

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7522189号
(P7522189)

(45)発行日 令和6年7月24日(2024.7.24)

(24)登録日 令和6年7月16日(2024.7.16)

(51)国際特許分類	F I			
C 1 2 Q	1/6851(2018.01)	C 1 2 Q	1/6851	Z
C 1 2 Q	1/6869(2018.01)	C 1 2 Q	1/6869	Z
C 1 2 Q	1/6816(2018.01)	C 1 2 Q	1/6816	Z
C 1 2 Q	1/6855(2018.01)	C 1 2 Q	1/6855	Z
C 1 2 Q	1/6876(2018.01)	C 1 2 Q	1/6876	Z
請求項の数 16 (全103頁) 最終頁に続く				

(21)出願番号	特願2022-525692(P2022-525692)	(73)特許権者	595117091
(86)(22)出願日	令和2年11月6日(2020.11.6)		ベクトン・ディキンソン・アンド・カン
(65)公表番号	特表2023-500679(P2023-500679		パニー
	A)		BECTON, DICKINSON A
(43)公表日	令和5年1月10日(2023.1.10)		ND COMPANY
(86)国際出願番号	PCT/US2020/059419		アメリカ合衆国 ニュー・ジャージー 0
(87)国際公開番号	WO2021/092386		7 4 1 7 - 1 8 8 0 フランクリン・レ
(87)国際公開日	令和3年5月14日(2021.5.14)		イクス ベクトン・ドライブ 1
審査請求日	令和5年10月4日(2023.10.4)	(74)代理人	100094569
(31)優先権主張番号	62/933,285		弁理士 田中 伸一郎
(32)優先日	令和1年11月8日(2019.11.8)	(74)代理人	100103610
(33)優先権主張国・地域又は機関	米国(US)		弁理士 吉 田 和彦
早期審査対象出願		(74)代理人	100109070
			弁理士 須田 洋之
		(74)代理人	100119013
			最終頁に続く

(54)【発明の名称】 免疫レパートリーシーケンシングのための完全長V(D)J情報を得るためのランダムブライミングの使用

(57)【特許請求の範囲】

【請求項1】

核酸標的の1つまたは複数の発生を含む試料中の核酸標的を標識するための方法であって、

核酸標的を複数のオリゴヌクレオチドバーコードと接触させるステップであって、各オリゴヌクレオチドバーコードが、第1のユニバーサル配列、分子標識、および核酸標的にハイブリダイズすることが可能な標的結合領域を含む、ステップと；

逆転写酵素と、標的結合領域またはその一部分を含むテンプレートスイッチオリゴヌクレオチドとの存在下で、核酸標的にハイブリダイズした複数のオリゴヌクレオチドバーコードを伸長させて、それぞれが核酸標的の少なくとも一部分に相補的な配列、第1の分子標識、標的結合領域、および該標的結合領域の相補体を含む、複数のバーコード化された核酸分子を生成し、それによって、それぞれが標的結合領域および該標的結合領域の相補体を含む、複数のバーコード化された核酸分子を生成するステップと；

各バーコード化された核酸分子の標的結合領域の相補体を、

(i) 複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコード、

(i i) バーコード化された核酸分子自体、および

(i i i) 複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子

のうちの1つまたは複数の標的結合領域とハイブリダイズさせるステップと；

複数のバーコード化された核酸分子の3'末端を伸長させて、それぞれが第1の分子標識

および第 2 の分子標識を含む、複数の伸長したバーコード化された核酸分子を生成するステップと；

核酸標的の配列にハイブリダイズすることが可能な標的特異的プライマーおよび第 1 のユニバーサル配列を含むプライマーを使用して、複数の伸長したバーコード化された核酸分子を増幅して、それによって、核酸標的の配列またはその一部分を含む、第 1 の複数のバーコード化されたアンプリコンを生成するステップと；

ランダムプライマーを第 1 の複数のバーコード化されたアンプリコンにハイブリダイズさせ、ランダムプライマーを伸長させて、複数の伸長産物を生成するステップであって、ランダムプライマーが、第 2 のユニバーサル配列またはその相補体を含む、ステップと；

第 1 のユニバーサル配列および第 2 のユニバーサル配列またはそれらの相補体にハイブリダイズすることが可能なプライマーを使用して、複数の伸長産物を増幅させて、それによって、第 2 の複数のバーコード化されたアンプリコンを生成するステップと、を含む、方法。

【請求項 2】

核酸標的の 1 つまたは複数の発生を含む試料中の核酸標的を標識するための方法であって、

核酸標的を複数のオリゴヌクレオチドバーコードと接触させるステップであって、各オリゴヌクレオチドバーコードが、第 1 のユニバーサル配列、分子標識、および核酸標的にハイブリダイズすることが可能な標的結合領域を含む、ステップと；

複数のオリゴヌクレオチドバーコードを使用して、核酸標的をバーコード化して、それぞれが核酸標的の少なくとも一部分に相補的な配列、第 1 の分子標識、および標的結合領域を含む、複数のバーコード化された核酸分子を生成するステップと；

以下の (i) または (i i) のステップと、

(i) 標的結合領域の相補体を含むオリゴヌクレオチドを複数のバーコード化された核酸分子に結合させて、それぞれが標的結合領域および該標的結合領域の相補体を含む、複数のバーコード化された核酸分子を生成するステップ、または

(i i) 複数のバーコード化された核酸分子を増幅させて、複数の増幅したバーコード化された核酸分子を生成するステップ、および、標的結合領域の相補体を含むオリゴヌクレオチドを複数の増幅したバーコード化された核酸分子に結合させて、それぞれが標的結合領域および該標的結合領域の相補体を含む、複数のバーコード化された核酸分子を生成するステップ；

各バーコード化された核酸分子の標的結合領域の相補体を、

(i) 複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコード、

(i i) バーコード化された核酸分子自体、および

(i i i) 複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子

のうちの 1 つまたは複数の標的結合領域とハイブリダイズさせるステップと；

複数のバーコード化された核酸分子の 3 ' 末端を伸長させて、それぞれが第 1 の分子標識および第 2 の分子標識を含む、複数の伸長したバーコード化された核酸分子を生成するステップと；

核酸標的の配列にハイブリダイズすることが可能な標的特異的プライマーおよび第 1 のユニバーサル配列を含むプライマーを使用して、複数の伸長したバーコード化された核酸分子を増幅して、それによって、核酸標的の配列またはその一部分を含む、第 1 の複数のバーコード化されたアンプリコンを生成するステップと；

ランダムプライマーを第 1 の複数のバーコード化されたアンプリコンにハイブリダイズさせ、ランダムプライマーを伸長させて、複数の伸長産物を生成するステップであって、ランダムプライマーが、第 2 のユニバーサル配列またはその相補体を含む、ステップと；

第 1 のユニバーサル配列および第 2 のユニバーサル配列またはそれらの相補体にハイブリダイズすることが可能なプライマーを使用して、複数の伸長産物を増幅させて、それによって、第 2 の複数のバーコード化されたアンプリコンを生成するステップと、

10

20

30

40

50

を含む、方法。

【請求項 3】

第 2 の複数のバーコード化されたアンプリコンまたはその産物と会合した、別個の配列を有する第 1 の分子標識、別個の配列を有する第 2 の分子標識、またはそれらの組合せの数に基づいて、試料中の核酸標的のコピー数を決定するステップを含み、

任意に、試料中の核酸標的のコピー数を決定するステップが、

(a) 第 2 の複数のバーコード化されたアンプリコンもしくはその産物と会合した、別個の配列を有する第 2 の分子標識の数に基づいて、試料中の核酸標的のコピー数を決定するステップ、および / または

(b) 第 2 の複数のバーコード化されたアンプリコンもしくはその産物と会合した、別個の配列を有する第 1 の分子標識の数に基づいて、試料中の核酸標的のコピー数を決定するステップ

を含んでもよい、

請求項 1 又は 2 のいずれか 1 項に記載の方法。

【請求項 4】

(a) 方法が、各バーコード化された核酸分子の標的結合領域の相補体を、(i) 複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコード、(i i) バーコード化された核酸分子自体、および / または (i i i) 複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子の標的結合領域とハイブリダイズさせる前に、複数のバーコード化された核酸分子を変性させるステップを含む、

(b) 方法が、複数の伸長したバーコード化された核酸分子を増幅させる前に、複数の伸長したバーコード化された核酸分子を変性させるステップを含む、

(c) 複数の核酸標的のそれぞれの配列が、複数の核酸標的のそれぞれの部分配列を含む、

(d) 複数のバーコード化された核酸分子中の核酸標的の配列が、核酸標的の部分配列を含む、

(e) 第 1 の分子標識が、複数のバーコード化された核酸分子の 3 ' 末端を伸長させた後に、第 2 の分子標識にハイブリダイズされる、

(f) 伸長したバーコード化された核酸分子がそれぞれ、第 1 の分子標識、第 2 の分子標識、標的結合領域、および該標的結合領域の相補体を含む、

(g) 標的結合領域の相補体が、標的結合領域の一部に相補的である、

(h) 標的結合領域が、遺伝子特異的配列、ポリ (d T) 配列、または両方を含む、
請求項 1 ~ 3 のいずれか 1 項に記載の方法。

【請求項 5】

(a) 試料中の核酸標的をバーコード化するステップが、核酸標的にハイブリダイズした複数のオリゴヌクレオチドバーコードを伸長させて、複数のバーコード化された核酸分子を生成するステップを含み、任意に、核酸標的にハイブリダイズした複数のオリゴヌクレオチドバーコードを伸長させるステップが、複数のオリゴヌクレオチドバーコードにハイブリダイズした核酸標的を逆転写するステップを含んでもよい、および / または

(b) 標的結合領域の相補体を含むオリゴヌクレオチドを結合させるステップが、標的結合領域の相補体を含むオリゴヌクレオチドを、複数のバーコード化された核酸分子および / または増幅したバーコード化された核酸分子にライゲーションさせるステップを含む、および / または

(c) 標的結合領域がポリ (d T) 配列を含み、該標的結合領域の相補体を含むオリゴヌクレオチドを結合させるステップが、末端デオキシヌクレオチド転移酵素を使用して、複数のアデノシンリン酸を複数のバーコード化された核酸分子および / または増幅したバーコード化された核酸分子に付加するステップを含む、

請求項 2 に記載の方法。

【請求項 6】

(a) 逆転写酵素が、末端転移酵素活性を有する、

(b) テンプレートスイッチオリゴヌクレオチドが、1つまたは複数の3'リボヌクレオチドを含み、任意に、3つの3'リボヌクレオチドを含んでもよく、任意に、3'リボヌクレオチドが、グアニンを含んでもよい、および/または

(c) 逆転写酵素が、ウイルス逆転写酵素を含み、任意に、マウス白血病ウイルス(MLV)逆転写酵素またはモロニーマウス白血病ウイルス(MMLV)逆転写酵素を含んでもよい、

請求項1または3~5のいずれか1項に記載の方法。

【請求項7】

(a) バーコード化された核酸分子の標的結合領域の相補体をバーコード化された核酸分子自体の標的結合領域とハイブリダイズさせるステップが、ステムループを形成する、バーコード化された核酸分子内の、標的結合領域と該標的結合領域の相補体との分子内ハイブリダイゼーションを含み、任意に、第2の分子標識が、第1の分子標識の相補体であってもよい、および/または

10

(b) バーコード化された核酸分子の標的結合領域の相補体を、複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコードの標的結合領域とハイブリダイズさせるステップが、バーコード化された核酸分子の標的結合領域の相補体と、複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコードの標的結合領域との分子間ハイブリダイゼーションを含み、任意に

(i) 第2の分子標識が、第1の分子標識と異なってもよく、第2の分子標識が、第1の分子標識の相補体でなくともよく、および/または

20

(ii) 方法が、バーコード化された核酸分子の標的結合領域の相補体にハイブリダイズしたオリゴヌクレオチドバーコードの3'末端を伸長させて、それぞれが第1の分子標識の相補体および第2の分子標識を含む、複数の伸長したバーコード化された核酸分子を生成するステップを含んでもよく、任意に、第2の分子標識の配列が、第1の分子標識の配列と異なってもよく、第2の分子標識が、第1の分子標識の相補体でなくともよい、請求項1~6のいずれか1項に記載の方法。

【請求項8】

(a) バーコード化された核酸分子の標的結合領域の相補体を、複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子の標的結合領域とハイブリダイズさせるステップが、バーコード化された核酸分子の標的結合領域の相補体と、複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子の標的結合領域との分子間ハイブリダイゼーションを含み、任意に、第2の分子標識の配列が、第1の分子標識の配列と異なってもよく、第2の分子標識が、第1の分子標識の相補体でなくともよい、

30

(b) 試料が、単一細胞、複数の細胞、複数の単一細胞、組織、腫瘍試料、またはこれらの任意の組合せを含み、任意に、単一細胞が、免疫細胞または循環腫瘍細胞を含んでもよく、任意に、免疫細胞がB細胞またはT細胞であってもよい、

(c) 標的・特異的プライマーが免疫受容体をコードする核酸に特異的にハイブリダイズし、任意に、標的・特異的プライマーが、免疫受容体の定常領域、免疫受容体の可変領域、免疫受容体の多様性領域、免疫受容体の可変領域と多様性領域の接合部、またはこれらの組合せをコードする核酸に特異的にハイブリダイズしてもよく、任意に、免疫受容体が、T細胞受容体(TCR)および/またはB細胞受容体(BCR)であってもよく、任意に、TCRが、TCRアルファ鎖、TCRベータ鎖、TCRガンマ鎖、TCRデルタ鎖、またはこれらの任意の組合せを含んでもよく、さらに、任意に、BCRが、BCR重鎖および/またはBCR軽鎖を含んでもよい、

40

(d) 複数のバーコード化された核酸分子の3'末端を伸長させるステップが、5'から3'へのエキソヌクレアーゼ活性および3'から5'へのエキソヌクレアーゼ活性のうち少なくとも1つを欠くDNAポリメラーゼを使用して、複数のバーコード化された核酸分子の3'末端を伸長させるステップを含み、任意に、DNAポリメラーゼが、Klenow断片を含んでもよい、および/または

50

(e) 複数の伸長産物を増幅させるステップが、シーケンシングプライマーおよび/またはシーケンシングアダプターの結合部位の配列、その相補配列、および/またはそれらの一部分を、複数の伸長産物に付加するステップを含み、任意に、シーケンシングアダプターが、P5配列、P7配列、それらの相補配列、またはそれらの一部分を含んでもよく、さらに、任意に、シーケンシングプライマーが、リード1シーケンシングプライマー、リード2シーケンシングプライマー、それらの相補配列、またはそれらの一部分を含んでもよい、

請求項1～7のいずれか1項に記載の方法。

【請求項9】

第2の複数のバーコード化されたアンプリコンまたはその産物の配列情報を得るステップを含み、任意に、配列情報を得るステップが、シーケンシングアダプターを第2の複数のバーコード化されたアンプリコンまたはその産物に結合させるステップを含んでもよく、任意に、

10

(a) 第2の複数のバーコード化されたアンプリコンまたはその産物の配列情報を得るステップが、

第2の複数のバーコード化されたアンプリコンまたはその産物の複数のシーケンシングリードを含むシーケンシングデータを得るステップであって、複数のシーケンシングリードのそれぞれが、(1)細胞標識配列、(2)分子標識配列、および/または(3)核酸標的の部分配列を含む、ステップ

を含んでもよく、任意に、

20

(b) 方法が、試料の単一細胞を示すそれぞれ固有の細胞標識配列について、

核酸標的の複数のシーケンシングリードのそれぞれをアライメントさせて、核酸標的のアライメントされた配列を生成するステップを含んでもよく、任意に、

(c) 核酸標的のアライメントされた配列が、核酸標的のcDNA配列の少なくとも50%、核酸標的のcDNA配列の少なくとも70%、核酸標的のcDNA配列の少なくとも90%、または核酸標的のcDNA配列の完全長を含んでもよい、

請求項1～8のいずれか1項に記載の方法。

【請求項10】

核酸標的が、免疫受容体をコードし、任意に、免疫受容体が、BCR軽鎖、BCR重鎖、TCRアルファ鎖、TCRベータ鎖、TCRガンマ鎖、TCRデルタ鎖、またはこれらの任意の組合せを含んでもよく、任意に、

30

(a) 核酸標的のアライメントされた配列が、相補性決定領域1(CDR1)、相補性決定領域2(CDR2)、相補性決定領域3(CDR3)、可変領域、可変領域の完全長、またはこれらの組合せを含んでもよい、

(b) 核酸標的のアライメントされた配列が、可変領域、多様性領域、可変領域、多様性領域および/または定常領域の接合部、またはこれらの任意の組合せを含んでもよい、および/または

(c) 方法が、断片化、タグメンテーション、またはその両方を含まなくてもよい、

請求項1～8のいずれか1項に記載の方法。

【請求項11】

40

(a) 配列情報を得るステップが、単一細胞のBCR軽鎖およびBCR重鎖の配列情報を得るステップを含み、任意に、

(i) BCR軽鎖およびBCR重鎖の配列情報が、BCR軽鎖および/またはBCR重鎖の相補性決定領域1(CDR1)、CDR2、CDR3、またはこれらの任意の組合せの配列を含んでもよい、

(i i) 方法が、得られた配列情報に基づいて、単一細胞のBCR軽鎖とBCR重鎖を対合させるステップを含んでもよい、および/または

(i i i) 試料が、複数の単一細胞を含んでもよく、方法が、得られた配列情報に基づいて、前記単一細胞の少なくとも50%のBCR軽鎖とBCR重鎖を対合させるステップを含んでもよい、ならびに/または

50

(b) 配列情報を得るステップが、単一細胞のTCRアルファ鎖およびTCRベータ鎖の配列情報を得るステップを含み、任意に、

(i) TCRアルファ鎖およびTCRベータ鎖の配列情報が、TCRアルファ鎖および/またはTCRベータ鎖の相補性決定領域1(CDR1)、CDR2、CDR3、またはこれらの任意の組合せの配列を含んでもよい、

(ii) 方法が、得られた配列情報に基づいて、単一細胞のTCRアルファ鎖とTCRベータ鎖を対合させるステップを含んでもよい、および/または

(iii) 試料が、複数の単一細胞を含んでもよく、方法が、得られた配列情報に基づいて、前記単一細胞の少なくとも50%のTCRアルファ鎖とTCRベータ鎖を対合させるステップを含んでもよい、ならびに/または

(c) 配列情報を得るステップが、単一細胞のTCRガンマ鎖およびTCRデルタ鎖の配列情報を得るステップを含み、任意に、

(i) TCRガンマ鎖およびTCRデルタ鎖の配列情報が、TCRガンマ鎖および/またはTCRデルタ鎖の、相補性決定領域1(CDR1)、CDR2、CDR3、またはこれらの任意の組合せの配列を含んでもよい、

(ii) 方法が、得られた配列情報に基づいて、単一細胞のTCRガンマ鎖とTCRデルタ鎖を対合させるステップを含んでもよい、および/または

(iii) 試料が、複数の単一細胞を含んでもよく、方法が、得られた配列情報に基づいて、前記単一細胞の少なくとも50%のTCRガンマ鎖とTCRデルタ鎖を対合させるステップを含んでもよい、

請求項9または10に記載の方法。

【請求項12】

(a) 標的結合領域の相補体が、標的結合領域の逆相補配列または標的結合領域の相補配列を含む、

(b) 分子標識の相補体が、分子標識の逆相補配列または分子標識の相補配列を含む、

(c) 複数のバーコード化された核酸分子が、バーコード化されたデオキシリボ核酸(DNA)分子、バーコード化されたリボ核酸(RNA)分子、またはこれらの組合せを含む、

(d) 核酸標的が、核酸分子を含み、任意に、核酸分子が、リボ核酸(RNA)、メッセンジャーRNA(mRNA)、マイクロRNA、低分子干渉RNA(siRNA)、RNA分解産物、ポリ(A)テールを含むRNA、またはこれらの任意の組合せを含んでもよく、さらに、任意に、mRNAが免疫受容体をコードしてもよく、任意に、核酸標的が、細胞成分結合試薬を含んでもよく、および/または核酸分子が、細胞成分結合試薬に会合しているともよく、任意に、核酸分子と細胞成分結合試薬を解離させるステップを含んでもよい、

(e) 複数のオリゴヌクレオチドバーコードのうちの少なくとも10個が、異なる分子標識配列を含む、

(f) 複数のオリゴヌクレオチドバーコードの各分子標識が、少なくとも6個のヌクレオチドを含む、ならびに/または

(g) 複数のオリゴヌクレオチドバーコードが、固体支持体と会合しており、任意に、

(i) 固体支持体が、合成粒子、平面表面、またはこれらの組合せを含んでもよい、および/または

(ii) 同じ固体支持体に会合した複数のオリゴヌクレオチドバーコードがそれぞれ、同一の試料標識を含んでもよく、任意に、複数のオリゴヌクレオチドバーコードの各試料標識が、少なくとも6個のヌクレオチドを含んでもよい、ならびに/または

(h) 複数のオリゴヌクレオチドバーコードがそれぞれ、細胞標識を含み、任意に、

(i) 複数のオリゴヌクレオチドバーコードの各細胞標識が、少なくとも6個のヌクレオチドを含んでもよい、

(ii) 同じ固体支持体に会合しているオリゴヌクレオチドバーコードが、同じ細胞標識を含んでもよい、

10

20

30

40

50

(i i i)異なる固体支持体に会合しているオリゴヌクレオチドバーコードが、異なる細胞標識を含んでもよい、および/または

(i v)複数の伸長したバーコード化された核酸分子がそれぞれ、細胞標識および該細胞標識の相補体を含んでもよく、任意に、細胞標識の相補体が、細胞標識の逆相補配列または細胞標識の相補配列を含んでもよい、ならびに/または

(i)方法が、エチレングリコール、ポリエチレングリコール、1,2-プロパンジオール、ジメチルスルホキシド(DMSO)、グリセロール、ホルムアミド、7-デアザ-GTP、アセトアミド、塩化テトラメチルアンモニウム塩、ベタイン、またはこれらの任意の組合せのうちの1つまたは複数の存在下で、核酸標的にハイブリダイズした複数のオリゴヌクレオチドバーコードを伸長させるステップを含む、
請求項1~11のいずれか1項に記載の方法。

10

【請求項13】

試料が、単一細胞を含み、方法が、複数のオリゴヌクレオチドバーコードを含む合成粒子を、試料中の単一細胞と会合させるステップを含む、請求項1~12のいずれか1項に記載の方法。

【請求項14】

(a)方法が、合成粒子を単一細胞と会合させた後に、単一細胞を溶解させるステップを含み、任意に、単一細胞を溶解させるステップが、試料を加熱するステップ、試料を界面活性剤と接触させるステップ、試料のpHを変更するステップ、またはこれらの任意の組合せを含んでもよい、

20

(b)合成粒子と単一細胞が、同じ区画中にあり、任意に、区画がウェルまたは液滴であってよい、

(c)複数のオリゴヌクレオチドバーコードのうちの少なくとも1つが、合成粒子上に固定化されているかもしくは部分的に固定化されている、または複数のオリゴヌクレオチドバーコードのうちの少なくとも1つが、合成粒子内に封入されているかもしくは部分的に封入されている、

(d)合成粒子が、崩壊可能であり、任意に、崩壊可能なハイドロゲル粒子であってよい、

(e)合成粒子が、ビーズを含み、任意に、ビーズが、セファロースビーズ、ストレプトアビジンビーズ、アガロースビーズ、磁性ビーズ、コンジュゲートビーズ、プロテインAコンジュゲートビーズ、プロテインGコンジュゲートビーズ、プロテインA/Gコンジュゲートビーズ、プロテインLコンジュゲートビーズ、オリゴ(dT)コンジュゲートビーズ、シリカビーズ、シリカ様ビーズ、抗ビオチンマイクロビーズ、抗蛍光色素マイクロビーズ、もしくはこれらの任意の組合せを含んでもよい、

30

(f)合成粒子が、ポリジメチルシロキサン(PDMS)、ポリスチレン、ガラス、ポリプロピレン、アガロース、ゼラチン、ハイドロゲル、常磁性体、セラミック、プラスチック、メチルスチレン、アクリルポリマー、チタン、ラテックス、セファロース、セルロース、ナイロン、シリコン、およびこれらの任意の組合せからなる群から選択される材料を含む、および/または

(g)複数のオリゴヌクレオチドバーコードのそれぞれが、リンカー官能基を含み、合成粒子が、固体支持体官能基を含み、支持体官能基およびリンカー官能基が、互いに会合しており、任意に、リンカー官能基および支持体官能基が、それぞれに、C6、ビオチン、ストレプトアビジン、第一級アミン、アルデヒド、ケトン、およびこれらの任意の組合せからなる群から選択されてもよい、

40

請求項13に記載の方法。

【請求項15】

複数のオリゴヌクレオチドバーコードであって、複数のオリゴヌクレオチドバーコードのそれぞれが、分子標識および標的結合領域を含み、複数のオリゴヌクレオチドバーコードのうちの少なくとも10個が、異なる分子標識配列を含む、複数のオリゴヌクレオチドバーコードと、

50

逆転写酵素と、

5'から3'へのエキソヌクレアーゼ活性および3'から5'へのエキソヌクレアーゼ活性のうち少なくとも1つを欠くDNAポリメラーゼと、

ランダムプライマーと、

標的結合領域またはその一部分を含むテンプレートスイッチングオリゴヌクレオチドと、を含む、キット。

【請求項16】

(a) 逆転写酵素が、ウイルス逆転写酵素を含み、任意に、ウイルス逆転写酵素が、マウス白血病ウイルス(MLV)逆転写酵素またはモロニー Maus 白血病ウイルス(MMLV)逆転写酵素であってもよく、

10

(b) テンプレートスイッチオリゴヌクレオチドが、1つまたは複数の3'リボヌクレオチドを含み、任意に、3つの3'リボヌクレオチドを含んでもよく、任意に、3'リボヌクレオチドが、グアニンを含んでもよい、

(c) キットが、エチレングリコール、ポリエチレングリコール、1,2-プロパンジオール、ジメチルスルホキシド(DMSO)、グリセロール、ホルムアミド、7-デアザ-GTP、アセトアミド、塩化テトラメチルアンモニウム塩、ベタイン、またはこれらの任意の組合せのうち1つまたは複数を含む、

(d) DNAポリメラーゼが、Klenow断片を含む、

(e) キットが、緩衝液、カートリッジ、または両方を含む、

(f) キットが、逆転写反応および/または増幅反応のための1つまたは複数の試薬を含む、

20

(g) 標的結合領域が、遺伝子特異的配列、オリゴ(dT)配列、ランダム多量体、またはこれらの任意の組合せを含む、

(h) オリゴヌクレオチドバーコードが、同一の試料標識および/または同一の細胞標識を含み、任意に、複数のオリゴヌクレオチドバーコードの各試料標識および/または細胞標識が、少なくとも6個のヌクレオチドを含んでもよい、

(i) 複数のオリゴヌクレオチドバーコードの各分子標識が、少なくとも6個のヌクレオチドを含む、

(j) 複数のオリゴヌクレオチドバーコードのうち少なくとも1つが、合成粒子上に固定化されているかもしくは部分的に固定化されている、または合成粒子内に封入されているかもしくは部分的に封入されている、

30

(k) 合成粒子が、崩壊可能であり、任意に、崩壊可能なハイドロゲル粒子であってもよい、

(l) 合成粒子が、ビーズを含み、任意に、ビーズが、セファロースビーズ、ストレプトアビジンビーズ、アガロースビーズ、磁性ビーズ、コンジュゲートビーズ、プロテインAコンジュゲートビーズ、プロテインGコンジュゲートビーズ、プロテインA/Gコンジュゲートビーズ、プロテインLコンジュゲートビーズ、オリゴ(dT)コンジュゲートビーズ、シリカビーズ、シリカ様ビーズ、抗ビオチンマイクロビーズ、抗蛍光色素マイクロビーズ、またはこれらの任意の組合せを含んでもよい、

(m) 合成粒子が、ポリジメチルシロキサン(PDMS)、ポリスチレン、ガラス、ポリプロピレン、アガロース、ゼラチン、ハイドロゲル、常磁性体、セラミック、プラスチック、メチルシロキサン、アクリルポリマー、チタン、ラテックス、セファロース、セルロース、ナイロン、シリコン、およびこれらの任意の組合せからなる群から選択される材料を含む、および/または

40

(n) 複数のオリゴヌクレオチドバーコードのそれぞれが、リンカー官能基を含み、合成粒子が、固体支持体官能基を含み、支持体官能基およびリンカー官能基が、互いに会合しており、任意に、リンカー官能基および支持体官能基が、それぞれに、C6、ビオチン、ストレプトアビジン、第一級アミン、アルデヒド、ケトン、およびこれらの任意の組合せからなる群から選択されてもよい、

請求項15に記載のキット。

50

【発明の詳細な説明】

【技術分野】

【0001】

関連出願

本出願は、35 U.S.C. § 119(e)のもとに、2019年11月8日に出願された米国仮特許出願第62/933285号の利益を主張し、この関連出願の内容は、あらゆる目的で参照によりその全体が本明細書に組み込まれる。

【背景技術】

【0002】

本開示は、概して、分子生物学の分野に関し、より具体的には、分子バーコード化を使用するマルチオミクス分析に関する。

10

分子バーコード化の方法および技法は、例えば、逆転写、ポリメラーゼ連鎖反応(PCR)増幅、および次世代シーケンシング(NGS)を使用して、細胞の状態を決定するために、遺伝子発現プロファイルを解読することなど、単一細胞トランスクリプトミクス分析に有用である。核酸標的(例えば、V(D)Jを含有する転写物)の完全長発現プロファイリングのための方法および技法に対する必要性が存在する。核酸標的(例えば、V(D)Jを含有する転写物)の同定と計数の両方を可能にし得る組成物および方法に対する必要性が存在する。

【発明の概要】

【0003】

本明細書の開示は、試料中の核酸標的を標識するための方法を含む。一部の実施形態では、本方法は、核酸標的のコピーを複数のオリゴヌクレオチドバーコードと接触させるステップであって、各オリゴヌクレオチドバーコードが、第1のユニバーサル配列、分子標識(例えば、第1の分子標識)、および核酸標的にハイブリダイズすることが可能な標的結合領域を含む、ステップと；それぞれが、標的結合領域および該標的結合領域の相補体を含む複数のバーコード化された核酸分子を生成するステップと；各バーコード化された核酸分子の標的結合領域の相補体を、(i)複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコード、(ii)バーコード化された核酸分子自体、および(iii)複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子のうちの1つまたは複数の標的結合領域とハイブリダイズさせるステップと；複数のバーコード化された核酸分子の3'末端を伸長させて、それぞれが、第1の分子標識および第2の分子標識を含む、複数の伸長したバーコード化された核酸分子を生成するステップと；核酸標的の配列にハイブリダイズすることが可能な標的的特異的プライマーおよび第1のユニバーサル配列を含むプライマーを使用して、複数の伸長したバーコード化された核酸分子を増幅させて、それによって、核酸標的の配列、またはその一部分を含む第1の複数のバーコード化されたアンプリコンを生成するステップと；ランダムプライマーを第1の複数のバーコード化されたアンプリコンにハイブリダイズさせ、ランダムプライマーを伸長させて、複数の伸長産物を生成するステップであって、ランダムプライマーが、第2のユニバーサル配列、またはその相補体を含む、ステップと；第1のユニバーサル配列および第2のユニバーサル配列、またはその相補体にハイブリダイズすることが可能なプライマーを使用して、複数の伸長産物を増幅させて、それによって、第2の複数のバーコード化されたアンプリコンを生成するステップとを含む。一部の実施形態では、本方法は、第2の複数のバーコード化されたアンプリコン、またはその産物と会合した、別個の配列を有する第1の分子標識、別個の配列を有する第2の分子標識、またはこれらの組合せの数に基づいて、試料中の核酸標的のコピー数を決定するステップを含む。一部の実施形態では、本方法は、断片化、タグメンテーション、またはその両方を含まない。

20

30

40

【0004】

本明細書の開示は、試料中の核酸標的の数を決定する方法を含む。一部の実施形態では、本方法は、核酸標的のコピーを複数のオリゴヌクレオチドバーコードと接触させるステップであって、各オリゴヌクレオチドバーコードが、第1のユニバーサル配列、分子標識

50

(例えば、第1の分子標識)、および核酸標的にハイブリダイズすることが可能な標的結合領域を含む、ステップと；それぞれが、標的結合領域および標的結合領域の相補体を含む複数のバーコード化された核酸分子を生成するステップと；各バーコード化された核酸分子の標的結合領域の相補体を、(i)複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコード、(ii)バーコード化された核酸分子自体、および(iii)複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子のうちの1つまたは複数の標的結合領域とハイブリダイズさせるステップと；複数のバーコード化された核酸分子の3'末端を伸長させて、それぞれが、第1の分子標識および第2の分子標識を含む、複数の伸長したバーコード化された核酸分子を生成するステップと；核酸標的の配列にハイブリダイズすることが可能な標的特定のプライマーおよび第1のユニバーサル配列を含むプライマーを使用して、複数の伸長したバーコード化された核酸分子を増幅させて、それによって、核酸標的の配列、またはその一部分を含む第1の複数のバーコード化されたアンプリコンを生成するステップと；ランダムプライマーを第1の複数のバーコード化されたアンプリコンにハイブリダイズさせ、ランダムプライマーを伸長させて、複数の伸長産物を生成するステップであって、ランダムプライマーが、第2のユニバーサル配列、またはその相補体を含む、ステップと；第1のユニバーサル配列および第2のユニバーサル配列、またはその相補体にハイブリダイズすることが可能なプライマーを使用して、複数の伸長産物を増幅させて、それによって、第2の複数のバーコード化されたアンプリコンを生成するステップと；第2の複数のバーコード化されたアンプリコン、またはその産物と会合した、別個の配列を有する第1の分子標識、別個の配列を有する第2の分子標識、またはこれらの組合せの数に基づいて、試料中の核酸標的のコピー数を決定するステップとを含む。

【0005】

一部の実施形態では、試料中の核酸標的のコピー数を決定するステップは、(a)第2の複数のバーコード化されたアンプリコン、またはその産物と会合した、別個の配列を有する第2の分子標識の数に基づいて、試料中の核酸標的のコピー数を決定するステップ、および/または(b)第2の複数のバーコード化されたアンプリコン、またはその産物と会合した、別個の配列を有する第1の分子標識の数に基づいて、試料中の核酸標的のコピー数を決定するステップを含む。一部の実施形態では、本方法は、各バーコード化された核酸分子の標的結合領域の相補体を、(i)複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコード、(ii)バーコード化された核酸分子自体、および/または(iii)複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子の標的結合領域とハイブリダイズさせる前に、複数のバーコード化された核酸分子を変性させるステップを含む。一部の実施形態では、本方法は、複数の伸長したバーコード化された核酸分子を増幅させる前に、複数の伸長したバーコード化された核酸分子を変性させるステップを含む。

【0006】

一部の実施形態では、本方法は、複数の伸長したバーコード化された核酸分子を増幅させる前に、複数の伸長したバーコード化された核酸分子を変性させるステップを含む。一部の実施形態では、複数の核酸標的のそれぞれの配列は、複数の核酸標的のそれぞれの部分配列を含む。一部の実施形態では、複数のバーコード化された核酸分子中の核酸標的の配列は、核酸標的の部分配列を含む。一部の実施形態では、第1の分子標識は、複数のバーコード化された核酸分子の3'末端を伸長させた後に、第2の分子標識にハイブリダイズされる。一部の実施形態では、伸長したバーコード化された核酸分子はそれぞれ、第1の分子標識、第2の分子標識、標的結合領域、および標的結合領域の相補体を含む。一部の実施形態では、標的結合領域の相補体は、標的結合領域の一部分に相補的である。一部の実施形態では、標的結合領域は、遺伝子特異的配列を含む。一部の実施形態では、標的結合領域は、遺伝子特異的配列、ポリ(dT)配列、または両方を含む。

【0007】

一部の実施形態では、それぞれが、標的結合領域および該標的結合領域の相補体を含む

複数のバーコード化された核酸分子を生成するステップは、(i)複数のオリゴヌクレオチドバーコードを使用して、試料中の核酸標的のコピーをバーコード化して、それぞれが、核酸標的の少なくとも一部分に相補的な配列、第1の分子標識、および標的結合領域を含む、複数のバーコード化された核酸分子を生成するステップと；(ii)標的結合領域の相補体を含むオリゴヌクレオチドを複数のバーコード化された核酸分子に結合させて、それぞれが、標的結合領域および該標的結合領域の相補体を含む複数のバーコード化された核酸分子を生成するステップとを含む。

【0008】

一部の実施形態では、それぞれが、標的結合領域および該標的結合領域の相補体を含む複数のバーコード化された核酸分子を生成するステップは、(i)複数のオリゴヌクレオチドバーコードを使用して、試料中の核酸標的のコピーをバーコード化して、それぞれが、核酸標的の少なくとも一部分に相補的な配列、第1の分子標識、および標的結合領域を含む、複数のバーコード化された核酸分子を生成するステップと；(ii)複数のバーコード化された核酸分子を増幅させて、複数の増幅したバーコード化された核酸分子を生成するステップと；(iii)標的結合領域の相補体を含むオリゴヌクレオチドを複数の増幅したバーコード化された核酸分子に結合させて、それぞれが、標的結合領域および該標的結合領域の相補体を含む複数のバーコード化された核酸分子を生成するステップとを含む。

10

【0009】

一部の実施形態では、試料中の核酸標的のコピーをバーコード化するステップは、核酸標的のコピーにハイブリダイズした複数のオリゴヌクレオチドバーコードを伸長させて、複数のバーコード化された核酸分子を生成するステップを含む。一部の実施形態では、核酸標的のコピーにハイブリダイズした複数のオリゴヌクレオチドバーコードを伸長させるステップは、複数のオリゴヌクレオチドバーコードにハイブリダイズした核酸標的のコピーを逆転写するステップを含む。一部の実施形態では、標的結合領域の相補体を含むオリゴヌクレオチドを結合させるステップは、該標的結合領域の相補体を含むオリゴヌクレオチドを、複数のバーコード化された核酸分子および/または増幅したバーコード化された核酸分子にライゲーションさせるステップを含む。一部の実施形態では、標的結合領域はポリ(dT)配列を含み、該標的結合領域の相補体を含むオリゴヌクレオチドを結合させるステップは、末端デオキシヌクレオチド転移酵素を使用して、複数のアデノシン-リン酸を複数のバーコード化された核酸分子および/または増幅したバーコード化された核酸分子に付加するステップを含む。

20

30

【0010】

一部の実施形態では、それぞれが、標的結合領域および該標的結合領域の相補体を含む複数のバーコード化された核酸分子を生成するステップは、逆転写酵素および標的結合領域またはその一部分を含むテンプレートスイッチオリゴヌクレオチドの存在下で、核酸標的のコピーにハイブリダイズした複数のオリゴヌクレオチドバーコードを伸長させて、それぞれが、核酸標的の少なくとも一部分に相補的な配列、第1の分子標識、標的結合領域、および該標的結合領域の相補体を含む、複数のバーコード化された核酸分子を生成するステップを含む。一部の実施形態では、逆転写酵素は、末端転移酵素活性の能力がある。一部の実施形態では、テンプレートスイッチオリゴヌクレオチドは、1つまたは複数の3'リボヌクレオチド、例えば、3つの3'リボヌクレオチドを含む。一部の実施形態では、3'リボヌクレオチドは、グアニンを含む。一部の実施形態では、逆転写酵素は、ウイルス逆転写酵素、例えば、マウス白血病ウイルス(MLV)逆転写酵素またはモロニーマウス白血病ウイルス(MMLV)逆転写酵素を含む。一部の実施形態では、本方法は、エチレングリコール、ポリエチレングリコール、1,2-プロパンジオール、ジメチルスルホキシド(DMSO)、グリセロール、ホルムアミド、7-デアザ-GTP、アセトアミド、塩化テトラメチルアンモニウム塩、ベタイン、またはこれらの任意の組合せのうちの1つまたは複数の存在下で、核酸標的のコピーにハイブリダイズした複数のオリゴヌクレオチドバーコードを伸長させるステップを含む。

40

50

一部の実施形態では、バーコード化された核酸分子の標的結合領域の相補体をバーコード化された核酸分子自体の標的結合領域とハイブリダイズさせるステップは、ステムループを形成する、バーコード化された核酸分子内の標的結合領域と該標的結合領域の相補体との分子内ハイブリダイゼーションを含む。一部の実施形態では、第2の分子標識は、第1の分子標識の相補体である。

【0011】

一部の実施形態では、バーコード化された核酸分子の標的結合領域の相補体を、複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコードの標的結合領域とハイブリダイズさせるステップは、バーコード化された核酸分子の標的結合領域の相補体の、複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコードの標的結合領域との分子間ハイブリダイゼーションを含む。一部の実施形態では、第2の分子標識は第1の分子標識と異なり、第2の分子標識は、第1の分子標識の相補体ではない。一部の実施形態では、本方法は、バーコード化された核酸分子の標的結合領域の相補体にハイブリダイズしたオリゴヌクレオチドバーコードの3'末端を伸長させて、それぞれが、第1の分子標識の相補体および第2の分子標識を含む、複数の伸長したバーコード化された核酸分子を生成するステップを含む。一部の実施形態では、第2の分子標識の配列は、第1の分子標識の配列と異なり、第2の分子標識は、第1の分子標識の相補体ではない。

10

【0012】

一部の実施形態では、バーコード化された核酸分子の標的結合領域の相補体を、複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子の標的結合領域とハイブリダイズさせるステップは、バーコード化された核酸分子の標的結合領域の相補体の、複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子の標的結合領域との分子間ハイブリダイゼーションを含む。一部の実施形態では、第2の分子標識の配列は第1の分子標識の配列と異なり、第2の分子標識は、第1の分子標識の相補体ではない。

20

【0013】

一部の実施形態では、試料は、単一細胞、複数の細胞、複数の単一細胞、組織、腫瘍試料、またはこれらの任意の組合せを含む。一部の実施形態では、単一細胞は、免疫細胞または循環腫瘍細胞を含む。一部の実施形態では、免疫細胞は、B細胞またはT細胞である。一部の実施形態では、標的特定のプライマーは、免疫受容体に特異的にハイブリダイズする。一部の実施形態では、標的特定のプライマーは、免疫受容体の定常領域、免疫受容体の可変領域、免疫受容体の多様性領域、免疫受容体の可変領域と多様性領域の接合部、またはこれらの組合せに特異的にハイブリダイズする。一部の実施形態では、免疫受容体は、T細胞受容体(TCR)および/またはB細胞受容体(BCR)受容体である。一部の実施形態では、TCRは、TCRアルファ鎖、TCRベータ鎖、TCRガンマ鎖、TCRデルタ鎖、またはこれらの任意の組合せを含む。一部の実施形態では、BCR受容体は、BCR重鎖および/またはBCR軽鎖を含む。

30

【0014】

一部の実施形態では、複数のバーコード化された核酸分子の3'末端を伸長させるステップは、5'から3'へのエキソヌクレアーゼ活性および3'から5'へのエキソヌクレアーゼ活性のうち少なくとも1つを欠くDNAポリメラーゼを使用して、複数のバーコード化された核酸分子の3'末端を伸長させるステップを含む。一部の実施形態では、DNAポリメラーゼは、Klenow断片を含む。一部の実施形態では、本方法は、第2の複数のバーコード化されたアンプリコン、またはその産物の配列情報を得るステップを含む。一部の実施形態では、配列情報を得るステップは、シーケンシングアダプターを、第2の複数のバーコード化されたアンプリコン、またはその産物に結合させるステップを含む。一部の実施形態では、複数の伸長産物を増幅させるステップは、シーケンシングプライマーおよび/もしくはシーケンシングアダプターの結合部位の配列、その相補配列、ならびに/またはその一部分を、複数の伸長産物に付加するステップを含む。一部の実施形態では、シーケンシングアダプターは、P5配列、P7配列、その相補配列、またはその一部分を含

40

50

む。一部の実施形態では、シーケンシングプライマーは、リード1シーケンシングプライマー、リード2シーケンシングプライマー、その相補配列、またはその一部分を含む。一部の実施形態では、第2の複数のバーコード化されたアンプリコン、またはその産物の配列情報を得るステップは、第2の複数のバーコード化されたアンプリコン、またはその産物の複数のシーケンシングリードを含むシーケンシングデータを得るステップであって、複数のシーケンシングリードのそれぞれが、(1)細胞標識配列、(2)分子標識配列、および/または(3)核酸標的の部分配列を含む、ステップを含む。

【0015】

一部の実施形態では、本方法は、試料の単一細胞を示す、それぞれ固有の細胞標識配列について、核酸標的の複数のシーケンシングリードのそれぞれをアライメントさせて、核酸標的のアライメントされた配列を生成するステップを含む。一部の実施形態では、核酸標的のアライメントされた配列は、核酸標的のcDNA配列の少なくとも50%、核酸標的のcDNA配列の少なくとも70%、核酸標的のcDNA配列の少なくとも90%、または核酸標的のcDNA配列の完全長を含む。一部の実施形態では、核酸標的は、免疫受容体である。一部の実施形態では、免疫受容体は、BCR軽鎖、BCR重鎖、TCRアルファ鎖、TCRベータ鎖、TCRガンマ鎖、TCRデルタ鎖、またはこれらの任意の組合せを含む。一部の実施形態では、核酸標的のアライメントされた配列は、相補性決定領域1(CDR1)、相補性決定領域2(CDR2)、相補性決定領域3(CDR3)、可変領域、可変領域の完全長、またはこれらの組合せを含む。一部の実施形態では、核酸標的のアライメントされた配列は、可変領域、多様性領域、可変領域、多様性領域および/もしくはは定常領域の接合部、またはこれらの任意の組合せを含む。

【0016】

一部の実施形態では、配列情報を得るステップは、単一細胞のBCR軽鎖およびBCR重鎖の配列情報を得るステップを含む。一部の実施形態では、BCR軽鎖およびBCR重鎖の配列情報は、BCR軽鎖および/またはBCR重鎖の、相補性決定領域1(CDR1)、CDR2、CDR3、またはこれらの任意の組合せの配列を含む。一部の実施形態では、方法は、得られた配列情報に基づいて、単一細胞のBCR軽鎖とBCR重鎖を対合させるステップを含む。一部の実施形態では、試料は、複数の単一細胞を含み、本方法は、得られた配列情報に基づいて、前記単一細胞の少なくとも50%のBCR軽鎖とBCR重鎖を対合させるステップを含む。

一部の実施形態では、配列情報を得るステップは、単一細胞のTCRアルファ鎖およびTCRベータ鎖の配列情報を得るステップを含む。一部の実施形態では、TCRアルファ鎖およびTCRベータ鎖の配列情報は、TCRアルファ鎖および/またはTCRベータ鎖の、相補性決定領域1(CDR1)、CDR2、CDR3、またはこれらの任意の組合せの配列を含む。一部の実施形態では、本方法は、得られた配列情報に基づいて、単一細胞のTCRアルファ鎖とTCRベータ鎖を対合させるステップを含む。一部の実施形態では、試料は、複数の単一細胞を含み、本方法は、得られた配列情報に基づいて、前記単一細胞の少なくとも50%のTCRアルファ鎖とTCRベータ鎖を対合させるステップを含む。

一部の実施形態では、配列情報を得るステップは、単一細胞のTCRガンマ鎖およびTCRデルタ鎖の配列情報を得るステップを含む。一部の実施形態では、TCRガンマ鎖およびTCRデルタ鎖の配列情報は、TCRガンマ鎖および/またはTCRデルタ鎖の、相補性決定領域1(CDR1)、CDR2、CDR3、またはこれらの任意の組合せの配列を含む。一部の実施形態では、本方法は、得られた配列情報に基づいて、単一細胞のTCRガンマ鎖とTCRデルタ鎖を対合させるステップを含む。一部の実施形態では、試料は、複数の単一細胞を含み、本方法は、得られた配列情報に基づいて、前記単一細胞の少なくとも50%のTCRガンマ鎖とTCRデルタ鎖を対合させるステップを含む。

【0017】

一部の実施形態では、標的結合領域の相補体は、標的結合領域の逆相補配列または標的結合領域の相補配列を含む。一部の実施形態では、分子標識の相補体は、分子標識の逆相補配列、または分子標識の相補配列を含む。一部の実施形態では、複数のバーコード化さ

れた核酸分子は、バーコード化されたデオキシリボ核酸（DNA）分子、バーコード化されたリボ核酸（RNA）分子、またはこれらの組合せを含む。一部の実施形態では、核酸標的は、核酸分子を含む。一部の実施形態では、核酸分子は、リボ核酸（RNA）、メッセンジャーRNA（mRNA）、マイクロRNA、低分子干渉RNA（siRNA）、RNA分解産物、ポリ（A）テールを含むRNA、またはこれらの任意の組合せを含む。一部の実施形態では、mRNAは、免疫受容体をコードする。一部の実施形態では、核酸標的は、細胞成分結合試薬を含む、および/または核酸分子は、細胞成分結合試薬に会合している。一部の実施形態では、本方法は、核酸分子と細胞成分結合試薬を解離させるステップを含む。

【0018】

一部の実施形態では、複数のオリゴヌクレオチドバーコードのうちの少なくとも10個は、異なる分子標識配列を含む。一部の実施形態では、複数のオリゴヌクレオチドバーコードの各分子標識は、少なくとも6個のヌクレオチドを含む。一部の実施形態では、複数のオリゴヌクレオチドバーコードは、固体支持体に会合している。一部の実施形態では、同じ固体支持体に会合した複数のオリゴヌクレオチドバーコードはそれぞれ、同一の試料標識を含む。一部の実施形態では、複数のオリゴヌクレオチドバーコードの各試料標識は、少なくとも6個のヌクレオチドを含む。一部の実施形態では、複数のオリゴヌクレオチドバーコードはそれぞれ、細胞標識を含む。一部の実施形態では、複数のオリゴヌクレオチドバーコードの各細胞標識は、少なくとも6個のヌクレオチドを含む。一部の実施形態では、同じ固体支持体に会合したオリゴヌクレオチドバーコードは、同じ細胞標識を含む。一部の実施形態では、異なる固体支持体に会合したオリゴヌクレオチドバーコードは、異なる細胞標識を含む。一部の実施形態では、複数の伸長したバーコード化された核酸分子はそれぞれ、細胞標識および該細胞標識の相補体を含む。一部の実施形態では、細胞標識の相補体は、細胞標識の逆相補配列、または細胞標識の相補配列を含む。

