

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号

特表2024-540168

(P2024-540168A)

(43)公表日 令和6年10月31日(2024.10.31)

(51)国際特許分類		F I		テーマコード (参考)	
C 1 2 Q	1/6869(2018.01)	C 1 2 Q	1/6869	Z Z N A	4 B 0 6 3
C 1 2 Q	1/6876(2018.01)	C 1 2 Q	1/6876	Z	
C 1 2 N	9/00 (2006.01)	C 1 2 N	9/00		
C 1 2 Q	1/6844(2018.01)	C 1 2 Q	1/6844	Z	

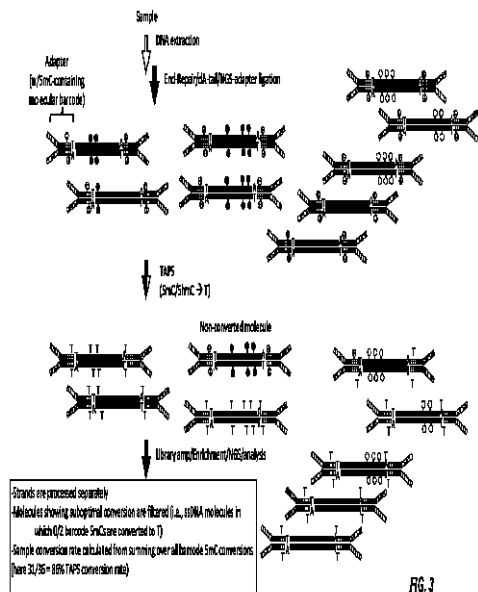
審査請求 未請求 予備審査請求 未請求 (全103頁)

(21)出願番号	特願2024-525613(P2024-525613)	(71)出願人	515059083
(86)(22)出願日	令和4年11月2日(2022.11.2)		ガーダント ヘルス, インコーポレイテッド
(85)翻訳文提出日	令和6年6月27日(2024.6.27)		アメリカ合衆国 カリフォルニア 9 4 3
(86)国際出願番号	PCT/US2022/079173		0 4, パロ アルト, ハノーバー ストリート 3 1 0 0
(87)国際公開番号	WO2023/081722	(74)代理人	100078282
(87)国際公開日	令和5年5月11日(2023.5.11)		弁理士 山本 秀策
(31)優先権主張番号	63/274,868	(74)代理人	100113413
(32)優先日	令和3年11月2日(2021.11.2)		弁理士 森下 夏樹
(33)優先権主張国・地域又は機関	米国(US)	(74)代理人	100181674
(81)指定国・地域	AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,最終頁に続く	(74)代理人	100181641
			弁理士 飯田 貴敏
		(74)代理人	100181641
			弁理士 石川 大輔
		(74)代理人	230113332
			最終頁に続く

(54)【発明の名称】 品質管理方法

(57)【要約】

本明細書に提供されるのは、塩基対変換手順を用いて、サンプル中のDNA中の修飾ヌクレオシドの偽陰性及び／又は偽陽性検出を監視するための方法である。方法では、DNAにライゲートされているアダプター内の既知のヌクレオシド同一性及び既知の修飾状態を有するヌクレオシドが使用される。特定の態様では、本開示は、こうした方法の品質管理を改善するための方法に関する。



10

20

【特許請求の範囲】

【請求項 1】

サンプル中の DNA 中の修飾ヌクレオシドの偽陰性及び／又は偽陽性検出を監視するための品質管理方法であって、

(a) 前記 DNA をオリゴヌクレオチドアダプターにライゲートすることであって、前記アダプターは修飾ヌクレオシドを含む品質管理ヌクレオシドを含み、前記品質管理ヌクレオシドは、前記 DNA 中の検出対象の修飾ヌクレオシドに対して同じヌクレオシド同一性及び同じ又は異なる修飾状態を有し、前記品質管理ヌクレオシドの前記修飾状態は既知である、ライゲートすることと；

(b) 前記適応した DNA、又はそのサブサンプルに、前記ヌクレオシドの前記修飾状態に応じて、前記品質管理ヌクレオシドの塩基対特異性を変更するか又は前記品質管理ヌクレオシドの塩基対特異性を変更しない変換手順であって、

(i) 前記アダプター内の品質管理ヌクレオシドと同じヌクレオシド同一性及び修飾状態を有する適応した DNA ヌクレオシドの塩基対特異性を変更し、且つ前記アダプター内の品質管理ヌクレオシドと同じヌクレオシド同一性であるが異なる修飾状態を有する適応した DNA ヌクレオシドの塩基対特異性を変更しない；及び／又は

(ii) 前記アダプター内の品質管理ヌクレオシドと同じヌクレオシド同一性及び修飾状態を有する適応した DNA ヌクレオシドの塩基対特異性を変更せず、且つ前記アダプター内の品質管理ヌクレオシドと同じ対同一性を有するが、異なる修飾状態を有する適応した DNA ヌクレオシドの塩基対特異性を変更する

ように選択される変換手順を施すことと；

(c) 変換ステップ (b) 後、前記適応した DNA を配列決定することと；

(d) ステップ (c) において得られた配列データを使用して、前記アダプター内の前記品質管理ヌクレオシドの塩基対特異性変換を判定することと；

(e) 前記アダプター内の前記品質管理ヌクレオシドの前記塩基対特異性変換を変換ステップ (b) のための品質管理尺度として使用することと、を含み、ここでステップ (b) (i) の変換手順後のアダプター品質管理ヌクレオシドの準最適な変換及び／又はステップ (b) (ii) の変換手順後のアダプター品質管理ヌクレオシドの誤った変換は、前記 DNA サンプル中の修飾ヌクレオシドの偽陰性及び／又は偽陽性検出を予測する、品質管理方法。

【請求項 2】

前記変換手順が、前記アダプター内の修飾された品質管理ヌクレオシドの前記塩基対特異性を変更するが、同じヌクレオシド同一性を有するが異なる修飾状態及び／又は非修飾を有する DNA サンプルのヌクレオシドの塩基対特異性を変更しないように選択され；前記修飾された品質管理ヌクレオシドの準最適な変換が、前記変換手順への曝露時、前記品質管理ヌクレオシドと同じヌクレオシド同一性及び修飾状態、又は異なる修飾状態及び塩基対特異性の同じ変化を有する DNA サンプルヌクレオシドの偽陰性検出を予測する、請求項 1 に記載の方法。

【請求項 3】

前記アダプター内の前記品質管理ヌクレオシドが、5-メチルシトシン (5 m C) 及び／又は 5-ヒドロキシメチルシトシン (5 h m C) を含む、請求項 1 又は請求項 2 に記載の方法。

【請求項 4】

前記変換手順が、置換されたボラン還元剤による T e t 補助変換を含み、任意選択的に、前記置換されたボラン還元剤が、2-ピコリンボラン、ボランピリジン、t e r t -ブチルアミンボラン、又はアンモニアボランである、請求項 3 に記載の方法。

【請求項 5】

サンプル中の DNA 中の修飾ヌクレオシドの偽陰性及び／又は偽陽性検出を監視するための品質管理方法であって、

(a) 前記 DNA をオリゴヌクレオチドアダプターにライゲートすることであって、前

10

20

30

40

50

記アダプターは修飾ヌクレオシドを含む品質管理ヌクレオシドを含み、前記品質管理ヌクレオシドは、前記DNA中の検出対象の修飾ヌクレオシドに対して同じヌクレオシド同一性及び同じ又は異なる修飾状態を有し、前記品質管理ヌクレオシドの前記修飾状態は既知である、ライゲートすることと；

(b) 前記適応したDNA、又はそのサブサンプルに、前記ヌクレオシドの前記修飾状態に応じて、前記品質管理ヌクレオシドの塩基対特異性を変更するか又は前記品質管理ヌクレオシドの塩基対特異性を変更しない変換手順であって、

(i) 前記アダプター内の品質管理ヌクレオシドと同じヌクレオシド同一性及び修飾状態を有する適応したDNAヌクレオシドの塩基対特異性を変更し、且つ前記アダプター内の品質管理ヌクレオシドと同じヌクレオシド同一性であるが異なる修飾状態を有する適応したDNAヌクレオシドの塩基対特異性を変更しない；及び／又は

(ii) 前記アダプター内の品質管理ヌクレオシドと同じヌクレオシド同一性及び修飾状態を有する適応したDNAヌクレオシドの塩基対特異性を変更せず、且つ前記アダプター内の品質管理ヌクレオシドと同じ対同一性を有するが、異なる修飾状態を有する適応したDNAヌクレオシドの塩基対特異性を変更するように選択される変換手順を施すことと；

(c) 変換ステップ(b)後、前記適応したDNAを配列決定することと；

(d) ステップ(c)において得られた前記配列データを使用して、前記アダプター内の前記品質管理ヌクレオシドの塩基対特異性変換を判定することと；

(e) 前記アダプター内の前記品質管理ヌクレオシドの前記塩基対特異性変換を変換ステップ(b)のための品質管理尺度として使用することと、を含み、ここでステップ(b)(i)の変換手順後のアダプターの品質管理ヌクレオシドの準最適な変換及び／又はステップ(b)(ii)の変換手順後のアダプターの品質管理ヌクレオシドの誤った変換は、前記DNAサンプル中の修飾ヌクレオシドの偽陰性及び／又は偽陽性検出を予測する、品質管理方法。

【請求項6】

前記アダプターの品質管理ヌクレオシドが、非修飾ヌクレオシドを含む、請求項5に記載の方法。

【請求項7】

前記変換手順が、前記アダプター内の品質管理ヌクレオシドの前記塩基対特異性を変更するが、同じヌクレオシド同一性及び異なる修飾状態を有するDNAサンプルヌクレオシドの塩基対特異性を変更しないように選択され；前記品質管理ヌクレオシドの準最適な変換が、前記異なる修飾状態を有するDNAサンプルヌクレオシドの偽陽性検出を予測する、請求項5又は請求項6に記載の方法。

【請求項8】

前記アダプター内の前記品質管理ヌクレオシドがシトシンを含み、且つ／又は変換手順がバイサルファイト変換を含む、請求項7に記載の方法。

【請求項9】

前記アダプターの品質管理ヌクレオシドが、第1の修飾状態を有する第1の品質管理ヌクレオシド及び前記第1の修飾状態とは異なる第2の修飾状態を有する第2の品質管理ヌクレオシドを含む、請求項1～8のいずれか一項に記載の方法。

【請求項10】

前記第1の品質管理ヌクレオシドが修飾され、前記第2の品質管理ヌクレオシドが修飾されない、請求項9に記載の方法。

【請求項11】

前記第1の品質管理ヌクレオシドが修飾シトシンであり、前記第2の品質管理ヌクレオシドが非修飾シトシンである、請求項9に記載の方法。

【請求項12】

前記第1の品質管理ヌクレオシドが、5-メチルシトシン(5mC)又は5-ヒドロキシメチルシトシン(5hmC)である、請求項11に記載の方法。

10

20

30

40

50

【請求項 13】

前記変換手順が、前記第1の品質管理ヌクレオシドの前記塩基対特異性を変更するが、前記第2の品質管理ヌクレオシドを変更しないように選択される、又は前記変換手順が、前記第2の品質管理ヌクレオシドの前記塩基対特異性を変更するが、前記第1の品質管理ヌクレオシドを変更しないように選択される、請求項9～11のいずれか一項に記載の方法。

【請求項 14】

前記変換手順が、前記アダプター内の修飾された品質管理ヌクレオシドの前記塩基対特異性を変更するが、同じヌクレオシド同一性を有するが異なる修飾状態及び／又は非修飾を有するDNAサンプルヌクレオシドの前記塩基対特異性を変更しないように選択され；前記修飾された品質管理ヌクレオシドの準最適な変換が、前記変換手順への曝露時、前記品質管理ヌクレオシドと同じヌクレオシド同一性及び修飾状態、又は異なる修飾状態及び塩基対特異性の同じ変化を有するDNAサンプルヌクレオシドの偽陰性検出を予測する、請求項9～11のいずれか一項に記載の方法。

【請求項 15】

ステップ(c)において得られた前記配列データを使用して、

(i) 前記アダプター配列内の品質管理ヌクレオシドの準最適な又は誤った変換を伴う適応したDNA分子を同定することと；

(ii) ステップ(i)において同定された前記全長分子における同じヌクレオシド同一性及び修飾状態を有するヌクレオシドの準最適な又は誤った変換を推測することと、をさらに含む、請求項1～14のいずれか一項に記載の方法。

【請求項 16】

前記適応したDNA又は個別の適応したDNA分子における品質管理ヌクレオシドの変換率を判定すること、

(i) 前記変換率に依存する加重を、

(A) 前記DNAサンプル；又は

(B) 個別の適応したDNA分子

における前記修飾ヌクレオシド検出の分析に適用すること；

(ii) 修飾ヌクレオシドを検出するためのさらなる分析からの

(A) アダプター品質管理ヌクレオシドの準最適な変換、若しくは所定の品質管理閾値を下回るアダプター品質管理ヌクレオシドの変換率；及び／又は

(B) アダプター品質管理ヌクレオシドの誤った変換、若しくは所定の品質管理閾値を上回るアダプター品質管理ヌクレオシドの変換率を有するDNAサンプルを除外すること；並びに／又は

(iii) 修飾ヌクレオシドを検出するためのさらなる分析からの

(A) アダプター品質管理ヌクレオシドの準最適な変換、若しくは所定の品質管理閾値を下回るアダプター品質管理ヌクレオシドの変換率；及び／又は

(B) アダプター品質管理ヌクレオシドの誤った変換、若しくは所定の品質管理閾値を上回るアダプターの品質管理ヌクレオシドの変換率を有する適応したDNA分子を除外すること

をさらに含む、請求項1～15のいずれか一項に記載の方法。

【請求項 17】

前記サンプルから標的領域セットを捕捉することによって前記DNAを富化することをさらに含み、前記捕捉ステップは、前記ライゲートステップ(a)及び前記変換ステップ(b)の前、後又はその間になされる、請求項1～16のいずれか一項に記載の方法。

【請求項 18】

(i) ステップ(c)において得られた前記配列データを、

(A) 所定の参照配列；及び／又は

(B) 前記変換手順を受けなかった前記DNAのサブサンプルを配列決定することによって得られた配列データ

10

20

30

40

50

と比較することと；

(i i) 前記変換されたDNA配列と前記参照配列(A)又は変換されていないDNA配列データ(B)との間のポイント差異を、前記変換手順への曝露時、塩基対特異性の変化を可能にする修飾状態を有するヌクレオシドとして同定することと、をさらに含む、請求項1～17のいずれか一項に記載の方法。

【請求項19】

前記DNAが、無細胞DNA(cfDNA)、任意選択的に、試験対象から得られたcfDNAを含み、任意選択的に、前記試験対象は、がんを有するか又は有することが疑われる患者である、請求項1～18のいずれか一項に記載の方法。

【請求項20】

前記DNAサンプル中の修飾ヌクレオシドの前記検出を用いて、がん細胞若しくは腫瘍によって生成されているDNAの存在を判定若しくは予測すること、試験対象が腫瘍若しくはがんを有する確率を判定すること、又は前記対象のがん若しくは腫瘍を特徴づけることをさらに含む、請求項1～19のいずれか一項に記載の方法。

【請求項21】

前記DNAのサブサンプルが、配列決定前に前記変換手順を受けず、前記変換されたサブサンプル及び前記変換されていないサブサンプルが、異なるアダプター配列を有し、且つ前記変換されたサブサンプル及び前記変換されていないサブサンプルが、配列決定ステップ(c)のために組み換えられる、請求項1～20のいずれか一項に記載の方法。

【請求項22】

前記DNAを分析して、コピー数多型、一塩基変異、挿入、欠失、メチル化、及び／又は融合を検出することをさらに含む、請求項1～21のいずれか一項に記載の方法。

【請求項23】

エピジェネティック標的領域を前記アダプターライゲートDNAから捕捉することと、前記エピジェネティック標的領域を増幅及び配列決定することと、をさらに含む、請求項1～22のいずれか一項に記載の方法。

【請求項24】

前記捕捉されたエピジェネティック標的領域が、エピジェネティック標的領域セットを形成する、請求項23に記載の方法。

【請求項25】

前記エピジェネティック標的領域セットが、複数の型特異的なエピジェネティック標的領域を含み、且つ前記型特異的なエピジェネティック標的領域が、型特異的な差次的にメチル化された領域及び／又は型特異的な断片である、請求項24に記載の方法。

【請求項26】

前記複数の型特異的なエピジェネティック標的領域が、型特異的な低メチル化領域を含む、請求項25に記載の方法。

【請求項27】

前記サンプルが血液サンプルである、請求項1～26のいずれか一項に記載の方法。

【請求項28】

前記複数の型特異的なエピジェネティック標的領域が、免疫細胞において前記血液サンプル中に存在する非免疫細胞型と比較して高メチル化されている；
結腸において他の組織型と比較して差次的にメチル化されている；
乳房において他の組織型と比較して差次的にメチル化されている；
肝臓において他の組織型と比較して差次的にメチル化されている；
腎臓において他の組織型と比較して差次的にメチル化されている；
膵臓において他の組織型と比較して差次的にメチル化されている；
前立腺において他の組織型と比較して差次的にメチル化されている；
皮膚において他の組織型と比較して差次的にメチル化されている；
又は膀胱において他の組織型と比較して差次的にメチル化されている
標的領域を含む、請求項25～27のいずれか一項に記載の方法。

10

20

30

40

50

【請求項 29】

前記複数の型特異的なエピジェネティック標的領域が、非免疫血球において前記サンプル中の異なる細胞型若しくは組織型における前記標的領域の前記メチル化レベルと比較して；免疫細胞に特異的な断片において前記血液サンプル中に存在する非免疫細胞型と比較して；又は結腸、肺、乳房、肝臓、腎臓、膵臓、前立腺、皮膚、若しくは膀胱に特異的な断片において他の組織型と比較して、低メチル化されている標的領域を含む、請求項 25～28 のいずれか一項に記載の方法。

【請求項 30】

前記型特異的なエピジェネティック標的領域の起源である少なくとも 1 つの細胞型又は組織型を同定することをさらに含む、請求項 23～29 のいずれか一項に記載の方法。

10

【請求項 31】

ある細胞型又は組織型に起源を持つ型特異的なエピジェネティック標的領域のレベルが決定される、請求項 23～30 のいずれか一項に記載の方法。

【請求項 32】

免疫細胞、非免疫血球、結腸、肺、乳房、肝臓、腎臓、前立腺、皮膚、膀胱、又は膵臓に起源を持つ型特異的なエピジェネティック標的領域のレベルが決定される、請求項 31 に記載の方法。

【請求項 33】

前記型特異的なエピジェネティック標的領域が、細胞型特異的、組織型特異的、及び／又はがんタイプ特異的なエピジェネティック標的領域を含む、請求項 32 に記載の方法。

20

【請求項 34】

前記血液サンプルが、少なくとも DNA のエピジェネティック標的領域セットを捕捉する前に分画される、請求項 27～33 のいずれか一項に記載の方法。

【請求項 35】

前記サンプル又はその一定分量を、第 1 の分割されたサブサンプル及び第 2 の分割されたサブサンプルを含む複数の分割されたサブサンプルに分割することと；

前記第 2 の分割されたサブサンプルをメチル化依存性ヌクレアーゼと接触させて、それにより前記第 2 のサブサンプル中の非特異的に分割された DNA を分解し、処理された第 2 のサブサンプルを生成すること、及び任意選択的に、前記第 1 の分割されたサブサンプルをメチル化感受性エンドヌクレアーゼと接触させて、それにより前記第 1 の分割されたサブサンプル中の非特異的に分割された DNA を分解し、処理された第 1 のサブサンプルを生成することと、

30

をさらに含み、ここで前記第 1 のサブサンプルは、シトシン修飾を前記第 2 のサブサンプルより大きい割合で有する DNA を含み、ここでエピジェネティック標的領域は、前記第 1 のサブサンプル又は前記処理された第 1 のサブサンプルの少なくとも一部分から捕捉され、

任意選択的に、前記 1 つ以上の捕捉プローブと接触されている前記対象からの前記 DNA は、前記第 1 の分割されたサブサンプル、前記処理された第 1 のサブサンプル、前記第 2 の分割されたサブサンプル、及び／又は前記処理された第 2 のサブサンプルからの DNA を含む、請求項 1～34 のいずれか一項に記載の方法。

40

【請求項 36】

前記シトシン修飾が、メチル化であり、任意選択的に、前記シトシン修飾が、シトシンの 5 位でのメチル化である、請求項 35 に記載の方法。

【請求項 37】

前記第 1 のサブサンプルが、メチル化感受性エンドヌクレアーゼと接触される、請求項 35 又は請求項 36 に記載の方法。

【請求項 38】

前記メチル化感受性エンドヌクレアーゼが、非メチル化 CpG 配列を切断する、請求項 37 に記載の方法。

【請求項 39】

50

前記メチル化感受性エンドヌクレアーゼが、AatII、AccII、AciI、Aor13HI、Aor15HI、BspT104I、BssHII、BstUI、Cfr1OI、ClaI、CpoI、Eco52I、HaeII、HapII、HhaI、Hin6I、HpaII、HpyCH4IV、MluI、NaeI、NotI、NruI、NsbI、PmaCI、Psp1406I、PvuI、SacII、SalI、SmaI、及びSnaBIの1つ以上である、請求項35～38のいずれか一項に記載の方法。

【請求項40】

前記変換手順が、バイサルファイト変換；5hmCの保護；Tet補助バイサルファイト変換；任意選択的に、2-ピコリンボラン、ボランピリジン、tert-ブチルアミンボラン、若しくはアンモニアボランである置換されたボラン還元剤によるTet補助変換；hmCの保護とその後の、任意選択的に、2-ピコリンボラン、ボランピリジン、tert-ブチルアミンボラン、若しくはアンモニアボランである置換されたボラン還元剤によるTet補助変換；hmCの保護とその後のmC及び／若しくはCの脱アミノ化；任意選択的に、2-ピコリンボラン、ボランピリジン、tert-ブチルアミンボラン、若しくはアンモニアボランである置換されたボラン還元剤による化学的補助変換；又は修飾シトシンの酵素的保護とその後の保護されていないシトシンのウラシルへの脱アミノ化を含む、請求項1～39のいずれか一項に記載の方法。

10

【請求項41】

前記置換されたボラン還元剤が、2-ピコリンボラン又はボランピリジンである、請求項40に記載の方法。

20

【請求項42】

mC及び／又はCの前記脱アミノ化が、AID／APOBECファミリーDNAデアミナーゼ酵素による処理を含む、請求項40に記載の方法。

【請求項43】

hmCの保護が、hmCのグルコシル化を含む、請求項40に記載の方法。

【請求項44】

配列多様性の存在若しくは不在を検出すること、及び／又は断片化パターンを判定することをさらに含み、ここで、品質管理ヌクレオシドの準最適な又は誤った変換を意味する品質管理ヌクレオシドを含む適応したDNAは、配列多様性の存在若しくは不在を検出し、及び／又は断片化パターンを判定することに含まれる、請求項1～43のいずれか一項に記載の方法。

30

【請求項45】

修飾ヌクレオシド、並びに変換試薬及びリガーゼのうち1つ以上である、品質管理ヌクレオシドを含むアダプターを含むキット。

【請求項46】

前記品質管理ヌクレオシドの前記塩基対特異性を変更することができる前記変換試薬を含む、請求項45に記載のキット。

【請求項47】

前記品質管理ヌクレオシドの前記塩基対特異性を実質的に変更することができない前記変換試薬を含む、請求項45に記載のキット。

40

【請求項48】

前記変換試薬が、バイサルファイト変換、酸化バイサルファイト変換、Tet補助（TAB）変換、置換されたボラン還元剤によるTet補助変換、置換されたボラン還元剤によるTet補助変換と組み合わせられたhmCの保護、APOBEC連結エピジェネティック（ACE）変換、又は核酸塩基の酵素変換において使用される変換試薬である、請求項45～47のいずれか一項に記載のキット。

【請求項49】

前記オリゴヌクレオチドアダプターが配列決定プライマー結合部位を含み、前記品質管理ヌクレオシドが前記配列決定プライマー結合部位の下流に位置する、請求項1～44のいずれか一項に記載の方法又は請求項45～48のいずれか一項に記載のキット。

50

【請求項 50】

前記アダプターを標的にするプライマーを使用して、前記DNAを増幅することをさらに含み、前記増幅ステップが、前記変換ステップ（b）と前記配列決定ステップ（c）との間になされる、請求項1～44又は49のいずれか一項に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の相互参照

本願は、あらゆる目的でその全体が参照により本明細書に援用される、2021年11月2日に出願された米国仮特許出願第63／274,868号明細書の優先権の利益を主張する。

10

【背景技術】

【0002】

アッセイでのエピジェネティック変異体又はヌクレオシド修飾の検出について解決する一塩基は、一般に、修飾ヌクレオシド又は対応する非修飾ヌクレオシドの、それらの塩基対特異性を変更するための変換を必要とする。次に、変換は、配列決定することによって検出される。そのような方法の例としては、バイサルファイト（bisulfite）及び酸化バイサルファイト（bisulfite）及びTet補助バイサルファイト（bisulfite）変換、EM-seq、TAPS及びTAPS β変換及びACE-seqが挙げられる。例えば、Moss et al., Nat Commun. 2018; 9:5068; Booth et al., Science 2012; 336:934-937; Yu et al., Cell 2012; 149:1368-80; Liu et al., Nature Biotechnology 2019; 37:424-429; Schutsky, E. K. et al.; 及び Vaisvila et al. Genome Research 2021 31(7):1280-1289を参照されたい。

20

【0003】

バイサルファイトベース及びEM-seqメチル化アッセイでは、非メチル化シトシンがウラシルに変換され、それはチミンとしてPCR増幅及びNGSリードがなされる。非メチル化塩基の不完全な（失敗した）変換の結果、その塩基は、メチル化された、即ち偽陽性シグナルとして不正確に同定される。逆に、メチル化塩基の誤った変換の結果、その塩基は、非メチル化された、即ち偽陰性シグナルとして不正確に同定される。非CpGシトシンの約99%がヒトゲノムにおいて非メチル化されることから、所与のDNA分子中のCHシトシン（即ち、シトシン、続いてグアニン以外の塩基）の低い変換／非変換を時として使用して、変換されていない分子は、（CpG）メチル化レベルを評価する前に濾過して除去される。

30

【0004】

他の方法では、エピジェネティック変換は、逆である、即ち、非修飾ヌクレオシドではなく、修飾ヌクレオシドが変換される。例えば、Ludwig Cancer InstituteのSong's LabからのTAPS法では、メチル化シトシン（5mC及び5hmC）は、DHUに変換され、チミンとしてPCR増幅及びNGSリードがなされる。TAPSにおけるメチル化残基の不完全な／失敗した変換の結果、その塩基は、非メチル化～即ち、偽陰性シグナルとして不正確に同定される。逆に、非メチル化塩基の誤った変換の結果、その塩基は、メチル化～即ち、偽陽性シグナルとして不正確に同定される。これらの方法では、変換されているヌクレオシドは、一般にサンプル中ではるかに希少であることから、無効な変換、そして結果としての偽陰性シグナルは、十分に検出するのがより困難である。

40

これらのアッセイにおける偽陽性及び偽陰性シグナルにまつわる課題を仮定すれば、これらの偽シグナルの推定を可能にし、それ故にデータを解釈することが可能であるような品質管理方法が求められている。

50

【先行技術文献】

【非特許文献】

【0005】

【非特許文献1】 Moss et al., Nat Commun. 2018; 9: 5068

【非特許文献2】 Booth et al., Science 2012; 336: 934-937

【非特許文献3】 Yu et al., Cell 2012; 149: 1368-80

【非特許文献4】 Liu et al., Nature Biotechnology 2019; 37: 424-429

【発明の概要】

【課題を解決するための手段】

【0006】

本明細書に記載されるのは、ヌクレオシドの修飾状態に対して感受性がある塩基対特異性変換手順を用いることに依存する、DNAサンプル中の修飾ヌクレオシドを検出及び／又は同定するための方法における変換ステップのための改善された品質管理を提供する方法である。本開示は、以下の例示的な実施形態を含む。

【0007】

実施形態1は、サンプル中のDNA中の修飾ヌクレオシドの偽陰性及び／又は偽陽性検出を監視するための品質管理方法であって、(a) DNAをオリゴヌクレオチドアダプターにライゲートすることであって、アダプターは修飾ヌクレオシドを含む品質管理ヌクレオシドを含み、品質管理ヌクレオシドは、DNA中の検出対象の修飾ヌクレオシドに対して同じヌクレオシド同一性及び同じ又は異なる修飾状態を有し、品質管理ヌクレオシドの修飾状態は既知である、ライゲートすることと；(b) 適応したDNA (adapted DNA)、又はそのサブサンプルに、ヌクレオシドの修飾状態に応じて、品質管理ヌクレオシドの塩基対特異性を変更するか又は品質管理ヌクレオシドの塩基対特異性を変更しない変換手順であって、(i) アダプター内の品質管理ヌクレオシドと同じヌクレオシド同一性及び修飾状態を有する適応したDNAヌクレオシドの塩基対特異性を変更し、アダプター内の品質管理ヌクレオシドと同じヌクレオシド同一性を有するが異なる修飾状態を有する適応したDNAヌクレオシドの塩基対特異性を変更しない；及び／又は(ii) アダプター内の品質管理ヌクレオシドと同じヌクレオシド同一性及び修飾状態を有する適応したDNAヌクレオシドの塩基対特異性を変更せず、且つアダプター内の品質管理ヌクレオシドと同じ対同一性 (pairing identity) を有するが、異なる修飾状態を有する適応したDNAヌクレオシドの塩基対特異性を変更するように選択される変換手順を施すことと；(c) 変換ステップ(b) 後、適応したDNAを配列決定することと；(d) ステップ(c) において得られた配列データを使用して、アダプター内の品質管理ヌクレオシドの塩基対特異性変換を判定することと；(e) アダプター内の品質管理ヌクレオシドの塩基対特異性変換を変換ステップ(b) のための品質管理尺度として使用することと、を含み、ここでステップ(b) (i) の変換手順後のアダプター品質管理ヌクレオシドの準最適な変換及び／又はステップ(b) (ii) の変換手順後のアダプター品質管理ヌクレオシドの誤った変換は、DNAサンプル中の修飾ヌクレオシドの偽陰性及び／又は偽陽性検出を予測する、品質管理方法である。

【0008】

実施形態2は、変換手順が、アダプター内の修飾された品質管理ヌクレオシドの塩基対特異性を変更するが、同じヌクレオシド同一性を有するが異なる修飾状態及び／又は非修飾を有するDNAサンプルのヌクレオシドの塩基対特異性を変更しないように選択され；且つ修飾された品質管理ヌクレオシドの準最適な変換が、変換手順への曝露時、品質管理ヌクレオシドと同じヌクレオシド同一性及び修飾状態又は異なる修飾状態並びに塩基対特異性における同じ変化を有するDNAサンプルヌクレオシドの偽陰性検出を予測する、実施形態1の方法である。

【0009】

実施形態 3 は、アダプター内の品質管理ヌクレオシドが、5-メチルシトシン（5 m C）及び／又は 5-ヒドロキシメチル-シトシン（5 h m C）を含む、実施形態 1 又は実施形態 2 の方法である。

【0010】

実施形態 4 は、変換手順が、置換されたボラン還元剤による T e t 補助変換を含み、任意選択的に、置換されたボラン還元剤が、2-ピコリンボラン、ボランピリジン、t e r t-ブチルアミンボラン、又はアンモニアボランである、実施形態 3 の方法である。

【0011】

実施形態 5 は、サンプル中の DNA 中の修飾ヌクレオシドの偽陰性及び／又は偽陽性検出を監視するための品質管理方法であって、（a）DNA をオリゴヌクレオチドアダプターにライゲートすることであって、アダプターは修飾ヌクレオシドを含む品質管理ヌクレオシドを含み、品質管理ヌクレオシドは、DNA 中の検出対象の修飾ヌクレオシドに対して同じヌクレオシド同一性及び同じ又は異なる修飾状態を有し、品質管理ヌクレオシドの修飾状態は既知である、ライゲートすることと；（b）適応した DNA、又はそのサブサンプルに、ヌクレオシドの修飾状態に応じて、品質管理ヌクレオシドの塩基対特異性を変更するか又は品質管理ヌクレオシドの塩基対特異性を変更しない変換手順であって、（i）アダプター内の品質管理ヌクレオシドと同じヌクレオシド同一性及び修飾状態を有する適応した DNA ヌクレオシドの塩基対特異性を変更し、且つアダプター内の品質管理ヌクレオシドと同じヌクレオシド同一性であるが異なる修飾状態を有する適応した DNA ヌクレオシドの塩基対特異性を変更しない；及び／又は（i i）アダプター内の品質管理ヌクレオシドと同じヌクレオシド同一性及び修飾状態を有する適応した DNA ヌクレオシドの塩基対特異性を変更せず、且つアダプター内の品質管理ヌクレオシドと同じ対同一性を有するが異なる修飾状態を有する適応した DNA ヌクレオシドの塩基対特異性を変更するように選択される変換手順を施すことと；（c）変換ステップ（b）後、適応した DNA を配列決定することと；（d）ステップ（c）において得られた配列データを使用して、アダプター内の品質管理ヌクレオシドの塩基対特異性変換を判定することと；（e）アダプター内の品質管理ヌクレオシドの塩基対特異性変換を変換ステップ（b）のための品質管理尺度として使用することと、を含み、ここでステップ（b）（i）の変換手順後のアダプターの品質管理ヌクレオシドの準最適な変換及び／又はステップ（b）（i i）の変換手順後のアダプターの品質管理ヌクレオシドの誤った変換は、DNA サンプル中の修飾ヌクレオシドの偽陰性及び／又は偽陽性検出を予測する、品質管理方法である。

【0012】

実施形態 6 は、アダプターの品質管理ヌクレオシドが非修飾ヌクレオシドを含む場合の実施形態 5 の方法である。

【0013】

実施形態 7 は、変換手順が、アダプター内の品質管理ヌクレオシドの塩基対特異性を変更するが、同じヌクレオシド同一性及び異なる修飾状態を有する DNA サンプルヌクレオシドの塩基対特異性を変更しないように選択され；品質管理ヌクレオシドの準最適な変換が、異なる修飾状態を有する DNA サンプルヌクレオシドの偽陽性検出を予測する、実施形態 5 又は実施形態 6 の方法である。

【0014】

実施形態 8 は、アダプター内の品質管理ヌクレオシドがシトシンを含み、且つ／又は変換手順がバイサルファイト変換を含む、実施形態 7 の方法である。

【0015】

実施形態 9 は、アダプターの品質管理ヌクレオシドが、第 1 の修飾状態を有する第 1 の品質管理ヌクレオシド及び第 1 の修飾状態とは異なる第 2 の修飾状態を有する第 2 の品質管理ヌクレオシドを含む、実施形態 1 ～ 8 のいずれか 1 つの方法である。

【0016】

実施形態 10 は、第 1 の品質管理ヌクレオシドが修飾され、第 2 の品質管理ヌクレオシドが修飾されない場合の実施形態 9 の方法である。

【0017】

実施形態11は、第1の品質管理ヌクレオシドが修飾シトシンであり、第2の品質管理ヌクレオシドが非修飾シトシンである場合の実施形態9の方法である。

【0018】

実施形態12は、第1の品質管理ヌクレオシドが、5-メチルシトシン(5mC)又は5-ヒドロキシメチルシトシン(5hmC)である場合の実施形態11の方法である。

【0019】

実施形態13は、変換手順が、第1の品質管理ヌクレオシドの塩基対特異性を変更するが、第2の品質管理ヌクレオシドを変更しないように選択される、又は変換手順が、第2の品質管理ヌクレオシドの塩基対特異性を変更するが、第1の品質管理ヌクレオシドを変更しないように選択される、実施形態9～11のいずれか1つの方法である。

10

【0020】

実施形態14は、変換手順が、アダプター内の修飾された品質管理ヌクレオシドの塩基対特異性を変更するが、同じヌクレオシド同一性を有するが、異なる修飾状態及び／又は非修飾を有するDNAサンプルヌクレオシドの塩基対特異性を変更しないように選択され；修飾された品質管理ヌクレオシドの準最適な変換が、変換手順への曝露時、品質管理ヌクレオシドと同じヌクレオシド同一性及び修飾状態、又は異なる修飾状態及び塩基対特異性の同じ変化を有するDNAサンプルヌクレオシドの偽陰性検出を予測する、実施形態9～11のいずれか1つの方法である。

20

【0021】

実施形態15は、ステップ(c)において得られた配列データを使用して、(i)アダプター配列内の品質管理ヌクレオシドの準最適な又は誤った変換を伴う適応したDNA分子を同定することと；(ii)ステップ(i)において同定された全長分子における同じヌクレオシド同一性及び修飾状態を有するヌクレオシドの準最適な又は誤った変換を推測することと、をさらに含む、実施形態1～14のいずれか1つの方法である。

30

【0022】

実施形態16は、適応したDNA又は個別の適応したDNA分子における品質管理ヌクレオシドの変換率を判定すること、(i)変換率に依存する加重を、(A)DNAサンプル；又は(B)個別の適応したDNA分子における修飾ヌクレオシド検出の分析に適用すること；(ii)修飾ヌクレオシドを検出するためのさらなる分析からの、(A)アダプター品質管理ヌクレオシドの準最適な変換若しくは所定の品質管理閾値を下回るアダプター品質管理ヌクレオシドの変換率；及び／又は(B)アダプター品質管理ヌクレオシドの誤った変換、若しくは所定の品質管理閾値を上回るアダプター品質管理ヌクレオシドの変換率を有するDNAサンプルを除去すること；並びに／又は(iii)修飾ヌクレオシドを検出するためのさらなる分析からの、(A)アダプター品質管理ヌクレオシドの準最適な変換若しくは所定の品質管理閾値を下回るアダプター品質管理ヌクレオシドの変換率；及び／又は(B)アダプター品質管理ヌクレオシドの誤った変換、若しくは所定の品質管理閾値を上回るアダプター品質管理ヌクレオシドの変換率を有する適応したDNA分子を除去すること、をさらに含む、実施形態1～15のいずれか1つの方法である。

40

【0023】

実施形態17は、サンプルから標的領域セットを捕捉することによって、DNAを富化することをさらに含み、捕捉ステップは、ライゲートステップ(a)及び変換ステップ(b)の前、後又はその間になされる、実施形態1～16のいずれか1つの方法である。

【0024】

実施形態18は、(i)ステップ(c)において得られた配列データを、(A)所定の参照配列；及び／又は(B)変換手順を受けなかったDNAのサブサンプルを配列決定することによって得られた配列データと比較することと；(ii)変換されたDNA配列と参照配列(A)又は変換されていないDNA配列データ(B)との間のポイント差異を、変換手順への曝露時、塩基対特異性の変化を可能にする修飾状態を有するヌクレオシドとして同定することと、をさらに含む、実施形態1～17のいずれか1つの方法である。

50

【0025】

実施形態19は、DNAが、無細胞DNA（cfDNA）、任意選択的に、試験対象から得られたcfDNAを含み、任意選択的に、試験対象は、がんを有するか又は有することが疑われる患者である、実施形態1～18のいずれか1つの方法である。

【0026】

実施形態20は、DNAサンプル中の修飾ヌクレオシドの検出を用いて、がん細胞若しくは腫瘍によって生成されているDNAの存在を判定又は予測すること、試験対象が腫瘍若しくはがんを有する確率を判定すること、又は対象のがん若しくは腫瘍を特徴づけることをさらに含む、実施形態1～19のいずれか1つの方法である。

【0027】

実施形態21は、DNAのサブサンプルが配列決定前に変換手順を受けず、変換されたサブサンプル及び変換されていないサブサンプルが異なるアダプター配列を有し、且つ変換されたサブサンプル及び変換されていないサブサンプルが配列決定ステップ（c）のために組み換えられる、実施形態1～20のいずれか1つの方法である。

【0028】

実施形態22は、DNAを分析して、コピー数多型、一塩基変異、挿入、欠失、メチル化、及び／又は融合を検出することをさらに含む、実施形態1～21のいずれか1つの方法である。

【0029】

実施形態23は、エピジェネティック標的領域をアダプターライゲートDNAから捕捉することと、エピジェネティック標的領域を増幅及び配列決定することと、をさらに含む、実施形態1～22のいずれか1つの方法である。

【0030】

実施形態24は、捕捉されたエピジェネティック標的領域が、エピジェネティック標的領域セットを形成する、実施形態23の方法である。

【0031】

実施形態25は、エピジェネティック標的領域セットが、複数の型特異的なエピジェネティック標的領域を含み、且つ型特異的なエピジェネティック標的領域が、型特異的な差次的にメチル化された領域及び／又は型特異的な断片である、実施形態24の方法である。

【0032】

実施形態26は、複数の型特異的なエピジェネティック標的領域が、型特異的な低メチル化領域を含む、実施形態25の方法である。

【0033】

実施形態27は、サンプルが血液サンプルである、実施形態1～26のいずれか1つの方法である。

【0034】

実施形態28は、複数の型特異的なエピジェネティック標的領域が、免疫細胞において血液サンプル中に存在する非免疫細胞型と比較して高メチル化されている；結腸において他の組織型と比較して差次的にメチル化されている；乳房において他の組織型と比較して差次的にメチル化されている；肝臓において他の組織型と比較して差次的にメチル化されている；腎臓において他の組織型と比較して差次的にメチル化されている；膵臓において他の組織型と比較して差次的にメチル化されている；前立腺において他の組織型と比較して差次的にメチル化されている；皮膚において他の組織型と比較して差次的にメチル化されている；又は膀胱において他の組織型と比較して差次的にメチル化されている標的領域を含む、実施形態25～27のいずれか1つの方法である。

【0035】

実施形態29は、複数の型特異的なエピジェネティック標的領域が、非免疫血球においてサンプル中の異なる細胞型若しくは組織型における標的領域のメチル化レベルと比較して；免疫細胞に特異的な断片において血液サンプル中に存在する非免疫細胞型と比較して；又は結腸、肺、乳房、肝臓、腎臓、膵臓、前立腺、皮膚、若しくは膀胱に特異的な断片

10

20

30

40

50

において他の組織型と比較して、低メチル化されている標的領域を含む、実施形態 25～28 のいずれか 1 つの方法である。

【0036】

実施形態 30 は、型特異的なエピジェネティック標的領域の起源である少なくとも 1 つの細胞型又は組織型を同定することをさらに含む、実施形態 23～29 のいずれか 1 つの方法である。

【0037】

実施形態 31 は、ある細胞型又は組織型に起源を持つ型特異的なエピジェネティック標的領域のレベルが決定される、実施形態 23～30 のいずれか 1 つの方法である。

【0038】

実施形態 32 は、免疫細胞、非免疫血球、結腸、肺、乳房、肝臓、腎臓、前立腺、皮膚、膀胱、又は脾臓に起源を持つ型特異的なエピジェネティック標的領域のレベルが決定される、実施形態 31 の実施形態の方法である。

【0039】

実施形態 33 は、型特異的なエピジェネティック標的領域が、細胞型特異的、組織型特異的、及び／又はがんタイプ特異的なエピジェネティック標的領域を含む、実施形態 32 の方法である。

【0040】

実施形態 34 は、血液サンプルが、少なくとも DNA のエピジェネティック標的領域セットを捕捉する前に分画される、実施形態 27～33 のいずれか 1 つの方法である。

【0041】

実施形態 35 は、サンプル又はその一定分量を、第 1 の分割されたサブサンプル及び第 2 の分割されたサブサンプルを含む複数の分割されたサブサンプルに分割することと；第 2 の分割されたサブサンプルをメチル化依存性ヌクレアーゼと接触させて、それにより第 2 のサブサンプル中の非特異的に分割された DNA を分解し、処理された第 2 のサブサンプルを生成すること、及び任意選択的に、第 1 の分割されたサブサンプルをメチル化感受性エンドヌクレアーゼと接触させて、それにより第 1 の分割されたサブサンプル中の非特異的に分割された DNA を分解し、処理された第 1 のサブサンプルを生成することと、をさらに含み、ここで第 1 のサブサンプルは、シトシン修飾を第 2 のサブサンプルより大きい割合で有する DNA を含み、ここでエピジェネティック標的領域は、第 1 のサブサンプル又は処理された第 1 のサブサンプルの少なくとも一部分から捕捉され、任意選択的に、1 つ以上の捕捉プローブと接触されている対象からの DNA は、第 1 の分割されたサブサンプル、処理された第 1 のサブサンプル、第 2 の分割されたサブサンプル、及び／又は処理された第 2 のサブサンプルからの DNA を含む、実施形態 1～34 のいずれか 1 つの方法である。

【0042】

実施形態 36 は、シトシン修飾がメチル化であり、任意選択的に、シトシン修飾がシトシンの 5 位でのメチル化である、実施形態 35 の方法である。

【0043】

実施形態 37 は、第 1 のサブサンプルが、メチル化感受性エンドヌクレアーゼと接触される、実施形態 35 又は実施形態 36 の方法である。

【0044】

実施形態 38 は、メチル化感受性エンドヌクレアーゼが、非メチル化 CpG 配列を切断する、実施形態 37 の方法である。

【0045】

実施形態 39 は、メチル化感受性エンドヌクレアーゼが、AatII、AccII、AciI、Aor13HI、Aor15HI、BspT104I、BssHII、BstUI、Cfr10I、ClaI、CpoI、Eco52I、HaeII、HapII、HhaI、Hin6I、HpaII、HpyCH4IV、MluI、NaeI、NotI、NruI、NsbI、PmaCI、Psp1406I、PvuI、SacII、SalI、

