は処理することができる。

30

【0241】

いくつかの態様において、個々の細胞試料、例えば、フローセル上に固定化された複数のインサイチュセルを有する試料について、試料のシーケンシング結果は、第1のデータファイル、例えば、従来のFastQファイル、に含まれ得、試料の対応する細胞特徴のデータは、同じ第1のデータファイル、又は第2の異なるデータファイルに含まれ得る。第2のデータファイルは、依然として、第1のデータファイルと互換性のある、又は同一のデータ形式を有し得、ユーザが、データファイル（複数可）を再形式すること、データファイル（複数可）を変換すること、又はそうでなければデータファイルを処理することを必要とせずに、第1のデータファイル及び第2のデータファイル内の情報を組み合わせることを容易にする。例えば、細胞特徴の従来のデータは、データ処理が細胞特徴の空間座標を抽出するために保証されるように、JPEG、PNG、TIFF、又は他のデジタル画像符号化形式であり得る。更に、抽出された空間座標は、従来のFastQと互換性のある8ビット又は他の符号化に変換される必要がある場合がある。本明細書に開示されるデータファイル及びその形式は、空間生物学的情報及びインサイチュシーケンシング結果を有利に含むことができる。本明細書に開示されるデータファイルは、そのようなデータファイルが、インサイチュシーケンシング情報と対応する細胞特徴とを組み合わせるインサイチュシーケンシング分析のために便利かつより効率的に利用され得るように、インサイチュシーケンシングに特有の情報を有利に含むことができる。本明細書に開示されるシステム及び方法の他の利点は、インサイチュ細胞又は組織形態をデータファイルから効

40

50

率的に再構築することを可能にすること、データファイル内のバイオマーカー情報及びそのデータ処理を簡素化すること、単一のファイルタイプからより簡単でより便利なデータ学習を可能にして、様々なアッセイタイプからの情報をシームレスに統合することができるようにすること、並びに異なるアッセイタイプにわたって重要な情報コンテンツを保持しながら大幅なデータ圧縮を可能にすることを含むが、これらに限定されない。

【0242】

いくつかの態様において、データファイルは、1つ以上の細胞特徴及びベースコーリング結果を含む画像の再構成のために構成され、1つ以上の細胞特徴及びベースコーリング結果は、互いにそれぞれに整列又は位置合わせされる。図28は、細胞膜及び核膜のセグメンテーションとシーケンシング結果を整列させる例示的な再構築画像（右下）を示す。したがって、データファイルは、有利には、従来のシーケンシング結果、例えば、ベースコールのユーザが、細胞特徴に基づいてシーケンシング分析を実行することを可能にすることができる。

【0243】

図29Aは、例示的な細胞及びその構造要素のうちのいくつかを示す。核、膜、細胞質、及びミトコンドリアを表すために、本明細書に開示されるデータファイルにおいて、固有の細胞特徴インジケータを使用することができる。本明細書に開示されるデータファイルはまた、シーケンシングシステムに関連する1つ以上の第1のパラメータ及びそれらの値と、イメージングパラメータに関連する1つ以上の第2のパラメータ及びそれらの値と、を含むことができる。そのような第1及び第2のパラメータ及びそれらの値を使用して、シーケンシングハードウェア、ソフトウェア、又は他のシーケンシング関連プロセスを一意に示すことができる。図29Bに示されるように、文字「n」及びそのASCIIコードを使用して、文字「n」の後に来るベースコーリング結果が細胞の核から得られたことを示すことができる。文字「m」は、この文字の後に含まれる情報が細胞形態に属することを示すことができる。そのような情報は、図29Bに示されるように、次の別個の「@」が現れるときに終了し得る。

【0244】

各細胞特徴インジケータについて、データファイルは、サブ特徴インジケータを更に含み、同じ細胞特徴インジケータ内に別個の特徴を更に含み得る。例えば、図29Bに示されるように、細胞質についての細胞特徴指標は、「c」である。タンパク質などの細胞質内の異なる要素について、データファイルは、サブ特徴インジケータを含み得る。サブ特徴インジケータは、異なる文字又は文字と数字の様々な一意の組み合わせであり得る。例えば、図29Bのデータファイルは、異なるヌクレオチドの組み合わせであるタンパク質マーカーを使用する。ヌクレオチドの組み合わせは、可変長又は固定長、例えば、5であり得る。同様に、図29Bのデータファイルは、異なるヌクレオチドの組み合わせである転写マーカーを使用する。

【0245】

データファイルが使用することができるインジケータの例示的なリストが、図29Aに列挙される。いくつかの態様において、各細胞特徴インジケータ、パラメータ、本明細書に開示されるパラメータの値は、1つ以上の文字、1つ以上の数字、1つ以上の記号、又はそれらの組み合わせを含むことができる。

【0246】

データファイルはまた、1つ以上の区切り符号を含み得る。各区切り符号を使用して、データファイル内の2つの異なるエントリを区切ることができる。例えば、従来の区切り符号「+」を使用して、図29Bに示されるように、ベースコーリング結果と品質スコアとを区切ることができる。別の例として、区切り符号は、図29Bに示されるように、「@」「:」及び1つ以上のスペースを含むことができる。

【0247】

データファイルは、図29A～29Bに示されるインデックス配列を更に含むことができる。

【0248】

本明細書には、データファイルを生成する方法が開示される。データファイルを生成する方法は、本明細書に開示される操作のうちのいくつか又は全てを含むことができる。

【0249】

データファイルを生成する方法は、本明細書に開示される操作のうちのいくつか又は全てを含むことができる。操作は、本明細書に記載される順序で実行され得るが、これに限定されない。

【0250】

データファイルを生成する方法は、本明細書に開示される1つ以上のプロセッサ（例えば、図4の404）によって実行することができる。いくつかの態様において、プロセッサは、プロセッシングユニット、集積回路、又はそれらの組み合わせのうちの1つ以上を含むことができる。例えば、プロセッシングユニットは、セントラルプロセッシングユニット（CPU）及び／又はグラフィックスプロセッシングユニット（GPU）を含むことができる。集積回路は、フィールドプログラマブルゲートアレイ（FPGA）などのチップを含むことができる。いくつかの態様において、プロセッサは、コンピューティングシステム400を含み得る。

10

【0251】

いくつかの態様において、データファイルを生成する方法におけるいくつかの、又は全ての操作は、FPGA（複数可）によって実行され得る。いくつかの操作がFPGA（複数可）によって実行される態様では、FPGA（複数可）によって実行される操作後のデータは、FPGA（複数可）によってCPU（複数可）に通信されることが可能であり、その結果、CPU（複数可）は、そのようなデータを使用してデータファイルを生成する方法において後続の操作（複数可）を実行することが可能である。同様に、データはまた、FPGA（複数可）による処理のために、CPU（複数可）からFPGA（複数可）に伝達され得る。いくつかの態様において、データファイルを生成する方法における全ての操作は、CPU（複数可）によって実行され得る。あるいは、CPU（複数可）によって実行される操作は、専用プロセッサ、又はGPU（複数可）などの他のプロセッサによって実行され得る。いくつかの態様において、データファイルを生成する方法における全ての操作は、FPGA（複数可）によって実行され得る。いくつかの態様において、データファイルを生成する方法のうちのいくつか又は全ての操作は、シーケンシングランにおけるシーケンシングサイクルMの間又はその前に実行される。いくつかの態様において、Mは、5、10、20、30、40、50、又は60未満である。

20

30

【0252】

操作は、本明細書に記載される順序で実行され得るが、これに限定されない。本明細書のデータファイルを生成する方法は、光学システムによって、中に1つ以上の細胞特徴を含む1つ以上の画像を生成することを含むことができる。1つ以上の画像は、1つ以上の細胞特徴からの光シグナルを含み得る。1つ以上の画像は、1つ以上の細胞特徴のコントラストの増加を含み得る。例えば、1つ以上の画像は、支持体上の細胞試料の染色の顕微鏡画像であり得る。いくつかの態様において、1つ以上の画像は、シーケンシングリードが完了した後に取得される。いくつかの態様において、1つ以上の画像は、図1に示されるように、シーケンシングシステムを使用してベースコールを生成するためのフローセル画像が取得された後に取得される。

40

【0253】

図28は、細胞及び核のシミュレートされた染色を伴う細胞試料のシミュレートされた画像（上部中央）を示す。

【0254】

データファイルを生成するための方法は、1つ以上の画像内の1つ以上の細胞特徴の各々の空間座標を判定する操作を含み得る。判定操作は、セグメンテーション、輪郭トレーシングなどのための様々な画像処理アルゴリズムを使用し得る。図28は、核及び細胞膜（上部中央）などの細胞特徴を示す画像に由来する例示的なセグメンテーション（上部左

50

）を示す。空間座標は、セグメンテーション画像を使用して判定することができる。

【0255】

本明細書におけるデータファイルを生成するための方法は、判定された空間座標に基づいて1つ以上の細胞特徴のデータを生成する操作を含み得る。1つ以上の細胞特徴の各々について、判定された空間座標に基づいてデータを生成する操作は、細胞特徴インジケータを生成することと、細胞特徴の空間座標を挿入することと、を含む。そのような生成及び挿入操作は、細胞特徴の複数のエントリに対してであり得、各エントリは、2D又は3Dでの非反復的な空間位置に対応する。そのような生成及び挿入操作は、特定の細胞特徴の複数のエントリのうちのいくつか又は全てを含むように反復的であり得る。図28は、境界のための2つの例示的なエントリを示し、一方は、それぞれ、ライン1及びライン134549において、膜からのものであり、他方は、核境界からのものである。エントリを含むための例示的な形式は、図28にも示される：

境界形式：@Instrument:RunID:FlowcellID:Lane:Title:X:Y:__:Paired:Filter:ControlBits:Feature(nor m):__

【0256】

データファイルは、次世代シーケンシング(NGS)システムによって生成され、ベースコールを含むデータファイルと同一又は互換性のあるデータ形式を含み得る。

【0257】

本明細書におけるデータファイルを生成するための方法は、ベースコールを含む次世代シーケンシング(NGS)システムによって生成されるデータ形式と同一又は互換性のあるデータ形式で、判定された空間座標に基づいて1つ以上の細胞特徴のデータを生成する操作を含み得る。いくつかの態様において、1つ以上の細胞特徴のデータを生成するそのような操作は、1つ以上の細胞特徴の各々について、数回の操作の反復を実行することを含み、(i)細胞特徴の細胞特徴インジケータを挿入すること、(ii)細胞特徴の空間座標の空間座標のセットを挿入すること、(iii)細胞特徴に対応する1つ以上のシーケンシングパラメータ及びその値を挿入すること、並びに(iv)1つ以上の区切り符号のうちの少なくとも1つ、又はそれらの組み合わせを挿入することを含む。いくつかの態様において、1つ以上の細胞特徴のデータを生成するそのような操作は、1つ以上の細胞特徴の各々について、数回の操作の反復を実行することを含み、(v)ベースコールを挿入すること、(vi)ベースコールに対応する品質インジケータを挿入すること、及び(vii)ベースコールの空間座標、又はそれらの組み合わせを挿入することを含む。図28は、ライン249307~249309におけるバーコード配列の少なくとも一部からのベースコールの例示的なエントリを示す。「PACRG」は、バーコードの標的タンパク質/遺伝子である。

例示的なバーコード形式は、@Instrument:RunID:FlowcellID:Lane:Title:X:Y:Z:Paired:Filter:ControlBits:Feature(nor c):AssignedTargetであり得る。

【0258】

品質スコアは、後続の行の区切り符号「+」の後に挿入され得る。

【0259】

いくつかの態様において、本明細書におけるデータファイルを生成するための方法は、シーケンシングシステムによって、操作510と同様に支持体上に固定化された細胞試料の第1の複数のフローセル画像を生成する操作を含み得る。

【0260】

いくつかの態様において、本明細書でデータファイルを生成するための方法は、本明細書で開示されるような細胞試料の第1の複数のフローセル画像に基づいてベースコールを生成する操作を含み得る。

【0261】

いくつかの態様において、バーコード又は他のベースコールを挿入するための正確な形

10

20

30

40

50

式、及び境界若しくは細胞特徴を挿入するための正確な形式には、本明細書に開示される例示的な形式が含まれるが、これらに限定されない。それらの形式は、例えば、シーケンシングされる細胞試料の種類、空間的位置を定義するために使用される座標系、フローセルの特性などを含む様々な要因に基づいてカスタマイズすることができる。

【0262】

コンピュータシステム

様々な態様は、例えば、図4に示されるコンピュータシステム400などの1つ以上のコンピュータシステムを使用して実施され得る。1つ以上のコンピュータシステム400は、例えば、本明細書で考察される態様のうちのいずれか、並びにその組み合わせ及びサブ組み合わせを実施するために使用され得る。

10

【0263】

コンピュータシステム400は、プロセッサ404などの1つ以上のプロセッサ（セントラルプロセッシングユニット、又はCPUとも呼ばれる）を含み得る。プロセッサ404は、バス又は通信インフラストラクチャ406に接続され得る。

【0264】

コンピュータシステム400はまた、ユーザ入力／出力インターフェース（複数可）402を介して通信インフラストラクチャ406と通信し得る、モニタ、キーボード、ポインティングデバイスなどのユーザ入力／出力デバイス（複数可）403を含み得る。ユーザ入力／出力デバイス403は、図1のユーザインターフェース124に結合され得る。

【0265】

プロセッサ404のうちの1つ以上は、グラフィックスプロセッシングユニット（GPU）であり得る。一態様では、GPUは、数学的に集中的なアプリケーションを処理するように設計された特殊な電子回路であるプロセッサであり得る。グラフィックスプロセッシングユニット（GP GPU）上の汎用コンピューティングの能力により、GPUは、少なくとも本明細書に記載される画像認識及び機械学習の態様において特に有用であり得る。

20

【0266】

加えて、プロセッサ404のうちの1つ以上は、ハードウェア加速された暗号化コプロセッサを含む、暗号化計算を加速するための論理のコプロセッサ又は他の実装、又は他の特殊な数学的関数を含み得る。そのような加速されたプロセッサは、そのような加速を促進するために、コプロセッサ及び／又は他の論理を使用して加速するための命令セット（複数可）を更に含み得る。

30

【0267】

コンピュータシステム400はまた、ランダムアクセスメモリ（RAM）などのメイン又はプライマリメモリ408を含み得る。メインメモリ408は、1つ以上のレベルのキャッシュを含み得る。メインメモリ408は、その中に制御論理（すなわち、コンピュータソフトウェア）及び／又はデータを格納していてもよい。

【0268】

コンピュータシステム400はまた、1つ以上の二次ストレージ又は二次メモリ410を含み得る。二次メモリ410は、例えば、メインストレージドライブ412及び／又はリムーバブルストレージデバイス若しくはドライブ414を含み得る。メインストレージドライブ412は、例えば、ハードディスクドライブ又はソリッドステートドライブであり得る。リムーバブルストレージドライブ414は、フロッピーディスクドライブ、磁気テープドライブ、コンパクトディスクドライブ、光学ストレージデバイス、テープバックアップデバイス、及び／又は任意の他のストレージデバイス／ドライブであり得る。

40

【0269】

リムーバブルストレージドライブ414は、リムーバブルストレージユニット418と相互作用し得る。

【0270】

リムーバブルストレージユニット418は、コンピュータソフトウェア（制御論理）及

50

び／又はデータを格納したコンピュータ使用可能又は可読ストレージデバイスを含み得る。リムーバブルストレージユニット 418 は、フロッピーディスク、磁気テープ、コンパクトディスク、DVD、光学ストレージディスク、及び／又は任意の他のコンピュータデータストレージデバイスであってもよい。リムーバブルストレージドライブ 414 は、リムーバブルストレージユニット 418 から読み出し、及び／又はリムーバブルストレージユニット 418 に書き込み得る。

【0271】

二次メモリ 410 は、コンピュータプログラム及び／又は他の命令及び／又はデータにコンピュータシステム 400 がアクセスすることを可能にするための他の手段、デバイス、構成要素、手段又は他のアプローチを含み得る。そのような手段、デバイス、構成要素、機器、又は他のアプローチは、例えば、リムーバブルストレージユニット 422 及びインターフェース 420 を含み得る。リムーバブルストレージユニット 422 及びインターフェース 420 の例には、プログラムカートリッジ及びカートリッジインターフェース（ビデオゲームデバイスに見られるものなど）、リムーバブルメモリチップ（EPROM 又は PROM など）及び関連するソケット、メモリスティック及び USB ポート、メモリカード及び関連するメモリカードスロット、並びに／又は任意の他のリムーバブルストレージユニット及び関連するインターフェースが含まれ得る。

【0272】

コンピュータシステム 400 は、通信又はネットワークインターフェース 424 を更に含み得る。通信インターフェース 424 は、コンピュータシステム 400 が、外部デバイス、外部ネットワーク、外部エンティティなどの任意の組み合わせ（参照番号 428 によって個別に及び集合的に参照される）と通信し、対話することを可能にし得る。例えば、通信インターフェース 424 は、コンピュータシステム 400 が、有線及び／又は無線（又はそれらの組み合わせ）であり得、LAN、WAN、インターネットなどの任意の組み合わせを含み得る、通信パス 426 を介して外部又はリモートデバイス 428 と通信することを可能にし得る。制御論理及び／又はデータは、通信経路 426 を介してコンピュータシステム 400 との間で伝送され得る。いくつかの態様において、通信経路 426 は、図 1 に示されるように、クラウド 130 への接続である。参照番号 428 によって参照される外部デバイスなどは、クラウド 130 内のデバイス、ネットワーク、エンティティなどであり得る。

【0273】

コンピュータシステム 400 はまた、いくつかの非限定的な例、又はそれらの任意の組み合わせを挙げると、パーソナルデジタルアシスタント（PDA）、デスクトップワークステーション、ラップトップ若しくはノート型コンピュータ、ネットブック、タブレット、スマートフォン、スマートウォッチ、又は他のウェアラブル、アプライアンス、Internet of Things（IoT）の一部、及び／又は埋め込みシステムのうちのいずれかであってもよい。

【0274】

本明細書に記載されるフレームワークは、非一時的コンピュータ可読媒体若しくはデバイスなどの方法、プロセス、装置、システム、又は製品として実装され得ることを理解されたい。例示の目的のために、本フレームワークは、分散型台帳が公的に利用可能であるか、又は少なくとも信頼できない第三者がそれを利用可能であるという文脈で説明される。最新のユースケースとして、ブロックチェーンベースのシステムが挙げられる。しかしながら、本フレームワークは、機密情報又は機密情報が信頼できない第三者の手を通過する必要がある他の設定にも適用され得、この技術は、分散型台帳又はブロックチェーンの使用に決して限定されないことを理解されたい。

【0275】

コンピュータシステム 400 は、リモート若しくは分散クラウドコンピューティングソリューション、ローカル若しくはオンプレミスのソフトウェア（例えば、「オンプレミス」クラウドベースのソリューション）、「サービスとしての」モデル（例えば、サービス

10

20

30

40

50

としてのコンテンツ（C a a S）、サービスとしてのデジタルコンテンツ（D C a a S）、サービスとしてのソフトウェア（S a a S）、サービスとしてのマネージドソフトウェア（M S a a S）、サービスとしてのプラットフォーム（P a a S）、サービスとしてのデスクトップ（D a a S）、サービスとしてのフレームワーク（F a a S）、サービスとしてのバックエンド（B a a S）、サービスとしてのモバイルバックエンド（M B a a S）、サービスとしてのインフラストラクチャ（I a a S）、サービスとしてのデータベース（D B a a S）など）、及び／又は上記の例若しくは他のサービス若しくは配信パラダイムの任意の組み合わせを含むハイブリッドモデルを含むが限定されない任意の配信パラダイムを通じて、任意のアプリケーション及び／又はデータにアクセスするか、又はホスティングするクライアント又はサーバであり得る。

10

【0276】

任意の適用可能なデータ構造、ファイル形式、及びスキーマは、J a v a S c r i p t オブジェクト表記法（J S O N）、拡張マークアップ言語（X M L）、更にもう1つのマークアップ言語（Y A M L）、拡張ハイパーテキストマークアップ言語（X H T M L）、ワイヤレスマークアップ言語（W M L）、M e s s a g e P a c k、X M L ユーザインターフェース言語（X U L）、又は単独で又は組み合わせての他の機能的に類似した表現を含むが、これらに限定されない標準から導出され得る。あるいは、独自のデータ構造、形式、又はスキーマを、排他的に、又は既知の標準又はオープン標準と組み合わせて使用することができる。

20

【0277】

任意の関連するデータ、ファイル、及び／又はデータベースは、他の可能な形式の中でも、様々なタイプのマークアップ言語を更に含む、数値、テキスト、グラフィック、又はマルチメディア形式のような人間が読み取り可能な形式で格納され、取得され、アクセスされ、及び／又は送信されてもよい。代替的に、又は上記の形式との組み合わせで、データ、ファイル、及び／又はデータベースは、バイナリ、符号化、圧縮、及び／若しくは暗号化された形式、又は任意の他の機械可読形式で格納、取得、アクセス、及び／又は送信されてもよい。

【0278】

様々なシステム及び層間のインターフェース又は相互接続は、任意の数のプロトコル、プログラマティックフレームワーク、フロアプラン、又はアプリケーションプログラミングインターフェース（A P I）などの任意の数のメカニズムを用い得、これらには、ドキュメントオブジェクトモデル（D O M）、ディスカバリーサービス（D S）、N S U s e r D e f a u l t s、ウェブサービス記述言語（W S D L）、メッセージ交換パターン（M E P）、ウェブ分散型データ交換（W D D X）、ウェブハイパーテキストアプリケーション技術ワーキンググループ（W H A T W G）、H T M L 5 ウェブメッセージング、表現状態転送（R E S T 又は R E S T f u l ウェブサービス）、拡張可能ユーザインターフェースプロトコル（X U P）、シンプルオブジェクトアクセスプロトコル（S O A P）、X M L スキーマ定義（X S D）、X M L リモートプロシージャコール（X M L - R P C）、又は同様の機能及び結果を達成し得る任意の他のメカニズムが含まれるが、これらに限定されない。

30

40

【0279】

そのようなインターフェース又は相互接続はまた、ユニフォームリソース識別子（U R I）を利用し得、これは、ユニフォームリソースロケータ（U R L）又はユニフォームリソース名（U R N）を更に含み得る。他の形態の一樣及び／又は一意の識別子、ロケータ、又は名前が、排他的に、又は上に記載されたものなどの形態と組み合わせて使用される。

【0280】

上記のプロトコル又は A P I のうちのいずれかは、任意のプログラミング言語、手続き型言語、機能型言語、又はオブジェクト指向言語とインターフェースをとり得るか、又はそれらで実装され得、コンパイル又は解釈され得る。非限定的な例としては、C、C++

50

、C#、Objective-C、Java、Scala、Clojure、Elixir、Swift、Go、Perl、PHP、Python、Ruby、JavaScript、WebAssembly、又は実質的にあらゆる他の言語（あらゆる他のライブラリ若しくはスキーマで、あらゆる種類のフレームワーク、ランタイム環境、仮想マシン、インタープリター、スタック、エンジン、若しくは同様のメカニズム（他の多くの非限定的な例の中で、Node.js、V8、Knockout、jQuery、Dojo、Dijit、OpenUI5、AngularJS、Express.js、Backbone.js、Ember.js、DHTMLX、Vue、React、Electronなどを含むが、これに限定されない）で）が挙げられる。

【0281】

10

いくつかの態様において、制御論理（ソフトウェア）が格納されたタンジブルな非一時的コンピュータ使用可能又は可読媒体を含むタンジブルな非一時的装置又は製品は、本明細書ではコンピュータプログラム製品又はプログラムストレージデバイスと呼ばれる場合がある。これは、コンピュータシステム400、メインメモリ408、二次メモリ410、及びリムーバブルストレージユニット418及び422、並びに前述の任意の組み合わせを具現化するタンジブルな製品を含むが、これらに限定されない。そのような制御論理は、1つ以上のデータ処理デバイス（コンピュータシステム400など）によって実行されると、そのようなデータ処理デバイスに、本明細書に記載されるように動作させ得る。

【0282】

20

本開示に含まれる教示に基づいて、図4に示されるもの以外のデータ処理デバイス、コンピュータシステム、及び／又はコンピュータアーキテクチャを使用して、本開示の態様を作成及び使用する方法は、関連技術分野（複数可）の当業者には明らかであろう。特に、態様は、本明細書に記載されるもの以外のソフトウェア、ハードウェア、及び／又はオペレーティングシステムの実装で動作し得る。

【0283】

他のセクションではなく、詳細な説明セクションが、特許請求の範囲を解釈するために使用されることが意図されることを理解されたい。他のセクションは、発明者（複数可）によって企図されるように、1つ以上の例示的な態様を記載し得るが、全てではなく、したがって、本開示又は添付の特許請求の範囲をいかなる方法でも限定することを意図するものではない。

30

【0284】

本開示は、例示的な分野及び用途のための例示的な態様を説明するが、本開示はそれに限定されるものではないことを理解されたい。他の態様及びその修正は可能であり、本開示の範囲及び精神内である。例えば、このパラグラフの一般性を限定することなく、態様は、図面に示される、及び／又は本明細書に記載されるソフトウェア、ハードウェア、ファームウェア、及び／又はエンティティに限定されない。更に、態様は（本明細書に明示的に記載されているかどうかにかかわらず）、本明細書に記載されている実施例を超える分野及び用途に対して重要な有用性を有する。

【0285】

40

態様は、特定の機能及びそれらの関係の実装を例示する機能的ビルディングブロックを用いて上で説明されている。これらの機能的ビルディングブロックの境界は、本明細書において、説明の便宜のために任意に定義されている。指定された機能及び関係（又はその等価物）が適切に実行される限り、代替の境界は定義され得る。また、代替的な態様は、本明細書に記載されるものとは異なる順序を使用して、機能的ブロック、ステップ、操作、方法などを実行し得る。

【0286】

「一態様」、「態様」、「例示的な態様」、「いくつかの態様」、又は同様のフレーズへの本明細書の言及は、説明される態様が特定の特徴、構造、又は特徴を含み得るが、全ての態様が必ずしも特定の特徴、構造、又は特性を含み得るわけではないことを示す。更に、そのような句は、必ずしも同じ態様を指すものではない。更に、特定の特徴、構造、

50

又は特性が態様に関連して記載されるとき、そのような特徴、構造、又は特性を、本明細書に明示的に言及又は記載されるかどうかにかかわらず、他の態様に組み込むことは、関連技術分野（複数可）の当業者の知識の範囲内であろう。

【0287】

加えて、いくつかの態様は、「カップリングされた」及び「接続された」という表現をそれらの派生語とともに使用して説明され得る。これらの用語は、必ずしも互いの同義語として意図されているわけではない。例えば、いくつかの態様は、2つ以上の要素が互いに直接物理的又は電氣的に接触していることを示すために、「接続された」及び／又は「カップリングされた」という用語を使用して説明され得る。しかしながら、「カップリングされた」という用語はまた、2つ以上の要素が互いに直接接触しないが、それでも互いに協働又は相互作用することを意味し得る。

10

【0288】

いくつかの態様において、検出可能に標識された多価分子、多価結合複合体が形成されるステップ（2）、及び結合された、検出可能に標識された多価分子がイメージング及び検出されるステップ（3）を用いてシーケンシングする場合、条件は、検出可能に標識された鎖終結ヌクレオチドを使用するシーケンシングワークフローと比較して穏やかである。例えば、ステップ（2）及び（3）は、約35～45℃、又は約39～42℃の穏やかな温度で行うことができる。ステップ（2）及び（3）は、細胞試料内のDNAナノボールのスポット画像のFWHM（全幅半値）を改善することができる複数のシーケンシングサイクル（例えば、最大30サイクル）中に、DNAナノボールのコンパクトなサイズ及び形状の保持を助けることができる穏やかな温度で行うことができる。いくつかの態様において、DNAナノボールは、複数のシーケンシングサイクルの間には解明されない。いくつかの態様において、DNAナノボールのスポット画像は、複数のシーケンシングサイクルの間には拡大しない。いくつかの態様において、DNAナノボールのスポット画像は、複数のシーケンシングサイクルの間、離散スポットのままである。スポット画像は、ガウススポットとして表され得、サイズは、FWHMとして測定され得る。より小さいFWHMによって示されるより小さいスポットサイズは、典型的には、スポットの改善された画像と相関する。いくつかの態様において、ナノボールスポットのFWHMは、約10µm以下であることができる。

20

【0289】

いくつかの態様において、非同期のフェージング及び／又は事前フェージング事象は、検出可能に標識された多価分子を用いる同期ポリメラーゼ触媒シーケンシング反応の間に生じ得る。シーケンシング中に、多価分子の相補的ヌクレオチド単位の複合型ポリメラーゼへの結合に対応する蛍光シグナルを検出し、それによって多価結合複合体を形成することができる。したがって、フェージング及び事前フェージング事象は、標識された多価分子の結合を使用して検出及び監視することができる。いくつかの態様において、検出可能に標識された多価分子を用いて最大30回のシーケンシングサイクルを行う場合、フェージング及び／又は事前フェージング率は、約5%未満、又は約1%未満、又は約0.01%未満、又は約0.001%未満であり得る。対照的に、標識されたチェーンターミネーターヌクレオチドを使用して最大30回のシーケンシングサイクルを行うためのフェージング及び／又は事前フェージング率は、約5%であり得る。

30

40

【0290】

本明細書に記載される方法のうちのいずれにおいても、細胞試料内の複数のRNA又はcDNAを増幅して、RNA又はcDNAのアンプリコンを生成することができ、アンプリコンは、コンカテマーを含む。いくつかの態様において、細胞試料内の複数のRNA又はcDNA分子は、パドロックプローブの環状化及びローリングサークル増幅ワークフローを行うことによって増幅され得る。いくつかの態様において、方法は、細胞試料内の複数のRNA又はcDNA分子を、第1の標的RNA又はcDNA分子とハイブリダイズする第1の複数の標的特異的パドロックプローブと、第2の標的RNA又はcDNA分子とハイブリダイズする第2の複数の標的特異的パドロックプローブと、を含む複数のパドロ

50

ックプローブと接触させることを含む。

【0291】

いくつかの態様において、パドロックプローブは、一本鎖オリゴヌクレオチドを含む。いくつかの態様において、パドロックプローブは、DNA、RNA、又はDNA及びRNAを含む。いくつかの態様において、個々のパドロックプローブは、第1の末端領域と第2の末端領域との間の内部領域を含み、内部領域は、試料バーコード配列、増幅プライマー結合部位、シーケンシングプライマー結合部位、コンパクションオリゴヌクレオチド結合部位及び／又は表面キャプチャプライマー結合部位を含む少なくとも1つのユニバーサルアダプター配列を含む（図6）。いくつかの態様において、パドロックプローブは、パドロックプローブが結合する所与の標的RNA又は標的cDNAに対応する少なくとも1つの標的バーコード配列を含む。いくつかの態様において、パドロックプローブは、少なくとも1つの固有の識別配列（例えば、固有の分子インデックス（UMI））を含む。いくつかの態様において、パドロックプローブは、少なくとも1つの制限酵素認識配列を含む。

10

【0292】

いくつかの態様において、個々のパドロックプローブは、標的RNA又は標的cDNA分子の部分にハイブリダイズして、複数のRNA-パドロックプローブ複合体又は複数のcDNA-パドロックプローブ複合体を形成する第1及び第2の末端領域（例えば、第1及び第2の結合アーム）を含み、個々の複合体は、RNA又はcDNA分子の近位領域にハイブリダイズして、第1及び第2の末端プローブ末端間にニック又はギャップを形成する第1及び第2の末端プローブ領域を有する。いくつかの態様において、個々のパドロックプローブの第1の末端領域は、標的RNA又はcDNA分子の第1の領域に選択的にハイブリダイズする第1の標的特異的配列を有し、個々のパドロックプローブの第2の末端領域は、同じ標的RNA又はcDNA分子の第2の領域に選択的にハイブリダイズする第2の標的特異的配列を有し、ハイブリダイズした第1及び第2の末端領域の間にニック又はギャップが形成され、それによってパドロックプローブを環状化する（例えば、図7）。

20

【0293】

いくつかの態様において、パドロックプローブは、正準ヌクレオチド及び／又はヌクレオチド類似体を含む。いくつかの態様において、パドロックプローブは、ヌクレアーゼ分解（例えば、リボヌクレアーゼ分解）に対する耐性を付与するように修飾される。例えば、パドロックプローブは、少なくとも1つのホスホロチオエートジエステル結合をそれらの5'末端において含み、これは、パドロックプローブがヌクレアーゼ分解に対する耐性を有するようにすることができる。いくつかの態様において、パドロックプローブは、2〜5つ以上の連続したホスホロチオエートジエステル結合をそれらの5'末端に含む。いくつかの態様において、パドロックプローブは、少なくとも1つのリボヌクレオチド、及び／又は少なくとも1つの2'-O-メチル、2'-O-メトキシエチル（MOE）、2'フルオロベースのヌクレオチドを含む。いくつかの態様において、パドロックプローブは、リン酸化3'末端を含む。いくつかの態様において、パドロックプローブは、少なくとも1つのロックド核酸（LNA）塩基を含む。いくつかの態様において、パドロックプローブは、（例えば、ポリヌクレオチドキナーゼを使用して）リン酸化5'末端を含む。

30

40

【0294】

いくつかの態様において、パドロックプローブのセット（例えば、複数のパドロックプローブ）内の個々のパドロックプローブは、標的RNA又はcDNA分子の同一の標的領域にハイブリダイズして、同一のRNA又はcDNA配列を有する複数のRNA-パドロックプローブ複合体又は複数のcDNA-パドロックプローブ複合体を形成する第1及び第2の末端領域を含む。

【0295】

いくつかの態様において、パドロックプローブのセット（例えば、複数のパドロックプローブ）は、パドロックプローブの少なくとも2つのサブセットを含む。いくつかの態様

50

において、パドロックプローブの第1のサブセット内の個々のパドロックプローブは、標的RNA又はcDNA分子の同じ標的領域（例えば、第1の標的領域）にハイブリダイズして、同じRNA又はcDNA配列を有する第1の複数のRNA－パドロックプローブ複合体又は第1の複数のcDNA－パドロックプローブ複合体を形成する第1及び第2の末端領域を含む。いくつかの態様において、パドロックプローブの第2のサブセットにおける個々のパドロックプローブは、標的RNA又はcDNA分子の同じ標的領域（例えば、第2の標的領域）にハイブリダイズして、同じcDNA配列を有する第2の複数のRNA－パドロックプローブ複合体又は第2の複数のcDNA－パドロックプローブ複合体を形成する第1及び第2の末端領域を含む。いくつかの態様において、パドロックプローブの第1及び第2のサブセットは、同じ標的RNA又はcDNA分子の異なる標的領域にハイブリダイズする。いくつかの態様において、パドロックプローブの第1及び第2のサブセットは、異なる標的RNA又はcDNA分子の異なる標的領域にハイブリダイズする。いくつかの態様において、パドロックプローブのセットは、パドロックプローブの2～10のサブセット、又はパドロックプローブの10～25のサブセット、又はパドロックプローブの25～50のサブセット、又はパドロックプローブの最大100のサブセットを含む。いくつかの態様において、パドロックプローブのセットは、パドロックプローブの少なくとも100のサブセット、パドロックプローブの少なくとも500のサブセット、パドロックプローブの少なくとも1000のサブセット、パドロックプローブの少なくとも10,000のサブセット、又はパドロックプローブのより多くのサブセットを含む。

10

20

【0296】

いくつかの態様において、ニックを酵素的にライゲーションして、共有結合性閉鎖環状パドロックプローブを生成することができる。いくつかの態様において、リガーゼ酵素は、マッチしたハイブリダイズされた末端とミスマッチしたハイブリダイズされた末端とを区別して、標的特異的ハイブリダイゼーションを確実にすることができる。いくつかの態様において、ライゲーション反応は、T3、T4、T7、又はTaq DNAリガーゼ酵素を含むリガーゼ酵素の使用を含む。

【0297】

いくつかの態様において、ハイブリダイズされた第1の末端領域と第2の末端領域との間のギャップのサイズは、1～25塩基である。ハイブリダイズされたパドロックプローブの3'OH末端は、標的cDNA分子（又は標的RNA分子）をテンプレートとして使用する、ポリメラーゼ触媒充填反応（例えば、ギャップ充填反応）のための開始部位として機能することができる。充填反応の後、残りのニックを酵素的にライゲーションして、共有結合性閉鎖環状パドロックプローブを生成することができる。

30

【0298】

いくつかの態様において、ギャップ充填反応は、環状化パドロックプローブをDNAポリメラーゼ及び複数のヌクレオチドと接触させることを含む。いくつかの態様において、DNAポリメラーゼは、E. coli DNAポリメラーゼI、E. coli DNAポリメラーゼIのクレノウ断片、T7 DNAポリメラーゼ、又はT4 DNAポリメラーゼを含む。いくつかの態様において、リガーゼ酵素は、マッチしたハイブリダイズされた末端とミスマッチしたハイブリダイズされた末端とを区別して、標的特異的ハイブリダイゼーションを確実にすることができる。いくつかの態様において、ライゲーション反応は、T3、T4、T7、又はTaq DNAリガーゼ酵素を含むリガーゼ酵素の使用を含む。

40

【0299】

本明細書に記載される方法のうちのいずれかにおいて、複数の共有結合性閉鎖環状パドロックプローブをローリングサークル増幅反応に供して、単位の2つ以上のタンデムコピーを各々有する複数のコンカテマー分子を生成することができ、単位は、標的RNA分子に対応する標的配列、並びにユニバーサルアダプター配列（複数可）、固有の分子インデックス配列（複数可）、及び／又は制限酵素認識配列（複数可）を含むパドロックプローブによって保有される任意の追加の配列（複数可）を含む。

【0300】

50

いくつかの態様において、ローリングサークル増幅反応は、共有結合性閉鎖環状化パドロックプローブを、増幅プライマー（例えば、ユニバーサルローリングサークル増幅プライマー）、鎖置換DNAポリメラーゼ、及び複数のヌクレオチドと、個々の増幅プライマーを共有結合性閉鎖パドロックプローブにハイブリダイズするのに好適な条件下で、並びに共有結合性閉鎖パドロックプローブをテンプレート分子として使用してプライマー伸長を行うのに好適な条件下で接触させて、核酸コンカテマーを生成することを含む。いくつかの態様において、ローリングサークル増幅反応における複数のヌクレオチドは、dATP、dGTP、dCTP、dTTP、及び／又はdUTPのうちの2つ以上のいずれかの混合物を含む。いくつかの態様において、本明細書に記載されるローリングサークル増幅反応のうちのは、複数のコンパクションオリゴヌクレオチドの存在下又は不在下で行われ得る。

10

【0301】

いくつかの態様において、ローリングサークル増幅反応がdUTPを含む複数のヌクレオチドを含む場合、得られるコンカテマーは、細胞試料をスクシンイミドエステル（NHS）、マレイミド（Sulfo-SMCC）、イミドエステル（DMP）、カルボジイミド（DCC、EDC）又はフェニルアジドで処理することによって、架橋反応性基に架橋され得る。いくつかの態様において、架橋反応性基の重合は、光又はUV光で開始され得る。いくつかの態様において、得られるコンカテマーは、架橋アガロース、架橋デキストラン、若しくは架橋ポリエチレングリコール（PEG）、ポリアクリルアミド、アルギン酸セルロース、又はポリアミドで細胞試料を処理することによって、マトリックスに架橋させることができる。いくつかの態様において、PEGは、一端又は両端にスルホ-NHSEステル部分、例えば、PEG化ビス（スルホスクシンイミジル）スベレート（例えば、Thermo Fisher Scientific、カタログ番号21582からのBS（PEG）9）を含む。

