

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号
特表2022-514010
(P2022-514010A)

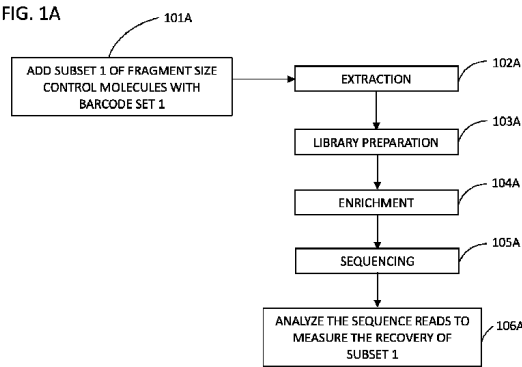
(43)公表日 令和4年2月9日(2022.2.9)

(51)国際特許分類		F I		テーマコード (参考)	
C 1 2 Q	1/24 (2006.01)	C 1 2 Q	1/24	4 B 0 6 3	
C 1 2 N	15/10 (2006.01)	C 1 2 N	15/10	Z	
C 1 2 N	15/11 (2006.01)	C 1 2 N	15/10	1 0 0 Z	
C 1 2 Q	1/6813(2018.01)	C 1 2 N	15/11	Z	
C 1 2 Q	1/6844(2018.01)	C 1 2 Q	1/6813	Z	
		審査請求	未請求	予備審査請求	未請求 (全78頁) 最終頁に続く
(21)出願番号	特願2021-535033(P2021-535033)	(71)出願人	515059083		
(86)(22)出願日	令和1年12月20日(2019.12.20)		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3年8月4日(2021.8.4)		アメリカ合衆国 カリフォルニア 9 4 0		
(86)国際出願番号	PCT/US2019/068146		6 3, レッドウッド シティ, ペノブ		
(87)国際公開番号	WO2020/132628		スコット ドライブ 5 0 5		
(87)国際公開日	令和2年6月25日(2020.6.25)	(74)代理人	100078282		
(31)優先権主張番号	62/783,046		弁理士 山本 秀策		
(32)優先日	平成30年12月20日(2018.12.20)	(74)代理人	100113413		
(33)優先権主張国・地域又は機関	米国(US)		弁理士 森下 夏樹		
(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,A T,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	100181674		
			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】 核酸分子の回収率を改善するための方法、組成物、およびシステム

(57)【要約】

一態様では、本開示は、ポリヌクレオチドの試料中の核酸分子を分析するための方法であって (a) 断片サイズ対照分子のサブセットを核酸分子に添加し第1のスパイクイン試料を産生するステップ (b) 第1のスパイクイン試料から核酸を抽出するステップ (c) 抽出された核酸の少なくともサブセットを処理し処理された試料を産生するステップであって処理することが第1のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含むステップ (d) 処理された試料の少なくともサブセットを富化し富化された試料を産生するステップ (e) 富化された試料の少なくともサブセットをシーケンシングして複数の配列リードを生成するステップ (f) 配列リードを分析して断片サイズ対照分子のサブセットの断片サイズスコアを生成するステップならびに (g) 断片サイズスコアを断片サイズ閾値と比較するステップを含む方法を提供する。



【特許請求の範囲】

【請求項 1】

既定の断片サイズ対照分子の少なくとも 1 つのサブセットを含む断片サイズ対照分子のセットであって、既定の断片サイズ対照分子の前記少なくとも 1 つのサブセットが、断片サイズ領域を含む複数の断片サイズ対照分子を含む、断片サイズ対照分子のセット。

【請求項 2】

前記少なくとも 1 つのサブセットが、断片サイズ対照分子の少なくとも 1 つの群を含む、請求項 1 に記載の断片サイズ対照分子のセット。

【請求項 3】

群における前記断片サイズ対照分子の前記断片サイズ領域が、同じ長さの領域である、請求項 2 に記載の断片サイズ対照分子のセット。 10

【請求項 4】

断片サイズ対照分子の第 1 の群における前記断片サイズ領域の長さが、断片サイズ対照分子の第 2 の群における前記断片サイズ領域の長さとは異なる、請求項 2 または 3 に記載の断片サイズ対照分子のセット。

【請求項 5】

前記断片サイズ対照分子が、識別子領域をさらに含む、請求項 1 に記載の断片サイズ対照分子のセット。

【請求項 6】

前記識別子領域が、前記断片サイズ領域の片側または両側に存在する、請求項 5 に記載の断片サイズ対照分子のセット。 20

【請求項 7】

前記識別子領域が分子バーコードを含む、請求項 5 に記載の断片サイズ対照分子のセット。

【請求項 8】

前記複数の断片対照分子が、1 つまたは複数のプライマー結合部位を含む、上記請求項のいずれかに記載の断片サイズ対照分子のセット。

【請求項 9】

前記 1 つまたは複数のプライマー結合部位が、前記識別子領域に存在する、請求項 8 に記載の断片サイズ対照分子のセット。 30

【請求項 10】

群における前記断片サイズ対照分子の前記断片サイズ領域が、同じオリゴヌクレオチド配列を含む、請求項 2 ～ 4 のいずれかに記載の断片サイズ対照分子のセット。

【請求項 11】

群における前記断片サイズ対照分子の前記断片サイズ領域が、少なくとも 2 つの区別可能なオリゴヌクレオチド配列を含む、請求項 2 ～ 9 のいずれかに記載の断片サイズ対照分子のセット。

【請求項 12】

既定の断片サイズ対照分子の第 1 のサブセットにおける前記断片サイズ対照分子の前記断片サイズ領域が、既定の断片サイズ対照分子の第 2 のサブセットにおける前記断片サイズ対照分子の前記断片サイズ領域のオリゴヌクレオチド配列とは区別可能なオリゴヌクレオチド配列を含む、上記請求項のいずれかに記載の断片サイズ対照分子のセット。 40

【請求項 13】

前記断片サイズ領域の長さが、少なくとも 10 b p、少なくとも 50 b p、少なくとも 60 b p、少なくとも 70 b p、少なくとも 80 b p、少なくとも 90 b p、少なくとも 100 b p、少なくとも 120 b p、少なくとも 150 b p、少なくとも 200 b p、少なくとも 250 b p、少なくとも 300 b p、少なくとも 400 b p、少なくとも 500 b p、少なくとも 600 b p、少なくとも 700 b p、少なくとも 800 b p、少なくとも 900 b p、または少なくとも 1000 b p である、上記請求項のいずれかに記載の断片サイズ対照分子のセット。 50

【請求項 14】

前記断片サイズ領域の長さが、10bp～1000bpの間である、上記請求項のいずれかに記載の断片サイズ対照分子のセット。

【請求項 15】

既定の断片サイズ対照分子の前記少なくとも1つのサブセットの各サブセットが、等モル濃度で存在する、上記請求項のいずれかに記載の断片サイズ対照分子のセット。

【請求項 16】

既定の断片サイズ対照分子の前記少なくとも1つのサブセットの各サブセットが、等しくないモル濃度で存在する、上記請求項のいずれかに記載の断片サイズ対照分子のセット。

【請求項 17】

前記少なくとも1つのサブセットにおける断片サイズ対照分子の前記少なくとも1つの群の各群が、等モル濃度で存在する、上記請求項のいずれかに記載の断片サイズ対照分子のセット。

【請求項 18】

前記少なくとも1つのサブセットにおける断片サイズ対照分子の前記少なくとも1つの群の各群が、等しくないモル濃度で存在する、請求項2～16のいずれかに記載の断片サイズ対照分子のセット。

【請求項 19】

a. 既定の断片サイズ対照分子の少なくとも1つのサブセットを含む断片サイズ対照分子のセットであって、既定の断片サイズ対照分子の前記少なくとも1つのサブセットが断片サイズ領域を含む複数の断片サイズ対照分子を含む、断片サイズ対照分子のセット；および

b. 対象からのポリヌクレオチドの試料中の核酸分子のセットを含む、核酸の集団。

【請求項 20】

前記少なくとも1つのサブセットが、断片サイズ対照分子の少なくとも1つの群を含む、請求項19に記載の核酸の集団。

【請求項 21】

群における前記断片サイズ対照分子の前記断片サイズ領域が、同じ長さの領域である、請求項20に記載の核酸の集団。

【請求項 22】

断片サイズ対照分子の第1の群における前記断片サイズ領域の長さが、断片サイズ対照分子の第2の群における前記断片サイズ領域の長さとは異なる、請求項20または21に記載の核酸の集団。

【請求項 23】

前記断片サイズ対照分子が、識別子領域をさらに含む、請求項19に記載の核酸の集団。

【請求項 24】

前記識別子領域が、前記断片サイズ領域の片側または両側に存在する、請求項23に記載の核酸の集団。

【請求項 25】

前記識別子領域が分子バーコードを含む、請求項23に記載の核酸の集団。

【請求項 26】

前記複数の断片対照分子が、1つまたは複数のプライマー結合部位を含む、請求項19～25のいずれかに記載の核酸の集団。

【請求項 27】

前記1つまたは複数のプライマー結合部位が、前記識別子領域に存在する、請求項26に記載の核酸の集団。

【請求項 28】

群における前記断片サイズ対照分子の前記断片サイズ領域が、同じオリゴヌクレオチド配列を含む、請求項20～22のいずれかに記載の核酸の集団。

10

20

30

40

50

【請求項 29】

群における前記断片サイズ対照分子の前記断片サイズ領域が、少なくとも2つの区別可能なオリゴヌクレオチド配列を含む、請求項20～27のいずれかに記載の断片サイズ対照分子のセット。

【請求項 30】

既定の断片サイズ対照分子の第1のサブセットにおける前記断片サイズ対照分子の前記断片サイズ領域が、既定の断片サイズ対照分子の第2のサブセットにおける前記断片サイズ対照分子の前記断片サイズ領域のオリゴヌクレオチド配列と区別可能なオリゴヌクレオチド配列を含む、請求項19～29のいずれかに記載の核酸の集団。

【請求項 31】

前記断片サイズ領域の長さが、少なくとも10bp、少なくとも50bp、少なくとも60bp、少なくとも70bp、少なくとも80bp、少なくとも90bp、少なくとも100bp、少なくとも120bp、少なくとも150bp、少なくとも200bp、少なくとも250bp、少なくとも300bp、少なくとも400bp、少なくとも500bp、少なくとも600bp、少なくとも700bp、少なくとも800bp、少なくとも900bp、または少なくとも1000bpである、請求項19～30のいずれかに記載の核酸の集団。

【請求項 32】

前記断片サイズ領域の長さが、10bp～1000bpの間である、請求項19～31のいずれかに記載の核酸の集団。

【請求項 33】

既定の断片サイズ対照分子の前記少なくとも1つのサブセットの各サブセットが、等モル濃度で存在する、請求項19～32のいずれかに記載の核酸の集団。

【請求項 34】

既定の断片サイズ対照分子の前記少なくとも1つのサブセットの各サブセットが、等しくないモル濃度で存在する、請求項19～33のいずれかに記載の核酸の集団。

【請求項 35】

前記少なくとも1つのサブセットにおける断片サイズ対照分子の前記少なくとも1つの群の各群が、等モル濃度で存在する、請求項19～34のいずれかに記載の核酸の集団。

【請求項 36】

前記少なくとも1つのサブセットにおける断片サイズ対照分子の前記少なくとも1つの群の各群が、等しくないモル濃度で存在する、請求項20～34のいずれかに記載の核酸の集団。

【請求項 37】

ポリヌクレオチドの試料中の核酸分子を分析するための方法であって：

- a) 断片サイズ対照分子のサブセットをポリヌクレオチドの前記試料中の前記核酸分子に添加し、それによって第1のスパイクイン試料を産生するステップ；
- b) 前記第1のスパイクイン試料から核酸を抽出するステップ；
- c) 前記抽出された核酸の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、前記処理することが前記第1のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；
- d) 前記処理された試料の少なくともサブセットを富化し、それによって富化された試料を産生するステップ；
- e) 前記富化された試料の少なくともサブセットをシーケンシングして、複数の配列リードを生成するステップ；ならびに
- f) 前記複数の配列リードを分析して、断片サイズ対照分子の前記サブセットの複数の断片サイズスコアを生成するステップを含む方法。

【請求項 38】

c) の前に、断片サイズ対照分子の第2のサブセットを添加し、それによって第2のスパ

イクイン試料を産生するステップをさらに含む、請求項 37 に記載の方法。

【請求項 39】

d) の前に、断片サイズ対照分子の第 3 のサブセットを添加し、それによって第 3 のスパイクイン試料を産生するステップをさらに含む、請求項 38 に記載の方法。

【請求項 40】

e) の前に、断片サイズ対照分子の第 4 のサブセットを添加し、それによって第 4 のスパイクイン試料を産生するステップをさらに含む、請求項 39 に記載の方法。

【請求項 41】

ポリヌクレオチドの試料中の核酸分子を分析するための方法であって：

a. 断片サイズ対照分子の第 1 のサブセットをポリヌクレオチドの前記試料中の前記核酸分子に添加し、それによって第 1 のスパイクイン試料を産生するステップ；

b. 前記第 1 のスパイクイン試料から核酸を抽出するステップ；

c. 断片サイズ対照分子の第 2 のサブセットを前記抽出された核酸に添加し、それによって第 2 のスパイクイン試料を産生するステップ；

d. 前記第 2 のスパイクイン試料の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、前記処理することが、前記第 2 のスパイクイン試料の前記少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；

e. 断片サイズ対照分子の第 3 のサブセットを前記処理された試料に添加し、それによって第 3 のスパイクイン試料を産生するステップ；

f. 前記第 3 のスパイクイン試料の少なくともサブセットを富化し、それによって富化された試料を産生するステップ；

g. 断片サイズ対照分子の第 4 のサブセットを前記富化された試料の少なくともサブセットに添加し、それによって第 4 のスパイクイン試料を産生するステップ；

h. 前記第 4 のスパイクイン試料をシーケンシングして、複数の配列リードを生成するステップ；ならびに

i. 前記複数の配列リードを分析して、断片サイズ対照分子の前記第 1 のサブセット、断片サイズ対照分子の前記第 2 のサブセット、断片サイズ対照分子の前記第 3 のサブセット、および／または断片サイズ対照分子の前記第 4 のサブセットの複数の断片サイズスコアを生成するステップ

を含む方法。

【請求項 42】

前記複数の断片サイズスコアを複数の断片サイズ閾値と比較するステップをさらに含む、請求項 37 または 41 に記載の方法。

【請求項 43】

前記複数の断片サイズスコアに基づいてポリヌクレオチドの前記試料中の核酸分子の前記分析を最適化するステップをさらに含む、請求項 42 に記載の方法。

【請求項 44】

前記複数の断片サイズスコアを使用して、ポリヌクレオチドの前記試料中の核酸分子の前記分析における断片サイズバイアスを補正するステップをさらに含む、請求項 42 に記載の方法。

【請求項 45】

前記方法を、(i) 前記複数の断片サイズスコアの少なくとも 1 つが前記複数の断片サイズ閾値の対応する断片サイズ閾値内にある場合、成功である；または(ii) 前記複数の断片サイズスコアの少なくとも 1 つが前記複数の断片サイズ閾値の前記対応する断片サイズ閾値内でない場合、失敗であると分類するステップをさらに含む、請求項 42 に記載の方法。

【請求項 46】

第 1 の試料の第 2 の試料による汚染を検出する方法であって、前記第 1 の試料および前記第 2 の試料の各々に関して：

10

20

30

40

50

a. 断片サイズ対照分子のサブセットを添加して第1のスパイクイン試料を生成するステップであって、前記第1の試料に添加された断片サイズ対照分子の前記サブセットが、前記第2の試料に添加された断片サイズ対照分子の前記サブセットと区別され得る、ステップ；

b. 前記第1のスパイクインから核酸を抽出するステップ；

c. 前記抽出された核酸の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、前記処理することが、前記第1のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；

d. 前記処理された試料の少なくともサブセットを富化し、それによって富化された試料を産生するステップ；

e. 前記富化された試料の少なくともサブセットをシーケンシングして、複数の配列リードを生成するステップ；ならびに

f. 前記複数の配列リードを分析して、断片サイズ対照分子の前記サブセットの1つまたは複数の汚染スコアを生成するステップを含む方法。

10

【請求項47】

c)の前に、断片サイズ対照分子の第2のサブセットを添加し、それによって第2のスパイクイン試料を産生するステップであって、前記第1の試料に添加された断片サイズ対照分子の前記サブセットが、前記第2の試料に添加された断片サイズ対照分子の前記サブセットと区別され得る、ステップをさらに含む、請求項46に記載の方法。

20

【請求項48】

d)の前に、断片サイズ対照分子の第3のサブセットを添加し、それによって第3のスパイクイン試料を産生するステップであって、前記第1の試料に添加された断片サイズ対照分子の前記サブセットが、前記第2の試料に添加された断片サイズ対照分子の前記サブセットと区別され得る、ステップをさらに含む、請求項47に記載の方法。

【請求項49】

e)の前に、断片サイズ対照分子の第4のサブセットを添加し、それによって第4のスパイクイン試料を産生するステップであって、前記第1の試料に添加された断片サイズ対照分子の前記サブセットが、前記第2の試料に添加された断片サイズ対照分子の前記サブセットと区別され得る、ステップをさらに含む、請求項48に記載の方法。

30

【請求項50】

第1の試料の第2の試料による汚染を検出する方法であって、前記第1の試料および前記第2の試料の各々に関して：

a. 断片サイズ対照分子の第1のサブセットを添加して第1のスパイクイン試料を生成するステップであって、前記第1の試料に添加された断片サイズ対照分子の前記第1のサブセットが、前記第2の試料に添加された断片サイズ対照分子の前記第1のサブセットと区別され得る、ステップ；

b. 前記第1のスパイクインから核酸を抽出するステップ；

c. 断片サイズ対照分子の第2のサブセットを前記抽出された核酸に添加し、それによって第2のスパイクイン試料を産生するステップであって、前記第1の試料に添加された断片サイズ対照分子の前記第2のサブセットが、前記第2の試料に添加された断片サイズ対照分子の前記第2のサブセットと区別され得る、ステップ；

40

d. 前記抽出された核酸の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、前記処理することが、前記第1のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；

e. 断片サイズ対照分子の第3のサブセットを前記抽出された核酸に添加し、それによって第3のスパイクイン試料を産生するステップであって、前記第1の試料に添加された断片サイズ対照分子の前記第3のサブセットが、前記第2の試料に添加された断片サイズ対照分子の前記第3のサブセットと区別され得る、ステップ；

f. 前記処理された試料の少なくともサブセットを富化するステップ；

50

g. 断片サイズ対照分子の第 4 のサブセットを前記抽出された核酸分子に添加し、それによって第 4 のスパイクイン試料を産生するステップであって、前記第 1 の試料に添加された断片サイズ対照分子の前記第 4 のサブセットが、前記第 2 の試料に添加された断片サイズ対照分子の前記第 4 のサブセットと区別され得る、ステップ；

h. 前記富化された試料の少なくともサブセットをシーケンシングして、複数の配列リードを生成するステップ；ならびに

i. 前記複数の配列リードを分析して、断片サイズ対照分子の前記サブセットの 1 つまたは複数の汚染スコアを生成するステップを含む方法。

【請求項 5 1】

少なくとも 1 つまたは複数の汚染スコアを少なくとも 1 つまたは複数の汚染閾値と比較するステップをさらに含む、請求項 4 6 ~ 5 0 に記載の方法。

【請求項 5 2】

前記第 1 の試料を、(i) 前記汚染スコアの少なくとも 1 つまたは複数の前記 1 つまたは複数の汚染閾値の対応する汚染閾値内にない場合、前記第 2 の試料によって汚染されている；または (i i) 前記汚染スコアの少なくとも 1 つまたは複数の前記 1 つまたは複数の汚染閾値の前記対応する汚染閾値内にある場合、前記第 2 の試料によって汚染されていないと分類するステップをさらに含む、請求項 5 1 に記載の方法。

【請求項 5 3】

断片サイズ対照分子の前記サブセットが、断片サイズ領域を含む複数の断片サイズ対照分子を含む、請求項 3 7 ~ 5 2 のいずれかに記載の方法。

【請求項 5 4】

前記断片サイズ対照分子が、識別子領域をさらに含む、請求項 5 3 に記載の方法。

【請求項 5 5】

前記サブセットが、断片サイズ対照分子の少なくとも 1 つの群を含む、請求項 5 3 または 5 4 に記載の方法。

【請求項 5 6】

群における前記複数の断片サイズ対照分子の前記断片サイズ領域が、同じ長さの領域である、請求項 5 5 に記載の方法。

【請求項 5 7】

断片サイズ対照分子の第 1 の群における前記断片サイズ領域の長さが、断片サイズ対照分子の第 2 の群における前記断片サイズ領域の長さとは異なる、請求項 5 5 に記載の方法。

【請求項 5 8】

前記識別子領域が、前記断片サイズ領域の片側または両側に存在する、請求項 5 4 に記載の方法。

【請求項 5 9】

前記識別子領域が分子バーコードを含む、請求項 5 4 に記載の方法。

【請求項 6 0】

前記断片サイズ対照分子が、1 つまたは複数のプライマー結合部位を含む、上記請求項のいずれかに記載の方法。

【請求項 6 1】

前記プライマー結合部位が、前記識別子領域に存在する、請求項 6 0 に記載の方法。

【請求項 6 2】

群における前記断片サイズ対照分子の前記断片サイズ領域が、同じオリゴヌクレオチド配列を含む、請求項 3 7 ~ 6 1 のいずれかに記載の方法。

【請求項 6 3】

群における前記断片サイズ対照分子の前記断片サイズ領域が、少なくとも 2 つの区別可能なオリゴヌクレオチド配列を含む、請求項 5 5 に記載の方法。

【請求項 6 4】

断片サイズ対照分子の第 1 のサブセットにおける前記断片サイズ対照分子の前記断片サイ

10

20

30

40

50

ズ領域が、断片サイズ対照分子の第2のサブセットにおける前記断片サイズ対照分子の前記断片サイズ領域のオリゴヌクレオチド配列と区別可能なオリゴヌクレオチド配列を含む、請求項37～63のいずれかに記載の方法。

【請求項65】

前記断片サイズ領域の長さが、少なくとも10bp、少なくとも50bp、少なくとも60bp、少なくとも70bp、少なくとも80bp、少なくとも90bp、少なくとも100bp、少なくとも120bp、少なくとも150bp、少なくとも200bp、少なくとも250bp、少なくとも300bp、少なくとも400bp、少なくとも500bp、少なくとも600bp、少なくとも700bp、少なくとも800bp、少なくとも900bp、または少なくとも1000bpである、請求項37～64のいずれかに記載の方法。

10

【請求項66】

前記断片サイズ領域の長さが、10bp～1000bpの間である、請求項37～64のいずれかに記載の方法。

【請求項67】

前記断片サイズ対照分子の前記サブセットの各々が、等モル濃度で存在する、請求項38～66のいずれかに記載の方法。

【請求項68】

前記断片サイズ対照分子の前記サブセットの各々が、等しくないモル濃度で存在する、請求項38～66のいずれかに記載の方法。

20

【請求項69】

前記サブセットにおける前記断片サイズ対照分子の前記群の各々が、等モル濃度で存在する、請求項37～68のいずれかに記載の方法。

【請求項70】

前記サブセットにおける前記断片サイズ対照分子の前記群の各々が、等しくないモル濃度で存在する、請求項37～68のいずれかに記載の方法。

【請求項71】

前記分画することが、前記第2のスパイクイン試料の前記少なくともサブセットの前記核酸分子を複数の分画化セットに分画することを含む、請求項37または41に記載の方法。

30

【請求項72】

前記複数の分画化セットが、前記第2のスパイクイン試料の前記核酸分子のエピジェネティック修飾レベルに基づいて分画された前記第2のスパイクイン試料の核酸分子を含む、請求項71に記載の方法。

【請求項73】

前記タグ付けすることが、タグのセットを前記核酸に付着させて、タグ付けされた核酸の集団を生成することであって、前記タグ付けされた核酸が1つまたは複数のタグを含む、ことを含む、請求項37または41に記載の方法。

【請求項74】

前記分画化に起因する複数の分画化セットの第1の分画化セットにおいて使用されるタグの前記セットが、前記複数の分画化セットの第2の分画化セットにおいて使用されるタグの前記セットとは異なる、請求項73に記載の方法。

40

【請求項75】

タグの前記セットが、アダプターの前記核酸へのライゲーションによって前記核酸に付着され、前記アダプターが1つまたは複数のタグを含む、請求項74に記載の方法。

【請求項76】

ポリヌクレオチドの試料のシーケンシングライブラリを生成するための方法であって：
a) 断片サイズ対照分子のサブセットを前記試料に添加し、それによって第1のスパイクイン試料を生成するステップ；
b) 前記第1のスパイクイン試料から核酸を抽出するステップ；

50

c) 前記抽出された核酸の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、前記処理することが、前記第1のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；ならびに

d) 前記処理された試料の少なくともサブセットを富化するステップを含む方法。

【請求項77】

c) の前に、断片サイズ対照分子の第2のサブセットを添加し、それによって第2のスパイクイン試料を産生するステップをさらに含む、請求項76に記載の方法。

【請求項78】

d) の前に、断片サイズ対照分子の第3のサブセットを添加し、それによって第3のスパイクイン試料を産生するステップをさらに含む、請求項77に記載の方法。

【請求項79】

e) 断片サイズ対照分子の第4のサブセットを添加し、それによって第4のスパイクイン試料を産生するステップをさらに含む、請求項78に記載の方法。

【請求項80】

前記断片サイズ対照分子の濃度が、1アトム〜10ピコモルの間である、請求項37〜79のいずれかに記載の方法。

【請求項81】

ポリヌクレオチドの前記試料が、無細胞ポリヌクレオチドの試料である、請求項37〜80のいずれかに記載の方法。

【請求項82】

無細胞ポリヌクレオチドの前記試料が、無細胞DNAの試料および無細胞RNAの試料からなる群から選択される、請求項81に記載の方法。

【請求項83】

無細胞ポリヌクレオチドの前記試料が、無細胞DNAの試料である、請求項81に記載の方法。

【請求項84】

前記無細胞DNAが、1ng〜500ngの間である、請求項81に記載の方法。

【請求項85】

少なくとも1つの電子プロセッサによって実行された場合に方法を実施する非一時的コンピュータ実行可能命令を含むコンピュータ可読媒体を含むか、またはそれにアクセスすることが可能なコントローラを含むシステムであって、前記方法が：

a. 断片サイズ対照分子のサブセットをポリヌクレオチドの試料中の核酸分子に添加し、それによって第1のスパイクイン試料を産生するステップ；

b. 前記第1のスパイクイン試料から核酸を抽出するステップ；

c. 前記抽出された核酸の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、前記処理することが、前記第1のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；

d. 前記処理された試料の少なくともサブセットを富化し、それによって富化された試料を産生するステップ；

e. 前記富化された試料の少なくともサブセットをシーケンシングして、複数の配列リードを生成するステップ；ならびに

f. 前記複数の配列リードを分析して、断片サイズ対照分子の前記サブセットの複数の断片サイズスコアを生成するステップを含む、システム。

【請求項86】

c) の前に、断片サイズ対照分子の第2のサブセットを添加し、それによって第2のスパイクイン試料を産生するステップをさらに含む、請求項85に記載のシステム。

【請求項87】

d) の前に、断片サイズ対照分子の第3のサブセットを添加し、それによって第3のスパ

10

20

30

40

50

イクイン試料を産生するステップをさらに含む、請求項 8 6 に記載のシステム。

【請求項 8 8】

前記方法が、e) の前に、断片サイズ対照分子の第 4 のサブセットを添加し、それによって第 4 のスパイクイン試料を産生するステップをさらに含む、請求項 8 7 に記載のシステム。

【請求項 8 9】

少なくとも 1 つの電子プロセッサによって実行された場合に方法を実施する非一時的コンピュータ実行可能命令を含むコンピュータ可読媒体を含むか、またはそれにアクセスすることが可能なコントローラを含むシステムであって、前記方法が：

a. 断片サイズ対照分子の第 1 のサブセットをポリヌクレオチドの試料中の核酸分子に添加し、それによって第 1 のスパイクイン試料を産生するステップ；

b. 前記第 1 のスパイクイン試料から核酸を抽出するステップ；

c. 断片サイズ対照分子の第 2 のサブセットを前記抽出された核酸に添加し、それによって第 2 のスパイクイン試料を産生するステップ；

d. 前記第 2 のスパイクイン試料の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、前記処理することが、前記第 2 のスパイクイン試料の前記少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；

e. 断片サイズ対照分子の第 3 のサブセットを前記処理された試料に添加し、それによって第 3 のスパイクイン試料を産生するステップ；

f. 前記第 3 のスパイクイン試料の少なくともサブセットを富化し、それによって富化された試料を産生するステップ；

g. 断片サイズ対照分子の第 4 のサブセットを前記富化された試料の少なくともサブセットに添加し、それによって第 4 のスパイクイン試料を産生するステップ；

h. 前記第 4 のスパイクイン試料をシーケンシングして、複数の配列リードを生成するステップ；ならびに

i. 前記複数の配列リードを分析して、断片サイズ対照分子の前記第 1 のサブセット、断片サイズ対照分子の前記第 2 のサブセット、断片サイズ対照分子の前記第 3 のサブセット、および／または断片サイズ対照分子の前記第 4 のサブセットの複数の断片サイズスコアを生成するステップ

を含む、システム。

【請求項 9 0】

前記方法が、前記複数の断片サイズスコアを複数の断片サイズ閾値と比較するステップをさらに含む、請求項 8 5 または 8 9 に記載のシステム。

【請求項 9 1】

前記方法が、前記複数の断片サイズスコアに基づいてポリヌクレオチドの前記試料中の核酸分子の前記分析を最適化するステップをさらに含む、請求項 9 0 に記載のシステム。

【請求項 9 2】

前記方法が、前記複数の断片サイズスコアを使用してポリヌクレオチドの前記試料中の核酸分子の前記分析における断片サイズバイアスを補正するステップをさらに含む、請求項 9 0 に記載のシステム。

【請求項 9 3】

前記方法が、前記方法を、(i) 前記複数の断片サイズスコアの各々が前記複数の断片サイズ閾値の対応する断片サイズ閾値内にある場合、成功である；または(ii) 前記複数の断片サイズスコアの少なくとも 1 つが前記複数の断片サイズ閾値の前記対応する断片サイズ閾値内でない場合、失敗であると分類するステップをさらに含む、請求項 9 0 に記載のシステム。

【請求項 9 4】

前記方法が、前記複数の断片サイズスコアの少なくとも 1 つを複数の汚染閾値の少なくとも 1 つと比較するステップをさらに含む、請求項 8 5 または 8 9 に記載のシステム。

【請求項 9 5】

前記試料を、(i) 前記複数の断片サイズスコアの少なくとも1つが前記複数の汚染閾値の対応する汚染閾値内にない場合、別の試料によって汚染されている；または(ii) 前記複数の断片サイズスコアの少なくとも1つが前記複数の汚染閾値の前記対応する汚染閾値内にある場合、別の試料によって汚染されていないと分類するステップをさらに含む、請求項 9 4 に記載のシステム。

【請求項 9 6】

断片サイズ対照分子の前記サブセットが、断片サイズ領域を含む複数の断片サイズ対照分子を含む、請求項 8 5 ～ 9 5 のいずれかに記載のシステム。

【請求項 9 7】

前記断片サイズ対照分子が、識別子領域をさらに含む、請求項 9 6 に記載のシステム。

【請求項 9 8】

断片サイズ対照分子の前記サブセットが、断片サイズ対照分子の少なくとも1つの群を含む、請求項 9 6 に記載のシステム。

【請求項 9 9】

群における前記複数の断片サイズ対照分子の前記断片サイズ領域が、同じ長さの領域である、請求項 9 8 に記載のシステム。

【請求項 1 0 0】

断片サイズ対照分子の第1の群における前記断片サイズ領域の長さが、断片サイズ対照分子の第2の群における前記断片サイズ領域の長さとは異なる、請求項 9 8 に記載のシステム。

【請求項 1 0 1】

前記識別子領域が、前記断片サイズ領域の片側または両側に存在する、請求項 9 7 に記載のシステム。

【請求項 1 0 2】

前記識別子領域が分子バーコードを含む、請求項 9 7 に記載のシステム。

【請求項 1 0 3】

前記断片サイズ対照分子が、1つまたは複数のプライマー結合部位を含む、請求項 8 5 ～ 1 0 2 のいずれかに記載のシステム。

【請求項 1 0 4】

前記プライマー結合部位が前記識別子領域に存在する、請求項 1 0 3 に記載のシステム。

【請求項 1 0 5】

群における前記断片サイズ対照分子の前記断片サイズ領域が、同じオリゴヌクレオチド配列を含む、請求項 8 5 ～ 1 0 4 のいずれかに記載のシステム。

【請求項 1 0 6】

群における前記断片サイズ対照分子の前記断片サイズ領域が、少なくとも2つの区別可能なオリゴヌクレオチド配列を含む、請求項 9 6 に記載のシステム。

【請求項 1 0 7】

断片サイズ対照分子の第1のサブセットにおける前記断片サイズ対照分子の前記断片サイズ領域が、断片サイズ対照分子の第2のサブセットにおける前記断片サイズ対照分子の前記断片サイズ領域のオリゴヌクレオチド配列とは区別可能なオリゴヌクレオチド配列を含む、請求項 9 6 に記載のシステム。