10

20

30

40

50

Sma I、及び SnaBI の 1 つ以上である、実施形態 35～38 のいずれか 1 つの方法である。

【0046】

実施形態 40 は、変換手順が、バイサルファイト変換；5hmC の保護；Tet 補助バイサルファイト変換；任意選択的に、2-ピコリンボラン、ボランピリジン、tert-ブチルアミンボラン、若しくはアンモニアボランである置換されたボラン還元剤による Tet 補助変換；hmC の保護とその後の、任意選択的に、2-ピコリンボラン、ボランピリジン、tert-ブチルアミンボラン、若しくはアンモニアボランである置換されたボラン還元剤による Tet 補助変換；hmC の保護とその後の mC 及び／若しくは C の脱アミノ化；任意選択的に、2-ピコリンボラン、ボランピリジン、tert-ブチルアミンボラン、若しくはアンモニアボランである置換されたボラン還元剤による化学的補助変換（chemical-assisted conversion）；又は修飾シトシンの酵素的保護とその後の保護されていないシトシンのウラシルへの脱アミノ化を含む、実施形態 1～39 のいずれか 1 つの方法である。

10

【0047】

実施形態 41 は、置換されたボラン還元剤が、2-ピコリンボラン又はボランピリジンである、実施形態 40 の方法である。

【0048】

実施形態 42 は、mC 及び／又は C の脱アミノ化が、AID／APOBEC ファミリー DNA デアミナーゼ 酵素による処理を含む、実施形態 40 の方法である。

20

【0049】

実施形態 43 は、hmC の保護が、hmC のグルコシル化を含む、実施形態 40 の方法である。

【0050】

実施形態 44 は、配列多様性の存在若しくは不在を検出すること、及び／又は断片化パターンを判定することをさらに含み、ここで、品質管理ヌクレオシドの準最適な又は誤った変換を意味する品質管理ヌクレオシドを含む適応した DNA は、配列多様性の存在若しくは不在を検出し、且つ／又は断片化パターンを判定することに含まれる、実施形態 1～43 のいずれか 1 つの方法である。

【0051】

実施形態 45 は、修飾ヌクレオシド、並びに変換試薬及びリガーゼのうち 1 つ以上である、品質管理ヌクレオシドを含むアダプターを含むキットである。

30

【0052】

実施形態 46 は、品質管理ヌクレオシドの塩基対特異性を変更することができる変換試薬を含む、実施形態 45 のキットである。

【0053】

実施形態 47 は、品質管理ヌクレオシドの塩基対特異性を実質的に変更することができない変換試薬を含む、実施形態 45 のキットである。

【0054】

実施形態 48 は、変換試薬が、バイサルファイト変換、酸化バイサルファイト変換、Tet 補助（TAB）変換、置換されたボラン還元剤による Tet 補助変換、置換されたボラン還元剤による Tet 補助変換と組み合わせられた hmC の保護、APOBEC 連結エピジェネティック（ACE）変換、又は核酸塩基の酵素変換において使用される変換試薬である、実施形態 45～47 のいずれか 1 つのキットである。

40

【0055】

実施形態 49 は、オリゴヌクレオチドアダプターが配列決定プライマー結合部位を含み、品質管理ヌクレオシドが配列決定プライマー結合部位の下流に位置する、実施形態 1～44 のいずれか 1 つの方法又は実施形態 45～48 のいずれか 1 つのキットである。鎖の 5' 末端のアダプターについてのプライマー結合部位「の下流」であることは、プライマー結合部位の 3' に位置することを意味し、鎖の 3' 末端のアダプターについては、プライ

50

マー結合部位の 5' に位置することを意味する。

【0056】

実施形態 50 は、アダプターを標的にするプライマーを使用して、DNA を増幅することをさらに含み、増幅ステップが、変換ステップ (b) と配列決定ステップ (c) との間になされる、実施形態 1 ~ 44 又は 49 のいずれか 1 つの方法である。

【0057】

ある態様では、本開示は、サンプル中の DNA 中の修飾ヌクレオシドの偽陰性及び／又は偽陽性検出を監視するための品質管理方法であって、(a) DNA をオリゴヌクレオチドアダプターにライゲートすることであって、アダプターは修飾ヌクレオシドを含む品質管理ヌクレオシドを含み、品質管理ヌクレオシドは、DNA 中の検出対象の修飾ヌクレオシドに対して同じヌクレオシド同一性及び同じ又は異なる修飾状態を有し、品質管理ヌクレオシドの修飾状態は既知である、ライゲートすることと；(b) 適応した DNA、又はそのサブサンプルに、ヌクレオシドの修飾状態に応じて、品質管理ヌクレオシドの塩基対特異性を変更するか又は品質管理ヌクレオシドの塩基対特異性を変更しない変換手順であって、(i) アダプター内の品質管理ヌクレオシドと同じヌクレオシド同一性及び修飾状態を有する適応した DNA ヌクレオシドの塩基対特異性を変更し、且つアダプター内の品質管理ヌクレオシドと同じヌクレオシド同一性を有するが異なる修飾状態を有する適応した DNA ヌクレオシドの塩基対特異性を変更しない；及び／又は(ii) アダプター内の品質管理ヌクレオシドと同じヌクレオシド同一性及び修飾状態を有する適応した DNA ヌクレオシドの塩基対特異性を変更せず、且つアダプター内の品質管理ヌクレオシドと同じ対同一性を有するが、異なる修飾状態を有する適応した DNA ヌクレオシドの塩基対特異性を変更するように選択される変換手順を施すことと；(c) 変換ステップ (b) 後、適応した DNA を配列決定することと；(d) ステップ (c) において得られた配列データを使用して、アダプター内の品質管理ヌクレオシドの塩基対特異性変換を判定することと；(e) アダプター内の品質管理ヌクレオシドの塩基対特異性変換を変換ステップ (b) のための品質管理尺度として使用することと、を含み、ここでステップ (b) (i) の変換手順後のアダプターの品質管理ヌクレオシドの準最適な変換及び／又はステップ (b) (ii) の変換手順後のアダプターの品質管理ヌクレオシドの誤った変換は、DNA サンプル中の修飾ヌクレオシドの偽陰性及び／又は偽陽性検出を予測する、品質管理方法を提供する。

【0058】

いくつかの実施形態では、方法は、ステップ (c) において得られた配列データを使用して、(i) アダプター配列内の品質管理ヌクレオシドの準最適な又は誤った変換を伴う適応した DNA 分子を同定することと；(ii) ステップ (i) において同定された全長分子における同じヌクレオシド同一性及び修飾状態を有するヌクレオシドの(追加的な)準最適な又は誤った変換を推測することと、をさらに含む。

【0059】

いくつかの実施形態では、方法は、適応した DNA 又は個別の適応した DNA 分子における品質管理ヌクレオシドの変換率を判定することと、(i) 変換率に依存する加重を、(A) DNA サンプル；又は(B) 個別の適応した DNA 分子における修飾ヌクレオシド検出の分析に適用することと；(ii) 修飾ヌクレオシドを検出するためのさらなる分析から、(A) アダプター品質管理ヌクレオシドの準最適な変換、若しくは品質管理閾値若しくは所定の品質管理閾値を下回るアダプター品質管理ヌクレオシドの変換率；及び／若しくは(B) アダプター品質管理ヌクレオシドの誤った変換、若しくは品質管理閾値若しくは所定の品質管理閾値を上回るアダプター品質管理ヌクレオシドの変換率を有する DNA サンプルを除去することと；並びに／又は(iii) 修飾ヌクレオシドを検出するためのさらなる分析から、(A) アダプター品質管理ヌクレオシドの準最適な変換、若しくは品質管理閾値若しくは所定の品質管理閾値を下回るアダプター品質管理ヌクレオシドの変換率；及び／又は(B) アダプター品質管理ヌクレオシドの誤った変換、若しくは品質管理閾値若しくは所定の品質管理閾値を上回るアダプター品質管理ヌクレオシドの変換率を

有する適応したDNA分子を除去することと、をさらに含む。

【0060】

いくつかの実施形態では、アダプター品質管理ヌクレオシドは修飾ヌクレオシドを含む。

【0061】

修飾されたアダプターの品質管理ヌクレオシドを使用して、修飾ヌクレオシドの塩基対特異性を変換する変換手順を用いる場合、偽陰性（即ち、修飾されていないとして不正確に同定されたヌクレオシド）を検出することができる。したがって、いくつかの実施形態では、変換手順は、アダプター内の修飾された品質管理ヌクレオシドの塩基対特異性を変更するが、品質管理ヌクレオシドと比較して、ヌクレオシド同一性が同じであるが異なる修飾状態及び／若しくは非修飾を有するDNAサンプルのヌクレオシドの塩基対特異性を変更しないように選択され；修飾された品質管理ヌクレオシドの準最適な変換は、品質管理ヌクレオシドと同じ同一性及び修飾状態、又は変換手順への曝露時、同じ塩基対特異性の変化を有するDNAサンプルのヌクレオシドの偽陰性検出を予測する。即ち、修飾されていない品質管理ヌクレオシドの塩基対特異性の変換は、同じヌクレオシド同一性を有するDNAサンプル中の他の非修飾ヌクレオシド（又は品質管理ヌクレオシドと同じヌクレオシド同一性を有するが、変換手順への曝露時、塩基対特異性の変化からヌクレオシドを保護する異なる修飾状態を有するDNAサンプルの修飾ヌクレオシド）が、品質管理ヌクレオシドの修飾状態（又は、適用可能な場合、変換手順への曝露時、塩基対特異性における同じ変化を依然として可能にする異なる修飾）を有するとして誤って同定されることを予測する。

【0062】

いくつかの実施形態では、アダプター内の品質管理ヌクレオシドは、5-メチルシトシン（5mC）及び／又は5-ヒドロキシメチル-シトシン（5hmC）を含む。いくつかの実施形態では、変換手順は、置換されたボラン還元剤によるTet補助変換を含み、任意選択的に、置換されたボラン還元剤は、2-ピコリンボラン、ボランピリジン、tert-ブチルアミンボラン、又はアンモニアボランである。

【0063】

また、修飾されたアダプターの品質管理ヌクレオシドを使用して、非修飾ヌクレオシドを変換する変換手順を用いる場合、偽陰性（即ち、修飾されていないとして不正確に同定されたヌクレオシド）を検出することができる。したがって、いくつかの実施形態では、変換手順は、アダプター内の品質管理ヌクレオシドの塩基対特異性を変更しないが、品質管理ヌクレオシドと比較して、ヌクレオシド同一性が同じであるが、異なる修飾状態及び／又は非修飾を有するDNAサンプルのヌクレオシドの塩基対特異性を変更するように選択され、ここで品質管理ヌクレオシドの誤った変換は、変換手順への曝露時、品質管理ヌクレオシドと同じヌクレオシド同一性及び修飾状態又は同じ塩基対特異性変化を有するDNAサンプルのヌクレオシドの偽陰性検出／同定を予測する。即ち、修飾された品質管理ヌクレオシドの塩基対特異性の変換は、同じヌクレオシド同一性及び同じ修飾（又は異なる修飾状態であるが、変換手順への曝露時の塩基対特異性の同じ予想される変化）を有する他のDNAサンプルのヌクレオシドが、非修飾（又は、適用可能な場合、変換手順への曝露時に塩基対特異性の変化からヌクレオシドを保護する異なる修飾）を有するとして誤って同定されることを予測する。いくつかの実施形態では、アダプター内の品質管理ヌクレオシドは、5-メチルシトシン（5mC）及び／又は5-ヒドロキシメチル-シトシン（5hmC）を含む。いくつかの実施形態では、変換手順は、バイサルファイト変換を含む。

【0064】

いくつかの実施形態では、アダプター品質管理ヌクレオシドは非修飾ヌクレオシドを含む。

【0065】

他の場合、修飾されていないアダプターの品質管理ヌクレオシドを使用して、非修飾ヌ

クレオシドを変換する変換手順を用いる場合、偽陽性（即ち、修飾されているとして不正確に同定されたヌクレオシド）を検出することができる。したがって、いくつかの実施形態では、変換手順は、アダプター内の品質管理ヌクレオシドの塩基対特異性を変更するが、品質管理ヌクレオシドと比較して、同じヌクレオシド同一性を有するが、異なる修飾状態を有するDNAサンプルのヌクレオシドの塩基対特異性を変更しないように選択され；ここで品質管理ヌクレオシドの準最適な変換は、異なる修飾状態を有するDNAサンプルのヌクレオシドの偽陽性検出を予測する。即ち、修飾されていない品質管理ヌクレオシドの塩基対特異性の非変換は、品質管理ヌクレオシドと同じヌクレオシド同一性（及び／又は、適用可能な場合、同じヌクレオシド同一性及び異なる修飾状態、具体的に、変換手順への曝露時に塩基対特異性の变化からヌクレオシドを保護しない場合）を有する他のDNAサンプルの非修飾ヌクレオシドが、修飾されていると（又は、適用可能な場合、変換手順への曝露時に塩基対特異性の变化からヌクレオシドを保護しない修飾を有すると）誤って同定されることを予測する。いくつかの実施形態では、アダプター内の品質管理ヌクレオシドは、シトシンを含む。いくつかの実施形態では、変換手順は、バイサルファイト変換を含む。

10

【0066】

また、修飾されていないアダプターの品質管理ヌクレオシドを使用して、修飾ヌクレオシドを変換する変換手順を用いる場合、偽陽性（即ち、修飾されているとして不正確に同定されたヌクレオシド）を検出することができる。したがって、いくつかの実施形態では、変換手順は、アダプター内の品質管理ヌクレオシドの塩基対特異性を変更しないが、品質管理ヌクレオシドと比較して、ヌクレオシド同一性が同じであるが、異なる修飾状態を有するDNAサンプルのヌクレオシドの塩基対特異性を変更するように選択され；ここで品質管理ヌクレオシドの誤った変換は、異なる修飾状態を有するDNAサンプルのヌクレオシドの偽陽性検出を予測する。即ち、修飾されていない品質管理ヌクレオシドの塩基対特異性の変換は、同じヌクレオシド同一性を有するDNAサンプル中の他の非修飾ヌクレオシド（又は品質管理ヌクレオシドと同じヌクレオシド同一性を有するが、変換手順への曝露時、塩基対特異性の变化からヌクレオシドを保護する異なる修飾状態を有するDNAサンプルの修飾ヌクレオシド）が、品質管理ヌクレオシドの修飾状態（又は、適用可能な場合、変換手順への曝露時、塩基対特異性における同じ変化を依然として可能にする異なる修飾）を有するとして誤って同定されることを予測する。いくつかの実施形態では、アダプター内の品質管理ヌクレオシドは、シトシンを含む。いくつかの実施形態では、変換手順は、置換されたボラン還元剤によるTet補助変換を含み、任意選択的に、置換されたボラン還元剤は、2-ピコリンボラン、ボランピリジン、tert-ブチルアミンボラン、又はアンモニアボランである。

20

30

【0067】

方法は、サンプルから標的領域セットを捕捉することによってDNAを富化することをさらに含んでもよい。捕捉ステップは、ライゲートステップ（a）及び変換ステップ（b）の前、後又はその間に行われてもよい。

【0068】

配列決定ステップ（c）は、初期サンプル中の修飾ヌクレオシドを変換されているものとして同定することを可能にする。例えば、方法は、（i）ステップ（c）において得られた配列データを、（A）（所定の）参照配列；及び／又は（B）変換手順を受けなかったDNAのサブサンプルを配列決定することによって得られた配列データと比較することと；（ii）変換されたDNA配列と参照配列（A）又は変換されていないDNA配列データ（B）との間のポイント差異を、変換手順への曝露時に塩基対特異性の变化を可能にする修飾状態を有するヌクレオチドとして同定することと、を含んでもよい。

40

【0069】

いくつかの実施形態では、DNAは、無細胞DNA（cfDNA）を含む。cfDNAは、例えば、試験対象から得てもよい。いくつかの場合、試験対象は、がんを有するか又は有することが疑われる患者である。

50

【0070】

いくつかの場合、方法は、DNAサンプル中の修飾ヌクレオシドの検出を用いて、がん細胞又は腫瘍によって生成されているDNAの存在を判定又は予測し、試験対象が腫瘍若しくはがんを有する確率を判定する、又は対象のがん若しくは腫瘍を特徴づけることをさらに含んでもよい。

【0071】

いくつかの実施形態では、DNAのサブサンプルは、配列決定前に変換手順を受けない。いくつかの場合、変換されたサブサンプル及び変換されていないサブサンプルは、異なるアダプター配列を有し得る。変換されたサブサンプル及び変換されていないサブサンプルは、配列決定ステップ(c)のために組み換えることができる。異なるアダプター配列を使用して、後のステップ又は分析において、曝露されたサブサンプル及び曝露されていないサブサンプルから配列又は分子を区別することができる。

10

【0072】

さらなる態様では、本開示は、サンプル中のDNAの修飾ヌクレオシドプロファイル进行分析するための方法であって、(a) DNAを、1つ以上の既知の修飾ヌクレオシドを含むオリゴヌクレオチドアダプターにライゲートすることと；(b) 適応したDNA、又はそのサブサンプルに、アダプター内の既知の修飾ヌクレオシドの塩基対特異性を変更する変換手順を施すことと；(c) ステップ(b)の適応したDNAを配列決定することと；(d) 配列データを使用して、アダプター内の既知の修飾ヌクレオシドの変換率を判定することと；(e) ステップ(d)で判定されている変換率を使用して、DNAサンプル中の修飾ヌクレオシドの変換率を推定することと、を含む方法を提供する。他の実施形態では、ステップ(d)は、ステップ(c)において得られた配列データを使用して、アダプター内の品質管理ヌクレオシドの塩基対特異性変換を判定することを含み；ステップ(e)は、アダプター内の品質管理ヌクレオシドの塩基対特異性変換を変換ステップ(b)における品質管理尺度として用いることを含む。配列決定ステップ(c)は、初期サンプル中の修飾ヌクレオシドが変換されているものとして同定することを可能にする。例えば、方法は、(f) ステップ(c)において得られた配列データを、(A) 所定の参照配列；及び／又は(B) 変換手順を受けなかったDNAのサブサンプルを配列決定することによって得られた配列データと比較することと；(g) 変換されたDNA配列と参照又は変換されていないDNA配列との間のポイント差異をDNAサンプル中の修飾ヌクレオチドとして同定することと、を含んでもよい。

20

30

【0073】

本開示の方法を用いて、サンプルレベル又は分子レベルのいずれかに対する変換率を推測することができる。それ故、さらなる態様では、本開示は、サンプル中のDNAの修飾ヌクレオシドプロファイル进行分析するための方法であって、(a) DNAを、1つ以上の既知の修飾ヌクレオシドを含むオリゴヌクレオチドアダプターにライゲートすることと；(b) 適応したDNA、又はそのサブサンプルに、アダプター内の既知の修飾ヌクレオシドの塩基対特異性を変更する変換手順を施すことと；(c) ステップ(b)の適応したDNAを配列決定することと；(d) 配列データを使用して、アダプター内の既知の修飾ヌクレオシドの準最適な変換を伴う適応したDNA分子を同定することと；(e) ステップ(d)で同定された全長の適応したDNA分子中の修飾ヌクレオチドの準最適な変換を推測／予測することと、を含む方法を提供する。方法は、配列データを使用して、ステップ(d)で同定された1つ以上の分子のアダプター内の既知の修飾ヌクレオシドの変換率を判定することと；判定された変換率を使用して、全長の適応したDNA分子の変換率を推定することと、をさらに含んでもよい。

40

【0074】

さらなる態様では、本開示は、DNAサンプル中の修飾ヌクレオシドを検出する方法を提供する。方法は、上に記載されているステップ(a)～(e)のいずれかのセットを含んでもよい。また、上に記載されている他の特徴は、必要に応じて、こうした方法に適用可能である。例えば、方法は、(f) ステップ(c)において得られた配列データを、(

50

A) 所定の参照配列；及び／又は (B) 変換手順を受けなかった DNA のサブサンプルを配列決定することによって得られた配列データと比較することと；(g) 変換された DNA 配列と参照又は変換されていない DNA 配列との間のポイント差異を DNA サンプル中の修飾ヌクレオチドとして同定することと、をさらに含んでもよい。ステップ (e) は、方法における品質管理尺度を提供する。

【0075】

ステップ (d) において判定されている変換率は、変換手順における品質管理尺度を提供し、それを使用して、偽陰性、即ち、変換手順によって有効に変換されず、それ故に誤って同定された初期サンプル中の修飾残基の比率を推定することができる。

【0076】

それ故、さらなる態様において、本開示は、サンプル中の DNA の修飾ヌクレオシドプロファイル进行分析するため、又は DNA サンプル中の修飾ヌクレオシドを検出するための品質管理方法を提供する。方法は、上に記載されているステップ (a) ~ (e) のいずれかのセットを含んでもよい。また、上に記載されている他の特徴は、必要に応じて、こうした方法に適用可能である。

【0077】

いくつかの実施形態では、方法は、分析における加重を、DNA サンプル又はサンプル中の個別の適応した DNA 分子に適用することをさらに含み、ここで加重は、ステップ (d) で判定されている変換率に依存する。典型的に、より大きい加重が、変換率がより高く判定されたサンプル又は分子に与えられ、より小さい加重が、変換率がより低く判定されたサンプル又は分子に与えられる。加重は、変換後にサンプルの DNA 分子を配列決定することによって判定されている修飾ヌクレオシドプロファイルに割り当てることができる確度のレベルを反映する。いくつかの場合、(所定の) 品質管理閾値を下回る準最適な変換又は変換率を伴うサンプル又は分子は、さらなる分析から除外してもよい。

【0078】

いくつかの実施形態では、本明細書で開示される方法の結果は、報告を作成するための入力として使用される。報告は、紙又は電子形式であってもよい。例えば、本明細書で開示される方法、又はそれから得られた情報によって得られるとおりの偽陽性及び／又は偽陰性の検出は、こうした報告中で直接的に提示することができる。代替的に、又は追加的に、少なくとも部分的に本明細書で開示される方法に基づく診断情報又は治療推奨は、報告に含めることができる。

【0079】

本明細書で開示される方法の様々なステップは、同じ若しくは異なる時点で、同じ若しくは異なる地理的位置、例えば国において、及び／又は同じ若しくは異なる人々によって実施してもよい。

【0080】

追加的利点は、一部には以下の説明において記載されることになり、又は実践によって学習することができる。利点は、特に添付の請求項に指摘される要素及び組み合わせを用いて実現され、達成されるであろう。

【図面の簡単な説明】

【0081】

【図1】図1は、バイサルファイト配列決定、NEB EM-seq 及び TAPS といった3つの一塩基分解メチル化アッセイ間の差異を図示する。3つ全ての方法では、メチルシトシン及びヒドロキシメチルシトシンが非メチル化シトシンと区別される。しかし、TAPS は、他の2つの方法に対する逆変換ロジックを有し、メチル化シトシン(「T」として解読されている配列)の塩基対を変更し、非メチル化シトシン(さらに「C」として解読されている)を変更しない一方で、バイサルファイト配列決定及び EM-seq は、非メチル化シトシン(「T」として解読されている配列)の塩基対を変更し、メチル化シトシン(さらに「C」として解読されている)を変更しない。

【図2】図2は、本開示のいくつかの実施形態での使用に適したシステムの例の模式図で

10

20

30

40

50

ある。

【図3】図3は、TAPS塩基変換手順を受けているDNAの偽陰性及び／又は偽陽性検出を監視するための品質管理方法の実施形態を図示する。（例えば、分子バーコード中の）5mCを含有するアダプターは、DNAにライゲートされ、次にTAPS変換手順を受け、メチル化シトシン（「T」として解読されている配列）の塩基対特異性を変更するが、非メチル化シトシン（さらに「C」として解読されている）を変更しない。アダプター内に変換されなかった5mCを含有するDNA分子を使用して、全長分子における5mCの準最適な変換を推測することができる。

【発明を実施するための形態】

【0082】

ここで、本開示の特定の実施形態に対して、詳細に参照がなされる。本開示がこうした実施形態と組み合わせて記載される一方で、それらが本発明をそれらの実施形態に限定することが意図されていないことは理解されるであろう。それに対して、本開示は、添付の特許請求の範囲によって定義されている本発明の中に含めてもよい、全ての代替物、修飾、及び均等物を包含することが意図されている。

【0083】

本教示内容を詳細に説明する前に、本開示が、具体的な組成物又はプロセスステップに、それらの変動し得ることから、限定されないことは理解されるべきである。本明細書及び添付の特許請求の範囲において使用する場合、単数形の「a」、「an」及び「the」が、文脈上そうでないとする明確な指示がない限り、複数の参照を含むことは注目すべきである。したがって、例えば、「核酸」は、複数の核酸を含む。

【0084】

数値範囲は、範囲を定義する数を包含する。測定された値及び測定可能な値は、有効数字及び測定に伴う誤差を考慮する近似値であると理解される。

【0085】

上の明細書中で具体的に記述されない限り、様々な成分を「含むこと（comprising）」として列記する明細書中の実施形態は、列記された成分「からなる（consisting of）」又は「から本質的になる（consisting essentially of）」ことも企図される。

【0086】

本明細書で用いられるセクション見出しは、構造化を目的としており、決して開示される主題を限定するものとして解釈されるべきでない。

【0087】

本明細書中で引用されている全ての特許、特許出願、ウェブサイト、他の出版物又は文書などは、上記又は下記のいずれであっても、あらゆる目的として、各個別の項目が参照により援用されることが具体的且つ個別に示される場合と同程度に、それら全体が参照により明示的に援用される。出版物、ウェブサイトなどの異なるバージョンが異なる時点で公開される場合、別段の指示がない限り、ごく最近、本願の有効出願日に公開されているバージョンを意味する。

【0088】

定義

本明細書で使用する場合、「塩基対特異性」は、所与の塩基が最も優先的に対合する場合の標準的なDNA塩基（A、C、G、又はT）を指す。したがって、例えば、非修飾シトシン及び5-メチルシトシンは、同じ塩基対特異性（即ち、Gに対する特異性）を有する一方で、ウラシル及びシトシンは、ウラシルがAに対する塩基対特異性を有する一方で、シトシンがGに対する塩基対特異性を有することから、異なる塩基対特異性を有する。例えば、ウラシルがGとゆらぎ対を形成する能力は、ウラシルがやはり4つの標準的なDNA塩基の中のAと最も優先的に対合することから、無関係である。

【0089】

「同じ同一性」のヌクレオシド又は「同じヌクレオシド同一性」は、同じ塩基を、その

10

20

30

40

50

塩基の修飾状態と無関係に有するヌクレオシドを指す。例えば、シトシンは、5-メチルシトシン（5mC）及び／又は5-ヒドロキシメチルシトシン（5hmC）と、それらが異なる修飾状態を有することと無関係に「同じ同一性」であると考えられる。

【0090】

「変換試薬」は、核酸中の少なくとも1つの修飾又は非修飾ヌクレオシドの塩基対特異性を変更するために使用することができる試薬を指す。例えば、バイサルファイトは、（Gに対する塩基対特異性を有する）非修飾シトシンを（Aに対する塩基対特異性を有する）ウラシルに変更するために使用することができる変換試薬であり、Tet酵素及びピリジンボランは、（Gに対する塩基対特異性を有する）メチル化又はヒドロキシメチル化シトシンを（Aに対する塩基対特異性を有する）ジヒドロウラシルに変更するために一緒に使用することができる変換試薬である。変換手順は、核酸中の少なくとも1つの修飾又は非修飾ヌクレオシドの塩基対特異性を変更するために1つ以上の変換試薬を使用する手順である。

10

【0091】

変換試薬又は手順は、変換試薬又は手順が、第2の異なる修飾状態を有するそのヌクレオシドの塩基対特異性を、第1の修飾状態を有するヌクレオシドの塩基対特異性への変更が誤っていると適切にみなされるポイントまで優先的に変更する場合、（修飾又は非修飾であり得る）第1の修飾状態を有するヌクレオシドの「塩基対特異性を実質的に変更することができない」。例えば、バイサルファイト変換は、5mC（第1の修飾状態を有するヌクレオシド）の塩基対特異性を実質的に変更することができないが、非修飾シトシン（第2の修飾状態を有するヌクレオシド）の塩基対特異性を優先的に変更する。それに対し、TAPSは、5mC（及び5hmC）（それらのいずれかは、第2の修飾状態を有すると考えることができる）の塩基対特異性を優先的に変更するが、TAPSは、非修飾C（第1の修飾状態を有するヌクレオシド）の塩基対特異性を実質的に変更することができない。

20

【0092】

1つ以上の標的核酸を「捕捉すること」は、1つ以上の標的核酸を非標的核酸から優先的に単離又は分離することを指す。

【0093】

核酸の「捕捉されたセット」は、捕捉を受けている核酸を指す。

30

【0094】

「標的領域セット」又は「標的領域のセット」は、捕捉のために標的化され、且つ／又は（例えば、配列相補性を通じて）プローブのセットによって標的化される複数のゲノム遺伝子座を指す。

【0095】

「標的領域セットに対応する」は、核酸、例えばcfDNAが、標的領域セットにおける遺伝子座に起源を持つ、又は標的領域セットに対する1つ以上のプローブに特異的に結合することを意味する。

【0096】

「配列可変標的領域」は、新生細胞（例えば、腫瘍細胞及びがん細胞）における、正常細胞に対する、ヌクレオチド置換（即ち、一塩基多様性）、挿入、欠失、又は遺伝子融合若しくは転位などの配列の変化を示し得る標的領域を指す。配列可変標的領域セットは、配列可変標的領域のセットである。いくつかの実施形態では、配列可変標的領域は、50以下の隣接ヌクレオチド、例えば、40、30、20、10、5、4、3、若しくは2以下のヌクレオチドに影響する、又は1ヌクレオチドに影響する変化を示し得る標的領域である。

40

【0097】

「エピジェネティック標的領域」は、組織型を通じた配列非依存性の変化を示し得る標的領域（例えば、固形組織型において造血細胞内と異なる程度のメチル化を有する標的領域）、又は新生細胞、例えば腫瘍細胞又はがん細胞における正常細胞との差異を示し得る

50

標的領域を指す。いくつかの実施形態では、エピジェネティック標的領域は、通常は c f DNA に実質的に寄与しない組織型、例えば、肺、結腸などに起源を持つ c f DNA における、バックグラウンド c f DNA、例えば、造血細胞に起源を持つ c f DNA に対する配列非依存性の差異を示す。いくつかの実施形態では、エピジェネティック標的領域は、がんを有する対象からの c f DNA における、健常対象からの c f DNA に対する配列非依存性の差異を示す。配列非依存性の変化の例としては、限定はされないが、メチル化の変化（増加又は減少）、ヌクレオソーム分布、c f DNA 断片化パターン、C C C T C 結合因子（「C T C F」）結合、転写開始部位、及び調節タンパク質結合領域が挙げられる。エピジェネティック標的領域セットは、エピジェネティック標的領域のセットである。それ故、エピジェネティック標的領域セットとして、限定はされないが、高メチル化可変標的領域セット、低メチル化可変標的領域セット、及び断片化可変標的領域セット、例えば、C T C F 結合部位及び転写開始部位が挙げられる。本目的として、新生物、腫瘍、若しくはがん関連の限局的増幅及び／又は遺伝子融合を受けやすい遺伝子座は、例えば、限局的増幅及び／又は遺伝子融合が、それらの検出が1つ又はいくつかの個別位置で塩基要求の正確性に依存しないことから、比較的浅い配列決定の深さで検出可能であるという点で、配列決定によるコピー数の変化又は参照ゲノム内の2つ以上の遺伝子座に位置づける融合された配列の検出が、ヌクレオチド置換、挿入、又は欠失の検出よりも、上で考察された代表的なエピジェネティック変化の検出に類似する傾向があるため、エピジェネティック標的領域セットに含めることもできる。

10

20

【0098】

本明細書で使用する場合、「エピジェネティックな特徴」は、主要な配列（即ち、A、C、G、及びT塩基の配列）以外のDNA又はクロマチンの任意の特徴を指す。エピジェネティックな特徴は、塩基の共有結合修飾、例えばメチル化、並びにヒストン及び他の安定なDNA関連タンパク質の修飾及びポジショニングを含む。

【0099】

本明細書で使用する場合、「差次的メチル化領域」は、組織の少なくとも1つのタイプにおいて、健常対象からの無細胞DNAにおけるメチル化の程度に対して検出可能に異なる程度のメチル化を有する領域を指す。いくつかの実施形態では、DNAの差次的メチル化領域は、組織の少なくとも1つのタイプにおいて、組織の別のタイプにおけるメチル化の程度に対して；又は健常対象からのサンプルにおいて、前がん、がん、又は新生物を有する対象におけるメチル化の程度に対して、検出可能に異なる程度のメチル化を有する。いくつかの実施形態では、差次的メチル化領域は、組織の少なくとも1つのタイプにおいて、健常対象からの無細胞DNAにおけるメチル化の程度に対して検出可能により高い程度のメチル化を有する。いくつかの実施形態では、差次的メチル化領域は、組織の少なくとも1つのタイプにおいて、健常対象からの無細胞DNAにおけるメチル化の程度に対して検出可能により低い程度のメチル化を有する。いくつかの実施形態では、差次的メチル化領域は、赤血球系統において、又は未熟赤血球（例えば、網状赤血球）において低メチル化され、少なくとも1つの非赤血球細胞型又は組織型（例えば、白血球又は固形組織細胞型、例えば、上皮細胞、筋細胞など）において高メチル化される。

30

40

【0100】

本明細書で使用する場合、エピジェネティック変異との関連での「型特異的な」は、1つの細胞型若しくは組織型において、又は複数の関連する細胞型若しくは組織型において、他の細胞型又は組織型に対して検出可能に異なる程度で存在するエピジェネティック変異を意味する。同様に、「型特異的なエピジェネティック標的領域」は、1つの細胞型若しくは組織型において、又は複数の関連する細胞型若しくは組織型において、他の細胞型又は組織型に対して検出可能に異なるエピジェネティックな特徴を有するエピジェネティック標的領域である。例示的なエピジェネティックな特徴は、上記のエピジェネティック標的領域の定義において考察されている。例えば、「型特異的な差次的メチル化領域」は、1つの細胞型若しくは組織型において、又は複数の関連する細胞型若しくは組織型において、他の細胞型又は組織型に対して検出可能に異なる程度のメチル化を有するDNAの

50

領域である。型特異的な差次的メチル化領域の例としては、組織特異的な差次的メチル化領域、例えば、早期がんにおけるコピー数増加を伴うものが挙げられる。いくつかの実施形態では、型特異的な差次的メチル化領域の捕捉、同定、及び／又は検出により、DNAの起源である細胞型又は組織型の同定が容易になる。型特異的な差次的メチル化領域の起源である細胞又は組織は、野生型細胞若しくは組織又は新生細胞若しくは組織であってもよい。別の例では、DNAの「型特異的断片」は、1つの細胞型若しくは組織型において、又は複数の関連する細胞型若しくは組織型において、他の細胞型又は組織型に対して検出可能に異なる程度で存在する型特異的な断片化パターンから生じるDNA断片である。いくつかの実施形態では、型特異的断片は、特定の細胞型又は組織型に限って存在する。いくつかの実施形態では、型特異的断片は、特定の細胞型又は組織型において検出可能により大きい程度で存在する。

10

【0101】

本明細書で使用する場合、「血液サンプル」は、全血又はその成分（例えば、血漿、血清、バフィーコート、血漿ペレット）を含むサンプルを指す。

【0102】

「バフィーコート」は、サンプルの白血球及び血小板の全部又は大部分を含有する、血液（全血など）又は骨髓サンプルの一部分を指す。サンプルのバフィーコート画分は、サンプル成分を密度によって分離する遠心分離を用いて、サンプルから調製することができる。例えば、全血サンプルの遠心分離後、バフィーコート画分は、血漿層と赤血球（erythrocyte）（赤血球（red blood cell））層との間に位置する。バフィーコートは、単核（例えば、T細胞、B細胞、NK細胞、樹状細胞、及び単球）と多形核（例えば、好中球及び好酸球などの顆粒球）白血球の両方を含有し得る。

20

【0103】

本明細書で使用する場合、「プライマーとアニールされたDNA」は、少なくとも1つのプライマーがアニールされる対象のDNAを意味する。

【0104】

本明細書中のDNAの「イントロン領域」は、イントロン又はその一部分をコードする。イントロン領域は、mRNAを生成するためのpre-mRNAの転写後スプライシング中に除去されるイントロンをコードする領域とともに、（例えば、免疫グロブリン又はT細胞受容体遺伝子座における）生殖系列J遺伝子セグメント間に介入し、体細胞V(D)J組換え中に除去される配列である「Jイントロン領域」を含む。本明細書中のDNAの「エクソン領域」は、pre-mRNA又はmRNAにおける少なくとも1つのエクソン又はその一部分をコードする。いくつかの実施形態では、エクソン領域は、VDJエクソン領域であり、それがpre-mRNA又はmRNAにおける1つ以上のV、D、又はJエクソンをコードすることを意味する。本明細書中のDNAの「エクソン-エクソン接合領域」は、任意のイントロンがスプライシングによって除去される前、pre-mRNA（例えば、免疫グロブリン又はT細胞受容体のpre-mRNA）における少なくとも1つのエクソン-エクソン接合部をコードする。エクソン-エクソン接合領域は、例えば、JセグメントがD若しくはVセグメントに連結される、又はVセグメントがD若しくはJセグメントに連結されるとき、V(D)J組換えによって形成され得る。

30

40

【0105】

本明細書で使用する場合、DNA「構造変異」又は「構造変異体」は、点突然変異以外の野生型ゲノム中に存在しないDNA配列を含む突然変異である（例えば、少なくとも5、10、20、又は50の隣接ヌクレオチドが対応する遺伝子座で野生型配列と異なる場合）。DNA構造変異の例として、再配列、例えば、転位、挿入、欠失、重複、コピー数多型、及び逆位が挙げられる。本明細書で使用する場合、「構造変異配列」又は「構造変異体配列」は、構造変異の一部分又は全体を含むか又はそれからなるDNA配列である。

【0106】

本明細書で使用する場合、DNA分子などの核酸の「分割」は、核酸のサンプル又は集団の、核酸の複数のサブサンプル又は亜集団への、複数のサブサンプル又は亜集団のそれ

50

それにおいて異なる割合である1つ以上の修飾又は特徴に基づく、分離、分画、選別、又は富化を意味する。分割は、核酸分子を、1つ以上のメチル化核酸塩基の存在又は不在に基づき、物理的に分割することを含んでもよい。サンプル又は集団は、遺伝的又はエピジェネティックな変化又は病態を意味する特徴に基づき、1つ以上の分割されたサブサンプル又は亜集団に分割してもよい。

【0107】

本明細書で使用する場合、「最初に単離された」サンプルの形態は、サンプルの、それが単離された時点、及び単離されたサンプルの化学構造を変更する任意の手順を受ける前の組成物又は化学構造を指す。同様に、分子中に「最初に存在する」特徴は、「元の分子」又はその分子が分子の化学構造を変更する任意の手順を受ける前に特徴を「最初を含む」分子に存在する特徴を指す。

10

【0108】

本明細書で使用する場合、所与の核酸塩基の「塩基対特異性の実質的変更を伴わない」は、配列決定可能であるその核酸塩基を含む分子の大半が、所与の核酸塩基の塩基対特異性の、最初に単離されたサンプル中に存在したときのその塩基対特異性に対する変更を有しないことを意味する。いくつかの実施形態では、配列決定可能であるその核酸塩基を含む分子の75%、90%、95%、又は99%は、塩基対特異性の、最初に単離されたサンプル中に存在したときのその塩基対特異性に対する変更を有しない。本明細書で使用する場合、所与の核酸塩基の「変更された塩基対特異性」は、配列決定可能であるその核酸塩基を含む分子の大半が、その核酸塩基での塩基対特異性の、最初に単離されたサンプル中のその塩基対特異性に対する変更を有することを意味する。

20

【0109】

本明細書で使用する場合、複数のメンバーを含む「組み合わせ」は、例えば、別々の容器又は区画であって、より大きい容器、例えば、マルチウェルプレート、チューブラック、冷蔵庫、冷凍庫、インキュベーター、水槽、アイスバケット、器具、又は他の貯蔵形態に内包される容器又は区画内に、メンバーを含む単一組成物又は近接した組成物のセットのいずれかを指す。

【0110】

本明細書で使用する場合、「標識」は、それが付着するものの検出、分離、又は単離を容易にする、捕捉部分、フルオロフォア、オリゴヌクレオチド、又は他の部分である。

30

【0111】

本明細書で使用する場合、「捕捉部分」は、捕捉部分に連結されている分子の、捕捉部分を欠いている分子からの親和性分離を可能にする分子である。例示的な捕捉部分は、固相に連結されているか又は連結可能であるストレプトアビジンに結合することによって親和性分離を可能にするビオチン、又は固相に連結されているか又は連結可能である相補的オリゴヌクレオチドへの結合を通じた親和性分離を可能にするオリゴヌクレオチドを含む。

【0112】

本明細書で使用する場合、「捕捉プローブ」は、捕捉部分を含み、増幅によって生成されることから、鋳型DNAのアンプリコンを含むプローブを意味する。いくつかの実施形態では、増幅は、ポリメラーゼ連鎖反応(PCR)を含む。

40

【0113】

本明細書で使用する場合、2つのプライマーの「逆平行方向」は、プライマーが互いに対して逆方向で核酸に(例えば、核酸の逆鎖に)、及び／又は指数関数的なPCR増幅に適合する方向で核酸にアニールすることを意味する。例えば、逆平行方向で再配列にアニールするプライマーは、再配列ブレークポイントを含むアンプリコンの増幅を促進し得る。

【0114】

本明細書で使用する場合、「タグ」は、タグの会合対象の分子の特徴を示す情報を含む、核酸、標識、フルオロフォア、又はペプチドなどの分子である。例えば、分子は、(1

50

つのサンプル中の分子を異なるサンプル中の分子と区別する) サンプルタグ、((固有及び非固有の両方のタグ付けシナリオにおいて) 異なる分子を互いから区別する分子タグ／分子バーコード／バーコード、精製タグ、及び／又は検出可能なタグ若しくは標識を有し得る。

【0115】

プローブ又は他のオリゴヌクレオチド及び標的配列との関連での「特異的に結合する」は、適切なハイブリダイゼーション条件下で、オリゴヌクレオチド又はプローブがその標的配列、又はその複製物とハイブリダイズして、安定なプローブ：標的ハイブリッドを形成する一方で、同時に安定なプローブ：非標的ハイブリッドの形成が最小化されることを意味する。したがって、プローブは、標的配列又はその複製物と、非標的配列より十分に大きい範囲までハイブリダイズして、標的配列の捕捉又は検出を可能にする。当該技術分野で周知である適切なハイブリダイゼーション条件は、配列組成物に基づいて予測することができ、又は通常の試験方法を用いることによって決定することができる(例えば、参照により本明細書に援用される、Sambrook et al., Molecular Cloning, A Laboratory Manual, 2nd ed. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) の §§ 1.90~1.91, 7.37~7.57, 9.47~9.51 及び 11.47~11.57、特に §§ 9.50~9.51, 11.12~11.13, 11.45~11.47 及び 11.55~11.57 を参照されたい)。

10

20

【0116】

核酸は、それが腫瘍細胞に起源を有した場合、「腫瘍によって生成される」、又は ctDNA 又は循環腫瘍 DNA である。腫瘍細胞は、それらが腫瘍内に留まるか又は腫瘍から分かれるようになるか(例えば、転移性がん細胞及び循環腫瘍細胞の症例などの場合)にかかわらず、腫瘍に起源を持つ新生細胞である。

【0117】

核酸との関連での「標的領域」は、例えば、(例えば、配列相補性を通じた) プローブを使用することによる同定及び／又は捕捉が標的とされる遺伝子座を指す。「標的領域セット」又は「標的領域のセット」は、例えば、(例えば、配列相補性を通じた) プローブのセットを使用することによる同定及び／又は捕捉が標的とされる複数のゲノム遺伝子座を指す。

30

【0118】

所与の標的領域セットに対するプローブの収集の「捕捉収率」は、プローブの収集物が典型的な条件下で捕捉する、標的領域セットに対応する核酸の量(例えば、別の標的領域セットに対する量又は絶対量)を指す。例示的な典型的な捕捉条件は、サンプル核酸及びプローブの、ストリンジェントなハイブリダイゼーション緩衝液を含有する小さい反応体積(約 20 µL)内、65℃で 10~18 時間のインキュベーションである。捕捉収率は、絶対項で表すことができ、又は複数のプローブの収集物の場合、相対項で表すことができる。複数の標的領域のセットについての捕捉収率が比較される場合、それらは(例えば、キロベース当たりを基準として)標的領域セットのフットプリントサイズに対して正規化される。それ故、例えば、第 1 及び第 2 の標的領域のフットプリントサイズのそれぞれが 50 kb 及び 500 kb である(0.1 の正規化係数を示す)場合、第 1 の標的領域セットに対応する捕捉された DNA の体積濃度当たりの質量が、第 2 の標的領域セットに対応する捕捉された DNA の体積濃度当たりの質量の 0.1 倍を超えると、第 1 の標的領域セットに対応する DNA は、第 2 の標的領域セットに対応する DNA より高い収率で捕捉される。さらなる例として、同じフットプリントサイズを用いると、第 1 の標的領域セットに対応する捕捉された DNA が、第 2 の標的領域セットに対応する捕捉された DNA の体積濃度当たりの質量の 0.2 倍の体積濃度当たりの質量を有する場合、第 1 の標的領域セットに対応する DNA は、第 2 の標的領域セットに対応する DNA より 2 倍大きい捕捉収率で捕捉された。