20

【0302】

いくつかの態様において、ローリングサークル増幅反応は、定温（例えば、等温）で行われ得、定温は、室温～約30℃、又は約30～40℃、又は約40～50℃、又は約50～65℃である。

【0303】

いくつかの態様において、鎖置換活性を有するDNAポリメラーゼは、phi29 DNAポリメラーゼ、Bst DNAポリメラーゼの大きい断片、Bsu DNAポリメラーゼの大きい断片、及びBca（エキソ）DNAポリメラーゼ、E. coli DNAポリメラーゼのクレノウ断片、T5ポリメラーゼ、M-MuLV逆転写酵素、HIVウイルス逆転写酵素、又はDeep Vent DNAポリメラーゼからなる群から選択され得る。いくつかの態様において、phi29 DNAポリメラーゼは、野生型phi29 DNAポリメラーゼ（例えば、ExpedeonからのMagniPhi）、又はバリエーションEquiPhi29 DNAポリメラーゼ（例えば、Thermo Fisher Scientificから）、又はキメラQualiPhi DNAポリメラーゼ（例えば、4basebioから）であることができる。

30

【0304】

いくつかの態様において、ローリングサークル増幅プライマーは、ヌクレアーゼ分解に対する耐性を増加させるように修飾され得る。いくつかの態様において、ローリングサークル増幅プライマーは、少なくとも1つのホスホロチオエートジエステル結合をそれらの5'末端において含み、これは、増幅プライマーがエキソヌクレアーゼ分解に対する耐性を有するようにすることができる。いくつかの態様において、ローリングサークル増幅プライマーは、2～5つ以上の連続したホスホロチオエートジエステル結合をそれらの5'末端に含む。いくつかの態様において、ローリングサークル増幅プライマーは、少なくとも1つのリボヌクレオチド、及び／又は少なくとも1つの2'-O-メチル、若しくは2'-O-メトキシエチル（MOE）のヌクレオチドを含む。

40

【0305】

50

いくつかの態様において、ローリングサークル増幅反応は、コンカテマー分子にハイブリダイズしたときにコンカテマーのサイズ及び／又は形状をコンパクトして、コンパクトなナノボールを形成する複数のコンパクションオリゴヌクレオチドの存在下で行われ得る。いくつかの態様において、コンパクションオリゴヌクレオチドは、コンカテマー分子の一部分にハイブリダイズする一方の端に第1の領域と、同じコンカテマー分子の別の部分にハイブリダイズする他方の端に第2の領域とを有する一本鎖オリゴヌクレオチドを含み、コンパクションオリゴヌクレオチドの所与のコンカテマーへのハイブリダイズは、コンカテマーのサイズ及び／又は形状をコンパクトする。

【0306】

コンパクションオリゴヌクレオチドは、5'領域、任意選択の内部領域（介在領域）、及び3'領域を含む。コンパクションオリゴヌクレオチドの5'及び3'領域は、コンカテマーの任意の部分にハイブリダイズすることができる。コンパクションオリゴヌクレオチドの5'及び3'領域は、コンカテマーの異なる部分にハイブリダイズして、コンカテマーの遠位部分を一緒にし、コンカテマーのコンパクションを引き起こして、DNAナノボールを形成することができる。例えば、コンパクションオリゴヌクレオチドの5'領域は、コンカテマー分子の第1の部分（例えば、ユニバーサルコンパクションオリゴヌクレオチド結合部位）にハイブリダイズするように設計され、コンパクションオリゴヌクレオチドの3'領域は、コンカテマー分子の第2の部分（例えば、ユニバーサルコンパクションオリゴヌクレオチド結合部位）にハイブリダイズするように設計される。RCA中のコンパクションオリゴヌクレオチドの包含は、コンパクションオリゴヌクレオチドの不在下で生成されるコンカテマーと比較して、より緊密なサイズ及び形状を有するDNAナノボールの形成を促進することができる。DNAナノボールのコンパクトで安定した特徴は、シグナル強度を増加させることによってインサイチュシーケンシング精度を向上させ、ナノボールは、複数のシーケンシングサイクルの間、その形状及びサイズを保持する。

【0307】

いくつかの態様において、コンパクションオリゴヌクレオチドは、DNA、RNA、又はDNA及びRNAの組み合わせを含む一本鎖オリゴヌクレオチドを含む。コンパクションオリゴヌクレオチドは、長さが20～150個のヌクレオチド、又は30～100個のヌクレオチド、又は40～80個のヌクレオチドを含むいずれかの長さであることができる。

【0308】

いくつかの態様において、コンパクションオリゴヌクレオチドは、5'領域及び3'領域、並びに任意選択で、5'領域と3'領域との間の中間領域を含む。介在領域は、いずれかの長さ、例えば、2～20個のヌクレオチドであることができる。介在領域は、連続した同一の塩基（例えば、AAA、GGG、CCC、TTT、又はUUU）を有するホモポリマーを含む。介在領域は、非ホモポリマー配列を含む。

【0309】

コンパクションオリゴヌクレオチドの5'領域は、その長さに沿って、コンカテマー分子の第1の部分と完全に相補的であってもよく、又は部分的に相補的であってもよい。コンパクションオリゴヌクレオチドの3'領域は、その長さに沿って、コンカテマー分子の第2の部分と完全に相補的であってもよく、又は部分的に相補的であってもよい。コンパクションオリゴヌクレオチドの5'領域は、コンカテマー分子の第1のユニバーサル配列部分にハイブリダイズすることができる。コンパクションオリゴヌクレオチドの3'領域は、コンカテマー分子の第2のユニバーサル配列部分にハイブリダイズすることができる。

【0310】

いくつかの態様において、コンパクションオリゴヌクレオチドの5'領域は、3'領域と同じ配列を有することができる。コンパクションオリゴヌクレオチドの5'領域は、3'領域とは異なる配列を有することができる。いくつかの態様において、コンパクションオリゴヌクレオチドの3'領域は、5'領域の逆配列である配列を有することができる。いくつ

かの態様において、コンパクションオリゴヌクレオチドの 5' 領域は、3' 領域の逆配列である配列を有することができる。

【0311】

いくつかの態様において、コンパクションオリゴヌクレオチドのうちのいずれかの 3' 領域は、2'-O-メチル RNA 塩基（例えば、指定された mUmUmU）を含む末端 3' 末端に追加の 3 つの塩基を含むことができるか、又は末端 3' 末端が追加の 2'-O-メチル RNA 塩基を欠くことができる。

【0312】

いくつかの態様において、コンパクションオリゴヌクレオチドは、ある特定の官能性を付与するために、それらの 5' 又は 3' 末端において 1 つ以上の修飾塩基又は結合を含む。いくつかの態様において、コンパクションオリゴヌクレオチドは、エキソヌクレアーゼ耐性を付与するために、少なくとも 1 つのホスホロチオエート結合をその 5' 及び／又は 3' 末端に含む。いくつかの態様において、3' 末端又はその付近の少なくとも 1 つのヌクレオチドは、エキソヌクレアーゼ耐性を付与する 2' フルオロ塩基を含む。いくつかの態様において、コンパクションオリゴヌクレオチドの 3' 末端は、ポリメラーゼ触媒伸長をブロックする少なくとも 1 つの 2'-O-メチル RNA 塩基を含む。例えば、コンパクションオリゴヌクレオチドの 3' 末端は、2'-O-メチル RNA 塩基（例えば、指定された mUmUmU）を含む 3 つの塩基を含む。いくつかの態様において、コンパクションオリゴヌクレオチドは、ポリメラーゼ触媒伸長をブロックするその 3' 末端に 3' 逆位 dT を含む。いくつかの態様において、コンパクションオリゴヌクレオチドは、ポリメラーゼ触媒伸長をブロックする 3' リン酸化を含む。いくつかの態様において、コンパクションオリゴヌクレオチドの内部領域は、少なくとも 1 つのロックド核酸（LNA）を含み、少なくとも 1 つの LNA は、コンパクションオリゴヌクレオチドをコンカテマー分子にハイブリダイズすることによって形成される二重鎖の熱安定性を増加させる。いくつかの態様において、コンパクションオリゴヌクレオチドは、（例えば、ポリヌクレオチドキナーゼを使用して）リン酸化された 5' 末端を含む。

【0313】

いくつかの態様において、コンパクションオリゴヌクレオチドは、
5'-CATGTAATGCACGTACTTTTCAGGGTAACAATGTAATG
CACGTACTTTTCAGGGT-3'、
（配列番号 1）を含む。いくつかの態様において、コンパクションオリゴヌクレオチドは、2'-O-メチル RNA 塩基（例えば、指定された mUmUmU）を含む末端 3' 末端に追加の 3 つの塩基を含むか、又は末端 3' 末端が追加の 2'-O-メチル RNA 塩基を欠く。

【0314】

いくつかの態様において、コンパクションオリゴヌクレオチドは、連続したグアニンを有する少なくとも 1 つの領域を含むことができる。例えば、コンパクションオリゴヌクレオチドは、2、3、4、5、又は 6 つ以上の連続したグアニンを有する少なくとも 1 つの領域を含むことができる。いくつかの態様において、コンパクションオリゴヌクレオチドは、グアニン四重鎖構造を形成することができる 4 つの連続したグアニンを含む（図 25 を参照されたい）。グアニン四重鎖構造は、フーグスティーン水素結合を介して安定化され得る。

グアニン四重鎖構造は、カリウム、ナトリウム、リチウム、ルビジウム、又はセシウムを含む中央カチオンによって安定化され得る。

【0315】

少なくとも 1 つのコンパクションオリゴヌクレオチドは、グアニン四重鎖（図 25）を形成し、コンカテマー内のユニバーサル結合配列にハイブリダイズすることができ、これにより、コンカテマーが折り畳まれて、分子内 G 四重鎖構造を形成することができる（図 26）。コンカテマーは、コンパクトなナノボールを形成するように自己崩壊することができる。ナノボールにおけるグアニン四重鎖及び G 四重鎖の形成は、ナノボールの安定性

を増加させて、細胞試料内でのシーケンシング中に pH、温度の変化及び／又は試薬の繰り返される流れに耐えられるコンパクトなサイズ及び形状を保持し得る。

【0316】

いくつかの態様において、ローリングサークル増幅反応における複数のコンパクションオリゴヌクレオチドは、同じ配列を有する。あるいは、ローリングサークル増幅反応における複数のコンパクションオリゴヌクレオチドは、異なる配列を有する2つ以上の異なるコンパクションオリゴヌクレオチド集団の混合物を含む。

【0317】

いくつかの態様において、固定化されたコンカテマーテンプレート分子は、コンパクトな核酸ナノボールに自己崩壊することができる。ナノボールはイメージングされ得、FWHM測定が得られて、ナノボールの形状／サイズが得られ得る。

【0318】

いくつかの態様において、ローリングサークル増幅反応におけるコンパクションオリゴヌクレオチドの包含は、DNAナノボールへのコンカテマーの崩壊を促進することができる。コンパクションオリゴヌクレオチドを用いて RCA を行うことは、細胞試料内の DNA ナノボールのスポット画像の FWHM（全幅半値）を改善することができる、複数のシーケンシングサイクル中に DNA ナノボールのコンパクトなサイズ及び形状を保持するのに役立つ。いくつかの態様において、DNA ナノボールは、複数のシーケンシングサイクルの間には解明されない。いくつかの態様において、DNA ナノボールのスポット画像は、複数のシーケンシングサイクルの間には拡大しない。いくつかの態様において、DNA ナノボールのスポット画像は、複数のシーケンシングサイクルの間、離散スポットのままである。スポット画像は、ガウススポットとして表され得、サイズは、FWHMとして測定され得る。より小さい FWHM によって示されるより小さいスポットサイズは、典型的には、スポットの改善された画像と相関する。いくつかの態様において、ナノボールスポットの FWHM は、約 10 μm 以下であることができる。

【0319】

一本鎖コンカテマーは、コンパクト DNA ナノボールに崩壊し、各ナノボールは、それらの長さに沿ってポリヌクレオチド単位の多数のタンデムコピーを保有し、ポリヌクレオチド単位は、目的の配列（例えば、標的 RNA 又は標的 cDNA に対応する）及び少なくともユニバーサルシーケンシングプライマー結合部位を含む。各ポリヌクレオチド単位は、シーケンシングプライマー、シーケンシングポリメラーゼ、及び検出可能に標識されたヌクレオチド試薬（例えば、検出可能に標識された多価分子）に結合して、検出可能なシーケンシング複合体（例えば、検出可能な三元複合体）を形成することができる。各ナノボールは、多数の検出可能なシーケンシング複合体を保有する。したがって、ナノボールのコンパクトな性質は、シーケンシングワークフロー中に使用される、検出可能に標識されたヌクレオチド試薬の局所濃度を増加させ、ナノボールから放出されるシグナル強度を増加させて、細胞試料内の蛍光スポットとしてイメージングすることができる別個の検出可能なシグナルを与える。各スポットは、コンカテマーに対応し、各コンカテマーは、細胞試料中の標的 RNA 分子に対応する。細胞試料中で複数のスポットを同時に検出及びイメージングすることができる。シーケンシング中に増加したシグナル強度及び色分化を生成するコンパクトな形状及びサイズを有する DNA ナノボール。

【0320】

本明細書に記載される方法のうちのいずれにおいても、細胞試料は、全細胞、複数の全細胞、無傷の組織、又は無傷の腫瘍を含む。いくつかの態様において、細胞試料は、新鮮な細胞試料、新鮮凍結細胞試料、切片化細胞試料、又は FFP E 細胞試料を含む。いくつかの態様において、細胞試料は、1つ以上の生細胞又は非生細胞を含む。

【0321】

いくつかの態様において、細胞試料は、ウイルス、真菌、原核生物又は真核生物から得ることができる。いくつかの態様において、細胞試料は、動物、昆虫又は植物から得ることができる。いくつかの態様において、細胞試料は、1つ以上のウイルス感染細胞を含む

10

20

30

40

50

。

【0322】

いくつかの態様において、細胞試料は、ヒト、サル、類人猿、イヌ、ネコ、ウシ、ウマ、マウス、ブタ、ヤギ、オオカミ、カエル、魚、植物、昆虫、又は細菌を含むいずれかの生物から得ることができる。

【0323】

いくつかの態様において、細胞試料は、頭部、頸部、脳、乳房、卵巣、子宮頸部、結腸、直腸、子宮内膜、胆嚢、腸、膀胱、前立腺、精巣、肝臓、肺、腎臓、食道、脾臓、甲状腺、下垂体、胸腺、皮膚、心臓、喉頭、又は他の臓器を含むいずれかの臓器から得ることができる。

10

【0324】

本明細書に記載される方法のいずれにおいても、「単細胞培養地」という用語又は関連用語は、典型的には、培養物中の細胞の成長及び／又は増殖を支持する成分を欠く細胞培養地を指す。単細胞培養地を使用して、例えば、細胞試料を洗浄、懸濁、又は希釈することができる。単細胞培養地をある特定の成分と混合して、培養物中の細胞の成長及び／又は増殖を支持することができる細胞培養地を調製することができる。単細胞培養地は、緩衝液、リン酸化合物、ナトリウム化合物、カリウム化合物、カルシウム化合物、マグネシウム化合物、及び／又はグルコースのうちのいずれか1つ又は2つ以上のいずれかの組み合わせを含む。いくつかの態様において、単細胞培養地は、PBS（リン酸緩衝生理食塩水）、DPBS（ダルベッコのリン酸緩衝生理食塩水）、HBSS（ハンクの平衡塩溶液）、DMEM（ダルベッコの改変イーグル培地）、EMEM（イーグルの最小必須培地）、及び／又はEBSSを含む。いくつかの態様において、細胞試料は、本明細書に記載される核酸方法のうちのいずれかを行うステップの前又はその間に、単細胞培養地中に配置することができる。

20

【0325】

本明細書に記載される方法のうちのいずれにおいても、「複細胞培養地」という用語又は関連用語は、補充物も添加物も用いずに培養物中の細胞の成長及び／又は増殖を支持するために使用できる細胞培養地を指す。複細胞培養地は、緩衝系（例えば、HEPES）、無機塩（複数可）、アミノ酸（複数可）、タンパク質（複数可）、ポリペプチド（複数可）、炭水化物（複数可）、脂肪酸（複数可）、脂質（複数可）、プリン（複数可）及びそれらの誘導体（例えば、ヒポキサンチン）、ピリミジン（複数可）及びそれらの誘導体、並びに／又は微量元素（複数可）のうちの2つ以上のいずれかの組み合わせを含むことができる。複細胞培養地は、流体又は組織抽出物から得られる流体を含む。複細胞培養地は、人工細胞培養地を含む。いくつかの態様において、複細胞培養地は、血清含有培地であってもよく、例えば、複細胞培養地は、胎児ウシ血清、血漿、血清、リンパ液、ヒト胎盤臍帯血清、及び羊水などの流体を含む。いくつかの態様において、複細胞培養地は、典型的には（必ずしもそうではないが）定義された細胞培養地である無血清培地であり得る。いくつかの態様において、複細胞培養地は、典型的には（必ずしもそうではないが）組換えポリペプチド、並びに超純粋な無機化合物及び／又は有機化合物を含む、化学的に定義された培地であり得る。いくつかの態様において、複細胞培養地は、例えば、MEM（最小必須培地）及びRPMI-1640（Roswell Park Memorial Institute）を含むタンパク質を含まない培地でありいくつかの態様において、複細胞培養地は、IMDM（Iscoe's Modified Dulbecco's Mediumを含む。いくつかの態様において、複細胞培養地は、DMEM（ダルベッコの改変イーグル培地）を含む。いくつかの態様において、細胞試料は、本明細書に記載される核酸方法のうちのいずれかを行うステップの前又はその間に、複細胞培養地中に配置することができる。

30

40

【0326】

本明細書に記載される方法のうちのいずれにおいても、細胞試料は、固定細胞試料を含む。いくつかの態様において、細胞試料は、細胞及びその内容物を保存して分解を阻害し

50

、細胞溶解を阻害することができる固定試薬（例えば、固定試薬）で処理することができる。例えば、固定試薬は、細胞試料によって保有されるRNAを保存することができる。いくつかの態様において、固定試薬は、細胞試料からの核酸の喪失を阻害する。

【0327】

いくつかの態様において、固定試薬は、RNAを架橋して、RNAが細胞試料から逃げるのを防ぐことができる。いくつかの態様において、架橋固定試薬は、アルデヒド、ホルムアルデヒド、パラホルムアルデヒド、ホルマリン、グルタルアルデヒド、イミドエステル、N-ヒドロキシスクシンイミドエステル（NHS）、及び／又はグリオキサール（二官能性アルデヒド）のいずれかの組み合わせを含む。

【0328】

いくつかの態様において、固定試薬は、少なくとも1つのアルコールを含み、少なくとも1つのアルコールは、メタノール又はエタノールを含む。いくつかの態様において、固定試薬は、アセトンを含む少なくとも1つのケトンを含む。いくつかの態様において、固定試薬は、酢酸、氷酢酸及び／又はピクリン酸を含む。いくつかの態様において、固定試薬は、塩化水銀を含む。いくつかの態様において、固定試薬は、硫酸亜鉛又は塩化亜鉛を含む亜鉛塩を含む。いくつかの態様において、固定試薬は、ポリペプチドを変性させることができる。

【0329】

いくつかの態様において、固定試薬は、4% w/vのパラホルムアルデヒドを水/PBSに対して含む。いくつかの態様において、固定試薬は、中性pHで10%の35%ホルムアルデヒドを含む。いくつかの態様において、固定試薬は、2% v/vのグルタルアルデヒドを水/PBSに対して含む。いくつかの態様において、固定試薬は、25%の37%ホルムアルデヒド溶液、70%のピクリン酸、及び5%の酢酸を含む。

【0330】

いくつかの態様において、細胞試料を、約30～60分間、4%のパラホルムアルデヒドで支持体上に固定し、PBSで洗浄することができる。

【0331】

いくつかの態様において、細胞試料は、染色されてもよく、脱染色されてもよく、又は染色されなくてもよい。

【0332】

本明細書に記載される方法のうちのいずれにおいても、細胞試料は、透過化された細胞試料を含む。いくつかの態様において、この方法は、細胞膜を変化させて、細胞への実験試薬の浸透を可能にする透過化試薬で細胞試料を処理することを含む。例えば、透過化試薬は、細胞膜から膜脂質を除去する。いくつかの態様において、細胞試料は、有機溶媒、洗剤、化合物、架橋剤、及び／又は酵素のいずれかの組み合わせを含む透過化試薬で処理することができる。いくつかの態様において、有機溶媒は、アセトン、エタノール、及びメタノールを含む。いくつかの態様において、洗浄剤は、サポニン、Triton X-100、Tween-20、ドデシル硫酸ナトリウム（SDS）、N-ラウロイルサルコシナトリウム塩溶液、又は非イオン性ポリオキシエチレン界面活性剤（例えば、NP40）を含む。いくつかの態様において、架橋剤は、パラホルムアルデヒドを含む。いくつかの態様において、酵素は、トリプシン、ペプシン又はプロテアーゼ（例えば、プロテイナーゼK）を含む。いくつかの態様において、細胞は、アルカリ性条件、又はプロテアーゼ酵素による酸性条件を使用して透過化され得る。いくつかの態様において、透過化試薬は、水及び／又はPBSを含む。

【0333】

例えば、固定細胞は、70%エタノールで約30～60分間透過化され得、透過化試薬は、PBS-T（例えば、0.05% Tween-20を有するPBS）と交換され得る。いくつかの態様において、細胞を、3%パラホルムアルデヒド及び0.1%グルタルアルデヒドで約30～60分間後固定し、PBS-Tで複数回洗浄することができる。

【0334】

本明細書に記載される方法のいずれにおいても、細胞試料に、膨潤性ポリ電解質ハイドロゲル（米国特許第10,309,879号及びChen 2015 Science 347:543、これらの文書の内容は参照によりその全体が組み込まれる）を注入する。いくつかの態様において、固定細胞試料及び透過化された細胞試料は、アクリル酸ナトリウム、アクリルアミド、及び架橋剤N-N'-メチレンビスアクリルアミドを注入することができる。いくつかの態様において、過硫酸アンモニウム（APS）開始剤及びテトラメチルエチレンジアミン（TEMED）促進剤を注入して重合を達成した。いくつかの態様において、細胞試料は、タンパク質分解のためにプロテイナーゼKを注入され、消化緩衝液中でインキュベートされ得る。いくつかの態様において、細胞試料内のゲルは、水の添加によって膨潤し得る。

10

【0335】

本明細書に記載される方法のうちのいずれにおいても、細胞試料内の複数のRNAは、cDNAに変換され得る。いくつかの態様において、方法は、固定細胞試料及び透過化された細胞試料内の複数のRNAを、(i)複数の逆転写プライマー、(ii)複数の逆転写酵素、及び(iii)複数のヌクレオチドと、細胞試料中で複数のcDNA分子（例えば、複数の第1の鎖cDNA分子）を生成するための逆転写反応を行うのに好適な条件下で接触させることを含む。いくつかの態様において、第2の鎖cDNA分子の合成は省略される。いくつかの態様において、細胞試料内のRNAは、cDNAに変換されず、RNAは、標的特異的パドロックプローブにハイブリダイズされる。

20

【0336】

いくつかの態様において、逆転写酵素は、RNA依存性DNAポリメラーゼ活性を示す。いくつかの態様において、逆転写酵素は、AMV（鳥骨髄芽細胞症ウイルス）、MuLV（モロニーマウス白血病ウイルス）、又はHIV（ヒト免疫不全ウイルス）からの逆転写酵素を含む。いくつかの態様において、逆転写酵素は、低減されたRNase H活性を示す組換え酵素、例えば、REVERTAID（例えば、Thermo Fisher Scientificから、カタログ番号EP0441）を含む。いくつかの態様において、逆転写酵素は、MULTISCRIBE（例えば、Thermo Fisher Scientificから、カタログ番号4311235）、THERMOSCRIPT（例えば、Thermo Fisher Scientificから、カタログ番号12236-014）、又はARRAYSCRIPT（例えば、Ambionから、AM2048）を含む、市販の酵素であり得る。いくつかの態様において、逆転写酵素は、上付き文字II（例えば、カタログ番号18064014）、上付き文字III（例えば、カタログ番号18080044）、又は上付き文字IV酵素（例えば、カタログ番号18090010）（全て、Invitrogenからの上付き文字酵素）を含む。いくつかの態様において、逆転写反応は、RNase阻害剤を含み得る。

30

【0337】

いくつかの態様において、逆転写プライマーは、DNA、RNA、又はキメラDNA/RNAを含む一本鎖オリゴヌクレオチドを含む。いくつかの態様において、逆転写プライマーは、アデニン（A）、チミン（T）、グアニン（G）、シトシン（C）、ウラシル（U）、及び／又はイノシン（I）の任意の組み合わせである。いくつかの態様において、逆転写プライマーは、いずれかの長さ、例えば、5～25塩基、又は25～50塩基、又は50～75塩基、又は75～100塩基、又はより長い長さであることができる。逆転写プライマーは、各々、5'末端及び3'末端を含む。いくつかの態様において、逆転写プライマーの3'末端は、ポリメラーゼ触媒プライマー伸長反応におけるヌクレオチド重合開始部位として機能する3'OH部分を含むことができる。いくつかの態様において、逆転写プライマーの3'末端は、ポリメラーゼ触媒プライマー伸長反応をブロックする鎖終結部分を有する。鎖終結部分は、3'糖位を伸長可能な3'OHに変換するために除去することができる。

40

【0338】

いくつかの態様において、逆転写プライマーは、ヌクレアーゼ分解（例えば、リボヌク

50

レアーゼ分解) に対する耐性を付与するように修飾される。例えば、逆転写プライマーは、少なくとも1つのホスホロチオエートジエステル結合をそれらの5'末端に含み、これは、逆転写プライマーがヌクレアーゼ分解に対する耐性を有するようにすることができる。いくつかの態様において、逆転写プライマーは、2~5つ以上の連続したホスホロチオエートジエステル結合をそれらの5'末端に含む。いくつかの態様において、複数の逆転写プライマーは、少なくとも1つのリボヌクレオチド、及び／又は少なくとも1つの2'-O-メチル、2'-O-メトキシエチル(MOE)、2'フルオロベースのヌクレオチドを含む。いくつかの態様において、逆転写プライマーは、リン酸化3'末端を含む。いくつかの態様において、逆転写プライマーは、ロックド核酸(LNA)塩基を含む。いくつかの態様において、逆転写プライマーは、(例えば、ポリヌクレオチドキナーゼを使用して)リン酸化5'末端を含む。

10

【0339】

いくつかの態様において、逆転写プライマーの全長は、RNA分子の一部分にハイブリダイズすることができる。いくつかの態様において、個々の逆転写プライマーは、RNA分子の一部分にハイブリダイズする配列を有する3'領域と、RNA分子にハイブリダイズしないテールを保有する5'領域と、を含む。いくつかの態様において、5'テールは、試料バーコード配列、増幅プライマー結合部位、シーケンシングプライマー結合部位、コンパクションオリゴヌクレオチド結合部位、及び／又は表面キャプチャプライマー結合部位のうちのいずれか1つ又は2つ以上の任意の組み合わせを含むユニバーサルアダプター配列を含む。いくつかの態様において、5'テールは、固有の識別配列(例えば、固有の分子インデックス(UMI))を含む。いくつかの態様において、5'テールは、制限酵素認識配列を含む。いくつかの態様において、個々の逆転写プライマーは、ホモポリマー配列、例えば、ポリA、ポリT、ポリC、ポリG又はポリUを有する3'領域の少なくとも一部分を含む。いくつかの態様において、逆転写プライマーは、RNA分子の5'若しくは3'末端、又はRNA分子の内部部分を含む、RNA分子の任意の部分にハイブリダイズすることができる。

20

【0340】

いくつかの態様において、複数の逆転写プライマーは、第1の標的RNAに選択的にハイブリダイズする標的特異的逆転写プライマーの第1のサブ集団(例えば、標的トランスクリプトミクス)を含む。いくつかの態様において、複数の逆転写プライマーは、第2の標的RNAに選択的にハイブリダイズする標的特異的逆転写プライマーの第2のサブ集団を更に含む。いくつかの態様において、標的特異的逆転写プライマーは、標的RNA分子にハイブリダイズする3'領域に所定の配列を含む。いくつかの態様において、逆転写プライマーの所定の配列部分は、長さが4~20塩基、又は20~40塩基、又は40~50塩基であり得る。

30

【0341】

いくつかの態様において、標的特異的逆転写プライマーの第1のサブ集団は、ハウスキーピング遺伝子によって細胞試料中で転写されるRNAに選択的にハイブリダイズすることができる。いくつかの態様において、ハウスキーピング遺伝子の選択は、本明細書に記載されるインサイチュ方法に使用される細胞試料の種類に依存し得る。例示的なハウスキーピング遺伝子としては、グリセルアルデヒド-3-リン酸デヒドロゲナーゼ(GAPDH)、ベーターアクチン(ACTB)、チューブリン、PPIA(ペプチジルプロリンシストランスイソメラーゼ)、NME4(NME/NM23ヌクレオシドジホスフェートキナーゼ4)、SMARCA11(SWI/SNF関連マトリックスアクチン依存性クロマチン調節因子、サブファミリーA、例えば1)、及びPOMK(タンパク質-O-マンノースキナーゼ)が挙げられる。当業者は、多数のハウスキーピング遺伝子のうちのいずれかからのRNA転写産物にハイブリダイズする標的特異的逆転写プライマーの第1のサブ集団を設計することができる。

40

【0342】

いくつかの態様において、標的特異的逆転写プライマーの第2のサブ集団は、検査され

50

る細胞試料中で発現される遺伝子（例えば、細胞特異的又は組織特異的RNA）から転写されるRNAに選択的にハイブリダイズすることができる。

【0343】

いくつかの態様において、複数の逆転写プライマーは、第1の標的RNAにハイブリダイズするランダム配列逆転写プライマーの第1のサブ集団（例えば、全トランスクリプトミクス）を含む。いくつかの態様において、複数の逆転写プライマーは、第2の標的RNAにハイブリダイズするランダム配列逆転写プライマーの第2のサブ集団を更に含む。いくつかの態様において、逆転写プライマーは、RNA分子にハイブリダイズする3'領域にランダム配列及び／又は縮重配列を含む。いくつかの態様において、逆転写プライマーのランダム配列又は縮重配列部分は、長さが4～20塩基、又は20～40塩基、又は40～50塩基であり得る。

10

【0344】

シーケンシングポリメラーゼ

本明細書に記載される方法のうちのいずれにおいても、シーケンシングポリメラーゼは、シーケンシング反応を行うために使用され得る。いくつかの態様において、シーケンシングポリメラーゼ（複数可）は、コンカテマーテンプレート分子におけるヌクレオチドの反対側の相補的ヌクレオチドに結合し、組み込むことが可能である。いくつかの態様において、シーケンシングポリメラーゼ（複数可）は、コンカテマーテンプレート分子におけるヌクレオチドの反対側の多価分子の相補的ヌクレオチド単位を結合することが可能である。いくつかの態様において、複数のシーケンシングポリメラーゼは、組換え変異体ポリメラーゼを含む。

20

【0345】

ヌクレオチド及び／又は多価分子でのシーケンシングにおける使用に好適であるポリメラーゼの例としては、クレノウDNAポリメラーゼ；*Thermus aquaticus* DNAポリメラーゼI（Taqポリメラーゼ）；*KlenTaq*ポリメラーゼ；*Candidatus altiarchaeales*古細菌；*Candidatus Hadarchaeum Yellowstoneense*；*Hadesarchaea*古細菌；*Euryarchaeota*古細菌；*Thermoplasma*古細菌；*Thermococcus*ポリメラーゼ、例えば、*Thermococcus litoralis*、バクテリオファージT7 DNAポリメラーゼ；ヒトアルファ、デルタ、及びエプシロンDNAポリメラーゼ；バクテリオファージポリメラーゼ、例えば、T4、RB69、及びphi29バクテリオファージDNAポリメラーゼ；*Pyrococcus furiosus* DNAポリメラーゼ（Pfuポリメラーゼ）；*Bacillus subtilis* DNAポリメラーゼIII；*E. coli* DNAポリメラーゼIIIIアルファ及びエプシロン；9 degree Nポリメラーゼ；逆転写酵素、例えば、HIVタイプM若しくはO逆転写酵素；トリ骨髄芽球症ウイルス逆転写酵素；モロニー Maus 白血病ウイルス（MMLV）逆転写酵素；又はテロメラーゼが挙げられるが、これらに限定されない。DNAポリメラーゼの更なる非限定的な例としては、様々な古細菌族、例えば、*Aeropyrum*、*Archaeoglobus*、*Desulfurococcus*、*Pyrobaculum*、*Pyrococcus*、*Pyrolobus*、*Pyrodictum*、*Staphylothermus*、*Stetteria*、*Sulfolobus*、*Thermococcus*、及び*Vulcanisaeta*などからのもの、又はそれらのバリエーションが挙げられ、これらは、当該技術分野で既知のそのようなポリメラーゼ、例えば、9 degrees N、VENT、DEEP VENT、THERMINATOR、Pfu、KOD、Pfx、Tgo、及びRB69ポリメラーゼを含む。

30

40

【0346】

結合によるシーケンシング

本明細書に記載される方法のうちのいずれにおいても、シーケンシングは、細胞試料内で結合によるシーケンシング（SBB）反応を行うことを含み、cDNAアンプリコンは、コンカテマー分子である。いくつかの態様において、結合によるシーケンシング（SBB

50

B) 手順は、非標識鎖終結ヌクレオチドを用いる。いくつかの態様において、シーケンシングによる結合 (SBB) は、(a) プライミングされたコンカテマー (例えば、複数のシーケンシングプライマーにアニーリングされたコンカテマー) を、三元複合体安定化条件下で少なくとも2つの別個の混合物と順次接触させるステップであって、少なくとも2つの別個の混合物は、各々、ポリメラーゼ及びヌクレオチドを含み、それによって、順次接触させるステップは、三元複合体安定化条件下で、プライミングされたコンカテマーが、テンプレートにおける第1、第2、及び第3の塩基の種類の塩基の種類のヌクレオチド同族体を用いて接触されるプライミングされたテンプレート核酸をもたらす、順次接触させるステップと、(b) 少なくとも2つの別個の混合物を調べて、三元複合体が形成されたかどうかを判定するステップと、(c) プライミングされたコンカテマーの次の正しいヌクレオチドを識別するステップであって、次の正しいヌクレオチドは、三元複合体がステップ(b)で検出された場合、第1、第2、又は第3の塩基の種類の同族体として識別され、次の正しいヌクレオチドは、ステップ(b)での三元複合体の不在に基づいて、第4の塩基の種類のヌクレオチド同族体であると推定される、識別するステップと、(d) 次の正しいヌクレオチドを、ステップ(b)の後、プライミングされたコンカテマーを添加し、それによって伸長プライマーを産生するステップと、(e) 伸長プライマーを含むプライミングされたコンカテマーに対するステップ(a)～(d)を繰り返すステップと、を含む。例示的な結合によるシーケンシング方法は、米国特許第10,246,744号及び同第10,731,141号(両方の特許の内容全体が参照により本明細書に組み込まれる)に記載されている。

10

20

【0347】

ヌクレオチド及び鎖終結ヌクレオチド

本明細書に記載される方法のうちのいずれにおいても、本明細書に記載されるシーケンシング方法のうちのいずれかは、少なくとも1つのヌクレオチドを用いることができる。ヌクレオチドは、塩基、糖、及び少なくとも1つのリン酸基を含む。いくつかの態様において、複数のヌクレオチドにおける少なくとも1つのヌクレオチドは、芳香族塩基、5炭素糖 (例えば、リボース又はデオキシリボース)、及び1つ以上のリン酸基 (例えば、1つ～10個のリン酸基) を含む。複数のヌクレオチドは、dATP、dGTP、dCTP、dTTP、及びdUTPからなる群から選択される少なくとも1つの種類のヌクレオチドを含むことができる。複数のヌクレオチドは、dATP、dGTP、dCTP、dTTP、及び/又はdUTPからなる群から選択される2つ以上の種類のヌクレオチドのいずれかの組み合わせの混合物において含むことができる。いくつかの態様において、複数のヌクレオチドにおける少なくとも1つのヌクレオチドは、ヌクレオチド類似体でない。いくつかの態様において、複数のヌクレオチドにおける少なくとも1つのヌクレオチドは、ヌクレオチド類似体を含む。

30

【0348】

いくつかの態様において、本明細書に記載されるシーケンシングのための方法のうちのいずれかにおいて、複数のヌクレオチドのうちの少なくとも1つのヌクレオチドは、1、2、又は3つのリン原子の鎖を含み、鎖は、典型的には、エステル又はホスホラミド結合を介して、糖部分の5'炭素に付着している。いくつかの態様において、複数のヌクレオチドにおける少なくとも1つのヌクレオチドは、リン鎖を有する類似体であり、リン鎖において、リン原子と一緒に、介在するO、S、NH、メチレン、又はエチレンと結合している。いくつかの態様において、鎖におけるリン原子は、O、S、又はBH₃を含む置換された側鎖基を含む。いくつかの態様において、鎖は、ホスホロアミデート、ホスホロチオエート、ホスホルジチオエート、及びO-メチルホスホロアミダイト基を含む類似体で置換されたリン酸基を含む。

40

【0349】

いくつかの態様において、本明細書に記載されるシーケンシングのための方法のうちのいずれかにおいて、複数のヌクレオチドにおける少なくとも1つのヌクレオチドは、ターミネーターヌクレオチド類似体を含み、ターミネーターヌクレオチド類似体は、鎖終結部

50

分（例えば、ブロッキング部分）を、糖 2' 位、糖 3' 位、又は糖 2' 及び 3' 位において有する。いくつかの態様において、鎖終結部分は、プライマー伸長反応中の新生の鎖における後続のヌクレオチド単位又は遊離ヌクレオチドのポリメラーゼ触媒組み込みを阻害することができる。いくつかの態様において、鎖終結部分は、3' 糖ヒドロキシル位に付着しており、3' 糖ヒドロキシル位において、糖は、リボース又はデオキシリボース糖部分を含む。いくつかの態様において、鎖終結部分は、ポリメラーゼ触媒ヌクレオチド組み込み反応において後続のヌクレオチドで伸長可能である 3' OH 糖基を有するヌクレオチドを生成するように、3' 糖ヒドロキシル位から除去可能／切断可能である。いくつかの態様において、鎖終結部分は、アルキル基、アルケニル基、アルキニル基、アリル基、アリール基、ベンジル基、アジド基、アミン基、アミド基、ケト基、イソシアネート基、リン酸基、チオ基、ジスルフィド基、カーボネート基、尿素基、又はシリル基を含む。いくつかの態様において、鎖終結部分は、例えば、鎖終結部分を化学剤、pH 変化、光、又は熱と反応させることによって、ヌクレオチドから切断可能／除去可能である。いくつかの態様において、鎖終結部分であるアルキル、アルケニル、アルキニル、及びアリルは、ペリジンを有する又は 2, 3-ジクロロ-5, 6-ジシアノ-1, 4-ベンゾキノン (DDQ) を有するテトラキス (トリフェニルホスフィン) パラジウム (0) (Pd(PPh₃)₄) で切断可能である。いくつかの態様において、鎖終結部分であるアリール及びベンジルは、H₂ Pd/C で切断可能である。いくつかの態様において、鎖終結部分であるアミン、アミド、ケト、イソシアネート、ホスフェート、チオ、ジスルフィドは、ホスフィン、又はベータメルカプトエタノール若しくはジチオトリートル (DTT) を含むチオール基で切断可能である。いくつかの態様において、鎖終結部分であるカーボネートは、MeOH 中の炭酸カリウム (K₂CO₃)、ピリジン中のトリエチルアミン、又は酢酸中の Zn (AcOH) で切断可能である。いくつかの態様において、鎖終結部分である尿素及びシリルは、フッ化テトラブチルアンモニウム、ピリジン-HF、フッ化アンモニウム、又はトリエチルアミントリヒドロフルオリドで切断可能である。

