

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2017-530693

(P2017-530693A)

(43) 公表日 平成29年10月19日(2017.10.19)

(51) Int. Cl.	F I	テーマコード (参考)
<b>C12Q 1/68 (2006.01)</b>	C12Q 1/68	A 2G045
<b>G01N 33/53 (2006.01)</b>	G01N 33/53	M 4B063
<b>G01N 33/48 (2006.01)</b>	G01N 33/48	Z

審査請求 未請求 予備審査請求 未請求 (全 39 頁)

(21) 出願番号 特願2017-506897 (P2017-506897)  
 (86) (22) 出願日 平成27年8月4日 (2015.8.4)  
 (85) 翻訳文提出日 平成29年3月29日 (2017.3.29)  
 (86) 国際出願番号 PCT/US2015/043609  
 (87) 国際公開番号 W02016/022559  
 (87) 国際公開日 平成28年2月11日 (2016.2.11)  
 (31) 優先権主張番号 62/035,019  
 (32) 優先日 平成26年8月8日 (2014.8.8)  
 (33) 優先権主張国 米国 (US)

(71) 出願人 513322707  
 ナノストリング テクノロジーズ, インコーポレイティド  
 アメリカ合衆国, ワシントン 98109  
 , シアトル, フェアビュー アベニュー ノース 530, スイート 2000  
 (74) 代理人 100099759  
 弁理士 青木 篤  
 (74) 代理人 100077517  
 弁理士 石田 敬  
 (74) 代理人 100087871  
 弁理士 福本 積  
 (74) 代理人 100087413  
 弁理士 古賀 哲次

最終頁に続く

(54) 【発明の名称】 遺伝子発現データを使用した混成細胞集団のデコンボリューション方法

(57) 【要約】

mRNA プロファイリングによる体液の同定は、法医学サンプルから状況的「活性レベル」情報の抽出を可能にし得る。従って、法医学的な体液/組織同定のための基本形の複合デジタル遺伝子発現法は、色分けされたプローブ（例えば、NanoString（登録商標））の溶液ハイブリダイゼーションに基づいて提供される。例えば、単一の体液からのサンプルにおける遺伝子発現に関するモデルが提供され、そして、体液の混合物へと拡張される。サンプル中の体液量の最尤推定値の算出が実施され、そしてサンプル中の各体液の存在について試験するための尤度比の使用が説明される。方法/アルゴリズムが、説明されるが、組織および細胞の検出のための在来型アルゴリズムと異なり、複数のサンプルを通じて偽陽性体液の同定をゼロにすることを可能にできる。斯かるプロトコールは、以前は信頼できなかったケースワーク（例えば、法医学的）検査室におけるmRNAプロファイリングの日常的な使用を容易にし得る。

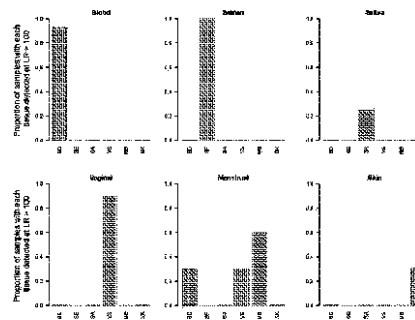


FIGURE 5

## 【特許請求の範囲】

## 【請求項 1】

法医学的な生物学的サンプルを同定するための方法であって：

分析のための少なくとも 1 つの生物学的サンプルを得；

生物学的サンプルから全 RNA を抽出し；

少なくとも 1 つのアッセイにおいて少なくとも 1 つのプローブと全 RNA とをハイブリダイズさせ；そして

複数のコードセットを使用した少なくとも 1 つのアッセイを分析すること、  
を含み、ここで、分析が、

サンプル中で定量化する 1 セットの遺伝子を決定し；

該遺伝子セットの各遺伝子について遺伝子発現対数関数を作成することによって該遺伝子セット中の各遺伝子の遺伝子発現をモデル化し；そして

該遺伝子セットの各遺伝子のモデル化された遺伝子発現に基づいて生物学的サンプル中の生物学的物質の量の最尤推定値を作成することを含む、方法。

10

## 【請求項 2】

前記生物学的サンプルが、組織サンプルである、請求項 1 に記載の方法。

## 【請求項 3】

前記物質が、皮膚、静脈血、腔分泌物、唾液、月経血、精液、およびバイオ粒子のうちの少なくとも 1 つである、請求項 1 に記載の方法。

20

## 【請求項 4】

前記生物学的サンプルが、少なくとも 2 つの生物学的物質を含み得る、請求項 1 に記載の方法。

## 【請求項 5】

前記全 RNA が、精製を伴った直接溶解および精製を伴わない直接溶解のうちの少なくとも一方を使用して生物学的サンプルから抽出される、請求項 1 に記載の方法。

## 【請求項 6】

生物学的サンプルからの全 RNA の抽出が、約 5 分間、75℃にて生物学的サンプルを溶解することを含む、請求項 5 に記載の方法。

## 【請求項 7】

前記少なくとも 1 つのプローブが、レポータープローブおよび捕捉プローブのうちの少なくとも 1 つを含む、請求項 1 に記載の方法。

30

## 【請求項 8】

前記複数のコードセットが、遺伝子セットを標的とするプローブ対を指定する、請求項 1 に記載の方法。

## 【請求項 9】

前記複数のコードセットが：

静脈血の遺伝子 ALAS2、ALOX5AP、AM1CA1、ANK1、AQP9、ARHGAP26、C1QR1、C5R1、CASP2、CD3G、GYPA、HBA、HBB、HMBS(PBGD)、MNDA、NCFS2、および SPTB；

月経血の遺伝子 LEFTY2、MMP7、MMP10、および MMP11；

40

唾液の遺伝子 HTN3、MUC7、S. ミュータンス 16S、S. ミュータンス proC、S. ミュータンス relA、S. ミュータンス rplA、S. ミュータンス rpoB、S. ミュータンス rpoS、S. サリバリウス 16S、S. サリバリウス proC、S. サリバリウス relA、S. サリバリウス rplA、S. サリバリウス rpoB、S. サリバリウス rpoS、SMR3B、および STATH；

精液の遺伝子 IZUMO1、MSP、PSA(KLK3)、PRM1、PRM2、SEMG1、SEMG2、および TGM4；

皮膚の遺伝子 CCL27、IL1F7、KRT9、LCE1C、および LCE2D；

腔分泌物の遺伝子 CYP2A7、CYP2B7P1、DKK4、FUT6、IL19、MYOZ1、および NOXO1；ならびに

50

参照遺伝子 B 2 M、C O X 1、H P R T 1、P G K 1、P P I H、S 1 5、T C E A 1、T F R C、U B C、および U B E 2 D 2、  
のうちの少なくとも 1 つを含む、請求項 1 に記載の方法。

【請求項 1 0】

前記複数のコードセットが、陽性対照プローブおよび陰性対照プローブのうちの少なくとも 1 つを含む、請求項 1 に記載の方法。

【請求項 1 1】

前記陰性対照プローブが、分析におけるバックグラウンドノイズを評価するために使用される、請求項 1 0 に記載の方法。

【請求項 1 2】

前記遺伝子発現対数関数が、以下の関数：

【化 1】

$$\log(y_i) \sim N(\log(X\beta_i), \sigma^2 I),$$

{ 式中、 $y_i$  は、生物学的サンプルの遺伝子発現プロフィールであり、 $N$  は、遺伝子セットの量であり、 $X$  は、複数の生物学的物質中の複数の遺伝子の予想される割合を表す行列であり、 $\beta_i$  は、生物学的物質  $i$  中のすべての生物学的物質の量を表すベクトルであり、 $\sigma^2$  は、複数の遺伝子における全遺伝子の対数スケールの一般分散であり、および  $I$  は、単位行列である } を使用してモデル化される、請求項 1 に記載の方法。

【請求項 1 3】

前記最尤推定値が、以下の関数：

【化 2】

$$\hat{\beta}_i = \underset{\beta}{\operatorname{argmin}} \|\log(y_i) - \log(X\beta)\|_2^2 \text{ s.t. } \beta \geq 0.$$

を使用して作成される、請求項 1 に記載の方法。

【請求項 1 4】

少なくとも 1 つの生物学的サンプル中の物質の存在を推定する方法であって：

生物学的サンプル内で検出するための 1 セットの生物学的物質を決定し；

該生物学的物質セット中の各生物学的物質について、生物学的物質の 1 セットの独特な遺伝子中の各遺伝子の発現をモデル化し；

生物学的物質の独特な遺伝子セット中の各遺伝子のモデル化された発現を使用して予想される遺伝子割合モデルを作成し；

生物学的サンプル内の生物学的物質セットの大量の各生物学的物質を含む物質モデルを作成し；

予想される遺伝子割合モデルと物質モデルとを使用することによって予想される遺伝子発現モデルを作成し；

予想される遺伝子発現モデルを使用して生物学的サンプル中で発現される遺伝子を推定し；

生物学的サンプルにおける推定遺伝子発現を使用して生物学的物質セットにおける各生物学的物質の最尤推定値 ( M L E ) に基づく推定サンプルプロフィールを作成し；

生物学的物質セットにおける各生物学的物質について尤度比を算出し、そして、該尤度比が、該生物学的物質が生物学的サンプル中に含まれている可能性がどの程度か示し；そして、

算出された尤度比に基づいて生物学的物質セットにおける各生物学的物質が生物学的サ

10

20

30

40

50

ンプル中に存在するか否か判定することを含む。

【請求項 15】

前記生物学的サンプルが、組織サンプルである、請求項 14 に記載の方法。

【請求項 16】

前記生物学的物質セット中の各生物学的物質が、皮膚、静脈血、腔分泌物、唾液、月経血、精液、およびバイオ粒子のうち少なくとも 1 つである、請求項 14 に記載の方法。

【請求項 17】

前記生物学的物質セット中の各生物学的物質の独特な遺伝子セットにおける各遺伝子のモデル化された発現が、生物学的物質セット中の各生物学的物質に関する遺伝子発現ベクトルとして表され、ここで、該遺伝子発現ベクトルが：

【化 3】

$$y_i = (y_{i1}, \dots, y_{ip})^T$$

{ 式中、 $y_{ij}$  は、生物学的物質  $i$  中の独特な遺伝子セット中の遺伝子  $j$  の発現に値する } と表される、請求項 14 に記載の方法。

【請求項 18】

前記予想される遺伝子割合モデルが、生物学的物質セット中の各生物学的物質に関する各遺伝子発現ベクトルを含む予想される遺伝子割合の行列である、請求項 17 に記載の方法。

【請求項 19】

前記物質モデルが物質ベクトルであり、かつここで、予想される遺伝子発現モデルを、予想される遺伝子割合モデルに物質ベクトルを乗じることによって作成している、請求項 14 に記載の方法。

【請求項 20】

前記遺伝子発現モデルが、関数：

【化 4】

$$\log(y_i) \sim N(\log(X\beta_i), \sigma^2 I),$$

{ 式中、 $y_i$  は、生物学的サンプル  $i$  中の生物学的物質セット中の各生物学的物質に関する独特な遺伝子セットの各遺伝子のモデル化された発現であり、 $N$  は、独特な遺伝子セットの遺伝子量であり、 $X$  は、予想される遺伝子割合モデルであり、 $\beta_i$  は、生物学的サンプル  $i$  の生物学的物質の割合のモデルであり、 $I$  は単位行列であり、および  $\sigma^2$  は、生物学的サンプルセット中の各生物学的サンプルに関する独特な遺伝子セットの各遺伝子の平均分散である } によって表される、請求項 14 に記載の方法。

【請求項 21】

前記生物学的物質セット中の各生物学的物質の MLE が、各生物学的サンプルに関する独特な遺伝子セットの各遺伝子に関する観察された遺伝子発現と、予想される遺伝子発現モデルから得られた各生物学的サンプルに関する独特な遺伝子セットの各遺伝子に関する予想される遺伝子発現の差の合計である、請求項 14 に記載の方法。

【請求項 22】

前記生物学的物質セット中の各生物学的物質の MLE が、関数：

10

20

30

40

【化 5】

$$\hat{\beta}_i = \operatorname{argmin}_{\beta} \|\log(\mathbf{y}_i) - \log(\mathbf{X}\beta)\|_2^2 \text{ s.t. } \beta \geq \mathbf{0},$$

{ 式中、生物学的物質セット中に負の量の各生物学的物質が存在しないとき、

【化 6】

10

$$\hat{\beta}_i$$

は、各生物学的サンプルに関する独特な遺伝子セットの各遺伝子について観察された遺伝子発現  $y_i$  と、各生物学的サンプルに関する独特な遺伝子セットの各遺伝子について予想される遺伝子発現  $X$  との間で 2 乗誤差の合計を最小化する } によって算出される、請求項 2 1 に記載の方法。

【請求項 2 3】

前記尤度比が、関数：

20

【化 7】

$$\loglik(\mathbf{y}_i | \hat{\beta}_i) =$$

$$-\frac{1}{2} \log(\det(\sigma^2 \mathbf{I})) - \frac{1}{2} (\log(\mathbf{y}_i) - \log(\mathbf{X}\hat{\beta}_i))^T \sigma^{-2} \mathbf{I} (\log(\mathbf{y}_i) - \log(\mathbf{X}\hat{\beta}_i))$$

30

を使用して、生物学的サンプル中の生物学的物質の存在の尤度と、生物学的サンプル中の生物学的物質の不存在の尤度との比を算出することによって表され、

ここで、生物学的サンプル中の生物学的物質の存在の尤度が、関数内の M L E を使用して算出され；

ここで、生物学的サンプル中の生物学的物質の不存在の尤度が、関数内の制約された M L E を使用して算出される、請求項 1 4 に記載の方法。

【請求項 2 4】

前記制約された M L E が、生物学的サンプル中の生物学的物質の量をゼロに設定したときに算出された M L E である、請求項 2 3 に記載の方法。

【請求項 2 5】

請求項 1 ~ 2 4 のいずれか 1 項に記載の方法を実施するために構成されたシステム。

40

【請求項 2 6】

前記システムが、請求項 1 ~ 2 4 のいずれか 1 項に記載の方法の 1 若しくは複数のステップを実施するためのコンピュータプロセッサを含む、請求項 2 5 に記載のシステム。