【請求項 1 0 8】

前記断片サイズ領域の長さが、少なくとも10 bp、少なくとも50 bp、少なくとも60 bp、少なくとも70 bp、少なくとも80 bp、少なくとも90 bp、少なくとも100 bp、少なくとも120 bp、少なくとも150 bp、少なくとも200 bp、少なくとも250 bp、少なくとも300 bp、少なくとも400 bp、少なくとも500 bp、少なくとも600 bp、少なくとも700 bp、少なくとも800 bp、少なくとも900 bp、または少なくとも1000 bpである、請求項 8 5 ～ 1 0 7 のいずれかに記載のシステム。

10

20

30

40

50

【請求項 109】

前記断片サイズ領域の長さが、10bp～1000bpの間である、請求項85～107のいずれかに記載のシステム。

【請求項 110】

前記断片サイズ対照分子の前記サブセットの各々が、等モル濃度で存在する、請求項85～109のいずれかに記載のシステム。

【請求項 111】

前記断片サイズ対照分子の前記サブセットの各々が、等しくないモル濃度で存在する、請求項85～109のいずれかに記載のシステム。

【請求項 112】

前記サブセットにおける前記断片サイズ対照分子の前記群の各々が、等しくないモル濃度で存在する、請求項85～111のいずれかに記載のシステム。

【請求項 113】

前記サブセットにおける前記断片サイズ対照分子の前記群の各々が、等モル濃度で存在する、請求項85～111のいずれかに記載のシステム。

【請求項 114】

前記分画することが、前記第2のスパイクイン試料の前記少なくともサブセットの前記核酸分子を、複数の分画化セットに分画することを含む、請求項85または89に記載のシステム。

【請求項 115】

前記複数の分画化セットが、前記第2のスパイクイン試料の前記核酸分子のエピジェネティック修飾レベルに基づいて分画された前記第2のスパイクイン試料の核酸分子を含む、請求項114に記載のシステム。

【請求項 116】

前記タグ付けすることが、タグのセットを前記核酸に付着させて、タグ付けされた核酸の集団を生成することであって、前記タグ付けされた核酸が1つまたは複数のタグを含む、ことを含む、請求項85または89に記載のシステム。

【請求項 117】

前記分画化に起因する複数の分画化セットの第1の分画化セットにおいて使用されるタグの前記セットが、前記複数の分画化セットの第2の分画化セットに使用されるタグの前記セットとは異なる、請求項116に記載のシステム。

【請求項 118】

タグの前記セットがアダプターの前記核酸へのライゲーションによって前記核酸に付着され、前記アダプターが1つまたは複数のタグを含む、請求項116に記載のシステム。

【請求項 119】

少なくとも1つの電子プロセッサによって実行された場合に方法を実施する非一時的コンピュータ実行可能命令を含むコンピュータ可読媒体を含むか、またはそれにアクセスすることが可能なコントローラを含むシステムであって、前記方法が：

a. 断片サイズ対照分子のサブセットをポリヌクレオチドの試料に添加し、それによって第1のスパイクイン試料を生成するステップ；

b. 前記第1のスパイクイン試料から核酸を抽出するステップ；

c. 前記抽出された核酸の少なくともサブセットを処理し、それによって処理された試料を生成するステップであって、前記処理することが、前記第1のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；ならびに

d. 前記処理された試料の少なくともサブセットを富化するステップを含む、システム。

【請求項 120】

前記方法が、c)の前に、断片サイズ対照分子の第2のサブセットを添加し、それによって第2のスパイクイン試料を生成するステップをさらに含む、請求項119に記載のシス

10

20

30

40

50

テム。

【請求項 1 2 1】

前記方法が、d)の前に、断片サイズ対照分子の第3のサブセットを添加し、それによって第3のスパイクイン試料を産生するステップをさらに含む、請求項 1 2 0 に記載のシステム。

【請求項 1 2 2】

前記方法が、e)断片サイズ対照分子の第4のサブセットを添加し、それによって第4のスパイクイン試料を産生するステップをさらに含む、請求項 1 2 1 に記載のシステム。

【請求項 1 2 3】

前記断片サイズ対照分子の濃度が、1アトモル～10ピコモルの間である、請求項 8 5 ～ 1 2 2 のいずれかに記載のシステム。

【請求項 1 2 4】

ポリヌクレオチドの前記試料が、無細胞ポリヌクレオチドの試料である、請求項 8 5 ～ 1 2 3 のいずれかに記載のシステム。

【請求項 1 2 5】

ポリヌクレオチドの前記試料が、無細胞DNAの試料および無細胞RNAの試料からなる群から選択される、請求項 1 2 3 に記載のシステム。

【請求項 1 2 6】

無細胞ポリヌクレオチドの前記試料が、無細胞DNAの試料である、請求項 1 2 3 に記載のシステム。

【請求項 1 2 7】

前記無細胞DNAが、1ng～500ngの間である、請求項 1 2 6 に記載のシステム。

【請求項 1 2 8】

必要に応じて、前記核酸分子の分析に関する情報および／または分析に由来する情報を含む報告書を作成するステップをさらに含む、または請求項 3 7 ～ 1 2 7 いずれか1つに記載のシステムまたは方法。

【請求項 1 2 9】

試料が由来する対象または医療従事者などの第3者に報告書を伝達するステップをさらに含む、請求項 1 2 8 に記載の方法またはシステム。

【請求項 1 3 0】

前記断片サイズ対照分子が合成分子である、上記請求項のいずれかに記載の方法またはシステム。

【請求項 1 3 1】

前記断片サイズ対照分子が、PCR増幅によってアンプリコンとして生成される、上記請求項のいずれかに記載の方法またはシステム。

【発明の詳細な説明】

【技術分野】

【0001】

相互参照

本出願は、その全体が参照により本明細書に援用される、2018年12月20日に提出された米国仮特許出願第62／783，046号に基づく利益および優先権を主張する。

【背景技術】

【0002】

背景

現行の無細胞核酸（例えば、無細胞DNAまたは無細胞RNA）のがん診断アッセイの方法は、単一ヌクレオチドバリエーション（SNV）、コピー数変異（CNV）、融合体、およびインデル（すなわち、挿入または欠失）を含む腫瘍関連体細胞バリエーションの検出に集中しており、これらは全てリキッドバイオロジーの主流の標的である。無細胞DNAにおけるサイズ分布および断片化パターンは、無細胞DNAの起源および疾患レベルに関する情報を提供することができるという証拠が増えつつある。無細胞DNAのサイズ分布および

10

20

30

40

50

断片化パターンを体細胞突然変異のコーリングと組み合わせると、いずれかのアプローチ単独から得られる評価よりも包括的な腫瘍の状況の評価を得ることができる。

【発明の概要】

【課題を解決するための手段】

【0003】

要旨

本開示は、断片サイズ分子を使用して核酸を分析するための方法、組成物、およびシステムを提供する。

【0004】

一態様では、本開示は、ポリヌクレオチドの試料中の核酸分子を分析するための方法であって、a)断片サイズ対照分子のサブセットを無細胞ポリヌクレオチドの試料中の核酸分子に添加し、それによって第1のスパイクイン試料を産生するステップ；b)第1のスパイクイン試料から核酸を抽出するステップ；c)抽出された核酸の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、処理することが、第1のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；d)処理された試料の少なくともサブセットを富化し、それによって富化された試料を産生するステップ；e)富化された試料の少なくともサブセットをシーケンシングして、複数の配列リードを生成するステップ；ならびにf)複数の配列リードを分析して、断片サイズ対照分子のサブセットの複数の断片サイズスコアを生成するステップを含む方法を提供する。一部の実施形態では、方法は、c)の前に、断片サイズ対照分子の第2のサブセットを添加し、それによって第2のスパイクイン試料を産生するステップをさらに含む。一部の実施形態では、方法は、d)の前に、断片サイズ対照分子の第3のサブセットを添加し、それによって第3のスパイクイン試料を産生するステップをさらに含む。一部の実施形態では、方法は、e)の前に、断片サイズ対照分子の第4のサブセットを添加し、それによって第4のスパイクイン試料を産生するステップを含む。

【0005】

一部の実施形態では、ポリヌクレオチドの試料は、無細胞ポリヌクレオチドの試料である。一部の実施形態では、ポリヌクレオチドの試料は、無細胞DNAの試料および無細胞RNAの試料からなる群から選択される。一部の実施形態では、無細胞ポリヌクレオチドの試料は、無細胞DNAである。一部の実施形態では、無細胞DNAは、1ng～500ngの間である。一部の実施形態では、断片サイズ対照分子の濃度は、1アtomol～10ピコmolの間である。

【0006】

別の態様では、本開示は、既定の断片サイズ対照分子の少なくとも1つのサブセットを含む断片サイズ対照分子のセットであって、既定の断片サイズ対照分子の少なくとも1つのサブセットが、断片サイズ領域を含む複数の断片サイズ対照分子を含む、断片サイズ対照分子のセットを提供する。一部の実施形態では、断片サイズ対照分子は、識別子領域をさらに含む。

【0007】

一部の実施形態では、少なくとも1つのサブセットは、断片サイズ対照分子の少なくとも1つの群を含む。

【0008】

一部の実施形態では、断片サイズ対照分子のサブセットは、断片サイズ領域を含む複数の断片サイズ対照分子を含む。一部の実施形態では、断片サイズ対照分子は、識別子領域をさらに含む。一部の実施形態では、サブセットは、断片サイズ対照分子の少なくとも1つの群を含む。

【0009】

一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、同じ長さの領域である。一部の実施形態では、断片サイズ対照分子の第1の群における断片サイズ領域

の長さは、断片サイズ対照分子の第2の群における断片サイズ領域の長さとは異なる。

【0010】

一部の実施形態では、識別子領域は、断片サイズ領域の片側または両側に存在する。一部の実施形態では、識別子領域は分子バーコードを含む。

【0011】

一部の実施形態では、複数の断片対照分子は、1つまたは複数のプライマー結合部位を含む。一部の実施形態では、1つまたは複数のプライマー結合部位は識別子領域に存在する。

【0012】

一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、同じオリゴヌクレオチド配列を含む。一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、少なくとも2つの区別可能なオリゴヌクレオチド配列を含む。一部の実施形態では、既定の断片サイズ対照分子の第1のサブセットにおける断片サイズ対照分子の断片サイズ領域は、既定の断片サイズ対照分子の第2のサブセットにおける断片サイズ対照分子の断片サイズ領域のオリゴヌクレオチド配列と区別可能なオリゴヌクレオチド配列を含む。

10

【0013】

一部の実施形態では、断片サイズ領域の長さは、少なくとも10bp、少なくとも50bp、少なくとも60bp、少なくとも70bp、少なくとも80bp、少なくとも90bp、少なくとも100bp、少なくとも120bp、少なくとも150bp、少なくとも200bp、少なくとも250bp、少なくとも300bp、少なくとも400bp、少なくとも500bp、少なくとも600bp、少なくとも700bp、少なくとも800bp、少なくとも900bp、または少なくとも1000bpである。一部の実施形態では、断片サイズ領域の長さは10bp～1000bpの間である。

20

【0014】

一部の実施形態では、既定の断片サイズ対照分子の少なくとも1つのサブセットの各サブセットは、等モル濃度で存在する。一部の実施形態では、既定の断片サイズ対照分子の少なくとも1つのサブセットの各サブセットは、等しくないモル濃度で存在する。一部の実施形態では、少なくとも1つのサブセットにおける断片サイズ対照分子の少なくとも1つの群の各群は、等モル濃度で存在する。一部の実施形態では、少なくとも1つのサブセットにおける断片サイズ対照分子の少なくとも1つの群の各群は、等しくないモル濃度で存在する。

30

【0015】

一部の実施形態では、断片サイズ対照分子のサブセットの各々は、等モル濃度で存在する。一部の実施形態では、断片サイズ対照分子のサブセットの各々は、等しくないモル濃度で存在する。一部の実施形態では、サブセットにおける断片サイズ対照分子の群の各々は、等モル濃度で存在する。一部の実施形態では、サブセットにおける断片サイズ対照分子の群の各々は、等しくないモル濃度で存在する。

【0016】

別の態様では、本開示は、無細胞ポリヌクレオチドの試料中の核酸分子の分析における断片サイズバイアスを評価するための方法であって：a)断片サイズ対照分子の第1のサブセットをポリヌクレオチドの試料中の核酸分子に添加し、それによって第1のスパイクイン試料を産生するステップ；b)第1のスパイクイン試料から核酸を抽出するステップ；c)断片サイズ対照分子の第2のサブセットを抽出された核酸に添加し、それによって第2のスパイクイン試料を産生するステップ；d)第2のスパイクイン試料の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、処理することが、第2のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；e)断片サイズ対照分子の第3のサブセットを処理された試料に添加し、それによって第3のスパイクイン試料を産生するステップ；f)第3のスパイクイン試料の少なくともサブセットを富化し、それによって富化された試料を

40

50

產生するステップ； g) 断片サイズ対照分子の第 4 のサブセットを富化された試料の少なくともサブセットに添加し、それによって第 4 のスパイクイン試料を產生するステップ； h) 第 4 のスパイクイン試料をシーケンシングして、複数の配列リードを生成するステップ； ならびに i) 複数の配列リードを分析して、断片サイズ対照分子の第 1 のサブセット、断片サイズ対照分子の第 2 のサブセット、断片サイズ対照分子の第 3 のサブセット、および／または断片サイズ対照分子の第 4 のサブセットの複数の断片サイズスコアを生成するステップを含む方法を提供する。一部の実施形態では、方法は、複数の断片サイズスコアを複数の断片サイズ閾値と比較するステップをさらに含む。一部の実施形態では、方法は、複数の断片サイズスコアに基づいてポリヌクレオチドの試料中の核酸分子の分析を最適化するステップをさらに含む。一部の実施形態では、方法は、複数の断片サイズスコアを使用して無細胞ポリヌクレオチドの試料中の核酸分子の分析における断片サイズバイアスを補正するステップをさらに含む。一部の実施形態では、方法は、方法を、(i) 複数の断片サイズスコアの少なくとも 1 つが複数の断片サイズ閾値の対応する断片サイズ閾値内にある場合、成功である；または (ii) 複数の断片サイズスコアの少なくとも 1 つが複数の断片サイズ閾値の対応する断片サイズ閾値内でない場合、失敗であると分類するステップをさらに含む。

10

【0017】

別の態様では、本開示は、第 1 の試料の第 2 の試料による汚染を検出する方法であって、第 1 の試料および第 2 の試料の各々に関して： (a) 断片サイズ対照分子のサブセットを添加して、第 1 のスパイクイン試料を生成するステップであって、第 1 の試料に添加された断片サイズ対照分子のサブセットが、第 2 の試料に添加された断片サイズ対照分子のサブセットと区別され得る、ステップ； (b) 第 1 のスパイクインから核酸を抽出するステップ； (c) 抽出された核酸の少なくともサブセットを処理し、それによって処理された試料を產生するステップであって、処理することが、第 1 のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ； (d) 処理された試料の少なくともサブセットを富化するステップ； (e) 富化された試料の少なくともサブセットをシーケンシングして、複数の配列リードを生成するステップ； ならびに (f) 複数の配列リードを分析して、断片サイズ対照分子のサブセットの 1 つまたは複数の汚染スコアを生成するステップを含む方法を提供する。一部の実施形態では、方法は、c) の前に、断片サイズ対照分子の第 2 のサブセットを添加し、それによって第 2 のスパイクイン試料を產生するステップであって、第 1 の試料に添加された断片サイズ対照分子のサブセットが、第 2 の試料に添加された断片サイズ対照分子のサブセットと区別され得る、ステップをさらに含む。一部の実施形態では、方法は、d) の前に、断片サイズ対照分子の第 3 のサブセットを添加し、それによって第 3 のスパイクイン試料を產生するステップであって、第 1 の試料に添加された断片サイズ対照分子のサブセットが、第 2 の試料に添加された断片サイズ対照分子のサブセットと区別され得る、ステップをさらに含む。一部の実施形態では、方法は、e) の前に、断片サイズ対照分子の第 4 のサブセットを添加し、それによって第 4 のスパイクイン試料を產生するステップであって、第 1 の試料に添加された断片サイズ対照分子のサブセットが、第 2 の試料に添加された断片サイズ対照分子のサブセットと区別され得る、ステップをさらに含む。

20

30

40

【0018】

別の態様では、本開示は、第 1 の試料の第 2 の試料による汚染を検出する方法であって、第 1 の試料および第 2 の試料の各々に関して： (a) 断片サイズ対照分子の第 1 のサブセットを添加して、第 1 のスパイクイン試料を生成するステップであって、第 1 の試料に添加された断片サイズ対照分子の第 1 のサブセットが、第 2 の試料に添加された断片サイズ対照分子の第 1 のサブセットと区別され得る、ステップ； (b) 第 1 のスパイクインから核酸を抽出するステップ； (c) 断片サイズ対照分子の第 2 のサブセットを抽出された核酸に添加し、それによって第 2 のスパイクイン試料を產生するステップであって、第 1 の試料に添加された断片サイズ対照分子の第 2 のサブセットが、第 2 の試料に添加された断片サイズ対照分子の第 2 のサブセットと区別され得る、ステップ； (d) 抽出された核酸

50

の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、処理することが、第1のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；（e）断片サイズ対照分子の第3のサブセットを抽出された核酸に添加し、それによって第3のスパイクイン試料を産生するステップであって、第1の試料に添加された断片サイズ対照分子の第3のサブセットが、第2の試料に添加された断片サイズ対照分子の第3のサブセットと区別され得る、ステップ；（f）処理された試料の少なくともサブセットを富化するステップ；（g）断片サイズ対照分子の第4のサブセットを抽出された核酸分子に添加し、それによって第4のスパイクイン試料を産生するステップであって、第1の試料に添加された断片サイズ対照分子の第4のサブセットが、第2の試料に添加された断片サイズ対照分子の第4のサブセットと区別され得る、ステップ；（h）富化された試料の少なくともサブセットをシーケンシングして、複数の配列リードを生成するステップ；ならびに（i）複数の配列リードを分析して、断片サイズ対照分子のサブセットの1つまたは複数の汚染スコアを生成するステップを含む方法を提供する。

10

【0019】

一部の実施形態では、方法は、少なくとも1つまたは複数の汚染スコアを、少なくとも1つまたは複数の汚染閾値と比較するステップをさらに含む。一部の実施形態では、方法は、第1の試料を、（i）汚染スコアの少なくとも1つまたは複数の1つまたは複数の汚染閾値の対応する汚染閾値内にない場合、第2の試料によって汚染されている；または（ii）汚染スコアの少なくとも1つまたは複数の1つまたは複数の汚染閾値の対応する汚染閾値内にある場合、第2の試料によって汚染されていないと分類するステップをさらに含む。

20

【0020】

一部の実施形態では、分画することは、第2のスパイクイン試料の少なくともサブセットの核酸分子を複数の分画化セットに分画することを含む。一部の実施形態では、複数の分画化セットは、第2のスパイクイン試料の核酸分子のエピジェネティック修飾のレベルに基づいて分画された第2のスパイクイン試料の核酸分子を含む。

【0021】

一部の実施形態では、タグ付けすることは、タグのセットを核酸に付着させて、タグ付けされた核酸の集団を産生することであって、タグ付けされた核酸が1つまたは複数のタグを含む、ことを含む。一部の実施形態では、分画化に起因する複数の分画化セットの第1の分画化セットにおいて使用されるタグのセットは、複数の分画化セットの第2の分画化セットにおいて使用されるタグのセットとは異なる。一部の実施形態では、タグのセットは、アダプターの核酸へのライゲーションによって核酸に付着され、アダプターは1つまたは複数のタグを含む。

30

【0022】

一部の実施形態では、ポリヌクレオチドの試料は、無細胞ポリヌクレオチドの試料である。一部の実施形態では、ポリヌクレオチドの試料は、無細胞DNAの試料および無細胞RNAの試料からなる群から選択される。一部の実施形態では、無細胞ポリヌクレオチドの試料は、無細胞DNAである。一部の実施形態では、無細胞DNAは、1 ng～500 ngの間である。一部の実施形態では、断片サイズ対照分子の濃度は、1 アトモル～10 ピコモルの間である。

40

【0023】

別の態様では、本開示は、既定の断片サイズ対照分子の少なくとも1つのサブセットを含む断片サイズ対照分子のセットであって、既定の断片サイズ対照分子の少なくとも1つのサブセットが、断片サイズ領域を含む複数の断片サイズ対照分子を含む、断片サイズ対照分子のセットを提供する。一部の実施形態では、断片サイズ対照分子は、識別子領域をさらに含む。

【0024】

一部の実施形態では、少なくとも1つのサブセットは、断片サイズ対照分子の少なくとも

50

1つの群を含む。

【0025】

一部の実施形態では、断片サイズ対照分子のサブセットは、断片サイズ領域を含む複数の断片サイズ対照分子を含む。一部の実施形態では、断片サイズ対照分子は、識別子領域をさらに含む。一部の実施形態では、サブセットは、断片サイズ対照分子の少なくとも1つの群を含む。

【0026】

一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、同じ長さの領域である。一部の実施形態では、断片サイズ対照分子の第1の群における断片サイズ領域の長さは、断片サイズ対照分子の第2の群における断片サイズ領域の長さとは異なる。

10

【0027】

一部の実施形態では、識別子領域は、断片サイズ領域の片側または両側に存在する。一部の実施形態では、識別子領域は分子バーコードを含む。

【0028】

一部の実施形態では、複数の断片対照分子は、1つまたは複数のプライマー結合部位を含む。一部の実施形態では、1つまたは複数のプライマー結合部位は識別子領域に存在する。

【0029】

一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、同じオリゴヌクレオチド配列を含む。一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、少なくとも2つの区別可能なオリゴヌクレオチド配列を含む。一部の実施形態では、既定の断片サイズ対照分子の第1のサブセットにおける断片サイズ対照分子の断片サイズ領域は、既定の断片サイズ対照分子の第2のサブセットにおける断片サイズ対照分子の断片サイズ領域のオリゴヌクレオチド配列とは区別可能なオリゴヌクレオチド配列を含む。

20

【0030】

一部の実施形態では、断片サイズ領域の長さは、少なくとも10bp、少なくとも50bp、少なくとも60bp、少なくとも70bp、少なくとも80bp、少なくとも90bp、少なくとも100bp、少なくとも120bp、少なくとも150bp、少なくとも200bp、少なくとも250bp、少なくとも300bp、少なくとも400bp、少なくとも500bp、少なくとも600bp、少なくとも700bp、少なくとも800bp、少なくとも900bp、または少なくとも1000bpである。一部の実施形態では、断片サイズ領域の長さは10bp～1000bpの間である。

30

【0031】

一部の実施形態では、既定の断片サイズ対照分子の少なくとも1つのサブセットの各サブセットは、等モル濃度で存在する。一部の実施形態では、既定の断片サイズ対照分子の少なくとも1つのサブセットの各サブセットは、等しくないモル濃度で存在する。一部の実施形態では、少なくとも1つのサブセットにおける断片サイズ対照分子の少なくとも1つの群の各群は、等モル濃度で存在する。一部の実施形態では、少なくとも1つのサブセットにおける断片サイズ対照分子の少なくとも1つの群の各群は、等しくないモル濃度で存在する。

40

【0032】

一部の実施形態では、断片サイズ対照分子のサブセットの各々は、等モル濃度で存在する。一部の実施形態では、断片サイズ対照分子のサブセットの各々は、等しくないモル濃度で存在する。一部の実施形態では、サブセットにおける断片サイズ対照分子の群の各々は、等モル濃度で存在する。一部の実施形態では、サブセットにおける断片サイズ対照分子の群の各々は、等しくないモル濃度で存在する。

【0033】

別の態様では、本開示は、無細胞ポリヌクレオチドの試料のシーケンシングライブラリを産生するための方法であって：a)断片サイズ対照分子のサブセットを試料に添加し、そ

50

れによって第1のスパイクイン試料を産生するステップ；b) 第1のスパイクイン試料から核酸を抽出するステップ；c) 抽出された核酸の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、処理することが、第1のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；ならびにd) 処理された試料の少なくともサブセットを富化するステップを含む方法を提供する。一部の実施形態では、方法は、c)の前に、断片サイズ対照分子の第2のサブセットを添加し、それによって第2のスパイクイン試料を産生するステップをさらに含む。一部の実施形態では、方法は、d)の前に、断片サイズ対照分子の第3のサブセットを添加し、それによって第3のスパイクイン試料を産生するステップをさらに含む。一部の実施形態では、方法は、e)断片サイズ対照分子の第4のサブセットを添加し、それによって第4のスパイクイン試料を産生するステップをさらに含む。

10

【0034】

一部の実施形態では、少なくとも1つのサブセットは、断片サイズ対照分子の少なくとも1つの群を含む。

【0035】

一部の実施形態では、断片サイズ対照分子のサブセットは、断片サイズ領域を含む複数の断片サイズ対照分子を含む。一部の実施形態では、断片サイズ対照分子は、識別子領域をさらに含む。一部の実施形態では、サブセットは、断片サイズ対照分子の少なくとも1つの群を含む。

【0036】

一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、同じ長さの領域である。一部の実施形態では、断片サイズ対照分子の第1の群における断片サイズ領域の長さは、断片サイズ対照分子の第2の群における断片サイズ領域の長さとは異なる。

20

【0037】

一部の実施形態では、識別子領域は、断片サイズ領域の片側または両側に存在する。一部の実施形態では、識別子領域は分子バーコードを含む。

【0038】

一部の実施形態では、複数の断片対照分子は、1つまたは複数のプライマー結合部位を含む。一部の実施形態では、1つまたは複数のプライマー結合部位は識別子領域に存在する。

30

【0039】

一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、同じオリゴヌクレオチド配列を含む。一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、少なくとも2つの区別可能なオリゴヌクレオチド配列を含む。一部の実施形態では、既定の断片サイズ対照分子の第1のサブセットにおける断片サイズ対照分子の断片サイズ領域は、既定の断片サイズ対照分子の第2のサブセットにおける断片サイズ対照分子の断片サイズ領域のオリゴヌクレオチド配列とは区別可能なオリゴヌクレオチド配列を含む。

【0040】

一部の実施形態では、断片サイズ領域の長さは、少なくとも10bp、少なくとも50bp、少なくとも60bp、少なくとも70bp、少なくとも80bp、少なくとも90bp、少なくとも100bp、少なくとも120bp、少なくとも150bp、少なくとも200bp、少なくとも250bp、少なくとも300bp、少なくとも400bp、または少なくとも500bpである。一部の実施形態では、断片サイズ領域の長さは10bp～1000bpの間である。

40

【0041】

一部の実施形態では、既定の断片サイズ対照分子の少なくとも1つのサブセットの各サブセットは、等モル濃度で存在する。一部の実施形態では、既定の断片サイズ対照分子の少なくとも1つのサブセットの各サブセットは、等しくないモル濃度で存在する。一部の実施形態では、少なくとも1つのサブセットにおける断片サイズ対照分子の少なくとも1つ

50

の群の各群は、等モル濃度で存在する。一部の実施形態では、少なくとも1つのサブセットにおける断片サイズ対照分子の少なくとも1つの群の各群は、等しくないモル濃度で存在する。

【0042】

なお別の態様では、本開示は、(a) 既定の断片サイズ対照分子の少なくとも1つのサブセットを含む断片サイズ対照分子のセットであって、既定の断片サイズ対照分子の少なくとも1つのサブセットが、断片サイズ領域を含む複数の断片サイズ対照分子を含む、断片サイズ対照分子のセット；および(b) 対象からのポリヌクレオチドの試料中の核酸分子のセットを含む、核酸の集団を提供する。一部の実施形態では、少なくとも1つのサブセットは、断片サイズ対照分子の少なくとも1つの群を含む。一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、同じ長さの領域である。一部の実施形態では、断片サイズ対照分子の第1の群における断片サイズ領域の長さは、断片サイズ対照分子の第2の群における断片サイズ領域の長さとは異なる。一部の実施形態では、断片サイズ対照分子は、識別子領域をさらに含む。一部の実施形態では、識別子領域は、断片サイズ領域の片側または両側に存在する。一部の実施形態では、識別子領域は分子バーコードを含む。一部の実施形態では、複数の断片対照分子は、1つまたは複数のプライマー結合部位を含む。一部の実施形態では、1つまたは複数のプライマー結合部位は識別子領域に存在する。一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、同じオリゴヌクレオチド配列を含む。一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、少なくとも2つの区別可能なオリゴヌクレオチド配列を含む。一部の実施形態では、既定の断片サイズ対照分子の第1のサブセットにおける断片サイズ対照分子の断片サイズ領域は、既定の断片サイズ対照分子の第2のサブセットにおける断片サイズ対照分子の断片サイズ領域のオリゴヌクレオチド配列とは区別可能なオリゴヌクレオチド配列を含む。一部の実施形態では、断片サイズ領域の長さは、少なくとも10bp、少なくとも50bp、少なくとも60bp、少なくとも70bp、少なくとも80bp、少なくとも90bp、少なくとも100bp、少なくとも120bp、少なくとも150bp、少なくとも200bp、少なくとも250bp、少なくとも300bp、少なくとも400bp、少なくとも500bp、少なくとも600bp、少なくとも700bp、少なくとも800bp、少なくとも900bp、または少なくとも1000bpである。一部の実施形態では、断片サイズ領域の長さは10bp~1000bpの間である。一部の実施形態では、既定の断片サイズ対照分子の少なくとも1つのサブセットの各サブセットは、等モル濃度で存在する。一部の実施形態では、少なくとも1つのサブセットにおける断片サイズ対照分子の少なくとも1つの群の各群は、等モル濃度で存在する。一部の実施形態では、少なくとも1つのサブセットにおける断片サイズ対照分子の少なくとも1つの群の各群は、等しくないモル濃度で存在する。

【0043】

別の態様では、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た場合に方法を実施する非一時的コンピュータ実行可能命令を含むコンピュータ可読媒体を含むか、またはそれにアクセスすることが可能なコントローラを含むシステムであって、方法が：a) 断片サイズ対照分子のサブセットをポリヌクレオチドの試料中の核酸分子に添加し、それによって第1のスパイクイン試料を産生するステップ；b) 第1のスパイクイン試料から核酸を抽出するステップ；c) 抽出された核酸の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、処理することが、第1のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；d) 処理された試料の少なくともサブセットを富化し、それによって富化された試料を産生するステップ；e) 富化された試料の少なくともサブセットをシーケンシングして、複数の配列リードを生成するステップ；ならびにf) 複数の配列リードを分析して、断片サイズ対照分子のサブセットの複数の断片サイズスコアを生成するステップを含む、システムを提供する。一部の実施形態では、方法は、c)の前に、断片サイズ対照分子の第2のサブセットを添加し、それによって第2のスパイクイン試料を産生する

ステップをさらに含む。一部の実施形態では、方法は、d)の前に、断片サイズ対照分子の第3のサブセットを添加し、それによって第3のスパイクイン試料を産生するステップをさらに含む。一部の実施形態では、方法は、e)の前に、断片サイズ対照分子の第4のサブセットを添加し、それによって第4のスパイクイン試料を産生するステップをさらに含む。

【0044】

一部の実施形態では、ポリヌクレオチドの試料は、無細胞ポリヌクレオチドの試料である。一部の実施形態では、ポリヌクレオチドの試料は、無細胞DNAの試料および無細胞RNAの試料からなる群から選択される。一部の実施形態では、無細胞ポリヌクレオチドの試料は、無細胞DNAである。一部の実施形態では、無細胞DNAは、1 ng ~ 500 ngの間である。一部の実施形態では、断片サイズ対照分子の濃度は、1アtomol ~ 10ピコmolの間である。

10

【0045】

別の態様では、本開示は、既定の断片サイズ対照分子の少なくとも1つのサブセットを含む断片サイズ対照分子のセットであって、既定の断片サイズ対照分子の少なくとも1つのサブセットが、断片サイズ領域を含む複数の断片サイズ対照分子を含む、断片サイズ対照分子のセットを提供する。一部の実施形態では、断片サイズ対照分子は、識別子領域をさらに含む。

【0046】

一部の実施形態では、少なくとも1つのサブセットは、断片サイズ対照分子の少なくとも1つの群を含む。

20

【0047】

一部の実施形態では、断片サイズ対照分子のサブセットは、断片サイズ領域を含む複数の断片サイズ対照分子を含む。一部の実施形態では、断片サイズ対照分子は、識別子領域をさらに含む。一部の実施形態では、サブセットは、断片サイズ対照分子の少なくとも1つの群を含む。

【0048】

一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、同じ長さの領域である。一部の実施形態では、断片サイズ対照分子の第1の群における断片サイズ領域の長さは、断片サイズ対照分子の第2の群における断片サイズ領域の長さとは異なる。

30

【0049】

一部の実施形態では、識別子領域は、断片サイズ領域の片側または両側に存在する。一部の実施形態では、識別子領域は分子バーコードを含む。

【0050】

一部の実施形態では、複数の断片対照分子は、1つまたは複数のプライマー結合部位を含む。一部の実施形態では、1つまたは複数のプライマー結合部位は識別子領域に存在する。

【0051】

一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、同じオリゴヌクレオチド配列を含む。一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、少なくとも2つの区別可能なオリゴヌクレオチド配列を含む。一部の実施形態では、既定の断片サイズ対照分子の第1のサブセットにおける断片サイズ対照分子の断片サイズ領域は、既定の断片サイズ対照分子の第2のサブセットにおける断片サイズ対照分子の断片サイズ領域のオリゴヌクレオチド配列と区別可能なオリゴヌクレオチド配列を含む。