10

20

【0350】

いくつかの態様において、本明細書に記載されるシーケンシングのための方法のうちのいずれかにおいて、複数のヌクレオチドにおける少なくとも 1 つのヌクレオチドは、ターミネーターヌクレオチド類似体を含み、ターミネーターヌクレオチド類似体は、鎖終結部分（例えば、ブロッキング部分）を、糖 2' 位、糖 3' 位、又は糖 2' 及び 3' 位において有する。いくつかの態様において、鎖終結部分は、アジド (azide)、アジド (azido)、及びアジドメチル基を含む。いくつかの態様において、鎖終結部分は、3'-O-アジド又は 3'-O-アジドメチル基を含む。いくつかの態様において、鎖終結部分であるアジド (azide)、アジド (azido)、及びアジドメチル基は、ホスフィン化合物で切断可能／除去可能である。いくつかの態様において、ホスフィン化合物は、誘導体化トリアルキルホスフィン部分又は誘導体化トリアルリルホスフィン部分を含む。いくつかの態様において、ホスフィン化合物は、トリス (2-カルボキシエチル) ホスフィン (TCEP)、又はビス-スルホトリフェニルホスフィン (BS-TPP)、又はトリス (ヒドロキシプロピル) ホスフィン (THPP) を含む。いくつかの態様において、切断剤は、4-ジメチルアミノピリジン (4-DMAP) を含む。

40

【0351】

いくつかの態様において、本明細書に記載されるシーケンシングのための方法のうちのいずれかにおいて、ヌクレオチド類似体は、3'-デオキシヌクレオチド、2', 3'-ジデオキシヌクレオチド、3'-メチル、3'-アジド、3'-アジドメチル、3'-O-アジドアルキル、3'-O-エチニル、3'-O-アミノアルキル、3'-O-フルオロアルキル、3'-フルオロメチル、3'-ジフルオロメチル、3'-トリフルオロメチル、3'-スルホニル、3'-マロニル、3'-アミノ、3'-O-アミノ、3'-スルフヒドラル、3'-アミノメチル、3'-エチル、3'-ブチル、3'-tert-ブチル、3'-フルオレニルメチルオキシカルボニル、3'-tert-ブチルオキシカルボニル、3'-O-アルキルヒドロキシルアミノ基、3'-ホスホロチオート、及び 3'-O-ベンジル、又はそれらの誘導体

50

からなる群から選択される鎖終結部分を含む。

【0352】

いくつかの態様において、本明細書に記載されるシーケンシングのための方法のうちのいずれかにおいて、複数のヌクレオチドは、検出可能なレポーター部分で標識された複数のヌクレオチドを含む。検出可能なレポーター部分は、フルオロフォアを含む。いくつかの態様において、フルオロフォアは、ヌクレオチド塩基に付着している。いくつかの態様において、フルオロフォアは、ヌクレオチド塩基にリンカーで付着しており、リンカーは、塩基から切断可能／除去可能である。いくつかの態様において、複数のヌクレオチドにおけるヌクレオチドのうちの少なくとも1つは、検出可能なレポーター部分で標識されていない。いくつかの態様において、ヌクレオチドに付着した特定の検出可能なレポーター部分（例えば、フルオロフォア）は、ヌクレオチド塩基の検出及び識別を可能にするように、ヌクレオチド塩基（例えば、dATP、dGTP、dCTP、dTTP、又はdUTP）に対応することができる。

10

【0353】

いくつかの態様において、本明細書に記載される核酸分子をシーケンシングするための方法のうちのいずれかにおいて、ヌクレオチド塩基上の切断可能なリンカーは、アルキル基、アルケニル基、アルキニル基、アリル基、アリール基、ベンジル基、アジド基、アミン基、アミド基、ケト基、イソシアネート基、リン酸基、チオ基、ジスルフィド基、カーボネート基、尿素基、又はシリル基を含む切断可能な部分を含む。いくつかの態様において、塩基上の切断可能なリンカーは、切断可能な部分を化学剤、pH変化、光、又は熱と反応させることによって、塩基から切断可能／除去可能である。いくつかの態様において、切断可能な部分であるアルキル、アルケニル、アルキニル、及びアリルは、ペペリジン有する又は2,3-ジクロロ-5,6-ジシアノ-1,4-ベンゾキノン(DDQ)有するテトラキス(トリフェニルホスフィン)パラジウム(0)(Pd(PPh₃)₄)で切断可能である。いくつかの態様において、切断可能な部分であるアリール及びベンジルは、H₂Pd/Cで切断可能である。いくつかの態様において、切断可能な部分であるアミン、アミド、ケト、イソシアネート、ホスフェート、チオ、ジスルフィドは、ホスフィン、又はベータメルカプトエタノール若しくはジチオトリートル(DTT)を含むチオール基で切断可能である。いくつかの態様において、切断可能な部分であるカーボネートは、MeOH中の炭酸カリウム(K₂CO₃)、ピリジン中のトリエチルアミン、又は酢酸中のZn(AcOH)で切断可能である。いくつかの態様において、切断可能な部分である尿素及びシリルは、フッ化テトラブチルアンモニウム、ピリジン-HF、フッ化アンモニウム、又はトリエチルアミントリヒドロフルオリドで切断可能である。

20

30

【0354】

いくつかの態様において、本明細書に記載されるシーケンシングのための方法のうちのいずれかにおいて、ヌクレオチド塩基上の切断可能なリンカーは、アジド(azide)、アジド(azido)、及びアジドメチル基を含む切断可能な部分を含む。いくつかの態様において、切断可能な部分であるアジド(azide)、アジド(azido)、及びアジドメチル基は、ホスフィン化合物で切断可能／除去可能である。いくつかの態様において、ホスフィン化合物は、誘導体化トリアルキルホスフィン部分又は誘導体化トリアリールホスフィン部分を含む。いくつかの態様において、ホスフィン化合物は、トリス(2-カルボキシエチル)ホスフィン(TCEP)、又はビス-スルホトリフェニルホスフィン(BS-TPP)、又はトリス(ヒドロキシプロピル)ホスフィン(THPP)を含む。いくつかの態様において、切断剤は、4-ジメチルアミノピリジン(4-DMAP)を含む。

40

【0355】

いくつかの態様において、本明細書に記載されるシーケンシングのための方法のうちのいずれかにおいて、(例えば、糖2'及び／又は糖3'位における)鎖終結部分、及びヌクレオチド塩基上の切断可能なリンカーは、同じ又は異なる切断可能な部分を有する。いくつかの態様において、(例えば、糖2'及び／又は糖3'位における)鎖終結部分、及び塩

50

基に結合した検出可能なレポーター部分は、同じ化学剤で化学的に切断可能／除去可能である。いくつかの態様において、（例えば、糖 2' 及び／又は糖 3' 位における）鎖終結部分、及び塩基に結合した検出可能なレポーター部位は、異なる化学剤で化学的に切断可能／除去可能である。

【0356】

多価分子

本明細書に記載される方法のうちのいずれかにおいて、シーケンシングは、コアに付着し、スターバースト、ヘルタースケルター、又はボトルブラシの構成を含む任意の構成を有する複数のヌクレオチドアームを含む少なくとも 1 つの多価分子を用いる（例えば、図 16）。多価分子は、（1）コア及び（2）複数のヌクレオチドアームを含み、複数のヌクレオチドアームは、（i）コア付着部分、（ii）PEG 部分を含むスペーサー、（iii）リンカー、及び（iv）ヌクレオチド単位を含み、コアは、複数のヌクレオチドアームに付着しており、スペーサーは、リンカーに付着しており、リンカーは、ヌクレオチド単位に付着している。いくつかの態様において、ヌクレオチド単位は、塩基、糖、及び少なくとも 1 つのリン酸基を含み、リンカーは、塩基を介して、ヌクレオチド単位に付着している。いくつかの態様において、リンカーは、脂肪族鎖又はオリゴエチレングリコール鎖を含み、両方のリンカー鎖は、2～6 つのサブユニットを有する。いくつかの態様において、リンカーはまた、芳香族部分を含む。例示的なヌクレオチドアームを図 20 に示す。例示的な多価分子は、図 16～20 に示される。例示的なスペーサーを図 21（上）に示し、例示的なリンカーを図 21（下）及び図 22 に示す。リンカーに付着した例示的なヌクレオチドを図 23 A～23 D に示す。例示的なビオチン化ヌクレオチドアームを図 24 に示す。

10

20

【0357】

いくつかの態様において、多価分子は、複数のヌクレオチドアームに結合するコアを含み、複数のヌクレオチドアームは、dATP、dGTP、dCTP、dTTP、及び dUTP からなる群から選択される同じ種類のヌクレオチド単位を有する。

【0358】

いくつかの態様において、多価分子は、複数のヌクレオチドアームに付着するコアを含み、各アームは、ヌクレオチド単位を含む。いくつかの態様において、ヌクレオチド単位は、芳香族塩基、5 炭素糖（例えば、リボース又はデオキシリボース）、及び 1 つ以上のリン酸基（例えば、1 つ～10 個のリン酸基）を含む。複数の多価分子は、dATP、dGTP、dCTP、dTTP、及び dUTP からなる群から選択される 1 つの種類のヌクレオチド単位を有する 1 つの種類の多価分子を含むことができる。複数の多価分子は、2 つ以上の種類の多価分子のいずれかの組み合わせの混合物において含むことができ、混合物における個々の多価分子は、dATP、dGTP、dCTP、dTTP、及び／又は dUTP からなる群から選択されるヌクレオチド単位を含む。

30

【0359】

いくつかの態様において、ヌクレオチド単位は、1、2、又は 3 つのリン原子の鎖を含み、鎖は、典型的には、エステル又はホスホラミド結合を介して、糖部分の 5' 炭素に付着している。いくつかの態様において、少なくとも 1 つのヌクレオチド単位は、リン鎖を有するヌクレオチド類似体であり、リン鎖において、リン原子と一緒に、介在する O、S、NH、メチレン、又はエチレンと結合している。いくつかの態様において、鎖におけるリン原子は、O、S、又は BH3 を含む置換された側鎖基を含む。いくつかの態様において、鎖は、ホスホロアミデート、ホスホロチオエート、ホスホルジチオエート、及び O-メチルホスホロアミダイト基を含む類似体で置換されたリン酸基を含む。

40

【0360】

いくつかの態様において、多価分子は、複数のヌクレオチドアームに付着したコアを含み、個々のヌクレオチドアームは、糖 2' 位、糖 3' 位、又は糖 2' 及び 3' 位に鎖終結部分（例えば、ブロッキング部分）を有するヌクレオチド類似体であるヌクレオチド単位を含む。いくつかの態様において、ヌクレオチド単位は、糖 2' 位、糖 3' 位、又は糖 2' 及び

50

3' 位に鎖終結部分（例えば、ブロッキング部分）を含む。いくつかの態様において、鎖終結部分は、プライマー伸長反応中の新生の鎖における後続のヌクレオチド単位又は遊離ヌクレオチドのポリメラーゼ触媒組み込みを阻害することができる。いくつかの態様において、鎖終結部分は、3' 糖ヒドロキシル位に付着しており、3' 糖ヒドロキシル位において、糖は、リボース又はデオキシリボース糖部分を含む。いくつかの態様において、鎖終結部分は、ポリメラーゼ触媒ヌクレオチド組み込み反応において後続のヌクレオチドで伸長可能である3' OH糖基を有するヌクレオチドを生成するように、3' 糖ヒドロキシル位から除去可能／切断可能である。いくつかの態様において、鎖終結部分は、アルキル基、アルケニル基、アルキニル基、アリル基、アリール基、ベンジル基、アジド基、アミン基、アミド基、ケト基、イソシアネート基、リン酸基、チオ基、ジスルフィド基、カーボネート基、尿素基、又はシリル基を含む。いくつかの態様において、鎖終結部分は、例えば、鎖終結部分を化学剤、pH変化、光、又は熱と反応させることによって、ヌクレオチド単位から切断可能／除去可能である。いくつかの態様において、鎖終結部分であるアルキル、アルケニル、アルキニル、及びアリルは、ペペリジン（D D Q）を有する又は2, 3-ジクロロ-5, 6-ジシアノ-1, 4-ベンゾキノン（D D Q）を有するテトラキス（トリフェニルホスフィン）パラジウム（0）（Pd（PPh₃）₄）で切断可能である。いくつかの態様において、鎖終結部分であるアリール及びベンジルは、H₂ Pd/Cで切断可能である。いくつかの態様において、鎖終結部分であるアミン、アミド、ケト、イソシアネート、ホスフェート、チオ、ジスルフィドは、ホスフィン、又はベータメルカプトエタノール若しくはジチオトリオール（D T T）を含むチオール基で切断可能である。いくつかの態様において、鎖終結部分であるカーボネートは、MeOH中の炭酸カリウム（K₂CO₃）、ピリジン中のトリエチルアミン、又は酢酸中のZn（AcOH）で切断可能である。いくつかの態様において、鎖終結部分である尿素及びシリルは、フッ化テトラブチルアンモニウム、ピリジン-HF、フッ化アンモニウム、又はトリエチルアミントリヒドロフルオリドで切断可能である。

10

20

【0361】

いくつかの態様において、ヌクレオチド単位は、糖2' 位、糖3' 位、又は糖2' 及び3' 位に鎖終結部分（例えば、ブロッキング部分）を含む。いくつかの態様において、鎖終結部分は、アジド（azide）、アジド（azido）、及びアジドメチル基を含む。いくつかの態様において、鎖終結部分は、3'-O-アジド又は3'-O-アジドメチル基を含む。いくつかの態様において、鎖終結部分であるアジド（azide）、アジド（azido）、及びアジドメチル基は、ホスフィン化合物で切断可能／除去可能である。いくつかの態様において、ホスフィン化合物は、誘導体化トリアルキルホスフィン部分又は誘導体化トリアリールホスフィン部分を含む。いくつかの態様において、ホスフィン化合物は、トリス（2-カルボキシエチル）ホスフィン（TCPEP）、又はビス-スルホトリフェニルホスフィン（BS-TPP）、又はトリス（ヒドロキシプロピル）ホスフィン（THPP）を含む。いくつかの態様において、切断剤は、4-ジメチルアミノピリジン（4-DMAP）を含む。

30

【0362】

いくつかの態様において、ヌクレオチド単位は、3'-デオキシヌクレオチド、2', 3'-ジデオキシヌクレオチド、3'-メチル、3'-アジド、3'-アジドメチル、3'-O-アジドアルキル、3'-O-エチニル、3'-O-アミノアルキル、3'-O-フルオロアルキル、3'-フルオロメチル、3'-ジフルオロメチル、3'-トリフルオロメチル、3'-スルホニル、3'-マロニル、3'-アミノ、3'-O-アミノ、3'-スルフィド、3'-アミノメチル、3'-エチル、3'-ブチル、3'-tert-ブチル、3'-フルオレニルメチルオキシカルボニル、3'-tert-ブチルオキシカルボニル、3'-O-アルキルヒドロキシルアミノ基、3'-ホスホロチオート、及び3'-O-ベンジル、又はそれらの誘導体からなる群から選択される鎖終結部分を含む。

40

【0363】

いくつかの態様において、多価分子は、複数のヌクレオチドアームに付着するコアを含

50

み、ヌクレオチドアームは、スペーサー、リンカー及びヌクレオチド単位を含み、コア、リンカー及び／又はヌクレオチド単位は、検出可能なレポーター部分で標識される。いくつかの態様において、検出可能なレポーター部分は、フルオロフォアを含む。いくつかの態様において、多価分子に付着した特定の検出可能なレポーター部分（例えば、フルオロフォア）は、ヌクレオチド塩基の検出及び識別を可能にするように、ヌクレオチド単位の塩基（例えば、d A T P、d G T P、d C T P、d T T P、又はd U T P）に対応することができる。

【0364】

いくつかの態様において、多価分子の少なくとも1つのヌクレオチドアームは、検出可能なレポーター部分に付着したヌクレオチド単位を有する。いくつかの態様において、検出可能なレポーター部分は、ヌクレオチド塩基に付着している。いくつかの態様において、検出可能なレポーター部分は、フルオロフォアを含む。いくつかの態様において、多価分子に付着した特定の検出可能なレポーター部分（例えば、フルオロフォア）は、ヌクレオチド塩基の検出及び識別を可能にするように、ヌクレオチド単位の塩基（例えば、d A T P、d G T P、d C T P、d T T P、又はd U T P）に対応することができる。

【0365】

いくつかの態様において、多価分子のコアは、アビジン様又はストレプトアビジン様部分を含み、コア付着部分は、ビオチンを含む。いくつかの態様において、コアは、アビジンタンパク質、並びに少なくとも1つのビオチン部分に結合することができる任意の誘導体、類似体、及び他の非天然形態のアビジンを含むストレプトアビジン型又はアビジン型部分を含む。アビジン部分の他の形態としては、天然及び組換えアビジン及びストレプトアビジン、並びに誘導体化分子、例えば、非グリコシル化アビジン及び切断ストレプトアビジンが挙げられる。例えば、アビジン部分は、アビジンの脱グリコシル化形態、*Streptomyces*（例えば、*Streptomyces avidinii*）によって産生される細菌ストレプトアビジン、並びに誘導体化形態、例えば、N-アシルアビジン、例えば、N-アセチル、N-フタリル、及びN-スクシニルアビジン、並びに市販の製品であるEXTRAVIDIN、CAPTAVIDIN、NEUTRAVIDIN、及びNEUTRALITE AVIDINを含む。

【0366】

いくつかの態様において、本明細書に記載される核酸分子をシーケンシングするための方法のうちのいずれかは、結合複合体を形成することを含んでもよく、結合複合体は、（i）ポリメラーゼ、プライマーと二重鎖を有する核酸コンカテマー分子、及びヌクレオチドを含むか、又は結合複合体は、（ii）ポリメラーゼ、プライマーと二重鎖を有する核酸コンカテマー分子、及び多価分子のヌクレオチド単位を含む。いくつかの態様において、結合複合体は、約0.1、0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9、又は1秒間超の持続時間を有する。結合複合体は、約0.1～0.25秒、若しくは約0.25～0.5秒、若しくは約0.5～0.75秒、若しくは約0.75～1秒、若しくは約1～2秒、若しくは約2～3秒、若しくは約3～4秒、若しくは約4～5秒を超える持続時間を有する、及び／又は本方法が15℃以上、20℃以上、25℃以上、35℃以上、37℃以上、42℃以上、55℃以上、60℃以上、若しくは72℃以上、若しくは80℃以上、又は前述のうちのいずれかによって定義される範囲内で実施されるか又は実施され得る。結合複合体（例えば、三元複合体）は、ポリメラーゼ、テンプレート分子、プライマー、及び／又はヌクレオチド単位若しくはヌクレオチドのうちのいずれかの間の相互作用の解離を引き起こす条件に供されるまで依然として安定している。例えば、解離条件は、結合複合体を、洗浄剤、EDTA、及び／若しくは水のうちのいずれか1つ、又はそれらのいずれかの組み合わせと接触させることを含む。いくつかの態様において、本開示は、結合複合体が、検出ステップにおいて20を超えるコントラストノイズ比を示す表面上に堆積されるか、それに付着するか、又はそれにハイブリダイズされる、当該方法を提供する。いくつかの態様において、本開示は、接触させることが、ヌクレオチド又はヌクレオチド単位がテンプレート核酸の次の塩基に相補的であるときに結合複

10

20

30

40

50

合体を安定化し、ヌクレオチド又はヌクレオチド単位がテンプレート核酸の次の塩基に相補的でないときに結合複合体を不安定化する条件下で実行される、当該方法を提供する。

【0367】

いくつかの態様において、多価分子を用いるシーケンシング方法のうちのいずれかにおいて、複数の第1の複合型ポリメラーゼを複数の多価分子と結合させることは、少なくとも1つのアビディティ複合体を形成し、方法は、以下のステップ、(a)第1の核酸プライマー、第1のシーケンシングポリメラーゼ、及び第1の多価分子をコンカテマーテンプレート分子の第1の部分に結合し、それによって第1の結合複合体を形成することであって、第1の多価分子の第1のヌクレオチド単位が、第1のシーケンシングポリメラーゼに結合する、形成することと、(b)第2の核酸プライマー、第2のシーケンシングポリメラーゼ、及び第1の多価分子を同じコンカテマーテンプレート分子の第2の部分に結合し、それによって第2の結合複合体を形成することであって、第1の多価分子の第2のヌクレオチド単位は、第2のシーケンシングポリメラーゼに結合し、同じ多価分子を含む第1及び第2の結合複合体は、アビディティ複合体を形成する、形成することと、を含む。いくつかの態様において、第1のシーケンシングポリメラーゼは、本明細書に記載されるいずれかの野生型又は変異体ポリメラーゼを含む。いくつかの態様において、第2のシーケンシングポリメラーゼは、本明細書に記載されるいずれかの野生型又は変異体ポリメラーゼを含む。コンカテマーテンプレート分子は、目的の配列及び少なくとも1つのユニバーサルシーケンシングプライマー結合部位のタンデムリピート配列を含む。第1及び第2の核酸プライマーは、コンカテマーテンプレート分子に沿ってシーケンシングプライマー結合部位に結合することができる。例示的な多価分子は、図16～20に示される。

10

20

【0368】

いくつかの態様において、多価分子を用いるシーケンシング方法のうちのいずれかにおいて、方法は、複数の第1の複合型ポリメラーゼを複数の多価分子と結合させて、少なくとも1つのアビディティ複合体を形成することを含み、方法は、以下のステップ、(a)複数のシーケンシングポリメラーゼ及び複数の核酸プライマーを、コンカテマー核酸コンカテマー分子の異なる部分と接触させて、同じコンカテマー分子上に少なくとも第1及び第2の複合ポリメラーゼを形成するステップと、(b)複数の多価分子を、同じコンカテマーテンプレート分子上の少なくとも第1及び第2の複合ポリメラーゼに、複数の多価分子からの単一の多価分子を第1及び第2の複合ポリメラーゼに結合するのに好適な条件下で接触させるステップであって、単一の多価分子の少なくとも第1のヌクレオチド単位は、コンカテマーテンプレート分子の第1の部分にハイブリダイズし、それによって第1の結合複合体(例えば、第1の三元複合体)を形成する第1のプライマーを含む第1の複合ポリメラーゼに結合し、単一の多価分子の少なくとも第2のヌクレオチド単位は、コンカテマーテンプレート分子の第2の部分にハイブリダイズし、それによって第2の結合複合体(例えば、第2の三元複合体)を形成する第2のプライマーを含む第2の複合ポリメラーゼに結合して、接触させることが、第1及び第2の結合複合体中の結合した第1及び第2のヌクレオチド単位のポリメラーゼ触媒による取り込みを阻害するのに好適な条件下で行われ、同じ多価分子に結合する第1及び第2の結合複合体がアビディティ複合体を形成する、接触させるステップと、(c)同じコンカテマーテンプレート分子上で第1及び第2の結合複合体を検出するステップと、(d)第1の結合複合体中の第1のヌクレオチド単位を識別し、それによってコンカテマーテンプレート分子の第1の部分の配列を決定し、第2の結合複合体中の第2のヌクレオチド単位を識別し、それによってコンカテマーテンプレート分子の第2の部分の配列を決定するステップと、を含む。いくつかの態様において、複数のシーケンシングポリメラーゼは、本明細書に記載されるいずれかの野生型又は変異体シーケンシングポリメラーゼを含む。コンカテマーテンプレート分子は、目的の配列及び少なくとも1つのユニバーサルシーケンシングプライマー結合部位のタンデムリピート配列を含む。複数の核酸プライマーは、コンカテマーテンプレート分子に沿ってシーケンシングプライマー結合部位に結合することができる。例示的な多価分子は、図16～20に示される。

30

40

50

【0369】

図16は、多価分子の様々な例示的な構成の概略図である。左（クラスI）：「スターバースト」又は「ヘルタースケルター」構成を有する多価分子の概略図。中央（クラスII）：デンドリマー構成を有する多価分子の概略図。右（クラスIII）：ストレプトアビジンを、ピオチンを有する4アーム又は8アーム型PEG-NHS及びdNTPと反応させることによって形成された複数の多価分子の概略図。ヌクレオチド単位は、「N」と表され、ピオチンは、「B」と表され、ストレプトアビジンは、「SA」と表される。

【0370】

図17は、複数のヌクレオチドアームに付着した一般的なコアを含む例示的な多価分子の概略図である。

10

【0371】

図18は、複数のヌクレオチドアームに付着したデンドリマーコアを含む例示的な多価分子の概略図である。

【0372】

図19は、複数のヌクレオチドアームに付着したコアを含む例示的な多価分子の概略図を示し、ヌクレオチドアームは、ピオチン、スペーサー、リンカー、及びヌクレオチド単位を含む。

【0373】

図20は、コア付着部分、スペーサー、リンカー、及びヌクレオチド単位を含む例示的なヌクレオチドアームの概略図である。

20

【0374】

図21は、例示的なスペーサーの化学構造（上）、並びに11原子リンカー、16原子リンカー、23原子リンカー、及びN3リンカーを含む様々な例示的なリンカーの化学構造（下）を示す。

【0375】

図22は、リンカー1～9を含む様々な例示的なリンカーの化学構造を示す。

【0376】

23Aは、ヌクレオチド単位に連結された／付着した様々な例示的なリンカーの化学構造を示す。

【0377】

図23Bは、ヌクレオチド単位に連結された／付着した様々な例示的なリンカーの化学構造を示す。

30

【0378】

図23Cは、ヌクレオチド単位に連結された／付着した様々な例示的なリンカーの化学構造を示す。

【0379】

図23Dは、ヌクレオチド単位に連結された／付着した様々な例示的なリンカーの化学構造を示す。

【0380】

図24は、例示的なヌクレオチドアームの化学構造を示す。この例において、ヌクレオチド単位は、ピリミジン塩基の5位又はプリン塩基の7位におけるプロパルギルアミン付着を介して、リンカーに接続されている。

40

【0381】

図25は、グアニン四重鎖（例えば、G四重鎖）の概略図である。

【0382】

図26は、例示的な分子内G四重鎖構造の概略図である。

【0383】

フローセル

本明細書に記載される方法のいずれかにおいて、細胞試料は、固体支持体（例えば、フローセル）上に堆積され得る。いくつかの態様において、細胞試料は、壁（例えば、上部

50

又は第1の壁、及び底部又は第2の壁)と、その間のギャップとを有するフローセル上に堆積され、ギャップは、流体で満たされ得、フローセルは、蛍光光学イメージングシステム内に配置される。細胞試料は厚さを有するため、従来のイメージングシステムを使用する場合、イメージングシステムを使用してフローセルの第1及び第2の表面に別々に合焦させる必要があり得る。細胞試料中のコンカテマーのシーケンシング反応の改善されたイメージングのために、フローセルは、2つ以上の蛍光波長でフローセルの第1及び第2の表面に最適なイメージング性能を提供するように設計された2つ以上のチューブレンズを含む、高性能蛍光イメージングシステム内に配置することができる。いくつかの態様において、高性能イメージングシステムは、フローセルの第1及び第2の表面の画像を取得する間に光学システムを再合焦させるように構成された合焦機構を更に備える。いくつかの態様において、高性能イメージングシステムは、第1のフローセル表面又は第2のフローセル表面のうちの少なくとも1つ上の2つ以上の視野をイメージングするように構成される。

10

【0384】

支持体及びコーティング

本明細書に記載される方法のうちのいずれにおいても、固体支持体は、細胞接着を促進するコーティングを有するフローセルを含む。いくつかの態様において、フローセルは、平面又は非平面の支持体であり得る支持体を含む。支持体は、固体又は半固体であり得る。いくつかの態様において、支持体は、多孔質、半多孔質、又は非多孔質であり得る。支持体は、ガラス、プラスチック、又はポリマー材料などのいずれかの材料から作製され得る。いくつかの態様において、支持体の表面は、1つ以上の化合物でコーティングされ、支持体上に不動態化層を生成することができる(図15)。いくつかの態様において、不動態化層は、多孔質又は半多孔質層を形成する。いくつかの態様において、支持体は、リジン化合物、ポリーリジン化合物、アルギニン化合物、又はアミノ末端化合物でコーティングされている。支持体は、非分枝化合物、分枝化合物、又は非分枝及び分枝化合物の混合物でコーティングすることができる。いくつかの態様において、支持体は、細胞試料から核酸をキャプチャするための表面プライマーでコーティングされている。あるいは、支持体は、表面プライマーを欠いている。

20

【0385】

図15は、ガラス基材及び親水性コーティングの交互の層を含む例示的な低結合支持体の概略図であり、親水性コーティングの交互の層は、ガラスに共有結合的に又は非共有結合的に接着しており、化学反応性官能基を更に含み、化学反応性官能基は、オリゴヌクレオチドプライマー(例えば、キャプチャオリゴヌクレオチド及び環状化オリゴヌクレオチド)のための付着部位として機能する。代替の態様において、支持体は、ガラス、プラスチック、又はポリマー材料などのいずれかの材料から作製され得る。

30

【0386】

支持体は、1つ以上の基材を含み得る。支持体は、ガラス又はプラスチック基材を含み得る。支持体は、光学システムの対物レンズに最も近い透明な上部基材を含むことができる。支持体は、1つ以上のマイクロ流体チャネル及びコンカテマー分子を含むことができ、細胞試料は、マイクロ流体チャネルの表面に固定化される。いくつかの態様において、支持体は、フローセルデバイスに含まれている。

40

【0387】

図1のイメージャ116は、1つ以上の光学システムを含むことができる。本明細書に更に開示されるのは、蛍光イメージングベースのゲノミクス用途のために改善された光学解像度及び画像品質を提供する光学システム設計ガイドライン並びに高性能蛍光イメージング方法及びシステムである。開示された光学イメージングシステム設計は、より大きな視野、増加した空間解像度、改善された変調転送、コントラスト対ノイズ比、及び画像品質、より高い空間サンプリング周波数、一連の画像(例えば、異なる視野の)をキャプチャするために試料面を再配置するときの画像キャプチャ間のより速い遷移、並びに改善されたイメージングシステムデューティサイクルを提供し、したがって、より高いスループ

50

ット画像取得及び分析を可能にする。

【0388】

いくつかの事例では、例えば、デュアルサイド（フローセル）イメージング用途のためのイメージング性能の改善は、対物レンズと組み合わせて電気光学位相板を使用して、フローセルの上部（近い）及び下部（遠い）内側表面を分離する流体層によって引き起こされる光学収差を補償することによって達成され得る。いくつかの事例では、この設計アプローチはまた、例えば、フローセルのどの表面が画像であるかに応じて、光路内又は光路外に移動される動き作動型コンペンセータによって導入される振動を補償し得る。

【0389】

いくつかの事例では、例えば、厚いフローセル壁（例えば、壁（又はカバースリップ）の厚さが700 μm を超える）及び流体チャネル（例えば、流体チャネルの高さ又は厚さが50～200 μm ）の使用を含むデュアルサイド（フローセル）イメージング用途のためのイメージング性能の改善は、厚いフローセル壁及び／又は中間の流体層によって引き起こされる光学収差を対物鏡と組み合わせて補正するチューブレンズ設計を使用することによって、市販の既製の対物鏡を使用する場合でも達成され得る。

【0390】

いくつかの事例では、例えば、マルチチャネル（例えば、2色又は4色）イメージングアプリケーションのためのイメージング性能の改善は、複数のチューブレンズ（各イメージングチャネルに対して1つ）を使用することによって達成され得、各チューブレンズ設計は、そのイメージングチャネルで使用される特定の波長範囲に対して最適化されている。

【0391】

本明細書に開示される例示的な態様は、蛍光イメージングシステムを含み得、当該システムは、a) 1つ以上の指定された波長範囲内の励起光を提供するように構成された少なくとも1つの光源、b) 試料面を励起光に曝露したときに試料面の指定された視野内から生じる蛍光を収集するように構成された対物レンズであって、対物レンズの開口数は、少なくとも0.1、少なくとも0.2、少なくとも0.3、少なくとも0.4、少なくとも0.5、少なくとも0.6、少なくとも0.7、少なくとも0.8、少なくとも0.9、又は前述のもののうちの任意の2つによって定義される範囲内に入る開口数であり、対物レンズの作動距離が、少なくとも400 μm 、少なくとも500 μm 、少なくとも600 μm 、少なくとも700 μm 、少なくとも800 μm 、少なくとも900 μm 、少なくとも1000 μm 、又は前述のもののうちの任意の2つによって定義される範囲内に入る作動距離であり、視野が、少なくとも0.1 mm^2 、少なくとも0.2 mm^2 、少なくとも0.5 mm^2 、少なくとも0.7 mm^2 、少なくとも1 mm^2 、少なくとも2 mm^2 、少なくとも3 mm^2 、少なくとも5 mm^2 、又は少なくとも10 mm^2 、又は前述のもののうちのいずれか2つによって定義される範囲内にある視野である、対物レンズ、並びにc) 少なくとも1つのイメージセンサであって、対物レンズによって収集された蛍光がイメージセンサ上にイメージングされ、蛍光イメージングシステムの空間サンプリング周波数が蛍光イメージングシステムの光学解像度の少なくとも2倍であるように、イメージセンサのピクセル寸法が選択される、少なくとも1つのイメージセンサを備え得る。

【0392】

いくつかの態様において、開口数は、少なくとも0.75であり得る。いくつかの態様において、開口数は、少なくとも1.0である。いくつかの態様において、作動距離は、少なくとも850 μm である。いくつかの態様において、作動距離は、少なくとも1,000 μm である。いくつかの態様において、視野は、少なくとも2.5 mm^2 の面積を有し得る。いくつかの態様において、視野は、少なくとも3 mm^2 の面積を有し得る。いくつかの態様において、空間サンプリング周波数は、蛍光イメージングシステムの光学解像度の少なくとも2.5倍であり得る。いくつかの態様において、空間サンプリング周波数は、蛍光イメージングシステムの光学解像度の少なくとも3倍であり得る。いくつかの態様において、システムは、システムが自動化された方法で一連の2つ以上の蛍光画像を取

得するように構成されるように、X-Y-Z並進ステージを更に備え得、一連の画像の各画像は、異なる視野に対して取得されるか、又は取得され得る。いくつかの態様において、試料面の位置は、異なる視野の画像を取得する間に対物レンズ焦点面の位置と一致するように、X方向、Y方向、及びZ方向に同時に調整されてもよい。いくつかの態様において、X方向、Y方向、及びZ方向における同時調整に必要な時間は、0.3秒未満、0.4秒未満、0.5秒未満、0.7秒未満、若しくは1秒未満、又は前述のうちのいずれか2つによって定義される範囲内にある時間であり得る。いくつかの態様において、システムは、誤差シグナルがZ方向における焦点面及び試料面の位置の差が指定された誤差閾値よりも大きいことを示す場合、異なる視野の画像を取得する前に焦点面の位置を調整するように構成されたオートフォーカス機構を更に備える。いくつかの態様において、指定された誤差閾値は、100nm以上である。いくつかの態様において、指定された誤差閾値は、50nm以下である。いくつかの態様において、システムは、3つ以上の画像センサを備え、システムは、3つ以上の波長範囲の各々における蛍光を異なる画像センサ上にイメージングするように構成される。いくつかの態様において、3つ以上の画像センサの各々と試料面との焦点面の位置の差は、100nm未満である。いくつかの態様において、3つ以上の画像センサの各々と試料面との焦点面の位置の差は、50nm未満である。いくつかの態様において、試料面を再配置し、必要に応じて焦点を調整し、画像を取得するのに必要な合計時間は、1視野当たり0.4秒未満である。いくつかの態様において、試料面を再配置し、必要に応じて焦点を調整し、画像を取得するのに必要な合計時間は、1視野当たり0.3秒未満である。

10

20

【0393】

本明細書の開示者はまた、フローセルのデュアルサイドイメージングのための蛍光イメージングシステムであって、a) フローセル内の試料面の特定の視野内から生じる蛍光を収集するように構成された対物レンズと、b) 対物レンズと少なくとも1つのイメージセンサとの間に配置された少なくとも1つのチューブレンズと、を含み、少なくとも1つのチューブレンズは、フローセルの内側表面をイメージングするときに、対物レンズ、少なくとも1つのチューブレンズ、及び少なくとも1つのイメージセンサの組み合わせについてのイメージング性能メトリックを補正するように構成され、フローセルは、少なくとも700 μm の壁厚と、少なくとも50 μm の上部内側表面と下部内側表面との間のギャップとを有し、イメージング性能メトリックは、フローセルの上部内側表面又は下部内側表面を、フローセルと少なくとも1つのイメージセンサとの間の光学経路内又は光学経路外に光学コンペンセータを移動させることなく、チューブレンズの1つ以上の光学素子を光学経路に沿って移動させることなく、かつチューブレンズの1つ以上の光学素子を光学経路内又は光学経路外に移動させることなく、イメージングするために実質的に同じである。

30

【0394】

いくつかの態様において、対物レンズは、市販の顕微鏡対物レンズであり得る。いくつかの態様において、市販の顕微鏡対物レンズは、少なくとも0.3の開口数を有し得る。いくつかの態様において、対物レンズは、少なくとも700 μm の作動距離を有し得る。いくつかの態様において、対物レンズは、0.17mmのカバースリップ厚さ（又はフローセル壁厚さ）、又は0.17mmよりも大きい又は小さい厚さのカバースリップ厚さ（又はフローセル壁厚さ）を補償するように補正され得る。いくつかの態様において、光学システムは、カバースリップ厚さ、フローセル厚さ、又は所望の焦点面間の距離を補正するように補正され得る。いくつかの態様において、当該補正は、レンズ又は光学アセンブリなどの補正光学システムを光学システムの光路に挿入することによって行われ得る。いくつかの態様において、当該補正は、レンズ又は光学アセンブリなどの補正光学システムを光学システムの光路に挿入することなく行われ得る。いくつかの態様において、蛍光イメージングシステムは、対物レンズに隣接して、対物レンズとチューブレンズとの間に配置された電気光学位相プレートに更に備え得、電気光学位相プレートは、フローセルの上部内側表面と下部内側表面との間のギャップを満たす流体によって引き起こされる光学収