【請求項 2 7】

前記生物学的物質セット中の各生物学的物質の M L E が、関数：

【化 8】

$$S = \operatorname{argmin}_{\beta} \{ \|(\log(y) - \log(X\beta))^T \Sigma^{-1} (\log(y) - \log(X\beta))\|_p + \text{ペナルティ}(\beta) \}$$

{ 式中、S は、生物学的サンプル中の生物学的物質セットに関する 1 セットの M L E 値であり、式中、ペナルティ( ) は、 の要素に対する更なるペナルティを表し、および式中、関数は、 の要素が負にならないように制約される } によって算出される、請求項 2 1 に記載の方法。

10

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の相互参照

本出願は、2014年8月8日に出願された米国仮出願番号第62/035,019号の利益を主張する。前述の特許出願の内容は、その全体が参照によって本明細書中に援用される。

【背景技術】

【0002】

本発明の背景

生物学的サンプルは、異なるタイプの物質（例えば、腫瘍細胞と正常細胞、複数の微生物の混合物、異なる生体液の混合物、免疫細胞の混合物、および/または同類のものなどの異なるタイプの細胞）の混合物を含んでいることが多い。一般的に、デコンボリューションは、物質内の既知の遺伝子発現パターンに基づいて所定のサンプル中の物質の割合を推定するために、および/または所定のサンプル中に既知の物質比で与えられた各タイプの物質内の平均遺伝子発現プロフィールを推定するために使用される。

20

【0003】

従来のデコンボリューション方法は、サンプル混合物データの加法モデル： $E(Y) = XB$ 、を仮定することが多く、ここで、Y は n 個のサンプルと p 個の遺伝子における遺伝子発現の  $n \times p$  行列であり、X は K 細胞型における p 個の遺伝子の典型的な遺伝子発現の  $p \times K$  行列であり、および B は各サンプル中の各細胞型の量の  $n \times K$  行列である。加法モデルでは通例、サンプル中の遺伝子転写物の量が各サンプルの細胞垂集団中の転写産物の量であると仮定する。さらに、加法モデルを使用することによって、事前の実験が細胞型の典型的な遺伝子発現プロフィール X に関する推定を可能にするのであれば、X と Y から細胞型の量 B の行列を推定することが可能である。あるいは、（例えば、発現プロファイリングの前にセルソーターを通してサンプルを流すことによって）B が既知であれば、各細胞型の平均発現プロフィールが推定され得る。

30

あるサンプルタイプに独特であると予想される遺伝子の同一性や同定可能性を確実にするためのパラメーターに対する制約条件のような事前情報の導入によって、一部の科学者は B と X を同時に推定するために伝統的にこのモデルを使用した。

40

【0004】

しかしながら、加法モデルは多くの方法で問題が多い。例えば、遺伝子発現データは、（既に対数スケールで存在している q P C R データを除いて）分析前には対数変化していることが多く、そして一般的に、示差的発現は加法的増加ではなく、倍数変化で計測される。データを変換することによって、および/またはそれを加法モデルに組み込むための様式でそれを利用することによって、精度が失われる可能性があり、そして不正確な結果（例えば、サンプル中の物質の偽陽性および/または偽陰性、あるいは混合割合および/または細胞型遺伝子発現プロフィールの非効率な推定）をもたらす。

【発明の概要】

【0005】

50

本明細書中に開示した方法は、より正確な遺伝子発現を算出するために加法モデルと対数ベースの算出の両方を使用したデコンポリューション方法を説明する。この利便は、これだけに限定されるものではないが、法医学分析で遭遇する体液混合物および/または類似したサンプル混合物を含めたサンプル混合物について分析するとき、著しく有益なものであると予想される。特に、対数または乗法スケールと加法モデルを使用した統計的方法を本明細書中に記載し、そしてそれは、サンプル中の種々の標的遺伝子の遺伝子発現に基づいてサンプル中の所定の体液の量を算出できる。

【0006】

いくつかの実施形態において、法医学的な生物学的サンプルを同定するための方法は、分析のための少なくとも1つの生物学的サンプルを得、生物学的サンプルから全RNAを抽出し、少なくとも1つのアッセイにおいて少なくとも1つのプローブと全RNAとをハイブリダイズさせ、そして複数のコードセットを使用した少なくとも1つのアッセイを分析することを含み得る。分析の実施の一部において、アッセイは、サンプル中で定量化するための1セットの遺伝子を決定し、該遺伝子セットの各遺伝子について遺伝子発現対数関数を作成することによって該遺伝子セット中の各遺伝子の遺伝子発現をモデル化し、そして該遺伝子セットの各遺伝子のモデル化された遺伝子発現に基づいて生物学的サンプル中の生物学的物質の量の最尤推定値を作成することを含み得る。

10

【0007】

いくつかの実施形態において、少なくとも1つの生物学的サンプル中の物質の存在を推定する方法は、生物学的サンプル内で検出するための1セットの生物学的物質を決定し、該生物学的物質セット中の各生物学的物質について生物学的物質の1セットの独特な遺伝子中の各遺伝子の発現をモデル化し、そして生物学的物質の独特な遺伝子セット中の各遺伝子のモデル化された発現を使用して予想される遺伝子割合モデルを作成することを含み得る。いくつかの実施形態において、該方法は、生物学的サンプル内の生物学的物質セットの大量の各生物学的物質を含む物質モデルを作成し、予想される遺伝子割合モデルと物質モデルとを使用することによって予想される遺伝子発現モデルを作成し、そして予想される遺伝子発現モデルを使用して生物学的サンプル中で発現される遺伝子を推定することをさらに含んでもよい。さらに、該方法は、生物学的サンプルにおける推定遺伝子発現を使用して生物学的物質セットにおける各生物学的物質の最尤推定値に基づく推定サンプルプロフィールを作成し、生物学的物質セットにおける各生物学的物質について尤度比(likelihood ratio)を算出し、ここで、該尤度比とは、該生物学的物質が生物学的サンプル中に含まれている可能性がどの程度か示しており、そして算出された尤度比に基づいて生物学的物質セットにおける各生物学的物質が生物学的サンプル中に存在するか否か判定することを含み得る。

20

30

【0008】

いくつかの実施形態において、本明細書中に記載した装置、方法、およびシステムは、5分間直接細胞溶解プロトコール(five minute direct lysis protocol)した特異的mRNA標的へのバーコードプローブの多重溶液ハイブリダイゼーションによって、一般的な法医学的に関連している体液および/またはさまざまなサンプル中に存在する可能性があるさまざまな物質を同定できる。最小限の実地の要件を有するこの簡易型プロトコールは、ケースワーク検査室におけるmRNAプロファイリングの日常的な使用を容易にし得る。大抵の遺伝子発現に基づく選別装置とは対照的に、アルゴリズムは、正しくサンプルを呼び出す能力を最適化するように機械学習アルゴリズムをトレーニングすることを伴ってはならない；むしろ、それは、体液サンプルにおける遺伝子発現の生物学的に合理的なモデルを規定してもよく、そして、特定の体液の存在を証明するサンプルの根拠力を評価するモデルを使用する。このアルゴリズムは、各体液タイプの検出のための対数尤度の算出を可能にし、法廷の場においてアルゴリズムの結果をより弁護可能にし得る。

40

【0009】

本開示のいくつかの実施形態によるアプローチの更なる恩恵は、それが、トレーニングに使用されるものを含めたすべてのサンプルに対してアルゴリズムの評価を可能にする点

50

である：アルゴリズムが体液混合物における遺伝子発現の従来モデルに基づく場合、そのパラメーターがモデル性能に関係ないと推定され得るので、該アルゴリズムは、トレーニングデータの過剰適合を最小限にし得る。

【0010】

いくつかの実施において、本明細書中に記載した装置、方法、およびシステムは、遺伝子発現データ、タンパク質データ、代謝産物データ、およびmiRNA発現データ、および/または対数スケールの変動性に伴うその他のデータにも適用され得る。いくつかの実施形態において、ここで記載した方法のアウトプットは、分類、クラスタリング、および/または他の機械学習の問題に使用できる。いくつかの実施形態において、ここで記載した方法は、サンプルまたはクラス間の示差的な遺伝子発現について試験するために使用できる。いくつかの実施形態において、ここで記載した方法は、サンプルタイプにおける遺伝子の発現について試験するために使用できる。

10

【0011】

好ましい実施形態において、NanoString Technologies (登録商標) のnCounter (登録商標) システムおよび方法が使用されている。特定のmRNA標的を結合および同定するためのプローブおよび方法は、例えばUS 2003/0013091、US 2007/0166708、US 2010/0015607、US 2010/0261026、US 2010/0262374、US 2010/0112710、US 2010/0047924、およびUS 2014/0371088に記載されており、そのそれぞれは、その全体が参照によって本明細書中に援用される。

20

【0012】

本明細書中に記載したいずれの態様または実施形態も、本明細書中に開示されるいずれの他の態様または実施形態とも組み合わせることができる。開示はその詳細な説明と関連づけて記載された一方で、上記の説明は例示することを意図したものであって、添付の請求項の範囲によって規定される開示の範囲を制限することを意図していない。他の態様、利点、および修飾も以下の請求項の範囲内にある。

【図面の簡単な説明】

【0013】

特許または出願ファイルは、カラーで仕上げられた少なくとも1枚の図面を含んでいる。カラー図面を伴った本願の特許または特許出願公開のコピーは、必要な料金の請求と支払いにより特許庁によって提供される。

30

【0014】

【図1】図1は、いくつかの例示的な実施形態における各組織についてのアルゴリズムの真陽性率(TPR)と偽陽性率(FPR)を示す代表的なROCカーブを示す。

【0015】

【図2】図2は、いくつかの例示的な実施形態における5つの混合物サンプルでのアルゴリズムの代表的な性能結果を示す。

【0016】

【図3】図3は、いくつかの例示的な実施形態におけるサンプル組成の算出を例示する論理フローチャートを示す。

40

【0017】

【図4】図4は、本明細書中に開示した直接溶解プロトコールに従って調製したサンプルの代表的な性能結果と、本明細書中に開示した精製プロトコールに従って調製したサンプルの代表的な性能結果との比較を示す。

【0018】

【図5】図5は、いくつかの例示的な実施形態における91個の単一起源サンプルのアルゴリズムの代表的な性能結果を示す。

【0019】

【図6】図6は、いくつかの例示的な実施形態における23個の単一起源の、適当なRNAサンプルのアルゴリズムの代表的な性能結果を示す。

50

## 【 0 0 2 0 】

【図 7 A - F】図 7 A ~ F は、同じ体液タイプの異なるサンプルの遺伝子発現プロファイルを示す一連のプロットを示す。図 7 A は、血液 ( B D ) の遺伝子発現プロファイルの一貫性を示す。図 7 B は、精液 ( S E ) の遺伝子発現プロファイルの一貫性を示す。図 7 C は、唾液 ( S A ) の遺伝子発現プロファイルの一貫性を示す。図 7 D は、腔分泌物 ( V S ) の遺伝子発現プロファイルの一貫性を示す。図 7 E は、月経血 ( M B ) の遺伝子発現プロファイルの一貫性を示す。図 7 F は、皮膚 ( S K ) の遺伝子発現プロファイルの一貫性を示す。各点は 1 つの遺伝子であり；遺伝子はそれらの固有の体液タイプによって着色される。名目上の血液の遺伝子は赤色であり、精液の遺伝子は青色であり、唾液の遺伝子は緑色であり、腔分泌物の遺伝子は黄色であり、月経血の遺伝子はピンク色であり、皮膚の遺伝子は紫色であり、およびすべての細胞型に現れるハウスキーピング遺伝子は黒色である。血液 ( B D ) 。

10

## 【 0 0 2 1 】

【図 8】図 8 は、互いの体液に対する各体液の平均遺伝子発現プロファイルをプロットする。遺伝子は、図 7 A ~ 7 F と同様に着色される。

## 【 発明を実施するための形態 】

## 【 0 0 2 2 】

発明の詳細な説明

## 【 0 0 2 3 】

いくつかの実施形態において、統計解析は、サンプルの組成およびサンプル内の遺伝子発現を判定するために、少なくとも 1 つの同定可能な物質を含むサンプルに対して実施されてもよい。いくつかの実施形態において、模範的な事例は、複数の物質 (例えば、皮膚、静脈血、腔分泌物、唾液、月経血、精液、およびバイオ粒子) を含む法医学サンプル、および / または複数の物質 (例えば、生物学的物質) を含むあらゆるサンプル (例えば、生物学的サンプル) を含んでもよく、そしてそれは、例えば、各物質中に存在することが知られている標的遺伝子の遺伝子発現を使用して、同定されるおよび / または定量化される必要があり得る。

20

## 【 0 0 2 4 】

いくつかの実施形態において、図 3 に関して、当業者は、サンプル 3 0 2 (例えば、複数の物質を含む生物学的サンプル) を入手することができ、そして、全 RNA 量は、精製を伴った直接溶解と精製を伴わない直接溶解の少なくとも 1 つを使用してサンプル 3 0 4 から抽出され得る。いくつかの実施において、直接溶解は、指定された期間、例えば約 5 分間、7 5 にてサンプルを溶解することを含んでもよい。RNA は、ユーザーによって指定されたプローブ (例えば、レポータープローブおよび捕捉プローブ)、あるいはサンプルおよび / またはサンプル内に存在すると推測される物質に特異的に設計されたコンピューターで作成した複数のコードセットとハイブリダイズされ得る 3 0 6。例えば、上記の法医学的な物質のいずれかを伴った法医学的な組織サンプルに関して、複数のコードセットは、例えば静脈血の遺伝子 A L A S 2、A L O X 5 A P、A M 1 C A 1、A N K 1、A Q P 9、A R H G A P 2 6、C 1 Q R 1、C 5 R 1、C A S P 2、C D 3 G、G Y P A、H B A、H B B、H M B S ( P B G D )、M N D A、N C F S 2、および S P T B、月経血の遺伝子 L E F T Y 2、M M P 7、M M P 1 0、および M M P 1 1、唾液の遺伝子 H T N 3、M U C 7、S . ミュータンス ( S . mutans ) 1 6 S、S . ミュータンス p r o C、S . ミュータンス r e l A、S . ミュータンス r p l A、S . ミュータンス r p o B、S . ミュータンス r p o S、S . サリバリウス ( S . salivarius ) 1 6 S、S . サリバリウス p r o C、S . サリバリウス r e l A、S . サリバリウス r p l A、S . サリバリウス r p o B、S . サリバリウス r p o S、S M R 3 B、および S T A T H、精液の遺伝子 I Z U M O 1、M S P、P S A ( K L K 3 )、P R M 1、P R M 2、S E M G 1、S E M G 2、および T G M 4、皮膚の遺伝子 C C L 2 7、I L 1 F 7、K R T 9、L C E 1 C、および L C E 2 D、腔分泌物の遺伝子 C Y P 2 A 7、C Y P 2 B 7 P 1、D K K 4、F U T 6、I L 1 9、M Y O Z 1、および N O X O 1、ならびに参照遺伝子 B 2 M