40

【0052】

一部の実施形態では、断片サイズ領域の長さは、少なくとも10 bp、少なくとも50 bp、少なくとも60 bp、少なくとも70 bp、少なくとも80 bp、少なくとも90 bp、少なくとも100 bp、少なくとも120 bp、少なくとも150 bp、少なくとも200 bp、少なくとも250 bp、少なくとも300 bp、少なくとも400 bp、少

50

なくとも500bp、少なくとも600bp、少なくとも700bp、少なくとも800bp、少なくとも900bp、または少なくとも1000bpである。一部の実施形態では、断片サイズ領域の長さは10bp~1000bpの間である。

【0053】

一部の実施形態では、既定の断片サイズ対照分子の少なくとも1つのサブセットの各サブセットは、等モル濃度で存在する。一部の実施形態では、既定の断片サイズ対照分子の少なくとも1つのサブセットの各サブセットは、等しくないモル濃度で存在する。一部の実施形態では、少なくとも1つのサブセットにおける断片サイズ対照分子の少なくとも1つの群の各群は、等モル濃度で存在する。一部の実施形態では、少なくとも1つのサブセットにおける断片サイズ対照分子の少なくとも1つの群の各群は、等しくないモル濃度で存在する。

10

【0054】

一部の実施形態では、断片サイズ対照分子のサブセットの各々は、等モル濃度で存在する。一部の実施形態では、断片サイズ対照分子のサブセットの各々は、等しくないモル濃度で存在する。一部の実施形態では、サブセットにおける断片サイズ対照分子の群の各々は、等モル濃度で存在する。一部の実施形態では、サブセットにおける断片サイズ対照分子の群の各々は、等しくないモル濃度で存在する。

【0055】

別の態様では、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た場合に方法を実施する非一時的コンピュータ実行可能命令を含むコンピュータ可読媒体を含むか、またはそれにアクセスすることが可能なコントローラを含むシステムであって、方法が：a)断片サイズ対照分子の第1のサブセットをポリヌクレオチドの試料中の核酸分子に添加し、それによって第1のスパイクイン試料を産生するステップ；b)第1のスパイクイン試料から核酸を抽出するステップ；c)断片サイズ対照分子の第2のサブセットを抽出された核酸に添加し、それによって第2のスパイクイン試料を産生するステップ；d)第2のスパイクイン試料の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、処理することが、第2のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；e)断片サイズ対照分子の第3のサブセットを処理された試料に添加し、それによって第3のスパイクイン試料を産生するステップ；f)第3のスパイクイン試料の少なくともサブセットを富化し、それによって富化された試料を産生するステップ；g)断片サイズ対照分子の第4のサブセットを富化された試料の少なくともサブセットに添加し、それによって第4のスパイクイン試料を産生するステップ；h)第4のスパイクイン試料をシーケンシングして、複数の配列リードを生成するステップ；ならびにi)複数の配列リードを分析して、断片サイズ対照分子の第1のサブセット、断片サイズ対照分子の第2のサブセット、断片サイズ対照分子の第3のサブセット、および／または断片サイズ対照分子の第4のサブセットの複数の断片サイズスコアを生成するステップを含む、システムを提供する。一部の実施形態では、方法は、複数の断片サイズスコアを複数の断片サイズ閾値と比較するステップをさらに含む。一部の実施形態では、方法は、複数の断片サイズスコアに基づいてポリヌクレオチドの試料中の核酸分子の分析を最適化するステップをさらに含む。一部の実施形態では、方法は、複数の断片サイズスコアを使用してポリヌクレオチドの試料中の核酸分子の分析における断片サイズバイアスを補正するステップをさらに含む。一部の実施形態では、方法は、方法を、(i)複数の断片サイズスコアの各々が複数の断片サイズ閾値の対応する断片サイズ閾値内にある場合、成功である；または(ii)複数の断片サイズスコアの少なくとも1つが複数の断片サイズ閾値の対応する断片サイズ閾値内にない場合、失敗であると分類するステップをさらに含む。一部の実施形態では、方法は、複数の断片サイズスコアの少なくとも1つを、複数の汚染閾値の少なくとも1つと比較するステップをさらに含む。一部の実施形態では、方法は、試料を、(i)複数の断片サイズスコアの少なくとも1つが複数の汚染閾値の対応する汚染閾値内にない場合、別の試料によって汚染されている；または(ii)複数の断片サイズスコアの少なくとも1つが複数の汚

20

30

40

50

染閾値の対応する汚染閾値内にある場合、別の試料によって汚染されていないと分類するステップをさらに含む。

【0056】

一部の実施形態では、ポリヌクレオチドの試料は、無細胞ポリヌクレオチドの試料である。一部の実施形態では、ポリヌクレオチドの試料は、無細胞DNAの試料および無細胞RNAの試料からなる群から選択される。一部の実施形態では、無細胞ポリヌクレオチドの試料は、無細胞DNAである。一部の実施形態では、無細胞DNAは、1 ng ~ 500 ngの間である。一部の実施形態では、断片サイズ対照分子の濃度は、1 アtomol ~ 10 ピコmolの間である。

【0057】

別の態様では、本開示は、既定の断片サイズ対照分子の少なくとも1つのサブセットを含む断片サイズ対照分子のセットであって、既定の断片サイズ対照分子の少なくとも1つのサブセットが、断片サイズ領域を含む複数の断片サイズ対照分子を含む、断片サイズ対照分子のセットを提供する。一部の実施形態では、断片サイズ対照分子は、識別子領域をさらに含む。

【0058】

一部の実施形態では、少なくとも1つのサブセットは、断片サイズ対照分子の少なくとも1つの群を含む。

【0059】

一部の実施形態では、断片サイズ対照分子のサブセットは、断片サイズ領域を含む複数の断片サイズ対照分子を含む。一部の実施形態では、断片サイズ対照分子は、識別子領域をさらに含む。一部の実施形態では、サブセットは、断片サイズ対照分子の少なくとも1つの群を含む。

【0060】

一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、同じ長さの領域である。一部の実施形態では、断片サイズ対照分子の第1の群における断片サイズ領域の長さは、断片サイズ対照分子の第2の群における断片サイズ領域の長さとは異なる。

【0061】

一部の実施形態では、識別子領域は、断片サイズ領域の片側または両側に存在する。一部の実施形態では、識別子領域は分子バーコードを含む。

【0062】

一部の実施形態では、複数の断片対照分子は、1つまたは複数のプライマー結合部位を含む。一部の実施形態では、1つまたは複数のプライマー結合部位は識別子領域に存在する。

【0063】

一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、同じオリゴヌクレオチド配列を含む。一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、少なくとも2つの区別可能なオリゴヌクレオチド配列を含む。一部の実施形態では、既定の断片サイズ対照分子の第1のサブセットにおける断片サイズ対照分子の断片サイズ領域は、既定の断片サイズ対照分子の第2のサブセットにおける断片サイズ対照分子の断片サイズ領域のオリゴヌクレオチド配列とは区別可能なオリゴヌクレオチド配列を含む。

【0064】

一部の実施形態では、断片サイズ領域の長さは、少なくとも10 bp、少なくとも50 bp、少なくとも60 bp、少なくとも70 bp、少なくとも80 bp、少なくとも90 bp、少なくとも100 bp、少なくとも120 bp、少なくとも150 bp、少なくとも200 bp、少なくとも250 bp、少なくとも300 bp、少なくとも400 bp、少なくとも500 bp、少なくとも600 bp、少なくとも700 bp、少なくとも800 bp、少なくとも900 bp、または少なくとも1000 bpである。一部の実施形態では、断片サイズ領域の長さは10 bp ~ 1000 bpの間である。

10

20

30

40

50

【0065】

一部の実施形態では、既定の断片サイズ対照分子の少なくとも1つのサブセットの各サブセットは、等モル濃度で存在する。一部の実施形態では、既定の断片サイズ対照分子の少なくとも1つのサブセットの各サブセットは、等しくないモル濃度で存在する。一部の実施形態では、少なくとも1つのサブセットにおける断片サイズ対照分子の少なくとも1つの群の各群は、等モル濃度で存在する。一部の実施形態では、少なくとも1つのサブセットにおける断片サイズ対照分子の少なくとも1つの群の各群は、等しくないモル濃度で存在する。

【0066】

一部の実施形態では、断片サイズ対照分子のサブセットの各々は、等モル濃度で存在する。一部の実施形態では、断片サイズ対照分子のサブセットの各々は、等しくないモル濃度で存在する。一部の実施形態では、サブセットにおける断片サイズ対照分子の群の各々は、等モル濃度で存在する。一部の実施形態では、サブセットにおける断片サイズ対照分子の群の各々は、等しくないモル濃度で存在する。

【0067】

別の態様では、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た場合に方法を実施する非一時的コンピュータ実行可能命令を含むコンピュータ可読媒体を含むか、またはそれにアクセスすることが可能なコントローラを含むシステムであって、方法が：a)断片サイズ対照分子のサブセットを試料に添加し、それによって第1のスパイクイン試料を産生するステップ；b)第1のスパイクイン試料から核酸を抽出するステップ；c)抽出された核酸の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、処理することが、第1のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；ならびにd)処理された試料の少なくともサブセットを富化するステップを含む、システムを提供する。一部の実施形態では、方法は、c)の前に、断片サイズ対照分子の第2のサブセットを添加し、それによって第2のスパイクイン試料を産生するステップをさらに含む。一部の実施形態では、方法は、d)の前に、断片サイズ対照分子の第3のサブセットを添加し、それによって第3のスパイクイン試料を産生するステップをさらに含む。一部の実施形態では、方法は、e)断片サイズ対照分子の第4のサブセットを添加し、それによって第4のスパイクイン試料を産生するステップをさらに含む。

【0068】

一部の実施形態では、ポリヌクレオチドの試料は、無細胞ポリヌクレオチドの試料である。一部の実施形態では、ポリヌクレオチドの試料は、無細胞DNAの試料および無細胞RNAの試料からなる群から選択される。一部の実施形態では、無細胞ポリヌクレオチドの試料は、無細胞DNAである。一部の実施形態では、無細胞DNAは、1 ng～500 ngの間である。一部の実施形態では、断片サイズ対照分子の濃度は、1アトモル～10ピコモルの間である。

【0069】

別の態様では、本開示は、既定の断片サイズ対照分子の少なくとも1つのサブセットを含む断片サイズ対照分子のセットであって、既定の断片サイズ対照分子の少なくとも1つのサブセットが、断片サイズ領域を含む複数の断片サイズ対照分子を含む、断片サイズ対照分子のセットを提供する。一部の実施形態では、断片サイズ対照分子は、識別子領域をさらに含む。

【0070】

一部の実施形態では、少なくとも1つのサブセットは、断片サイズ対照分子の少なくとも1つの群を含む。

【0071】

一部の実施形態では、断片サイズ対照分子のサブセットは、断片サイズ領域を含む複数の断片サイズ対照分子を含む。一部の実施形態では、断片サイズ対照分子は、識別子領域をさらに含む。一部の実施形態では、サブセットは、断片サイズ対照分子の少なくとも1つ

の群を含む。

【0072】

一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、同じ長さの領域である。一部の実施形態では、断片サイズ対照分子の第1の群における断片サイズ領域の長さは、断片サイズ対照分子の第2の群における断片サイズ領域の長さとは異なる。

【0073】

一部の実施形態では、識別子領域は、断片サイズ領域の片側または両側に存在する。一部の実施形態では、識別子領域は分子バーコードを含む。

【0074】

一部の実施形態では、複数の断片対照分子は、1つまたは複数のプライマー結合部位を含む。一部の実施形態では、1つまたは複数のプライマー結合部位は識別子領域に存在する。

10

【0075】

一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、同じオリゴヌクレオチド配列を含む。一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、少なくとも2つの区別可能なオリゴヌクレオチド配列を含む。一部の実施形態では、既定の断片サイズ対照分子の第1のサブセットにおける断片サイズ対照分子の断片サイズ領域は、既定の断片サイズ対照分子の第2のサブセットにおける断片サイズ対照分子の断片サイズ領域のオリゴヌクレオチド配列とは区別可能なオリゴヌクレオチド配列を含む。

20

【0076】

一部の実施形態では、断片サイズ領域の長さは、少なくとも10bp、少なくとも50bp、少なくとも60bp、少なくとも70bp、少なくとも80bp、少なくとも90bp、少なくとも100bp、少なくとも120bp、少なくとも150bp、少なくとも200bp、少なくとも250bp、少なくとも300bp、少なくとも400bp、少なくとも500bp、少なくとも600bp、少なくとも700bp、少なくとも800bp、少なくとも900bp、または少なくとも1000bpである。一部の実施形態では、断片サイズ領域の長さは10bp～1000bpの間である。

【0077】

一部の実施形態では、断片サイズ対照分子は、合成分子である。一部の実施形態では、断片サイズ対照分子は、PCR増幅によってアンプリコンとして生成される。

30

【0078】

一部の実施形態では、既定の断片サイズ対照分子の少なくとも1つのサブセットの各サブセットは、等モル濃度で存在する。一部の実施形態では、既定の断片サイズ対照分子の少なくとも1つのサブセットの各サブセットは、等しくないモル濃度で存在する。一部の実施形態では、少なくとも1つのサブセットにおける断片サイズ対照分子の少なくとも1つの群の各群は、等モル濃度で存在する。一部の実施形態では、少なくとも1つのサブセットにおける断片サイズ対照分子の少なくとも1つの群の各群は、等しくないモル濃度で存在する。

【0079】

一部の実施形態では、断片サイズ対照分子のサブセットの各々は、等モル濃度で存在する。一部の実施形態では、断片サイズ対照分子のサブセットの各々は、等しくないモル濃度で存在する。一部の実施形態では、サブセットにおける断片サイズ対照分子の群の各々は、等モル濃度で存在する。一部の実施形態では、サブセットにおける断片サイズ対照分子の群の各々は、等しくないモル濃度で存在する。

40

【0080】

本開示はまた、上記の方法のいずれかを実施するためのキットも提供する。例示的なキットは：(a) 既定の断片サイズ対照分子の少なくとも1つのサブセットを含む断片サイズ対照分子のセットであって、既定の断片サイズ対照分子の少なくとも1つのサブセットが、断片サイズ領域を含む複数の断片サイズ対照分子を含む、断片サイズ対照分子のセット

50

を含む。

【0081】

一部の実施形態では、方法またはシステムは、必要に応じて核酸分子の分析に関する情報、および／またはそれに由来する情報を含む報告書を作成するステップをさらに含む。

【0082】

一部の実施形態では、本明細書に開示されるシステムおよび／または方法の結果を、インプットとして使用して報告書を作成する。報告書は、書面または電子フォーマットであり得る。例えば、本明細書に開示される方法またはシステムによって決定された核酸分子の分析に関する情報および／または分析に由来する情報を、そのような報告書に表示することができる。本明細書に開示される方法またはシステムは、試料が由来する対象または医療従事者などの第3者に報告書を伝達するステップをさらに含む得る。

10

【0083】

本明細書に開示される方法の様々なステップまたは本明細書に開示されるシステムによって実行されるステップは、同じ時間もしくは異なる時間、および／または同じ地理的場所もしくは異なる地理的場所、例えば国で実行され得る。本明細書に開示される方法の様々なステップは、同じ人または異なる人々が実施することができる。

【0084】

本開示のさらなる態様および利点は、本開示の単なる例示的な実施形態を示し、記載する以下の詳細な説明から当業者に容易に明らかとなるであろう。認識されるように、本開示は、他のおよび異なる実施形態が可能であり、そのいくつかの詳細は、いずれも本開示から逸脱することなく、様々な明白な点で、変更することが可能である。したがって、図面および説明は、本質的に例示的であり、制限的ではないと考えるべきである。

20

【0085】

本明細書に組み込まれ、その一部を構成する添付の図面は、ある特定の実施形態を例証し、書かれている説明と共に本明細書で開示される方法、コンピュータ可読媒体、およびシステムのある特定の原理を説明するために役立つ。本明細書に提供される説明は、限定ではなく例として含まれる添付の図面と併せて読むとよりよく理解される。文脈により特に示していない限り、図面全体を通して類似の参照記号は類似の構成成分を特定すると理解される。同様に、図の一部または全ては、説明目的のための概略図であり得、実際の相対的なサイズまたは示された要素の位置を必ずしも描写していないと理解される。

30

【図面の簡単な説明】

【0086】

【図1-1】図1Aは、本開示の実施形態に従うポリヌクレオチドの試料中の核酸分子を分析するための方法の概略図である。図1Bは、図1Aに示す実施形態に従うポリヌクレオチドの試料中の核酸分子を分析するための方法のフローチャート図である。

【図1-2】図1Aは、本開示の実施形態に従うポリヌクレオチドの試料中の核酸分子を分析するための方法の概略図である。図1Bは、図1Aに示す実施形態に従うポリヌクレオチドの試料中の核酸分子を分析するための方法のフローチャート図である。

【0087】

【図2-1】図2Aは、本開示の実施形態に従うポリヌクレオチドの試料中の核酸分子を分析するための方法の概略図である。図2Bは、図2Aに示す実施形態に従うポリヌクレオチドの試料中の核酸分子を分析するための方法のフローチャート図である。

40

【図2-2】図2Aは、本開示の実施形態に従うポリヌクレオチドの試料中の核酸分子を分析するための方法の概略図である。図2Bは、図2Aに示す実施形態に従うポリヌクレオチドの試料中の核酸分子を分析するための方法のフローチャート図である。

【0088】

【図3】図3は、本開示の一部の実施形態との使用に適した断片サイズ対照分子の概略図である。

【0089】

【図4】図4は、本開示の一部の実施形態との使用に適した断片サイズ対照分子の概略図

50

である。

【0090】

【図5】図5は、本開示の一部の実施形態との使用に適したシステムの例の模式図である。

【発明を実施するための形態】

【0091】

定義

本開示をより容易に理解するために、ある特定の用語を最初に以下に定義する。以下の用語および他の用語の追加の定義は、本明細書を通して記載され得る。以下に記載される用語の定義が、参照により組み込まれる出願または特許における定義と一貫しない場合、本出願に記載される定義を使用して、用語の意味を理解すべきである。

【0092】

本明細書および添付の特許請求の範囲において使用される場合、単数形「1つの(a)」、「1つの(an)」、および「その(the)」は、文脈によりそれ以外であることが明確に規定されない限り、複数の参照対象を含む。したがって、例えば、「1つの(a)方法」への言及は、本明細書に記載のタイプのおよび／または本開示を読むことにより明らかになる1つまたは複数の方法および／またはステップなどを含む。

【0093】

同様に、本明細書で使用される用語法は、単に特定の実施形態を説明する目的のためであり、限定することを意図していないことも理解される。さらに、特に定義されていない限り、本明細書に使用される全ての科学技術用語は、本開示が関係する技術分野の当業者によって一般に理解されるものと同じ意味を有する。方法、コンピュータ可読媒体、およびシステムの説明および特許請求では、以下の用語法およびその文法上の変化形は以下に記載の定義に従って使用される。

【0094】

約：本明細書で使用される場合、目的の1つまたは複数の値または要素に適用される場合の「約」または「およそ」は、明記された参照値または要素と類似である値または要素を指す。ある特定の実施形態では、用語「約」または「およそ」は、特に明記されていない限り、または文脈からそれ以外であることが明らかでない限り、明記された参照値または要素のいずれかの方向に（より大きいまたはより小さい）25%、20%、19%、18%、17%、16%、15%、14%、13%、12%、11%、10%、9%、8%、7%、6%、5%、4%、3%、2%、1%、またはそれ未満以内に入る値または要素の範囲を指す（そのような数が、可能性がある値または要素の100%を超える場合を除く）。

【0095】

アダプター：本明細書で使用される場合、「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ヌクレオチド未満、約100ヌクレオチド未満、または約50ヌクレオチド未満の長さ）を指し、アダプターは、所定の試料核酸分子の一方または両方の末端に付着することができる。アダプターは、両末端にアダプターが隣接する核酸分子の増幅を可能にするための核酸プライマー結合部位、および／またはシーケンシング適用、例えば様々な次世代シーケンシング（NGS）適用のためのプライマー結合部位を含むシーケンシングプライマー結合部位を含み得る。アダプターはまた、フローセル支持体などに付着したオリゴヌクレオチドなどの捕捉プローブの結合部位も含み得る。アダプターはまた、本明細書に記載される核酸タグも含み得る。核酸タグは典型的には、核酸タグが所定の核酸分子のアンプリコンおよび配列リードに含まれるように、増幅プライマーおよびシーケンシングプライマー結合部位に対して配置される。同じまたは異なる配列のアダプターを、核酸分子のそれぞれの末端に連結することができる。一部の実施形態では、核酸タグが異なることを除いて、同じ配列のアダプターは、核酸分子のそれぞれの末端に連結される。一部の実施形態では、アダプターは、一方の末端が平滑末端であるかまたは1つもしくは複数の相補的ヌクレオチドによるテールを有する核酸分子に

10

20

30

40

50

接合するために、本明細書に記載されるように一方の末端が平滑末端であるかまたはテールを有するY字形状のアダプターであり、Y字形状のアダプターの他方の末端は、ハイブリダイズして二本鎖を形成しない非相補配列を含む。なお他の例としての実施形態では、アダプターは、分析される核酸分子に接合するための平滑またはテールを有する末端を含むベル形状のアダプターである。アダプターの他の例としては、TテールおよびCテールアダプターが挙げられる。

【0096】

増幅する：本明細書で使用される場合、核酸の文脈における「増幅する」または「増幅」は、典型的には、増幅産物またはアンプリコンが一般的に検出可能である少量のポリヌクレオチド（例えば、単一のポリヌクレオチド分子）から始まるポリヌクレオチドまたはポリヌクレオチドの一部の複数のコピーの産生を指す。ポリヌクレオチドの増幅は多様な化学的および酵素的プロセスを包含する。増幅は、ポリメラーゼ連鎖反応（PCR）を含むがこれらに限定されない。

10

【0097】

がんの型：本明細書で使用される場合、「がんの型」は、例えば組織病理学によって定義されたがんの型または亜型を指す。がんの型は、任意の従来の基準によって、例えば所定の組織における発生（例えば、血液のがん、中枢神経系（CNS）、脳がん、肺がん（小細胞および非小細胞）、皮膚癌、鼻のがん、喉のがん、肝がん、骨がん、リンパ腫、膵がん、腸がん（bowel cancer）、直腸がん、甲状腺がん、膀胱がん、腎臓がん、口腔がん、胃がん、乳がん、前立腺がん、卵巣がん、肺がん、腸がん（intestinal cancer）、軟部組織がん、神経内分泌がん、消化器がん、頭頸部がん、婦人科がん、結腸直腸がん、尿路上皮がん、固形状態のがん、不均一性がん、均一性がん）、原発不明など、ならびに／または同じ細胞系列（例えば、癌腫、肉腫、リンパ腫、胆管癌、白血病、中皮腫、黒色腫、または神経膠芽腫）、ならびに／または、例えばHer2、CA15-3、CA19-9、CA-125、CEA、AFP、PSA、HCG、ホルモン受容体、およびNMP-22などの、しかしこれらに限定されないがんマーカーを示すがんに基づいて定義することができる。がんはまた、ステージ（例えば、ステージ1、2、3、または4）、および原発性であるか続発性であるかによっても分類することができる。

20

【0098】

無細胞核酸：本明細書で使用される場合、「無細胞核酸」は、細胞内に含有されていないもしくはそうでなければ細胞に結合していない核酸、または一部の実施形態では、インタクト細胞の除去後に試料中に残っている核酸を指す。無細胞核酸は、例えば対象からの体液（例えば、血液、血漿、血清、尿、脳脊髄液（CSF）など）を供給源とする全ての封入されていない核酸を含み得る。無細胞核酸は、ゲノムDNA、ミトコンドリアDNA、循環DNA、siRNA、miRNA、循環RNA（cRNA）、tRNA、rRNA、核小体低分子RNA（snRNA）、Piwi相互作用RNA（piRNA）、長鎖非コードRNA（lncRNA）、および／またはこれらのいずれかの断片を含む、DNA（cfDNA）、RNA（cfRNA）、およびそのハイブリッドを含む。無細胞核酸は、二本鎖、一本鎖、またはそのハイブリッドであり得る。無細胞核酸は、分泌または細胞死プロセス、例えば細胞の壊死、アポトーシスなどを通して体液に放出され得る。一部の無細胞核酸、がん細胞、例えば循環腫瘍DNA（ctDNA）から体液に放出される。他は、健康な細胞から放出される。ctDNAは、封入されていない腫瘍由来断片化DNAであり得る。無細胞核酸は、1つまたは複数のエピジェネティック修飾を有することができ、例えば無細胞核酸は、アセチル化、5-メチル化、および／またはヒドロキシメチル化され得る。

30

40

【0099】

細胞核酸：本明細書で使用される場合、「細胞核酸」は、核酸がその後所定の分析プロセスの一部として除去される（例えば、細胞溶解を介して）場合であっても、少なくとも試料が対象から採取または収集された時点で、核酸が起源とする1つまたは複数の細胞内に

50

配置されている核酸を意味する。

【0100】

汚染：本明細書で使用される場合、用語「汚染」または「試料の汚染」は、1つの試料の別の試料による任意の化学的またはデジタルの汚染を指す。汚染は、試料間の液体の物理的キャリアオーバー（例えば、ピペッティング、試料調製またはシーケンサーシステムを介した自動液体取り扱い、増幅された材料の操作）；逆多重化アーチファクト（例えば、限定されたペアワイズハミング距離を有するベースコールエラー交絡試料インデックス；限定されたペアワイズ編集距離を有する挿入／欠失交絡試料インデックス）、および試薬不純物（例えば、別の試料インデックスを含有するオリゴによって汚染された（合成エラーのいずれかのキャリアオーバーを通して）試料インデックスオリゴ）を含むがこれらに限定されない多様な起源によるものであり得る。

10

【0101】

汚染スコア：本明細書で使用される場合、「汚染スコア」は、第2の試料に添加された断片サイズ対照分子の存在を表す第1の試料中のスコアを指す。一部の実施形態では、1つの試料中の1つのサブセットにおいて使用されるサブセット識別子バーコードは、同じ試料の他のサブセットおよび他の試料のサブセットにおいて使用されるサブセット識別子とは異なり得る。これらの実施形態では、断片サイズ対照分子に存在するサブセット識別子バーコードの配列から、異なる第2の試料に属する断片サイズ対照分子の存在を識別することができる。一部の実施形態では、汚染スコアは、サブセットにおいて使用される断片サイズ対照分子の各型／群に対して特異的であり得るか（すなわち、サブセットにおいて使用される断片サイズ対照分子の異なる長さ／群に関して個別の汚染スコア）、または断片サイズ対照分子の異なる長さ／群を表す総スコアであり得る。一部の実施形態では、汚染スコアは、異なる第2の試料に属する断片サイズ対照分子の数に基づいて推定することができる。一部の実施形態では、汚染スコアは、異なる第2の試料に属する断片サイズ対照分子のシーケンシングリードの数に基づいて推定することができる。一部の実施形態では、試料の汚染スコアは、他の試料に属する断片サイズ対照分子のシーケンシングリードの数の、その試料に添加された断片サイズ対照分子のシーケンシングリードの数に対する画分または百分率に基づいて推定することができる。一部の実施形態では、試料の汚染スコアは、他の試料に属する断片サイズ対照分子の分子の数の、その試料に添加された断片サイズ対照分子の分子の数に対する画分または百分率に基づいて推定することができる。これらの実施形態では、その分子に添加された断片サイズ対照分子は、断片サイズ対照分子のサブセット識別子バーコードから識別することができる。

20

30

【0102】

汚染閾値：本明細書で使用される場合、「汚染閾値」は、試料中の核酸分子の分析において試料の汚染を評価するために使用される既定の閾値の値または範囲を指す。これらの閾値はまた、抽出、ライブラリ調製、富化、洗浄／浄化、およびシーケンシングなどのステップのいずれかにおいてアッセイを最適化するためにも使用することができる。一部の実施形態では、汚染閾値は、サブセットにおいて使用される断片サイズ対照分子の各型／群に関して特異的（すなわち、サブセットにおいて使用される断片サイズ対照分子の異なる長さ／群に関して個別の汚染閾値）であり得る。一部の実施形態では、汚染閾値は、サブセットにおける断片サイズ対照分子の異なる長さ／群の総閾値、またはアッセイで添加した1つもしくは複数のサブセットにおいて使用される断片サイズ対照分子の総閾値であり得る。一部の実施形態では、サブセットにおける各群は、特定の汚染閾値を有する。例えば、断片サイズ対照分子のセットは、2つのサブセット、サブセット1およびサブセット2を含む。各サブセットが2つの群、G11およびG12（サブセット1に関して）ならびにG21およびG22（サブセット2に関して）を含む場合には、4つの群の各々は、断片サイズ対照分子の回収率に基づいて汚染スコア（S）、G11に関してS11、G12に関してS12、G21に関してS21、およびG22に関してS22を有し得る。サブセットにおける各群は、個別の既定の汚染閾値を有し得る。この例では、T11、T12、T21、およびT22はそれぞれ、G11、G12、G21、およびG22の汚染閾

40

50

値である。閾値は、百分率または画分に関する閾値であり得、閾値は、特定の閾値の値の代わりに閾値の範囲であってもよい。方法が成功であると考えられるためには、これらの汚染スコアの少なくとも1つは、その対応する汚染閾値内でなければならない。一部の実施形態では、複数の汚染スコアの各々が対応する汚染閾値内にある場合には、方法は実施に成功したと分類される。

【0103】

カバレッジ：本明細書で使用される場合、用語「カバレッジ」、「総分子数」、または「総対立遺伝子数」は、互換的に使用される。それらは、所定の試料中の特定のゲノム位置でのDNA分子の総数を指す。

【0104】

デオキシリボ核酸またはリボ核酸：本明細書で使用される場合、「デオキシリボ核酸」または「DNA」は、糖部分の2'位に水素基を有する天然または改変ヌクレオチドを指す。DNAは典型的に、4つの型のヌクレオチド塩基：アデニン（A）、チミン（T）、シトシン（C）、およびグアニン（G）を含むヌクレオチドの鎖を含む。本明細書で使用される場合、「リボ核酸」または「RNA」は、糖部分の2'位にヒドロキシル基を有する天然または改変ヌクレオチドを指す。RNAは典型的に、4つの型のヌクレオチド塩基；A、ウラシル（U）、G、およびCを含むヌクレオチドの鎖を含む。本明細書で使用される場合、用語「ヌクレオチド」は、天然のヌクレオチドまたは改変ヌクレオチドを指す。ある特定のヌクレオチドの対は互いに相補的な様式で特異的に結合する（相補的塩基対形成と呼ばれる）。DNAでは、アデニン（A）はチミン（T）と対を形成し、シトシン（C）はグアニン（G）と対を形成する。RNAでは、アデニン（A）はウラシル（U）と対を形成し、シトシン（C）はグアニン（G）と対を形成する。第1の核酸鎖が第1の鎖のヌクレオチドと相補的なヌクレオチドで構成される第2の核酸鎖に結合すると、2つの鎖は結合して二本鎖を形成する。本明細書で使用される場合、「核酸シーケンシングデータ」、「核酸シーケンシング情報」、「配列情報」、「核酸配列」、「ヌクレオチド配列」、「ゲノム配列」、「遺伝子配列」、または「断片配列、または「核酸シーケンシングリード」は、DNAまたはRNAなどの核酸の分子（例えば、全ゲノム、全トランスクリプトーム、エクソーム、オリゴヌクレオチド、ポリヌクレオチド、または断片）中のヌクレオチド塩基（例えば、アデニン、グアニン、シトシン、およびチミンまたはウラシル）の順序および同一性を示す任意の情報またはデータを指す。本教示は、キャピラリー電気泳動、マイクロアレイ、ライゲーションに基づくシステム、ポリメラーゼに基づくシステム、ハイブリダイゼーションに基づくシステム、直接または間接的ヌクレオチド識別システム、パイロシーケンシング、イオンまたはpHに基づく検出システム、および電子シグネチャーに基づくシステムを含むがこれらに限定されない、全ての利用可能な多様な技法、プラットフォーム、または技術を使用して得られる配列情報を企図すると理解すべきである。

【0105】

DNA配列：本明細書で使用される場合、「DNA配列」または「配列」は、「未加工配列リード」、および／または「コンセンサス配列」を指す。未加工配列リードは、DNAシーケンサーの出力であり、典型的には、例えば増幅後の同じ親分子の重複配列を含む。

「コンセンサス配列」は、元の親分子の配列を表すことが意図される親分子の重複配列に由来する配列である。コンセンサス配列は、投票（voting）（配列の中で、各大多数のヌクレオチド、例えば所定の塩基位置で最も一般的に観察されるヌクレオチドは、コンセンサスヌクレオチドである）、または参照ゲノムと比較するなどの他のアプローチによって產生することができる。コンセンサス配列は、元の親分子を一意的または非一意的分子タグによってタグ付けすることによって產生することができ、これによりタグの追跡および／または配列リード内部情報の使用によって子孫配列（例えば、増幅後）の追跡が可能となる。タグ付けまたはバーコード化、およびタグまたはバーコードの使用の例は、例えば、その各々の全体が参照により本明細書に組み込まれる、米国特許出願公開第2015/0368708号、第2015/0299812号、第2016/0040229