40

50

【0119】

「メチル化」又は「DNAメチル化」という用語は、メチル基の核酸分子中のヌクレオチド塩基への付加を指す。いくつかの実施形態では、メチル化は、メチル基のシトシンへのCpG部位（シトシンーリン酸塩ーグアニン部位（即ち、核酸配列の5'→3'方向でのシトシンとそれに続くグアニン））での付加を指す。いくつかの実施形態では、DNAメチル化は、メチル基のアデニンへの付加であって、例えば、N6-メチルアデニンなどの場合を指す。いくつかの実施形態では、DNAメチル化は、5-メチル化（シトシン環の5位における炭素の修飾）である。いくつかの実施形態では、5-メチル化は、5-メチルシトシン（5mC）を生成するためのメチル基のシトシンの5C位への付加を指す。いくつかの実施形態では、メチル化は、5mCの誘導体を含む。5mCの誘導体として、限定はされないが、5-ヒドロキシメチルシトシン（5-hmC）、5-ホルミルシトシン（5-fC）、及び5-カルボキシルシトシン（carboxyl cytosine）（5-caC）が挙げられる。いくつかの実施形態では、DNAメチル化は、3Cメチル化（シトシン環の3位における炭素の修飾）である。いくつかの実施形態では、3Cメチル化は、3-メチルシトシン（3mC）を生成するためのメチル基のシトシンの3C位への付加を含む。メチル化はまた、非CpG部位で行うことができ、例えば、メチル化は、CpA、CpT、又はCpC部位で行うことができる。DNAメチル化は、メチル化DNA領域の活性を変更し得る。例えば、プロモーター領域内のDNAがメチル化される場合、遺伝子の転写は抑制され得る。DNAメチル化は、正常な発現のために決定的であり、メチル化の異常は、エピジェネティック調節を破壊し得る。エピジェネティック調節における破壊、例えば抑制は、がんなどの疾患を引き起こすことがある。DNAにおけるプロモーターメチル化は、がんを意味し得る。

10

20

【0120】

「DNAの修飾ヌクレオシドプロファイル」は、DNA配列内のヌクレオシドの位置及び同一性並びにヌクレオシドの修飾状態、例えばメチル化を意味する。上記のとおり、変換とそれに続く配列決定の異なる方法により、修飾又は非修飾ヌクレオシドの1つ以上の異なるタイプが検出される。例えば、TAPS法は、5-メチルシトシン（5mC）及び5-ヒドロキシメチルシトシン（5hmC）を検出するが、それらの間を区別することができない。それ故、サンプル中のDNAの修飾ヌクレオシドプロファイルを分析するための方法は、典型的に、特定の修飾又は修飾基、例えば5mC及び／又は5hmCを同定することを意味する。修飾ヌクレオシドは、上記のように用いられている特定の方法／変換手順に従って同定される。これは一般に、変換手順を受けているDNAから得られた配列データを参照配列と比較することを含む。典型的に、方法は、（i）配列データを、（A）典型的には、例えば、がんの診断、予後又は特徴づけにおいて、修飾ヌクレオシドプロファイルに特定の有意性が加わる場合の1つ以上のエピジェネティック標的領域に対応する、1つ以上の所定の参照配列；又は（B）変換手順を受けなかったDNAのサブサンプル、例えば、例えば本明細書に記載のとおり、別々のサブサンプルを変換手順にかける前に分離されたサブサンプルを配列決定することによって得られた配列データと比較することと；（ii）変換されたDNA配列と参照配列（A）又は変換されていないDNA配列（B）との間のポイント差異を、変換手順への曝露時の塩基対特異性の変化を可能にする修飾状態を有する（初期サンプル中の）ヌクレオシドとして同定することと、を含む。

30

40

【0121】

標準変換及び配列決定方法によって判定されている修飾ヌクレオシドプロファイルが、サンプル中の修飾又は非修飾ヌクレオシドの不完全な又は誤った変換に起因するエラーを含有してもよいことは理解される。本開示の方法は、サンプル又は分子基盤のいずれかに対する変換率を評価するための手段を提供する。

【0122】

「高メチル化」という用語は、核酸分子の、核酸分子の集団（例えば、サンプル）中の他の核酸分子に対するメチル化のレベル又は程度の増加を指す。いくつかの実施形態では、高メチル化DNAは、少なくとも1つのメチル化残基、少なくとも2つのメチル化残基

50

、少なくとも3つのメチル化残基、少なくとも5つのメチル化残基、又は少なくとも10のメチル化残基を含むDNA分子を含み得る。本明細書で使用する場合、「型特異的な高メチル化」は、1つの細胞型若しくは組織型における、又は複数の関連する細胞型若しくは組織型における、他の細胞型又は組織型に対する、DNAのメチル化のレベル又は程度の増加を意味する。いくつかの実施形態では、型特異的な高メチル化領域の捕捉、同定、及び／又は検出により、DNAの起源である細胞型又は組織型の同定が促進される。型特異的な高メチル化領域の起源である細胞又は組織は、野生型細胞若しくは組織又は新生細胞若しくは組織であってもよい。

【0123】

「低メチル化」という用語は、核酸分子の、核酸分子の集団（例えば、サンプル）中の他の核酸分子に対するメチル化のレベル又は程度の低下を指す。いくつかの実施形態では、低メチル化DNAは、非メチル化DNA分子を含む。いくつかの実施形態では、低メチル化DNAは、0のメチル化残基、最大で1つのメチル化残基、最大で2つのメチル化残基、最大で3つのメチル化残基、最大で4つのメチル化残基、又は最大で5つのメチル化残基を含むDNA分子を含み得る。本明細書で使用する場合、「型特異的な低メチル化」は、1つの細胞型若しくは組織型における、又は複数の関連する細胞型若しくは組織型における、他の細胞型又は組織型に対する、DNAのメチル化のレベル又は程度の低下を意味する。いくつかの実施形態では、型特異的な低メチル化領域の捕捉、同定、及び／又は検出により、DNAの起源である細胞型又は組織型の同定が促進される。型特異的な低メチル化領域の起源である細胞又は組織は、野生型細胞若しくは組織又は新生細胞若しくは組織であってもよい。

【0124】

「DNA中の修飾核酸塩基を認識する薬剤」、例えば「DNA中の修飾シトシンを認識する薬剤」という用語は、DNA中の1つ以上の修飾核酸塩基、例えばメチルシトシンに結合する、又はそれを検出する分子又は試薬を指す。

【0125】

「修飾ヌクレオシド」は、非修飾ヌクレオシドとの化学構造の差異を含むヌクレオシドである。DNAの場合、非修飾ヌクレオシドは、デオキシリボシル及びアデニン、シトシン、グアニン、又はチミンの1つを含む。いくつかの実施形態では、修飾ヌクレオシドは、修飾シトシンを含む。いくつかの実施形態では、修飾ヌクレオシドは、メチル化核酸塩基を含む。いくつかの実施形態では、修飾シトシンは、メチルシトシン、例えば、5-メチルシトシンである。かかる実施形態では、シトシン修飾は、メチルである。DNA中のメチルシトシンを認識する薬剤は、限定はされないが、本明細書中でメチルシトシンに結合する試薬を指す「メチル結合試薬」を含む。メチル結合試薬は、限定はされないが、メチル結合ドメイン（MBD）及びメチル結合タンパク質（MBP）並びにメチルシトシンに特異的な抗体を含む。いくつかの実施形態では、そのような抗体は、DNA中の5-メチルシトシンに結合する。いくつかのそのような実施形態では、DNAは、一本鎖又は二本鎖であってもよい。

【0126】

「又はその組み合わせ（or a combination thereof）」及び「又はその組み合わせ（or combinations thereof）」は、本明細書で使用する場合、その用語に先行する列記された用語のありとあらゆる並べ替え及び組み合わせを指す。例えば、「A、B、C、又はその組み合わせ」は、A、B、C、AB、AC、BC、又はABCの少なくとも1つを含む、そして特定の状況下で順序が重要である場合、BA、CA、CB、ACB、CBA、BCA、BAC、又はCABも含むことが意図されている。この例に続き、明示的に含まれるのは、例えば、BB、AAA、AAB、BBC、AAABCCCC、CBBAAAA、CABABBBなどの1つ以上の項目又は用語の繰り返しを含む組み合わせである。当業者は、文脈から特に明白でない限り、典型的には任意の組み合わせでの項目又は用語の数に対する制限がないことを理解するであろう。

【0127】

「又は」は、文脈上他の意味に解すべき場合を除き、包括の意味で、即ち、「及び／又は」と同等に使用される。

【0128】

サンプル及び対象

本開示は、核酸の修飾ヌクレオシドプロファイル进行分析する方法、又はサンプル中のDNA、例えばサンプル中のDNA中の修飾ヌクレオシドの偽陰性及び／若しくは偽陽性検出を監視するための品質管理方法に関する。いくつかの場合、核酸は、対象から得られる又は得られた。いくつかの実施形態では、核酸サンプルは、対象から得られる生体サンプルからの核酸、例えばDNAを含んでもよい、又はそれから構成されてもよい。対象は、ヒト、哺乳動物、動物、霊長類、齧歯類（マウス及びラットを含む）、又は他の通常の実験、家庭用、コンパニオン、サービス若しくは農業動物、例えば、ウサギ、イヌ、ネコ、ウマ、雌ウシ、ヒツジ、ヤギ又はブタであってもよい。対象は、いくつかの場合、がん、腫瘍又は新生物を有し得る、又は有することが疑われ得る。他の場合、対象は、がん又は検出可能ながん症状を有しないことがある。対象は、1つ以上のがん療法、例えば、化学療法、抗体、ワクチン又は生物剤（*biologics*）のいずれか1つ以上で治療されているかもしれない。対象は、例えば、（例えば、化学療法、外科的切除、放射線、又はその組み合わせなどの治療後）腫瘍、がん、又は新生物から寛解状態であり得る。対象は、がん又は任意のがん関連遺伝子の突然変異／障害に罹りやすいことが診断されているか又はされていないかのいずれかであってもよい。いくつかの実施形態では、サンプルは、腫瘍組織生検から得られるポリヌクレオチドサンプルである。がん、腫瘍又は新生物は、一般に、任意のタイプ、例えば、肺、結腸、直腸（又は結腸直腸）、腎臓、乳房、前立腺、若しくは肝臓のがん、腫瘍若しくは新生物、又は本明細書に記載されるがんの他のタイプであってもよい。いくつかの実施形態では、サンプルは、（例えば、化学療法、外科的切除、放射線、又はその組み合わせの後）腫瘍、がん、又は新生物からの寛解状態にある対象から得られる。前記実施形態のいずれかにおいて、前がん、がん、腫瘍、若しくは新生物又は疑われる前がん、がん、腫瘍、若しくは新生物は、膀胱、頭頸部、肺、結腸、直腸、腎臓、乳房、前立腺、皮膚、又は肝臓に属することがある。いくつかの実施形態では、前がん、がん、腫瘍、若しくは新生物又は疑われる前がん、がん、腫瘍、若しくは新生物は、肺に属する。いくつかの実施形態では、前がん、がん、腫瘍、若しくは新生物又は疑われる前がん、がん、腫瘍、若しくは新生物は、結腸又は直腸に属する。いくつかの実施形態では、前がん、がん、腫瘍、若しくは新生物又は疑われる前がん、がん、腫瘍、若しくは新生物は、乳房に属する。いくつかの実施形態では、前がん、がん、腫瘍、若しくは新生物又は疑われる前がん、がん、腫瘍、若しくは新生物は、前立腺に属する。前記実施形態のいずれかにおいて、対象は、ヒト対象であってもよい。いくつかの実施形態では、サンプルは、ステージⅠのがん、ステージⅡのがん、ステージⅢのがん又はステージⅣのがんを有する対象から得られる。

【0129】

サンプルは、対象から単離されている任意の生体サンプルであり得る。サンプルは、身体サンプルであり得る。サンプルは、身体組織、例えば既知の又は疑われた固形腫瘍、全血、血小板、血清、血漿、便、赤血球、白血球（*white blood cells*）若しくは白血球（*leukocytes*）、内皮細胞、組織生検、脳脊髄液、滑液、リンパ液、腹水、間質若しくは細胞外液、細胞間の空間内の体液、例えば、歯肉溝滲出液、骨髓、胸水貯留液、脳脊髄液、唾液、粘液、痰、精液、汗、尿を含み得る。サンプルは、好ましくは、体液、特に血液及びその画分、並びに尿である。サンプルは、対象から最初に単離される形態であり得る、又は細胞などの成分を除去若しくは添加する、若しくは1つの成分を別の成分に対して富化するため、さらなる処理を受けている可能性がある。

【0130】

いくつかの実施形態では、核酸の集団は、新生物、腫瘍、前がん、若しくはがんを有することが疑われる、又は以前に新生物、腫瘍、前がん、若しくはがんと診断された対象か

10

20

30

40

50

らの血清、血漿又は血液サンプルから得られる。集団は、異なるレベルの配列多様性、エピジェネティック変異、及び／又は複製後修飾若しくは転写修飾を有する核酸を含む。複製後修飾は、特に核酸塩基の5位に、シトシンの修飾、例えば、5-メチルシトシン、5-ヒドロキシメチルシトシン、5-ホルミルシトシン及び5-カルボキシルシトシンを含む。

【0131】

サンプルは、対象から単離又は取得し、サンプル分析の場所に輸送することができる。サンプルは、望ましい温度、例えば、室温、4℃、-20℃、及び／又は-80℃で保存し、輸送することができる。サンプルは、サンプル分析の場所で、対象から単離又は取得することができる。対象は、ヒト、哺乳動物、動物、コンパニオン動物、サービス動物、又はペットであり得る。対象は、がん、前がん、感染、移植片拒絶、又は免疫系の変化に関連する他の疾患又は障害を有し得る。対象は、がん又は検出可能ながん症状を有しない場合がある。対象は、1つ以上のがん療法、例えば、化学療法、抗体、ワクチン又は生物剤のいずれか1つ以上で治療されている場合がある。対象は、寛解状態であってもよい。対象は、がん又は任意のがん関連遺伝子の突然変異／障害に罹りやすいと診断されているか又はされていないかのいずれであってもよい。

10

【0132】

血漿の体積は、配列決定された領域の所望のリード深度に依存し得る。例示的な体積は、0.4~40mL、5~20mL、10~20mLである。例えば、体積は、0.5mL、1mL、5mL、10mL、20mL、30mL、又は40mLであり得る。採取された血漿の体積は、5~20mLであってもよい。

20

【0133】

サンプルは、様々な量のゲノム等価物を含む核酸を含み得る。例えば、約30ngのDNAのサンプルは、約10,000(10⁴)の一倍体ヒトゲノム等価物、そして無細胞DNA(cfDNA)の場合、約2000億(2×10¹¹)の個別のポリヌクレオチド分子を含むし得る。同様に、約100ngのDNAのサンプルは、約30,000の一倍体ヒトゲノム等価物、そしてcfDNAの場合、約6000億の個別分子を含むし得る。

【0134】

サンプルは、異なる供給源から、例えば、同じ対象の細胞及び無細胞から、異なる対象の細胞及び無細胞からの核酸を含み得る。サンプルは、突然変異を有する核酸を含み得る。例えば、サンプルは、生殖系列突然変異及び／又は体細胞突然変異を有するDNAを含み得る。生殖系列突然変異は、対象の生殖系列DNA中に存在する突然変異を指す。体細胞突然変異は、対象の体細胞内、例えばがん細胞内に起源を持つ突然変異を指す。サンプルは、がん関連突然変異(例えば、がん関連体細胞突然変異)を有するDNAを含み得る。サンプルは、エピジェネティック変異体(即ち、化学又はタンパク質修飾)を含み得て、ここでエピジェネティック変異体は、がん関連突然変異などの遺伝子変異体の存在に関連する。いくつかの実施形態では、サンプルは、遺伝子変異体の存在に関連するエピジェネティック変異体を含み、ここでサンプルは、遺伝子変異体を含まない。

30

【0135】

サンプルは、無細胞核酸又はcfDNAであってもよく、又はそれを含んでもよい。cfDNAは、例えば上記のような試験対象から得てもよい。例えば、分析用のサンプルは、無細胞核酸を含む血漿又は血清であってもよい。「無細胞DNA」「cfDNA分子」、又は「cfDNA」は、例えば、対象において細胞外形態(例えば、血液、血清、血漿、又は他の体液、例えば、リンパ、脳脊髄液、尿、若しくは痰)で天然に存在するDNA分子を含む。cfDNAが大規模な複雑な生物、例えば哺乳動物内の1つ以上の細胞内に最初から存在した一方で、それは細胞からインビボで生物内に見出される体液への放出を経ており、体液のサンプルを、インビトロ細胞溶解ステップを実施する必要なく取得することによって得ることができる。換言すれば、無細胞核酸又はDNAは、細胞内に含まれない、若しくはそうでなければ細胞に結合されている核酸若しくはDNA、又は無傷

40

50

細胞の除去後にサンプル中に残存している核酸若しくはDNAである。無細胞核酸は、DNA、RNA、及びそのハイブリッド、例えば、ゲノムDNA、ミトコンドリアDNA、siRNA、miRNA、循環RNA (cRNA)、tRNA、rRNA、核小体低分子RNA (snRNA)、Piwi相互作用RNA (piRNA)、長鎖非翻訳RNA (長鎖ncRNA)、又はこれらのいずれかの断片を含む。無細胞核酸は、二本鎖、一本鎖、又はそのハイブリッドであり得る。無細胞核酸は、分泌又は細胞死プロセス、例えば、細胞壊死及びアポトーシスを通じて、体液に放出され得る。一部の無細胞核酸は、がん細胞から体液へ放出され、例えば、循環腫瘍DNA、(ctDNA)が挙げられる。その他は、健常細胞から放出される。いくつかの実施形態では、cfDNAは、無細胞胎児DNA (cffDNA)である。いくつかの実施形態では、無細胞核酸は、腫瘍細胞によって生成される。いくつかの実施形態では、無細胞核酸は、腫瘍細胞と非腫瘍細胞の混合物によって生成される。

10

【0136】

増幅前のサンプル中の無細胞核酸の例示的な量は、約1fg～約1pg、例えば、1pg～200ng、1ng～100ng、10ng～1000ngの範囲である。例えば、その量は、最大で約600ng、最大で約500ng、最大で約400ng、最大で約300ng、最大で約200ng、最大で約100ng、最大で約50ng、又は最大で約20ngの無細胞核酸分子であり得る。量は、少なくとも1fg、少なくとも10fg、少なくとも100fg、少なくとも1pg、少なくとも10pg、少なくとも100pg、少なくとも1ng、少なくとも10ng、少なくとも100ng、少なくとも150ng、又は少なくとも200ngの無細胞核酸分子であり得る。量は、最大で1フェムトグラム (fg)、10fg、100fg、1ピコグラム (pg)、10pg、100pg、1ng、10ng、100ng、150ng、又は200ngの無細胞核酸分子であり得る。方法は、サンプルから1フェムトグラム (fg)～200ngの無細胞核酸分子を得ることを含み得る。

20

【0137】

無細胞DNAは、対象からのその単離時に細胞内に含有されていないDNAを指す。例えば、cfDNAは、細胞を溶解せず、又はそうでなければ細胞内DNAを抽出することなく、無傷細胞を除去した後にサンプル中に残存しているDNAとして、サンプルから単離することができる。無細胞核酸は、DNA、RNA、及びそのハイブリッド、例えば、ゲノムDNA、ミトコンドリアDNA、siRNA、miRNA、循環RNA (cRNA)、tRNA、rRNA、核小体低分子RNA (snRNA)、Piwi相互作用RNA (piRNA)、長鎖非翻訳RNA (長鎖ncRNA)、又はこれらのいずれかの断片を含む。無細胞核酸は、二本鎖、一本鎖、又はそのハイブリッドであり得る。無細胞核酸は、分泌又は細胞死プロセス、例えば、細胞壊死及びアポトーシスを通じて体液中に放出され得る。一部の無細胞核酸は、がん細胞から体液中に放出され、例えば、循環腫瘍DNA、(ctDNA)が挙げられる。その他は、健常細胞から放出される。いくつかの実施形態では、cfDNAは、無細胞胎児DNA (cffDNA)である。いくつかの実施形態では、無細胞核酸は、腫瘍細胞によって生成される。いくつかの実施形態では、無細胞核酸は、腫瘍細胞と非腫瘍細胞の混合物によって生成される。

30

40

【0138】

無細胞核酸は、約100～500ヌクレオチドの代表的なサイズ分布を有し、110～約230ヌクレオチドの分子は分子の約90%を表し、約168ヌクレオチドのモードと240～440ヌクレオチドの範囲内に2番目のマイナーピークとを有する。

【0139】

無細胞核酸は、溶液中で見出される無細胞核酸が体液の無傷細胞及び他の非可溶性成分から分離される場合の分画又は分割ステップを通じて、体液から単離することができる。分割は、遠心分離又は濾過などの技術を含んでもよい。或いは、体液中の細胞を溶解させ、無細胞及び細胞核酸と一緒に処理することができる。一般に、緩衝液の添加及び洗浄ステップ後、核酸をアルコールで沈殿させることができる。さらに、シリカベースのカラム

50

など、クリーンアップステップを使用して、汚染物質又は塩を除去することができる。非特異的な大量の担体の核酸、例えば、C1 DNA、バイサルファイト配列決定、ハイブリダイゼーション、及び／又はライゲーションのためのDNA又はタンパク質を、反応全体を通じて添加して、収率などの手順の特定の態様を最適化することができる。

【0140】

このような処理後、サンプルは、二本鎖DNA、一本鎖DNA及び一本鎖RNAを含む、核酸の様々な形態を含み得る。いくつかの実施形態では、一本鎖DNA及びRNAは、後の処理及び分析ステップに含められるように、二本鎖形態に変換することができる。

【0141】

サンプル中の二本鎖DNA分子及び二本鎖DNA分子に変換されている一本鎖核酸分子は、一方の末端又は両方の末端のいずれかでアダプターに連結することができる。典型的に、二本鎖分子は、全部で4つの標準ヌクレオチドの存在下で、5' - 3' ポリメラーゼ及び3' - 5' エキソヌクレアーゼ（又はプルーフリーディング機能）を有するポリメラーゼによる処理によって平滑末端化される。クレノウ大断片及びT4ポリメラーゼは、好適なポリメラーゼの例である。平滑末端化DNA分子は、少なくとも部分的に、二本鎖アダプター（例えば、Y形状又は釣鐘状のアダプター）とライゲートすることができる。或いは、相補的なヌクレオチドをサンプル核酸及びアダプターの平滑末端に付加して、ライゲーションを促進することができる。本明細書で企図されるのは、平滑末端ライゲーションと粘着末端ライゲーションの両方である。平滑末端ライゲーションでは、核酸分子とアダプタータグの両方が平滑末端を有する。粘着末端ライゲーションでは、典型的に、核酸分子は、「A」オーバーハングを有し、アダプターは、「T」オーバーハングを有する。

【0142】

アダプターへのライゲーション

二本鎖核酸、例えば、サンプル中のDNA分子、及び二本鎖分子に変換されている一本鎖核酸分子は、一方の末端又は両方の末端のいずれかでアダプターに連結することができる。いくつかの場合、DNAは、例えば、DNA分子の末端オーバーハングを伸長させ、アデノシン残基を断片の3'末端に付加し、各DNA断片の5'末端をリン酸化することによって、ライゲート可能になる。典型的に、二本鎖分子は、全部で4つの標準ヌクレオチドの存在下で、5' - 3' ポリメラーゼ及び3' - 5' エキソヌクレアーゼ（又はプルーフリーディング機能）を有するポリメラーゼによる処理によって平滑末端化される。クレノウ大断片及びT4ポリメラーゼは、好適なポリメラーゼの例である。

【0143】

平滑末端DNA分子は、少なくとも部分的に、二本鎖のアダプター（例えば、Y形状又は釣鐘状のアダプター）とライゲートすることができる。或いは、相補的なヌクレオチドをサンプル核酸及びアダプターの平滑末端に付加して、ライゲーションを促進することができる。本明細書で企図されるのは、平滑末端ライゲーションと粘着末端ライゲーションの両方である。平滑末端ライゲーションでは、核酸分子とアダプタータグの両方が平滑末端を有する。粘着末端ライゲーションでは、典型的に、核酸分子は、「A」オーバーハングを有し、アダプターは、「T」オーバーハングを有する。

【0144】

DNAリガーゼ及びアダプターを付加して、サンプル中のDNA分子を一方の末端又は両方の末端でアダプターとライゲートする、即ち、適応したDNAを形成する。本明細書で使用する場合、「アダプター」は、典型的に、少なくとも部分的に二本鎖である短い核酸（例えば、約500未満、約100未満若しくは約50ヌクレオチド未満の長さ、又は末端から末端まで20～30、20～40、30～50、30～60、40～60、40～70、50～60、50～70、20～500、若しくは30～100塩基である）を指し、所与のサンプル核酸分子の末端にライゲートすることができる。いくつかの場合、2つのアダプターを単一のサンプル核酸分子にライゲートし、一方のアダプターをサンプル核酸分子の各末端にライゲートすることができる。

【0145】

アダプターは、両方の末端でアダプターによって隣接されているサンプル核酸分子の増幅を可能にする核酸プライマー結合部位、及び／又は配列決定適用、例えば様々な次世代配列決定（NGS）適用におけるプライマー結合部位を含む、配列決定プライマー結合部位を含み得る。アダプターは、固体支持体にハイブリダイズするための配列、例えば、フローセル配列を含み得る。アダプターはまた、捕捉プローブの結合部位、例えば、フローセル支持体などに付着されたオリゴヌクレオチドを含み得る。アダプターはまた、サンプルインデックス及び／又は分子バーコードを含み得る。これらは、典型的に、増幅プライマー及び配列決定プライマー結合部位に対して、サンプルインデックス及び／又は分子バーコードが所与の核酸分子のアンプリコン及び配列決定リードに含まれるように位置づけられる。同じ又は異なる配列のアダプターは、サンプル核酸分子の各末端に連結され得る。いくつかの場合、同じ又は異なる配列のアダプターは、サンプルインデックス及び／又は分子バーコードがその配列において異なる点を除き、核酸分子の各末端に連結される。いくつかの実施形態では、アダプターは、Y形状アダプターであり、ここで一末端は、本明細書に記載のように平滑末端化又はテール化され、それは、アダプターのテール内に対して1つ以上の相補的なヌクレオチドでさらに平滑末端化又はテール化されている核酸分子に連結する場合である。別の例示的な実施形態では、アダプターは、分析対象の核酸分子に連結するための平滑又はテール末端を含む釣鐘状のアダプターである。他の例示的なアダプターは、Tテール、Cテールアダプター又はヘアピン形状のアダプターを含む。例えば、ヘアピン形状のアダプターは、二本鎖部分が二本鎖ポリヌクレオチドに付着（例えば、ライゲート）可能である場合、相補的な二本鎖部分及びループ部分を含み得る。ヘアピン形状の配列決定アダプターは、ポリヌクレオチド断片の両方の末端に付着して、複数回配列決定可能である環状分子を生成することができる。本開示の方法において使用されるアダプターは、1つ以上の既知の修飾ヌクレオシド、例えば、メチル化ヌクレオシドを含む。2つのアダプターがサンプル核酸に（各末端に1つ）ライゲートされる場合、アダプターの一方又は両方は、1つ以上の既知の修飾ヌクレオシドを含んでもよい。典型的に、プライマー結合部位、配列決定プライマー結合部位、サンプルインデックス及び／又は分子バーコードは、存在する場合、変換手順の結果として塩基対特異性を変更する既知の修飾ヌクレオシドを含まない。

10

20

30

40

【0146】

いくつかの実施形態では、アダプターによって隣接されているサンプル核酸は、PCR及び他の増幅方法によって増幅することができる。こうした増幅は、配列決定などの後続ステップに利用可能なDNAの量を増加させるために用いることができ、再構成された配列を含むDNAの選択的増幅に加えて実施することができる。例えば、この種の増幅は、ライブラリー調製の一部として（例えば、再構成された配列を含むDNAの選択的増幅前）及び／又は（再構成された配列を含むDNAの選択的増幅後となる）標的化ライブラリーの調製後、実施することができる。増幅は、増幅対象のDNA分子に隣接するアダプター内のプライマー結合部位に結合するプライマーによって刺激することができる。増幅方法は、サーモサイクリングから得られる、変性、アニーリング及び伸長のサイクルを含み得る、又は転写媒介性増幅などの場合、等温であり得る。他の増幅方法は、リガーゼ連鎖反応、鎖置換増幅、核酸配列に基づく増幅、及び自己持続的配列に基づく複製を含む。

【0147】

いくつかの実施形態では、本方法では、アダプターへの連結前に二本鎖核酸の少なくとも50、60、70又は80%の増幅をもたらす、Tテールアダプター及びCテールアダプターとのdsDNAライゲーションが実施される。好ましくは、本方法は、増幅された分子の量又は数を、Tテールアダプター単独によって実施される対照方法と比較して、少なくとも10、15又は20%増加させる。

【0148】

いくつかの実施形態では、アダプターをDNA又はそのサブサンプルに付加することができる。アダプターは、本明細書の方法において、いずれかのポイントでDNAにライゲートすることができる。いくつかの実施形態では、アダプターは、捕捉プローブを生成す

50

るためのDNAへのプライマーのアニーリング前に、そのサンプル又はサブサンプルのDNAにライゲートされる。いくつかのそのような実施形態では、アダプターライゲートDNAは、捕捉プローブを生成するためのDNAへのプライマーのアニーリング前に増幅される。いくつかの実施形態では、アダプターは、DNAが捕捉プローブと接触される前に、そのサンプル又はサブサンプルのDNAにライゲートされる。いくつかの実施形態では、アダプターがライゲートされる対象のDNAは、同じサンプル又はサブサンプル中に、捕捉プローブを生成するための鑄型として使用されるDNAとして存在する。いくつかの実施形態では、アダプターがライゲートされる対象のDNAは、捕捉プローブを生成するための鑄型として使用されるDNAと異なるサンプル又はサブサンプル、例えば、第2のサンプル又は第1のサンプルの第2のサブサンプル中に存在する。いくつかの実施形態では、DNAにライゲートされているアダプターは、捕捉プローブによって捕捉される。

10

【0149】

いくつかの実施形態では、捕捉プローブを生成するために使用されるプライマーは、アダプターに対して相補的でなく、それ故、得られる捕捉プローブは、アダプターを含まない。したがって、アダプターライゲートDNAは、アダプターを含まない捕捉プローブの存在下で、選択的に増幅することができる。同様に、アダプターライゲートDNAは、アダプターを含まないDNAから分離することができる。

【0150】

いくつかの実施形態では、開示される方法は、サンプル中のDNAを分析することを含む。そのような方法では、アダプターをDNAに付加することができる。これは、例えば、プライマーの5'部分にアダプターを提供することによって、増幅手順（PCRが使用される場合、これはライブラリーprep-PCR又はLP-PCRと称することができる）と同時に、増幅ステップの前、又は後に行ってもよい。いくつかの実施形態では、アダプターは、他の手法、例えばライゲーションによって付加される。いくつかのそのような方法では、第1のアダプターは、一本鎖DNAへのライゲーションを含み得るライゲーションによって、核酸の3'末端に付加される。いくつかの実施形態では、任意の分割又は捕捉ステップの前に、第1のアダプターは、一本鎖DNAへの（例えば、その3'末端への）ライゲーションを含み得るライゲーションによって、核酸に付加される。いくつかの実施形態では、捕捉プローブは、分割及びライゲーション後に単離することができる。例えば、低メチル化画分は、アダプターとライゲート可能であり、次にライゲートされた低メチル化画分の一部を使用して、再配列のための捕捉プローブを生成することができる。アダプターは、例えば、ユニバーサルプライマー及びDNAポリメラーゼを使用する、二本鎖合成のためのプライミング部位として使用することができる。次に、第2のアダプターは、少なくともこの二本鎖分子の第2鎖の3'末端にライゲートすることができる。いくつかの実施形態では、第1のアダプターは、ビオチンなどのアフィニティータグを含み、第1のアダプターにライゲートされている核酸は、ストレプトアビジンなどのアフィニティータグに対する結合パートナーを含み得る固体支持体（例えば、ビーズ）に結合される。関連手順のさらなる考察においては、Gansauge et al., Nature Protocols 8: 737-748 (2013) を参照されたい。一本鎖核酸に適合できる配列決定ライブラリー調製のための市販キット、例えば、Swift Biosciences製のAccel-NGS（登録商標）Methyl-Seq DNAライブラリーキットが利用可能である。いくつかの実施形態では、アダプターライゲーション後、核酸は増幅される、

20

30

40

【0151】

いくつかの実施形態では、アダプターは、タグの組み合わせの数において、同じ開始及び停止ポイントを有する2つの核酸が低い確率、例えば、95、99又は99.9%でタグの同じ組み合わせを受ける結果となるのに十分な数の異なるタグを含む。アダプターは、同じタグを有するか又は異なるタグを有するかにかかわらず、同じ又は異なるプライマー結合部位を含み得るが、好ましくは、アダプターは同じプライマー結合部位を含む。

【0152】

50

いくつかの実施形態では、アダプターの付着後、核酸は、増幅を受ける。増幅では、例えば、アダプター内のプライマー結合部位を認識するユニバーサルプライマーを使用することができる。

【0153】

いくつかの実施形態では、アダプターの付着後、DNA又はDNAのサブサンプル又は一部分は分割され、DNAを、エピジェネティック修飾を有する核酸に優先的に結合する薬剤と接触させることが含まれる。核酸は、核酸が薬剤への結合からの修飾を有する程度が異なる、少なくとも2つの分割されたサブサンプルに分割される。例えば、薬剤が修飾を有する核酸に対して親和性を有する場合、修飾において（集団における中央値表現と比較して）過剰表現されている核酸は、薬剤に優先的に結合する一方で、修飾において過小表現されている核酸は、薬剤に結合しない、又は薬剤からより容易に溶出される。次に、核酸は、アダプター内のプライマー結合部位に結合しているプライマーから増幅され得る。分割は、アダプター付着前に代わりに実施することができ、いずれの場合でも、アダプターは、いずれの分割が分子で生じたかを特定する成分を含む差次的タグを含んでもよい。

10

【0154】

いくつかの実施形態では、核酸は、プライマー結合部位及びタグを含むY形状アダプターの両方の末端側で連結される。分子は増幅される。

【0155】

分子タグ付け

20

いくつかの実施形態では、サンプルの核酸分子は、サンプルインデックス及び／又は分子バーコード（一般に「タグ」と称される）でタグ付けされ得る。

【0156】

タグ又はインデックスは、タグが会合される相手の分子の特徴を示す情報を含む分子、例えば核酸であり得る。例えば、分子は、（ある1つのサンプル中の分子を異なるサンプル中の分子と区別する）サンプルタグ又はサンプルインデックス、（ある1つの分割における分子を異なる分割における分子と区別する）分割タグ及び／又は（（固有及び非固有両方のタグ付けシナリオにおいて）異なる分子を互いから区別する分子タグ／分子バーコード／バーコードを有し得る。

【0157】

30

タギング戦略は、固有のタギング戦略及び非固有のタギング戦略に分割することができる。固有のタギングでは、サンプル中の分子の全て又は実質的に全ては、異なるタグを有することで、タグ情報単独に基づき、リードを元の分子に割り当てることができる。こうした方法において使用されるタグは、時として「固有タグ」と称される。非固有のタギングでは、同じサンプル中の異なる分子は、同じタグを有し得ることで、タグ情報に加わる他の情報を使用して、配列リードが元の分子に割り当てられる。こうした情報は、開始及び停止座標、分子を位置づける座標、開始又は停止座標単独などを含んでもよい。こうした方法において使用されるタグは、時として、「非固有タグ」と称される。したがって、サンプル中の全ての分子を固有にタグ付けすることは必要でない。それは、サンプル中の同定可能なクラス内に入る分子を固有にタグ付けするのに十分である。したがって、異なる同定可能なファミリー内の分子は、タグ付け分子の同一性についての情報を欠いていない、同じタグを有し得る。

40

【0158】

特定の実施形態では、タグは、1つのバーコード又はバーコードの組み合わせを含み得る。本明細書で使用する場合、「バーコード」という用語は、特定のヌクレオチド配列を有する核酸分子、又は状況に応じて、ヌクレオチド配列それ自体を指す。バーコードは、例えば、10～100ヌクレオチドを有し得る。バーコードの収集物は、特定の目的として所望されるとおり、縮重配列を有し得る、又は特定のハミング距離を有する配列を有し得る。それ故、例えば、分子バーコードは、1つのバーコード又は各々が分子の異なる末端に付着されている2つのバーコードの組み合わせで構成され得る。追加的に、又は代替

50

的に、異なる画分及び／又はサンプルの場合、分子バーコード、分子タグ、又は分子インデックスの異なるセットを使用して、バーコードがそれら個別の配列を通じた分子タグとして役立つだけでなく、それらがメンバーであるセットに基づいてそれらが対応する画分及び／又はサンプルを同定するために役立つことが可能である。

【0159】

いくつかの実施形態では、2つ以上の画分、例えば、各画分は、差次的にタグ付けされる。タグを使用して、個別のポリヌクレオチド集団の画分を標識することで、タグ (tag) (又はタグ (tags)) を特定の画分と関連させることができる。或いは、タグは、分割ステップを用いない本開示の実施形態において使用することができる。いくつかの実施形態では、単一のタグを使用して、特定の画分を標識することができる。いくつかの実施形態では、複数の異なるタグを使用して、特定の画分を標識することができる。複数の異なるタグを使用して、特定の画分を標識する実施形態では、1つの画分を標識するために使用されるタグのセットは、他の画分を標識するために使用されるタグのセットに対して容易に区別することができる。いくつかの実施形態では、タグは、追加的機能を有してもよく、例えば、タグを使用して、サンプル源を示すことができる、又は固有の分子識別子として使用することができる (使用することで、配列決定エラーを突然変異と区別することによって配列決定データの質を改善することができる、例えば、Kind et al., Proc Nat'l Acad Sci USA 108:9530-9535 (2011), Kou et al., PLoS ONE, 11:e0146638 (2016)) などの場合、又は例えば、米国特許第9,598,731号明細書に記載されるとおり、非固有の分子識別子として使用することができる。同様に、いくつかの実施形態では、タグは、追加的機能を有し得る、例えば、タグを使用して、サンプル源を示すことができる、又は非固有の分子識別子として使用することができる (使用することで、配列決定エラーを突然変異と区別することによって配列決定データの質を改善することができる)。

【0160】

タグは、中でも、化学合成、ライゲーション (例えば、上記のとおり、例えば、平滑末端ライゲーション又は粘着末端ライゲーションによる)、又は重複伸長ポリメラーゼ連鎖反応 (PCR) によって、アダプターに組み込むか、又はそうでなければアダプターに連結してもよい。こうしたアダプターは、結局は標的核酸分子に連結される。他の実施形態では、1ラウンド以上の増幅サイクル (例えば、PCR増幅) を適用して、従来の核酸増幅方法を用いて、サンプルインデックスを核酸分子に導入することができる。増幅は、1つ以上の反応混合物 (例えば、アレイ内の複数のマイクロウェル) において実施してもよい。分子バーコード及び／又はサンプルインデックスは、同時に、又は任意の逐次的順序で導入してもよい。いくつかの実施形態では、分子バーコード及び／又はサンプルインデックスは、変換手順の前及び／又は後に導入される。いくつかの実施形態では、分子バーコード及び／又はサンプルインデックスは、配列捕捉ステップ (存在する場合) が実施される前及び／又は後に導入される。いくつかの実施形態では、分子バーコードに限り、プローブ捕捉前に導入され、サンプルインデックスは、配列捕捉ステップが実施された後に導入される。いくつかの実施形態では、分子バーコードとサンプルインデックスの両方は、プローブベースの捕捉ステップ (存在する場合) を実施する前に導入される。いくつかの実施形態では、サンプルインデックスは、配列捕捉ステップ (存在する場合) が実施された後に導入される。いくつかの実施形態では、サンプルインデックスは、重複伸長ポリメラーゼ連鎖反応 (PCR) を通じて組み入れられる。

【0161】

いくつかの実施形態では、タグは、サンプル核酸分子の一方の末端又は両方の末端に位置し得る。いくつかの実施形態では、タグは、所定の又は無作為若しくは準無作為なオリゴヌクレオチド配列である。いくつかの実施形態では、タグは、共に約500、200、100、50、20、10、9、8、7、6、5、4、3、2、又は1ヌクレオチド未満の長さであってもよい。典型的に、タグは、約5~20又は6~15ヌクレオチドの長さ

である。タグは、無作為に又は非無作為にサンプル核酸に連結されてもよい。

【0162】

いくつかの実施形態では、各サンプル又は画分（以下に考察される）は、サンプルインデックス又はサンプルインデックスの組み合わせで固有にタグ付けされる。いくつかの実施形態では、サンプル又はサブサンプルの各核酸分子は、分子バーコード又は分子バーコードの組み合わせで固有にタグ付けされる。他の実施形態では、複数の分子バーコードを使用して、分子バーコードが複数において必ずしも互いに固有でない（例えば、非固有の分子バーコード）ようにしてもよい。これらの実施形態では、分子バーコードは、一般に（例えば、ライゲーションによって）個別分子に付着されることで、分子バーコード及びそれが付着され得る配列の組み合わせにより、個別に追跡可能な固有の配列が作出される。内因性配列情報（例えば、初期（開始）及び／若しくは終了（停止）ゲノム位置／サンプル中の元の核酸分子の配列に対応する位置、サンプル中の元の核酸分子の配列に対応する開始及び停止ゲノム位置、初期（開始）及び／若しくは終了（停止）ゲノム位置／参照配列に位置づけられている配列リードの位置、参照配列に位置づけられている配列リードの開始及び停止ゲノム位置、一方若しくは両方の末端の配列リードの下位配列、配列リードの長さ、並びに／又はサンプル中の元の核酸分子の長さ）と組み合わせた非固有の分子バーコードの検出は、典型的に、固有の同一性の特定分子への割り当てを可能にする。いくつかの実施形態では、初期領域は、参照配列に整列する配列決定リードの5'末端での最初の1、最初の2、最初の5、最初の10、最初の15、最初の20、最初の25、最初の30又は少なくとも最初の30塩基位置を含む。いくつかの実施形態では、末端領域は、参照配列に整列する配列決定リードの3'末端での最終の1、最終の2、最終の5、最終の10、最終の15、最終の20、最終の25、最終の30又は少なくとも最終の30塩基位置を含む。また、個別の配列リードの長さ、又は塩基対の数を任意選択的に使用して、固有の同一性が所与の分子に割り当てられる。それにより、本明細書に記載のとおり、固有の同一性が割り当てられている核酸の一本鎖からの断片は、親鎖、及び／又は相補鎖からの断片のその後の同定を可能にし得る。

【0163】

非固有のタグ付けの特定の実施形態では、使用される異なるタグの数は十分であり得る、特定基の全てのDNA分子が異なるタグを有する可能性は非常に高い（例えば、少なくとも99%、少なくとも99.9%、少なくとも99.99%又は少なくとも99.999%）。バーコードがタグとして使用される場合、またバーコードが例えば無作為に分子の両方の末端に付着される場合、バーコードの組み合わせが一緒にタグを構成し得ることは注目されるべきである。この数は、項において、要件に該当する分子の数の関数である。例えば、クラスは、参照ゲノム上の同じ開始-停止位置に位置している全ての分子であってもよい。クラスは、特定の遺伝子座、例えば、特定の塩基又は特定の領域（例えば、最大で100の塩基又は遺伝子又は遺伝子のエクソン）を通じて位置している全ての分子であってもよい。特定の実施形態では、クラス内の分子の数 z を固有に特定するために使用される異なるタグの数は、 $2 * z$ 、 $3 * z$ 、 $4 * z$ 、 $5 * z$ 、 $6 * z$ 、 $7 * z$ 、 $8 * z$ 、 $9 * z$ 、 $10 * z$ 、 $11 * z$ 、 $12 * z$ 、 $13 * z$ 、 $14 * z$ 、 $15 * z$ 、 $16 * z$ 、 $17 * z$ 、 $18 * z$ 、 $19 * z$ 、 $20 * z$ 又は $100 * z$ のいずれか（例えば、下限）と100、 $000 * z$ 、 $10,000 * z$ 、 $10000 * z$ 又は $100 * z$ のいずれか（例えば、上限）との間であり得る。いくつかの実施形態では、分子バーコードは、識別子のセット（例えば、固有又は非固有の分子バーコードの組み合わせ）のサンプル中の分子に対する予想される比で導入される。フォーマットの一例では、標的分子の両方の末端にライゲートされている、約2～約1,000,000の異なる分子バーコード配列、又は約5～約150の異なる分子バーコード配列、又は約20～約50の異なる分子バーコード配列が使用される。或いは、約25～約1,000,000の異なる分子バーコード配列を使用してもよい。例えば、 $20-50 \times 20-50$ の分子バーコード配列（即ち、 $20-50$ の異なる分子バーコード配列の1つが標的分子の各末端に付着され得る）を使用することができる。識別子のかかる数は、典型的に、識別子の異なる組み合わせを受ける高い確率（

例えば、少なくとも94%、99.5%、99.99%、又は99.999%)を有するように、同じ開始及び停止ポイントを有する異なる分子にとって十分である。いくつかの実施形態では、分子の約80%、約90%、約95%、又は約99%は、分子バーコードの同じ組み合わせを有する。

【0164】

例えば、約5 ng～30 ngの無細胞DNAのサンプル中で、特定のヌクレオチド座標に位置づけるためにおよそ3000個の分子、また同じ停止座標を共有するために任意の開始座標を有する約3～10個の分子が予想される。したがって、約50～約50,000の異なるタグ(例えば、約6～220のバーコードの組み合わせ)は、全てのかかる分子を固有にタグ付けするのに十分であり得る。ヌクレオチド座標を通じて位置づけている全部で3000個の分子を固有にタグ付けするため、約100万～約2000万の異なるタグが必要となる。