30

40

50

、COX1、HPRT1、PGK1、PPIH、S15、TCEA1、TFRC、UBC、およびUBE2D2など、各物質について複数の独特な遺伝子が指定されてもよい308。また、複数のコードセットには、前記の代表的な遺伝子を追跡するために複数のプローブおよび/または類似物質を指定してもよい。同様の複数のコードセットは、様々なタイプのサンプルに関して、任意数の物質の任意数の遺伝子について作成され得る。

いくつかの実施において、複数のコードセットは、例えば、遺伝子を検出するため（例えば、陽性対照プローブ）、およびサンプルの分析におけるバックグラウンドノイズを計測するため（例えば、陰性対照プローブ）の両方のために、陽性対照プローブおよび陰性対照プローブのうち少なくとも1つを含み得る。

統計的方法

10

#### 【0025】

ケースワークサンプルの3つの代表的な特性は次のものを含む：それらは、(i) 2個以上の体液の混合物を含み、(ii) サイズが制限され、および(iii) 部分的にまたは高度に分解される可能性があることが多い。これにより、ケースワークサンプルに対処するための1つの代表的なアプローチは、以下のとおりである：

- 体液サンプルにおける遺伝子発現の確率分布をモデル化する。

- サンプル中の各体液のレベルについて最尤推定値(MLE)を算出するため、および各体液の推定レベルを提供するサンプルのプロフィールの対数尤度を算出するためにモデルを使用する。

- 所定の体液が存在するまたは存在しない所定のサンプルプロフィールの尤度を比較する尤度比を構成する。

20

特定の体液がモデル内に含まれるときに、サンプルプロフィールがはるかに高い可能性がある場合には、我々は、体液がサンプル中に存在すると結論づけてもよい。

混合サンプルにおける遺伝子発現のモデル化

#### 【0026】

いくつかの実施形態において、遺伝子発現は、対数(乗法)スケールに対して最も良くモデル化され得る。例えば、遺伝子の発現レベルの倍増は、一般的に、その発現レベルを半分に減らす規模に匹敵する変化と考えられ得るので、200個から400個のmRNA転写物への遺伝子の増加は、遺伝子が2000カウントから4000カウントに増加する場合の遺伝子発現と大きく異なる。しかしながら、混合物の数学的処理は加法的であつてもよい。例えば、サンプルが、半分は血液であり、半分は唾液であった場合、遺伝子の累積発現レベルは、各組織サンプルにおける発現レベルの加法からもたらされ得る。そのため、混合物に対する各体液の寄与は、線形スケールに対してモデル化され得るが、観察された発現と予測された発現の間の矛盾は、対数スケールにより計測され得る。

30

#### 【0027】

いくつかの実施形態において、単一体液からのサンプルにおける遺伝子発現のモデルが規定され、次に、体液の混合物に拡張され得る。いくつかの実施において、さまざまなモデルが、コンピュータデバイス上で実行、作成、保存、および/または利用され得る。そこから、サンプル中の体液量の最尤推定量(MLE)の算出、およびサンプル中の体液の存在に関する試験に対する尤度比の使用が記載され得る。

40

単一体液からのサンプルにおける遺伝子発現のモデル化

#### 【0028】

いくつかの実施形態において、各遺伝子は、各体液中で全遺伝子発現の所定の割合を示す。例えば、平均的な血液サンプルでは、当業者は、全RNAの15%がHBBであり、1%がALAS1である、などと予想されるであろう。いくつかの実施形態において、これらは、予想される割合 $X_{HBB}$ 、 $X_{ALAS1}$ 、などと呼ぶこともできる。そのため、所定の血液サンプルでは、予想される遺伝子発現のベクトルは、 $(X_{HBB}, X_{ALAS1}, \dots)^T$ となり得、ここで、はサンプル中のRNAの総量である。

#### 【0029】

生物学的ノイズおよび技術的ノイズの両方に起因して、実際の発現はその期待値付近で

50

変動し得る。遺伝子発現の乗法的な自然挙動に従って、変動性は、対数正規分布から生じるものとしてモデル化され得、ここで、各遺伝子は等しく可変であると仮定されてもよい。次に、サンプル中の単一遺伝子の発現は、次の代表的な関数を使用してモデル化できる

【化 1】

$$\log(y_{HBB}) \sim N(\log(X_{HBB} \beta), \sigma^2),$$

10

ここで、 $y_{HBB}$  は、サンプル中の HBB の発現であり得、そして、 $\sigma^2$  は、その期待値付近の HBB 発現の（対数スケールでの）分散であり得る。

体液の混合物における遺伝子発現のモデル化

【0030】

混合物のモデルは、単一体液サンプルのモデルから誘導され得る 312。表記目的のために、行列はボールド体の大文字で、ベクトルはボールド体の小文字で、およびスカラー量は小文字で表され得る。サンプルは  $i$  (1, . . . , n)、遺伝子は  $j$  (1, . . . , p)、および組織は  $k$  (1, . . . , K) の指数を付され得る。所定のサンプルの遺伝子発現プロフィールは、 $y_i = (y_{i1}, \dots, y_{ip})^T$  となり得、ここで、 $y_{ij}$  はサンプル  $i$  中の遺伝子  $j$  の発現を表す。 $x_{ik}$  は、サンプル  $i$  中の体液  $k$  の量であり得、 $x_i = (x_{i1}, \dots, x_{iK})$  はサンプル  $i$  中のすべての体液の量のベクトルであり得る 316。最後に、行列  $X$  は、 $X$  の  $j^{\text{th}}$  列および  $k^{\text{th}}$  列の要素である  $x_{jk}$  を用いて、各体液タイプ  $k$  中の各遺伝子  $j$  の予想される割合を表すために規定され得る 314、体液  $k$  からのサンプル中の遺伝子  $j$  の予想される割合を表す。いくつかの実施において、 $p$  遺伝子の対数変換した発現レベルの共分散行列は、 $\Sigma$  と記号で表され得る。さらに、行列  $A$  の  $L_p$  ノルムは  $\|A\|_p$  (例えば、ここで、いくつかの実施において  $p = 2$ ) と表され得る。

20

【0031】

図 3 に関して、体液の混合物中の mRNA 分子の数が混合物の各成分中の mRNA 分子の数であり得ると仮定すると、当業者は、サンプル  $I$  中の遺伝子  $j$  の予想されるカウントを：

30

【化 2】

$$E(y_{ij}) = \sum_{k=1}^K \beta_{ik} x_{jk},$$

と書き表すことができ、そして、いくつかの実施形態において、サンプルの全体の予想される遺伝子発現ベクトルの式は：

【化 3】

40

$$E(y_i) = X\beta_i.$$

となり得る 320。

【0032】

一方、遺伝子発現の変動性が対数スケール上で起こると仮定すると、サンプル中の遺伝子発現は：

50

【化4】

$$\log(y_i) \sim N(\log(X\beta_i), \sigma^2 I),$$

とモデル化でき318、ここで、 $I$ は単位行列であり、および $\sigma^2$ は全遺伝子の（対数スケールの）一般分散である。（ $E(y_i) = X_i$ である場合、 $(E(\log(y_i)) = \log(X_i))$ であることに留意すべきである）。しかしながら、本出願で考慮された値の下では、 $E(\log(y_i))$ は、 $\log(X_i)$ に非常に近く近似する。いくつかの実施形態において、遺伝子の共分散行列を完全に推定するために必要なデータが失われているおよび/または存在しない場合、当業者は $\sigma^2 I$ でその近似値を見だし得る。

【0033】

体液中の遺伝子発現に上記モデルを適用する前に、当業者は、2つのパラメーター： $X$ 、例えば、遺伝子発現の予想される割合の行列、および $\sigma^2$ 、例えば、遺伝子発現の分散を推定し得る。 $X$ 行列の推定は、先に記載されている。 $\sigma^2$ 、全遺伝子に共通の対数スケールでの分散は、各体液中の各遺伝子の平均分散と推定され得る。いくつかの実施において、 $X$ は、合計して1になる列を有するようにスケールされ得る；他の実施において、 $X$ の代わりに $\log(X)$ がスケールされてもよく、さまざまな別の値に対して、行列がスケールされることも、および/または行列の一方もしくは両方がスケールされることもない。

サンプル中の各体液量の最尤推定値

【0034】

通常、対数遺伝子発現がその期待値の対数付近まで拡張され、各遺伝子が等しく変動性であるという仮定の下で、 $\beta_i$ のMLE322は、次のように算出できる：

【化5】

$$\hat{\beta}_i = \operatorname{argmin}_{\beta} \|\log(y_i) - \log(X\beta)\|_2^2 \text{ s.t. } \beta \geq 0,$$

30

すなわち、

【化6】

$$\hat{\beta}_i$$

は、観察された遺伝子発現 $y_i$ と予測された遺伝子発現 $X$ との間で対数スケールにおける2乗誤差の合計を最小化し、 $\beta_i$ のすべての要素が負でない（サンプルは負の量の体液を有することができない）という制約条件に依存する。この発現に対する閉形式解が存在しない場合、数値法がそれを最適化するために使用され得る（Byrd et al, SIAM J. Scientific Computing, 1995）。発現は、 $\beta_i$ において凸状でないが；しかしながら、その推定は、異なる初期条件に対して合理的にロバストであり得、非常に類似した対数尤度を有する同様の推定を返した。

【0035】

アルゴリズムが、アッセイのバックグラウンドの遺伝子発現値を当てはめようと試みるそれ自体が努力しすぎる危険性があり得る、いくつかの実施形態において、その後のレイヤーの複雑性がモデルに追加され得る。例えば、各体液についてタームを当てはめるこ

50

とに加えて、全遺伝子に対して等しい加重で X 行列の対応する列と共に、 はバックグラウンドに加えられ得る。バックグラウンド タームは、各遺伝子に対してある数以下（例えば、15 カウント）にしか寄与しないようにさらに制約され得る。同じ理由から、全遺伝子の発現値は、平均バックグラウンドカウントの合理的な推定を引き出すために 5 カウントで切り捨てられ得る 3 2 4。

体液の存在を試験するための尤度比の使用

【 0 0 3 6 】

いずれかの所定のサンプル  $y_i$  では、当業者は、体液が存在するか判定することもできる。いくつかの実施形態において、これは、  $y_i$  の各要素が 0 と等しいか否かを試験することを伴ってもよい。代表的なアプローチの 1 つは、0 に近づけるように問題となっている組織に対応する  $y_{i,j}$  タームを用いて、

10

【化 7】

MLE  $\hat{\beta}_i$

下および制約された

【化 8】

20

$\hat{\beta}_{i,-j}$

下でデータの尤度を算出することである 3 2 6。完全な MLE および制約された MLE 下での尤度比は問題の組織の存在の証拠を要約し得る。

【 0 0 3 7 】

MLE を提供するデータの対数尤度の算出は、予測された遺伝子発現の対数付近に通常分布する対数遺伝子発現に参与し得る。定数までは、

30

【化 9】

$\hat{\beta}_i$

を提供する  $y_i$  の対数尤度は：

【化 10】

40

$$\text{loglik}(y_i|\hat{\beta}_i) =$$

$$-\frac{1}{2}\log(\det(\sigma^2\mathbf{I})) - \frac{1}{2}(\log(y_i) - \log(X\hat{\beta}_i))^T \sigma^{-2}\mathbf{I}(\log(y_i) - \log(X\hat{\beta}_i)).$$

である。

【 0 0 3 8 】

体液 j がサンプル i 中に存在するか否か試験するために、我々は、  $y_i$  および

【化 1 1】

$$\hat{\beta}_i$$

を使用し、また一方で  $y_i$  および制約された

【化 1 2】

10

$$\text{MLE } \hat{\beta}_{i,-j}$$

を使用して上式を評価し、さらに我々は尤度比を算出する。尤度比に由来する得られた値は、サンプル組成物による包含が予想されるものを示し得る 3 2 8。いくつかの実施において、上記の算出のすべてが電子計算デバイスにより処理され得る。いくつかの実施において、次に、電子計算デバイスは、例えば、電子計算デバイスに作動的に連結された、およびデジタルグラフィカルユーザーインターフェース、および/または同類のものにアウトプットを表示するように構成されたディスプレイ・モジュールによって、ユーザーに対してサンプル組成アウトプットが提供され得る 3 3 0。

20

【0 0 3 9】

いくつかの実施において、電子計算デバイスは、例えば、推定される  $X$  または 行列、および任意の  $X$  または 行列、および/または同類のものとの間の対数尤度比に基づいて、推定される  $X$  または 値付近の信頼区間を決定し、実装してもよい。

推定された遺伝子発現に基づくサンプル中の物質割合の推定

【0 0 4 0】

いくつかの実施において、電子計算デバイスは、例えば、ペナルティー値および/または同様の定数を使用して、サンプル（例えば、組織サンプル、および/または同類のもの）中の各物質（例えば、細胞型および/または同類のもの）の割合を算出し得る。推定は、以下の代表的な関数に類似した関数を使用して算出され得る：

30

【化 1 3】

$$S = \text{argmin}_{\beta} \{ \|(\log(y) - \log(X\beta))\|^T \Sigma^{-1} (\log(y) - \log(X\beta))\|_p + \text{ペナルティー}(\beta) \}$$