10

20

30

40

50

号、および第2016／0046986号に提供される。

【0106】

エピジェネティック修飾：本明細書で使用される場合、「エピジェネティック修飾」は、その特定の核酸配列の調節および／または遺伝子発現に影響を及ぼす核酸分子におけるヌクレオチドの塩基の修飾を指す。修飾は、ヌクレオチドの塩基の化学修飾であり得る。一部の例では、修飾は、ヌクレオチドの塩基のメチル化であり得る。例えば、修飾はシトシンのメチル化であり得、それによって5-メチルシトシンをもたらす。

【0107】

エピジェネティック状態：本明細書で使用される場合、「エピジェネティック状態」は、核酸分子のエピジェネティック修飾のレベル／程度を指す。例えば、エピジェネティック修飾がDNAメチル化（またはヒドロキシメチル化）である場合には、エピジェネティック状態は、DNA塩基（例えば、シトシン）におけるメチル化の存在もしくは非存在、または核酸配列におけるメチル化の程度（例えば、高度のメチル化、低度のメチル化、中程度のメチル化、または非メチル化核酸分子）を指し得る。エピジェネティック状態はまた、エピジェネティック修飾を有するヌクレオチドの数も指し得る。例えば、エピジェネティック修飾がDNAメチル化である場合には、エピジェネティック状態はまた、核酸分子のメチル化ヌクレオチドの数を指し得る。

【0108】

富化された試料：本明細書で使用される場合、「富化された試料」は、目的の特定の領域が富化されている試料を指す。試料は、目的の領域を選択的に増幅することによって、または二本鎖DNA／RNAプローブ（例えば、Twist Biosciencesのプローブ）もしくは目的の核酸分子にハイブリダイズすることができる一本鎖DNA／RNAプローブ（例えば、SureSelect（登録商標）プローブ、Agilent Technology）を使用することによって富化することができる。一部の実施形態では、富化された試料は、富化される処理された試料のサブセットであって、富化される処理された試料のサブセットが、無細胞ポリヌクレオチドの試料、ならびに断片サイズ対照分子の第1および／または第2のサブセットからの核酸分子を含有する、サブセットを指す。他の実施形態では、富化された試料は、富化される第3のスパイクイン試料のサブセットであって、富化される第3のスパイクイン試料のサブセットが、無細胞ポリヌクレオチドの試料からの、ならびに断片サイズ対照分子の第1、第2、および／または第3のサブセットからの核酸分子を含有する、サブセットを指す。

【0109】

第1のスパイクイン試料：本明細書で使用される場合、「第1のスパイクイン試料」は、断片サイズ対照分子の第1のサブセットが、対象からの無細胞ポリヌクレオチドの試料に添加されている試料である。

【0110】

第4のスパイクイン試料：本明細書で使用される場合、「第4のスパイクイン試料」は、断片サイズ対照分子の第4のサブセットが、富化された試料のサブセットに添加されている試料を指す。

【0111】

断片サイズバイアス：本明細書で使用される場合、「断片サイズバイアス」は、アッセイにおいて分析される核酸分子の断片の長さまたはサイズにおける任意の人為的バイアスを指す。アーチファクトバイアスは、オペレーターの取り扱いまたはアッセイに使用した試薬および／もしくはステップが原因であり得る。このバイアスは、健康な対象と疾患を有する対象の間で観察される断片サイズの真の生物学的バイアスを含まない。アッセイは、液体の取り扱い、抽出、ライブラリ調製、富化、洗浄／浄化、およびシーケンシングなどの、しかしこれらに限定されない1つまたは複数のステップを含み得る。一部の実施形態では、試料毎に異なる断片サイズバイアスが存在し得る。反応条件およびこれらのステップの各々の手順に基づいて、特定の断片サイズに属する核酸分子の回収率はバイアスを受け得る。フラグメントーム分析では、特定のサイズのcfDNA分子の正確な定量的測定

10

20

30

40

50

を有することは、無細胞DNAの断片長分布および断片化パターンの推定において有用である。アッセイにおける断片サイズバイアスを推定することによって、アッセイワークフローにおける各ステップによって導入される断片サイズバイアス（人為的バイアス）を補正し、cfDNA分子の元の断片サイズ分布（真の生物学的バイアスを反映する）のより良好な推定値を有することができる。

【0112】

断片サイズ対照分子：本明細書で使用される場合、「断片サイズ対照分子」は、試料中の核酸分子の分析を評価および／または最適化するためにポリヌクレオチドの試料に添加される核酸分子のセットを指す。断片サイズ対照分子は、2つの領域、断片サイズ領域および識別子領域を有し得る。一部の実施形態では、断片サイズ対照分子は、断片サイズ領域のみからなる。一部の実施形態では、断片サイズ対照分子は、断片サイズ領域および識別子領域の両方を有し得る。断片サイズ対照分子のセットを、断片サイズ対照分子のサブセットに分類することができ、断片サイズ対照分子の各サブセットを、断片長分布、QC指標、ならびに／またはポリヌクレオチド試料の抽出、ライブラリ調製、富化、および／もしくはシーケンシングのいずれかのステップにおける試料の汚染を評価するために、アッセイにおける1つまたは複数の異なるステップで添加することができる。一部の実施形態では、断片サイズ対照分子を使用して、試料中のポリヌクレオチドの元の末端がアッセイワークフローを通してどのように良好に保存されているかを推定することができる。これらの断片サイズ対照分子の長さは既定であり、無細胞DNAの天然のサイズ分布を模倣することができる。断片サイズ対照分子の各サブセットは、断片サイズ領域の長さに基づいて断片サイズ対照分子の群にさらに分類することができ、各群は異なる長さの断片サイズ領域を有し得る。例えば、断片サイズ対照分子のサブセットを断片サイズ領域の長さに基づいて3つの群に分類することができ、第1の群は、長さ120bpの断片サイズ領域を有し得、第2の群は、長さ160bpの断片サイズ領域を有し得、第3の群は、長さ320bpの断片サイズ領域を有し得る。一部の実施形態では、断片サイズ対照分子は、合成分子であり得、すなわち酵素を使用してまたは使用せずに*in vitro*で合成することができる。一部の実施形態では、断片サイズ対照分子は、天然に存在しない核酸配列を有し得る。一部の実施形態では、断片サイズ対照分子は、天然に存在する核酸配列を有し得る。一部の実施形態では、任意のゲノム、プラスミド、もしくはベクター、またはその一部における特定の領域（すなわち、特定の長さの）の増幅を介して生成されたアンプリコンを、断片サイズ対照分子として使用することができる。一部の実施形態では、プラスミド、ベクター、または任意のゲノムを、制限酵素を使用して消化することができ、消化産物を、断片サイズ対照分子として使用することができる。一部の実施形態では、断片サイズ対照分子は、非ヒトゲノムに対応する核酸配列を有し得る。例えば、これらの分子は、(i) ラムダファージDNAの領域に対応する配列、(ii) 天然に存在しない配列、および／または(iii) (i)と(ii)の組合せのいずれかを有し得る。一部の実施形態では、断片サイズ対照分子は、天然に存在しないヌクレオチドアナログを含み得る。

【0113】

断片サイズ領域：本明細書で使用される場合、用語「断片サイズ領域」は、断片サイズ対照分子の長さを表す断片サイズ対照分子の領域を指す。断片サイズ対照分子の各群は、断片サイズ領域の長さが異なる。例えば、断片サイズ対照分子のサブセットを、断片サイズ領域の長さに基づいて3つの群に分類することができ、第1の群は、長さ120bpの断片サイズ領域を有し得、第2の群は、長さ160bpの断片サイズ領域を有し得、第3の群は、長さ320bpの断片サイズ領域を有し得る。断片サイズ領域の長さは10bp～1000bpの間であり得る。

【0114】

断片サイズスコア：本明細書で使用される場合、「断片サイズスコア」は、特定の群および特定のサブセットに属する断片サイズ対照分子の回収率を表すスコアを指す。断片サイズ対照分子および断片サイズ対照分子が属するサブセットの同一性は、バーコードの使用を通して維持される。一部の実施形態では、断片サイズスコアは、特定の群および特定の

サブセットに属する断片サイズ対照分子の数に基づいて推定することができる。一部の実施形態では、断片サイズスコアは、特定の群および特定のサブセットに属する断片サイズ対照分子のシーケンシングリードの数に基づいて推定することができる。一部の実施形態では、断片サイズスコアは、特定のサブセットにおける特定の群に関して推定することができる。一部の実施形態では、断片サイズスコアは、あらゆるサブセットにおける群の各々に関して推定することができる。一部の実施形態では、断片サイズスコアは、特定のサブセットにおいて使用される断片サイズ対照分子の総スコア（すなわち、サブセット内の断片サイズ対照分子を表す全ての単一スコアに関して）、またはアッセイの1つもしくは複数のステップで添加された断片サイズ対照分子の総スコア（すなわち、アッセイで添加した1つまたは複数のサブセットにおける断片サイズ対照分子を表す単一スコア）であり得る。一部の実施形態では、断片サイズスコアは、特定のステップで添加された断片サイズ対照分子のサブセットの量と、異なるステップで添加された断片サイズ対照分子の別のサブセットの量との差として測定することができ、一部の実施形態では、量は、質量（例えば、fg、pg、ng、μg）、シーケンシングリードの数または断片サイズ分子の数（サブセットに属する、または特定の群に属する）のいずれかに関する量であり得る。一部の実施形態では、断片サイズスコアは、断片サイズ対照分子（サブセットに属する、または特定の群に属する）の回収率（画分または百分率）として測定することができ、回収率は、試料に添加された特定のサブセット内の断片サイズ対照分子の相対的存在量のオルトゴナル測定を利用して計算することができる。一部の実施形態では、オルトゴナル測定は、ゲル電気泳動、qPCR、ddPCR、またはPCRフリーシーケンシングから得ることができる。一部の実施形態では、断片サイズスコアは、特定の長さの断片サイズ対照分子の量と、異なる長さの断片サイズ対照分子の量との差として測定することができる。

10

20

30

40

【0115】

断片サイズ閾値：本明細書で使用される場合、「断片サイズ閾値」は、試料中の核酸分子の分析における断片サイズバイアスを評価するために使用される既定の閾値の値または範囲を指す。これらの閾値を使用して、抽出、ライブラリ調製、富化、洗浄／浄化、およびシーケンシングなどのステップのいずれかにおける任意の断片長分布を補正することもできる。一部の実施形態では、断片サイズ閾値は、特定のサブセットにおける特定の群に関して推定することができる。一部の実施形態では、サブセットにおける各群は、特定の断片サイズ閾値を有する。一部の実施形態では、断片サイズ閾値は、あらゆるサブセットにおける群の各々に関して推定することができる。一部の実施形態では、断片サイズ閾値は、特定のサブセットにおいて使用される断片サイズ対照分子の総閾値（すなわち、サブセット内の断片サイズ対照分子の全ての単一閾値に関して）、またはアッセイの1つもしくは複数のステップで添加された断片サイズ対照分子の総閾値（すなわち、アッセイで添加した1つまたは複数のサブセットにおける断片サイズ対照分子の単一閾値）であり得る。例えば、断片サイズ対照分子のセットは、2つのサブセット、サブセット1およびサブセット2を含む。各サブセットが2つの群、G11およびG12（サブセット1に関して）ならびにG21およびG22（サブセット2に関して）を含む場合には、4つの群の各々は、断片サイズ対照分子の回収率に基づいて断片サイズスコア（S）、G11に関してS11、G12に関してS12、G21に関してS21、およびG22に関してS22を有し得る。サブセットにおける各群は、個別の既定の断片サイズ閾値を有し得る。この例では、T11、T12、T21、およびT22はそれぞれ、G11、G12、G21、およびG22の断片サイズ閾値である。閾値は、百分率または画分に関する閾値であり得、閾値は、特定の閾値の値の代わりに閾値の範囲であり得る。方法が成功であると考えられるためには、これらの断片サイズスコアの少なくとも1つは、その対応する断片サイズ閾値内でなければならない。一部の実施形態では、複数の断片サイズスコアの各々が対応する断片サイズ閾値内にある場合には、方法は実施に成功したと分類される。

【0116】

ゲノム領域：本明細書で使用される場合、「ゲノム領域」は、ゲノム、例えば全ゲノム、染色体、遺伝子、またはエクソンの任意の領域（例えば、塩基対位置の範囲）を指す。ゲ

50

ノム領域は、連続または不連続領域であり得る。「遺伝子座」（または「座」）は、ゲノム領域の一部または全体（例えば、遺伝子、遺伝子の一部、または遺伝子の単一ヌクレオチド）であり得る。

【0117】

識別子領域：本明細書で使用される場合、「識別子領域」は、断片サイズ対照分子を他の断片サイズ対照分子から区別するために使用される断片サイズ対照分子の領域を指す。識別子領域はまた、1つのサブセットからの断片サイズ対照分子を、別のサブセットからの断片サイズ対照分子と区別するためにも使用される。識別子領域は、分子バーコードを有し得る。識別子領域は、断片サイズ領域の片側または両側に存在し得る。識別子領域は、連続または不連続のいずれかであり得る。分子バーコードは、断片サイズ対照分子の識別子として役立つ。識別子領域は、1つまたは複数のプライマーの結合を容易にする追加の領域（プライマー結合部位）を有し得る。一部の実施形態では、識別子領域はまた、断片サイズ対照分子が次世代シーケンサー（例えば、Illuminaシーケンサー）のフローセル表面に付着するのを可能にする追加のフローセル結合部位、例えばP5およびP7も有し得る。一部の実施形態では、識別子領域は、（i）分子識別子バーコード、例えば各断片サイズ対照分子を識別し、1つの断片サイズ対照分子を別の断片サイズ対照分子と識別するために使用されるバーコード、および（ii）サブセット識別子バーコード、例えば断片サイズ対照分子が属するサブセットを識別するために使用されるバーコード（すなわち、断片サイズ対照分子がサブセット1またはサブセット2のいずれに属するか）として作用するバーコードを含む。サブセット識別子バーコードは、サブセットにおける全ての断片サイズ対照分子に関して同じであってもよく、1つのサブセットのサブセット識別子バーコードは、他のサブセットのサブセット識別子バーコードとは異なってもよい。一部の実施形態では、1つの試料中の1つのサブセットのサブセット識別子バーコードは、他の試料中の対応するサブセットのサブセット識別子バーコードとは異なり得る。例えば、断片サイズ対照分子を使用して試料の汚染を評価する場合、1つのサブセットのサブセット識別子バーコードは、同じフローセル／バッチ内の全ての他の試料において使用されるその対応するサブセットのサブセット識別子バーコードとは異なる。一部の実施形態では、識別子領域は、1つの試料の断片サイズ対照分子を他の試料の断片サイズ対照分子と区別するための試料インデックス、およびシーケンサーのフローセルに付着させるための配列を含み得る。例えば、富化ステップの後に添加される断片サイズ対照分子のサブセットの識別子領域は、試料インデックス配列を含み得る。一部の実施形態では、識別子領域を、ライゲーションを介して断片サイズ領域に付着させることができる。一部の実施形態では、識別子領域を、PCRを介して断片サイズ領域に付加することができる。

【0118】

突然変異：本明細書で使用される場合、「突然変異」は、公知の参照配列からの変異を指し、例えば単一ヌクレオチドバリエーション（SNV）などの突然変異、および挿入または欠失（インデル）を含む。突然変異は、生殖系列または体細胞突然変異であり得る。一部の実施形態では、比較目的のための参照配列は、試験試料を提供する対象の種の野生型ゲノム配列、典型的にはヒトゲノムである。

【0119】

新生物：本明細書で使用される場合、用語「新生物」および「腫瘍」は互換的に使用される。それらは、対象における細胞の異常な成長を指す。新生物または腫瘍は、良性、潜在的に悪性、または悪性であり得る。悪性腫瘍は、がんまたはがん性腫瘍と呼ばれる。

【0120】

次世代シーケンシング：本明細書で使用される場合、「次世代シーケンシング」または「NGS」は、従来のサンガーおよびキャピラリー電気泳動に基づくアプローチと比較して増加したスループットを有し、例えば何十万もの比較的小さい配列リードを一度に生成する能力を有するシーケンシング技術を指す。次世代シーケンシング技法の一部の例としては、合成によるシーケンシング、ライゲーションによるシーケンシング、およびハイブリダイゼーションによるシーケンシングが挙げられるがこれらに限定されない。一部の実施

形態では、次世代シーケンシングは、単一分子をシーケンシングすることが可能な機器の使用を含む。

【0121】

核酸タグ：本明細書で使用される場合、「核酸タグ」は、核酸を、異なる試料からの核酸（例えば、試料インデックスを表す）または同じ試料中の異なる核酸分子（例えば、分子バーコードを表す）、異なる型の核酸、または異なる処理を受けている核酸と区別するために使用される短い核酸（例えば、約500ヌクレオチド未満、約100ヌクレオチド未満、約50ヌクレオチド未満、または約10ヌクレオチド未満の長さ）を指す。核酸タグは、既定の、固定された、ランダムでない、ランダムな、またはセミランダムなオリゴヌクレオチド配列を含む。そのような核酸タグを使用して、異なる核酸分子または異なる核酸試料またはサブサンプルを標識することができる。核酸タグは、一本鎖、二本鎖、または少なくとも部分的に二本鎖であり得る。核酸タグは、必要に応じて同じ長さまたは多様な長さを有する。核酸タグは、1つもしくは複数の平滑末端を有する二本鎖分子を含み、5'もしくは3'一本鎖領域（例えば、オーバーハング）を含み、および／または所定の分子内の他の位置に1つもしくは複数の他の一本鎖領域を含み得る。核酸タグは、他の核酸（例えば、増幅および／またはシーケンシングされる試料核酸）の一方の末端または両方の末端に付着させることができる。核酸タグを復号して、所定の核酸の起源試料、形態、または処理などの情報を明らかにすることができる。例えば、核酸タグを使用して、異なる分子バーコードおよび／または試料インデックスを有する核酸を含む複数の試料のプールおよび／または並行処理を可能にすることもでき、この場合、その後、核酸タグを検出すること（例えば、読み取ること）によって核酸をデコンボリューションする。核酸タグ（分子バーコード）はまた、識別子（例えば、分子識別子、試料識別子）とも呼ばれ得る。加えてまたはあるいは、核酸タグを分子識別子として使用することができる（例えば、同じ試料またはサブサンプル中の異なる分子または異なる親分子のアンプリコンを区別するために）。これは、例えば、所定の試料中の異なる核酸分子を一意的にタグ付けすること、またはそのような分子を非一意的にタグ付けすることを含む。非一意的にタグ付けする適用の場合、有限数のタグ（すなわち、分子バーコード）を使用して各核酸分子をタグ付けすることができ、その結果、異なる分子を、少なくとも1つの分子バーコードと組み合わせたそれらの内因性配列情報（例えば、選択された参照ゲノムにそれらがマッピングされる開始および／もしくは終止位置、配列の一方もしくは両方の末端の部分配列、および／または配列の長さ）に基づいて区別することができる。典型的には、任意の2つの分子が同じ内因性配列情報（例えば、開始および／または終止位置、配列の一方もしくは両方の末端の部分配列、および／または長さ）を有し、同じ分子バーコードも有し得る確率が低くなるように（例えば、約10%未満、約5%未満、約1%未満、または約0.1%未満の可能性）、十分な数の異なる分子バーコードを使用する。

【0122】

分画化：本明細書で使用される場合、「分画化」および「エピジェネティック分画化」は、互換的に使用される。この用語は、核酸分子を、核酸分子の特徴（例えば、エピジェネティック修飾のレベル／程度）に基づいて分離または分別することを指す。分画化は、分子の物理的分画化であり得る。分画化は、エピジェネティック修飾のレベル（すなわち、エピジェネティック状態）に基づいて、核酸分子を群またはセットに分離することを伴い得る。例えば、核酸分子のメチル化のレベルに基づいて核酸分子を分画することができる。一部の実施形態では、分画化のために使用される方法およびシステムは、その全体が参照によって組み込まれるPCT特許出願第PCT/US2017/068329号に見出され得る。

【0123】

分画化セット：本明細書で使用される場合、「分画化セット」は、核酸分子の結合剤に対する示差的な結合親和性に基づいてセット／群に分画された核酸分子のセットを指す。結合剤は、エピジェネティック修飾を有するヌクレオチドを含む核酸分子に優先的に結合する。例えば、エピジェネティック修飾がメチル化である場合、結合剤は、メチル結合ドメ

イン (MBD) タンパク質であり得る。一部の実施形態では、分画化セットは、特定のレベル／程度のエピジェネティック修飾 (すなわち、エピジェネティック状態) に属する核酸分子を含み得る。例えば、核酸分子を3つのセット: 高メチル化分画化セットまたは高分画化セットと呼ぶことができる高度にメチル化された核酸分子 (または高メチル化核酸分子) の1つのセット、低メチル化分画化セットまたは低分画化セットと呼ぶことができる低度メチル化された核酸分子 (または低メチル化核酸分子) の別のセット、および中程度にメチル化された分画化セットまたは中分画化セットと呼ぶことができる中程度にメチル化された核酸分子の第3のセットに分画することができる。別の例では、核酸分子を、エピジェネティック修飾を有するヌクレオチドの数に基づいて分画することができ、1つの分画化セットは、9個のメチル化ヌクレオチドを有する核酸分子を有し得、別の分画化セットは、非メチル化核酸分子 (ゼロ個のメチル化ヌクレオチド) を有し得る。

10

【0124】

ポリヌクレオチド: 本明細書で使用される場合、「ポリヌクレオチド」、「核酸」、「核酸分子」、または「オリゴヌクレオチド」は、ヌクレオシド間連結によって接合したヌクレオシドの直鎖状ポリマー (デオキシリボヌクレオシド、リボヌクレオシド、またはその類似体を含む) を指す。典型的には、ポリヌクレオチドは、少なくとも3つのヌクレオシドを含む。オリゴヌクレオチドのサイズは、多くの場合、数個の単量体単位、例えば3~4個から数百の単量体単位の範囲である。ポリヌクレオチドが「ATGCCCTG」などの文字列によって表される場合は常に、特に明記していない限り、ヌクレオチドは左から右に5'→3'の順序であること、およびDNAの場合では、「A」はデオキシアデノシンを示し、「C」はデオキシシチジンを示し、「G」はデオキシグアノシンを示し、「T」はデオキシチミジンを示すことが理解される。文字A、C、G、およびTは、塩基そのもの、ヌクレオシド、または塩基を含むヌクレオチドを指すために使用され得る。

20

【0125】

処理: 本明細書で使用される場合、「処理」は、シーケンシングに適した核酸のライブラリを生成するために使用されるステップのセットを指す。ステップのセットは、核酸の分画化、末端修復、シーケンシングアダプターの付加、タグ付け、および／またはPCR増幅を含み得るがこれらに限定されない。

【0126】

処理された試料: 本明細書で使用される場合、「処理された試料」は、本明細書において他所で記載されるように処理されている試料を指す。一部の実施形態では、処理された試料は、処理された第1のスパイクイン試料から抽出された核酸のサブセットを指し得、第1のスパイクイン試料から抽出された核酸のサブセットは、無細胞ポリヌクレオチドの試料からの、および断片サイズ対照分子の第1のサブセットからの核酸分子を含有する。他の実施形態では、処理された試料は、処理された第2のスパイクイン試料のサブセットを指し、処理された第2のスパイクイン試料のサブセットは、無細胞ポリヌクレオチドの試料からの、ならびに断片サイズ対照分子の第1のサブセットおよび第2のサブセットからの核酸分子を含有する。

30

【0127】

定量的測定: 本明細書で使用される場合、「定量的測定」は、絶対的または相対的測定を指す。定量的測定は、数、統計学的測定 (例えば、頻度、平均値、中央値、標準偏差、また四分位数)、または程度もしくは相対量 (例えば、高い、中等度、または低い) であり得るがこれらに限定されない。定量的測定は、2つの定量的測定の比であり得る。定量的測定は、定量的測定の線形の組合せであり得る。定量的測定は、正規化された測定であり得る。

40

【0128】

参照配列: 本明細書で使用される場合、「参照配列」は、実験により決定された配列との比較目的のために使用される公知の配列を指す。例えば、公知の配列は、全ゲノム、染色体、またはそれらの任意のセグメントであり得る。一部の実施形態では、参照配列は、少なくとも約20、少なくとも約50、少なくとも約100、少なくとも約200、少なく

50

とも約250、少なくとも約300、少なくとも約350、少なくとも約400、少なくとも約450、少なくとも約500、少なくとも約1000、または1000よりも多くのヌクレオチドであり得る。参照配列は、ゲノムもしくは染色体の単一の連続した配列と整列することができるか、またはゲノムもしくは染色体の異なる領域と整列する不連続のセグメントを含み得る。一部の実施形態では、参照配列は全ゲノムであり得る。参照配列の例としては、例えば、hG19およびhG38などのヒトゲノムが挙げられる。

【0129】

試料：本明細書で使用される場合、「試料」は、本明細書に開示される方法および／またはシステムによって分析することが可能な任意のものを意味する。

【0130】

第2のスパイクイン試料：本明細書で使用される場合、「第2のスパイクイン試料」は、断片サイズ対照分子の第2のサブセットが、第1のスパイクイン試料から抽出された核酸に添加されている試料であり、第1のスパイクイン試料から抽出された核酸は、無細胞ポリヌクレオチドの試料からの、および断片サイズ対照分子の第1のサブセットからの核酸分子を含有する。

【0131】

シーケンシング：本明細書で使用される場合、「シーケンシング」は、生体分子、例えば、DNAまたはRNAなどの核酸の配列（例えば、単量体単位の同一性および順序）を決定するために使用されるいくつかの技術のいずれかを指す。シーケンシング方法の例としては、標的化シーケンシング、単一分子リアルタイムシーケンシング、エクソンまたはエキソームシーケンシング、イントロンシーケンシング、電子顕微鏡に基づくシーケンシング、パネルシーケンシング、トランジスタ媒介性シーケンシング、直接シーケンシング、ランダムショットガンシーケンシング、サンガージデオキシターミネーションシーケンシング、全ゲノムシーケンシング、ハイブリダイゼーションによるシーケンシング、パイロシーケンシング、2重鎖シーケンシング、サイクルシーケンシング、一塩基伸長シーケンシング、固相シーケンシング、ハイスループットシーケンシング、大規模並列処理シグネチャーシーケンシング、エマルジョンPCR、低変性温度での共増幅PCR（COLD-PCR）、多重PCR、可逆的ダイターミネーターによるシーケンシング、ペアエンドシーケンシング、短期シーケンシング、エキソヌクレアーゼシーケンシング、ライゲーションによるシーケンシング、ショートリードシーケンシング、単一分子シーケンシング、合成によるシーケンシング、リアルタイムシーケンシング、リバースターミネーターシーケンシング、ナノポアシーケンシング、454シーケンシング、Solexa Genome Analyzerシーケンシング、SOLiD（商標）シーケンシング、MS-PETシーケンシング、およびこれらの組合せが挙げられるがこれらに限定されない。一部の実施形態では、シーケンシングは、例えば、多くの他のものの中でも、Illumina, Inc.、Pacific Biosciences, Inc.、またはApplied Biosystems/Thermo Fisher Scientificから市販されている遺伝子分析機器などの遺伝子分析機器によって実施することができる。

【0132】

配列情報：本明細書で使用される場合、核酸ポリマーの文脈における「配列情報」は、そのポリマー内の単量体単位（例えば、ヌクレオチドなど）の順序および同一性を意味する。

【0133】

体細胞突然変異：本明細書で使用される場合、用語「体細胞突然変異」または「体細胞変異」は、互換的に使用される。これらの用語は、受胎後に起こるゲノム内の突然変異を指す。体細胞突然変異は、生殖細胞以外の体の任意の細胞において起こり得、したがって、後代には受け渡されない。

【0134】

対象：本明細書で使用される場合、「対象」は、哺乳動物種（例えば、ヒト）もしくは鳥類（例えば、トリ）の種などの動物、または植物などの他の生物体を指す。より具体的に

10

20

30

40

50

は、対象は、脊椎動物、例えば、マウス、霊長類、サルまたはヒトなどの哺乳動物であり得る。動物は、農場動物（例えば、肉用牛、乳牛、家禽、ウマ、ブタなど）、競技動物、およびコンパニオン動物（例えば、愛玩動物または支援動物）を含む。対象は、健康な個体、疾患もしくは疾患に対する素因を有するもしくはそれを有することが疑われる個体、または治療を必要とするもしくは治療を必要とすることが疑われる個体であり得る。用語「個体」または「患者」は、「対象」と互換的であると意図される。

【0135】

例えば、対象は、がんを有すると診断されている、がん治療を受ける予定の、および／または少なくとも1つのがん治療を受けたことがある個体であり得る。対象は、がんが寛解した状態であり得る。別の例として、対象は、自己免疫疾患を有すると診断された個体であり得る。別の例として、対象は、疾患、例えば、がん、自己免疫疾患を有すると診断されているかまたはそれを有することが疑われ得る妊娠中または妊娠する計画がある女性個体であり得る。

10

【0136】

第3のスパイクイン試料：本明細書で使用される場合、「第3のスパイクイン試料」は、断片サイズ対照分子の第3のサブセットが処理された試料に添加されている試料である。

I. 概要

【0137】

循環無細胞DNAにおける断片サイズ分布および断片化パターンは、無細胞DNAの起源に関する情報、および疾患レベルに関する情報も提供することができる。しかし、断片長およびパターンを読み取るまたは分析する（例えば、次世代シーケンシングのライブラリ調製）ために使用される様々な抽出プロセスおよび方法を通して、断片サイズ分布および断片化パターンにバイアスが導入され得る。したがって、バイアスの起源、プロセスのどのステップでバイアスが起こるかを理解するために使用することができる対照を有すること、ならびにまたバイアスを補正するためまたは試料毎におよびバッチレベルの断片サイズバイアスに関して正規化するためにこれらの対照を使用することが必須である。

20

【0138】

本開示は、断片長の回収率を正規化するための、所望のサイズの断片の全て／大部分を回収するためにアッセイを最適化するための、およびQC指標としての方法、組成物、およびシステムを提供する。そのような方法は、例えばシーケンシングを通しての抽出からの異なるステップで無細胞ポリヌクレオチドの試料に添加することができる対照として断片サイズ対照分子を使用することを含み得る。これらの対照は、無細胞DNAの天然のサイズ分布を模倣するために既定の断片長を有し得る。異なるステップで添加される断片サイズ対照分子の同一性は、特定のステップで断片サイズバイアスを把握することができるように、異なるバーコードを使用することによって維持することができる。これらの対照分子は、これらに限定されないが、以下などの多くの適用において使用することができる：

30

(i) 断片サイズバイアスのモニタリング、(ii) 断片サイズバイアスを低減するためのプロセスの最適化、(iii) 断片サイズ対照分子の回収率を分析することによってプロセス中に導入された断片サイズバイアスの補正または正規化、(iv) QC指標としての断片サイズ対照分子の回収率に基づくアッセイの性能の評価、(v) 試料の任意の汚染の分析するためのQC指標として、および(v) 任意の汚染を最小限にするためのアッセイの最適化。

40

【0139】

がんの形成および進行は、デオキシリボ核酸(DNA)の遺伝子修飾およびエピジェネティック修飾の両方から生じ得る。本開示は、DNA、例えば無細胞DNA(cfDNA)のエピジェネティック修飾の分析方法を提供する。そのようなエピジェネティック分析は、断片長分布の変化またはゲノム位置にマッピングされる断片エンドポイントの頻度を測定することにより識別可能なメチル化およびDNA断片パターンを含む。そのような「フラグメントーム」分析を単独でまたは既存の技術と組み合わせて使用して、疾患もしくは状態の有無、診断される疾患もしくは状態の予後、診断される疾患もしくは状態の治療的

50

処置、または疾患もしくは状態の予想される処置の転帰を決定することができる。そのような疾患または状態の例はがんである。

【0140】

循環無細胞DNA (cfDNA) は、死につつある組織細胞から体液、例えば末梢血 (血漿または血清) 中に脱落した主に短いDNA断片 (例えば、長さ約100~400塩基対、最頻値約165bpを有する) であり得る。cfDNAの分析により、がん関連遺伝子バリエーションに加えて、死につつある細胞のエピジェネティックフットプリントおよび食細胞による除去のシグネチャーが明らかとなり得、これによって本悪性腫瘍 (例えば、腫瘍) ならびにその微小環境構成成分の凝集ヌクレオソーム占有プロファイルがもたらされ得る。

10

【0141】

血漿フラグメントームシグナル (例えば、cfDNA断片の分析から得たシグナル) に寄与し得る構成成分または要因としては、悪性の固形腫瘍が腫瘍関連の正常胞、上皮、および間質細胞、免疫細胞、ならびに血管細胞を含み、そのありとあらゆるものがcfDNA試料 (例えば、対象の体液から得られ得る) に寄与し、その中で表され得ることから、(i) DNAの破壊中の細胞死の型および関連するクロマチン凝縮事象、(ii) 対象の免疫系によって調節される様々な型の貪食の仕組みを伴い得るクリアランス機構、(iii) 循環中の細胞型の基礎となる組合せによって影響を受け得る血液組成の非悪性の変異、(iv) 所定の型の臓器または組織における非悪性の細胞死の複数の起源または原因、ならびに(v) がん内の細胞型の不均一性が挙げられる。