10

【0165】

いくつかの実施形態では、反応物における固有又は非固有の分子バーコードの割り当ては、例えば、米国特許出願公開第20010053519号明細書、米国特許出願公開第20030152490号明細書、及び米国特許出願公開第20110160078号明細書、及び米国特許第6,582,908号明細書、米国特許第7,537,898号明細書、米国特許第9,598,731号明細書、及び米国特許第9,902,992号明細書(これらのそれぞれは、その全体が参照により本明細書に援用される)に記載されている方法及びシステムを用いて実施される。或いは、いくつかの実施形態では、サンプルの異なる核酸分子は、内因性配列情報(例えば、開始及び/若しくは停止位置、配列の一方若しくは両方の末端の下位配列、並びに/又は長さ)のみを用いて同定してもよい。

20

【0166】

いくつかの実施形態では、タグ付け核酸は、マイクロウェルプレートへの負荷後、配列決定される。マイクロウェルプレートは、96、384、又は1536マイクロウェルを有し得る。いくつかの場合、それらは、固有のタグ対マイクロウェルの予想される比で導入される。例えば、固有のタグは、ゲノムサンプル当たり、約1、2、3、4、5、6、7、8、9、10、20、50、100、500、1000、5000、10000、50,000、100,000、500,000、1,000,000、10,000,000、50,000,000又は1,000,000,000を超える固有のタグが負荷されるように負荷されてもよい。いくつかの場合、固有のタグは、ゲノムサンプル当たり、約2、3、4、5、6、7、8、9、10、20、50、100、500、1000、5000、10000、50,000、100,000、500,000、1,000,000、10,000,000、50,000,000又は1,000,000,000未満の固有のタグが負荷されるように負荷されてもよい。いくつかの場合、ゲノムサンプル当たり負荷されている固有のタグの平均数は、ゲノムサンプル当たり、約1、2、3、4、5、6、7、8、9、10、20、50、100、500、1000、5000、10000、50,000、100,000、500,000、1,000,000、10,000,000、50,000,000又は1,000,000,000未満の、又はそれを超える固有のタグである。

30

40

【0167】

好ましいフォーマットでは、標的核酸の両方の末端にライゲートされている20～50の異なるタグ(例えば、バーコード)が使用される。例えば、標的分子の両方の末端にライゲートされている35の異なるタグ(例えば、バーコード)は、35のタグに対して1225に相当する、35×35の並べ替えを生じさせる。タグのかかるメンバーは、同じ開始及び停止ポイントを有する異なる分子が、タグの異なる組み合わせを受ける高い確率(例えば、少なくとも94%、99.5%、99.99%、99.999%)を有するのに十分である。他のバーコードの組み合わせは、10～500の任意の数、例えば、約15×15、約35×35、約75×75、約100×100、約250×250、約500×500を含む。

50

【0168】

いくつかの場合、固有のタグは、所定の又は無作為な若しくは準無作為なオリゴヌクレオチド配列であってもよい。他の場合、複数のバーコードを、バーコードが複数において必ずしも互いに固有でないように使用してもよい。この例において、バーコードが個別の分子にライゲートされ得ることで、バーコード及びそれがライゲートされ得る配列の組み合わせにより、個別に追跡可能な固有の配列が作出される。本明細書に記載のとおり、配列リードの初期（開始）及び終了（停止）部分の配列データと組み合わせた非固有のバーコードの検出は、固有の同一性の特定分子への割り当てを可能にし得る。また、個別の配列リードの長さ又は塩基対の数を使用して、固有の同一性をこうした分子に割り当てることができる。それにより、本明細書に記載のとおり、固有の同一性が割り当てられている核酸の一本鎖からの断片は、親鎖からの断片のその後の同定を可能にし得る。

10

【0169】

いくつかの実施形態では、方法は、内部対照DNAを増幅するため、1つ以上の内部対照DNA並びにフォワード及びリバースプライマーを付加することを含む。内部対照DNAは、再配列ブレイクポイントの上流及び下流をアニールするプライマーを使用する増幅前に付加してもよい。内部対照DNAを増幅するためのフォワード及びリバースプライマーは、再配列ブレイクポイントの上流及び下流をアニールするプライマーとともに含めてもよい、又はそれと同時に付加してもよい。内部対照DNAは、対象のゲノム中に存在しない、又は対象がメンバーである種のゲノム（例えば、ヒトゲノム）中に存在しない配列を含んでもよい、又はそれらから構成されてもよい。内部対照DNAを増幅するためのフォワード及び／又はリバースプライマーは、対象のゲノム、例えばヒトゲノム中の任意の配列に相補的でない配列を含んでもよい。内部対照DNAを使用することで、増幅プロセスが設計通りに進行したことを保証することができる。こうしたことから、方法は、1つ以上の内部対照DNAから増幅されている分子及び／又はそれによって捕捉されている分子を検出する（例えば、配列決定する）ことを含んでもよい。方法は、内部対照DNAの量（例えば、内部対照DNA配列に対応する、検出されている分子又はリードの数）を所定の閾値と比較することと、所定の閾値が満たされない場合に配列決定結果を拒絶するか又は所定の閾値が満たされる場合に配列決定結果を受け入れることと、を含み得る。所定の閾値は、例えば、ヒストリカルデータに基づき、又は試験対象、例えば健常ボランティアからのDNAのサンプルに対して方法を試験することによって確立することができる。例えば、1つ以上の内部対照DNAの増幅及び検出は、増幅プロセスが適切に進行し、それ故、偽陰性の可能性が低減されることの確証をもたらす。

20

30

【0170】

変換手順

本明細書で開示される方法に記載されるとおり、アダプター内の品質管理ヌクレオシドの使用は、修飾ヌクレオシドの塩基対特異性を変換する変換手順（例えば、置換されたボラン還元剤によるTet補助変換）又は非修飾ヌクレオシドの塩基対特異性を変換する変換手順（例えば、バイサルファイト変換）によって有利に用いることができる。例えば、いくつかの実施形態では、2つ以上の品質管理ヌクレオシドを含有するアダプターを含む分子が、品質管理ヌクレオシドの塩基対特異性を変更するために選択される変換手順に曝露される場合、品質管理ヌクレオシドの第1の部分（例えば、少なくとも1つ）の塩基対特異性は変更されるが、アダプター内の品質管理ヌクレオシドの第2の部分（例えば、少なくとも1つ）の塩基対特異性は影響を受けず、それは準最適な変換を示し得る。本明細書で開示される方法に記載されるとおり、アダプター内の品質管理ヌクレオシドの使用は、DNAサンプル中の修飾ヌクレオシドの偽陰性検出及び／若しくは同定（即ち、塩基を修飾されていないとして不正確に同定する）並びに／又はDNAサンプル中の修飾ヌクレオシドの偽陽性検出及び／若しくは同定（即ち、塩基を修飾されているとして不正確に同定する）を予測／推測／指示するために有利に用いることができる。修飾ヌクレオシドの偽陽性検出の発生を検出するための使用を意図した本明細書に記載される品質管理ヌクレオシドは、「偽陽性品質管理ヌクレオシド」と称してもよい。修飾ヌクレオシドの偽陰性

40

50

検出の発生を検出するための使用を意図した本明細書に記載される品質管理ヌクレオシドは、「偽陰性品質管理ヌクレオシド」と称してもよい。修飾状態を有するヌクレオシドであって、特定の変換手順に曝露されるときにその塩基対特異性に変更されないことを意味するものは、いくつかの場合、「保護された」ヌクレオシド又は「保護された修飾状態」又は類するものを有すると称してもよい。

【0171】

修飾ヌクレオシドの塩基対特異性を変換する変換手順を用いて偽陰性を検出する場合、アダプター内の品質管理ヌクレオシドは、変換手順／準最適な変換の変換効率が測定でき、それ故、偽陰性の頻度が予測できるような修飾ヌクレオシドを含んでもよい。準最適な変換は、通常変換する変換手順；例えば、Tet 酵素及びピリジンボラン（例えば、TAPS 中）による準最適な変換において使用される試薬が、一部であるが全部でない 5mC 及び 5hmC のジヒドロウラシルへの変換をもたらすようなタイプの全てに満たないヌクレオシドの変換を指す。準最適 (sub-optimal) 及び準最適な (suboptimal) という用語は、等価な意味を有する。準最適な変換はまた、完全反応における変換手順によって変換されているはずである一部のヌクレオシド（修飾又は非修飾）が実際には変換されなかったという意味で、不完全変換と称してもよい。

10

【0172】

非修飾ヌクレオシドの塩基対特異性を変換する変換手順を用いて偽陽性を検出する場合、アダプター内の品質管理ヌクレオシドは、修飾ヌクレオシドの誤った変換頻度が測定でき、それ故、偽陽性の頻度が予測できるような修飾ヌクレオシドを含んでもよい。誤った変換は、典型的に変換手順によって変換されるヌクレオシド以外のヌクレオシドの変換を指す。（典型的に非修飾シトシンのみを変換する）バイサルファイト変換によるメチル化シトシンの変換は、誤った変換の例である。

20

【0173】

修飾ヌクレオシドの塩基対特異性を変換する変換手順を用いて偽陽性を検出する場合、アダプター内の品質管理ヌクレオシドは、非修飾ヌクレオシドの誤った変換頻度が測定でき、それ故、偽陽性の頻度が予測できるような非修飾ヌクレオシドを含んでもよい。

【0174】

非修飾ヌクレオシドの塩基対特異性を変換する変換手順を用いて偽陽性を検出する場合、アダプター内の品質管理ヌクレオシドは、変換手順／準最適な変換の変換効率が測定でき、それ故、偽陽性の頻度が予測できるような非修飾ヌクレオシドを含んでもよい。

30

【0175】

ヌクレオシドの修飾状態に基づく、ヌクレオシドの塩基対特異性を変更する変換手順に応じて、修飾ヌクレオシドを検出し、且つ／又は同定する様々な方法が存在する。次に、配列決定により、塩基対特異性のこれらの変更が検出でき、それ故、ヌクレオシドの修飾状態が推測できる。

【0176】

いくつかの場合、本開示の方法において使用される変換手順は、修飾ヌクレオシド（例えば、メチル化シトシン）の塩基対特異性を変更するが、対応する非修飾ヌクレオシド（例えば、シトシン）の塩基対特異性を変更しない、又は任意の非修飾ヌクレオシド（例えば、シトシン、アデノシン、グアノシン及びチミジン（又はウラシル））の塩基対特異性を変更しないものである。非修飾ヌクレオシドの塩基対特異性を変換しない方法の利点は、配列複雑性の喪失の減少、より高い配列決定効率及びアラインメント欠損の減少を含む。加えて、TAPS などの方法は、いくつかの場合、破壊性がより低く（cf DNA などの低収率のサンプルにとって特に重要である）、変性を必要とせず、非変換のエラーが理論的に無作為である可能性がより高いことを意味することから、バイサルファイト配列決定及び EM-seq などの方法よりも好ましいことがある（図 1 参照）。変換のために変性を必要とする方法では、DNA 分子を変性しないことから、DNA 分子中の全ての塩基の非変換がもたらされる。メチル化における生物学的変化が主に目的の局在化領域に協調されることから、これらの非無作為（局在化）変換は、偽陰性（非メチル化領域）として

40

50

認めることができる。無作為な非変換方法は、領域内の低いパーセントの塩基に最大限影響する可能性があり、それ故、メチル化変化検出の特異性は、メチル化／非メチル化されている領域内の塩基の％で閾値を設けることによって最大化される（偽陽性を低減する）可能性がある。それ故、いくつかの場合、変性を含まない変換手順が好ましい。

【0177】

いくつかの実施形態では、アダプターは、（例えば、修飾された、例えばメチル化された）第1の修飾状態を有する第1の品質管理ヌクレオシド及び（例えば、修飾されていない）第2の修飾状態を有する第2の品質管理ヌクレオシドを含む。そのようなアダプターを使用して、準最適な変換及び誤った変換の両方を検出することができる。

【0178】

図3は、TAPS塩基変換手順を受けているDNAの偽陰性及び／又は偽陽性検出を監視するための品質管理方法の実施形態を示す。（例えば、分子バーコードにおいて）5mCを含有するアダプターは、DNAにライゲートされ、次にTAPS変換手順を受けることで、メチル化シトシン（「T」としての配列リード）の塩基対合が変更され、非メチル化シトシン（さらに「C」としてのリード）の場合は変更されない。各鎖は、配列決定される。準最適な変換を受けた分子が同定され、少なくともメチル化の判定を目的として、濾過・除去され得る。いくつかの実施形態では、準最適な変換を受けた分子は、0／2のバーコード5mCがTに変換される場合のssDNA分子を含む。サンプル変換率は、全ての変換されたバーコード5mCをバーコード5mCの合計で割ることによって計算することができる（この場合、 $31 / 36 = 86\%$ のTAPS変換率）。

【0179】

他の場合、本開示の方法において使用される変換手順は、非修飾ヌクレオシド（例えば、シトシン）の塩基対特異性を変更するが、対応する修飾ヌクレオシド（例えば、メチル化シトシン）の塩基対特異性を変更しないものである。こうした方法は、例えば、バイサルファイト配列決定を含む。

【0180】

当業者は、好適な方法を、どのヌクレオシド修飾が検出及び／又は同定されるべきであるかを含む、その必要性に応じて選択することができる。

【0181】

いくつかの実施形態では、変換手順は、修飾ヌクレオシドを変換する。いくつかの実施形態では、修飾ヌクレオシドを変換する変換手順は、置換されたボラン還元剤によるTet補助変換を含み、任意選択的に、置換されたボラン還元剤は、2-ピコリンボラン、ボランピリジン、tert-ブチルアミンボラン、アンモニアボラン又はピリジンボランである。置換されたボラン還元剤変換によるTet補助pic-ボラン変換では、TETタンパク質を使用して、mC及びhmCが、非修飾Cに影響することなく、caCに変換される。次いで、caC、及びfCは、存在する場合、2-ピコリンボラン（pic-ボラン）又はボランピリジン、tert-ブチルアミンボラン、若しくはアンモニアボランなどの別の置換されたボラン還元剤による処理により、やはり非修飾Cに影響することなく、ジヒドロウラシル（DHU）に変換される。例えば、Liu et al., Nature Biotechnology 2019; 37: 424-429（例えば、付録の図1及び付録の注7）を参照されたい。DHUは、配列決定においてTとして解読される。したがって、このタイプの変換が用いられる場合、第1の核酸塩基は、mC、fC、caC、又はhmCの1つ以上を含み、第2の核酸塩基は、非修飾シトシンを含む。変換されたDNAの配列決定では、非修飾Cの位置であるシトシンとして解読されている位置が同定される。一方で、Tとして解読されている位置は、T、mC、fC、caC、又はhmCであるとして同定される。それ故、TAP変換の実施により、得られた配列リードを用いて、非修飾Cを含有する位置の特定が容易になる。それ故、これらの実施形態では、方法において使用されるアダプター内の品質管理ヌクレオシドは、メチル-シトシン（5mC）及び／又はヒドロキシメチル-シトシン（hmC）を含む。この手順は、Liu et al., 2019、上記にさらに詳細に記載されている、Tet補助ピリジンボラン

配列決定 (TAPS) を包含する。この方法では、Tet 酵素を使用して、5mC 及び 5hmC を 5fC 又は 5CaC に次第に酸化して、次にピリジンボランが 5fC、5CaC を DHU に脱アミノ化する (T として増幅される)。この手順は、Liu et al., 2019、上記にさらに詳細に記載されている、Tet 補助ピリジンボラン配列決定 (TAPS) を包含する。

【0182】

或いは、(例えば、βGT を用いる) hmC の保護は、例えば、上記のとおり、置換されたボラン還元剤による Tet 補助変換と組み合わせることができる。この方法 (TAPS-β) では、hmC は、例えば、ghmC を形成する (5-グルコシルヒドロキシメチルシトシンを形成する) β-グルコシルトランスフェラーゼ (βGT) を用いるグルコシル化を通じて、変換から保護することができる。これは、Yuet al., Cell 2012; 149: 1368-80 に記載されている。次に、mTet1 などの TET タンパク質による処理は、mC を caC に変換するが、C 又は ghmC を変換しない。次に、caC は、pic-ボラン又はボランピリジン、tert-ブチルアミンボラン、若しくはアンモニアボランなどの別の置換されたボラン還元剤による処理によって DHU に変換されるが、やはり非修飾 C 又は ghmC に影響しない。変換された DNA の配列決定では、hmC 又は非修飾 C のいずれかの位置であるシトシンとして解読されている位置が同定される。一方、T として解読されている位置は、T、fC、caC、又は mC であるとして同定される。それ故、本明細書に記載されるサンプルに対する TAPS β 変換の実施により、非修飾 C 又は hmC を含有する位置を、一方では mC を含有する位置と、得られた配列リードを用いて区別することが容易になる。それ故、これらの実施形態では、方法において使用されるアダプター内の品質管理ヌクレオシドは、典型的に、メチル-シトシン (5mC) 及びヒドロキシメチル-シトシン (hmC) の両方を含んでもよい。これは、2つのステップのそれぞれの効率を別々に決定することを可能にする。例えば、アダプターの配列決定が 5mC 及び hmC ヌクレオシドの両方が塩基対特異性を変換していることを示す場合、これは hmC の保護ステップが無効であったことを示す。5mC ヌクレオシドも hmC ヌクレオシドのいずれも塩基対特異性を変換していない場合、これは (少なくとも) Tet 補助変換が無効であったことを示す。アダプター内の mC ヌクレオシドの塩基対合が塩基対特異性を変換しているが、hmC ヌクレオシドの場合には変換していない場合、両方のステップが有効であった。或いは、ちょうど一方のステップ又は他方の効率は、ちょうど mC 又は hmC をアダプター内の品質管理ヌクレオシドとして含めることによって判定することができた。変換の効率及び/又は保護ステップが無効であったか否かの判定を用いて、本明細書で開示される方法において定義する場合、準最適な変換が存在するか否かを判定することができる。このタイプの変換の例示的説明については、例えば、Liu et al., Nature Biotechnology 2019; 37: 424-429 を参照されたい。

【0183】

修飾ヌクレオシドの準最適な変換のために制御することに加えて、アダプター内の品質管理ヌクレオシドを使用して、偽陽性 (即ち、修飾されていると誤って分類されているヌクレオシド) を予測することもできる。この場合、アダプター内の品質管理ヌクレオシドは、適切な変換手順の場合、非修飾 C を含む。アダプターの配列決定で、品質管理ヌクレオシドが塩基対特異性を変換していることを示す場合、これは非修飾塩基 (例えば、非修飾 C) が誤って変換されていることを示す。次に、この情報を用いて、DNA サンプル中の修飾ヌクレオシド (例えば、修飾 C) の偽陽性検出を予測することができる。

【0184】

いくつかの実施形態では、修飾ヌクレオシドを変換する変換手順は、置換されたボラン還元剤による化学的な補助変換を含み、任意選択的に、置換されたボラン還元剤は、2-ピコリンボラン、ボランピリジン、tert-ブチルアミンボラン、ボランピリジン又はアンモニアボランである。置換されたボラン還元剤による化学的な補助変換では、(ox-B S 変換における使用にも適する) 過ルテニウム酸カリウム (KRuO₄) などの酸化

剤を使用して、h m C が f C に特異的に酸化される。p i c - ボラン又はボランピリジン、t e r t - ブチルアミンボラン、若しくはアンモニアボランなどの別の置換されたボラン還元剤による処理は、f C 及び c a C を D H U に変換するが、m C 又は非修飾 C に影響しない。変換された D N A の配列決定により、m C 又は非修飾 C のいずれかの位置であるシトシンとして解読されている位置が同定される。一方で、T として解読されている位置は、T、f C、c a C、又は h m C であるとして同定される。それ故、本明細書に記載されるこのタイプの変換の実施により、一方では非修飾 C 又は m C を含有する位置を、得られた配列リードを用いて、h m C を含有する位置と区別することが容易になる。それ故、これらの実施形態において、そして偽陰性（即ち、修飾されていないと誤って分類されているヌクレオシド）について制御する場合、方法において使用されるアダプターは、ヒドロキシメチルシトシン（h m C）を含む。このタイプの変換の例示的説明については、例えば、L i u e t a l., N a t u r e B i o t e c h n o l o g y 2 0 1 9 ; 3 7 : 4 2 4 - 4 2 9 を参照されたい。

10

20

30

40

50

【0185】

修飾シトシンの塩基対特異性を変更する例示的な変換手順が記載されている。しかし、本明細書に記載の方法では、原則として、任意の修飾ヌクレオシド及び修飾ヌクレオシドの塩基対特異性を変更する好適な変換手順（即ち、一塩基のエピジェネティック変換アッセイ）を使用することができたが、それにより、配列決定の際、修飾塩基を、対応する非修飾ヌクレオシド及び／又は修飾の他のタイプと区別することを可能にする。例えば、任意の変換手順を使用して、N⁶-メチルアデニン（m A）、N⁶-ヒドロキシメチルアデニン（h m A）、又は N⁶-ホルミルアデニン（f A）のいずれか1つを、N⁶-メチルアデニン（m A）、N⁶-ヒドロキシメチルアデニン（h m A）、又は N⁶-ホルミルアデニン（f A）の任意の他の1つ以上、及び／又は対応する非修飾ヌクレオシド、即ち、非修飾アデノシンと区別することを可能にできた。

【0186】

いくつかの実施形態では、変換手順は、非修飾ヌクレオシドを変換する。いくつかの実施形態では、非修飾ヌクレオシドを変換する変換手順は、バイサルファイト変換を含む。バイサルファイトによる処理は、非修飾シトシン及び特定の修飾シトシンヌクレオチド（例えば、5-ホルミルシトシン（f C）又は5-カルボキシルシトシン（c a C））をウラシルに変換する一方で、他の修飾シトシン（例えば、5-メチルシトシン、5-ヒドロキシメチルシトシン（h y d r o x y l m e t h y l c y s t o s i n e））は変換されない。したがって、バイサルファイト変換が用いられる場合、変換された核酸塩基は、バイサルファイトによる影響を受ける、非修飾シトシン、5-ホルミルシトシン、5-カルボキシルシトシン、又は他のシトシン形態の1つ以上を含むことが推測される。変換されていない核酸塩基は、m C 及び h m C の1つ以上を含むことが推測される。バイサルファイト処理された D N A の配列決定では、m C 又は h m C 位置であるシトシンとして解読されている位置が同定される。一方で、T として解読されている位置は、T 又は C のバイサルファイト感受性形態、例えば、非修飾シトシン、5-ホルミルシトシン、又は5-カルボキシルシトシンであるとして同定される。それ故、本明細書に記載されるバイサルファイト変換の実施により、m C 又は h m C を含有する位置を特定することが容易になる。バイサルファイト変換の例示的説明については、例えば、M o s s e t a l., N a t C o m m u n. 2 0 1 8 ; 9 : 5 0 6 8 を参照されたい。

【0187】

いくつかの実施形態では、非修飾ヌクレオシドを変換する手順は、酸化バイサルファイト（O x - B S）変換を含む。この手順では、最初に h m C を f C（バイサルファイト感受性である）に変換し、続いてバイサルファイト変換を行う。したがって、酸化バイサルファイト変換が用いられる場合、変換された核酸塩基は、バイサルファイトによる影響を受ける、非修飾シトシン、f C、c a C、h m C、又は他のシトシン形態の1つ以上を含むことが推測される。変換されていない核酸塩基は、m C を含むことが推測される。O x - B S 変換された D N A の配列決定では、m C 位置であるシトシンとして解読されている

位置が同定される。一方で、Tとして解読されている位置は、T、hmC、又はCのバイサルファイト感受性形態、例えば、非修飾シトシン、fC、又はhmCであるとして同定される。それ故、Ox-B S変換の実施により、mCを含有する位置を特定することが容易になる。酸化バイサルファイト変換の例示的説明については、例えば、Booth et al., Science 2012; 336: 934-937を参照されたい。

【0188】

いくつかの実施形態では、非修飾ヌクレオシドを変換する手順は、Tet補助バイサルファイト (TAB) 変換を含む。TAB変換では、hmCが変換から保護され、mCがバイサルファイト処理に先立って酸化されることで、mCによって最初に占められた位置がUに変換される一方で、hmCによって最初に占められた位置がシトシンの保護形態として残る。例えば、Yuet al., Cell 2012; 149: 1368-80に記載のとおり、β-グルコシルトランスフェラーゼを使用して、(5-グルコシルヒドロキシメチルシトシン (ghmC) を形成する) hmCを保護することができ、次にmTet 1などのTETタンパク質を使用して、mCをcaCに変換することができ、次にバイサルファイト処理を使用して、C及びcaCをUに変換することができる一方で、ghmCは影響を受けないままである。したがって、TAB変換が使用される場合、変換された核酸塩基は、バイサルファイトによる影響を受ける、非修飾シトシン、fC、caC、mC、又は他のシトシン形態の1つ以上を含むことが推測される。変換されていない核酸塩基は、hmCを含むことが推測される。TAB変換されたDNAの配列決定では、hmC位置であるシトシンとして解読されている位置が同定される。一方で、Tとして解読されている位置は、T、mC、又はCのバイサルファイト感受性形態、例えば、非修飾シトシン、fC、又はcaCであるとして同定される。それ故、本明細書に記載される第1のサブサンプルに対するTAB変換の実施により、hmCを含有する位置を特定することが容易になる。

【0189】

いくつかの実施形態では、非修飾ヌクレオシドを変換する変換手順は、APOBEC連結エピジェネティック (ACE) 変換を含む。ACE変換では、APOBEC3A (A3A) などのAID/APOBECファミリーDNAデアミナーゼ酵素を使用して、非修飾シトシン及びmCが脱アミノ化されるが、hmC、fC、又はcaCは脱アミノ化されない。ACE変換されたDNAの配列決定では、hmC、fC、又はcaC位置であるシトシンとして解読されている位置が同定される。一方で、Tとして解読されている位置は、T、非修飾C、又はmCであるとして同定される。それ故、本明細書に記載されるACE変換の実施により、hmCを含有する位置を、第1のサブサンプルから得られた配列リードを用いて、mC又は非修飾Cを含有する位置と区別することが容易になる。ACE変換の例示的説明については、例えば、Schutsky et al., Nature Biotechnology 2018; 36: 1083-1090を参照されたい。

【0190】

いくつかの実施形態では、変換手順は、修飾シトシンの酵素的保護とそれに続く保護されていないシトシンのウラシルへの脱アミノ化 (次いでTとして解読される) を含む。修飾シトシンの酵素的保護及び脱アミノ化の前、ヘアピンがDNAにライゲートされ、ここでヘアピンは、相補的DNA鎖に連結する。次に、DNA鎖は分離され、欠損している鎖が合成される。アダプターがDNAにライゲートされ、ヘアピンが開かれ、PCR増幅がなされる。ヘアピンの各側は、解読され、DNAの同じストレッチを表し、分解ルールのセットは、リードを、A、非修飾C、修飾C、G、又はTの1つに分解する。この変換手順の例示的説明については、例えば、Fullgrabe et al., bioRxiv 2022; 07.08.499285を参照されたい。

【0191】

特定の実施形態では、本開示の方法は、任意の配列コンテキスト内に存在しないメチル化シトシン (即ち、CpG及びCpHシトシン) を同定するための品質管理方法を提供することにおいて有用性を有する。メチル化CpH又は非CpGシトシンはまれであり、そ

れ故、確実に検出するための高レベルの感度を必要とする。加えて、メチル化非 CpG と同位置にあるメチル化 CpG は、準最適な分子変換の指標として非 CpG シトシンのメチル化状態を用いる方法によって検出することができない。本開示の方法は、特定の修飾状態を有することが知られている品質管理ヌクレオシドを提供することによってこれを達成し、それ故、誤った変換及び／又は準最適な変換の頻度の信頼できる尺度を提供する。

【0192】

いくつかの実施形態では、本開示の方法は、配列多様性及び／又は断片化パターンの分析を含み、配列多様性及び／又は断片化パターンの分析からの品質管理ヌクレオシドの準最適な又は誤った変換を伴う適応した DNA を除外しない。例えば、方法は、配列多様性の存在若しくは不在を検出し、且つ／又は断片化パターンを判定することを含み得て、ここで品質管理ヌクレオシドの準最適な又は誤った変換を意味する、品質管理ヌクレオシドを含む適応した DNA は、配列多様性の存在若しくは不在を検出し、且つ／又は断片化パターンを判定することに含まれる。このようにして、本方法は、修飾ヌクレオシド（例えば、mC 又は hmC）の検出における偽陰性及び／又は偽陽性の可能性を、準最適な又は誤った変換に起因する、その目的に適さない適応した DNA を除去することによって低減し得る一方で、こうした適応した DNA を、（準最適な又は誤った変換による影響を受けない）配列多様性及び／又は断片化パターンの分析のために保持し、それ故、感受性への影響を回避することができる。

【0193】

DNA の分析及び／又は分割

いくつかの場合、異質性核酸サンプルが、2 つ以上の画分（サブサンプル）に分割される。いくつかの実施形態では、各画分は、差次的にタグ付けされる。次に、タグ付けされた画分は、集団サンプルの prep 及び／又は配列決定のため、一緒にプールすることができる。分割－タグ付け－プーリングステップは、2 回以上行うことができ、分割の各ラウンドは、異なる特徴に基づいて行い、タグ付けでは、他の画分及び分割手段と区別されている差次的タグを使用する。

【0194】

分割するために使用することができる特徴の例としては、配列の長さ、メチル化レベル、ヌクレオソーム結合、配列ミスマッチ、免疫沈降、及び／又は DNA に結合するタンパク質が挙げられる。得られる画分は、以下の核酸形態の 1 つ以上を含み得る：一本鎖 DNA (ssDNA)、二本鎖 DNA (dsDNA)、より短い DNA 断片及びより長い DNA 断片。いくつかの実施形態では、一般にシトシン修飾（例えば、シトシンメチル化）又はメチル化に基づく分割が実施され、任意選択的に、DNA の前述の特徴又は形態のいずれかに基づき得る、少なくとも 1 つの追加的な分割ステップと組み合わせられる。いくつかの実施形態では、核酸の異質な集団が、1 つ以上のエピジェネティック修飾の存在下及び 1 つ以上のエピジェネティック修飾の不在下で核酸に分割される。エピジェネティック修飾の例としては、メチル化の存在又は不在；メチル化のレベル；メチル化のタイプ（例えば、5-メチルシトシンとメチル化の他のタイプ、例えば、アデニンのメチル化及び／又はシトシンのヒドロキシメチル化）；並びに 1 つ以上のタンパク質、例えばヒストンとの会合及び会合のレベルが挙げられる。代替的又は追加的に、核酸の異質な集団は、ヌクレオソームと会合した核酸分子及びヌクレオソームを欠く核酸分子に分割することができる。代替的又は追加的に、核酸の異質な集団は、一本鎖 DNA (ssDNA) 及び二本鎖 DNA (dsDNA) に分割してもよい。代替的に、又は追加的に、核酸の異質な集団は、核酸の長さ（例えば、最大 160 bp の分子及び 160 bp より大きい長さを有する分子）に基づいて分割してもよい。

【0195】

いくつかの場合、異なる手順を異なる画分に適用して、初期サンプルの異なる特徴が判定される。少なくとも 1 つの画分の核酸、例えば DNA は、本明細書に記載される開示の方法に従って変換手順を受ける。いくつかの実施形態では、少なくとも 1 つの画分は、変換手順を受けない。変換された画分及び変換されていない画分からの対応する配列を比較

して、変換を受けている一塩基を同定し、それ故、初期サンプル中の対応する修飾ヌクレオシドを同定することができる。

【0196】

分割ステップを含む方法の場合、分割タグ（1つの画分中の分子を異なる画分中の分子と区別する）をアダプター内に含めてもよく、又はサンプル分子に付加してもよい。

【0197】

いくつかの実施形態では、2つ以上の画分、例えば、各画分は、差次的にタグ付けされる。タグを使用して、個別のポリヌクレオチド集団の画分を標識することで、タグ（tag）（又はタグ（tags））を特定の画分と相関させることができる。いくつかの実施形態では、単一のタグを使用して、特定の画分を標識することができる。いくつかの実施形態では、複数の異なるタグを使用して、特定の画分を標識することができる。複数の異なるタグを使用して、特定の画分を標識する実施形態では、1つの画分を標識するために使用されるタグのセットは、他の画分を標識するために使用されるタグのセットに対して容易に区別することができる。いくつかの実施形態では、タグは、追加的機能を有してもよく、例えば、タグを使用して、サンプル源を示すことができる、又は固有の分子識別子として使用することができる（使用することで、配列決定エラーを突然変異と区別することによって配列決定データの質を改善することができる、例えば、Kind et al., Proc Nat'l Acad Sci USA 108:9530-9535 (2011), Kou et al., PLoS ONE, 11:e0146638 (2016)）などの場合、又は例えば、米国特許第9,598,731号明細書に記載されるとおり、非固有の分子識別子として使用することができる。同様に、いくつかの実施形態では、タグは、追加的機能を有し得る、例えば、タグを使用して、サンプル源を示すことができる、又は非固有の分子識別子として使用することができる（使用することで、配列決定エラーを突然変異と区別することによって配列決定データの質を改善することができる）。

10

20

【0198】

いくつかの実施形態では、分割タグ付けは、各画分中の分子を分割タグでタグ付けすることを含む。画分を再結合させ（例えば、必要な配列決定ランの数を低減し、不必要なコストを回避するため）、分子を配列決定した後、分割タグは、源の画分を同定する。別の実施形態では、異なる画分は、例えば、バーコードのペアから構成される、分子タグの異なるセットでタグ付けされる。このようにして、各分子バーコードは、源の画分を示すとともに、画分中の分子を識別するのに有用である。例えば、35のバーコードの第1のセットを使用して、第1の画分中の分子をタグ付けすることができる一方で、35のバーコードの第2のセットを使用して、第2の画分中の分子をタグ付けすることができる。

30

【0199】

いくつかの実施形態では、分割及び分割タグによるタグ付け後、分子は、単一ランでの配列決定のためにプールしてもよい。いくつかの実施形態では、例えば、分割タグの付加及びプーリングに続くステップにおいて、サンプルタグが分子に付加される。サンプルタグは、単一の配列決定ランにおける配列決定のため、複数のサンプルから生成されている材料をプールすることを容易にし得る。

40

【0200】

或いは、いくつかの実施形態では、分割タグは、サンプル及び画分と相関させることができる。単純な例としては、第1のタグは、第1のサンプルの第1の画分を示すことができ；第2のタグは、第1のサンプルの第2の画分を示すことができ；第3のタグは、第2のサンプルの第1の画分を示すことができ；そして第4のタグは、第2のサンプルの第2の画分を示すことができる。

【0201】

タグが、1つ以上の特徴に基づいて既に分割されている分子に付着され得る一方で、ライブラリー内の最終的なタグ付け分子は、その特徴をもはや有しないことがある。例えば、一本鎖DNA分子が分割及びタグ付けされ得る一方で、ライブラリー内の最終的なタグ

50

付け分子は、二本鎖である可能性が高い。同様に、DNAが最終ライブラリー内のメチル化の異なるレベルに基づいて分割を受け得る一方で、これら分子由来のタグ付け分子は、メチル化されない可能性が高い。したがって、ライブラリー内の分子に付着されているタグは、典型的に、最終的なタグ付け分子の由来元である「親分子」の特徴を示し、必ずしもタグ付け分子自体の特徴を示さない。

【0202】

例としては、バーコード1、2、3、4などを使用して、第1の画分中の分子がタグ付け及び標識され；バーコードA、B、C、Dなどを使用して、第2の画分中の分子がタグ付け及び標識され；バーコードa、b、c、dなどを使用して、第3の画分中の分子がタグ付け及び標識される。差次的にタグ付けされた画分は、配列決定前にプールすることができる。差次的にタグ付けされた画分は、例えば、Illuminaシーケンサーの同じフローセルにおいて、別々に配列決定する、又は一緒に同時に配列決定することができる。

10

【0203】

配列決定後、リードの分析は、画分毎のレベル、及び全DNA集団レベルで実施することができる。タグを使用して、リードが異なる画分から選別される。分析は、配列情報、ゲノム配座の長さ、カバー率、及び／又はコピー数を用いて、遺伝子及びエピジェネティック変異（メチル化、クロマチン構造などの1つ以上）を判定するためのインシリコ分析を含み得る。いくつかの実施形態では、より高いカバー率は、ゲノム領域内のより高いヌクレオソーム占有率と相関し得る一方で、より低いカバー率は、より低いヌクレオソーム占有率又はヌクレオソーム枯渇領域（NDR）と相関し得る。

20

【0204】

本明細書で開示される方法は、サンプル中のDNAを分析することを含む。本明細書に記載されるいくつかの実施形態では、開示される方法は、DNAを分割することを含む。こうした方法では、DNAの異なる形態（例えば、高メチル化及び低メチル化DNA）は、DNAの1つ以上の特徴に基づき、物理的に分割することができる。この手法を用いて、例えば、特定の配列が高メチル化されるか又は低メチル化されるかを判定することができる。いくつかの実施形態では、サンプルの第1のサブサンプル又は一定分量は、本明細書中の他の箇所に記載される捕捉プローブを生成するためのステップを受け、サンプルの第2のサブサンプル又は一定分量は、分割を受ける。いくつかの実施形態では、サンプル又はそのサブサンプル若しくは一定分量は、分割及び差次的タグ付けとそれに続く再構成された配列に対する捕捉プローブ及び任意選択的には例えば配列変数及び／又はエピジェネティック標的領域に対する追加的な捕捉プローブを用いる捕捉ステップを受ける。

30

【0205】

メチル化のプロファイリングは、ゲノムの異なる領域を通じたメチル化パターンを判定することを含み得る。例えば、分子をメチル化の程度（例えば、1分子当たりのメチル化核酸塩基の相対数）に基づいて分割し、配列決定した後、異なる画分中の分子の配列は、参照ゲノムに位置づけることができる。これは、他の領域と比較して、より高度にメチル化されるか又は大して高度にメチル化されないゲノムの領域を示し得る。このようにして、ゲノム領域は、個別の分子に対して、そのメチル化の程度が異なり得る。

40

【0206】

サンプル中の核酸分子の分割により、例えば、サンプルの1つの画分中でより優位である希少な核酸分子を富化することによって、希少なシグナルが増加し得る。例えば、高メチル化DNA中に存在するが、低メチル化DNA中により少なく存在する（又は存在しない）遺伝的変異は、サンプルを高メチル化及び低メチル化核酸分子に分割することによって、より容易に検出することができる。サンプルの複数の画分を分析することにより、単一分子の多次元分析を実施することができ、それ故、より大きい感受性を達成することができる。分割は、1つ以上のメチル化核酸塩基の存在又は不在に基づき、核酸分子を画分又はサブサンプルに物理的に分割することを含んでもよい。サンプルは、画分又はサブサンプルに、差次的な遺伝子発現又は病態を意味する特徴に基づいて分割してもよい。サン

50

ブルは、核酸、例えば、無細胞DNA (cfDNA)、非cfDNA、腫瘍DNA、循環腫瘍DNA (ctDNA) 及び無細胞核酸 (cfNA) の分析中、正常及び病的状態の間のシグナルの差異をもたらす特徴、又はその組み合わせに基づいて分割してもよい。

【0207】

いくつかの実施形態では、高メチル化及び／又は低メチル化可変エピジェネティック標的領域を分析して、それらが腫瘍細胞又は通常は分析中のDNAサンプル (cfDNAなど) に寄与しないタイプの細胞、及び／又は特定の免疫細胞型の差次的メチル化の特徴を示すか否かが判定される。

【0208】

いくつかの場合、サンプル中の異質なDNAは、2つ以上の画分 (例えば、少なくとも3、4、5、6又は7つの画分) に分割される。いくつかの実施形態では、各画分は、差次的にタグ付けされる。次に、タグ付けされた画分は、集団サンプルのprep及び／又は配列決定のために一緒にプールすることができる。分割－タグ付け－プーリングステップは、2回以上行うことができ、分割の各ラウンドは、異なる特徴に基づいて行い (本明細書に提供される例)、タグ付けでは、他の画分及び分割手段と区別されている差次的タグを使用する。その他の場合、差次的にタグ付けされた画分は、別々に配列決定される。

【0209】

いくつかの実施形態では、差次的にタグ付けされ、プールされたDNAからの配列リードは、取得されて、インシリコで分析される。タグを使用して、リードが異なる画分から選別される。遺伝子変異体を検出するための分析は、画分毎のレベル、及び全核酸集団レベルで実施することができる。例えば、分析は、各画分におけるCNV、SNV、インデル、核酸の融合などの遺伝子変異体を判定するためのインシリコ分析を含み得る。いくつかの場合、インシリコ分析は、クロマチン構造を判定することを含み得る。例えば、配列リードのカバー率を用いて、クロマチン中のヌクレオソームポジショニングを判定することができる。より高いカバー率は、ゲノム領域内のより高いヌクレオソーム占有率と相関し得る一方で、より低いカバー率は、より低いヌクレオソーム占有率又はヌクレオソーム枯渇領域 (NDR) と相関し得る。

【0210】

分割に使用可能である特徴の例としては、配列の長さ、メチル化レベル、ヌクレオソーム結合、配列ミスマッチ、免疫沈降、及び／又はDNAに結合するタンパク質が挙げられる。得られる画分は、以下の核酸形態の1つ以上を含み得る：一本鎖DNA (ssDNA)、二本鎖DNA (dsDNA)、より短いDNA断片及びより長いDNA断片。いくつかの実施形態では、一般にシトシン修飾 (例えば、シトシンメチル化) 又はメチル化に基づく分割が実施され、任意選択的に、DNAの前述の特徴又は形態のいずれかに基づき得る、少なくとも1つの追加的な分割ステップと組み合わせられる。いくつかの実施形態では、核酸の異質な集団が、1つ以上のエピジェネティック修飾の存在下及び1つ以上のエピジェネティック修飾の不在下で核酸に分割される。エピジェネティック修飾の例としては、メチル化の存在又は不在；メチル化のレベル；メチル化のタイプ (例えば、5-メチルシトシンとメチル化の他のタイプ、例えば、アデニンのメチル化及び／又はシトシンのヒドロキシメチル化)；並びに1つ以上のタンパク質、例えばヒストンとの会合及び会合のレベルが挙げられる。代替的又は追加的に、核酸の異質な集団は、ヌクレオソームと会合した核酸分子及びヌクレオソームを欠く核酸分子に分割することができる。代替的又は追加的に、核酸の異質な集団は、一本鎖DNA (ssDNA) 及び二本鎖DNA (dsDNA) に分割してもよい。代替的に、又は追加的に、核酸の異質な集団は、核酸の長さ (例えば、最大160bpの分子及び160bpより大きい長さを有する分子) に基づいて分割してもよい。

【0211】

サンプル中の核酸の集団を分割するために使用される薬剤は、親和性薬剤、例えば、所望の特異性を有する抗体、天然結合パートナー又はその変異体 (Bock et al., Nat Biotech 28:1106-1114 (2010); Song et

10

20

30

40

50

a l. , Nat Biotech 29 : 68 - 72 (2011)) 、又は所与の標的に
対する特異性を有するように、例えばファージディスプレイによって選択される人工ペプ
チドであり得る。いくつかの実施形態では、分割において使用される薬剤は、修飾核酸塩
基を認識する薬剤である。いくつかの実施形態では、薬剤によって認識される修飾核酸塩
基は、修飾シトシン、例えばメチルシトシン（例えば、5-メチルシトシン）である。いく
つかの実施形態では、薬剤によって認識される修飾核酸塩基は、DNA中の第1の核酸
塩基に、サンプルのDNA中の第2の核酸塩基と異なるように影響する手順の産物である
。いくつかの実施形態では、修飾核酸塩基は、その塩基対特異性が手順によって変更され
たことを意味する、「変換された核酸塩基」であってもよい。例えば、特定の手順は、非
メチル化又は非修飾シトシンをジヒドロウラシルに変換し、又はより一般に、シトシンの
少なくとも1つの修飾又は非修飾形態は、脱アミノ化を受け、ウラシル（DNAとの関連
において修飾核酸塩基と考えられる）又はウラシルのさらなる修飾形態をもたらす。分割
剤の例としては、抗体、例えば、修飾シトシン、例えばメチルシトシン（例えば、5-メ
チルシトシン）であり得る修飾核酸塩基を認識する抗体が挙げられる。いくつかの実施形
態では、分割剤は、5-メチルシトシン以外の修飾シトシン、例えば、5-カルボキシル
シトシン（5 c a C）を認識する抗体である。代替的な分割剤は、MeCP2などのタン
パク質を含む、本明細書に記載のようなメチル結合ドメイン（MBD）及びメチル結合タン
パク質（MBP）を含む。

10

【0212】

追加的な分割剤の非限定例が、ヒストンに結合している核酸を遊離又は非結合核酸から
分離し得る、ヒストン結合タンパク質である。本明細書で開示される方法において使用可
能であるヒストン結合タンパク質の例としては、RBBP4、RbAp48及びSANT
ドメインペプチドが挙げられる。

20

【0213】

いくつかの実施形態では、分割は、二成分分割及び修飾の程度／レベルに基づく分割の
両方を含み得る。例えば、メチル化断片は、メチル化DNA免疫沈降（MeDIP）によ
って分割することができ、又は全てのメチル化断片は、メチル結合ドメインタンパク質（
例えば、MethylMinderメチル化DNA富化キット（ThermoFisher
Scientific））を使用して、非メチル化断片から分割することができる。その
後、追加的な分割は、異なるレベルのメチル化を有する断片を、メチル結合ドメイン及
び結合断片を有する溶液中の塩濃度を調節することによって溶出させることを含み得る。
塩濃度が増加すると、より大きいメチル化レベルを有する断片が溶出される。