ここで、 $S$  = サンプル中の物質の割合であり、ここで、関数は の要素がすべて負でないという制約条件を受け、そしてここで、ペナルティー ( ) は、 の要素に対するペナルティー ( ) は、（これだけに限定されるものではないが、「弾性ネット」ペナルティー、Dantzig セレクタ、 $L_p$  ペナルティー、グループ型または融合型  $l_{asso}$  ペナルティー、適当である場合には、それらの組み合わせ、および/または同類のものを含めた） の要素に対する更なるペナルティーを表す。いくつかの実施において、 は  $K * 1$  行列であり得る。

40

サンプル中の物質割合に基づく各物質の遺伝子発現プロファイルの推定

【0 0 4 1】

いくつかの実施において、サンプル中の物質の推定割合の先の等式は、電子計算デバイスがまた、サンプル中に存在することが推定される各物質の遺伝子発現プロファイルも推定できるように、電子計算デバイスによって改変され得る。例えば、遺伝子  $j$  に関して、その発現は  $y' = (y_{1,j}, \dots, y_{n,j})^T$  として  $n$  サンプルにおいて書き表され

50

得る。各物質の遺伝子  $j$  の予想される発現は、 $x' = (X_{j,1}, \dots, X_{j,k})^T$  として表されることができ、ここで、 $X$  は、先の方程式に類似した、遺伝子発現の予想される割合の行列と規定される。  $n$  サンプル中の  $K$  細胞型のそれぞれの推定される割合の行列を  $(\quad)^T)_{n \times k}$  とする。いくつかの実施において、 $(\quad)^T)_{n \times k}$  は、複合サンプルの包摂による  $K \times n$  行列であり得る。

【0042】

上記の値を使用して、 $x'$  は、以下の代表的な関数：

【化14】

$$GE = \operatorname{argmin}_{x'} \{ \|(\log(y') - \log(\beta^T x'))^T \Sigma^{-1} (\log(y') - \log(\beta^T x'))\|_p + \text{ペナルティ}(x') \}$$

10

に類似した関数を使用して算出され得、ここで、 $GE$  = 各物質の遺伝子発現プロフィールであり、およびここで、関数は、 $x'$  の要素がすべて負でないという制約条件を受ける。

さらなる応用

【0043】

いくつかの実施において、 $X$  および  $S$  が未知である場合には、 $GE$  と  $S$  が、共同で両方の行列を一緒に推定するために組み合わせられてもよい。例えば、 $X$  または  $S$  のいずれかについて最も合理的な推定を用いて開始することが可能であれば、当業者は、両方の行列の値において推定値が収束するまで、 $S$  から  $X$  を推定することとその逆を繰り返し得る。

20

【0044】

いくつかの実施において、(例えば、癌細胞が正常組織と混合されたとき、正常細胞における遺伝子発現よりも、癌における遺伝子発現がはるかに変動性であることに起因して)  $X$  の1つの列が未知であり、かつ、他の列が既知である場合には、統計的方法は、 $X$  行列の最も利用可能な推定を使用して  $S$  を推定し得る(例えば、癌細胞と正常細胞が分析される場合には、当業者は、 $X$  の未知の列に癌細胞の平均的な遺伝子発現プロフィールを使用してもよい)。次に、不確実な発現プロフィール(例えば、 $X$  の未知の列)を有する物質における発現は、以下の代表的な関数に類似した関数を使用して推定され得る：

【化15】

30

$$y - X_{\cdot k} \beta_k$$

ここで、 $X_{\cdot k}$  は、不確実な列を有していない  $X$  行列であり、およびここで、 $\beta_k$  は、不確実な物質型のタームを有していないベクトルである。

【0045】

いくつかの実施において、当業者はまた、各物質の共分散行列  $\Sigma$  を推定することも可能にし得る。次に、物質特異的な共分散行列  $\Sigma_1, \dots, \Sigma_k$  を使用して、統計的方法は物質特異的な行列に基づく大局的共分散行列  $\Sigma$  を洗練することも可能にし得る。例えば、 $\Sigma$  を推定するために、(例えば、最尤推定値、罰則付き最尤推定値、実証的な共分散行列、および/または同類のものに基づいて) 適当な大局的共分散行列  $\Sigma$  を選んだ後に、電子計算デバイスは、サンプルのための新しい共分散行列  $\Sigma$  を決定するために推定される  $\Sigma$  および  $\Sigma_1, \dots, \Sigma_k$  を使用し得る。電子計算デバイスは、収束および/または同類のものまで、共分散行列  $\Sigma$  を算出するために、 $\Sigma$  を推定し続け、そして、それと物質特異的な行列を使用し得る。

40

【0046】

この明細書および添付の請求項で使用される場合、単数形の「 $a$ 」、「 $a_n$ 」、および「 $the$ 」は、別段の明確な文脈指示のない限り、複数の指示物も含む。

50

## 【0047】

具体的に記載されるかまたは文脈から明白でない限り、本明細書中に使用される場合、「または」という語句は、包括的であると理解されるので、「または」と「および」の両方を網羅する。

## 【0048】

具体的に記載されるかまたは文脈から明白でない限り、本明細書中に使用される場合、「約」という語句は、例えば、当該技術分野における標準的な許容性の範囲内、例えば、平均の2標準偏差の範囲内にあると理解される。約は、記載された値の10%、9%、8%、7%、6%、5%、4%、3%、2%、1%、0.5%、0.1%、0.05%、または0.01%の範囲内と理解され得る。別段文脈から明確でない限り、本明細書中に提供されたすべての数値が、「約」という語句によって修飾される。

10

## 【0049】

別段の規定がない限り、本明細書中に使用されるすべての学術用語および科学用語は、本発明が関係する当業者によって一般的に理解されるのと同じ意味を有する。本明細書中に記載したものと類似したまたは同等の他のプローブ、組成物、方法、およびキットが本発明の実施において使用可能であるが、好ましい材料および方法を本明細書中に記載している。本明細書中に使用される専門用語は、特定の実施形態だけを記載する目的のためのものであって、制限することを意図するものではないことは、理解されるべきである。

## 【実施例】

## 【0050】

20

## mRNAバイオマーカーの選択

いくつかの実施形態において、57種の体液/組織特異的遺伝子、および10種のハウスキーピング遺伝子対照の「コードセット」（例えば、複数のコードセット）（表1）（そしてそれは、800種を標的化するシステムの技術的能力の十分な範囲内である）を利用し得る。システムの高い多重処理能力を利用するために、特定の体液に高度に特異的であることが実証されているバイオマーカー（例えば、精液のPRM2およびSEMG1）を、より低い程度の組織特異性が示されているもの（例えば、膣分泌物のMYOZ1および唾液のMUC7）と同様に含み得る。また、表2および表3も参照のこと。

## 【0051】

表1. NanoString（登録商標）Custom Codesetにおける体液特異的遺伝子およびハウスキーピング遺伝子

30

【表 1 - 1】

遺伝子	体液標的	
ALAS2	血液	
ALOX5AP	血液	
AM1CA1	血液	
ANK1	血液	
AQP9	血液	
ARHGAP26	血液	10
C1QR1	血液	
C5R1	血液	
CASP2	血液	
CD3G	血液	
GYP A	血液	
HBA,	血液	
HBB	血液	
HMBS (PBGD)	血液	
MNDA	血液	
NCFS2	血液	20
SPTB	血液	
LEFTY2	月経血	
MMP7	月経血	
MMP10	月経血	
MMP11	月経血	
HTN3	唾液	
MUC7	唾液	
<i>S.</i> ミュータンス 16S	唾液	
<i>S.</i> ミュータンス proC	唾液	
<i>S.</i> ミュータンス relA	唾液	30
<i>S.</i> ミュータンス rplA	唾液	
<i>S.</i> ミュータンス rpoB	唾液	
<i>S.</i> ミュータンス rpoS	唾液	
<i>S.</i> サリバリウス 16S	唾液	
<i>S.</i> サリバリウス proC	唾液	
<i>S.</i> サリバリウス relA	唾液	
<i>S.</i> サリバリウス rplA	唾液	
<i>S.</i> サリバリウス rpoB	唾液	
<i>S.</i> サリバリウス rpoS	唾液	
SMR3B	唾液	40
STATH	唾液	

【表 1 - 2】

遺伝子	体液標的
IZUM01	精液
MSP	精液
PSA (KLK3)	精液
PRM1	精液
PRM2	精液
SEMG1	精液
SEMG2	精液
TGM4	精液
GCL27	皮膚
IL1F7	皮膚
KRT9	皮膚
LCE1C	皮膚
LCE2D	皮膚
CYP2A7	腔
CYP2B7P1	腔
DKK4	腔
FUT6	腔
IL19	腔
MYOZ1	腔
NOX01	腔
B2M	参照遺伝子
COX1	参照遺伝子
HPRT1	参照遺伝子
PGK1	参照遺伝子
PPIH	参照遺伝子
S15	参照遺伝子
TCEA1	参照遺伝子
TFRC	参照遺伝子
UBC	参照遺伝子
UBE2D2	参照遺伝子

10

20

30

【 0 0 5 2 】

表 2 : 試験したサンプルの一覧

【表 2】

サンプルのタイプ	N	説明
血液	14	
有機溶媒抽出	7	綿布上の血痕 (乾燥後に-47 °Cにて保管)
	1	環境 (屋外 (FL) - 熱, 日光, 湿度, 雨 (1カ月))
	1	環境 (屋外 (FL) - 熱, 日光, 湿度, 遮蔽 (3日間))
直接溶解 (RLT)	5	綿布上の血痕 (乾燥後に-47 °Cにて保管)
精液	17	
有機溶媒抽出	7	綿棒上で乾燥させた (乾燥後に-47°Cにて保管)
	2	環境 (屋外 (FL) - 熱, 日光, 湿度, 遮蔽 (1週間))
	3	感度: 25ng, 12.5ng, 6.25ng (5 $\mu$ lの抽出物の使用によって達成された入力)
直接溶解 (RLT)	5	綿棒上で乾燥させた (乾燥後に-47 °Cにて保管)
唾液	17	
有機溶媒抽出	7	綿棒上で頬側サンプルを乾燥させた (乾燥後に-47 °Cにて保管)
	1	環境 (屋外 (FL) - 熱, 日光, 湿度, 雨 (1週間))
	1	環境 (屋外 (FL) - 熱, 日光, 湿度, 遮蔽 (1カ月))
	3	感度: 25ng, 12.5ng, 6.25ng (5 $\mu$ lの抽出物の使用によって達成された入力)
直接溶解 (RLT)	5	綿棒上で頬側サンプルを乾燥させた (乾燥後に-47 °Cにて保管)
腔分泌物	10	
有機溶媒抽出	6	綿棒上でサンプルを乾燥させた (乾燥後に-47 °Cにて保管)
	1	環境 (屋外 (FL) - 熱, 日光, 湿度, 雨 (3日間))
直接溶解 (RLT)	3	綿棒上でサンプルを乾燥させた (乾燥後に-47 °Cにて保管)
月経血	10	
有機溶媒抽出	7	綿棒上でサンプルを乾燥させた (乾燥後に-47 °Cにて保管)
直接溶解 (RLT)	3	綿棒上でサンプルを乾燥させた (乾燥後に-47 °Cにて保管)
皮膚	14	
有機溶媒抽出	1	表皮 (男性の手) のスワブ; 滅菌水で湿らせたスワブ
	1	コーヒーカップ表面のスワブ; 滅菌水で湿らせたスワブ
	1	コンピューターマウスのスワブ; 滅菌水で湿らせたスワブ
直接溶解 (RLT)	1	表皮 (男性の手) のスワブ; 滅菌水で湿らせたスワブ
	1	コーヒーカップ表面のスワブ; 滅菌水で湿らせたスワブ
	1	コンピューターマウスのスワブ; 滅菌水で湿らせたスワブ
直接溶解 (RWAGEM)	1	25個のバイオ粒子 (塊); シャツのカラー (男性)
	1	50個のバイオ粒子 (塊); シャツのカラー (男性)
直接溶解 (forensicGEM)	1	100個のバイオ粒子 (55個の塊/45個の単体) (塊); シャツのカラー (男性)
なし	5	皮膚の全RNA (商業的供給源)
混合物	5	
有機溶媒抽出	2	腔/精液 (各ドナーの1/2のスワブを同じチューブ内で抽出した)
	2	血液/唾液 (各ドナーの1/2のスワブを同じチューブ内で抽出した)
	1	精液/唾液/腔 (各ドナーの1/2のスワブを同じチューブ内で抽出した)
対照	3	
有機溶媒抽出	2	清潔な無菌のスワブ (陰性対照)
なし	1	脳的全RNA (商業的供給源)

染み付け = 50  $\mu$ l の染み付け; スワブ - 体液を染み込ませたスワブ (無菌の綿)

環境サンプル (血液、精液、唾液) - 綿布上

全RNA - 商業的供給源 (方法を参照のこと)

【0053】

10

20

30

40

50

表3. サンプルの説明およびアッセイ入力(完全なサンプルセット)