20

【0142】

ヒストン保護複体の形態での無細胞DNAは、好中球、マクロファージ、好酸球、ならびに腫瘍細胞を含む様々な宿主細胞によって放出され得る。循環DNAは、典型的に半減期が短く (例えば、約10~15分)、肝臓は、典型的に循環DNA断片を血液循環から除去する主要な臓器である。循環中のcfDNAの蓄積は、細胞死および／もしくは活性化の増加、cfDNAのクリアランス障害、ならびに／または内因性のDNアーゼ酵素レベルの減少に起因し得る。対象の血流中を循環する無細胞DNAは、典型的には膜で覆われた構造 (例えば、アポトーシス体) に充填され得るか、または生体高分子 (例えば、ヒストンまたはDNA結合血漿タンパク質) と複合体を形成し得る。DNA断片化およびその後の輸送のプロセスを、フラグメントーム分析によって検出される無細胞DNAシグナルの特徴に対するその効果に関して分析することができる。

30

【0143】

細胞核 (例えば、ヒトの) において、DNAは典型的にはヌクレオソームに存在し、これはコアヒストン8量体の周囲を包む約145塩基対 (bp) のDNAを含む構造へと組織化される。DNAおよびヌクレオソームの静電気的および水素結合相互作用は、タンパク質表面上でのDNAのエネルギー的に不都合な湾曲をもたらし得る。そのような湾曲は、他のDNA結合タンパク質の立体障害を引き起こすことがあり得、したがって細胞核におけるDNAに対するアクセスを調節するために役立ち得る。細胞におけるヌクレオソームの位置決めは、動的に (例えば、時間と共におよび様々な細胞状態および条件にわたって)、例えばDNAのアンラッピングおよびリラッピングによって変動し得る。フラグメントームシグナルは、ヌクレオソーム単位によって影響を受ける立体配置に由来するヌクレオソーム保護DNA断片を反映し得ることから、ヌクレオソームの安定性およびダイナミクスは、そのようなフラグメントームシグナルに影響を及ぼし得る。これらのヌクレオソームダイナミクスは、多様な要因、例えば：(i) ATP加水分解のエネルギーを使用して、ヌクレオソームをスライドさせ、クロマチン繊維からヒストンを交換または排除し得るATP依存的リモデリング複合体、(ii) 正規のヒストンの特性とは異なる特性を有し、クロマチン繊維内で局所的な特異的ドメインを作製し得るヒストンバリエーション、(iii) 遊離のヒストンの供給を制御し、ヒストン沈着および排除においてクロマチンリモデリングと協調し得るヒストンシャペロン、(iv) クロマチンの構造に直接または間接的に影響を及ぼし得るヒストンの翻訳後修飾 (PTM) (例えば、アセチル化、メチル化、

40

50

リン酸化、およびユビキチン化)、ならびに(v)転写因子およびRNAポリメラーゼによる能動的転写に由来し得る。

【0144】

したがって、cfDNAにおける断片化シグナルまたはパターンは、ゲノム全体のクロマチン組織化の不均一性に関連する複数の事象に由来する凝集したcfDNAシグナルを示し得る。そのようなクロマチン組織化は、死につつある細胞における全体的な細胞の同一性、代謝状態、限局的調節状態、局所遺伝子活性、およびDNAクリアランスの機構などの要因に応じて異なり得る。その上、無細胞DNAフラグメントームシグナルは、寄与する細胞の基礎となるクロマチン構造に部分的にのみ起因し得る。そのようなcfDNAフラグメントームシグナルは、細胞死の間にクロマチンの圧縮のより複雑なフットプリントおよび酵素消化からのDNA保護を示し得る。したがって、所定の細胞型または細胞系列型に対して特異的なクロマチンマップは、細胞死または破片の輸送の様々な段階でのヌクレオソームの安定性、立体構造、および組成の変化により、DNAのアクセシビリティの固有の不均一性に部分的にのみ寄与し得る。その結果、一部のヌクレオソームは、無細胞DNAに優先的に存在するまたは存在しないようになり得るが(例えば、cfDNAクリアランスに影響を及ぼし、血液循環に放出する濾過機構が存在し得る)、これらは死の様式および機構ならびに細胞死体のクリアランスなどの要因に依存し得る。

10

【0145】

フラグメントームシグナルは、細胞において生成され、アポトーシスおよび壊死などの細胞プロセス中の核DNA断片化の結果として血液循環にcfDNAとして放出され得る。そのような断片化は、細胞の異なる段階におけるDNAに作用する異なるヌクレアーゼ酵素の結果として産生され、cfDNAフラグメントームシグナルにおいて分析され得る配列特異的DNA切断パターンをもたらし得る。そのような切断パターンを分類することは、細胞環境の臨床的に関連するマーカー(例えば、腫瘍の微小環境、炎症、疾患状態、腫瘍形成など)であり得る。

20

【0146】

本開示は、ポリヌクレオチド(一部の実施形態では、ポリヌクレオチドは無細胞ポリヌクレオチドであり得る)の試料中の核酸分子の分析における断片サイズバイアスを評価および補正するための方法、組成物、およびシステムを提供する。これらの方法は、様々な適用において、例えば、疾患の予後判定、診断、および/またはモニタリングのために使用することができる。

30

【0147】

ポリヌクレオチドの試料中の核酸分子の分析は、断片サイズ対照分子の回収率を測定することから断片サイズバイアスに関して最適化および補正することができる。断片サイズ対照分子は、既定の断片長を有する合成核酸分子であり得る。断片サイズ対照分子のセットは、特定のステップでポリヌクレオチドの試料に添加された断片サイズ対照分子の少なくとも1つのサブセットを含む核酸分子を含み得る。一部の実施形態では、断片サイズ対照分子のセットは、断片サイズ対照分子の2つまたはそれより多くのサブセットを含み得、各サブセットは、アッセイの異なるステップで添加される。各サブセットは、断片サイズ対照分子の1つまたは複数の群を含み、各群は、異なる断片長またはサイズを有する。例えば、断片サイズバイアスを分析すべきアッセイにおいて3つのステップが存在する場合、断片サイズ対照分子のセットは、断片サイズ対照分子の3つのサブセット(S1、S2、およびS3)を含み得、各サブセットは、各ステップの前に試料に添加される(すなわち、S1はステップ1の前に添加され、S2はステップ2の前に添加され、S3はステップ3の前に添加される)。

40

【0148】

一部の実施形態では、断片サイズ対照分子は、合成核酸分子であり得る。一部の実施形態では、任意のゲノム、プラスミド、もしくはベクター、またはそれらの一部における特定の領域(すなわち、特定の長さの)を増幅することを介して生成されたアンプリコンを、断片サイズ対照分子として使用することができる。断片サイズ対照分子の配列および長さ

50

は、分析の前に公知であり得る。一部の実施形態では、断片サイズ対照分子は、これらの分子がいかなる二次構造も形成しないように設計される。一部の実施形態では、断片サイズ対照分子の配列は、いかなるヒトゲノム領域とも重複しないように設計される。したがって、断片サイズ対照分子をポリヌクレオチドの試料に添加することによって、およびサブセットにおける断片サイズ対照分子を追跡することによって、断片サイズ対照分子の回収率を分析し、それによって一部の実施形態では、断片サイズバイアスを推定することができる。

【0149】

したがって、一態様では、本開示は、ポリヌクレオチドの試料中の核酸分子を分析する方法であって：(a)断片サイズ対照分子のサブセットをポリヌクレオチドの試料中の核酸分子に添加し、それによって第1のスパイクイン試料を生成するステップ；(b)第1のスパイクイン試料から核酸を抽出するステップ；(c)抽出された核酸の少なくともサブセットを処理し、それによって処理された試料を生成するステップであって、処理することが、第1のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；(d)処理された試料の少なくともサブセットを富化し、それによって富化された試料を生成するステップ；(e)富化された試料の少なくともサブセットをシーケンシングして、複数の配列リードを生成するステップ；ならびに(f)複数の配列リードを分析して、断片サイズ対照分子のサブセットの複数の断片サイズスコアを生成するステップを含む方法を提供する。一部の実施形態では、ポリヌクレオチドの試料は、無細胞ポリヌクレオチド（例えば、無細胞DNA）の試料であり得る。

【0150】

一部の実施形態では、方法は、複数の断片サイズスコアを複数の断片サイズ閾値と比較するステップをさらに含む。一部の実施形態では、方法は、複数の断片サイズスコアに基づいてポリヌクレオチドの試料中の核酸分子の分析を最適化するために使用することができる。一部の実施形態では、方法は、複数の断片サイズスコアを使用してポリヌクレオチドの試料中の核酸分子の分析における断片サイズバイアスを補正するために使用することができる。一部の実施形態では、方法は、品質管理（QC）指標を使用することができ、方法は、(i)複数の断片サイズスコアの少なくとも1つが複数の断片サイズ閾値の対応する断片サイズ閾値内にある場合、成功である；または(ii)複数の断片サイズスコアの少なくとも1つが複数の断片サイズ閾値の対応する断片サイズ閾値内にない場合、失敗であると分類される。

【0151】

図1Aは、本開示の実施形態に従う無細胞ポリヌクレオチドの試料中の核酸分子を分析するための方法の概略図である。一部の実施形態では、断片サイズ対照分子のセットは、既定の断片サイズ対照分子の少なくとも1つのサブセットを含み得る。一部の実施形態では、断片サイズ対照分子のセットは、断片サイズ対照分子の1つのサブセットを含み得る。101Aでは、断片サイズ対照分子のサブセット（サブセット1）は、その断片サイズバイアスを分析することができるポリヌクレオチドの試料に添加されて、ポリヌクレオチドの抽出前に第1のスパイクイン試料を生成することができる。一部の実施形態では、断片サイズ対照分子は、サブセット内の各個々の断片サイズ対照分子（分子識別子）および同様にそれが属するサブセット、すなわちサブセット1（サブセット識別子）を識別するために役立ち得る1つまたは複数のタグまたは分子バーコードによってタグ付けされる。

【0152】

一部の実施形態では、断片サイズ対照分子のサブセットは、断片サイズ対照分子の1つまたは複数の群を含み得る。一部の実施形態では、断片サイズ対照分子の1つまたは複数の群は、異なる長さおよび／または異なる配列の核酸分子を含み得る。一部の実施形態では、断片サイズ対照分子の群は、同じ長さおよび同じ配列の断片サイズ対照分子を含み得る。一部の実施形態では、各群は、同じ長さであるが異なる配列の断片サイズ対照分子を含み得る。

【0153】

102 Aでは、第1のスパイクイン試料の核酸分子を抽出する。103 Aでは、抽出された核酸分子を処理して、末端修復、シーケンシングアダプターの付加、タグ付け、および／またはPCR増幅などの、しかしこれらに限定されないステップを含むシーケンシングに適した核酸分子のライブラリを生成する。一部の実施形態では、核酸分子のライブラリを生成するための抽出された核酸分子の処理は、分画化、末端修復、シーケンシングアダプターの付加、タグ付け、洗浄／浄化、および／またはPCR増幅などの、しかしこれらに限定されないステップを含む。104 Aでは、処理された試料中の核酸は、(i) 目的の特定の領域に属する無細胞ポリヌクレオチドの試料中の核酸、および(ii) サブセット1の断片サイズ対照分子に関して富化される。105 Aでは、サブセット1における断片サイズ対照分子の回収率を分析することができるように、富化された試料をシーケンシングする。106 Aでは、シーケンサーから生成された配列リードを分析して、サブセット1断片サイズ対照分子の回収率を測定する。一部の実施形態では、断片サイズ対照分子の回収率は、試料に添加される特定のサブセット内の断片サイズ対照分子の相対的存在量のオルトゴナル測定を利用して計算することができる。一部の実施形態では、オルトゴナル測定は、ゲル電気泳動、qPCR、ddPCR、またはPCRフリーシーケンシングから得ることができる。

10

【0154】

図1 Bは、ポリヌクレオチドの試料（例えば、試料は無細胞DNA試料であり得る）中の核酸分子を分析するための方法100 Bの例としての実施形態を例証する。101 Bでは、断片サイズ対照分子のサブセット（サブセット1）を、抽出ステップの前にポリヌクレオチドの試料に添加して、第1のスパイクイン試料を生成する。断片サイズ対照分子は、ステップのいずれかにおいて導入され得る断片サイズバイアスをモニターするために試料に添加される。一部の実施形態では、断片サイズ対照分子を使用して推定された断片サイズバイアスを使用して、試料中のポリヌクレオチドの計算された断片長分布を最適化することができる。一部の実施形態では、断片サイズ対照分子のセットは、既定の断片サイズ対照分子の少なくとも1つのサブセットを含み得る。一部の実施形態では、断片サイズ対照分子のセットは、断片サイズ対照分子の1つのサブセット（例えば、サブセット1）を含み得る。一部の実施形態では、各サブセットは、断片サイズ対照分子の少なくとも1つの群を含む。一部の実施形態では、各群は、同じ長さおよび同じ配列の断片サイズ対照分子を含む。一部の実施形態では、1つの群の断片サイズ対照分子の長さは、他の群の断片サイズ対照分子の長さとは異なる。一部の実施形態では、各群は、同じ長さであるが異なる配列の断片サイズ対照分子を含む。一部の実施形態では、断片サイズ対照分子は、サブセット内の各個々の断片サイズ対照分子（分子識別子）、および同様にそれが属するサブセット、すなわちサブセット1（サブセット識別子）を識別するために役立ち得る1つまたは複数のタグまたは分子バーコードによってタグ付けされる。一部の実施形態では、断片サイズ対照分子のサブセットの各々は、等モル濃度で存在する。一部の実施形態では、断片サイズ対照分子のサブセットの各々は、等しくないモル濃度で存在する。

20

30

【0155】

102 Bでは、第1のスパイクイン試料の核酸分子を抽出する。103 Bでは、抽出された核酸を処理して処理された試料を生成し、処理された試料は、(i) 無細胞ポリヌクレオチドの試料中の核酸分子、および(ii) サブセット1の断片サイズ対照分子を含む。

40

【0156】

核酸分子は、103 Bにおいて処理されて、シーケンシング（例えば、次世代シーケンサーによる）に適した核酸のライブラリを生成する。処理は、核酸の末端修復、シーケンシングアダプターの付加、タグ付け、洗浄／浄化、および／または増幅などの、しかしこれらに限定されないステップを含み得る。一部の実施形態では、処理は、核酸の分画化、末端修復、シーケンシングアダプターの付加、タグ付け、洗浄／浄化、および／または増幅などの、しかしこれらに限定されないステップを含み得る。

【0157】

一部の実施形態では、分画化は、化学修飾（例えば、メチル化）を有するヌクレオチドを

50

含む核酸分子に優先的に結合する結合剤に対する核酸分子の示差的結合親和性に基づいて核酸分子を分画することを含む。結合剤の例としては、メチル結合ドメイン（MBD）およびメチル結合タンパク質（MBP）が挙げられるがこれらに限定されない。本明細書において企図されるMBPの例としては：

（a）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MeCP2；

（b）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RPL26、PRP8、およびDNAミスマッチ修復タンパク質MHS6；

（c）非修飾シトシンよりも5-ホルミルシトシンに好ましくは結合するFOXK1、FOXK2、FOX P1、FOX P4、およびFOX I3（例えば、その全体が参照により本明細書に組み込まれる、Iurlaro et al., Genome Biol. 14, R119（2013）に記載される）；ならびに

（d）1つまたは複数のメチル化ヌクレオチド塩基に対して特異的な抗体が挙げられるがこれらに限定されない。

【0158】

分画化は、核酸分子の特徴に基づく核酸分子の分離または分別を指し得る。分画化は、分子の物理的分画化であり得る。分画化は、エピジェネティック修飾（例えば、エピジェネティック状態）のレベルに基づいて核酸分子を群またはセットに分離することを伴い得る。例えば、核酸分子を、核酸分子のメチル化のレベルに基づいて分画することができる。一部の実施形態では、分画化のために使用される方法およびシステムは、その全体が参照により本明細書に組み込まれる、PCT特許出願第PCT/US2017/068329号に見出され得る。それらの実施形態では、核酸は、メチル化の異なるレベル（メチル化ヌクレオチドの異なる数）に基づいて分画される。一部の実施形態では、核酸は、2つまたはそれより多くの分画化セット（例えば、少なくとも3、4、5、6、または7つの分画化セット）に分画することができる。一部の実施形態では、分画化セットは、異なる程度の修飾を有する（修飾の過剰提示または過小提示）核酸を表す。過剰提示および過小提示は、集団における鎖当たりの修飾の数の中央値と比較して核酸が有する修飾の数によって定義することができる。例えば、試料中の核酸分子における5-メチルシトシンヌクレオチドの数の中央値が2である場合、2つより多くの5-メチルシトシン残基を含む核酸分子は、この修飾で過剰提示され、1またはゼロ個の5-メチルシトシン残基を有する核酸分子は、過小提示される。親和性分離の効果は、結合相中の修飾で過剰提示される核酸および非結合相中（すなわち、溶液中）の修飾で過小提示される核酸を富化することである。結合相中の核酸は、その後の処理の前に溶出させることができる。一部の実施形態では、複数の分画化セットの各々は、示差的にタグ付けされる。次いで、タグ付けされた分画化セットを集成的な試料調製、富化、および／またはシーケンシングのために共にプールする。分画化セットの示差的タグ付けは、特定の分画化セットに属する核酸分子を把握するために役立つ。タグは、アダプターの構成成分として提供され得る。異なる分画化セットにおける核酸分子は、1つの分画化セットのメンバーを別のメンバーと区別することができる異なるタグを与えられる。同じ分画化セットの核酸分子に連結されたタグは、互いに同じまたは異なり得る。しかし、互いに異なる場合であっても、タグは、それらが付着する分子が特定の分画化セットの分子であると識別されるように、その配列の一部を共通に有し得る。例えば、スパイクイン試料の分子が2つの分画化セット、P1およびP2に分画されている場合には、P1内の分子をA1、A2、A3などでタグ付けすることができ、P2内の分子にB1、B2、B3などでタグ付けすることができる。そのようなタグ付けシステムは、分画化セットおよび分画化セット内の分子間を区別することを可能にする。

【0159】

104Bでは、処理された試料を富化して、富化された試料を生成する。富化された試料は、（i）目的の特定の領域に属する無細胞ポリヌクレオチドの試料中の核酸分子、および（ii）サブセット1の断片サイズ対照分子を含む。105Bでは、断片サイズバイアスを計算するために、富化された試料をシーケンシングして複数の配列リードを生成し、

10

20

30

40

50

サブセット 1 における断片サイズ対照分子の回収率を分析することができる。得られた配列情報は、核酸分子および核酸分子に付着したタグの配列を含む。断片サイズ対照分子に付着したタグの配列から、タグを、個々の断片サイズ対照分子および断片サイズ対照分子が属するサブセットと関連させることができる。この情報を使用して、サブセットにおける断片サイズ対照分子の回収率を分析する。

【0160】

106Bでは、配列リードを分析して、断片サイズ対照分子の断片サイズスコアを生成する。断片サイズスコアは、特定の群および特定のサブセットに属する断片サイズ対照分子の回収率を表す。断片サイズ対照分子および断片サイズ対照分子が属するサブセットの同一性は、タグまたはバーコードの使用を通して維持される。一部の実施形態では、断片サイズスコアは、特定の群および特定のサブセットに属する断片サイズ対照分子の数に基づいて推定することができる。一部の実施形態では、断片サイズスコアは、特定の群および特定のサブセットに属する断片サイズ対照分子のシーケンシングリードの数に基づいて推定することができる。一部の実施形態では、断片サイズスコアは、特定のサブセットにおける特定の群に関して推定することができる。一部の実施形態では、断片サイズスコアは、あらゆるサブセットにおける群の各々に関して推定することができる。一部の実施形態では、断片サイズスコアは、特定のサブセットにおいて使用される断片サイズ対照分子の総スコア（すなわち、サブセット内の断片サイズ対照分子を表す全ての単一スコアに関して）、またはアッセイの1つもしくは複数のステップで添加された断片サイズ対照分子の総スコア（すなわち、アッセイで添加した1つまたは複数のサブセットにおける断片サイズ対照分子を表す単一スコア）であり得る。一部の実施形態では、断片サイズスコアは、特定のステップで添加された断片サイズ対照分子のサブセットの量と、異なるステップで添加された断片サイズ対照分子の別のサブセットの量との差として測定することができる。一部の実施形態では、量は、質量（例えば、fg、pg、ng、μg）、シーケンシングリードの数または断片サイズ分子の数（サブセットに属する、または特定の群に属する）のいずれかに関する量であり得る。一部の実施形態では、断片サイズスコアは、断片サイズ対照分子（サブセットに属する、または特定の群に属する）の回収率（画分または百分率）として測定することができる。回収率は、試料に添加された特定のサブセット内の断片サイズ対照分子の相対的存在量のオルトゴナル測定を利用して計算することができる。一部の実施形態では、オルトゴナル測定は、ゲル電気泳動、qPCR、ddPCR、またはPCRフリーシーケンシングから得ることができる。一部の実施形態では、断片サイズスコアは、特定の長さの断片サイズ対照分子の量と、異なる長さの断片サイズ対照分子の量との差として測定することができる。

【0161】

107Bでは、断片サイズスコアを、アッセイにおける断片サイズバイアスを決定するために対応する断片サイズ閾値と比較する。断片サイズ閾値は、試料中の核酸分子の分析における断片サイズバイアスを評価または最適化するために使用される既定の閾値の値または範囲である。これらの閾値を使用して、抽出、ライブラリ調製、富化、洗浄／浄化、およびシーケンシングなどのステップのいずれかにおける任意の断片長分布を補正することもできる。一部の実施形態では、断片サイズ閾値は、特定のサブセットにおける特定の群に関して推定することができる。一部の実施形態では、サブセットにおける各群は、特定の断片サイズ閾値を有する。一部の実施形態では、断片サイズ閾値は、あらゆるサブセットにおける群の各々に関して推定することができる。一部の実施形態では、断片サイズ閾値は、特定のサブセットにおいて使用される断片サイズ対照分子の総閾値（すなわち、サブセット内の断片サイズ対照分子の全ての単一閾値に関して）、またはアッセイの1つもしくは複数のステップで添加された断片サイズ対照分子の総閾値（すなわち、アッセイで添加した1つまたは複数のサブセットにおける断片サイズ対照分子の単一閾値）であり得る。例えば、断片サイズ対照分子のセットは、2つのサブセット、サブセット1およびサブセット2を含む。各サブセットが2つの群、G11およびG12（サブセット1に関して）ならびにG21およびG22（サブセット2に関して）を含む場合には、4つの群の

各々は、断片サイズ対照分子の回収率に基づいて断片サイズスコア（S）、G 1 1に関してS 1 1、G 1 2に関してS 1 2、G 2 1に関してS 2 1、およびG 2 2に関してS 2 2を有し得る。サブセットにおける各群は、個別の既定の断片サイズ閾値を有し得る。この例では、T 1 1、T 1 2、T 2 1、およびT 2 2はそれぞれ、G 1 1、G 1 2、G 2 1、およびG 2 2の断片サイズ閾値である。閾値は、百分率または画分に関する閾値であり得、閾値は、特定の閾値の値の代わりに閾値の範囲であり得る。方法が成功であると考えられるためには、これらの断片サイズスコアの少なくとも1つは、その対応する断片サイズ閾値内であればならない。一部の実施形態では、複数の断片サイズスコアの各々が対応する断片サイズ閾値内にある場合には、方法は実施に成功したと分類される。一部の実施形態では、断片サイズ閾値は、百分率または画分に関する閾値であり得る。一部の実施形態では、断片サイズ閾値は、特定の閾値の値の代わりに閾値の範囲であり得る。

10

【0162】

一部の実施形態では、103Bの前に、断片サイズ対照分子の第2のサブセット（サブセット2）を抽出された核酸に添加して、第2のスパイクイン試料を生成してもよい。これらの実施形態では、第2のスパイクイン試料を処理して、処理された試料を生成する。処理された試料は、（i）無細胞ポリヌクレオチドの試料中の核酸分子、ならびに（ii）サブセット1およびサブセット2の断片サイズ対照分子を含む。一部の実施形態では、104Bの前に、断片サイズ対照分子の第3のサブセット（サブセット3）を処理された試料に添加して、第3のスパイクイン試料を生成してもよい。これらの実施形態では、第3のスパイクイン試料を富化して、富化された試料を生成する。富化された試料は、（i）

20

【0163】

別の態様では、本開示は、ポリヌクレオチドの試料中の核酸分子を分析するための方法であって：（a）断片サイズ対照分子の第1のサブセットをポリヌクレオチドの試料中の核酸分子に添加し、それによって第1のスパイクイン試料を産生するステップ；（b）第1のスパイクイン試料から核酸を抽出するステップ；（c）断片サイズ対照分子の第2のサブセットを抽出された核酸に添加し、それによって第2のスパイクイン試料を産生するステップ；（d）第2のスパイクイン試料の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、処理することが、第2のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；（e）断片サイズ対照分子の第3のサブセットを処理された試料に添加し、それによって第3のスパイクイン試料を産生するステップ；（f）第3のスパイクイン試料の少なくともサブセットを富化し、それによって富化された試料を産生するステップ；（g）断片サイズ対照分子の第4のサブセットを富化された試料の少なくともサブセットに添加し、それによって第4のスパイクイン試料を産生するステップ；（h）第4のスパイクイン試料をシーケンシングして、複数の配列リードを生成するステップ；（i）複数の配列リードを分析して、断片サイズ対照分子の第1のサブセット、断片サイズ対照分子の第2のサブセット、断片サイズ対照分子の第3のサブセット、および／または断片サイズ対照分子の第4のサブセットの複数の断片サイズスコアを生成するステップ；ならびに（j）複数の断片サイズスコアを複数の断片サイズ閾値と比較するステップを含む方法を提供する。

30

40

【0164】

一部の実施形態では、方法は、複数の断片サイズスコアを複数の断片サイズ閾値と比較す

50

るステップをさらに含む。一部の実施形態では、方法は、複数の断片サイズスコアに基づいてポリヌクレオチドの試料中の核酸分子の分析を最適化するために使用することができる。一部の実施形態では、方法は、複数の断片サイズスコアを使用してポリヌクレオチドの試料中の核酸分子の分析における断片サイズバイアスを補正するために使用することができる。一部の実施形態では、方法は、品質管理（QC）指標を使用することができ、方法は、（i）複数の断片サイズスコアの少なくとも1つが複数の断片サイズ閾値の対応する断片サイズ閾値内にある場合、成功である；または（ii）複数の断片サイズスコアの少なくとも1つが複数の断片サイズ閾値の対応する断片サイズ閾値内にない場合、失敗であると分類される。一部の実施形態では、ポリヌクレオチドの試料は、無細胞ポリヌクレオチド（例えば、無細胞DNA）の試料であり得る。

10

【0165】

図2Aは、本開示の実施形態に従うポリヌクレオチド（例えば、無細胞ポリヌクレオチド）の試料中の核酸分子を分析するための方法の概略図である。一部の実施形態では、断片サイズ対照分子のセットは、既定の断片サイズ対照分子の少なくとも1つのサブセットを含み得る。201Aでは、断片サイズ対照分子の第1のサブセット（サブセット1）を、その断片サイズバイアスを分析すべきであるポリヌクレオチドの試料に添加して、ポリヌクレオチドの抽出前に第1のスパイクイン試料を生成する。一部の実施形態では、断片サイズ対照分子は、サブセット内の各個々の断片サイズ対照分子（分子識別子）および同様にそれが属するサブセット、すなわちサブセット1（サブセット識別子）を識別するために役立ち得る1つまたは複数のタグまたはバーコードによってタグ付けされる。

20

【0166】

一部の実施形態では、断片サイズ対照分子のサブセットは、断片サイズ対照分子の1つまたは複数の群を含み得る。一部の実施形態では、断片サイズ対照分子の1つまたは複数の群は、異なる長さおよび／または異なる配列の核酸分子を含み得る。一部の実施形態では、断片サイズ対照分子の群は、同じ長さおよび同じ配列の断片サイズ対照分子を含み得る。一部の実施形態では、各群は、同じ長さであるが異なる配列の断片サイズ対照分子を含み得る。

【0167】

202Aでは、第1のスパイクイン試料の無細胞核酸分子を抽出する。203Aでは、断片サイズ対照分子の第2のサブセット（サブセット2）を抽出された核酸分子に添加して、第2のスパイクイン試料を生成する。204Aでは、第2のスパイクイン試料を処理し、末端修復、シーケンシングアダプターの付加、タグ付け、洗浄／浄化、および／またはPCR増幅などの、しかしこれらに限定されないステップを含むシーケンシングに適した核酸分子のライブラリを生成する。一部の実施形態では、核酸分子のライブラリを生成するための第1のスパイクイン試料の処理は、分画化、末端修復、シーケンシングアダプターの付加、タグ付け、洗浄／浄化、および／またはPCR増幅などの、しかしこれらに限定されないステップを含む。205Aでは、断片サイズ対照分子の第3のサブセット（サブセット3）を処理された試料に添加して、第3のスパイクイン試料を生成する。206Aでは、第3のスパイクイン試料中の核酸を、（i）目的の特定の領域に属する無細胞ポリヌクレオチドの試料中の核酸分子、ならびに（ii）サブセット1、サブセット2、およびサブセット3の断片サイズ対照分子に関して富化する。207Aでは、断片サイズ対照分子の第4のサブセット（サブセット4）を富化された試料に添加して、第4のスパイクイン試料を生成する。208Aでは、断片サイズバイアスを決定するために、第4のスパイクイン試料をシーケンシングして、サブセット1、サブセット2、サブセット3、およびサブセット4における断片サイズ対照分子の回収率を分析することができる。209Aでは、シーケンサーから生成された配列リードを分析して、サブセット1、サブセット2、サブセット3、およびサブセット4断片サイズ対照分子の回収率を測定する。

30

40

【0168】

図2Bは、ポリヌクレオチド（例えば、無細胞ポリヌクレオチド）の試料中の核酸分子を分析するための方法200Bの例としての実施形態を例証する。201Bでは、断片サイ

50

ズ対照分子のサブセット（サブセット 1）を、抽出ステップの前に無細胞ポリヌクレオチドの試料に添加して、第 1 のスパイクイン試料を生成する。断片サイズ対照分子を、ステップのいずれかで導入された断片サイズバイアスをモニターおよび／または補正するために、試料に添加する。一部の実施形態では、断片サイズ対照分子のセットは、既定の断片サイズ対照分子の少なくとも 1 つのサブセットを含み得る。一部の実施形態では、断片サイズ対照分子のセットは、断片サイズ対照分子の 1 つのサブセット（例えば、サブセット 1）を含み得る。一部の実施形態では、各サブセットは、断片サイズ対照分子の少なくとも 1 つの群を含む。一部の実施形態では、各群は、同じ長さおよび同じ配列の断片サイズ対照分子を含む。一部の実施形態では、1 つの群における断片サイズ対照分子の長さは、他の群における断片サイズ対照分子の長さとは異なる。一部の実施形態では、各群は、同じ長さであるが異なる配列の断片サイズ対照分子を含む。一部の実施形態では、断片サイズ対照分子は、サブセット内の各個々の断片サイズ対照分子（分子識別子）および同様にそれが属するサブセット、すなわちサブセット 1（サブセット識別子）を識別するために役立ち得る 1 つまたは複数のタグまたはバーコードによってタグ付けされる。一部の実施形態では、断片サイズ対照分子のサブセットの各々は、等モル濃度で存在する。一部の実施形態では、断片サイズ対照分子のサブセットの各々は、等しくないモル濃度で存在する。

10

【0169】

202B では、第 1 のスパイクイン試料の核酸分子を抽出する。203B では、断片サイズ対照分子の第 2 のサブセット（サブセット 2）を抽出された核酸分子に添加して、第 2 のスパイクイン試料を生成する。一部の実施形態では、第 1 のサブセットにおける断片サイズ対照分子は、異なるタグを使用することによって第 2 のサブセットにおける断片サイズ対照分子から識別され、すなわちサブセット 1 における断片サイズ対照分子のタグ（サブセット識別子）は、サブセット 2 における断片サイズ対照分子のタグとは異なる。例えば、サブセット 1 における断片サイズ対照分子は全て、サブセット識別子タグ（S1）を有し得、サブセット 2 における断片サイズ対照分子は全て、サブセット識別子タグ（S2）を有し得る。サブセット識別子タグとは別に、各断片サイズ対照分子は、各個々の断片サイズ対照分子を識別するために使用される分子識別子タグを含む。

20

【0170】

204B では、第 2 のスパイクイン試料を処理して、処理された試料を生成し、処理された試料は、(i) 無細胞ポリヌクレオチドの試料中の核酸分子、ならびに (ii) サブセット 1 およびサブセット 2 の断片サイズ対照分子を含む。第 2 のスパイクイン試料を 204B において処理して、シーケンシング（例えば、次世代シーケンサーによる）に適した核酸のライブラリを生成する。処理は、核酸の末端修復、シーケンシングアダプターの付加、タグ付け、洗浄／浄化、および／または増幅などの、しかしこれらに限定されないステップを含み得る。