【0019】

一部の実施形態では、固体支持体は、合成粒子、平面表面、またはこれらの組合せを含む。一部の実施形態では、試料は、単一細胞を含み、本方法は、複数のオリゴヌクレオチドバーコードを含む合成粒子を、試料中の単一細胞と会合させるステップを含む。本方法は、合成粒子を単一細胞と会合させた後に、単一細胞を溶解させるステップを含んでもよい。一部の実施形態では、単一細胞を溶解させるステップは、試料を加熱するステップ、試料を界面活性剤と接触させるステップ、試料のpHを変更するステップ、またはこれらの任意の組合せを含む。一部の実施形態では、合成粒子と単一細胞は、同じ区画中に存在する。一部の実施形態では、区画は、ウェルまたは液滴である。一部の実施形態では、複数のオリゴヌクレオチドバーコードのうちの少なくとも1つは、合成粒子上に固定化されているかもしくは部分的に固定化されている、または複数のオリゴヌクレオチドバーコードのうちの少なくとも1つは、合成粒子内に封入されているかもしくは部分的に封入されている。一部の実施形態では、合成粒子は、崩壊可能である。一部の実施形態では、合成粒子は、崩壊可能ハイドロゲル粒子である。一部の実施形態では、合成粒子は、ビーズを含む。一部の実施形態では、ビーズは、セファロースビーズ、ストレプトアビジンビーズ、アガロースビーズ、磁性ビーズ、コンジュゲートビーズ、プロテインAコンジュゲートビーズ、プロテインGコンジュゲートビーズ、プロテインA/Gコンジュゲートビーズ、プロテインLコンジュゲートビーズ、オリゴ（dT）コンジュゲートビーズ、シリカビーズ、シリカ様ビーズ、抗ピオチンマイクロビーズ、抗蛍光色素マイクロビーズ、またはこれらの任意の組合せを含む。一部の実施形態では、合成粒子は、ポリジメチルシロキサン（PDMS）、ポリスチレン、ガラス、ポリプロピレン、アガロース、ゼラチン、ハイドロゲル、常磁性体（paramagnetic）、セラミック、プラスチック、ガラス、メチルスチレン、アクリルポリマー、チタン、ラテックス、セファロース、セルロース、ナイロン、シリコン、またはこれらの任意の組合せからなる群から選択される材料を含む。一部の実施形態では、複数のオリゴヌクレオチドバーコードのそれぞれは、リンカー官能基を含み、合成粒子は、固体支持体官能基を含み、ならびに支持体官能基およびリンカー官能基

10

20

30

40

50

は互いに会合している。一部の実施形態では、リンカー官能基および支持体官能基は、独立して、C6、ピオチン、ストレプトアビジン、第一級アミン、アルデヒド、ケトン、およびこれらの任意の組合せからなる群から選択される。

【0020】

本明細書の開示は、キットを含む。一部の実施形態では、キットは、複数のオリゴヌクレオチドバーコードであって、複数のオリゴヌクレオチドバーコードのそれぞれが、分子標識および標的結合領域を含み、複数のオリゴヌクレオチドバーコードのうち少なくとも10個が異なる分子標識配列を含む、複数のオリゴヌクレオチドバーコード；逆転写酵素；標的結合領域またはその一部分を含むプレートスイッチングオリゴヌクレオチド；ならびに5'から3'へのエキソヌクレアーゼ活性および3'から5'へのエキソヌクレアーゼ活性のうち少なくとも1つを欠くDNAポリメラーゼを含む。

【0021】

一部の実施形態では、逆転写酵素は、ウイルス逆転写酵素、例えば、マウス白血病ウイルス(MLV)逆転写酵素またはモロニー Maus 白血病ウイルス(MMLV)逆転写酵素を含む。一部の実施形態では、プレートスイッチオリゴヌクレオチドは、1つまたは複数の3'リボヌクレオチドを含む。一部の実施形態では、プレートスイッチオリゴヌクレオチドは、3つの3'リボヌクレオチドを含む。一部の実施形態では、3'リボヌクレオチドは、グアニンを含む。キットは、エチレングリコール、ポリエチレングリコール、1,2-プロパンジオール、ジメチルスルホキシド(DMSO)、グリセロール、ホルムアミド、7-デアザ-GTP、アセトアミド、塩化テトラメチルアンモニウム塩、ベタイン、またはこれらの任意の組合せのうち1つまたは複数を含み得る。

20

【0022】

本明細書の開示は、キットを含む。一部の実施形態では、キットは、複数のオリゴヌクレオチドバーコードであって、複数のオリゴヌクレオチドバーコードのそれぞれが、分子標識および標的結合領域を含み、複数のオリゴヌクレオチドバーコードのうち少なくとも10個が異なる分子標識配列を含む、複数のオリゴヌクレオチドバーコード；末端デオキシヌクレオチド転移酵素；ならびに5'から3'へのエキソヌクレアーゼ活性および3'から5'へのエキソヌクレアーゼ活性のうち少なくとも1つを欠くDNAポリメラーゼを含む。一部の実施形態では、DNAポリメラーゼは、Klenow断片を含む。キットは、緩衝液、カートリッジ、または両方を含み得る。キットは、逆転写反応および/または増幅反応のための1種または複数の試薬を含み得る。一部の実施形態では、標的結合領域は、遺伝子特異的配列、オリゴ(dT)配列、ランダム多量体、またはこれらの任意の組合せを含む。一部の実施形態では、オリゴヌクレオチドバーコードは、同一の試料標識および/または同一の細胞標識を含む。一部の実施形態では、複数のオリゴヌクレオチドバーコードの各試料標識および/または細胞標識は、少なくとも6個のヌクレオチドを含む。一部の実施形態では、複数のオリゴヌクレオチドバーコードの各分子標識は、少なくとも6個のヌクレオチドを含む。

30

【0023】

一部の実施形態では、複数のオリゴヌクレオチドバーコードのうち少なくとも1つは、合成粒子上に固定化されているかもしくは部分的に固定化されている、または合成粒子内に封入されているかもしくは部分的に封入されている。一部の実施形態では、合成粒子は、崩壊可能(例えば、崩壊可能ハイドロゲル粒子)である。一部の実施形態では、合成粒子は、ビーズを含む。一部の実施形態では、ビーズは、セファロースビーズ、ストレプトアビジンビーズ、アガロースビーズ、磁性ビーズ、コンジュゲートビーズ、プロテインAコンジュゲートビーズ、プロテインGコンジュゲートビーズ、プロテインA/Gコンジュゲートビーズ、プロテインLコンジュゲートビーズ、オリゴ(dT)コンジュゲートビーズ、シリカビーズ、シリカ様ビーズ、抗ピオチンマイクロビーズ、抗蛍光色素マイクロビーズ、またはこれらの任意の組合せを含む。一部の実施形態では、合成粒子は、ポリジメチルシロキサン(PDMS)、ポリスチレン、ガラス、ポリプロピレン、アガロース、ゼラチン、ハイドロゲル、常磁性体、セラミック、プラスチック、ガラス、メチルステレ

40

50

ン、アクリルポリマー、チタン、ラテックス、セファロース、セルロース、ナイロン、シリコン、およびこれらの任意の組合せからなる群から選択される材料を含む。一部の実施形態では、複数のオリゴヌクレオチドバーコードのそれぞれは、リンカー官能基を含み、合成粒子は、固体支持体官能基を含み、ならびに支持体官能基およびリンカー官能基は互いに会合している。一部の実施形態では、リンカー官能基および支持体官能基は、独立して、C6、ピオチン、ストレプトアビジン、第一級アミン、アルデヒド、ケトン、およびこれらの任意の組合せからなる群から選択される。

【図面の簡単な説明】

【0024】

【図1】非限定的な例示的バーコードを示す図である。

10

【図2】バーコード化および電子計数の非限定的な例示的ワークフローを示す図である。

【図3】複数の標的から3'末端でバーコード化された標的のインデックス化されたライブラリーを生成するための非限定的な例示的プロセスを示す略図である。

【図4A】5'末端において核酸標的を遺伝子特異的に標識する非限定的な例示的方法の略図を示す。

【図4B】5'末端において核酸標的を遺伝子特異的に標識する非限定的な例示的方法の略図を示す。

【図5A】全トランスクリプトーム分析のために5'末端において核酸標的を標識する非限定的な例示的方法の略図を示す。

【図5B】全トランスクリプトーム分析のために5'末端において核酸標的を標識する非限定的な例示的方法の略図を示す。

20

【図6A】核酸標的（例えば、免疫受容体mRNA）の完全長配列を決定する非限定的な例示的ワークフローの略図を示す。

【図6B】核酸標的（例えば、免疫受容体mRNA）の完全長配列を決定する非限定的な例示的ワークフローの略図を示す。

【図6C】核酸標的（例えば、免疫受容体mRNA）の完全長配列を決定する非限定的な例示的ワークフローの略図を示す。

【図6D】核酸標的（例えば、免疫受容体mRNA）の完全長配列を決定する非限定的な例示的ワークフローの略図を示す。

【図6E】核酸標的（例えば、免疫受容体mRNA）の完全長配列を決定する非限定的な例示的ワークフローの略図を示す。

30

【図6F】核酸標的（例えば、免疫受容体mRNA）の完全長配列を決定する非限定的な例示的ワークフローの略図を示す。

【図6G】核酸標的（例えば、免疫受容体mRNA）の完全長配列を決定する非限定的な例示的ワークフローの略図を示す。

【図6H】核酸標的（例えば、免疫受容体mRNA）の完全長配列を決定する非限定的な例示的ワークフローの略図を示す。

【図6I】核酸標的（例えば、免疫受容体mRNA）の完全長配列を決定する非限定的な例示的ワークフローの略図を示す。

【図6J】核酸標的（例えば、免疫受容体mRNA）の完全長配列を決定する非限定的な例示的ワークフローの略図を示す。

40

【図6K】核酸標的（例えば、免疫受容体mRNA）の完全長配列を決定する非限定的な例示的ワークフローの略図を示す。

【図6L】核酸標的（例えば、免疫受容体mRNA）の完全長配列を決定する非限定的な例示的ワークフローの略図を示す。

【図6M】核酸標的（例えば、免疫受容体mRNA）の完全長配列を決定する非限定的な例示的ワークフローの略図を示す。

【図6N】核酸標的（例えば、免疫受容体mRNA）の完全長配列を決定する非限定的な例示的ワークフローの略図を示す。

【図6O】核酸標的（例えば、免疫受容体mRNA）の完全長配列を決定する非限定的な

50

例示的ワークフローの略図を示す。

【図 7】完全長発現プロファイリングを実施する非限定的な例示的ワークフローの略図である。

【図 8】本明細書において提供される方法に従って生成された非限定的な例示的バイオアナライザーによる追跡を示す。

【図 9 A】免疫受容体 mRNA の発現プロファイリングの非限定的な例示的略図を示す。

【図 9 B】免疫受容体 mRNA の発現プロファイリングの非限定的な例示的略図を示す。

【図 10 A】本明細書において提供されるランダムプライミングベースの方法に関するデータを示す。現在利用可能な方法（図 10 B および 10 D）と比較した、本明細書において提供されるランダムプライミングベースのアプローチ（図 10 A および 10 C）で得られた TCR 産物（図 10 A ~ 10 B）および BCR 産物（図 10 C ~ 10 D）のバイオアナライザーによる追跡が示される。

10

【図 10 B】本明細書において提供されるランダムプライミングベースの方法に関するデータを示す。現在利用可能な方法（図 10 B および 10 D）と比較した、本明細書において提供されるランダムプライミングベースのアプローチ（図 10 A および 10 C）で得られた TCR 産物（図 10 A ~ 10 B）および BCR 産物（図 10 C ~ 10 D）のバイオアナライザーによる追跡が示される。

【図 10 C】本明細書において提供されるランダムプライミングベースの方法に関するデータを示す。現在利用可能な方法（図 10 B および 10 D）と比較した、本明細書において提供されるランダムプライミングベースのアプローチ（図 10 A および 10 C）で得られた TCR 産物（図 10 A ~ 10 B）および BCR 産物（図 10 C ~ 10 D）のバイオアナライザーによる追跡が示される。

20

【図 10 D】本明細書において提供されるランダムプライミングベースの方法に関するデータを示す。現在利用可能な方法（図 10 B および 10 D）と比較した、本明細書において提供されるランダムプライミングベースのアプローチ（図 10 A および 10 C）で得られた TCR 産物（図 10 A ~ 10 B）および BCR 産物（図 10 C ~ 10 D）のバイオアナライザーによる追跡が示される。

【図 11 A - 11 B】本明細書において提供されるランダムプライミングベースの方法に関するデータを示す。全長である（図 11 B）および細胞に由来する（図 11 A）VDJ 配列のパーセンテージは、現在利用可能な方法と比較して、本明細書において提供されるランダムプライミングベースのアプローチ（- Pre t）に関して示される。

30

【図 12 A - 12 D】本明細書において提供されるランダムプライミングベースの方法に関するデータを示す。現在利用可能な方法（図 12 B および 12 D）と比較した、本明細書において提供されるランダムプライミングベースのアプローチ（図 12 A および 12 C）で得られた BCR 対合（図 12 A ~ 12 B）および TCR 対合（図 12 C ~ 12 D）が示される。

【発明を実施するための形態】

【0025】

以下の詳細な説明において、本明細書の一部を形成する添付の図面への参照がなされる。図面では、類似の記号は、文脈により別途示されない限り、典型的には類似の構成要素を特定する。詳細な説明、図面、および特許請求の範囲に記載される例示的な実施形態は、制限することを意味するものではない。本明細書に提示される主題の趣旨または範囲から逸脱することなく、他の実施形態を用いることができ、他の変更を行うことができる。概して本明細書に記載され、図面で図示されている本開示の態様は、広範な異なる構成で配置、置換え、組合せ、分離、および設計がなされてもよく、それらのすべてが、本明細書において明示的に企図され、本明細書の開示の一部をなすことが、容易に理解される。

40

本明細書において参照されるすべての特許、公開特許出願、他の刊行物、および Gen Bank からの配列、ならびに他のデータベースは、関連技術に関してその全体が参照により組み込まれる。

【0026】

50

少数の核酸、例えば、メッセンジャーリボヌクレオチド酸 (mRNA) 分子を定量することは、例えば、異なる発生段階にあるかまたは異なる環境条件下にある細胞において発現される遺伝子を決定するために、臨床上重要である。しかしながら、特に分子の数が非常に少ない場合には、核酸分子 (例えば、mRNA 分子) の絶対数を決定することも非常に困難であり得る。試料中の分子の絶対数を決定する 1 つの方法は、デジタルポリメラーゼ連鎖反応 (PCR) である。理想的には、PCR により、それぞれのサイクルで同一の分子コピーが得られる。しかしながら、PCR は、それぞれの分子が確率的な見込みで複製し、この見込みは、PCR サイクルおよび遺伝子配列によって変動し、増幅バイアスおよび不正確な遺伝子発現測定値をもたらすため、欠点を有し得る。固有の分子標識 (分子インデックス (MI) とも称される) を有する確率的バーコードを使用して、分子数を計数し、増幅バイアスを補正することができる。Precise (商標) アッセイ (Cellular Research, Inc. (Palo Alto, CA)) および Rhapsody (商標) アッセイ (Becton, Dickinson and Company (Franklin Lakes, NJ)) などの確率的バーコード化によって、PCR および逆転写 (RT) の間に mRNA を標識するために分子標識 (ML) を使用することによるライブラリー調製ステップによって誘導されるバイアスを補正することができる。

10

【0027】

Precise (商標) アッセイは、ポリ (T) オリゴヌクレオチドに、多数の、例えば、6561 ~ 65536 個の固有の分子標識配列を有する確率的バーコードの枯渇しないプールを利用して、RT ステップの間に試料中のすべてのポリ (A) - mRNA にハイブリダイズさせることができる。確率的バーコードは、ユニバーサル PCR プライミング部位を含み得る。RT の間に、標的遺伝子分子は、確率的バーコードとランダムに反応する。それぞれの標的分子は、確率的バーコードにハイブリダイズし、その結果、確率的にバーコード化された相補リボヌクレオチド酸 (cDNA) 分子を生成し得る。標識した後、マイクロウェルプレートのマイクロウェルから得られた確率的にバーコード化された cDNA 分子を、PCR 増幅およびシーケンシングのために単一のチューブにプールしてもよい。生のシーケンシングデータを分析して、リードの数、固有の分子標識配列を有する確率的バーコードの数、および mRNA 分子の数を得ることができる。

20

【0028】

本明細書の開示は、試料中の核酸標的を標識する方法を含む。一部の実施形態では、本方法は、核酸標的のコピーを複数のオリゴヌクレオチドバーコードと接触させるステップであって、各オリゴヌクレオチドバーコードが、第 1 のユニバーサル配列、分子標識、および核酸標的にハイブリダイズすることが可能な標的結合領域を含む、ステップと；それぞれが、標的結合領域および該標的結合領域の相補体を含む複数のバーコード化された核酸分子を生成するステップと；各バーコード化された核酸分子の標的結合領域の相補体を、(i) 複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコード、(ii) バーコード化された核酸分子自体、および (iii) 複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子のうちの 1 つまたは複数の標的結合領域とハイブリダイズさせるステップと；複数のバーコード化された核酸分子の 3' 末端を伸長させて、それぞれが、第 1 の分子標識および第 2 の分子標識を含む、複数の伸長したバーコード化された核酸分子を生成するステップと；核酸標的の配列にハイブリダイズすることが可能な標的特定のプライマーおよび第 1 のユニバーサル配列を含むプライマーを使用して、複数の伸長したバーコード化された核酸分子を増幅させて、それによって、核酸標的の配列、またはその一部分を含む第 1 の複数のバーコード化されたアンプリコンを生成するステップと；ランダムプライマーを第 1 の複数のバーコード化されたアンプリコンにハイブリダイズさせ、ランダムプライマーを伸長させて、複数の伸長産物を生成するステップであって、ランダムプライマーが、第 2 のユニバーサル配列、またはその相補体を含む、ステップと；第 1 のユニバーサル配列および第 2 のユニバーサル配列、またはその相補体にハイブリダイズすることが可能なプライマーを使用して、複数の伸長産物を増

30

40

50

幅させて、それによって、第2の複数のバーコード化されたアンプリコンを生成するステップとを含む。

【0029】

本明細書の開示は、試料中の核酸標的の数を決定する方法を含む。一部の実施形態では、本方法は、核酸標的のコピーを複数のオリゴヌクレオチドバーコードと接触させるステップであって、各オリゴヌクレオチドバーコードが、第1のユニバーサル配列、分子標識（例えば、第1の分子標識）、および核酸標的にハイブリダイズすることが可能な標的結合領域を含む、ステップと；それぞれが、標的結合領域および該標的結合領域の相補体を含む複数のバーコード化された核酸分子を生成するステップと；各バーコード化された核酸分子の標的結合領域の相補体を、(i)複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコード、(ii)バーコード化された核酸分子自体、および(iii)複数のバーコード化された核酸分子のうち異なるバーコード化された核酸分子のうち1つまたは複数の標的結合領域とハイブリダイズさせるステップと；複数のバーコード化された核酸分子の3'末端を伸長させて、それぞれが、第1の分子標識および第2の分子標識を含む、複数の伸長したバーコード化された核酸分子を生成するステップと；核酸標的の配列にハイブリダイズすることが可能な標的特定のプライマーおよび第1のユニバーサル配列を含むプライマーを使用して、複数の伸長したバーコード化された核酸分子を増幅させて、それによって、核酸標的の配列、またはその一部分を含む第1の複数のバーコード化されたアンプリコンを生成するステップと；ランダムプライマーを第1の複数のバーコード化されたアンプリコンにハイブリダイズさせ、ランダムプライマーを伸長させて、複数の伸長産物を生成するステップであって、ランダムプライマーが、第2のユニバーサル配列、またはその相補体を含む、ステップと；第1のユニバーサル配列および第2のユニバーサル配列、またはその相補体にハイブリダイズすることが可能なプライマーを使用して、複数の伸長産物を増幅させて、それによって、第2の複数のバーコード化されたアンプリコンを生成するステップと；第2の複数のバーコード化されたアンプリコン、またはその産物と会合した、別個の配列を有する第1の分子標識、別個の配列を有する第2の分子標識、またはこれらの組合せの数に基づいて、試料中の核酸標的のコピー数を決定するステップとを含む。

【0030】

本明細書の開示は、キットを含む。一部の実施形態では、キットは、複数のオリゴヌクレオチドバーコードであって、複数のオリゴヌクレオチドバーコードのそれぞれが、分子標識および標的結合領域を含み、複数のオリゴヌクレオチドバーコードのうち少なくとも10個が異なる分子標識配列を含む、複数のオリゴヌクレオチドバーコード；逆転写酵素；標的結合領域またはその一部分を含むプレートスイッチングオリゴヌクレオチド；ならびに5'から3'へのエキソヌクレアーゼ活性および3'から5'へのエキソヌクレアーゼ活性のうち少なくとも1つを欠くDNAポリメラーゼを含む。

【0031】

本明細書の開示は、キットを含む。一部の実施形態では、キットは、複数のオリゴヌクレオチドバーコードであって、複数のオリゴヌクレオチドバーコードのそれぞれが、分子標識および標的結合領域を含み、複数のオリゴヌクレオチドバーコードのうち少なくとも10個が異なる分子標識配列を含む、複数のオリゴヌクレオチドバーコード；末端デオキシヌクレオチド転移酵素；ならびに5'から3'へのエキソヌクレアーゼ活性および3'から5'へのエキソヌクレアーゼ活性のうち少なくとも1つを欠くDNAポリメラーゼを含む。

【0032】

定義

別途定義されない限り、本明細書に使用される技術用語および科学用語は、本開示が属する分野の当業者によって広く理解されているものと同じ意味を有する。例えば、Singleton et al., Dictionary of Microbiology and Molecular Biology 2nd ed., J. Wiley & Sons (New York, NY 1994)、Sambrook et al., Molecular Cloning, A Laboratory

10

20

30

40

50

ry Manual, Cold Spring Harbor Press (Cold Spring Harbor, NY 1989)を参照されたい。本開示の目的で、次の用語を、以下に定義する。

【0033】

本明細書で使用される場合、「アダプター」という用語は、会合した核酸の増幅またはシーケンシングを容易にするための配列を意味し得る。会合した核酸は、標的核酸を含み得る。会合した核酸は、空間標識、標的標識、試料標識、インデックス化標識、またはバーコード配列（例えば、分子標識）のうちの1つまたは複数を含み得る。アダプターは、線形であり得る。アダプターは、事前にアデニル化されたアダプターであってもよい。アダプターは、二本鎖であっても一本鎖であってもよい。1つまたは複数のアダプターが、核酸の5'末端または3'末端に配置され得る。アダプターが5'末端および3'末端に公知の10配列を含む場合、公知の配列は、同じ配列であっても異なる配列であってもよい。ポリヌクレオチドの5'末端および/または3'末端に配置されるアダプターは、表面に固定化された1つまたは複数のオリゴヌクレオチドにハイブリダイズすることが可能であり得る。アダプターは、一部の実施形態では、ユニバーサル配列を含み得る。ユニバーサル配列は、2つ以上の核酸分子に共通しているヌクレオチド配列の領域であり得る。2つ以上の核酸分子はまた、異なる配列の領域を有し得る。したがって、例えば、5'アダプターは、同一および/またはユニバーサル核酸配列を含み得、3'アダプターは、同一および/またはユニバーサル配列を含み得る。複数の核酸分子の異なるメンバーにユニバーサル配列が存在し得ることにより、ユニバーサル配列に相補的である単一のユニバーサルプライマーを使用した複数の異なる配列の複製または増幅が可能となり得る。同様に、少なくとも1つ、2つ（例えば、1対）、またはそれ以上のユニバーサル配列が、核酸分子の集合体の異なるメンバーに存在し得ることにより、ユニバーサル配列に相補的である少なくとも1つ、2つ（例えば、1対）、またはそれ以上の単一のユニバーサルプライマーを使用した複数の異なる配列の複製または増幅が可能となり得る。したがって、ユニバーサルプライマーには、そのようなユニバーサル配列にハイブリダイズすることができる配列が含まれる。標的核酸配列を有する分子を、異なる標的核酸配列の一方または両方の末端にユニバーサルアダプター（例えば、非標的核酸配列）を付加するように改変してもよい。標的核酸に結合した1つまたは複数のユニバーサルプライマーは、ユニバーサルプライマーのハイブリダイゼーションのための部位を提供し得る。標的核酸に結合した1つまたは複数のユニバーサルプライマーは、互いに同じであっても異なってもよい。

20

30

【0034】

本明細書で使用される場合、「会合した」または「と会合した」という用語は、2つ以上の種が、ある時点で一緒に位置しているとして特定可能であることを意味し得る。会合は、2つ以上の種が、類似の容器内にあるかまたはそこにあったことを意味し得る。会合は、情動的会合であってもよい。例えば、2つ以上の種に関するデジタル情報は、記憶され得、種のうちの1つまたは複数がある時点で一緒に位置していたことを決定するために使用され得る。会合はまた、物理的会合であってもよい。一部の実施形態では、2つ以上の会合した種は、互いにまたは共通の固体もしくは半固体表面に「係留されている」、「結合している」、または「固定化されている」。会合は、ビーズなどの固体または半固体支持体に標識を結合させるための共有結合的または非共有結合的手段を指し得る。会合は、標的と標識との間の共有結合であってもよい。会合は、2つの分子（例えば、標的分子および標識）間のハイブリダイゼーションを含み得る。

40

【0035】

本明細書で使用される場合、「相補的」という用語は、2つのヌクレオチド間の正確な対合の能力を指し得る。例えば、核酸の所与の位置にあるヌクレオチドが、別の核酸のヌクレオチドと水素結合することができる場合、2つの核酸は、その位置において互いに相補的であると考えられる。2つの一本鎖核酸分子間の相補性は、ヌクレオチドのうちのいくつかが結合する「部分的」であってもよく、または一本鎖分子間に全相補性が存在する完全なものであってもよい。第1のヌクレオチド配列は、第1のヌクレオチド配列が第2のヌクレオチド配列に相補的である場合、第2の配列の「相補体」とであると称され得

50

る。第1のヌクレオチド配列は、第1のヌクレオチド配列が第2の配列の逆転物である（すなわち、ヌクレオチドの順序が逆転している）配列に相補的である場合、第2の配列の「逆相補体」と称され得る。本明細書で使用される場合、「相補」配列は、配列の「相補体」または「逆相補体」を指し得る。ある分子が別の分子にハイブリダイズすることができる場合、それは、ハイブリダイズしている分子に相補的であるか、または部分的に相補的であり得ることが、本開示から理解される。

【0036】

本明細書で使用される場合、「デジタル計数」という用語は、試料中の標的分子の数を概算するための方法を指し得る。デジタル計数は、試料中の標的と会合している固有の標識の数を決定するステップを含み得る。本質的に確率的であり得るこの方法論は、分子計数の問題を、同一の分子の位置決定および同定のうちの1つを、事前に定義した標識セットの検出に関する一連のはい/いい式のデジタル質問に変換する。

10

【0037】

本明細書で使用される場合、「標識」または「複数の標識」という用語は、試料内の標的と会合した核酸コードを指し得る。標識は、例えば、核酸標識であり得る。標識は、全体的または部分的に増幅可能な標識であってもよい。標識は、全体的または部分的にシーケンシング可能な標識であってもよい。標識は、異なるものとして特定可能な天然の核酸の一部であってもよい。標識は、公知の配列であってもよい。標識は、核酸配列の接合部、例えば、天然および非天然の配列の接合部を含み得る。本明細書で使用される場合、「標識」という用語は、「インデックス」、「タグ」、または「標識タグ」という用語と互換可能に使用することができる。標識は、情報を運ぶことができる。例えば、様々な実施形態では、標識は、試料の同一性、試料の供給源、細胞の同一性、および/または標的を判定するために使用することができる。

20

【0038】

本明細書で使用される場合、「枯渇しないリザーバ」という用語は、多数の異なる標識で構成されるバーコード（例えば、確率的バーコード）のプールを指し得る。枯渇しないリザーバは、枯渇しないリザーバを標的のプールと会合させたときに、それぞれの標的が固有のバーコードと会合する可能性が高くなるように、多数の異なるバーコードを含み得る。それぞれの標識した標的分子の固有性は、ランダム選択の統計学によって決定することができ、標識の多様性と比較して、集団内の同一な標的分子のコピー数に依存する。結果として得られる標識した標的分子のサイズは、バーコード化プロセスの確率的性質によって決定することができ、次いで、検出されたバーコードの数の分析により、もともとの集団または試料中に存在している標的分子数の計算が可能となる。固有のバーコードの数に対する存在する標的分子の数の比が、低い場合、標識した標的分子は、固有性が高い（すなわち、1つを上回る標的分子が1つの所与の標識で標識されることになる確率は非常に低い）。

30

【0039】

本明細書で使用される場合、「核酸」という用語は、ポリヌクレオチド配列またはその断片を指す。核酸は、ヌクレオチドを含み得る。核酸は、細胞に対して外因性であっても内因性であってもよい。核酸は、無細胞環境で存在してもよい。核酸は、遺伝子またはその断片であり得る。核酸は、DNAであり得る。核酸は、RNAであり得る。核酸は、1つまたは複数のアナログ（例えば、改変された骨格、糖、または核酸塩基）を含み得る。アナログの一部の非限定的な例としては、5-プロモウラシル、ペプチド核酸、ゼノ核酸、モルホリノ、ロックド核酸、グリコール核酸、トレオース核酸、ジデオキシヌクレオチド、コーディセピン、7-デアザ-GTP、フルオロフォア（例えば、糖に連結したローダミンまたはフルオレセイン）、チオール含有ヌクレオチド、ビオチン連結ヌクレオチド、蛍光塩基アナログ、CpGアイランド、メチル-7-グアノシン、メチル化ヌクレオチド、イノシン、チオウリジン、シュードウリジン、ジヒドロウリジン、キューオシン、およびワイオシンが挙げられる。「核酸」、「ポリヌクレオチド」、「標的ポリヌクレオチド」、および「標的核酸」は、互換可能に使用することができる。

40

50

【0040】

核酸は、新しいかまたは亢進された特性（例えば、改善された安定性）を有する核酸をもたらし、1つまたは複数の改変（例えば、塩基改変、骨格改変）を含んでもよい。核酸は、核酸親和性タグを含んでもよい。ヌクレオシドは、塩基-糖の組合せであり得る。ヌクレオシドの塩基部分は、複素環式塩基であり得る。そのような複素環式塩基の2つのもっとも一般的なクラスは、プリンおよびピリミジンである。ヌクレオチドは、ヌクレオシドの糖部分に共有結合で連結されたリン酸基をさらに含んだヌクレオシドであり得る。ペントフラノシル糖を含むヌクレオシドについては、リン酸基は、糖の2'、3'、または5'ヒドロキシル部分に連結され得る。核酸を形成する際、リン酸基は、隣接するヌクレオシドを互いに共有結合的に連結させて、線形ポリマー化合物を形成し得る。次に、この線形ポリマー化合物のそれぞれの末端が、さらに結合されて、環状化合物が形成されるが、線形化合物が、一般的には好適である。加えて、線形化合物は、内部ヌクレオチド塩基相補性を有し得、したがって、完全または部分的に二本鎖の化合物が生じるような様式でフォールディングし得る。核酸内において、リン酸基は、一般に、核酸のヌクレオチド間骨格を形成すると称され得る。連結または骨格は、3'と5'のホスホジエステル連結であり得る。

10

【0041】

核酸は、改変された骨格および/または改変されたヌクレオチド間連結を含み得る。改変された骨格には、骨格にリン原子を保持するものおよび骨格にリン原子を有さないものが含まれ得る。リン原子を中に含む好適な改変された核酸骨格としては、例えば、ホスホロチオエート、キラルホスホロチオエート、ホスホロジチオエート、ホスホトリエステル、アミノアルキルホスホトリエステル、メチルおよび他のアルキルホスホネート、例えば、3'-アルキレンホスホネート、5'-アルキレンホスホネート、キラルホスホネート、ホスフィネート、ホスホロアミデート、例えば、3'-アミノホスホロアミデートおよびアミノアルキルホスホロアミデート、ホスホロジアミデート、チオノホスホロアミデート、チオノアルキルホスホネート、チオノアルキルホスホトリエステル、セレノホスフェート、およびノルマル3'-5'連結を有するボラノホスフェート、2'-5'連結アナログ、ならびに逆極性を有するものを挙げることができ、ここで、1つまたは複数のヌクレオチド間連結は、3'と3'、5'と5'、または2'と2'の連結である。

20

【0042】

核酸は、短鎖アルキルもしくはシクロアルキルヌクレオチド間連結、混合ヘテロ原子およびアルキルもしくはシクロアルキルヌクレオチド間連結、または1つもしくは複数のヘテロ原子もしくは複素環式ヌクレオチド間連結によって形成されるポリヌクレオチド骨格を含み得る。これらには、モルホリノ連結（ヌクレオチドの糖部分から部分的に形成される）を有するもの、シロキサ骨格、スルフィド、スルホキシド、およびスルホン骨格、ホルムアセチルおよびチオホルムアセチル骨格、メチレンホルムアセチルおよびチオホルムアセチル骨格、リポアセチル骨格、アルケン含有骨格、スルファメート骨格、メチレンイミノおよびメチレンヒドラジノ骨格、スルホネートおよびスルホンアミド骨格、アミド骨格、ならびに混合N、O、S、およびCH₂成分部分を有するその他のものを挙げることができる。

30

【0043】

核酸は、核酸模倣体を含み得る。「模倣体」という用語は、フラノース環のみ、またはフラノース環およびヌクレオチド間連結の両方が、非フラノース基で置き換えられたポリヌクレオチドを含むことを意図し得、フラノース環のみの置換えはまた、糖代理物であると称され得る。複素環式塩基部分または改変された複素環式塩基部分は、適切な標的核酸とのハイブリダイゼーションのために維持され得る。1つのそのような核酸は、ペプチド核酸（PNA）であり得る。PNAでは、ポリヌクレオチドの糖骨格は、アミド含有骨格、特に、アミノエチルグリシン骨格で置き換えられ得る。ヌクレオチドは保持されてもよく、骨格のアミド部分のアザ窒素原子に直接的または間接的に結合される。PNA化合物における骨格は、2つ以上の連結されたアミノエチルグリシン単位を含み得、これにより

40

50

、PNAにアミド含有骨格が提供される。複素環式塩基部分は、骨格のアミド部分のアザ窒素原子に直接的または間接的に結合され得る。

【0044】

核酸は、モルホリノ骨格構造を含み得る。例えば、核酸は、リボース環の代わりに6員モルホリノ環を含み得る。これらの実施形態のうちいくつかでは、ホスホロジアミデートまたは他の非ホスホジエステルヌクレオシド間連結により、ホスホジエステル連結を置き換えることができる。

【0045】

核酸は、モルホリノ環に結合した複素環式塩基を有する連結されたモルホリノ単位（例えば、モルホリノ核酸）を含み得る。連結基は、モルホリノ核酸においてモルホリノモノマー単位を連結し得る。非イオン性モルホリノベースのオリゴマー化合物は、細胞内タンパク質との望ましくない相互作用をあまり有さない可能性がある。モルホリノベースのポリヌクレオチドは、核酸の非イオン性模倣体であり得る。モルホリノクラス内の様々な化合物は、異なる連結基を使用して結合することができる。ポリヌクレオチド模倣体のさらなるクラスは、シクロヘキセニル核酸（CeNA）と称され得る。核酸分子に通常存在するフラノース環を、シクロヘキセニル環で置き換えることができる。CeNA DMTで保護されたホスホロアミダイトモノマーを調製し、ホスホロアミダイトケミストリーを使用したオリゴマー化合物の合成に使用することができる。CeNAモノマーを核酸鎖に組み込むことにより、DNA/RNAハイブリッドの安定性を増加させることができる。CeNAオリゴアデニレートは、天然の複合体に類似する安定性で、核酸相補体と複合体を形成することができる。さらなる改変には、2'-ヒドロキシル基が糖環の4'炭素原子に連結され、それによって2'-C, 4'-C-オキシメチレン連結が形成され、それによって二環式糖部分が形成される、ロックド核酸（LNA）を挙げることができる。連結は、2'酸素原子および4'炭素原子を架橋するメチレン（-C₂H）基であり得、ここで、nは、1または2である。LNAおよびLNAアナログは、相補性核酸との非常に高い二重鎖熱安定性（ $T_m = +3 \sim +10$ ）、3'-エクソヌクレアーゼ分解に対する安定性、および良好な可溶性特性を示し得る。

【0046】

核酸はまた、核酸塩基（単純に「塩基」と称されることが多い）改変または置換も含み得る。本明細書で使用される場合、「改変されていない」または「天然の」核酸塩基としては、プリン塩基（例えば、アデニン（A）およびグアニン（G））、ならびにピリミジン塩基（例えば、チミン（T）、シトシン（C）、およびウラシル（U））を挙げることができる。改変された核酸塩基としては、他の合成および天然の核酸塩基、例えば、5-メチルシトシン（5-me-C）、5-ヒドロキシメチルシトシン、キサントシン、ヒポキサントシン、2-アミノアデニン、アデニンおよびグアニンの6-メチルおよび他のアルキル誘導体、アデニンおよびグアニンの2-プロピルおよび他のアルキル誘導体、2-チオウラシル、2-チオチミンおよび2-チオシトシン、5-ハロウラシルおよびシトシン、ピリミジン塩基の5-プロピニル（-C=C-CH₃）ウラシルおよびシトシン、ならびに他のアルキル誘導体、6-アゾウラシル、シトシンおよびチミン、5-ウラシル（シュードウラシル）、4-チオウラシル、8-ハロ、8-アミノ、8-チオール、8-チオールキル、8-ヒドロキシル、および他の8-置換アデニンおよびグアニン、5-ハロ、特に5-プロモ、5-トリフルオロメチル、および他の5-置換ウラシルおよびシトシン、7-メチルグアニンおよび7-メチルアデニン、2-F-アデニン、2-アミノアデニン、8-アザグアニンおよび8-アザアデニン、7-デアザグアニンおよび7-デアザアデニン、ならびに3-デアザグアニンおよび3-デアザアデニンを挙げることができる。改変された核酸塩基としては、三環式ピリミジン、例えば、フェノキサジンシチジン（1H-ピリミド（5, 4-b）（1, 4）ベンゾキサジン-2（3H）-オン）、フェノチアジンシチジン（1H-ピリミド（5, 4-b）（1, 4）ベンゾチアジン-2（3H）-オン）、G-クランプ、例えば、置換フェノキサジンシチジン（例えば、9-（2-アミノエトキシ）-H-ピリミド（5, 4-（b）（1, 4）ベンゾキサジン-2（3H）-

10

20

30

40

50

オン)、フェノチアジンシチジン(1H-ピリミド(5,4-b)(1,4)ベンゾチアジン-2(3H)-オン)、G-クランプ、例えば、置換フェノキサジンシチジン(例えば、9-(2-アミノエトキシ)-H-ピリミド(5,4-(b)(1,4)ベンゾキサジン-2(3H)-オン)、カルバゾールシチジン(2H-ピリミド(4,5-b)インドール-2-オン)、ピリドインドールシチジン(H-ピリド(3',2':4,5)ピロロ[2,3-d]ピリミジン-2-オン)を挙げることができる。

【0047】

本明細書で使用される場合、「試料」という用語は、標的を含む組成物を指し得る。開示される方法、デバイス、およびシステムによる分析に好適な試料としては、細胞、組織、器官、または生物が挙げられる。

10

本明細書で使用される場合、「試料採取デバイス」または「デバイス」という用語は、試料の切片を採取し、かつ/または基質上に切片を配置することができる、デバイスを指し得る。試料デバイスは、例えば、蛍光活性化細胞分取(FACS)機、細胞分取機、生検ニードル、生検デバイス、組織切片化デバイス、マイクロ流体デバイス、ブレードグリッド、および/またはマイクロトームを指し得る。

【0048】

本明細書で使用される場合、「固体支持体」という用語は、複数のバーコード(例えば、確率的バーコード)が結合され得る別々の固体または半固体表面を指し得る。固体支持体は、核酸を(例えば、共有結合的または非共有結合的に)固定化することができる、プラスチック、セラミック、金属、またはポリマー材料(例えば、ハイドロゲル)で構成される任意の種類(例えば、多孔質、または中空の球体、ボール、ベアリング、シリンダー、または他の類似の構成)を包含し得る。固体支持体は、球状(例えば、マイクロスフェア)であり得る別個の粒子を含み得るか、または非球状もしくは不規則な形状、例えば、立方体、立方体様、ピラミッド形、円筒形、円錐形、長方形、もしくは円盤形などを有し得る。ビーズは、非球体の形状であり得る。アレイに離間配置された複数の固体支持体は、基質を含まない場合がある。固体支持体は、「ビーズ」という用語と互換可能に使用することができる。

20

【0049】

本明細書で使用される場合、「確率的バーコード」という用語は、本開示の標識を含むポリヌクレオチド配列を指し得る。確率的バーコードは、確率的バーコード化に使用することができるポリヌクレオチド配列であり得る。確率的バーコードは、試料内の標的を定量するために使用することができる。確率的バーコードは、標識を標的に会合させた後に発生し得るエラーを制御するために使用することができる。例えば、確率的バーコードは、増幅またはシーケンシングエラーを評価するために使用することができる。標的と会合した確率的バーコードは、確率的バーコード-標的または確率的バーコード-タグ-標的と称され得る。

30

【0050】

本明細書で使用される場合、「遺伝子特異的確率的バーコード」という用語は、標識および遺伝子特異的である標的結合領域を含むポリヌクレオチド配列を指し得る。確率的バーコードは、確率的バーコード化に使用することができるポリヌクレオチド配列であり得る。確率的バーコードは、試料内の標的を定量するために使用することができる。確率的バーコードは、標識を標的に会合させた後に発生し得るエラーを制御するために使用することができる。例えば、確率的バーコードは、増幅またはシーケンシングエラーを評価するために使用することができる。標的と会合した確率的バーコードは、確率的バーコード-標的または確率的バーコード-タグ-標的と称され得る。

40

【0051】

本明細書で使用される場合、「確率的バーコード化」という用語は、核酸のランダムな標識化(例えば、バーコード化)を指し得る。確率的バーコード化は、ポアソン再帰戦略を利用して、標識を会合させ、標的と会合した標識を定量することができる。本明細書で使用される場合、「確率的バーコード化」という用語は、「確率的標識化」と互換可能に

50

使用することができる。

本明細書で使用される場合、「標的」という用語は、バーコード（例えば、確率的バーコード）と会合され得る組成物を指し得る。本開示の方法、デバイス、およびシステムによる分析に好適な例示的な標的としては、オリゴヌクレオチド、DNA、RNA、mRNA、マイクロRNA、tRNAなどが挙げられる。標的は、一本鎖であっても二本鎖であってもよい。一部の実施形態では、標的は、タンパク質、ペプチド、またはポリペプチドであってもよい。一部の実施形態では、標的は、脂質である。本明細書で使用される場合、「標的」は、「種」と互換可能に使用することができる。

【0052】

本明細書で使用される場合、「逆転写酵素」という用語は、逆転写酵素活性を有する（すなわち、RNA鋳型からのDNAの合成を触媒する）酵素の群を指し得る。一般に、そのような酵素としては、レトロウイルス逆転写酵素、レトロトランスポゾン逆転写酵素、レトロプラスミド逆転写酵素、レトロン逆転写酵素、細菌逆転写酵素、II群イントロン由来逆転写酵素、およびこれらの変異体、バリエーション、または誘導体が挙げられるが、これらに限定されない。非レトロウイルス逆転写酵素としては、非LTRレトロトランスポゾン逆転写酵素、レトロプラスミド逆転写酵素、レトロン逆転写酵素、およびII群イントロン逆転写酵素が挙げられる。II群イントロン逆転写酵素の例としては、ラクトコッカス・ラクティス (*Lactococcus lactis*) LI.LtrBイントロン逆転写酵素、サーモシネココッカス・エロンガタス (*Thermosynechococcus elongatus*) TeI4cイントロン逆転写酵素、またはゲオバチルス・ステアロサーモフィルス (*Geobacillus stearothermophilus*) GsI-IIcイントロン逆転写酵素が挙げられる。逆転写酵素の他のクラスとしては、非レトロウイルス逆転写酵素の多数のクラス（すなわち、とりわけ、レトロン、II群イントロン、および多様性発生レトロエレメント）を挙げるができる。

【0053】

「ユニバーサルアダプタープライマー」、「ユニバーサルプライマーアダプター」または「ユニバーサルアダプター配列」という用語は、バーコード（例えば、確率的バーコード）にハイブリダイズし、遺伝子特異的バーコードを生成するために使用することができるヌクレオチド配列を指すために、互換可能に使用される。ユニバーサルアダプター配列は、例えば、本開示の方法において使用されるすべてのバーコードにわたりユニバーサルである公知の配列であり得る。例えば、複数の標的が、本明細書に開示される方法を使用して標識される場合、標的特異的配列のそれぞれは、同じユニバーサルアダプター配列に連結され得る。一部の実施形態では、2つ以上のユニバーサルアダプター配列は、本明細書に開示される方法において使用することができる。例えば、複数の標的が、本明細書に開示される方法を使用して標識される場合、標的特異的配列のうち少なくとも2つは、異なるユニバーサルアダプター配列に連結される。ユニバーサルアダプタープライマーおよびその相補体は、2つのオリゴヌクレオチド中に含まれてもよく、そのうちの一方は標的特異的配列を含み、他方はバーコードを含む。例えば、ユニバーサルアダプター配列は、標的核酸に相補的であるヌクレオチド配列を生成するために、標的特異的配列を含むオリゴヌクレオチドの一部であってもよい。バーコードおよびユニバーサルアダプター配列の相補配列を含む第2のオリゴヌクレオチドは、ヌクレオチド配列とハイブリダイズし、標的特異的バーコード（標的特異的な確率的バーコード）を生成し得る。一部の実施形態では、ユニバーサルアダプタープライマーは、本開示の方法において使用されるユニバーサルPCRプライマーと異なる配列を有する。

【0054】

バーコード

バーコード化、例えば、確率的バーコード化は、例えば、米国特許出願公開第US2015/0299784号、国際公開第WO2015/031691号、およびFu et al, *Proc Natl Acad Sci U.S.A.* 2011 May 31;108(22):9026-31に記載されており、これらの刊行物の内容は、その全体が本明細書に組み込まれる。一部の実施形態では、本明細

10

20

30

40

50

書に開示されるバーコードは、標的を確率的に標識する（例えば、バーコード化する、タグ付けする）ために使用することができるポリヌクレオチド配列であり得る、確率的バーコードであり得る。バーコードは、確率的バーコードの異なるバーコード配列の数の、標識しようとする標的のうちの任意のものの発生の数に対する比が、1 : 1、2 : 1、3 : 1、4 : 1、5 : 1、6 : 1、7 : 1、8 : 1、9 : 1、10 : 1、11 : 1、12 : 1、13 : 1、14 : 1、15 : 1、16 : 1、17 : 1、18 : 1、19 : 1、20 : 1、30 : 1、40 : 1、50 : 1、60 : 1、70 : 1、80 : 1、90 : 1、100 : 1、またはこれらの値のうちのいずれか2つの間の数もしくは範囲であり得るか、またはおおよそこれらの値もしくはそのような数もしくは範囲であり得る場合に、確率的バーコードと称され得る。標的は、同一またはほぼ同一な配列を有するmRNA分子を含むmRNA種であり得る。バーコードは、確率的バーコードの異なるバーコード配列の数の、標識しようとする標的のうちの任意のものの発生の数に対する比が、少なくとも、または最大で1 : 1、2 : 1、3 : 1、4 : 1、5 : 1、6 : 1、7 : 1、8 : 1、9 : 1、10 : 1、11 : 1、12 : 1、13 : 1、14 : 1、15 : 1、16 : 1、17 : 1、18 : 1、19 : 1、20 : 1、30 : 1、40 : 1、50 : 1、60 : 1、70 : 1、80 : 1、90 : 1、100 : 1である場合に、確率的バーコードと称され得る。確率的バーコードのバーコード配列は、分子標識と称され得る。

【0055】

バーコード、例えば、確率的バーコードは、1つまたは複数の標識を含み得る。例示的な標識としては、ユニバーサル標識、細胞標識、バーコード配列（例えば、分子標識）、試料標識、プレート標識、空間標識、および/または空間前（pre-spatial）標識を挙げることができる。図1は、空間標識を有する例示的なバーコード104を示す。バーコード104は、バーコードを固体支持体105に連結させることができる5'アミンを含み得る。バーコードは、ユニバーサル標識、次元標識（dimension label）、空間標識、細胞標識、および/または分子標識を含み得る。バーコード内の異なる標識（ユニバーサル標識、次元標識、空間標識、細胞標識、および分子標識を含むがこれらに限定されない）の順序は、変動し得る。例えば、図1に示されるように、ユニバーサル標識がもっとも5'側の標識であってもよく、分子標識がもっとも3'側の標識であってもよい。空間標識、次元標識、および細胞標識は、任意の順序であってもよい。一部の実施形態では、ユニバーサル標識、空間標識、次元標識、細胞標識、および分子標識は、任意の順序である。バーコードは、標的結合領域を含み得る。標的結合領域は、試料内の標的（例えば、標的核酸、RNA、mRNA、DNA）と相互作用し得る。例えば、標的結合領域は、mRNAのポリ(A)尾部と相互作用することができるオリゴ(dT)配列を含み得る。一部の事例では、バーコードの標識（例えば、ユニバーサル標識、次元標識、空間標識、細胞標識、およびバーコード配列）は、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20ヌクレオチド、またはそれ以上離間していてもよい。

【0056】

標識、例えば、細胞標識は、エラー補正能力を提供するように設計され得る、定義された長さ、例えば、それぞれが7ヌクレオチドずつの（一部のハミングエラー補正コードにおいて使用されるビット数と同等の）固有の核酸部分配列セットを含み得る。7ヌクレオチドの配列を含むエラー補正用部分配列セットは、セット内の配列の任意の対での組合せが、定義された「遺伝子距離」（またはミスマッチ塩基数）を呈するように設計することができ、例えば、エラー補正用部分配列セットは、3ヌクレオチドの遺伝子距離を呈するように設計することができる。この場合、標識した標的核酸分子の配列データセットにおけるエラー補正配列の考察（以下により詳細に記載される）により、増幅またはシーケンシングエラーを検出または補正することが可能となり得る。一部の実施形態では、エラー補正コードを作成するために使用される核酸部分配列の長さは、変動し得、例えば、1、2、3、4、5、6、7、8、9、10、15、20、30、31、40、50ヌクレオチド、またはこれらの値のうちのいずれか2つの間の数もしくは範囲、またはおおよそこ

10

20

30

40

50

これらの値もしくはそのような数もしくは範囲のヌクレオチドの長さであり得る。一部の実施形態では、他の長さの核酸部分配列が、エラー補正コードを作成するために使用されてもよい。

【0057】

バーコードは、標的結合領域を含み得る。標的結合領域は、試料内の標的と相互作用し得る。標的は、リボ核酸 (RNA)、メッセンジャーRNA (mRNA)、マイクロRNA、低分子干渉RNA (siRNA)、RNA分解産物、それぞれがポリ(A)尾部を含むRNA、またはこれらの任意の組合せであり得るか、またはこれらを含み得る。一部の実施形態では、複数の標的は、デオキシリボ核酸 (DNA) を含み得る。

【0058】

一部の実施形態では、標的結合領域は、mRNAのポリ(A)尾部と相互作用することができるオリゴ(dT)配列を含み得る。バーコードの標識(例えば、ユニバーサル標識、次元標識、空間標識、細胞標識、およびバーコード配列(例えば、分子標識))のうちの1つまたは複数は、バーコードの残りの標識のうちの別の1つまたは2つから、スペーサーによって離間していてもよい。スペーサーは、例えば、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20ヌクレオチド、またはそれ以上であり得る。一部の実施形態では、バーコードの標識のいずれも、スペーサーによって離間していない。