【表3-1】

サンプル	説明	抽出のタイプ	入力 ( $\mu$ l)	入力 (ng)	
1	50 $\mu$ lの綿布上の血痕; ドナー1	標準	5 $\mu$ l	50 ng	
2	50 $\mu$ lの綿布上の血痕; ドナー2	標準	5 $\mu$ l	50 ng	
3	50 $\mu$ lの綿布上の血痕; ドナー3	標準	5 $\mu$ l	50 ng	
4	50 $\mu$ lの綿布上の血痕; ドナー4	標準	5 $\mu$ l	50 ng	
5	Env. 血痕: 屋外, 遮蔽 3日間 (ドナー5)	標準	5 $\mu$ l	NA	10
6	50 $\mu$ lの綿布上の血痕; ドナー4	直接溶解 (RLT)	5 $\mu$ l	NA	
7	染み込ませた精液スワブ(綿, 乾燥); ドナー1	標準	5 $\mu$ l	50 ng	
8	染み込ませた精液スワブ(綿, 乾燥); ドナー2	標準	5 $\mu$ l	50 ng	
9	染み込ませた精液スワブ(綿, 乾燥); ドナー3	標準	5 $\mu$ l	50 ng	
10	染み込ませた精液スワブ(綿, 乾燥); ドナー4	標準	5 $\mu$ l	50 ng	
11	Env: 50 $\mu$ lの綿布上の精液; 屋外, 遮蔽 1週間 (ドナー5)	標準	5 $\mu$ l	NA	
12	1/2の染み込ませた精液スワブ(綿, 乾燥); ドナー1	直接溶解 (RLT)	5 $\mu$ l	NA	
13	頬側スワブ(綿, 乾燥); ドナー1	標準	5 $\mu$ l	50 ng	
14	頬側スワブ(綿, 乾燥); ドナー2	標準	5 $\mu$ l	50 ng	
15	頬側スワブ(綿, 乾燥); ドナー3	標準	5 $\mu$ l	50 ng	20
16	頬側スワブ(綿, 乾燥); ドナー4	標準	5 $\mu$ l	50 ng	
17	Env: 50 $\mu$ lの綿布上の唾液; 屋外, 遮蔽 1カ月 (ドナー5)	標準	5 $\mu$ l	50 ng	
18	1/2の頬側スワブ(綿, 乾燥); ドナー6	直接溶解 (RLT)	5 $\mu$ l	NA	
19	1/2の膣スワブ(綿, 乾燥); ドナー1	標準	5 $\mu$ l	50 ng	
20	1/2の膣スワブ(綿, 乾燥); ドナー2	標準	5 $\mu$ l	50 ng	
21	1/2の膣スワブ(綿, 乾燥); ドナー3	標準	5 $\mu$ l	50 ng	
22	1/2の膣スワブ(綿, 乾燥); ドナー4	標準	5 $\mu$ l	50 ng	
23	Env: 1/2の膣スワブ; 屋外, 遮蔽なしで3日間 (ドナー5)	標準	5 $\mu$ l	50 ng	
24	1/2の膣スワブ(綿, 乾燥); ドナー2	直接溶解 (RLT)	5 $\mu$ l	NA	30
25	1/2の月経血スワブ(綿; 乾燥) ドナー1, 生理 2日目	標準	5 $\mu$ l	50 ng	
26	1/2の月経血スワブ(綿; 乾燥) ドナー2	標準	5 $\mu$ l	50 ng	
27	1/2の月経血スワブ(綿; 乾燥) ドナー3, 生理 1日目	標準	5 $\mu$ l	50 ng	
28	1/2の月経血スワブ(綿; 乾燥) ドナー4, 生理 2日目	標準	5 $\mu$ l	50 ng	
29	1/2の月経血スワブ(綿; 乾燥) ドナー5, 生理 3日目	標準	5 $\mu$ l	50 ng	
30	1/2の月経血スワブ(綿; 乾燥) ドナー1	直接溶解 (RLT)	5 $\mu$ l	NA	
31	皮膚 - 全RNA (商業的供給源)	なし	5 $\mu$ l	50 ng	
32	皮膚 - 全RNA (商業的供給源)	なし	5 $\mu$ l	50 ng	
33	皮膚 - 全RNA (商業的供給源)	なし	5 $\mu$ l	50 ng	
34	皮膚 - 全RNA (商業的供給源)	なし	5 $\mu$ l	50 ng	
35	コンピューターマウスの表面スワブ(全体)	標準	5 $\mu$ l	NA	40
36	コンピューターマウスの表面スワブ(全体)	直接溶解 (RLT)	5 $\mu$ l	NA	
37	精液 (ドナー2) - 希釈系列	標準	5 $\mu$ l	25 ng	
38	精液 (ドナー2) - 希釈系列	標準	5 $\mu$ l	12.5 ng	
39	精液 (ドナー2) - 希釈系列	標準	5 $\mu$ l	6.25 ng	
40	唾液 (ドナー1) - 希釈系列	標準	5 $\mu$ l	25 ng	
41	唾液 (ドナー1) - 希釈系列	標準	5 $\mu$ l	12.5 ng	

【表 3 - 2】

サンプル	説明	抽出のタイプ	入力 ( $\mu$ l)	入力 (ng)
42	唾液 (ドナー1) - 希釈系列	標準	5 $\mu$ l	6.25 ng
43	ヒト脳 - 全RNA (商業的供給源)	なし	5 $\mu$ l	50 ng
44	抽出ブランク (ブランク/清潔なスワブ)	標準	5 $\mu$ l	NA
45	100個のバイオ粒子 (55個の塊/45個の単体) ; 男性のシャツのカラー	直接溶解 (FG)	5 $\mu$ l	NA
46	腔 (ドナー3) - 精液 (ドナー1) 混合物 (各スワブの1/2)	標準	5 $\mu$ l	50 ng
47	血液 (ドナー1) - saliva (ドナー2) 混合物 (各スワブの1/2)	標準	5 $\mu$ l	50 ng
48	精液 (ドナー1) - saliva (ドナー2) - 腔 (ドナー3) (各スワブの1/2)	標準	5 $\mu$ l	50 ng
49	$\frac{1}{2}$ 50 $\mu$ lの綿布上の血痕 ; ドナー6	標準	10 $\mu$ l	60 ng
50	$\frac{1}{2}$ 50 $\mu$ lの綿布上の血痕 ; ドナー6	直接溶解 (RLT)	5 $\mu$ l	NA
51	#50の技術の反復	直接溶解 (RLT)	10 $\mu$ l	NA
52	$\frac{1}{2}$ 50 $\mu$ lの綿布上の血痕 ; ドナー7	標準	8 $\mu$ l	104 ng
53	$\frac{1}{2}$ 50 $\mu$ lの綿布上の血痕 ; ドナー7	直接溶解 (RLT)	5 $\mu$ l	NA
54	$\frac{1}{2}$ 50 $\mu$ lの綿布上の血痕 ; ドナー8	直接溶解 (RLT)	5 $\mu$ l	NA
55	$\frac{1}{2}$ 50 $\mu$ lの綿布上の血痕 ; ドナー8	直接溶解 (RLT)	10 $\mu$ l	NA
56	$\frac{1}{2}$ の染み込ませた精液スワブ (綿、乾燥) ; ドナー6	標準	4 $\mu$ l	108 ng
57	$\frac{1}{2}$ の染み込ませた精液スワブ (綿、乾燥) ; ドナー6	直接溶解 (RLT)	5 $\mu$ l	NA
58	$\frac{1}{2}$ の染み込ませた精液スワブ (綿、乾燥) ; ドナー7	標準	5.3 $\mu$ l	101 ng
59	$\frac{1}{2}$ の染み込ませた精液スワブ (綿、乾燥) ; ドナー7	直接溶解 (RLT)	5 $\mu$ l	NA
60	#59の技術の反復	直接溶解 (RLT)	10 $\mu$ l	NA
61	$\frac{1}{2}$ の染み込ませた精液スワブ (綿、乾燥) ; ドナー8	直接溶解 (RLT)	5 $\mu$ l	NA
62	$\frac{1}{2}$ の染み込ませた精液スワブ (綿、乾燥) ; ドナー8	直接溶解 (RLT)	10 $\mu$ l	NA
63	$\frac{1}{2}$ の新鮮な頬側スワブ (綿) ; ドナー7	標準	5 $\mu$ l	610 ng
64	$\frac{1}{2}$ の新鮮な頬側スワブ (綿) ; ドナー7	直接溶解 (RLT)	5 $\mu$ l	NA
65	$\frac{1}{2}$ の新鮮な頬側スワブ (綿) ; ドナー8	標準	10 $\mu$ l	470 ng
66	$\frac{1}{2}$ の新鮮な頬側スワブ (綿) ; ドナー8	直接溶解 (RLT)	5 $\mu$ l	NA
67	#66の技術の反復	直接溶解 (RLT)	10 $\mu$ l	NA
68	$\frac{1}{2}$ の新鮮な頬側スワブ (綿) ; ドナー9	直接溶解 (RLT)	5 $\mu$ l	NA
69	$\frac{1}{2}$ の新鮮な頬側スワブ (綿) ; ドナー9	直接溶解 (RLT)	10 $\mu$ l	NA
70	$\frac{1}{2}$ の新鮮な頬側スワブ (綿) ; ドナー9	直接溶解 (RLT)	5 $\mu$ l	NA
71	$\frac{1}{2}$ の新鮮な頬側スワブ (綿) ; ドナー9	直接溶解 (RLT)	10 $\mu$ l	NA
72	$\frac{1}{2}$ の腔スワブ (綿 ; 乾燥) ; ドナー6	標準	1 $\mu$ l	332 ng
73	$\frac{1}{2}$ の腔スワブ (綿 ; 乾燥) ; ドナー6	直接溶解 (RLT)	5 $\mu$ l	NA
74	$\frac{1}{2}$ の腔スワブ (綿 ; 乾燥) ; ドナー7	標準	1 $\mu$ l	255 ng
75	$\frac{1}{2}$ の腔スワブ (綿 ; 乾燥) ; ドナー7	直接溶解 (RLT)	5 $\mu$ l	NA
76	$\frac{1}{2}$ の月経血スワブ (綿 ; 乾燥) ; ドナー6, 生理2日目	標準	1 $\mu$ l	118 ng
77	$\frac{1}{2}$ の月経血スワブ (綿 ; 乾燥) ; ドナー6, 生理2日目	直接溶解 (RLT)	5 $\mu$ l	NA
78	$\frac{1}{2}$ の月経血スワブ (綿 ; 乾燥) ; ドナー7	標準	3.6 $\mu$ l	101 ng
79	$\frac{1}{2}$ の月経血スワブ (綿 ; 乾燥) ; ドナー7	直接溶解 (RLT)	5 $\mu$ l	NA
80	#79の技術の反復	直接溶解 (RLT)	10 $\mu$ l	NA
81	ヒト皮膚のスワブ (男性の手、左)	標準	10 $\mu$ l	80 ng

10

20

30

40

【表 3 - 3】

サンプル	説明	抽出のタイプ	入力 ( $\mu$ l)	入力 (ng)
82	ヒト皮膚のスワブ(男性の手、右)	直接溶解 (RLT)	5 $\mu$ l	NA
83	#88の技術の反復	直接溶解 (RLT)	10 $\mu$ l	NA
84	金属製コーヒーカップの表面(側面1)のスワブ	標準	8.3 $\mu$ l	100 ng
85	金属製コーヒーカップの表面(側面2)のスワブ	直接溶解 (RLT)	5 $\mu$ l	NA
86	#85の技術の反復	直接溶解 (RLT)	10 $\mu$ l	NA
87	25個のバイオ粒子 (塊) ; 男性のシャツのカラー	直接溶解 (RG)	5 $\mu$ l	NA
88	50個のバイオ粒子 (塊) ; 男性のシャツのカラー	直接溶解 (RG)	5 $\mu$ l	NA
89	Env: 50 $\mu$ lの綿布上の精液: 屋外, 遮蔽して1週間 (ドナー9)	標準	1.3 $\mu$ l	100 ng
90	50 $\mu$ lの綿布上の血痕 ; ドナー9	標準	7.1 $\mu$ l	99 ng
91	腔 (ドナー4)-精液 (ドナー9) 混合物 (各スワブの1/2)	標準	1.0 $\mu$ l	164 ng
92	Env: 50 $\mu$ lの綿布上の唾液: 屋外, 遮蔽して1週間 (ドナー10)	標準	7.7 $\mu$ l	100 ng
93	1/2の染み込ませた精液スワブ(綿、乾燥) ; ドナー10	標準	4.3 $\mu$ l	99 ng
94	血液 (ドナー10)-唾液 (ドナー7) 混合物 (各スワブの1/2)	標準	2.0 $\mu$ l	98 ng
95	抽出ブランク (ブランク/清潔なスワブ)	標準	5.0 $\mu$ l	0 ng
96	乾燥させた頬側スワブ (綿) ; ドナー1	標準	1.0 $\mu$ l	133 ng
97	Env: 50 $\mu$ lの綿布上の血液: 屋外, 遮蔽なしで1カ月 (ドナー11)	標準	2.0 $\mu$ l	106 ng
98	皮膚 - 全RNA (商業的供給源)	標準	2.0 $\mu$ l	100 ng

10

20

Env = 環境 ; 直接溶解 (FG) = forensicGEM(商標) ; 直接溶解 (RG) = RNAGEM(商標)

予想される体液プロフィールの推定

## 【0054】

いくつかの実施形態において、データセットは、大きく変動するRNA濃度のサンプルを含んでもよく、また、頻繁にアッセイのバックグラウンドノイズになるより低い濃度のサンプル中に遺伝子を含んでもよい。各体液の平均遺伝子発現プロフィールの正確な推定を確実にするために、高発現レベルのハウスキーピング遺伝子を有するサンプルがさらに処理するために保持され得る。

30

## 【0055】

体液の混合物中の遺伝子発現のモデルに関する開示に記載したモデルによれば、いくつかの実施形態において、各体液中の遺伝子の相対発現レベルが得られる ; 言い換えると、総シグネチャー遺伝子発現の割合は、所定の体液中の各遺伝子から予想した。これは、ほとんどの遺伝子発現に基づく選別装置とは対照的であり、そして、その選別装置は、得ることが不可能ではないにしても、困難であり得る、各遺伝子の絶対発現レベルにより注目する。そのため、各サンプルは、大局的に正規化され、そして、すべての発現値の合計が1つの値(例えば、1)となり得るように、および各遺伝子の発現値が合計シグネチャー遺伝子発現の割合となり得るように、再スケールされる。次に、各体液内の平均正規化発現値を伴った各体液中の各遺伝子の予想される発現割合を推定し得る。

40

## 【0056】

いくつかの実施形態において、5個の代表的な体液および皮膚は、大きく異なる遺伝子発現プロフィールを実証する可能性があり、シグネチャー遺伝子は同じ体液のサンプル間で異なる可能性はあるが、体液間でのそれらの差は更に大きい可能性がある。少なくともいくつかの体液では、平均発現プロフィールは、体液の推定上の特徴遺伝子の高い発現を