40

50

差の補正を提供し得る。いくつかの態様において、少なくとも1つのチューブレンズは、3つ以上の光学部品を含む複合レンズであり得る。いくつかの態様において、少なくとも1つのチューブレンズは、第1の非対称凸凸レンズ、第2の凸平面レンズ、第3の非対称凹凹レンズ、及び第4の非対称凸凹レンズのうちの1つ以上を含み得、上記の順序で、又は任意の代替の順序で存在し得る、4つの光学部品を含む複合レンズである。いくつかの態様において、少なくとも1つのチューブレンズは、少なくとも1mmの壁厚を有するフローセルの内側表面をイメージングするときに、対物レンズ、少なくとも1つのチューブレンズ、及び少なくとも1つのイメージング素子の組み合わせについてのイメージング性能メトリックを補正するように構成される。いくつかの態様において、少なくとも1つのチューブレンズは、少なくとも100 μm のギャップを有するフローセルの内側表面をイメージングするときに、対物レンズ、少なくとも1つのチューブレンズ、及び少なくとも1つのイメージング素子の組み合わせのイメージング性能メトリックを補正するように構成される。いくつかの態様において、少なくとも1つのチューブレンズは、少なくとも200 μm のギャップを有するフローセルの内側表面をイメージングするときに、対物レンズ、少なくとも1つのチューブレンズ、及び少なくとも1つのイメージング素子の組み合わせについてのイメージング性能メトリックを補正するように構成される。いくつかの態様において、システムは、単一の対物レンズ、2つのチューブレンズ、及び2つのイメージセンサを含み、2つのチューブレンズの各々は、異なる蛍光波長において最適なイメージング性能を提供するように設計される。いくつかの態様において、システムは、単一の対物レンズ、3つのチューブレンズ、及び3つのイメージセンサを含み、3つのチューブレンズの各々は、異なる蛍光波長において最適なイメージング性能を提供するように設計される。いくつかの態様において、システムは、単一の対物レンズ、4つのチューブレンズ、及び4つのイメージセンサを含み、4つのチューブレンズの各々は、異なる蛍光波長において最適なイメージング性能を提供するように設計される。いくつかの態様において、対物レンズ又は少なくとも1つのチューブレンズの設計は、中～高空間周波数範囲における変調伝達関数を最適化するように構成される。いくつかの態様において、イメージング性能メトリックは、1つ以上の指定された空間周波数、脱焦点、球面収差、色収差、コマ、乱視、視野曲率、画像歪み、コントラスト対ノイズ比(CNR)、又はそれらのいずれかの組み合わせにおける変調伝達関数(MTF)の測定を含む。いくつかの態様において、フローセルの上部内側表面及び下部内側表面をイメージングするためのイメージング性能メトリックの差は、10%未満である。いくつかの態様において、フローセルの上部内側表面及び下部内側表面をイメージングするためのイメージング性能メトリックの差は、5%未満である。いくつかの態様において、少なくとも1つのチューブレンズを使用することにより、対物レンズ、動き作動型コンペンセータ、及びイメージセンサを備える従来のシステムについてのものと比較して、デュアルサイドイメージングについてのイメージング性能メトリックを少なくとも同等又はより良好に改善する。いくつかの態様において、少なくとも1つのチューブレンズを使用することにより、対物レンズ、動き作動型コンペンセータ、及びイメージセンサを備える従来のシステムについてのものと比較して、デュアルサイドイメージングについてのイメージング性能メトリックを少なくとも10%改善する。

10

20

30

40

【0395】

本明細書に開示されるのは、イメージングベースの固相ジェノタイピング及びシーケンシング用途で使用するための照明システムであり、照明システムは、a)光源と、b)光源によって放出される光を収集し、それをテザーされた生体巨大分子を含む支持体表面上の指定された照明野に送達するように構成された液体ライトガイドと、を備える。

【0396】

いくつかの態様において、照明システムは、集光レンズを更に備える。いくつかの態様において、指定された照明野は、少なくとも2mm²の面積を有する。いくつかの態様において、指定された照明野に送達される光は、支持体表面の画像を取得するために使用されるイメージングシステムに対して指定された視野にわたって均一な強度のものである。

50

いくつかの態様において、指定された視野は、少なくとも 2 mm^2 の面積を有する。いくつかの態様において、指定された照明野に送達される光は、光強度の変動係数（CV）が 10% 未満であるとき、指定された視野にわたって均一な強度のものである。いくつかの態様において、指定された照明野に送達される光は、光強度の変動係数（CV）が 5% 未満であるとき、指定された視野にわたって均一な強度のものである。いくつかの態様において、指定された照明野に送達される光は、0.1 未満のスペックルコントラスト値を有する。いくつかの態様において、指定された照明野に送達される光は、0.05 未満のスペックルコントラスト値を有する。

【0397】

イメージングモジュール及びシステム：開示される光学システム、イメージングシステム、又はモジュールは、いくつかの事例では、試料又は基材表面をイメージングするように設計されたスタンドアロン光学システムであり得ることが当業者によって理解されるであろう。いくつかの事例では、それらは、1つ以上のプロセッサ又はコンピュータを備え得る。いくつかの事例では、それらは、器具制御機能及び／又は画像処理機能を提供する1つ以上のソフトウェアパッケージを備え得る。いくつかの事例では、光源（例えば、固体レーザ、染料レーザ、ダイオードレーザ、アークランプ、タングステンハロゲンランプなど）、レンズ、プリズム、ミラー、ダイクロイックリフレクタ、光学フィルタ、光バンドパスフィルタ、開口、及び画像センサ（例えば、相補的金属酸化物半導体（CMOS）画像センサ及びカメラ、電荷結合デバイス（CCD）画像センサ及びカメラなど）などの光学部品に加えて、それらはまた、X-Y 並進ステージ、X-Y-Z 並進ステージ、圧電気集束機構などの機械的及び／又は光機械的構成要素を備え得る。いくつかの事例では、それらは、ゲノミクス用途（例えば、遺伝子検査及び／又は核酸シーケンシング用途）のために設計されたより大きなシステムのモジュール、構成要素、サブアセンブリ、又はサブシステムとして機能し得る。例えば、いくつかの事例では、それらは、遮光及び／又は他の環境制御ハウジング、温度制御モジュール、流体制御モジュール、流体分配ロボット工学、ピックアンドプレイスロボット工学、1つ以上のプロセッサ若しくはコンピュータ、1つ以上のローカル及び／若しくはクラウドベースのソフトウェアパッケージ（例えば、機器／システム制御ソフトウェアパッケージ、画像処理ソフトウェアパッケージ、データ分析ソフトウェアパッケージ）、データ格納モジュール、データ通信モジュール（例えば、Bluetooth、WiFi、イントラネット、若しくはインターネット通信ハードウェア及び関連するソフトウェア）、ディスプレイモジュール、又はそれらの任意の組み合わせを更に備える、より大きなシステムのモジュール、構成要素、サブアセンブリ、又はサブシステムとして機能し得る。

【0398】

シーケンシングのための方法

本開示のいくつかの態様は、固定化されたテンプレート分子又は固定化されていないテンプレート分子をシーケンシングするための方法を提供する。方法は、システム 100 において、例えば、シーケンサ 114 において操作することができる。いくつかの態様において、固定化されたテンプレート分子は、目的の標的配列の1つのコピーを有する複数の核酸テンプレート分子を含む。いくつかの態様において、目的の標的配列の1つのコピーを有する核酸テンプレート分子は、線形ライブラリ分子を使用してブリッジ増幅を行うことによって生成することができる。いくつかの態様において、固定化されたテンプレート分子は、目的の標的配列の2つ以上のタンデムコピーを各々が有する複数の核酸テンプレート分子（例えば、コンカテマー）を含む。いくつかの態様において、コンカテマー分子を含む核酸テンプレート分子は、環状化直鎖ライブラリ分子のローリングサークル増幅を行うことによって生成され得る。いくつかの態様において、固定化されていないテンプレート分子は、環状分子を含む。いくつかの態様において、シーケンシングのための方法は、可溶性（例えば、固定化されていない）シーケンシングポリメラーゼ又は支持体に固定化されたシーケンシングポリメラーゼを用いる。

【0399】

いくつかの態様において、シーケンシング反応は、検出可能に標識されたヌクレオチド類似体を用いる。いくつかの態様において、シーケンシング反応は、検出可能に標識された多価分子に結合することと、ヌクレオチド類似体を組み込むことと、を含む2段階シーケンシング反応を用いる。いくつかの態様において、シーケンシング反応は、非標識ヌクレオチド類似体を用いる。いくつかの態様において、シーケンシング反応は、リン酸塩鎖標識ヌクレオチドを用いる。

【0400】

いくつかの態様において、固定化コンカテマーは、各々、目的の配列（例えば、挿入領域）及び任意のアダプター配列のタンデムリピート単位を含む。例えば、タンデムリピート単位は、（i）第1の表面プライマー（720）に対する結合配列を有する左のユニバーサルアダプター配列（例えば、表面ピニングプライマー）、（ii）第1のシーケンシングプライマー（740）（例えば、フォワードシーケンシングプライマー）に対する結合配列を有する左のユニバーサルアダプター配列、（iii）目的の配列（710）、（iv）第2のシーケンシングプライマー（750）（例えば、リバースシーケンシングプライマー）に対する結合配列を有する右のユニバーサルアダプター配列、（v）第2の表面プライマー（730）（例えば、表面キャプチャプライマー）に対する結合配列を有する右のユニバーサルアダプター配列、及び（vi）左の試料インデックス配列（760）及び／又は右の試料インデックス配列（770）を含む。いくつかの態様において、タンデムリピート単位は、左の固有の識別配列（780）及び／又は右の固有の識別配列（790）を更に含む。いくつかの態様において、タンデムリピート単位は、コンパクションオリゴヌクレオチドに対する少なくとも1つの結合配列を更に含む。いくつかの態様において、図6及び7は、線形ライブラリ分子又はコンカテマー分子の単位を示す。

【0401】

固定化されたコンカテマーは、コンパクトな核酸ナノボールに自己崩壊することができる。RCA反応中に1つ以上のコンパクションオリゴヌクレオチドを含めることは、ナノボールのサイズ及び／又は形状を更にコンパクトにすることができる。所与のコンカテマーにおけるタンデムリピート単位の数における増加は、ポリメラーゼ触媒シーケンシング反応のための複数の開始部位として機能する複数のシーケンシングプライマー（例えば、ユニバーサル配列を有するシーケンシングプライマー）にハイブリダイズするためのコンカテマーに沿う部位の数を増加させる。シーケンシング反応が、検出可能に標識されたヌクレオチド及び／又は（例えば、ヌクレオチド単位を有する）検出可能に標識された多価分子を用いるときに、コンカテマーに沿う並列シーケンシング反応に関与するヌクレオチド又はヌクレオチド単位によって放出されるシグナルは、各々のコンカテマーについて増加したシグナル強度をもたらす。所与のコンカテマーの複数の部分が、同時にシーケンシングされ得る。更に、複数の結合複合体は、特定のコンカテマー分子に沿って形成することができ、各々の結合複合体は、テンプレート／プライマー二重鎖に結合した及び多価分子に結合したシーケンシングポリメラーゼを含み、複数の結合複合体は、解離することなく依然として安定しており、増加した持続時間をもたらす、増加した持続時間は、シグナル強度を増加させ、イメージング時間を低減する。

【0402】

ヌクレオチド類似体を使用したシーケンシングのための方法

本開示のいくつかの態様は、本明細書に記載される固定化されたテンプレート分子のうちのいずれかをシーケンシングするための方法であって、（a）シーケンシングポリメラーゼを、（i）核酸テンプレート分子及び（ii）核酸シーケンシングプライマーに接触させるステップを含み、接触させることが、シーケンシングポリメラーゼを、核酸プライマーにハイブリダイズする核酸テンプレート分子に結合するのに好適な条件下で行われ、核酸プライマーにハイブリダイズする核酸テンプレート分子が核酸二本鎖を形成する、方法を提供する。いくつかの態様において、シーケンシングポリメラーゼは、ヌクレオチド類似体を結合し、組み込むことができる組換え変異体シーケンシングポリメラーゼを含む。

【0403】

いくつかの態様において、テンプレート分子をシーケンシングするための方法では、シーケンシングプライマーは、3'伸長可能な末端又は3'伸長不可能な末端を含む。いくつかの態様において、複数の核酸テンプレート分子は、増幅されたテンプレート分子（例えば、クローン的に増幅されたテンプレート分子）を含む。いくつかの態様において、複数の核酸テンプレート分子は、目的の標的配列の1つのコピーを含む。いくつかの態様において、複数の核酸分子は、目的の標的配列（例えば、コンカテマー）の2つ以上のタンデムコピーを含む。いくつかの態様において、複数の核酸テンプレート分子は、同一の目的の標的配列又は異なる目的の標的配列を含む。いくつかの態様において、複数の核酸プライマーは、溶液中にあるか、又は支持体に固定化される。いくつかの態様において、複数の核酸テンプレート分子及び／又は複数の核酸プライマーが支持体に固定化される場合、第1のシーケンシングポリメラーゼとの結合は、複数の固定化された第1の複合型ポリメラーゼを生成する。いくつかの態様において、複数の核酸テンプレート分子及び／又は核酸プライマーは、支持体上の102～1015個の異なる部位に固定化される。いくつかの態様において、複数のテンプレート分子及び核酸プライマーと複数の第1のシーケンシングポリメラーゼとの結合は、支持体上の102～1015個の異なる部位に固定化される複数の第1の複合型ポリメラーゼを生成する。いくつかの態様において、支持体上の複数の固定化された第1の複合型ポリメラーゼは、事前に判定された部位又は支持体上のランダムな部位に固定化される。いくつかの態様において、複数の固定化された第1の複合型ポリメラーゼは、互いに流体連通して、試薬（例えば、シーケンシングポリメラーゼ、多価分子、ヌクレオチド、及び／又は二価カチオンを含む酵素）の溶液を支持体上に流すことを可能にし、その結果、支持体上の複数の固定化された複合型ポリメラーゼを超並列方法で試薬の溶液と反応させる。

10

20

【0404】

いくつかの態様において、シーケンシングのための方法は、ステップ（b）：シーケンシングポリメラーゼを、少なくとも1つのヌクレオチドを、核酸二重鎖に結合するシーケンシングポリメラーゼに結合するのに好適であり、1つのヌクレオチドによってシーケンシングプライマーを伸長するポリメラーゼ触媒ヌクレオチド組み込みに好適である条件下で、複数のヌクレオチドと接触させることを更に含む。いくつかの態様において、シーケンシングポリメラーゼは、マグネシウム及び／又はマンガンを含む少なくとも1つの触媒カチオンの存在下で、複数のヌクレオチドと接触させる。いくつかの態様において、複数のヌクレオチドは、糖2'又は3'位に鎖終結部分を有する少なくとも1つのヌクレオチド類似体を含む。いくつかの態様において、鎖終結部分は、鎖終結部分をOH又はH基に変換するために、糖2'又は3'位から除去可能である。いくつかの態様において、複数のヌクレオチドは、鎖終結部分を欠く少なくとも1つのヌクレオチドを含む。いくつかの態様において、少なくともヌクレオチド上には、検出可能なシグナルを放出する検出可能なレポーター部分（例えば、フルオロフォア）で標識されている。検出可能なレポーター部分は、フルオロフォアを含む。いくつかの態様において、フルオロフォアは、抜くレオベースに付着している。いくつかの態様において、フルオロフォアは、ヌクレオ塩基にリンカーで付着しており、リンカーは、塩基から切断可能／除去可能である。いくつかの態様において、複数のヌクレオチドにおけるヌクレオチドのうちの少なくとも1つは、検出可能なレポーター部分で標識されていない。いくつかの態様において、ヌクレオチドに付着した特定の検出可能なレポーター部分（例えば、フルオロフォア）は、ヌクレオ塩基の検出及び識別を可能にするように、ヌクレオチド塩基（例えば、dATP、dGTP、dCTP、dTTP、又はdUTP）に対応することができる。組み込まれた鎖終結ヌクレオチドが検出可能に標識されるとき、ステップ（b）は、組み込まれた鎖終結ヌクレオチドから放出されたシグナルを検出することを更に含む。いくつかの態様において、ステップ（b）は、組み込まれた鎖終結ヌクレオチドのヌクレオ塩基を識別することを更に含む。

30

40

【0405】

いくつかの態様において、シーケンシングのための方法は、ステップ（c）：組み込ま

50

れた鎖終結ヌクレオチドから鎖終結部分を除去して、伸長可能な3'OH基を生成することを更に含む。いくつかの態様において、ステップ(c)は、組み込まれた鎖終結ヌクレオチドから検出可能な標識を除去することを更に含む。いくつかの態様において、シーケンシングポリメラーゼは、1つのヌクレオ塩基によって伸長されるシーケンシングプライマーにハイブリダイズされるテンプレート分子に結合したままである。

【0406】

いくつかの態様において、シーケンシングのための方法は、ステップ(d)：ステップ(b)及び(c)を少なくとも1回繰り返すことを更に含む。

【0407】

核酸シーケンシングのための2段階方法

本開示のいくつかの態様は、本明細書に記載される固定化されたテンプレート分子のうちのいずれかをシーケンシングするための2段階方法を提供する。いくつかの態様において、第1の段階は、概して、多価分子を複合型ポリメラーゼに結合させて、多価複合型ポリメラーゼを形成し、多価複合型ポリメラーゼを検出することを含む。

【0408】

いくつかの態様において、第1の段階は、(a)複数の第1のシーケンシングポリメラーゼを、(i)複数の核酸テンプレート分子及び(ii)複数の核酸シーケンシングプライマーに接触させることを含み、接触させることは、複数の第1のシーケンシングポリメラーゼを複数の核酸テンプレート分子及び複数の核酸プライマーに結合するのに好適な条件下で行われ、それによって複数の第1の複合型ポリメラーゼであって、各々が、核酸二重鎖に結合する第1のシーケンシングポリメラーゼを含む複数の第1の複合型ポリメラーゼを形成し、核酸二重鎖は、核酸プライマーにハイブリダイズする核酸テンプレート分子を含む。いくつかの態様において、第1のポリメラーゼは、組換え変異体シーケンシングポリメラーゼを含む。

【0409】

いくつかの態様において、テンプレート分子をシーケンシングするための方法では、シーケンシングプライマーは、3'伸長可能な末端又は3'伸長不可能な末端を有するオリゴヌクレオチドを含む。いくつかの態様において、複数の核酸テンプレート分子は、増幅されたテンプレート分子(例えば、クローン的に増幅されたテンプレート分子)を含む。いくつかの態様において、複数の核酸テンプレート分子は、目的の標的配列の1つのコピーを含む。いくつかの態様において、複数の核酸分子は、目的の標的配列(例えば、コンカテマー)の2つ以上のタンデムコピーを含む。いくつかの態様において、複数の核酸テンプレート分子中の核酸テンプレート分子は、同一の目的の標的配列又は異なる目的の標的配列を含む。いくつかの態様において、複数の核酸テンプレート分子及び/又は複数の核酸プライマーは、溶液中にあるか、又は支持体に固定化される。いくつかの態様において、複数の核酸テンプレート分子及び/又は複数の核酸プライマーが支持体に固定化される場合、第1のシーケンシングポリメラーゼとの結合は、複数の固定化された第1の複合型ポリメラーゼを生成する。いくつかの態様において、複数の核酸テンプレート分子及び/又は核酸プライマーは、支持体上の102~1015個の異なる部位に固定化される。いくつかの態様において、複数のテンプレート分子及び核酸プライマーと複数の第1のシーケンシングポリメラーゼとの結合は、支持体上の102~1015個の異なる部位に固定化される複数の第1の複合型ポリメラーゼを生成する。いくつかの態様において、支持体上の複数の固定化された第1の複合型ポリメラーゼは、事前に判定された部位又は支持体上のランダムな部位に固定化される。いくつかの態様において、複数の固定化された第1の複合型ポリメラーゼは、互いに流体連通して、試薬(例えば、シーケンシングポリメラーゼ、多価分子、ヌクレオチド、及び/又は二価カチオンを含む酵素)の溶液を支持体上に流すことを可能にし、その結果、支持体上の複数の固定化された複合型ポリメラーゼを超並列方法で試薬の溶液と反応させる。

【0410】

いくつかの態様において、シーケンシングするための方法は、ステップ(b)：複数の

10

20

30

40

50

第1の複合型ポリメラーゼを、複数の多価分子と接触させて、複数の多価複合型ポリメラーゼ（例えば、結合複合体）を形成することを更に含む。いくつかの態様において、複数の多価分子における個々の多価分子は、複数のヌクレオチドアームに付着したコアを含み、ヌクレオチドアームの各々は、ヌクレオチド（例えば、ヌクレオチド単位）に付着している（図9～13を参照されたい）。いくつかの態様において、ステップ（b）の接触させることは、多価分子の相補的ヌクレオチド単位を複数の第1の複合型ポリメラーゼのうちの少なくとも2つに結合させるのに好適である条件下で行われ、それによって、複数の多価複合型ポリメラーゼを形成する。いくつかの態様において、条件は、複数の多価複合型ポリメラーゼのプライマーへの相補的ヌクレオチド単位のポリメラーゼ触媒組み込みを阻害するのに好適である。いくつかの態様において、複数の多価分子は、複数のヌクレオチドアームを有する少なくとも1つの多価分子であって（例えば、図9～12）、各々が、ヌクレオチド類似体（例えば、ヌクレオチド類似体単位）で結合する少なくとも1つの多価分子を含み、ヌクレオチド類似体は、糖2'及び／又は3'位に鎖終結部分を含む。いくつかの態様において、複数の多価分子は、複数のヌクレオチドアームを含む少なくとも1つの多価分子を含み、複数のヌクレオチドアームは各々、鎖終結部分を欠くヌクレオチド単位と付着している。いくつかの態様において、複数の多価分子における多価分子のうちの少なくとも1つは、シグナルを放出する検出可能なレポーター部分で標識されている。いくつかの態様において、検出可能なレポーター部分は、フルオロフォアを含む。いくつかの態様において、ステップ（b）の接触させることは、ストロンチウム、バリウム、及び／又はカルシウムを含む少なくとも1つの非触媒カチオンの存在下で行われる。

10

20

【0411】

いくつかの態様において、シーケンシングのための方法は、ステップ（c）：複数の多価複合型ポリメラーゼを検出することを更に含む。いくつかの態様において、検出することは、複合型ポリメラーゼに結合する多価分子によって放出されるシグナルを検出することを含み、多価分子の相補的ヌクレオチド単位は、プライマーに結合するが、相補的ヌクレオチド単位の組み込みは、阻害される。いくつかの態様において、多価分子は、検出を可能にするように、検出可能なレポーター部分で標識されている。いくつかの態様において、標識された多価分子は、多価分子のコア、リンカー、及び／又はヌクレオチド単位に付着したフルオロフォアを含む。

【0412】

いくつかの態様において、シーケンシングのための方法は、ステップ（d）：複数の第1の複合型ポリメラーゼに結合した相補的ヌクレオチド単位の抜くレオ塩基を識別し、それによって、核酸テンプレートの配列を判定することを更に含む。いくつかの態様において、多価分子は、複数の第1の複合型ポリメラーゼに結合した相補的ヌクレオチド単位（例えば、ヌクレオチド塩基であるアデニン、グアニン、シトシン、チミン、又はウラシル）の特定を可能にするように、ヌクレオチドアームに付着した特定のヌクレオチド単位に対応する検出可能なレポーター部分で標識されている。

30

【0413】

いくつかの態様において、シーケンシングのための方法は、ステップ（e）：複数の多価複合型ポリメラーゼを解離させ、複数の第1のシーケンシングポリメラーゼ及びそれらの結合した多価分子を除去し、複数の核酸二重鎖を保持することを更に含む。

40

【0414】

いくつかの態様において、2段階シーケンシング方法の第2の段階は、概して、ヌクレオチド組み込みを含む。いくつかの態様において、シーケンシングのための方法は、ステップ（f）：ステップ（e）の複数の保持された核酸二重鎖を、複数の第2のシーケンシングポリメラーゼと接触させることであって、接触させることが、複数の第2のシーケンシングポリメラーゼを複数の保持された核酸二重鎖に結合させるのに好適である条件下で行われ、それによって、複数の第2の複合型ポリメラーゼを形成し、複数の第2の複合型ポリメラーゼが各々、核酸二重鎖に結合した第2のシーケンシングポリメラーゼを含む、接触させることを更に含む。いくつかの態様において、第2のシーケンシングポリメラー

50

ぜは、組換え変異体ポリメラーゼを含む。

【0415】

いくつかの態様において、ステップ（a）の複数の第1のシーケンシングポリメラーゼは、ステップ（f）の複数の第2のシーケンシングポリメラーゼとしてアミノ酸配列と100%同一であるアミノ酸配列を有する。いくつかの態様において、ステップ（a）の複数の第1のシーケンシングポリメラーゼは、ステップ（f）の複数の第2のシーケンシングポリメラーゼのアミノ酸配列と異なるアミノ酸配列を有する。

【0416】

いくつかの態様において、シーケンシングのための方法は、ステップ（g）：複数の第2の複合型ポリメラーゼを、複数のヌクレオチドと接触させることであって、接触させることが、複数のヌクレオチドからの相補的ヌクレオチドを、第2の複合型ポリメラーゼのうちの少なくとも2つに結合させるのに好適である条件下で行われ、それによって、複数のヌクレオチド複合型ポリメラーゼを形成する、接触させることを更に含む。いくつかの態様において、ステップ（g）の接触させることは、結合した相補的ヌクレオチドのヌクレオチド複合型ポリメラーゼのプライマーへのポリメラーゼ触媒組み込みを促進するのに好適な条件下で行われ、それによってシーケンシングプライマーを1ヌクレオ塩基だけ伸長させる。いくつかの態様において、ステップ（g）においてシーケンシングプライマーの3'末端にヌクレオチドを組み込むことは、プライマー伸長反応を含む。いくつかの態様において、ステップ（g）の接触させることは、マグネシウム及び／又はマンガンを含む少なくとも1つの触媒カチオンの存在下で行われる。いくつかの態様において、複数のヌクレオチドは、天然ヌクレオチド（例えば、非類似体ヌクレオチド）又はヌクレオチド類似体を含む。いくつかの態様において、複数のヌクレオチドは、除去可能であるか又は除去可能でない2'及び／又は3'鎖終結部分を含む。いくつかの態様において、複数のヌクレオチドにおけるヌクレオチドのうちの少なくとも1つは、検出可能なレポーター部分で標識されていない。いくつかの態様において、複数のヌクレオチドは、非標識である。いくつかの態様において、複数のヌクレオチドは、検出可能なレポーター部分で標識された複数のヌクレオチドを含む。検出可能なレポーター部分は、フルオロフォアを含む。いくつかの態様において、フルオロフォアは、ヌクレオチド塩基に付着している。いくつかの態様において、フルオロフォアは、ヌクレオチド塩基にリンカーで付着しており、リンカーは、塩基から切断可能／除去可能であるか、又は塩基から除去可能でない。いくつかの態様において、ヌクレオチドに付着した特定の検出可能なレポーター部分（例えば、フルオロフォア）は、ヌクレオチド塩基の検出及び識別を可能にするように、ヌクレオチド塩基（例えば、dATP、dGTP、dCTP、dTTP、又はdUTP）に対応することができる。

【0417】

いくつかの態様において、ステップ（g）における複数のヌクレオチドが検出可能に標識されるとき、シーケンシングのための方法は、ステップ（h）：ヌクレオチド複合型ポリメラーゼのプライマーに組み込まれる相補的ヌクレオチドを検出することを更に含む。いくつかの態様において、複数のヌクレオチドは、検出を可能にするように、検出可能なレポーター部分で標識されている。いくつかの態様において、ステップ（g）における複数のヌクレオチドが非標識である場合、ステップ（h）の検出することは省略される。

【0418】

いくつかの態様において、ステップ（g）における複数のヌクレオチドが検出可能に標識されるとき、シーケンシングのための方法は、ステップ（i）：ヌクレオチド複合型ポリメラーゼのプライマーに組み込まれる相補的ヌクレオチドの塩基を識別することを更に含む。いくつかの態様において、ステップ（i）における組み込まれた相補的ヌクレオチドの識別は、ステップ（d）における複数の第1の複合型ポリメラーゼに結合する多価分子の相補的ヌクレオチドの同一性を確認するために使用することができる。いくつかの態様において、ステップ（i）の特定することは、核酸テンプレート分子の配列を判定するために使用され得る。いくつかの態様において、ステップ（g）における複数のヌクレオ

10

20

30

40

50

チドが非標識である場合、ステップ（i）の識別することは省略される。

【0419】

いくつかの態様において、シーケンシングのための方法は、ステップ（j）：ステップ（g）が、複数の第2の複合型ポリメラーゼを、2'及び／又は3'鎖終結部分を有する少なくとも1つのヌクレオチドを含む複数のヌクレオチドと接触させることによって行われたときに、鎖終結部分を、組み込まれたヌクレオチドから除去することを更に含む。

【0420】

いくつかの態様において、シーケンシングのための方法は、ステップ（k）：ステップ（a）～（j）を少なくとも1回繰り返すことを更に含む。いくつかの態様において、核酸テンプレート分子の配列は、ステップ（c）及び（d）において、シーケンシングポリメラーゼに結合するが、プライマーの3'末端内に組み込まれない多価分子を検出及び識別することによって判定され得る。いくつかの態様において、核酸テンプレート分子の配列は、ステップ（h）及び（i）において、プライマーの3'末端内に組み込まれるヌクレオチドを検出及び識別することによって判定（又は確認）され得る。

【0421】

いくつかの態様において、核酸分子をシーケンシングための方法のうちのいずれかにおいて、複数の第1の複合型ポリメラーゼを複数の多価分子と結合させることは、少なくとも1つのアビディティ複合体を形成し、方法は、以下のステップ、（a）第1の核酸プライマー、第1のシーケンシングポリメラーゼ、及び第1の多価分子をコンカテマーテンプレート分子の第1の部分に結合し、それによって第1の結合複合体を形成することであって、第1の多価分子の第1のヌクレオチド単位が、第1のシーケンシングポリメラーゼに結合する、形成することと、（b）第2の核酸プライマー、第2のシーケンシングポリメラーゼ、及び第1の多価分子を同じコンカテマーテンプレート分子の第2の部分に結合し、それによって第2の結合複合体を形成することであって、第1の多価分子の第2のヌクレオチド単位は、第2のシーケンシングポリメラーゼに結合し、同じ多価分子を含む第1及び第2の結合複合体は、アビディティ複合体を形成する、形成することと、を含む。いくつかの態様において、第1のシーケンシングポリメラーゼは、本明細書に記載されるいずれかの野生型又は変異体ポリメラーゼを含む。いくつかの態様において、第2のシーケンシングポリメラーゼは、本明細書に記載されるいずれかの野生型又は変異体ポリメラーゼを含む。コンカテマーテンプレート分子は、目的の配列及び少なくとも1つのユニバーサルシーケンシングプライマー結合部位のタンデムリピート配列を含む。第1及び第2の核酸プライマーは、コンカテマーテンプレート分子に沿ってシーケンシングプライマー結合部位に結合することができる。例示的な多価分子は、図9～12に示される。

【0422】

いくつかの態様において、核酸分子をシーケンシングための方法のうちのいずれかにおいて、方法は、複数の第1の複合型ポリメラーゼを複数の多価分子と結合して、少なくとも1つのアビディティ複合体を形成することを含み、方法は、以下のステップ、（a）複数のシーケンシングポリメラーゼ及び複数の核酸プライマーを、コンカテマー核酸コンカテマー分子の異なる部分と接触させて、同じコンカテマー分子上に少なくとも第1及び第2の複合ポリメラーゼを形成するステップと、（b）複数の多価分子を、同じコンカテマーテンプレート分子上の少なくとも第1及び第2の複合ポリメラーゼに、複数の多価分子からの単一の多価分子を第1及び第2の複合ポリメラーゼに結合するのに好適な条件下で接触させるステップであって、単一の多価分子の少なくとも第1のヌクレオチド単位は、コンカテマーテンプレート分子の第1の部分にハイブリダイズし、それによって第1の結合複合体（例えば、第1の三元複合体）を形成する第1のプライマーを含む第1の複合ポリメラーゼに結合し、単一の多価分子の少なくとも第2のヌクレオチド単位は、コンカテマーテンプレート分子の第2の部分にハイブリダイズし、それによって第2の結合複合体（例えば、第2の三元複合体）を形成する第2のプライマーを含む第2の複合ポリメラーゼに結合して、接触させることが、第1及び第2の結合複合体中の結合した第1及び第2のヌクレオチド単位のポリメラーゼ触媒による取り込みを阻害するのに好適な条件下で行

われ、同じ多価分子に結合する第1及び第2の結合複合体がアビディティ複合体を形成する、接触させるステップと、(c) 同じコンカテマーテンプレート分子上で第1及び第2の結合複合体を検出するステップと、(d) 第1の結合複合体中の第1のヌクレオチド単位を識別し、それによってコンカテマーテンプレート分子の第1の部分の配列を決定し、第2の結合複合体中の第2のヌクレオチド単位を識別し、それによってコンカテマーテンプレート分子の第2の部分の配列を決定するステップと、を含む。いくつかの態様において、複数のシーケンシングポリメラーゼは、本明細書に記載されるいずれかの野生型又は変異体シーケンシングポリメラーゼを含む。コンカテマーテンプレート分子は、目的の配列及び少なくとも1つのユニバーサルシーケンシングプライマー結合部位のタンデムリピート配列を含む。複数の核酸プライマーは、コンカテマーテンプレート分子に沿ってシーケンシングプライマー結合部位に結合することができる。例示的な多価分子は、図10～13に示される。

10

【0423】

結合によるシーケンシング

本開示のいくつかの態様は、本明細書に記載される固定化されたテンプレート分子のうちのいずれかをシーケンシングするための方法であって、シーケンシング方法は、非標識鎖終結ヌクレオチドを用いる結合によるシーケンシング(SBB)手順を含む、方法を提供する。いくつかの態様において、シーケンシングによる結合(SBB)は、(a) プライミングされたテンプレート核酸を、三元複合体安定化条件下で少なくとも2つの別個の混合物と順次接触させるステップであって、少なくとも2つの別個の混合物は、各々、ポリメラーゼ及びヌクレオチドを含み、それによって、順次接触させるステップは、三元複合体安定化条件下で、プライミングされたテンプレート核酸が、テンプレートにおける第1、第2、及び第3の塩基の種類塩基の種類のヌクレオチド同族体を用いて接触されるプライミングされたテンプレート核酸をもたらす、順次接触させるステップと、(b) 少なくとも2つの別個の混合物を調べて、三元複合体が形成されたかどうかを判定するステップと、(c) プライミングされたテンプレート核酸分子の次の正しいヌクレオチドを識別するステップであって、次の正しいヌクレオチドは、三元複合体がステップ(b)で検出された場合、第1、第2、又は第3の塩基の種類の同族体として識別され、次の正しいヌクレオチドは、ステップ(b)での三元複合体の不在に基づいて、第4の塩基の種類のヌクレオチド同族体であると推定される、識別するステップと、(d) 次の正しいヌクレオチドを、ステップ(b)の後、プライミングされたテンプレート核酸を添加し、それによって伸長プライマーを産生するステップと、(e) 伸長プライマーを含むプライミングされたテンプレート核酸に対してステップ(a)～(d)を繰り返すステップと、を含む。例示的な結合によるシーケンシング方法は、米国特許第10,246,744号及び同第10,731,141号(両方の特許の内容全体が参照により本明細書に組み込まれる)に記載されている。

20

30

【0424】

リン酸塩鎖標識ヌクレオチドを使用したシーケンシングのための方法

本開示のいくつかの態様は、固定化されていないテンプレート分子を結合する固定化されたシーケンシングポリメラーゼを使用したシーケンシングのための方法であって、シーケンシング反応が、リン酸塩鎖標識ヌクレオチドで行われる、方法を提供する。いくつかの態様において、シーケンシング方法は、ステップ(a): 複数のシーケンシングポリメラーゼが固定化された支持体を提供することを含む。いくつかの態様において、シーケンシングポリメラーゼは、プロセッシブDNAポリメラーゼを含む。いくつかの態様において、シーケンシングポリメラーゼは、例えば、Phi29DNAポリメラーゼを含む、野生型又は変異体DNAポリメラーゼを含む。いくつかの態様において、支持体は、複数の別個のコンパートメントを含み、シーケンシングポリメラーゼは、コンパートメントの底部に固定化される。いくつかの態様において、別個のコンパートメントは、光が浸透することができるシリカ底部を含む。いくつかの態様において、別個のコンパートメントは、金属クラディングフィルム(例えば、アルミニウムクラディングフィルム)の穴を含

40

50

むナノフォトニック拘束構造で構成されたシリカ底部を含む。いくつかの態様において、金属クラディングの穴は、例えば、およそ70 nmの小さな開口部を有する。いくつかの態様において、ナノフォトニック拘束構造の高さは、およそ100 nmである。いくつかの態様において、ナノフォトニック拘束構造は、ゼロモード導波路（ZMW）を備える。いくつかの態様において、ナノフォトニック拘束構造は、液体を含む。

【0425】

いくつかの態様において、シーケンシング方法は、ステップ（b）：複数の固定化シーケンシングポリメラーゼを、複数の一本鎖環状核酸テンプレート分子及び複数のオリゴヌクレオチドシーケンシングプライマーと、個々の固定化シーケンシングポリメラーゼが一本鎖環状テンプレート分子に結合するのに好適であり、個々のシーケンシングプライマーが個々の一本鎖環状テンプレート分子にハイブリダイズするのに好適である条件下で接触させ、それによって複数のポリメラーゼ／テンプレート／プライマー複合体を生成することを更に含む。いくつかの態様において、個々のシーケンシングプライマーは、一本鎖環状テンプレート分子上のユニバーサルシーケンシングプライマー結合部位にハイブリダイズする。

【0426】

いくつかの態様において、シーケンシング方法は、ステップ（c）：複数のポリメラーゼ／テンプレート／プライマー複合体を、芳香族塩基、5炭素糖（例えば、リボース又はデオキシリボース）、及び3～20個のリン酸基を含むリン酸塩鎖を各々が含む複数のリン酸塩鎖標識ヌクレオチドと接触させることであって、末端リン酸基が、検出可能なレポーター部分（例えば、フルオロフォア）に結合している、接触させることを更に含む。第1、第2、及び第3のリン酸基は、アルファ、ベータ、及びガンマホスフェート基と称され得る。いくつかの態様において、末端リン酸基に付着した特定の検出可能なレポーター部分は、ヌクレオ塩基の検出及び識別を可能にするように、ヌクレオ塩基（例えば、dATP、dGTP、dCTP、dTTP、又はdUTP）に対応する。いくつかの態様において、複数のポリメラーゼ／テンプレート／プライマー複合体を、ポリメラーゼ触媒ヌクレオチド組み込みに好適な条件下で、複数のリン酸塩鎖標識ヌクレオチドと接触させる。いくつかの態様において、シーケンシングポリメラーゼは、相補的リン酸塩鎖標識ヌクレオチドを結合し、テンプレート分子におけるヌクレオチドの反対側の相補的ヌクレオチドを組み込むことが可能である。いくつかの態様において、ポリメラーゼ触媒ヌクレオチド組み込み反応は、アルファリン酸基とベータリン酸基との間で切断し、それによって、フルオロフォアに連結された多リン酸塩鎖を放出する。

【0427】

いくつかの態様において、シーケンシング方法は、ステップ（d）：シーケンシングポリメラーゼによって結合され、シーケンシングプライマーの末端に組み込まれる、リン酸塩鎖標識ヌクレオチドによって放出される蛍光シグナルを検出することを更に含む。いくつかの態様において、ステップ（d）は、シーケンシングポリメラーゼによって結合され、シーケンシングプライマーの末端に組み込まれるリン酸塩鎖標識ヌクレオチドを識別することを更に含む。