【0059】

ユニバーサル標識

バーコードは、1つまたは複数のユニバーサル標識を含み得る。一部の実施形態では、1つまたは複数のユニバーサル標識は、所与の固体支持体に結合したバーコードのセットにおいて、すべてのバーコードについて同じであり得る。一部の実施形態では、1つまたは複数のユニバーサル標識は、複数のピースに結合したすべてのバーコードについて同じであり得る。一部の実施形態では、ユニバーサル標識は、シーケンシングプライマーにハイブリダイズすることができる核酸配列を含み得る。シーケンシングプライマーは、ユニバーサル標識を含むバーコードをシーケンシングするために使用することができる。シーケンシングプライマー(例えば、ユニバーサルシーケンシングプライマー)は、高スループットのシーケンシングプラットフォームと会合したシーケンシングプライマーを含み得る。一部の実施形態では、ユニバーサル標識は、PCRプライマーにハイブリダイズすることができる核酸配列を含み得る。一部の実施形態では、ユニバーサル標識は、シーケンシングプライマーおよびPCRプライマーにハイブリダイズすることができる核酸配列を含み得る。シーケンシングプライマーまたはPCRプライマーにハイブリダイズすることができるユニバーサル標識の核酸配列は、プライマー結合部位と称され得る。ユニバーサル標識は、バーコードの転写を開始させるために使用することができる配列を含み得る。ユニバーサル標識は、バーコードまたはバーコード内の領域の伸長のために使用することができる配列を含み得る。ユニバーサル標識は、1、2、3、4、5、10、15、20、25、30、35、40、45、50ヌクレオチド、またはこれらの値のうちのいずれか2つの間の数もしくは範囲、またはおよそこれらの値もしくはそのような数もしくは範囲のヌクレオチドの長さであり得る。例えば、ユニバーサル標識は、少なくとも約10個のヌクレオチドを含み得る。ユニバーサル標識は、例えば、少なくとも、または最大で1、2、3、4、5、10、15、20、25、30、35、40、45、50、100、200、または300ヌクレオチドの長さであり得る。一部の実施形態では、バーコードを支持体から切断することを可能にするために、切断可能なリンカーまたは改変されたヌクレオチドが、ユニバーサル標識配列の一部となってもよい。

【0060】

次元標識

バーコードは、1つまたは複数の次元標識を含み得る。一部の実施形態では、次元標識は、標識化(例えば、確率的標識化)が発生した次元に関する情報を提供する核酸配列を含み得る。例えば、次元標識は、標的がバーコード化された時間に関する情報を提供する

10

20

30

40

50

ことができる。次元標識は、試料におけるバーコード化（例えば、確率的バーコード化）の時間と関連付けられてもよい。次元標識は、標識化の時点で活性化され得る。異なる次元標識は、異なる時点で活性化され得る。次元標識は、標的、標的の群、および/または試料がバーコード化された順序に関する情報を提供する。例えば、細胞集団は、細胞周期のG 0 期に、バーコード化され得る。細胞は、細胞周期のG 1 期に、バーコード（例えば、確率的バーコード）で再びパルスされ得る。細胞は、細胞周期のS 期に、バーコードで再びパルスされ得るなどとなり得る。それぞれのパルス（例えば、細胞周期のそれぞれの段階）におけるバーコードは、異なる次元標識を含み得る。このようにすると、次元標識により、細胞周期のどの段階でどの標的が標識されたかに関する情報が得られる。次元標識は、多くの異なる生物学的時間を調べることができる。例示的な生物学的時間としては、細胞周期、転写（例えば、転写開始）、および転写物分解を挙げることができるが、これらに限定されない。別の例では、試料（例えば、細胞、細胞集団）は、薬物および/または治療法での処置の前および/または後に、確率的に標識され得る。別個の標的のコピー数の変化により、薬物および/または治療法に対する試料の応答を示すことができる。

【0061】

次元標識は、活性化可能であり得る。活性化可能な次元標識は、特定の時点で活性化させることができる。活性化可能な標識は、例えば、継続的に活性化させることができる（例えば、停止されない）。活性化可能な次元標識は、例えば、可逆的に活性化可能であってもよい（例えば、活性化可能な次元標識は、作動させることができ、停止させることができる）。次元標識は、例えば、少なくとも1回、2回、3回、4回、5回、6回、7回、8回、9回、10回、またはそれ以上、可逆的に活性化可能であってもよい。次元標識は、例えば、少なくとも1回、2回、3回、4回、5回、6回、7回、8回、9回、10回、またはそれ以上、可逆的に活性化可能であってもよい。一部の実施形態では、次元標識は、蛍光、光、化学的事象（例えば、切断、別の分子のライゲーション、改変の付加（例えば、ペグ化、SUMO化、アセチル化、脱アセチル化、脱メチル化）、光化学的事象（例えば、光ケージング）、および非天然のヌクレオチドの導入により活性化させることができる。

【0062】

次元標識は、一部の実施形態では、所与の固体支持体（例えば、ビーズ）に結合したすべてのバーコード（例えば、確率的バーコード）について同一であってもよいが、異なる固体支持体（例えば、ビーズ）については異なり得る。一部の実施形態では、同じ固体支持体上のバーコードのうちの少なくとも60%、70%、80%、85%、90%、95%、97%、99%、または100%が、同じ次元標識を含み得る。一部の実施形態では、同じ固体支持体上のバーコードのうちの少なくとも60%が、同じ次元標識を含み得る。一部の実施形態では、同じ固体支持体上のバーコードのうちの少なくとも95%が、同じ次元標識を含み得る。

複数の固体支持体（例えば、ビーズ）に、 10^6 個以上もの多くの固有の次元標識配列が提示されてもよい。次元標識は、1、2、3、4、5、10、15、20、25、30、35、40、45、50ヌクレオチド、またはこれらの値のうちのいずれか2つの間の数もしくは範囲、またはおおよそこれらの値もしくはそのような数もしくは範囲のヌクレオチドの長さであり得る。次元標識は、少なくとも、または最大で1、2、3、4、5、10、15、20、25、30、35、40、45、50、100、200、または300ヌクレオチド長であり得る。次元標識は、約5~約200個のヌクレオチドを含み得る。次元標識は、約10~約150個のヌクレオチドを含み得る。次元標識は、約20~約125ヌクレオチドの長さを含み得る。

【0063】

空間標識

バーコードは、1つまたは複数の空間標識を含み得る。一部の実施形態では、空間標識は、バーコードと会合した標的分子の空間配向に関する情報を提供する、核酸配列を含み得る。空間標識は、試料における座標と関連付けられ得る。座標は、固定座標であっても

10

20

30

40

50

よい。例えば、座標は、基質に関連して固定され得る。空間標識は、二次元または三次元グリッドを参照するものであってもよい。座標は、ランドマークに対して固定され得る。ランドマークは、空間内で特定可能であり得る。ランドマークは、画像化することができる構造体であり得る。ランドマークは、生物学的構造体、例えば、解剖学的ランドマークであり得る。ランドマークは、細胞ランドマーク、例えば、オルガネラであり得る。ランドマークは、非天然のランドマーク、例えば、特定可能な識別子、例えば、カラーコード、バーコード、磁気特性、蛍光、放射性、もしくは固有のサイズもしくは形状を有する構造体であってもよい。空間標識は、物理的区画（例えば、ウェル、容器、または液滴）と関連付けられていてもよい。一部の実施形態では、複数の空間標識が、空間内の1つまたは複数の位置をコードするために、一緒に使用される。

10

【0064】

空間標識は、所与の固体支持体（例えば、ビーズ）に結合したすべてのバーコードについて同一であってもよいが、異なる固体支持体（例えば、ビーズ）については異なり得る。一部の実施形態では、同じ固体支持体上の同じ空間標識を含むバーコードのパーセンテージは、60%、70%、80%、85%、90%、95%、97%、99%、100%、またはこれらの値のうちのいずれか2つの間の数もしくは範囲であり得るか、またはおおよそこれらの値もしくはそのような数もしくは範囲であり得る。一部の実施形態では、同じ固体支持体上の同じ空間標識を含むバーコードのパーセンテージは、少なくとも、または最大で60%、70%、80%、85%、90%、95%、97%、99%、または100%であり得る。一部の実施形態では、同じ固体支持体上のバーコードのうちの少なくとも60%が、同じ空間標識を含み得る。一部の実施形態では、同じ固体支持体上のバーコードのうちの少なくとも95%が、同じ空間標識を含み得る。

20

【0065】

複数の固体支持体（例えば、ビーズ）に、 10^6 個以上もの多くの固有の空間標識配列が提示されてもよい。空間標識は、1、2、3、4、5、10、15、20、25、30、35、40、45、50ヌクレオチド、またはこれらの値のうちのいずれか2つの間の数もしくは範囲、またはおおよそこれらの値もしくはそのような数もしくは範囲のヌクレオチドの長さであり得る。空間標識は、例えば、少なくとも、または最大で1、2、3、4、5、10、15、20、25、30、35、40、45、50、100、200、または300ヌクレオチドの長さであり得る。空間標識は、約5～約200の間のヌクレオチド、例えば、約10～約150の間のヌクレオチド、または約20～約125の間のヌクレオチドの長さを含み得る。

30

【0066】

細胞標識

バーコード（例えば、確率的バーコード）は、1つまたは複数の細胞標識を含み得る。一部の実施形態では、細胞標識は、どの標的核酸がどの細胞を起源としているかを決定するための情報を提供する、核酸配列を含み得る。一部の実施形態では、細胞標識は、所与の固体支持体（例えば、ビーズ）に結合したすべてのバーコードについて同一であるが、異なる固体支持体（例えば、ビーズ）については異なる。一部の実施形態では、同じ固体支持体上の同じ細胞標識を含むバーコードのパーセンテージは、60%、70%、80%、85%、90%、95%、97%、99%、100%、またはこれらの値のうちのいずれか2つの間の数もしくは範囲、またはおおよそこれらの値もしくはそのような数もしくは範囲であり得る。一部の実施形態では、同じ固体支持体上の同じ細胞標識を含むバーコードのパーセンテージは、60%、70%、80%、85%、90%、95%、97%、99%、または100%であり得るか、またはおおよそこれらの値であり得る。例えば、同じ固体支持体上のバーコードのうちの少なくとも60%、または少なくとも95%が、同じ細胞標識を含み得る。

40

複数の固体支持体（例えば、ビーズ）に、 10^6 個以上もの多くの固有の細胞標識配列が提示されてもよい。細胞標識は、1、2、3、4、5、10、15、20、25、30、35、40、45、50ヌクレオチド、またはこれらの値のうちのいずれか2つの間の数

50

もしくは範囲、またはおよそでこれらの値もしくはそのような数もしくは範囲のヌクレオチドの長さであり得る。細胞標識は、例えば、少なくとも、または最大で1、2、3、4、5、10、15、20、25、30、35、40、45、50、100、200、または300ヌクレオチドの長さであり得る。例えば、細胞標識は、約5～約200の間のヌクレオチド、約10～約150の間のヌクレオチド、または約20～約125の間のヌクレオチドの長さを含み得る。

【0067】

バーコード配列

バーコードは、1つまたは複数のバーコード配列を含み得る。一部の実施形態では、バーコード配列は、バーコードにハイブリダイズした標的核酸種の特定の種類に関する情報の特定をもたらす、核酸配列を含み得る。バーコード配列は、バーコード（例えば、標的結合領域）にハイブリダイズした標的核酸種の特定の発生に関するカウンタを提供する（例えば、およその近似値を提供する）、核酸配列を含む。

10

【0068】

一部の実施形態では、バーコード配列の多様なセットが、所与の固体支持体（例えば、ビーズ）に結合される。一部の実施形態では、 10^2 個、 10^3 個、 10^4 個、 10^5 個、 10^6 個、 10^7 個、 10^8 個、 10^9 個、またはこれらの値のうちのいずれか2つの間の数もしくは範囲、またはおよそでこれらの値もしくはそのような数もしくは範囲の固有の分子標識配列が存在し得る。例えば、複数のバーコードは、別個の配列を有する約6561個のバーコード配列を含み得る。別の例として、複数のバーコードは、別個の配列を有する約65536個のバーコード配列を含み得る。一部の実施形態では、少なくとも、または最大で 10^2 個、 10^3 個、 10^4 個、 10^5 個、 10^6 個、 10^7 個、 10^8 個、または 10^9 個の固有のバーコード配列が存在し得る。固有の分子標識配列は、所与の固体支持体（例えば、ビーズ）に結合していてもよい。一部の実施形態では、固有の分子標識配列は、粒子（例えば、ハイドロゲルビーズ）によって部分的または全体的に包含される。

20

【0069】

バーコードの長さは、異なるインプリメンテーションにおいて異なってもよい。例えば、バーコードは、1、2、3、4、5、10、15、20、25、30、35、40、45、50、もしくはこれらの値のうちのいずれか2つの間の数もしくは範囲のヌクレオチド長であってもよく、または約1、2、3、4、5、10、15、20、25、30、35、40、45、50、もしくはこれらの値のうちのいずれか2つの間の数もしくは範囲のヌクレオチド長であってもよい。別の例として、バーコードは、少なくとも、または最大で1、2、3、4、5、10、15、20、25、30、35、40、45、50、100、200、または300ヌクレオチド長であってもよい。

30

【0070】

分子標識

バーコード（例えば、確率的バーコード）は、1つまたは複数の分子標識を含み得る。分子標識は、バーコード配列を含み得る。一部の実施形態では、分子標識は、バーコードにハイブリダイズした標的核酸種の特定の種類に関する情報の特定をもたらす、核酸配列を含み得る。分子標識は、バーコード（例えば、標的結合領域）にハイブリダイズした標的核酸種の特定の発生に対するカウンタを提供する、核酸配列を含む。

40

一部の実施形態では、分子標識の多様なセットが、所与の固体支持体（例えば、ビーズ）に結合される。一部の実施形態では、 10^2 個、 10^3 個、 10^4 個、 10^5 個、 10^6 個、 10^7 個、 10^8 個、 10^9 個、もしくはこれらの値のうちのいずれか2つの間の数もしくは範囲の固有の分子標識配列、または約 10^2 個、 10^3 個、 10^4 個、 10^5 個、 10^6 個、 10^7 個、 10^8 個、 10^9 個、もしくはこれらの値のうちのいずれか2つの間の数もしくは範囲の固有の分子標識配列が存在し得る。例えば、複数のバーコードは、別個の配列を有する約6561個の分子標識を含み得る。別の例として、複数のバーコードは、別個の配列を有する約65536個の分子標識を含み得る。一部の実施形態では、少なくとも、または最大で 10^2 個、 10^3 個、 10^4 個、 10^5 個、 10^6 個、 10^7 個、 10^8 個、

50

もしくは 10^9 個の固有の分子標識配列が存在し得る。固有の分子標識配列を有するバーコードは、所与の固体支持体（例えば、ビーズ）に結合していてもよい。

【0071】

複数の確率的バーコードを使用するバーコード化（例えば、確率的バーコード化）については、異なる分子標識配列の数と、標的のうちのいずれかの発生の数との比は、 $1:1$ 、 $2:1$ 、 $3:1$ 、 $4:1$ 、 $5:1$ 、 $6:1$ 、 $7:1$ 、 $8:1$ 、 $9:1$ 、 $10:1$ 、 $11:1$ 、 $12:1$ 、 $13:1$ 、 $14:1$ 、 $15:1$ 、 $16:1$ 、 $17:1$ 、 $18:1$ 、 $19:1$ 、 $20:1$ 、 $30:1$ 、 $40:1$ 、 $50:1$ 、 $60:1$ 、 $70:1$ 、 $80:1$ 、 $90:1$ 、 $100:1$ 、もしくはこれらの値のうちのいずれか2つの間の数もしくは範囲であってもよく、または約 $1:1$ 、 $2:1$ 、 $3:1$ 、 $4:1$ 、 $5:1$ 、 $6:1$ 、 $7:1$ 、 $8:1$ 、 $9:1$ 、 $10:1$ 、 $11:1$ 、 $12:1$ 、 $13:1$ 、 $14:1$ 、 $15:1$ 、 $16:1$ 、 $17:1$ 、 $18:1$ 、 $19:1$ 、 $20:1$ 、 $30:1$ 、 $40:1$ 、 $50:1$ 、 $60:1$ 、 $70:1$ 、 $80:1$ 、 $90:1$ 、 $100:1$ 、もしくはこれらの値のうちのいずれか2つの間の数もしくは範囲であってもよい。標的は、同一またはほぼ同一な配列を有するmRNA分子を含むmRNA種であり得る。一部の実施形態では、異なる分子標識配列の数と、標的のうちのいずれかの発生の数との比は、少なくとも、または最大で $1:1$ 、 $2:1$ 、 $3:1$ 、 $4:1$ 、 $5:1$ 、 $6:1$ 、 $7:1$ 、 $8:1$ 、 $9:1$ 、 $10:1$ 、 $11:1$ 、 $12:1$ 、 $13:1$ 、 $14:1$ 、 $15:1$ 、 $16:1$ 、 $17:1$ 、 $18:1$ 、 $19:1$ 、 $20:1$ 、 $30:1$ 、 $40:1$ 、 $50:1$ 、 $60:1$ 、 $70:1$ 、 $80:1$ 、 $90:1$ 、もしくは $100:1$ である。

【0072】

分子標識は、 1 、 2 、 3 、 4 、 5 、 10 、 15 、 20 、 25 、 30 、 35 、 40 、 45 、 50 ヌクレオチド、またはこれらの値のうちのいずれか2つの間の数もしくは範囲、またはおおよそこれらの値もしくはそのような数もしくは範囲のヌクレオチドの長さであり得る。分子標識は、例えば、少なくとも、または最大で 1 、 2 、 3 、 4 、 5 、 10 、 15 、 20 、 25 、 30 、 35 、 40 、 45 、 50 、 100 、 200 、または 300 ヌクレオチドの長さであり得る。

【0073】

標的結合領域

バーコードは、1つまたは複数の標的結合領域、例えば、捕捉プローブを含み得る。一部の実施形態では、標的結合領域は、目的の標的とハイブリダイズし得る。一部の実施形態では、標的結合領域は、標的（例えば、標的核酸、標的分子、例えば、分析しようとする細胞核酸）、例えば、特定の遺伝子配列に特異的にハイブリダイズする核酸配列を含み得る。一部の実施形態では、標的結合領域は、特定の標的核酸の特定の位置に結合（例えば、ハイブリダイズ）することができる核酸配列を含み得る。一部の実施形態では、標的結合領域は、制限酵素部位オーバーハング（例えば、EcoRI粘着末端オーバーハング）への特異的ハイブリダイゼーションが可能である核酸配列を含み得る。バーコードは、次いで、制限部位オーバーハングに相補的な配列を含む任意の核酸分子にライゲーションすることができる。

【0074】

一部の実施形態では、標的結合領域は、非特異的標的核酸配列を含み得る。非特異的標的核酸配列は、標的核酸の特定の配列とは独立して、複数の標的核酸に結合することができる配列を指し得る。例えば、標的結合領域は、ランダム多量体配列、またはmRNA分子上のポリ(A)尾部にハイブリダイズするオリゴ(dT)配列を含み得る。ランダム多量体配列は、例えば、ランダム二量体、三量体、四量体、五量体、六量体、七量体、八量体、九量体、十量体、または任意の長さのそれ以上の多量体配列であり得る。一部の実施形態では、標的結合領域は、所与のビーズに結合したすべてのバーコードについて、同じである。一部の実施形態では、所与のビーズに結合した複数のバーコードの標的結合領域は、2つ以上の異なる標的結合配列を含み得る。標的結合領域は、 5 、 10 、 15 、 20 、 25 、 30 、 35 、 40 、 45 、 50 ヌクレオチド、またはこれらの値のうちのいずれ

か 2 つの間の数もしくは範囲、またはおよそこれらの値もしくはそのような数もしくは範囲のヌクレオチドの長さであり得る。標的結合領域は、最大で約 5、10、15、20、25、30、35、40、45、50ヌクレオチド、またはそれ以上のヌクレオチドの長さであり得る。

【0075】

一部の実施形態では、標的結合領域は、ポリアデニル化末端を含む mRNA とハイブリダイズすることができるオリゴ (dT) を含み得る。標的結合領域は、遺伝子特異的であり得る。例えば、標的結合領域は、標的の特定の領域にハイブリダイズするように構成され得る。標的結合領域は、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、またはこれらの値のうちのいずれか 2 つの間の数もしくは範囲、またはおよそこれらの値もしくはそのような数もしくは範囲のヌクレオチドの長さであり得る。標的結合領域は、少なくとも、または最大で 1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、または 30ヌクレオチドの長さであり得る。標的結合領域は、約 5 ~ 30ヌクレオチドの長さであり得る。バーコードが遺伝子特異的標的結合領域を含む場合、バーコードは、本明細書において、遺伝子特異的バーコードと称され得る。

10

【0076】

配向特性

確率的バーコード (例えば、確率的バーコード) は、バーコードを配向する (例えば、アライメント) するために使用することができる 1 つまたは複数の配向特性を含み得る。バーコードは、等電点電気泳動のための部分を含み得る。異なるバーコードは、異なる等電点電気泳動点を含み得る。これらのバーコードを試料に導入すると、試料は、バーコードを公知の様式で配向するために、等電点電気泳動を受け得る。このようにして、配向特性を使用して、試料におけるバーコードの公知のマップを展開することができる。例示的な配向特性としては、電気泳動移動度 (例えば、バーコードのサイズに基づく)、等電点、スピン、導電性、および/または自己アセンブリを挙げることができる。例えば、自己アセンブリの配向特性を有するバーコードは、活性化されると、特定の配向 (例えば、核酸ナノ構造) に自己アセンブルし得る。

20

【0077】

親和性特性

バーコード (例えば、確率的バーコード) は、1 つまたは複数の親和性特性を含み得る。例えば、空間標識は、親和性特性を含み得る。親和性特性としては、別の実体 (例えば、細胞受容体) へのバーコードの結合を促進することができる化学的および/または生物学的部分を挙げることができる。例えば、親和性特性は、抗体、例えば、試料上の特定の部分 (例えば、受容体) に特異的な抗体を含み得る。一部の実施形態では、抗体は、バーコードを、特定の細胞種または分子へと誘導し得る。特定の細胞種または分子にあるおよび/またはその近傍にある標的を、標識 (例えば、確率的に標識) することができる。親和性特性は、一部の実施形態では、抗体がバーコードを特定の位置へと誘導し得るため、空間標識のヌクレオチド配列に加えて、空間情報を提供することができる。抗体は、治療用抗体、例えば、モノクローナル抗体またはポリクローナル抗体であり得る。抗体は、ヒト化またはキメラであってもよい。抗体は、ネイキッド抗体または融合抗体であってもよい。

30

40

【0078】

抗体は、全長 (すなわち、天然に存在するかもしくは通常の免疫グロブリン遺伝子断片組換えプロセスによって形成される) 免疫グロブリン分子 (例えば、IgG 抗体)、または免疫グロブリン分子の免疫学的に活性な (すなわち、特異的結合) 部分、例えば、抗体断片であり得る。

抗体断片は、例えば、抗体の一部、例えば、F(ab')₂、Fab'、Fab、Fv

50

、sFvなどであり得る。一部の実施形態では、抗体断片は、全長抗体によって認識されるものと同じ抗原に結合することができる。抗体断片としては、抗体の可変領域からなる単離された断片、例えば、重鎖および軽鎖の可変領域からなる「Fv」断片、ならびに軽鎖および重鎖の可変領域がペプチドリinkerによって接続された組換え一本鎖ポリペプチド分子（「scFvタンパク質」）を挙げることができる。例示的な抗体としては、がん細胞に対する抗体、ウイルスに対する抗体、細胞表面受容体（CD8、CD34、CD45）に結合する抗体、および治療用抗体を挙げることができるが、これらに限定されない。

【0079】

ユニバーサルアダプタープライマー

バーコードは、1つまたは複数のユニバーサルアダプタープライマーを含み得る。例えば、遺伝子特異的バーコード、例えば、遺伝子特異的確率的バーコードは、ユニバーサルアダプタープライマーを含み得る。ユニバーサルアダプタープライマーは、すべてのバーコードにわたってユニバーサルなヌクレオチド配列を指し得る。ユニバーサルアダプタープライマーは、遺伝子特異的バーコードを構築するために使用することができる。ユニバーサルアダプタープライマーは、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30ヌクレオチド、またはこれらのうちのいずれか2つの間の数もしくは範囲、またはおよそこれらの値もしくはそのような数もしくは範囲のヌクレオチドの長さであり得る。ユニバーサルアダプタープライマーは、少なくとも、または最大で1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、または30ヌクレオチドの長さであり得る。ユニバーサルアダプタープライマーは、5～30ヌクレオチドの長さであり得る。

【0080】

リンカー

バーコードが、1つを上回る種類の標識（例えば、1つを上回る細胞標識または1つを上回るバーコード配列、例えば、1つの分子標識）を含む場合、標識には、リンカー標識配列が散在していてもよい。リンカー標識配列は、少なくとも約5、10、15、20、25、30、35、40、45、50ヌクレオチド、またはそれ以上のヌクレオチドの長さであり得る。リンカー標識配列は、最大で約5、10、15、20、25、30、35、40、45、50ヌクレオチド、またはそれ以上のヌクレオチドの長さであり得る。一部の事例では、リンカー標識配列は、12ヌクレオチドの長さである。リンカー標識配列は、バーコードの合成を容易にするために使用することができる。リンカー標識は、エラー補正（例えば、ハミング）コードを含み得る。

【0081】

固体支持体

本明細書において開示されるバーコード、例えば、確率的バーコードは、一部の実施形態では、固体支持体と会合していてもよい。固体支持体は、例えば、合成粒子であり得る。一部の実施形態では、バーコード配列の一部またはすべて、例えば、固体支持体上の複数のバーコード（例えば、第1の複数のバーコード）の確率的バーコード（例えば、第1のバーコード配列）の分子標識は、少なくとも1つのヌクレオチドが異なる。同じ固体支持体上のバーコードの細胞標識は、同じであり得る。異なる固体支持体上のバーコードの細胞標識は、少なくとも1つのヌクレオチドが異なり得る。例えば、第1の固体支持体上の第1の複数のバーコードの第1の細胞標識は、同じ配列を有してもよく、第2の固体支持体上の第2の複数のバーコードの第2の細胞標識は、同じ配列を有してもよい。第1の固体支持体上の第1の複数のバーコードの第1の細胞標識および第2の固体支持体上の第2の複数のバーコードの第2の細胞標識は、少なくとも1つのヌクレオチドが異なり得る。細胞標識は、例えば、約5～20ヌクレオチド長であり得る。バーコード配列は、例えば、約5～20ヌクレオチド長であり得る。合成粒子は、例えば、ビーズであり得る。

【0082】

10

20

30

40

50

ビーズは、例えば、シリカゲルビーズ、細孔制御ガラスビーズ、磁性ビーズ、Dyna bead、sephadex/sepharoseビーズ、セルロースビーズ、ポリスチレンビーズ、またはこれらの任意の組合せであり得る。ビーズは、ポリジメチルシロキサン(PDMS)、ポリスチレン、ガラス、ポリプロピレン、アガロース、ゼラチン、ハイドロゲル、常磁性体、セラミック、プラスチック、ガラス、メチルスチレン、アクリルポリマー、チタン、ラテックス、セファロース、セルロース、ナイロン、シリコン、またはこれらの任意の組合せなどの材料を含み得る。

一部の実施形態では、ビーズは、ポリマービーズ、例えば、バーコードまたは確率的バーコードで官能化された変形可能なビーズまたはゲルビーズ(例えば、10X Genomics(San Francisco, CA)からのゲルビーズ)であり得る。一部の
10
実装形態では、ゲルビーズは、ポリマー系ゲルを含み得る。ゲルビーズは、例えば、1つまたは複数のポリマー前駆体を液滴に封入することによって、生成することができる。ポリマー前駆体を、促進剤(例えば、テトラメチルエチレンジアミン(TEMED))に曝露すると、ゲルビーズが生成され得る。

【0083】

一部の実施形態では、粒子は、崩壊可能(例えば、溶解可能、分解可能)であり得る。例えば、ポリマービーズは、例えば、所望される条件下において、溶解、融解、または分解し得る。所望される条件には、環境条件が含まれ得る。所望される条件は、制御された様式で、ポリマービーズの溶解、融解、または分解をもたらし得る。ゲルビーズは、化学的
20
刺激、物理的刺激、生物学的刺激、熱刺激、磁気刺激、電気刺激、光刺激、またはこれらの任意の組合せに起因して、溶解、融解、または分解し得る。

【0084】

分析物および/または試薬、例えば、オリゴヌクレオチドバーコードは、例えば、ゲルビーズの内部表面(例えば、オリゴヌクレオチドバーコードおよび/またはオリゴヌクレオチドバーコードを生成するために使用される材料の拡散を通じてアクセス可能な内部)、および/またはゲルビーズもしくは本明細書に記載の任意の他のマイクロカプセルの外部表面に連結/固定化されていてもよい。連結/固定化は、任意の形態の化学結合(例えば、共有結合、イオン結合)または物理的現象(例えば、ファンデルワールス力、双極子間相互作用など)を介してもよい。一部の実施形態では、試薬の、ゲルビーズまたは本明細書に記載の任意の他のマイクロカプセルへの連結/固定化は、例えば、不安定性部分を
30
介する(例えば、本明細書に記載の化学的架橋剤を含む、化学的架橋剤を介する)ものなど、可逆的であってもよい。刺激を適用すると、不安定性部分は、切断され、固定化された試薬が遊離され得る。一部の実施形態では、不安定性部分は、ジスルフィド結合である。例えば、オリゴヌクレオチドバーコードがジスルフィド結合を介してゲルビーズに固定化されている事例では、ジスルフィド結合を還元剤に曝露することにより、ジスルフィド結合が切断され、オリゴヌクレオチドバーコードがビーズから遊離され得る。不安定性部分は、ゲルビーズもしくはマイクロカプセルの一部として、試薬もしくは分析物をゲルビーズもしくはマイクロカプセルに連結させる化学的リンカーの一部として、および/または試薬もしくは分析物の一部として含まれてもよい。一部の実施形態では、複数のバーコードのうち少なくとも1つのバーコードは、粒子に固定化されていてもよく、粒子に部分的に固定化されていてもよく、粒子に封入されていてもよく、粒子に部分的に封入されていてもよく、またはこれらの任意の組合せであってもよい。
40

【0085】

一部の実施形態では、ゲルビーズは、ポリマー、熱感受性ポリマー、光感受性ポリマー、磁性ポリマー、pH感受性ポリマー、塩感受性ポリマー、化学感受性ポリマー、高分子電解質、多糖、ペプチド、タンパク質、および/またはプラスチックを含むがこれらに限定されない、広範な異なるポリマーを含み得る。ポリマーとしては、ポリ(N-イソプロピルアクリルアミド)(PNIPAAm)、ポリ(スチレンスルホネート)(PSS)、ポリ(アリルアミン)(PAAm)、ポリ(アクリル酸)(PAA)、ポリ(エチレンジアミン)(PEI)、ポリ(ジアリルジメチル-アンモニウムクロリド)(PDADMAC
50

)、ポリ(ピロール)(PPy)、ポリ(ビニルピロリドン)(PVPON)、ポリ(ビニルピリジン)(PVP)、ポリ(メタクリル酸)(PMAA)、ポリ(メチルメタクリレート)(PMMA)、ポリスチレン(PS)、ポリ(テトラヒドロフラン)(PTHF)、ポリ(フタルアルデヒド)(PTHF)、ポリ(ヘキシルピオロゲン)(PHV)、ポリ(L-リシン)(PLL)、ポリ(L-アルギニン)(PARG)、ポリ(乳酸-コ-グリコール酸)(PLGA)などの材料を挙げることができるが、これらに限定されない。

【0086】

多数の化学的刺激を使用して、ビーズの崩壊、溶解、または分解をトリガーすることができる。これらの化学変化の例としては、ビーズ壁へのpHに媒介される変化、架橋結合の化学的切断を介したビーズ壁の崩壊、ビーズ壁の脱重合のトリガー、およびビーズ壁のスイッチング反応を挙げることができるが、これらに限定されない。バルク変化を使用して、ビーズの崩壊をトリガーすることもできる。

10

様々な刺激を通じたマイクロカプセルに対するバルク変化または物理的变化はまた、試薬を放出するためのカプセルの設計において、多くの利点を提供する。バルク変化または物理的变化は、巨視的規模で発生し、ビーズの破裂は、刺激によって誘導される機械-物理力の結果である。これらのプロセスとしては、圧力に誘導される破裂、ビーズ壁の融解、またはビーズ壁の多孔性の変化を挙げることができるが、これらに限定されない。

【0087】

生物学的刺激もまた、ビーズの崩壊、溶解、または分解をトリガーするために使用することができる。一般に、生物学的トリガーは、化学的トリガーに似ているが、多くの例では、生体分子、または酵素、ペプチド、糖、脂肪酸、核酸などといった生体系において一般に見出される分子を使用する。例えば、ビーズは、特定のプロテアーゼによる切断に対して感受性であるペプチド架橋を有するポリマーを含み得る。より具体的には、1つの例では、GFLGKペプチド架橋を含むマイクロカプセルを含み得る。プロテアーゼカテプシンBなどの生物学的トリガーを添加すると、シェル壁のペプチド架橋が切断され、ビーズの内容物が放出される。他の事例では、プロテアーゼは、熱活性化され得る。別の例では、ビーズは、セルロースを含むシェル壁を含む。加水分解性酵素であるキトサンの添加が、セルロース結合の切断、シェル壁の脱重合、およびその内部の内容物の放出の生物学的トリガーとしての機能を果たす。

20

30

【0088】

ビーズはまた、熱刺激を適用するとその内容物を放出するように誘導することもできる。温度の変化は、ビーズに様々な変化を引き起こし得る。熱の変化は、ビーズ壁が崩壊するようにビーズの融解を引き起こし得る。他の事例では、熱は、ビーズが崩壊または爆発するように、ビーズの内部成分の内圧を増加させ得る。なおも他の事例では、熱は、ビーズを圧縮され脱水された状態へと変換させ得る。熱はまた、ビーズの壁内の熱感受性ポリマーに対して作用して、ビーズの崩壊を引き起こすことができる。

マイクロカプセルのビーズ壁に磁性ナノ粒子を含めることにより、ビーズの崩壊をトリガーすること、ならびにビーズをアレイにおいて誘導することができる。本開示のデバイスは、いずれかの目的で磁性ビーズを含み得る。1つの例では、Fe₃O₄ナノ粒子を高分子電解質含有ビーズに組み込むことにより、振動磁界刺激の存在下において崩壊がトリガーされる。

40

【0089】

ビーズはまた、電気刺激の結果として、崩壊、溶解、または分解され得る。前の節に記載されている磁性粒子と同様に、電気感受性ビーズにより、ビーズの崩壊、ならびに電界におけるアライメント、電気伝導、または酸化還元反応などの他の機能の両方をトリガーすることが可能となり得る。1つの例では、電気感受性材料を含有するビーズは、内部試薬の放出を制御することができるように、電界においてアライメントされる。他の例では、電界は、ビーズ壁自体内における酸化還元反応を誘導し得、これにより多孔性が増加し得る。

50

光刺激もまた、ビーズを崩壊させるために使用することができる。多数の光トリガーが、可能であり、これには、特定の範囲の波長の光子を吸収することができるナノ粒子および発色団などの様々な分子を使用するシステムが含まれ得る。例えば、金属酸化物コーティングを、カプセルトリガーとして使用することができる。SiO₂をコーティングした高分子電解質カプセルのUV照射は、ビーズ壁の崩壊をもたらす得る。さらに別の例では、光スイッチ可能な材料、例えば、アゾベンゼン基を、ビーズ壁に組み込んでよい。UVまたは可視光を適用すると、これらのような化学物質は、光子を吸収して可逆的なシスからトランスへの異性化を受ける。この態様では、光スイッチの組み込みにより、光トリガーを適用すると崩壊し得るかまたはより多孔性となり得るビーズ壁が得られる。

【0090】

例えば、図2に示されるバーコード化（例えば、確率的バーコード化）の非限定的な例では、ブロック208において、細胞、例えば、単一細胞を、マイクロウェルアレイの複数のマイクロウェルに導入した後、ブロック212において、ビーズを、マイクロウェルアレイの複数のマイクロウェルに導入することができる。それぞれのマイクロウェルは、1つのビーズを含み得る。ビーズは、複数のバーコードを含み得る。バーコードは、ビーズに結合した5'アミン領域を含み得る。バーコードは、ユニバーサル標識、バーコード配列（例えば、分子標識）、標的結合領域、またはこれらの任意の組合せを含み得る。

【0091】

本明細書において開示されるバーコードは、固体支持体（例えば、ビーズ）と会合している（例えば、それに結合している）（例えば、それに結合している）。固体支持体と会合しているバーコードは、それぞれが、固有の配列を有する少なくとも100または1000個のバーコード配列を含む群から選択される、バーコード配列を含み得る。一部の実施形態では、固体支持体と会合した異なるバーコードは、異なる配列を有するバーコードを含み得る。一部の実施形態では、固体支持体と会合しているあるパーセンテージのバーコードは、同じ細胞標識を含む。例えば、パーセンテージは、60%、70%、80%、85%、90%、95%、97%、99%、100%、またはこれらの値のうちのいずれか2つの間の数もしくは範囲であり得るか、またはおよそこれらの値もしくはそのような数もしくは範囲であり得る。別の例として、パーセンテージは、少なくとも、または最大で60%、70%、80%、85%、90%、95%、97%、99%、または100%であり得る。一部の実施形態では、固体支持体と会合しているバーコードは、同じ細胞標識を有し得る。異なる固体支持体と会合しているバーコードは、固有の配列を有する少なくとも100または1000個の細胞標識を含む群から選択される、異なる細胞標識を有し得る。

【0092】

本明細書において開示されるバーコードは、固体支持体（例えば、ビーズ）に会合している（例えば、それに結合している）（例えば、それに結合している）。一部の実施形態では、試料における複数の標識をバーコード化するステップは、複数のバーコードと会合した複数の合成粒子を含む固体支持体を用いて行うことができる。一部の実施形態では、固体支持体は、複数のバーコードと会合した複数の合成粒子を含み得る。異なる固体支持体上の複数のバーコードの空間標識は、少なくとも1つのヌクレオチドが異なり得る。固体支持体は、例えば、二次元または三次元で複数のバーコードを含み得る。合成粒子は、ビーズであり得る。ビーズは、シリカゲルビーズ、細孔制御ガラスビーズ、磁性ビーズ、Dynabead、Sephadex/Sepharoseビーズ、セルロースビーズ、ポリスチレンビーズ、またはこれらの任意の組合せであり得る。固体支持体としては、ポリマー、マトリックス、ハイドロゲル、ニードルアレイデバイス、抗体、またはこれらの任意の組合せを挙げることができる。一部の実施形態では、固体支持体は、自由に浮動することができる。一部の実施形態では、固体支持体は、半固体または固体アレイに包埋されていてもよい。バーコードは、固体支持体と会合していなくてもよい。バーコードは、個別のヌクレオチドであってもよい。バーコードは、基質と会合している（例えば、それに結合している）。

【0093】

本明細書で使用される場合、「係留された」、「結合した」、および「固定化された」

10

20

30

40

50

という用語は、互換可能に使用され、バーコードを固体支持体に結合させるための共有結合的または非共有結合的手段を指し得る。様々な異なる固体支持体のうちのいずれかを、事前に合成したバーコードを結合させるため、またはバーコードのインサイツでの固相合成のための固体支持体として使用することができる。

一部の実施形態では、固体支持体は、ビーズである。ビーズは、核酸を（例えば、共有結合的または非共有結合的に）固定化することができる1つまたは複数の種類の固体、多孔質、または中空の球体、ボール、ベアリング、シリンダー、または他の類似の構成を含み得る。ビーズは、例えば、プラスチック、セラミック、金属、ポリマー材料、またはこれらの任意の組合せから構成され得る。ビーズは、球状（例えば、マイクロスフェア）である別個の粒子であるか、もしくはそれを含み得るか、または非球状もしくは不規則な形状、例えば、立方体、立方体様、ピラミッド形、円筒形、円錐形、長方形、もしくは円盤形などを有し得る。一部の実施形態では、ビーズは、非球体の形状であり得る。

10

【0094】

ビーズは、常磁性材料（例えば、マグネシウム、モリブデン、リチウム、およびタンタル）、超常磁性材料（例えば、フェライト（ Fe_3O_4 、マグネタイト）ナノ粒子）、強磁性材料（例えば、鉄、ニッケル、コバルト、これらの一部の合金、および一部の希土類金属化合物）、セラミック、プラスチック、ガラス、ポリスチレン、シリカ、メチルスチレン、アクリルポリマー、チタン、ラテックス、セファロース、アガロース、ハイドロゲル、ポリマー、セルロース、ナイロン、またはこれらの任意の組合せを含むがこれらに限定されない、様々な材料を含み得る。

20

一部の実施形態では、ビーズ（例えば、標識が結合されるビーズ）は、ハイドロゲルビーズである。一部の実施形態では、ビーズは、ハイドロゲルを含む。

【0095】

本明細書において開示される一部の実施形態は、1つまたは複数の粒子（例えば、ビーズ）を含む。粒子のそれぞれは、複数のオリゴヌクレオチド（例えば、バーコード）を含み得る。複数のオリゴヌクレオチドのそれぞれは、バーコード配列（例えば、分子標識配列）、細胞標識、および標的結合領域（例えば、オリゴ（dT）配列、遺伝子特異的配列、ランダム多量体、またはこれらの組合せ）を含み得る。複数のオリゴヌクレオチドのそれぞれの細胞標識配列は、同じであってもよい。異なる粒子上のオリゴヌクレオチドの細胞標識配列は、異なる粒子上のオリゴヌクレオチドを特定することができるように、異なってもよい。異なる細胞標識配列の数は、異なる実装形態では異なり得る。一部の実施形態では、細胞標識配列の数は、10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、10000個、20000個、30000個、40000個、50000個、60000個、70000個、80000個、90000個、100000個、 10^6 個、 10^7 個、 10^8 個、 10^9 個、これらの値のうちのいずれか2つの間の数もしくは範囲、またはおよそこれらの値もしくはそのような数もしくは範囲、またはそれ以上であってもよい。一部の実施形態では、細胞標識配列の数は、少なくとも、または最大で10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、10000個、20000個、30000個、40000個、50000個、60000個、70000個、80000個、90000個、100000個、 10^6 個、 10^7 個、 10^8 個、または 10^9 個であり得る。一部の実施形態では、複数の粒子のうちの1個以下、2個以下、3個以下、4個以下、5個以下、6個以下、7個以下、8個以下、9個以下、10個以下、20個以下、30個以下、40個以下、50個以下、60個以下、70個以下、80個以下、90個以下、100個以下、200個以下、300個以下、400個以下、500個以下、600個以下、700個以下、800個以下、900個以下、1000個以下、またはそれ以上を上回らないものが、同じ細胞配列を有するオリゴヌクレオチドを含む。一部の実施形態では、同じ細胞配列を有するオリゴ

30

40

50

ヌクレオチドを含む複数の粒子は、最大で0.1%、0.2%、0.3%、0.4%、0.5%、0.6%、0.7%、0.8%、0.9%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、またはそれ以上であり得る。一部の実施形態では、複数の粒子のうちのいずれもが、同じ細胞標識配列を有さない。

【0096】

それぞれの粒子上の複数のオリゴヌクレオチドは、異なるバーコード配列（例えば、分子標識）を含み得る。一部の実施形態では、バーコード配列の数は、10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、10000個、20000個、30000個、40000個、50000個、60000個、70000個、80000個、90000個、100000個、 10^6 個、 10^7 個、 10^8 個、 10^9 個、またはこれらの値のうちのいずれか2つの間の数もしくは範囲であり得るか、またはおおよそこれらの値もしくはそのような数もしくは範囲であり得る。一部の実施形態では、バーコード配列の数は、少なくとも、または最大で10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、10000個、20000個、30000個、40000個、50000個、60000個、70000個、80000個、90000個、100000個、 10^6 個、 10^7 個、 10^8 個、または 10^9 個であり得る。例えば、複数のオリゴヌクレオチドのうちの少なくとも100個は、異なるバーコード配列を含む。別の例として、単一の粒子において、複数のオリゴヌクレオチドのうちの少なくとも100個、500個、1000個、5000個、10000個、15000個、20000個、50000個、これらの値のうちのいずれか2つの間の数もしくは範囲、またはそれよりも多くが、異なるバーコード配列を含む。一部の実施形態は、バーコードを含む複数の粒子を提供する。一部の実施形態では、標識しようとする標的の発生（またはコピーもしくは数）と、異なるバーコード配列との比は、少なくとも1:1、1:2、1:3、1:4、1:5、1:6、1:7、1:8、1:9、1:10、1:11、1:12、1:13、1:14、1:15、1:16、1:17、1:18、1:19、1:20、1:30、1:40、1:50、1:60、1:70、1:80、1:90、またはそれ以上であり得る。一部の実施形態では、複数のオリゴヌクレオチドのそれぞれは、試料標識、ユニバーサル標識、または両方をさらに含む。粒子は、例えば、ナノ粒子またはマイクロ粒子であり得る。

【0097】

ビーズのサイズは、変動し得る。例えば、ビーズの径は、0.1マイクロメートル~50マイクロメートルの範囲に及び得る。一部の実施形態では、ビーズの径は、0.1、0.5、1、2、3、4、5、6、7、8、9、10、20、30、40、50マイクロメートル、またはこれらの値のうちのいずれか2つの間の数もしくは範囲であり得るか、またはおおよそこれらの値もしくはそのような数もしくは範囲であり得る。

ビーズの径は、基質のウェルの径に関連し得る。一部の実施形態では、ビーズの径は、ウェルの径よりも、10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、またはこれらの値のうちのいずれか2つの間の数もしくは範囲、またはおおよそこれらの値もしくはそのような数もしくは範囲、長いまたは短くてもよい。ビーズの径は、細胞（例えば、基質のウェルによって包囲される単一細胞）の径と関連し得る。一部の実施形態では、ビーズの径は、ウェルの径よりも、少なくとも、または最大で10%、20%、30%、40%、50%、60%、70%、80%、90%、または100%、長いまたは短くてもよい。ビーズの径は、細胞（例えば、基質のウェルによって包囲される単一細胞）の径と関連し得る。一部の実施形態では、ビーズの径は、細胞の径よりも、10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、150%、200%、250%、300%、またはこれらの値のうちのいずれか2つの間の数もしくは範囲、またはおおよそこれらの値もしくはそのような数も

10

20

30

40

50

しくは範囲、長いかまたは短くてもよい。一部の実施形態では、ビーズの径は、細胞の径よりも、少なくとも、または最大で10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、150%、200%、250%、または300%、長いまたは短くてもよい。

【0098】

ビーズは、基質に結合および/または包埋されていてもよい。ビーズは、ゲル、ハイドロゲル、ポリマー、および/またはマトリックスに結合および/または包埋されていてもよい。基質（例えば、ゲル、マトリックス、スキャフォールド、またはポリマー）内のビーズの空間位置は、位置アドレスとしての機能を果たし得るビーズ上のバーコードに存在する空間標識を使用して特定することができる。

10

ビーズの例としては、ストレプトアビジンビーズ、アガロースビーズ、磁性ビーズ、Dyna bead（登録商標）、MACS（登録商標）マイクロビーズ、抗体コンジュゲートビーズ（例えば、抗免疫グロブリンマイクロビーズ）、プロテインAコンジュゲートビーズ、プロテインGコンジュゲートビーズ、プロテインA/Gコンジュゲートビーズ、プロテインLコンジュゲートビーズ、オリゴ(dT)コンジュゲートビーズ、シリカビーズ、シリカ様ビーズ、抗ビオチンマイクロビーズ、抗蛍光色素マイクロビーズ、およびBcMag（商標）カルボキシル末端化磁性ビーズを挙げることができるが、これらに限定されない。

【0099】

ビーズは、量子ドットまたは蛍光色素と会合（例えば、それを含浸）させて、それを1つの蛍光光学チャネルまたは複数の光学チャネルにおいて蛍光性にすることができる。ビーズは、酸化鉄または酸化クロムと会合させて、それを常磁性または強磁性にすることができる。ビーズは、特定可能であり得る。例えば、ビーズは、カメラを使用してイメージングすることができる。ビーズは、ビーズと会合した検出可能なコードを有し得る。例えば、ビーズは、バーコードを含み得る。ビーズは、例えば、有機または無機溶液中での膨張に起因して、サイズが変化し得る。ビーズは、疎水性であり得る。ビーズは、親水性であり得る。ビーズは、生体適合性であり得る。

20

固体支持体（例えば、ビーズ）は、視覚化することができる。固体支持体は、視覚化タグ（例えば、蛍光色素）を含み得る。固体支持体（例えば、ビーズ）は、識別子（例えば、数字）をエッチングすることができる。識別子は、ビーズのイメージングを通じて視覚化することができる。

30

【0100】

固体支持体は、不溶性、半可溶性、または不溶性材料を含み得る。固体支持体は、そこに結合したリンカー、スキャフォールド、構築ブロック、または他の反応性部分が含まれる場合、「官能化」されたと称され得るが、固体支持体は、そこに結合したそのような反応性部分が欠如している場合、「非官能化」であり得る。固体支持体は、溶液中、例えばマイクロタイターウェル形式において、フロースルー形式、例えばカラムにおいて、またはディップスティックにおいて、自由に用いることができる。

【0101】

固体支持体は、膜、紙、プラスチック、コーティング表面、平坦面、ガラス、スライド、チップ、またはこれらの任意の組合せを含み得る。固体支持体は、樹脂、ゲル、マイクロスフェア、または他の幾何構成の形態をとり得る。固体支持体は、シリカチップ、マイクロ粒子、ナノ粒子、プレート、アレイ、キャリアー、平坦支持体、例えば、ガラス繊維フィルタ、ガラス面、金属面、金属面（鋼、金銀、アルミニウム、シリコン、および銅）、ガラス支持体、プラスチック支持体、シリコン支持体、チップ、フィルタ、膜、マイクロウェルプレート、スライド、多重ウェルプレートもしくは膜を含むプラスチック材料（例えば、ポリエチレン、ポリプロピレン、ポリアミド、ニフッ化ポリビニリデンが形成される）、および/またはウエハ、コーム、ピンもしくはニードル（例えば、コンビナトリアル合成もしくは分析に好適なピンのアレイ）、または平坦面の穴もしくはナノリットルウェルのアレイ、例えば、ウエハ（例えば、シリコンウエハ）、底部がフィルタになって

40

50

いるか、もしくはなっていない穴を有するウエハにおけるビーズを含み得る。

固体支持体は、ポリマーマトリックス（例えば、ゲル、ハイドロゲル）を含み得る。ポリマーマトリックスは、細胞内空間（例えば、オルガネラの周囲）に浸透することが可能であり得る。ポリマーマトリックスは、循環系を通して送り出すことが可能であり得る。

【0102】

基質およびマイクロウェルアレイ

本明細書で使用される場合、基質は、ある種の固体支持体を指し得る。基質は、本開示のバーコードまたは確率的バーコードを含み得る固体支持体を指し得る。基質は、例えば、複数のマイクロウェルを含み得る。例えば、基質は、2つ以上のマイクロウェルを含むウェルアレイであり得る。一部の実施形態では、マイクロウェルは、容積が定義されている小さな反応チャンバを含み得る。一部の実施形態では、マイクロウェルは、1つまたは複数の細胞を取り込むことができる。一部の実施形態では、マイクロウェルは、1つのみの細胞を取り込むことができる。一部の実施形態では、マイクロウェルは、1つまたは複数の固体支持体を取り込むことができる。一部の実施形態では、マイクロウェルは、1つのみの固体支持体を取り込むことができる。一部の実施形態では、マイクロウェルは、単一の細胞および単一の固体支持体（例えば、ビーズ）を取り込む。マイクロウェルは、本開示のバーコード試薬を含み得る。