30

【0171】

一部の実施形態では、処理は、核酸の分画化、末端修復、シーケンシングアダプターの付加、タグ付け、洗浄／浄化、および／または増幅などの、しかしこれらに限定されないステップを含み得る。一部の実施形態では、分画化は、化学修飾（例えば、メチル化）を有するヌクレオチドを含む核酸に優先的に結合する結合剤に対する核酸分子の示差的な結合親和性に基づいて核酸分子を分画することを含む。一部の実施形態では、核酸を、2 つまたはそれより多くの分画化セット（例えば、少なくとも 3、4、5、6、または 7 つの分画化セット）に分画することができる。一部の実施形態では、複数の分画化セットの各々は、複数の分画化セットの第 1 の分画化セットにおいて使用したタグのセットが、複数の分画化セットの第 2 の分画化セットにおいて使用したタグのセットとは異なるように、示差的にタグ付けされる。次いで、タグ付けされた分画化セットを、集合的な試料調製、富化、および／またはシーケンシングのために共にプールする。分画化セットの示差的タグ付けは、特定の分画化セットに属する核酸分子を把握するために役立つ。タグは、一部の実施形態では、アダプターの構成成分として提供され得る。異なる分画化セットにおける

40

50

核酸分子は、1つの分画化セットのメンバーを別のメンバーと区別することができる異なるタグを与えられる。同じ分画化セットの核酸分子に連結されたタグは、互いに同じまたは異なり得る。しかし、互いに異なる場合であっても、タグは、それらが付着する分子が特定の分画化セットの分子であると識別されるように、その配列の一部を共通に有し得る。

【0172】

一部の実施形態では、タグ付けすることは、タグのセットを核酸に付着させて、タグ付けされた核酸の集団を生成することであって、タグ付けされた核酸が1つまたは複数のタグを含む、ことを含む。一部の実施形態では、タグのセットを、アダプターの核酸へのライゲーションによって核酸に付着させ、アダプターは1つまたは複数のタグを含む。

【0173】

205Bでは、断片サイズ対照分子の第3のサブセット（サブセット3）を処理された試料に添加して、第3のスパイクイン試料を生成する。206Bでは、第3のスパイクイン試料を富化して、富化された試料を生成する。富化された試料は、(i) 目的の特定の領域に属する無細胞ポリヌクレオチドの試料中の核酸分子、ならびに(ii) サブセット1、サブセット2、およびサブセット3の断片サイズ対照分子を含む。207Bでは、断片サイズ対照分子の第4のサブセット（サブセット4）を富化された試料に添加して、第4のスパイクイン試料を生成する。208Bでは、断片サイズバイアスを計算するために、第4のスパイクイン試料をシーケンシングして、複数の配列リードを生成し、サブセット1、サブセット2、サブセット3、およびサブセット4における断片サイズ対照分子の回収率を分析することができる。得られた配列情報は、核酸分子および核酸分子に付着したタグの配列を含む。断片サイズ対照分子に付着したタグの配列から、タグを、個々の断片サイズ対照分子および分子が属するサブセットと関連させることができる。この情報を使用して、サブセットにおける断片サイズ対照分子の回収率を分析する。

【0174】

209Bでは、配列リードを分析して、断片サイズ対照分子の断片サイズスコアを生成する。断片サイズスコアは、特定の群および特定のサブセットに属する断片サイズ対照分子の回収率を表す。一部の実施形態では、断片サイズ対照分子の回収率は、試料に添加された特定のサブセット内の断片サイズ対照分子の相対的存在量のオルトゴナル測定を利用して計算することができる。一部の実施形態では、オルトゴナル測定は、ゲル電気泳動、qPCR、ddPCR、またはPCRフリーシーケンシングから得ることができる。断片サイズ対照分子、ならびに断片サイズ対照分子が属する群およびサブセットの同一性は、タグまたはバーコードの使用を通して維持される。一部の実施形態では、断片サイズスコアは、特定の群および特定のサブセットに属する断片サイズ対照分子の数に基づいて推定することができる。一部の実施形態では、断片サイズスコアは、特定の群および特定のサブセットに属する断片サイズ対照分子のシーケンシングリードの数に基づいて推定することができる。一部の実施形態では、断片サイズスコアは、特定のサブセットにおける特定の群に関して推定することができる。一部の実施形態では、断片サイズスコアは、あらゆるサブセットにおける群の各々に関して推定することができる。一部の実施形態では、断片サイズスコアは、特定のサブセットにおいて使用される断片サイズ対照分子の総スコア（すなわち、サブセット内の断片サイズ対照分子を表す単一スコア）、またはアッセイの1つもしくは複数のステップに添加された断片サイズ対照分子の総スコア（すなわち、アッセイで添加した1つまたは複数のサブセットにおける断片サイズ対照分子を表す単一スコア）であり得る。一部の実施形態では、断片サイズスコアは、特定のステップで添加された断片サイズ対照分子のサブセットの量と、異なるステップで添加された断片サイズ対照分子の別のサブセットの量との差として測定することができ、一部の実施形態では、量は、質量（例えば、fg、pg、ng、μg）、シーケンシングリードの数または断片サイズ分子の数（サブセットに属する、または特定の群に属する）のいずれかに関する量であり得る。一部の実施形態では、断片サイズスコアは、断片サイズ対照分子（サブセットに属する、または特定の群に属する）の回収率（画分または百分率）として測定することができ、回収率は、試料に添加された特定のサブセット内の断片サイズ対照分子の相対的存

10

20

30

40

50

在量のオルトゴナル測定を利用して計算することができる。一部の実施形態では、オルトゴナル測定は、ゲル電気泳動、qPCR、ddPCR、またはPCRフリーシーケンシングから得ることができる。一部の実施形態では、断片サイズスコアは、特定の長さの断片サイズ対照分子の量と、異なる長さの断片サイズ対照分子の量との差として測定することができる。

【0175】

210Bでは、断片サイズスコアを、アッセイにおける断片サイズバイアスを決定するために、対応する断片サイズ閾値と比較する。断片サイズ閾値は、試料中の核酸分子の分析における断片サイズバイアスを評価または最適化するために使用される既定の閾値の値または範囲である。これらの閾値を使用して、抽出、ライブラリ調製、富化、洗浄／浄化、およびシーケンシングなどのステップのいずれかにおける任意の断片長分布を補正することもできる。一部の実施形態では、断片サイズ閾値は、特定のサブセットにおける特定の群に関して推定することができる。一部の実施形態では、サブセットにおける各群は、特定の断片サイズ閾値を有する。一部の実施形態では、断片サイズ閾値は、あらゆるサブセットにおける群の各々に関して推定することができる。一部の実施形態では、断片サイズ閾値は、特定のサブセットにおいて使用される断片サイズ対照分子の総閾値（すなわち、サブセット内の断片サイズ対照分子の単一閾値）、またはアッセイの1つもしくは複数のステップで添加された断片サイズ対照分子の総閾値（すなわち、アッセイで添加した1つまたは複数のサブセットにおける断片サイズ対照分子の単一閾値）であり得る。例えば、断片サイズ対照分子のセットは、2つのサブセット、サブセット1およびサブセット2を含む。各サブセットが2つの群、G11およびG12（サブセット1に関して）ならびにG21およびG22（サブセット2に関して）を含む場合には、4つの群の各々は、断片サイズ対照分子の回収率に基づいて断片サイズスコア（S）、G11に関してS11、G12に関してS12、G21に関してS21、およびG22に関してS22を有し得る。サブセットにおける各群は、個別の既定の断片サイズ閾値を有し得る。この例では、T11、T12、T21、およびT22はそれぞれ、G11、G12、G21、およびG22の断片サイズ閾値である。閾値は、百分率または画分に関する閾値であり得、閾値は、特定の閾値の値の代わりに閾値の範囲であり得る。方法が成功であると考えられるためには、これらの断片サイズスコアの少なくとも1つは、その対応する断片サイズ閾値内であればならない。一部の実施形態では、複数の断片サイズスコアの各々が対応する断片サイズ閾値内にある場合には、方法は実施に成功したと分類される。一部の実施形態では、断片サイズ閾値は、百分率または画分に関する閾値であり得る。一部の実施形態では、断片サイズ閾値は、特定の閾値の値の代わりに閾値の範囲であり得る。

【0176】

一部の実施形態では、断片サイズ対照分子を使用して推定される断片サイズバイアスに関する核酸分子の分析を使用して、試料中のポリヌクレオチドの計算された断片長分布を最適化することができる。一部の実施形態では、複数の断片サイズスコアを使用して、核酸分子の分析における断片サイズバイアスを補正する。一部の実施形態では、断片サイズバイアスを補正するステップは、断片サイズスコアの、サブセット内の群の断片サイズ閾値に対する比に由来する定量的測定を使用するステップ、ならびに／またはサブセット内の2つの群の断片サイズスコアの比および／もしくは2つのサブセットの断片サイズスコアの比に由来する定量的測定を使用するステップを含む。一部の実施形態では、方法は、（i）複数の断片サイズスコアの少なくとも1つが複数の断片サイズ閾値の対応する断片サイズ閾値内にある場合、成功である；または（ii）複数の断片サイズスコアの少なくとも1つが複数の断片サイズ閾値の対応する断片サイズ閾値内にない場合、失敗であると分類される。

【0177】

一部の実施形態では、全てのサブセットにおける全ての群に関する断片サイズ対照分子の断片サイズスコアが、全ての群に関する対応する断片サイズ閾値内にある場合には、核酸分子を分析する方法は成功であると分類され得る。それ以外の場合、断片サイズスコアの

いずれか1つがその対応する断片サイズ閾値の外側である場合、核酸分子を分析する方法は失敗であると分類され得る。例えば、断片サイズ対照分子のセットは、断片サイズ対照分子の2つのサブセット、サブセット1およびサブセット2を含む。各サブセットが2つの群、G11およびG12（サブセット1に関して）ならびにG21およびG22（サブセット2に関して）を含む場合には、4つの群の各々は、断片サイズ対照分子の回収率に基づいて断片サイズスコア（S）、G11に関してS11、G12に関してS12、G21に関してS21、およびG22に関してS22を有し得る。サブセットにおける各群は、個別の断片サイズ閾値を有する。この例では、T11、T12、T21、およびT22はそれぞれ、G11、G12、G21、およびG22の断片サイズ閾値である。S11 < T11、S12 < T12、S21 < T21、およびS22 < T22である場合には、方法は成功であると分類され得る。断片サイズスコアのいずれか1つがその対応する断片サイズ閾値の外側である（すなわち、内部にない）場合には、方法は失敗であると分類され得る。

10

【0178】

一部の実施形態では、断片サイズ対照分子は、真の生物学的原因ではなく技術的原因（人為的バイアス）に起因する観察された断片長分布における断片サイズバイアスを補正するために、個々の試料内で使用され得る。例えば、腫瘍細胞に由来するcfDNA分子は、典型的に造血細胞に由来する断片よりも短い平均長を示すことから、断片長分布の平均値、中央値、最頻値、およびIQRを含むがこれらに限定されない要約統計量を、単独または他の特色との組合せのいずれかで悪性腫瘍の検出のための特色として使用してもよい。しかし、これらの要約統計量は、技術的要因、輸送、試料および液体取り扱い、ならびにステップのいずれか（例えば、抽出、分画化、タグ付け、増幅、洗浄／浄化、富化およびシーケンシング）中に導入される人為的バイアスによって影響を受け得、これらの影響は、悪性腫瘍の検出を交絡させ得る。この潜在的交絡源を回避するために、観察された断片長分布を、異なる長さの断片サイズ対照分子の相対的回収率に基づいて調整してもよい。例えば、特定の長さの断片サイズ対照分子の予想される回収率が75%であるが、所定の試料中のその特定の長さの断片サイズ対照分子の観察された回収率がわずか25%である場合、断片サイズ対照分子の長さに対応する試料ポリヌクレオチドの断片長分布の構成成分の密度を3倍増加させて、低い回収率を補正してもよい。逆に、特定の長さの断片サイズ対照分子の予想される回収率が60%であるが、所定の試料中のその特定の長さの断片サイズ対照分子の観察された回収率が80%である場合、試料ポリヌクレオチドの断片長分布の密度を、断片サイズ対照分子の長さに対応して25%減少させてもよい。

20

30

【0179】

観察された断片長分布に対する調整は、回収された断片サイズ対照分子の長さ周囲のウィンドウで起こり得る。例えば、断片サイズ対照分子が200塩基対の長さを有する場合、観察された長さの分布密度を、180bp～220bp、または170bp～230bpの間の長さの試料ポリヌクレオチドに対して調整してもよい。ウィンドウは、対称性である必要はなく、適用される調整は、ウィンドウ内に入るポリヌクレオチドのあらゆる長さに関して同じである必要はない。一例として、適用される調整は、均一なカーネルではなく、断片サイズ対照分子の長さを中心とするガウスカーネルを使用し得る。

40

【0180】

異なる長さの断片サイズ対照分子の予想される回収率は、1つまたは複数の対照実験を実施することによって決定することができる。断片サイズ閾値を、特定の長さの任意の断片サイズ対照分子の回収率に関して設定してもよい。断片サイズ対照分子の観察された回収率が、断片サイズ閾値の値より下であるか、または断片サイズ閾値の範囲内に入らない場合、試料は品質管理不良であると考えられ得る。これらの断片サイズ閾値は、異なる長さの断片サイズ対照分子に関して異なり得る。

【0181】

別の態様では、本開示は、無細胞ポリヌクレオチドの試料のシーケンシングライブラリを産生するための方法であって：（a）断片サイズ対照分子のサブセットを試料に添加し、

50

それによって第1のスパイクイン試料を産生するステップ；（b）第1のスパイクイン試料から核酸を抽出するステップ；（c）抽出された核酸の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、処理することが、第1のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；ならびに（d）処理された試料の少なくともサブセットを富化するステップを含む方法を提供する。一部の実施形態では、方法は、処理することの前に、断片サイズ対照分子の第2のサブセットを添加し、それによって第2のスパイクイン試料を産生するステップをさらに含む。一部の実施形態では、方法は、富化するステップの前に、断片サイズ対照分子の第3のサブセットを添加し、それによって第3のスパイクイン試料を産生するステップをさらに含む。一部の実施形態では、方法は、e）断片サイズ対照分子の第4のサブセットを添加し、それによって第4のスパイクイン試料を産生するステップをさらに含む。

10

【0182】

別の態様では、本開示は、第1の試料の第2の試料による汚染を検出するための方法であって、各第1の試料および第2の試料に関して：（a）断片サイズ対照分子のサブセットを添加して、第1のスパイクイン試料を生成するステップであって、第1の試料に添加された断片サイズ対照分子のサブセットが、第2の試料に添加された断片サイズ対照分子のサブセットと区別され得る、ステップ；（b）第1のスパイクインから核酸を抽出するステップ；（c）抽出された核酸の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、処理することが、第1のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；（d）処理された試料の少なくともサブセットを富化するステップ；（e）富化された試料の少なくともサブセットをシーケンシングして、複数の配列リードを生成するステップ；ならびに（f）複数の配列リードを分析して、断片サイズ対照分子のサブセットの1つまたは複数の汚染スコアを生成するステップを含む方法を提供する。一部の実施形態では、方法は、処理することの前に、断片サイズ対照分子の第2のサブセットを添加し、それによって第2のスパイクイン試料を産生するステップであって、第1の試料に添加された断片サイズ対照分子のサブセットが、第2の試料に添加された断片サイズ対照分子のサブセットと区別され得る、ステップをさらに含む。一部の実施形態では、方法は、富化するステップの前に、断片サイズ対照分子の第3のサブセットを添加し、それによって第3のスパイクイン試料を産生するステップであって、第1の試料に添加された断片サイズ対照分子のサブセットが、第2の試料に添加された断片サイズ対照分子のサブセットと区別され得る、ステップをさらに含む。一部の実施形態では、方法は、シーケンシングするステップの前に、断片サイズ対照分子の第4のサブセットを添加し、それによって第4のスパイクイン試料を産生するステップであって、第1の試料に添加された断片サイズ対照分子のサブセットが、第2の試料に添加された断片サイズ対照分子のサブセットと区別され得る、ステップをさらに含む。一部の実施形態では、第1の試料に添加された断片サイズ対照分子のサブセットは、他の試料において使用されるサブセット識別子バーコードとは異なる第1の試料中のサブセット識別子バーコードを使用することによって、第2の試料に添加された断片サイズ対照分子のサブセットと区別することができる。

20

30

40

【0183】

汚染スコアは、第2の試料に添加された断片サイズ対照分子の存在を表す第1の試料におけるスコアを指す。一部の実施形態では、1つの試料中の1つのサブセットにおいて使用されるサブセット識別子バーコードは、他の試料中の他のサブセットにおいて使用されるサブセット識別子とは異なり得る。これらの実施形態では、断片サイズ対照分子に存在するサブセット識別子バーコードの配列から、異なる第2の試料に属する断片サイズ対照分子の存在を識別することができる。一部の実施形態では、汚染スコアは、サブセットにおいて使用される断片サイズ対照分子の各型／群に対して特異的であり得るか（すなわち、サブセットにおいて使用される断片サイズ対照分子の異なる長さ／群に関して個別の汚染スコア）、または断片サイズ対照分子の異なる長さ／群を表す総スコアであり得る。一部

50

の実施形態では、汚染スコアは、異なる第2の試料に属する断片サイズ対照分子の数に基づいて推定することができる。一部の実施形態では、汚染スコアは、異なる第2の試料に属する断片サイズ対照分子のシーケンシングリードの数に基づいて推定することができる。一部の実施形態では、汚染スコアは、異なる第2の試料に属する断片サイズ対照分子のシーケンシングリードの数に基づいて推定することができる。一部の実施形態では、試料の汚染スコアは、他の試料に属する断片サイズ対照分子のシーケンシングリードの数の、その試料に添加された断片サイズ対照分子のシーケンシングリードの数に対する画分または百分率に基づいて推定することができる。一部の実施形態では、試料の汚染スコアは、他の試料に属する断片サイズ対照分子の分子の数の、その試料に添加された断片サイズ対照分子の分子の数に対する画分または百分率に基づいて推定することができる。これらの実施形態では、その分子に添加された断片サイズ対照分子は、断片サイズ対照分子のサブセット識別子バーコードから識別することができる。

10

【0184】

一部の実施形態では、方法は、少なくとも1つまたは複数の汚染スコアを少なくとも1つまたは複数の汚染閾値と比較するステップをさらに含む。汚染閾値は、試料中の核酸分子の分析における試料の汚染を評価するために使用される既定の閾値の値または範囲を指す。これらの閾値を使用して、抽出、ライブラリ調製、富化、洗浄／浄化、およびシーケンシングなどのステップのいずれかにおいてアッセイを最適化することもできる。一部の実施形態では、汚染閾値は、サブセットにおいて使用される断片サイズ対照分子の各型／群に対して特異的であり得る（すなわち、サブセットにおいて使用される断片サイズ対照分子の異なる長さ／群に関する個別の汚染閾値）。一部の実施形態では、汚染閾値は、サブセットにおける断片サイズ対照分子の異なる長さ／群の総閾値、またはアッセイで添加した1つもしくは複数のサブセットにおいて使用される断片サイズ対照分子の総閾値であり得る。一部の実施形態では、サブセットにおける各群は、特定の汚染閾値を有する。例えば、断片サイズ対照分子のセットは、2つのサブセット、サブセット1およびサブセット2を含む。各サブセットが2つの群、G11およびG12（サブセット1に関して）ならびにG21およびG22（サブセット2に関して）を含む場合には、4つの群の各々は、断片サイズ対照分子の回収率に基づいて汚染スコア（S）、G11に関してS11、G12に関してS12、G21に関してS21、およびG22に関してS22を有し得る。サブセットにおける各群は、個別の既定の汚染閾値を有し得る。この例では、T11、T12、T21、およびT22はそれぞれ、G11、G12、G21、およびG22の汚染閾値である。閾値は、百分率または画分に関する閾値であり得、閾値は、特定の閾値の値の代わりに閾値の範囲であり得る。方法が成功であると考えられるためには、これらの汚染スコアの少なくとも1つは、その対応する汚染閾値内でなければならない。一部の実施形態では、複数の汚染スコアの各々が対応する汚染閾値内にある場合には、方法は実施に成功したと分類される。一部の実施形態では、方法は、第1の試料を、（i）汚染スコアの少なくとも1つまたは複数の汚染閾値の対応する汚染閾値内にない場合、第2の試料によって汚染されている；または（ii）汚染スコアの少なくとも1つまたは複数の汚染閾値の対応する汚染閾値内にある場合、第2の試料によって汚染されていないと分類するステップをさらに含む。

20

30

40

II. 断片サイズ対照分子

【0185】

断片サイズ対照分子は、試料中の核酸分子の分析を評価および／または最適化するためにポリヌクレオチドの試料に添加される核酸分子である。断片サイズ対照分子は2つの領域、断片サイズ領域および識別子領域を有し得る。断片サイズ対照分子のセットを、断片サイズ対照分子のサブセット（複数可）に分類することができ、断片サイズ対照分子の各サブセットを、無細胞ポリヌクレオチドの試料の断片長分布、QC指標、ならびに／またはステップの、抽出、ライブラリ調製、富化、および／もしくはシーケンシングのいずれかでの試料の汚染を評価するために、アッセイにおける1つまたは複数の異なるステップで添加することができる。これらの断片サイズ対照分子の長さは既定であり、一部の実施形

50

態では、無細胞DNAの天然のサイズ分布を模倣することができる。断片サイズ対照分子の各サブセットを、断片サイズ領域の長さに基づいて断片サイズ対照分子の群にさらに分類することができ、各群は、異なる長さの断片サイズ領域を有し得る。例えば、断片サイズ対照分子のサブセットを、断片サイズ領域の長さに基づいて3つの群に分類することができ、第1の群は、長さ120bpの断片サイズ領域を有し得、第2の群は、長さ160bpの断片サイズ領域を有し得、第3の群は、長さ320bpの断片サイズ領域を有し得る。一部の実施形態では、断片サイズ対照分子は、合成オリゴヌクレオチドであり得る。一部の実施形態では、断片サイズ対照分子は、天然に存在しない核酸配列を有し得る。一部の実施形態では、断片サイズ対照分子は、天然に存在する核酸配列を有し得る。一部の実施形態では、任意のゲノム、プラスミド、もしくはベクター、またはその一部における特定の領域（すなわち、特定の長さの）を増幅することを介して生成されたアンプリコンを、断片サイズ対照分子として使用することができる。一部の実施形態では、断片サイズ対照分子は、非ヒトゲノムに対応する核酸配列を有し得る。例えば、これらの分子は、(i)ラムダファージDNAの領域に対応する配列、(ii)天然に存在しない配列、および／または(iii)(i)と(ii)の組合せのいずれかを有し得る。一部の実施形態では、断片サイズ対照分子は、天然に存在しないヌクレオチドアナログを含み得る。

10

【0186】

別の態様では、本開示は、既定の断片サイズ対照分子の少なくとも1つのサブセットを含む断片サイズ対照分子のセットであって、既定の断片サイズ対照分子の少なくとも1つのサブセットが、断片サイズ領域を含む複数の断片サイズ対照分子を含む、断片サイズ対照分子のセットを提供する。一部の実施形態では、断片サイズ対照分子は、識別子領域をさらに含む。断片サイズ領域は、断片サイズ対照分子の長さを表す断片サイズ対照分子の領域である。一部の実施形態では、少なくとも1つのサブセットは、断片サイズ対照分子の少なくとも1つの群を含む。一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、同じ長さの領域である。断片サイズ対照分子の各群は、断片サイズ領域の長さが異なり得る。例えば、断片サイズ対照分子のサブセットを、断片サイズ領域の長さに基づいて3つの群に分類することができ、第1の群は、長さ120bpの断片サイズ領域を有し得、第2の群は、長さ160bpの断片サイズ領域を有し得、第3の群は、長さ320bpの断片サイズ領域を有し得る。断片サイズ領域の長さは10bp～1000bpの間であり得る。一部の実施形態では、群における断片サイズ領域の長さは、他の群における断片サイズ領域の長さとは異なる。

20

30

【0187】

識別子領域は、断片サイズ対照分子を他の断片サイズ対照分子と区別するために使用される断片サイズ対照分子の領域である。識別子領域はまた、1つのサブセットからの断片サイズ対照分子を別のサブセットからの断片サイズ対照分子と区別するためにも使用される。一部の実施形態では、識別子領域は、断片サイズ領域の片側または両側に存在する。一部の実施形態では、識別子領域は分子バーコードを含む。分子バーコードは、断片サイズ対照分子の識別子として役立つ。識別子領域は、断片サイズ領域の片側または両側に存在し得る。分子バーコードは、断片サイズ対照分子の識別子として役立つ。一部の実施形態では、識別子領域は、(i)分子識別子バーコード、例えば各断片サイズ対照分子を識別し、1つの断片サイズ対照分子を別の断片サイズ対照分子と識別するために使用されるバーコード、および(ii)サブセット識別子バーコード、例えば断片サイズ対照分子が属するサブセットを識別するために使用されるバーコード（すなわち、断片サイズ対照分子がサブセット1またはサブセット2のいずれに属するか）として作用するバーコードを含む。サブセット識別子バーコードは、サブセットにおける全ての断片サイズ対照分子に関して同じであってもよく、1つのサブセットのサブセット識別子バーコードは、他のサブセットのサブセット識別子バーコードと異なってもよい。例えば、サブセット1における全ての断片サイズ対照分子が、サブセット識別子タグS1を有することができ、サブセット2における全ての断片サイズ対照分子が、サブセット識別子タグS2を有することができる。一部の実施形態では、1つの試料中の1つのサブセットのサブセット識別子バ

40

50

ーコードは、他の試料中の対応するサブセットのサブセット識別子バーコードとは異なり得る。例えば、断片サイズ対照分子を使用して試料の汚染を評価する場合、1つのサブセットのサブセット識別子バーコードは、同じフローセル／バッチ内の全ての他の試料において使用されるその対応するサブセットのサブセット識別子バーコードとは異なる。一部の実施形態では、識別子領域は、1つの試料の断片サイズ対照分子を他の試料の断片サイズ対照分子と区別するための試料インデックスを含み得る。例えば、富化ステップ後に添加された断片サイズ対照分子のサブセットの識別子領域は、試料インデックス配列を含み得る。一部の実施形態では、識別子領域は、ライゲーションを介して断片サイズ領域に付着させることができる。

【0188】

一部の実施形態では、断片サイズ対照分子は、1つまたは複数のプライマー結合部位を含む。一部の実施形態では、プライマー結合部位は識別子領域に存在する。一部の実施形態では、識別子領域はまた、断片サイズ対照分子を次世代シーケンサー（例えば、Illuminaシーケンサー）のフローセル表面に付着させるP5およびP7などの追加のフローセル結合部位を有し得る。

【0189】

一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、同じオリゴヌクレオチド配列を含む。一部の実施形態では、群における断片サイズ対照分子の断片サイズ領域は、少なくとも2つの区別可能なオリゴヌクレオチド配列を含む。一部の実施形態では、第1のサブセットにおける断片サイズ対照分子の断片サイズ領域は、第2のサブセットにおける断片サイズ対照分子の断片サイズ領域のオリゴヌクレオチド配列と区別可能なオリゴヌクレオチド配列を含む。

【0190】

一部の実施形態では、断片サイズ領域は、少なくとも10bp、少なくとも50bp、少なくとも60bp、少なくとも70bp、少なくとも80bp、少なくとも90bp、少なくとも100bp、少なくとも120bp、少なくとも150bp、少なくとも200bp、少なくとも250bp、少なくとも300bp、少なくとも400bp、少なくとも500bp、少なくとも600bp、少なくとも700bp、少なくとも800bp、少なくとも900bp、または少なくとも1000bpの長さであり得る。一部の実施形態では、断片サイズ領域の長さは10bp～1000bpの間であり得る。一部の実施形態では、断片サイズ対照分子のサブセットの各々は、等モル濃度で存在する。一部の実施形態では、断片サイズ対照分子のサブセットの各々は、等しくないモル濃度で存在する。一部の実施形態では、サブセットにおける断片サイズ対照分子の群の各々は、等モル濃度で存在する。一部の実施形態では、サブセットにおける断片サイズ対照分子の群の各々は、等しくないモル濃度で存在する。

【0191】

図3は、本開示の一部の実施形態との使用に適した断片サイズ対照分子の概略図である。ここで記載される断片サイズ対照分子のセットは、無細胞DNAの天然のサイズ分布のそれと類似の長さを有する。図3では、一例として、断片サイズ対照分子のセットは、2つのサブセット、サブセット1およびサブセット2に大別されている。図3における断片サイズ対照分子は、二本鎖DNA分子である。説明目的のために、二本鎖断片サイズ対照分子の1つの鎖のみを示す。この実施形態では、各サブセットは、3つの群、群1、群2、および群3に大別される。図3では、各群内の全ての断片サイズ対照分子の断片サイズ領域は同じ配列を有し、1つの群における断片サイズ領域の長さは他の群における断片サイズ領域の長さとは異なる。図3では、「—————」の領域は、断片サイズ対照分子の断片サイズ領域を表す。この実施形態では、識別子領域は、断片サイズ領域の両側に存在する。識別子領域は、断片サイズ領域の両末端にプライマー結合部位を有する。同様に、本明細書において、サブセット1における群1の断片サイズ領域の配列は、サブセット2における群1の断片サイズ領域の配列と同じである。同様に、サブセット1における群2および群3の断片サイズ領域の配列はそれぞれ、サブセット2における群2および群3の

10

20

30

40

50

断片サイズ領域の配列と同じである。

【0192】

断片サイズ領域の両側の識別子領域は、分子識別子バーコード（MB）を有するが、サブセット識別子バーコード（S）は片側のみに存在する。分子バーコードは、個々の断片サイズ対照分子の識別子として使用され、各断片サイズ対照分子は、一意的な分子バーコードを有する（すなわち、分子1はMB1およびMB2を有し、分子2はMB3およびMB4を有し、分子3はMB5およびMB6を有するなど）。サブセット識別子バーコードは、断片サイズ対照分子が属するサブセットの識別子として使用され得る。ここでは、サブセット1およびサブセット2の全ての断片サイズ対照分子はそれぞれ、S1およびS2のサブセット識別子バーコードを有する。この例では、サブセット識別子バーコードは、断片サイズ領域の片側に存在する。

10

【0193】

一部の実施形態では、分子バーコードは、断片サイズ領域の片側または両側に存在し得る。一部の実施形態では、サブセット識別子バーコードは、断片サイズ領域の片側または両側に存在し得る。

【0194】

本実施形態では、識別子領域は、断片サイズ領域の両側にプライマー結合部位を有する。ここでは、サブセット1に関して、群1、群2、および群3の断片サイズ対照分子はそれぞれ、断片サイズ領域の両側にPr1およびPr2、Pr3およびPr4、ならびにPr5およびPr6プライマー結合部位を有する。一部の実施形態では、識別子領域は、1つまたは複数のプライマーの結合を容易にする追加の領域（プライマー結合部位）を有し得る。一部の実施形態では、1つのサブセットにおける識別子領域のプライマー結合部位は、他のサブセットにおけるプライマー結合部位とは異なる。一部の実施形態では、これらのプライマー結合部位は、断片サイズ対照分子の回収率を分析するために使用される。一部の実施形態では、シーケンシングによって断片サイズ対照分子の回収率を分析する代わりに、断片サイズ対照分子の回収率を、これらのプライマー結合部位に結合するプライマーを使用して、デジタルドロップレットPCR（ddPCR）、定量的qPCR、またはゲル電気泳動によって分析することができる。

20

【0195】

図4は、本開示の一部の実施形態との使用に適した断片サイズ対照分子の概略図である。図4では、一例として、断片サイズ対照分子のセットは、2つのサブセット、サブセット1およびサブセット2に大別されている。図4における断片サイズ対照分子は、二本鎖DNA分子である。例示目的のために、二本鎖断片サイズ対照分子の1つの鎖のみを示す。この実施形態では、各サブセットは3つの群、群1、群2、および群3に大別される。図4では、各群内の全ての断片サイズ対照分子の断片サイズ領域は、同じ長さの領域であるが、群内の断片サイズ領域の配列は異なり得る。図4では、「————」および「~~~~」領域は、断片サイズ対照分子の断片サイズ領域を表す。「————」および「~~~~」領域は、各群が2つの異なる配列を含むが、それらが同じ長さの配列であることを示す。1つの群における断片サイズ領域の長さは、他の群における断片サイズ領域の長さとは異なる。この実施形態では、識別子領域は、断片サイズ領域の両側に存在する。同様に、ここでは、サブセット1における群1の断片サイズ領域の2つの配列は、サブセット2における群1の断片サイズ領域の2つの配列と同じである。同様に、サブセット1における群2および群3の断片サイズ領域の2つの配列はそれぞれ、サブセット2における群2および群3の断片サイズ領域の2つの配列と同じである。

30

40

【0196】

両側の識別子領域は分子識別子バーコード（MB）を有するが、サブセット識別子バーコード（S）は片側のみに存在する。分子バーコードは、個々の断片サイズ対照分子の識別子として使用され、各断片サイズ対照分子は、一意的な分子バーコードを有する（すなわち、分子1はMB1およびMB2を有し、分子2はMB3およびMB4を有し、分子3はMB5およびMB6を有するなど）。サブセット識別子バーコードは、断片サイズ対照分

50

子が属するサブセットの識別子として使用され得る。ここでは、サブセット 1 およびサブセット 2 の全ての断片サイズ対照分子はそれぞれ、S 1 および S 2 のサブセット識別子バーコードを有する。この例では、サブセット識別子バーコードは、断片サイズ領域の片側に存在する。