【0428】

いくつかの態様において、シーケンシング方法は、ステップ（d）：ステップ（c）～（d）を少なくとも1回繰り返すことを更に含む。いくつかの態様において、リン酸塩鎖標識ヌクレオチドを用いるシーケンシング方法は、米国特許第7,170,050号、第7,302,146号、及び／又は第7,405,281号に記載の方法に従って行うことができる。

【0429】

シーケンシングポリメラーゼ

本開示のいくつかの態様は、核酸分子をシーケンシングするための方法であって、本明細書に記載されるシーケンシング方法のうちのいずれかが、少なくとも1つの種類のシーケンシングポリメラーゼ及び複数のヌクレオチドを使用するか、又は少なくとも1つの種

類のシーケンシングポリメラーゼ及び複数のヌクレオチド及び複数の多価分子を用いる、方法を提供する。いくつかの態様において、シーケンシングポリメラーゼ（複数可）は、テンプレート分子におけるヌクレオチドの反対側の相補的ヌクレオチドを組み込むことが可能である。いくつかの態様において、シーケンシングポリメラーゼ（複数可）は、テンプレート分子におけるヌクレオチドの反対側の多価分子の相補的ヌクレオチド単位を結合することが可能である。いくつかの態様において、複数のシーケンシングポリメラーゼは、組換え変異体ポリメラーゼを含む。

【0430】

ヌクレオチド及び／又は多価分子でのシーケンシングにおける使用に好適であるポリメラーゼの例としては、クレノウDNAポリメラーゼ；*Thermus aquaticus* DNAポリメラーゼI（*Taq*ポリメラーゼ）；*KlenTaq*ポリメラーゼ；*Candidatus altiarcheales*古細菌；*Candidatus Hadarchaeum Yellowstoneense*；*Hadesarchaea*古細菌；*Euryarchaeota*古細菌；*Thermoplasma*古細菌；*Thermococcus*ポリメラーゼ、例えば、*Thermococcus litoralis*、バクテリオファージT7 DNAポリメラーゼ；ヒトアルファ、デルタ、及びエプシロンDNAポリメラーゼ；バクテリオファージポリメラーゼ、例えば、T4、RB69、及びphi29バクテリオファージDNAポリメラーゼ；*Pyrococcus furiosus* DNAポリメラーゼ（*Pfu*ポリメラーゼ）；*Bacillus subtilis* DNAポリメラーゼIII；*E. coli* DNAポリメラーゼIIIIアルファ及びエプシロン；9 degree Nポリメラーゼ；逆転写酵素、例えば、HIVタイプM若しくはO逆転写酵素；トリ骨髄芽球症ウイルス逆転写酵素；モロニー Maus 白血病ウイルス（MMLV）逆転写酵素；又はテロメラーゼが挙げられるが、これらに限定されない。DNAポリメラーゼの更なる非限定的な例としては、様々な古細菌族、例えば、*Aeropyrum*、*Archaeoglobus*、*Desulfurococcus*、*Pyrobaculum*、*Pyrococcus*、*Pyrolobus*、*Pyrodictum*、*Staphylothermus*、*Stetteria*、*Sulfolobus*、*Thermococcus*、及び*Vulcanisaeta*などからのもの、又はそれらのバリエーションが挙げられ、これらは、当該技術分野で既知のそのようなポリメラーゼ、例えば、9 degrees N、VENT、DEEP VENT、THERMINATOR、*Pfu*、KOD、*Pfx*、*Tgo*、及びRB69ポリメラーゼを含む。

【0431】

ヌクレオチド

本開示のいくつかの態様は、本明細書に記載されるシーケンシング方法のうちのいずれかが、少なくとも1つのヌクレオチドを用いる、核酸分子をシーケンシングするための方法を提供する。ヌクレオチドは、塩基、糖、及び少なくとも1つのリン酸基を含む。いくつかの態様において、複数のヌクレオチドにおける少なくとも1つのヌクレオチドは、芳香族塩基、5炭素糖（例えば、リボース又はデオキシリボース）、及び1つ以上のリン酸基（例えば、1つ～10個のリン酸基）を含む。複数のヌクレオチドは、dATP、dGTP、dCTP、dTTP、及びdUTPからなる群から選択される少なくとも1つの種類のヌクレオチドを含むことができる。複数のヌクレオチドは、dATP、dGTP、dCTP、dTTP、及び／又はdUTPからなる群から選択される2つ以上の種類のヌクレオチドのいずれかの組み合わせの混合物において含むことができる。いくつかの態様において、複数のヌクレオチドにおける少なくとも1つのヌクレオチドは、ヌクレオチド類似体でない。いくつかの態様において、複数のヌクレオチドにおける少なくとも1つのヌクレオチドは、ヌクレオチド類似体を含む。

【0432】

いくつかの態様において、本明細書に記載される核酸分子をシーケンシングするための方法のうちのいずれかにおいて、複数のヌクレオチドのうちの少なくとも1つのヌクレオチドは、1、2、又は3つのリン原子の鎖を含み、鎖は、典型的には、エステル又はホス

ホラミド結合を介して、糖部分の5'炭素に付着している。いくつかの態様において、複数のヌクレオチドにおける少なくとも1つのヌクレオチドは、リン鎖を有する類似体であり、リン鎖において、リン原子と一緒に、介在するO、S、NH、メチレン、又はエチレンと結合している。いくつかの態様において、鎖におけるリン原子は、O、S、又はBH₃を含む置換された側鎖基を含む。いくつかの態様において、鎖は、ホスホロアミデート、ホスホロチオエート、ホスホルジチオエート、及びO-メチルホスホロアミダイト基を含む類似体で置換されたリン酸基を含む。

【0433】

いくつかの態様において、本明細書に記載される核酸分子をシーケンシングするための方法のうちのいずれかにおいて、複数のヌクレオチドにおける少なくとも1つのヌクレオチドは、ターミネーターヌクレオチド類似体を含み、ターミネーターヌクレオチド類似体は、鎖終結部分（例えば、ブロッキング部分）を、糖2'位、糖3'位、又は糖2'及び3'位において有する。いくつかの態様において、鎖終結部分は、プライマー伸長反応中の新生の鎖における後続のヌクレオチド単位又は遊離ヌクレオチドのポリメラーゼ触媒組み込みを阻害することができる。いくつかの態様において、鎖終結部分は、3'糖ヒドロキシル位に付着しており、3'糖位において、糖は、リボース又はデオキシリボース糖部分を含む。いくつかの態様において、鎖終結部分は、ポリメラーゼ触媒ヌクレオチド組み込み反応において後続のヌクレオチドで伸長可能である3'OH糖基を有するヌクレオチドを生成するように、3'糖位から除去可能／切断可能である。いくつかの態様において、鎖終結部分は、アルキル基、アルケニル基、アルキニル基、アリル基、アリール基、ベンジル基、アジド基、アミン基、アミド基、ケト基、イソシアネート基、リン酸基、チオ基、ジスルフィド基、カーボネート基、尿素基、シリル基、又はアセタール基を含む。いくつかの態様において、鎖終結部分は、例えば、鎖終結部分を化学剤、pH変化、光、又は熱と反応させることによって、ヌクレオチドから切断可能／除去可能である。いくつかの態様において、鎖終結部分であるアルキル、アルケニル、アルキニル、及びアリルは、ピペリジンを有する又は2,3-ジクロロ-5,6-ジシアノ-1,4-ベンゾキノン（DDQ）を有するテトラキス（トリフェニルホスフィン）パラジウム（0）（Pd（PPh₃）₄）で切断可能である。いくつかの態様において、鎖終結部分であるアリール及びベンジルは、H₂ Pd/Cで切断可能である。いくつかの態様において、鎖終結部分であるアミン、アミド、ケト、イソシアネート、ホスフェート、チオ、ジスルフィドは、ホスフィン、又はベータメルカプトエタノール若しくはジチオトリート（DTT）を含むチオール基で切断可能である。いくつかの態様において、鎖終結部分であるカーボネートは、MeOH中の炭酸カリウム（K₂CO₃）、ピリジン中のトリエチルアミン、又は酢酸中のZn（AcOH）で切断可能である。いくつかの態様において、鎖終結部分である尿素及びシリルは、フッ化テトラブチルアンモニウム、ピリジン-HF、フッ化アンモニウム、又はトリエチルアミントリヒドロフルオリドで切断可能である。いくつかの態様において、鎖終結部分は、亜硝酸で切断可能／除去可能であり得る。いくつかの態様において、鎖終結部分は、亜硝酸塩を含む溶液、例えば、亜硝酸塩と、酸、例えば、酢酸、硫酸、又は硝酸との組み合わせを使用して切断可能／除去可能であり得る。いくつかの更なる態様において、当該溶液は、有機酸を含み得る。

【0434】

いくつかの態様において、本明細書に記載される核酸分子をシーケンシングするための方法のうちのいずれかにおいて、複数のヌクレオチドにおける少なくとも1つのヌクレオチドは、ターミネーターヌクレオチド類似体を含み、ターミネーターヌクレオチド類似体は、鎖終結部分（例えば、ブロッキング部分）を、糖2'位、糖3'位、又は糖2'及び3'位において有する。いくつかの態様において、鎖終結部分は、アジド（azide）、アジド（azido）、及びアジドメチル基を含む。いくつかの態様において、鎖終結部分は、3'-O-アジド又は3'-O-アジドメチル基を含む。いくつかの態様において、鎖終結部分であるアジド（azide）、アジド（azido）、及びアジドメチル基は、ホスフィン化合物で切断可能／除去可能である。いくつかの態様において、ホスフィン化

合物は、誘導体化トリアルキルホスフィン部分又は誘導体化トリアリールホスフィン部分を含む。いくつかの態様において、ホスフィン化合物は、トリス（２－カルボキシエチル）ホスフィン（ＴＣＥＰ）、又はビス－スルホトリフェニルホスフィン（ＢＳ－ＴＰＰ）、又はトリス（ヒドロキシプロピル）ホスフィン（ＴＨＰＰ）を含む。いくつかの態様において、切断剤は、４－ジメチルアミノピリジン（４－ＤＭＡＰ）を含む。いくつかの態様において、３’－Ｏ－アミノ基、３’－Ｏ－アミノメチル基、３’－Ｏ－メチルアミノ基、又はそれらの誘導体のうちの１つ以上を含む鎖終結部分は、亜硝酸を利用する機構を介して、又は亜硝酸を含む溶液を使用して、亜硝酸を用いて切断され得る。いくつかの態様において、３’－Ｏ－アミノ基、３’－Ｏ－アミノメチル基、３’－Ｏ－メチルアミノ基、又はそれらの誘導体のうちの１つ以上を含む鎖終結部分は、亜硝酸塩を含む溶液を使用して切断され得る。いくつかの態様において、例えば、亜硝酸塩は、酢酸、硫酸、又は硝酸などの酸と組み合わせられ得るか、又はそれらと接触させ得る。いくつかの更なる態様において、例えば、亜硝酸塩は、例えば、ギ酸、酢酸、プロピオン酸、酪酸、イソ酪酸などの有機酸と組み合わせられ得るか、又はそれらと接触させ得る。いくつかの態様において、鎖終結部分は、パラジウムデブロッキング試薬（例えば、Ｐｄ（０））で切断することができる３’－アセタール部分を含む。

10

【０４３５】

いくつかの態様において、本明細書に記載される核酸分子をシーケンシングするための方法のうちのいずれかにおいて、ヌクレオチド類似体は、３’－デオキシヌクレオチド、２’，３’－ジデオキシヌクレオチド、３’－メチル、３’－アジド、３’－アジドメチル、２’，３’－Ｏ－アジドアルキル、３’－Ｏ－エチニル、３’－Ｏ－アミノアルキル、３’－フルオロアルキル、３’－フルオロメチル、３’－ジフルオロメチル、３’－トリフルオロメチル、３’－スルホニル、３’－マロニル、３’－アミノ、３’－Ｏ－アミノ、３’－スルフヒドラル、３’－アミノメチル、３’－エチル、３’－ブチル、３’－tertブチル、３’－フルオレニルメチルオキシカルボニル、３’－tertブチルオキシカルボニル、３’－Ｏ－アルキルヒドロキシアルキル、３’－ホスホロチオート、３’－Ｏ－ベンジル、及び３’－Ｏ－ベンジル、３’－アセタール部分、又はそれらの誘導体からなる群から選択される鎖終結部分を含む。

【０４３６】

いくつかの態様において、本明細書に記載される核酸分子をシーケンシングするための方法のうちのいずれかにおいて、複数のヌクレオチドは、検出可能なレポーター部分で標識された複数のヌクレオチドを含む。検出可能なレポーター部分は、フルオロフォアを含む。いくつかの態様において、フルオロフォアは、ヌクレオチド塩基に付着している。いくつかの態様において、フルオロフォアは、ヌクレオチド塩基にリンカーで付着しており、リンカーは、塩基から切断可能／除去可能である。いくつかの態様において、複数のヌクレオチドにおけるヌクレオチドのうちの少なくとも１つは、検出可能なレポーター部分で標識されていない。いくつかの態様において、ヌクレオチドに付着した特定の検出可能なレポーター部分（例えば、フルオロフォア）は、ヌクレオチド塩基の検出及び識別を可能にするように、ヌクレオチド塩基（例えば、ｄＡＴＰ、ｄＧＴＰ、ｄＣＴＰ、ｄＴＴＰ、又はｄＵＴＰ）に対応することができる。

30

40

【０４３７】

いくつかの態様において、本明細書に記載される核酸分子をシーケンシングするための方法のうちのいずれかにおいて、ヌクレオチド塩基上の切断可能なリンカーは、アルキル基、アルケニル基、アルキニル基、アリル基、アリール基、ベンジル基、アジド基、アミン基、アミド基、ケト基、イソシアネート基、リン酸基、チオ基、ジスルフィド基、カーボネート基、尿素基、又はシリル基を含む切断可能な部分を含む。いくつかの態様において、塩基上の切断可能なリンカーは、切断可能な部分を化学剤、ｐＨ変化、光、又は熱と反応させることによって、塩基から切断可能／除去可能である。いくつかの態様において、切断可能な部分であるアルキル、アルケニル、アルキニル、及びアリルは、ピペリジンを有する又は２，３－ジクロロ－５，６－ジシアノ－１，４－ベンゾ－キノン（ＤＤＱ）

50

を有するテトラキス（トリフェニルホスフィン）パラジウム（0）（ $\text{Pd}(\text{PPh}_3)_4$ ）で切断可能である。いくつかの態様において、切断可能な部分であるアリール及びベンジルは、 $\text{H}_2 \text{ Pd} / \text{C}$ で切断可能である。いくつかの態様において、切断可能な部分であるアミン、アミド、ケト、イソシアネート、ホスフェート、チオ、ジスルフィドは、ホスフィン、又はベータメルカプトエタノール若しくはジチオトリートール（DTT）を含むチオール基で切断可能である。いくつかの態様において、切断可能な部分であるカーボネートは、 MeOH 中の炭酸カリウム（ K_2CO_3 ）、ピリジン中のトリエチルアミン、又は酢酸中の $\text{Zn}(\text{AcOH})$ で切断可能である。いくつかの態様において、切断可能な部分である尿素及びシリルは、フッ化テトラブチルアンモニウム、ピリジン- HF 、フッ化アンモニウム、又はトリエチルアミントリヒドロフルオリドで切断可能である。

10

【0438】

いくつかの態様において、本明細書に記載される核酸分子をシーケンシングするための方法のうちのいずれかにおいて、ヌクレオチド塩基上の切断可能なリンカーは、アジド（azide）、アジド（azido）、及びアジドメチル基を含む切断可能な部分を含む。いくつかの態様において、切断可能な部分であるアジド（azide）、アジド（azido）、及びアジドメチル基は、ホスフィン化合物で切断可能／除去可能である。いくつかの態様において、ホスフィン化合物は、誘導体化トリアルキルホスフィン部分又は誘導体化トリアリールホスフィン部分を含む。いくつかの態様において、ホスフィン化合物は、トリス（2-カルボキシエチル）ホスフィン（TCPE）、又はビス-スルホトリフェニルホスフィン（BS-TPP）、又はトリス（ヒドロキシプロピル）ホスフィン（THPP）を含む。いくつかの態様において、切断剤は、4-ジメチルアミノピリジン（4-DMAP）を含む。

20

【0439】

いくつかの態様において、本明細書に記載される核酸分子をシーケンシングのための方法のうちのいずれかにおいて、（例えば、糖2'及び／又は糖3'位における）鎖終結部分、及びヌクレオチド塩基上の切断可能なリンカーは、同じ又は異なる切断可能な部分を有する。いくつかの態様において、（例えば、糖2'及び／又は糖3'位における）鎖終結部分、及び塩基に結合した検出可能なレポーター部分は、同じ化学剤で化学的に切断可能／除去可能である。いくつかの態様において、（例えば、糖2'及び／又は糖3'位における）鎖終結部分、及び塩基に結合した検出可能なレポーター部位は、異なる化学剤で化学的に切断可能／除去可能である。

30

【0440】

多価分子

本開示のいくつかの態様は、本明細書に記載されるシーケンシング方法のうちのいずれかが、少なくとも1つの多価分子を用いる、核酸分子をシーケンシングするための方法を提供する。いくつかの態様において、多価分子は、コアに付着し、スターバースト、ヘルタースケルター、又はボトルブラシ構成を含む任意の構成を有する複数のヌクレオチドアームを含む（例えば、図9）。多価分子は、（1）コア及び（2）複数のヌクレオチドアームを含み、複数のヌクレオチドアームは、（i）コア付着部分、（ii）PEG部分を含むスペーサー、（iii）リンカー、及び（iv）ヌクレオチド単位を含み、コアは、複数のヌクレオチドアームに付着しており、スペーサーは、リンカーに付着しており、リンカーは、ヌクレオチド単位に付着している。いくつかの態様において、ヌクレオチド単位は、塩基、糖、及び少なくとも1つのリン酸基を含み、リンカーは、塩基を介して、ヌクレオチド単位に付着している。いくつかの態様において、リンカーは、脂肪族鎖又はオリゴエチレングリコール鎖を含み、両方のリンカー鎖は、2～6つのサブユニットを有する。いくつかの態様において、リンカーはまた、芳香族部分を含む。例示的なヌクレオチドアームを図13に示す。例示的な多価分子は、図9～12に示される。例示的なスペーサーを図14（上）に示し、例示的なリンカーを図15（下）及び図15に示す。リンカーに付着した例示的なヌクレオチドを図16～19に示す。例示的なビオチン化ヌクレオチドアームを図20に示す。

40

50

【0441】

いくつかの態様において、多価分子は、複数のヌクレオチドアームに結合するコアを含み、複数のヌクレオチドアームは、 $dATP$ 、 $dGTP$ 、 $dCTP$ 、 $dTTP$ 、及び $dUTP$ からなる群から選択される同じ種類のヌクレオチド単位を有する。

【0442】

いくつかの態様において、多価分子は、複数のヌクレオチドアームに付着するコアを含み、各アームは、ヌクレオチド単位を含む。いくつかの態様において、ヌクレオチド単位は、芳香族塩基、5炭素糖（例えば、リボース又はデオキシリボース）、及び1つ以上のリン酸基（例えば、1つ～10個のリン酸基）を含む。複数の多価分子は、 $dATP$ 、 $dGTP$ 、 $dCTP$ 、 $dTTP$ 、及び $dUTP$ からなる群から選択される1つの種類のヌクレオチド単位を有する1つの種類の多価分子を含むことができる。複数の多価分子は、2つ以上の種類の多価分子のいずれかの組み合わせの混合物において含むことができ、混合物における個々の多価分子は、 $dATP$ 、 $dGTP$ 、 $dCTP$ 、 $dTTP$ 、及び／又は $dUTP$ からなる群から選択されるヌクレオチド単位を含む。

10

【0443】

いくつかの態様において、ヌクレオチド単位は、1、2、又は3つのリン原子の鎖を含み、鎖は、典型的には、エステル又はホスホラミド結合を介して、糖部分の5'炭素に付着している。いくつかの態様において、少なくとも1つのヌクレオチド単位は、リン鎖を有するヌクレオチド類似体であり、リン鎖において、リン原子と一緒に、介在するO、S、NH、メチレン、又はエチレンと結合している。いくつかの態様において、鎖におけるリン原子は、O、S、又はBH3を含む置換された側鎖基を含む。いくつかの態様において、鎖は、ホスホロアミデート、ホスホロチオエート、ホスホルジチオエート、及びO-メチルホスホロアミダイト基を含む類似体で置換されたリン酸基を含む。

20

【0444】

いくつかの態様において、多価分子は、複数のヌクレオチドアームに付着したコアを含み、個々のヌクレオチドアームは、糖2'位、糖3'位、又は糖2'及び3'位に鎖終結部分（例えば、ブロッキング部分）を有するヌクレオチド類似体であるヌクレオチド単位を含む。いくつかの態様において、ヌクレオチド単位は、糖2'位、糖3'位、又は糖2'及び3'位に鎖終結部分（例えば、ブロッキング部分）を含む。いくつかの態様において、鎖終結部分は、プライマー伸長反応中の新生の鎖における後続のヌクレオチド単位又は遊離ヌクレオチドのポリメラーゼ触媒組み込みを阻害することができる。いくつかの態様において、鎖終結部分は、3'糖ヒドロキシル位に付着しており、3'糖位において、糖は、リボース又はデオキシリボース糖部分を含む。いくつかの態様において、鎖終結部分は、ポリメラーゼ触媒ヌクレオチド組み込み反応において後続のヌクレオチドで伸長可能である3'OH糖基を有するヌクレオチドを生成するように、3'糖位から除去可能／切断可能である。いくつかの態様において、鎖終結部分は、アルキル基、アルケニル基、アルキニル基、アリル基、アリール基、ベンジル基、アジド基、アミン基、アミド基、ケト基、イソシアネート基、リン酸基、チオ基、ジスルフィド基、カーボネート基、尿素基、又はシリル基を含む。いくつかの態様において、鎖終結部分は、例えば、鎖終結部分を化学剤、pH変化、光、又は熱と反応させることによって、ヌクレオチド単位から切断可能／除去可能である。いくつかの態様において、鎖終結部分であるアルキル、アルケニル、アルキニル、及びアリルは、ペペリジン有する又は2,3-ジクロロ-5,6-ジシアノ-1,4-ベンゾキノン(DDQ)を有するテトラキス(トリフェニルホスフィン)パラジウム(0)(Pd(PPh₃)₄)で切断可能である。いくつかの態様において、鎖終結部分であるアリール及びベンジルは、H₂Pd/Cで切断可能である。いくつかの態様において、鎖終結部分であるアミン、アミド、ケト、イソシアネート、ホスフェート、チオ、ジスルフィドは、ホスフィン、又はベータメルカプトエタノール若しくはジチオトリートール(DTT)を含むチオール基で切断可能である。いくつかの態様において、鎖終結部分であるカーボネートは、MeOH中の炭酸カリウム(K₂CO₃)、ピリジン中のトリエチルアミン、又は酢酸中のZn(AcOH)で切断可能である。いくつかの態様におい

30

40

50

て、鎖終結部分である尿素及びシリルは、フッ化テトラブチルアンモニウム、ピリジン-HF、フッ化アンモニウム、又はトリエチルアミントリヒドロフルオリドで切断可能である。

【0445】

いくつかの態様において、ヌクレオチド単位は、糖2'位、糖3'位、又は糖2'及び3'位に鎖終結部分（例えば、ブロック部分）を含む。いくつかの態様において、鎖終結部分は、アジド（azide）、アジド（azido）、及びアジドメチル基を含む。いくつかの態様において、鎖終結部分は、3'-O-アジド又は3'-O-アジドメチル基を含む。いくつかの態様において、鎖終結部分であるアジド（azide）、アジド（azido）、及びアジドメチル基は、ホスフィン化合物で切断可能／除去可能である。いくつかの態様において、ホスフィン化合物は、誘導体化トリアルキルホスフィン部分又は誘導体化トリアリールホスフィン部分を含む。いくつかの態様において、ホスフィン化合物は、トリス（2-カルボキシエチル）ホスフィン（TCEP）、又はビス-スルホトリフェニルホスフィン（BS-TPP）、又はトリス（ヒドロキシプロピル）ホスフィン（THPP）を含む。いくつかの態様において、切断剤は、4-ジメチルアミノピリジン（4-DMAP）を含む。

10

【0446】

いくつかの態様において、ヌクレオチド単位は、3'-デオキシヌクレオチド、2', 3'-ジデオキシヌクレオチド、3'-メチル、3'-アジド、3'-アジドメチル、3'-O-アジドアルキル、3'-O-エチニル、3'-O-アミノアルキル、3'-O-フルオロアルキル、3'-フルオロメチル、3'-ジフルオロメチル、3'-トリフルオロメチル、3'-スルホニル、3'-マロニル、3'-アミノ、3'-O-アミノ、3'-スルフヒドラル、3'-アミノメチル、3'-エチル、3'-ブチル、3'-tert-ブチル、3'-フルオレニルメチルオキシカルボニル、3'-tert-ブチルオキシカルボニル、3'-O-アルキルヒドロキシルアミノ基、3'-ホスホロチオート、及び3'-O-ベンジル、又はそれらの誘導体からなる群から選択される鎖終結部分を含む。

20

【0447】

いくつかの態様において、多価分子は、複数のヌクレオチドアームに付着するコアを含み、ヌクレオチドアームは、スパーサー、リンカー及びヌクレオチド単位を含み、コア、リンカー及び／又はヌクレオチド単位は、検出可能なレポーター部分で標識される。いくつかの態様において、検出可能なレポーター部分は、フルオロフォアを含む。いくつかの態様において、多価分子に付着した特定の検出可能なレポーター部分（例えば、フルオロフォア）は、ヌクレオチド塩基の検出及び識別を可能にするように、ヌクレオチド単位の塩基（例えば、dATP、dGTP、dCTP、dTTP、又はdUTP）に対応することができる。

30

【0448】

いくつかの態様において、多価分子の少なくとも1つのヌクレオチドアームは、検出可能なレポーター部分に付着したヌクレオチド単位を有する。いくつかの態様において、検出可能なレポーター部分は、ヌクレオチド塩基に付着している。いくつかの態様において、検出可能なレポーター部分は、フルオロフォアを含む。いくつかの態様において、多価分子に付着した特定の検出可能なレポーター部分（例えば、フルオロフォア）は、ヌクレオチド塩基の検出及び識別を可能にするように、ヌクレオチド単位の塩基（例えば、dATP、dGTP、dCTP、dTTP、又はdUTP）に対応することができる。

40

【0449】

いくつかの態様において、多価分子のコアは、アビジン様又はストレプトアビジン様部分を含み、コア付着部分は、ビオチンを含む。いくつかの態様において、コアは、アビジンタンパク質、並びに少なくとも1つのビオチン部分に結合することができる任意の誘導体、類似体、及び他の非天然形態のアビジンを含むストレプトアビジン型又はアビジン型部分を含む。アビジン部分の他の形態としては、天然及び組換えアビジン及びストレプトアビジン、並びに誘導体化分子、例えば、非グリコシル化アビジン及び切断ストレプトア

50

ビジンが挙げられる。例えば、アビジン部分は、アビジンの脱グリコシル化形態、Streptomyces（例えば、Streptomyces avidinii）によって産生される細菌ストレプトアビジン、並びに誘導体化形態、例えば、N-アシルアビジン、例えば、N-アセチル、N-フタリル、及びN-スクシニルアビジン、並びに市販の製品であるEXTRAVIDIN、CAPTAVIDIN、NEUTRAVIDIN、及びNEUTRALITE AVIDINを含む。

【0450】

いくつかの態様において、本明細書に記載される核酸分子をシーケンシングするための方法のうちのいずれかは、結合複合体を形成することを含んでもよく、結合複合体は、（i）ポリメラーゼ、プライマーと二重鎖を有する核酸テンプレート分子、及びヌクレオチドを含むか、又は結合複合体は、（ii）ポリメラーゼ、プライマーと二重鎖を有する核酸テンプレート分子、及び多価分子のヌクレオチド単位を含む。いくつかの態様において、結合複合体は、約0.1、0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9、又は1秒間超の持続時間を有する。結合複合体は、約0.1～0.25秒、若しくは約0.25～0.5秒、若しくは約0.5～0.75秒、若しくは約0.75～1秒、若しくは約1～2秒、若しくは約2～3秒、若しくは約3～4秒、若しくは約4～5秒を超える持続時間を有する、及び／又は本方法が15℃以上、20℃以上、25℃以上、35℃以上、37℃以上、42℃以上、55℃以上、60℃以上、若しくは72℃以上、若しくは80℃以上、又は前述のうちのいずれかによって定義される範囲内で実施されるか又は実施され得る。結合複合体（例えば、三元複合体）は、ポリメラーゼ、テンプレート分子、プライマー、及び／又はヌクレオチド単位若しくはヌクレオチドのうちのいずれかの間の相互作用の解離を引き起こす条件に供されるまで依然として安定している。例えば、解離条件は、結合複合体を、洗浄剤、EDTA、及び／若しくは水のうちのいずれか1つ、又はそれらのいずれかの組み合わせと接触させることを含む。いくつかの態様において、本開示は、結合複合体が、検出ステップにおいて20を超えるコントラストノイズ比を示す表面上に堆積されるか、それに付着するか、又はそれにハイブリダイズされる、当該方法を提供する。いくつかの態様において、本開示は、接触させることが、ヌクレオチド又はヌクレオチド単位がテンプレート核酸の次の塩基に相補的であるときに結合複合体を安定化し、ヌクレオチド又はヌクレオチド単位がテンプレート核酸の次の塩基に相補的でないときに結合複合体を不安定化する条件下で実行される、当該方法を提供する。

【0451】

コンパクションオリゴヌクレオチド

コンパクションオリゴヌクレオチドは、コンカテマー分子の第1の部分にハイブリダイズすることができる5'領域を有する一本鎖直鎖オリゴヌクレオチド、及びコンカテマー分子（例えば、同じコンカテマー分子）の第2の部分にハイブリダイズすることができる3'領域を有するコンパクションオリゴヌクレオチドを含む。いくつかの態様において、コンパクションオリゴヌクレオチドの個々のコンカテマー分子へのハイブリダイゼーションは、コンカテマー分子を、非崩壊DNA分子と比較して形状及びサイズがよりコンパクトなDNAナノボールに崩壊又は折り畳ませる。DNAナノボールのスポット画像は、ガウススポットとして表され得、サイズは、全幅半値（FWHM）として測定され得る。より小さいFWHMによって示されるより小さいスポットサイズは、典型的には、スポットの改善された画像と相関する。いくつかの態様において、DNAナノボールスポットのFWHMは、約10µm以下であることができる。DNAナノボールは、DNAナノボールに崩壊／折り畳まれていないコンカテマーと比較して小さい全幅半値（FWHM）を有するコンパクトな核酸構造であり得る。

【0452】

いくつかの態様において、コンパクションオリゴヌクレオチドは、DNA、RNA、又はDNA及びRNAの組み合わせを含む一本鎖オリゴヌクレオチドを含む。コンパクションオリゴヌクレオチドは、長さが20～150個のヌクレオチド、又は30～100個のヌクレオチド、又は40～80個のヌクレオチドを含むいずれかの長さであることができ

る。

【0453】

いくつかの態様において、コンパクションオリゴヌクレオチドは、5'領域及び3'領域、並びに任意選択で、5'領域と3'領域との間の中間領域を含む。介在領域は、いずれかの長さ、例えば、2～20個のヌクレオチドであることができる。介在領域は、連続した同一の塩基（例えば、AAA、GGG、CCC、TTT、又はUUU）を有するホモポリマーを含む。介在領域は、非ホモポリマー配列を含む。

【0454】

コンパクションオリゴヌクレオチドの5'領域は、その長さに沿って、コンカテマー分子の第1の部分と完全に相補的であってもよく、又は部分的に相補的であってもよい。コンパクションオリゴヌクレオチドの3'領域は、その長さに沿って、コンカテマー分子の第2の部分と完全に相補的であってもよく、又は部分的に相補的であってもよい。コンパクションオリゴヌクレオチドの5'領域は、コンカテマー分子の第1のユニバーサル配列部分にハイブリダイズすることができる。コンパクションオリゴヌクレオチドの3'領域は、コンカテマー分子の第2のユニバーサル配列部分にハイブリダイズすることができる。コンパクションオリゴヌクレオチドの5'及び3'領域は、コンカテマーにハイブリダイズして、コンカテマーの遠位部分を一緒にし、コンカテマーのコンパクションを引き起こして、DNAナノボールを形成することができる。

【0455】

コンパクションオリゴヌクレオチドの5'領域は、3'領域と同じ配列を有することができる。コンパクションオリゴヌクレオチドの5'領域は、3'領域とは異なる配列を有することができる。コンパクションオリゴヌクレオチドの3'領域は、5'領域の逆配列である配列を有することができる。

【0456】

いくつかの態様において、配列データは、ナノ細孔シーケンシングを介して導出され得、ナノポアシーケンシングは、膜を横切って、例えば、細孔を介して、当該核酸をトランスロケーションすることによって核酸をシーケンシングすることを含み、配列リード又はベースコールは、トランスロケーション事象中に、インピーダンス、電流、電圧、又はキャパシタンスなどの1つ以上のシグナルを測定することによって行われる。いくつかの態様において、ヌクレオチドのアイデンティティは、電流ブロック、インピーダンス変化、電圧変化、又はキャパシタンス変化のタイミング、持続時間、範囲、又は線形などの独特の電氣的署名によって判定され得る。膜を横切る及び／又は細孔を通るトランスロケーションによる核酸のシーケンシングは、光学的、化学的、生化学的、蛍光的、発光的、磁氣的、電磁的、音響的、又は電気音響的検出などの代替の検出方法を排除しない。

【0457】

支持体及び低非特異的コーティング

いくつかの態様において、図1のフローセル112は、本明細書に開示される支持体、例えば、固体支持体を含むことができる。本開示のいくつかの態様は、ペアワイズシーケンシング組成物及び方法であって、複数のオリゴヌクレオチド表面プライマーが上に固定化された支持体を用いる、組成物及び方法を提供する。いくつかの態様において、支持体は、低非特異的結合コーティングで不動態化されている。本明細書に記載される表面コーティングは、核酸キャプチャ、増幅、及びシーケンシングワークフローのために典型的に使用される試薬、例えば、色素、ヌクレオチド、酵素、及び核酸プライマーへの非常に低い非特異的結合を示す。表面コーティングは、従来の表面コーティングと比較して、低いバックグラウンド蛍光シグナル又は高いコントラスト対ノイズ（CNR）比を示す。

【0458】

低非特異的結合コーティングは、1つの層又は複数の層を含む（図20）。いくつかの態様において、複数の表面プライマーが、低非特異的結合コーティングに固定化されている。いくつかの態様において、少なくとも1つの表面プライマーが、低非特異的結合コーティング内に埋め込まれている。低非特異的結合コーティングは、改善された核酸ハイブ

リダイゼーション及び増幅性能を可能にする。概して、支持体は、基材（又は支持体構造）と、共有結合的に又は非共有結合的に付着した低結合化学修飾層、例えば、シラン層、ポリマーフィルムの中の1つ以上の層と、一本鎖核酸ライブラリ分子を支持体にデザイするために使用され得る1つ以上の共有結合的に又は非共有結合的に付着した表面プライマーと、を含む。いくつかの態様において、コーティングの配合、例えば、1つ以上の層の化学組成、1つ以上の層を支持体にかつ／又は互いに架橋させるために使用されるカップリング化学、並びに層の総数は、コーティングへのタンパク質、核酸分子、並びに他のハイブリダイゼーション及び増幅反応構成要素の非特異的結合が、同等の単層に対して最小化又は低減されるように変動し得る。本明細書に記載されるコーティングの配合は、コーティング上での非特異的ハイブリダイゼーションが、同等の単層に対して最小化又は低減されるように変動し得る。コーティングの配合は、コーティング上での非特異的増幅が、同等の単層に対して最小化又は低減されるように変動し得る。コーティングの配合は、コーティング上での特異的増幅率及び／又は収率が最大化されるように変動し得る。検出に好適である増幅レベルは、本明細書に開示されるいくつかの場合において、2、3、4、5、6、7、8、9、10、15、20、25、30回以下、又は30回超の増幅サイクルにおいて達成される。

10

【0459】

1つ以上の化学修飾層、例えば、低非特異的結合ポリマーの層を含む支持体構造は、独立したものであってもよく、又は別の構造若しくはアセンブリ内に統合されたものであってもよい。例えば、いくつかの態様において、支持体構造は、統合又はアセンブルされたマイクロ流体フローセル内の1つ以上の表面を含み得る。支持体構造は、マイクロプレートフォーマット内の1つ以上の表面、例えば、マイクロプレートにおけるウェルの底面を含み得る。いくつかの態様において、支持体構造は、キャピラリーの内側表面（例えば、ルーメン表面）を含む。いくつかの態様において、支持体構造は、平面チップ内にエッチングされたキャピラリーの内側表面（例えば、ルーメン表面）を含む。

20

【0460】

第1の化学修飾層を支持体の表面にグラフトするために使用される付着化学は、概して、表面が製造される材料及び層の化学的性質の両方に依存する。いくつかの態様において、第1の層は、表面に共有結合的に付着し得る。いくつかの態様において、第1の層は、支持体に、支持体と第1の層の分子構成要素との間の非共有結合相互作用、例えば、静電相互作用、水素結合、又はファンデルワールス相互作用を介して非共有結合的に付着、例えば、吸着し得る。いずれの場合においても、支持体は、第1の層の付着又は堆積前に処理され得る。当業者に既知の様々な表面調製技法のうちのいずれかが、表面を清浄化又は処理するために使用され得る。例えば、ガラス又はシリコン表面は、ピラニア溶液（硫酸（ H_2SO_4 ）及び過酸化水素（ H_2O_2 ）の混合物）、 KOH 及び NaOH 中の塩基処理を使用して酸洗浄され得、かつ／又は酸素プラズマ処理方法を使用して清浄化され得る。

30

【0461】

シラン化学は、ガラス又はシリコン表面上のシラノール基を共有結合的に修飾して、より反応性の官能基（例えば、アミン又はカルボキシル基）を付着させるための非限定的な手法を構築し、次いで、これは、表面へのリンカー分子（例えば、様々な長さの直鎖炭化水素分子、例えば、 C_6 、 C_{12} 、 C_{18} 炭化水素、又は直鎖ポリエチレングリコール（PEG）分子）又は層分子（例えば、分岐PEG分子又は他のポリマー）のカップリングにおいて使用され得る。開示される低結合コーティングのうちのいずれかの作製において使用され得る好適なシランの例としては、（3-アミノプロピル）トリメトキシシラン（APTMS）、（3-アミノプロピル）トリエトキシシラン（APTES）、様々なPEGシラン（例えば、1K、2K、5K、10K、20Kなどの分子量を含むもの）のうちのいずれか、アミノ-PEGシラン（すなわち、遊離アミノ官能基を含むもの）、マレイミド-PEGシラン、及びビオチン-PEGシランなどが挙げられるが、これらに限定されない。