10

【0103】

バーコード化の方法

本開示は、身体試料（例えば、組織、器官、腫瘍、細胞）の別個の位置における別個の標的の数を概算するための方法を提供する。本方法は、試料に近接してバーコード（例えば、確率的バーコード）を配置するステップ、試料を溶解させるステップ、別個の標的をバーコードと会合させるステップ、標的を増幅させるステップ、および/または標的をデジタル計数するステップを含み得る。本方法は、バーコードの空間標識から得られる情報を分析および/または視覚化するステップをさらに含み得る。一部の実施形態では、方法は、試料中の複数の標的を視覚化するステップを含む。複数の標的を試料のマッピングするステップは、試料の二次元マップまたは三次元マップを作成するステップを含み得る。二次元マップおよび三次元マップは、試料中の複数の標的をバーコード化（例えば、確率的バーコード化）する前または後に作成され得る。試料中の複数の標的を視覚化するステップは、複数の標的を試料のマッピングするステップを含み得る。複数の標的を試料のマッピングするステップは、試料の二次元マップまたは三次元マップを作成するステップを含み得る。二次元マップおよび三次元マップは、試料中の複数の標的をバーコード化する前または後に作成され得る。一部の実施形態では、二次元マップおよび三次元マップは、試料を溶解させるステップの前または後に作成され得る。二次元マップまたは三次元マップを作成する前または後に試料を溶解させるステップは、試料を加熱するステップ、試料を界面活性剤と接触させるステップ、試料のpHを変更するステップ、またはこれらの任意の組合せを含み得る。

20

30

【0104】

一部の実施形態では、複数の標的をバーコード化するステップは、複数のバーコードを複数の標的にハイブリダイズさせ、バーコード化された標的（例えば、確率的にバーコード化された標的）を作製するステップを含む。複数の標的をバーコード化するステップは、バーコード化された標的のインデックス化されたライブラリーを生成するステップを含み得る。バーコード化された標的のインデックス化されたライブラリーを生成するステップは、複数のバーコード（例えば、確率的バーコード）を含む固体支持体を用いて実施され得る。

40

【0105】

試料とバーコードを接触させるステップ

本開示は、試料（例えば、細胞）を本開示の基質に接触させるための方法を提供する。例えば、細胞、臓器、または組織の薄切片を含む試料を、バーコード（例えば、確率的バーコード）に接触させることができる。例えば、重力流によって、細胞を接触させること

50

ができ、その場合、細胞は定着して単層を形成し得る。試料は、組織薄片であってもよい。薄片を基質上に配置することができる。試料は、一次元であってもよい（例えば、平面表面を形成してもよい）。例えば、基質上で細胞を増殖させる／培養することによって、試料（例えば、細胞）を基質全体に広げることができる。

バーコードが標的と近接して存在する場合、標的は、バーコードにハイブリダイズし得る。それぞれ別個の標的が、本開示の別個のバーコードと会合し得るように、バーコードを、非枯渴的比率で接触させることができる。標的とバーコードとの間の効率的な会合を確実にするために、標的をバーコードと架橋させてもよい。

【0106】

細胞溶解

細胞およびバーコードの分配後に、細胞を溶解させて、標的分子を遊離させることができる。細胞溶解は、種々の手段のいずれかによって、例えば、化学的もしくは生化学的手段によって、浸透圧ショックによって、または熱溶解、機械溶解、もしくは光学溶解によって達成することができる。界面活性剤（例えば、S D S、L i d デシルスルフェート、T r i t o n X - 1 0 0、T w e e n - 2 0、またはN P - 4 0）、有機溶媒（例えば、メタノールまたはアセトン）、もしくは消化酵素（例えば、プロテイナーゼK、ペプシン、またはトリプシン）、またはこれらの任意の組合せを含む細胞溶解緩衝液を添加することによって、細胞を溶解させてもよい。標的とバーコードとの会合を増加させるために、例えば、温度を低下させるおよび／またはライセートの粘度を増加させることによって、標的分子の拡散速度を変化させてもよい。

【0107】

一部の実施形態では、濾紙を使用することによって、試料を溶解させてもよい。濾紙は、濾紙の上を溶解緩衝液で浸漬可能である。試料の溶解および基質への試料の標的のハイブリダイゼーションを促進可能な圧力で試料に適用することができる。

一部の実施形態では、溶解は、機械溶解、熱溶解、光学溶解、および／または化学溶解によって実施することができる。化学溶解には、プロテイナーゼK、ペプシン、およびトリプシンなどの消化酵素の使用が含まれてもよい。溶解は、基質への溶解緩衝液の添加によって実施することができる。溶解緩衝液は、トリスH C l を含んでもよい。溶解緩衝液は、少なくとも約0.01、0.05、0.1、0.5、もしくは1 Mまたはそれより多いトリスH C l を含んでもよい。溶解緩衝液は、最大で約0.01、0.05、0.1、0.5、もしくは1 Mまたはそれより多いトリスH C l を含んでもよい。溶解緩衝液は、約0.1 MのトリスH C l を含んでもよい。溶解緩衝液のp Hは、少なくとも約1、2、3、4、5、6、7、8、9、10、またはそれより高い値であってもよい。溶解緩衝液のp Hは、最大約1、2、3、4、5、6、7、8、9、10、またはそれより高い値であってもよい。一部の実施形態では、溶解緩衝液のp Hは、約7.5である。溶解緩衝液は、塩（例えば、L i C l）を含んでもよい。溶解緩衝液中の塩濃度は、少なくとも約0.1、0.5、もしくは1 Mまたはそれより高くてもよい。溶解緩衝液中の塩濃度は、最大約0.1、0.5、もしくは1 Mまたはそれより高くてもよい。一部の実施形態では、溶解緩衝液中の塩濃度は、約0.5 Mである。溶解緩衝液は、界面活性剤（例えば、S D S、L i d デシルスルフェート、t r i t o n X、t w e e n、N P - 4 0）を含んでもよい。溶解緩衝液中の界面活性剤の濃度は、少なくとも約0.0001%、0.0005%、0.001%、0.005%、0.01%、0.05%、0.1%、0.5%、1%、2%、3%、4%、5%、6%、もしくは7%、またはそれより高い割合であってもよい。溶解緩衝液中の界面活性剤の濃度は、最大約0.0001%、0.0005%、0.001%、0.005%、0.01%、0.05%、0.1%、0.5%、1%、2%、3%、4%、5%、6%、もしくは7%、またはそれより高い割合であってもよい。一部の実施形態では、溶解緩衝液中の界面活性剤の濃度は、約1%のL i d デシルスルフェートである。溶解のための方法において使用される時間は、使用される界面活性剤の量に依存し得る。一部の実施形態では、使用される界面活性剤が多いほど、溶解のために必要とされる時間は少ない。溶解緩衝液は、キレート剤（例えば、E D T A、E G T A）を含

10

20

30

40

50

んでもよい。溶解緩衝液中のキレート剤の濃度は、少なくとも約 1、5、10、15、20、25、もしくは 30 mM、またはそれより高くてもよい。溶解緩衝液中のキレート剤の濃度は、最大約 1、5、10、15、20、25、もしくは 30 mM、またはそれより高くてもよい。一部の実施形態では、溶解緩衝液中のキレート剤の濃度は、約 10 mM である。溶解緩衝液は、還元試薬（例えば、ベータ - メルカプトエタノール、DTT）を含んでもよい。溶解緩衝液中の還元試薬の濃度は、少なくとも約 1、5、10、15、もしくは 20 mM またはそれより高くてもよい。溶解緩衝液中の還元試薬の濃度は、最大約 1、5、10、15、もしくは 20 mM またはそれより高くてもよい。一部の実施形態では、溶解緩衝液中の還元試薬の濃度は、約 5 mM である。一部の実施形態では、溶解緩衝液は、約 0.1 M のトリス HCl（約 pH 7.5）、約 0.5 M の LiCl、約 1% のリチウムドデシルスルフェート、約 10 mM の EDTA、および約 5 mM の DTT を含んでもよい。

10

【0108】

溶解は、約 4、10、15、20、25、または 30 の温度で実施することができる。溶解は、約 1、5、10、15、20 分、またはそれより長い時間実施することができる。溶解された細胞は、少なくとも約 100000、200000、300000、400000、500000、600000、700000 個、またはそれより多い標的核酸分子を含んでもよい。溶解された細胞は、最大約 100000、200000、300000、400000、500000、600000、700000 個、またはそれより多い標的核酸分子を含んでもよい。

20

【0109】

標的核酸分子へのバーコードの結合

細胞の溶解およびそれからの核酸分子の放出の後に、核酸分子は、共局在化された固体支持体のバーコードとランダムに会合し得る。会合には、標的核酸分子の相補的部分へのバーコードの標的認識領域のハイブリダイゼーションが含まれ得る（例えば、バーコードのオリゴ（dT）は、標的のポリ（A）テールと相互作用し得る）。ハイブリダイゼーションに使用されるアッセイ条件（例えば、緩衝液の pH、イオン強度、温度など）は、特定の安定なハイブリッドの形成を促進するように選択することができる。一部の実施形態では、溶解された細胞から放出された核酸分子は、基質上の複数のプローブと会合し得る（例えば、基板上のプローブとハイブリダイズする）。プローブが、オリゴ（dT）を含む場合、mRNA 分子は、プローブにハイブリダイズし、逆転写され得る。オリゴヌクレオチドのオリゴ（dT）部分は、cDNA 分子の第 1 鎖合成のためのプライマーとして作用し得る。例えば、図 2 のブロック 216 に示すバーコード化の非限定的な例において、mRNA 分子は、ビーズ上のバーコードにハイブリダイズし得る。例えば、一本鎖ヌクレオチド断片は、バーコードの標的結合領域にハイブリダイズし得る。

30

【0110】

結合は、バーコードの標的認識領域と標的核酸分子の一部とのライゲーションをさらに含み得る。例えば、標的結合領域は、制限部位オーバーハング（例えば、EcoRI 付着末端オーバーハング）への特異的ハイブリダイゼーションが可能となり得る核酸配列を含み得る。アッセイの手順は、制限酵素（例えば、EcoRI）で標的核酸を処置して、制限部位オーバーハングを生成することをさらに含み得る。次いで、バーコードは、制限部位オーバーハングに相補的な配列を含む任意の核酸分子にライゲートされ得る。2 つの断片を接続するために、リガーゼ（例えば、T4 DNA リガーゼ）を使用することができる。

40

【0111】

例えば、図 2 のブロック 220 に示すバーコード化の非限定的な例では、複数の細胞（または複数の試料）に由来する標識標的（例えば、標的 - バーコード分子）を、次に、例えば、チューブ中にプールすることができる。例えば、バーコードおよび / または標的 - バーコード分子が結合したビーズを回収することによって、標識標的をプールすることができる。

50

結合した標的 - バーコード分子の固体支持体ベースのコレクションの回収は、磁性ビーズおよび外部から印加された磁界の使用によって実現され得る。標的 - バーコード分子がプールのされたら、すべてのさらなる処理を単一反応槽内で進行させることができる。さらなる処理としては、例えば、逆転写反応、増幅反応、切断反応、解離反応、および/または核酸伸長反応が挙げることができる。さらなる処理反応は、マイクロウェル内で、すなわち、複数の細胞から標識標的核酸分子を最初にプールすることなく、実施することができる。

【0112】

逆転写

本開示は、逆転写（例えば、図2のブロック224）を使用して、標的 - バーコードコンジュゲートを作製するための方法を提供する。標的 - バーコードコンジュゲートは、バーコードと標的核酸のすべてまたは一部分の相補配列（すなわち、バーコード化されたcDNA分子、例えば、確率的にバーコード化されたcDNA分子）とを含み得る。会合したRNA分子の逆転写は、逆転写酵素と一緒に逆転写プライマーを添加することによって起こり得る。逆転写プライマーは、オリゴ（dT）プライマー、ランダムヘキサヌクレオチドプライマー、または標的特異的オリゴヌクレオチドプライマーであってもよい。オリゴ（dT）プライマーは、12～18ヌクレオチド長であってもよく、または約12～18ヌクレオチド長であってもよく、哺乳動物mRNAの3'末端の内在性ポリ（A）テールに結合する。ランダムヘキサヌクレオチドプライマーは、種々の相補部位でmRNAに結合し得る。標的特異的オリゴヌクレオチドプライマーは、典型的には、目的のmRNAを

10

20

【0113】

一部の実施形態では、標識RNA分子の逆転写は、逆転写プライマーの添加によって起こり得る。一部の実施形態では、逆転写プライマーは、オリゴ（dT）プライマー、ランダムヘキサヌクレオチドプライマー、または標的特異的オリゴヌクレオチドプライマーである。一般的に、オリゴ（dT）プライマーは、12～18ヌクレオチド長であり、哺乳動物mRNAの3'末端の内在性ポリ（A）テールに結合する。ランダムヘキサヌクレオチドプライマーは、種々の相補部位でmRNAに結合し得る。標的特異的オリゴヌクレオチドプライマーは、典型的には、目的のmRNAを選択的にプライミングする。

逆転写は、繰返し起こり、複数の標識cDNA分子を生じ得る。本明細書に開示される方法は、少なくとも約1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、または20回の逆転写反応を行うステップを含み得る。本方法は、少なくとも約25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、または100回の逆転写反応を行うステップを含み得る。

30

【0114】

増幅

標識標的核酸分子の複数のコピーを作製するために、1または複数回の核酸増幅反応（例えば、図2のブロック228）を実施することができる。増幅は、複数の標的核酸配列が同時に増幅される、多重方式で実施することができる。増幅反応を使用して、核酸分子にシーケンシングアダプターを付加することができる。増幅反応は、存在する場合、試料標識のうちの少なくとも一部分を増幅することを含み得る。増幅反応は、細胞標識および/またはバーコード配列（例えば、分子標識）の少なくとも一部分を増幅させるステップを含み得る。増幅反応は、試料タグ、細胞標識、空間標識、バーコード配列（例えば、分子標識）、標的核酸、またはこれらの組合せのうちの少なくとも一部分を増幅することを含み得る。増幅反応は、複数の核酸の0.5%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、97%、100%、またはこれらの値のいずれか2つの間の範囲もしくは数を増幅することを含み得る。本方法は、試料標識、細胞標識、空間標識、および/またはバーコード

40

50

配列（例えば、分子標識）を含む標的 - バーコード分子の cDNA コピーを 1 つまたは複数生成するために、1 または複数回の cDNA 合成反応を行うステップをさらに含み得る。

【0115】

一部の実施形態では、ポリメラーゼ連鎖反応（PCR）を使用して、増幅を実施することができる。本明細書で使用する場合、PCR は、DNA の相補鎖の同時プライマー伸長により特定の DNA 配列を *in vitro* で増幅するための反応を指し得る。本明細書で使用する場合、PCR は、以下に限定されないが、RT-PCR、リアルタイム PCR、ネステッド PCR、定量 PCR、多重 PCR、デジタル PCR、およびアセンブリ PCR を含む、反応の派生形を包含し得る。

【0116】

標識核酸の増幅は、非 PCR ベースの方法を含み得る。非 PCR ベースの方法の例としては、以下に限定されないが、多重置換増幅（MDA）、転写媒介増幅（TMA）、核酸配列ベースの増幅（NASBA）、鎖置換増幅（SDA）、リアルタイム SDA、ローリングサークル増幅、またはサークル - サークル増幅が挙げられる。他の非 PCR ベースの増幅方法としては、DNA または RNA 標的を増幅するための DNA 依存性 RNA ポリメラーゼ駆動 RNA 転写増幅または RNA 指向 DNA 合成および転写の多重サイクル、リガーゼ連鎖反応（LCR）、ならびに Q レプリカーゼ（Q）法、パリンドロームプローブの使用、鎖置換増幅、制限エンドヌクラーゼを使用するオリゴヌクレオチド駆動増幅、プライマーが核酸配列にハイブリダイズされかつ得られた二本鎖が伸長反応および増幅の前に切断される増幅方法、5' エキソヌクラーゼ活性を欠く核酸ポリメラーゼを使用する鎖置換増幅、ローリングサークル増幅、ならびに分岐伸長増幅（RAM）が挙げられる。一部の実施形態では、増幅は、環化転写物を生成しない。

【0117】

一部の実施形態では、本明細書に開示される方法は、標識核酸（例えば、標識 RNA、標識 DNA、標識 cDNA）上でポリメラーゼ連鎖反応を行い、標識アンプリコン（例えば、確率標識アンプリコン）を生成するステップをさらに含む。標識アンプリコンは、二本鎖分子であってもよい。二本鎖分子は、二本鎖 RNA 分子、二本鎖 DNA 分子、または DNA 分子にハイブリダイズされた RNA 分子を含み得る。二本鎖分子の一方または両方の鎖は、試料標識、空間標識、細胞標識、および/またはバーコード配列（例えば、分子標識）を含み得る。標識アンプリコンは、一本鎖分子であってもよい。一本鎖分子は、DNA、RNA、またはこれらの組合せを含み得る。本開示の核酸は、合成核酸または改変核酸を含み得る。

【0118】

増幅は、1 つまたは複数の非天然ヌクレオチドの使用を含み得る。非天然ヌクレオチドは、光不安定性またはトリガー性のヌクレオチドを含み得る。非天然ヌクレオチドの例としては、以下に限定されないが、ペプチド核酸（PNA）、モルホリノ核酸およびロックド核酸（LNA）、ならびにグリコール核酸（GNA）およびトレオース核酸（TNA）を挙げることができる。非天然ヌクレオチドを増幅反応の 1 または複数回のサイクルに添加することができる。非天然ヌクレオチドの添加を使用して、増幅反応の特定のサイクルまたは時点として産物を同定することができる。

【0119】

1 または複数回の増幅反応を行うことは、1 つまたは複数のプライマーの使用を含み得る。1 つまたは複数のプライマーは、例えば、1、2、3、4、5、6、7、8、9、10、11、12、13、14、もしくは 15 またはそれより多いヌクレオチドを含み得る。1 つまたは複数のプライマーは、少なくとも 1、2、3、4、5、6、7、8、9、10、11、12、13、14、もしくは 15 またはそれより多いヌクレオチドを含み得る。1 つまたは複数のプライマーは、12 未満 ~ 15 ヌクレオチドを含み得る。1 つまたは複数のプライマーは、複数の標識標的（例えば、確率標識標的）のうちの少なくとも一部分にアニールし得る。1 つまたは複数のプライマーは、複数の標識標的の 3' 末端または 5' 末端にアニールし得る。1 つまたは複数のプライマーは、複数の標識標的の内部領域にア

10

20

30

40

50

ニールし得る。内部領域は、複数の標識標的の3'末端から少なくとも約50、100、150、200、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400、410、420、430、440、450、460、470、480、490、500、510、520、530、540、550、560、570、580、590、600、650、700、750、800、850、900または1000ヌクレオチドであってもよい。1つまたは複数のプライマーは、プライマーの一定パネルを含み得る。1つまたは複数のプライマーは、少なくとも1つまたは複数のカスタムプライマーを含み得る。1つまたは複数のプライマーは、少なくとも1つまたは複数の対照プライマーを含み得る。1つまたは複数のプライマーは、少なくとも1つまたは複数の遺伝子特異的プライマーを含み得る。

10

【0120】

1つまたは複数のプライマーは、ユニバーサルプライマーを含み得る。ユニバーサルプライマーは、ユニバーサルプライマー結合部位にアニールし得る。1つまたは複数のカスタムプライマーは、第1の試料標識、第2の試料標識、空間標識、細胞標識、バーコード配列（例えば、分子標識）、標的、またはこれらの任意の組合せにアニールし得る。1つまたは複数のプライマーは、ユニバーサルプライマーおよびカスタムプライマーを含み得る。カスタムプライマーは、1つまたは複数の標的を増幅するように設計され得る。標的は、1つまたは複数の試料中の全核酸のサブセットを含み得る。標的は、1つまたは複数の試料中の全標識標的のサブセットを含み得る。1つまたは複数のプライマーは、少なくとも96個またはそれより多いカスタムプライマーを含み得る。1つまたは複数のプライマーは、少なくとも960個またはそれより多いカスタムプライマーを含み得る。1つまたは複数のプライマーは、少なくとも9600個またはそれより多いカスタムプライマーを含み得る。1つまたは複数のカスタムプライマーは、2つ以上の異なる標識核酸にアニールし得る。2つ以上の異なる標識核酸は、1つまたは複数の遺伝子に対応し得る。

20

【0121】

任意の増幅スキームを本開示の方法において使用することができる。例えば、一スキームでは、1回目のPCRによって、遺伝子特異的プライマーおよびユニバーサルIlluminaシーケンシングプライマー1の配列に対するプライマーを使用して、ビーズに結合した分子を増幅することができる。2回目のPCRでは、Illuminaシーケンシングプライマー2の配列に隣接するネストド遺伝子特異的プライマー、およびユニバーサルIlluminaシーケンシングプライマー1の配列に対するプライマーを使用して、第1のPCR産物を増幅することができる。3回目のPCRでは、P5およびP7ならびに試料インデックスが付加されて、PCR産物をIlluminaシーケンシングライブラリーに入れる。150bp×2シーケンシングを使用するシーケンシングは、リード1上の細胞標識およびバーコード配列（例えば、分子標識）、リード2上の遺伝子、ならびにインデックス1リード上の試料インデックスを明らかにし得る。

30

【0122】

一部の実施形態では、化学的切断を使用して基質から核酸を除去することができる。例えば、核酸中に存在する化学基または修飾塩基を使用して、固体支持体からのその除去を促進することができる。例えば、酵素を使用して、核酸を基質から除去することができる。例えば、核酸は、制限エンドヌクレアーゼ消化によって、基質から除去することができる。例えば、dUTPまたはddUTPを含有する核酸のウラシル-d-グリコシラーゼ(UDG)による処理を使用して、核酸を基質から除去することができる。例えば、ヌクレオチド切除を実施する酵素、例えば、塩基除去修復酵素、例えば、脱プリン/脱ピリミジン(AP)エンドヌクレアーゼを使用して、核酸を基質から除去することができる。一部の実施形態では、光切断可能な基と光とを使用して、核酸を基質から除去することができる。一部の実施形態では、切断可能なリンカーを、核酸を基質から除去するために使用することができる。例えば、切断可能なリンカーは、ピオチン/アビジン、ピオチン/ストレプトアビジン、ピオチン/ニュートラアビジン、Ig-プロテインA、光不安定性リ

40

50

ンカー、酸もしくは塩基不安定性リンカー基、またはアプタマーのうちの少なくとも1つを含み得る。

【0123】

プローブが遺伝子特異的である場合、分子は、プローブにハイブリダイズし、逆転写および/または増幅され得る。一部の実施形態では、核酸が合成された(例えば、逆転写された)後に、増幅され得る。増幅は、複数の標的核酸配列が同時に増幅される多重方式で実施することができる。増幅によって、核酸にシーケンシングアダプターを付加し得る。

【0124】

一部の実施形態では、増幅は、例えば、ブリッジ増幅を用いて基質上で実施することができる。基質上でオリゴ(dT)プローブを使用するブリッジ増幅に適合する末端を生成するために、cDNAにホモポリマーテールを付加することができる。ブリッジ増幅では、テンプレート核酸の3'末端に相補的なプライマーは、固体粒子に共有結合される各対の第1のプライマーであってもよい。テンプレート核酸を含有する試料が粒子に接触しており、かつ1回の熱サイクルが実施される場合、テンプレート分子は第1のプライマーにアニールされ、かつ第1のプライマーはヌクレオチドの添加によって順方向に伸長して、テンプレート分子とテンプレートに相補的な新たに形成されたDNA鎖とからなる二本鎖分子を形成し得る。次のサイクルの加熱ステップでは、二本鎖分子は変性されて、粒子からテンプレート分子を放出し、第1のプライマーを通じて粒子に結合した相補的DNA鎖を残存させ得る。続くアニーリングおよび伸長ステップのアニーリング段階では、相補鎖は、第1のプライマーから除去された位置の相補鎖のセグメントに相補的な第2のプライマーにハイブリダイズし得る。このハイブリダイゼーションにより、相補鎖は、共有結合により第1のプライマーにかつハイブリダイゼーションにより第2のプライマーに固定化されたブリッジを第1および第2のプライマー間に形成し得る。伸長段階では、第2のプライマーは、同一の反応混合物中にヌクレオチドを添加することにより逆方向に伸長し、それにより、ブリッジを二本鎖ブリッジに変換し得る。次いで、次のサイクルが開始し、二本鎖ブリッジは変性されて、それぞれ第1および第2のプライマーを通じて粒子表面に結合した一方の末端と、それぞれ未結合の他方の末端とを有する2つの一本鎖核酸分子を得ることができる。この第2のサイクルのアニーリングおよび伸長ステップでは、各鎖は同一の粒子上のこれまで未使用であったさらなる相補的プライマーにハイブリダイズして、新たな一本鎖ブリッジを形成し得る。この時点でハイブリダイズされる2つのこれまで未使用であったプライマーは、伸長して2つの新たなブリッジを二本鎖ブリッジに変換する。

【0125】

増幅反応は、複数の核酸の少なくとも1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、95%、97%、または100%を増幅することを含み得る。

標識核酸の増幅は、PCRベースの方法または非PCRベースの方法を含み得る。標識核酸の増幅は、標識核酸の指数関数的増幅を含み得る。標識核酸の増幅は、標識核酸の線形増幅を含み得る。増幅は、ポリメラーゼ連鎖反応(PCR)によって実施することができる。PCRは、DNAの相補鎖の同時プライマー伸長による特定のDNA配列の*in vitro*増幅のための反応を指し得る。PCRは、以下に限定されないが、RT-PCR、リアルタイムPCR、ネステッドPCR、定量PCR、多重PCR、デジタルPCR、サプレッションPCR、セミサプレッシブPCR、およびアセンブリPCRを含む、反応の派生形を包含し得る。

【0126】

一部の実施形態では、標識核酸の増幅は、非PCRベースの方法を含む。非PCRベースの方法の例としては、以下に限定されないが、多重置換増幅(MDA)、転写媒介増幅(TMA)、核酸配列ベースの増幅(NASBA)、鎖置換増幅(SDA)、リアルタイムSDA、ローリングサークル増幅、またはサークル-サークル増幅が挙げられる。他の非PCRベースの増幅方法としては、DNAもしくはRNA標的を増幅するためのDNA

10

20

30

40

50

依存性RNAポリメラーゼ駆動RNA転写増幅もしくはRNA指向DNA合成および転写の多重サイクル、リガーゼ連鎖反応(LCR)、にQレプリカーゼ(Q)法、パリンドロームプローブの使用、鎖置換増幅、制限エンドヌクレアーゼを使用するオリゴヌクレオチド駆動増幅、プライマーが核酸配列にハイブリダイズされかつ得られた二本鎖が伸長反応および増幅の前に切断される増幅方法、5'エキソヌクレアーゼ活性を欠く核酸ポリメラーゼを使用する鎖置換増幅、ローリングサークル増幅、ならびにノまたは分岐伸長増幅(RAM)が挙げられる。

【0127】

一部の実施形態では、本明細書に開示される方法は、増幅アンプリコン(例えば、標的)上でネステッドポリメラーゼ連鎖反応を行うステップをさらに含む。アンプリコンは二本鎖分子であってもよい。二本鎖分子は、二本鎖RNA分子、二本鎖DNA分子、またはDNA分子にハイブリダイズされたRNA分子を含み得る。二本鎖分子の一方または両方の鎖は、試料タグまたは分子識別子標識を含み得る。代替的に、アンプリコンは、一本鎖分子であってもよい。一本鎖分子は、DNA、RNA、またはこれらの組合せを含み得る。本発明の核酸は、合成核酸または改変核酸を含み得る。

10

一部の実施形態では、本方法は、多数のアンプリコンを生成するために標識核酸を繰返し増幅するステップを含む。本明細書に開示される方法は、少なくとも約1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、または20回の増幅反応を行うステップを含み得る。代替的に、本方法は、少なくとも約25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、または100回の増幅反応を行うステップを含む。

20

【0128】

増幅は、複数の核酸を含む1つまたは複数の試料に1つまたは複数の対照核酸を添加することをさらに含み得る。増幅は、複数の核酸に1つまたは複数の対照核酸を添加することをさらに含み得る。対照核酸は、対照標識を含み得る。

増幅は、1つまたは複数の非天然ヌクレオチドの使用を含み得る。非天然ヌクレオチドは、光不安定性およびノまたはトリガー性ヌクレオチドを含み得る。非天然ヌクレオチドの例としては、以下に限定されないが、ペプチド核酸(PNA)、モルホリノ核酸およびロックド核酸(LNA)、ならびにグリコール核酸(GNA)およびトレース核酸(TNA)が挙げられる。非天然ヌクレオチドを増幅反応の1または複数回のサイクルに添加することができる。非天然ヌクレオチドの添加を使用して、増幅反応の特定のサイクルまたは時点として産物を同定することができる。

30

【0129】

増幅反応を1または複数回行うことは、1つまたは複数のプライマーの使用を含み得る。1つまたは複数のプライマーは、1つまたは複数のオリゴヌクレオチドを含み得る。1つまたは複数のオリゴヌクレオチドは、少なくとも約7~9ヌクレオチドを含み得る。1つまたは複数のオリゴヌクレオチドは、12未満~15ヌクレオチドを含み得る。1つまたは複数のプライマーは、複数の標識核酸のうちの少なくとも一部分にアニールし得る。

1つまたは複数のプライマーは、複数の標識核酸の3'末端およびノまたは5'末端にアニールし得る。1つまたは複数のプライマーは、複数の標識核酸の内部領域にアニールし得る。内部領域は、複数の標識核酸の3'末端から少なくとも約50、100、150、200、220、230、240、250、260、270、280、290、300、310、320、330、340、350、360、370、380、390、400、410、420、430、440、450、460、470、480、490、500、510、520、530、540、550、560、570、580、590、600、650、700、750、800、850、900、または1000ヌクレオチドであってもよい。1種または複数のプライマーは、プライマーの一定パネル、少なくとも1種もしくは複数のカスタムプライマー、少なくとも1種もしくは複数の対照プライマー、またはこれらの組合せを含み得る。1つまたは複数のプライマーは、少なくとも1つまたは複数のハウスキーピング遺伝子プライマーを含み得る。1つまたは複数のプライマーは、ユニバ

40

50

ーサルプライマーを含み得る。ユニバーサルプライマーは、ユニバーサルプライマー結合部位にアニールし得る。1つまたは複数のカスタムプライマーは、第1の試料タグ、第2の試料タグ、分子識別子標識、核酸またはその産物にアニールし得る。1つまたは複数のプライマーは、ユニバーサルプライマーおよびカスタムプライマーを含み得る。カスタムプライマーは、1つまたは複数の標的核酸を増幅するように設計され得る。標的核酸は、1つまたは複数の試料中の全核酸のサブセットを含み得る。一部の実施形態では、プライマーは、本開示のアレイに結合したプローブである。

【0130】

一部の実施形態では、試料中の複数の標的をバーコード化するステップ（例えば、確率的にバーコード化するステップ）は、バーコード化された標的（例えば、確率的にバーコード化された標的）またはそれらの標的のバーコード化された断片のインデックス化されたライブラリーを生成するステップをさらに含む。異なるバーコードのバーコード配列（例えば、異なる確率的バーコードの分子標識）は、互いに異なってもよい。バーコード化された標的のインデックス化されたライブラリーを生成するステップは、試料中の複数の標的から複数のインデックス化されたポリヌクレオチドを生成するステップを含む。例えば、第1のインデックス化された標的と第2のインデックス化された標的とを含むバーコード化された標的のインデックス化されたライブラリーについて、第1のインデックス化されたポリヌクレオチドの標識領域は、第2のインデックス化されたポリヌクレオチドの標識領域と、1、2、3、4、5、6、7、8、9、10、20、30、40、50ヌクレオチド、またはこれらの値のうちのいずれか2つの間の数もしくは範囲、またはおおよそこれらの値もしくはそのような数もしくは範囲、または少なくともこれらの値もしくはそのような数もしくは範囲、または最大でこれらの値もしくはそのような数もしくは範囲で異なってもよい。一部の実施形態では、バーコード化された標的のインデックス化されたライブラリーを生成するステップは、複数の標的、例えば、mRNA分子を、ポリ(T)領域および標識領域を含む複数のオリゴヌクレオチドと接触させるステップと、各々がcDNA領域および標識領域を含む一本鎖標識cDNA分子を生成するために、逆転写酵素を使用して第1鎖合成を行うステップとを含み、ここで、複数の標的は、異なる配列のうちの少なくとも2つのmRNA分子を含み、複数のオリゴヌクレオチドは、異なる配列のうちの少なくとも2つのオリゴヌクレオチドを含む。バーコード化された標的のインデックス化されたライブラリーを生成するステップは、一本鎖標識cDNA分子を増幅して、二本鎖標識cDNA分子を生成するステップと、二本鎖標識cDNA分子上でネステッドPCRを行って、標識アンプリコンを生成するステップとをさらに含む得る。一部の実施形態では、本方法は、アダプター-標識アンプリコンを生成するステップを含み得る。

【0131】

バーコード化するステップ（例えば、確率的にバーコード化するステップ）は、核酸バーコードまたはタグを使用して、個々の核酸（例えば、DNAまたはRNA）分子を標識するステップを含み得る。一部の実施形態では、これは、DNAバーコードまたはタグがmRNAから生成される際に、cDNA分子にこれらを付加するステップを含む。ネステッドPCRは、PCR増幅バイアスの最小化を実施することができる。例えば、次世代シーケンシング(NGS)を使用するシーケンシングのためにアダプターを付加することができる。例えば、図2のブロック232の標的の1つまたは複数のコピーの細胞標識、分子標識、およびヌクレオチド断片の配列を決定するために、シーケンシング結果を使用することができる。

【0132】

図3は、バーコード化された標的（例えば、確率的にバーコード化された標的）、例えば、バーコード化されたmRNAまたはその断片のインデックス化されたライブラリーを生成する非限定的な例示のプロセスを示す略図である。ステップ1に示されるように、逆転写プロセスによって、固有の分子標識配列、細胞標識配列、およびユニバーサルPCR部位を含む各mRNA分子がコードされ得る。特に、RNA分子302のポリ(A)ター

10

20

30

40

50

ル領域 308 への、バーコード（例えば、確率的バーコード）310 のセットのハイブリダイゼーション（例えば、確率的ハイブリダイゼーション）によって、RNA 分子 302 を逆転写して、cDNA 領域 306 を含む標識 cDNA 分子 304 を生成することができる。バーコード 310 のそれぞれは、標的結合領域、例えば、ポリ（dT）領域 312、標識領域 314（例えば、バーコード配列または分子）、およびユニバーサル PCR 領域 316 を含み得る。

【0133】

一部の実施形態では、細胞標識配列は、3 ~ 20 ヌクレオチドを含み得る。一部の実施形態では、分子標識配列は、3 ~ 20 ヌクレオチドを含み得る。一部の実施形態では、複数の確率的バーコードのそれぞれは、ユニバーサル標識および細胞標識のうちの 1 つまたは複数のさらに含み、ユニバーサル標識は、固体支持体上の複数の確率的バーコードについて同じであり、細胞標識は、固体支持体上の複数の確率的バーコードについて同じである。一部の実施形態では、ユニバーサル標識は、3 ~ 20 ヌクレオチドを含み得る。一部の実施形態では、細胞標識は、3 ~ 20 ヌクレオチドを含む。

10

【0134】

一部の実施形態では、標識領域 314 は、バーコード配列または分子標識 318 および細胞標識 320 を含み得る。一部の実施形態では、標識領域 314 は、ユニバーサル標識、次元標識、および細胞標識のうちの 1 つまたは複数を含み得る。バーコード配列または分子標識 318 は、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、またはこれらの値のうちのいずれかの間の数もしくは範囲のヌクレオチド長であってもよく、またはおおよそこれらの値もしくはそのような数もしくは範囲のヌクレオチド長であってもよく、または少なくともこれらの値もしくはそのような数もしくは範囲のヌクレオチド長であってもよく、または最大でこれらの値もしくはそのような数もしくは範囲のヌクレオチド長であってもよい。細胞標識 320 は、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、またはこれらの値のうちのいずれかの間の数もしくは範囲のヌクレオチド長であってもよく、またはおおよそこれらの値もしくはそのような数もしくは範囲のヌクレオチド長であってもよく、または少なくともこれらの値もしくはそのような数もしくは範囲のヌクレオチド長であってもよく、または最大でこれらの値もしくはそのような数もしくは範囲のヌクレオチド長であってもよい。ユニバーサル標識は、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、またはこれらの値のうちのいずれかの間の数もしくは範囲のヌクレオチド長であってもよく、またはおおよそこれらの値もしくはそのような数もしくは範囲のヌクレオチド長であってもよく、または少なくともこれらの値もしくはそのような数もしくは範囲のヌクレオチド長であってもよく、または最大でこれらの値もしくはそのような数もしくは範囲のヌクレオチド長であってもよい。ユニバーサル標識は、固体支持体上の複数の確率的バーコードについて同じであってもよく、細胞標識は、固体支持体上の複数の確率的バーコードについて同じであってもよい。次元標識は、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、またはこれらの値のうちのいずれかの間の数もしくは範囲のヌクレオチド長であってもよく、またはおおよそこれらの値もしくはそのような数もしくは範囲のヌクレオチド長であってもよく、または少なくともこれらの値もしくはそのような数もしくは範囲のヌクレオチド長であってもよく、または最大でこれらの値もしくはそのような数もしくは範囲のヌクレオチド長であってもよい。

20

30

40

【0135】

一部の実施形態では、標識領域 314 は、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、200、300、400、500、600、700、800、900、1000、もしくはこれらの値のいずれかの間の数もしくは範囲の異なる標識、例えば、バーコード配列もしくは分子標識 318 および細胞標識 320 を含むか、またはおおよそこれらの値もしくはそのような数もしくは

50

範囲の異なる標識、例えば、バーコード配列もしくは分子標識 3 1 8 および細胞標識 3 2 0 を含むか、または少なくともこれらの値もしくはそのような数もしくは範囲の異なる標識、例えば、バーコード配列もしくは分子標識 3 1 8 および細胞標識 3 2 0 を含むか、または最大でこれらの値もしくはそのような数もしくは範囲の異なる標識、例えば、バーコード配列もしくは分子標識 3 1 8 および細胞標識 3 2 0 を含み得る。各標識は、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、90、100、またはこれらの値のうちのいずれかの間の数もしくは範囲のヌクレオチド長であってもよく、またはおよそでこれらの値もしくはそのような数もしくは範囲のヌクレオチド長であってもよく、または少なくともこれらの値もしくはそのような数もしくは範囲のヌクレオチド長であってもよく、または最大でこれらの値もしくはそのような数もしくは範囲のヌクレオチド長であってもよい。バーコードまたは確率的バーコード 3 1 0 のセットは、10、20、40、50、70、80、90、10²、10³、10⁴、10⁵、10⁶、10⁷、10⁸、10⁹、10¹⁰、10¹¹、10¹²、10¹³、10¹⁴、10¹⁵、10²⁰、またはこれらの値のうちのいずれかの間の数もしくは範囲のバーコードもしくは確率的バーコード 3 1 0 を含有するか、またはおよそでこれらの値もしくはそのような数もしくは範囲のバーコードもしくは確率的バーコード 3 1 0 を含有するか、または少なくともこれらの値もしくはそのような数もしくは範囲のバーコードもしくは確率的バーコード 3 1 0 を含有するか、または最大でこれらの値もしくはそのような数もしくは範囲のバーコードもしくは確率的バーコード 3 1 0 であり得る。また、バーコードまたは確率的バーコード 3 1 0 のセットは、例えば、それぞれ、特有の標識領域 3 1 4 を含有し得る。過剰なバーコードまたは確率的バーコード 3 1 0 を除去するために、標識 c D N A 分子 3 0 4 は精製されてもよい。精製には、Ampure ビーズ精製が含まれてもよい。

10

20

【0136】

ステップ 2 に示すように、ステップ 1 の逆転写プロセスからの産物は、1 チューブ中にプールされ、第 1 の P C R プライマープールおよび第 1 のユニバーサル P C R プライマーを用いて P C R 増幅され得る。プールすることは、特有の標識領域 3 1 4 によって可能である。特に、標識 c D N A 分子 3 0 4 を増幅して、ネステッド P C R 標識アンプリコン 3 2 2 を生成することができる。増幅は、多重 P C R 増幅を含み得る。増幅は、単一反応量で 9 6 多重プライマーを用いる多重 P C R 増幅を含み得る。一部の実施形態では、多重 P C R 増幅は、単一反応量で、10、20、40、50、70、80、90、10²、10³、10⁴、10⁵、10⁶、10⁷、10⁸、10⁹、10¹⁰、10¹¹、10¹²、10¹³、10¹⁴、10¹⁵、10²⁰、またはこれらの値のうちのいずれかの間の数もしくは範囲の多重プライマーを利用するか、またはおよそでこれらの値もしくはそのような数もしくは範囲の多重プライマーを利用するか、または少なくともこれらの値もしくはそのような数もしくは範囲の多重プライマーを利用するか、または最大でこれらの値もしくはそのような数もしくは範囲の多重プライマーを利用することができる。増幅は、特定の遺伝子を標的化するカスタムプライマー 3 2 6 A ~ C を含む第 1 の P C R プライマープール 3 2 4 およびユニバーサルプライマー 3 2 8 を含み得る。カスタムプライマー 3 2 6 は、標識 c D N A 分子 3 0 4 の c D N A 部分 3 0 6 ' 内の領域とハイブリダイズし得る。ユニバーサルプライマー 3 2 8 は、標識 c D N A 分子 3 0 4 のユニバーサル P C R 領域 3 1 6 とハイブリダイズし得る。

30

40

【0137】

図 3 のステップ 3 に示すように、ステップ 2 の P C R 増幅からの産物は、ネステッド P C R プライマープールおよび第 2 のユニバーサル P C R プライマーを用いて増幅され得る。ネステッド P C R は、P C R 増幅バイアスを最小限に抑えることができる。特に、ネステッド P C R 標識アンプリコン 3 2 2 は、ネステッド P C R によってさらに増幅することができる。ネステッド P C R は、単一反応量でネステッド P C R プライマー 3 3 2 a ~ c のネステッド P C R プライマープール 3 3 0 と、第 2 のユニバーサル P C R プライマー 3 2 8 ' とを含む多重 P C R を含み得る。ネステッド P C R プライマープール 3 2 8 は、1、2、3、4、5、6、7、8、9、10、20、30、40、50、60、70、80、

50

90、100、200、300、400、500、600、700、800、900、1000、またはこれらの値のうちのいずれかの間の数もしくは範囲の異なるネステッドPCRプライマー330を含有するか、またはおよそこれらの値もしくはそのような数もしくは範囲の異なるネステッドPCRプライマー330を含有するか、または少なくともこれらの値もしくはそのような数もしくは範囲の異なるネステッドPCRプライマー330を含有するか、または最大でこれらの値もしくはそのような数もしくは範囲の異なるネステッドPCRプライマー330を含有し得る。ネステッドPCRプライマー332は、アダプター334を含有し、標識アンプリコン322のcDNA部分306'内の領域とハイブリダイズし得る。ユニバーサルプライマー328'は、アダプター336を含有し、標識アンプリコン322のユニバーサルPCR領域316とハイブリダイズし得る。よって、ステップ3は、アダプター標識アンプリコン338を生成する。一部の実施形態では、ネステッドPCRプライマー332と第2のユニバーサルPCRプライマー328'は、アダプター334および336を含有しなくてもよい。代わりに、アダプター334および336は、ネステッドPCRの産物とライゲートして、アダプター標識アンプリコン338を生成することができる。

【0138】

ステップ4に示すように、ステップ3からのPCR産物は、ライブラリー増幅プライマーを使用するシーケンシングのためにPCR増幅され得る。特に、アダプター334および336を使用して、アダプター標識アンプリコン338に関する1または複数回の追加のアッセイを行うことができる。アダプター334および336は、プライマー340および342とハイブリダイズされ得る。1つまたは複数のプライマー340および342は、PCR増幅プライマーであってもよい。1つまたは複数のプライマー340および342は、シーケンシングプライマーであってもよい。アダプター標識アンプリコン338のさらなる増幅のために、1つまたは複数のアダプター334および336を使用することができる。アダプター標識アンプリコン338のシーケンシングのために、1つまたは複数のアダプター334および336を使用することができる。プライマー342は、プレートインデックス344を含有することができ、これによって、バーコードまたは確率的バーコード310の同じセットを使用して生成されたアンプリコンを、次世代シーケンシング(NGS)を使用する1回のシーケンシング反応でシーケンシングすることができる。

【0139】

核酸標的の5'末端のバーコード化

本明細書の開示は、分子標識(または分子インデックス)を有するバーコード(例えば、確率的バーコード)の、バーコード化または標識化されている核酸標的(例えば、デオキシリボ核酸分子、およびリボ核酸分子)の5'末端への結合のための、システム、方法、組成物、およびキットを含む。本明細書に開示される5'ベースの転写物計数方法は、例えば、3'ベースの転写物計数方法(例えば、Rhapsody(商標)アッセイ(Becton, Dickinson and Company(Franklin Lakes, NJ))、Chromium(商標)Single Cell 3' Solution(10X Genomics(San Francisco, CA)))を補うか、または補足することができる。バーコード化された核酸標的は、配列同定、転写物の計数、選択的スプライシング分析、変異スクリーニング、および/またはハイスループット方式の全長シーケンシングのために使用され得る。5'末端(標識されている標的核酸標的に対して5'側)での転写物計数によって、核酸分子の5'末端における、またはそれにより近い選択的スプライシングアイソフォームおよびバリエーション(以下に限定されないが、スプライスバリエーション、一塩基多型(SNP)、挿入、欠失、置換を含む)が明らかとなり得る。一部の実施形態では、本方法は、分子内ハイブリダイゼーションに参与し得る。

【0140】

図4A~4Bは、5'末端において核酸標的を遺伝子特異的に標識する非限定的な例示的方法400の略図を示す。標的結合領域(例えば、ポリ(dT)テール422)を含むバ

10

20

30

40

50

ーコード 4 2 0 (例えば、確率的バーコード)は、標識化またはバーコード化(例えば、固有の標識化)するために、ポリ(dA)テール 4 2 6 を介してポリアデニル化 RNA 転写物 4 2 4、または他の核酸標的に結合し得る。バーコード 4 2 0 は、後続反応のために各バーコード 4 2 0 の分子標識 (ML) 4 2 8 / 試料標識 (SL) 4 3 0 領域に隣接する、1つまたは複数の追加の配列(例えば、アダプター配列 4 3 2 などのコンセンサス配列)と一緒に、それぞれ、転写物 4 2 4 を標識し、RNA 転写物 4 2 4 の試料起源を追跡するための分子標識 4 2 8 および試料標識 4 3 0 を含み得る。試料ごとのバーコードにおける分子標識の配列のレパートリーは、RNA 転写物の確率的標識化に対して十分に多量であり得る。

【0141】

RNA 転写物 4 2 4 (またはその一部分)を含むバーコード化された cDNA 分子 4 3 4 を生成するためのブロック 4 0 2 における cDNA 合成後に、遺伝子特異的方法を 5' 分子バーコード化のために使用することができる。選択的であり得るブロック 4 0 4 における遺伝子特異的増幅の後に、ブロック 4 0 6 において末端転写酵素およびデオキシアデノシン三リン酸 (dATP) を添加し、3' ポリ(dA)テールを促進して、ポリ(A)テール 4 3 8 を含むアンプリコン 4 3 6 を生成し得る。ブロック 4 0 8 における短い変性ステップによって、アンプリコン 4 3 6 のフォワード鎖 4 3 6 m およびリバース鎖 4 3 6 c (ポリ(dA)テールを含むバーコード化された cDNA 分子)の離間が可能となる。アンプリコン 4 3 6 のリバース鎖 4 3 6 c は、鎖の 3' 末端およびポリ(dT)領域 4 2 2 末端におけるそのポリ(dA)テール 4 3 8 を介して分子内でハイブリダイズし、ブロック 4 1 0 において、ヘアピンまたはステムループ 4 4 0 を形成し得る。次いで、ポリメラーゼ(例えば、Klenow 断片)を使用して、バーコードを複製するためにポリ(dA)テール 4 3 8 から伸長させて、ブロック 4 1 2 において、伸長したバーコード化されたりリバース鎖 4 4 2 を形成することができる。次いで、ブロック 4 1 4 (例えば、任意に)における遺伝子特異的増幅を、目的の遺伝子を増幅するために実施して、ブロック 4 1 6 において、シーケンシングのために、5' 末端 (RNA 転写物 4 2 4 に対して)においてバーコードを含むアンプリコン 4 4 4 を生成することができる。一部の実施形態では、方法 4 0 0 は、ブロック 4 0 4 におけるバーコード化された cDNA 分子 4 3 4 の遺伝子特異的増幅とブロック 4 1 4 における伸長したバーコード化されたりリバース鎖 4 4 2 の遺伝子特異的増幅のうち的一方または両方を含む。

【0142】

図 5 A ~ 5 B は、全トランスクリプトーム分析のために 5' 末端において核酸標的を標識する非限定的な例示的方法 5 0 0 の略図を示す。標的結合領域(例えば、ポリ(dT)テール 4 2 2)を含むバーコード 4 2 0 (例えば、確率的バーコード)は、標識化またはバーコード化(例えば、固有の標識化)するために、ポリ(dA)テール 4 2 6 を介してポリアデニル化された RNA 転写物 4 2 4、または他の核酸標的に結合し得る。例えば、標的結合領域を含むバーコード 4 2 0 は、標識化またはバーコード化のために、核酸標的に結合し得る。バーコード 4 2 0 は、分子標識 (ML) 4 2 8 および試料標識 (SL) 4 3 0 を含み得る。分子標識 4 2 8 および試料標識 4 3 0 は、転写物 4 2 4、または核酸標的(例えば、抗体と会合したかまたは抗体から解離した、抗体オリゴヌクレオチド)を標識し、転写物 4 2 4 の試料起源を追跡するために、それぞれ、後続反応のための各バーコード 4 2 0 の分子標識 4 2 8 / 試料標識 4 3 0 領域に隣接する、1つまたは複数の追加の配列(例えば、アダプター配列 4 3 2 などのコンセンサス配列)と一緒に、使用することができる。試料ごとのバーコードにおける分子標識 4 2 8 の配列のレパートリーは、RNA 転写物 4 2 4、または核酸標的の確率的標識化に対して十分に多量であり得る。

【0143】

ブロック 4 0 2 におけるバーコード化された cDNA 分子 4 3 4 を生成するための cDNA 合成の後に、末端転写酵素を、バーコード化された cDNA 分子 4 3 4 の 3' 末端(標識された RNA 転写物の 5' 末端に等しい)の A-テールリングのために使用し、ブロック 4 0 6 において、3' ポリ(dA)テール 4 3 8 をそれぞれ含む cDNA 分子 4 3 6 c を生成