30

【0214】

DNAの分析は、目的のDNAを検出又は定量化することを含んでもよい。DNAの分
析は、遺伝子変異体及び／又はエピジェネティックな特徴（例えば、DNAメチル化及び
／又はDNA断片化）を検出することを含み得る。

【0215】

いくつかの実施形態では、メチル化レベルは、分割、バイサルファイト変換などの修飾
感受性の変換、配列決定中の直接検出、メチル化感受性制限酵素の消化、メチル化依存性
制限酵素の消化、又は任意の他の好適な手法を用いて判定することができる。例えば、D
NAの異なる形態（例えば、高メチル化及び低メチル化DNA）は、DNAの1つ以上の
特徴に基づいて物理的に分割することができる。例えば、メチル化DNA結合タンパク質
（例えば、MBD2、MBD4、又はMeCP2などのMBD）又は5-メチルシトシン
に特異的な抗体（MeDIPなどの場合）を使用して、DNAを分割することができる。
この手法を使用して、例えば、特定の配列が高メチル化されるか又は低メチル化されるか
を判定することができる。いくつかの実施形態では、DNA分子、例えばcfDNA分子
の終点及び／又は中央点に基づき、DNA断片化パターンを判定することができる。

40

【0216】

いくつかの場合、最終画分は、修飾の異なる程度（修飾の過剰表現又は過小表現）を有
する核酸において富化される。過剰表現及び過小表現は、集団内の鎖当たりの修飾の中央

50

値数に対する、核酸によって生じている修飾の数によって定義することができる。例えば、サンプル中の核酸における5-メチルシトシン残基の中央値数が2である場合、3つ以上の5-メチルシトシン残基を含む核酸は、この修飾において過剰表現され、1又は0の5-メチルシトシン残基を有する核酸は、過小表現される。親和性分離の効果は、結合相中の修飾において過剰表現されている核酸及び非結合相中（即ち、溶液中）の修飾において過小表現されている核酸を富化することである。結合相中の核酸は、後続処理の前に溶出させることができる。

【0217】

MeDIP又はMethylMiner（登録商標）メチル化DNA富化キット（ThermoFisher Scientific）を使用するとき、連続的溶出を用いて、様々なレベルのメチル化を分割することができる。例えば、低メチル化画分（非メチル化）は、核酸集団を、磁気ビーズに付着されている、キットからのMBDと接触させることによって、メチル化画分から分離することができる。ビーズを使用して、メチル化核酸が非メチル化核酸から分離・除去される。その後、1つ以上の溶出ステップを連続的に実施することで、異なるレベルのメチル化を有する核酸が溶出される。例えば、メチル化核酸の第1のセットは、160mM以上、例えば、少なくとも150mM、少なくとも200mM、300mM、400mM、500mM、600mM、700mM、800mM、900mM、1000mM、又は2000mMの塩濃度で溶出することができる。こうしたメチル化核酸が溶出された後、磁気分離を再び使用して、より高いレベルのメチル化核酸が、メチル化のレベルがより低い核酸から分離される。溶出及び磁気分離ステップを繰り返して、低メチル化画分（非メチル化を含む核酸において富化されている）、メチル化画分（低いレベルのメチル化を含む核酸において富化されている）、及び高メチル化画分（高いレベルのメチル化を含む核酸において富化されている）などの様々な画分を作出することができる。

【0218】

いくつかの方法では、親和性分離に基づく分割に使用される薬剤に結合されている核酸は、洗浄ステップを受ける。洗浄ステップでは、親和性薬剤に弱く結合している核酸が洗い流される。こうした核酸は、修飾を有する核酸において、平均値又は中央値（即ち、サンプルの薬剤との初期接触時の、固相に結合されたままである核酸と固相に結合していない核酸との間の中間）に近い程度まで富化することができる。

【0219】

親和性分離は、異なる程度の修飾を有する核酸の、少なくとも2つ、時として3つ以上の画分をもたらす。画分が依然として分離される一方で、少なくとも1つの画分、及び通常は2つ又は3つ（以上）の画分の核酸は、通常はアダプターの成分として提供されている核酸タグに連結されることで、異なる画分中の核酸は、一方の画分のメンバーを他方と識別する異なるタグを受ける。同じ画分の核酸分子に連結されているタグは、互いに同一又は異なる可能性がある。しかし、互いに異なる場合、タグは、それらのコードの一部を共有することで、それらが特定の画分であるように付着されている分子を特定することができる。

【0220】

メチル化などの特徴に基づく核酸サンプルの分割に関するさらなる詳細については、参照により本明細書に援用される国際公開第2018/119452号パンフレットを参照されたい。

【0221】

いくつかの実施形態では、分割は、DNAをメチル化感受性制限酵素（MSRE）及び／又はメチル化依存性制限酵素（MDRE）と接触させることを含む。DNAのMSRE又はMDREによる処理後、DNAをサイズに基づいて分割して、高メチル化（MSRE処理後の最長のDNA分子及びMDRE処理後の最短のDNA断片）、中間（MSRE又はMDRE処理後の中間長のDNA分子）、及び低メチル化（MSRE処理後の最短のDNA分子及びMDRE処理後の最長のDNA断片）サブサンプルを生成することができる。

。

【0222】

いくつかの実施形態では、分割は、核酸をメチル結合タンパク質（「MBP」）のメチル結合ドメイン（「MBD」）と接触させることによって実施される。いくつかのそのような実施形態では、核酸は、全MBPと接触される。いくつかの実施形態では、MBDは、5-メチルシトシン（5mC）に結合し、MBPは、MBDを含み、本明細書で互換的にメチル結合タンパク質又はメチル結合ドメインタンパク質と称される。いくつかの実施形態では、MBDは、ビオチンリンカーを介して、常磁性ビーズ、例えば、Dynabeads（登録商標）M-280ストレプトアビジンにカップリングされる。異なる程度のメチル化を伴う画分への分割は、NaCl濃度を増加させることにより画分を溶出することによって実施することができる。

10

【0223】

いくつかの実施形態では、結合されたDNAは、抗体又はMBDをプロテアーゼ、例えばプロテイナーゼKと接触させることによって溶出される。これは、上で考察されるとおり、NaClを使用する溶出ステップの代わりに又はそれに加えて実施することができる。

。

【0224】

本明細書で企図される修飾核酸塩基を認識する薬剤の例としては、限定はされないが、以下が挙げられる：

（a）MeCP2は、非修飾シトシンよりも5-メチルシトシンに優先的に結合するタンパク質である。

20

（b）RPL26、PRP8及びDNAミスマッチ修復タンパク質MHS6は、非修飾シトシンよりも5-ヒドロキシメチルシトシンに優先的に結合する。

（c）FOXK1、FOXK2、FOXP1、FOXP4及びFOXI3は、好ましくは、非修飾シトシンよりも5-ホルミルシトシンに結合する（Iurlaro et al., Genome Biol. 14:R119 (2013)）。

（d）1つ以上のメチル化又は修飾核酸塩基又はその変換産物、例えば、5mC、5caC、又はDHUに特異的な抗体。

【0225】

一般に、溶出は、修飾の数、例えば、1分子当たりのメチル化部位の数の関数であり、より多くのメチル化を有する分子が増加した塩濃度下で溶出する。DNAをメチル化の程度に基づいて異なる集団に溶出するため、NaCl濃度を増加させる一連の溶出緩衝液を使用することができる。塩濃度は、約100nM～約2500mMのNaClの範囲であり得る。一実施形態では、プロセスは、3つの画分をもたらす。分子は、第1の塩濃度で溶液と接触され、修飾核酸塩基を認識する薬剤を含む分子を含み、分子は、捕捉部分、例えばストレプトアビジンに付着され得る。第1の塩濃度で、分子の集団は、薬剤に結合することになり、集団は、非結合のままとなる。非結合集団は、「低メチル化」集団として分離され得る。例えば、DNAの低メチル化形態で富化されている第1の画分は、低い塩濃度、例えば、100mM又は160mMで非結合のままであるものである。中程度にメチル化されたDNAにおいて富化されている第2の画分は、中程度の塩濃度、例えば、1000mM～2000mMの濃度を用いて溶出される。これもまた、サンプルから分離される。DNAの高メチル化形態で富化されている第3の画分は、高い塩濃度、例えば少なくとも約2000mMを用いて溶出される。

30

40

【0226】

いくつかの実施形態では、5-メチルシチジン（5mC）に対して産生されているモノクローナル抗体を使用して、メチル化DNAが精製される。DNAは、一本鎖DNA断片を得るため、例えば95℃で変性される。標準又は磁気ビーズに共役されたプロテインG及び抗5mC抗体とのインキュベーション後の洗浄を用いて、抗体に結合したDNAを免疫沈降させる。次に、こうしたDNAを溶出することができる。画分は、沈殿していないDNA及びビーズから溶出した1つ以上の画分を含んでもよい。

50

【0227】

いくつかの実施形態では、DNAの画分は、ライブラリー調製の酵素ステップのための調製において脱塩及び富化される。

【0228】

異常に高いコピー数を含む配列は、高メチル化される傾向があり得る。したがって、いくつかの実施形態では、型特異的な差次的にメチル化された領域及びコピー数多型の両方である複数の標的領域を含むエピジェネティック標的領域セットのメンバーに特異的な捕捉プローブと接触されているDNAは、高メチル化画分の少なくとも一部分を含む。高メチル化画分の少なくとも一部分に由来するか又はそれを含むDNAは、1つ以上の他の画分、例えば、中間の画分又は低メチル化画分の少なくとも一部分に由来するか又はそれを含むDNAと組み合わせてもよく、又は組み合わせなくてもよい。

10

【0229】

いくつかの実施形態では、メチル化は、変換手順を用いて検出される。変換手順は、修飾依存性に、第1の核酸塩基を差次的に変更するが、第2の核酸塩基を差次的に変更しない、例えば、非修飾に対してメチル化されている（又はヒドロキシメチル化、又はホルミル化、又はカルボキシル化など）、且つ／又はある1つの様式で、別の様式に対して修飾されている（例えば、メチル化対ヒドロキシメチル化）、任意の技術を含む。こうした変換手順の例としては、非修飾シトシン及び特定の修飾シトシン（例えば、5-ホルミルシトシン（fC）又は5-カルボキシルシトシン（caC））をウラシルに変換する一方で、他の修飾シトシン（例えば、5-メチルシトシン、5-ヒドロキシメチルシトシン（hydroxymethylcystosine））が変換されない、バイサルファイト変換が挙げられる。バイサルファイト変換の実施は、配列リードを用いて、mC又はhmCを含有する位置を特定することを容易にし得る。バイサルファイト変換の例示的説明については、例えば、Moss et al., Nat Commun. 2018; 9: 5068を参照されたい。

20

【0230】

こうした変換手順の例としては、酸化バイサルファイト（Ox-B S）変換も挙げられる。Ox-B S変換の実施は、配列リードを用いて、mCを含有する位置を特定することを容易にし得る。酸化バイサルファイト変換の例示的説明については、例えば、Booth et al., Science 2012; 336: 934-937を参照されたい。

30

【0231】

こうした変換手順の例としては、Tet補助バイサルファイト（TAB）変換も挙げられる。例えば、Yuet al., Cell 2012; 149: 1368-80に記載のとおり、β-グルコシルトランスフェラーゼを使用して、（5-グルコシルヒドロキシメチルシトシン（ghmC）を形成する）hmCを保護することができ、次にmTet1などのTETタンパク質を使用して、mCをcaCに変換することができ、次にバイサルファイト処理を使用して、C及びcaCをUに変換することができる一方で、ghmCは影響を受けないままである。したがって、TAB変換が使用される場合、第1の核酸塩基は、バイサルファイトによる影響を受ける、非修飾シトシン、fC、caC、mC、又は他のシトシン形態の1つ以上を含み、第2の核酸塩基は、hmCを含む。TAB変換の実施は、配列リードを用いて、hmCを含有する位置を特定することを容易にし得る。

40

【0232】

こうした変換手順の例としては、置換されたボラン還元剤によるTet補助変換も挙げられ、任意選択的に、置換されたボラン還元剤は、2-ピコリンボラン、ボランピリジン、tert-ブチルアミンボラン、又はアンモニアボランである。例えば、Liu et al., Nature Biotechnology 2019; 37: 424-429（例えば、付録の図1及び付録の注7）を参照されたい。TAP変換の実施は、配列リードを用いて、非修飾Cを含有する位置を特定することを容易にし得る。この手順は、Liu et al., 2019、上記にさらに詳細に記載されている、Tet補助ピリジンボラン配列決定（TAPS）を包含する。

50

【0233】

或いは、（例えば、 β GTを使用する）hmCの保護は、置換されたボラン還元剤によるTet補助変換と組み合わせることができる。こうしたTAPSB変換の実施は、配列リードを用いて、非修飾C又はhmCを含有する位置を、一方ではmCを含有する位置と区別することを容易にし得る。このタイプの変換の例示的説明については、例えば、Liu et al., Nature Biotechnology 2019; 37: 424-429を参照されたい。

【0234】

こうした変換手順の例としては、APOBEC連結エピジェネティック（ACE）変換も挙げられる。ACE変換の実施は、配列リードを用いて、hmCを含有する位置を、mC又は非修飾Cを含有する位置と区別することを容易にし得る。ACE変換の例示的説明については、例えば、Schutskey et al., Nature Biotechnology 2018; 36: 1083-1090を参照されたい。

【0235】

こうした変換手順の例としては、核酸塩基の酵素変換、例えば、EM-Seqなどの場合も挙げられる。例えば、www.biorxiv.org/content/10.1101/2019.12.20.884692で利用可能である、Vaisvila R, et al. (2019) EM-seq: Detection of DNA methylation at single base resolution from picograms of DNA. bioRxiv; DOI: 10.1101/2019.12.20.884692を参照されたい。

【0236】

いくつかの実施形態では、メチル化は、メチル化感受性制限酵素（MSRE）を用いて検出される。例えば、サンプル、サブサンプル、又はサンプルの一部は、非メチル化配列を切断するため、1つ以上のMSREによる消化を受け得る。例示的なMSREとして、AatII、AccII、AciI、AorI3HI、AorI5HI、BspT104I、BssHII、BstUI、CfrI0I、ClaI、CpoI、Eco52I、HaeII、HapII、HhaI、Hin6I、HpaII、HpyCH4IV、MluI、MspI、NaeI、NotI、NruI、NsbI、PmaCI、Psp1406I、PvuI、SacII、SalI、SmaI、及びSnaBIが挙げられる。いくつかの実施形態では、少なくとも2つのメチル化感受性ヌクレアーゼが使用される。いくつかの実施形態では、少なくとも3つのメチル化感受性ヌクレアーゼが使用される。いくつかの実施形態では、メチル化感受性ヌクレアーゼは、BstUI及びHpaIIを含む。いくつかの実施形態では、2つのメチル化感受性ヌクレアーゼは、HhaI及びAccIIを含む。いくつかの実施形態では、メチル化感受性ヌクレアーゼは、BstUI、HpaII及びHin6Iを含む。いくつかの実施形態では、1つ以上のMSREと接触されているサンプルの一部は、高メチル化DNAを含む、又は高メチル化DNA画分であるか若しくはそれを含み、それは本明細書中の他の箇所に記載されるとおりに得てもよい。

【0237】

いくつかの実施形態では、DNA断片化は、DNA（例えば、cfDNA）の配列決定された断片の終点及び／又は中間点を判定することによって検出される。例えば、断片化パターンにおける差異は、断片が腫瘍に起源を有するか又は健常細胞に起源を有するかに応じて生じ得る。cfDNAの腫瘍細胞由来DNAを断片化に基づいて検出するため、異常な断片のレベル増加の存在又は不在は、例えば、増加及び異常が対照又は健常サンプルに対する場合、（例えば、増幅の程度に比例する）コピー数増幅を伴う領域で判定することができる。

【0238】

いくつかの実施形態では、サンプル又はサブサンプル（例えば、本明細書に記載されるサンプルを、例えば、メチル化などのシトシン修飾、例えば、シトシンなどの5-メチル

化のレベルに基づいて分割することによって調製されている第1、第2、又は第3のサブサンプル）は、メチル化依存性ヌクレアーゼ又はメチル化感受性ヌクレアーゼと接触される。別段の指示がない限り、分割がシトシン修飾に基づいて実施される場合、第1のサブサンプルは、より高いレベルの修飾を有するサブサンプルであり；第2のサブサンプルは、より低いレベルの修飾を有するサブサンプルであり；また、存在する場合、第3のサブサンプルは、第1及び第2のサブサンプル間の中間レベルの修飾を有する。

【0239】

上で考察したように、分割手順は、サブサンプル中のDNA分子の不完全な選別を生じ得る。メチル化依存性ヌクレアーゼ又はメチル化感受性ヌクレアーゼの選別は、非特異的に分割されたDNAを分解するように行うことができる。例えば、第2のサブサンプルは、メチル化依存性ヌクレアーゼ、例えば、メチル化依存性制限酵素と接触させることができる。これは、第2のサブサンプル中の非特異的に分割されたDNA（例えば、メチル化DNA）を分解して、処理された第2のサブサンプルを生成することができる。代替的に又は追加的に、第1のサブサンプルは、メチル化感受性エンドヌクレアーゼ、例えばメチル化感受性制限酵素と接触させて、それにより第1のサブサンプル中の非特異的に分割されたDNAを分解して、処理された第1のサブサンプルを生成することができる。第1又は第2のサブサンプルのいずれか又は両方における非特異的に分割されたDNAの分解は、例えば、サンプル中の異常に修飾されたDNAの存在を検出する、DNAの起源の組織を判定する、且つ／又は対象ががんを有するか否かを判定するための、シトシン修飾に基づくDNAの正確な分割に依存する方法の性能に対する改善として提起される。例えば、こうした分解は、改善された感受性を提起し、且つ／又は下流分析を簡素化し得る。一般に、非特異的に分割されたDNAが、例えば低メチル化画分において高メチル化される場合、メチル化依存性ヌクレアーゼ、例えばメチル化依存性制限酵素が使用される必要がある。逆に、非特異的に分割されたDNAが、例えば高メチル化画分において低メチル化される場合、メチル化感受性ヌクレアーゼ、例えばメチル化感受性制限酵素が使用される必要がある。メチル化依存性ヌクレアーゼ、例えばメチル化依存性制限酵素は、メチル化DNAを非メチル化DNAに対して優先的に切断する一方で、メチル化感受性ヌクレアーゼ、例えばメチル化感受性制限酵素は、非メチル化DNAをメチル化DNAに対して優先的に切断する。

【0240】

サブサンプルをヌクレアーゼと接触させる場合、1つ以上のヌクレアーゼを使用することができる。いくつかの実施形態では、サブサンプルは、複数のヌクレアーゼと接触される。サブサンプルは、逐次的に又は同時にヌクレアーゼと接触させてもよい。ヌクレアーゼの同時使用は、ヌクレアーゼが不必要なサンプル操作を回避するための類似条件下（例えば、緩衝液組成物）で活性である場合、有利であり得る。第2のサブサンプルを2つ以上のメチル化依存性制限酵素と接触させることにより、非特異的に分割された高メチル化DNAがより完全に分解され得る。同様に、第1のサブサンプルを2つ以上のメチル化感受性制限酵素と接触させることにより、非特異的に分割された低メチル化及び／又は非メチル化DNAがより完全に分解され得る。

【0241】

いくつかの実施形態では、メチル化依存性ヌクレアーゼは、MspJI、LpnPI、FspEI、又はMcrBCの1つ以上を含む。いくつかの実施形態では、少なくとも2つのメチル化依存性ヌクレアーゼが使用される。いくつかの実施形態では、少なくとも3つのメチル化依存性ヌクレアーゼが使用される。いくつかの実施形態では、メチル化依存性ヌクレアーゼは、FspEIを含む。いくつかの実施形態では、メチル化依存性ヌクレアーゼは、例えば逐次的に使用される、FspEI及びMspJIを含む。

【0242】

いくつかの実施形態では、メチル化感受性ヌクレアーゼは、AatII、AccII、AciI、Aorl3HI、Aorl5HI、BspT104I、BssHII、BstUI、Cfr1OI、ClaI、CpoI、Eco52I、HaeII、HapII、H

h a I、H i n 6 I、H p a I I、H p y C H 4 I V、M l u I、M s p I、N a e I、N o t I、N r u I、N s b I、P m a C I、P s p 1 4 0 6 I、P v u I、S a c I I、S a l I、S m a I、及び S n a B I の 1 つ以上を含む。いくつかの実施形態では、少なくとも 2 つのメチル化感受性ヌクレアーゼが使用される。いくつかの実施形態では、少なくとも 3 つのメチル化感受性ヌクレアーゼが使用される。いくつかの実施形態では、メチル化感受性ヌクレアーゼは、B s t U I 及び H p a I I を含む。いくつかの実施形態では、2 つのメチル化感受性ヌクレアーゼは、H h a I 及び A c c I I を含む。いくつかの実施形態では、メチル化感受性ヌクレアーゼは、B s t U I、H p a I I 及び H i n 6 I を含む。

【0243】

いくつかの実施形態では、F s p E I は、少なくとも 1 つのサブサンプル（例えば、低メチル化画分）中の核酸分子を消化するために使用される。いくつかの実施形態では、B s t U I、H p a I I 及び H i n 6 I は、少なくとも 1 つのサブサンプル（例えば、高メチル化画分）中の核酸分子を消化するために使用され、F s p E I は、少なくとも 1 つの他のサブサンプル（例えば、低メチル化画分）中の核酸分子を消化するために使用される。中程度にメチル化された画分を含む実施形態では、その中の核酸分子は、メチル化感受性ヌクレアーゼ又はメチル化依存性ヌクレアーゼで消化され得る。いくつかの実施形態では、中程度にメチル化された画分中の核酸分子は、高メチル化画分と同じヌクレアーゼで消化される。例えば、中程度にメチル化された画分は、高メチル化画分とともにプールすることができ、次にプールされた画分は、消化を受け得る。いくつかの実施形態では、中程度にメチル化された画分中の核酸分子は、低メチル化画分と同じヌクレアーゼで消化される。例えば、中程度にメチル化された画分は、低メチル化画分とともにプールすることができ、次にプールされた画分は、消化を受け得る。

【0244】

いくつかの実施形態では、サブサンプルは、アダプターをタグ付けするか又はアダプターを DNA の両方の末端に付着させるステップ後、上記のヌクレアーゼと接触される。タグ又はアダプターは、上記の手法のいずれかを用いる、ヌクレアーゼによる切断に対して抵抗性であり得る。この手法では、切断産物が両方の末端でタグ又はアダプターを欠いていることから、切断は非特異的に分割された分子に対する分析実行を阻止し得る。

【0245】

或いは、アダプターをタグ付け又は付着させるステップは、上記のヌクレアーゼによる切断後、実施することができる。次に、切断された分子は、ヌクレアーゼ認識部位に対応する末端（タグ又はアダプターへの付着点）を有することに基づき、配列リードにおいて同定することができる。また、分子をこのように処理することで、例えば、体細胞突然変異の観察といった切断された分子からの情報の取得が可能であり得る。サブサンプルをヌクレアーゼと接触後にアダプターをタグ付け又は付着し、c f DNA などの低分子量 DNA が分析中であるとき、接触ステップ前にサンプルから高分子量 DNA（汚染しているゲノム DNA など）を除去することが望ましいことがある。また、比較的低い温度（例えば、65℃以下、又は60℃以下）で熱不活性化され得るヌクレアーゼを使用して、変性がその後のライゲーションステップに干渉し得ることから、DNA の変性を回避することが望ましいことがある。

【0246】

サンプルが、中程度にメチル化された分子を含有する第3のサブサンプルを含む3つのサブサンプルに分割される場合、第3のサブサンプルは、いくつかの実施形態において、メチル化感受性ヌクレアーゼと接触される。こうしたステップは、接触ステップに関連する本明細書中の他の箇所に記載の特徴のいずれかを有してもよく、上で考察したようなアダプターをタグ付け又は付着させるステップの前又は後に実施してもよい。いくつかの実施形態では、第1及び第3のサブサンプルは、メチル化感受性ヌクレアーゼと接触させる前に組み合わせられる。こうしたステップは、接触ステップに関連する本明細書中の他の箇所に記載の特徴のいずれかを有してもよく、上で考察したようなアダプターをタグ付け又

10

20

30

40

50

は付着させるステップの前又は後に実施してもよい。いくつかの実施形態では、第1及び第3のサブサンプルは、組み合わせられる前に、差次的にタグ付けされる。

【0247】

或いは、サンプルが、中程度にメチル化された分子を含有する第3のサブサンプルを含む3つのサブサンプルに分割される場合、第3のサブサンプルは、いくつかの実施形態において、メチル化依存性ヌクレアーゼと接触される。こうしたステップは、接触ステップに関連する本明細書中の他の箇所に記載の特徴のいずれかを有してもよく、上で考察したようなアダプターをタグ付け又は付着させるステップの前又は後に実施してもよい。いくつかの実施形態では、第2及び第3のサブサンプルは、メチル化依存性ヌクレアーゼと接触させる前に組み合わせられる。こうしたステップは、接触ステップに関連する本明細書中の他の箇所に記載の特徴のいずれかを有してもよく、上で考察したようなアダプターをタグ付け又は付着させるステップの前又は後に実施してもよい。いくつかの実施形態では、第2及び第3のサブサンプルは、組み合わせられる前に、差次的にタグ付けされる。

10

【0248】

いくつかの実施形態では、DNAは、例えばSPRIビーズを使用して、ヌクレアーゼと接触された後に精製される。こうした精製は、ヌクレアーゼの熱不活性化後に行ってもよい。或いは、精製は、省略することができ；それ故、例えば、増幅などの後続ステップは、熱不活性化ヌクレアーゼ (heat-inactivated nuclease) を含有するサブサンプルに対して実施することができる。別の実施形態では、接触ステップは、SPRIビーズなどの精製試薬の存在下で行い、例えば、チューブ交換に伴う減少を最小化することができる。切断及び熱不活性化後、SPRIビーズは、分子クラウディング試薬 (例えば、PEG) 及び塩を加えることによって、精製のために再使用することができる。

20

【0249】

いくつかの実施形態では、変換手順がサンプル又はサブサンプルに対して実施される場合、1つ以上の標的領域セット (例えば、少なくともエピジェネティック標的領域セット) のそのサンプル又はサブサンプルからの後続する捕捉において、(例えば、プローブがハイブリダイズする配列内の少なくとも1つの塩基の) 修飾状態に特異的である、例えば、所望のとおり、変換 (例えば、修飾又は非修飾シトシンの、アデニンと優先的に対合する、ウラシル又はその類似体、例えばDHUへの変換) を受けている、又は変換を受けていない標的配列に対して相補的である、プローブを含む捕捉プローブが使用される。そのようなものとして、プローブは、目的の修飾、例えばメチル化が存在した又は不在であった配列に特異的であり得る。いくつかの実施形態では、修飾感受性変換 (modification sensitive conversion) がサンプル又はサブサンプルに対して実施される場合、1つ以上の標的領域セット (例えば、少なくともエピジェネティック標的領域セット) のそのサンプル又はサブサンプルからの後続する捕捉において、修飾状態と無関係に標的配列とハイブリダイズし得るプローブを含む (例えば、修飾又は非修飾シトシンの、アデニンと優先的に対合する (例えば、イノシンはC又はUと対合し得る)、ウラシル又はその類似体、例えばDHUへの変換を受けているか又は受けていない可能性がある位置で乱雑に対合する核酸塩基を含む) 捕捉プローブが使用される。

30

40

【0250】

いくつかの実施形態では、方法は、(低メチル化画分とも称される) 第2のサブサンプルのDNAの少なくとも一部分及び(高メチル化画分とも称される) 第1のサブサンプルのDNAの少なくとも一部分を含むプールを調製することを含む。例えば、エピジェネティック標的領域及び／又は配列可変標的領域を含む標的領域は、プールから捕捉されてもよい。標的領域セットを本明細書中の他の箇所に記載のサブサンプルの少なくとも一部分から捕捉するステップは、第1及び第2のサブサンプルからのDNAを含むプールに対して実施される捕捉ステップを包含する。プール中のDNAを増幅するステップは、標的領域をプールから捕捉する前に実施してもよい。捕捉ステップは、本明細書中の他の箇所に記載の特徴のいずれかを有してもよい。

50

【0251】

エピジェネティック標的領域は、本明細書中の別の箇所で考察のとおり、それらが腫瘍に起源を持つか若しくは健常細胞に起源を持つか、又はそれらの起源である組織のタイプが何であるかに応じて、メチル化レベル及び／又は断片化パターンにおける差異を示し得る。配列可変標的領域は、それらが腫瘍に起源を持つか又は健常細胞に起源を持つかに応じて、配列における差異を示し得る。

【0252】

低メチル化画分からのエピジェネティック標的領域の分析は、いくつかの適用において、高メチル化画分及び低メチル化画分からの配列可変標的領域並びに高メチル化画分からのエピジェネティック標的領域の分析よりも有用でない可能性がある。そのようなものとして、配列可変標的領域及びエピジェネティック標的領域が捕捉中である場合の方法では、後者は、高メチル化画分及び低メチル化画分からの配列可変標的領域並びに高メチル化画分からのエピジェネティック標的領域の1つ以上より少ない程度まで捕捉され得る。例えば、配列可変標的領域は、高メチル化画分とともにプールされていない低メチル化画分の一部分から捕捉することができ、プールは、高メチル化画分からのDNAの一部（例えば、大半、実質的に全て、又は全て）、及び低メチル化画分からのDNAを含めないか、又はその一部（none or some）（例えば、少数）によって調製することができる。こうした手法は、低メチル化画分からのエピジェネティック標的領域を配列決定することを低減又は除去し、それによりさらなる分析にとって十分である配列決定データの量を低減することができる。

【0253】

いくつかの実施形態では、プール中の低メチル化画分のDNAの少数を含めることで、1つ以上のエピジェネティックな特徴（例えば、本明細書中の別の箇所で詳細に考察されているメチル化又は他のエピジェネティックな特徴）の、例えば相対ベースでの定量化が容易になる。

【0254】

いくつかの実施形態では、プールは、低メチル化画分のDNAの少数、例えば、低メチル化画分のDNAの約50%未満、例えば、低メチル化画分のDNAの約45%、40%、35%、30%、25%、20%、15%、10%、又は5%以下を含む。いくつかの実施形態では、プールは、低メチル化画分のDNAの約5%～25%を含む。いくつかの実施形態では、プールは、低メチル化画分のDNAの約10%～20%を含む。いくつかの実施形態では、プールは、低メチル化画分のDNAの約10%を含む。いくつかの実施形態では、プールは、低メチル化画分のDNAの約15%を含む。いくつかの実施形態では、プールは、低メチル化画分のDNAの約20%を含む。

【0255】

いくつかの実施形態では、プールは、高メチル化画分の一部を含み、それは高メチル化画分のDNAの少なくとも約50%であってもよい。例えば、プールは、高メチル化画分のDNAの少なくとも約55%、60%、65%、70%、75%、80%、85%、90%、又は95%を含んでもよい。いくつかの実施形態では、プールは、高メチル化画分のDNAの50～55%、55～60%、60～65%、65～70%、70～75%、75～80%、80～85%、85～90%、90～95%、又は95～100%を含む。いくつかの実施形態では、第2のプールは、高メチル化画分の全て又は実質的に全てを含む。

【0256】

いくつかの実施形態では、方法は、低メチル化画分のDNAの少なくとも一部分を含む第1のプールを調製することを含む。いくつかの実施形態では、方法は、高メチル化画分のDNAの少なくとも一部分を含む第2のプールを調製することを含む。いくつかの実施形態では、第1のプールは、高メチル化画分のDNAの一部をさらに含む。いくつかの実施形態では、第2のプールは、低メチル化画分のDNAの一部をさらに含む。いくつかの実施形態では、第1のプールは、低メチル化画分のDNAの大半、及び任意選択的に

高メチル化画分のDNAの少数を含む。いくつかの実施形態では、第2のプールは、高メチル化画分のDNAの大半及び低メチル化画分のDNAの少数を含む。中程度にメチル化された画分を含むいくつかの実施形態では、第2のプールは、中程度にメチル化された画分のDNAの少なくとも一部分、例えば、中程度にメチル化された画分のDNAの大半を含む。いくつかの実施形態では、第1のプールは、低メチル化画分のDNAの大半を含み、第2のプールは、高メチル化画分のDNAの大半及び中程度にメチル化された画分のDNAの大半を含む。

【0257】

いくつかの実施形態では、方法は、少なくとも標的領域の第1のセットを第1のプールから捕捉することを含み、例えば、第1のプールは、上の実施形態のいずれかに記載されたとおりである。いくつかの実施形態では、第1のセットは、配列可変標的領域を含む。いくつかの実施形態では、第1のセットは、低メチル化可変標的領域及び／又は断片化可変標的領域を含む。いくつかの実施形態では、第1のセットは、配列可変標的領域及び断片化可変標的領域を含む。いくつかの実施形態では、第1のセットは、配列可変標的領域、低メチル化可変標的領域及び断片化可変標的領域を含む。第1のプール中のDNAを増幅するステップは、この捕捉ステップ前に実施してもよい。いくつかの実施形態では、標的領域の第1のセットを第1のプールから捕捉することは、第1のプールのDNAを捕捉プローブの第1のセットと接触させることを含む。いくつかの実施形態では、捕捉プローブの第1のセットは、配列可変標的領域に特異的な標的結合プローブを含む。いくつかの実施形態では、捕捉プローブの第1のセットは、配列可変標的領域、低メチル化可変標的領域及び／又は断片化可変標的領域に特異的な標的結合プローブを含む。

【0258】

いくつかの実施形態では、方法は、標的領域の第2のセット又は標的領域の複数のセットを第2のプールから捕捉することを含み、例えば、第1のプールは、上の実施形態のいずれかに記載されたとおりである。いくつかの実施形態では、第2の複数のセットは、エピジェネティック標的領域、例えば、高メチル化可変標的領域及び／又は断片化可変標的領域を含む。いくつかの実施形態では、第2の複数のセットは、配列可変標的領域及びエピジェネティック標的領域、例えば、高メチル化可変標的領域及び／又は断片化可変標的領域を含む。第2のプール中のDNAを増幅するステップは、この捕捉ステップ前に実施してもよい。いくつかの実施形態では、標的領域の第2の複数のセットを第2のプールから捕捉することは、第1のプールのDNAを捕捉プローブの第2のセットと接触させることを含み、捕捉プローブの第2のセットは、配列可変標的領域に特異的な標的結合プローブ及びエピジェネティック標的領域に特異的な標的結合プローブを含む。いくつかの実施形態では、標的領域の第1のセット及び標的領域の第2のセットは、同一でない。例えば、標的領域の第1のセットは、標的領域の第2のセット内に存在しない1つ以上の標的領域を含んでもよい。代替的に又は追加的に、標的領域の第2のセットは、標的領域の第1のセット内に存在しない1つ以上の標的領域を含んでもよい。いくつかの実施形態では、少なくとも1つの高メチル化可変標的領域は、第2のプールから捕捉されるが、第1のプールから捕捉されない。いくつかの実施形態では、複数の高メチル化可変標的領域は、第2のプールから捕捉されるが、第1のプールから捕捉されない。いくつかの実施形態では、標的領域の第1のセットは、配列可変標的領域を含み、且つ／又は標的領域の第2のセットは、エピジェネティック標的領域を含む。いくつかの実施形態では、標的領域の第1のセットは、配列可変標的領域、及び断片化可変標的領域を含み；且つ標的領域の第2のセットは、エピジェネティック標的領域、例えば、高メチル化可変標的領域及び断片化可変標的領域を含む。いくつかの実施形態では、標的領域の第1のセットは、配列可変標的領域、断片化可変標的領域を含み、低メチル化可変標的領域を含み；且つ標的領域の第2のセットは、エピジェネティック標的領域、例えば、高メチル化可変標的領域及び断片化可変標的領域を含む。

【0259】

いくつかの実施形態では、第1のプールは、低メチル化画分のDNAの大半及び高メチ

ル化画分のDNAの一部（例えば、約半分）を含み、第2のプールは、高メチル化画分のDNAの一部（例えば、約半分）を含む。いくつかのそのような実施形態では、標的領域の第1のセットは、配列可変標的領域を含み、且つ／又は標的領域の第2のセットは、エピジェネティック標的領域を含む。配列可変標的領域及び／又はエピジェネティック標的領域は、本明細書中の他の箇所に記載の実施形態のいずれかに記載のとおりであってもよい。

【0260】

いくつかの実施形態では、DNAの画分は、ライブラリー調製の酵素ステップにおける調製において脱塩及び富化される。構造変異を含む配列は、低メチル化される傾向があり得る。したがって、いくつかの実施形態では、DNAの再構成された配列に対して逆平行方向にアニールする少なくとも2つのプライマーを含む複数のプライマーと接触されているDNAは、低メチル化画分の少なくとも一部分に由来するか又はそれを含む。低メチル化画分の少なくとも一部分に由来するか又はそれを含むDNAは、中程度の画分又は高メチル化画分などの1つ以上の他の画分の少なくとも一部分に由来するか又はそれを含むDNAと組み合わせてもよく、又は組み合わせなくてもよい。

【0261】

増幅

アダプターによって隣接されているサンプル核酸は、PCR及び他の増幅方法によって増幅することができる。増幅は、典型的に、増幅対象のDNA分子に隣接するアダプター内のプライマー結合部位に結合しているプライマーによって刺激される。増幅方法は、サーモサイクリングに起因する、変性、アニーリング及び伸長のサイクルを含み得る、又は転写媒介性増幅などの場合、等温であり得る。他の増幅方法は、リガーゼ連鎖反応、鎖置換増幅、核酸配列に基づく増幅、及び自己持続的配列に基づく複製を含む。

【0262】

いくつかの実施形態では、本方法では、二本鎖核酸の少なくとも50、60、70又は80%の増幅をもたらす、Tテールアダプター及びCテールアダプターとのdsDNAライゲーションが実施される。好ましくは、本方法により、増幅された分子の量又は数が、Tテールアダプター単独によって実施される対照方法と比較して、少なくとも10、15又は20%増加する。

【0263】

いくつかの実施形態では、サンプル核酸は、配列決定前に増幅される。増幅は、変換ステップの前及び／又は後であってもよい。増幅は、いくつかの場合、1つ以上の捕捉ステップの前であってもよい。いくつかの実施形態では、ライゲートは、増幅前に又は増幅と同時にされる。

【0264】

いくつかの実施形態では、増幅は、アダプターオリゴヌクレオチドにおけるプライマー結合部位に結合しているプライマーによって刺激される。アダプターの既知の修飾ヌクレオチドは、プライマー結合部位の外側に存在してもよい、即ち、品質管理方法に使用される既知の修飾ヌクレオチドは、プライマー結合部位に存在しない、又は増幅プライマーによって結合されない。

【0265】

捕捉プローブ；標的領域の富化、捕捉及び使用

サンプル中の核酸は、標的配列を有する分子がその後の分析のために捕捉される場合の捕捉ステップを受け得る。捕捉は、当該技術分野で公知の任意の好適な手法を用いて実施してもよい。標的捕捉は、捕捉部分、例えば、ビオチン又は以下に記される他の例で標識されているオリゴヌクレオチドベイトを含むベイトセットの使用を含み得る。プローブは、遺伝子などの領域のパネルを通じて並べて表示するために選択されている配列を有し得る。こうしたベイトセットは、標的分子のベイトとのハイブリダイゼーションを可能にする条件下でサンプルと組み合わせられる。次に、捕捉された分子は、捕捉部分を用いて単離される。例えば、ビオチン捕捉部分は、ビーズベースのストレプトアビジンを介する。こ

うした方法は、例えば、参照により本明細書に援用される、2017年12月26日交付の米国特許第9,850,523号明細書にさらに記載されている。

【0266】

捕捉部分は、限定はされないが、ビオチン、アビジン、ストレプトアビジン、特定のヌクレオチド配列を含む核酸、抗体によって認識されるハプテン、及び磁気吸引力性粒子を含む。抽出部分は、ビオチン／ストレプトアビジン又はハプテン／抗体などの結合ペアのメンバーであり得る。いくつかの実施形態では、分析物に付着されている捕捉部分は、単離可能な部分、例えば、磁気吸引力性粒子又は遠心分離を通じて沈降され得る大きい粒子に付着されているその結合ペアによって捕捉される。捕捉部分は、捕捉部分を有する核酸の捕捉部分を欠いている核酸からの親和性分離を可能にする分子の任意のタイプであり得る。例示的な捕捉部分は、固相又はオリゴヌクレオチドに連結されているか又は連結可能であるストレプトアビジンに結合することによって親和性分離を可能にするビオチンであり、それは固相に連結されているか又は連結可能である相補的オリゴヌクレオチドに結合することによって親和性分離を可能にする。

【0267】

いくつかの実施形態では、型特異的なエピジェネティック変異を含む領域を含むDNAが捕捉される。いくつかの実施形態では、変異は、健常細胞内に存在するが、通常は血液サンプルなどのサンプルタイプ中に存在しない。いくつかの実施形態では、変異は、異常細胞（例えば、過形成、異形成、又は新生物細胞）内に存在する。

【0268】

いくつかの実施形態では、サンプル又は第1のサブサンプルから捕捉されている第1の捕捉されたエピジェネティック標的領域セットは、高メチル化可変標的領域を含む。いくつかの実施形態では、高メチル化可変標的領域は、1つ以上の関連する細胞型又は組織型からの健常cfDNAにおいて型特異的な高メチル化を示す。任意の特定の理論によって拘束されることは望まないが、がん細胞の存在は、（例えば、がん及び／又は周囲組織からの）血流へのDNAの排出を増加させ得る。そのようなものとして、cfDNAの起源の組織の分布は、発がん時に変化し得る。したがって、第1のサブサンプル中の高メチル化可変標的領域のレベル増加は、がんの存在（又は対象の病歴に応じては、再発）の指標であり得る。

【0269】

いくつかの実施形態では、本明細書中の方法は、第2の捕捉されたエピジェネティック標的領域セットをサンプル又は第2のサブサンプルから捕捉することを含む。いくつかの実施形態では、第2のエピジェネティック標的領域セットは、低メチル化可変標的領域を含む。任意の特定の理論によって拘束されることは望まないが、がん細胞は、同じ組織型の健常細胞よりも、より多くのDNAを血流に排出し得る。そのようなものとして、cfDNAの起源の組織の分布は、発がん時に変化し得る。したがって、第2のサブサンプル中の低メチル化可変標的領域のレベル増加は、がんの存在（又は対象の病歴に応じては、再発）の指標であり得る。

【0270】

加えて、捕捉された標的領域セットは、配列可変標的領域セットに対応するDNAを含んでもよい。捕捉されたセットを組み合わせ、組み合わせられた捕捉されたセットを提供してもよい。

【0271】

配列可変標的領域セット及びエピジェネティック標的領域セットに対応するDNAを含む捕捉されたセットが上で考察されている組み合わせられた捕捉されたセットを含むいくつかの実施形態では、配列可変標的領域セットに対応するDNAは、エピジェネティック標的領域セットに対応するDNAより大きい濃度、例えば、1.1～1.2倍大きい濃度、1.2～1.4倍大きい濃度、1.4～1.6倍大きい濃度、1.6～1.8倍大きい濃度、1.8～2.0倍大きい濃度、2.0～2.2倍大きい濃度、2.2～2.4倍大きい濃度、2.4～2.6倍大きい濃度、2.6～2.8倍大きい濃度、2.8～3.0倍

10

20

30

40

50

大きい濃度、3.0～3.5倍大きい濃度、3.5～4.0倍大きい濃度、4.0～4.5倍大きい濃度、4.5～5.0倍大きい濃度、5.0～5.5倍大きい濃度、5.5～6.0倍大きい濃度、6.0～6.5倍大きい濃度、6.5～7.0倍大きい濃度、7.0～7.5倍大きい濃度、7.5～8.0倍大きい濃度、8.0～8.5倍大きい濃度、8.5～9.0倍大きい濃度、9.0～9.5倍大きい濃度、9.5～10.0倍大きい濃度、10～11倍大きい濃度、11～12倍大きい濃度、12～13倍大きい濃度、13～14倍大きい濃度、14～15倍大きい濃度、15～16倍大きい濃度、16～17倍大きい濃度、17～18倍大きい濃度、18～19倍大きい濃度、19～20倍大きい濃度、20～30倍大きい濃度、30～40倍大きい濃度、40～50倍大きい濃度、50～60倍大きい濃度、60～70倍大きい濃度、70～80倍大きい濃度、80～90倍大きい濃度、又は90～100倍大きい濃度で存在し得る。濃度における差異の程度は、定義セクションで考察のとおり、標的領域のフットプリントサイズに対する正規化を説明する。

10

【0272】

いくつかの実施形態では、捕捉されているDNAは、イントロン領域を含む。いくつかの実施形態では、イントロン領域は、DNAを、新生（例えば、腫瘍又はがん）細胞及び健常細胞、例えば非新生の循環細胞と区別する可能性が高い1つ以上のイントロンを含む。例えば、一部の新生細胞内に存在し、健常細胞では不在であることが知られている再配列を含むイントロンを使用して、DNAを新生（例えば、腫瘍又はがん）細胞及び健常細胞と区別することができる。いくつかの実施形態では、再配列は、転位である。