【0197】

一部の実施形態では、分子バーコードは、断片サイズ領域の片側または両側に存在し得る。一部の実施形態では、サブセット識別子バーコードは、断片サイズ領域の片側または両側に存在し得る。

【0198】

一部の実施形態では、断片サイズ対照分子は、天然に存在しない核酸配列を有し得る。一部の実施形態では、断片サイズ対照分子は、天然に存在する核酸配列を有し得る。一部の実施形態では、断片サイズ対照分子は、非ヒトゲノムに対応する核酸配列を有し得る。例えば、これらの分子は、(i) ラムダファージ DNA の領域に対応する配列、(ii) 天然に存在しない配列、および／または (iii) (i) と (ii) の組合せのいずれかを有し得る。一部の実施形態では、断片サイズ対照分子は、天然に存在しないヌクレオチドアナログを含み得る。

【0199】

一部の実施形態では、ポリヌクレオチドの試料は、DNA の試料、RNA の試料、無細胞ポリヌクレオチドの試料、無細胞 DNA の試料、または無細胞 RNA の試料である。一部の実施形態では、ポリヌクレオチドの試料は、無細胞 DNA の試料である。

【0200】

一部の実施形態では、無細胞 DNA は、少なくとも 1 ng、少なくとも 5 ng、少なくとも 10 ng、少なくとも 15 ng、少なくとも 20 ng、少なくとも 30 ng、少なくとも 50 ng、少なくとも 75 ng、少なくとも 100 ng、少なくとも 150 ng、少なくとも 200 ng、少なくとも 250 ng、少なくとも 300 ng、少なくとも 350 ng、少なくとも 400 ng、少なくとも 450 ng、または少なくとも 500 ng である。

【0201】

一部の実施形態では、断片サイズ対照分子の量は、少なくとも 1 アトモル、少なくとも 2 アトモル、少なくとも 5 アトモル、少なくとも 10 アトモル、少なくとも 15 アトモル、少なくとも 20 アトモル、少なくとも 50 アトモル、少なくとも 75 アトモル、少なくとも 100 アトモル、少なくとも 1 フェムトモル、少なくとも 2 フェムトモル、少なくとも 5 フェムトモル、少なくとも 10 フェムトモル、少なくとも 15 フェムトモル、少なくとも 20 フェムトモル、少なくとも 50 フェムトモル、少なくとも 75 フェムトモル、少なくとも 100 フェムトモル、少なくとも 125 フェムトモル、少なくとも 150 フェムトモル、または少なくとも 200 フェムトモル、少なくとも 300 フェムトモル、少なくとも 400 フェムトモル、少なくとも 500 フェムトモル、少なくとも 600 フェムトモル、少なくとも 700 フェムトモル、少なくとも 800 フェムトモル、少なくとも 900 フェムトモル、少なくとも 1 ピコモル、少なくとも 2 ピコモル、少なくとも 5 ピコモル、または少なくとも 10 ピコモルである。一部の実施形態では、断片サイズ対照分子の量は、1 アトモル～10 ピコモルの間であり得る。

【0202】

本開示の追加の実施形態は、サイズ断片対照分子が添加されている組成物を含む。例えば、サイズ断片対照分子が添加されている無細胞 DNA 試料は、本開示の一実施形態である。同様に、本方法の実践の間に産生されたサイズ断片対照分子の 1 つまたは複数の異なるサブセットを含む多数の組成物が、本開示の実施形態であると考えられる。

III. 方法の全般的特色

A. 試料

【0203】

試料は、対象から単離された任意の生体試料であり得る。試料は、体組織、全血、血小板

10

20

30

40

50

、血清、血漿、便、赤血球、白血球（white blood cell）または白血球（leucocyte）、内皮細胞、組織生検（例えば、固形腫瘍であることが公知のまたは疑われる生検）、脳脊髄液、滑液、リンパ液、腹水、間質液または細胞外液（例えば、細胞間隙からの滲出液）、歯肉滲出液、歯肉溝滲出液、骨髓、胸水、脳脊髄液、唾液、粘液、喀痰、精液、汗、および尿を含み得る。試料は、血液およびその画分、ならびに尿などの体液であり得る。そのような試料は、腫瘍から放出された核酸を含み得る。核酸は、DNAおよびRNAを含み得、二本鎖および一本鎖形態であり得る。試料は、対象から最初に単離された形態であり得るか、または細胞などの構成成分を除去もしくは添加する、1つの構成成分を別の構成成分と比べて富化させる、もしくは1つの形態の核酸を別の形態に変換する、例えば、RNAをDNAにもしくは一本鎖核酸を二本鎖に変換するさらなる処理に供されたものであり得る。このように、例えば、分析用の体液は、無細胞核酸、例えば、無細胞離DNA（cfDNA）を含有する血漿または血清であり得る。

10

【0204】

一部の実施形態では、対象から採取した体液の試料体積は、シーケンシングされる領域に対する所望のリードの深さに依存する。体積の例は、約0.4~40ミリリットル（mL）、約5~20mL、約10~20mLである。例えば、体積は、約0.5mL、約1mL、約5mL、約10mL、約20mL、約30mL、約40mL、またはそれよりも多くであり得る。試料採取された血漿の体積は、典型的に約5mL~約20mLの間である。

【0205】

試料は、様々な量の核酸を含み得る。典型的には、所定の試料中の核酸の量は多数のゲノム等価物と同等とみなされる。例えば、DNA約30ナノグラム（ng）の試料は一倍体ヒトゲノム等価物を約10,000（10⁴）個、cfDNAの場合では、個々のポリヌクレオチド分子を約2000億（2×10¹¹）個含有し得る。同様に、DNA約100ngの試料は、一倍体ヒトゲノム等価物を約30,000個、cfDNAの場合では、個々の分子を約6000億個含有し得る。

20

【0206】

一部の実施形態では、試料は、異なる供給源由来、例えば、細胞由来および無細胞供給源（例えば、血液試料など）由来の核酸を含む。典型的には、試料は、突然変異を有する核酸を含む。例えば試料は、必要に応じて、生殖細胞系列突然変異および／または体細胞突然変異を有するDNAを含む。典型的には試料は、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を含む。

30

【0207】

増幅前の試料中の無細胞核酸の例としての量は、典型的には約1フェムトグラム（fg）~約1マイクログラム（μg）、例えば、約1ピコグラム（pg）~約200ナノグラム（ng）、約1ng~約100ng、約10ng~約1000ngの範囲である。一部の実施形態では、試料は、無細胞核酸分子の最大約600ng、最大約500ng、最大約400ng、最大約300ng、最大約200ng、最大約100ng、最大約50ng、または最大約20ngを含む。必要に応じて、量は、無細胞核酸分子の少なくとも約1fg、少なくとも約10fg、少なくとも約100fg、少なくとも約1pg、少なくとも約10pg、少なくとも約100pg、少なくとも約1ng、少なくとも約10ng、少なくとも約100ng、少なくとも約150ng、または少なくとも約200ngである。一部の実施形態では、量は、無細胞核酸分子の最大約1fg、約10fg、約100fg、約1pg、約10pg、約100pg、約1ng、約10ng、約100ng、約150ng、または約200ngである。一部の実施形態では、方法は、試料から約1fg~約200ngの間の無細胞核酸を得るステップを含む。

40

【0208】

無細胞核酸は、典型的には、約100ヌクレオチド長~約500ヌクレオチド長の間のサイズ分布を有し、約110ヌクレオチド長~約230ヌクレオチド長の分子が試料中の分子の約90%を表し、最頻値は約168ヌクレオチド長（ヒト対象からの試料中）および

50

第2の軽微なピークは約240ヌクレオチド～約440ヌクレオチド長の範囲である。一部の実施形態では、無細胞核酸は、約160ヌクレオチド～約180ヌクレオチド長、または約320ヌクレオチド～約360ヌクレオチド長、または約440ヌクレオチド～約480ヌクレオチド長である。

【0209】

一部の実施形態では、無細胞核酸は、溶液中で見出される無細胞核酸がインタクト細胞および体液の他の不溶性の構成成分から分離される分画化ステップを通して体液から単離される。一部の実施形態では、分画化は、遠心分離または濾過などの技法を含む。あるいは、体液中の細胞を溶解し得、無細胞および細胞核酸を共に処理してもよい。一般的には、緩衝液の添加および洗浄ステップの後、無細胞核酸を例えばアルコールによって沈殿させてもよい。一部の実施形態では、シリカに基づくカラムなどの追加の浄化ステップを使用して、夾雑物または塩を除去する。例えば非特異的な大量の担体核酸を必要に応じて反応全体に添加して、例としての手順の態様、例えば収量を最適化する。そのような処理後、試料は、典型的には、二本鎖DNA、一本鎖DNA、および／または一本鎖RNAを含む様々な形態の核酸を含む。必要に応じて、一本鎖DNAおよび／または一本鎖RNAは、その後の処理および分析ステップに含まれるように、二本鎖型に変換される。

10

B. タグ付け

【0210】

一部の実施形態では、核酸分子（ポリヌクレオチドおよび断片サイズ対照分子の試料から）を、試料インデックスおよび／または分子バーコード（一般的に「タグ」と呼ばれる）によってタグ付けしてもよい。タグは、他の方法の中でも化学合成、ライゲーション（例えば、平滑末端ライゲーションまたは付着末端ライゲーション）、またはオーバーラップ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まれ得るか、またはそうでなければ接合される。そのようなアダプターは、標的核酸分子に最終的に接合され得る。他の実施形態では、増幅サイクルの1つまたは複数のラウンド（例えば、PCR増幅）を一般的に適用して、従来の核酸増幅法を使用して試料インデックスを核酸分子に導入する。増幅は、1つまたは複数の反応混合物（例えば、アレイ中の複数のマイクロウェル）中で実施され得る。分子バーコードおよび／または試料インデックスは、同時に、または任意の順序で導入してもよい。一部の実施形態では、分子バーコードおよび／または試料インデックスを、配列捕捉ステップを実施する前および／または後に導入する。一部の実施形態では、分子バーコードのみをプローブ捕捉の前に導入し、試料インデックスは配列捕捉ステップを実施した後に導入する。一部の実施形態では、分子バーコードおよび試料インデックスの両方を、プローブに基づく捕捉ステップを実施する前に導入する。一部の実施形態では、試料インデックスを、配列捕捉ステップを実施した後に導入する。一部の実施形態では、分子バーコードを、ライゲーション（例えば、平滑末端ライゲーションまたは付着末端ライゲーション）を介してアダプターを通して試料中の核酸分子（例えば、cfDNA分子）に組み込む。一部の実施形態では、試料インデックスを、オーバーラップ伸長ポリメラーゼ連鎖反応（PCR）を通して試料中の核酸分子（例えば、cfDNA分子）に組み込む。典型的には、配列捕捉プロトコルは、標的化核酸配列、例えばゲノム領域のコード配列と相補的な一本鎖核酸分子を導入するステップを伴い、そのような領域の突然変異は、がんの型に関連する。

20

30

40

【0211】

一部の実施形態では、タグは、試料核酸分子の一方の末端または両方の末端に位置し得る。一部の実施形態では、タグは、既定の、またはランダムもしくはセミランダム配列オリゴヌクレオチドである。一部の実施形態では、タグは、約500、200、100、50、20、10、9、8、7、6、5、4、3、2、または1ヌクレオチド長未満であり得る。タグは、試料核酸にランダムにまたは非ランダムに連結され得る。

【0212】

一部の実施形態では、各試料は、試料インデックスまたは試料インデックスの組合せによって一意的にタグ付けされる。一部の実施形態では、試料またはサブサンプルの各核酸分

50

子は、分子バーコードまたは分子バーコードの組合せによって一意的にタグ付けされる。他の実施形態では、分子バーコードが必ずしも複数の中で互いに対して一意的ではないように（例えば、非一意的分子バーコード）、複数の分子バーコードを使用してもよい。これらの実施形態では、分子バーコードを、一般的に、個々の分子に（例えば、ライゲーションによって）付着させて、その結果、分子バーコードとそれを付着させ得る配列の組合せにより、個別に追跡することができる一意的な配列が作製される。非一意的にタグ付けされた分子バーコードの、内因性配列情報（例えば、試料中の元の核酸分子の配列に対応する始まりの（開始）および／または終わりの（終止）ゲノム場所／位置、一方もしくは両方の末端における配列リードの部分配列、配列リードの長さ、ならびに／または試料中の元の核酸分子の長さ）と組み合わせた検出により、典型的には、一意的な同一性の、特定の分子への割り当てが可能となる。10 一部の実施形態では、非一意的にタグ付けされた分子バーコードの、内因性配列情報（例えば、配列リードの参照配列へのアライメントの始まりの（開始）および／または終わりの（終止）領域、一方もしくは両方の末端における配列リードの部分配列、配列リードの長さ、および／または試料中の元の核酸分子の長さ）と組み合わせた検出により、典型的には、一意的な同一性の、特定の分子への割り当てが可能となる。一部の実施形態では、始まりの領域は、シーケンシングリードの5'末端が参照配列に対する整列を開始することが決定されるシーケンシングリードのゲノム開始位置を含み、終わりの領域は、シーケンシングリードの3'末端が参照配列に対する整列を終止することが決定されるシーケンシングリードのゲノム終止位置を含む。一部の20 実施形態では、始まりの領域は、参照配列に対して整列するシーケンシングリードの5'末端の最初の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、または少なくとも最初の30塩基位置を含む。一部の実施形態では、終わりの領域は、参照配列に対して整列するシーケンシングリードの3'末端の最後の1、最後の2、最後の5、最後の10、最後の15、最後の20、最後の25、最後の30、または少なくとも最後の30塩基位置を含む。

【0213】

個々の配列リードの長さまたは塩基対の数もまた、一意的な同一性を所定の分子に割り当てるために必要に応じて使用される。本明細書に記載されるように、一意的な同一性が割り当てられている核酸の一本鎖由来の断片により、その後の親鎖および／または相補鎖由来の断片の識別が可能となり得る。30

【0214】

一部の実施形態では、分子バーコードは、識別子のセット（例えば、一意的または非一意的分子バーコードの組合せ）の試料中の分子に対する予想される比で導入される。一例のフォーマットは、標的分子の両末端にライゲーションされた約2～約1,000,000個の異なる分子バーコード配列、または約5～約150個の異なる分子バーコード配列、または約20～約50個の異なる分子バーコード配列を使用する。あるいは、約25～約1,000,000個の異なる分子バーコード配列を使用してもよい。例えば、20～50×20～50個の分子バーコード配列（すなわち、20～50個の異なる分子バーコード配列のうちの1つを標的分子の各末端に付着させることができる）を使用することができる。そのような識別子の数は、典型的には、同じ開始点および終止点を有する異なる分子が、識別子の異なる組合せを受ける確率が高くなる（例えば、少なくとも94%、99.5%、99.99%、または99.999%）ために十分である。一部の実施形態では、分子の約80%、約90%、約95%、または約99%が分子バーコードの同じ組合せを有する。40

【0215】

一部の実施形態では、反応における一意的または非一意的分子バーコードの割り当ては、例えば、その各々の全体が参照により本明細書に組み込まれる、米国特許出願第20010053519号、第20030152490号、および第20110160078号、ならびに米国特許第6,582,908号、第7,537,898号、第9,598,731号、および第9,902,992号に記載される方法およびシステムを使用して実施50

される。あるいは、一部の実施形態では、内因性配列情報（例えば、開始および／または終止位置、配列の一方または両方の末端の部分配列、および／または長さ）のみを使用して、試料の異なる核酸分子を識別してもよい。

【0216】

サブセット識別子バーコードは、断片サイズ対照分子の識別子領域の一部であり得る。サブセット識別子バーコードを使用して、断片サイズ対照分子が属するサブセット（すなわち、断片サイズ対照分子がサブセット1またはサブセット2のいずれに属するか）を識別する。サブセット識別子バーコードは、サブセットにおける全ての断片サイズ対照分子に関して同じであり得、同様に、1つのサブセットのサブセット識別子バーコードは、同じ試料中および／または同じバッチ中で使用される他の試料中の他のサブセットのサブセット識別子バーコードとは異なり得る。一部の実施形態では、1つの試料中の1つのサブセットのサブセット識別子バーコードは、他の試料中の対応するサブセットのサブセット識別子バーコードとは異なり得る。例えば、断片サイズ対照分子を使用して試料の汚染を評価する場合、1つのサブセットのサブセット識別子バーコードは、同じフローセル／バッチ内の全ての他の試料において使用されるその対応するサブセットのサブセット識別子バーコードとは異なる。一部の実施形態では、識別子領域は、1つの試料の断片サイズ対照分子を他の試料の断片サイズ対照分子と区別するために試料インデックスを含み得る。例えば、富化ステップの後に添加された断片サイズ対照分子のサブセットの識別子領域は、試料インデックス配列を含み得る。

10

【0217】

一部の実施形態では、反応における一意的または非一意的分子バーコードの割り当ては、例えば、その各々の全体が参照により本明細書に組み込まれる、米国特許出願第20010053519号、第20030152490号、および第20110160078号、ならびに米国特許第6,582,908号、第7,537,898号、第9,598,731号、および第9,902,992号に記載される方法およびシステムを使用して実施される。

20

C. 増幅

【0218】

試料核酸および断片サイズ対照分子を、アダプターに隣接させ、増幅されるDNA分子に隣接するアダプター中のプライマー結合部位に結合する核酸プライマーを使用してPCRおよび他の増幅方法によって増幅することができる。一部の実施形態では、増幅方法は、サーモサイクリングに起因する伸長、変性、およびアニーリングのサイクルを伴うか、または例えば転写媒介増幅において等温的であり得る。必要に応じて利用され得る増幅方法の他の例としては、リガーゼ連鎖反応、鎖置換増幅、核酸配列に基づく増幅、および自家持続配列に基づく複製が挙げられる。

30

【0219】

典型的には、増幅反応は、約150ヌクレオチド(nt)～約700nt、250nt～約350nt、または約320nt～約550ntの範囲のサイズの、分子バーコードおよび試料インデックスを有する複数の非一意的または一意的にタグ付けされた核酸分子を生成する。一部の実施形態では、アンプリコンは約180ntのサイズを有する。一部の実施形態では、アンプリコンは約200ntのサイズを有する。

40

D. 富化

【0220】

一部の実施形態では、配列は核酸をシーケンシングする前に富化される。富化は、必要に応じて特異的標的領域についてまたは非特異的に行われる（「標的配列」）。一部の実施形態では、目的の標的化領域は、示差的タイリングおよび捕捉スキームを使用して1つまたは複数のベイトセットパネルのために選択された核酸捕捉プローブ（「ベイト」）によって富化され得る。示差的タイリングおよび捕捉スキームは一般的に、様々な相対的濃度のベイトセットを使用して、一連の制約（例えば、シーケンシング負荷などのシーケンサーの制約、各ベイトの有用性）を受けながら、ベイトに関連するゲノム領域にわたって示

50

差的にタイリングし（例えば、異なる「分解能」で）、下流のシーケンシングのために所望のレベルで標的化核酸を捕捉する。目的のこれらの標的化ゲノム領域は、必要に応じて核酸構築物の天然または合成ヌクレオチド配列を含む。一部の実施形態では、1つまたは複数の目的の領域に対するプローブを有するビオチン標識ビーズを使用して、標的配列を捕捉し、必要に応じて、その後それらの領域を増幅させて、目的の領域を富化することができる。

【0221】

配列捕捉は、典型的には、標的核酸配列にハイブリダイズするオリゴヌクレオチドプローブの使用を伴う。一部の実施形態では、プローブセット戦略は、目的の領域にわたってプローブをタイリングすることを伴う。そのようなプローブは、例えば、約60～約120ヌクレオチド長であり得る。セットは、約2×、3×、4×、5×、6×、7×、8×、9×、10×、15×、20×、50×、または50×を超える深さ（例えば、カバレッジの深さ）を有し得る。配列捕捉の有効性は一般的に、一部にはプローブの配列と相補的な（またはほぼ相補的な）標的分子内の配列の長さに依存する。

10

E. シーケンシング

【0222】

必要に応じてアダプターに隣接させ、予め増幅させたまたはさせていない試料核酸および断片サイズ対照分子を、一般的にシーケンシングに供する。必要に応じて利用されるシーケンシング方法または市販のフォーマットとしては、例えば、サンガーシーケンシング、ハイスループットシーケンシング、パイロシーケンシング、合成によるシーケンシング、単一分子シーケンシング、ナノポアに基づくシーケンシング、半導体シーケンシング、ライゲーションによるシーケンシング、ハイブリダイゼーションによるシーケンシング、RNA-Seq (Illumina)、Digital Gene Expression (Helicos)、次世代シーケンシング (NGS)、Molecule Sequencing by Synthesis (SMSS) (Helicos)、大規模並列シーケンシング、Clonal Single Molecule Array (Solexa)、ショットガンシーケンシング、Ion Torrent、Oxford Nanopore、Roche Genia、Maxim-Gilbertシーケンシング、プライマーウォーキング、PacBio、SOLiD、Ion Torrent、またはNanoporeプラットフォームを使用するシーケンシングが挙げられる。シーケンシング反応は、複数の試料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、または他の手段を含み得る多様な試料処理装置で実施することができる。試料処理装置は、複数の動作を同時に処理することを可能にするために複数の試料チャンバーも含み得る。

20

30

【0223】

シーケンシング反応は、がんまたは他の疾患のマーカーを含有する1つまたは複数の核酸断片型または領域に対して実施することができる。シーケンシング反応はまた、試料中に存在する任意の核酸断片に対して実施することもできる。シーケンシング反応は、ゲノムの少なくとも約5%、10%、15%、20%、25%、30%、40%、50%、60%、70%、80%、90%、95%、99%、99.9%、または100%について実施することができる。他の場合では、シーケンシング反応を、ゲノムの約5%、10%、15%、20%、25%、30%、40%、50%、60%、70%、80%、90%、95%、99%、99.9%、または100%未満について実施することができる。

40

【0224】

同時シーケンシング反応を、マルチプレックスシーケンシング技法を使用して実施することができる。一部の実施形態では、無細胞ポリヌクレオチドを、少なくとも約1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、50000、または100,000のシーケンシング反応でシーケンシングする。他の実施形態では、無細胞ポリヌクレオチドを、約1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、500

50

00、または100、000未満のシーケンシング反応でシーケンシングする。シーケンシング反応は、典型的には、逐次的または同時に実施される。その後のデータ分析は、一般的に、シーケンシング反応の全部または一部について実施される。一部の実施形態では、データ分析を、少なくとも約1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、50000、または100、000のシーケンシング反応について実施する。他の実施形態では、データ分析を、約1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、50000、または100、000未満のシーケンシング反応について実施してもよい。リードの深さの例は、座（例えば、塩基位置）当たり約1000～約50000リードである。

10

F. 分析

【0225】

シーケンシングにより、複数の配列リードまたはリードが生成され得る。シーケンシングリードまたはリードは、約150塩基未満の長さ、または約90塩基未満の長さのヌクレオチド配列データを含み得る。一部の実施形態では、リードは、約80塩基～約90塩基の間、例えば、約85塩基の長さである。一部の実施形態では、本開示の方法を非常に短いリード、例えば、約50塩基未満または約30塩基の長さに適用する。配列リードデータは、配列データならびにメタ情報を含み得る。配列リードデータは、例えば、VCFファイル、FASTAファイル、またはFASTQファイルを含む任意の適切なファイルフォーマットに保存することができる。

20

【0226】

FASTAは、配列データベースを検索するためのコンピュータプログラムを指し得、FASTAという名称は、標準ファイルフォーマットも指し得る。FASTAは、例えば、その全体が参照により本明細書に組み込まれる、Pearson & Lipman, 1988, Improved tools for biological sequence comparison, PNAS 85:2444-2448によって記載されている。FASTAフォーマットの配列は、一行の説明行で始まり、その後に配列データの行が続く。説明行は最初の列の大なり（「>」）記号によって配列データと区別される。「>」記号に続くワードが配列の識別子であり、行の残りは説明である（いずれも必要に応じて）。「>」と識別子の最初の文字の間にスペースがなくてもよい。文字列の全ての行が80文字未満であることが推奨される。「>」から始まる別の行が出現すると配列が終わり、これは、別の配列の始まりを示す。

30

【0227】

FASTQフォーマットは、生物学的配列（通常ヌクレオチド配列）およびその対応するクオリティスコアの両方を保存するためのテキストに基づくフォーマットである。FASTQフォーマットはFASTAフォーマットと同様であるが、配列データの後にクオリティスコアが続く。簡潔にするために配列文字およびクオリティスコアの両方を単一のASCII文字で符号化する。FASTQフォーマットは、例えば、その全体が参照により本明細書に組み込まれる、Cock et al. (“The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants,” Nucleic Acids Res 38 (6): 1767-1771, 2009)によって記載されるように、Illumina Genome Analyzerなどのハイスループットシーケンシング機器の出力を保存するための事実上の標準である。

40

【0228】

FASTAおよびFASTQファイルに関しては、メタ情報は説明行を含み、配列データの行は含まない。一部の実施形態では、FASTQファイルに関して、メタ情報はクオリティスコアを含む。FASTAおよびFASTQファイルに関して、配列データは説明行の後に始まり、一般的には、IUPAC多義性コードのいくつかのサブセットを使用し、

50

必要に応じて「-」を伴って存在する。ある実施形態では、配列データは、A、T、C、G、およびN文字を使用することができ、必要に応じて「-」または必要であればUを含む（例えば、ギャップまたはウラシルを表すために）。

【0229】

一部の実施形態では、少なくとも1つのマスター配列リードファイルおよび出力ファイルは、プレーンテキストファイルとして保存される（例えば、ASCII；ISO／IEC 646；EBCDIC；UTF-8；またはUTF-16などのエンコーディングを使用する）。本開示によって提供されるコンピュータシステムは、プレーンテキストファイルを開くことが可能なテキストエディタプログラムを含み得る。テキストエディタプログラムは、テキストファイル（例えば、プレーンテキストファイルなど）の内容をコンピュータスクリーン上に提示することが可能で、人がテキストを編集することができる（例えば、モニター、キーボード、およびマウスを使用して）コンピュータプログラムを指し得る。テキストエディタの例としては、Microsoft Word、emacs、pico、vi、BBEdit、およびTextWranglerが挙げられるがこれらに限定されない。テキストエディタプログラムは、プレーンテキストファイルをコンピュータスクリーン上に表示し、メタ情報および配列リードを人が読むことが可能なフォーマット（例えば、バイナリコードされているのではなく、印刷または人が書く場合に使用され得る英数字を使用する）で示すことが可能であり得る。

10

【0230】

方法を、FASTAまたはFASTQファイルに関連して考察してきたが、本開示の方法およびシステムを使用して、例えば、Variant Call Format（VCF）フォーマットのファイルを含む任意の適切な配列ファイルフォーマットを圧縮することができる。典型的なVCFファイルは、ヘッダーセクションおよびデータセクションを含み得る。ヘッダーは、任意の数のメタ情報行を含有し、各々は「##」文字で始まり、TABで区切られるフィールド定義行が単一の「#」文字で始まる。フィールド定義行により8つの必須列が指名され、本文セクションは、フィールド定義行によって定義される列を集合させたデータの行を含有する。VCFフォーマットは、例えば、その全体が参照により本明細書に組み込まれる、Danecek et al.（“The variant call format and VCF tools,” Bioinformatics 27（15）：2156-2158，2011）によって記載されている。ヘッダーセクションは圧縮されたファイルに書き込むためのメタ情報として扱うことができ、データセクションは行として扱うことができ、その各々が、一意的な場合に限りマスターファイルに保存され得る。

20

30

【0231】

一部の実施形態は、配列リードのアセンブリを提供する。アライメントによるアセンブリでは、例えば、配列リードを互いに整列させるかまたは参照配列に対して整列させる。各リードを参照ゲノムに対して順番に整列させることにより、リードの全てが互に関連して配置されて、アセンブリを作出する。加えて、配列リードを参照配列に対して整列またはマッピングすることを使用して、配列リード内のバリエーション配列を識別することもできる。疾患もしくは状態の診断もしくは予後判定をさらに補助するため、または処置決定を手引きするために、バリエーション配列を識別することを、本明細書に記載の方法およびシステムと組み合わせ使用することができる。

40

【0232】

一部の実施形態では、ステップのいずれかまたは全てを自動化する。あるいは、本開示の方法を、例えば、各々が必要に応じてC++などのコンパイル型言語で書かれた1つまたは複数の専用プログラムに完全にまたは部分的に具体化し、次いで、コンパイルし、バイナリとして配布することができる。本開示の方法は、完全にまたは部分的に、既存の配列分析プラットフォーム内のモジュールとして、または既存の配列分析プラットフォーム内のファンクションを呼び出すことによってインプリメントすることができる。一部の実施形態では、本開示の方法は、全て単一の開始キュー（例えば、人による作業、別のコンピ

50

ュータプログラム、または機械から供給される誘発事象の1つまたは組合せ)に応答して自動的に呼び出されるいくつかのステップを含む。このように、本開示は、任意のまたはステップまたはステップの任意の組合せがキューに応答して自動的に起こり得る方法を提供する。「自動的に」とは、一般に、人による入力、影響、または相互作用を介さないことを意味する(例えば、元のまたはキュー前の人による作業にのみ応答する)。

【0233】

本開示の方法は、対象の核酸試料に関する正確かつ高感度の解釈を含む、様々な形態の出力も包含し得る。検索の出力は、コンピュータファイルのフォーマットで提供することができる。一部の実施形態では、出力は、FASTAファイル、FASTQファイル、またはVCFファイルである。出力を処理して、参照ゲノムの配列に対して整列させた核酸の配列などの配列データを含むテキストファイル、またはXMLファイルを作成することができる。他の実施形態では、処理により、参照ゲノムと比較して対象核酸における1つまたは複数の突然変異を記述する座標またはストリングを含む出力がもたらされる。アライメントストリングは、Simple UnGapped Alignment Report (SUGAR)、Verbose Useful Labeled Gapped Alignment Report (VULGAR)、およびCompact Idiosyncratic Gapped Alignment Report (CIGAR) (例えば、その全体が参照により本明細書に組み込まれる、Ning et al., Genome Research 11 (10): 1725-9, 2001によって記載されている)を含み得る。これらのストリングは、例えば、European Bioinformatics Institute (Hinxton, UK)のExonerate配列アライメントソフトウェアにおいてインプリメントすることができる。

10

20

【0234】

一部の実施形態では、例えば、CIGARストリングを含む、配列アライメントマップ(SAM)またはバイナリアライメントマップ(BAM)ファイルなどの配列アライメントを作成する(SAMフォーマットは、例えば、その全体が参照により本明細書に組み込まれる、Li et al., "The Sequence Alignment/Map format and SAMtools," Bioinformatics, 25 (16): 2078-9, 2009によって記載されている)。一部の実施形態では、CIGARは、行当たり1つのギャップを有するアライメントを表示するかまたは含む。CIGARは、CIGARストリングとして報告される圧縮されたペアワイズアライメントフォーマットである。CIGARストリングは、長い(例えば、ゲノム)ペアワイズアライメントを表すために有用であり得る。CIGARストリングをSAMフォーマットで使用して、参照ゲノム配列に対するリードのアライメントを表してもよい。

30

【0235】

CIGARストリングは、確立されたモチーフに従い得る。各文字の前には、事象の塩基数を与える数が示される。使用される文字は、M、I、D、N、およびSを含み得る(M=マッチ; I=挿入; D=欠失; N=ギャップ; S=置換)。CIGARストリングは、マッチおよび/またはミスマッチおよび欠失(またはギャップ)の配列を定義する。例えば、CIGARストリング2MD3M2D2Mは、アライメントが2つのマッチ、1つの欠失(スペースを節約するために数字の1は省略される)、3つのマッチ、2つの欠失、および2つのマッチを含むことを示し得る。

40

【0236】

一部の実施形態では、一方の末端または両方の末端に一本鎖オーバーハングを有する二本鎖核酸上に平滑末端を酵素により形成することによって、核酸集団をシーケンシングのために調製する。これらの実施形態では、集団を典型的には、ヌクレオチド(例えば、A、C、G、およびTまたはU)の存在下で、5'-3' DNAポリメラーゼ活性および3'-5'エキソヌクレアーゼ活性を有する酵素によって処理する。必要に応じて使用することができる酵素またはその触媒断片の例としては、クレノウ大断片およびT4ポリメラーゼ

50

が挙げられる。5'オーバーハングでは、酵素は典型的に、反対鎖上の陥凹3'末端を5'末端と平滑になるまで伸長させて平滑末端を産生する。3'オーバーハングでは、酵素は一般的に、3'末端から反対鎖の5'末端まで、および時にはそれを越えて消化する。この消化が反対鎖の5'末端を越えて進行する場合、5'オーバーハングのために使用される同じポリメラーゼ活性を有する酵素によってギャップを埋めることができる。二本鎖核酸上での平滑末端の形成により、例えば、アダプターの付着およびその後の増幅が容易となる。

【0237】

一部の実施形態では、核酸集団を、一本鎖核酸から二本鎖核酸への変換および／またはRNAからDNA（例えば、相補DNAまたはcDNA）への変換などの追加的な処理に供する。これらの形態の核酸もまた、必要に応じてアダプターに連結し、増幅させる。