10

20

30

40

50

することができる。cDNA分子436cの3'ポリ(dA)テール438との分子内ハイブリダイゼーションが開始され(例えば、加熱および冷却サイクルによるか、またはバーコード化されたcDNA分子436cをポリ(dA)テール438で希釈することによって)、その結果、新たな3'ポリ(dA)テール438が、同じ標識されたcDNA分子のポリ(dT)テール422とアニーリングし、ブロック410において、バーコード化されたcDNA分子のヘアピンまたはステムループ構造440を生成することができる。ブロック412では、新たな3'ポリ(dA)テール438を超える3'伸長を促進するために、dNTPと共にポリメラーゼ(例えば、Klenow酵素)を添加し、バーコード(例えば、ステムループ440を有する標識されたcDNA分子の5'末端に存在する分子標識428)を複製することができる。ブロック414では、鏡像化したアダプター432、432rcまたはアダプター432、432rcの配列(または部分配列)を含有するプライマーを使用して、全トランスクリプトーム増幅(WTA)を実施することができる。ブロック418では、シーケンシングのために(例えば、Illumina(San Diego, CA, U.S.)シーケンサーを使用して)、タグメンテーションまたはランダムプライミングなどの方法を使用して、シーケンシングアダプター(例えば、P5446およびP7448配列)を含むより小さなアンプリコン444の断片を生成することができる。一部の実施形態では、他のシーケンシング方法またはシーケンサー(例えば、Pacific Biosciences of California, Inc.(Menlo Park, CA, US)またはOxford Nanopore Technologies Limited(Oxford, UK)からのシーケンサー)のためのシーケンシングアダプターを直接ライゲーションさせて、シーケンシングのためのアンプリコンを生成することができる。

10

20

【0144】

本明細書の開示は、試料中の核酸標的の数を決定するための方法を含む。一部の実施形態では、本方法は、試料中の核酸標的424のコピーを複数のオリゴヌクレオチドバーコード420に接触させるステップであって、複数のオリゴヌクレオチドバーコード420のそれぞれが、分子標識配列428および核酸標的424にハイブリダイズすることが可能な標的結合領域(例えば、ポリ(dT)配列422)を含み、複数のオリゴヌクレオチドバーコード420のうち少なくとも10個が、異なる分子標識配列428を含む、ステップと、オリゴヌクレオチドバーコード420にハイブリダイズした核酸標的424のコピーを伸長させて、ブロック402において、それぞれが、核酸標的424の少なくとも一部分に相補的な配列450cを含む、複数の核酸分子434を生成するステップと、ブロック404において、複数のバーコード化された核酸分子434を増幅させて、複数の増幅したバーコード化された核酸分子436を生成するステップと、標的結合領域422の相補体438を含むオリゴヌクレオチドを複数の増幅したバーコード化された核酸分子436に結合させて、ブロック406において、それぞれが、標的結合領域422および標的結合領域の相補体438を含む、複数のバーコード化された核酸分子436cを生成するステップと、複数のバーコード化された核酸分子436cのそれぞれにおいて、標的結合領域422と標的結合領域422の相補体438とをハイブリダイズさせて、ブロック410において、ステムループ440を形成するステップと、ブロック412において、それぞれ、ステムループ440を含む複数のバーコード化された核酸分子の3'末端を伸長させて、ステムループ440を伸長させ、それぞれが、分子標識428および該分子標識の相補体428rcを含む、複数の伸長したバーコード化された核酸分子442を生成するステップと、ブロック414において、複数の伸長したバーコード化された核酸分子442を増幅させて、それぞれが、分子標識の相補体428rcを含む、複数の単一標識核酸分子444cを生成するステップと、複数の単一標識核酸分子と会合した別個の配列を有する分子標識の相補体428rcの数に基づいて、試料中の核酸標的の数を決定するステップとを含む。

30

40

【0145】

一部の実施形態では、分子標識428は、ステムループ440を有する複数のバーコー

50

ド化された核酸分子の3'末端を伸長させた後に、分子標識の相補体428rcにハイブリダイズされる。本方法は、複数の伸長したバーコード化された核酸分子442を増幅させる前に、複数の伸長したバーコード化された核酸分子442を変性させて、複数の単一標識核酸分子444c（アンプリコン444cの一部であってもよい）を生成するステップを含み得る。試料中の核酸標的424のコピーを接触させるステップは、複数の核酸標的424のコピーを複数のオリゴヌクレオチドバーコード420に接触させるステップを含み得る。核酸標的424のコピーを伸長させるステップは、オリゴヌクレオチドバーコード420にハイブリダイズした複数の核酸標的424のコピーを伸長させて、それぞれが、複数の核酸標的424のうちの1つの少なくとも一部分に相補的な配列450cを含む、複数のバーコード化された核酸分子436cを生成するステップを含み得る。核酸標的424の数を決定するステップは、複数の核酸標的424のそれぞれの配列452cを含む、複数の単一標識核酸分子444cのうちの単一標識核酸分子と会合した、別個の配列を有する分子標識の相補体428rcの数に基づいて、試料中の複数の核酸標的424のそれぞれの数を決定するステップを含み得る。複数の核酸標的のそれぞれの配列452cは、複数の核酸標的424のそれぞれの部分配列（相補体または逆相補体を含む）を含み得る。

10

【0146】

本明細書の開示は、試料中の標的の数を決定するための方法を含む。一部の実施形態では、本方法は、複数のオリゴヌクレオチドバーコード420を使用して、試料中の核酸標的424のコピーをバーコード化して（402）、それぞれが、核酸標的424の配列450c（例えば、相補配列、逆相補配列、またはこれらの組合せ）、分子標識428、および標的結合領域（例えば、ポリ（dT）領域422）を含む、複数のバーコード化された核酸分子434を生成するステップであって、複数のオリゴヌクレオチドバーコード420のうちの少なくとも10個が、異なる分子標識配列428を含む、ステップと、標的結合領域422の相補体438を含むオリゴヌクレオチドを複数のバーコード化された核酸分子434に結合して（406）、それぞれが、標的結合領域422および該標的結合領域422の相補体438を含む、複数のバーコード化された核酸分子436を生成する、ステップと、複数のバーコード化された核酸分子436cのそれぞれにおいて、標的結合領域422と該標的結合領域の相補体438とをハイブリダイズさせて（410）、ステムループ440を形成するステップと、複数のバーコード化された核酸分子の3'末端を伸長させて（412）、ステムループ440を伸長させ、それぞれが、分子標識428および該分子標識の相補体428rcを含む、複数の伸長したバーコード化された核酸分子を生成するステップと、複数の伸長したバーコード化された核酸分子442と会合した別個の配列を有する分子標識の相補体428rcの数に基づいて、試料中の核酸標的424の数を決定するステップとを含む。

20

30

【0147】

本明細書の開示は、オリゴヌクレオチドバーコードを試料中の標的に結合させるための方法を含む。一部の実施形態では、本方法は、複数のオリゴヌクレオチドバーコード420を使用して、試料中の核酸標的424のコピーをバーコード化して（402）、それぞれが、核酸標的424の配列450c、分子標識428、および標的結合領域422を含む、複数のバーコード化された核酸分子434を生成するステップであって、複数のオリゴヌクレオチドバーコード420のうちの少なくとも10個が、異なる分子標識配列428を含む、ステップと、標的結合領域422の相補体438を含むオリゴヌクレオチドを複数のバーコード化された核酸分子434に結合させて、それぞれが、標的結合領域422および該標的結合領域422の相補体438を含む、複数のバーコード化された核酸分子436cを生成するステップと、複数のバーコード化された核酸分子436cのそれぞれにおいて、標的結合領域422と該標的結合領域422の相補体438とをハイブリダイズさせて（410）、ステムループ440を形成するステップと、複数のバーコード化された核酸分子の3'末端を伸長させて（412）、ステムループ440を伸長させ、それぞれが、分子標識428および該分子標識428の相補体428rcを含む、複数の伸長

40

50

したバーコード化された核酸分子 4 4 2 を生成するステップとを含む。一部の実施形態では、本方法は、複数の伸長したバーコード化された核酸分子 4 4 2 と会合した、別個の配列を有する分子標識 4 2 8、その相補体 4 2 8 r c、またはこれらの組合せの数に基づいて、試料中の核酸標的 4 2 4 の数を決定するステップを含む。例えば、核酸標的 4 2 4 の数は、別個の配列を有する分子標識 4 2 8、その相補体 4 2 8 r c のうち的一方または両方に基づいて決定され得る。

【 0 1 4 8 】

一部の実施形態では、複数の標的 4 2 4 のコピーをバーコード化するステップ (4 0 2) は、核酸標的 4 2 4 のコピーを複数のオリゴヌクレオチドバーコード 4 2 0 に接触させるステップであって、複数のオリゴヌクレオチドバーコード 4 2 0 のそれぞれが、核酸標的 4 2 4 にハイブリダイズさせることが可能な標的結合領域 4 2 2 を含む、ステップと、オリゴヌクレオチドバーコード 4 2 0 にハイブリダイズした核酸標的 4 2 4 のコピーを伸長させて (4 0 2)、複数のバーコード化された核酸分子 4 3 4 を生成するステップとを含む。

10

【 0 1 4 9 】

一部の実施形態では、本方法は、複数のバーコード化された核酸分子 4 3 4 を増幅させて (4 0 4)、複数の増幅したバーコード化された核酸分子 4 3 6 c を生成するステップであって、標的結合領域 4 2 2 の相補体 4 3 8 を含むオリゴヌクレオチドを結合させるステップが、標的結合領域の相補体 4 3 8 を含むオリゴヌクレオチドを複数の増幅したバーコード化された核酸分子に結合させて、それぞれが、標的結合領域 4 2 2 および該標的結合領域の相補体 4 3 8 を含む、複数のバーコード化された核酸分子 4 3 6 r を生成するステップを含む、ステップを含む。

20

【 0 1 5 0 】

遺伝子特異的分析。一部の実施形態では、本方法 (例えば、方法 4 0 0) は、複数の伸長したバーコード化された核酸分子 4 4 2 を増幅させて (4 1 4)、それぞれが、分子標識 4 2 8 の相補体 4 2 8 r c を含む、複数の単一標識核酸分子 4 4 4 c を生成するステップを含む。単一標識核酸分子 4 4 4 c は、それらを含むアンプリコン 4 4 4 が変性される場合に生成され得る。試料中の核酸標的 4 2 4 の数を決定するステップは、複数の単一標識核酸分子 4 4 4 c と会合した、別個の配列を有する分子標識 4 2 8 の相補体 4 2 8 r c の数に基づいて、試料中の核酸標的 4 2 4 の数を決定するステップを含み得る。

30

【 0 1 5 1 】

全トランスクリプトーム分析。一部の実施形態では、本方法 (例えば、方法 5 0 0) は、複数の伸長したバーコード化された核酸分子 4 4 2 を増幅させて (4 1 4)、複数の伸長したバーコード化された核酸分子のコピー 4 4 4 c を生成するステップを含む。試料中の核酸標的 4 2 4 の数を決定するステップは、複数の伸長したバーコード化された核酸分子のコピー 4 4 4 c と会合した、別個の配列を有する分子標識 4 2 8 の相補体 4 2 8 r c の数に基づいて、試料中の核酸標的 4 2 4 の数を決定するステップを含む。複数の伸長したバーコード化された核酸分子のコピー 4 4 4 c は、それらを含むアンプリコン 4 4 4 が変性される場合に形成され得る。

一部の実施形態では、複数のバーコード化された核酸分子中の核酸標的の配列は、核酸標的の部分配列 4 5 2 c を含む。標的結合領域は、遺伝子特異的配列を含み得る。標的結合領域 4 2 2 の相補体 4 3 8 を含むオリゴヌクレオチドを結合させるステップ (4 0 6) は、標的結合領域 4 2 2 の相補体 4 3 8 を含むオリゴヌクレオチドを複数のバーコード化された核酸分子 4 3 4 にライゲーションさせるステップを含み得る。

40

【 0 1 5 2 】

一部の実施形態では、標的結合領域は、ポリ (d T) 配列 4 2 2 を含み得る。標的結合領域 4 2 2 の相補体 4 3 8 を含むオリゴヌクレオチドを結合させるステップは、末端デオキシヌクレオチド転移酵素を使用して、複数のアデノシンリン酸を複数のバーコード化された核酸分子 4 3 4 に付加するステップを含む。

一部の実施形態では、オリゴヌクレオチドバーコード 4 2 0 にハイブリダイズした核酸

50

標的 4 2 4 のコピーを伸長させるステップは、オリゴヌクレオチドバーコード 4 2 0 にハイブリダイズした核酸標的 4 2 4 のコピーを逆転写して、複数のバーコード化された相補デオキシリボ核酸 (c D N A) 分子 4 3 4 を生成するステップを含み得る。オリゴヌクレオチドバーコード 4 2 0 にハイブリダイズした核酸標的 4 2 4 のコピーを伸長させるステップは、5 ' から 3 ' へのエキソヌクレアーゼ活性および 3 ' から 5 ' へのエキソヌクレアーゼ活性のうちの少なくとも 1 つを欠く DNA ポリメラーゼを使用して、オリゴヌクレオチドバーコード 4 2 0 にハイブリダイズした核酸標的 4 2 4 のコピーを伸長させるステップ (4 0 2) を含み得る。DNA ポリメラーゼは、K l e n o w 断片を含んでもよい。

【 0 1 5 3 】

一部の実施形態では、本方法は、複数の伸長したバーコード化された核酸分子 4 4 2 の配列情報を得るステップを含む。配列情報を得るステップは、シーケンシングアダプター (例えば、P 5 4 4 6 および P 7 4 4 8 アダプター) を複数の伸長したバーコード化された核酸分子 4 4 2 に結合させるステップを含み得る。

10

一部の実施形態では、標的結合領域の相補体 4 3 8 は、標的結合領域の逆相補配列を含み得る。標的結合領域の相補体 4 3 8 は、標的結合領域の相補配列を含み得る。分子標識の相補体 4 2 8 r c は、分子標識の逆相補配列を含み得る。分子標識の相補体は、分子標識の相補配列を含み得る。

一部の実施形態では、複数のバーコード化された核酸分子 4 3 4 は、バーコード化されたデオキシリボ核酸 (D N A) 分子を含み得る。バーコード化された核酸分子 4 3 4 は、バーコード化されたリボ核酸 (R N A) 分子を含み得る。核酸標的 4 2 4 は、核酸分子を含み得る。核酸分子は、リボ核酸 (R N A) 、メッセンジャー R N A (m R N A) 、マイクロ R N A 、低分子干渉 R N A (s i R N A) 、R N A 分解産物、ポリ (A) テールを含む R N A 、またはこれらの任意の組合せを含み得る。

20

【 0 1 5 4 】

抗体オリゴヌクレオチド。一部の実施形態では、核酸標的は、細胞成分結合試薬を含み得る。核酸標的 (例えば、試料インデックス化オリゴヌクレオチドなどの抗体オリゴヌクレオチド) に会合した細胞結合試薬については、U S 2 0 1 8 / 0 0 8 8 1 1 2 、および 2 0 1 8 年 3 月 2 7 日に出願された米国出願第 1 5 / 9 3 7 , 7 1 3 号に記載されており、これらの出願のそれぞれの内容は、参照によりその全体が本明細書に組み込まれる。一部の実施形態では、本開示の 5 ' バーコード化方法を使用して、ゲノミクス、クロマチンアクセシビリティ、メチロミクス、トランスクリプトミクス、およびプロテオミクスなどの単一細胞のマルチオミクス情報を得ることができる。核酸分子は、細胞成分結合試薬に会合していてもよい。本方法は、核酸分子と細胞成分結合試薬を解離させるステップを含み得る。

30

【 0 1 5 5 】

一部の実施形態では、複数のオリゴヌクレオチドバーコード 4 2 0 のうちの各分子標識 4 2 8 は、少なくとも 6 個のヌクレオチドを含む。オリゴヌクレオチドバーコード 4 2 0 は、同一の試料標識 4 3 0 を含み得る。複数のオリゴヌクレオチドバーコード 4 2 0 のうちの各試料標識 4 3 0 は、少なくとも 6 個のヌクレオチドを含み得る。オリゴヌクレオチドバーコード 4 2 0 は、同一の細胞標識を含み得る。複数のオリゴヌクレオチドバーコード 4 2 0 のうちの各細胞標識は、少なくとも 6 個のヌクレオチドを含み得る。

40

【 0 1 5 6 】

一部の実施形態では、複数のバーコード化された核酸分子 4 3 6 c のうちの少なくとも 1 つは、該複数のバーコード化された核酸分子のそれぞれにおいて、標的結合領域と該標的結合領域の相補体とをハイブリダイズさせて (4 1 0) 、ステムループを形成する場合に、固体支持体と会合している。複数のバーコード化された核酸分子 4 3 6 c のうちの少なくとも 1 つは、該複数のバーコード化された核酸分子 4 3 6 c のそれぞれにおいて、標的結合領域 4 2 2 と該標的結合領域 4 2 2 の相補体 4 3 8 とをハイブリダイズさせて (4 1 0) 、ステムループ 4 4 0 を形成する場合に、固体支持体から解離し得る。複数のバーコード化された核酸分子 4 3 6 c のうちの少なくとも 1 つは、該複数のバーコード化され

50

た核酸分子436cのそれぞれにおいて、標的結合領域422と該標的結合領域の相補体438とをハイブリダイズさせて(410)、ステムループ440を形成する場合に、固体支持体と会合していてもよい。

【0157】

一部の実施形態では、複数のバーコード化された核酸分子のうちの少なくとも1つは、複数のバーコード化された核酸分子の3'末端を伸長させて(412)、ステムループ440を伸長させ、それぞれが、分子標識428および該分子標識の相補体428rcを含む、複数の伸長したバーコード化された核酸分子442を生成する場合に、固体支持体と会合している。複数のバーコード化された核酸分子のうちの少なくとも1つは、複数のバーコード化された核酸分子の3'末端を伸長させて(412)、ステムループ440を伸長させ、それぞれが、分子標識428および該分子標識の相補体428rcを含む、複数の伸長したバーコード化された核酸分子442を生成する場合に、固体支持体から解離し得る。複数のバーコード化された核酸分子436cのうちの少なくとも1つは、複数のバーコード化された核酸分子の3'末端を伸長させて(412)、ステムループ440を伸長させ、それぞれが、分子標識428および該分子標識の相補体428rcを含む、複数の伸長したバーコード化された核酸分子442を生成する場合に、固体支持体と会合していてもよい。固体支持体は、合成粒子454を含み得る。固体支持体は、平面表面または実質的平面表面(例えば、顕微鏡スライドまたはカバーガラスなどのスライド)を含み得る。

10

【0158】

一部の実施形態では、複数のバーコード化された核酸分子436cのうちの少なくとも1つは、複数のバーコード化された核酸分子436cのそれぞれにおいて、標的結合領域422と該標的結合領域422の相補体438とをハイブリダイズさせて(410)、ステムループ440を形成する場合に、溶液中にある。例えば、溶液中の複数のバーコード化された核酸分子436cの濃度が十分に低い場合に、このような分子内ハイブリダイゼーションが起こり得る。複数のバーコード化された核酸分子のうちの少なくとも1つは、複数のバーコード化された核酸分子の3'末端を伸長させて(412)、ステムループ440を伸長させ、それぞれが、分子標識428および該分子標識の相補体428rcを含む、複数の伸長したバーコード化された核酸分子442を生成する場合に、溶液中にあり得る。

20

【0159】

一部の実施形態では、試料は、単一細胞を含み、本方法は、複数のオリゴヌクレオチドバーコード420を含む合成粒子454を、試料中の単一細胞と会合させるステップを含む。本方法は、合成粒子454を単一細胞と会合させた後に、単一細胞を溶解させるステップを含み得る。単一細胞を溶解させるステップは、試料を加熱するステップ、試料を界面活性剤と接触させるステップ、試料のpHを変更するステップ、またはこれらの任意の組合せを含み得る。合成粒子と単一細胞は、同じウェル中にあり得る。合成粒子と単一細胞は、同じ液滴中にあり得る。

30

【0160】

一部の実施形態では、複数のオリゴヌクレオチドバーコード420のうちの少なくとも1つは、合成粒子454上に固定化されていてもよい。複数のオリゴヌクレオチドバーコード420のうちの少なくとも1つは、合成粒子454上に部分的に固定化されていてもよい。複数のオリゴヌクレオチドバーコード420のうちの少なくとも1つは、合成粒子454内に封入されていてもよい。複数のオリゴヌクレオチドバーコード420のうちの少なくとも1つは、合成粒子454内に部分的に封入されていてもよい。合成粒子454は、崩壊可能であり得る。合成粒子454は、ビーズを含み得る。ビーズは、セファロースビーズ、ストレプトアビジンビーズ、アガロースビーズ、磁性ビーズ、コンジュゲートビーズ、プロテインAコンジュゲートビーズ、プロテインGコンジュゲートビーズ、プロテインA/Gコンジュゲートビーズ、プロテインLコンジュゲートビーズ、オリゴ(dT)コンジュゲートビーズ、シリカビーズ、シリカ様ビーズ、抗ビオチンマイクロビーズ、抗蛍光色素マイクロビーズ、またはこれらの任意の組合せを含み得る。合成粒子454は

40

50

、ポリジメチルシロキサン（PDMS）、ポリスチレン、ガラス、ポリプロピレン、アガロース、ゼラチン、ハイドロゲル、常磁性体、セラミック、プラスチック、ガラス、メチルスチレン、アクリルポリマー、チタン、ラテックス、セファロース、セルロース、ナイロン、シリコン、およびこれらの任意の組合せからなる群から選択される材料を含み得る。合成粒子454は、崩壊可能ハイドロゲル粒子を含み得る。複数のオリゴヌクレオチドバーコード420のそれぞれは、リンカー官能基を含み得る。合成粒子454は、固体支持体官能基を含み得る。支持体官能基およびリンカー官能基は、互いに会合していてもよい。リンカー官能基および支持体官能基は、独立して、C6、ビオチン、ストレプトアビジン、第一級アミン、アルデヒド、ケトン、およびこれらの任意の組合せからなる群から選択され得る。

10

【0161】

核酸標的の5'末端におけるバーコード化のためのキット

本明細書の開示は、オリゴヌクレオチドバーコード420を試料中の標的424に結合させる、試料中の標的424の数を決定する、および/または試料中の核酸標的424の数を決定するためのキットを含む。一部の実施形態では、キットは、複数のオリゴヌクレオチドバーコード420であって、複数のオリゴヌクレオチドバーコード420のそれぞれが、分子標識428および標的結合領域（例えば、ポリ(dT)配列422）を含み、複数のオリゴヌクレオチドバーコード420のうちの少なくとも10個が異なる分子標識配列428を含む、複数のオリゴヌクレオチドバーコード420；末端デオキシヌクレオチド転移酵素またはリガーゼ；ならびに5'から3'へのエキソヌクレアーゼ活性および3'から5'へのエキソヌクレアーゼ活性のうちの少なくとも1つを欠くDNAポリメラーゼを含む。DNAポリメラーゼは、Klenow断片を含み得る。キットは、緩衝液および/またはカートリッジを含み得る。キットは、逆転写反応のための1つまたは複数の試薬を含み得る。キットは、増幅反応のための1つまたは複数の試薬を含み得る。

20

【0162】

一部の実施形態では、標的結合領域は、遺伝子特異的配列、オリゴ(dT)配列、ランダム多量体、またはこれらの任意の組合せを含む。オリゴヌクレオチドバーコードは、同一の試料標識および/または同一の細胞標識を含み得る。複数のオリゴヌクレオチドバーコードの各試料標識および/または細胞標識は、少なくとも6個のヌクレオチドを含み得る。複数のオリゴヌクレオチドバーコードの各分子標識は、少なくとも6個のヌクレオチドを含み得る。

30

【0163】

一部の実施形態では、複数のオリゴヌクレオチドバーコード420のうちの少なくとも1つは、合成粒子454上に部分的に固定化されている。複数のオリゴヌクレオチドバーコード420のうちの少なくとも1つは、合成粒子454上に部分的に固定化されていてもよい。複数のオリゴヌクレオチドバーコード420のうちの少なくとも1つは、合成粒子454内に封入されていてもよい。複数のオリゴヌクレオチドバーコード420のうちの少なくとも1つは、合成粒子454内に部分的に封入されていてもよい。合成粒子454は、崩壊可能であり得る。合成粒子454は、ビーズを含み得る。ビーズは、セファロースビーズ、ストレプトアビジンビーズ、アガロースビーズ、磁性ビーズ、コンジュゲートビーズ、プロテインAコンジュゲートビーズ、プロテインGコンジュゲートビーズ、プロテインA/Gコンジュゲートビーズ、プロテインLコンジュゲートビーズ、オリゴ(dT)コンジュゲートビーズ、シリカビーズ、シリカ様ビーズ、抗ビオチンマイクロビーズ、抗蛍光色素マイクロビーズ、またはこれらの任意の組合せを含み得る。合成粒子は、ポリジメチルシロキサン（PDMS）、ポリスチレン、ガラス、ポリプロピレン、アガロース、ゼラチン、ハイドロゲル、常磁性体、セラミック、プラスチック、ガラス、メチルスチレン、アクリルポリマー、チタン、ラテックス、セファロース、セルロース、ナイロン、シリコン、およびこれらの任意の組合せからなる群から選択される材料を含み得る。合成粒子454は、崩壊可能ハイドロゲル粒子を含み得る。複数のオリゴヌクレオチドバーコードのそれぞれは、リンカー官能基を含み得る。合成粒子454は、固体支持体官能

40

50

基を含み得る。支持体官能基およびリンカー官能基は、互いに会合していてもよい。リンカー官能基および支持体官能基は、独立して、C6、ビオチン、ストレプトアビジン、第一級アミン、アルデヒド、ケトン、およびこれらの任意の組合せからなる群から選択され得る。

【0164】

完全長発現プロファイリング

ハイスループット単一細胞RNAシーケンシングによって、複雑かつ不均一な生体試料についての理解は変化した。しかし、ほとんどの方法はmRNA転写物情報の3'のみの分析しか可能にせず、スプライスバリエーション、選択的転写開始部位およびT細胞およびB細胞の受容体と抗体とのVDJ接合などの再編成による可変性の高い遺伝子座の分析が制限される場合がある。5'バーコード化および/または3'バーコード化を使用して、核酸標的(例えば、免疫受容体のV(D)J領域)の配列を決定するための方法は、その内容が参照によりその全体として本明細書に組み込まれる、2019年9月30日に出願された、米国特許出願第16/588,405号に記載されている。図9Aは、免疫受容体mRNAの5'発現プロファイリングの非限定的な例示的略図である。免疫受容体のV(D)J領域の5'発現プロファイリングによって、TCR/BCR被覆率を改善することができ、CDR3配列情報を得ることができるが、このアプローチは、V領域を失う場合がある(図9Aにおいて点線のボックスによって示される)。T細胞とB細胞の両方に関して、現在利用可能なCプライミングベースのアプローチは、V(D)J中へと読み進むことができるが、上流のV領域を失う。よって、現在利用可能な方法は、完全長核酸標的(例えば、V(D)Jを含有する転写物)の情報を得るための能力を限定する場合がある。当技術分野における特定の問題は、多数の再構成事象が起こり得ることにより多くのVDJが存在するため、より長いリードとしてVDJ配列を知る必要があるということである。配列(例えば、V(D)Jを含有する転写物)の計数と前記配列の同定(特に、完全長配列の同定)の両方の方法に対する必要性が存在する。

【0165】

一部の実施形態では、ランダムプライミングアプローチを使用して、完全長V(D)J情報を得るための方法(例えば、RhapsodyシステムでのIlluminaシーケンシングによる)が提供される。TおよびB細胞受容体は、Vセグメント、Dセグメント(TCRベータおよびBCR重鎖のみに関する)、JセグメントおよびmRNAの3'末端の定常領域を含有する。V(D)J接合部により構成されたCDR3は、レパートリー多様性のバルクを含有し、Illuminaショートリードプラットフォームでシーケンシングされるのに十分短い。しかし、Illuminaショートリードの性能では完全長V(D)J情報を得るための能力が限定されるため、完全長Vセグメント情報も有用であり、これは、ロングリードシーケンシング技術を用いなければ容易に得ることができない。本明細書において提供される方法は、ランダムプライミングおよび伸長を用いて、完全長V(D)Jアンプリコンと、部分的なVセグメント配列しか含有しないより短いアンプリコンの両方を含有するライブラリーを作成することができ、これにより、ユーザーは、単一のライブラリーおよびIlluminaシーケンサーに対応するシーケンシングのランからCDR3情報と完全長Vセグメント配列の両方を得ることが可能になる。よって、本明細書において提供される方法の一部の実施形態により、完全長免疫受容体mRNA配列が得られる。

【0166】

現在利用可能な5'VDJアッセイによって、断片化アプローチを使用して完全長の情報を取得することができ、これはRhapsodyでも可能であるが、よりコストのかかる酵素と試薬が必要とされる。さらに、本明細書において提供されるランダムプライミングアプローチ(ランダムプライミングおよび伸長のみを含む)に比べて、より多くの酵素ステップが、この断片化ベースのアプローチ(例えば、断片化、末端修復、a-テーリング、ライゲーション)に関与する。断片化は、プライマー部位を付加するために酵素の使用を必要とする場合もあるが、本明細書において提供されるランダムプライミングベースの

10

20

30

40

50

方法の一部の実施形態は公知のプライマー部位のライゲーションを必要としない。

【0167】

図7は、完全長発現プロファイリングを実施する非限定的な例示的ワークフローの概略図である。本方法は、TCR/BCR濃縮PCR産物に関するランダムプライミングを含み得る。TCR/BCR濃縮PCR産物は、本明細書において提供される伸長したバーコード化された核酸分子の増幅の産物であり得る（例えば、第1の複数のバーコード化されたアンプリコン）。TCR/BCR濃縮PCR産物は、核酸標的の配列にハイブリダイズすることが可能な標的的特異的プライマーおよび第1のユニバーサル配列を含むプライマーを使用して、複数の伸長したバーコード化された核酸分子を増幅する結果であり得る。ランダムプライミングの性能により、より短い産物が生成され得る。センス鎖プライミングおよびアンチセンス鎖プライミングにより、ランダムプライミング伸長産物を得ることができる。本方法は、PCR3を実施することを含み得る。細胞バーコードおよびUMIを含有するセンス鎖産物のPCR3増幅によって、シーケンシングライブラリーが作成され得る。しかし、細胞バーコードとUMIがないアンチセンス鎖産物の増幅は生じない。ランダムプライマーの結合部位に応じて、様々な伸長産物が生成され得る。図8は、本明細書において提供される方法に従って生成された非限定的な例示的バイオアナライザーによる追跡を示す。したがって、本発明では、V領域のより長いリードを得るための手段が提供される（特定の試料のVDJを同定するために）。IlluminaはC領域から読み取り、正常なリード長は約150bpであり、Vの末端に達するシーケンシングデータが得られる（本明細書において提供される伸長したバーコード化された核酸分子で開始する場合）。PCR3プライマーは、オーバーハングによって、シーケンシングアダプター（例えば、P5およびP7）および試料インデックス（例えば、i5、i7）を付加し得る。本方法は、完全長V(D)Jをシーケンシングするステップと、生物情報学的に再構築するステップ（例えば、図9Bに示されるように複数のシーケンシングリードをアライメントするステップ）を含み得る。一部の実施形態では、本明細書において提供される方法は、核酸標的（例えば、V(D)Jを含有する転写物）を同定し、かつ前記核酸標的のコピー数を計数する。

【0168】

一部の実施形態では、全長核酸標的（例えば、転写物）の情報を得るための組成物、方法、システム、およびキットが提供される。一部の実施形態では、本発明において提供される方法に従って生成された伸長したバーコード化された核酸分子が、ランダムプライミングおよび伸長を実施するためのテンプレートとして用いられる。一部の実施形態では、伸長したバーコード化された核酸分子が増幅され（例えば、標的的特異的プライマーおよび第1のユニバーサル配列を含むプライマーを用いて）、第1の複数のバーコード化されたアンプリコンを生成する。一部の実施形態では、第1の複数のバーコード化されたアンプリコンをランダムプライミングおよび伸長を実施するためのテンプレートとして使用して（例えば、第2のユニバーサル配列、またはその相補体を含むランダムプライマーを使用して）、複数の伸長産物を生成する。ランダムプライマーは、すべての転写物のコード配列に沿って様々な位置に結合し、伸長して、複数の伸長産物（例えば、直線的に増幅された産物）を生成することができる。伸長産物は、ランダムプライマーの結合部位に応じて様々な長さであるcDNAを含み得る。伸長産物は、シーケンシングライブラリー増幅プライマーを用いて増幅され、第2の複数のバーコード化されたアンプリコンを生成することができる。シーケンシングライブラリー増幅は、オーバーハングによってシーケンシングアダプターおよび/またはライブラリーインデックスを付加するライブラリーフォワードおよびリバースプライマーを使用することを含んでもよい。ライブラリー増幅は、ライブラリーフォワードプライマーおよびライブラリーリバースプライマーにおいて、オーバーハングによってシーケンシングアダプター（例えば、P5およびP7配列）および試料インデックス（例えば、i5、i7）を付加し得る。本方法は、第2の複数のバーコード化されたアンプリコン（またはその産物）の配列情報を得るステップを含んでもよい。第2の複数のバーコード化されたアンプリコンをシーケンシングし、本開示の下流方法に供すること

10

20

30

40

50

ができる。150bp×2シーケンシングリードを生成するために使用するペアエンドシーケンシングによって、細胞標識、固有の分子インデックス、ポリ(A)テール、および/またはリード1上の遺伝子(または遺伝子の部分的配列)、リード2上の遺伝子(または遺伝子の部分的配列)および/またはポリ(A)テール、ならびにインデックス1リード上の試料インデックスが明らかになり得る。第2の複数のバーコード化されたアンプリコン(またはその産物)の配列情報を得るステップは、第2の複数のバーコード化されたアンプリコン(またはその産物)の複数のシーケンシングリードを含むシーケンシングデータを得ることを含み得る。本方法は、複数のシーケンシングリードのそれぞれをアライメントすることによって、核酸標的の完全長配列を生成することを含んでもよい。

【0169】

B細胞受容体(BCR)、T細胞受容体(TCR)、および抗体のVDJ領域を同定するために、本開示の方法を使用することができる。体細胞組換えとしても公知の、VDJ組換えは、免疫系の免疫グロブリン(Ig)(例えば、BCR)およびT細胞受容体(TCR)産生の初期段階における遺伝子組換えの機序である。VDJ組換えによって、可変(V)、多様性(D)および接合性(J)遺伝子セグメントが、ほぼランダムに組み合わせられ得る。様々な遺伝子を選択する際のそのランダム性のために、細菌、ウイルス、寄生生物、腫瘍細胞などの機能不全細胞および花粉に由来する抗原に適合するタンパク質を多様にコードすることが可能となる。

【0170】

VDJ領域は、可変(V)遺伝子、多様性(D)遺伝子および接合性(J)遺伝子を含む3Mbの大きな遺伝子座を含み得る。これらは、VDJ組換えに関与し得るセグメントである。VDJ組換えを受けない可能性もある定常遺伝子も存在し得る。この遺伝子座のVDJ組換えにおける最初の事象は、D遺伝子のうちの1つをJ遺伝子のうちの1つに対して再編成することであり得る。この後に、V遺伝子のうちの1つがこのDJ再編成に付加されて、後に重鎖タンパク質の可変セグメントをコードする機能的なVDJ再編成遺伝子を形成することができる。これらのステップの両方が、介在DNAを欠失し得るリコンビナーゼ酵素によって触媒され得る。

【0171】

この組換えプロセスは、前駆体B細胞において段階的に起こり、抗体レパートリーに必要とされる多様性をもたらす。各B細胞は、1つの抗体(例えば、BCR)のみを生じ得る。この特異性は、1つの対立遺伝子の機能的再編成によって、第2の対立遺伝子のさらなる組換えを防ぐためのシグナルが伝達されるような、対立遺伝子排除によって達成され得る。

一部の実施形態では、試料は、免疫細胞を含む。免疫細胞としては、例えば、T細胞、B細胞、リンパ球系幹細胞、骨髄前駆細胞、リンパ球、顆粒球、B細胞前駆体、T細胞前駆体、ナチュラルキラー細胞、Tc細胞、Th細胞、形質細胞、メモリー細胞、好中球、好酸球、好塩基球、マスト細胞、単球、樹状細胞および/もしくはマクロファージ、またはこれらの任意の組合せを挙げることができる。

【0172】

T細胞は、単一T細胞に由来するか、または同一のTCRを有するT細胞を指し得るT細胞クローンであってもよい。T細胞は、T細胞クローンおよび異なるTCRであるが、そのすべてが同じ標的(例えば、抗原、腫瘍、ウイルス)を認識することができる、TCRを有するT細胞の混合集団を含み得るT細胞株の一部であってもよい。末梢血単核細胞、骨髄、リンパ節組織、脾組織、および腫瘍を含む、いくつかの供給源からT細胞を得ることができる。Ficoll分離を使用するなどして、対象から採取された血液単位からT細胞を得ることができる。アフエーシスまたは白血球アフエーシスによって、個体の循環血に由来する細胞を得ることができる。アフエーシス産物は、T細胞、単球、顆粒球、B細胞を含むリンパ球、他の有核白血球、赤血球、および血小板を含み得る。細胞を洗浄し、培地中に再懸濁させて、目的の細胞を単離することができる。

【0173】

10

20

30

40

50

赤血球を溶解させ、例えば、PERCOLL（商標）勾配による遠心分離によって単球を枯渇させることによって、末梢血リンパ球からT細胞を単離することができる。ポジティブまたはネガティブ選択技法によって、T細胞の特異的亜集団、例えば、CD28⁺、CD4⁺、CDC、CD45RA⁺、およびCD45RO⁺ T細胞をさらに単離することができる。例えば、所望のT細胞のポジティブ選択に十分な時間、抗CD3/抗CD28（すなわち、3×28）コンジュゲートビーズ、例えば、DYNABEADS（登録商標）M-450 CD3/CD28 T、またはXCYTE DYNABEADS（商標）とのインキュベーションによって、T細胞を単離することができる。免疫細胞（例えば、T細胞およびB細胞）は、抗原特異的（例えば、腫瘍に対して特異的）であり得る。

一部の実施形態では、細胞は、B細胞、リンパ節由来の活性化B細胞、リンパ芽球様細胞、休止B細胞、または例えば、リンパ腫由来の新生物B細胞などの抗原提示細胞（APC）であってもよい。APCは、表面にBCRCタンパク質のうちの少なくとも1つを発現するB細胞または濾胞樹状細胞を指し得る。

【0174】

本開示の方法を使用して、単一T細胞の分子表現型を追跡することができる。T細胞の様々なサブタイプは、様々な分子マーカーの発現によって識別することができる。T細胞は、TCRの多様なレパートリーから固有のT細胞受容体（TCR）を発現する。ほとんどのT細胞では、TCRは、および鎖のヘテロ二量体から構成されてもよく、各機能鎖は、T細胞発生前の体細胞DNA組換え事象の産物であってもよく、単一の個体において100万を超える異なるTCRの発現を可能にする。TCRを使用して、個々のT細胞の同一性を定義することができ、免疫応答中のT細胞クローン増殖の系統追跡が可能となる。本開示の免疫学的方法は、以下に限定されないが、単一T細胞における固有のTCR鎖とTCR鎖の対合を特定すること、単一細胞レベルでTCRおよびマーカーの発現を定量すること、個体におけるTCRの多様性を特定すること、異なるT細胞集団において発現されるTCRレパートリーを特徴付けること、TCRのアルファ鎖およびベータ鎖の対立遺伝子の機能性を決定すること、および免疫応答の間のT細胞のクローン増殖を特定することを含む、様々な方法で使用され得る。

【0175】

T細胞受容体鎖の対合

T細胞受容体（TCR）は、Tリンパ球表面に存在する認識分子である。T細胞表面に見られるT細胞受容体は、アルファ鎖およびベータ鎖と称される2つの糖タンパク質サブユニットから構成され得る。両鎖は、約40kDaの分子量を含み、可変ドメインおよび定常ドメインを有し得る。アルファ鎖とベータ鎖をコードする遺伝子は、遺伝子再編成によって遺伝子が形成されるV、DおよびJ領域のライブラリーにおいて編成され得る。TCRは、組織適合遺伝子によってコードされる特異的自己分子との複合体の一部としての抗原提示細胞によって提示される抗原を認識し得る。ほとんどの有力な組織適合遺伝子は、主要組織適合複合体（MHC）として公知である。したがって、T細胞受容体によって認識される複合体は、MHC/ペプチドリガンドからなる。

【0176】

一部の実施形態では、本開示の方法、デバイス、およびシステムは、T細胞受容体のシーケンシングおよび対合に使用することができる。本開示の方法、デバイス、およびシステムは、T細胞受容体アルファ鎖およびベータ鎖をシーケンシングする、アルファ鎖およびベータ鎖を対合させる、ならびに/またはT細胞受容体アルファ鎖の機能的コピーを決定するために使用することができる。単一細胞を単一の固体支持体（例えば、ビーズ）を含む単一の区画（例えば、ウェル）に含有させることができる。細胞は溶解されてもよい。ビーズは、TCRのアルファ鎖および/またはベータ鎖内の特定の位置に結合することができる確率的標識を含み得る。固体支持体に会合したTCRのアルファ分子およびベータ分子は、逆転写、増幅、およびシーケンシングを含む本開示の分子生物学的方法に供することができる。同じ細胞標識を含むTCRのアルファ鎖およびベータ鎖は、同じ単一細胞に由来すると考えられ、それによって、TCRのアルファ鎖とベータ鎖を対合させるこ

10

20

30

40

50

とができる。

【0177】

抗体レパートリーにおける重鎖と軽鎖の対合

本開示の方法、デバイスおよびシステムは、BCRの受容体および抗体の重鎖と軽鎖の対合に使用することができる。本開示の方法は、決定される個々の生物または細胞集団における免疫受容体および抗体のレパートリーを可能にする。本開示の方法は、免疫受容体を構成するポリペプチド鎖の対を決定するのを助け得る。B細胞およびT細胞はそれぞれ、免疫受容体を発現し、B細胞は、免疫グロブリンおよびBCRを発現し、T細胞は、T細胞受容体(TCR)を発現する。両タイプの免疫受容体は、2つのポリペプチド鎖を含み得る。免疫グロブリンは、可変重(VH)鎖および可変軽(VL)鎖を含み得る。2つのタイプのTCRが存在し、1つは、アルファ鎖およびベータ鎖からなり、1つは、デルタ鎖およびガンマ鎖からなる。免疫受容体におけるポリペプチドは、定常領域および可変領域を含み得る。可変領域は、B細胞またはT細胞の染色体における遺伝子断片の組換えおよび末端接合再編成から生じ得る。B細胞では、可変領域のさらなる多様化が体細胞超変異によって生じ得る。

10

【0178】

免疫系は、受容体の大きなレパートリーを有し、リンパ球によって発現される任意の所与の受容体対は、別々の、固有の転写物の対によってコードされ得る。単一細胞において発現される免疫受容体鎖の対の配列についての知識を使用して、所与の個体または細胞集団の免疫レパートリーを確認することができる。

20

一部の実施形態では、本開示の方法、デバイス、およびシステムは、抗体のシーケンシングおよび対合に使用することができる。本開示の方法、デバイス、およびシステムは、抗体重鎖および軽鎖のシーケンシング(例えば、B細胞における)、ならびに/または重鎖および軽鎖の対合に使用することができる。単一細胞を単一の固体支持体(例えば、ビーズ)を含む単一の区画(例えば、ウェル)に含有させることができる。細胞は溶解されてもよい。

ビーズは、抗体の重鎖および/または軽鎖(例えば、B細胞における)内の特異的位置に結合することができる確率的標識を含み得る。固体支持体に会合した重鎖および軽鎖の分子は、逆転写、増幅、およびシーケンシングを含む本開示の分子生物学的方法に供することができる。同じ細胞標識を含む抗体の重鎖および軽鎖は、同じ単一細胞に由来すると考えられ、それによって、抗体の重鎖と軽鎖を対合させることができる。

30

【0179】

完全長発現プロファイリングの方法

一部の実施形態では、試料中の核酸標的を標識する方法が提供される。一部の実施形態では、本方法は、核酸標的のコピーを複数のオリゴヌクレオチドバーコードと接触させるステップであって、各オリゴヌクレオチドバーコードが、第1のユニバーサル配列、分子標識(例えば、第1の分子標識)、および核酸標的にハイブリダイズすることが可能な標的結合領域を含む、ステップと;それぞれが、標的結合領域および該標的結合領域の相補体を含む複数のバーコード化された核酸分子を生成するステップと;各バーコード化された核酸分子の標的結合領域の相補体を、(i)複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコード、(ii)バーコード化された核酸分子自体、および(iii)複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子のうちの1つまたは複数の標的結合領域とハイブリダイズさせるステップと;複数のバーコード化された核酸分子の3'末端を伸長させて、それぞれが、第1の分子標識および第2の分子標識を含む、複数の伸長したバーコード化された核酸分子を生成するステップと;核酸標的の配列にハイブリダイズすることが可能な標的的特異的プライマーおよび第1のユニバーサル配列を含むプライマーを使用して、複数の伸長したバーコード化された核酸分子を増幅させて、それによって、核酸標的の配列、またはその一部分を含む第1の複数のバーコード化されたアンプリコンを生成するステップと;ランダムプライマーを第1の複数のバーコード化されたアンプリコンにハイブリダイズさせ、ランダムプライマーを伸

40

50

長させて、複数の伸長産物を生成するステップであって、ランダムプライマーが、第2のユニバーサル配列、またはその相補体を含む、ステップと；第1のユニバーサル配列および第2のユニバーサル配列、またはその相補体にハイブリダイズすることが可能なプライマーを使用して、複数の伸長産物を増幅させて、それによって、第2の複数のバーコード化されたアンプリコンを生成するステップとを含む。一部の実施形態では、本方法は、第2の複数のバーコード化されたアンプリコン、またはその産物と会合した、別個の配列を有する第1の分子標識、別個の配列を有する第2の分子標識、またはこれらの組合せの数に基づいて、試料中の核酸標的のコピー数を決定するステップを含む。一部の実施形態では、第1のユニバーサル配列と第2のユニバーサル配列は同一である。一部の実施形態では、第1のユニバーサル配列と第2のユニバーサル配列は異なる。

10

【0180】

一部の実施形態では、試料中の核酸標的の数を決定する方法が提供される。一部の実施形態では、本方法は、核酸標的のコピーを複数のオリゴヌクレオチドバーコードと接触させるステップであって、各オリゴヌクレオチドバーコードが、第1のユニバーサル配列、分子標識（例えば、第1の分子標識）、および核酸標的にハイブリダイズすることが可能な標的結合領域を含む、ステップと；それぞれが、標的結合領域および該標的結合領域の相補体を含む複数のバーコード化された核酸分子を生成するステップと；各バーコード化された核酸分子の標的結合領域の相補体を、(i) 複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコード、(ii) バーコード化された核酸分子自体、および(iii) 複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子のうちの1つまたは複数の標的結合領域とハイブリダイズさせるステップと；複数のバーコード化された核酸分子の3'末端を伸長させて、それぞれが、第1の分子標識および第2の分子標識を含む、複数の伸長したバーコード化された核酸分子を生成するステップと；核酸標的の配列にハイブリダイズすることが可能な標的特定のプライマーおよび第1のユニバーサル配列を含むプライマーを使用して、複数の伸長したバーコード化された核酸分子を増幅させて、それによって、核酸標的の配列、またはその一部分を含む第1の複数のバーコード化されたアンプリコンを生成するステップと；ランダムプライマーを第1の複数のバーコード化されたアンプリコンにハイブリダイズさせ、ランダムプライマーを伸長させて、複数の伸長産物を生成するステップであって、ランダムプライマーが、第2のユニバーサル配列、またはその相補体を含む、ステップと；第1のユニバーサル配列および第2のユニバーサル配列、またはその相補体にハイブリダイズすることが可能なプライマーを使用して、複数の伸長産物を増幅させて、それによって、第2の複数のバーコード化されたアンプリコンを生成するステップと；第2の複数のバーコード化されたアンプリコン、またはその産物と会合した、別個の配列を有する第1の分子標識、別個の配列を有する第2の分子標識、またはこれらの組合せの数に基づいて、試料中の核酸標的のコピー数を決定するステップとを含む。

20

30

【0181】

試料中の核酸標的のコピー数を決定するステップは、(a) 第2の複数のバーコード化されたアンプリコン、もしくはその産物と会合した、別個の配列を有する第2の分子標識の数に基づいて、試料中の核酸標的のコピー数を決定するステップ、および/または(b) 第2の複数のバーコード化されたアンプリコン、もしくはその産物と会合した、別個の配列を有する第1の分子標識の数に基づいて、試料中の核酸標的のコピー数を決定するステップを含み得る。一部の実施形態では、本方法は、各バーコード化された核酸分子の標的結合領域の相補体を、(i) 複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコード、(ii) バーコード化された核酸分子自体、および/または(iii) 複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子の標的結合領域とハイブリダイズさせる前に、複数のバーコード化された核酸分子を変性させるステップを含む。

40

【0182】

本方法は、複数の伸長したバーコード化された核酸分子を増幅させる前に、複数の伸長

50

したバーコード化された核酸分子を変性させるステップを含み得る。複数の核酸標的のそれぞれの配列は、複数の核酸標的のそれぞれの部分配列を含み得る。複数のバーコード化された核酸分子中の核酸標的の配列は、核酸標的の部分配列を含み得る。一部の実施形態では、第1の分子標識は、複数のバーコード化された核酸分子の3'末端を伸長させた後に、第2の分子標識にハイブリダイズされる。伸長したバーコード化された核酸分子はそれぞれ、第1の分子標識、第2の分子標識、標的結合領域、および該標的結合領域の相補体を含み得る。一部の実施形態では、標的結合領域の相補体は、標的結合領域の一部に相補的である。標的結合領域は、遺伝子特異的配列、ポリ(dT)配列、または両方を含み得る。

それぞれが、標的結合領域および該標的結合領域の相補体を含む複数のバーコード化された核酸分子を生成するステップは、(i)複数のオリゴヌクレオチドバーコードを使用して、試料中の核酸標的のコピーをバーコード化して、それぞれが、核酸標的の少なくとも一部分に相補的な配列、第1の分子標識、および標的結合領域を含む、複数のバーコード化された核酸分子を生成するステップと；(ii)標的結合領域の相補体を含むオリゴヌクレオチドを複数のバーコード化された核酸分子に結合させて、それぞれが、標的結合領域および該標的結合領域の相補体を含む複数のバーコード化された核酸分子を生成するステップとを含み得る。

【0183】

それぞれが、標的結合領域および該標的結合領域の相補体を含む複数のバーコード化された核酸分子を生成するステップは、(i)複数のオリゴヌクレオチドバーコードを使用して、試料中の核酸標的のコピーをバーコード化して、それぞれが、核酸標的の少なくとも一部分に相補的な配列、第1の分子標識、および標的結合領域を含む、複数のバーコード化された核酸分子を生成するステップと；(ii)複数のバーコード化された核酸分子を増幅させて、複数の増幅したバーコード化された核酸分子を生成するステップと；(iii)標的結合領域の相補体を含むオリゴヌクレオチドを複数の増幅したバーコード化された核酸分子に結合させて、それぞれが、標的結合領域および該標的結合領域の相補体を含む複数のバーコード化された核酸分子を生成するステップとを含み得る。