50

呈し得るが、この傾向は、一部の状況下で唾液サンプルでは明らかに弱くなり得る（図5～8を参照のこと）。

#### 【0057】

いくつかの実施形態において、HBB発現は、血液プロフィールを支配することができ、血中のALAS2レベルが他の遺伝子のレベルをはるかに超えて得るが、例えば、ALAS2、ALOX5AP、AM1CA1、ANK1、AQP9、ARHGAP26、C1QR1、C5R1、CASP2、CD3G、GYPA、HBA、HMBS(PBGD)、MNDA、NCF52、SPTBなどの他の血液マーカーをはるかに超越している。推定上の血液マーカーANK1は、血液サンプル中では豊富であり得ないが、唾液サンプルにおいて最も顕著に現れ得る。いくつかの状況において、精液サンプル中での発現は、精液特異的遺伝子IZUMO1、MSP、PSA(KLK3)、PRM1、PRM2、SEMG1、SEMG2、およびTGM4から主として生じ得るが、他の遺伝子、特にHBBもまた、検出可能であり得る。唾液サンプルは、例えばHTN3、MUC7、S.ミュータンス16S、S.ミュータンスproC、S.ミュータンスrelA、S.ミュータンスrp1A、S.ミュータンスrpoB、S.ミュータンスrpoS、S.サリパリウス16S、S.サリパリウスproC、S.サリパリウスrelA、S.サリパリウスrp1A、S.サリパリウスrpoB、S.サリパリウスrpoS、SMR3B、およびSTATHなどの唾液特異的遺伝子による最も広汎なプロフィールを有することができ、いくつかの状況において、計測した発現の28%しか寄与していない。腔分泌物サンプルは、高く上昇したレベルの、例えばDKK4、CYP2B7P1およびある程度のFUT6などの腔マーカーを有し得る。月経血サンプルは、LEFTY2、MMP7、MMP10、およびMMP11を含めたそれらの独特の遺伝子の高い発現を示し得る。また、月経血サンプルは、血液(HBB、ALAS2)および腔分泌物(CYP2B7P1)バイオマーカーも含み得る。皮膚サンプルは、例えばLCE1C、IL1F7およびCCL27などの皮膚遺伝子の高い発現を示し得るが、これらの遺伝子はまた、腔分泌物および月経血においてもわずかに上昇し得る。いくつかの状況において、一つには斯かる調製物における内皮組織混入の潜在的な存在に起因して、HBBは、市販の皮膚調製物において最も普及している遺伝子であり得る。

10

20

30

40

50

#### 【0058】

遺伝子の少なくとも一部が、唾液サンプル中の総発現のうちの無視できない割合で存在し得る。唾液中で高度に発現した遺伝子を計測した場合、唾液中の他の体液特徴的な遺伝子の相対発現は劇的に縮小し得る。

サンプルの体液組成物を予測するための遺伝子発現の使用

#### 【0059】

先に記載したように、体液検出法のためのいくつかの実施形態による代表的なアルゴリズムを提供する。以下は、サンプルの体液組成を予測する性能の概要である。100の尤度比カットオフは、体液が所定のサンプル中で検出されるか否かを明らかにするために使用し得る。いくつかの実施形態において、それらの尤度比が100を超えている場合には、体液は検出されたものが呼び出され得る。アルゴリズムは、正しい体液を同定するのに成功し得る。所定の物質の独特の遺伝子に一般的に情報提供がない（例えば、物質に関して独特かつ容易に検出される遺伝子がわずかしかない）場合、情報提供のある遺伝子データ不存在下での算出を改善する方法を決めるために、アルゴリズムの洗練が実施されてもよい。いくつかの実施形態において、サンプルが分解されず、および/または非非常に小さい場合に、アルゴリズムの感度は改善され得る。

#### 【0060】

いくつかの実施形態において、アルゴリズムは、LR>100カットオフを変えることによってより良い性能を達成し得る。図1は、いくつかの実施形態による代表的な法医学的な体液タイプの検出に関する真陽性率(TPR)および偽陽性率(FPR)の代表的なROC曲線を示す。LR閾値がアルゴリズムを緩和し、そして、偽陽性と偽陰性の両方をより多く返し得る。例えば、月経血、唾液および皮膚などのいくつかの物質について、R

OC 曲線は、LR 閾値の控え目な緩和が、FPR を少しも増やすことなく TPR の大きな増加をもたらす得ることを明らかにする。いくつかの実施形態において、点は、100 の LR カットオフを使用して達成された成果を示す。よって、LR カットオフの変更は、他の誤差の増大をもたらすことなく、サンプル物質の検出を改善し得る。

体液混合物

#### 【0061】

体液の混合物について判別する方法の能力の予備的な表示として、5 個の混合物は、各体液からの 50  $\mu$ l を染み込ませたまたは単独の綿棒の 1/2 を組み合わせることによって調製し得る。代表的な混合物は、4 種類のバイナリ (2  $\times$  腔分泌物 / 精液、2  $\times$  血液 / 唾液) および 1 種の三成分混合物 (精液 / 唾液 / 腔分泌物) を含む場合がある。血液 / 唾液および腔分泌物 / 精液は、技術的なものとは対照的に生物学的な複製物であり得る。識別閾値として 100 の LR を使用すると、混合物のいくつかは、言い換えれば、腔分泌物 / 精液の 1 つおよび血液 / 唾液サンプルの 1 つを正確に呼び出し得る (例えば、図 2)。いくつかの実施形態において、5 個の代表的な混合物サンプルのそれぞれについて、棒グラフのプロットは、各体液タイプの存在の尤度比を示す。点線は 100 の LR を示す。意義深いことには、代表的なサンプルに対して本明細書中に開示した統計的方法を利用するとき、偽陽性は全く観察できなかった。

10

日常的に使用する 5 分間 RNA 直接溶解法の開発

#### 【0062】

日常的な分析を容易にするために、5 分間の室温細胞溶解プロトコールを、先に概説した手順を用いる法医学サンプル処理のための標準的な RNA 単離に代わる手段として利用し得る。その方法は、高濃度のグアニジンチオシアナート、ならびに専用の界面活性剤ミックスを含む QIAGEN の RLT バッファーに基づく。また、 $\beta$ -メルカプトエタノール (1% v/v) が、溶解物中の RNAses を不活性化するために使用前に加えられてもよい。多くの直接溶解試薬と異なって、RLT バッファーは、多くの生化学反応、例えばハイブリダイゼーションなどが起こるのを許容する。放出された核酸は、主に一本鎖 RNA および二本鎖 DNA の形態であり得、その後者は、そのため、一本鎖プローブにハイブリダイズできない。この事実は、相同 DNA 配列および他の試薬に対するアッセイプローブの DNA 力価の不足を伴い、よって、RNA 分析感度および特異性を増強し得る。

20

#### 【0063】

同じ原料物質からの標準的な RNA 単離 / 精製と直接溶解プロトコールとのアッセイの再現性は、同等であり得る。一般に、高度の発現を加減した全遺伝子に関する 2 つのプロトコール間で、素晴らしい一致を観察し得る。該プロトコール間の相関関係は、わずかしが発現されていない遺伝子については破壊される可能性があり、消失標的を計測するときの、アッセイにおけるより大きなノイズを反映している。溶解物濃度が現行手法によって信頼性高く測定できないので、反復の間での最も劇的な相違は、溶解および精製 RNA 間の RNA 入力量における予想される変動性に起因し得る。溶解および精製プロトコール間で観察された一致は、単純に、5 分間溶解プロトコールが、日常的な法医学的なケースワークのワークフローの効果的な選択肢となるであろうことを示唆している (図 4 を参照のこと)。

30

40

#### 【0064】

さらに、トレーニングから除かれたサンプルは、過剰適合を受けてはならない。いくつかの実施形態において、アルゴリズムは、すべての体液タイプの識別閾値として LR > 100 を利用し得る；他の実施形態において、体液に特異的な閾値を使用した代替アプローチを利用してもよい。

#### 【0065】

いくつかの実施において、コードセットの更なる最適化が可能であり得る。例えば、完全なコードセットとのハイブリダイゼーションの前に、HBB RNA に相補的な特異的に設計した非標識オリゴヌクレオチドの正確な所定の量を添加して HBB シグナルを減弱することは、低レベルの血管組織生産物の汚染から生じる偽陽性を回避する助けとなり得

50

る。これらは、標識プローブとのハイブリダイゼーション反応を拮抗阻害する。血液バイオマーカーの1つを減弱する必要性とは対照的に、唾液バイオマーカーのシグナルは賦活されてもよい。シグナル強化は、単一のHTN3 mRNAに結合する複数のプローブを設計することによって達成し得る。加えて、最新のプローブは、HTN3とHTN1の両方にハイブリダイズするように設計されてもよく、その後者はまた、唾液中に特異的でもある。HTN3強化ストラテジーが期待に添わずに終わった場合、RNA-Seq試験によって同定した代替の新規バイオマーカーが用いられてもよい。いくつかの実施形態において、ANKIプローブを再合成または再設計してもよく、同様のアプローチを、任意の非最適実施バイオマーカーを用いて実施し得る。いくつかの実施形態において、追加の体液特異的バイオマーカー（例えば、腔由来の共生細菌、例えばラクトバチルス属（*Lactobacillus* sp.）など）もまた、分析成果を改善するために組み込まれ得る。

10

#### 【0066】

いくつかの実施形態において、アルゴリズムは、例えば、図2に示されているように、体液の混合物を判別し得る。いくつかの混合物が、誤った陽性結果を伴わないアッセイアルゴリズムを使用して完全に呼び出され得る、そして、一部の成分体液が、任意の「偽陰性」混合物中で同定され得る。偽陰性混合物では、誤った体液である、唾液が、他のサンプルをはるかに超えたレベルで検出され得る。ハウスキーピング遺伝子が、分析に十分な特質および量のRNAが存在することを示すために、および正規化目的のために、遺伝子発現アッセイに加えられ得る（Hanson et al, *Forensic Sci Rev.*, 2010; Haas et al, *Forensic Sci Int Genet.*, 2014; Juusola and Ballantyne, *J Forensic Sci.*, 2007）。ハウスキーピング遺伝子の不均等な発現のため、正規化群としてのそれらの値は疑わしい（Moreno et al, *J. Forensic Sci.*, 2012; Vandesompele et al, *Genome Biol.*, 2002）。いくつかの実施形態において、開示したアルゴリズムは、ハウスキーピング遺伝子を用いた正規化を必要としないので、この目的のために必要とされない。しかしながら、それらの存在は、分析のために好適なRNAの回収を示すことができ、そのため、アッセイにおける特定の有用性を有し続けている。

20

#### 【0067】

当該出願内に存在する、これだけに限定されるものではないが、特許、特許出願、文献、ウェブページ、書籍などを含めた刊行物または他の文書に対する任意のまたはすべての参照は、内容が本開示の実施形態と矛盾しない限り（その場合、本明細書中に存在するものが優勢であるものとする）、その全体を参照によって本明細書中に援用する。参照したものは、本願の出願日前のそれらの開示を単に提供する。本明細書中に開示したいずれの発明も先願発明の効力によって斯かる構成要素がもっと前に起こったものとする権利を与えることがないことは、認められた事実と解釈されるべきである。

30

#### 【0068】

本明細書中に装置、方法、およびシステムの代表的な実施形態を記載したが、斯かる実施形態に対する他の修飾が可能である。これらの実施形態は、説明のためだけに記載されているものであって、制限するものではない。他の実施形態も可能なので、開示によって網羅されており、そしてそれは、本明細書中に含まれた教示から明らかである。よって、開示の幅と範囲は、先に記載した実施形態のいずれによっても制限されるものではなく、本開示およびそれらの同等物によって支持された請求項によってのみ規定されるものとする。さらに、先の開示で示されたフローチャートおよび/または添付の図面は、望ましい結果を達成するために、示された具体的な程度または連続した順序を求められない。そのうえ、対象の開示の実施形態は、遺伝子発現やサンプルの利用に対応する任意のおよびすべての要素を含めた、その他の開示された方法、システム、およびデバイスからの任意のおよびすべての要素をさらに含み得る、方法、システム、およびデバイスを含み得る。言い換えれば、一方のおよび/または別の開示された実施形態は、他の開示された実施形態由来の要素と交換可能であり得る。加えて、開示されている実施形態の1若しくは複数の特徴/要素は、取り出され、特許可能な対象をさらにもたらし得る（そして、これにより、対象の開示に関するより一層さらなる実施形態をもたらす）。より一層さらに、本開示

40

50

のいくつかの実施形態は、従来技術で開示されたある特徴および/または別の特徴をわざわざ必要とせずに、従来技術から識別され得る（例えば、いくつかの実施形態は、負の制限を含み得る）。本明細書中に開示した実施形態の一部は、提示され得る本開示によって支持される多数の請求項から成る以下の代表的な請求項の少なくとも一部の範囲内にある。

【 0 0 6 9 】

【 化 1 6 】

#### 参考文献

[1] J. Butler, *Advanced Topics in Forensic DNA Typing: Methodology*, Elsevier/Academic Press, San Diego, CA, 2012. 10

[2] R. Cook, I. Evett, G. Jackson, P. Jone, A. Lambert, A hierarchy of propositions: deciding which level to address in casework, *Science & Justice*. 38 (1998) 231-239.

[3] J. Juusola, J. Ballantyne, Messenger RNA profiling: a prototype method to supplant conventional methods for body fluid identification, *Forensic Sci Int*. 135 (2003) 85-96.

[4] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, J.D. Watson, *Molecular Biology of the Cell*, 2nd, Garland Publishing, New York, NY, 1994. 20

[5] C. Haas, E. Hanson, J. Ballantyne, Capillary electrophoresis of a multiplex reverse transcription-polymerase chain reaction to target messenger RNA markers for body fluid identification, *Methods Mol. Biol.* 830 (2012) 169-183.

[6] E. Hanson, J. Ballantyne, RNA Profiling for the Identification of the Tissue Origin of Dried Stains in Forensic Biology, *Forensic Sci Rev*. 22 (2010) 145-157.