20

【0273】

いくつかの実施形態では、捕捉されたイントロン領域は、少なくとも30bp、例えば、少なくとも100bp、少なくとも200bp、少なくとも500bp、少なくとも1kb、少なくとも2kb、少なくとも5kb、少なくとも10kb、少なくとも20kb、少なくとも50kb、少なくとも200kb、少なくとも300kb、又は少なくとも400kbのフットプリントを有する。いくつかの実施形態では、イントロン標的領域セットは、30bp～1000kb、例えば、30bp～100bp、100bp～200bp、200bp～500bp、500bp～1kb、1kb～2kb、2kb～5kb、5kb～10kb、10kb～20kb、20kb～50kb、50kb～100kb、100～200kb、200～300kb、300～400kb、400～500kb、500～600kb、600～700kb、700～800kb、800～900kb、及び900～1,000kbの範囲内のフットプリントを有する。

30

【0274】

例示的な再配列、例えば、本明細書に記載の方法を用いて検出可能であるイントロン転位としては、限定はされないが、転位に関与する2つの遺伝子の少なくとも1つが受容体チロシンキナーゼである転位が挙げられる。例示的な転位産物は、BCR-ABL融合体、及びALK、FGFR2、FGFR3、NTRK1、RET、又はROS1のいずれかを含む融合体である。

【0275】

いくつかの実施形態では、捕捉されているDNAは、型特異的なエピジェネティック変異を有する標的領域を含む。いくつかの実施形態では、エピジェネティック標的領域セットは、型特異的なエピジェネティック変異を有する標的領域からなる。いくつかの実施形態では、型特異的なエピジェネティック変異、例えば、差次的メチル化又は型特異的な断片化パターンは、1つ以上の関連する細胞型又は組織型の細胞からのDNAをサンプル中又は対象中に存在する他の細胞型又は組織型からのDNAと区別する可能性が高い。

40

【0276】

いくつかの実施形態では、本明細書に記載の方法を用いて捕捉又は富化されている核酸は、捕捉されたDNA、例えば、1つ以上の捕捉されたDNAのセットを含む。いくつかの実施形態では、捕捉されたDNAは、異なる免疫細胞型において差次的にメチル化されている標的領域を含む。いくつかの実施形態では、免疫細胞型は、希少な又は密接に関連

50

した免疫細胞型、例えば、異なる分化のステージでの活性化及びナイーブリンパ球又は骨髓系細胞を含む。

【0277】

いくつかの実施形態では、サンプル又は第1のサブサンプルから捕捉されている捕捉されたエピジェネティック標的領域セットは、高メチル化可変標的領域を含む。いくつかの実施形態では、高メチル化可変標的領域は、1つ以上の関連する細胞型又は組織型において差次的又は排他的に高メチル化される。いくつかの実施形態では、高メチル化可変標的領域は、1つの細胞型又は1つの免疫細胞型、又はクラスター内の1つの免疫細胞型において差次的又は排他的に高メチル化される。いくつかの実施形態では、高メチル化可変標的領域は、識別可能により高い、又は1つの細胞型若しくは1つの免疫細胞型若しくはクラスター内の1つの免疫細胞型において排他的に存在する程度まで高メチル化される。こうした高メチル化可変標的領域は、他の細胞型又は組織型において高メチル化され得るが、1つ以上の関連する細胞型又は組織型において認められる程度には至らない。いくつかの実施形態では、高メチル化可変標的領域は、健常なcfDNAにおいて、少なくとも1つの他の組織型における場合より低いメチル化を示す。いくつかの実施形態では、高メチル化可変標的領域は、1つ以上の関連する細胞型又は組織型の異常細胞からのcfDNAにおいて、さらにより高いメチル化を示す。いくつかの実施形態では、標的領域は、異常に高いコピー数を有する高メチル化領域を含む。いくつかのそのような実施形態では、標的領域は、健常組織及び罹患した結腸組織において高メチル化され、前がん又は癌性結腸組織において異常に高いコピー数を有する。そのような標的領域の例が下の表1に示される。

【0278】

10

20

30

40

50

【表 1】

表1: 結腸がん又は前がんにおける異常に高いコピー数を伴う高メチル化標的領域

コピー数増加を伴う 染色体領域	染色体領域内にDMRを含む遺伝子
7p21.3-15.1	VWDE, TWIST1, DNAH1, GPNMB, NPY, GSDME, NFE2L3, HOXA1, HOXA7, EVX1, CREB5
8q13.3-24.3	PRDM14, MSC, TRPA1, SBSPON, CRISPLD1, ZFHX4, RALYL, MMP16, C8orf88, RUNX1T1, CDH17, GDF6, SDC2, OSR2, PABPC1, ZNF706, RIMS2, ZFPM2, SYBU, CSMD3, TRPS1, AARD, TNFRSF11B, MAL2, COL14A1, POU5F1B, ADCY8, ST3GAL1, TRAPPC9, AGO2, ADGRB1, LY6E, LY6H, GLI4, MAFA, TSTA3, PLEC, OPLAH, FOXH1
13q14.1-33.3	ELF1, RGCC, TNFSF11, SERP2, RB1, CNMD, PCDH17, EDNRB, POU4F1, SLITRK5, GPC6, SOX21, CLDN10, STK24, ZIC2, NALCN, FGF14, MYO16, COL4A1, COL4A2, SOX1, F7
20q12-13.33	MAFB, LPIN3, EMILIN3, PTPRT, TOX2, GDAP1L1, HNF4A, MMP9, CDH22, PREX1, ZNFX1, PARD6B, ZFP64, DOK5, FAM210B, TFAP2C, RBM38, CTCFL, GNAS, PHACTR3, CDH4, GATA5, BHLHE23, YTHDF1, NKAIN4, CHRNA4, KCNQ2, GMEB2

【 0 2 7 9 】

【表 2】

表2. 肺がん試験に基づく例示的な高メチル化標的領域

遺伝子名	染色体
MARCH11	chr5
TAC1	chr7
TCF21	chr6
SHOX2	chr3
p16	chr3
Casp8	chr2
CDH13	chr16
MGMT	chr10
MLH1	chr3
MSH2	chr2
TSLC1	chr11
APC	chr5
DKK1	chr10
DKK3	chr11
LKB1	chr11
WIF1	chr12
RUNX3	chr1
GATA4	chr8
GATA5	chr20
PAX5	chr9
E-カドヘリン	chr16
H-カドヘリン	chr16

10

20

30

【0280】

いくつかの実施形態では、サンプル又はサブサンプルから捕捉されている捕捉されたエピジェネティック標的領域セットは、低メチル化可変標的領域を含む。いくつかの実施形態では、低メチル化可変標的領域は、1つ以上の関連する細胞型又は組織型において排他的に低メチル化される。いくつかの実施形態では、低メチル化可変標的領域は、1つの細胞型又は1つの免疫細胞型、又はクラスター内の1つの免疫細胞型において排他的に低メチル化される。いくつかの実施形態では、低メチル化可変標的領域は、1つの細胞型又は1つの免疫細胞型、又はクラスター内の1つの免疫細胞型において排他的に存在する程度まで低メチル化される。こうした低メチル化可変標的領域は、他の細胞型又は組織型において低メチル化され得るが、1つ以上の細胞型又は組織型において認められる程度には至らない。いくつかの実施形態では、低メチル化可変標的領域は、健常なcfDNAにおいて、少なくとも1つの他の組織型における場合より高いメチル化を示す。

40

【0281】

任意の特定の理論に拘束されることを望むものではないが、がんを有する個体において

50

、増殖する又は活性化された免疫細胞及び／又は死滅するがん細胞は、それぞれ、健常者における免疫細胞及び／又は同じ組織型の健常細胞よりも多くのDNAを血流に排出し得る。そのようなものとして、cfDNAの起源の細胞型及び／又は組織の分布は、発がん時に変化し得る。したがって、特定の細胞型又は組織型に起源を持つcfDNAの存在及び／又はレベルは、疾患の指標であり得る。高メチル化及び／又は低メチル化における変異は、疾患の指標であり得る。例えば、分割ステップ後のサブサンプル中の高メチル化可変標的領域及び／又は低メチル化可変標的領域のレベル増加は、がんの存在（又は対象の病歴に応じては、再発）の指標であり得る。

【0282】

様々な細胞型の間を区別するのに有用である、限定はされないが、免疫細胞型を含む、例示的な高メチル化可変標的領域及び低メチル化可変標的領域は、様々な細胞型から得られたDNAを、例えば、Scott, C. A., Duryea, J. D., MacKay, H. et al., "Identification of cell type-specific methylation signals in bulk whole genome bisulfite sequencing data," *Genome Biol* 21, 156 (2020) (doi.org/10.1186/s13059-020-02065-5)に記載されるような全ゲノムバイサルファイト配列決定を介して分析することによって同定された。全ゲノムバイサルファイト配列決定データは、インターネット上のdcc.blueprint-epigenome.euで閲覧可能なBlueprintコンソーシアムから入手可能である。

【0283】

いくつかの実施形態では、例えば、国際公開第2020/160414号パンフレットに記載のとおり、第1及び第2の捕捉された標的領域セットはそれぞれ、配列可変標的領域セットに対応するDNA及びエピジェネティック標的領域セットに対応するDNAを含む。第1及び第2の捕捉されたセットを組み合わせて、組み合わせられた捕捉されたセットを提供してもよい。配列可変標的領域セット及びエピジェネティック標的領域セットは、その全体が参照により本明細書に援用される、国際公開第2020/160414号パンフレットにおけるこうしたセットについて記載された特徴のいずれかを有してもよい。いくつかの実施形態では、エピジェネティック標的領域セットは、高メチル化可変標的領域セットを含む。いくつかの実施形態では、エピジェネティック標的領域セットは、低メチル化可変標的領域セットを含む。いくつかの実施形態では、エピジェネティック標的領域セットは、CTCF結合領域を含む。いくつかの実施形態では、エピジェネティック標的領域セットは、断片化可変標的領域を含む。いくつかの実施形態では、エピジェネティック標的領域セットは、転写開始部位を含む。いくつかの実施形態では、エピジェネティック標的領域セットは、がんにおける限局的増幅を示し得る領域、例えば、AR、BRAF、CCND1、CCND2、CCNE1、CDK4、CDK6、EGFR、ERBB2、FGFR1、FGFR2、KIT、KRAS、MET、MYC、PDGFRA、PIK3CA、及びRAF1の1つ以上を含む。例えば、いくつかの実施形態では、エピジェネティック標的領域セットは、前記標的の少なくとも2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、又は18を含む。

【0284】

いくつかの実施形態では、配列可変標的領域セットは、がんにおける体細胞突然変異を受けることが知られている複数の領域を含む。いくつかの態様では、配列可変標的領域セットは、複数の異なる遺伝子又はゲノム領域（「パネル」）であって、がんを有する対象の判定された割合がパネルにおける1つ以上の異なる遺伝子又はゲノム領域内の遺伝子変異体又は腫瘍マーカーを示すように選択されるものを標的にする。パネルは、一定数の塩基対に対する配列決定のための領域を制限するために選択されてもよい。パネルを選択して、例えば、本明細書中の他の箇所に記載されるプローブの親和性及び／又は量を調節することによって、所望量のDNAを配列決定することができる。パネルをさらに選択して、所望の配列リード深度を達成することができる。パネルを選択して、配列決定された塩

基対の量に対する所望の配列リード深度又は配列リードカバー率を達成することができる。パネルを選択して、サンプル中の1つ以上の遺伝子変異体を検出するための理論上の感受性、理論上の特異性、及び／又は理論上の正確性を達成することができる。

【0285】

領域のパネルを検出するためのプローブは、目的のゲノム領域（ホットスポット領域）を検出するためのものを含み得る。クロマチン構造についての情報は、プローブの設計において考慮することができ、且つ／又はプローブは、特定部位（例えば、KRASコドン12及び13）が捕捉され得る可能性を最大化するように設計することができ、且つヌクレオソーム結合パターン及びGC配列組成物によって影響を受けるcfDNAのカバー率及び断片サイズ変異の分析に基づいて捕捉を最適化するように設計してもよい。本明細書

10

【0286】

領域のパネルを検出するためのプローブは、目的のゲノム領域（ホットスポット領域）を検出するためのものを含み得る。クロマチン構造についての情報は、プローブの設計において考慮することができ、且つ／又はプローブは、特定部位（例えば、KRASコドン12及び13）が捕捉され得る可能性を最大化するように設計することができ、且つヌクレオソーム結合パターン及びGC配列組成物によって影響を受けるcfDNAのカバー率及び断片サイズ変異の分析に基づいて捕捉を最適化するように設計してもよい。本明細書

20

【0287】

目的のゲノム位置のリストの例は、国際公開第2020/160414号パンフレットの表3及び表4に見出すことができる。いくつかの実施形態では、本開示の方法において使用されている配列可変標的領域セットは、国際公開第2020/160414号パンフレットの表3の遺伝子の少なくとも5、少なくとも10、少なくとも15、少なくとも20、少なくとも25、少なくとも30、少なくとも35、少なくとも40、少なくとも45、少なくとも50、少なくとも55、少なくとも60、少なくとも65、又は70の少なくとも一部分を含む。いくつかの実施形態では、本開示の方法において使用されている配列可変標的領域セットは、国際公開第2020/160414号パンフレットの表4の

30

【0288】

いくつかの実施形態では、配列可変標的領域セットは、少なくとも10、20、30、又は35のがん関連遺伝子、例えば、上に列記されたがん関連遺伝子及び国際公開第2020/160414号パンフレット中のがん関連遺伝子からの標的領域を含む。

【0289】

いくつかの実施形態では、捕捉プローブの収集は、本明細書に記載の方法において使用され、例えば、そのように行うための本明細書で開示される任意の方法によって調製され

40

50

ている捕捉プローブが含まれる。いくつかの実施形態では、捕捉プローブの収集は、配列可変標的領域セットに特異的な標的結合プローブ及び／又はエピジェネティック標的領域セットに特異的な標的結合プローブをさらに含む。いくつかの実施形態では、配列可変標的領域セットに特異的な標的結合プローブの捕捉収率は、エピジェネティック標的領域セットに特異的な標的結合プローブの捕捉収率より高い（例えば、少なくとも2倍高い）。いくつかの実施形態では、捕捉プローブの収集は、配列可変標的領域セットに特異的な捕捉収率であって、エピジェネティック標的領域セットに特異的なその捕捉収率より高い（例えば、少なくとも2倍高い）捕捉収率を有するように設計される。

【0290】

いくつかの実施形態では、配列可変標的領域セットに特異的な標的結合プローブの捕捉収率は、エピジェネティック標的領域セットに特異的な標的結合プローブの捕捉収率よりも少なくとも1.25倍、1.5倍、1.75倍、2倍、2.25倍、2.5倍、2.75倍、3倍、3.5倍、4倍、4.5倍、5倍、6倍、7倍、8倍、9倍、10倍、11倍、12倍、13倍、14倍、又は15倍高い。いくつかの実施形態では、配列可変標的領域セットに特異的な標的結合プローブの捕捉収率は、エピジェネティック標的領域セットに特異的な標的結合プローブの捕捉収率よりも1.25～1.5倍、1.5～1.75倍、1.75～2倍、2～2.25倍、2.25～2.5倍、2.5～2.75倍、2.75～3倍、3～3.5倍、3.5～4倍、4～4.5倍、4.5～5倍、5～5.5倍、5.5～6倍、6～7倍、7～8倍、8～9倍、9～10倍、10～11倍、11～12倍、13～14倍、又は14～15倍高い。

【0291】

いくつかの実施形態では、捕捉プローブの収集は、配列可変標的領域セットに特異的な捕捉収率であって、エピジェネティック標的領域セットに対するその捕捉収率よりも少なくとも1.25倍、1.5倍、1.75倍、2倍、2.25倍、2.5倍、2.75倍、3倍、3.5倍、4倍、4.5倍、5倍、6倍、7倍、8倍、9倍、10倍、11倍、12倍、13倍、14倍、又は15倍高い捕捉収率を有するように設計される。いくつかの実施形態では、捕捉プローブの収集は、配列可変標的領域セットに特異的な捕捉収率であって、エピジェネティック標的領域セットに特異的なその捕捉収率よりも1.25～1.5倍、1.5～1.75倍、1.75～2倍、2～2.25倍、2.25～2.5倍、2.5～2.75倍、2.75～3倍、3～3.5倍、3.5～4倍、4～4.5倍、4.5～5倍、5～5.5倍、5.5～6倍、6～7倍、7～8倍、8～9倍、9～10倍、10～11倍、11～12倍、13～14倍、又は14～15倍高い捕捉収率を有するように設計される。

【0292】

プローブの収集は、配列可変標的領域セットに対するより高い捕捉収率を、濃度、異なる長さ及び／又はケミストリー（例えば、親和性に影響する）、並びにそれらの組み合わせを含む様々な方法において提供するように設計することができる。親和性は、プローブ長さを調節し、且つ／又は以下で考察されるようなヌクレオチド修飾を含めることによって調節することができる。

【0293】

いくつかの実施形態では、配列可変標的領域セットに特異的な捕捉プローブは、エピジェネティック標的領域セットに特異的な捕捉プローブより高い濃度で存在する。いくつかの実施形態では、配列可変標的領域セットに特異的な標的結合プローブの濃度は、エピジェネティック標的領域セットに特異的な標的結合プローブの濃度よりも少なくとも1.25倍、1.5倍、1.75倍、2倍、2.25倍、2.5倍、2.75倍、3倍、3.5倍、4倍、4.5倍、5倍、6倍、7倍、8倍、9倍、10倍、11倍、12倍、13倍、14倍、又は15倍高い。いくつかの実施形態では、配列可変標的領域セットに特異的な標的結合プローブの濃度は、エピジェネティック標的領域セットに特異的な標的結合プローブの濃度よりも1.25～1.5倍、1.5～1.75倍、1.75～2倍、2～2.25倍、2.25～2.5倍、2.5～2.75倍、2.75～3倍、3～3.5倍、

3.5～4倍、4～4.5倍、4.5～5倍、5～5.5倍、5.5～6倍、6～7倍、7～8倍、8～9倍、9～10倍、10～11倍、11～12倍、13～14倍、又は14～15倍高い。かかる実施形態では、濃度は、各セットにおける個別プローブの体積濃度当たりの平均質量を指してもよい。

【0294】

いくつかの実施形態では、配列可変標的領域セットに特異的な捕捉プローブは、エピジェネティック標的領域セットに特異的な捕捉プローブよりも、それらの標的に対して高い親和性を有する。親和性は、当業者に公知の任意の方法で、例えば、異なるプローブケミストリーを使用することによって調節することができる。例えば、特定のヌクレオチド修飾、例えば、（特定の配列内容における）シトシン5-メチル化、2'糖位置にヘテロ原子を提供する修飾、及びLNAヌクレオチドは、二本鎖核酸の安定性を増加させることができ、そのような修飾を有するオリゴヌクレオチドがそれらの相補的配列に対して比較的高い親和性を有することが示される。例えば、Severin et al., Nucleic Acids Res. 39:8740-8751 (2011); Freier et al., Nucleic Acids Res. 25:4429-4443 (1997); 米国特許第9,738,894号明細書を参照されたい。また、より長い配列の長さは、一般に親和性の増加をもたらすことになる。他のヌクレオチド修飾、例えば、核酸塩基ヒポキサンチンのグアニンとの置換は、オリゴヌクレオチドとその相補的配列との間の水素結合の量を低減することによって親和性を低下させる。いくつかの実施形態では、配列可変標的領域セットに特異的な捕捉プローブは、それらの標的に対するそれらの親和性を増加させる修飾を有する。いくつかの実施形態では、代替的又は追加的に、エピジェネティック標的領域セットに特異的な捕捉プローブは、それらの標的に対するそれらの親和性を低下させる修飾を有する。いくつかの実施形態では、配列可変標的領域セットに特異的な捕捉プローブは、エピジェネティック標的領域セットに特異的な捕捉プローブより長い平均長さ及び／又はより高い平均融解温度を有する。これらの実施形態は、捕捉収率における所望の倍数変化、例えば、上に記載の任意の倍数変化又はその範囲を達成するため、相互に及び／又は上で考察されるような濃度の差異と組み合わせてもよい。

【0295】

いくつかの実施形態では、捕捉プローブは、捕捉部分を含む。捕捉部分は、本明細書に記載の捕捉部分のいずれか、例えばビオチンであってもよい。いくつかの実施形態では、捕捉プローブは、例えば、共有結合的に又は非共有結合的に、例えば、捕捉部分の結合ペアの相互作用を通じて、固体支持体に連結される。いくつかの実施形態では、固体支持体は、磁気ビーズなどのビーズである。

【0296】

いくつかの実施形態では、配列可変標的領域セットに特異的な捕捉プローブ及び／又はエピジェネティック標的領域セットに特異的な捕捉プローブは、上で考察された捕捉プローブセット、例えば、捕捉部分及び遺伝子などの領域のパネルを通じて並べて表示するように選択されている配列を含むプローブである。

【0297】

いくつかの実施形態では、捕捉プローブは、単一組成物で提供される。単一組成物は、溶液（液体又は凍結物）であってもよい。或いは、それは凍結乾燥物であってもよい。

【0298】

或いは、捕捉プローブは、例えば、エピジェネティック標的領域セットに特異的なプローブを含む第1の組成物及び配列可変標的領域セットに特異的なプローブを含む第2の組成物を含む、複数の組成物として提供することができる。これらのプローブを適切な割合で混合して、濃度及び／又は捕捉収率における前述の倍数変化のいずれかを伴う組み合わせたプローブ組成物を提供することができる。或いは、それらを（例えば、サンプルの一定分量による、又は連続的に同じサンプルによる）別々の捕捉手順において使用して、捕捉されたエピジェネティック標的領域及び配列可変標的領域をそれぞれ含む第1及び第2の組成物を提供することができる。

【0299】

エピジェネティック標的領域に特異的なプローブ

エピジェネティック標的領域セットに対するプローブは、新生（例えば、腫瘍又はがん）細胞からのDNAを健常細胞、例えば非新生物循環細胞からのDNAと区別する可能性が高い標的領域の1つ以上のタイプに特異的なプローブを含んでもよい。そのような領域の例示的なタイプは、本明細書中の、例えば捕捉されたセットに関連する上のセクションで詳細に考察されている。エピジェネティック標的領域セットに対するプローブはまた、例えば本明細書に記載される、1つ以上の対照領域に対するプローブを含んでもよい。

【0300】

いくつかの実施形態では、エピジェネティック標的領域セットに対するプローブは、少なくとも100 kbp、例えば、少なくとも200 kbp、少なくとも300 kbp、又は少なくとも400 kbpのフットプリントを有する。いくつかの実施形態では、エピジェネティック標的領域セットは、100~20 Mbp、例えば、100~200 kbp、200~300 kbp、300~400 kbp、400~500 kbp、500~600 kbp、600~700 kbp、700~800 kbp、800~900 kbp、900~1,000 kbp、1~1.5 Mbp、1.5~2 Mbp、2~3 Mbp、3~4 Mbp、4~5 Mbp、5~6 Mbp、6~7 Mbp、7~8 Mbp、8~9 Mbp、9~10 Mbp、又は10~20 Mbpの範囲のフットプリントを有する。いくつかの実施形態では、エピジェネティック標的領域セットは、少なくとも20 Mbpのフットプリントを有する。

【0301】

高メチル化可変標的領域

いくつかの実施形態では、エピジェネティック標的領域セットに対するプローブは、1つ以上の高メチル化可変標的領域に特異的なプローブを含む。高メチル化可変標的領域は、本明細書中で高メチル化DMR（差次的にメチル化された領域）と称してもよい。高メチル化可変標的領域は、上に記載されているもののいずれかであってもよい。例えば、いくつかの実施形態では、高メチル化可変標的領域に特異的なプローブは、表1に列記されている遺伝子座の複数、例えば、表1に列記されている遺伝子座の少なくとも10%、20%、30%、40%、50%、60%、70%、80%、90%、又は100%に特異的なプローブを含む。いくつかの実施形態では、高メチル化可変標的領域に特異的なプローブは、表2に列記されている遺伝子座の複数、例えば、表2に列記されている遺伝子座の少なくとも10%、20%、30%、40%、50%、60%、70%、80%、90%、又は100%に特異的なプローブを含む。いくつかの実施形態では、高メチル化可変標的領域に特異的なプローブは、表1又は表2に列記されている遺伝子座の複数、例えば、表1又は表2に列記されている遺伝子座の少なくとも10%、20%、30%、40%、50%、60%、70%、80%、90%、又は100%に特異的なプローブを含む。いくつかの実施形態では、標的領域として含まれる各遺伝子座の場合、遺伝子の転写開始部位と停止コドン（交互スプライスされている遺伝子の最後の停止コドン）との間で結合するハイブリダイゼーション部位を有する1つ以上のプローブが存在し得る。いくつかの実施形態では、1つ以上のプローブは、列記された位置の300 bp以内、例えば、200又は100 bp以内で結合する。いくつかの実施形態では、プローブは、上に列記されている位置と重複しているハイブリダイゼーション部位を有する。いくつかの実施形態では、高メチル化標的領域に特異的なプローブは、乳がん、結腸がん、腎臓がん、肝臓がん、及び肺がんの1つ、2つ、3つ、4つ、又は5つにおける高メチル化を集合的に示す、高メチル化標的領域の1つ、2つ、3つ、4つ、又は5つのサブセットに特異的なプローブを含む。

【0302】

低メチル化可変標的領域

いくつかの実施形態では、エピジェネティック標的領域セットに対するプローブは、1つ以上の低メチル化可変標的領域に特異的なプローブを含む。低メチル化可変標的領域は

、本明細書中で低メチル化DMR（差次的にメチル化された領域）と称してもよい。低メチル化可変標的領域は、上に記載されているもののいずれかであってもよい。例えば、1つ以上の低メチル化可変標的領域に特異的なプローブは、反復配列、例えば、LINE1要素、Alu要素、セントロメアのタンデムリピート、ペリセントロメアのタンデムリピート、及びサテライトDNAなどの領域に対するプローブを含んでもよく、健常細胞内で通常はメチル化されている遺伝子間領域は、腫瘍細胞内でメチル化の低下を示し得る。

【0303】

いくつかの実施形態では、低メチル化可変標的領域に特異的なプローブは、反復配列及び／又は遺伝子間領域に特異的なプローブを含む。いくつかの実施形態では、反復配列に特異的なプローブは、LINE1要素、Alu要素、セントロメアのタンデムリピート、ペリセントロメアのタンデムリピート、及び／又はサテライトDNAの1つ、2つ、3つ、4つ、又は5つに特異的なプローブを含む。

10

【0304】

がん関連低メチル化を示すゲノム領域に特異的な例示的なプローブは、ヒト染色体1のヌクレオチド8403565-8953708及び／又は151104701-151106035に特異的なプローブを含む。いくつかの実施形態では、低メチル化可変標的領域に特異的なプローブは、ヒト染色体1のヌクレオチド8403565-8953708及び／又は151104701-151106035と重複している又はそれを含む領域に特異的なプローブを含む。

20

【0305】

CTCF結合領域

いくつかの実施形態では、エピジェネティック標的領域セットに対するプローブは、CTCF結合領域に特異的なプローブを含む。いくつかの実施形態では、CTCF結合領域に特異的なプローブは、少なくとも10、20、50、100、200、若しくは500のCTCF結合領域、又は10～20、20～50、50～100、100～200、200～500、若しくは500～1000のCTCF結合領域、例えば、上に記載されている、又は上で引用されているCTCFBSDB若しくはCuddapahら、Martinら、若しくはRheeらの論文の1つ以上におけるCTCF結合領域などに特異的なプローブを含む。いくつかの実施形態では、エピジェネティック標的領域セットに対するプローブは、CTCF結合部位の少なくとも100bp、少なくとも200bp、少なくとも300bp、少なくとも400bp、少なくとも500bp、少なくとも750bp、又は少なくとも1000bpの上流及び下流領域を含む。

30

【0306】

転写開始部位

いくつかの実施形態では、エピジェネティック標的領域セットに対するプローブは、転写開始部位に特異的なプローブを含む。いくつかの実施形態では、転写開始部位に特異的なプローブは、少なくとも10、20、50、100、200、若しくは500の転写開始部位、又は10～20、20～50、50～100、100～200、200～500、若しくは500～1000の転写開始部位、例えば、DBTSSに列記されている転写開始部位などに特異的なプローブを含む。いくつかの実施形態では、エピジェネティック標的領域セットに対するプローブは、転写開始部位の少なくとも100bp、少なくとも200bp、少なくとも300bp、少なくとも400bp、少なくとも500bp、少なくとも750bp、又は少なくとも1000bpの上流及び下流の配列に対するプローブを含む。

40

【0307】

限局的増幅

上記のとおり、限局的増幅は体細胞突然変異であるが、それらは、メチル化における変化などの特定のエピジェネティックな変化を検出するための手法と類似する様式でリード頻度に基づいて配列決定することによって検出することができる。そのようなものとして、がんにおける限局的増幅を示し得る領域は、上で考察のとおり、エピジェネティック標

50

的領域セットに含めることができる。いくつかの実施形態では、エピジェネティック標的領域セットに特異的なプローブは、限局的増幅に特異的なプローブを含む。いくつかの実施形態では、限局的増幅に特異的なプローブは、AR、BRAF、CCND1、CCND2、CCNE1、CDK4、CDK6、EGFR、ERBB2、FGFR1、FGFR2、KIT、KRAS、MET、MYC、PDGFRA、PIK3CA、及びRAF1の1つ以上に特異的なプローブを含む。例えば、いくつかの実施形態では、限局的増幅に特異的なプローブは、前述の標的の少なくとも2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、又は18の1つ以上に特異的なプローブを含む。

【0308】

対照領域

データ検証を容易にするため、対照領域を含むことは有用であり得る。いくつかの実施形態では、エピジェネティック標的領域セットに特異的なプローブは、本質的に全てのサンプル中のメチル化されることが予想される対照メチル化領域に特異的なプローブを含む。いくつかの実施形態では、エピジェネティック標的領域セットに特異的なプローブは、本質的に全てのサンプル中の低メチル化されることが予想される対照低メチル化領域に特異的なプローブを含む。

【0309】

配列可変標的領域に特異的なプローブ

配列可変標的領域セットに対するプローブは、がんにおける体細胞突然変異を受けることが知られている複数の領域に特異的なプローブを含んでもよい。プローブは、本明細書に記載される任意の配列可変標的領域セットに特異的であってもよい。例示的な配列可変標的領域セットは、本明細書中の、例えば捕捉されたセットに関連する上のセクションで詳細に考察されている。

【0310】

いくつかの実施形態では、配列可変標的領域プローブセットは、少なくとも0.5 kb、例えば、少なくとも1 kb、少なくとも2 kb、少なくとも5 kb、少なくとも10 kb、少なくとも20 kb、少なくとも30 kb、又は少なくとも40 kbのフットプリントを有する。いくつかの実施形態では、エピジェネティック標的領域プローブセットは、0.5～100 kb、例えば、0.5～2 kb、2～10 kb、10～20 kb、20～30 kb、30～40 kb、40～50 kb、50～60 kb、60～70 kb、70～80 kb、80～90 kb、及び90～100 kbの範囲内のフットプリントを有する。いくつかの実施形態では、配列可変標的領域プローブセットは、少なくとも50 kbp、例えば、少なくとも100 kbp、少なくとも200 kbp、少なくとも300 kbp、又は少なくとも400 kbpのフットプリントを有する。いくつかの実施形態では、配列可変標的領域プローブセットは、100～2000 kbp、例えば、100～200 kbp、200～300 kbp、300～400 kbp、400～500 kbp、500～600 kbp、600～700 kbp、700～800 kbp、800～900 kbp、900～1,000 kbp、1～1.5 Mbp又は1.5～2 Mbpの範囲内のフットプリントを有する。いくつかの実施形態では、配列可変標的領域セットは、少なくとも2 Mb

【0311】

いくつかの実施形態では、配列可変標的領域セットに特異的なプローブは、表3の遺伝子の少なくとも5つ、少なくとも10、少なくとも15、少なくとも20、少なくとも25、少なくとも30、少なくとも35、少なくとも40、少なくとも45、少なくとも50、少なくとも55、少なくとも60、少なくとも65、又は少なくとも70の少なくとも一部分に特異的なプローブを含む。いくつかの実施形態では、配列可変標的領域セットに特異的なプローブは、表3のSNVの少なくとも5つ、少なくとも10、少なくとも15、少なくとも20、少なくとも25、少なくとも30、少なくとも35、少なくとも40、少なくとも45、少なくとも50、少なくとも55、少なくとも60、少なくとも6

10

20

30

40

50

5、又は70に特異的なプローブを含む。いくつかの実施形態では、配列可変標的領域セットに特異的なプローブは、表3の融合の少なくとも1つ、少なくとも2つ、少なくとも3つ、少なくとも4つ、少なくとも5つ、又は6つに特異的なプローブを含む。いくつかの実施形態では、配列可変標的領域セットに特異的なプローブは、表3のインデルの少なくとも1つ、少なくとも2つ、又は3つの少なくとも一部分に特異的なプローブを含む。いくつかの実施形態では、配列可変標的領域セットに特異的なプローブは、表4の遺伝子の少なくとも5つ、少なくとも10、少なくとも15、少なくとも20、少なくとも25、少なくとも30、少なくとも35、少なくとも40、少なくとも45、少なくとも50、少なくとも55、少なくとも60、少なくとも65、少なくとも70、又は73の少なくとも一部分に特異的なプローブを含む。いくつかの実施形態では、配列可変標的領域セットに特異的なプローブは、表4のSNVの少なくとも5つ、少なくとも10、少なくとも15、少なくとも20、少なくとも25、少なくとも30、少なくとも35、少なくとも40、少なくとも45、少なくとも50、少なくとも55、少なくとも60、少なくとも65、少なくとも70、又は73に特異的なプローブを含む。いくつかの実施形態では、配列可変標的領域セットに特異的なプローブは、表4の融合の少なくとも1つ、少なくとも2つ、少なくとも3つ、少なくとも4つ、少なくとも5つ、又は6つに特異的なプローブを含む。いくつかの実施形態では、配列可変標的領域セットに特異的なプローブは、表4のインデルの少なくとも1つ、少なくとも2つ、少なくとも3つ、少なくとも4つ、少なくとも5つ、少なくとも6つ、少なくとも7つ、少なくとも8つ、少なくとも9つ、少なくとも10、少なくとも11、少なくとも12、少なくとも13、少なくとも14、少なくとも15、少なくとも16、少なくとも17、又は18の少なくとも一部分に特異的なプローブを含む。いくつかの実施形態では、配列可変標的領域セットに特異的なプローブは、表5の遺伝子の少なくとも1つ、少なくとも2つ、少なくとも3つ、少なくとも4つ、少なくとも5つ、少なくとも6つ、少なくとも7つ、少なくとも8つ、少なくとも9つ、少なくとも10、少なくとも11、少なくとも12、少なくとも13、少なくとも14、少なくとも15、少なくとも16、少なくとも17、少なくとも18、少なくとも19、又は少なくとも20の少なくとも一部分に特異的なプローブを含む。

10

20

【0312】

【表3】

表3

30

点突然変異(SNV)						融合
AKT1	ALK	APC	AR	ARAF	ARID1A	ALK
ATM	BRAF	BRCA1	BRCA2	CCND1	CCND2	FGFR2
CCNE1	CDH1	CDK4	CDK6	CDKN2A	CDKN2B	FGFR3
CTNNB1	EGFR	ERBB2	ESR1	EZH2	FBXW7	NTRK1
FGFR1	FGFR2	FGFR3	GATA3	GNA11	GNAQ	RET
GNAS	HNF1A	HRAS	IDH1	IDH2	JAK2	ROS1
JAK3	KIT	KRAS	MAP2K1	MAP2K2	MET	
MLH1	MPL	MYC	NF1	NFE2L2	NOTCH1	
NPM1	NRAS	NTRK1	PDGFRA	PIK3CA	PTEN	
PTPN11	RAF1	RB1	RET	RHEB	RHOA	
RIT1	ROS1	SMAD4	SMO	SRC	STK11	
TERT	TP53	TSC1	VHL			

40

【0313】

50

【表 4】

表4

点突然変異(SNV)						融合
AKT1	ALK	APC	AR	ARAF	ARID1A	ALK
ATM	BRAF	BRCA1	BRCA2	CCND1	CCND2	FGFR2
CCNE1	CDH1	CDK4	CDK6	CDKN2A	DDR2	FGFR3
CTNNB1	EGFR	ERBB2	ESR1	EZH2	FBXW7	NTRK1
FGFR1	FGFR2	FGFR3	GATA3	GNA11	GNAQ	RET
GNAS	HNF1A	HRAS	IDH1	IDH2	JAK2	ROS1
JAK3	KIT	KRAS	MAP2K1	MAP2K2	MET	
MLH1	MPL	MYC	NF1	NFE2L2	NOTCH1	
NPM1	NRAS	NTRK1	PDGFRA	PIK3CA	PTEN	
PTPN11	RAF1	RB1	RET	RHEB	RHOA	
RIT1	ROS1	SMAD4	SMO	MAPK1	STK11	
TERT	TP53	TSC1	VHL	MAPK3	MTOR	
NTRK3						

【 0 3 1 4 】

【表 5 - 1】

表5

遺伝子	染色体	開始位置	停止位置	長さ (bp)	カバーされたエクソン	決定的特徴
ALK	chr2	29446405	29446655	250	イントロン19	融合
ALK	chr2	29446062	29446197	135	イントロン20	融合
ALK	chr2	29446198	29446404	206	20	融合
ALK	chr2	29447353	29447473	120	イントロン19	融合
ALK	chr2	29447614	29448316	702	イントロン19	融合
ALK	chr2	29448317	29448441	124	19	融合
ALK	chr2	29449366	29449777	411	イントロン18	融合
ALK	chr2	29449778	29449950	172	18	融合
BRAF	chr7	140453064	140453203	139	15	BRAF V600
CTNNB1	chr3	41266007	41266254	247	3	S37
EGFR	chr7	55240528	55240827	299	18及び19	G719及び欠失
EGFR	chr7	55241603	55241746	143	20	挿入/T790M
EGFR	chr7	55242404	55242523	119	21	L858R
ERBB2	chr17	37880952	37881174	222	20	挿入
ESR1	chr6	152419857	152420111	254	10	V534, P535, L536, Y537, D538
FGFR2	chr10	123279482	123279693	211	6	S252
GATA3	chr10	8111426	8111571	145	5	SS/インデル
GATA3	chr10	8115692	8116002	310	6	SS/インデル
GNAS	chr20	57484395	57484488	93	8	R844
IDH1	chr2	209113083	209113394	311	4	R132
IDH2	chr15	90631809	90631989	180	4	R140, R172
KIT	chr4	55524171	55524258	87	1	
KIT	chr4	55561667	55561957	290	2	

10

20

30

40

50

【表 5 - 2】

KIT	chr4	55564439	55564741	302	3	
KIT	chr4	55565785	55565942	157	4	
KIT	chr4	55569879	55570068	189	5	
KIT	chr4	55573253	55573463	210	6	
KIT	chr4	55575579	55575719	140	7	
KIT	chr4	55589739	55589874	135	8	
KIT	chr4	55592012	55592226	214	9	
KIT	chr4	55593373	55593718	345	10及び11	557, 559, 560, 576
KIT	chr4	55593978	55594297	319	12及び13	V654
KIT	chr4	55595490	55595661	171	14	T670, S709
KIT	chr4	55597483	55597595	112	15	D716
KIT	chr4	55598026	55598174	148	16	L783
KIT	chr4	55599225	55599368	143	17	C809, R815, D816, L818, D820, S821F, N822, Y823
KIT	chr4	55602653	55602785	132	18	A829P
KIT	chr4	55602876	55602996	120	19	
KIT	chr4	55603330	55603456	126	20	
KIT	chr4	55604584	55604733	149	21	
KRAS	chr12	25378537	25378717	180	4	A146
KRAS	chr12	25380157	25380356	199	3	Q61
KRAS	chr12	25398197	25398328	131	2	G12/G13
MET	chr7	116411535	116412255	720	13, 14, イントロン13, イントロン14	METエクソン14SS
NRAS	chr1	115256410	115256609	199	3	Q61
NRAS	chr1	115258660	115258791	131	2	G12/G13
PIK3CA	chr3	178935987	178936132	145	10	E545K
PIK3CA	chr3	178951871	178952162	291	21	H1047R
PTEN	chr10	89692759	89693018	259	5	R130
SMAD4	chr18	48604616	48604849	233	12	D537
TERT	chr5	1294841	1295512	671	プロモーター	chr5:1295228
TP53	chr17	7573916	7574043	127	11	Q331, R337, R342
TP53	chr17	7577008	7577165	157	8	R273
TP53	chr17	7577488	7577618	130	7	R248

10

20

30

40

50

【表 5 - 3】

TP53	chr17	7578127	7578299	172	6	R213/Y220
TP53	chr17	7578360	7578564	204	5	R175/欠失
TP53	chr17	7579301	7579600	299	4	
				12574 (全標的 領域)		
				16330 (全プローブ カバー率)		

10

【0315】

いくつかの実施形態では、配列可変標的領域セットに特異的なプローブは、AKT1、ALK、BRAF、CCND1、CDK2A、CTNNB1、EGFR、ERBB2、ESR1、FGFR1、FGFR2、FGFR3、FOXL2、GATA3、GNA11、GNAQ、GNAS、HRAS、IDH1、IDH2、KIT、KRAS、MED12、MET、MYC、NFE2L2、NRAS、PDGFRA、PIK3CA、PPP2R1A、PTEN、RET、STK11、TP53、及びU2AF1などの少なくとも10、20、30、又は35のがん関連遺伝子からの標的領域に特異的なプローブを含む。

20

【0316】

配列決定

一般に、事前の増幅の存在下又は不在下で、アダプターによって隣接されるサンプル核酸は、配列決定を施すことができる。配列決定方法としては、例えば、サンガー配列決定、ハイスループット配列決定、パイロシークエンシング、合成による配列決定、単一分子配列決定、ナノポア配列決定、半導体配列決定、ライゲーションによる配列決定、ハイブリダイゼーションによる配列決定、デジタル遺伝子発現 (Helicos)、次世代配列決定 (NGS)、合成による単一分子配列決定 (SMSS) (Helicos)、大規模並列配列決定、クローン単一分子アレイ (Solexa)、ショットガン配列決定、Ion Torrent、Oxford Nanopore、Roche Genia、Maxim-Gilbert配列決定、プライマー歩行、及びPacBio、SOLID、Ion Torrent、又はNanoporeプラットフォームを使用する配列決定が挙げられる。配列決定反応は、複数のサンプルセットを実質的に同時に処理する、複数のレーン、複数のチャンネル、複数のウェル、又は他の手段を含んでもよい、種々のサンプル処理ユニットで実施することができる。サンプル処理ユニットは、複数ランの処理を同時に可能にする複数のサンプルチャンバーも含み得る。

30

40

【0317】

いくつかの実施形態では、ゲノムの配列カバー率は、例えば、5%、10%、15%、20%、25%、30%、40%、50%、60%、70%、80%、90%、95%、99%、99.9%又は100%未満であってもよい。いくつかの実施形態では、配列反応は、例えば、ゲノムの少なくとも5%、10%、15%、20%、25%、30%、40%、50%、60%、70%、又は80%の配列カバー率を提供し得る。配列カバー率は、例えば、少なくとも5、10、20、70、100、200若しくは500の異なる遺伝子、又は最大で、例えば、5000、2500、1000、500若しくは100の異なる遺伝子に対して実施することができる。

【0318】

50

同時配列決定反応は、多重配列決定を用いて実施することができる。いくつかの場合、無細胞核酸は、少なくとも、例えば、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、50000、又は100,000の配列決定反応で配列決定することができる。他の場合、無細胞核酸は、例えば、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、50000、又は100,000未満の配列決定反応で配列決定することができる。配列決定反応は、連続的又は同時に実施することができる。その後のデータ分析は、配列決定反応の全部又は一部に対して実施することができる。いくつかの場合、データ分析は、少なくとも、例えば、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、50000、又は100,000の配列決定反応に対して実施することができる。他の場合、データ分析は、例えば、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、50000、又は100,000未満の配列決定反応に対して実施することができる。例示的なリード深度は、遺伝子座（塩基）当たり、1000～50000又は1000～10000又は1000～20000リードである。

10

20

30

40

50

【0319】

一般に、例えば、DNAの修飾ヌクレオシドプロファイルを分析するためのエピジェネティック標的領域の配列決定は、例えば、突然変異の分析のため、配列可変標的領域の配列決定より小さい深度の配列決定を必要とする。それ故、本明細書に記載される、より小さい配列決定深度は、いくつかの場合、本明細書に記載の方法にとって十分であり得る。

【0320】

分析

修飾ヌクレオシドを変換する変換手順を伴う偽陰性の検出

これらの実施形態では、開示の方法において使用されるアダプター配列は、既知の修飾を有する品質管理ヌクレオシドを含む。修飾された品質管理ヌクレオシドが、修飾ヌクレオシドの塩基対特異性を変更する変換手順と組み合わせて使用される場合、適用される変換手順の結果として、品質管理ヌクレオシドの塩基対特異性を変更することが予想される。それ故、変換手順が特定の既知の修飾ヌクレオシドにとって有効であったか否かは、単にどの程度の修飾ヌクレオシドが配列決定されているかを解読することによって判定される。例えば、品質管理ヌクレオシドにmCを含むアダプター分子がTAPS手順を受けており、Tとして配列決定される際に解読される場合、変換は有効であった。品質管理ヌクレオシドにおける既知のmCがCとして解読される場合、変換は無効であった。それ故、アダプター内の品質変換ヌクレオシド（quality conversion nucleosides）の準最適な変換を伴う（個別の／単一の）適応したDNA分子を同定することは可能である。アダプター（又は複数のアダプターで、1つが各末端に含まれる場合）が複数の既知の修飾された品質管理ヌクレオシドを含む場合、修飾された品質管理ヌクレオシドの一部は正確に変換されており、それ以外はそうではない可能性がある。したがって、分子アダプターの品質管理ヌクレオシドにおける既知の修飾ヌクレオシドの総数で割られた、正確に変換された既知の修飾された品質管理ヌクレオシドの数に等しい、特定分子の変換率を判定することは可能である。同様に、全サンプル（又はその画分／サブサンプル）の変換率は、サンプルの固有に同定された分子中の正確に変換された既知の修飾された品質管理ヌクレオシドの総数を、固有に同定されたサンプル分子の同じセット中の既知の修飾ヌクレオシドの総数で割ることによって判定することができる。換言すれば、未知の修飾状態を有するヌクレオシドを含む、サンプル又は分子の変換率は、アダプター配列内の既知の修飾ヌクレオシドの変換率から推測される。