【0184】

試料中の核酸標的のコピーをバーコード化するステップは、核酸標的のコピーにハイブリダイズした複数のオリゴヌクレオチドバーコードを伸長させて、複数のバーコード化された核酸分子を生成するステップを含み得る。核酸標的のコピーにハイブリダイズした複数のオリゴヌクレオチドバーコードを伸長させるステップは、複数のオリゴヌクレオチドバーコードにハイブリダイズした核酸標的のコピーを逆転写するステップを含み得る。標的結合領域の相補体を含むオリゴヌクレオチドを結合させるステップは、該標的結合領域の相補体を含むオリゴヌクレオチドを、複数のバーコード化された核酸分子および/または増幅したバーコード化された核酸分子にライゲーションさせるステップを含み得る。一部の実施形態では、標的結合領域はポリ(dT)配列を含み、該標的結合領域の相補体を含むオリゴヌクレオチドを結合させるステップは、末端デオキシヌクレオチド転移酵素を使用して、複数のアデノシンリン酸を複数のバーコード化された核酸分子および/または増幅したバーコード化された核酸分子に付加するステップを含む。それぞれが、標的結合領域および該標的結合領域の相補体を含む複数のバーコード化された核酸分子を生成するステップは、逆転写酵素および標的結合領域またはその一部分を含むプレートスィッチオリゴヌクレオチドの存在下で、核酸標的のコピーにハイブリダイズした複数のオリゴヌクレオチドバーコードを伸長させて、それぞれが、核酸標的の少なくとも一部分に相補的な配列、第1の分子標識、標的結合領域、および該標的結合領域の相補体を含む、複数のバーコード化された核酸分子を生成するステップを含む。一部の実施形態では、複数のバーコード化された核酸分子の3'末端を伸長させるステップは、5'から3'へのエキソヌクレアーゼ活性および3'から5'へのエキソヌクレアーゼ活性のうちの少なくとも1つを欠くDNAポリメラーゼを使用して、複数のバーコード化された核酸分子の3'末端を伸長させるステップを含む。一部の実施形態では、DNAポリメラーゼは、Klenow断

10

20

30

40

50

片を含む。

【0185】

ランダムプライミングおよび伸長のサイクル数は、異なる実装形態では異なり得る。一部の実施形態では、ランダムプライミングおよび伸長のサイクル数は、1回、2回、3回、4回、5回、6回、7回、8回、9回、10回、11回、12回、13回、14回、15回、16回、17回、18回、19回、20回、21回、22回、23回、24回、25回、26回、27回、28回、29回、30回、31回、32回、33回、34回、35回、36回、37回、38回、39回、40回、41回、42回、43回、44回、45回、46回、47回、48回、49回、50回、60回、70回、80回、90回、100回の、もしくはこれらの値のうちのいずれか2つの間の数もしくは範囲の、ランダムプライミングおよび伸長のサイクルを含み得るか、または約1回、2回、3回、4回、5回、6回、7回、8回、9回、10回、11回、12回、13回、14回、15回、16回、17回、18回、19回、20回、21回、22回、23回、24回、25回、26回、27回、28回、29回、30回、31回、32回、33回、34回、35回、36回、37回、38回、39回、40回、41回、42回、43回、44回、45回、46回、47回、48回、49回、50回、60回、70回、80回、90回、100回の、もしくはこれらの値のうちのいずれか2つの間の数もしくは範囲の、ランダムプライミングおよび伸長のサイクルを含み得る。一部の実施形態では、ランダムプライミングおよび伸長のサイクル数は、少なくとも、または最大で1回、2回、3回、4回、5回、6回、7回、8回、9回、10回、11回、12回、13回、14回、15回、16回、17回、18回、19回、20回、21回、22回、23回、24回、25回、26回、27回、28回、29回、30回、31回、32回、33回、34回、35回、36回、37回、38回、39回、40回、41回、42回、43回、44回、45回、46回、47回、48回、49回、50回、60回、70回、80回、90回、もしくは100回の、ランダムプライミングおよび伸長のサイクルを含み得る。第1の複数のバーコード化されたアンプリコンにハイブリダイズしたランダムプライマーを伸長させるステップは、5'から3'へのエキソヌクレアーゼ活性および3'から5'へのエキソヌクレアーゼ活性（例えば、Klenow断片）のうちの少なくとも1つを欠くDNAポリメラーゼを使用することを含み得る。一部の実施形態では、伸長酵素は、KlenowまたはKlenow exo-である。

【0186】

ランダムプライマーは、ヌクレオチドのランダム配列を含み得る。ヌクレオチドのランダム配列は、約4～約30ヌクレオチドの長さであり得る。一部の実施形態では、ヌクレオチドの前記ランダム配列は、6または9ヌクレオチドの長さである。ヌクレオチドのランダム配列は、異なる実装形態では異なる長さを有し得る。一部の実施形態では、ランダムプライマー内のヌクレオチドのランダム配列は、1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、60、70、80、90、100、もしくはこれらの値のうちのいずれか2つの間の数もしくは範囲、または約1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、60、70、80、90、100、もしくはこれらの値のうちのいずれか2つの間の数もしくは範囲のヌクレオチドの長さである。一部の実施形態では、ランダムプライマー内のヌクレオチドのランダム配列は、少なくとも、または最大で1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、60、70、80、90、100ヌクレオチドの長さである。ランダムプライマーは、異なる実装形態におけるランダムプライミングステップの間に異なる濃度を有し得る。一部の実施形態では、ランダムプライマーは、ランダムプライミングの間、少なくとも、または最大で1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、60、70、80、90、100、110、

120、128、もしくはこれらの値のうちのいずれか2つの間の数もしくは範囲の μM の濃度である。

【0187】

バーコード化された核酸分子の標的結合領域の相補体をバーコード化された核酸分子自体の標的結合領域とハイブリダイズさせるステップは、ステムループを形成するために、バーコード化された核酸分子内の標的結合領域と該標的結合領域の相補体との分子内ハイブリダイゼーションを含み得る。一部の実施形態では、第2の分子標識は、第1の分子標識の相補体である。

【0188】

バーコード化された核酸分子の標的結合領域の相補体を、複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコードの標的結合領域とハイブリダイズさせるステップは、バーコード化された核酸分子の標的結合領域の相補体の、複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコードの標的結合領域との分子間ハイブリダイゼーションを含み得る。一部の実施形態では、第2の分子標識は第1の分子標識と異なり、第2の分子標識は、第1の分子標識の相補体ではない。一部の実施形態では、本方法は、バーコード化された核酸分子の標的結合領域の相補体にハイブリダイズしたオリゴヌクレオチドバーコードの3'末端を伸長させて、それぞれが、第1の分子標識の相補体および第2の分子標識を含む、複数の伸長したバーコード化された核酸分子を生成するステップを含む。一部の実施形態では、第2の分子標識の配列は、第1の分子標識の配列と異なり、第2の分子標識は、第1の分子標識の相補体ではない。

【0189】

バーコード化された核酸分子の標的結合領域の相補体を、複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子の標的結合領域とハイブリダイズさせるステップは、バーコード化された核酸分子の標的結合領域の相補体の、複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子の標的結合領域との分子間ハイブリダイゼーションを含み得る。一部の実施形態では、第2の分子標識の配列は第1の分子標識の配列と異なり、第2の分子標識は、第1の分子標識の相補体ではない。

【0190】

一部の実施形態では、本方法は、標的結合領域の相補体の、バーコード化された核酸分子の末端（例えば、3'末端）への付加（例えば、テンプレートスイッチング反応による）を含む。一部の実施形態では、本方法は、i) 分子内ハイブリダイゼーションおよび/またはii) オリゴヌクレオチドバーコード（またはその産物、例えば、別のバーコード化された核酸分子、またはそのアンプリコンなど）の標的結合領域の分子間ハイブリダイゼーションと、それに続く、伸長したバーコード化された核酸分子を生成するための伸長、を含む。伸長したバーコード化された核酸分子は、3'末端と5'末端の両方でバーコード化され得る。一部の実施形態では、バーコード化された分子の分子内ハイブリダイゼーションによって、3'ポリ(dT)捕捉ビーズ上に捕捉mRNA転写物を含むヘアピンループが形成される。mRNA分子は、オリゴヌクレオチドバーコードの標的結合領域に結合しているポリ(A)テールによってビーズ上に捕捉され得る。ハイブリダイゼーション後に、テンプレートスイッチングを使用して、捕捉された転写物の5'末端にポリ(dA)テールを結合させることができる。次いで、新たなポリ(dA)テールがハイブリダイズし、同じビーズ上の捕捉オリゴヌクレオチド（例えば、確率的バーコードなどのバーコード）を遊離し得る。伸長後に、mRNA分子は、3'末端と5'末端の両方でバーコード化され得る。これにより、例えば、Illuminaのシーケンシングプラットフォームでシーケンシングされ得る3'および5'でバーコード化された転写物の生成が可能となる。バーコード化された5'配列へのアクセスにより、T細胞受容体(TCR)およびB細胞受容体(BCR)の変領域、ならびに転写物の5'末端において生じるスプライスバリエーションおよび配列多様性の検出を可能とすることができる。

【0191】

図6A~6Oは、5'バーコード化および/または3'バーコード化を使用して、核酸標

的（例えば、免疫受容体のV(D)J領域）の完全長配列を決定する非限定的な例示的ワークフローの略図を示す。BD（登録商標）Rhapsody（商標）ビーズは、幅広い物理的および化学的操作によって完全性を維持する、固体のバーコード化されたビーズである。ビーズ上でのmRNAのポリ(A)捕捉後に、逆転写およびテンプレートスイッチングを実施して、ポリ(dA)テールをバーコード化されたcDNAの3'末端に付加することができる。付加されたポリ(dA)テールによって、ビーズに結合したcDNAが同じビーズ上のバーコード（例えば、確率的バーコード）のオリゴ(dT)領域に自己ハイブリダイズし、架橋ループ構造を形成することが可能となる。架橋ループのKlenow伸長によって、反対方向に第1のバーコード化されたcDNAを有する、同じmRNA転写物に由来する新たなバーコード化されたcDNA分子を生成することができ、3'末端と5'末端の両方を分子バーコードに連結させることが可能となる。

10

【0192】

本明細書に開示される方法によって、3'ベースおよび/または5'ベースの完全長配列決定を可能にすることができる。この方法は、配列決定に柔軟性を与えることを可能にすることができる。一部の実施形態では、本方法は、使用されるプライマーは別として、プロトコルまたは産物の立体構造を変化させずに、試料、例えば、マウスおよびヒトの試料について、Rhapsody（商標）システムにおけるT細胞とB細胞の両方の免疫レパートリーのプロファイリングを可能にすることができる。一部の実施形態では、V(D)Jの3'および/または5'ベースの完全長遺伝子発現プロファイリングが実施され得る。一部の実施形態では、表現型マーカーと単一細胞プラットフォームにおけるT細胞およびB細胞のV(D)J配列の両方が調査され得る。一部の実施形態では、それらの転写物の3'と5'の両方の情報が、単回の実験で捕捉され得る。本明細書に開示される方法は、T細胞とB細胞の両方のV(D)J検出を可能にすることができる（例えば、超変異）。

20

【0193】

本明細書に記載の方法およびシステムは、オリゴヌクレオチド（本明細書において、AbOまたはAbOligoとも称される）と会合した（例えば、それに結合した、またはそれとコンジュゲートした）抗体を使用する、方法およびシステムに関して使用することができる。単一細胞におけるタンパク質発現プロファイルを決定するためにAbOを使用し、試料起源を追跡する実施形態については、米国特許出願公開第2018/0088112号として公開された米国特許出願第15/715,028号、および米国特許出願第15/937,713号に記載されており、それぞれの内容は、参照によりその全体が本明細書に組み込まれる。一部の実施形態では、本明細書に開示される方法は、T細胞およびB細胞の完全長V(D)Jプロファイリング、3'標的化、5'標的化、3'全トランスクリプトーム増幅(WTA)、5'WTA、AbOを用いるタンパク質発現プロファイリング、ならびに/または単回実験での試料多重化を可能にする。

30

【0194】

テンプレートスイッチング反応

図6A~6Oは、5'バーコード化および/または3'バーコード化を使用して、核酸標的（例えば、免疫受容体のV(D)J領域）の完全長配列を決定する非限定的な例示的ワークフローの略図を示す。バーコード（例えば、確率的バーコード、オリゴヌクレオチドバーコード602）は、標識化またはバーコード化（例えば、固有の標識化）のために、ポリ(dA)テール608によって核酸標的（例えば、ポリアデニル化RNA転写物606）に、または他の核酸標的に結合し得る標的結合領域（例えば、ポリ(dT)604）を含み得る。標的結合領域は、遺伝子特異的配列、オリゴ(dT)配列、ランダム多量体、またはこれらの任意の組合せを含み得る。一部の実施形態では、バーコードは、固体支持体（例えば、粒子610）に会合している。複数のバーコード602は、粒子610と会合していてもよい。一部の実施形態では、粒子は、ビーズである。ビーズは、ポリマービーズ、例えば、バーコードまたは確率的バーコードで機能化された変形可能なビーズまたはゲルビーズ（例えば、10X Genomics (San Francisco, CA)からのゲルビーズ）であり得る。一部の実施形態では、ゲルビーズは、ポリマー系ゲル

40

50

を含み得る。ゲルビーズは、例えば、1つまたは複数のポリマー前駆体を液滴中に封入することによって、生成することができる。ポリマー前駆体を、促進剤（例えば、テトラメチルエチレンジアミン（TEMED））に曝露すると、ゲルビーズが生成され得る。

【0195】

図6Aは、逆転写反応600aの非限定的な例示の実施形態を示す。逆転写600a中に、オリゴヌクレオチドバーコード602の末端に到達すると、酵素（例えば、モロニーマウス白血病ウイルス（MMLV）などの逆転写酵素）の末端転移酵素の活性によって、新たに合成されたcDNA配列鎖614c（RNA配列614rのアンチセンス配列）の3'末端に、いくつかの追加のヌクレオチド（例えば、デオキシシチジン、CCC612）が付加される。これらのCCC塩基612は、テールを付加した配列に相補的な配列（例えば、rGrGrG618）を含むテンプレートスイッチオリゴヌクレオチド616（例えば、テンプレートスイッチングオリゴヌクレオチド）のアンカー部位として機能し得る。テンプレートスイッチオリゴヌクレオチド616は、標的結合領域604の少なくとも一部を含み得る。rGrGrG618と付加されたデオキシシチジンストレッチ612との間の塩基対合の際に、酵素は、オリゴヌクレオチドバーコード602からテンプレートスイッチオリゴヌクレオチド616にテンプレート鎖を「切り替え」、テンプレートスイッチオリゴヌクレオチド616の5'末端への複製を継続する。よって、得られた第1の鎖で標識されたcDNA（例えば、バーコード化された核酸分子620）は、テンプレートスイッチオリゴヌクレオチド616の逆相補配列を含有し、よって、標的結合領域（例えば、ポリ(dA)608）の相補体（例えば、逆相補体）を含み得る。バーコード化された核酸分子620は、cDNA614c（RNA配列614rの逆相補配列）を含み得る。反応は、二次構造（例えば、エチレングリコール）を低減するよう構成された1つまたは複数の添加剤の存在下で実施することができる。バーコード化された核酸分子620は、いくつかの標識も含み得る。オリゴヌクレオチドバーコード602は、転写物606を標識し、RNA転写物606（または例えば、抗体と会合したかもしくは抗体から解離した、抗体オリゴヌクレオチドなどの核酸標的）の試料起源を追跡するための第1の分子標識（ML1）622および試料標識（例えば、区画標識、細胞標識（CL）624）を、それぞれ、後続反応のための各バーコード602の第1の分子標識622/細胞標識624領域に隣接する、1つまたは複数の追加の配列、例えば、第1のユニバーサル配列626（例えば、リード1配列）などと一緒に、含み得る。試料ごとのオリゴヌクレオチドバーコードにおける分子標識の配列のレパートリーは、RNA転写物の確率的標識化に対して十分に多量であり得る。一部の実施形態では、試料標識は、区画標識である。一部の実施形態では、試料標識は、細胞標識である。バーコード化された核酸分子620は、変性ステップ600b（例えば、変性するステップ）を受け、それによって、一本鎖のバーコード化された核酸分子621を生じ得る。

【0196】

一部の実施形態では、第1の分子標識は、複数のバーコード化された核酸分子の3'末端を伸長させた後に、第2の分子標識にハイブリダイズされる。一部の実施形態では、伸長したバーコード化された核酸分子はそれぞれ、第1の分子標識、第2の分子標識、標的結合領域、および標的結合領域の相補体を含む。一部の実施形態では、標的結合領域の相補体は、標的結合領域の一部分に相補的である。一部の実施形態では、標的結合領域は、遺伝子特異的配列を含む。一部の実施形態では、標的結合領域は、ポリ(dT)配列を含む。

【0197】

用語「テンプレートスイッチング」は、初期核酸配列テンプレートから、初期テンプレートから合成された核酸の3'末端に対してわずかな相補性を有するかまたは相補性を有さない新たな核酸配列テンプレートの3'末端に切り替える、逆転写酵素の能力を指し得る。テンプレートスイッチングの例は、初期核酸配列テンプレート/プライマー基質から、核酸プライマー鎖の3'末端に対してわずかな相補性を有するかまたは相補性を有さない新たな核酸配列テンプレートの3'末端に切り替える、逆転写酵素の能力である。テンプレートスイッチングは、例えば、DNAコピーが、初期核酸配列テンプレートから、初期テンプレ

10

20

30

40

50

レートから合成されたDNAの3'末端に対してわずかな相補性を有するかまたは相補性を有さない新たな核酸配列テンプレートの3'末端に切り替える、逆転写酵素を使用して調製されるのを可能にし、それによって、アダプター配列をライゲーションなしで標的オリゴヌクレオチド配列に直接的に連結させる、連続的産物DNAの合成を可能にする。テンプレートスイッチングは、アダプターのライゲーション、ホモポリマーのテーリング（例えば、ポリアデニル化）、ランダムプライマー、またはポリメラーゼが会合し得るオリゴヌクレオチドを含み得る。上述の実施形態のいずれかでは、テンプレートスイッチングを使用して、標的結合領域またはその相補体を導入してもよい。

【0198】

一部の実施形態では、逆転写酵素は、末端転移酵素活性の能力がある。一部の実施形態では、テンプレートスイッチオリゴヌクレオチドは、1つまたは複数の3'リボヌクレオチドを含む。一部の実施形態では、テンプレートスイッチオリゴヌクレオチドは、3つの3'リボヌクレオチドを含む。一部の実施形態では、3'リボヌクレオチドは、グアニンを含む。一部の実施形態では、逆転写酵素は、ウイルス逆転写酵素を含む。一部の実施形態では、ウイルス逆転写酵素は、マウス白血病ウイルス（MLV）逆転写酵素である。一部の実施形態では、ウイルス逆転写酵素は、モロニー Maus 白血病ウイルス（MMLV）逆転写酵素である。

標的結合領域の相補体は、該標的結合領域の逆相補配列を含み得るか、または該標的結合領域の相補配列を含み得る。分子標識の相補体は、該分子標識の逆相補配列を含み得るか、または該分子標識の相補配列を含み得る。一部の実施形態では、複数のバーコード化された核酸分子は、バーコード化されたデオキシリボ核酸（DNA）分子および/またはバーコード化されたリボ核酸（RNA）分子を含み得る。一部の実施形態では、核酸標的は、核酸分子（例えば、リボ核酸（RNA）、メッセンジャーRNA（mRNA）、マイクロRNA、低分子干渉RNA（siRNA）、RNA分解産物、ポリ（A）テールを含むRNA、またはこれらの任意の組合せ）を含む。一部の実施形態では、mRNAは、免疫受容体をコードする。核酸標的は、細胞成分結合試薬を含み得る。一部の実施形態では、核酸分子は、細胞成分結合試薬に会合している。本方法は、核酸分子と細胞成分結合試薬を解離させるステップを含み得る。一部の実施形態では、複数のオリゴヌクレオチドバーコードのうちの少なくとも10個は、異なる分子標識配列を含む。複数のオリゴヌクレオチドバーコードの各分子標識は、少なくとも6個のヌクレオチドを含み得る。

【0199】

一部の実施形態では、複数のオリゴヌクレオチドバーコードは、固体支持体に会合している。同じ固体支持体に会合した複数のオリゴヌクレオチドバーコードはそれぞれ、同一の試料標識を含み得る。複数のオリゴヌクレオチドバーコードの各試料標識は、少なくとも6個のヌクレオチドを含み得る。複数のオリゴヌクレオチドバーコードはそれぞれ、細胞標識を含み得る。複数のオリゴヌクレオチドバーコードの各細胞標識は、少なくとも6個のヌクレオチドを含み得る。同じ固体支持体に会合したオリゴヌクレオチドバーコードは、同じ細胞標識を含み得る。異なる固体支持体に会合したオリゴヌクレオチドバーコードは、異なる細胞標識を含み得る。複数の伸長したバーコード化された核酸分子はそれぞれ、細胞標識および該細胞標識の相補体を含み得る。細胞標識の相補体は、細胞標識の逆相補配列、または細胞標識の相補配列を含み得る。本方法は、エチレングリコール、ポリエチレングリコール、1,2-プロパンジオール、ジメチルスルホキシド（DMSO）、グリセロール、ホルムアミド、7-デアザ-GTP、アセトアミド、塩化テトラメチルアンモニウム塩、ベタイン、またはこれらの任意の組合せのうちの1つまたは複数の存在下で、核酸標的のコピーにハイブリダイズした複数のオリゴヌクレオチドバーコードを伸長させるステップを含み得る。一部の実施形態では、固体支持体は、合成粒子を含み得る。一部の実施形態では、固体支持体は、平面表面を含み得る。

【0200】

試料は、単一細胞を含んでもよく、本方法は、複数のオリゴヌクレオチドバーコードを含む合成粒子を、試料中の単一細胞と会合させるステップを含み得る。本方法は、合成粒

子を単一細胞と会合させた後に、単一細胞を溶解させるステップを含み得る。単一細胞を溶解させるステップは、試料を加熱するステップ、試料を界面活性剤と接触させるステップ、試料のpHを変更するステップ、またはこれらの任意の組合せを含み得る。一部の実施形態では、合成粒子と単一細胞は、同じウェル中にある。一部の実施形態では、合成粒子と単一細胞は、同じ液滴中にある。一部の実施形態では、複数のオリゴヌクレオチドバーコードのうちの少なくとも1つは、合成粒子上に固定化されている。一部の実施形態では、複数のオリゴヌクレオチドバーコードのうちの少なくとも1つは、合成粒子上に部分的に固定化されている。複数のオリゴヌクレオチドバーコードのうちの少なくとも1つは、合成粒子内に封入されていてもよい。一部の実施形態では、複数のオリゴヌクレオチドバーコードのうちの少なくとも1つは、合成粒子内に部分的に封入されている。一部の実施形態では、合成粒子は、崩壊可能である。合成粒子は、ビーズを含み得る。ビーズは、セファロースビーズ、ストレプトアビジンビーズ、アガロースビーズ、磁性ビーズ、コンジュゲートビーズ、プロテインAコンジュゲートビーズ、プロテインGコンジュゲートビーズ、プロテインA/Gコンジュゲートビーズ、プロテインLコンジュゲートビーズ、オリゴ(dT)コンジュゲートビーズ、シリカビーズ、シリカ様ビーズ、抗ビオチンマイクロビーズ、抗蛍光色素マイクロビーズ、またはこれらの任意の組合せを含み得る。合成粒子は、ポリジメチルシロキサン(PDMS)、ポリスチレン、ガラス、ポリプロピレン、アガロース、ゼラチン、ハイドロゲル、常磁性体、セラミック、プラスチック、ガラス、メチルスチレン、アクリルポリマー、チタン、ラテックス、セファロース、セルロース、ナイロン、シリコン、およびこれらの任意の組合せからなる群から選択される材料を含み得る。一部の実施形態では、合成粒子は、崩壊可能ハイドロゲル粒子を含み得る。複数のオリゴヌクレオチドバーコードのそれぞれは、リンカー官能基を含んでもよく、合成粒子は、固体支持体官能基を含んでもよく、ならびに/または支持体官能基およびリンカー官能基は互いに会合していてもよい。一部の実施形態では、リンカー官能基および支持体官能基は、独立して、C6、ビオチン、ストレプトアビジン、第一級アミン、アルデヒド、ケトン、およびこれらの任意の組合せからなる群から選択される。

10

20

【0201】

バーコード化された核酸分子の分子内ハイブリダイゼーション

一部の実施形態では、バーコード化された核酸分子の標的結合領域の相補体をバーコード化された核酸分子自体の標的結合領域とハイブリダイズさせるステップは、ステムループを形成する、バーコード化された核酸分子内の標的結合領域と該標的結合領域の相補体との分子内ハイブリダイゼーションを含む。一部の実施形態では、第2の分子標識は、第1の分子標識の相補体である。

30

【0202】

ワークフローは、図6Bの非限定的な例示的略図に示されるように、一本鎖のバーコード化された核酸分子621の分子内ハイブリダイゼーションを含み得る。ワークフローは、ステムループを形成するための、一本鎖のバーコード化された核酸分子621内の標的結合領域604および標的結合領域608の相補体の分子内ハイブリダイゼーション600c1を含み得る。ワークフローは、一本鎖のバーコード化された核酸分子621のステムループの3'末端を伸長させて(600c2)、伸長したバーコード化された核酸分子620cを生成することを含み得る。伸長したバーコード化された核酸分子620cは、第1の分子標識622rcの相補体(例えば、逆相補体)、細胞標識624rcの相補体(例えば、逆相補体)、および/または第1のユニバーサル配列626rcの相補体(例えば、逆相補体)を含み得る。ワークフローは、伸長したバーコード化された核酸分子620cを変性させて(600c3)、一本鎖の伸長したバーコード化された核酸分子620cdを生成することを含み得る。一部の実施形態では、分子間ハイブリダイゼーション600c1および/または伸長600c2は、高塩濃度緩衝液および/またはPEGの存在下で実施される。一部の実施形態では、伸長は、5'から3'へのエキソヌクレアーゼ活性および3'から5'へのエキソヌクレアーゼ活性(例えば、Klenow断片)のうちの少なくとも1つを欠くDNAポリメラーゼを使用して実施される。

40

50

【0203】

一本鎖の伸長したバーコード化された核酸分子620cdは、標的核酸分子（例えば、転写物）の5'末端と3'末端の両方にバーコード（例えば、細胞標識および分子標識）を含んでもよく、それによって、配列同定、転写物の計数、選択的スプライシング分析、変異スクリーニング、および/または全長シーケンシングに関して、一方の末端に1つのバーコードのみを有する標的核酸分子の分析と比較して、標的核酸分子のより広範な分析が可能となる。一本鎖の伸長したバーコード化された核酸分子620cdは、例えば、図6C~6Eに示される非限定的な例示的スキームなどの1つまたは複数の伸長反応（例えば、ランダムプライミングおよび伸長）および/または増幅反応（例えば、PCR）に関するテンプレートとしての役割を果たし得る。増幅は、標的的特異的（例えば、遺伝子特異的）cDNA増幅を含み得る。例えば、一本鎖の伸長したバーコード化された核酸分子620cdは、第1のユニバーサル配列（またはその相補体）の配列を含むユニバーサルオリゴヌクレオチドプライマー646および標的的特異的プライマー（例えば、標的的特異的プライマー650）を用いる1回目の増幅（「PCR1」）600c4を受けて、それによって、それぞれが、分子標識（例えば、第1の分子標識）、細胞標識、第1のユニバーサル配列および部分的アンチセンスcDNA614cas1（その長さは、cDNA614c内の標的的特異的プライマー650の結合部位に応じて変わる）を含む第1の複数のアンプリコン620cas1を生成することができる。PCR1 600c4は、1~30サイクル（例えば、15サイクル）を含み得る。

10

【0204】

ワークフローは、ランダムプライミングおよび伸長600c5を含み得る。ランダムプライマー670は、第1の複数のバーコード化されたアンプリコン620cas1とハイブリダイズされ得、ランダムプライマー670は、伸長されて、複数の伸長産物620e2c1および620e2c2を生成し得る。ランダムプライマー670は、例えば、第2のユニバーサル配列638（またはその相補体、例えば、逆相補体638rc）を含み得るか、またはそれであり得るオーバーハングを含み得る（例えば、ユニバーサルPCRハンドルである、リード2配列）。伸長産物620e2c1は、第1のユニバーサル配列、第2のユニバーサル配列、細胞標識、および分子標識（例えば、第1の分子標識）、またはこれらの相補体を含み得る。伸長産物620e2c1は、例えば、部分的cDNA614c2a、614c2b、614c2c、および614c2d（そのそれぞれの長さは、部分的cDNA内のランダムプライマー670の結合部位に応じて変わる）を含み得る。伸長産物620e2c2は、第2のユニバーサル配列（またはその相補体）を含み得る。伸長産物620e2c2は、例えば、部分的アンチセンスcDNA614cas2e、614cas2f、614cas2g、および614cas2h（そのそれぞれの長さは、部分的cDNA内のランダムプライマー670の結合部位に応じて変わる）を含み得る。

20

30

【0205】

ワークフローは、ライブラリー増幅（「ライブラリーPCR」）600c6を含み得る。ライブラリーPCR600c6は、シーケンシングライブラリー増幅プライマー656および658を用いる伸長産物620e2c1のライブラリー増幅を含み得る。シーケンシングライブラリー増幅プライマー656および658は、それぞれが、第1のユニバーサル配列626および第2のユニバーサル配列638（またはそれらの相補体）にアニリングし得る。ライブラリーPCR600c6は、シーケンシングライブラリー増幅プライマー656および658におけるオーバーハングによって、シーケンシングアダプター（例えば、P5 640およびP7 642）および試料インデックス644（例えば、i5、i7）を付加し得る。ライブラリーPCRアンプリコン620cL（例えば、第2の複数のバーコード化されたアンプリコン）は、核酸標的のmRNA配列の一部またはすべてを合わせて包含し得る様々なcDNAの長さ（例えば、部分的アンチセンスcDNA614cas2a、614cas2b、614cas2c、614cas2d）を含み得る。ライブラリーアンプリコン620cLをシーケンシングし、本開示の下流方法に供することができる。150bp x 2シーケンシングを使用するシーケンシング600c7によ

40

50

って、細胞標識、固有の分子標識および/またはリード1上の遺伝子(または遺伝子の部分的配列)、リード2上の遺伝子(または遺伝子の部分的配列)、ならびにインデックス1リードおよび/またはインデックス2リード上の試料インデックスが明らかになり得る。ライブラリーPCR 600c6は、1~30サイクル(例えば、15サイクル)を含み得る。本方法は、図9Bに示されるように、複数のリード1 910およびリード2 920のリードをアライメントすることによって、核酸標的(例えば、免疫受容体mRNA 930)の完全長配列の生物情報学的再構築900を含み得る。

【0206】

一部の実施形態では、免疫受容体のV(D)J領域の3'および/または5'ベースの完全長発現プロファイリングが実施され得る。一部の実施形態では、表現型マーカーと単一細胞プラットフォームにおけるT細胞および/またはB細胞の免疫受容体V(D)J配列の両方が調査され得る。本明細書に開示される方法は、T細胞とB細胞の両方のV(D)J検出を可能にすることができる(例えば、超変異)。一部の実施形態では、伸長したバーコード化された核酸分子620cdの3'領域と5'領域の両方が増幅される。一部の実施形態では、伸長したバーコード化された核酸分子620cdの5'領域のみが増幅される。一部の実施形態では、伸長したバーコード化された核酸分子620cdの3'領域のみが増幅される。一部の実施形態では、増幅反応の1つまたは複数が多重PCRを含む。例えば、伸長したバーコード化された核酸分子620cdの3'領域と5'領域の両方が同時に増幅され得る(例えば、多重PCR)。一部の実施形態では、ワークフローは、標的的特異的PCRプライマーのパネルを用いる多重PCRを含む。一部の実施形態では、標的は、BCR、TCR、および/または免疫関連転写物を含む。

【0207】

バーコード化された核酸分子のバーコード化された核酸分子との分子間ハイブリダイゼーション

一部の実施形態では、バーコード化された核酸分子の標的結合領域の相補体を、複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子の標的結合領域とハイブリダイズさせるステップは、バーコード化された核酸分子の標的結合領域の相補体の、複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子の標的結合領域との分子間ハイブリダイゼーションを含む。一部の実施形態では、第2の分子標識の配列は第1の分子標識の配列と異なり、第2の分子標識は、第1の分子標識の相補体ではない。

【0208】

ワークフローは、図6F~6Gの非限定的な例示的略図に示されるように、一本鎖のバーコード化された核酸分子621の、別個のバーコード化された核酸分子628との分子間ハイブリダイゼーションを含み得る。別個のバーコード化された核酸分子628は、cDNA 630c、第2の分子標識632、細胞標識624、および第1のユニバーサル配列626を含み得る。バーコード化された核酸分子628の第2の分子標識632の配列は、一本鎖のバーコード化された核酸分子621(例えば、相補体ではない)の第1の分子標識622の配列と異なってもよい。バーコード化された核酸分子628の標的結合領域604、細胞標識624および/または第1のユニバーサル配列626は、一本鎖のバーコード化された核酸分子621の標的結合領域604、細胞標識624および/または第1のユニバーサル配列626と同じ(またはその相補体)であってもよい。ワークフローは、一部の実施形態では、一本鎖のバーコード化された核酸分子621の標的結合領域608の相補体の、バーコード化された核酸分子628の標的結合領域604との分子間ハイブリダイゼーション600d1を含み得る。ワークフローは、一本鎖のバーコード化された核酸分子621の3'末端を伸長させて(600d2)、伸長したバーコード化された核酸分子620dを生成することを含み得る。伸長したバーコード化された核酸分子620dは、第2の分子標識の相補体(例えば、逆相補体)632rc、細胞標識の相補体(例えば、逆相補体)624rc、および/または第1のユニバーサル配列の相補体(例えば、逆相補体)626rcを含み得る。ワークフローは、伸長したバーコード化さ

れた核酸分子 6 2 0 d を変性させて (6 0 0 d 3)、一本鎖の伸長したバーコード化された核酸分子 6 2 0 d d を生成することを含み得る。一部の実施形態では、分子間ハイブリダイゼーション 6 0 0 d 1 および / または伸長 6 0 0 d 2 は、高塩濃度緩衝液および / または P E G の存在下で実施される。一部の実施形態では、伸長は、5 ' から 3 ' へのエキソヌクレアーゼ活性および 3 ' から 5 ' へのエキソヌクレアーゼ活性 (例えば、K l e n o w 断片) のうちの少なくとも 1 つを欠く D N A ポリメラーゼを使用して実施される。

【 0 2 0 9 】

一本鎖の伸長したバーコード化された核酸分子 6 2 0 d d は、標的核酸分子 (例えば、転写物) の 5 ' 末端と 3 ' 末端の両方にバーコード (例えば、細胞標識および分子標識) を含み、それによって、配列同定、転写物の計数、選択的スプライシング分析、変異スクリーニング、および / または全長シーケンシングに関して、一方の末端に 1 つのバーコードのみを有する標的核酸分子の分析と比較して、標的核酸分子のより広範な分析を可能とすることができる。一本鎖の伸長したバーコード化された核酸分子 6 2 0 d d は、例えば、図 6 H ~ 6 J に示される非限定的な例示的スキームなどの 1 つまたは複数の伸長反応 (例えば、ランダムプライミングおよび伸長) および / または増幅反応 (例えば、P C R) に関するテンプレートとしての役割を果たし得る。増幅は、標的的特異的 (例えば、遺伝子特異的) c D N A 増幅を含み得る。例えば、一本鎖の伸長したバーコード化された核酸分子 6 2 0 d d は、第 1 のユニバーサル配列 (またはその相補体) の配列を含むユニバーサルオリゴヌクレオチドプライマー 6 4 6 および標的的特異的プライマー (例えば、標的的特異的プライマー 6 5 0) を用いる 1 回目の増幅 (「P C R 1」) 6 0 0 d 4 を受け、それによって、それぞれが、分子標識 (例えば、第 2 の分子標識)、細胞標識、第 1 のユニバーサル配列および部分的アンチセンス c D N A 6 1 4 c a s 1 (その長さは、c D N A 6 1 4 c 内の標的的特異的プライマー 6 5 0 の結合部位に応じて変わる) を含む第 1 の複数のアンプリコン 6 2 0 d a s 1 を生成することができる。P C R 1 6 0 0 d 4 は、1 ~ 3 0 サイクル (例えば、1 5 サイクル) を含み得る。

【 0 2 1 0 】

ワークフローは、ランダムプライミングおよび伸長 6 0 0 d 5 を含み得る。ランダムプライマー 6 7 0 は、第 1 の複数のバーコード化されたアンプリコン 6 2 0 d a s 1 とハイブリダイズされ得、ランダムプライマー 6 7 0 は、伸長されて、複数の伸長産物 6 2 0 e 2 d 1 および 6 2 0 e 2 d 2 を生成し得る。ランダムプライマー 6 7 0 は、例えば、第 2 のユニバーサル配列 6 3 8 (またはその相補体、例えば、逆相補体 6 3 8 r c) を含み得るか、またはそれであり得るオーバーハングを含み得る (例えば、ユニバーサル P C R ハンドルである、リード 2 配列)。伸長産物 6 2 0 e 2 d 1 は、第 1 のユニバーサル配列、第 2 のユニバーサル配列、細胞標識、および分子標識 (例えば、第 1 の分子標識)、またはこれらの相補体を含み得る。伸長産物 6 2 0 e 2 d 1 は、例えば、部分的 c D N A 6 1 4 c 2 a、6 1 4 c 2 b、6 1 4 c 2 c、および 6 1 4 c 2 d (そのそれぞれの長さは、部分的 c D N A 内のランダムプライマー 6 7 0 の結合部位に応じて変わる) を含み得る。伸長産物 6 2 0 e 2 d 2 は、第 2 のユニバーサル配列 (またはその相補体) を含み得る。伸長産物 6 2 0 e 2 d 2 は、例えば、部分的アンチセンス c D N A 6 1 4 c a s 2 e、6 1 4 c a s 2 f、6 1 4 c a s 2 g、および 6 1 4 c a s 2 h (そのそれぞれの長さは、部分的 c D N A 内のランダムプライマー 6 7 0 の結合部位に応じて変わる) を含み得る。

【 0 2 1 1 】

ワークフローは、ライブラリー増殖 (「ライブラリー P C R」) 6 0 0 d 6 を含み得る。ライブラリー P C R 6 0 0 d 6 は、シーケンシングライブラリー増幅プライマー 6 5 6 および 6 5 8 を用いる伸長産物 6 2 0 e 2 d 1 のライブラリー増幅を含み得る。シーケンシングライブラリー増幅プライマー 6 5 6 および 6 5 8 は、それぞれが、第 1 のユニバーサル配列 6 2 6 および第 2 のユニバーサル配列 6 3 8 (またはそれらの相補体) にアニーリングし得る。ライブラリー P C R 6 0 0 d 6 は、シーケンシングライブラリー増幅プライマー 6 5 6 および 6 5 8 におけるオーバーハングによって、シーケンシングアダプター (例えば、P 5 6 4 0 および P 7 6 4 2) および試料インデックス 6 4 4 (例えば

10

20

30

40

50

、i5、i7)を付加し得る。ライブラリーPCRアンプリコン620dL(例えば、第2の複数のバーコード化されたアンプリコン)は、核酸標的のmRNA配列の一部またはすべてを合わせて包含し得る様々なcDNAの長さ(例えば、部分的アンチセンスcDNA614cas2a、614cas2b、614cas2c、614cas2d)を含み得る。ライブラリーアンプリコン620dLをシーケンシングし、本開示の下流方法に供することができる。150bp×2シーケンシングを使用するシーケンシング600d7によって、細胞標識、固有の分子標識および/またはリード1上の遺伝子(または遺伝子の部分的配列)、リード2上の遺伝子(または遺伝子の部分的配列)、ならびにインデックス1リードおよび/またはインデックス2リード上の試料インデックスが明らかになり得る。ライブラリーPCR600d6は、1~30サイクル(例えば、15サイクル)を含み得る。本方法は、図9Bに示されるように、複数のリード1910およびリード2920のリードをアライメントすることによって、核酸標的(例えば、免疫受容体mRNA930)の完全長配列の生物情報学的再構築900を含み得る。

10

【0212】

一部の実施形態では、免疫受容体のV(D)J領域の3'および/または5'ベースの完全長発現プロファイリングが実施され得る。一部の実施形態では、表現型マーカーと単一細胞プラットフォームにおけるT細胞および/またはB細胞の免疫受容体V(D)J配列の両方が調査され得る。本明細書に開示される方法は、T細胞とB細胞の両方のV(D)J検出を可能にすることができる(例えば、超変異)。一部の実施形態では、伸長したバーコード化された核酸分子620ddの3'領域と5'領域の両方が増幅される。一部の実施形態では、伸長したバーコード化された核酸分子620ddの5'領域のみが増幅される。一部の実施形態では、伸長したバーコード化された核酸分子620ddの3'領域のみが増幅される。一部の実施形態では、増幅反応の1つまたは複数が多重PCRを含む。例えば、伸長したバーコード化された核酸分子620ddの3'領域と5'領域の両方が同時に増幅され得る(例えば、多重PCR)。一部の実施形態では、ワークフローは、標的特定のPCRプライマーのパネルを用いる多重PCRを含む。一部の実施形態では、標的は、BCR、TCR、および/または免疫関連転写物を含む。

20

【0213】

バーコード化された核酸分子のオリゴヌクレオチドバーコードとの分子間ハイブリダイゼーション

30

一部の実施形態では、バーコード化された核酸分子の標的結合領域の相補体を、複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコードの標的結合領域とハイブリダイズさせるステップは、バーコード化された核酸分子の標的結合領域の相補体の、複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコードの標的結合領域との分子間ハイブリダイゼーションを含む。一部の実施形態では、第2の分子標識は第1の分子標識と異なり、第2の分子標識は、第1の分子標識の相補体ではない。一部の実施形態では、本方法は、バーコード化された核酸分子の標的結合領域の相補体にハイブリダイズしたオリゴヌクレオチドバーコードの3'末端を伸長させて、それぞれが、第1の分子標識の相補体および第2の分子標識を含む、複数の伸長したバーコード化された核酸分子を生成するステップを含む。一部の実施形態では、第2の分子標識の配列は、第1の分子標識の配列と異なり、第2の分子標識は、第1の分子標識の相補体ではない。

40

【0214】

ワークフローは、図6K~6Lの非限定的な例示的略図に示されるように、一本鎖のバーコード化された核酸分子621の、別個のオリゴヌクレオチドバーコード634との分子間ハイブリダイゼーションを含み得る。別個のオリゴヌクレオチドバーコード634は、第2の分子標識636、細胞標識624、および第1のユニバーサル配列626を含み得る。オリゴヌクレオチドバーコード634の第2の分子標識636の配列は、一本鎖のバーコード化された核酸分子621(例えば、相補体ではない)の第1の分子標識622の配列と異なってもよい。オリゴヌクレオチドバーコード634の標的結合領域604、細胞標識624および/または第1のユニバーサル配列626は、一本鎖のバーコー

50

ド化された核酸分子 6 2 1 の標的結合領域 6 0 4、細胞標識 6 2 4 および / または第 1 のユニバーサル配列 6 2 6 と同じ (またはその相補体) であってもよい。ワークフローは、一部の実施形態では、一本鎖のバーコード化された核酸分子 6 2 1 の標的結合領域 6 0 8 の相補体の、オリゴヌクレオチドバーコード 6 3 4 の標的結合領域 6 0 4 との分子間ハイブリダイゼーション 6 0 0 e 1 を含み得る。ワークフローは、一本鎖のバーコード化された核酸分子 6 2 1 の 3 ' 末端を伸長させて (6 0 0 e 2)、伸長したバーコード化された核酸分子 6 2 0 e 1 を生成することを含み得る。伸長したバーコード化された核酸分子 6 2 0 e 1 は、第 2 の分子標識の相補体 (例えば、逆相補体) 6 3 6 r c、細胞標識の相補体 (例えば、逆相補体) 6 2 4 r c、第 1 のユニバーサル配列の相補体 (例えば、逆相補体) 6 2 6 r c、および / または c D N A 6 1 4 c を含み得る。ワークフローは、伸長したバーコード化された核酸分子 6 2 0 e 1 を変性させて (6 0 0 e 3)、一本鎖の伸長したバーコード化された核酸分子 6 2 0 e 1 d を生成することを含み得る。ワークフローは、オリゴヌクレオチドバーコード 6 3 4 の 3 ' 末端を伸長させて (6 0 0 e 2)、伸長したバーコード化された核酸分子 6 2 0 e 2 を生成することを含み得る。伸長したバーコード化された核酸分子 6 2 0 e 2 は、第 1 の分子標識の相補体 (例えば、逆相補体) 6 2 2 r c、細胞標識の相補体 (例えば、逆相補体) 6 2 4 r c、第 1 のユニバーサル配列の相補体 (例えば、逆相補体) 6 2 6 r c、および / またはアンチセンス c D N A 6 1 4 c a s を含み得る。ワークフローは、伸長したバーコード化された核酸分子 6 2 0 e 2 を変性させて (6 0 0 e 3)、一本鎖の伸長したバーコード化された核酸分子 6 2 0 e 2 d を生成することを含み得る。一部の実施形態では、分子間ハイブリダイゼーション 6 0 0 e 1 および / または伸長 6 0 0 e 2 は、高塩濃度緩衝液および / または P E G の存在下で実施される。一部の実施形態では、伸長は、5 ' から 3 ' へのエキソヌクレアーゼ活性および 3 ' から 5 ' へのエキソヌクレアーゼ活性 (例えば、K l e n o w 断片) のうちの少なくとも 1 つを欠く D N A ポリメラーゼを使用して実施される。

【 0 2 1 5 】

一本鎖の伸長したバーコード化された核酸分子 6 2 0 e 1 d および一本鎖の伸長したバーコード化された核酸分子 6 2 0 e 2 d は、標的核酸分子 (例えば、転写物) の 5 ' 末端と 3 ' 末端の両方にバーコード (例えば、細胞標識および分子標識) を含み、それによって、配列同定、転写物の計数、選択的スプライシング分析、変異スクリーニング、および / または全長シーケンシングに関して、一方の末端に 1 つのバーコードのみを有する標的核酸分子の分析と比較して、標的核酸分子のより広範な分析を可能にすることができる。一本鎖の伸長したバーコード化された核酸分子 6 2 0 e 1 d および一本鎖の伸長したバーコード化された核酸分子 6 2 0 e 2 d は、例えば、図 6 M ~ 6 O に示される非限定的な例示的スキームなどの 1 つまたは複数の伸長反応 (例えば、ランダムプライミングおよび伸長) および / または増幅反応 (例えば、P C R) に関するテンプレートとしての役割を果たし得る。増幅は、標的特定の (例えば、遺伝子特定の) c D N A 増幅を含み得る。例えば、一本鎖の伸長したバーコード化された核酸分子 6 2 0 e 2 d は、第 1 のユニバーサル配列 (またはその相補体) の配列を含むユニバーサルオリゴヌクレオチドプライマー 6 4 6 および標的特定のプライマー (例えば、標的特定のプライマー 6 5 0) を用いる 1 回目の増幅 (「 P C R 1 」) 6 0 0 e 4 を受け、それによって、それぞれが、分子標識 (例えば、第 2 の分子標識)、細胞標識、第 1 のユニバーサル配列および部分的アンチセンス c D N A 6 1 4 c a s 1 (その長さは、c D N A 6 1 4 c a s 内の標的特定のプライマー 6 5 0 の結合部位に応じて変わる) を含む第 1 の複数のアンプリコン 6 2 0 e a s 1 を生成することができる。P C R 1 6 0 0 e 4 は、1 ~ 3 0 サイクル (例えば、1 5 サイクル) を含み得る。

【 0 2 1 6 】

ワークフローは、ランダムプライミングおよび伸長 6 0 0 e 5 を含み得る。ランダムプライマー 6 7 0 は、第 1 の複数のバーコード化されたアンプリコンとハイブリダイズされ得、ランダムプライマー 6 7 0 は、伸長されて、複数の伸長産物 6 2 0 e 2 e 1 および 6 2 0 e 2 e 2 を生成し得る。ランダムプライマー 6 7 0 は、例えば、第 2 のユニバーサル

10

20

30

40

50

配列 6 3 8 (またはその相補体、例えば、逆相補体 6 3 8 r c) を含み得るか、またはそれであり得るオーバーハングを含み得る (例えば、ユニバーサル P C R ハンドルである、リード 2 配列)。伸長産物 6 2 0 e 2 1 は、第 1 のユニバーサル配列、第 2 のユニバーサル配列、細胞標識、および分子標識 (例えば、第 1 の分子標識)、またはこれらの相補体を含み得る。伸長産物 6 2 0 e 2 e 1 は、例えば、部分的 c D N A 6 1 4 c 2 a、6 1 4 c 2 b、6 1 4 c 2 c、および 6 1 4 c 2 d (そのそれぞれの長さは、部分的 c D N A 内のランダムプライマー 6 7 0 の結合部位に応じて変わる) を含み得る。伸長産物 6 2 0 e 2 e 2 は、第 2 のユニバーサル配列 (またはその相補体) を含み得る。伸長産物 6 2 0 e 2 e 2 は、例えば、部分的アンチセンス c D N A 6 1 4 c a s 2 e、6 1 4 c a s 2 f、6 1 4 c a s 2 g、および 6 1 4 c a s 2 h (そのそれぞれの長さは、部分的 c D N A 内のランダムプライマー 6 7 0 の結合部位に応じて変わる) を含み得る。

10

【 0 2 1 7 】

ワークフローは、ライブラリー増殖 (「ライブラリー P C R」) 6 0 0 e 6 を含み得る。ライブラリー P C R 6 0 0 e 6 は、シーケンシングライブラリー増幅プライマー 6 5 6 および 6 5 8 を用いる伸長産物 6 2 0 e 2 e 1 のライブラリー増幅を含み得る。シーケンシングライブラリー増幅プライマー 6 5 6 および 6 5 8 は、それぞれ、第 1 のユニバーサル配列 6 2 6 および第 2 のユニバーサル配列 6 3 8 (またはそれらの相補体) にアニーリングし得る。ライブラリー P C R 6 0 0 e 6 は、シーケンシングライブラリー増幅プライマー 6 5 6 および 6 5 8 におけるオーバーハングによって、シーケンシングアダプター (例えば、P 5 6 4 0 および P 7 6 4 2) および試料インデックス 6 4 4 (例えば、i 5、i 7) を付加し得る。ライブラリー P C R アンプリコン 6 2 0 e 2 L (例えば、第 2 の複数のバーコード化されたアンプリコン) は、核酸標的の m R N A 配列の一部またはすべてを合わせて包含し得る様々な c D N A の長さ (例えば、部分的アンチセンス c D N A 6 1 4 c a s 2 a、6 1 4 c a s 2 b、6 1 4 c a s 2 c、6 1 4 c a s 2 d) を含み得る。ライブラリーアンプリコン 6 2 0 e 2 L をシーケンシングし、本開示の下流方法に供することができる。1 5 0 b p × 2 シーケンシングを使用するシーケンシング 6 0 0 e 7 によって、細胞標識、固有の分子標識および / またはリード 1 上の遺伝子 (または遺伝子の部分的配列)、リード 2 上の遺伝子 (または遺伝子の部分的配列)、ならびにインデックス 1 リードおよび / またはインデックス 2 リード上の試料インデックスが明らかになり得る。ライブラリー P C R 6 0 0 e 6 は、1 ~ 3 0 サイクル (例えば、1 5 サイクル) を含み得る。本方法は、図 9 B に示されるように、複数のリード 1 9 1 0 およびリード 2 9 2 0 のリードをアライメントすることによって、核酸標的 (例えば、免疫受容体 m R N A 9 3 0) の完全長配列の生物情報学的再構築 9 0 0 を含み得る。