40

50

【0462】

アミノ酸、ペプチド、ヌクレオチド、オリゴヌクレオチド、他のモノマー、若しくはポリマー、又はそれらの組み合わせを含むが、これらに限定されない、当業者に既知の様々な分子のうちのいずれかが、支持体上の1つ以上の化学修飾層の作製において使用され得、使用される構成要素の選択は、層の1つ以上の特性、例えば、官能基及び／若しくはテザーされたオリゴヌクレオチドプライマーの表面密度、層の親水性／疎水性、又は層の三次元性質（すなわち、「厚さ」）を改変するように変動し得る。開示されるコーティングのいずれかにおける低非特異的結合材料の1つ以上の層を作製するために使用され得るポリマーの例としては、様々な分子量及び分岐構造のポリエチレングリコール（PEG）、ストレプトアビジン、ポリアクリルアミド、ポリエステル、デキストラン、ポリーリジン、及びポリーリジンコポリマー、又はそれらのいずれかの組み合わせが挙げられるが、これらに限定されない。材料の1つ以上の層（例えば、ポリマー層）を表面にグラフトするために、かつ／又は層を互いに架橋させるために使用され得るコンジュゲーション化学の例としては、ビオチン-ストレプトアビジン相互作用（又はそれらの変形形態）、his タグ-Ni²⁺/NTAコンジュゲーション化学、メトキシエーテルコンジュゲーション化学、カルボキシレートコンジュゲーション化学、アミンコンジュゲーション化学、NHSエステル、マレイミド、チオール、エポキシ、アジド、ヒドラジド、アルキン、イソシアネート、及びシランが挙げられるが、これらに限定されない。

10

【0463】

低非特異的結合表面コーティングは、支持体全体にわたって均一に適用され得る。代替的に、表面コーティングは、化学修飾層が支持体の1つ以上の個別の領域に制限されるようにパターン化され得る。例えば、コーティングは、フォトリソグラフィ技法を使用して、化学修飾領域の整列したアレイ又はランダムなパターンを支持体上で作製するようにパターン化され得る。代替的に又は組み合わせで、コーティングは、例えば、コンタクトプリンティング及び／又はインクジェットプリンティング技法を使用してパターン化され得る。いくつかの態様において、化学修飾領域の整列したアレイ又はランダムなパターンは、少なくとも1つ、5つ、10個、20個、30個、40個、50個、60個、70個、80個、90個、100個、200個、300個、400個、500個、600個、700個、800個、900個、1000個、2000個、3000個、4000個、5000個、6000個、7000個、8000個、9000個、又は10,000個以上の個別の領域を含み得る。

20

30

【0464】

いくつかの態様において、低非特異性結合コーティングは、支持体に非特異的に吸着しているか又は共有結合的にグラフトされている親水性ポリマーを含む。典型的には、不動態化は、ポリ（エチレングリコール）（PEG、これは、ポリエチレンオキシド（PEO）又はポリオキシエチレンとしても既知である）、又は支持体に、例えば、シラン化学を使用して結合した異なる分子量及び末端基を有する他の親水性ポリマーを使用して実行される。表面から遠位の末端基は、ビオチン、メトキシエーテル、カルボキシレート、アミン、NHSエステル、マレイミド、及びビス-シランを含むことができるが、これらに限定されない。いくつかの態様において、親水性ポリマー、例えば、直鎖ポリマー、分岐ポリマー、又は多分岐ポリマーの2つ以上の層が、表面上に堆積し得る。いくつかの態様において、2つ以上の層は、得られるコーティングの安定性を改善するために、互いに共有結合的にカップリングしてもよく、又は内部的に架橋してもよい。いくつかの態様において、異なるヌクレオチド配列及び／又は塩基修飾（又は他の生体分子、例えば、酵素若しくは抗体）を有する表面プライマーは、得られた層に、様々な表面密度でテザーされ得る。いくつかの態様において、例えば、表面官能基密度及び表面プライマー濃度の両方が、所望の表面プライマー密度範囲を達成するように変動し得る。加えて、表面プライマー密度は、表面プライマーを、同じ官能基を保有する他の分子で希釈することによって制御され得る。例えば、最終プライマー密度を低減するために、アミン標識表面プライマーが、NHSエステルコーティング表面との反応において、アミン標識ポリエチレングリコール

40

50

で希釈され得る。また、ハイブリダイゼーション領域と表面付着官能基との間の異なる長さのリンカーを有する表面プライマーが、表面密度を制御するために適用され得る。好適なリンカーの例としては、プライマーの5'末端におけるポリT及びポリA鎖（例えば、0～20個の塩基）、PEGリンカー（例えば、3つ～20個のモノマー単位）、及び炭素鎖（例えば、C6、C12、C18など）が挙げられる。プライマー密度を測定するために、蛍光標識プライマーが、表面にテザーされ得、次いで、蛍光リードが、既知の濃度の色素溶液についてのものと比較され得る。

【0465】

いくつかの態様において、低非特異性結合コーティングは、少なくとも支持体の一部分に、支持体上の化学基を介して共有結合的に結合した官能化ポリマーコーティング層と、官能化ポリマーコーティングにグラフトされたプライマーと、プライマー及び官能化ポリマーコーティング上の水溶性保護コーティングと、を含む。いくつかの態様において、官能化ポリマーコーティングは、ポリ（N-（5-アジドアセトアミジルペンチル）アクリルアミド-コ-アクリルアミド（PAZAM）を含む。

【0466】

プライマー表面密度を調整し、追加の次元性を親水性又は両性コーティングに追加するために、PEG及び他の親水性ポリマーの多層コーティングを含む支持体が開発されている。以下に記載されるポリマー／コポリマー材料を含むが、これらに限定されない親水性及び両性表面層化手法を使用することによって、支持体上のプライマーロード密度を、顕著に増加させることが可能である。従来のPEGコーティング手法は、単層プライマー堆積を使用し、単層プライマー堆積は、概して、単一の分子の用途について報告されているが、核酸増幅用途について高いコピー数をもたらさない。本明細書に記載されるように、「層化」は、従来の架橋手法を使用して、2つ以上の高度に架橋した層を含む表面が順次構築され得るように、いずれかの適合性ポリマー又はモノマーサブユニットで達成され得る。好適なポリマーの例としては、ストレプトアビジン、ポリアクリルアミド、ポリエステル、デキストラン、ポリ-リジン、並びにポリ-リジン及びPEGのコポリマーが挙げられるが、これらに限定されない。いくつかの態様において、異なる層は、様々なコンジュゲーション反応のうちのいずれかを介して互いに付着し得、コンジュゲーション反応は、ビオチン-ストレプトアビジン結合、アジド-アルキンクリック反応、アミン-NHSエステル反応、チオール-マレイミド反応、及び正電荷を有するポリマーと負電荷を有するポリマーとの間のイオン相互作用を含むが、これらに限定されない。いくつかの態様において、高プライマー密度材料は、溶液中で構築され得、その後、複数のステップにおいて、表面上に層化され得る。

【0467】

支持体構造が製造され得る材料の例としては、ガラス、熔融シリカ、シリコン、ポリマー（例えば、ポリスチレン（PS）、マクロ多孔質ポリスチレン（MPPS）、ポリメチルメタクリレート（PMMA）、ポリカーボネート（PC）、ポリプロピレン（PP）、ポリエチレン（PE）、高密度ポリエチレン（HDPE）、環状オレフィンポリマー（COP）、環状オレフィンコポリマー（COC）、ポリエチレンテレフタレート（PET））、又はそれらのいずれかの組み合わせが挙げられるが、これらに限定されない。ガラス及びプラスチック支持体構造の両方の様々な組成が企図される。

【0468】

支持体構造は、当業者に既知の様々な幾何学形状及び寸法のうちのいずれかにされ得、当業者に既知の様々な材料のうちのいずれかを含み得る。例えば、支持体構造は、局所的に平面であり得る（例えば、顕微鏡スライド又は顕微鏡スライドの表面を含む）。全般的に、支持体構造は、円柱状（例えば、キャピラリー又はキャピラリーの内側表面を含む）、球状（例えば、非多孔質ビーズの外側表面を含む）、又は不規則（例えば、不規則な形状の非多孔質ビーズ又は粒子の外側表面を含む）であり得る。いくつかの態様において、核酸ハイブリダイゼーション及び増幅のために使用される支持体構造の表面は、固体非多孔質表面であり得る。いくつかの態様において、核酸ハイブリダイゼーション及び増幅の

ために使用される支持体構造の表面は、多孔質であり得、これにより、本明細書に記載されるコーティングは、多孔質表面を浸透し、その上で実行される核酸ハイブリダイゼーション及び増幅反応は、細孔内で発生し得る。

【0469】

1つ以上の化学修飾層、例えば、低非特異的結合ポリマーの層を含む支持体構造は、独立したものであってもよく、又は別の構造若しくはアセンブリ内に統合されたものであってもよい。例えば、支持体構造は、統合又はアセンブルされたマイクロ流体フローセル内の1つ以上の表面を含み得る。支持体構造は、マイクロプレートフォーマット内の1つ以上の表面、例えば、マイクロプレートにおけるウェルの底面を含み得る。いくつかの態様において、支持体構造は、キャピラリーの内側表面（例えば、ルーメン表面）を含む。いくつかの態様において、支持体構造は、平面チップ内にエッチングされたキャピラリーの内側表面（例えば、ルーメン表面）を含む。

10

【0470】

上記のように、本開示のいくつかの態様の低非特異的結合支持体は、固相核酸増幅のために使用されるハイブリダイゼーション及び／又は増幅配合物のタンパク質、核酸、及び他の構成要素の低減された非特異的結合を示す。所与の支持体表面によって示される非特異的結合の程度は、定性的又は定量的のいずれかで評価され得る。表面を、標準化された条件のセット下で、例えば、蛍光色素（例えば、シアニン、例えば、Cy3若しくはCy5など、フルオレセイン、クマリン、ローダミンなど、又は本明細書に開示される他の色素）、蛍光標識ヌクレオチド、蛍光標識オリゴヌクレオチド、及び／又は蛍光標識タンパク質（例えば、ポリメラーゼ）に曝露し、続いて、特定のリンスプロトコル及び蛍光イメージングを行うことは、異なる表面配合物を含む支持体上での非特異的結合の比較のための定性的ツールとして使用され得る。いくつかの態様において、表面を、標準化された条件のセット下で、例えば、蛍光色素、蛍光標識ヌクレオチド、蛍光標識オリゴヌクレオチド、及び／又は蛍光標識タンパク質（例えば、ポリメラーゼ）に曝露し、続いて、特定のリンスプロトコル及び蛍光イメージングを行うことは、異なる表面配合物を含む支持体上での非特異的結合の比較のための定量的ツールとして使用され得るが、蛍光イメージングが、蛍光シグナルが支持体上のフルオロフォアの数に直線的に関係する（又は予測可能な様式で関係する）条件下（例えば、フルオロフォアのシグナル飽和及び／又は自己消光が問題でない条件下）で実行され、好適な較正基準が使用されることを確実にするように配慮されていることを条件とする。いくつかの態様において、当業者に既知の他の技法、例えば、放射性同位体標識及びカウント方法が、非特異的結合が本開示のいくつかの態様の異なる支持体表面配合物によって示される程度の定量的評価のために使用され得る。

20

30

【0471】

本明細書に開示されるいくつかの表面は、少なくとも2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、50、75、100、若しくは100超、又は本明細書に記載される範囲によって包含されるいずれかの中間値のCy3などのフルオロフォアの特異的対非特異的結合の比を示す。本明細書に開示されるいくつかの表面は、少なくとも2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、50、75、100、若しくは100超、又は本明細書に記載される範囲によって包含されるいずれかの中間値のCy3などのフルオロフォアの特異的対非特異的蛍光の比を示す。

40

【0472】

開示される低結合支持体によって示される非特異的結合の程度は、表面を、標識タンパク質（例えば、ウシ血清アルブミン（BSA）、ストレプトアビジン、DNAポリメラーゼ、逆転写酵素、ヘリカーゼ、一本鎖結合タンパク質（SSB）など、又はそれらのいずれかの組み合わせ）、標識ヌクレオチド、標識オリゴヌクレオチドなどと接触させるための標準化されたプロトコルを使用して、標準化されたインキュベーション及びリンス条件のセットの下で評価され得、続いて、表面上に残存する標識の量の検出、及びそれから得

50

られたシグナルと適切な較正標準との比較が行われ得る。いくつかの態様において、標識は、蛍光標識を含み得る。いくつかの態様において、標識は、放射性同位体を含み得る。いくつかの態様において、標識は、当業者に既知のいずれかの他の検出可能な標識を含み得る。したがって、いくつかの態様において、所与の支持体表面配合物によって示される非特異的結合の程度は、1単位面積当たりの非特異的結合タンパク質分子（若しくは核酸分子、又は他の分子）の数の観点から評価され得る。いくつかの態様において、本開示の低結合支持体は、1 μm^2 当たり0.001分子未満、1 μm^2 当たり0.01分子未満、1 μm^2 当たり0.1分子未満、1 μm^2 当たり0.25分子未満、1 μm^2 当たり0.5分子未満、1 μm^2 当たり1分子未満、1 μm^2 当たり10分子未満、1 μm^2 当たり100分子未満、又は1 μm^2 当たり1,000分子未満の非特異的タンパク質結合（又は他の特定の分子（例えば、シアニン、例えば、Cy3若しくはCy5など、フルオレセイン、クマリン、ローダミンなど、又は本明細書に開示される他の色素）の非特異的結合）を示し得る。本開示の態様の所与の支持体表面は、この範囲内のいずれかの値である非特異的結合、例えば、1 μm^2 当たり86分子未満の非特異的結合を示し得ることが、当業者には認識されよう。例えば、本明細書に開示されるいくつかの修飾表面は、リン酸緩衝生理食塩水（PBS）緩衝液中のCy3標識ストレプトアビジン（GE Amersham）の1 μM の溶液と15分間接触させ、続いて、脱イオン水で3回リンスした後に、0.5分子/ μm^2 未満の非特異的タンパク質結合を示す。本明細書に開示されるいくつかの修飾表面は、1 μm^2 当たり0.25分子未満のCy3色素分子の非特異的結合を示す。独立した非特異的結合アッセイにおいて、1 μM の標識Cy3 SA（ThermoFisher）、1 μM のCy5 SA色素（ThermoFisher）、10 μM のアミノアリル-dUTP-ATTO-647N（Jena Biosciences）、10 μM のアミノアリル-dUTP-ATTO-Rho1（Jena Biosciences）、10 μM のアミノアリル-dUTP-ATTO-Rho1（Jena Biosciences）、10 μM の7-プロパルギルアミノ-7-デアザ-dGTP-Cy5（Jena Biosciences、及び10 μM の7-プロパルギルアミノ-7-デアザ-dGTP-Cy3（Jena Biosciences）が、低結合コーティング支持体上で、37℃で15分間、384ウェルプレートフォーマットにおいてインキュベートされた。各々のウェルは、50 μl の脱イオン化RNase/DNaseフリー水で2～3回され、pH7.4の25mMのACES緩衝液で2～3回リンスされた。384ウェルプレートは、GE Typhoon機器において、製造業者によって指定されるように（実行された色素試験に従って）Cy3、AF555、又はCy5フィルタセットを使用して、800のPMTゲイン設定及び50～100 μm の解像度でイメージングされた。より高い解像度のイメージングのために、画像は、全反射照明蛍光（TIRF）対物レンズ（100 $\times$ 、1.5NA、Olympus）、CCDカメラ（例えば、Olympus EM-CCDモノクロカメラ、Olympus XM-10モノクロカメラ、又はOlympus DP80カラー及びモノクロカメラ）、照明源（例えば、Olympus 100W Hgランプ、Olympus 75W Xeランプ、又はOlympus U-HGLGPS蛍光光源）、及び532nm又は635nmの励起波長を有するOlympus IX83顕微鏡（例えば、倒立蛍光顕微鏡）（Olympus Corp.、Center Valley, Pa.）において収集された。ダイクロイックミラーは、Semrock（IDEX Health & Science, LLC、Rochester, N.Y.）から購入され、例えば、405、488、532、又は633nmのダイクロイックリフレクタ／ビームスプリッタであり、バンドパスフィルタは、適切な励起波長と一致する532LP又は645LPとして選択された。本明細書に開示されるいくつかの修飾表面は、1 μm^2 当たり0.25分子未満の色素分子の非特異的結合を示す。いくつかの態様において、コーティング支持体は、画像が取得されつつ、緩衝液（例えば、25mMのACES、pH7.4）中に浸漬された。

【0473】

いくつかの態様において、本明細書に開示される表面は、少なくとも2、3、4、5、

10

20

30

40

50

6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、50、75、100、若しくは100超、又は本明細書に記載される範囲によって包含されるいずれかの中間値のCy3などのフルオロフォアの特異的対非特異的結合の比を示す。いくつかの態様において、本明細書に開示される表面は、少なくとも2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、50、75、100、若しくは100超、又は本明細書に記載される範囲によって包含されるいずれかの中間値のCy3などのフルオロフォアについての特異的対非特異的蛍光シグナルの比を示す。

【0474】

本明細書の開示と一致する低バックグラウンド表面は、少なくとも4：1、5：1、6：1、7：1、8：1、9：1、10：1、15：1、20：1、30：1、40：1、50：1の特異的色素付着（例えば、Cy3付着）対非特異的色素吸着（例えば、Cy3色素吸着）比、又は非特異的に吸着した1分子当たり50超の付着した特異的色素分子を示し得る。同様に、励起エネルギーに供されたときに、フルオロフォア、例えば、Cy3が付着した、本明細書の開示と一致する低バックグラウンド表面は、少なくとも4：1、5：1、6：1、7：1、8：1、9：1、10：1、15：1、20：1、30：1、40：1、50：1、又は50：1超の（例えば、表面に付着したCy3標識オリゴヌクレオチドから生じる）特異的蛍光シグナル対非特異的吸着色素蛍光シグナルの比を示し得る。

【0475】

いくつかの態様において、開示される支持体表面の親水性の程度（又は水溶液との「濡れ性」）は、例えば、水接触角の測定を介して評価され得、この測定において、小さい水滴が、表面上に配置され、表面とのその接触角が、例えば、光学テンシオメータを使用して測定される。いくつかの態様において、静的接触角が判定され得る。いくつかの態様において、前進又は後退接触角が判定され得る。いくつかの態様において、本明細書に開示される表面処理された親水性低結合支持体についての水接触角は、約0度～約30度の範囲であり得る。いくつかの態様において、本明細書に開示される表面処理された親水性低結合支持体についての水接触角は、50度、40度、30度、25度、20度、18度、16度、14度、12度、10度、8度、6度、4度、2度、又は1度以下であり得る。多くの場合において、接触角は、40度以下である。本開示のいくつかの態様の所与の親水性低結合支持体表面は、この範囲内のいずれかにおける値を有する水接触角を示し得ることが、当業者には認識されよう。

【0476】

いくつかの態様において、本明細書に開示される親水性表面は、しばしば、低結合表面への生体分子の低減された非特異的結合に起因して、バイオアッセイのため低減された洗浄時間を容易にする。いくつかの態様において、適切な洗浄ステップは、60、50、40、30、20、15、10秒間未満、又は10秒間未満内で実行され得る。例えば、適切な洗浄ステップは、30秒間未満内で実行され得る。

【0477】

本開示のいくつかの低結合表面は、溶媒及び高温への長期曝露に対する、又は溶媒曝露若しくは温度における変化の繰り返されるサイクルに対する安定性又は耐久性における顕著な改善を示す。例えば、開示される表面の安定性は、表面上の官能基、又は表面上のテザーされた生体分子（例えば、オリゴヌクレオチドプライマー）を蛍光標識し、溶媒及び高温への、又は溶媒曝露若しくは温度における変化の繰り返されるサイクルへの長期曝露前に、それ中に、かつその後、に、蛍光シグナルを監視することによって試験され得る。いくつかの態様において、表面の品質を評価するために使用される蛍光における変化の程度は、溶媒及び／又は高温への曝露の1分間、2分間、3分間、4分間、5分間、10分間、20分間、30分間、40分間、50分間、60分間、2時間、3時間、4時間、5時間、6時間、7時間、8時間、9時間、10時間、15時間、20時間、25時間、30時間、35時間、40時間、45時間、50時間、又は100時間の期間にわたって、1

10

20

30

40

50

%、2%、3%、4%、5%、10%、15%、20%、又は25%未満（又はこれらの期間にわたって測定された場合のこれらの百分率のいずれかの組み合わせ）であり得る。いくつかの態様において、表面の品質を評価するために使用される蛍光における変化の程度は、5回のサイクル、10回のサイクル、20回のサイクル、30回のサイクル、40回のサイクル、50回のサイクル、60回のサイクル、70回のサイクル、80回のサイクル、90回のサイクル、100回のサイクル、200回のサイクル、300回のサイクル、400回のサイクル、500回のサイクル、600回のサイクル、700回のサイクル、800回のサイクル、900回のサイクル、又は1,000回のサイクルの溶媒変化及び／又は温度における変化への繰り返される曝露にわたって、1%、2%、3%、4%、5%、10%、15%、20%、又は25%未満（又はサイクルのこの範囲にわたって測定された場合のこれらの百分率のいずれかの組み合わせ）であり得る。

10

【0478】

いくつかの態様において、本明細書に開示される表面は、特異的シグナル対非特異的シグナル又は他のバックグラウンドの高い比を示し得る。例えば、核酸増幅のために使用されるときに、いくつかの表面は、表面の隣接する非集合領域のシグナルよりも少なくとも4、5、6、7、8、9、10、15、20、30、40、50、75、100、又は100倍超大きい増幅シグナルを示し得る。同様に、いくつかの表面は、表面の隣接する増幅された核酸集合領域のシグナルよりも少なくとも4、5、6、7、8、9、10、15、20、30、40、50、75、100、又は100倍超大きい増幅シグナルを示す。

【0479】

いくつかの態様において、核酸ハイブリダイゼーション又は増幅用途において、（例えば、フルオロフォアで直接又は間接的に標識された）ハイブリダイズした又はクローン的に増幅された核酸分子のポロニーを作製するために使用されるときの開示される低バックグラウンド表面の蛍光画像は、少なくとも10、20、30、40、50、60、70、80、90、100、110、120、130、140、150、160、170、180、190、200、210、220、230、240、250、又は250超のコントラスト対ノイズ比（CNR）を示す。

20

【0480】

1つ以上のタイプのプライマーが、支持体表面に付着してもよく又はテザーされてもよい。いくつかの態様において、1つ以上のタイプのアダプター又はプライマーは、スパーサー配列、アダプターライゲーションされた標的ライブラリ核酸配列へのハイブリダイゼーションのためのアダプター配列、フォワード増幅プライマー、リバーズ増幅プライマー、シーケンシングプライマー、及び／若しくは分子バーコード配列、又はそれらのいずれかの組み合わせを含み得る。いくつかの態様において、1つのプライマー又はアダプター配列が、表面の少なくとも1つの層にテザーされ得る。いくつかの態様において、少なくとも2つ、3つ、4つ、5つ、6つ、7つ、8つ、9つ、10個、又は10個超の異なるプライマー又はアダプター配列が、表面の少なくとも1つの層にテザーされ得る。

30

【0481】

いくつかの態様において、テザーされたアダプター及び／又はプライマー配列は、約10個のヌクレオチド～約100個のヌクレオチドの範囲の長さであり得る。いくつかの態様において、テザーされたアダプター及び／又はプライマー配列は、少なくとも10個、少なくとも20個、少なくとも30個、少なくとも40個、少なくとも50個、少なくとも60個、少なくとも70個、少なくとも80個、少なくとも90個、又は少なくとも100個のヌクレオチド長であり得る。いくつかの態様において、テザーされたアダプター及び／又はプライマー配列は、最大100個、最大90個、最大80個、最大70個、最大60個、最大50個、最大40個、最大30個、最大20個、又は最大10個のヌクレオチド長であり得る。この段落に記載される下限及び上限値のうちのいずれかが、本開示内に含まれる範囲を形成するように組み合わせられ得、例えば、いくつかの態様において、テザーされたアダプター及び／又はプライマー配列の長さは、約20個のヌクレオチド～約80個のヌクレオチドの範囲であり得る。テザーされたアダプター及び／又はプライマ

40

50

一配列の長さは、この範囲内のいずれかの値、例えば、約24個のヌクレオチドを有し得ることが、当業者には認識されよう。

【0482】

いくつかの態様において、本開示の低結合支持体表面上のプライマー（例えば、キャプチャプライマー）の得られる表面密度は、 $1\mu\text{m}^2$ 当たり約100個のプライマー分子～ $1\mu\text{m}^2$ 当たり約100,000個のプライマー分子の範囲であり得る。いくつかの態様において、本開示の低結合支持体表面上のプライマーの得られる表面密度は、 $1\mu\text{m}^2$ 当たり約1,000個のプライマー分子～ $1\mu\text{m}^2$ 当たり約1,000,000個のプライマー分子の範囲であり得る。いくつかの態様において、プライマーの表面密度は、 $1\mu\text{m}^2$ 当たり少なくとも1,000個、少なくとも10,000個、少なくとも100,000個、又は少なくとも1,000,000個の分子であり得る。いくつかの態様において、プライマーの表面密度は、 $1\mu\text{m}^2$ 当たり最大1,000,000個、最大100,000個、最大10,000個、又は最大1,000個の分子であり得る。この段落に記載される下限及び上限値のうちのいずれかが、本開示内に含まれる範囲を形成するように組み合わせられ得、例えば、いくつかの態様において、プライマーの表面密度は、 $1\mu\text{m}^2$ 当たり約10,000分子～ $1\mu\text{m}^2$ 当たり約100,000分子の範囲であり得る。プライマー分子の表面密度は、この範囲内のいずれかの値、例えば、 $1\mu\text{m}^2$ 当たり約455,000個の分子を有し得ることが、当業者には認識されよう。いくつかの態様において、支持体表面上のアダプター又はプライマー配列に最初にハイブリダイズした標的ライブラリ核酸配列の表面密度は、テザーされたプライマーの表面密度について示されるもの以下であり得る。いくつかの態様において、支持体表面上のアダプター又はプライマー配列にハイブリダイズしたクローン的に増幅された標的ライブラリ核酸配列の表面密度は、テザーされたプライマーの表面密度について示されるものと同じ範囲を包含する。

10

20

【0483】

上記で列挙される局所密度は、表面が、例えば、 $500,000/\mu\text{m}^2$ のオリゴ密度を有する領域を含みつつ、また、実質的に異なる局所密度を有する少なくとも第2の領域を含み得るように、密度が表面にわたって変動することを排除しない。

【0484】

いくつかの態様において、開示される反応配合物及び低結合支持体を使用する核酸ハイブリダイゼーション及び／又は増幅反応の性能は、蛍光イメージング技法を使用して評価され得、画像のコントラスト対ノイズ比（CNR）は、増幅特異性及び支持体上での非特異的結合の評価における重要なメトリックを提供する。CNRは、一般に、 $\text{CNR} = (\text{シグナル} - \text{バックグラウンド}) / \text{ノイズ}$ として定義される。バックグラウンド項は、一般に、特定の関心領域（ROI）における特定の特徴（回折限界スポット、DLS）を取り囲むインターステイシャル領域について測定されたシグナルであると考えられる。シグナル対ノイズ比（SNR）は、しばしば、全体的なシグナル品質のベンチマークであると考えられるが、以下の例に示されるように、改善されたCNRは、迅速な画像キャプチャを必要とする用途（例えば、サイクル時間が最小化されなければならないシーケンシング用途）におけるシグナル品質についてのベンチマークとして、SNRよりも顕著な利点を提供できることが示され得る。高いCNRで、正確な識別に到達するために必要とされるイメージング時間（したがって、シーケンシング用途の場合において、正確なベースコーリング）は、CNRにおける中程度の改善であっても著しく低減され得る。イメージング統合時間についてのイメージングデータにおける改善されたCNRは、支持体表面上のクローン的に増幅された核酸コロニーなどの特徴をより正確に検出するための方法を提供する。

30

40

【0485】

ほとんどのアンサンブルベースのシーケンシング手法において、バックグラウンド項は、典型的には、「インターステイシャル」領域に関連するシグナルとして測定される。「インターステイシャル」バックグラウンド（B i n t e r）に加えて、「イントラスティシャル」バックグラウンド（B i n t r a）は、増幅されたDNAコロニーによって占められる領域内に存在する。これらの2つのバックグラウンドシグナルの組み合わせは、達

50

成可能なCNRを決定し、その後、光学機器要件、アーキテクチャコスト、試薬コスト、実行時間、コスト／ゲノムに直接影響し、最終的に、環状アレイベースのシーケンシング用途についての精度及びデータ品質に直接影響する。B i n t e rバックグラウンドシグナルは、様々な供給源から生じ、いくつかの例としては、消耗フローセルからの自己蛍光、ROIからのシグナルを不明瞭にし得るスプリアス蛍光シグナルをもたらす検出分子の非特異的吸着、非特異的DNA増幅産物（例えば、プライマー二量体から生じるもの）の存在が挙げられる。典型的な次世代シーケンシング（NGS）用途において、現在の視野（FOV）におけるこのバックグラウンドシグナルは、経時的に平均化され、減算される。個々のDNAコロニーから生じるシグナル（すなわち、FOVにおける（シグナル）－B（インタースティシャル））は、分類され得る識別可能な特徴をもたらす。いくつかの態様において、イントラスティシャルバックグラウンド（B（イントラスティシャル））は、目的の標的に特異的でないが、同じROIにおいて存在する交絡蛍光シグナルに寄与し得、したがって、平均化及び減算をはるかにより困難にする。

10

【0486】

本明細書に記載される低結合コーティング支持体上での核酸増幅は、B（インタースティシャル）バックグラウンドシグナルを、非特異的結合を低減することによって減少させ得、特異的核酸増幅における改善をもたらし得、インタースティシャル及びイントラスティシャル領域の両方から生じるバックグラウンドシグナルに影響し得る非特異的増幅における減少をもたらし得る。いくつかの態様において、任意選択で、開示されるハイブリダイゼーション及び／又は増幅反応配合物との組み合わせで使用される、開示される低結合コーティング支持体は、CNRにおける改善を、従来の支持体、並びにハイブリダイゼーション、増幅、及び／又はシーケンシングプロトコルを使用して達成されるものよりも2、5、10、100、250、500、又は1000倍もたらし得る。蛍光イメージングをリードアウト又は検出モードとして使用するという文脈において本明細書に記載されるが、同じ原理が、光学及び非光学検出モードの両方を含む他の検出モードのための、開示される低結合コーティング支持体、並びに核酸ハイブリダイゼーション及び増幅配合物の使用にも適用される。

20

【0487】

本明細書で提供される表題は、本開示の様々な態様の限定ではなく、これらの態様は、本明細書全体を参照することによって理解され得る。

30

【0488】

別途定義されない限り、本明細書で使用される全ての技術及び科学用語は、別途定義されない限り、当業者によって一般に理解される意味を有する。概して、本明細書に記載される分子生物学、核酸化学、タンパク質化学、遺伝学、微生物学、遺伝子導入細胞産生、及びハイブリダイゼーションの技法に関する用語は、当該技術分野で周知であり一般に使用されるものである。本明細書に記載される技法及び手順は、概して、当該技術分野で周知の従来の方法に従って実行され、本明細書全体にわたって引用及び考察される様々な一般的な及びより具体的な参照文献に記載されているように実行される。例えば、S a m b r o o k e t a l . , M o l e c u l a r C l o n i n g : A L a b o r a t o r y M a n u a l (T h i r d e d . , C o l d S p r i n g H a r b o r L a b o r a t o r y P r e s s , C o l d S p r i n g H a r b o r , N . Y . 2 0 0 0) を参照されたい。また、A u s u b e l e t a l . , C u r r e n t P r o t o c o l s i n M o l e c u l a r B i o l o g y , G r e e n e P u b l i s h i n g A s s o c i a t e s (1 9 9 2) を参照されたい。本明細書に記載される実験手順及び技法に関連して使用される命名法、並びに実験手順及び技法は、当該技術分野で周知であり一般に使用されるものである。

40

【0489】

別途本明細書で文脈によって必要とされない限り、単数形用語は、複数形を含み、複数形用語は、単数形を含む。1つという言葉及び明解に限定されない限り、単数形の「a」、「an」、及び「the」、並びにいずれもの単語の単数形の使用は、複

50

数の言及を含む。

【0490】

代替の用語（例えば、「又は」）の使用は、代替形態の一方若しくは両方、又はそれらのいずれかの組み合わせのいずれかを意味すると理解される。

【0491】

本明細書で使用される場合、「及び／又は」という用語は、特定の特徴又は構成要素の各々が、他方を有するか又は有しないという具体的な開示を意味すると理解されるべきである。例えば、本明細書で「A 及び／又は B」という語句などにおいて使用される場合、「及び／又は」という用語は、「A 及び B」、「A 又は B」、「A」（A 単独で）、及び「B」（B 単独で）を含むことが意図されている。同様の様式で、「A、B、及び／又は C」という語句などにおいて使用される場合、「及び／又は」という用語は、以下の態様、「A、B、及び C」、「A、B、又は C」、「A 又は C」、「A 又は B」、「B 又は C」、「A 及び B」、「B 及び C」、「A 及び C」、「A」（A 単独で）、「B」（B 単独で）、及び「C」（C 単独で）の各々を包含することが意図されている。

10

【0492】

本明細書及び添付の特許請求の範囲で使用される場合、「含む (comprising)」、「含む (including)」、「有する (having)」、及び「含有する (containing)」という用語、並びに本明細書で使用される場合、それらの文法上の変形は、列挙内の 1 つの項目又は複数の項目が、列挙された項目に置換又は追加され得る他の項目を排除しないように非限定的であることが意図されている。態様が、「含む (comprising)」という用語で本明細書に記載されるいかなる場合も、「からなる」及び／又は「から本質的になる」という用語で記載される別法の類似の態様もまた提供されることが理解される。

20

【0493】

本明細書で使用される場合、「約」、「およそ」、及び「実質的に」という用語は、当業者によって判定される場合の、特定の値又は組成物についての許容されるエラー範囲内である値又は組成物を指し、許容されるエラー範囲は、どのように値又は組成物が測定又は判定されるか、すなわち、測定システムの制限に部分的に依存する。例えば、「約」、「およそ」、又は「実質的に」とは、当該技術分野における実施当たり 1 つの又は 2 つ以上の標準偏差内を意味することができる。代替的に、「約」又は「およそ」とは、測定システムの制限に依存して、最大 10 %（すなわち、±10 %）以上の範囲を意味することができる。例えば、約 5 mg は、4.5 mg ~ 5.5 mg のいずれかの数を含むことができる。更に、特に、生物学的システム又はプロセスに関して、この用語は、値の最大 1 桁又は最大 5 倍を意味することができる。特定の値又は組成物が本開示で提供されるときに、別途記載されない限り、「約」、「およそ」、「実質的に」の意味は、この特定の値又は組成物についての許容されるエラー範囲内であると考えられるべきである。また、値の範囲及び／又はサブ範囲が提供される場合、範囲及び／又はサブ範囲は、範囲及び／又はサブ範囲の終点を含むことができる。

30

【0494】

本明細書で使用される「ポロニー」という用語は、核酸ライブラリ分子を溶液内又は支持体上でクローン的に増幅して、シーケンシングのためのテンプレート分子として機能することができるアンプリコンを生成することができることを指す。いくつかの態様において、線形ライブラリ分子を環状化して、環状化ライブラリ分子を生成することができ、環状化ライブラリ分子を溶液内又は支持体上でクローン的に増幅して、コンカテマーを生成することができる。いくつかの態様において、コンカテマーは、シーケンシングされ得る核酸テンプレート分子として機能し得る。コンカテマーは、ポロニーと呼ばれることもある。いくつかの態様において、ポロニーは、ヌクレオチド鎖を含む。

40

【0495】

「ペプチド」、「ポリペプチド」、及び「タンパク質」という用語並びに本明細書で使用される他の関連用語は、同義的に使用され、アミノ酸のポリマーを指し、任意の特定の

50

長さに限定されない。ポリペプチドは、天然及び非天然アミノ酸を含み得る。ポリペプチドとしては、組換え型又は化学合成型が挙げられる。また、ポリペプチドには、タンパク質分解切断、リボソームスキッピングによる切断、ヒドロキシル化、メチル化、脂質化、アセチル化、SUMO化、ユビキチン化、グリコシル化、リン酸化、及び／又はジスルフィド結合形成などの翻訳後修飾をまだ受けていない前駆体分子も含まれる。これらの用語は、タンパク質配列の天然及び人工タンパク質、タンパク質断片及びポリペプチド類似体（変異タンパク質、バリエーション、キメラタンパク質、及び融合タンパク質など）、並びに翻訳後、又はそうでなければ共有結合的若しくは非共有結合的に修飾されたタンパク質を包含する。

【0496】

「ポリメラーゼ」という用語及びそのバリエーションは、本明細書で使用される場合、ヌクレオチド（その類似体を含む）の核酸鎖への重合を触媒することができる任意の酵素を含む。典型的には、必ずしもそうではないが、そのようなヌクレオチド重合は、テンプレートに依存的な様式で発生することができる。典型的には、ポリメラーゼは、1つ以上の活性部位を含み、1つ以上の活性部位において、ヌクレオチド結合及び／又はヌクレオチド重合の触媒が発生することができる。いくつかの態様において、ポリメラーゼは、他の酵素活性、例えば、3'から5'エキソヌクレアーゼ活性又は5'から3'エキソヌクレアーゼ活性を含む。いくつかの態様において、ポリメラーゼは、鎖置換活性を有する。ポリメラーゼは、自然発生のポリメラーゼ並びにそれらのいずれかのサブユニット及び切断化、変異体ポリメラーゼ、バリエーションポリメラーゼ、組換え、融合又は別法で操作されたポリメラーゼ、化学修飾ポリメラーゼ、合成分子又はアセンブリ、並びにヌクレオチド重合を触媒する能力を保持するそれらのいずれかの類似体、誘導体、又は断片（例えば、触媒活性断片）を含むことができるが、これらに限定されない。いくつかの態様において、ポリメラーゼは、細胞から単離されてもよく、又は組換えDNA技術若しくは化学合成方法を使用して生成されてもよい。いくつかの態様において、ポリメラーゼは、原核生物、真核生物、ウイルス、又はファージ生物において発現され得る。いくつかの態様において、ポリメラーゼは、翻訳後修飾タンパク質又はそれらの断片であることができる。ポリメラーゼは、原核生物、真核生物、ウイルス、又はファージに由来することができる。ポリメラーゼは、DNA指向性DNAポリメラーゼ及びRNA指向性DNAポリメラーゼを含む。

【0497】

本明細書で使用される場合、「忠実度」という用語は、テンプレート依存性DNAポリメラーゼによるDNA重合の精度を指す。DNAポリメラーゼの忠実度は、典型的には、エラー率（不正確なヌクレオチド、すなわち、テンプレートヌクレオチドと相補的でないヌクレオチドを組み込む頻度）によって測定される。DNA重合の精度又は忠実度は、DNAポリメラーゼのポリメラーゼ活性及び3'から5'エキソヌクレアーゼ活性の両方によって維持される。

【0498】

本明細書で使用される場合、「結合複合体」という用語は、核酸二重鎖、ポリメラーゼ、及び多価分子の遊離ヌクレオチド又はヌクレオチド単位と一緒に結合させることによって形成される複合体であって、核酸二重鎖が、核酸プライマーにハイブリダイズした核酸テンプレート分子を含む、複合体を指す。結合複合体において、遊離ヌクレオチド又はヌクレオチド単位は、核酸プライマーの3'末端に、核酸テンプレート分子における相補的ヌクレオチドの反対側である位置において結合してもよく又は結合しなくてもよい。「三元複合体」は、核酸二重鎖、ポリメラーゼ、及び多価分子の遊離ヌクレオチド又はヌクレオチド単位と一緒に結合させることによって形成される結合複合体の例であり、遊離ヌクレオチド又はヌクレオチド単位は、（核酸二重鎖の一部として）核酸プライマーの3'末端に、核酸テンプレート分子における相補的ヌクレオチドと反対側である位置において結合している。