【0321】

変換率は、サンプルレベル又は分子レベルでの変換率の品質管理尺度として用いることができる。いくつかの場合、準最適な変換率、例えば、所定の閾値を下回る変換率を有する個別の分子又はサンプルは、さらなる分析から除外することができる。他の場合、変換率は、部分的又は準最適であり得るが、分子又はサンプル中の修飾ヌクレオシドの既知の

不完全な変換を考慮して、やはり分析に使用することができる。例えば、スコア又は加重を、判定された変換率に依存している各分子及び／又はサンプルに適用することができる。より高い変換率を有する分子又はサンプルには、より高い加重が割り当てられる、即ち、サンプルDNAの修飾ヌクレオシドプロファイルの下流分析又は判定において、より高い有意性が帰せられる。より低い相対変換率を有する分子又はサンプルには、より低い加重が割り当てられる、即ち、下流分析において、より低い有意性が帰せられる。例えば、DNAは、がんを有することが疑われる対象から得てもよい。分析は、変更された修飾ヌクレオシドプロファイルががんに関連している場合、複数のエピジェネティック標的領域の修飾ヌクレオシドプロファイルを判定することを含む。開示の方法では、標的領域のそれぞれからのDNA分子における異なる変換率を判定することができる。より低い変換率を有する領域内の分子を配列決定することによって判定されている修飾ヌクレオシドプロファイルの診断値は、より高い変換率を有する領域内の分子を配列決定することによって判定されている修飾ヌクレオシドプロファイルの診断値と比較して減少する。それ故、より低い変換率領域は、領域の全てからの修飾ヌクレオシドプロファイル分析が、典型的に、がん関連配列突然変異の検出などの他の指標とも組み合わせられる場合、より低い加重が割り当てられ、例えばリスクスコアの総和として、全体的な診断判定が提供される。

10

【0322】

非修飾ヌクレオシドを変換する変換手順を伴う偽陰性の検出

修飾された品質管理ヌクレオシドが、非修飾ヌクレオシドの塩基対特異性を変更する変換手順と組み合わせて使用される場合、典型的に、品質管理ヌクレオシドの塩基対特異性を、適用されている変換手順の結果として変更することは予想されない。それ故、変換手順が特定の既知の修飾ヌクレオシドに対して有効であったか否かは、単にどの程度の修飾ヌクレオシドが配列決定されているかを解読することによって判定される。例えば、品質管理ヌクレオシドにmCを含むアダプター分子がバイサルファイト手順を受けており、Cとして配列決定される際に解読される場合、変換手順はmCを誤って変換することはなく、既知のmCはメチル化されていると正確に同定される。品質管理ヌクレオシドにおける既知のmCがTとして解読される場合、変換手順はmCを誤って変換し、既知のmCはメチル化されていない（即ち、偽陰性）と不正確に同定される。それ故、アダプター内の品質変換ヌクレオシドの誤った変換を伴う（個別の／単一の）適応したDNA分子を同定することは可能である。アダプター（又は複数のアダプターで、1つが各末端に含まれる場合）が複数の既知の修飾ヌクレオシドを含む場合、修飾ヌクレオシドの一部は誤って変換されており、それ以外はそうではない可能性がある。したがって、分子アダプターの品質管理ヌクレオシドにおける既知の修飾ヌクレオシドの総数で割られた、誤って変換された既知の修飾ヌクレオシドの数に等しい、特定分子の誤った変換率を判定することは可能である。同様に、全サンプル（又はその画分／サブサンプル）の誤った変換率は、サンプルの固有に同定された分子中の誤って変換された既知の修飾ヌクレオシドの総数を、固有に同定されたサンプル分子の同じセット中の既知の修飾ヌクレオシドの総数で割ることによって判定することができる。換言すれば、未知の修飾状態を有するヌクレオシドを含む、全サンプル又は分子の誤った変換率は、アダプター配列内の既知の修飾ヌクレオシドの誤った変換率から推測される。

20

30

40

【0323】

誤った変換率は、サンプルレベル又は分子レベルでの誤った変換率の品質管理尺度として用いることができる。いくつかの場合、高い誤った変換率、例えば、所定の閾値を上回る変換率を有する個別の分子又はサンプルは、さらなる分析から除外することができる。他の場合、誤った変換率は、所定の閾値を上回り得るが、分子又はサンプル中の修飾ヌクレオシドの既知の誤った変換を考慮して、やはり分析に使用することができる。例えば、スコア又は加重を、判定された誤った変換率に依存している各分子及び／又はサンプルに適用することができる。より高い誤った変換率を有する分子又はサンプルには、より低い加重が割り当てられる、即ち、サンプルDNAの修飾ヌクレオシドプロファイルの下流分析又は判定において、より低い有意性が帰せられる。より低い相対的な誤った変換率を有

50

する分子又はサンプルには、より高い加重が割り当てられる、即ち、下流分析において、より高い有意性が帰せられる。例えば、DNAは、がんを有することが疑われる対象から得てもよい。分析は、変更された修飾ヌクレオシドプロファイルががんに関連している場合、複数のエピジェネティック標的領域の修飾ヌクレオシドプロファイルを判定することを含む。開示の方法では、標的領域のそれぞれからのDNA分子における異なる誤った変換率を判定することができる。より高い誤った変換率を有する領域内の分子を配列決定することによって判定されている修飾ヌクレオシドプロファイルの診断値は、より低い誤った変換率を有する領域内の分子を配列決定することによって判定されている修飾ヌクレオシドプロファイルの診断値と比較して減少する。それ故、より高い誤った変換率領域は、領域の全てからの修飾ヌクレオシドプロファイル分析が、典型的に、がん関連配列突然変異の検出などの他の指標とも組み合わされる場合、より低い加重が割り当てられ、例えばリスクスコアの総和として、全体的な診断判定が提供される。

10

【0324】

修飾ヌクレオシドを変換する変換手順を伴う偽陽性の検出

修飾されていない品質管理ヌクレオシドが、修飾ヌクレオシドの塩基対特異性を変更する変換手順と組み合わせて使用される場合、品質管理ヌクレオシドの塩基対特異性を、適用されている変換手順の結果として変更することは予想されない。それ故、変換手順が特定の既知の非修飾ヌクレオシドに対して有効であったか否かは、単にどの程度の非修飾ヌクレオシドが配列決定されているかを解読することによって判定される。例えば、品質管理ヌクレオシドに非修飾Cを含むアダプター分子がTAPS手順を受けており、Cとして配列決定される際に解読される場合、変換手順は非修飾Cを誤って変換することはなく、既知の非修飾Cはメチル化されていないと正確に同定される。品質管理ヌクレオシドにおける既知の非修飾CがTとして解読される場合、変換手順は非修飾Cを誤って変換し、既知の非修飾Cはメチル化されている（即ち、偽陽性）と不正確に同定される。それ故、アダプター内の品質変換ヌクレオシドの誤った変換を伴う（個別の／単一の）適応したDNA分子を同定することは可能である。アダプター（又は複数のアダプターで、1つが各末端に含まれる場合）が複数の既知の非修飾ヌクレオシドを含む場合、非修飾ヌクレオシドの一部は誤って変換されており、それ以外はそうではない可能性がある。したがって、分子アダプターの品質管理ヌクレオシドにおける既知の非修飾ヌクレオシドの総数で割られた、誤って変換された既知の非修飾ヌクレオシドの数に等しい、特定分子の誤った変換率を判定することは可能である。同様に、全サンプル（又はその画分／サブサンプル）の誤った変換率は、サンプルの固有に同定された分子中の誤って変換された既知の非修飾ヌクレオシドの総数を、固有に同定されたサンプル分子の同じセット中の既知の非修飾ヌクレオシドの総数で割ることによって判定することができる。換言すれば、未知の修飾状態を有するヌクレオシドを含む、全サンプル又は分子の誤った変換率は、アダプター配列内の既知の非修飾ヌクレオシドの誤った変換率から推測される。

20

30

【0325】

誤った変換率は、サンプルレベル又は分子レベルでの誤った変換率の品質管理尺度として用いることができる。いくつかの場合、高い誤った変換率、例えば、所定の閾値を上回る変換率を有する個別の分子又はサンプルは、さらなる分析から除外することができる。他の場合、誤った変換率は、所定の閾値を上回り得るが、分子又はサンプル中の修飾ヌクレオシドの既知の誤った変換を考慮して、やはり分析に使用することができる。例えば、スコア又は加重を、判定された誤った変換率に依存している各分子及び／又はサンプルに適用することができる。より高い誤った変換率を有する分子又はサンプルには、より低い加重が割り当てられる、即ち、サンプルDNAの修飾ヌクレオシドプロファイルの下流分析又は判定において、より低い有意性が帰せられる。より低い相対的な誤った変換率を有する分子又はサンプルには、より高い加重が割り当てられる、即ち、下流分析において、より高い有意性が帰せられる。例えば、DNAは、がんを有することが疑われる対象から得てもよい。分析は、変更された修飾ヌクレオシドプロファイルががんに関連している場合、複数のエピジェネティック標的領域の修飾ヌクレオシドプロファイルを判定すること

40

50

を含む。開示の方法では、標的領域のそれぞれからのDNA分子における異なる誤った変換率を判定することができる。より高い誤った変換率を有する領域内の分子を配列決定することによって判定されている修飾ヌクレオシドプロファイルの診断値は、より低い誤った変換率を有する領域内の分子を配列決定することによって判定されている修飾ヌクレオシドプロファイルの診断値と比較して減少する。それ故、より高い誤った変換率領域は、領域の全てからの修飾ヌクレオシドプロファイル分析が、典型的に、がん関連配列突然変異の検出などの他の指標とも組み合わされる場合、より低い加重が割り当てられ、例えばリスクスコアの総和として、全体的な診断判定が提供される。

【0326】

非修飾ヌクレオシドを変換する変換手順を伴う偽陽性の検出

修飾されていない品質管理ヌクレオシドが、非修飾ヌクレオシドの塩基対特異性を変更する変換手順と組み合わせて使用される場合、品質管理ヌクレオシドの塩基対特異性を、適用されている変換手順の結果として変更することは予想されない。それ故、変換手順が特定の既知の非修飾ヌクレオシドに対して有効であったか否かは、単にどの程度の非修飾ヌクレオシドが配列決定されているかを解読することによって判定される。例えば、品質管理ヌクレオシドに非修飾Cを含むアダプター分子がバイサルファイト手順を受けており、Tとして配列決定される際に解読される場合、変換は有効であった。品質管理ヌクレオシドにおける既知の非修飾CがCとして解読される場合、変換は無効であった。そのような場合、これは通常であれば、非修飾Cが修飾されている（即ち、偽陽性）と同定されることになる。それ故、アダプター内の品質変換ヌクレオシドの準最適な変換を伴う（個別の／単一の）適応したDNA分子を同定することは可能である。アダプター（又は複数のアダプターで、1つが各末端に含まれる場合）が複数の既知の非修飾ヌクレオシドを含む場合、非修飾ヌクレオシドの一部は正確に変換されており、それ以外はそうではない可能性がある。したがって、分子アダプターの品質管理ヌクレオシドにおける既知の非修飾ヌクレオシドの総数で割られた、正確に変換された既知の非修飾ヌクレオシドの数に等しい、特定分子の変換率を判定することは可能である。同様に、全サンプル（又はその画分／サブサンプル）の変換率は、サンプルの固有に同定された分子中の正確に変換された既知の非修飾ヌクレオシドの総数を、固有に同定されたサンプル分子の同じセット中の既知の非修飾ヌクレオシドの総数で割ることによって判定することができる。換言すれば、未知の修飾状態を有するヌクレオシドを含む、全サンプル又は分子の変換率は、アダプター配列内の既知の非修飾ヌクレオシドの変換率から推測される。変換率は、サンプルレベル又は分子レベルでの変換率の品質管理尺度として用いることができる。いくつかの場合、準最適な変換率、例えば、所定の閾値を下回る変換率を有する個別の分子又はサンプルは、さらなる分析から除外することができる。他の場合、変換率は、部分的又は準最適であり得るが、分子又はサンプル中の非修飾ヌクレオシドの既知の不完全な変換を考慮して、やはり分析に使用することができる。例えば、スコア又は加重を、判定された変換率に依存している各分子及び／又はサンプルに適用することができる。より高い変換率を有する分子又はサンプルには、より高い加重が割り当てられる、即ち、サンプルDNAの修飾ヌクレオシドプロファイルの下流分析又は判定において、より高い有意性が帰せられる。より低い相対変換率を有する分子又はサンプルには、より低い加重が割り当てられる、即ち、下流分析において、より低い有意性が帰せられる。例えば、DNAは、がんを有することが疑われる対象から得てもよい。分析は、変更された修飾ヌクレオシドプロファイルががんに関連している場合、複数のエビジェネティック標的領域の修飾ヌクレオシドプロファイルを判定することを含む。開示の方法では、標的領域のそれぞれからのDNA分子における異なる変換率を判定することができる。より低い変換率を有する領域内の分子を配列決定することによって判定されている修飾ヌクレオシドプロファイルの診断値は、より高い変換率を有する領域内の分子を配列決定することによって判定されている修飾ヌクレオシドプロファイルの診断値と比較して減少する。それ故、より低い変換率領域は、領域の全てからの修飾ヌクレオシドプロファイル分析が、典型的に、がん関連配列突然変異の検出などの他の指標とも組み合わされる場合、より低い加重が割り当てられ、例えばリスク

10

20

30

40

50

スコアの総和として、全体的な診断判定が提供される。

【0327】

例示的な適用

本明細書で提示される方法は、任意のサンプル中のDNAの正確な修飾ヌクレオシドプロファイルを得ることから利益を得る任意の方法の一部として使用してもよい。

【0328】

サンプル

サンプルは、対象から単離されている任意の生体サンプルであり得る。サンプルは、身体サンプルであり得る。サンプルは、身体組織又は体液、例えば、既知の又は疑われた固形腫瘍、全血、血小板、血清、血漿、便、赤血球、白血球 (white blood cells) 又は白血球 (leucocytes)、内皮細胞、組織生検、脳脊髄液、滑液、リンパ液、腹水、間質又は細胞外液、細胞間の空間内の体液、歯肉溝滲出液、骨髓、胸水液、胸膜液、脳脊髄液、唾液、粘液、痰、精液、汗、及び尿を含み得る。サンプルは、好ましくは、体液、特に血液及びその画分、脳脊髄液、胸膜液、唾液、痰、又は尿である。サンプルは、対象から最初に単離されている形態であり得る、又は細胞などの成分を除去若しくは添加する、若しくは1つの成分を別の成分に対して富化するため、さらなる処理を受けている可能性がある。したがって、分析にとって好ましい体液は、無細胞核酸を含有する血漿又は血清である。

10

【0329】

いくつかの実施形態では、核酸の集団は、新生物、腫瘍、前がん、若しくはがんを有することが疑われる、又は新生物、腫瘍、前がん、若しくはがんと以前に診断されている対象からの血清、血漿又は血液サンプルから得られる。集団は、異なるレベルの配列多様性、エピジェネティック変異、及び／又は複製後修飾若しくは転写修飾を有する核酸を含む。複製後修飾は、特に核酸塩基の5位に、シトシンの修飾、例えば、5-メチルシトシン、5-ヒドロキシメチルシトシン、5-ホルミルシトシン及び5-カルボキシルシトシンを含む。

20

【0330】

サンプルは、対象から単離又は取得し、サンプル分析の場所に輸送することができる。サンプルは、望ましい温度、例えば、室温、4℃、-20℃、及び／又は-80℃で保存し、輸送することができる。サンプルは、サンプル分析の場所で、対象から単離又は取得することができる。対象は、ヒト、哺乳動物、動物、コンパニオン動物、サービス動物、又はペットであり得る。対象は、がん、前がん、感染、移植片拒絶、又は免疫系の変化に関連する他の疾患又は障害を有し得る。対象は、がん又は検出可能ながん症状を有しない場合がある。対象は、1つ以上のがん療法、例えば、化学療法、抗体、ワクチン又は生物剤のいずれか1つ以上で治療されている場合がある。対象は、寛解状態であってもよい。対象は、がん又は任意のがん関連遺伝子突然変異／障害に罹りやすいと診断されているか又はされていないかのいずれであってもよい。

30

【0331】

いくつかの実施形態では、サンプルは、血漿を含む。得られた血漿の体積は、配列決定された領域の所望のリード深度に依存し得る。例示的な体積は、0.4~40 mL、5~20 mL、10~20 mLである。例えば、体積は、0.5 mL、1 mL、5 mL、10 mL、20 mL、30 mL、又は40 mLであり得る。採取された血漿の体積は、5~20 mLであってもよい。

40

【0332】

サンプルは、様々な量のゲノム等価物を含有する核酸を含み得る。例えば、約30 ngのDNAのサンプルは、約10,000 (10^4) の一倍体ヒトゲノム等価物、そしてcfDNAの場合、約2000億 (2×10^{11}) の個別のポリヌクレオチド分子を含有し得る。同様に、約100 ngのDNAのサンプルは、約30,000 の一倍体ヒトゲノム等価物、そしてcfDNAの場合、約6000億の個別分子を含有し得る。

【0333】

50

サンプルは、異なる供給源から、例えば、同じ対象の細胞及び無細胞から、異なる対象の細胞及び無細胞からの核酸を含み得る。サンプルは、突然変異を有する核酸を含み得る。例えば、サンプルは、生殖系列突然変異及び／又は体細胞突然変異を有するDNAを含み得る。生殖系列突然変異は、対象の生殖系列DNA中に存在する突然変異を指す。体細胞突然変異は、対象の体細胞内、例えば、がん細胞内に起源を持つ突然変異を指す。サンプルは、がん関連突然変異（例えば、がん関連体細胞突然変異）を有するDNAを含み得る。サンプルは、エピジェネティック変異体（即ち、化学又はタンパク質修飾）を含み得て、ここでエピジェネティック変異体は、がん関連突然変異などの遺伝子変異体の存在に関連する。いくつかの実施形態では、サンプルは、遺伝子変異体の存在に関連するエピジェネティック変異体を含み、ここでサンプルは、遺伝子変異体を含まない。

10

【0334】

増幅前のサンプル中の無細胞核酸の例示的な量は、約1fg～約1pg、例えば、1pg～200ng、1ng～100ng、10ng～1000ngの範囲である。例えば、量は、最大で約600ng、最大で約500ng、最大で約400ng、最大で約300ng、最大で約200ng、最大で約100ng、最大で約50ng、又は最大で約20ngの無細胞核酸分子であり得る。量は、少なくとも1fg、少なくとも10fg、少なくとも100fg、少なくとも1pg、少なくとも10pg、少なくとも100pg、少なくとも1ng、少なくとも10ng、少なくとも100ng、少なくとも150ng、又は少なくとも200ngの無細胞核酸分子であり得る。量は、最大で1フェムトグラム（fg）、10fg、100fg、1ピコグラム（pg）、10pg、100pg、1ng、10ng、100ng、150ng、又は200ngの無細胞核酸分子であり得る。方法は、1フェムトグラム（fg）～200ngの無細胞核酸分子を得ることを含み得る。

20

【0335】

無細胞DNAは、対象からのその単離時に細胞内に含有されていないDNAを指す。例えば、cfDNAは、細胞を溶解せず、又はそうでなければ細胞内DNAを抽出することなく、無傷細胞を除去した後にサンプル中に残存しているDNAとして、サンプルから単離することができる。無細胞核酸は、DNA、RNA、及びそのハイブリッド、例えば、ゲノムDNA、ミトコンドリアDNA、siRNA、miRNA、循環RNA（cRNA）、tRNA、rRNA、核小体低分子RNA（snRNA）、Piwi相互作用RNA（piRNA）、長鎖非翻訳RNA（長鎖ncRNA）、又はこれらのいずれかの断片を含む。無細胞核酸は、二本鎖、一本鎖、又はそのハイブリッドであり得る。無細胞核酸は、分泌又は細胞死プロセス、例えば、細胞壊死及びアポトーシスを通じて体液中に放出され得る。一部の無細胞核酸は、がん細胞から体液中に放出され、例えば、循環腫瘍DNA、（ctDNA）が挙げられる。その他は、健常細胞から放出される。いくつかの実施形態では、cfDNAは、無細胞胎児DNA（cffDNA）である。いくつかの実施形態では、無細胞核酸は、腫瘍細胞によって生成される。いくつかの実施形態では、無細胞核酸は、腫瘍細胞と非腫瘍細胞の混合物によって生成される。

30

【0336】

無細胞核酸は、約100～500ヌクレオチドの代表的なサイズ分布を有し、110～約230ヌクレオチドの分子は分子の約90%を表し、約168ヌクレオチドのモードと240～440ヌクレオチドの範囲内に2番目のマイナーピークとを有する。

40

【0337】

無細胞核酸は、溶液中で見出される無細胞核酸が体液の無傷細胞及び他の非可溶性成分から分離される場合の分画又は分割ステップを通じて、体液から単離することができる。分割は、遠心分離又は濾過などの技術を含んでもよい。或いは、体液中の細胞を溶解させ、無細胞及び細胞核酸と一緒に処理することができる。一般に、緩衝液の添加及び洗浄ステップ後、核酸をアルコールで沈殿させることができる。さらに、シリカベースのカラムなど、クリーンアップステップを使用して、汚染物質又は塩を除去することができる。非特異的な大量の担体の核酸、例えば、ニシン精子DNAを、反応全体を通じて添加して、

50

収率などの手順の特定の態様を最適化することができる。

【0338】

このような処理後、サンプルは、二本鎖DNA、一本鎖DNA及び一本鎖RNAを含む、核酸の様々な形態を含み得る。いくつかの実施形態では、一本鎖DNA及びRNAは、後の処理及び分析ステップに含められるように、二本鎖形態に変換することができる。

【0339】

サンプル中の二本鎖DNA分子及び二本鎖DNA分子に変換されている一本鎖核酸分子は、一方の末端又は両方の末端のいずれかでアダプターに連結することができる。典型的に、二本鎖分子は、全部で4つの標準ヌクレオチドの存在下で、5'-3'ポリメラーゼ及び3'-5'エキソヌクレアーゼ（又はブルーフリーディング機能）を有するポリメラーゼ 10 による処理によって平滑末端化される。クレノウ大断片及びT4ポリメラーゼは、好適なポリメラーゼの例である。平滑末端化DNA分子は、少なくとも部分的に、二本鎖アダプター（例えば、Y形状又は釣鐘状のアダプター）とライゲートすることができる。或いは、相補的なヌクレオチドをサンプル核酸及びアダプターの平滑末端に付加して、ライゲーションを促進することができる。本明細書で企図されるのは、平滑末端ライゲーションと粘着末端ライゲーションの両方である。平滑末端ライゲーションでは、核酸分子とアダプタータグの両方が平滑末端を有する。粘着末端ライゲーションでは、典型的に、核酸分子は、「A」オーバーハングを有し、アダプターは、「T」オーバーハングを有する。

【0340】

アダプターを含むDNAの増幅

アダプターによって隣接されているサンプル核酸は、PCR及び他の増幅方法によって増幅することができる。こうした増幅を用いて、配列決定などの後続ステップに利用可能なDNAの量を増加させることができる。例えば、この種の増幅は、ライブラリー調製の一部として、及び／又は標的化ライブラリーの調製後、実施することができる。増幅は、変換手順後に実施することができる。増幅は、増幅対象のDNA分子に隣接するアダプター内のプライマー結合部位に結合するプライマーによって刺激することができる。増幅方法は、サーモサイクリングから得られる、変性、アニーリング及び伸長のサイクルを含み得る、又は転写媒介性増幅などの場合、等温であり得る。他の増幅方法は、リガーゼ連鎖反応、鎖置換増幅、核酸配列に基づく増幅、及び自己持続的配列に基づく複製を含む。

【0341】

いくつかの実施形態では、本方法では、アダプターへの連結前に二本鎖核酸の少なくとも50、60、70又は80%の増幅をもたらす、Tテールアダプター及びCテールアダプターとのdsDNAライゲーションが実施される。好ましくは、本方法は、増幅された分子の量又は数を、Tテールアダプター単独によって実施される対照方法と比較して、少なくとも10、15又は20%増加させる。

【0342】

適用

開示の方法の1つの重要な例示的な適用は、がん又は他の遺伝的疾患若しくは症状の診断及び予後における修飾ヌクレオシドプロファイルの使用である。

【0343】

それ故、いくつかの実施形態では、本明細書に記載の方法は、腫瘍（又は新生細胞、又はがん細胞）によって生成されているDNAの存在若しくは不在を同定若しくは予測すること、試験対象が腫瘍若しくはがんを有する確率を判定すること、及び／又は本明細書に記載される腫瘍、新生細胞若しくはがんを特徴づけることを含む。

【0344】

本方法を用いて、対象における病態、特にがんの存在又は不在を診断して、病態を特徴づけて（例えば、がんをステージ分類する、又はがんの異質性を判定する）、病態の治療に対する応答を監視して、病態又はその後の病態の経過を発現する予後リスクをもたらすことができる。本開示はまた、特定の治療オプションの有効性の判定に有用であり得る。より多くのがんが死滅し、DNAを排出し得ることで、治療が成功している場合、治療オ 50

ブションの成功により、対象の血液中で検出されているコピー数多型又は希少突然変異の量が増加し得る。他の例では、これが生じることはない。別の例では、おそらく特定の治療オプションが、がんの遺伝的プロファイルと経時的に相関し得る。この相関は、治療法を選択するのに有用であり得る。いくつかの実施形態では、高メチル化可変エピジェネティック標的領域を分析して、それらが腫瘍細胞若しくは通常はcfDNAに有意に寄与することがない細胞の特徴を持つ高メチル化を示すか否かが判定され、且つ／又は低メチル化可変エピジェネティック標的領域を分析して、それらが腫瘍細胞若しくは通常はcfDNAに有意に寄与することがない細胞の特徴を持つ低メチル化を示すか否かが判定される。

【0345】

いくつかの実施形態では、本方法は、がんについてスクリーニングするため、又はがんをスクリーニングするための方法において使用される。例えば、サンプルは、以前にがんと診断されていない対象から得ることができる。いくつかの実施形態では、対象は、がんを有するか又は有しないかのいずれでもよい。いくつかの実施形態では、対象は、初期がんを有するか又は有しないかのいずれでもよい。いくつかの実施形態では、対象は、タバコ使用（例えば、喫煙）、過体重若しくは肥満であること、高い肥満度指数（BMI）を有すること、高齢であること、栄養不良、高いアルコール消費、又はがんの家族歴など、1つ以上のがんのリスク因子を有する。

【0346】

いくつかの実施形態では、対象は、例えば、少なくとも1、5、10、又は15年間、タバコを使用している。いくつかの実施形態では、対象は、高いBMI、例えば、25以上、26以上、27以上、28以上、29以上、又は30以上のBMIを有する。いくつかの実施形態では、対象は、少なくとも40、45、50、55、60、65、70、75、又は80歳である。いくつかの実施形態では、対象は、栄養不良、例えば、レッドミート及び／若しくは加工肉、トランス脂肪、飽和脂肪、及び精製糖の1つ以上の高消費、並びに／又は果実及び野菜、複合炭水化物、及び／若しくは不飽和脂肪の低消費を有する。高消費及び低消費は、例えば、それぞれ、以下、www.dietaryguidelines.gov/sites/default/files/2021-03/Dietary_Guidelines_for_Americans-2020-2025.pdfで入手可能な米国民向け食事ガイドライン2020-2025（Dietary Guidelines for Americans 2020-2025）における推奨を上回ること又は下回ることとして定義することができる。いくつかの実施形態では、対象は、例えば、平均で1日に少なくとも3回、4回、又は5回の飲酒（飲料が約1オンス又は30mLの80プルーフハードリカー又は相当物である場合）といった高いアルコール消費を有する。いくつかの実施形態では、対象は、がんの家族歴を有し、例えば、少なくとも1名、2名、又は3名の血縁者が以前にがんと診断された。いくつかの実施形態では、近親者は、少なくとも第三度近親者（例えば、高祖父母、大叔母又は大叔父、第1のいとこ）、少なくとも第二度近親者（例えば、祖父母、叔母又は叔父、又は半同胞）、又は第一度近親者（例えば、親又は全同胞）である。

【0347】

加えて、がんが治療後に寛解状態であることが認められる場合、本方法を用いて、残存疾患又は疾患の再発を監視することができる。

【0348】

典型的に、検討中の疾患は、例えば、本明細書で参照されるいずれかのがんのタイプである。検出可能であるがんのタイプ及び数は、血液がん、脳がん、肺がん、皮膚がん、鼻がん、咽頭がん、肝がん、骨がん、リンパ腫、膵がん、皮膚がん、腸がん、直腸がん、甲状腺がん、膀胱がん、腎がん、口がん、胃がん、固形腫瘍、異質性腫瘍、均質性腫瘍などを含んでもよい。こうしたがんの具体例としては、胆道がん、膀胱がん、移行上皮がん、尿路上皮がん、脳がん、神経膠腫、星状細胞腫、乳がん、化生がん、子宮頸がん、子宮頸部扁平上皮がん、直腸がん、結腸直腸がん、結腸がん、遺伝性非ポリープ性結腸直腸がん

10

20

30

40

50

、結腸直腸腺がん、消化管間質性腫瘍（GIST）、子宮内膜がん、子宮内膜間質肉腫、食道がん、食道扁平上皮がん、食道腺がん、眼内黒色腫、ぶどう膜黒色腫、胆嚢がん、胆嚢腺がん、腎細胞がん、明細胞腎細胞がん、移行上皮がん、尿路上皮がん、ウィルムス腫瘍、白血病、急性リンパ球性白血病（ALL）、急性骨髄性白血病（AML）、慢性リンパ性白血病（CLL）、慢性骨髄性白血病（CML）、慢性骨髄単球性白血病（CMML）、肝臓がん、肝がん、肝腫瘍、肝細胞がん、胆管細胞がん、肝芽腫、肺がん、非小細胞肺がん（NSCLC）、中皮腫、B細胞リンパ腫、非ホジキンリンパ腫、びまん性大細胞型B細胞リンパ腫、マントル細胞リンパ腫、T細胞リンパ腫、非ホジキンリンパ腫、前駆体Tリンパ芽球性リンパ腫／白血病、末梢性T細胞リンパ腫、多発性骨髄腫、鼻咽頭がん（NPC）、神経芽細胞腫、中咽頭がん、口腔扁平上皮がん、骨肉腫、卵巣がん、膵がん、膵管腺癌、偽乳頭状新生物、腺房細胞がん、前立腺がん、前立腺腺がん、皮膚がん、メラノーマ、悪性メラノーマ、皮膚メラノーマ、小腸がん、胃がん（stomach cancer）、胃がん（gastric carcinoma）、消化管間質性腫瘍（GIST）、子宮がん、又は子宮肉腫が挙げられる。

10

【0349】

がんのタイプ及び／又はステージは、突然変異、希少突然変異、インデル、コピー数多型、トランスポージョン、転位、反転、欠失、異数性、部分的異数性、倍数性、染色体不安定性、染色体構造変化、遺伝子融合、染色体融合、遺伝子トランケーション、遺伝子増幅、遺伝子重複、染色体病変、DNA病変、核酸化学修飾の異常な変化、エピジェネティックパターンの異常な変化、及び核酸5-メチルシトシンの異常な変化を含む遺伝的変異から検出することができる。それ故、本方法は、いくつかの場合、例えば、がんを検出するか若しくは特徴づける方法又は本明細書に記載の他の方法における、他の遺伝的／エピジェネティック変異を検出するために用いられる方法と組み合わせて用いることができる。

20

【0350】

いくつかの実施形態では、本明細書に記載の方法は、腫瘍（又は新生細胞、又はがん細胞）によって、又は前がん細胞によって生成されている標的領域及び／又はDNAの存在を同定することを含む。いくつかの実施形態では、本明細書に記載の方法は、標的領域のレベルを判定し、且つ／又は腫瘍（又は新生細胞、又はがん細胞）によって、若しくは前がん細胞によって生成されているDNAの存在を同定することを含む。いくつかの実施形態では、標的領域のレベルの判定は、標的領域の増加レベル又は減少レベルのいずれかを判定することを含み、ここで標的領域の増加又は低下レベルは、標的領域のレベルを閾値レベル／値と比較することによって判定される。

30

【0351】

また、遺伝的及び／又はエピジェネティックデータは、がんの特定の形態を特徴づけるために使用することができる。がんは、組成物及びステージングの両方において異質性であることが多い。遺伝的及び／又はエピジェネティックプロファイルデータは、がんの特定のサブタイプの特徴づけ（その特定のサブタイプの診断又は治療において重要であり得る）を可能にし得る。また、この情報は、対象又は施術者に、がんの特定タイプの予後に関する手がかりを提供し、対象又は施術者のいずれかが疾患の進行に応じて治療オプションを適応させ得る場合がある。一部のがんは、進行して、より侵襲的に、そして遺伝的に不安定になり得る。他のがんは、良性、不活性又は休止状態を維持し得る。本開示のシステム及び方法は、疾患進行を判定するのに有用であり得る。

40

【0352】

さらに、本開示の方法を用いて、対象における異常な状態の異質性を特徴づけることができる。こうした方法は、例えば、対象由来の細胞外ポリヌクレオチドの遺伝的及び／又はエピジェネティックプロファイルを作成することを含むことができ、ここで遺伝的及び／又はエピジェネティックプロファイルは、コピー数多型及び希少突然変異分析から得られた複数のデータを含む。いくつかの実施形態では、異常な状態は、例えば本明細書に記載されるがんである。いくつかの実施形態では、異常な状態は、異質ゲノム集団をもたら

50

すものであり得る。がんの例において、一部の腫瘍は、がんの異なるステージにある腫瘍細胞を含むことは既知である。他の例において、異質性は、疾患の複数の病巣を含んでもよい。再び、がんの例において、複数の腫瘍病巣が存在することがあり、その場合、おそらくは、1つ以上の病巣は、原発性部位から伝播している転移の結果である。

【0353】

本方法を用いて、異質性疾患における異なる細胞から得られた遺伝的及び／又はエピジェネティック情報の総和であるプロファイル、フィンガープリント又はデータセットを作成することができる。このデータセットは、コピー数多型、エピジェネティック変異、及び突然変異分析を単独で又は組み合わせて含んでもよい。

【0354】

本方法を用いて、がん、又は他の疾患を診断、予後、監視又は観察することができる。いくつかの実施形態では、本明細書中の方法は、胎児を診断、予後又は監視することを含まず、それ故、非侵襲的な出生前検査を対象としない。他の実施形態では、これらの方法を妊娠対象に用いて、DNA及び他のポリヌクレオチドが母体分子と共循環し得る胎児対象におけるがん又は他の疾患を診断、予後、監視又は観察することができる。

【0355】

任意選択的に、本明細書で開示される方法及びシステムを用いて評価されている他の遺伝子ベースの疾患、障害、又は病態の非限定例としては、軟骨無形成症、 α -1アンチトリプシン欠乏、抗リン脂質症候群、自閉症、常染色体優性多発性嚢胞腎疾患、シャルコー・マリー・トゥース（CMT）、クリ・デュ・チャット、クローン病、嚢胞性線維症、ダーカム病、ダウン症候群、デュアン症候群、デュシェンヌ型筋ジストロフィー、第V因子ライデン栓友病、家族性高コレステロール血症、家族性地中海熱、脆弱X症候群、ゴーシェ病、ヘモクロマトーシス、血友病、全前脳症、ハンチントン病、クライنفエルター症候群、マルファン症候群、筋緊張性ジストロフィー、神経線維腫症、ヌーナン症候群、骨形成不全症、パーキンソン病、フェニルケトン尿症、ポーランド異常、ポルフィリン症、早老症、網膜色素変性症、重症複合免疫不全（SCID）、鎌状赤血球症、脊髄性筋萎縮、テイ・サックス、サラセミア、トリメチルアミン尿症、ターナー症候群、口蓋心臓顔面症候群、WAGR症候群、ウィルソン病などが挙げられる。

【0356】

いくつかの実施形態では、本明細書に提供される方法は、対象におけるがん再発のリスクを判定する方法である。いくつかの実施形態では、本明細書に提供される方法は、対象を後のがん治療の候補であるとして分類する方法である。

【0357】

こうした方法のいずれかは、対象に対する1つ以上の以前のがん治療後の1つ以上の前選択された時点で、がんを診断されている対象からサンプルを収集することを含んでもよい。対象は、本明細書に記載の対象のいずれかであってもよい。サンプルは、DNA、例えば、cfDNAを含んでもよい。DNAは、組織サンプル又は液体サンプルから得てもよい。

【0358】

こうした方法のいずれかは、本明細書に記載される実施形態のいずれかに従い、サンプル又はそのサブサンプルを複数のプライマーと接触させて、捕捉プローブを生成して、少なくとも1つの構造変異の存在又はレベルを捕捉及び検出することを含んでもよい。いくつかの実施形態では、方法は、本明細書に記載される実施形態のいずれかに従い、サンプル又はそのサブサンプルを、エピジェネティック標的領域セットのメンバーに特異的な複数の捕捉プローブと接触させることを含んでもよい。いくつかの実施形態では、捕捉プローブは、より早い時点で同じ対象から得られたサンプルを用いて生成されている捕捉プローブを含む。方法は、対象からのDNAからの複数の標的領域セットを捕捉することをさらに含んでもよく、ここで複数の標的領域セットは、それによってDNA分子の捕捉されたセットが生成される、配列可変標的領域セット、及び／又はエピジェネティック標的領域セットを含む。1つ又は複数の捕捉ステップは、本明細書中の他の箇所に記載される実

10

20

30

40

50

施形態のいずれかに従って実施することができる。こうした方法のいずれかは、それによって配列情報のセットが生成される、捕捉されたDNA分子を配列決定することを含んでもよい。配列可変標的領域セットの捕捉されたDNA分子は、エピジェネティック標的領域セットの捕捉されたDNA分子よりも配列決定の大きい深度まで配列決定することができる。こうした方法のいずれかは、配列情報のセットを使用して、腫瘍細胞が起源であるか又はそれに由来するDNAの存在又は不在を前選択された時点で検出することを含んでもよい。腫瘍細胞が起源であるか又はそれに由来するDNAの存在又は不在の検出は、本明細書中の他の箇所に記載されるその実施形態のいずれかに従って実施することができる。

【0359】

こうした方法のいずれかにおいて、以前のがん治療は、手術、治療組成物の投与、及び／又は化学療法を含んでもよい。

【0360】

対象におけるがん再発のリスクを判定する方法は、少なくとも1つの再構成されたDNA配列、型特異的な標的領域、及び／又は対象における腫瘍細胞が起源であるか若しくはそれに由来するDNAの存在若しくは不在、又は量を意味するがん再発スコアを判定することを含んでもよい。がん再発スコアをさらに使用して、がん再発状態を判定することができる。がん再発状態は、例えば、がん再発スコアが所定の閾値を上回るとき、がん再発のリスクであり得る。がん再発状態は、例えば、がん再発スコアが所定の閾値を上回るとき、がん再発の低い又はより低いリスクであり得る。特定の実施形態では、所定の閾値に等しいがん再発スコアは、がん再発のリスク、又はがん再発の低い若しくはより低いリスクのいずれかであるがん再発状態をもたらし得る。

【0361】

後のがん治療の候補であるとして対象を分類する方法は、対象のがん再発スコアを所定のがん再発閾値と比較して、それにより、がん再発スコアががん再発閾値を上回るときに後のがん治療の候補として、又はがん再発スコアががん再発閾値を下回るときに治療法の候補でないとして対象を分類することを含んでもよい。特定の実施形態では、がん再発閾値に等しいがん再発スコアは、後のがん治療の候補であるか、又は治療法の候補でないかのいずれかの分類をもたらし得る。いくつかの実施形態では、後のがん治療は、化学療法又は治療組成物の投与を含む。

【0362】

こうした方法のいずれかは、がん再発スコアに基づき、対象の無疾患生存（DFS）期間を判定することを含んでもよく；例えば、DFS期間は、1年、2年、3年、4年、5年、又は10年であってもよい。

【0363】

いくつかの実施形態では、配列情報のセットは、配列可変標的領域配列を含み、がん再発スコアの判定は、配列可変標的領域配列内に存在する特定の免疫細胞型、SNV、挿入／欠失、CNV及び／又は融合のレベルを意味する、少なくとも第1のサブスコアを判定することを含んでもよい。

【0364】

いくつかの実施形態では、1、2、3、4、又は5から選択される、配列可変標的領域内の突然変異の数は、第1のサブスコアが、がん再発について陽性と分類されているがん再発スコアをもたらすのに十分である。いくつかの実施形態では、突然変異の数は、1、2、又は3から選択される。

【0365】

いくつかの実施形態では、配列情報のセットは、エピジェネティック標的領域配列を含み、がん再発スコアの判定は、健常対象からの対応するサンプル中に見出されるDNA（例えば、健常対象からの血液サンプル中に見出されるcfDNA、又は組織サンプルが試験対象から得られた場合と同じタイプの組織である場合、健常対象からの組織サンプル中に見出されるDNA）と異なるエピジェネティック状態を表す（エピジェネティック標的

10

20

30

40

50

領域配列から得られた) 分子の量を意味する第2のサブスコアを判定することを含む。これらの異常な分子(即ち、健常対象からの対応するサンプル中に見出されるDNAと異なるエピジェネティック状態を有する分子)は、がんに関連するエピジェネティック変化、例えば、高メチル化可変標的領域のメチル化及び／又は断片化可変標的領域の摂動した断片化に一致することがあり、ここで「摂動した」は、健常対象からの対応するサンプル中に見出されるDNAと異なることを意味する。

【0366】

いくつかの実施形態では、高メチル化可変標的領域セットにおける高メチル化及び／又は断片化可変標的領域セットにおける異常な断片化を示す、高メチル化可変標的領域セット及び／又は断片化可変標的領域セットに対応する分子の、0.001%~10%の範囲内の値以上の割合は、第2のサブスコアががん再発について陽性と分類されるのに十分である。範囲は、0.001%~1%、0.005%~1%、0.01%~5%、0.01%~2%、又は0.01%~1%であってもよい。

10

【0367】

いくつかの実施形態では、こうした方法のいずれかは、腫瘍DNAの画分を、腫瘍細胞に起源を持つことを意味する1つ以上の特徴を示す配列情報のセットにおける分子の画分から判定することを含んでもよい。これは、例えば、高メチル化可変標的領域及び断片化可変標的領域の一方又は両方を含む、エピジェネティック標的領域の一部又は全部に対応する分子について行うことができる(高メチル化可変標的領域の高メチル化及び／又は断片化可変標的領域の異常な断片化は、腫瘍細胞に起源を持つことを意味すると考えることができる)。これは、配列可変標的領域に対応する分子、例えば、SNV、インデル、CNV、及び／又は融合などのがんに合致する変更を含む分子について行うことができる。腫瘍DNAの画分は、エピジェネティック標的領域に対応する分子と配列可変標的領域に対応する分子との組み合わせに基づいて判定することができる。

20

【0368】

がん再発スコアの判定は、少なくとも部分的に、腫瘍DNAの画分に基づいてもよく、ここで 10^{-11} ~1又は 10^{-10} ~1の範囲内の閾値より大きい腫瘍DNAの画分は、がん再発スコアががん再発について陽性と分類されるのに十分である。いくつかの実施形態では、 10^{-10} ~ 10^{-9} 、 10^{-9} ~ 10^{-8} 、 10^{-8} ~ 10^{-7} 、 10^{-7} ~ 10^{-6} 、 10^{-6} ~ 10^{-5} 、 10^{-5} ~ 10^{-4} 、 10^{-5} ~ 10^{-3} 、 10^{-5} ~ 10^{-2} 、又は 10^{-2} ~ 10^{-1} の範囲内の閾値以上の腫瘍DNAの画分は、がん再発スコアががん再発について陽性と分類されるのに十分である。いくつかの実施形態では、少なくとも 10^{-7} の閾値より大きい腫瘍DNAの画分は、がん再発スコアががん再発について陽性と分類されるのに十分である。腫瘍DNAの画分が閾値、例えば、前述の実施形態のいずれかに対応する閾値より大きいという判定は、累積確率に基づいてなされてもよい。例えば、サンプルは、腫瘍画分が前述の範囲のいずれかにおける閾値より大きかったという累積確率が、少なくとも0.5、0.75、0.9、0.95、0.98、0.99、0.995、又は0.999の確率閾値を超える場合、陽性とみなされた。いくつかの実施形態では、確率閾値は、少なくとも0.95、例えば0.99である。