20

30

【 0 2 1 8 】

一部の実施形態では、免疫受容体の V (D) J 領域の 3 ' および / または 5 ' ベースの完全長発現プロファイリングが実施され得る。一部の実施形態では、表現型マーカーと単一細胞プラットフォームにおける T 細胞および / または B 細胞の免疫受容体 V (D) J 配列の両方が調査され得る。本明細書に開示される方法は、T 細胞と B 細胞の両方の V (D) J 検出を可能にすることができる (例えば、超変異)。一部の実施形態では、伸長したバーコード化された核酸分子 6 2 0 e 1 d および / または 6 2 0 e 2 d の 3 ' 領域と 5 ' 領域の両方が増幅される。一部の実施形態では、伸長したバーコード化された核酸分子 6 2 0 e 1 d および / または 6 2 0 e 2 d の 5 ' 領域のみが増幅される。一部の実施形態では、伸長したバーコード化された核酸分子 6 2 0 e 1 d および / または 6 2 0 e 2 d の 3 ' 領域のみが増幅される。一部の実施形態では、増幅反応の 1 つまたは複数が多重 P C R を含む。例えば、伸長したバーコード化された核酸分子 6 2 0 e 1 d および / または 6 2 0 e 2 d の 3 ' 領域と 5 ' 領域の両方が同時に増幅され得る (例えば、多重 P C R)。一部の実施形態では、ワークフローは、標的的特異的 P C R 1 プライマーのパネルを用いる多重 P C R を含む。一部の実施形態では、標的は、B C R、T C R、および / または免疫関連転写物を含む。

40

【 0 2 1 9 】

50

免疫レパートリープロファイリング

一部の実施形態では、免疫受容体のV(D)J領域の完全長発現プロファイリングのための方法が提供される。一部の実施形態では、試料は、単一細胞を含む。一部の実施形態では、試料は、複数の細胞、複数の単一細胞、組織、腫瘍試料、またはこれらの任意の組合せを含む。単一細胞は、免疫細胞を含み得る。一部の実施形態では、免疫細胞は、B細胞またはT細胞である。一部の実施形態では、単一細胞は、循環腫瘍細胞を含み得る。一部の実施形態では、各オリゴヌクレオチドバーコードは、第1のユニバーサル配列を含み得る。一部の実施形態では、複数の伸長したバーコード化された核酸分子は、第1のユニバーサル配列および該第1のユニバーサル配列の相補体を含む。複数の伸長したバーコード化された核酸分子を増幅させステップは、核酸標的の配列にハイブリダイズすることが可能な標的特異的プライマーおよび第1のユニバーサル配列を含むプライマーを使用することを含み得る。一部のこのような実施形態では、標的特異的プライマーは、免疫受容体に特異的にハイブリダイズする。例えば、標的特異的プライマーは、免疫受容体の定常領域、免疫受容体の可変領域、免疫受容体の多様性領域、免疫受容体の可変領域と多様性領域の接合部、またはこれらの任意の組合せに特異的にハイブリダイズし得る。免疫受容体は、T細胞受容体(TCR)および/またはB細胞受容体(BCR)受容体であり得る。TCRは、TCRアルファ鎖、TCRベータ鎖、TCRガンマ鎖、TCRデルタ鎖、またはこれらの任意の組合せを含み得る。BCRは、BCR重鎖および/またはBCR軽鎖を含み得る。

10

【0220】

20

本方法は、第2の複数のバーコード化されたアンプリコン、またはその産物の配列情報を得るステップを含んでもよい。一部の実施形態では、配列情報を得るステップは、シーケンシングアダプターを、第2の複数のバーコード化されたアンプリコン、またはその産物に結合させるステップを含む。複数の伸長産物を増幅させるステップは、シーケンシングプライマーおよび/もしくはシーケンシングアダプターの結合部位の配列、その相補配列、ならびに/またはその一部分を、複数の伸長産物に付加するステップを含み得る。シーケンシングアダプターは、P5配列、P7配列、その相補配列、またはその一部分を含み得る。シーケンシングプライマーは、リード1シーケンシングプライマー、リード2シーケンシングプライマー、その相補配列、またはその一部分を含み得る。

【0221】

30

配列情報を得るステップは、単一細胞のBCR軽鎖およびBCR重鎖の配列情報を得るステップを含み得る。BCR軽鎖およびBCR重鎖の配列情報は、BCR軽鎖および/またはBCR重鎖の、相補性決定領域1(CDR1)、CDR2、CDR3、またはこれらの任意の組合せの配列を含み得る。本方法は、得られた配列情報に基づいて、単一細胞のBCR軽鎖とBCR重鎖を対合させるステップを含み得る。試料は、複数の単一細胞を含んでもよく、本方法は、得られた配列情報に基づいて、単一細胞の少なくとも50%のBCR軽鎖とBCR重鎖を対合させるステップを含み得る。一部の実施形態では、BCR軽鎖とBCR重鎖が本明細書において提供される方法に従って対合される試料の単一細胞のパーセンテージは、0.0000000001%、0.000000001%、0.00000001%、0.0000001%、0.000001%、0.00001%、0.0001%、0.001%、0.01%、0.1%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100

40

50

%、もしくはこれらの値のうちのいずれか2つの間の数もしくは範囲であってもよく、または約0.000000001%、0.000000001%、0.00000001%、0.0000001%、0.000001%、0.00001%、0.0001%、0.001%、0.01%、0.1%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%、もしくはこれらの値のうちのいずれか2つの間の数もしくは範囲であってもよい。一部の実施形態では、BCR軽鎖とBCR重鎖が本明細書において提供される方法に従って対合される試料の単一細胞のパーセンテージが、少なくとも、または最大で0.000000001%、0.000000001%、0.00000001%、0.0000001%、0.000001%、0.00001%、0.0001%、0.001%、0.01%、0.1%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、もしくは100%であってもよい。

10

20

【0222】

配列情報を得るステップは、単一細胞のTCRアルファ鎖およびTCRベータ鎖の配列情報を得るステップを含み得る。一部の実施形態では、TCRアルファ鎖およびTCRベータ鎖の配列情報は、TCRアルファ鎖および/またはTCRベータ鎖の、相補性決定領域1(CDR1)、CDR2、CDR3、またはこれらの任意の組合せの配列を含み得る。一部の実施形態では、本方法は、得られた配列情報に基づいて、単一細胞のTCRアルファ鎖とTCRベータ鎖を対合させるステップを含み得る。一部の実施形態では、試料は、複数の単一細胞を含んでもよく、本方法は、得られた配列情報に基づいて、単一細胞の少なくとも50%のTCRアルファ鎖とTCRベータ鎖を対合させるステップを含み得る。一部の実施形態では、TCRアルファ鎖とTCRベータ鎖が本明細書において提供される方法に従って対合される試料の単一細胞のパーセンテージは、0.000000001%、0.000000001%、0.00000001%、0.0000001%、0.000001%、0.00001%、0.0001%、0.001%、0.01%、0.1%、1%、2%、3%、4%、5%、6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、26%、27%、28%、29%、30%、31%、32%、33%、34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、50%、51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%

30

40

50

、 9 6 %、 9 7 %、 9 8 %、 9 9 %、 1 0 0 %、 もしくはこれらの値のうちのいずれか 2 つの間の数もしくは範囲であってもよく、または約 0 . 0 0 0 0 0 0 0 0 1 %、 0 . 0 0 0 0 0 0 0 1 %、 0 . 0 0 0 0 0 0 0 1 %、 0 . 0 0 0 0 0 1 %、 0 . 0 0 0 0 1 %、 0 . 0 0 0 1 %、 0 . 0 1 %、 0 . 1 %、 1 %、 2 %、 3 %、 4 %、 5 %、 6 %、 7 %、 8 %、 9 %、 1 0 %、 1 1 %、 1 2 %、 1 3 %、 1 4 %、 1 5 %、 1 6 %、 1 7 %、 1 8 %、 1 9 %、 2 0 %、 2 1 %、 2 2 %、 2 3 %、 2 4 %、 2 5 %、 2 6 %、 2 7 %、 2 8 %、 2 9 %、 3 0 %、 3 1 %、 3 2 %、 3 3 %、 3 4 %、 3 5 %、 3 6 %、 3 7 %、 3 8 %、 3 9 %、 4 0 %、 4 1 %、 4 2 %、 4 3 %、 4 4 %、 4 5 %、 4 6 %、 4 7 %、 4 8 %、 4 9 %、 5 0 %、 5 1 %、 5 2 %、 5 3 %、 5 4 %、 5 5 %、 5 6 %、 5 7 %、 5 8 %、 5 9 %、 6 0 %、 6 1 %、 6 2 %、 6 3 %、 6 4 %、 6 5 %、 6 6 %、 6 7 %、 6 8 %、 6 9 %、 7 0 %、 7 1 %、 7 2 %、 7 3 %、 7 4 %、 7 5 %、 7 6 %、 7 7 %、 7 8 %、 7 9 %、 8 0 %、 8 1 %、 8 2 %、 8 3 %、 8 4 %、 8 5 %、 8 6 %、 8 7 %、 8 8 %、 8 9 %、 9 0 %、 9 1 %、 9 2 %、 9 3 %、 9 4 %、 9 5 %、 9 6 %、 9 7 %、 9 8 %、 9 9 %、 1 0 0 %、 もしくはこれらの値のうちのいずれか 2 つの間の数もしくは範囲であってもよい。一部の実施形態では、 T C R アルファ鎖と T C R ベータ鎖が本明細書において提供される方法に従って対合される試料の単一細胞のパーセンテージが、少なくとも、または最大で 0 . 0 0 0 0 0 0 0 0 1 %、 0 . 0 0 0 0 0 0 0 1 %、 0 . 0 0 0 0 0 1 %、 0 . 0 0 0 1 %、 0 . 0 0 1 %、 0 . 0 1 %、 0 . 1 %、 1 %、 2 %、 3 %、 4 %、 5 %、 6 %、 7 %、 8 %、 9 %、 1 0 %、 1 1 %、 1 2 %、 1 3 %、 1 4 %、 1 5 %、 1 6 %、 1 7 %、 1 8 %、 1 9 %、 2 0 %、 2 1 %、 2 2 %、 2 3 %、 2 4 %、 2 5 %、 2 6 %、 2 7 %、 2 8 %、 2 9 %、 3 0 %、 3 1 %、 3 2 %、 3 3 %、 3 4 %、 3 5 %、 3 6 %、 3 7 %、 3 8 %、 3 9 %、 4 0 %、 4 1 %、 4 2 %、 4 3 %、 4 4 %、 4 5 %、 4 6 %、 4 7 %、 4 8 %、 4 9 %、 5 0 %、 5 1 %、 5 2 %、 5 3 %、 5 4 %、 5 5 %、 5 6 %、 5 7 %、 5 8 %、 5 9 %、 6 0 %、 6 1 %、 6 2 %、 6 3 %、 6 4 %、 6 5 %、 6 6 %、 6 7 %、 6 8 %、 6 9 %、 7 0 %、 7 1 %、 7 2 %、 7 3 %、 7 4 %、 7 5 %、 7 6 %、 7 7 %、 7 8 %、 7 9 %、 8 0 %、 8 1 %、 8 2 %、 8 3 %、 8 4 %、 8 5 %、 8 6 %、 8 7 %、 8 8 %、 8 9 %、 9 0 %、 9 1 %、 9 2 %、 9 3 %、 9 4 %、 9 5 %、 9 6 %、 9 7 %、 9 8 %、 9 9 %、 もしくは 1 0 0 % であってもよい。

10

20

【 0 2 2 3 】

30

配列情報を得るステップは、単一細胞の T C R ガンマ鎖および T C R デルタ鎖の配列情報を得るステップを含み得る。 T C R ガンマ鎖および T C R デルタ鎖の配列情報は、 T C R ガンマ鎖および / または T C R デルタ鎖の、相補性決定領域 1 (C D R 1)、 C D R 2、 C D R 3、またはこれらの任意の組合せの配列を含み得る。本方法は、得られた配列情報に基づいて、単一細胞の T C R ガンマ鎖と T C R デルタ鎖を対合させるステップを含み得る。試料は、複数の単一細胞を含んでもよく、本方法は、得られた配列情報に基づいて、単一細胞の少なくとも 5 0 % の T C R ガンマ鎖と T C R デルタ鎖を対合させるステップを含み得る。一部の実施形態では、 T C R デルタ鎖と T C R ガンマ鎖が本明細書において提供される方法に従って対合される試料の単一細胞のパーセンテージは、 0 . 0 0 0 0 0 0 0 0 1 %、 0 . 0 0 0 0 0 0 0 1 %、 0 . 0 0 0 0 0 0 1 %、 0 . 0 0 0 0 0 1 %、 0 . 0 0 0 0 1 %、 0 . 0 0 1 %、 0 . 0 1 %、 0 . 1 %、 1 %、 2 %、 3 %、 4 %、 5 %、 6 %、 7 %、 8 %、 9 %、 1 0 %、 1 1 %、 1 2 %、 1 3 %、 1 4 %、 1 5 %、 1 6 %、 1 7 %、 1 8 %、 1 9 %、 2 0 %、 2 1 %、 2 2 %、 2 3 %、 2 4 %、 2 5 %、 2 6 %、 2 7 %、 2 8 %、 2 9 %、 3 0 %、 3 1 %、 3 2 %、 3 3 %、 3 4 %、 3 5 %、 3 6 %、 3 7 %、 3 8 %、 3 9 %、 4 0 %、 4 1 %、 4 2 %、 4 3 %、 4 4 %、 4 5 %、 4 6 %、 4 7 %、 4 8 %、 4 9 %、 5 0 %、 5 1 %、 5 2 %、 5 3 %、 5 4 %、 5 5 %、 5 6 %、 5 7 %、 5 8 %、 5 9 %、 6 0 %、 6 1 %、 6 2 %、 6 3 %、 6 4 %、 6 5 %、 6 6 %、 6 7 %、 6 8 %、 6 9 %、 7 0 %、 7 1 %、 7 2 %、 7 3 %、 7 4 %、 7 5 %、 7 6 %、 7 7 %、 7 8 %、 7 9 %、 8 0 %、 8 1 %、 8 2 %、 8 3 %、 8 4 %、 8 5 %、 8 6 %、 8 7 %、 8 8 %、 8 9 %、 9 0 %、 9 1 %、 9 2 %、 9 3 %、 9 4 %

40

50

、 9 5 %、 9 6 %、 9 7 %、 9 8 %、 9 9 %、 1 0 0 %、 もしくはこれらの値のうちのいずれか 2 つの間の数もしくは範囲であってもよく、または約 0 . 0 0 0 0 0 0 0 0 1 %、 0 . 0 0 0 0 0 0 0 0 1 %、 0 . 0 0 0 0 0 0 1 %、 0 . 0 0 0 0 0 1 %、 0 . 0 0 0 0 1 %、 0 . 0 0 1 %、 0 . 0 1 %、 0 . 1 %、 1 %、 2 %、 3 %、 4 %、 5 %、 6 %、 7 %、 8 %、 9 %、 1 0 %、 1 1 %、 1 2 %、 1 3 %、 1 4 %、 1 5 %、 1 6 %、 1 7 %、 1 8 %、 1 9 %、 2 0 %、 2 1 %、 2 2 %、 2 3 %、 2 4 %、 2 5 %、 2 6 %、 2 7 %、 2 8 %、 2 9 %、 3 0 %、 3 1 %、 3 2 %、 3 3 %、 3 4 %、 3 5 %、 3 6 %、 3 7 %、 3 8 %、 3 9 %、 4 0 %、 4 1 %、 4 2 %、 4 3 %、 4 4 %、 4 5 %、 4 6 %、 4 7 %、 4 8 %、 4 9 %、 5 0 %、 5 1 %、 5 2 %、 5 3 %、 5 4 %、 5 5 %、 5 6 %、 5 7 %、 5 8 %、 5 9 %、 6 0 %、 6 1 %、 6 2 %、 6 3 %、 6 4 %、 6 5 %、 6 6 %、 6 7 %、 6 8 %、 6 9 %、 7 0 %、 7 1 %、 7 2 %、 7 3 %、 7 4 %、 7 5 %、 7 6 %、 7 7 %、 7 8 %、 7 9 %、 8 0 %、 8 1 %、 8 2 %、 8 3 %、 8 4 %、 8 5 %、 8 6 %、 8 7 %、 8 8 %、 8 9 %、 9 0 %、 9 1 %、 9 2 %、 9 3 %、 9 4 %、 9 5 %、 9 6 %、 9 7 %、 9 8 %、 9 9 %、 1 0 0 %、 もしくはこれらの値のうちのいずれか 2 つの間の数もしくは範囲であってもよい。一部の実施形態では、 T C R デルタ鎖と T C R ガンマ鎖が本明細書において提供される方法に従って対合される試料の単一細胞のパーセンテージが、少なくとも、または最大で 0 . 0 0 0 0 0 0 0 0 1 %、 0 . 0 0 0 0 0 0 0 1 %、 0 . 0 0 0 0 0 0 1 %、 0 . 0 0 0 0 1 %、 0 . 0 0 0 1 %、 0 . 0 1 %、 0 . 1 %、 1 %、 2 %、 3 %、 4 %、 5 %、 6 %、 7 %、 8 %、 9 %、 1 0 %、 1 1 %、 1 2 %、 1 3 %、 1 4 %、 1 5 %、 1 6 %、 1 7 %、 1 8 %、 1 9 %、 2 0 %、 2 1 %、 2 2 %、 2 3 %、 2 4 %、 2 5 %、 2 6 %、 2 7 %、 2 8 %、 2 9 %、 3 0 %、 3 1 %、 3 2 %、 3 3 %、 3 4 %、 3 5 %、 3 6 %、 3 7 %、 3 8 %、 3 9 %、 4 0 %、 4 1 %、 4 2 %、 4 3 %、 4 4 %、 4 5 %、 4 6 %、 4 7 %、 4 8 %、 4 9 %、 5 0 %、 5 1 %、 5 2 %、 5 3 %、 5 4 %、 5 5 %、 5 6 %、 5 7 %、 5 8 %、 5 9 %、 6 0 %、 6 1 %、 6 2 %、 6 3 %、 6 4 %、 6 5 %、 6 6 %、 6 7 %、 6 8 %、 6 9 %、 7 0 %、 7 1 %、 7 2 %、 7 3 %、 7 4 %、 7 5 %、 7 6 %、 7 7 %、 7 8 %、 7 9 %、 8 0 %、 8 1 %、 8 2 %、 8 3 %、 8 4 %、 8 5 %、 8 6 %、 8 7 %、 8 8 %、 8 9 %、 9 0 %、 9 1 %、 9 2 %、 9 3 %、 9 4 %、 9 5 %、 9 6 %、 9 7 %、 9 8 %、 9 9 %、 もしくは 1 0 0 % であってもよい。

10

20

【 0 2 2 4 】

30

第 2 の複数のバーコード化されたアンプリコン、またはその産物の配列情報を得るステップは、第 2 の複数のバーコード化されたアンプリコンの複数のシーケンシングリード、またはその産物を含むシーケンシングデータを得るステップであって、複数のシーケンシングリードのそれぞれが、 (1) 細胞標識配列、 (2) 分子標識配列、 および / または (3) 核酸標的の部分配列を含む、ステップを含み得る。一部の実施形態では、本方法は、試料の単一細胞を示す、それぞれ固有の細胞標識配列について、核酸標的の複数のシーケンシングリードのそれぞれをアライメントさせて、核酸標的のアライメントされた配列を生成するステップを含む。図 9 B に示されるように、シーケンシングデータは、複数のシーケンシングリード、例えば、リード 1 9 1 0 およびリード 2 9 2 0 のリードなどを含み得る。一部の実施形態では、ランダムプライミングステップの結果として、リード 1 9 1 0 および / またはリード 2 9 2 0 のリードは、核酸標的 (例えば、免疫受容体転写物) 全体を合わせて包含し得る。本方法は、複数のリード 1 9 1 0 およびリード 2 9 2 0 をアライメントすることによって、核酸標的 (例えば、免疫受容体 m R N A 9 3 0) の完全長配列の生物情報学的再構築 9 0 0 を含み得る。有利には、本明細書において提供される組成物および方法は、核酸標的 (例えば、 V (D) J を含有する転写物) の同定および計数の両方を提供することができる。

40

【 0 2 2 5 】

一部の実施形態では、核酸標的のアライメントされた配列は、核酸標的の c D N A 配列の少なくとも 5 0 %、核酸標的の c D N A 配列の少なくとも 7 0 %、核酸標的の c D N A 配列の少なくとも 9 0 %、または核酸標的の c D N A 配列の完全長を含む。一部の実施形

50

態では、核酸標的のアライメントされた配列は、核酸標的の c D N A 配列の 1 %、2 %、3 %、4 %、5 %、6 %、7 %、8 %、9 %、10 %、11 %、12 %、13 %、14 %、15 %、16 %、17 %、18 %、19 %、20 %、21 %、22 %、23 %、24 %、25 %、26 %、27 %、28 %、29 %、30 %、31 %、32 %、33 %、34 %、35 %、36 %、37 %、38 %、39 %、40 %、41 %、42 %、43 %、44 %、45 %、46 %、47 %、48 %、49 %、50 %、51 %、52 %、53 %、54 %、55 %、56 %、57 %、58 %、59 %、60 %、61 %、62 %、63 %、64 %、65 %、66 %、67 %、68 %、69 %、70 %、71 %、72 %、73 %、74 %、75 %、76 %、77 %、78 %、79 %、80 %、81 %、82 %、83 %、84 %、85 %、86 %、87 %、88 %、89 %、90 %、91 %、92 %、93 %、94 %、95 %、96 %、97 %、98 %、99 %、100 %、もしくはこれらの値のうちのいずれか 2 つの間の数もしくは範囲であり得るか、または約 1 %、2 %、3 %、4 %、5 %、6 %、7 %、8 %、9 %、10 %、11 %、12 %、13 %、14 %、15 %、16 %、17 %、18 %、19 %、20 %、21 %、22 %、23 %、24 %、25 %、26 %、27 %、28 %、29 %、30 %、31 %、32 %、33 %、34 %、35 %、36 %、37 %、38 %、39 %、40 %、41 %、42 %、43 %、44 %、45 %、46 %、47 %、48 %、49 %、50 %、51 %、52 %、53 %、54 %、55 %、56 %、57 %、58 %、59 %、60 %、61 %、62 %、63 %、64 %、65 %、66 %、67 %、68 %、69 %、70 %、71 %、72 %、73 %、74 %、75 %、76 %、77 %、78 %、79 %、80 %、81 %、82 %、83 %、84 %、85 %、86 %、87 %、88 %、89 %、90 %、91 %、92 %、93 %、94 %、95 %、96 %、97 %、98 %、99 %、100 %、もしくはこれらの値のうちのいずれか 2 つの間の数もしくは範囲であり得る。一部の実施形態では、核酸標的のアライメントされた配列は、核酸標的の c D N A 配列の少なくとも、または最大で 1 %、2 %、3 %、4 %、5 %、6 %、7 %、8 %、9 %、10 %、11 %、12 %、13 %、14 %、15 %、16 %、17 %、18 %、19 %、20 %、21 %、22 %、23 %、24 %、25 %、26 %、27 %、28 %、29 %、30 %、31 %、32 %、33 %、34 %、35 %、36 %、37 %、38 %、39 %、40 %、41 %、42 %、43 %、44 %、45 %、46 %、47 %、48 %、49 %、50 %、51 %、52 %、53 %、54 %、55 %、56 %、57 %、58 %、59 %、60 %、61 %、62 %、63 %、64 %、65 %、66 %、67 %、68 %、69 %、70 %、71 %、72 %、73 %、74 %、75 %、76 %、77 %、78 %、79 %、80 %、81 %、82 %、83 %、84 %、85 %、86 %、87 %、88 %、89 %、90 %、91 %、92 %、93 %、94 %、95 %、96 %、97 %、98 %、99 %、もしくは 100 % であり得る。核酸標的は、免疫受容体であってもよい。

10

20

30

【0226】

一部の実施形態では、核酸標的のアライメントされた配列は、相補性決定領域 1 (C D R 1)、相補性決定領域 2 (C D R 2)、相補性決定領域 3 (C D R 3)、可変領域、可変領域の完全長、またはこれらの組合せを含む。核酸標的のアライメントされた配列は、可変領域、多様性領域、可変領域、多様性領域および/もしくは定常領域の接合部、またはこれらの任意の組合せを含み得る。一部の実施形態では、本方法は、断片化も、タグメンテーションも、またはその両方も含まない。

40

【0227】

核酸標的の完全長情報を得るためのキット

本明細書の開示は、キットを含む。一部の実施形態では、キットは、複数のオリゴヌクレオチドバーコードであって、複数のオリゴヌクレオチドバーコードのそれぞれが分子標識および標的結合領域を含み、複数のオリゴヌクレオチドバーコードのうちの少なくとも 10 個が異なる分子標識配列を含む、複数のオリゴヌクレオチドバーコード；逆転写酵素；標的結合領域またはその一部分を含むテンプレートスイッチングオリゴヌクレオチド；

および 5' から 3' へのエキソヌクレアーゼ活性および 3' から 5' へのエキソヌクレアーゼ活 50

性のうちの少なくとも1つを欠くDNAポリメラーゼを含む。本明細書の開示は、キットを含む。一部の実施形態では、キットは、複数のオリゴヌクレオチドバーコードであって、複数のオリゴヌクレオチドバーコードのそれぞれが、分子標識および標的結合領域を含み、複数のオリゴヌクレオチドバーコードのうちの少なくとも10個が異なる分子標識配列を含む、複数のオリゴヌクレオチドバーコード；末端デオキシヌクレオチド転移酵素；ならびに5'から3'へのエキソヌクレアーゼ活性および3'から5'へのエキソヌクレアーゼ活性のうちの少なくとも1つを欠くDNAポリメラーゼを含む。

【0228】

キットは、ランダムプライマーを含み得る。ランダムプライマーは、ヌクレオチドのランダム配列を含み得る。ヌクレオチドのランダム配列は、約4~約30ヌクレオチドの長さであり得る。一部の実施形態では、DNAポリメラーゼは、Klenow断片を含む。一部の実施形態では、逆転写酵素は、ウイルス逆転写酵素を含む。一部の実施形態では、ウイルス逆転写酵素は、マウス白血病ウイルス(MLV)逆転写酵素である。一部の実施形態では、ウイルス逆転写酵素は、モロニーマウス白血病ウイルス(MMLV)逆転写酵素である。一部の実施形態では、テンプレートスイッチオリゴヌクレオチドは、1つまたは複数の3'リボヌクレオチド、例えば、3つの3'リボヌクレオチドを含む。一部の実施形態では、3'リボヌクレオチドは、グアニンを含む。一部の実施形態では、キットは、エチレングリコール、ポリエチレングリコール、1,2-プロパンジオール、ジメチルスルホキシド(DMSO)、グリセロール、ホルムアミド、7-デアザ-GTP、アセトアミド、塩化テトラメチルアンモニウム塩、ベタイン、またはこれらの任意の組合せのうちの1つまたは複数を含む。

【0229】

一部の実施形態では、キットは、緩衝液を含む。一部の実施形態では、キットは、カートリッジを含む。一部の実施形態では、キットは、逆転写反応のための1つまたは複数の試薬を含む。一部の実施形態では、キットは、増幅反応のための1つまたは複数の試薬を含む。一部の実施形態では、標的結合領域は、遺伝子特異的配列、オリゴ(dT)配列、ランダム多量体、またはこれらの任意の組合せを含む。一部の実施形態では、オリゴヌクレオチドバーコードは、同一の試料標識および/または同一の細胞標識を含む。一部の実施形態では、複数のオリゴヌクレオチドバーコードの各試料標識および/または細胞標識は、少なくとも6個のヌクレオチドを含む。一部の実施形態では、複数のオリゴヌクレオチドバーコードの各分子標識は、少なくとも6個のヌクレオチドを含む。一部の実施形態では、複数のオリゴヌクレオチドバーコードのうちの少なくとも1つは、合成粒子上に固定化されている。一部の実施形態では、複数のオリゴヌクレオチドバーコードのうちの少なくとも1つは、合成粒子上に部分的に固定化されている。一部の実施形態では、複数のオリゴヌクレオチドバーコードのうちの少なくとも1つは、合成粒子内に封入されている。一部の実施形態では、複数のオリゴヌクレオチドバーコードのうちの少なくとも1つは、合成粒子内に部分的に封入されている。一部の実施形態では、合成粒子は、崩壊可能である。一部の実施形態では、合成粒子は、ビーズ、例えば、セファロースビーズ、ストレプトアビジンビーズ、アガロースビーズ、磁性ビーズ、コンジュゲートビーズ、プロテインAコンジュゲートビーズ、プロテインGコンジュゲートビーズ、プロテインA/Gコンジュゲートビーズ、プロテインLコンジュゲートビーズ、オリゴ(dT)コンジュゲートビーズ、シリカビーズ、シリカ様ビーズ、抗ビオチンマイクロビーズ、抗蛍光色素マイクロビーズ、またはこれらの任意の組合せを含む。一部の実施形態では、合成粒子は、ポリジメチルシロキサン(PDMS)、ポリスチレン、ガラス、ポリプロピレン、アガロース、ゼラチン、ハイドロゲル、常磁性体、セラミック、プラスチック、ガラス、メチルスチレン、アクリルポリマー、チタン、ラテックス、セファロース、セルロース、ナイロン、シリコン、およびこれらの任意の組合せからなる群から選択される材料を含む。一部の実施形態では、合成粒子は、崩壊可能ハイドロゲル粒子を含む。一部の実施形態では、複数のオリゴヌクレオチドバーコードのそれぞれは、リンカー官能基を含み、合成粒子は、固体支持体官能基を含み、ならびに/または支持体官能基およびリンカー官能基は互いに

10

20

30

40

50

会合している。一部の実施形態では、リンカー官能基および支持体官能基は、独立して、C6、ピオチン、ストレプトアビジン、第一級アミン、アルデヒド、ケトン、およびこれらの任意の組合せからなる群から選択される。

【実施例】

【0230】

本明細書に記載の実施形態のいくつかの態様を以下の実施例でさらに詳しく開示するが、これらの実施例は、本開示の範囲を制限することを何ら意図しない。

【0231】

(実施例1)

完全長V(D)J配列のランダムプライミングに基づく決定

この実施例は、本明細書において提供される核酸標的(例えば、V(D)Jを含有する転写物)の完全長発現プロファイリングのためのランダムプライミングに基づく方法について実証する。TCR産物(図10A~10B)およびBCR産物(図10C~10D)のバイオアナライザーによる追跡において示されているように、現在利用可能な方法(図10Bおよび10D)と比較して、より小さな断片が、本明細書において提供されるランダムプライミングベースのアプローチ(図10Aおよび10C)で得られる。次に、シーケンシングを実施して、完全長であり(図11B)、細胞に由来する(図11A)V(D)J配列のパーセンテージを決定した。VDJに対してアライメントされ、コンティグへとアセンブルされるのに成功し、かつそのコンティグが有効なVDJ配列である(CDRを有する)リードのうち、図11Aに示されるパーセンテージは、推定上の細胞に由来するこれらのリードの数に関する。推定上の細胞では、優勢であると同定されたコンティグに関して、完全長であるこれらのパーセンテージは、図11Bに示される。そのコンティグが、鎖:FR1、CDR1、FR2、CDR2、FR3、CDR3、FR4のすべての主要な部分について同定された配列を有するかどうかによって、完全長を決定した。さらに、図12A~12Dに示されるように、完全長の対合した鎖が正しい細胞型(B細胞におけるBCR、T細胞におけるTCR)で見られ、本明細書において提供されるランダムプライミングベースのアプローチで得られたBCR対合(図12A~12B)およびTCR対合(図12C~12D)は、現在利用可能な方法(図12Bおよび12D)よりも優れていた。これらの結果は、本明細書において提供されるランダムプライミングベースの方法および組成物に関する原理証明を与える。

【0232】

様々な態様および実施形態が本明細書に開示されているが、他の態様および実施形態が当業者にとって明らかであろう。本明細書に開示される様々な態様および実施形態は、例示を目的とし、限定を意図するものではなく、以下の請求項によって示されている真の範囲および精神を有する。

【0233】

当業者は、本明細書に開示されるこのおよび他のプロセスおよび方法について、そのプロセスおよび方法において実施される機能が、異なる順序で実現され得ることを認識するであろう。さらに、概説されるステップおよび操作は例として与えられるに過ぎず、ステップおよび操作のいくつかは、より少ないステップおよび操作へと任意に組み合わせられてもよく、または本開示の実施形態の根本的要素を減じることなく追加のステップおよび操作に拡大されてもよい。

【0234】

本明細書の実質的に任意の複数形および/または単数形の用語の使用に関連して、文脈上および/または適用上適切であれば、当業者は、複数形から単数形へおよび/または単数形から複数形への変換が可能である。明確にするために種々の単数形/複数形の入替えを本明細書に明示的に記述し得る。

【0235】

一般的には、本明細書において、特に添付の特許請求の範囲(例えば、添付の特許請求の範囲の本文)で使用される用語は、全般的に、「オープン」用語であることが意図され

10

20

30

40

50

ることが当業者によって理解されるであろう（例えば、「含む（including）」という用語は、「以下に限定されないが、～を含む（including but not limited to）」と解釈されるべきであり、「有する（having）」という用語は、「少なくとも～を有する（having at least）」と解釈されるべきであり、「含む（includes）」という用語は「以下に限定されないが、～を含む（includes but is not limited to）」と解釈されるべきであるなど）。さらに、導入された請求項の記載の特定数が意図される場合、このような意図は請求項で明示的に記載され、このような記載の非存在下では、このような意図は存在しないことが、当業者によって理解されるであろう。例えば、理解の一助として、以下の添付の特許請求の範囲は、請求項の記載を導入するために、導入語句「少なくとも1つの（at least one）」および「1つまたは複数の（one or more）」の使用を含み得る。しかし、このような語句の使用は、たとえ同一の請求項が、「1つまたは複数の（one or more）」または「少なくとも1つの（at least one）」という導入語句および「a」または「an」のような不定冠詞を含む場合でも、「a」または「an」という不定冠詞による請求項の記載の導入がこのような導入された請求項の記載を含むいかなる特定の請求項もそのような記載を1つのみ含む実施形態に限定することを意味するものと解釈させるべきではなく（例えば、「a」および/または「an」は、「少なくとも1つの（at least one）」または「1つまたは複数の（one or more）」を意味すると解釈されるべきである）、請求項の記載を導入するために使用される定冠詞の使用についても同様である。さらに、たとえ特定数の導入された請求項の記載が明示的に記載されたとしても、このような記載は少なくとも記載された数を意味すると解釈されるべきであることを当業者は認識するであろう（例えば、「2つの記載」という他の修飾語を含まない無修飾の記載は、少なくとも2つの記載、または2つ以上の記載を意味する）。さらに、「A、B、およびCのうちの少なくとも1つなど」に類似した慣例が使用される場合、一般的には、このような構成は当業者がその慣例を理解する意味であることが意図される（例えば、「A、B、およびCのうちの少なくとも1つを有するシステム」は、以下に限定されないが、A単独、B単独、C単独、AとBの両方、AとCの両方、BとCの両方、および/またはAとBとCの全部などを有するシステムを含むことになる）。「A、B、またはCのうちの少なくとも1つなど」に類似した慣例が使用される場合、一般的には、このような構成は当業者がその慣例を理解する意味であることが意図される（例えば、「A、B、またはCのうちの少なくとも1つを有するシステム」は、以下に限定されないが、A単独、B単独、C単独、AとBの両方、AとCの両方、BとCの両方、および/またはAとBとCの全部などを有するシステムを含むことになる）。さらに、2つ以上の代替用語を表す実質上任意の選言的な語および/または語句は、明細書、請求項、または図面にかかわらず、用語の1つ、用語のいずれか、または用語の両方を含む可能性が企図されると理解されるべきであることが当業者によって理解されるであろう。例えば、語句「AまたはB」は、「A」もしくは「B」または「AおよびB」の可能性を含むことが理解される。

【0236】

さらに、本開示の特徴または態様がマーカッシュ群により記述される場合、それにより、本開示は、マーカッシュ群の任意の個別のメンバーまたはメンバーの下位群によっても記述されることが、当業者に認識されるであろう。

【0237】

当業者に理解されるように、あらゆる目的で、例えば、明細書の提供に関して、本明細書に開示された範囲はすべて、あらゆる可能な下位範囲およびその下位範囲の組合せも包含する。いずれの列挙された範囲も、十分に記述されたものとしてかつその範囲が少なくとも2等分、3等分、4等分、5等分、10等分などされ得るものとして容易に認識可能である。非限定的な例として、本明細書で考察した各範囲は、下3分の1、中3分の1、上3分の1などに容易に分解可能である。同様に、当業者に理解されるように、「最大」、「少なくとも」などの表現はすべて、記載された数を含み、以上で考察したように後続

10

20

30

40

50

的に下位範囲に分解可能な範囲を指す。最後に、当業者によって理解されるように、範囲は、それぞれ個々のメンバーを含む。よって、例えば、1～3個の細胞を有する群は、1、2、または3個の細胞を有する群を指す。同様に、1～5個の細胞を有する群は、1、2、3、4、または5個の細胞を有する群を指し、他も同様である。

【0238】

前記内容から、本開示の様々な実施形態が例証のために本明細書に記載されており、本開示の範囲および精神を逸脱することなく様々な修正がなされ得ることが認識されよう。したがって、本明細書に開示される様々な実施形態は、限定を意図するものではなく、以下の請求項によって示されている真の範囲および精神を有する。

本発明の好ましい態様は、下記の通りである。

〔1〕試料中の核酸標的を標識するための方法であって、

核酸標的のコピーを複数のオリゴヌクレオチドバーコードと接触させるステップであって、各オリゴヌクレオチドバーコードが、第1のユニバーサル配列、分子標識、および核酸標的にハイブリダイズすることが可能な標的結合領域を含む、ステップと；

それぞれが標的結合領域および該標的結合領域の相補体を含む、複数のバーコード化された核酸分子を生成するステップと；

各バーコード化された核酸分子の標的結合領域の相補体を、

(i) 複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコード、

(ii) バーコード化された核酸分子自体、および

(iii) 複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子

のうちの1つまたは複数の標的結合領域とハイブリダイズさせるステップと；

複数のバーコード化された核酸分子の3'末端を伸長させて、それぞれが第1の分子標識および第2の分子標識を含む、複数の伸長したバーコード化された核酸分子を生成するステップと；

核酸標的の配列にハイブリダイズすることが可能な標的的特異的プライマーおよび第1のユニバーサル配列を含むプライマーを使用して、複数の伸長したバーコード化された核酸分子を増幅して、それによって、核酸標的の配列またはその一部分を含む、第1の複数のバーコード化されたアンプリコンを生成するステップと；

ランダムプライマーを第1の複数のバーコード化されたアンプリコンにハイブリダイズさせ、ランダムプライマーを伸長させて、複数の伸長産物を生成するステップであって、ランダムプライマーが、第2のユニバーサル配列またはその相補体を含む、ステップと；

第1のユニバーサル配列および第2のユニバーサル配列またはそれらの相補体にハイブリダイズすることが可能なプライマーを使用して、複数の伸長産物を増幅させて、それによって、第2の複数のバーコード化されたアンプリコンを生成するステップと、を含む、方法。

〔2〕第2の複数のバーコード化されたアンプリコンまたはその産物と会合した、別個の配列を有する第1の分子標識、別個の配列を有する第2の分子標識、またはそれらの組合せの数に基づいて、試料中の核酸標的のコピー数を決定するステップを含む、前記〔1〕に記載の方法。

〔3〕試料中の核酸標的の数を決定するための方法であって、

核酸標的のコピーを複数のオリゴヌクレオチドバーコードと接触させるステップであって、各オリゴヌクレオチドバーコードが、第1のユニバーサル配列、分子標識、および核酸標的にハイブリダイズすることが可能な標的結合領域を含む、ステップと；

それぞれが標的結合領域および該標的結合領域の相補体を含む、複数のバーコード化された核酸分子を生成するステップと；

各バーコード化された核酸分子の標的結合領域の相補体を、

(i) 複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコード、

(ii) バーコード化された核酸分子自体、および

(iii) 複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸

10

20

30

40

50

分子

のうちの1つまたは複数の標的結合領域とハイブリダイズさせるステップと；

複数のバーコード化された核酸分子の3'末端を伸長させて、それぞれが第1の分子標識および第2の分子標識を含む、複数の伸長したバーコード化された核酸分子を生成するステップと、

核酸標的の配列にハイブリダイズすることが可能な標的特異的プライマーおよび第1のユニバーサル配列を含むプライマーを使用して、複数の伸長したバーコード化された核酸分子を増幅して、それによって、核酸標的の配列またはその一部分を含む、第1の複数のバーコード化されたアンプリコンを生成するステップと；

ランダムプライマーを第1の複数のバーコード化されたアンプリコンにハイブリダイズさせ、ランダムプライマーを伸長させて、複数の伸長産物を生成するステップであって、ランダムプライマーが、第2のユニバーサル配列またはその相補体を含む、ステップと；

第1のユニバーサル配列および第2のユニバーサル配列またはそれらの相補体にハイブリダイズすることが可能なプライマーを使用して、複数の伸長産物を増幅させて、それによって、第2の複数のバーコード化されたアンプリコンを生成するステップと；

第2の複数のバーコード化されたアンプリコンまたはその産物と会合した、別個の配列を有する第1の分子標識、別個の配列を有する第2の分子標識、またはそれらの組合せの数に基づいて、試料中の核酸標的のコピー数を決定するステップと、を含む、方法。

〔4〕試料中の核酸標的のコピー数を決定するステップが、

（a）第2の複数のバーコード化されたアンプリコンもしくはその産物と会合した、別個の配列を有する第2の分子標識の数に基づいて、試料中の核酸標的のコピー数を決定するステップ、および/または

（b）第2の複数のバーコード化されたアンプリコンもしくはその産物と会合した、別個の配列を有する第1の分子標識の数に基づいて、試料中の核酸標的のコピー数を決定するステップ

を含む、前記〔2〕～〔3〕のいずれか1項に記載の方法。

〔5〕各バーコード化された核酸分子の標的結合領域の相補体を、（i）複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコード、（ii）バーコード化された核酸分子自体、および/または（iii）複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子の標的結合領域とハイブリダイズさせる前に、複数のバーコード化された核酸分子を変性させるステップを含む、前記〔1〕～〔4〕のいずれか1項に記載の方法。

〔6〕複数の伸長したバーコード化された核酸分子を増幅させる前に、複数の伸長したバーコード化された核酸分子を変性させるステップを含む、前記〔1〕～〔5〕のいずれか1項に記載の方法。

〔7〕複数の核酸標的のそれぞれの配列が、複数の核酸標的のそれぞれの部分配列を含む、前記〔1〕～〔6〕のいずれか1項に記載の方法。

〔8〕複数のバーコード化された核酸分子中の核酸標的の配列が、核酸標的の部分配列を含む、前記〔1〕～〔7〕のいずれか1項に記載の方法。

〔9〕第1の分子標識が、複数のバーコード化された核酸分子の3'末端を伸長させた後に、第2の分子標識にハイブリダイズされる、前記〔1〕～〔8〕のいずれか1項に記載の方法。

〔10〕伸長したバーコード化された核酸分子がそれぞれ、第1の分子標識、第2の分子標識、標的結合領域、および該標的結合領域の相補体を含む、前記〔1〕～〔9〕のいずれか1項に記載の方法。

〔11〕標的結合領域の相補体が、標的結合領域の一部分に相補的である、前記〔1〕～〔10〕のいずれか1項に記載の方法。

〔12〕標的結合領域が、遺伝子特異的配列、ポリ（dT）配列、または両方を含む、前記〔1〕～〔11〕のいずれか1項に記載の方法。

10

20

30

40

50

〔13〕それぞれが標的結合領域および該標的結合領域の相補体を含む、複数のバーコード化された核酸分子を生成するステップが、

（i）複数のオリゴヌクレオチドバーコードを使用して、試料中の核酸標的のコピーをバーコード化して、それぞれが核酸標的の少なくとも一部分に相補的な配列、第1の分子標識、および標的結合領域を含む、複数のバーコード化された核酸分子を生成するステップ；および

（ii）標的結合領域の相補体を含むオリゴヌクレオチドを複数のバーコード化された核酸分子に結合させて、それぞれが標的結合領域および該標的結合領域の相補体を含む、複数のバーコード化された核酸分子を生成するステップ

を含む、前記〔1〕～〔12〕のいずれか1項に記載の方法。

10

〔14〕それぞれが標的結合領域および該標的結合領域の相補体を含む、複数のバーコード化された核酸分子を生成するステップが、

（i）複数のオリゴヌクレオチドバーコードを使用して、試料中の核酸標的のコピーをバーコード化して、それぞれが核酸標的の少なくとも一部分に相補的な配列、第1の分子標識、および標的結合領域を含む、複数のバーコード化された核酸分子を生成するステップ；

（ii）複数のバーコード化された核酸分子を増幅させて、複数の増幅したバーコード化された核酸分子を生成するステップ；ならびに

（iii）標的結合領域の相補体を含むオリゴヌクレオチドを複数の増幅したバーコード化された核酸分子に結合させて、それぞれが標的結合領域および該標的結合領域の相補体を含む、複数のバーコード化された核酸分子を生成するステップ

20

を含む、前記〔1〕～〔12〕のいずれか1項に記載の方法。

〔15〕試料中の核酸標的のコピーをバーコード化するステップが、核酸標的のコピーにハイブリダイズした複数のオリゴヌクレオチドバーコードを伸長させて、複数のバーコード化された核酸分子を生成するステップを含む、前記〔13〕～〔14〕のいずれか1項に記載の方法。

〔16〕核酸標的のコピーにハイブリダイズした複数のオリゴヌクレオチドバーコードを伸長させるステップが、複数のオリゴヌクレオチドバーコードにハイブリダイズした核酸標的のコピーを逆転写するステップを含む、前記〔15〕に記載の方法。

〔17〕標的結合領域の相補体を含むオリゴヌクレオチドを結合させるステップが、標的結合領域の相補体を含むオリゴヌクレオチドを、複数のバーコード化された核酸分子および/または増幅したバーコード化された核酸分子にライゲーションさせるステップを含む、前記〔13〕～〔16〕のいずれか1項に記載の方法。

30

〔18〕標的結合領域がポリ（dT）配列を含み、該標的結合領域の相補体を含むオリゴヌクレオチドを結合させるステップが、末端デオキシヌクレオチド転移酵素を使用して、複数のアデノシンリン酸を複数のバーコード化された核酸分子および/または増幅したバーコード化された核酸分子に付加するステップを含む、前記〔13〕～〔16〕のいずれか1項に記載の方法。

〔19〕それぞれが標的結合領域および該標的結合領域の相補体を含む、複数のバーコード化された核酸分子を生成するステップが、

40

逆転写酵素と、標的結合領域またはその一部分を含むテンプレートスイッチオリゴヌクレオチドとの存在下で、核酸標的のコピーにハイブリダイズした複数のオリゴヌクレオチドバーコードを伸長させて、それぞれが核酸標的の少なくとも一部分に相補的な配列、第1の分子標識、標的結合領域、および該標的結合領域の相補体を含む、複数のバーコード化された核酸分子を生成するステップ

を含む、前記〔1〕～〔12〕のいずれか1項に記載の方法。

〔20〕逆転写酵素が、末端転移酵素活性の能力がある、前記〔19〕に記載の方法。

〔21〕テンプレートスイッチオリゴヌクレオチドが、1つまたは複数の3'リボヌクレオチド、任意に、3つの3'リボヌクレオチドを含む、前記〔19〕～〔20〕のいずれか1項に記載の方法。

50

〔 2 2 〕 3' リボヌクレオチドが、グアニンを含む、前記〔 2 1 〕に記載の方法。

〔 2 3 〕 逆転写酵素が、ウイルス逆転写酵素、任意に、マウス白血病ウイルス (M L V) 逆転写酵素またはモロニーマウス白血病ウイルス (M M L V) 逆転写酵素を含む、前記〔 1 9 〕 ~ 〔 2 2 〕のいずれか 1 項に記載の方法。

〔 2 4 〕 バーコード化された核酸分子の標的結合領域の相補体をバーコード化された核酸分子自体の標的結合領域とハイブリダイズさせるステップが、ステムループを形成する、バーコード化された核酸分子内の、標的結合領域と該標的結合領域の相補体との分子内ハイブリダイゼーションを含む、前記〔 1 〕 ~ 〔 2 3 〕のいずれか 1 項に記載の方法。

〔 2 5 〕 第 2 の分子標識が、第 1 の分子標識の相補体である、前記〔 2 4 〕に記載の方法。

〔 2 6 〕 バーコード化された核酸分子の標的結合領域の相補体を、複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコードの標的結合領域とハイブリダイズさせるステップが、バーコード化された核酸分子の標的結合領域の相補体と、複数のオリゴヌクレオチドバーコードのうちのオリゴヌクレオチドバーコードの標的結合領域との分子間ハイブリダイゼーションを含む、前記〔 1 〕 ~ 〔 2 5 〕のいずれか 1 項に記載の方法。

〔 2 7 〕 第 2 の分子標識が、第 1 の分子標識と異なり、第 2 の分子標識が、第 1 の分子標識の相補体ではない、前記〔 2 6 〕に記載の方法。

〔 2 8 〕 バーコード化された核酸分子の標的結合領域の相補体にハイブリダイズしたオリゴヌクレオチドバーコードの 3' 末端を伸長させて、それぞれが第 1 の分子標識の相補体および第 2 の分子標識を含む、複数の伸長したバーコード化された核酸分子を生成するステップを含む、前記〔 2 6 〕 ~ 〔 2 7 〕のいずれか 1 項に記載の方法。

〔 2 9 〕 第 2 の分子標識の配列が、第 1 の分子標識の配列と異なり、第 2 の分子標識が、第 1 の分子標識の相補体ではない、前記〔 2 8 〕に記載の方法。

〔 3 0 〕 バーコード化された核酸分子の標的結合領域の相補体を、複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子の標的結合領域とハイブリダイズさせるステップが、バーコード化された核酸分子の標的結合領域の相補体と、複数のバーコード化された核酸分子のうちの異なるバーコード化された核酸分子の標的結合領域との分子間ハイブリダイゼーションを含む、前記〔 1 〕 ~ 〔 2 9 〕のいずれか 1 項に記載の方法。

〔 3 1 〕 第 2 の分子標識の配列が、第 1 の分子標識の配列と異なり、第 2 の分子標識が、第 1 の分子標識の相補体ではない、前記〔 3 0 〕に記載の方法。