【0499】

「持続時間」という用語及び関連用語は、結合複合体が、構成要素のうちのいずれもが

10

20

30

40

50

解離することなく依然として安定している時間の長さであって、結合複合体の構成要素が、核酸テンプレート及び核酸プライマー、ポリメラーゼ、多価分子のヌクレオチド単位、又は遊離（例えば、非コンジュゲート型の）ヌクレオチドを含む、時間の長さを指す。ヌクレオチド単位又は遊離ヌクレオチドは、テンプレート分子におけるヌクレオチド残基と相補的であってもよく、又は非相補的であってもよい。ヌクレオチド単位又は遊離ヌクレオチドは、核酸プライマーの3'末端に、核酸テンプレート分子における相補的ヌクレオチド残基の反対側である位置において結合することができる。持続時間は、結合複合体の安定性及び結合相互作用の強度を示す。持続時間は、結合複合体の開始及び／又は持続時間を観察することによって、例えば、結合複合体の標識構成要素からのシグナルを観察することによって測定され得る。例えば、標識ヌクレオチド、又は1つ以上のヌクレオチドを含む標識試薬は、結合複合体において存在し得、したがって、標識からのシグナルが、結合複合体の持続時間中に検出されることを可能にする。1つの例示的な標識は、蛍光標識である。結合複合体（例えば、三元複合体）は、ポリメラーゼ、テンプレート分子、プライマー、及び／又はヌクレオチド単位若しくはヌクレオチドのうちのいずれかの間の相互作用の解離を引き起こす条件に供されるまで依然として安定している。例えば、解離条件は、結合複合体を、洗浄剤、EDTA、及び／若しくは水のうちのいずれか1つ、又はそれらのいずれかの組み合わせと接触させることを含む。

10

【0500】

本明細書で使用される「核酸」、「ポリヌクレオチド」、及び「オリゴヌクレオチド」という用語、並びに他の関連用語は、互換的に使用され、ヌクレオチドのポリマーを指し、いずれもの特定の長さに限定されない。核酸は、組換え及び化学合形成態を含む。核酸としては、DNA分子（例えば、cDNA又はゲノムDNA）、RNA分子（例えば、mRNA）、ヌクレオチド類似体を使用して生成されたDNA又はRNAの類似体（例えば、ペプチド核酸及び天然に存在しないヌクレオチド類似体）、並びにDNA及びRNAを含有するキメラ型が挙げられる。核酸は、一本鎖又は二本鎖であり得る。核酸は、ヌクレオチドのポリマーを含み、ヌクレオチドは、自然又は非自然塩基及び／又は糖を含む。核酸は、自然発生のヌクレオシド間結合、例えば、ホスホジエステル結合を含む。核酸は、非自然ヌクレオシド間結合を含み、非自然ヌクレオシド間結合は、ホスホロチオエート、ホスホロチオレート、又はペプチド核酸（PNA）結合を含む。いくつかの態様において、核酸は、1つの種類のポリヌクレオチド、又は2つ以上の異なる種類のポリヌクレオチドの混合物を含む。

20

30

【0501】

本明細書で使用される「プライマー」という用語及び関連する用語は、DNA及び／又はRNAポリヌクレオチドテンプレートとハイブリダイズし、二本鎖分子を形成することができる天然又は合成のうちのいずれかのオリゴヌクレオチドを指す。プライマーは、いずれかの長さを有し得るが、典型的には、4つ～50個のヌクレオチドの範囲であり得る。典型的なプライマーは、5'末端及び3'末端を含む。プライマーの3'末端は、ポリメラーゼ媒介プライマー伸長反応におけるヌクレオチド重合開始部位として機能する3'OH部分を含むことができる。代替的に、プライマーの3'末端は、3'OH部分を欠くことができるか、又はポリメラーゼ媒介反応においてヌクレオチド重合を阻害する末端3'ブロック基を含むことができる。プライマーの長さに沿ういずれか1つのヌクレオチド又は2つ以上のヌクレオチドは、検出可能なレポーター部分で標識され得る。プライマーは、溶液中にあってもよく（例えば、可溶性プライマー）、又は支持体に固定化されてもよい（例えば、キャプチャプライマー）。

40

【0502】

「テンプレート核酸」、「テンプレートポリヌクレオチド」、「標的核酸」、「標的ポリヌクレオチド」、「テンプレート鎖」、及び他の変形は、相補的な核酸鎖を生成するための基礎核酸分子として機能する核酸鎖を指す。テンプレート核酸は、一本鎖若しくは二本鎖であってもよく、又はテンプレート核酸は、一本鎖若しくは二本鎖部分を有してもよい。テンプレート核酸の配列は、相補鎖の配列に部分的又は完全に相補的であり得る。テ

50

ンプレート核酸は、自然発生の供給源若しくは組換え形態から得られてもよく、又はいずれかの種類の核酸類似体を含むように化学的に合成されてもよい。テンプレート核酸は、直鎖、環状、又は他の形態であることができる。テンプレート核酸は、目的の配列としても既知である挿入配列を有する挿入領域を含むことができる。テンプレート核酸はまた、少なくとも1つのアダプター配列を含むことができる。テンプレート核酸は、目的の配列及び少なくとも1つのアダプター配列の2つの又はタンデムコピーを有するコンカテマーであることができる。挿入領域は、いずれかの形態で単離され得、形態は、染色体、ゲノム、オルガネラ（例えば、ミトコンドリア、葉緑体、又はリボソーム）、組換え分子、クローニング、増幅、cDNA、前駆体mRNA若しくはmRNAなどのRNA、オリゴヌクレオチド、新鮮凍結パラフィン包埋組織から得られる全ゲノムDNA、針生検、無細胞循環DNA、又はいずれかのタイプの核酸ライブラリの形態を含む。挿入領域は、いずれかの供給源から単離され得、供給源は、原核生物、真核生物（例えば、ヒト、植物、及び動物）、真菌などの生物、ウイルス細胞、組織、正常若しくは疾患細胞若しくは組織、血液、尿、血清、リンパ液、腫瘍、唾液、肛門及び膺分泌物、羊水試料、汗、精液、環境試料、培養試料を含む体液、又は組換え分子生物学若しくは化学合成方法を使用して調製された合成核酸分子を含む。挿入領域は、いずれかの臓器から単離され得、臓器は、頭部、頸部、脳、乳房、卵巣、子宮頸部、結腸、直腸、子宮内膜、胆嚢、腸、膀胱、前立腺、精巣、肝臓、肺、腎臓、食道、脾臓、甲状腺、下垂体、胸腺、皮膚、心臓、喉頭、又は他の臓器を含む。テンプレート核酸は、シーケンシング及び組成分析を含む核酸分析に供され得る。

10

20

【0503】

核酸分子に関して使用されるときに、「ハイブリダイズする」若しくは「ハイブリダイズすること」若しくは「ハイブリダイゼーション」という用語、又は他の関連用語は、二重鎖核酸を形成するための2つの異なる核酸間の水素結合を指す。ハイブリダイゼーションはまた、二重鎖領域を有する自己ハイブリダイゼーション分子を形成するための単一の核酸分子の2つの異なる領域間の水素結合を含む。ハイブリダイゼーションは、二重鎖二本鎖核酸、又は核酸分子内の二本鎖領域を形成するためのワトソンクリック又はフーグステイン結合を含むことができる。二本鎖核酸、又は単一の核酸の2つの異なる領域は、完全に相補的であってもよく又は部分的に相補的であってもよい。相補的核酸鎖は、それらの長さ全体にわたって互いにハイブリダイズする必要はない。相補的塩基対形成は、標準的なA-T若しくはC-G塩基対形成であってもよく、又は塩基対形成相互作用の他の形態であってもよい。二重鎖核酸は、ミスマッチ塩基対形成ヌクレオチドを含み得る。

30

【0504】

「ヌクレオチド」という用語及び関連用語は、芳香族塩基、5炭素糖（例えば、リボース又はデオキシリボース）、及び少なくとも1つのリン酸基を含む分子を指す。標準又は非標準ヌクレオチドは、この用語の使用と一致する。いくつかの態様において、リン酸は、一リン酸、二リン酸、若しくは三リン酸、又は対応するリン酸類似体を含む。いくつかの態様において、ヌクレオチドは、1、2、3、4、5、6、7、8、9、又は10個のリン酸基を含む。「ヌクレオシド」という用語は、芳香族塩基及び糖を含む分子を指す。

【0505】

ヌクレオチド（及びヌクレオシド）は、典型的には、置換された又は非置換の窒素含有親ヘテロ芳香族環を含むヘテロ環状塩基を含み、これらは、核酸において一般に見出され、自然発生の、置換された、修飾された、若しくは操作されたバリエーション、又はそれらの類似体を含む。ヌクレオチド（又はヌクレオシド）の塩基は、適切な相補的塩基とのワトソンクリック及び／又はフーグステイン水素結合を形成することが可能である。例示的な塩基としては、プリン及びピリミジン、例えば、2-アミノプリン、2,6-ジアミノプリン、アデニン（A）、エテノアデニン、N6- Δ^2 -イソペンテニルアデニン（6iA）、N6- Δ^2 -イソペンテニル-2-メチルチオアデニン（2ms6iA）、N6-メチルアデニン、グアニン（G）、イソグアニン、N2-ジメチルグアニン（dmG）、7-メチルグアニン（7mG）、2-チオピリミジン、6-チオグアニン（6sG）、ヒポ

40

50

キサンチン、及び O 6 -メチルグアニン；7-デアザプリン、例えば、7-デアザアデニン（7-デアザ-A）及び7-デアザグアニン（7-デアザ-G）；ピリミジン、例えば、シトシン（C）、5-プロピニルシトシン、イソシトシン、チミン（T）、4-チオチミン（4sT）、5, 6-ジヒドロチミン、O 4 -メチルチミン、ウラシル（U）、4-チオウラシル（4sU）、及び5, 6-ジヒドロウラシル（ジヒドロウラシル；D）；インドール、例えば、ニトロインドール及び4-メチルインドール；ピロール、例えば、ニトロピロール；ネブラリン；イノシン；ヒドロキシメチルシトシン；5-メチシトシン；塩基（Y）；並びにメチル化、グリコシル化、及びアシル化塩基部分などが挙げられるが、これらに限定されない。追加の例示的な塩基は、Fasman, 1989, in “Practical Handbook of Biochemistry and Molecular Biology”, pp. 385-394, CRC Press, Boca Raton, Flaに見出され得る。

10

【0506】

ヌクレオチド（及びヌクレオシド）は、典型的には、糖部分、例えば、炭素環式部分（Ferraro and Gotor 2000 Chem. Rev. 100: 4319-48）、非環式部分（Martinez, et al., 1999 Nucleic Acids Research 27: 1271-1274、Martinez, et al., 1997 Bioorganic & Medicinal Chemistry Letters vol. 7: 3013-3016）、及び他の糖部分（Joeng, et al., 1993 J. Med. Chem. 36: 2627-2638、Kim, et al., 1993 J. Med. Chem. 36: 30-7、Eschenmosser 1999 Science 284: 2118-2124 及び米国特許第5, 558, 991号）を含む。糖部分は、リボシル；2'-デオキシリボシル；3'-デオキシリボシル；2', 3'-ジデオキシリボシル；2', 3'-ジデヒドロジデオキシリボシル；2'-アルコキシリボシル；2'-アジドリボシル；2'-アミノリボシル；2'-フルオロリボシル；2'-メルカプトリボシル；2'-アルキルチオリボシル；3'-アルコキシリボシル；3'-アジドリボシル；3'-アミノリボシル；3'-フルオロリボシル；3'-メルカプトリボシル；3'-アルキルチオリボシル炭素環状；非環状、又は他の修飾糖を含む。

20

【0507】

いくつかの態様において、ヌクレオチドは、1、2、又は3つのリン原子の鎖を含み、鎖は、典型的には、エステル又はホスホラミド結合を介して、糖部分の5'炭素に付着している。いくつかの態様において、ヌクレオチドは、リン鎖を有する類似体であり、リン鎖において、リン原子と一緒に、介在するO、S、NH、メチレン、又はエチレンと結合している。いくつかの態様において、鎖におけるリン原子は、O、S、又はBH₃を含む置換された側鎖基を含む。いくつかの態様において、鎖は、ホスホロアミデート、ホスホロチオエート、ホスホルジチオエート、及びO-メチルホスホロアミダイト基を含む類似体で置換されたリン酸基を含む。

30

【0508】

核酸に関して使用されるときに、「伸長する」、「伸長すること」、「伸長」という用語及び他のバリエーションは、核酸分子内への1つ以上のヌクレオチドの組み込みを指す。ヌクレオチドの組み込みは、核酸鎖の末端3'OH末端への1つ以上のヌクレオチドの重合を含み、核酸鎖の伸長をもたらす。ヌクレオチド組み込みは、自然ヌクレオチド及び／又はヌクレオチド類似体で行われ得る。典型的には、必ずしもそうではないが、ヌクレオチド組み込みは、テンプレートに依存的な様式で発生する。核酸分子を伸長させるいずれかの好適な方法が使用され得、好適な方法は、DNAポリメラーゼ又はRNAポリメラーゼによって触媒されるプライマー伸長を含む。

40

【0509】

「レポーター部分（reporter moiety）」、「レポーター部分（reporter moieties）」という用語又は関連用語は、検出可能なシグナルを生成するか、又はそれを生成することを引き起こす化合物を指す。レポーター部分は、しば

50

しば、「標識」と称される。いずれかの好適なレポーター部分が使用され得、好適なレポーター部分は、ルミネセンス、光ルミネセンス、エレクトロルミネセンス、生物ルミネセンス、化学ルミネセンス、蛍光、リン光、発色団、放射性同位体、電気化学、質量分析、ラマン、ハプテン、親和性タグ、原子、又は酵素を含む。レポーター部分は、化学的又は物理的变化（例えば、熱、光、電気、pH、塩濃度、酵素活性、又は近接事象）からもたらされる検出可能なシグナルを生成する。近接事象は、2つのレポーター部分が、互いに接近するか、又は互いに会合するか、又は互いに結合することを含む。レポーター部分を、各々が、励起放射線を吸収し、かつ／又は蛍光を、他のレポーター部分から区別可能な波長で発光して、同じ反応又は異なる反応における異なるレポーター部分の存在を監視することを可能にするように選択することは、当業者に周知である。スペクトル的に異なる発光プロファイルを有するか、又は最小のオーバーラップスペクトル発光プロファイルを有する2つ以上の異なるレポーター部分が選択され得る。レポーター部分は、ヌクレオチド、ヌクレオシド、核酸、酵素（例えば、ポリメラーゼ又は逆転写酵素）、又は支持体（例えば、表面）に結合（例えば、動作可能に結合）し得る。

10

20

30

40

50

【0510】

レポーター部分（又は標識）は、蛍光標識又はフルオロフォアを含む。蛍光標識又はフルオロフォアとして機能し得る例示的な蛍光部分としては、フルオレセイン及びフルオレセイン誘導体、例えば、カルボキシフルオレセイン、テトラクロロフルオレセイン、ヘキサクロロフルオレセイン、カルボキシナフトフルオレセイン、フルオレセインイソチオシアネート、NHS-フルオレセイン、ヨードアセトアミドフルオレセイン、フルオレセインマレイミド、SAMSA-フルオレセイン、フルオレセインチオセミカルバジド、カルボヒドラジノメチルチオアセチル-アミノフルオレセイン、ローダミン及びローダミン誘導体、例えば、TRITC、TMR、Lissamineローダミン、Texas Red、ローダミンB、ローダミン6G、ローダミン10、NHS-ローダミン、TMR-ヨードアセトアミド、LissamineローダミンBスルホニルクロリド、LissamineローダミンBスルホニルヒドラジン、Texas Redスルホニルクロリド、Texas Redヒドラジド、クマリン及びクマリン誘導体、例えば、AMCA、AMCA-NHS、AMCA-スルホ-NHS、AMCA-HPDP、DCIA、AMCE-ヒドラジド、BODIPY及び誘導体、例えば、BODIPY FL C3-SE、BODIPY 530/550 C3、BODIPY 530/550 C3-SE、BODIPY 530/550 C3ヒドラジド、BODIPY 493/503 C3ヒドラジド、BODIPY FL C3ヒドラジド、BODIPY FL IA、BODIPY 530/551 IA、Br-BODIPY 493/503、Cascade Blue及び誘導体、例えば、Cascade Blueアセチルアジド、Cascade Blueカダベリン、Cascade Blueエチレンジアミン、Cascade Blueヒドラジド、Lucifer Yellow及び誘導体、例えば、Lucifer Yellowヨードアセトアミド、Lucifer Yellow CH、シアニン及び誘導体、例えば、インドリウムベースのシアニン色素、ベンゾ-インドリウムベースのシアニン色素、ピリジウムベースのシアニン色素、チオゾリウムベースのシアニン色素、キノリニウムベースのシアニン色素、イミダゾリウムベースのシアニン色素、Cy3、Cy5、ランタニドキレート及び誘導体、例えば、BCPDA、TBP、TMT、BHHCT、BCOT、ユーロピウムキレート、テルビウムキレート、Alexa Fluor色素、DyLight色素、Atto色素、LightCycler Red色素、CAL Fluor色素、JOE及びその誘導体、Oregon Green色素、WellRED色素、IRD色素、フィコエリトリン及びフィコビル色素、マラカイトグリーン、スチルベン、DEG色素、NR色素、近赤外色素、並びに当該技術分野で既知の他のもの、例えば、Haugland, Molecular Probes Handbook, (Eugene, Oreg.) 6th Edition, Lakowicz, Principles of Fluorescence Spectroscopy, 2nd Ed., Plenum Press New York (1999)、若しくはHerman

son, Bioconjugate Techniques, 2nd Editionに記載されているもの又はそれらの誘導体、あるいはそれらのいずれかの組み合わせが挙げられるが、これらに限定されない。シアニン色素は、スルホン化又は非スルホン化形態のいずれかで存在し得、2つの窒素原子間のポリメチン架橋によって分離された2つのインドレニン、ベンゾインドリウム、ピリジウム、チオゾリウム、及び／又はキノリニウム基からなる。市販のシアニンフルオロフォアは、例えば、Cy 3（これは、1-[6-(2, 5-ジオキソピロリジン-1-イルオキシ)-6-オキソヘキシル]-2-(3-{1-[6-(2, 5-ジオキソピロリジン-1-イルオキシ)-6-オキソヘキシル]-3, 3-ジメチル-1, 3-ジヒドロ-2H-インドール-2-イリデン}プロプ-1-エン-1-イル)-3, 3-ジメチル-3H-インドリウム、又は1-[6-(2, 5-ジオキソピロリジン-1-イルオキシ)-6-オキソヘキシル]-2-(3-{1-[6-(2, 5-ジオキソピロリジン-1-イルオキシ)-6-オキソヘキシル]-3, 3-ジメチル-5-スルホ-1, 3-ジヒドロ-2H-インドール-2-イリデン}プロプ-1-エン-1-イル)-3, 3-ジメチル-3H-インドリウム-5-スルホネートを含み得る)、Cy 5（これは、1-(6-(2, 5-ジオキソピロリジン-1-イル)オキシ)-6-オキソヘキシル)-2-(1E, 3E)-5-(E)-1-(6-(2, 5-ジオキソピロリジン-1-イル)オキシ)-6-オキソヘキシル)-3, 3-ジメチル-5-インドリン-2-イリデン)ペンタ-1, 3-ジエン-1-イル)-3, 3-ジメチル-3H-インドール-1-イウム、又は1-(6-(2, 5-ジオキソピロリジン-1-イル)オキシ)-6-オキソヘキシル)-2-(1E, 3E)-5-(E)-1-(6-(2, 5-ジオキソピロリジン-1-イル)オキシ)-6-オキソヘキシル)-3, 3-ジメチル-5-スルホインドリン-2-イリデン)ペンタ-1, 3-ジエン-1-イル)-3, 3-ジメチル-3H-インドール-1-イウム-5-スルホネートを含み得る)、及びCy 7（1-(5-カルボキシペンチル)-2-[(1E, 3E, 5E, 7Z)-7-(1-エチル-1, 3-ジヒドロ-2H-インドール-2-イリデン)ヘプタ-1, 3, 5-トリエン-1-イル]-3H-インドリウム、又は1-(5-カルボキシペンチル)-2-[(1E, 3E, 5E, 7Z)-7-(1-エチル-5-スルホ-1, 3-ジヒドロ-2H-インドール-2-イリデン)ヘプタ-1, 3, 5-トリエン-1-イル]-3H-インドリウム-5-スルホネートを含み得る)を含み、「Cy」は、「シアニン」を表し、最初の数字は、2つのインドレニン基の間の炭素原子の数を特定する。インドレニンではなくオキサゾール誘導体であるCy 2、並びにベンゾ誘導体化Cy 3. 5、Cy 5. 5、及びCy 7. 5は、この規則に対する例外である。

10

20

30

【0511】

いくつかの態様において、レポーター部分は、FRET対であることができ、これにより、複数の分類が、単一の励起及びイメージングステップ下で実行され得る。本明細書で使用される場合、FRETは、励起交換(Forster)移動、又は電子交換(Dexter)移動を含み得る。

【0512】

「連結した(linked)」、「結合した(joined)」、「結合した(attached)」という用語及びそれらのバリエーションは、特定の手順での使用に耐えるのに十分な安定性を有する化合物又は分子の任意の組み合わせ間の任意のタイプの融合、結合、接着、又は会合を含む。手順は、ヌクレオチド過性結合、ヌクレオチドの取り込み、デプロッキング、洗浄、除去、流動、検出、イメージング、及び／又は特定を含むことができるが、これらに限定されない。そのような結合は、例えば、共有結合、イオン、水素、双極子-双極子、親水性、疎水性、又は親和性結合、ファンデルワールス力を含む結合又は会合、及び機械的結合などを含むことができる。いくつかの態様において、そのような結合は、分子内で、例えば、一本鎖又は二本鎖直鎖核酸分子の末端と一緒に結合させて環状分子を形成するように発生する。いくつかの態様において、そのような結合は、異なる分子の組み合わせ間、又は分子と非分子との間で発生することができ、核酸分子と固体表面との間の結合、タンパク質と検出可能なレポーター部分との間の結合、及びヌクレオ

40

50

チドと検出可能なレポーター部分との間の結合などを含むが、これらに限定されない。結合のいくつかの例は、例えば、Hermanson, G., "Bioconjugate Techniques", Second Edition (2008)、Aslam, M., Dent, A., "Bioconjugation: Protein Coupling Techniques for the Biomedical Sciences", London: Macmillan (1998)、Aslam, M., Dent, A., "Bioconjugation: Protein Coupling Techniques for the Biomedical Sciences", London: Macmillan (1998)に見出され得る。

【0513】

本明細書で使用される場合、「動作可能に結合した」及び「動作可能に連結された」という用語、又は関連用語は、構成要素の並列を指す。並列した構成要素は、共有結合的に一緒に結合し得る。例えば、2つの核酸構成要素は、酵素的に一緒にライゲーションされ得、2つの構成要素と一緒に連結する結合は、ホスホジエステル結合を含む。第1及び第2の核酸構成要素は、一緒に結合することができ、第1の核酸構成要素は、機能を第2の核酸構成要素に付与することができる。例えば、プライマー結合配列と目的の配列との間の結合は、プライマーに結合することができる一部分を有する核酸ライブラリ分子を形成する。別の例において、導入遺伝子（例えば、目的のポリペプチド又は核酸配列をコードする核酸）は、ベクターにライゲーションされ得、結合は、ベクター内に含有された導入遺伝子配列の発現又は機能を可能にする。いくつかの態様において、導入遺伝子は、導入遺伝子の発現に影響する宿主細胞調節配列（例えば、プロモーター配列）に動作可能に結合している。いくつかの態様において、ベクターは、少なくとも1つの宿主細胞調節配列を含み、少なくとも1つの宿主細胞調節配列は、プロモーター配列、エンハンサー、転写及び／又は翻訳開始配列、転写及び／又は翻訳終結配列、並びにポリペプチド分泌シグナル配列などを含む。いくつかの態様において、宿主細胞調節配列は、導入遺伝子のレベル、タイミング、及び／又は位置の発現を制御する。

【0514】

「アダプター」という用語及び関連用語は、標的ポリヌクレオチドに動作可能に結合し得る（付加され得る）オリゴヌクレオチドを指し、アダプターは、機能を共連結されたアダプター-標的分子に付与する。アダプターは、DNA、RNA、キメラDNA/RNA、又はそれらの類似体を含む。アダプターは、少なくとも1つのリボヌクレオシド残基を含むことができる。アダプターは、一本鎖若しくは二本鎖であってもよく、又は一本鎖及び／若しくは二本鎖部分を有してもよい。アダプターは、直鎖、ステムループ、ヘアピン、又はY字形状の形態であるように構成され得る。アダプターは、4つ〜100個のヌクレオチド以上を含むいずれかの長さであることができる。アダプターは、平滑末端、オーバーハング末端、又は両方の組み合わせを有することができる。オーバーハング末端は、5'オーバーハング及び3'オーバーハング末端を含む。一本鎖アダプターの5'末端、又は二本鎖アダプターの一方向の鎖は、5'リン酸基を有してもよく、又は5'リン酸基を欠いてもよい。アダプターは、標的ポリヌクレオチドにハイブリダイズしない5'テールを含んでもよく（例えば、テール付きアダプター）、又はアダプターは、テールなしであってもよい。アダプターは、プライマー、例えば、増幅プライマー、シーケンシングプライマー、又はキャプチャプライマー（例えば、可溶性又は固定化されたキャプチャプライマー）の少なくとも一部分と相補的である配列を含むことができる。アダプターは、ランダム配列又は縮重配列を含むことができる。アダプターは、少なくとも1つのイノシン残基を含むことができる。アダプターは、少なくとも1つのホスホロチオエート、ホスホロチオレート、及び／又はホスホロアミデート結合を含むことができる。アダプターは、バーコード配列を含むことができ、バーコード配列は、マルチプレックスアッセイにおいて、異なる試料供給源からのポリヌクレオチド（例えば、挿入配列）を区別するために使用され得る。アダプターは、固有の識別配列（例えば、固有の分子インデックス、UMI、又は固有の分子タグ）を含むことができ、固有の識別配列は、アダプターが付加された核酸分

10

20

30

40

50

子を固有に特定するために使用され得る。いくつかの態様において、固有の識別配列は、エラー訂正及び精度を増加させ、偽陽性バリエーションコールの率を低減し、かつ／又はバリエーション検出の感度を増加させるために使用され得る。アダプターは、少なくとも1つの制限酵素認識配列を含むことができ、少なくとも1つの制限酵素認識配列は、I型、II型、III型、IV型、Hs型、若しくはIIB型からなる群から選択されるいずれか1つ、又は2つ以上のいずれかの組み合わせを含む。

【0515】

「ユニバーサル配列」、「ユニバーサルアダプター配列」という用語、及び関連用語は、2つ以上のポリヌクレオチド分子間で共通する、核酸分子における配列を指す。例えば、同じユニバーサル配列を有するアダプターは、複数のポリヌクレオチドに連結され得、これにより、共連結された分子の集団は、同じユニバーサルアダプター配列を保有する。ユニバーサルアダプター配列の例としては、増幅プライマー配列、シーケンシングプライマー配列、又はキャプチャプライマー配列（例えば、可溶性又は支持体に固定化されたキャプチャプライマー）が挙げられる。

【0516】

本開示のいくつかの態様は、支持体に固定化された複数の（例えば、2つ以上の）核酸テンプレートを提供する。いくつかの態様において、固定化された複数の核酸テンプレートは、同じ配列を有するか、又は異なる配列を有する。いくつかの態様において、複数の核酸テンプレートにおける個々の核酸テンプレート分子は、支持体上の異なる部位に固定化されている。いくつかの態様において、複数の核酸テンプレートにおける2つ以上の個々の核酸テンプレート分子は、支持体上の部位に固定化されている。いくつかの態様において、支持体は、アレイに配置された複数の部位を含む。「アレイ」という用語は、支持体上の所定の位置において位置する複数の部位を、部位のアレイを形成するように含む支持体を指す。部位は、分散し得、インターステイシャル領域によって分離され得る。いくつかの態様において、支持体上の所定の部位は、一次元で行又は列に配置されてもよく、又は二次元で行及び列に配置されてもよい。いくつかの態様において、複数の所定の部位は、支持体上に、組織化された様式で配置されている。いくつかの態様において、複数の所定の部位は、いずれかの組織化されたパターンで配置されており、パターンは、直線、六角形パターン、格子パターン、反射対称性を有するパターン、又は回転対称性を有するパターンなどを含む。異なる対の部位間のピッチは、同じであってもよく又は変動してもよい。いくつかの態様において、支持体は、核酸テンプレートアレイを形成するために、 1 mm^2 当たり約102～1015個の部位の表面密度で複数の部位に固定化された核酸テンプレート分子を有し得る。いくつかの態様において、支持体は、少なくとも102個の部位、少なくとも103個の部位、少なくとも104個の部位、少なくとも105個の部位、少なくとも106個の部位、少なくとも107個の部位、少なくとも108個の部位、少なくとも109個の部位、少なくとも1010個の部位、少なくとも1011個の部位、少なくとも1012個の部位、少なくとも1013個の部位、少なくとも1014個の部位、又は少なくとも1015個の部位以上を含み、部位は、支持体上の所定の位置において位置する。いくつかの態様において、支持体上の複数の所定の部位（例えば、102～1015個の部位以上）は、核酸テンプレートが固定化されており、核酸テンプレートアレイを形成する。いくつかの態様において、固定化された表面キャプチャプライマーへのハイブリダイゼーションによって複数の所定の部位に固定化される核酸テンプレート、又は核酸テンプレートは、表面キャプチャプライマーに共有結合している。いくつかの態様において、複数の所定の部位において固定化されており、例えば、102～1015個の部位以上において固定化されている核酸テンプレート。いくつかの態様において、支持体上の複数の部位に固定化された核酸テンプレートは、線形若しくは環状核酸テンプレート分子、又は線形分子と環状分子の両方の混合物を含む。いくつかの態様において、固定化された核酸テンプレートは、クローン的に増幅されて、固定化された核酸ポロニーを、複数の所定の部位において生成する。いくつかの態様において、個々の固定化核酸テンプレート分子は、目的の標的配列の1つのコピーを含むか、又は目的の標的配列の2つ

10

20

30

40

50

以上のタンデムコピーを有するコンカテマーを含む。

【0517】

いくつかの態様において、支持体上のランダムな位置において位置する複数の部位を含む支持体は、ランダムに位置する部位を上にも有する支持体と本明細書で称される。支持体上のランダムに位置する部位の位置は、所定の位置でない。複数のランダムに位置する部位は、支持体上に、不整列にかつ／又は予測不可能な様式で配置されている。いくつかの態様において、支持体は、少なくとも102個の部位、少なくとも103個の部位、少なくとも104個の部位、少なくとも105個の部位、少なくとも106個の部位、少なくとも107個の部位、少なくとも108個の部位、少なくとも109個の部位、少なくとも1010個の部位、少なくとも1011個の部位、少なくとも1012個の部位、少なくとも1013個の部位、少なくとも1014個の部位、又は少なくとも1015個の部位以上を含み、部位は、支持体上にランダムに位置する。いくつかの態様において、支持体上の複数のランダムに位置する部位（例えば、102～1015個の部位以上）は、核酸テンプレートが固定化されており、核酸テンプレートが固定化された支持体を形成する。いくつかの態様において、固定化された表面キャプチャプライマーへのハイブリダイゼーションによって複数のランダムに位置する部位において固定化された核酸テンプレート、又は核酸テンプレートは、表面キャプチャプライマーに共有結合的に付着している。いくつかの態様において、複数のランダムに位置する部位において固定化されており、例えば、102～1015個の部位以上において固定化されている核酸テンプレート。いくつかの態様において、支持体上の複数の部位に固定化された核酸テンプレートは、線形若しくは環状核酸テンプレート分子、又は線形分子と環状分子の両方の混合物を含む。いくつかの態様において、固定化された核酸テンプレートは、クローン的に増幅されて、固定化された核酸ポロニーを、複数のランダムに位置する部位において生成する。いくつかの態様において、個々の固定化核酸テンプレート分子は、目的の標的配列の1つのコピーを含むか、又は目的の標的配列の2つ以上のタンデムコピーを有するコンカテマーを含む。

【0518】

いくつかの態様において、支持体上の所定の又はランダムな部位に固定化された核酸テンプレート分子に関して、支持体上の複数の固定化された核酸テンプレート分子は互いに流体連通して、試薬（例えば、ポリメラーゼ、多価分子、ヌクレオチド、二価カチオン及び／又は緩衝液などを含む酵素）の溶液を支持体上に流すことを可能にし、その結果、支持体上の複数の固定化された核酸テンプレート分子を超並列方法で試薬と反応させることができる。いくつかの態様において、複数の固定化された核酸テンプレート分子の流体連通を使用して、ヌクレオチド結合アッセイを行い、及び／又は複数の固定化された核酸テンプレート分子上でヌクレオチド重合反応（例えば、プライマー伸長若しくはシーケンシング）を行い、並びに超並列シーケンシングのための検出及びイメージングを行うことができる。いくつかの態様において、「固定化された」という用語及び関連用語は、核酸分子又は酵素が共有結合又は非共有結合相互作用を介して支持体に直接結合されるか、又は核酸分子又は酵素が支持体上のコーティングに結合される、所定の位置又はランダムな配置で支持体に結合される核酸分子又は酵素（例えば、ポリメラーゼ）を指す。

【0519】

低結合表面コーティングに関して使用されるときに、多層表面コーティングの1つ以上の層は、分岐ポリマーを含んでもよく、又は直鎖であってもよい。好適な分岐ポリマーの例としては、分岐PEG、分岐ポリ（ビニルアルコール）（分岐PVA）、分岐ポリ（ビニルピリジン）、分岐ポリ（ビニルピロリドン）（分岐PVP）、分岐、ポリ（アクリル酸）（分岐PAA）、分岐ポリアクリルアミド、分岐ポリ（N-イソプロピルアクリルアミド）（分岐PNIPAM）、分岐ポリ（メチルメタクリレート）（分岐PMA）、分岐ポリ（2-ヒドロキシルエチルメタクリレート）（分岐HEMA）、分岐ポリ（オリゴ（エチレングリコール）メチルエーテルメタクリレート）（分岐POEGMA）、分岐ポリグルタミン酸（分岐PGA）、分岐ポリーリジン、分岐ポリーグルコシド、及びデキストランが挙げられるが、これらに限定されない。

【0520】

いくつかの態様において、本明細書に開示される多層表面のうちのいずれかの1つ以上の層を作製するために使用される分岐ポリマーは、少なくとも4つの分岐、少なくとも5つの分岐、少なくとも6つの分岐、少なくとも7つの分岐、少なくとも8つの分岐、少なくとも9つの分岐、少なくとも10個の分岐、少なくとも12個の分岐、少なくとも14個の分岐、少なくとも16個の分岐、少なくとも18個の分岐、少なくとも20個の分岐、少なくとも22個の分岐、少なくとも24個の分岐、少なくとも26個の分岐、少なくとも28個の分岐、少なくとも30個の分岐、少なくとも32個の分岐、少なくとも34個の分岐、少なくとも36個の分岐、少なくとも38個の分岐、又は少なくとも40個の分岐を含み得る。

10

【0521】

本明細書に開示される多層表面のうちのいずれかの1つ以上の層を作製するために使用される直鎖、分岐、又は多分岐ポリマーは、少なくとも500、少なくとも1,000、少なくとも2,000、少なくとも3,000、少なくとも4,000、少なくとも5,000、少なくとも10,000、少なくとも15,000、少なくとも20,000、少なくとも25,000、少なくとも30,000、少なくとも35,000、少なくとも40,000、少なくとも45,000、又は少なくとも50,000ダルトンの分子量を有し得る。

【0522】

いくつかの態様において、例えば、多層表面の少なくとも1つの層が、分岐ポリマーを含む場合、堆積させる層の分岐ポリマー分子と前の層の分子との間の共有結合の数は、1分子当たり約1つの共有結合～1分子当たり約32個の共有結合の範囲であり得る。いくつかの態様において、新しい層の分岐ポリマー分子と前の層の分子との間の共有結合の数は、1分子当たり少なくとも1つ、少なくとも2つ、少なくとも3つ、少なくとも4つ、少なくとも5つ、少なくとも6つ、少なくとも7つ、少なくとも8つ、少なくとも9つ、少なくとも10個、少なくとも12個、少なくとも14個、少なくとも16個、少なくとも18個、少なくとも20個、少なくとも22個、少なくとも24個、少なくとも26個、少なくとも28個、少なくとも30個、又は少なくとも32個の共有結合であり得る。

20

【0523】

表面への材料層のカップリング後に残存するいずれもの反応性官能基が、任意選択で、高収率カップリング化学を使用して小さい不活性分子をカップリングさせることによってブロックされ得る。例えば、アミンカップリング化学が、新しい材料層を前のものに付着させるために使用される場合、いずれもの残留アミン基が、その後、グリシンなどの小さいアミノ酸とカップリングすることによってアセチル化又は不活性化され得る。

30

【0524】

表面上に堆積した低非特異的結合材料、例えば、親水性ポリマー材料の層の数は、1～約10の範囲であり得る。いくつかの態様において、層の数は、少なくとも1、少なくとも2、少なくとも3、少なくとも4、少なくとも5、少なくとも6、少なくとも7、少なくとも8、少なくとも9、又は少なくとも10である。いくつかの態様において、層の数は、最大10、最大9、最大8、最大7、最大6、最大5、最大4、最大3、最大2、又は最大1であり得る。この段落に記載される下限及び上限値のうちのいずれかが、本開示内に含まれる範囲を形成するように組み合わせられ得、例えば、いくつかの態様において、層の数は、約2～約4の範囲であり得る。いくつかの態様において、層の全てが、同じ材料を含み得る。いくつかの態様において、各々の層は、異なる材料を含み得る。いくつかの態様において、複数の層は、複数の材料を含み得る。いくつかの態様において、少なくとも1つの層は、分岐ポリマーを含み得る。いくつかの態様において、層の全てが、分岐ポリマーを含み得る。

40

【0525】

低非特異的結合材料の1つ以上の層は、いくつかの場合において、極性プロトン性溶媒、極性若しくは極性非プロトン性溶媒、非極性溶媒、又はそれらのいずれかの組み合わせ

50

を使用して、基材表面上に堆積し得、かつ／又はそれにコンジュゲートし得る。いくつかの態様において、層堆積及び／又はカップリングのために使用される溶媒は、アルコール（例えば、メタノール、エタノール、プロパノールなど）、別の有機溶媒（例えば、アセトニトリル、ジメチルスルホキシド（DMSO）、ジメチルホルムアミド（DMF）など）、水、緩衝水溶液（例えば、リン酸緩衝液、リン酸緩衝生理食塩水、3-（N-モルホリノ）プロパンスルホン酸（MOPS）など）、又はそれらのいずれかの組み合わせを含み得る。いくつかの態様において、使用される溶媒混合物の有機構成要素は、総計の少なくとも1%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、70%、75%、80%、85%、90%、95%、98%、又は99%を含み得、バランスは、水又は緩衝水溶液からなされる。いくつかの態様において、使用される溶媒混合物の水性構成要素は、総計の少なくとも1%、5%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、70%、75%、80%、85%、90%、95%、98%、又は99%を含み得、バランスは、有機溶媒からなされる。使用される溶媒混合物のpHは、6未満、約6、6.5、7、7.5、8、8.5、9、又はpH9超であり得る。