10

【0238】

事前の増幅の有無にかかわらず、上記の平滑末端を形成するプロセスに供された核酸および必要に応じて試料中の他の核酸をシーケンシングして、シーケンシングされた核酸を産生することができる。シーケンシングされた核酸は、核酸の配列（例えば、配列情報）またはその配列が決定されている核酸のいずれかを指し得る。シーケンシングは、試料中の個々の核酸分子の配列データを直接、または試料中の個々の核酸分子の増幅産物のコンセンサス配列から間接的に提供するように実施することができる。

【0239】

一部の実施形態では、平滑末端形成後の試料中の一本鎖オーバーハングを有する二本鎖核酸を、両末端でバーコードを含むアダプターに連結し、シーケンシングにより、核酸配列ならびにアダプターによって導入されたインラインバーコードを決定する。平滑末端DNA分子を必要に応じて少なくとも部分的に二本鎖のアダプター（例えば、Y形状またはベル形状アダプター）の平滑末端にライゲーションする。あるいは、ライゲーションを容易にするために（例えば、付着末端ライゲーションのために）、試料核酸の平滑末端およびアダプターに、相補的ヌクレオチドによるテールを付加することができる。

20

【0240】

核酸試料は、典型的には、同じ核酸の任意の2つのコピーが、両末端に連結したアダプター由来のアダプターバーコードの同じ組合せを受ける確率が低い（例えば、約1%または0.1%未満）、十分数のアダプターと接触させる。アダプターのこのような使用により、参照核酸における開始点および終止点が同じであり、バーコードの同じ組合せに連結された核酸配列のファミリーの識別が可能になり得る。そのようなファミリーは、増幅前の試料中の核酸の増幅産物の配列を表し得る。ファミリーメンバーの配列をコンパイルして、平滑末端形成およびアダプター付着によって修飾された元の試料中の核酸分子のコンセンサスヌクレオチドまたは完全なコンセンサス配列を導出することができる。言い換えれば、試料中の核酸の特定の位置を占有するヌクレオチドを、ファミリーメンバー配列のその対応する位置を占有するヌクレオチドのコンセンサスであると決定することができる。ファミリーは、二本鎖核酸の一方または両方の鎖の配列を含み得る。ファミリーのメンバーが二本鎖核酸の両方の鎖の配列を含む場合、配列をコンパイルしてコンセンサスヌクレオチド（複数可）または配列を導出する目的で、1つの鎖の配列をその相補体に変換することができる。一部のファミリーは、単一のメンバー配列のみを含む。この場合、この配列を、増幅前の試料中の核酸の配列とみなすことができる。あるいは、単一のメンバー配列のみを有するファミリーをその後の分析から除外することができる。

30

40

【0241】

シーケンシングされた核酸におけるヌクレオチド変異（例えば、SNVまたはインデル）は、シーケンシングされた核酸を参照配列と比較することによって決定することができる。参照配列は、多くの場合、公知の配列、例えば、対象由来の公知の全または部分的ゲノム配列（例えば、ヒト対象由来の全ゲノム配列）である。参照配列は、例えばhG19またはhG38であり得る。シーケンシングされた核酸は、上記のように、試料中の核酸について直接決定された配列、またはそのような核酸の増幅産物の配列のコンセンサスを表

50

し得る。比較は、参照配列の1つまたは複数の指定位置で実施することができる。それぞれの配列を最大限に整列させた場合に、参照配列の指定位置に対応する位置を含むシーケンシングされた核酸のサブセットを識別することができる。そのようなサブセット内では、もしあれば、どのシーケンシングされた核酸が指定位置にヌクレオチド変異を含むか、および必要に応じて、もしあれば、どれが参照ヌクレオチド（例えば、参照配列におけるものと同じ）を含むかを決定することができる。ヌクレオチドバリエーションを含むサブセットにおけるシーケンシングされた核酸の数が選択された閾値を超える場合には、指定位置でバリエーションヌクレオチドをコールすることができる。閾値は、ヌクレオチドバリエーションを含むサブセット内のシーケンシングされた核酸の単純な数、例えば、少なくとも1、2、3、4、5、6、7、8、9、もしくは10個であり得るか、または他の可能性の中でも、ヌクレオチドバリエーションを含むサブセット内のシーケンシングされた核酸の比、例えば、少なくとも0.5、1、2、3、4、5、10、15、もしくは20であり得る。参照配列内の目的の任意の指定位置について比較を繰り返すことができる。時には、参照配列の少なくとも約20、100、200、または300の連続した位置、例えば、約20～500、または約50～300の連続した位置を占有する指定位置について比較を実施することができる。

10

【0242】

本明細書に記載されるフォーマットおよび適用を含む核酸シーケンシングに関する追加の詳細はまた、例えば、その各々の全体が参照により本明細書に組み込まれる、Levy et al., Annual Review of Genomics and Human Genetics, 17: 95-115 (2016)、Liu et al., J. of Biomedicine and Biotechnology, Volume 2012, Article ID 251364:1-11 (2012)、Voelkerding et al., Clinical Chem., 55: 641-658 (2009)、MacLean et al., Nature Rev. Microbiol., 7: 287-296 (2009)、Astier et al., J Am Chem Soc., 128(5):1705-10 (2006)、米国特許第6,210,891号、米国特許第6,258,568号、米国特許第6,833,246号、米国特許第7,115,400号、米国特許第6,969,488号、米国特許第5,912,148号、米国特許第6,130,073号、米国特許第7,169,560号、米国特許第7,282,337号、米国特許第7,482,120号、米国特許第7,501,245号、米国特許第6,818,395号、米国特許第6,911,345号、米国特許第7,501,245号、米国特許第7,329,492号、米国特許第7,170,050号、米国特許第7,302,146号、米国特許第7,313,308号、および米国特許第7,476,503号にも提供される。

20

30

IV. コンピュータシステム

【0243】

本開示の方法は、コンピュータシステムを使用してまたはその助けを借りてインプリメントすることができる。例えば、(a)断片サイズ対照分子の第1のサブセットを、無細胞ポリヌクレオチドの試料中の核酸分子に添加し、それによって第1のスプライクイン試料を産生するステップ；(b)第1のスプライクイン試料から核酸を抽出するステップ；(c)断片サイズ対照分子の第2のサブセットを抽出された核酸に添加し、それによって第2のスプライクイン試料を産生するステップ；(d)第2のスプライクイン試料の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、処理することが、第2のスプライクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；(e)断片サイズ対照分子の第3のサブセットを処理された試料に添加し、それによって第3のスプライクイン試料を産生するステップ；(f)第3のスプライクイン試料の少なくともサブセットを富化し、それによって富化された試料の

40

50

少なくともサブセットに添加し、それによって第4のスパイクイン試料を産生するステップ；(h)第4のスパイクイン試料をシーケンシングして、複数の配列リードを生成するステップ；(i)複数の配列リードを分析して、断片サイズ対照分子の複数の断片サイズスコアを生成するステップ；ならびに(j)複数の断片サイズスコアを複数の断片サイズ閾値と比較するステップ、を含み得るそのような方法は、コンピュータプロセッサによって実施することができる。この実施形態では、システムは、断片サイズ対照分子を付加する、分画する、増幅する、富化する、およびシーケンシングするための構成成分を含む。

【0244】

図5は、本開示の方法をインプリメントするようにプログラムされたまたはそれ以外の方法で構成されたコンピュータシステム501を示す。コンピュータシステム501は、試料調製、シーケンシング、および／または分析の様々な態様を調節することができる。一部の例では、コンピュータシステム501は、核酸シーケンシングを含む試料調製および試料分析を実施するように構成される。

【0245】

コンピュータシステム501は、中央演算処理装置(CPU、本明細書において同様に「プロセッサ」、および「コンピュータプロセッサ」)505を含み、これは、シングルコアもしくはマルチコアプロセッサ、または並列処理のための複数のプロセッサであり得る。コンピュータシステム501はまた、メモリまたはメモリ位置510(例えば、ランダムアクセスメモリ、読み取り専用メモリ、フラッシュメモリ)、電子ストレージユニット515(例えば、ハードディスク)、1つまたは複数の他のシステムと通信するための通信インターフェース520(例えば、ネットワークアダプター)、ならびにキャッシュ、他のメモリ、データストレージ、および／または電子ディスプレイアダプターなどの周辺機器525も含む。メモリ510、ストレージユニット515、インターフェース520、および周辺機器525は、CPU505と通信ネットワークまたはマザーボードなどのバス(実線)を通して通信する。ストレージユニット515は、データを記憶するためのデータストレージユニット(またはデータリポジトリ)であり得る。コンピュータシステム501は、通信インターフェース520の助けを借りてコンピュータネットワーク530に作動可能にカップリングすることができる。コンピュータネットワーク530は、インターネット、インターネットおよび／もしくはエクストラネット、またはインターネットと通信するイントラネットおよび／も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530は、一部の 경우에는、電気通信および／また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530は、1つまたは複数のコンピュータサーバーを含むことができ、それにより、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することができる。コンピュータネットワーク530は、一部の 경우에는、コンピュータシステム501の助けを借りてピアツーピアネットワークをインプリメントすることができ、それにより、コンピュータシステム501にカップリングしたデバイスがクライアントまたはサーバーとして動作することが可能となり得る。

【0246】

CPU505は、プログラムまたはソフトウェアに具体化することができる一連の機械可読命令を実行することができる。命令は、メモリ510などのメモリ位置に記憶させることができる。CPU505によって実施される操作の例は、フェッチ、復号、実行、およびライトバックを含み得る。

【0247】

ストレージユニット515は、ドライバー、ライブラリ、および保存されたプログラムなどのファイルを記憶し得る。ストレージユニット515は、ユーザーによって作成されたプログラムおよび記録されたセッション、ならびにプログラムに関連する出力(複数可)を記憶し得る。ストレージユニット515は、ユーザーデータ、例えば、ユーザーの好みおよびユーザープログラムを記憶し得る。コンピュータシステム501は、一部の 경우에는、例えば、イントラネットまたはインターネットを通じてコンピュータシステム501

と通信する遠隔サーバー上に位置するものなど、コンピュータシステム501に対して外付けの1つまたは複数の追加的なデータストレージユニットを含み得る。例えば通信ネットワークまたは物理的なデータ移行（例えば、ハードドライブ、サムドライブ、または他のデータストレージ機構を使用する）を使用して、データを1つの場所から別の場所に移行することができる。

【0248】

コンピュータシステム501は、ネットワーク530を通して1つまたは複数の遠隔コンピュータシステムと通信することができる。具体化のために、コンピュータシステム501は、ユーザー（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は、パーソナルコンピュータ（例えば、携10帯型PC）、スレートまたはタブレットPC（例えば、Apple（登録商標）iPad（登録商標）、Samsung（登録商標）Galaxy Tab）、電話機、スマートフォン（例えば、Apple（登録商標）iPhone（登録商標）、Android対応デバイス、Blackberry（登録商標））、または携帯情報端末が挙げられる。ユーザーは、ネットワーク530を介してコンピュータシステム501にアクセスすることができる。

【0249】

本明細書に記載の方法は、コンピュータシステム501の電子ストレージ位置、例えば、メモリ510または電子ストレージユニット515などに記憶された機械（例えば、コン20ピュータプロセッサ）により実行可能なコードによってインプリメントすることができる。機械により実行可能なまた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505により実行することができる。一部の場20合では、コードをストレージユニット515から検索し、プロセッサ505による容易なアクセスのためにメモリ510に記憶させることができる。いくつかの状況では、電子ストレージユニット515を除外することができ、機械により実行可能な命令はメモリ510に記憶される。

【0250】

一態様では、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、方法の少なくとも一部を実施する、コンピュータで実行可能な命令を含む非一時的コンピュ30ータ可読媒体であって、方法が（a）断片サイズ対照分子の第1のサブセットを、無細胞ポリヌクレオチドの試料中の核酸分子に添加し、それによって第1のスパイクイン試料を産生するステップ；（b）第1のスパイクイン試料から核酸を抽出するステップ；（c）断片サイズ対照分子の第2のサブセットを核酸に添加し、それによって第2のスパイクイン試料を産生するステップ；（d）第2のスパイクイン試料の少なくともサブセットを処理し、それによって処理された試料を産生するステップであって、処理することが、第2のスパイクイン試料の少なくともサブセットを分画、タグ付け、および／または増幅することを含む、ステップ；（e）断片サイズ対照分子の第3のサブセットを処理された試料に30添加し、それによって第3のスパイクイン試料を産生するステップ；（f）第3のスパイクイン試料の少なくともサブセットを富化し、それによって富化された試料を産生するステップ；（g）断片サイズ対照分子の第4のサブセットを富化された試料の少なくともサブ30セットに添加し、それによって第4のスパイクイン試料を産生するステップ；（h）第4のスパイクイン試料をシーケンシングして、複数の配列リードを生成するステップ；（i）複数の配列リードを分析して、断片サイズ対照分子の複数の断片サイズスコアを生成するステップ；ならびに（j）複数の断片サイズスコアを複数の断片サイズ閾値と比較するステップを含む、非一時的コンピュータ可読媒体を提供する。

【0251】

コードを、プリコンパイルし、コードを実行するように適合させたプロセッサを有する機械での使用のために構成することができるか、または実行時間の間にコンパイルすることができる。コードは、コードをプリコンパイル様式またはそのつどコンパイル（a s c o m p i l e d）様式で実行することが可能になるように選択することができるプログ50

ラミング言語で供給され得る。

【0252】

コンピュータシステム501などの本明細書に提供されるシステムおよび方法の態様は、プログラミングに具体化することができる。技術の様々な態様は、典型的には、機械可読媒体の一種で行われるまたは具体化される機械（またはプロセッサ）により実行可能なコードおよび／または関連データの形態の「製品」または「製造品」と考えることができる。機械により実行可能なコードは、電子ストレージユニット、そのようなメモリ（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）またはハードディスクに記憶させることができる。「ストレージ」型媒体は、ソフトウェアプログラミングのためにいつでも非一時的ストレージを提供し得る、コンピュータ、プロセッサな

10

【0253】

ソフトウェアの全てまたは一部を、時にインターネットまたは種々の他の電気通信ネットワークを通じて通信させることができる。そのような通信により、例えば、1つのコンピュータまたはプロセッサから別のコンピュータまたはプロセッサに、例えば、管理サーバーまたはホストコンピュータから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にソフトウェアをロードすることが可能となり得る。このように、ソフトウェア要素を有し得る別の型の媒体としては、有線および光固定電話ネットワークを通じて、ならびに様々なエアリンクによってローカルデバイス間で物理的インターフェースを横切っ

20

【0254】

したがって、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、有形記憶媒体、搬送波媒体または物理的伝送媒体を含むがこれらに限定されない多くの形態を取り得る。非揮発性記憶媒体としては、例えば光学または磁気ディスク、例えば図面に示されているデータベースなどをインプリメントするために使用することができるものなどの、任意のコンピュータなどにおけるストレージデバイスのいずれかが挙げられる。揮発性記憶媒体としては、そのようなコンピュータプラットフォームのメインメモリなどのダイナミックメモリが挙げられる。有形伝達媒体としては、同軸ケーブル；コンピュータシステム内のバスを含む電線を含む銅線および光ファイバーが挙げられる。搬送波伝送媒体は、高周波（RF）および赤外（IR）データ通信中に生成するものなどの、電気シグナルもしくは電磁気シグナル、または音波もしくは光波の形態を取り得る。したがって、コンピュータ可読媒体の一般的な形態としては、例えば、フロッピー（登録商標）ディスク、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もしくは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記憶媒体、RAM、ROM、PROMおよびEPROM、FLASH-EPROM、任意の他のメモリチップもしくはカートリッジ、搬送波伝達データもしくは命令、そのような搬送波を輸送するケーブルもしくはリンク、またはコンピュータがプログラミングコードおよび／もしくはデータを読み取ることができる任意の他の媒体が挙げられる。これらの形態のコンピュータ可読媒体の多くは、実行のためにプロセッサに1つまたは複数の命令の1つまたは複数のシーケンスを運ぶことに関与し得る。

30

40

【0255】

コンピュータシステム501は、例えば、試料分析の1つまたは複数の結果を提供するための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むか、またはそれと通信することができる。UIの例としては、グラフィカルユーザーインターフェース（GUI）およびウェブに基づくユーザーインターフェースが挙げられるがこれらに限定され

50

ない。

【0256】

コンピュータシステムおよびネットワーク、データベース、およびコンピュータプログラム製品に関する追加的な詳細は、例えば、各々の全体が参照により本明細書に組み込まれる、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)、およびRhoton, Cloud Computing Architected: Solution Design Handbook, Recursive Press (2011)においても提供される。

10

V. 適用

A. がんおよび他の疾患

【0257】

一部の実施形態では、本明細書に開示される方法およびシステムを使用して、カスタマイズされたまたは標的化治療を識別して、核酸バリエーションが体細胞または生殖細胞系列起源であるかの分類に基づいて、患者における所定の疾患または状態を処置することができる。典型的には、考慮される疾患は、がんの一種である。そのようながんの非限定的な例としては、胆道がん、膀胱がん、移行上皮癌、尿路上皮癌、脳がん、神経膠腫、星状細胞腫、乳癌、化生性癌、子宮頸がん、子宮頸部扁平上皮癌、直腸がん、結腸直腸癌、結腸がん、遺伝性非ポリポーシス結腸直腸がん、結腸直腸腺癌、消化管間質腫瘍（GIST）、子宮内膜癌、子宮内膜間質肉腫、食道がん、食道扁平上皮癌、食道腺癌、眼内黒色腫、ブドウ膜黒色腫、胆嚢癌、胆嚢腺癌、腎細胞癌、腎明細胞癌、移行上皮癌、尿路上皮癌、ウィルムス腫瘍、白血病、急性リンパ性白血病（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）、肝がん、肝癌、ヘパトーマ、肝細胞癌、胆管癌、肝芽腫、肺がん、非小細胞肺癌（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細胞性リンパ腫、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、上咽頭癌（NPC）、神経芽細胞腫、中咽頭がん、口腔扁平上皮癌、骨肉腫、卵巣癌、膵がん、膵管腺癌、偽乳頭状新生物、腺房細胞癌、前立腺がん、前立腺腺癌、皮膚がん、黒色腫、悪性黒色腫、皮膚黒色腫、小腸癌、胃がん、胃癌、消化管間質腫瘍（GIST）、子宮がん、または子宮肉腫が挙げられる。

20

30

【0258】

本明細書に開示される方法およびシステムを使用して必要に応じて評価される他の遺伝子に基づく疾患、障害、または状態の非限定的な例としては、軟骨無形成症、アルファ1アンチトリプシン欠損症、抗リン脂質症候群、自閉症、常染色体優性多発性嚢胞腎、シャルコー・マリー・トゥース病（CMT）、ネコ鳴き症候群、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因子ライデン血栓形成傾向、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نفelter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、重症複合免疫不全症（scid）、鎌状赤血球症、脊髄

40

50

性筋萎縮症、テイ・サックス病、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、W A G R 症候群、ウィルソン病などが挙げられる。

B. 治療および関連する投与

【0259】

ある特定の実施形態では、本明細書に開示される方法は、体細胞または生殖細胞系列起源である核酸バリエーションの状況を考慮して、患者に対するカスタマイズされた治療を識別し、投与することに関する。一部の実施形態では、本質的に任意のがん治療（例えば、外科療法、放射線療法、化学療法、および／または同様のもの）が、これらの方法の一部として含まれ得る。典型的には、カスタマイズされた治療は、少なくとも1つの免疫療法（または免疫療法剤）を含む。免疫療法は一般的に、所定のがんの型に対する免疫応答を増強する方法を指す。ある特定の実施形態では、免疫療法は、腫瘍またはがんに対するT細胞応答を増強する方法を指す。

10

【0260】

ある特定の実施形態では、体細胞または生殖細胞系列起源である対象の試料由来の核酸バリエーションの状況を、参照集団からの対照薬の結果のデータベースと比較して、その対象に関するカスタマイズされたまたは標的化治療を識別することができる。典型的には、参照集団は、試験対象と同じがんもしくは疾患の型を有する患者、および／または試験対象と同じ治療を受けているもしくは受けたことがある患者を含む。核バリエーションおよび対照薬の結果がある特定の分類基準を満たす（例えば、実質的にまたはほぼマッチする）場合、カスタマイズされたまたは標的化治療（または複数の治療）が識別され得る。

20

【0261】

ある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は、典型的には非経口投与される（例えば、静脈内または皮下）。免疫療法剤を含有する医薬組成物は、典型的には静脈内投与される。ある特定の治療剤は経口投与される。しかし、カスタマイズされた治療（例えば、免疫療法剤など）をまた、例えば頬側、舌下、直腸、膣、尿道内、局部、眼内、鼻腔内、および／または耳介内などの方法によって投与してもよく、投与は、錠剤、カプセル剤、顆粒剤、水性懸濁剤、ゲル剤、噴霧剤、坐剤、膏薬、軟膏剤などを含み得る。

【0262】

本発明の好ましい実施形態が本明細書において示され、記載されているが、そのような実施形態が単に例として提供されていることは当業者には明白であろう。本発明が、本明細書内に提供される特定の実施例によって限定されることは意図されない。本発明を前述の明細書に関連して説明してきたが、本明細書の実施形態の説明および図表は、限定の意味で解釈されることを意味するものではない。当業者は、本発明から逸脱することなく多数の変形形態、変化および置換をすぐに思いつくであろう。さらに、本発明の全ての態様は、多様な条件および変数に依存する本明細書に記載の特定の描写、構成または相対的な割合に限定されないと理解すべきである。本明細書に記載の本開示の実施形態に対する様々な代替を本発明の実施に用いることができると理解すべきである。したがって、本開示はまた、任意のそのような代替、改変、変形形態、または均等物も網羅するものとすることが企図される。以下の特許請求の範囲は、本発明の範囲を定義し、これらの特許請求の範囲の範囲内の方法および構造ならびにそれらの均等物がそれにより網羅されると意図される。

30

40

【0263】

前述の開示は、明瞭さおよび理解の目的のために実例としておよび例としていくぶん詳細に説明してきたが、本開示を読むことにより、形態および詳細の様々な変化を本開示の真の範囲から逸脱することなく行うことができ、添付の特許請求の範囲の範囲内で実施することができることが当業者には明らかであろう。例えば、方法、システム、コンピュータ可読媒体、および／または構成要素の特色、ステップ、要素、またはそれらの他の態様は全て、様々な組合せで使用することができる。

【0264】

50

本明細書において引用されている特許、特許出願、ウェブサイト、他の刊行物または文書、受託番号などは全て、各個々の項目が参照により組み込まれることが具体的にかつ個別に示されるのと同程度に、全ての目的に関してその全体が参照により組み込まれる。異なるバージョンの配列が異なる時期に受託番号に関連付けられている場合、本出願の有効な出願日に受託番号に関連付けられたバージョンが意図される。有効な出願日とは、実際の出願日または該当する場合には受託番号を参照する優先出願の出願日のいずれか早いほうを意味する。同様に、異なるバージョンの刊行物、ウェブサイトなどが異なる時期に公開されている場合、特に指定のない限り、本出願の有効な出願日に最も近い日に公開されたバージョンが意図される。

【図面】

【図 1 - 1】

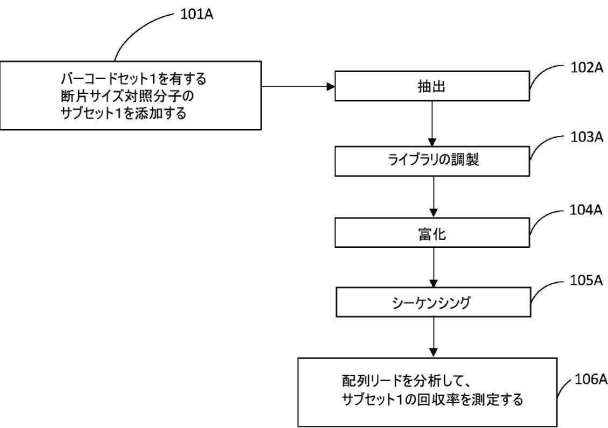


FIG. 1A

【図 1 - 2】

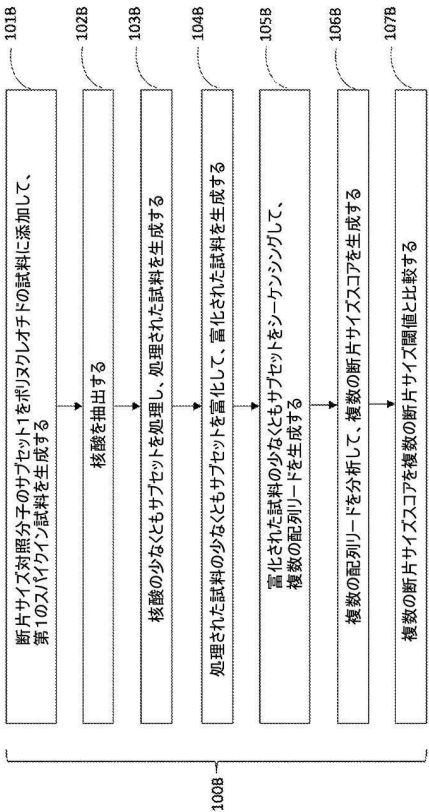


FIG. 1B

10

20

30

40

50

【図 2 - 1】

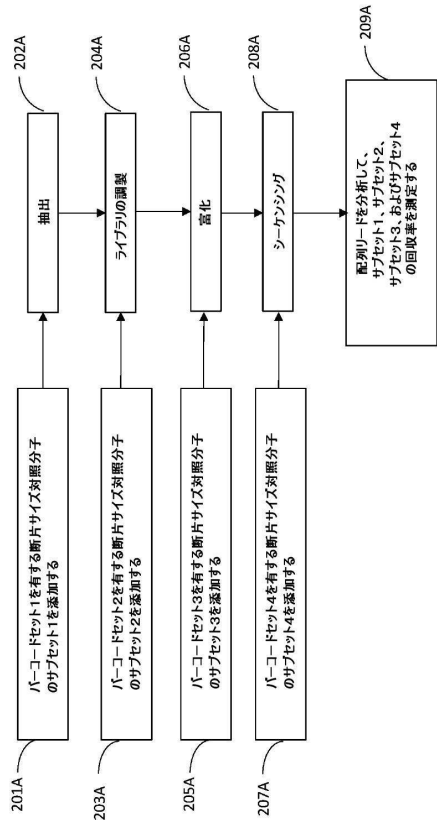


FIG. 2A

【図 2 - 2】

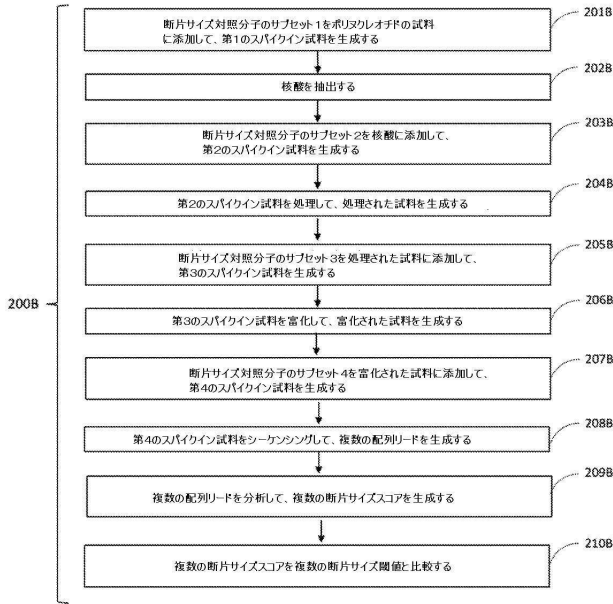


FIG. 2B

【図 3】

サブセット 1	群 1	MB1-S1-Pr1 MB3-S1-Pr1 MB2 Pr2-MB2 Pr2-MB4
	群 2	MB5-S1-Pr3 MB7-S1-Pr3 Pr4-MB6 Pr4-MB8
	群 3	MB9-S1-Pr5 MB11-S1-Pr5 Pr6-MB10 Pr6-MB12
サブセット 2	群 1	MB13-S2-Pr7 MB15-S2-Pr7 Pr8-MB14 Pr8-MB16
	群 2	MB17-S2-Pr9 MB19-S2-Pr9 Pr10-MB18 Pr10-MB20
	群 3	MB21-S2-Pr11 MB23-S2-Pr11 Pr12-MB22 Pr12-MB24

FIG. 3

【図 4】

サブセット 1	群 1	MB1-S1 MB3-S1 MB2 MB4
	群 2	MB5-S1 MB7-S1 MB6 MB8
	群 3	MB9-S1 MB11-S1 MB10 MB12
サブセット 2	群 1	MB13-S2 S2-MB15 MB14 MB16
	群 2	MB17-S2 MB19-S2 MB18 MB20
	群 3	MB21-S2 MB23-S2 MB22 MB24

FIG. 4

10

20

30

40

50

【 図 5 】

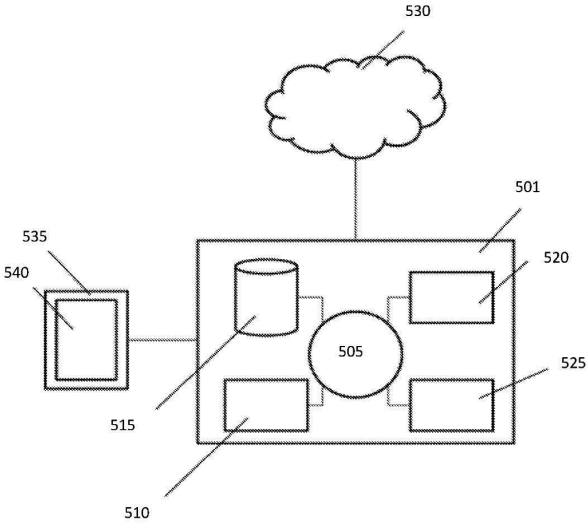


FIG. 5

10

20

30

40

50

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2019/068146

A. CLASSIFICATION OF SUBJECT MATTER INV. C12N15/10 C12Q1/6806 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) C12Q C40B C12N		
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, BIOSIS, Sequence Search, EMBASE, 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 2017/165864 A1 (KARIUS INC [US]) 28 September 2017 (2017-09-28) see whole doc., esp. claims, Figure 4 and examples 2-7, par. 191 -----	1-131
X	WO 2015/073080 A1 (LIFE TECHNOLOGIES CORP [US]) 21 May 2015 (2015-05-21) see whole doc. esp. claims, par. 56 ----- -/-	1-131
☒ 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 25 May 2020		Date of mailing of the international search report 03/06/2020
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 Mueller, Frank

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

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2019/068146

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MICHAEL A QUAIL ET AL: "SASI-Seq: sample assurance Spike-Ins, and highly differentiating 384 barcoding for Illumina sequencing", BMC GENOMICS, BIOMED CENTRAL, vol. 15, no. 1, 7 February 2014 (2014-02-07), page 110, XP021182710, ISSN: 1471-2164, DOI: 10.1186/1471-2164-15-110 the whole document	1-131
X	MAURO D. LOCATI ET AL: "Improving small RNA-seq by using a synthetic spike-in set for size-range quality control together with a set for data normalization", NUCLEIC ACIDS RESEARCH, vol. 43, no. 14, 18 August 2015 (2015-08-18), pages e89-e89, XP55386708, ISSN: 0305-1048, DOI: 10.1093/nar/gkv303 the whole document	1-131

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2019/068146

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
W0 2017165864 A1	28-09-2017	AU 2017237199 A1	16-08-2018
		BR 112018069557 A2	29-01-2019
		CA 3014070 A1	28-09-2017
		CN 109312400 A	05-02-2019
		EP 3433382 A1	30-01-2019
		JP 2019509068 A	04-04-2019
		KR 20180129844 A	05-12-2018
		US 2017275691 A1	28-09-2017
		US 2018237851 A1	23-08-2018
		W0 2017165864 A1	28-09-2017

W0 2015073080 A1	21-05-2015	EP 3068896 A1	21-09-2016
		US 2015133314 A1	14-05-2015
		US 2019284632 A1	19-09-2019
		W0 2015073080 A1	21-05-2015

フロントページの続き

(51)国際特許分類	F I	テーマコード (参考)
C 1 2 Q 1/686(2018.01)	C 1 2 Q 1/6844	Z
C 1 2 Q 1/6851(2018.01)	C 1 2 Q 1/686	Z
C 1 2 Q 1/6869(2018.01)	C 1 2 Q 1/6851	Z
	C 1 2 Q 1/6869	Z

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,
CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IR,IS,JO,JP,KE,K
G,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,N
I,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,SM,ST,SV,SY,TH,TJ,TM,
TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,ZA,ZM,ZW

(特許庁注：以下のものは登録商標)

1. ANDROID

弁護士 山本 健策

(72)発明者 モーティマー, ステファニー アン ウォード

アメリカ合衆国 カリフォルニア 9 5 0 3 7, モーガン ヒル, ウィロー スプリングス ロード
2 0 0 0

(72)発明者 ハイト, ダスティン ハワード

アメリカ合衆国 カリフォルニア 9 4 0 6 3, レッドウッド シティ, ペノブスコット ドライブ
5 0 5

(72)発明者 メルロイ, ローラ ミシェル

アメリカ合衆国 カリフォルニア 9 4 0 6 3, レッドウッド シティ, ペノブスコット ドライブ
5 0 5

F ターム (参考) 4B063 QA01 QA13 QA18 QA20 QQ42 QQ52 QR06 QR08 QR32 QR42
QR56 QR62 QS10 QS24 QS25 QS34 QS39 QX02 QX04