[7] C. Haas, B. Klessner, C. Maake, W. Bar, A. Kratzer, mRNA profiling for body fluid identification by reverse transcription endpoint PCR and realtime PCR, *Forensic Sci Int Genet*. 3 (2009) 80-88. 30

[8] M. Setzer, J. Juusola, J. Ballantyne, Recovery and stability of RNA in vaginal swabs and blood, semen, and saliva stains, *J Forensic Sci*. 53 (2008) 296-305.

[9] D. Zubakov, E. Hanekamp, M. Kokshoorn, I.W. van, M. Kayser, Stable RNA markers for identification of blood and saliva stains revealed from whole genome expression analysis of time-wise degraded samples, *Int. J. Legal Med*. 122 (2008) 135-142.

[10] D. Zubakov, M. Kokshoorn, A. Kloosterman, M. Kayser, New markers for old stains: stable mRNA markers for blood and saliva identification from up to 16-year-old stains, *Int J. Legal Med*. 123 (2009) 71-74.

[11] C. Haas, E. Hanson, W. Bar, R. Banemann, A.M. Bento, A. Berti, E. Borges, C. Bouakaze, A. Carracedo, M. Carvalho, A. Choma, M. Dotsch, M. Duriancikova, P. Hof-f-Olsen, C. Hohoff, P. Johansen, P.A. Lindenbergh, B. Loddenkotter, B. Ludes, O. Maronas, N. Morling, H. Niederstatter, W. Parson, G. Patel, C. Popielarz, E. Salata, P.M. Schneider, T. Sijen, B. Sviezena, L. Zatkalikova, J. Ballantyne, mRNA profiling for the identification of blood—results of a collaborative EDNAP exercise, *Forensic Sci Int Genet*. 5 (2011) 21-26. 40

【化 1 7】

[12] C. Haas, E. Hanson, N. Morling, J. Ballantyne, Collaborative EDNAP exercises on messenger RNA/DNA co-analysis for body fluid identification (blood, saliva, semen) and STR profiling, *Forensic Sci. Int. Genet. Supp. Ser.* 3 (2011) e5–e6.

[13] C. Haas, E. Hanson, M.J. Anjos, W. Bar, R. Banemann, A. Berti, E. Borges, C. Bouakaze, A. Carracedo, M. Carvalho, V. Castella, A. Choma, C.G. De, M. Dotsch, P. Hoff-Olsen, P. Johansen, F. Kohlmeier, P.A. Lindenbergh, B. Ludes, O. Maronas, D. Moore, M.L. Morerod, N. Morling, H. Niederstatter, F. Noel, W. Parson, G. Patel, C. Popielarz, E. Salata, P.M. Schneider, T. Sijen, B. Sviezena, M. Turanska, L. Zatkalikova, J. Ballantyne, RNA/DNA co-analysis from blood stains—results of a second collaborative EDNAP exercise, *Forensic Sci Int Genet.* 6 (2012) 70–80.

[14] C. Haas, E. Hanson, M.J. Anjos, R. Banemann, A. Berti, E. Borges, A. Carracedo, M. Carvalho, C. Courts, C.G. De, M. Dotsch, S. Flynn, I. Gomes, C. Hollard, B. Hjort, P. Hoff-Olsen, K. Hribikova, A. Lindenbergh, B. Ludes, O. Maronas, N. McCallum, D. Moore, N. Morling, H. Niederstatter, F. Noel, W. Parson, C. Popielarz, C. Rapone, A.D. Roeder, Y. Ruiz, E. Sauer, P.M. Schneider, T. Sijen, Court DS, B. Sviezena, M. Turanska, A. Vidaki, L. Zatkalikova, J. Ballantyne, RNA/DNA co-analysis from human saliva and semen stains—results of a third collaborative EDNAP exercise, *Forensic Sci Int Genet.* 7 (2013) 230–239.

[15] C. Haas, E. Hanson, M.J. Anjos, K.N. Ballantyne, R. Banemann, B. Bhoelai, E. Borges, M. Carvalho, C. Courts, C.G. De, K. Drobnic, M. Dotsch, R. Fleming, C. Franchi, I. Gomes, G. Hadzic, S.A. Harbison, J. Hartevelde, B. Hjort, C. Hollard, P. Hoff-Olsen, C. Huls, C. Keyser, O. Maronas, N. McCallum, D. Moore, N. Morling, H. Niederstatter, F. Noel, W. Parson, C. Phillips, C. Popielarz, A.D. Roeder, L. Salvaderi, E. Sauer, P.M. Schneider, G. Shanthan, Court DS, M. Turanska, R.A. van Oorschot, M. Vennemann, A. Vidaki, L. Zatkalikova, J. Ballantyne, RNA/DNA co-analysis from human menstrual blood and vaginal secretion stains: results of a fourth and fifth collaborative EDNAP exercise, *Forensic Sci Int Genet.* 8 (2014) 203–212.

[16] C. Courts, B. Madea, Specific micro-RNA signatures for the detection of saliva and blood in forensic body-fluid identification, *J. Forensic Sci.* 56 (2011) 1464–1470.

[17] E. Hanson, K. Rekab, J. Ballantyne, Binary logistic regression models enable miRNA profiling to provide accurate identification of forensically relevant body fluids and tissues, *For Sci Int Genet Supp Ser.* 4 (2013) e127–e128.

10

20

30

40

【化 1 8】

[18] E. Hanson, H. Lubenow, J. Ballantyne, Identification of forensically relevant body fluids using a panel of differentially expressed microRNAs, *Forensic Sci. Int. Genet. Supplement Series 2* (2009) 503–504.

[19] E.K. Hanson, H. Lubenow, J. Ballantyne, Identification of Forensically Relevant Body Fluids Using a Panel of Differentially Expressed microRNAs, *Anal. Biochem.* 387 (2009) 303–314.

[20] Z. Wang, H. Luo, X. Pan, M. Liao, Y. Hou, A model for data analysis of microRNA expression in forensic body fluid identification, *Forensic Sci. Int. Genet.* 6 (2012) 419–423. 10

[21] Z. Wang, J. Zhang, H. Luo, Y. Ye, J. Yan, Y. Hou, Screening and confirmation of microRNA markers for forensic body fluid identification, *Forensic Sci. Int. Genet.* 7 (2013) 116–123.

[22] D. Zubakov, A.W. Boersma, Y. Choi, P.F. van Kuijk, E.A. Wiemer, M. Kayser, MicroRNA markers for forensic body fluid identification obtained from microarray screening and quantitative RT-PCR confirmation, *Int. J. Legal Med.* 124 (2010) 217–226. 20

[23] J.H. An, A. Choi, K.J. Shin, W.I. Yang, H.Y. Lee, DNA methylation-specific multiplex assays for body fluid identification, *Int. J. Legal Med.* 127 (2013) 35–43.

[24] A. Choi, K.J. Shin, W.I. Yang, H.Y. Lee, Body fluid identification by integrated analysis of DNA methylation and body fluid-specific microbial DNA, *Int. J. Legal Med.* 128 (2014) 33–41.

[25] D. Frumkin, A. Wasserstrom, B. Budowle, A. Davidson, DNA methylation-based forensic tissue identification, *Forensic Sci. Int. Genet.* 5 (2011) 517–524.

[26] B.L. LaRue, J.L. King, B. Budowle, A validation study of the Nucleix DSI-Semen kit—a methylation-based assay for semen identification, *Int. J. Legal Med.* 127 (2013) 299–308. 30

[27] H.Y. Lee, M.J. Park, A. Choi, J.H. An, W.I. Yang, K.J. Shin, Potential forensic application of DNA methylation profiling to body fluid identification, *Int. J. Legal Med.* 126 (2012) 55–62.

[28] T. Madi, K. Balamurugan, R. Bombardi, G. Duncan, B. McCord, The determination of tissue-specific DNA methylation patterns in forensic biofluids using bisulfite modification and pyrosequencing, *Electrophoresis.* 33 (2012) 1736–1745.

[29] A. Wasserstrom, D. Frumkin, A. Davidson, M. Shpitzen, Y. Herman, R. Gafny, Demonstration of DSI-semen—A novel DNA methylation-based forensic semen identification assay, *Forensic Sci. Int. Genet.* 7 (2013) 136–142. 40

【化 1 9】

[30] J.L. Simons, S.K. Vintiner, Efficacy of several candidate protein biomarkers in the differentiation of vaginal from buccal epithelial cells, *J. Forensic Sci.* 57 (2012) 1585–1590.

[31] S.K. Van, C.M. De, M. Dhaenens, H.D. Van, D. Deforce, Mass spectrometry-based proteomics as a tool to identify biological matrices in forensic science, *Int. J. Legal Med.* 127 (2013) 287–298.

[32] H. Yang, B. Zhou, M. Prinz, D. Siegel, Proteomic analysis of menstrual blood, *Mol. Cell Proteomics.* 11 (2012) 1024–1035.

[33] E. Hanson, C. Haas, R. Jucker, J. Ballantyne, Specific and sensitive mRNA biomarkers for the identification of skin in 'touch DNA' evidence, *Forensic Sci Int Genet.* 6 (2012) 548–558.

[34] J. Juusola, J. Ballantyne, Multiplex mRNA profiling for the identification of body fluids, *Forensic Sci Int.* 152 (2005) 1–12.

[35] M.L. Richard, K.A. Harper, R.L. Craig, A.J. Onorato, J.M. Robertson, J. Donack, Evaluation of mRNA marker specificity for the identification of five human body fluids by capillary electrophoresis, *Forensic Sci Int Genet.* 6 (2012) 452–460.

[36] A.D. Roeder, C. Haas, mRNA profiling using a minimum of five mRNA markers per body fluid and a novel scoring method for body fluid identification, *Int J Legal Med.* 127 (2013) 707–721.

[37] M. Bauer, D. Patzelt, Identification of menstrual blood by real time RT-PCR: technical improvements and the practical value of negative test results, *Forensic Sci Int.* 174 (2008) 55–59.

[38] J. Juusola, J. Ballantyne, mRNA profiling for body fluid identification by multiplex quantitative RT-PCR, *J Forensic Sci.* 52 (2007) 1252–1262.

[39] C. Nussbaumer, E. Gharehbaghi-Schnell, I. Korschineck, Messenger RNA profiling: a novel method for body fluid identification by real-time PCR, *Forensic Sci Int.* 157 (2006) 181–186.

[40] E.K. Hanson, J. Ballantyne, Rapid and inexpensive body fluid identification by RNA profiling-based multiplex High Resolution Melt (HRM) analysis, *F1000Res.* 2 (2013) 281.

[41] S. Audic, J.M. Claverie, The significance of digital gene expression profiles, *Genome Res.* 7 (1997) 986–995.

[42] Z. Wang, M. Gerstein, M. Snyder, RNA-Seq: a revolutionary tool for transcriptomics, *Nat. Rev. Genet.* 10 (2009) 57–63.

10

20

30

40

【化 2 0】

[43] G.K. Geiss, R.E. Bumgarner, B. Birditt, T. Dahl, N. Dowidar, D.L. Dunaway, H.P. Fell, S. Ferree, R.D. George, T. Grogan, J.J. James, M. Maysuria, J.D. Mitton, P. Oliveri, J.L. Osborn, T. Peng, A.L. Ratcliffe, P.J. Webster, E.H. Davidson, L. Hood, K. Dimitrov, Direct multiplexed measurement of gene expression with color-coded probe pairs, *Nat. Biotechnol.* 26 (2008) 317–325.

[44] E.K. Hanson, J. Ballantyne, “Getting blood from a stone”: ultrasensitive forensic DNA profiling of microscopic bio-particles recovered from “touch DNA” evidence, *Methods Mol. Biol.* 1039 (2013) 3–17.

10

[45] E.K. Hanson, J. Ballantyne, Highly specific mRNA biomarkers for the identification of vaginal secretions in sexual assault investigations, *Sci Justice.* 53 (2013) 14–22.

[46] E. Hanson, G. Haas, R. Jucker, J. Ballantyne, Identification of skin in touch/contact forensic samples by messenger RNA profiling, *Forensic Sci Int Genet. Suppl Series.* 3 (2011) e305–e306.

[47] R.H. Byrd, P. Lu, J. N. Cedal, C. Zhu, A limited memory algorithm for bound constrained optimization, *SIAM J. Scientific Computing.* 1995) 1190–1208.

20

[48] L.I. Moreno, C.M. Tate, E.L. Knott, J.E. McDaniel, S.S. Rogers, B.W. Koons, M.F. Kavlick, R.L. Craig, J.M. Robertson, Determination of an effective housekeeping gene for the quantification of mRNA for forensic applications, *J. Forensic Sci.* 57 (2012) 1051–1058.

[49] J. Vandesompele, P.K. De, F. Pattyn, B. Poppe, R.N. Van, P.A. De, F. Speleman, Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes, *Genome Biol.* 3 (2002).