10

20

30

40

50

【0526】

「分岐ポリマー」という用語及び関連用語は、ヌクレオチドなどの生物学的に活性な分子にコンジュゲートするのを助長する複数の官能基を有するポリマーを指し、官能基は、ポリマーの側鎖上にあってもよく、又はポリマーの中央コア若しくは中央骨格に直接付着してもよい。分岐ポリマーは、コンジュゲーションのために骨格から解離する1つ以上の官能基を有する直鎖骨格を有することができる。分岐ポリマーはまた、1つ以上の側鎖を有するポリマーあることができ、側鎖は、コンジュゲーションに好適である部位を有する。官能基の例としては、ヒドロキシル、エステル、アミン、カーボネート、アセタール、アルデヒド、アルデヒド水和物、アルケニル、アクリレート、メタクリレート、アクリルアミド、活性スルホン、ヒドラジド、チオール、アルカン酸、酸ハロゲン化物、イソシアネート、イソチオシアネート、マレイミド、ビニルスルホン、ジチオピリジン、ビニルピリジン、ヨードアセトアミド、エポキシド、グリオキサール、ジオン、メシレート、トシレート、及びトレシレートが挙げられるが、これらに限定される。

【0527】

本明細書で使用される場合、「クローン的に増幅された」という用語及びそのバリエーションは、溶液中又は支持体上のいずれかの1つ以上の増幅反応に供された核酸テンプレート分子を指す。溶液中で増幅されたテンプレート分子の場合、得られたアンプリコンは、支持体上に分配される。増幅の前に、テンプレート分子は、目的の配列及び少なくとも1つのユニバーサルアダプター配列を含む。いくつかの態様において、クローン増幅が、ポリメラーゼ連鎖反応（PCR）、多重置換増幅（MDA）、転写媒介増幅（TMA）、核酸配列ベース増幅（NASBA）、鎖置換増幅（SDA）、リアルタイムSDA、ブリッジ増幅、等温ブリッジ増幅、ローリングサークル増幅（RCA）、円-円増幅、ヘリカーゼ依存性増幅、リコンビナーゼ依存性増幅、一本鎖結合（SSB）タンパク質依存性増幅、又はそれらの任意の組み合わせの使用を含む、当該方法を提供する。

【0528】

本明細書で使用される場合、「シーケンシング」という用語及びそのバリエーションは、典型的には、核酸テンプレート分子内の少なくともいくつかのヌクレオチド（それらの核酸塩基成分を含む）の同一性を決定することによって、核酸鎖から配列情報を取得することを含む。いくつかの態様において、核酸分子の所与の領域を「シーケンシング」は、シーケンシングされる領域内の各ヌクレオチド及び全てのヌクレオチドを識別する／読み取ることを含むが、いくつかの態様において、「シーケンシング」は、領域内の一部のヌクレオチドのみの同一性が判定される一方で、一部のヌクレオチドの同一性がはっきりしないままであるか、又は誤って判定される方法を含む。任意の好適なシーケンシング方法が、使用され得る。例示的な態様において、シーケンシングは、無標識又はイオンベースのシーケンシング方法を含むことができる。いくつかの態様において、シーケンシングは、標

識又は染料含有ヌクレオチド又は蛍光ベースのヌクレオチドシーケンシング方法を含むことができる。いくつかの態様において、シーケンシングは、ポロニーベースのシーケンシング又はブリッジシーケンシング方法を含むことができる。いくつかの態様において、シーケンシングには、合成による配列、ハイブリダイゼーションによる配列、又は結合による配列の手順を用いる、超並列したシーケンシングプラットフォームを含む。超並列した合成による配列の手順の例としては、ポロニーシーケンシング、パイロシーケンシング（例えば、454 Life Sciencesから、米国特許第7,211,390号、同第7,244,559号、及び同第7,264,929号）、鎖ターミネーターシーケンシング（例えば、Illuminaから、米国特許第7,566,537号、Bentley 2006 Current Opinion Genetics and Development 16:545-552、及びBentley, et al., 2008 Nature 456:53-59、イオン感受性シーケンシング（例えば、Ion Torrentから）、プローブアンカーライゲーションシーケンシング（例えば、Complete Genomics）、DNAナノボールシーケンシング、ナノ細孔DNAシーケンシングが挙げられる。単一分子シーケンシングの例としては、Heliscope単一分子シーケンシング、及びPacific Biosciences（Levene, et al., 2003 Science 299(5607):682-686、Eid, et al., 2009 Science 323(5910):133-138、米国特許第7,170,050号、第7,302,146号、及び第7,405,281号を参照されたい）からの単一分子リアルタイム（SMRT）シーケンシングが挙げられる。ハイブリダイゼーションによる配列の例としては、SOLIDシーケンシング（例えば、Life Technologiesから、WO2006/084132）が挙げられる。結合によるシーケンシングの例としては、Omniomeシーケンシング（例えば、米国特許第10,246,744号）が挙げられる。

10

20

30

40

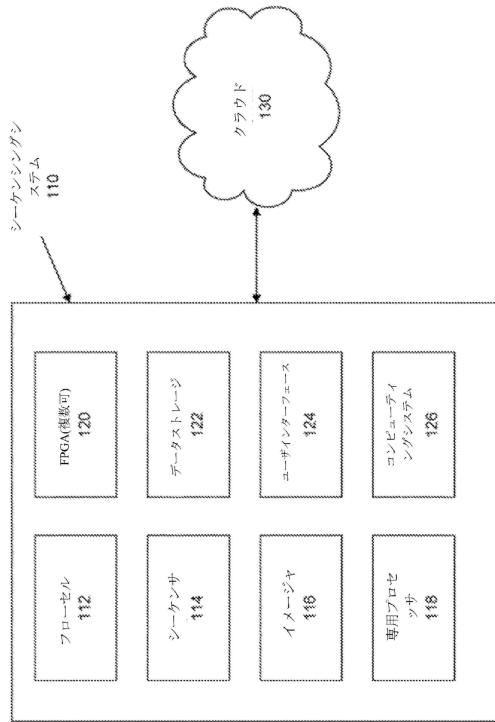
50

【0529】

本開示のいくつかの態様が本明細書に示され、記載されているが、そのような態様が例としてのみ提供されることは、当業者には明白であろう。本開示の幅及び範囲は、上記の例示的な実施形態のいずれかによって制限されるべきではなく、以下の態様及びそれらの均等物に従ってのみ定義されるべきである。本開示から逸脱することなく、多数の変形、変更、及び置換がここで当業者に想起されるであろう。本明細書に記載される本開示の態様に対する様々な代替案が、本開示の実施において用いられ得ることは理解されるべきである。以下の特許請求の範囲は、本開示の範囲を定義し、これらの特許請求の範囲及びそれらの同等物の範囲内の方法及び構造が、それらによって包含されることが意図される。

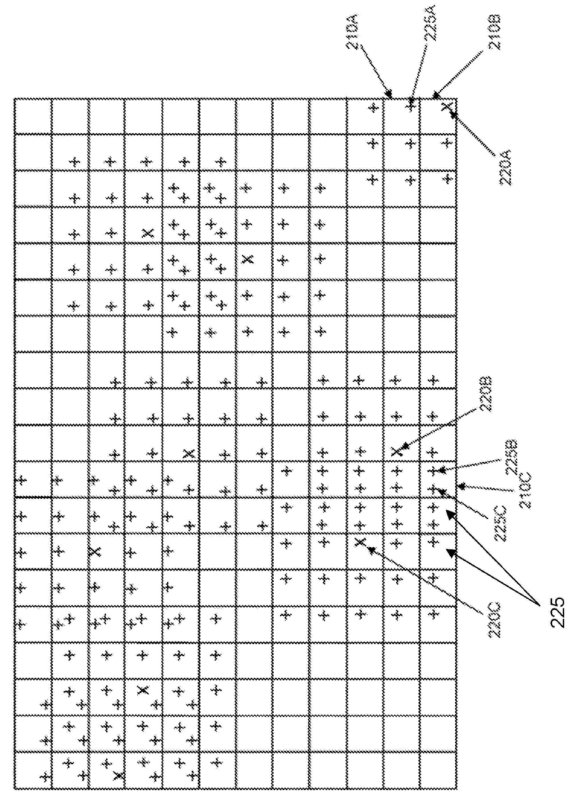
【図面】
【図 1】

100



【図 2】

200

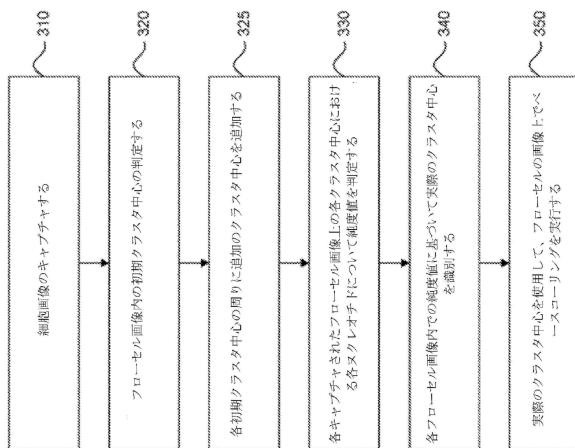


10

20

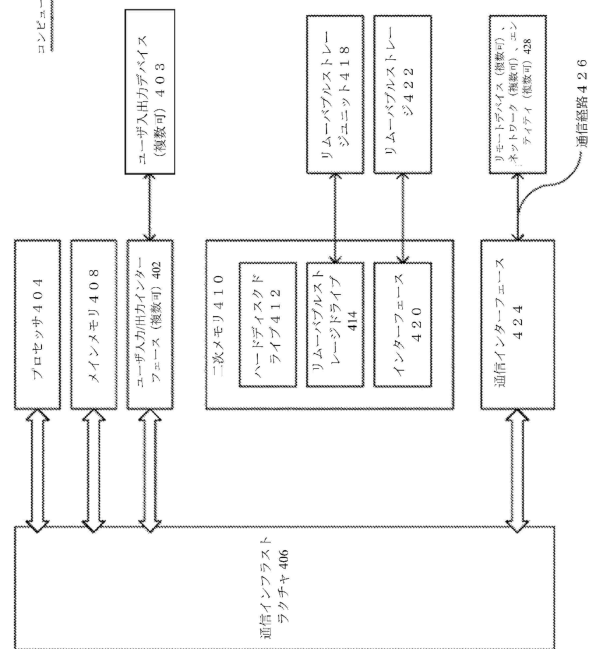
【図 3】

300



【図 4】

コンピュータシステム 400

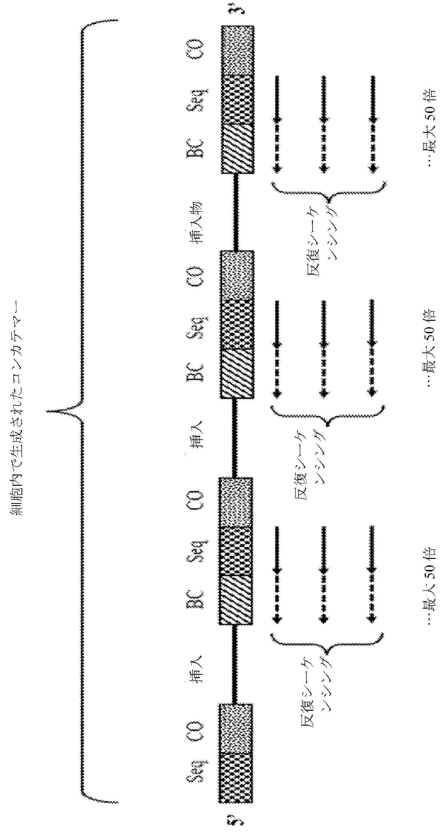


30

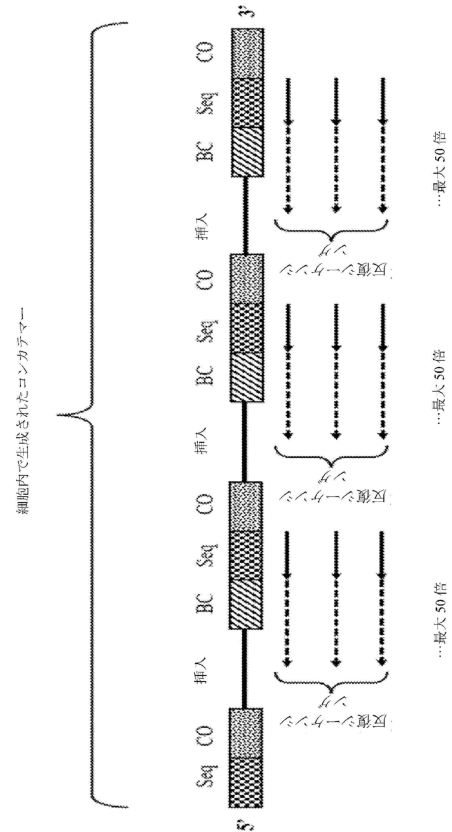
40

50

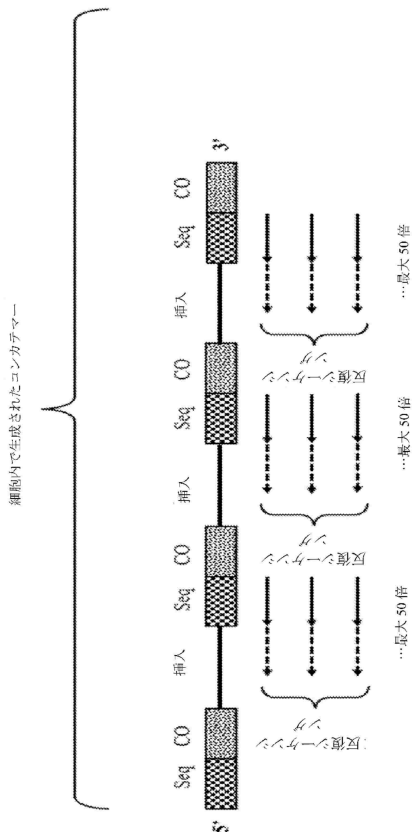
【図 9】



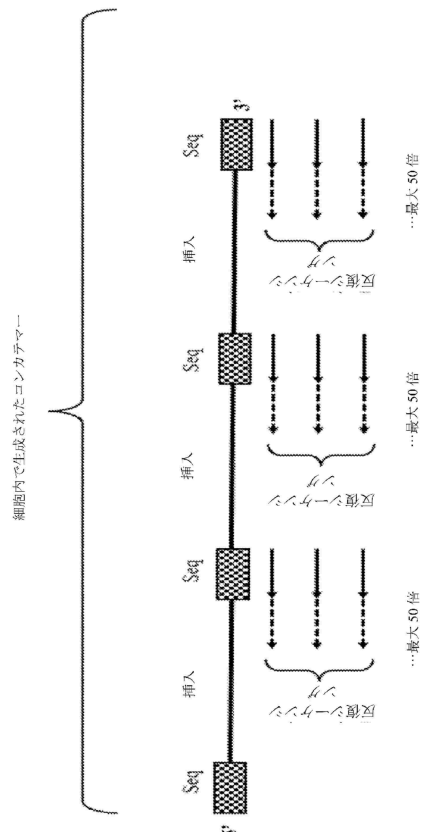
【図 10】



【図 11】



【図 12】



10

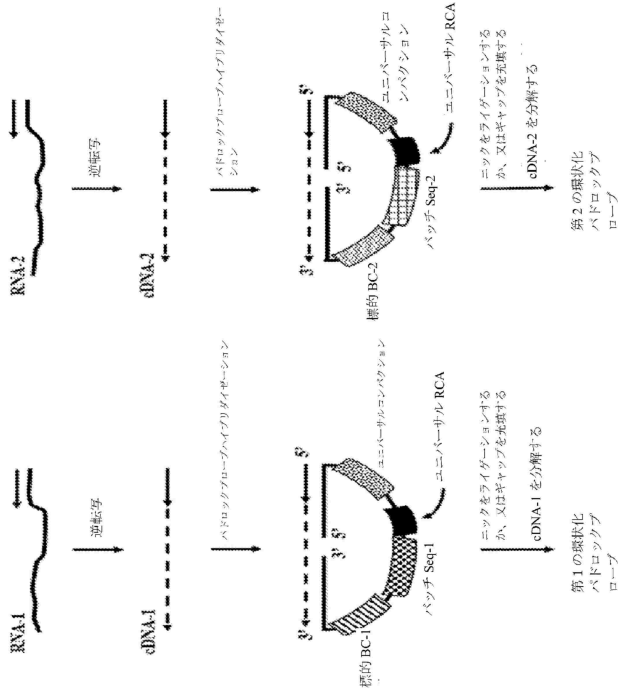
20

30

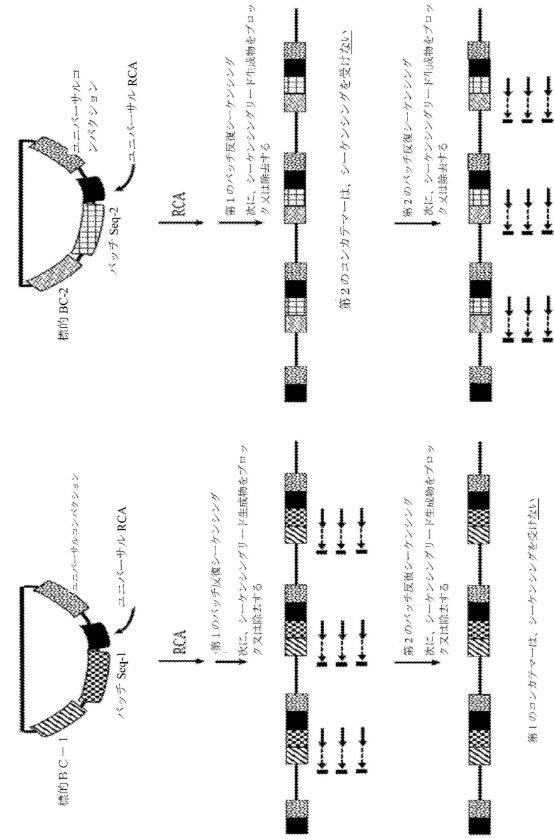
40

50

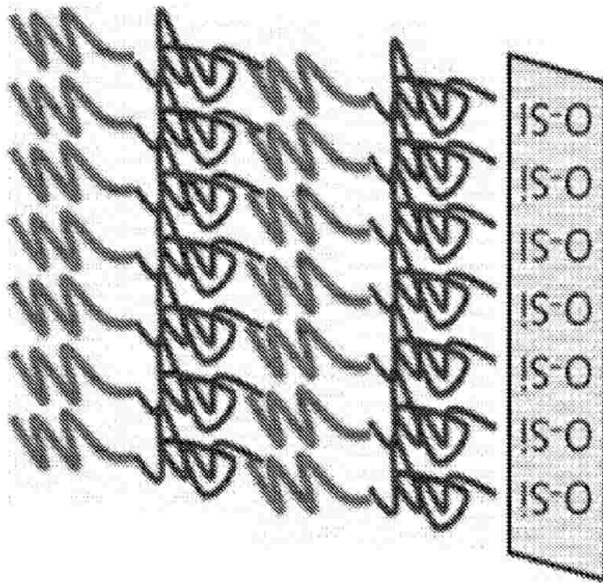
【図 13】



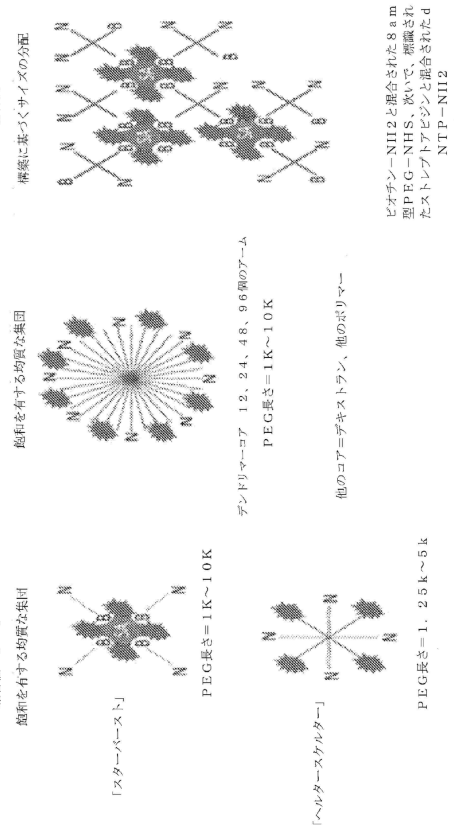
【図 14】



【図 15】



【図 16】



10

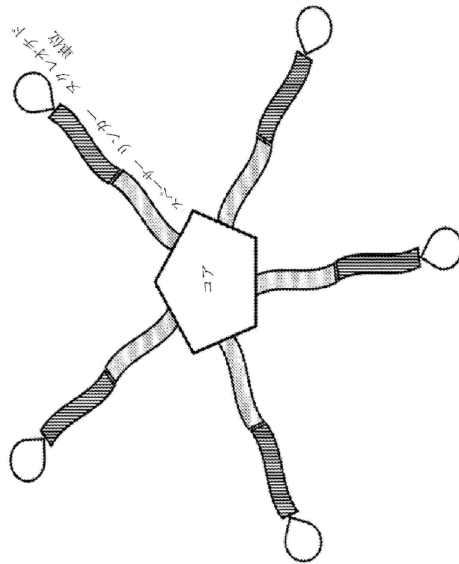
20

30

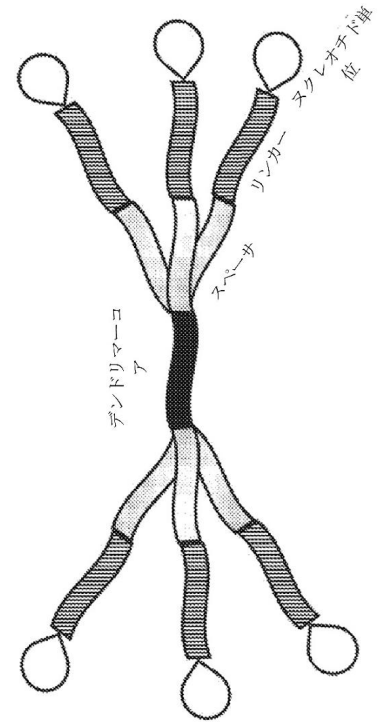
40

50

【図 17】



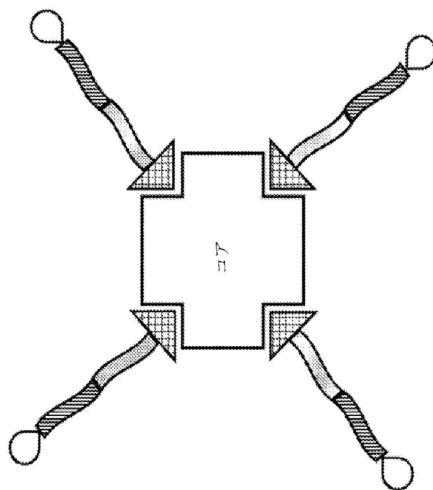
【図 18】



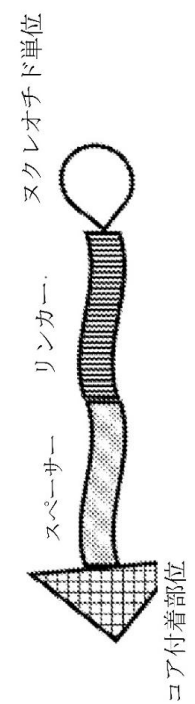
10

20

【図 19】



【図 20】



30

40

50

【図 2 1】

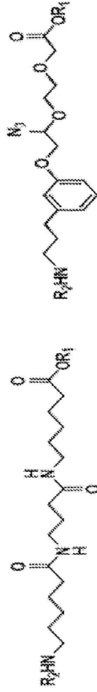


リンカー：

11 原子リンカー：

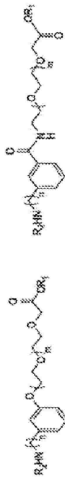


23 原子リンカー：

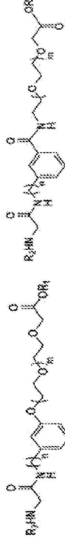


【図 2 2】

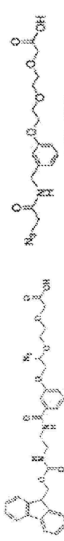
リンカー-1：



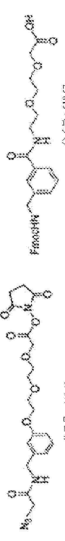
リンカー-3：



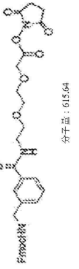
リンカー-5：



リンカー-7：



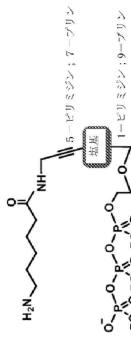
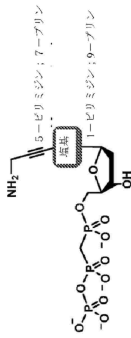
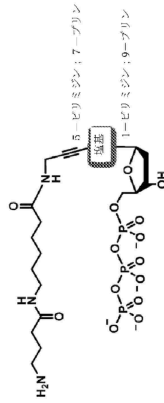
リンカー-9：



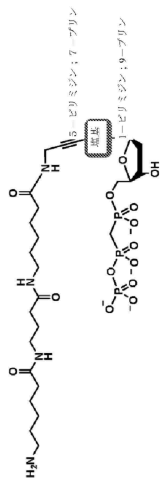
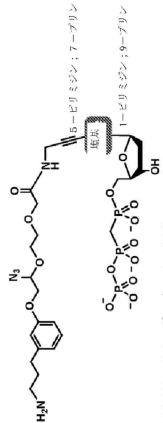
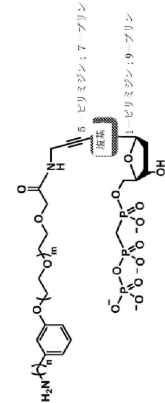
10

20

【図 2 3 A】

dNTP-PA-NH₂：dNTP-PA-23 原子リンカー-NH₂：

【図 2 3 B】

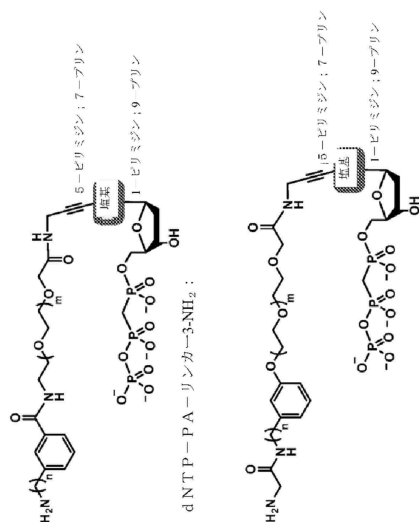
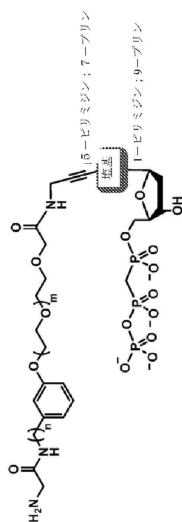
dNTP-PA-23 原子リンカー-NH₂：dNTP-PA-N3 原子リンカー-NH₂：dNTP-PA-リンカー-1-NH₂：

30

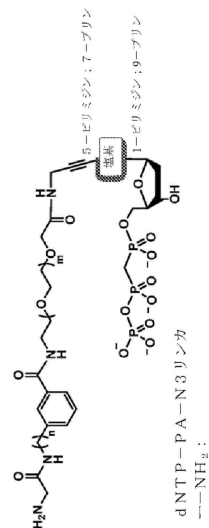
40

50

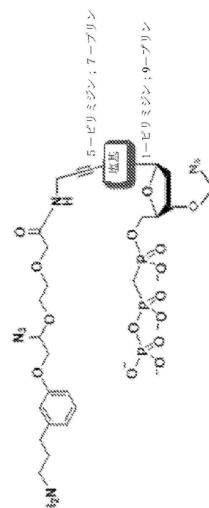
【図 23 C】

dNTP-PA-リンカー-2-NH₂ :dNTP-PA-リンカー-3-NH₂ :

【図 23 D】

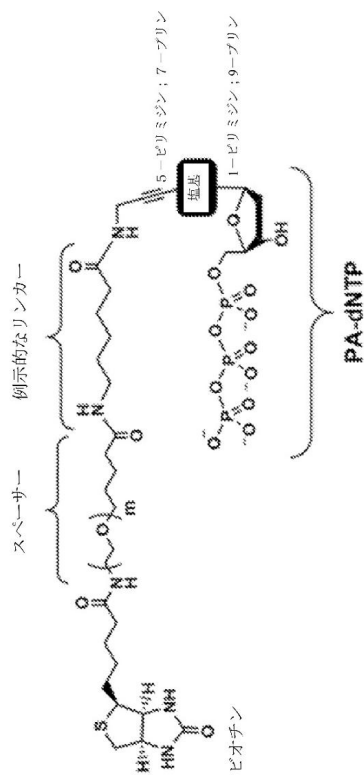
dNTP-PA-リンカー-4-NH₂ :

dNTP-PA-N3リンカ

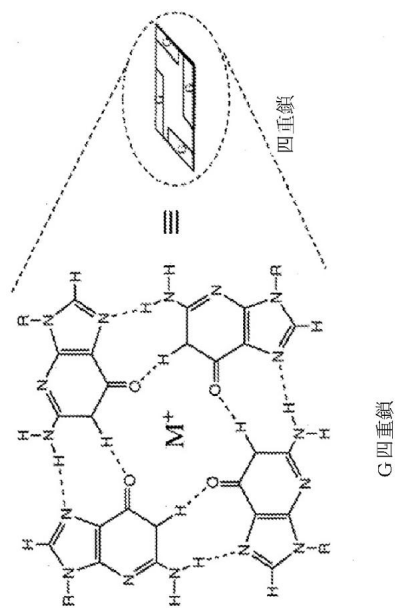
---NH₂ :

10

【図 24】



【図 25】



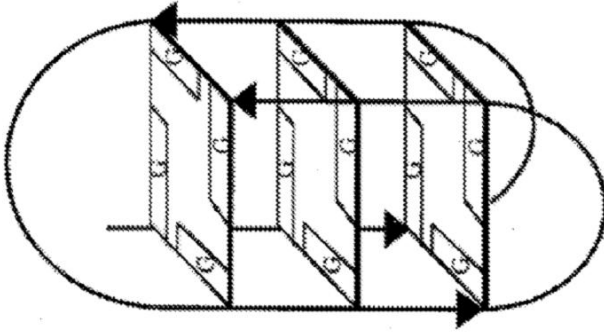
20

30

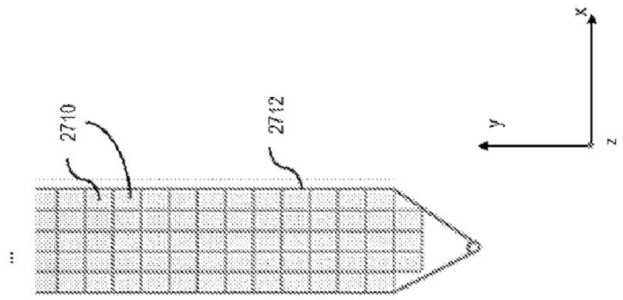
40

50

【図 26】

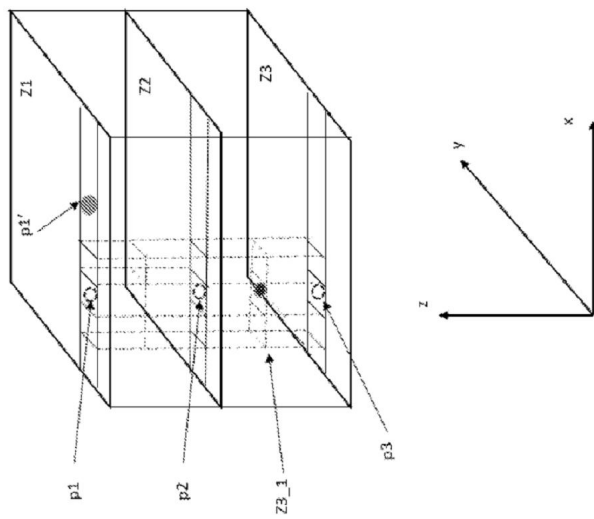


【図 27A】

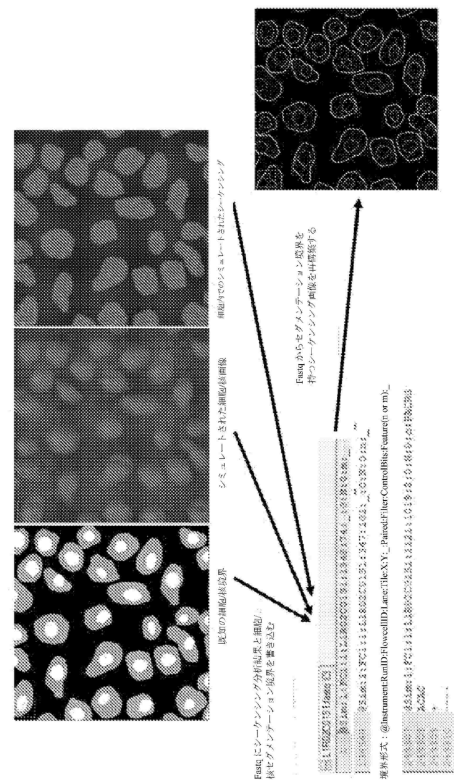


10

【図 27B】



【図 28】



20

30

40

50

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 22/52448

A. CLASSIFICATION OF SUBJECT MATTER IPC - INV. C12Q 1/6869, C12Q 1/6872, G06T 7/00, G16B 25/10, G06T 7/246, G01N 33/49 (2023.01) ADD. G01N 15/14, G16B 15/00 (2023.01) CPC - INV. C12Q 1/6869, C12Q 1/6872, G06T 7/0012, G16B 25/10, G06T 7/246, G01N 33/4915 ADD. G01N 15/14, G16B 15/00, G01N 15/1404, G01N 15/1436 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) See Search History document		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched See Search History document		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) See Search History document		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2020/0392574 A1 (Complete Genomics, Inc), 17 December 2020 (17.12.2020), entire document	1-8, 101-107, 109-115
Y	US 2021/0150715 A1 (Illumina, Inc.), 20 May 2021 (20.05.2021), entire document	1-8, 101-107, 109-115
A	US 2020/0080141 A1 (ACCURAGEN HOLDINGS LIMITED), 12 March 2020 (12.03.2020), entire document	1-8, 101-107, 109-115
A	US 2003/0224419 A1 (Corcoran et al.), 04 December 2003 (04.12.2003), entire document	1-8, 101-107, 109-115
A	US 2016/0102348 A1 (LIFE TECHNOLOGIES CORPORATION), 14 April 2016 (14.04.2016), entire document	1-8, 101-107, 109-115
☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "D" document cited by the applicant in the international application "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "I" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 25 April 2023		Date of mailing of the international search report MAY 15 2023
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300		Authorized officer Karl Rodriguez Telephone No. PCT Helpdesk: 571-272-4300

Form PCT/ISA/210 (second sheet) (July 2022)

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 22/52448

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

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

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
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.: 9-100, 108, 116 and 121-160
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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 Continuation in Supplemental Box) ---

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
Claims 1-8, 101-107 and 109-115

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 SEARCH REPORT

International application No.

PCT/US 22/52448

Continuation of:

Box III. Observations where unity of invention is lacking

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I - Claims 1-8, 101-107 and 109-115 are directed to a method and system for base calling of plurality of flow cell images.

Group II - Claims 117-120 are directed to a data file and methods for generating the same.

The inventions listed as Groups I-II do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Special Technical Features:

The invention of Group I included the features of identifying base calling locations in the first plurality of flow cell images; determinate a base calling template comprising the identified base calling locations; registering a second plurality of flow cell images of the support in one or more subsequent cycles of the one or more cycles to the base calling template; performing base calling of the second plurality of flow cell images based on the base calling locations of the base calling template, not required by any other group.

The invention of Group II included the features of determining spatial coordinates of each of the one or more cellular features in the one or more images; and generating data of one or more cellular features, performing a predetermined number of iterations of operations, for each of the one or more cellular features, inserting a cellular feature indicator of the cellular feature; Inserting a set of spatial coordinates of the spatial coordinates of the cellular feature; and Inserting one or more sequencing parameters and values thereof corresponding to the cellular feature; wherein the data file is configured for reconstruction of images comprising the one or more cellular features and the base calling results, not required by any other group.

Common Technical Features

Groups I and II share the features of base calls; generating a first plurality of flow cell images of a cellular sample.

However, the shared technical features do not represent a contribution over prior art as being anticipated by US 2003/0224419 A1 (Corcoran et al.) (hereinafter Corcoran), 04 December 2003 (04.12.2003).

Corcoran teaches base calls (para [0026]- generating a base call for position i in the jth group based on results of the adjusting.); generating a first plurality of flow cell images of a cellular sample (para [0065]- fluidic system 12 and detection system 14 are provided for collecting and imaging optical signals which are used to determine the sequences of the free ends of the cloned templates on each microbead in a flow cell.).

As the common features were known in the art at the time of the invention, this cannot be considered a common technical feature that would otherwise unify the groups. Therefore, Groups I-II lack unity under PCT Rule 13.

フロントページの続き

(32)優先日 令和4年6月30日(2022.6.30)

(33)優先権主張国・地域又は機関
米国(US)

(31)優先権主張番号 63/388,183

(32)優先日 令和4年7月11日(2022.7.11)

(33)優先権主張国・地域又は機関
米国(US)

(81)指定国・地域 AP(BW,CV,GH,GM,KE,LR,LS,MW,MZ,NA,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,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,ME,MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),OA(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,BB,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CV,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IQ,IR,IS,IT,JM,JO,JP,KE,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,LU,LY,MA,MD,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,ZM,ZW

(特許庁注：以下のものは登録商標)

1. J A V A S C R I P T
2. S W I F T
3. P Y T H O N
4. T R I T O N
5. T W E E N
6. B L U E T O O T H
7. J A V A

ン・ロード10055、エレメント・バイオサイエンス・インコーポレイテッド内

(72)発明者 リウ, ツン-リー

アメリカ合衆国92121カリフォルニア州サンディエゴ、バーンズ・キャニオン・ロード10055、エレメント・バイオサイエンス・インコーポレイテッド内

(72)発明者 クルグリャック, セミヨン

アメリカ合衆国92121カリフォルニア州サンディエゴ、バーンズ・キャニオン・ロード10055、エレメント・バイオサイエンス・インコーポレイテッド内

(72)発明者 グオ, ミンハオ

アメリカ合衆国92121カリフォルニア州サンディエゴ、バーンズ・キャニオン・ロード10055、エレメント・バイオサイエンス・インコーポレイテッド内

Fターム(参考) 4B029 AA07 AA23 BB20 FA12 FA15

4B063 QA13 QQ08 QQ42 QQ52 QR32 QR35 QS36 QX02