30

【0369】

いくつかの実施形態では、配列情報のセットは、配列可変標的領域配列及びエピジェネティック標的領域配列を含み、がん再発スコアの判定は、配列可変標的領域配列内に存在するSNV、挿入／欠失、CNV及び／又は融合の量を意味する第1のサブスコア、並びにエピジェネティック標的領域配列内の異常な分子の量を意味する第2のサブスコアを判定することと、第1のサブスコア及び第2のサブスコアを組み合わせ、がん再発スコアを提供することと、を含む。第1のサブスコア及び第2のサブスコアが組み合わせられる場合、それらは、閾値を各サブスコアに独立して適用する(例えば、配列可変標的領域において突然変異の所定の数より大きい(例えば、 >1)、及びエピジェネティック標的領域において異常な分子(即ち、健常対象からの対応するサンプル中に見出されるDNAと異なるエピジェネティックな状態を有する分子; 例えば、腫瘍)の所定の画分より大きい)

40

50

、又は機械学習の分類子を訓練して、複数の陽性及び陰性の訓練サンプルに基づいて状態を判定することによって組み合わせることができる。

【0370】

いくつかの実施形態では、 $-4 \sim 2$ 又は $-3 \sim 1$ の範囲内の組み合わせスコアの値は、がん再発スコアががん再発について陽性と分類されるのに十分である。

【0371】

いずれかの実施形態では、がん再発スコアががん再発について陽性と分類される場合、対象のがん再発状態は、がん再発のリスクの可能性がある、且つ／又は対象は、後のがん治療の候補として分類され得る。

【0372】

いくつかの実施形態では、がんは、本明細書中の他の箇所に記載のがんのタイプのいずれか1つ、例えば、結腸直腸がんである。

【0373】

特定の実施形態では、本明細書で開示される方法は、患者に対するカスタマイズされた治療法を特定し、施すことに関連する。いくつかの実施形態では、特定の核酸のレベルの決定により、適切な治療の選択が容易になる。いくつかの実施形態では、患者又は対象は、所与の疾患、障害、又は病態を有する。本質的に任意のがん療法（例えば、外科的療法、放射線療法、化学療法など）は、これらの方法の一部に含めてもよい。特定の実施形態では、対象に施される治療法は、少なくとも1つの化学療法薬を含む。いくつかの実施形態では、化学療法薬は、アルキル化剤（例えば、限定はされないが、クロラムブシル、シクロホスファミド、シスプラチン及びカルボプラチン）、ニトロソウレア（例えば、限定はされないが、カルムスチン及びロムスチン）、抗代謝産物（例えば、限定はされないが、フルオロウラシル（Fluorouracil）、メトトレキサート及びフルダラビン）、植物アルカロイド及び天然物（例えば、限定はされないが、ビンクリスチン、パクリタキセル及びトポテカン）、抗腫瘍抗生物質（例えば、限定はされないが、ブレオマイシン、ドキソルビシン及びミトキサントロン）、ホルモン剤（例えば、限定はされないが、プレドニゾン、デキサメタゾン、タモキシフェン及びリュープロリド）及び生物学的応答修飾因子（例えば、限定はされないが、ハーセプチン及びアバスタチン、アービタックス及びリツキサン）を含んでもよい。いくつかの実施形態では、対象に投与される化学療法は、FOLFOX又はFOLFIRIを含んでもよい。特定の実施形態では、少なくとも1つのPARP阻害剤を含む療法が対象に投与されてもよい。特定の実施形態では、PARP阻害剤は、特に、オラパリブ、タラゾパリブ、ルカパリブ、ニラパリブ（商品名ゼジューラ）を含んでもよい。典型的に、治療法は、少なくとも1つの免疫療法（又は免疫療法剤）を含む。免疫療法は、一般に、所与のがんタイプに対する免疫応答を増強する方法を指す。特定の実施形態では、免疫療法は、腫瘍又はがんに対するT細胞応答を増強する方法を指す。

【0374】

いくつかの実施形態では、治療法は、体細胞又は生殖系列が起源であるような核酸変異体の状態に基づいてカスタマイズされる。いくつかの実施形態では、本質的に任意のがん療法（例えば、外科的療法、放射線療法、化学療法など）は、これらの方法の一部として含めてもよい。

【0375】

いくつかの実施形態では、免疫療法又は免疫療法剤は、免疫チェックポイント分子を標的にする。特定の腫瘍は、免疫チェックポイント経路を利用することによって免疫系を回避することができる。したがって、免疫チェックポイントの標的化は、腫瘍の免疫系を回避する能力に対抗し、特定のがんに対する抗腫瘍免疫を活性化させるための有効な手法として認められている。Pardoll, Nature Reviews Cancer, 2012, 12: 252-264。

【0376】

特定の実施形態では、免疫チェックポイント分子は、抗原に対するT細胞応答に関与す

10

20

30

40

50

るシグナルを低減する阻害分子である。例えば、CTLA4は、T細胞上で発現され、抗原提示細胞上のCD80（別名B7.1）又はCD86（別名B7.2）に結合することによって、T細胞活性化の下方制御における役割を発揮する。PD-1は、T細胞上で発現される別の阻害チェックポイント分子である。PD-1は、炎症性応答中、末梢組織内でのT細胞の活性を制限する。さらに、PD-1に対するリガンド（PD-L1又はPD-L2）は、一般に多種の腫瘍の表面上で上方制御され、腫瘍微小環境における抗腫瘍免疫応答の下方制御をもたらす。特定の実施形態では、阻害免疫チェックポイント分子は、CTLA4又はPD-1である。他の実施形態では、阻害免疫チェックポイント分子は、PD-L1又はPD-L2などのPD-1に対するリガンドである。他の実施形態では、阻害免疫チェックポイント分子は、CD80又はCD86などのCTLA4に対するリガンドである。他の実施形態では、阻害免疫チェックポイント分子は、リンパ球活性化遺伝子3（LAG3）、キラー細胞免疫グロブリン様受容体（KIR）、T細胞膜タンパク質3（TIM3）、ガレクチン9（GAL9）、又はアデノシンA2a受容体（A2aR）である。

10

20

30

40

50

【0377】

これらの免疫チェックポイント分子を標的にするアンタゴニストを使用して、特定のがんに対する抗原特異的T細胞応答を増強することができる。したがって、特定の実施形態では、免疫療法又は免疫療法剤は、阻害免疫チェックポイント分子のアンタゴニストである。特定の実施形態では、阻害免疫チェックポイント分子は、PD-1である。特定の実施形態では、阻害免疫チェックポイント分子は、PD-L1である。特定の実施形態では、阻害免疫チェックポイント分子のアンタゴニストは、抗体（例えば、モノクローナル抗体）である。特定の実施形態では、抗体又はモノクローナル抗体は、抗CTLA4、抗PD-1、抗PD-L1、又は抗PD-L2抗体である。特定の実施形態では、抗体は、モノクローナル抗PD-1抗体である。いくつかの実施形態では、抗体は、モノクローナル抗PD-L1抗体である。特定の実施形態では、モノクローナル抗体は、抗CTLA4抗体及び抗PD-1抗体、抗CTLA4抗体及び抗PD-L1抗体、又は抗PD-L1抗体及び抗PD-1抗体の組み合わせである。特定の実施形態では、抗PD-1抗体は、ペンブロリズマブ（キイトルーダ（登録商標））又はニボルマブ（オプジーボ（登録商標））の1つ以上である。特定の実施形態では、抗CTLA4抗体は、イピリムマブ（ヤーボイ（登録商標））である。特定の実施形態では、抗PD-L1抗体は、アテゾリズマブ（テセントリク（登録商標））、アベルマブ（バベンチオ（登録商標））、又はデュルバルマブ（イミフィンジ（登録商標））の1つ以上である。

【0378】

特定の実施形態では、免疫療法又は免疫療法剤は、CD80、CD86、LAG3、KIR、TIM3、GAL9、又はA2aRに対するアンタゴニスト（例えば、抗体）である。他の実施形態では、アンタゴニストは、阻害免疫チェックポイント分子の可溶性バージョン、例えば、阻害免疫チェックポイント分子の細胞外ドメイン及び抗体のFcドメインを含む可溶性融合タンパク質である。特定の実施形態では、可溶性融合タンパク質は、CTLA4、PD-1、PD-L1、又はPD-L2の細胞外ドメインを含む。いくつかの実施形態では、可溶性融合タンパク質は、CD80、CD86、LAG3、KIR、TIM3、GAL9、又はA2aRの細胞外ドメインを含む。一実施形態では、可溶性融合タンパク質は、PD-L2又はLAG3の細胞外ドメインを含む。

【0379】

特定の実施形態では、免疫チェックポイント分子は、抗原に対するT細胞応答に関与するシグナルを増幅する共刺激分子である。例えば、CD28は、T細胞上に発現される共刺激受容体である。T細胞がそのT細胞受容体を通じて抗原に結合する場合、CD28は、抗原提示細胞上のCD80（別名B7.1）又はCD86（別名B7.2）に結合して、T細胞受容体シグナル伝達を増幅し、T細胞活性化を促進する。CD28がCTLA4と同じリガンド（CD80及びCD86）に結合することから、CTLA4は、CD28によって媒介される共刺激シグナル伝達に対抗するか又はそれを調節することができる。

特定の実施形態では、免疫チェックポイント分子は、CD28、誘導性T細胞共刺激因子（ICOS）、CD137、OX40、又はCD27から選択される共刺激分子である。他の実施形態では、免疫チェックポイント分子は、例えば、CD80、CD86、B7RP1、B7-H3、B7-H4、CD137L、OX40L、又はCD70を含む、共刺激分子のリガンドである。

【0380】

これらの共刺激チェックポイント分子を標的にするアゴニストを使用して、特定のがんに対する抗原特異的T細胞応答を増強することができる。したがって、特定の実施形態では、免疫療法又は免疫療法剤は、共刺激チェックポイント分子のアゴニストである。特定の実施形態では、共刺激チェックポイント分子のアゴニストは、アゴニスト抗体であり、好ましくは、モノクローナル抗体である。特定の実施形態では、アゴニスト抗体又はモノクローナル抗体は、抗CD28抗体である。他の実施形態では、アゴニスト抗体又はモノクローナル抗体は、抗ICOS、抗CD137、抗OX40、又は抗CD27抗体である。他の実施形態では、アゴニスト抗体又はモノクローナル抗体は、抗CD80、抗CD86、抗B7RP1、抗B7-H3、抗B7-H4、抗CD137L、抗OX40L、又は抗CD70抗体である。

【0381】

特定の実施形態では、体細胞又は生殖系列が起源である、対象からのサンプルからの核酸変異体の状態を、参照集団からのコンパレータの結果のデータベースと比較して、その対象に対するカスタマイズされた又は標的化された治療法を特定することができる。典型的に、参照集団は、対象と同じ治療法を受けている、又は受けてきている対象及び／又は患者と同じがん又は疾患タイプを有する患者を含む。カスタマイズされた又は標的化された治療法（therapy）（又は治療法（therapies））は、核酸変異体（nucleic variant）及びコンパレータの結果が特定の分類基準を満たす（例えば、実質的に又は概ね一致する）場合、特定することができる。

【0382】

特定の実施形態では、本明細書に記載のカスタマイズされた治療法は、典型的に非経口的に（例えば、静脈内又は皮下に）投与される。免疫療法剤を含有する医薬組成物は、典型的に静脈内に投与される。特定の治療薬は、経口的に投与される。しかし、カスタマイズされた治療法（例えば、免疫療法剤など）はまた、例えば、頬側、舌下、直腸、膣、尿道内、局所、眼内、鼻腔内、及び／又は耳介内といった当該技術分野で公知の任意の方法によって投与されてもよく、投与は、錠剤、カプセル剤、顆粒剤、水性懸濁液、ゲル剤、噴霧剤、坐剤、軟膏、軟膏剤などを含んでもよい。

【0383】

がん以外の特定の遺伝子ベースの疾患、障害、又は病態を治療するための治療選択肢は、一般に当業者に周知であり、検討中の特定の疾患、障害、又は病態を仮定すれば、明らかであろう。

【0384】

いくつかの実施形態では、例えば、遺伝子変異体が検出される場合、治療法は、体細胞又は生殖系列が起源である核酸変異体の状態に基づいてカスタマイズされる。いくつかの実施形態では、本質的に任意のがん療法（例えば、外科的療法、放射線療法、化学療法など）は、これらの方法の一部として含めてもよい。典型的に、カスタマイズされた治療法は、少なくとも1つの免疫療法（又は免疫療法剤）を含む。免疫療法は、一般に所与のがんタイプに対する免疫応答を増強する方法を指す。特定の実施形態では、免疫療法は、腫瘍又はがんに対するT細胞応答を増強する方法を指す。

【0385】

特定の実施形態では、体細胞又は生殖系列が起源である、対象からのサンプルからの核酸変異体の状態を、参照集団からのコンパレータの結果のデータベースと比較して、その対象に対するカスタマイズされた又は標的化された治療法を特定することができる。典型的に、参照集団は、対象と同じ治療法を受けている、又は受けてきている対象及び／又は

患者と同じがん又は疾患タイプを有する患者を含む。カスタマイズされた又は標的化された治療法 (therapy) (又は治療法 (therapies)) は、核酸変異体 (nucleic variant) 及びコンパレータの結果が特定の分類基準を満たす (例えば、実質的に又は概ね一致する) 場合、特定することができる。

【0386】

特定の実施形態では、本明細書に記載のカスタマイズされた治療法は、典型的に非経口的に (例えば、静脈内又は皮下に) 投与される。免疫療法剤を含有する医薬組成物は、典型的に静脈内に投与される。特定の治療薬は、経口的に投与される。しかし、カスタマイズされた治療法 (例えば、免疫療法剤など) はまた、例えば、頬側、舌下、直腸、膣、尿道内、局所、眼内、鼻腔内、及び／又は耳介内などの方法によって投与されてもよく、投与は、錠剤、カプセル剤、顆粒剤、水性懸濁液、ゲル剤、噴霧剤、坐剤、軟膏、軟膏剤などを含んでもよい。

10

【0387】

キット

また提供されるのは、本明細書に記載されるアダプターを含むキットである。アダプターは、品質管理ヌクレオシドを含む。キットは、本明細書に記載される方法の実施における使用を意図することができる。いくつかの実施形態では、キットは、複数のアダプターを含む。

【0388】

いくつかの実施形態では、キットは、1つ以上の変換試薬を含む。いくつかの実施形態では、変換試薬は、品質管理ヌクレオシドの塩基対特異性を変更する。いくつかの実施形態では、変換試薬は、品質管理ヌクレオシドの塩基対特異性を変更しない。いくつかの実施形態では、キットは、バイサルファイト変換、酸化バイサルファイト変換、Tet補助 (TAB) 変換、置換されたボラン還元剤によるTet補助変換、APOBEC連結エピジェネティック (ACE) 変換、又は核酸塩基の酵素変換において使用される変換試薬を含む。hmCを保護するための試薬 (例えば、βGT) は、変換試薬、例えば、Tet酵素及び／又は置換されたボラン還元剤と組み合わせ可能である。いくつかの実施形態では、1つ以上の変換試薬は、Tet酵素及び／又は置換されたボラン還元剤、例えば、本明細書中の他の箇所に記載のもののいずれかを含む。

20

【0389】

いくつかの実施形態では、キットは、複数のプライマーを含む。いくつかの実施形態では、キットは、複数の捕捉プローブを含む。いくつかの実施形態では、キットは、DNAポリメラーゼを含む。いくつかの実施形態では、キットは、デオキシヌクレオシド三リン酸を含む。いくつかの実施形態では、キットは、キットに含まれるアダプターにアニールするPCRプライマーを含む。いくつかの実施形態では、キットは、本明細書中の他の箇所に記載の追加的要素を含む。いくつかの実施形態では、キットは、本明細書に記載の方法を実施するための使用説明書を含む。

30

【0390】

いくつかの実施形態では、キットは、DNA中のメチルシトシンを認識する薬剤をさらに含む。いくつかのそのような実施形態では、薬剤は、抗体又はメチル結合タンパク質又はメチル結合ドメインである。いくつかの実施形態では、キットは、エピジェネティック及び／又は配列可変標的領域セットに特異的に結合する捕捉プローブを含む。いくつかのそのような実施形態では、捕捉プローブは、捕捉部分を含む。いくつかの実施形態では、キットは、捕捉部分の結合パートナーに連結されている固体支持体を含む。

40

【0391】

キットは、ALK、APC、BRAF、CDKN2A、EGFR、ERBB2、FBXW7、KRAS、MYC、NOTCH1、NRAS、PIK3CA、PTEN、RBI、TP53、MET、AR、ABL1、AKT1、ATM、CDH1、CSF1R、CTNNB1、ERBB4、EZH2、FGFR1、FGFR2、FGFR3、FLT3、GNA11、GNAQ、GNAS、HNF1A、HRAS、IDH1、IDH2、JAK2、

50

JAK3、KDR、KIT、MLH1、MPL、NPM1、PDGFRA、PROC、PTPN11、RET、SMAD4、SMARCB1、SMO、SRC、STK11、VHL、TERT、CCND1、CDK4、CDKN2B、RAF1、BRCA1、CCND2、CDK6、NF1、TP53、ARID1A、BRCA2、CCNE1、ESR1、RIT1、GATA3、MAP2K1、RHEB、ROS1、ARAF、MAP2K2、NFE2L2、RHOA、及びNTRK1からなる群から選択される少なくとも5、6、7、8、9、10、20、30、40又は全ての遺伝子と選択的にハイブリダイズする複数のオリゴヌクレオチドプローブをさらに含んでもよい。オリゴヌクレオチドプローブが選択的にハイブリダイズし得る対象の遺伝子の数は、変動し得る。例えば、遺伝子の数は、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、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、又は54を含み得る。キットは、本明細書に記載の方法のいずれかを実施するための複数のオリゴヌクレオチドプローブを含む容器及び使用説明書を含み得る。

10

【0392】

キットは、異なる分子バーコード及び同一のサンプルバーコードを有する、少なくとも4、5、6、7、又は8つの異なるライブラリーアダプターを含み得る。サンプルバーコードは、5mCなどの修飾ヌクレオチドを含んでもよい。ライブラリーアダプターは、配列決定アダプターでなくてもよい。例えば、ライブラリーアダプターは、フローセル配列又は配列決定のためのヘアピンループの形成を可能にする配列を含まない。分子バーコード及びサンプルバーコードの異なるバリエーション及び組み合わせは、全体を通じて記載され、キットに適用可能である。さらに、いくつかの場合、アダプターは、配列決定アダプターでない。加えて、キットとともに提供されるアダプターは、配列決定アダプターも含み得る。配列決定アダプターは、1つ以上の配列決定プライマーにハイブリダイズする配列を含み得る。配列決定アダプターは、固体支持体にハイブリダイズする配列、例えばフローセル配列をさらに含み得る。例えば、配列決定アダプターは、フローセルアダプターであり得る。配列決定アダプターは、ポリヌクレオチド断片の一方又は両方の末端に付着され得る。いくつかの場合、キットは、異なる分子バーコード及び同一のサンプルバーコードを有する少なくとも8つの異なるライブラリーアダプターを含み得る。ライブラリーアダプターは、配列決定アダプターでなくてもよい。キットは、ライブラリーアダプターに選択的にハイブリダイズする第1の配列及びフローセル配列に選択的にハイブリダイズする第2の配列を有する配列決定アダプターをさらに含み得る。別の例では、配列決定アダプターは、ヘアピン形状であり得る。例えば、ヘアピン形状のアダプターは、相補的な二本鎖部分及びループ部分を含み得て、ここで二本鎖部分は、二本鎖ポリヌクレオチドに付着（例えば、ライゲート）され得る。ヘアピン形状の配列決定アダプターは、ポリヌクレオチド断片の両方の末端に付着され、環状分子を生成することができ、それは複数回配列決定され得る。配列決定アダプターは、1つ以上のバーコードを含み得る。例えば、配列決定アダプターは、サンプルバーコードを含み得る。サンプルバーコードは、所定の配列を含み得る。サンプルバーコードを使用して、ポリヌクレオチドの供給源を同定することができる。サンプルバーコードは、少なくとも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、又はそれ以上（又は全体を通じて記載される任意の長さ）の核酸塩基、例えば、少なくとも8塩基であり得る。バーコードは、上記のような隣接又は非隣接配列であり得る。

20

30

40

【0393】

ライブラリーアダプターは、平滑末端及びY形状であり得て、長さが40以下の核酸塩基長であり得る。ライブラリーアダプターの他のバリエーションは、全体を通じて見出すことができ、キットに適用可能である。いくつかの実施形態では、アダプターは、品質管理ヌクレオシドを含む。

50

【0394】

コンピュータシステム

本開示の方法は、コンピュータシステムを使用して、又はその支援によって実行することができる。図2は、本開示の方法を実行するようにプログラムされている、又はそうでなければ設計されているコンピュータシステム201を示す。コンピュータシステム201は、様々な態様としてサンプル調製、配列決定、及び／又は分析を制御することができる。いくつかの例では、コンピュータシステム201は、例えば、本明細書で開示される方法のいずれかに従い、（適用可能である場合）核酸配列決定を含む、サンプル調製及びサンプル分析を実施するように設計されている。

【0395】

コンピュータシステム201は、中央処理ユニット（CPU、本明細書中では「プロセッサ」及び「コンピュータプロセッサ」でもある）205を含み、それはシングルコア若しくはマルチコアプロセッサ、又は並列処理用の複数のプロセッサであり得る。コンピュータシステム201はまた、メモリー又はメモリーロケーション210（例えば、ランダムアクセスメモリー、リードオンリーメモリー、フラッシュメモリー）、電子ストレージユニット215（例えば、ハードディスク）、1つ以上の他のシステムと通信するための通信インタフェース220（例えば、ネットワークアダプター）、及び周辺装置225、例えば、キャッシュ、その他のメモリー、データストレージ、及び／又は電子ディスプレイアダプターを含む。メモリー210、ストレージユニット215、インタフェース220、及び周辺装置225は、マザーボードなどの通信ネットワーク又はバス（実線）を通じてCPU205と通信する。ストレージユニット215は、データを記憶するためのデータストレージユニット（又はデータリポジトリ）であり得る。コンピュータシステム201は、通信インタフェース220の支援によってコンピュータネットワーク230と動作可能に連携され得る。コンピュータネットワーク230は、インターネット、インターネット及び／若しくはエクストラネット、又はインターネットと通信するイントラネット及び／若しくはエクストラネットであり得る。コンピュータネットワーク230は、いくつかの場合、電気通信及び／又はデータネットワークである。コンピュータネットワーク230は、クラウドコンピューティングなどの分散コンピューティングを可能にし得る1つ以上のコンピュータサーバーを含み得る。コンピュータネットワーク230は、いくつかの場合、コンピュータシステム201の支援によって、ピアツーピアネットワークを実行することができ、コンピュータシステム201と連携した装置がクライアント又はサーバーとして機能することを可能にし得る。

【0396】

CPU205は、プログラム又はソフトウェアに実装され得る、機械可読命令のシーケンスを実行することができる。命令は、メモリー210などのメモリーロケーションに記憶され得る。CPU205によって実施される演算の例は、フェッチ、デコード、実行、及びライトバックを含み得る。

【0397】

ストレージユニット215は、ドライバー、ライブラリー、及び保存されたプログラムなどのファイルを記憶することができる。ストレージユニット215は、ユーザによって作成されたプログラム及び記録されたセッション、並びにプログラムに関連する出力を記憶することができる。ストレージユニット215は、ユーザデータ、例えば、ユーザの好み及びユーザプログラムを記憶することができる。コンピュータシステム201は、いくつかの場合、例えば、イントラネット又はインターネットを通じてコンピュータシステム201と通信するリモートサーバー上に位置する、コンピュータシステム201に対して外部である1つ以上の追加的なデータストレージユニットを含み得る。データは、例えば、通信ネットワーク又は物理的なデータ移動を用いて（例えば、ハードドライブ、サムドライブ、又はその他のデータ記憶機構を使用して）、1つのロケーションから別のロケーションに移動され得る。

【0398】

コンピュータシステム 201 は、ネットワーク 230 を通じて、1 つ以上の遠隔コンピュータシステムと通信することができる。実装として、コンピュータシステム 201 は、ユーザ（例えば、オペレーター）の遠隔コンピュータシステムと通信することができる。遠隔コンピュータシステムの例として、パーソナルコンピュータ（例えば、ポータブル PC）、スレート若しくはタブレット PC（例えば、Apple（登録商標）iPad（登録商標）、Samsung（登録商標）Galaxy Tab）、電話、スマートフォン（例えば、Apple（登録商標）iPhone（登録商標）、アンドロイド（登録商標）対応機器、Blackberry（登録商標））、又はパーソナルデジタルアシスタントが挙げられる。ユーザは、ネットワーク 230 を介して、コンピュータシステム 201 にアクセスすることができる。

10

【0399】

本明細書に記載される方法は、コンピュータシステム 201 の電子記憶ロケーション、例えば、メモリ 210 又は電子ストレージユニット 215 などに記憶されているマシン（例えば、コンピュータプロセッサ）で実行可能なコードを用いて実行することができる。機械実行可能又は機械可読コードは、ソフトウェアの形態で提供することができる。使用中、コードは、プロセッサ 205 によって実行することができる。いくつかの場合、コードは、ストレージユニット 215 から検索し、プロセッサ 205 による迅速なアクセスのためにメモリ 210 に記憶することができる。いくつかの状況では、電子ストレージユニット 215 は、排除される可能性があり、機械実行可能命令は、メモリ 210 に記憶される。

20

【0400】

ある態様では、本開示は、少なくとも 1 つの電子プロセッサによって実行されるとき、本明細書に記載の方法の少なくとも一部分を実施する、コンピュータで実行可能な命令を含む非一時的なコンピュータ可読媒体を提供する。例えば、方法は、DNA をオリゴヌクレオチドアダプターにライゲートすることと；適応した DNA、又はそのサブサンプルに変換手順を施すことと；変換ステップ後に適応した DNA を配列決定することと；配列データを使用して、アダプター内の品質管理ヌクレオシドの塩基対特異性変換を判定することと；アダプター内の品質管理ヌクレオシドの塩基対特異性変換を変換ステップについての品質管理尺度として使用することと、を含んでもよく、ここでアダプターは、修飾ヌクレオシドを含む品質管理ヌクレオシドを含み、品質管理ヌクレオシドは、DNA 中の検出対象の修飾ヌクレオシドと同じヌクレオシド同一性及び同じ又は異なる修飾状態を有し、且つ品質管理ヌクレオシドの修飾状態は既知である。

30

【0401】

コードは、コードを実行するように適応させたプロセッサを有するマシンによる使用のため、プリコンパイルし、設計することができ、又は実行中にコンパイルすることができる。コードは、コードがプリコンパイルされた、又はコンパイルされたような様式で実行できるように選択され得るプログラミング言語で供与することができる。

【0402】

コンピュータシステム 201 などの本明細書に提供されるシステム及び方法の態様は、プログラミングで実装することができる。技術の様々な態様は、典型的に、機械可読媒体のタイプで実行又は実装されている、マシン（又はプロセッサ）で実行可能なコード及び／又は関連データの形態における「製品」又は「製造品」として考えることができる。マシンで実行可能なコードは、メモリ（例えば、リードオンリーメモリ、ランダムアクセスメモリ、フラッシュメモリ）又はハードディスクのような電子ストレージユニットに記憶させることができる。「記憶」型媒体は、コンピュータ、プロセッサなどの有形メモリのいずれか又は全て、或いはソフトウェアプログラミング用に随時非一時的記憶を提供可能である、様々な半導体メモリ、テープドライブ、ディスクドライブなどのその関連モジュールを含み得る。

40

【0403】

ソフトウェアの全部又は一部は、時としてインターネット又は様々な他の電気通信ネッ

50

トワークを通じて通信されることがある。例えば、こうした通信では、一方のコンピュータ又はプロセッサをもう一方へ、例えば、管理サーバー又はホストコンピュータをアプリケーションサーバーのコンピュータプラットフォームへ、ソフトウェアをロードすることが可能な場合がある。したがって、ソフトウェアエレメントを有し得る媒体の別タイプは、光波、電波、及び電磁波、例えば、有線及び光回線ネットワークを通じた、及び様々なエアリンクに及ぶ、ローカルデバイス間の物理的なインタフェースを通じて使用されるものを含む。こうした波を運ぶ物理エレメント、例えば、有線又は無線リンク、光リンク、又はそれに類するものは、ソフトウェアを有する媒体と考えることもできる。本明細書で使用する場合、非一時的な有形の「記憶」媒体に制限しない限り、コンピュータ又はマシンの「可読媒体」などの用語は、命令を実行用プロセッサに提供することに関与する任意の媒体を指す。

10

【0404】

それ故、コンピュータで実行可能なコードなどの機械可読媒体は、限定はされないが、有形の記憶媒体、搬送波媒体又は物理的伝達媒体を含む、多くの形態をなすことがある。非揮発性記憶媒体は、例えば、図面に示される、例えば、データベースなどを実行するために使用可能である、光又は磁気ディスク、例えば、任意のコンピュータ又はそれに類するものにおける記憶デバイスのいずれかを含む。揮発性記憶媒体は、動的メモリー、例えば、こうしたコンピュータプラットフォームのメインメモリーを含む。有形の伝達媒体は、同軸ケーブル；銅ワイヤー及び光ファイバー、例えば、コンピュータシステム内のバスを含むワイヤーを含む。搬送波伝達媒体は、電気若しくは電磁気シグナル、又は高周波（RF）及び赤外（IR）データ通信の間に生成されるような音波若しくは光波の形態をなし得る。したがって、コンピュータ可読媒体の一般的形態として、例えば、フロッピーディスク（登録商標）、フレキシブルディスク、ハードディスク、磁気テープ、任意の他の磁気媒体、CD-ROM、DVD若しくはDVD-ROM、任意の他の光学媒体、パンチカード、紙テープ、穴のパターンを有する任意の他の物理的な記憶媒体、RAM、ROM、PROM及びEPROM、FLASH（登録商標）-EPROM、任意の他のメモリーチップ若しくはカートリッジ、搬送波輸送データ若しくは命令、こうした搬送波を輸送するケーブル若しくはリンク、又はコンピュータがプログラミングコード及び／若しくはデータを読み取ることができる元となる任意の他の媒体が挙げられる。コンピュータ可読媒体のこれらの形態の多くは、1つ以上の命令の1つ以上のシーケンスの実行用プロセッサへの運搬に関与し得る。

20

30

【0405】

コンピュータシステム201は、例えば、サンプル分析の1つ以上の結果を提供するため、ユーザーインターフェース（UI）を含む電子ディスプレイを含み得る、又はそれと通信することができる。UIの例としては、限定はされないが、グラフィカルユーザーインターフェース（GUI）及びウェブベースのユーザーインターフェースが挙げられる。

【0406】

また、コンピュータシステム及びネットワーク、データベース、及びコンピュータプログラムプロダクトに関するさらなる詳細は、例えば、Peterson, Computer Networks: A Systems Approach, Morgan Kaufmann, 5th Ed. (2011), Kurose, Computer Networking: A Top-Down Approach, Pearson, 7th Ed. (2016), Elmasri, Fundamentals of Database Systems, Addison Wesley, 6th Ed. (2010), Coronel, Database Systems: Design, Implementation, & Management, Cengage Learning, 11th Ed. (2014), Tucker, Programming Languages, McGraw-Hill Science/Engineering/Math, 2nd Ed. (2006), 及びRhoton, Cloud Computing Architected: Solution Design Handbook, Recursive

40

50

Press (2011) (これらのそれぞれはその全体が参照により本明細書に援用される) に提供されている。

【0407】

本発明の好ましい実施形態が本明細書で示され、記載されている一方で、そのような実施形態があくまで例として提供されることは当業者にとって明白であろう。本発明が本明細書内に提供されている具体例によって限定されることは意図されない。本発明が前述の明細書に関連して記載されている一方で、本明細書中の実施形態の説明及び例示は、限定的意味で解釈されることは意図されていない。ここで当業者は、多数のバリエーション、変更、及び置換を本発明から逸脱することなく設けることになる。さらに、本発明の全ての態様が、種々の条件及び変数に依存する、本明細書に記載される具体的な描写、立体配置又は相対的比率に限定されないことは理解されるべきである。本明細書に記載されている本開示の実施形態に対する様々な代替物が本発明の実施において利用できることは理解されるべきである。したがって、本開示が任意のそのような代替物、修飾、バリエーション又は均等物も包含することが企図される。以下の請求項が本発明の範囲を定義し、且つこれら請求項の範囲内の方法及び構造並びにそれらの等価物がそれによって包含されることが意図される。

10

【0408】

前述の開示が、明確さや理解を目的として、図面及び実施例を用いてある程度詳細に説明されている一方で、形態及び詳細における様々な変更が、本開示の真の範囲から逸脱することなく設けることができ、添付の請求項の範囲内で実施できることは、当業者にとって本開示を通読することから明白となろう。例えば、全ての方法、システム、コンピュータ可読媒体、及び／又は成分特徴、ステップ、要素、又はそれらの他の態様は、様々な組み合わせで使用することができる。

20

【0409】

本明細書中で引用される全ての特許、特許出願、ウェブサイト、その他の出版物又は文書、受入番号などは、あらゆる目的でそれら全体が参照により援用されるが、それは各個別の項目が具体的且つ個別に参照によりそのように援用されることが示された場合と同程度である。配列の異なるバージョンが異なる時点での受入番号を伴う場合、本願の有効出願日での受入番号を伴うバージョンが意図される。有効出願日は、適用可能である場合、受入番号を参照する実際の出願日又は優先権出願の出願日より早期であることを意味する。同様に、出版物、ウェブサイト又はそれに類するものの異なるバージョンが異なる時点で公開される場合、別段の指示がない限り、出願の有効出願日にごく最近公開されたバージョンが意図される。

30

【実施例】

【0410】

サンプル及びライブラリー調製における核酸の修飾ヌクレオシドプロファイル进行分析するための例示的なワークフローが本明細書に提供される。いくつかの実施形態では、分割及びライブラリー調製ワークフローの一部又は全部の特徴は、組み合わせてもよい。

【0411】

40

例示的な方法は、以下のとおりである：

1. 例えば、下流処理のためのプロセスからの全ての溶出物を保存する、メチル結合ドメインタンパク質-ビーズ精製キットを使用する、抽出されたDNAサンプル（例えば、任意選択的に本明細書に記載される標的捕捉を受けている、ヒトサンプルからの抽出された血漿DNA）の任意選択的な物理的分割。

2. 差次的な分子タグ及びNGSを可能にするアダプター配列のサンプルDNA又は各画分への（並行）適用。例えば、高メチル化、残留メチル化（「洗浄」）、及び低メチル化画分を、分子タグとともにNGS-アダプターにライゲートする。サンプル、又はサンプルの少なくとも1つの画分、例えば、高メチル化画分に対するアダプターは、既知のヌクレオシド同一性及び既知の修飾状態を有する、本明細書に記載される、1つ以上の品質

50

管理ヌクレオシドを含む；

3. サンプル、又はサンプルの少なくとも1つの画分、例えば、高メチル化画分に、アダプター内の既知の修飾ヌクレオシドの塩基対特異性を変更する変換手順を施す；

4. 適用可能な場合、全ての分子タグ付け画分の再結合、及びアダプターに特異的なDNAプライマー配列を使用する任意選択的なその後の増幅。

5. 目的のゲノム領域（例えば、がん特異的な遺伝子変異体及び／又は差次的メチル化領域）を標的にする、（再結合及び／又は増幅された）全ライブラリーの捕捉／ハイブリダイゼーション。

6. サンプルタグを付加している、捕捉されたDNAライブラリーの任意選択的な増幅。異なるサンプルをプールし、NGS機器で多重にアッセイする。

7. 分子タグを任意選択的に使用して固有の分子が同定される、NGSデータのバイオフィンフォーマティクス解析、及び差次的に分割された分子へのサンプルのデコンボリューション。遺伝子配列決定／変異体検出では、正確に変換されているか又は変換されていないアダプター領域内の品質管理ヌクレオシドを同定する。サンプル又は分子レベルの変換率を判定し、分析のための品質管理尺度として使用する。

8. 任意選択的に、スコア又は加重をそれぞれの同定された固有の分子に、アダプター領域内の既知の修飾ヌクレオシドの変換率に応じて適用する。スコア又は加重を、例えば、診断又は予後において使用されるとき、サンプルデータからの下流の分析又は推測において適用する。

【0412】

任意選択的に、DNAを分割及び／又は濃縮した後、例えば、DNA分子の末端オーバーハングを伸長し、そしてアデノシン残基を断片の3'末端に付加し、各DNA断片の5'末端をリン酸化することによって、DNAをライゲート可能にすることができる。DNAリガーゼ及びアダプターを付加して、分割された各DNA分子を一方の末端又は両方の末端でアダプターとライゲートする。これらのアダプターは、他の画分において使用されるアダプター内の分割タグと区別可能である分割タグ（例えば、非無作為な非固有のバーコード）を含有してもよい。DNAをライゲート可能にして、ライゲーションを実施した後、サンプル、又はその画分（例えば、高メチル化画分）は、本明細書に記載のとおり、例えば、TAPS法を用いて、修飾状態に依存しているアダプター内のヌクレオシドの塩基対特異性を変更する変換手順を受ける。サンプルが分割されている場合、画分は、増幅前に一緒にプールすることができる。増幅は、アダプターに特異的なプライマーを用いるPCRによって行ってもよい。

【0413】

増幅後、DNAは、富化前に、洗浄し、濃縮することができる。増幅されたDNAは、目的の特定領域を標的にする、例えば本明細書に記載される、プローブの収集物（例えば、ビオチン化RNAプローブであってもよい）と接触させることができる。混合物は、例えば一晩、例えば塩緩衝液中でインキュベートすることができる。プローブは、（例えば、ストレプトアビジン磁気ビーズを使用して）捕捉し、例えば一連の塩洗浄によって、捕捉されなかった増幅されたDNAから分離し、それによりサンプルを富化することができる。富化後、富化されたサンプルは、例えばPCRによって（さらに）増幅することができる。いくつかの実施形態では、PCRプライマーは、サンプルタグを含有し、それにより、サンプルタグをDNA分子に取り込ませる。いくつかの実施形態では、異なるサンプルからのDNAと一緒にプールし、次に例えば、Illumina NovaSeqシーケンサーを使用して多重配列決定する。

10

20

30

40

50

【 図 面 】
【 図 1 】

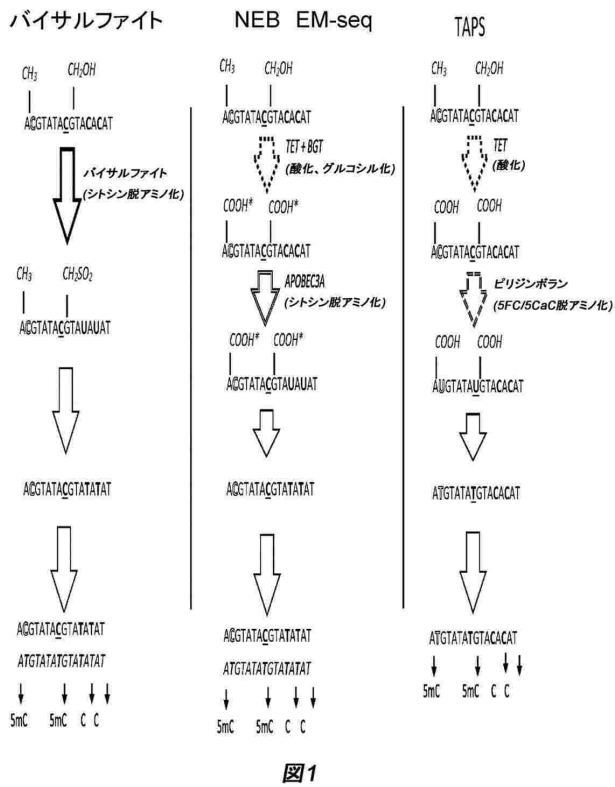


図 1

【 図 2 】

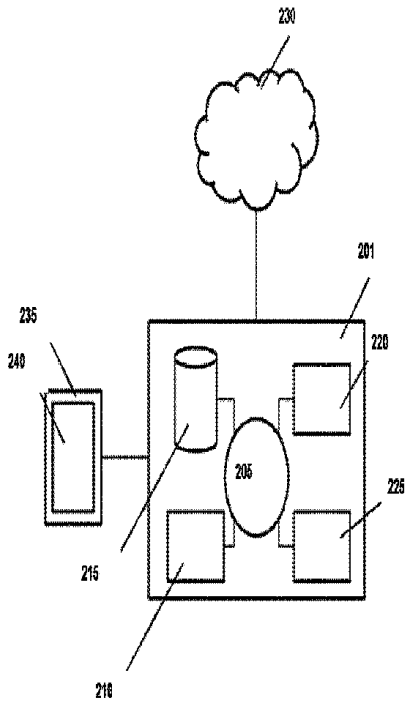


FIG. 2

【 図 3 】

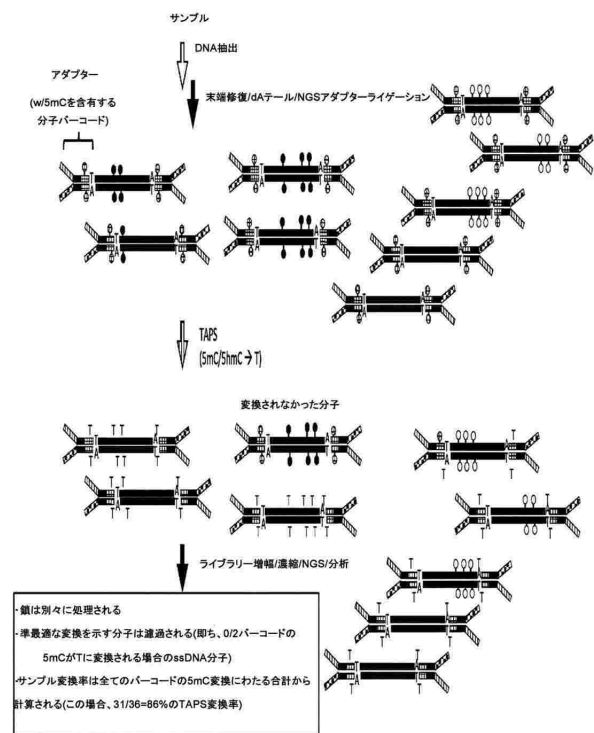


図 3

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2022/079173

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12Q1/6855
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2021/178893 A2 (SINGULAR GENOMICS SYSTEMS INC [US]) 10 September 2021 (2021-09-10) paragraph [0297]; figures 15A, 17A; example 1 -----	1-50
X	WO 2020/243609 A1 (FREEMOME HOLDINGS INC [US]) 3 December 2020 (2020-12-03) paragraph [0183]; figures 4, 6 -----	45-49
X	WO 2015/104302 A1 (FUNDACIÓ PRIVADA INST DE MEDICINA PREDICTIVA I PERSONALITZADA DEL C NC) 16 July 2015 (2015-07-16) page 44; figure 1 -----	45-49
X	WO 2016/016639 A1 (CAMBRIDGE EPIGENETIX LTD [GB]) 4 February 2016 (2016-02-04) claims 1, 16 -----	45-49
-/-		

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

13 June 2023

Date of mailing of the international search report

22/06/2023

Name and mailing address of the ISA/
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Aslund, Fredrik

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2022/079173

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2021/005537 A1 (THE CHANCELLOR MASTERS AND SCHOLARS OF THE UNIV OF OXFORD [GB]) 14 January 2021 (2021-01-14) figure 5 -----	1-50

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2022/079173

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
- a. ☐ forming part of the international application as filed.
- b. ☒ furnished subsequent to the international filing date for the purposes of international search (Rule 13ter.1(a)).
- ☒ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.
2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.
3. Additional comments:

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2022/079173

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2021178893 A2	10-09-2021	EP 4114978 A2	11-01-2023
		US 2021277461 A1	09-09-2021
		US 2022033895 A1	03-02-2022
		US 2022119878 A1	21-04-2022
		US 2022282324 A1	08-09-2022
		US 2022325341 A1	13-10-2022
		WO 2021178893 A2	10-09-2021
WO 2020243609 A1	03-12-2020	AU 2020283042 A1	01-04-2021
		CA 3111019 A1	03-12-2020
		CN 112823213 A	18-05-2021
		EP 3830285 A1	09-06-2021
		JP 2022539443 A	09-09-2022
		KR 20220015367 A	08-02-2022
		SG 11202101998U A	30-03-2021
		US 2021230684 A1	29-07-2021
		WO 2020243609 A1	03-12-2020
WO 2015104302 A1	16-07-2015	AU 2015205612 A1	21-07-2016
		CA 2936564 A1	16-07-2015
		CN 106103743 A	09-11-2016
		EP 3092308 A1	16-11-2016
		JP 6895753 B2	30-06-2021
		JP 2017501739 A	19-01-2017
		KR 20160111403 A	26-09-2016
		RU 2016132476 A	13-02-2018
		US 2017002404 A1	05-01-2017
		US 2019177776 A1	13-06-2019
		US 2022213533 A1	07-07-2022
		WO 2015104302 A1	16-07-2015
WO 2016016639 A1	04-02-2016	EP 3174996 A1	07-06-2017
		WO 2016016639 A1	04-02-2016
WO 2021005537 A1	14-01-2021	AU 2020310613 A1	17-02-2022
		BR 112022000361 A2	10-05-2022
		CA 3146212 A1	14-01-2021
		CN 114174534 A	11-03-2022
		EP 3997245 A1	18-05-2022
		IL 289669 A	01-03-2022
		JP 2022540453 A	15-09-2022
		KR 20220038367 A	28-03-2022
		US 2022275424 A1	01-09-2022
		WO 2021005537 A1	14-01-2021

フロントページの続き

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

弁護士 山本 健策

(72)発明者 ケネディ, アンドリュー

アメリカ合衆国 カリフォルニア 94304, パロ アルト, ハノーバー ストリート 3100

Fターム(参考) 4B063 QA20 QQ42 QR90 QS39 QX10