30



【 図 5 】

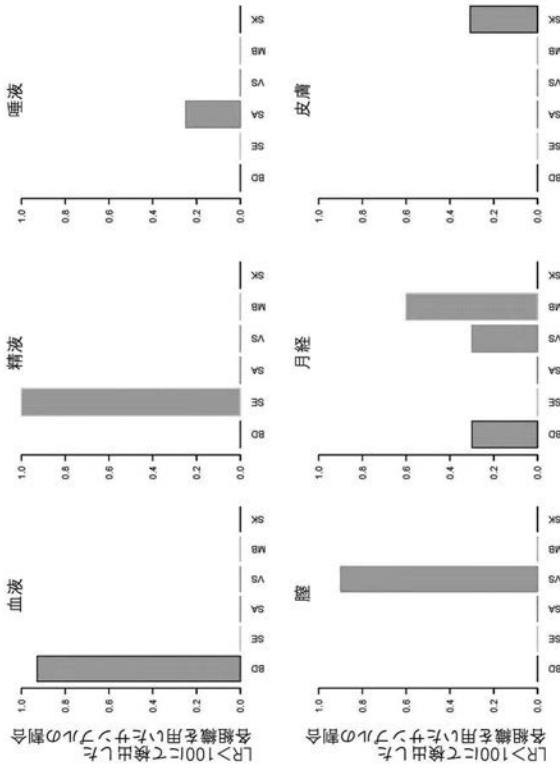


FIGURE 5

【 図 6 】

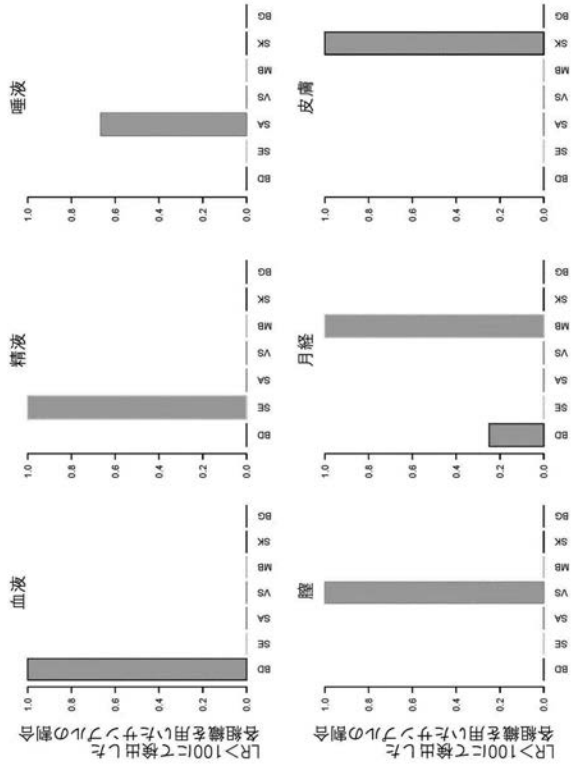


FIGURE 6

【 図 7 A 】

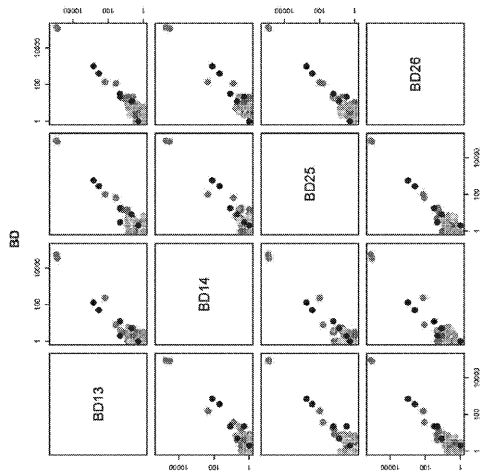


FIGURE 7A

【 図 7 B 】

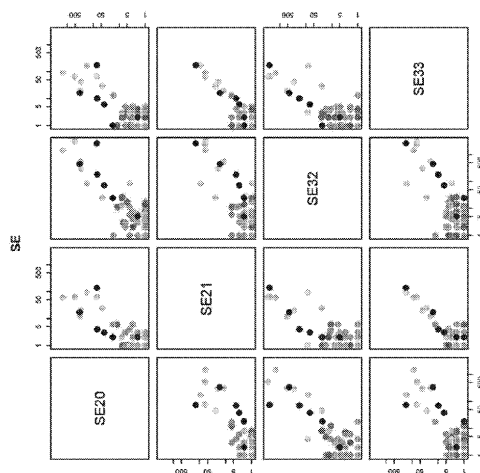


FIGURE 7B

【 7 C 】

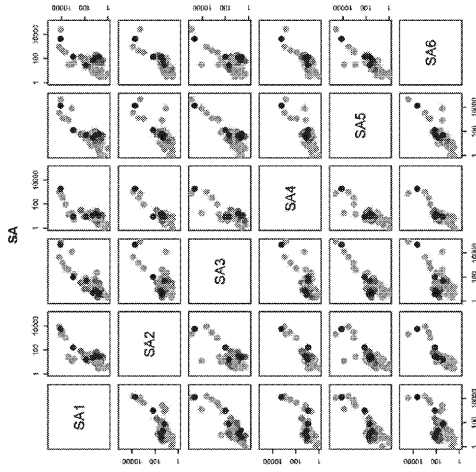


FIGURE 7C

【 7 D 】

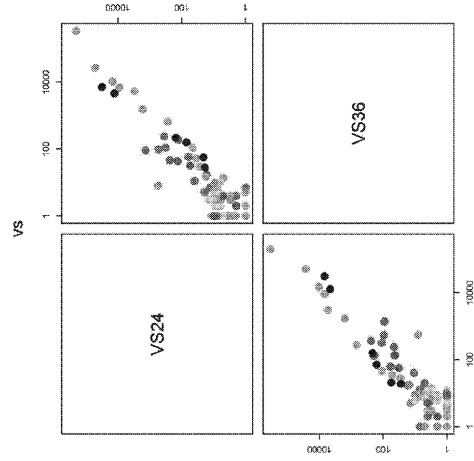


FIGURE 7D

【 7 E 】

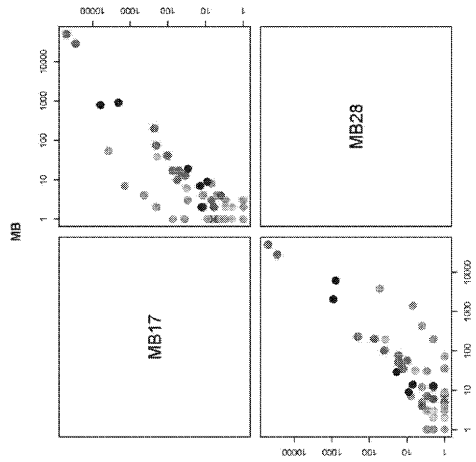


FIGURE 7E

【 7 F 】

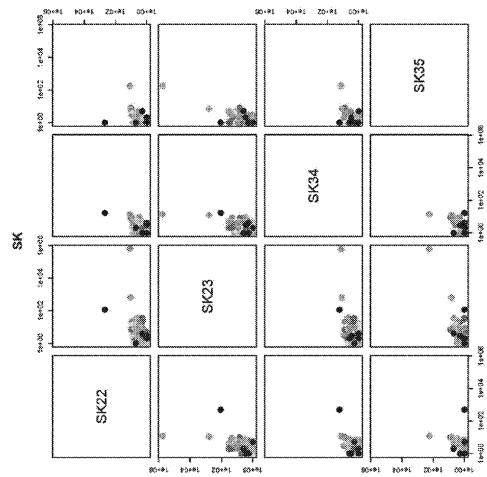


FIGURE 7F

【 8 】

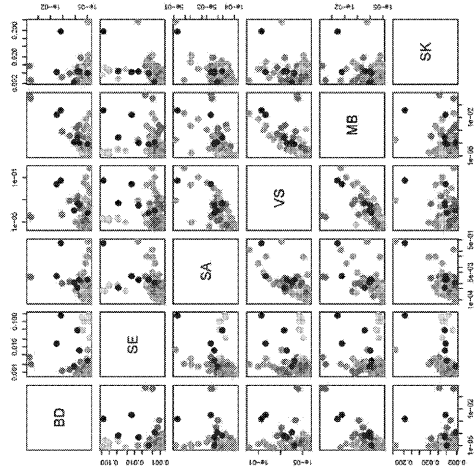


FIGURE 8

## 【 国際調査報告 】

## INTERNATIONAL SEARCH REPORT

International application No PCT/US2015/043609
---

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> INV. C12Q1/68 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols) C12Q		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, WPI Data, BIOSIS, EMBASE		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GEISS GARY K ET AL: "Direct multiplexed measurement of gene expression with color-coded probe pairs", NATURE BIOTECHNOLOGY, NATURE PUBLISHING GROUP, US, vol. 26, no. 3, 17 February 2008 (2008-02-17), pages 317-325, XP002505107, ISSN: 1087-0156, DOI: 10.1038/NBT1385 cited in the application	25,26
Y	the whole document ----- -/--	1-13
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents :		
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	
"E" earlier application or patent but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family	
"P" document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search 22 September 2015	Date of mailing of the international search report 09/10/2015	
Name and mailing address of the ISA/ European Patent Office, P.B. 6818 Patentlaan 2 NL-2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer Eveleigh, Anna	

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

## INTERNATIONAL SEARCH REPORT

International application No PCT/US2015/043609
---

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PARK SEONG-MIN ET AL: "Genome-wide mRNA profiling and multiplex quantitative RT-PCR for forensic body fluid identification", FORENSIC SCIENCE INTERNATIONAL: GENETICS, vol. 7, no. 1, 2013, pages 143-150, XP028959159, ISSN: 1872-4973, DOI: 10.1016/J.FSIGEN.2012.09.001	1-13
A	the whole document	25,26
A	JA-HYUN AN ET AL: "Body fluid identification in forensics", BMB REPORTS, vol. 45, no. 10, 31 October 2012 (2012-10-31), pages 545-553, XP055214907, ISSN: 1976-6696, DOI: 10.5483/BMBRep.2012.45.10.206 the whole document	1-13,25, 26
A	BALDI PIERRE ET AL: "A Bayesian framework for the analysis of microarray expression data: Regularized t-test and statistical inferences of gene changes", BIOINFORMATICS, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 17, no. 6, 2001, pages 509-519, XP002321472, ISSN: 1367-4803, DOI: 10.1093/BIOINFORMATICS/17.6.509 page 511, left-hand column, paragraph 2-3	1-13
A	HANED H ET AL: "The predictive value of the maximum likelihood estimator of the number of contributors to a DNA mixture", FORENSIC SCIENCE INTERNATIONAL: GENETICS, ELSEVIER BV, NETHERLANDS, vol. 5, no. 4, 21 April 2010 (2010-04-21), pages 281-284, XP028222027, ISSN: 1872-4973, DOI: 10.1016/J.FSIGEN.2010.04.005 [retrieved on 2010-04-28] the whole document	1-13
X,P	DANAHER PATRICK ET AL: "Facile semi-automated forensic body fluid identification by multiplex solution hybridization of NanoString(R) barcode probes to specific mRNA targets", FORENSIC SCIENCE INTERNATIONAL: GENETICS MAY 2012,, vol. 14, January 2015 (2015-01), pages 18-30, XP002743201, ISSN: 1878-0326 the whole document	1-13,25, 26

2

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

**INTERNATIONAL SEARCH REPORT**International application No.  
PCT/US2015/043609**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

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

1.  Claims Nos.: **14-24, 27(completely); 25(partially)**  
because they relate to subject matter not required to be searched by this Authority, namely:  
**see FURTHER INFORMATION sheet PCT/ISA/210**
2.  Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3.  Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

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

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of additional fees.
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

International Application No. PCT/US2015/043609

**FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210**

Continuation of Box II.1

Claims Nos.: 14-24, 27(completely); 25(partially)

Claims 14-24 and 27 do not appear to contain any technical features, since no gene expression is actually measured. The subject-matter of the claims therefore appears to relate merely to a method of mathematical calculations and has therefore not been searched, in accordance with Rule 39.1(i) and (iii) PCT. For the same reason, the subject-matter of claim 25 was not searched insofar as referring to the method of claims 14-24.

## フロントページの続き

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

(74)代理人 100117019

弁理士 渡辺 陽一

(74)代理人 100141977

弁理士 中島 勝

(74)代理人 100196977

弁理士 上原 路子

(72)発明者 パトリック ジョン ダナハー

アメリカ合衆国, ワシントン 9 8 1 0 5 , シアトル, ノースイースト フィフティセブンス ス  
トリート 4 0 5 0

Fターム(参考) 2G045 AA40 CA25 CB01 CB07 CB09 CB14 CB15 CB30 DA14 FB02

JA01

4B063 QA01 QA18 QQ03 QQ42 QQ53 QR56 QS34

专利名称(译)	使用基因表达数据对混合细胞群进行解卷积的方法		
公开(公告)号	<a href="#">JP2017530693A</a>	公开(公告)日	2017-10-19
申请号	JP2017506897	申请日	2015-08-04
申请(专利权)人(译)	纳米技术的字符串, Incorporated的雷开球德		
[标]发明人	パトリックジョンダナハー		
发明人	パトリック ジョン ダナハー		
IPC分类号	C12Q1/68 G01N33/53 G01N33/48		
CPC分类号	G16B25/00 C12Q1/6813 G16B35/00 G16B40/00 G16C20/60 C12Q2537/165		
FI分类号	C12Q1/68.A G01N33/53.M G01N33/48.Z		
F-TERM分类号	2G045/AA40 2G045/CA25 2G045/CB01 2G045/CB07 2G045/CB09 2G045/CB14 2G045/CB15 2G045/CB30 2G045/DA14 2G045/FB02 2G045/JA01 4B063/QA01 4B063/QA18 4B063/QQ03 4B063/QQ42 4B063/QQ53 4B063/QR56 4B063/QS34		
代理人(译)	青木 笃 石田 敬 渡边洋一 中岛胜		
优先权	62/035019 2014-08-08 US		
外部链接	<a href="#">Espacenet</a>		

摘要(译)

通过mRNA谱图鉴定体液可允许从法医样本中提取上下文“活动水平”信息。因此,基于颜色编码的探针(例如)的溶液杂交,提供了用于法医流体/组织鉴定的基本的杂交数字基因表达方法。例如,提供了用于在来自单一体液的样品中进行基因表达的模型,并将其扩展至体液的混合物。执行样品中流体体积的最大似然估计的计算,并描述了使用似然比来测试样品中每种流体的存在。描述了该方法/算法,但是与用于检测组织和细胞的常规算法不同,该方法/算法可以零识别多个样品之间的假阳性液体。这样的方案可以促进在以前不可靠的案例研究(例如法医学)实验室中mRNA谱的常规使用。

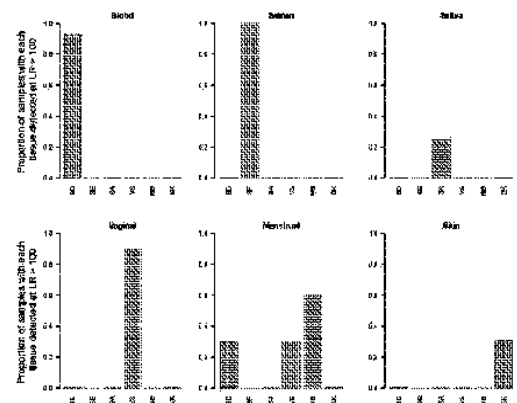


FIGURE 5