〔 3 2 〕 試料が、単一細胞、複数の細胞、複数の単一細胞、組織、腫瘍試料、またはこれらの任意の組合せを含む、前記〔 1 〕 ~ 〔 3 1 〕のいずれか 1 項に記載の方法。

〔 3 3 〕 単一細胞が、免疫細胞または循環腫瘍細胞を含み、免疫細胞が B 細胞または T 細胞であってもよい、前記〔 3 2 〕に記載の方法。

〔 3 4 〕 標的特定のプライマーが免疫受容体に特異的にハイブリダイズし、任意に、標的特定のプライマーが、免疫受容体の定常領域、免疫受容体の可変領域、免疫受容体の多様性領域、免疫受容体の可変領域と多様性領域の接合部、またはこれらの組合せに特異的にハイブリダイズする、前記〔 1 〕 ~ 〔 3 3 〕のいずれか 1 項に記載の方法。

〔 3 5 〕 免疫受容体が、T 細胞受容体 (T C R) および / または B 細胞受容体 (B C R) 受容体であり、T C R が、T C R アルファ鎖、T C R ベータ鎖、T C R ガンマ鎖、T C R デルタ鎖、またはこれらの任意の組合せを含んでもよく、さらに、B C R 受容体が、B C R 重鎖および / または B C R 軽鎖を含んでもよい、前記〔 3 4 〕に記載の方法。

〔 3 6 〕 複数のバーコード化された核酸分子の 3' 末端を伸長させるステップが、5' から 3' へのエキソヌクレアーゼ活性および 3' から 5' へのエキソヌクレアーゼ活性のうち少なくとも 1 つを欠く DNA ポリメラーゼを使用して、複数のバーコード化された核酸分子の 3' 末端を伸長させるステップを含み、DNA ポリメラーゼが、K l e n o w 断片を含んでもよい、前記〔 1 〕 ~ 〔 3 5 〕のいずれか 1 項に記載の方法。

〔 3 7 〕 第 2 の複数のバーコード化されたアンプリコンまたはその産物の配列情報を得るステップを含み、配列情報を得るステップが、シーケンシングアダプターを第 2 の複数のバーコード化されたアンプリコンまたはその産物に結合させるステップを含んでもよい、

10

20

30

40

50

前記〔1〕～〔36〕のいずれか1項に記載の方法。

〔38〕複数の伸長産物を増幅させるステップが、シーケンシングプライマーおよび/またはシーケンシングアダプターの結合部位の配列、その相補配列、および/またはそれら的一部分を、複数の伸長産物に付加するステップであって；シーケンシングアダプターが、P5配列、P7配列、それらの相補配列、またはそれら的一部分を含んでもよく；さらに、シーケンシングプライマーが、リード1シーケンシングプライマー、リード2シーケンシングプライマー、それらの相補配列、またはそれら的一部分を含んでもよい、ステップを含む、前記〔1〕～〔37〕のいずれか1項に記載の方法。

〔39〕第2の複数のバーコード化されたアンプリコンまたはその産物の配列情報を得るステップが、

第2の複数のバーコード化されたアンプリコンまたはその産物の複数のシーケンシングリードを含むシーケンシングデータを得るステップであって、複数のシーケンシングリードのそれぞれが、(1)細胞標識配列、(2)分子標識配列、および/または(3)核酸標的の部分配列を含む、ステップ

を含む、前記〔37〕～〔38〕のいずれか1項に記載の方法。

〔40〕試料の単一細胞を示す、それぞれ固有の細胞標識配列について、

核酸標的の複数のシーケンシングリードのそれぞれをアライメントさせて、核酸標的のアライメントされた配列を生成するステップを含む、前記〔39〕に記載の方法。

〔41〕核酸標的のアライメントされた配列が、核酸標的のcDNA配列の少なくとも50%、核酸標的のcDNA配列の少なくとも70%、核酸標的のcDNA配列の少なくとも90%、または核酸標的のcDNA配列の完全長を含む、前記〔40〕に記載の方法。

〔42〕核酸標的が、免疫受容体であり、免疫受容体が、BCR軽鎖、BCR重鎖、TCRアルファ鎖、TCRベータ鎖、TCRガンマ鎖、TCRデルタ鎖、またはこれらの任意の組合せを含んでもよい、前記〔1〕～〔41〕のいずれか1項に記載の方法。

〔43〕核酸標的のアライメントされた配列が、相補性決定領域1(CDR1)、相補性決定領域2(CDR2)、相補性決定領域3(CDR3)、可変領域、可変領域の完全長、またはこれらの組合せを含む、前記〔42〕に記載の方法。

〔44〕核酸標的のアライメントされた配列が、可変領域、多様性領域、可変領域、多様性領域および/または定常領域の接合部、またはこれらの任意の組合せを含む、前記〔42〕～〔43〕のいずれか1項に記載の方法。

〔45〕断片化、タグメンテーション、またはその両方を含まない、前記〔42〕～〔44〕のいずれか1項に記載の方法。

〔46〕配列情報を得るステップが、単一細胞のBCR軽鎖およびBCR重鎖の配列情報を得るステップを含み、BCR軽鎖およびBCR重鎖の配列情報が、BCR軽鎖および/またはBCR重鎖の相補性決定領域1(CDR1)、CDR2、CDR3、またはこれらの任意の組合せの配列を含んでもよい、前記〔37〕～〔45〕のいずれか1項に記載の方法。

〔47〕得られた配列情報に基づいて、単一細胞のBCR軽鎖とBCR重鎖を対合させるステップを含む、前記〔46〕に記載の方法。

〔48〕試料が、複数の単一細胞を含み、方法が、得られた配列情報に基づいて、前記単一細胞の少なくとも50%のBCR軽鎖とBCR重鎖を対合させるステップを含む、前記〔46〕～〔47〕のいずれか1項に記載の方法。

〔49〕配列情報を得るステップが、単一細胞のTCRアルファ鎖およびTCRベータ鎖の配列情報を得るステップを含み、TCRアルファ鎖およびTCRベータ鎖の配列情報が、TCRアルファ鎖および/またはTCRベータ鎖の相補性決定領域1(CDR1)、CDR2、CDR3、またはこれらの任意の組合せの配列を含んでもよい、前記〔37〕～〔48〕のいずれか1項に記載の方法。

〔50〕得られた配列情報に基づいて、単一細胞のTCRアルファ鎖とTCRベータ鎖を対合させるステップを含む、前記〔49〕に記載の方法。

〔51〕試料が、複数の単一細胞を含み、方法が、得られた配列情報に基づいて、前記単

10

20

30

40

50

一細胞の少なくとも50%のTCRアルファ鎖とTCRベータ鎖を対合させるステップを含む、前記〔49〕～〔50〕のいずれか1項に記載の方法。

〔52〕配列情報を得るステップが、単一細胞のTCRガンマ鎖およびTCRデルタ鎖の配列情報を得るステップを含む、前記〔37〕～〔51〕のいずれか1項に記載の方法。

〔53〕TCRガンマ鎖およびTCRデルタ鎖の配列情報が、TCRガンマ鎖および/またはTCRデルタ鎖の、相補性決定領域1(CDR1)、CDR2、CDR3、またはこれらの任意の組合せの配列を含む、前記〔52〕に記載の方法。

〔54〕得られた配列情報に基づいて、単一細胞のTCRガンマ鎖とTCRデルタ鎖を対合させるステップを含む、前記〔52〕～〔53〕のいずれか1項に記載の方法。

〔55〕試料が、複数の単一細胞を含み、方法が、得られた配列情報に基づいて、前記単一細胞の少なくとも50%のTCRガンマ鎖とTCRデルタ鎖を対合させるステップを含む、前記〔52〕～〔54〕のいずれか1項に記載の方法。

〔56〕標的結合領域の相補体が、標的結合領域の逆相補配列または標的結合領域の相補配列を含む、前記〔1〕～〔55〕のいずれか1項に記載の方法。

〔57〕分子標識の相補体が、分子標識の逆相補配列または分子標識の相補配列を含む、前記〔1〕～〔56〕のいずれか1項に記載の方法。

〔58〕複数のバーコード化された核酸分子が、バーコード化されたデオキシリボ核酸(DNA)分子、バーコード化されたリボ核酸(RNA)分子、またはこれらの組合せを含む、前記〔1〕～〔57〕のいずれか1項に記載の方法。

〔59〕核酸標的が、核酸分子を含み、核酸分子が、リボ核酸(RNA)、メッセンジャーRNA(mRNA)、マイクロRNA、低分子干渉RNA(siRNA)、RNA分解産物、ポリ(A)テールを含むRNA、またはこれらの任意の組合せを含んでもよく、さらに、mRNAが免疫受容体をコードしてもよい、前記〔1〕～〔58〕のいずれか1項に記載の方法。

〔60〕核酸標的が、細胞成分結合試薬を含む、および/または核酸分子が、細胞成分結合試薬に会合している、前記〔59〕に記載の方法。

〔61〕核酸分子と細胞成分結合試薬を解離させるステップを含む、前記〔60〕に記載の方法。

〔62〕複数のオリゴヌクレオチドバーコードのうちの少なくとも10個が、異なる分子標識配列を含む、前記〔1〕～〔61〕のいずれか1項に記載の方法。

〔63〕複数のオリゴヌクレオチドバーコードの各分子標識が、少なくとも6個のヌクレオチドを含む、前記〔1〕～〔62〕のいずれか1項に記載の方法。

〔64〕複数のオリゴヌクレオチドバーコードが、固体支持体と会合している、前記〔1〕～〔63〕のいずれか1項に記載の方法。

〔65〕同じ固体支持体に会合した複数のオリゴヌクレオチドバーコードがそれぞれ、同一の試料標識を含む、前記〔64〕に記載の方法。

〔66〕複数のオリゴヌクレオチドバーコードの各試料標識が、少なくとも6個のヌクレオチドを含む、前記〔65〕に記載の方法。

〔67〕複数のオリゴヌクレオチドバーコードがそれぞれ、細胞標識を含む、前記〔1〕～〔66〕のいずれか1項に記載の方法。

〔68〕複数のオリゴヌクレオチドバーコードの各細胞標識が、少なくとも6個のヌクレオチドを含む、前記〔67〕に記載の方法。

〔69〕同じ固体支持体に会合しているオリゴヌクレオチドバーコードが、同じ細胞標識を含む、前記〔67〕～〔68〕のいずれか1項に記載の方法。

〔70〕異なる固体支持体に会合しているオリゴヌクレオチドバーコードが、異なる細胞標識を含む、前記〔67〕～〔69〕のいずれか1項に記載の方法。

〔71〕複数の伸長したバーコード化された核酸分子がそれぞれ、細胞標識および該細胞標識の相補体を含む、前記〔67〕～〔70〕のいずれか1項に記載の方法。

〔72〕細胞標識の相補体が、細胞標識の逆相補配列または細胞標識の相補配列を含む、前記〔71〕に記載の方法。

10

20

30

40

50

〔 7 3 〕エチレングリコール、ポリエチレングリコール、1, 2 - プロパンジオール、ジメチルスルホキシド (D M S O)、グリセロール、ホルムアミド、7 - デアザ - G T P、アセトアミド、塩化テトラメチルアンモニウム塩、ベタイン、またはこれらの任意の組合せのうちの1つまたは複数の存在下で、核酸標的のコピーにハイブリダイズした複数のオリゴヌクレオチドバーコードを伸長させるステップを含む、前記〔 1 〕 ~〔 7 2 〕のいずれか1項に記載の方法。

〔 7 4 〕固体支持体が、合成粒子、平面表面、またはこれらの組合せを含む、前記〔 6 4 〕 ~〔 7 3 〕のいずれか1項に記載の方法。

〔 7 5 〕試料が、単一細胞を含み、方法が、複数のオリゴヌクレオチドバーコードを含む合成粒子を、試料中の単一細胞と会合させるステップを含む、前記〔 1 〕 ~〔 7 4 〕のいずれか1項に記載の方法。

10

〔 7 6 〕合成粒子を単一細胞と会合させた後に、単一細胞を溶解させるステップを含み、単一細胞を溶解させるステップが、試料を加熱するステップ、試料を界面活性剤と接触させるステップ、試料のpHを変更するステップ、またはこれらの任意の組合せを含んでもよい、前記〔 7 5 〕に記載の方法。

〔 7 7 〕合成粒子と単一細胞が、同じ区画中にあり、区画がウェルまたは液滴であってもよい、前記〔 7 5 〕 ~〔 7 6 〕のいずれか1項に記載の方法。

〔 7 8 〕複数のオリゴヌクレオチドバーコードのうちの少なくとも1つが、合成粒子上に固定化されているかもしくは部分的に固定化されている、または複数のオリゴヌクレオチドバーコードのうちの少なくとも1つが、合成粒子内に封入されているかもしくは部分的に封入されている、前記〔 7 5 〕 ~〔 7 7 〕のいずれか1項に記載の方法。

20

〔 7 9 〕合成粒子が、崩壊可能であり、任意に、崩壊可能なハイドロゲル粒子である、前記〔 7 5 〕 ~〔 7 8 〕のいずれか1項に記載の方法。

〔 8 0 〕合成粒子が、ビーズを含み、ビーズが、セファロースビーズ、ストレプトアビジンビーズ、アガロースビーズ、磁性ビーズ、コンジュゲートビーズ、プロテインAコンジュゲートビーズ、プロテインGコンジュゲートビーズ、プロテインA / Gコンジュゲートビーズ、プロテインLコンジュゲートビーズ、オリゴ (d T) コンジュゲートビーズ、シリカビーズ、シリカ様ビーズ、抗ピオチンマイクロビーズ、抗蛍光色素マイクロビーズ、もしくはこれらの任意の組合せを含んでもよい、前記〔 7 5 〕 ~〔 7 9 〕のいずれか1項に記載の方法。

30

〔 8 1 〕合成粒子が、ポリジメチルシロキサン (P D M S)、ポリスチレン、ガラス、ポリプロピレン、アガロース、ゼラチン、ハイドロゲル、常磁性体、セラミック、プラスチック、ガラス、メチルスチレン、アクリルポリマー、チタン、ラテックス、セファロース、セルロース、ナイロン、シリコーン、およびこれらの任意の組合せからなる群から選択される材料を含む、前記〔 7 5 〕 ~〔 8 0 〕のいずれか1項に記載の方法。

〔 8 2 〕複数のオリゴヌクレオチドバーコードのそれぞれが、リンカー官能基を含み、合成粒子が、固体支持体官能基を含み、支持体官能基およびリンカー官能基が、互いに会合しており、リンカー官能基および支持体官能基が、それぞれに、C 6、ピオチン、ストレプトアビジン、第一級アミン、アルデヒド、ケトン、およびこれらの任意の組合せからなる群から選択されてもよい、前記〔 7 5 〕 ~〔 8 1 〕のいずれか1項に記載の方法。

40

〔 8 3 〕複数のオリゴヌクレオチドバーコードであって、複数のオリゴヌクレオチドバーコードのそれぞれが、分子標識および標的結合領域を含み、複数のオリゴヌクレオチドバーコードのうちの少なくとも10個が、異なる分子標識配列を含む、複数のオリゴヌクレオチドバーコードと、

逆転写酵素と、

標的結合領域またはその一部分を含むテンプレートスイッチングオリゴヌクレオチドと、

5'から3'へのエキソヌクレアーゼ活性および3'から5'へのエキソヌクレアーゼ活性のうちの少なくとも1つを欠くDNAポリメラーゼと、

を含む、キット。

50

〔 8 4 〕逆転写酵素が、ウイルス逆転写酵素を含み、ウイルス逆転写酵素が、マウス白血病ウイルス (M L V) 逆転写酵素またはモロニーマウス白血病ウイルス (M M L V) 逆転写酵素であってもよい、前記〔 8 3 〕に記載のキット。

〔 8 5 〕テンプレートスイッチオリゴヌクレオチドが、1つまたは複数の3'リボヌクレオチド、任意に、3つの3'リボヌクレオチドを含む、前記〔 8 3 〕～〔 8 4 〕のいずれか1項に記載のキット。

〔 8 6 〕3'リボヌクレオチドが、グアニンを含む、前記〔 8 5 〕に記載のキット。

〔 8 7 〕エチレングリコール、ポリエチレングリコール、1,2-プロパンジオール、ジメチルスルホキシド (D M S O)、グリセロール、ホルムアミド、7-デアザ-G T P、アセトアミド、塩化テトラメチルアンモニウム塩、ベタイン、またはこれらの任意の組合せのうちの一つまたは複数を含む、前記〔 8 3 〕～〔 8 6 〕のいずれか1項に記載のキット。

10

〔 8 8 〕複数のオリゴヌクレオチドバーコードであって、複数のオリゴヌクレオチドバーコードのそれぞれが、分子標識および標的結合領域を含み、複数のオリゴヌクレオチドバーコードのうち少なくとも10個が、異なる分子標識配列を含む、複数のオリゴヌクレオチドバーコードと、

末端デオキシヌクレオチド転移酵素と、

5'から3'へのエキソヌクレアーゼ活性および3'から5'へのエキソヌクレアーゼ活性のうち少なくとも1つを欠くDNAポリメラーゼと、を含む、キット。

20

〔 8 9 〕DNAポリメラーゼが、K l e n o w断片を含む、前記〔 8 3 〕～〔 8 8 〕のいずれか1項に記載のキット。

〔 9 0 〕緩衝液、カートリッジ、または両方を含む、前記〔 8 3 〕～〔 8 9 〕のいずれか1項に記載のキット。

〔 9 1 〕逆転写反応および/または増幅反応のための1つまたは複数の試薬を含む、前記〔 8 3 〕～〔 9 0 〕のいずれか1項に記載のキット。

〔 9 2 〕標的結合領域が、遺伝子特異的配列、オリゴ (d T) 配列、ランダム多量体、またはこれらの任意の組合せを含む、前記〔 8 3 〕～〔 9 1 〕のいずれか1項に記載のキット。

〔 9 3 〕オリゴヌクレオチドバーコードが、同一の試料標識および/または同一の細胞標識を含む、前記〔 8 3 〕～〔 9 2 〕のいずれか1項に記載のキット。

30

〔 9 4 〕複数のオリゴヌクレオチドバーコードの各試料標識および/または細胞標識が、少なくとも6個のヌクレオチドを含む、前記〔 9 3 〕に記載のキット。

〔 9 5 〕複数のオリゴヌクレオチドバーコードの各分子標識が、少なくとも6個のヌクレオチドを含む、前記〔 8 3 〕～〔 9 3 〕のいずれか1項に記載のキット。

〔 9 6 〕複数のオリゴヌクレオチドバーコードのうち少なくとも1つが、合成粒子上に固定化されているかもしくは部分的に固定化されている、または合成粒子内に封入されているかもしくは部分的に封入されている、前記〔 8 3 〕～〔 9 5 〕のいずれか1項に記載のキット。

〔 9 7 〕合成粒子が、崩壊可能であり、任意に、崩壊可能なハイドロゲル粒子である、前記〔 8 3 〕～〔 9 6 〕のいずれか1項に記載のキット。

40

〔 9 8 〕合成粒子が、ビーズを含み、ビーズが、セファロースビーズ、ストレプトアビジンビーズ、アガロースビーズ、磁性ビーズ、コンジュゲートビーズ、プロテインAコンジュゲートビーズ、プロテインGコンジュゲートビーズ、プロテインA/Gコンジュゲートビーズ、プロテインLコンジュゲートビーズ、オリゴ (d T) コンジュゲートビーズ、シリカビーズ、シリカ様ビーズ、抗ビオチンマイクロビーズ、抗蛍光色素マイクロビーズ、またはこれらの任意の組合せを含んでもよい、前記〔 8 3 〕～〔 9 7 〕のいずれか1項に記載のキット。

〔 9 9 〕合成粒子が、ポリジメチルシロキサン (P D M S)、ポリスチレン、ガラス、ポリプロピレン、アガロース、ゼラチン、ハイドロゲル、常磁性体、セラミック、プラスチ

50

ック、ガラス、メチルスチレン、アクリルポリマー、チタン、ラテックス、セファロース、セルロース、ナイロン、シリコン、およびこれらの任意の組合せからなる群から選択される材料を含む、前記〔 8 3 〕～〔 9 8 〕のいずれか 1 項に記載のキット。

〔 1 0 0 〕複数のオリゴヌクレオチドバーコードのそれぞれが、リンカー官能基を含み、合成粒子が、固体支持体官能基を含み、

支持体官能基およびリンカー官能基が、互いに会合しており、リンカー官能基および支持体官能基が、それぞれに、C 6、ビオチン、ストレプトアビジン、第一級アミン、アルデヒド、ケトン、およびこれらの任意の組合せからなる群から選択されてもよい、前記〔 8 3 〕～〔 9 9 〕のいずれか 1 項に記載のキット。

【 図 面 】

【 図 1 】

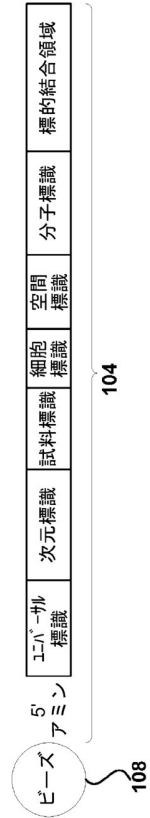


FIG. 1

【 図 2 】

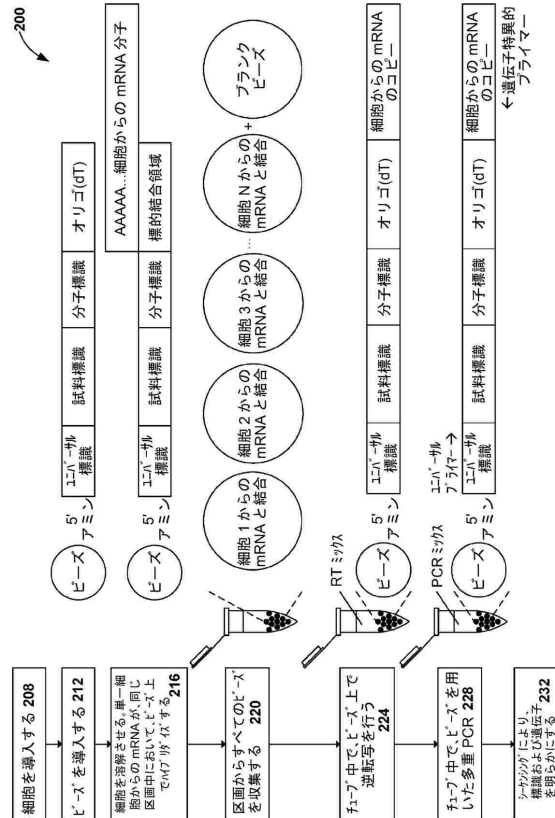


FIG. 2

10

20

30

40

50

【 図 3 】

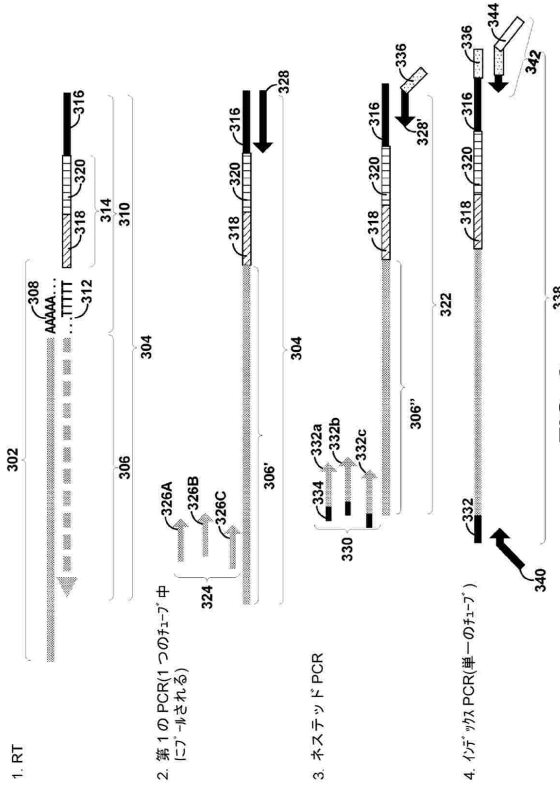


FIG. 3

【 図 4 A 】

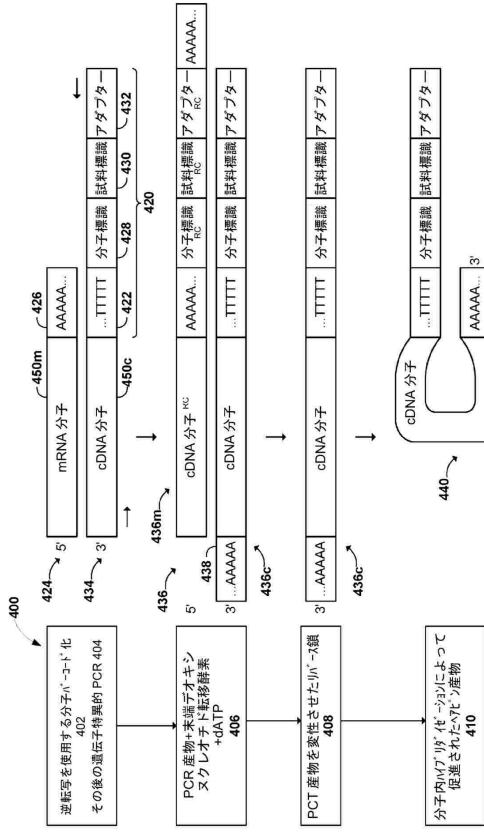


FIG. 4A

【 図 4 B 】

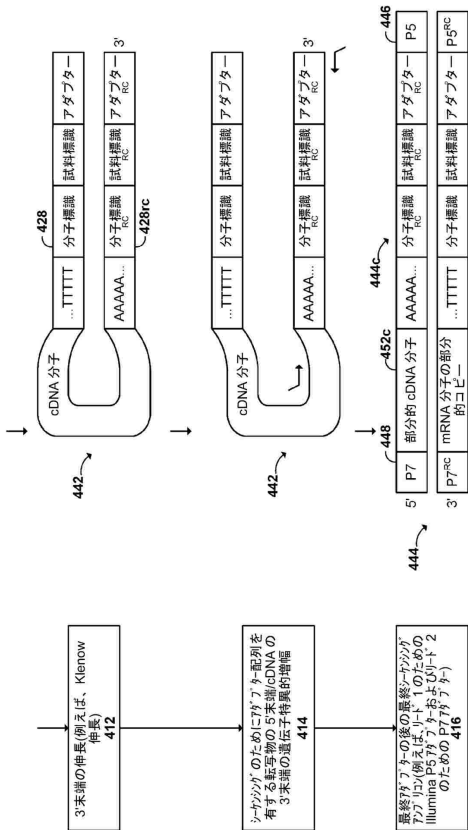


FIG. 4B

【 図 5 A 】

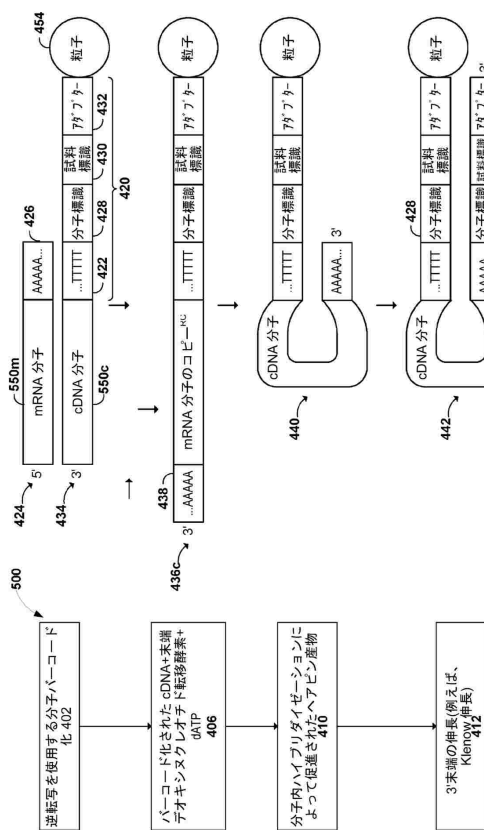


FIG. 5A

【 図 5 B 】

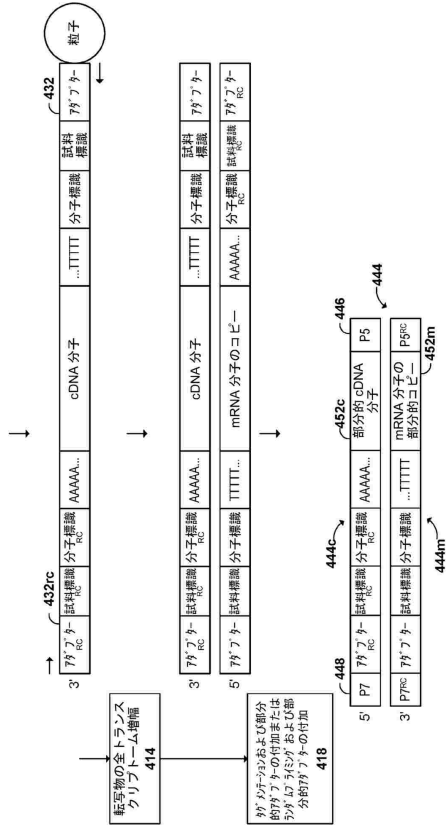


FIG. 5B

【 図 6 A 】

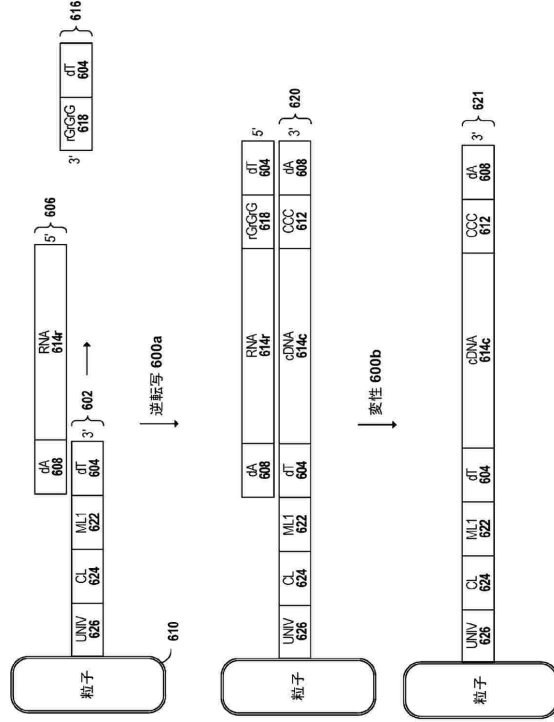


FIG. 6A

【 図 6 B 】

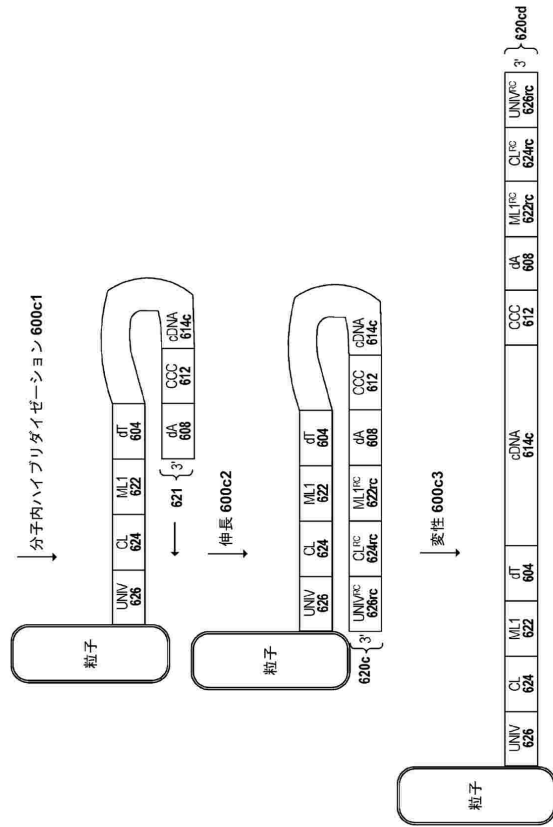


FIG. 6B

【 図 6 C 】

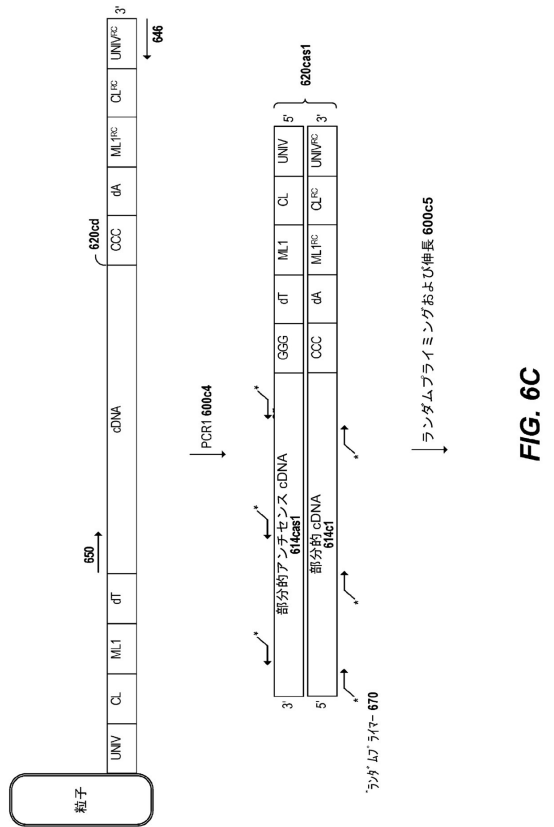


FIG. 6C

【 6 D 】

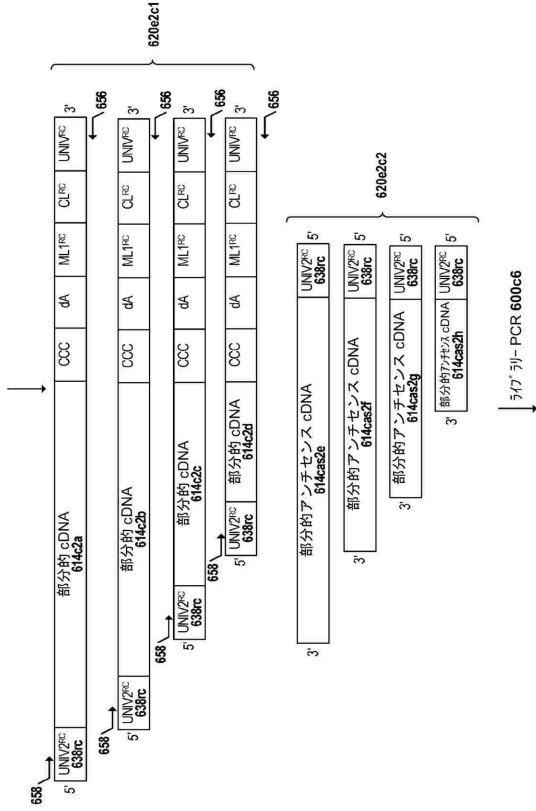


FIG. 6D

【 6 E 】

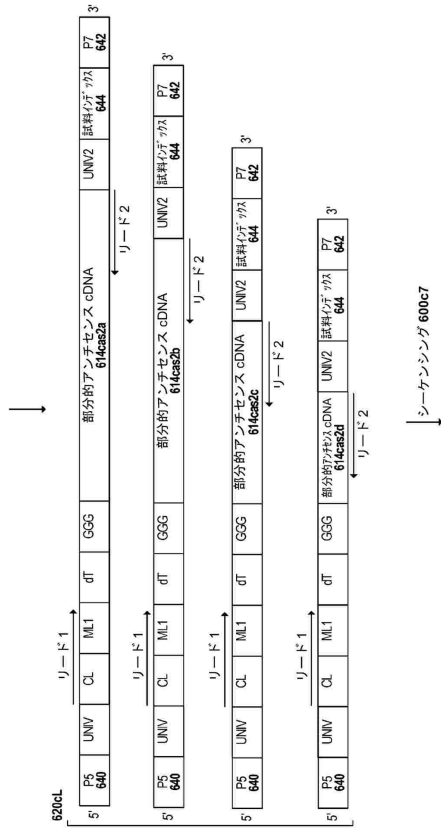


FIG. 6E

【 6 F 】

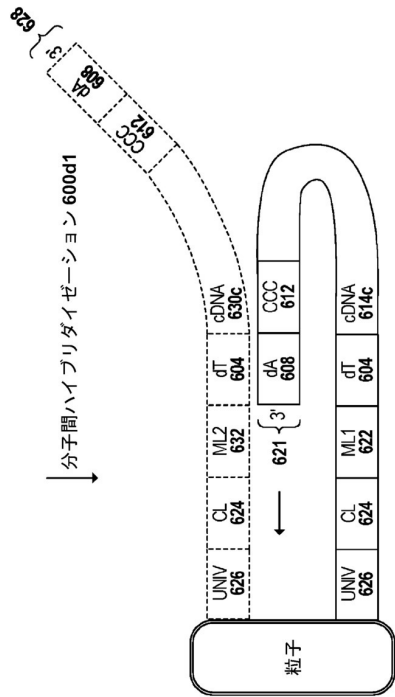


FIG. 6F

【 6 G 】

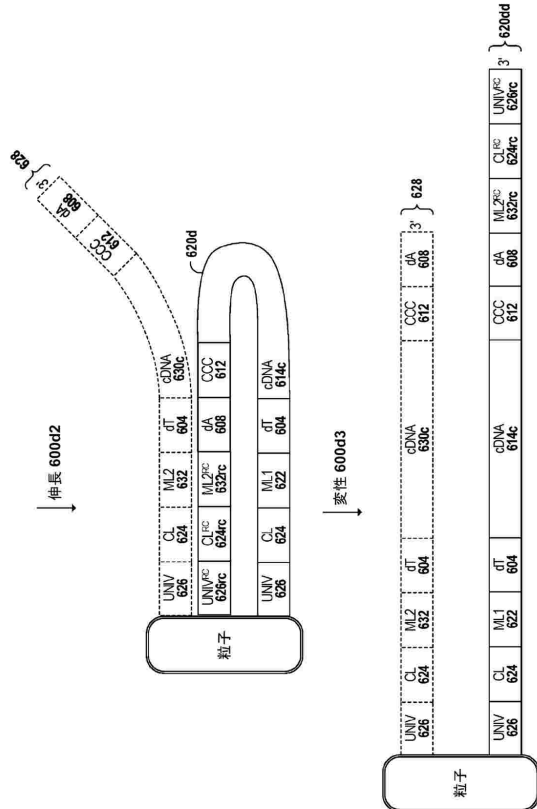


FIG. 6G

【 図 6 H 】

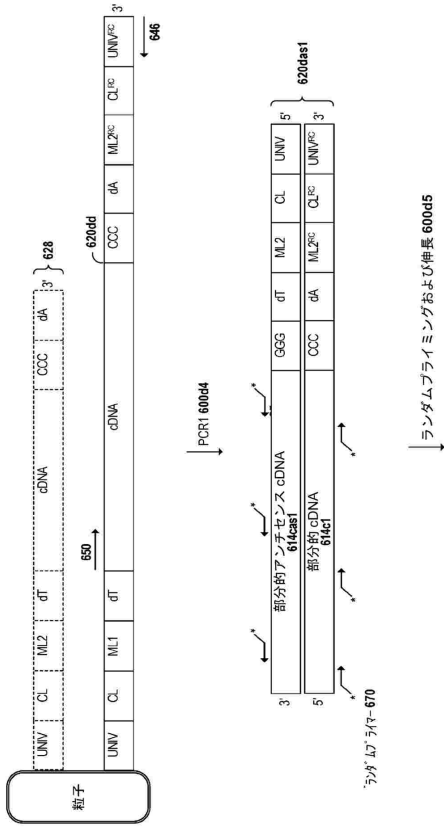


FIG. 6H

【 図 6 I 】

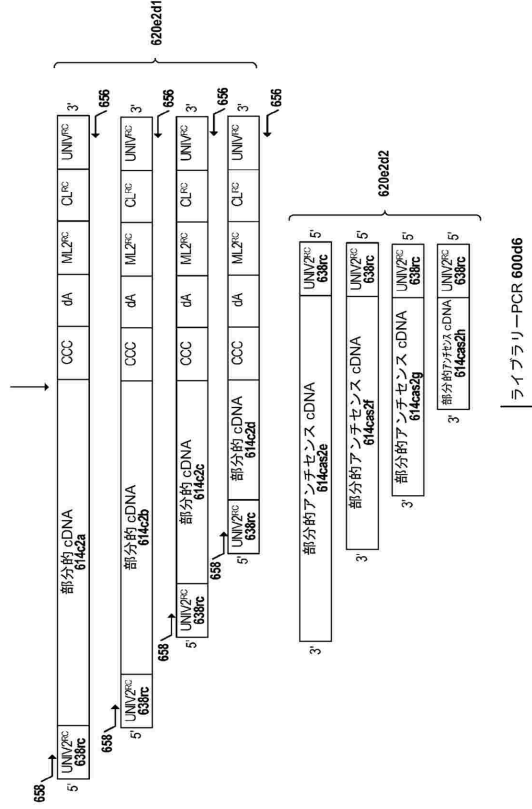


FIG. 6I

【 図 6 J 】

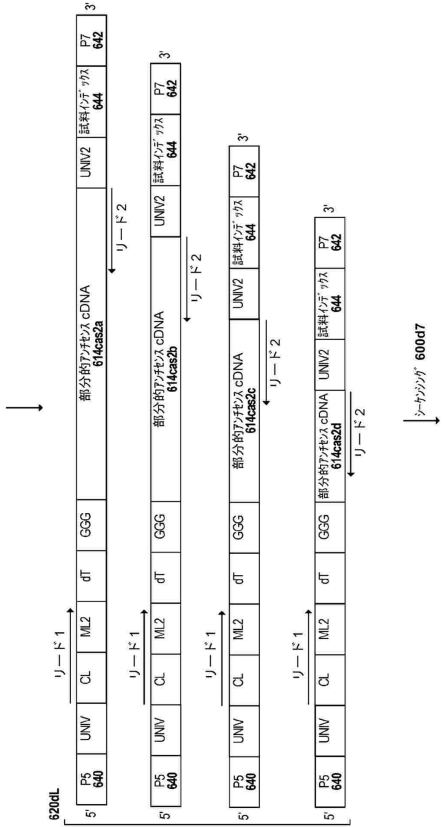


FIG. 6J

【 図 6 K 】

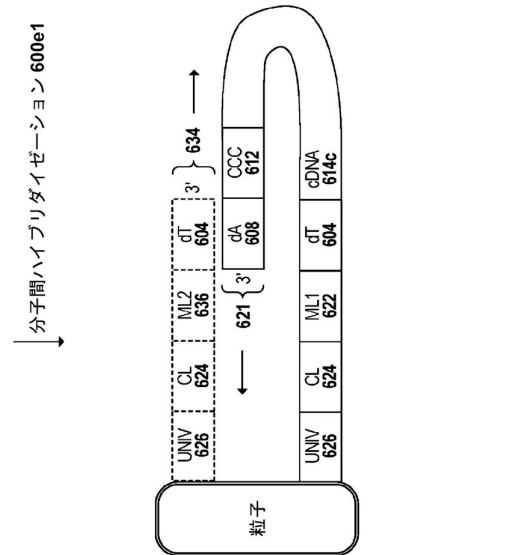


FIG. 6K

【 図 7 】

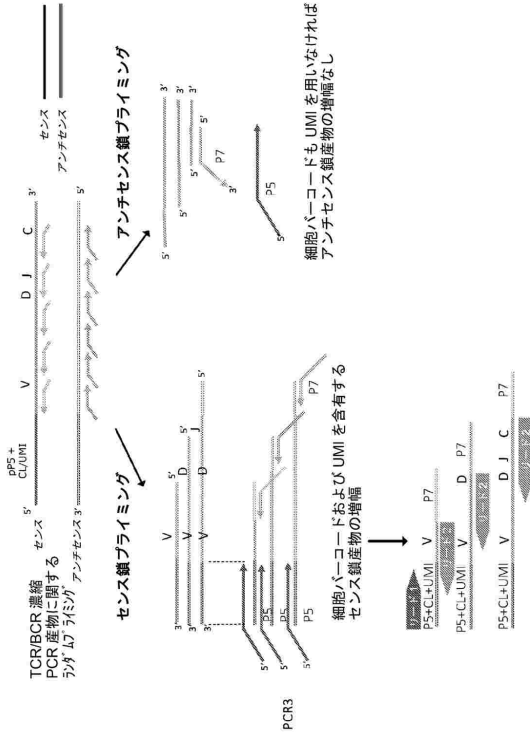


FIG. 7

【 図 8 】

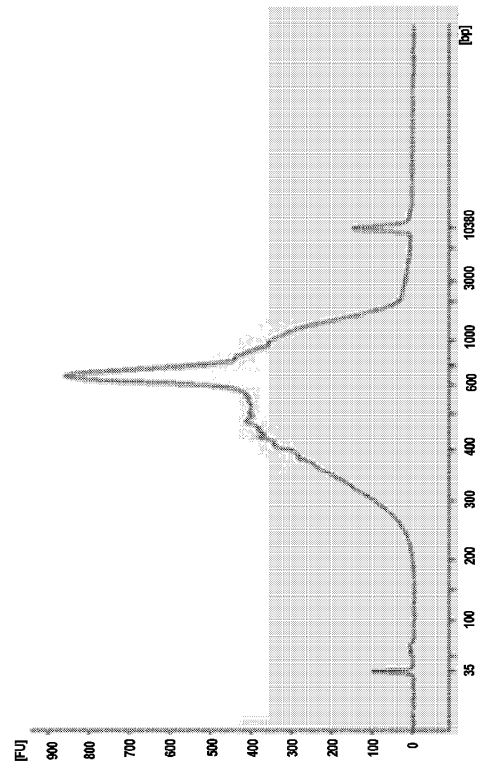


FIG. 8

【 図 9 A 】

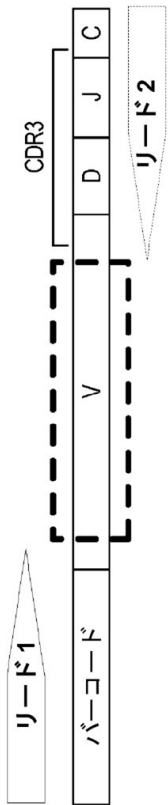


FIG. 9A

【 図 9 B 】

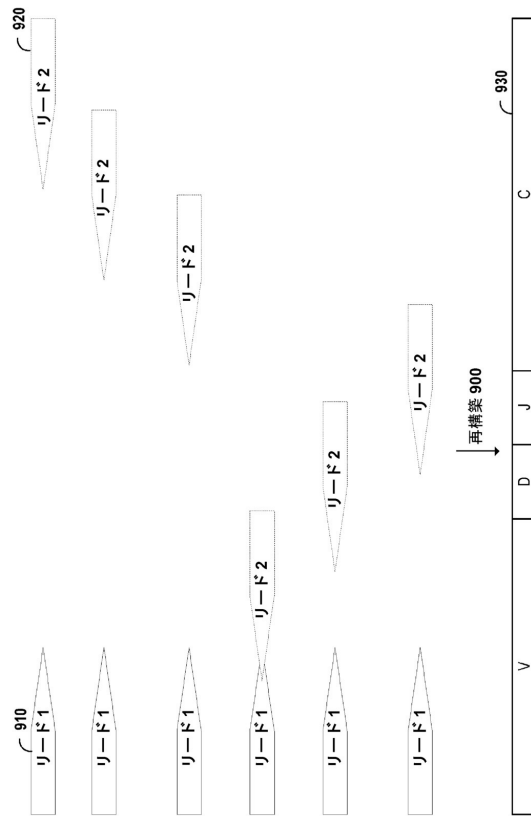


FIG. 9B

【 10 A 】

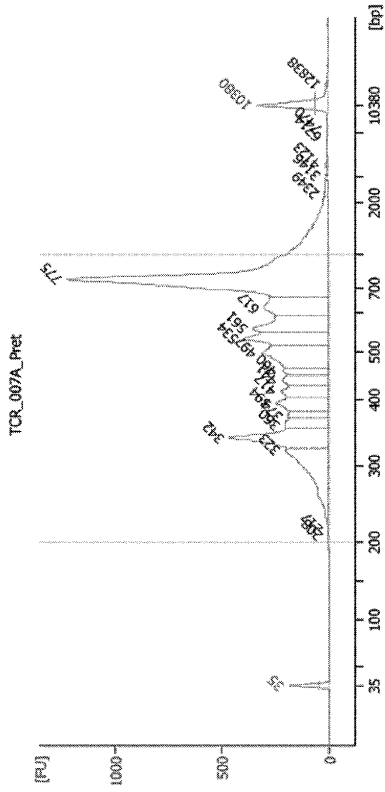
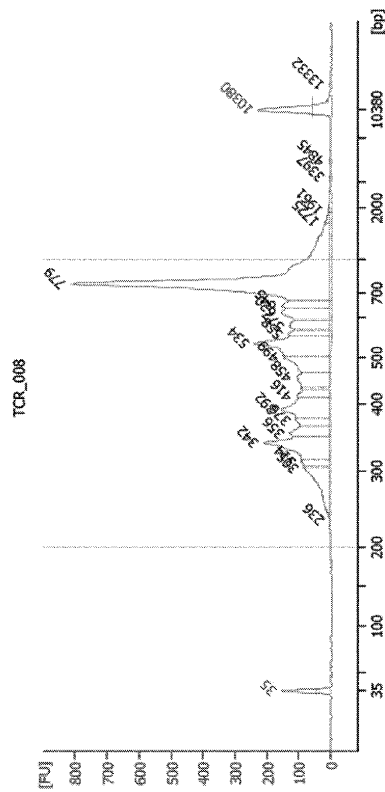


FIG. 10A

【 10 B 】



フロントページの続き

(51)国際特許分類

C 1 2 Q 1/6834(2018.01)
C 1 2 N 9/12 (2006.01)

F I

C 1 2 Q 1/6834 Z
C 1 2 N 9/12

弁理士 山崎 一夫

(74)代理人 100123777

弁理士 市川 さつき

(74)代理人 100111796

弁理士 服部 博信

(74)代理人 100123766

弁理士 松田 七重

(74)代理人 100183379

弁理士 藤代 昌彦

(72)発明者 チャン クリスティーナ

アメリカ合衆国 ニュージャージー州 0 7 4 1 7 - 1 8 8 0 フランクリン レイクス ベクトン
ドライブ 1

(72)発明者 ナカモト マーガレット

アメリカ合衆国 ニュージャージー州 0 7 4 1 7 - 1 8 8 0 フランクリン レイクス ベクトン
ドライブ 1

審査官 天野 皓己

(56)参考文献 国際公開第 2 0 1 9 / 2 1 3 2 3 7 (W O , A 1)

国際公開第 2 0 1 8 / 2 1 8 2 2 2 (W O , A 1)

国際公開第 2 0 2 0 / 0 7 2 3 8 0 (W O , A 1)

(58)調査した分野 (Int.Cl., D B 名)

C 1 2 Q 1 / 0 0 - 3 / 0 0

C 1 2 N 9 / 1 2

C A p l u s / M E D L I N E / E M B A S E / B I O S I S (S T N)