

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2017-104105

(P2017-104105A)

(43) 公開日 平成29年6月15日(2017.6.15)

(51) Int.Cl.	F I	テーマコード (参考)
C 1 2 N 5/074 (2010.01)	C 1 2 N 5/074 Z N A	2 G 0 4 5
C 1 2 N 5/0735 (2010.01)	C 1 2 N 5/0735	4 B 0 2 9
C 1 2 Q 1/02 (2006.01)	C 1 2 Q 1/02	4 B 0 6 3
C 1 2 Q 1/68 (2006.01)	C 1 2 Q 1/68 A	4 B 0 6 5
C 1 2 M 1/00 (2006.01)	C 1 2 Q 1/68 Z	4 C 0 8 7

審査請求 有 請求項の数 16 O L (全 202 頁) 最終頁に続く

(21) 出願番号 特願2016-233779 (P2016-233779)
 (22) 出願日 平成28年12月1日(2016.12.1)
 (62) 分割の表示 特願2013-529361 (P2013-529361) の分割
 原出願日 平成23年9月16日(2011.9.16)
 (31) 優先権主張番号 61/429,965
 (32) 優先日 平成23年1月5日(2011.1.5)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 61/384,030
 (32) 優先日 平成22年9月17日(2010.9.17)
 (33) 優先権主張国 米国 (US)

(特許庁注：以下のものは登録商標)

1. WINDOWS
2. UNIX
3. JAVA

(71) 出願人 507244910
 プレジデント・アンド・フェロウズ・オブ
 ・ハーバード・カレッジ
 アメリカ合衆国 マサチューセッツ州 ケ
 ンブリッジ クインシー ストリート 1
 7
 (74) 代理人 100102978
 弁理士 清水 初志
 (74) 代理人 100102118
 弁理士 春名 雅夫
 (74) 代理人 100160923
 弁理士 山口 裕孝
 (74) 代理人 100119507
 弁理士 刑部 俊

最終頁に続く

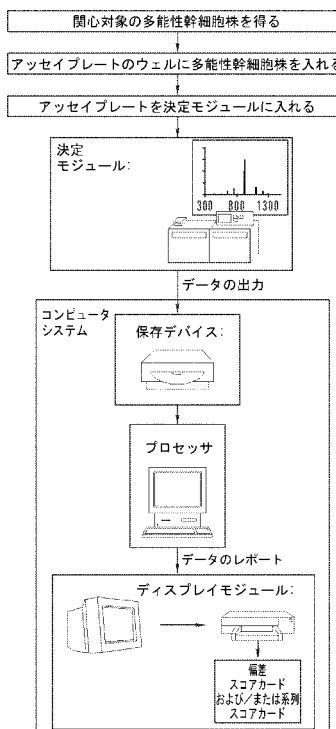
(54) 【発明の名称】 多能性幹細胞の有用性および安全性の特徴決定を行うための機能的ゲノミクスアッセイ

(57) 【要約】 (修正有)

【課題】 多能性幹細胞株の機能性及び適切性を予測するための多能性幹細胞に関する参照データセット、スコアカードを作成するための方法、システム及びキットの提供。

【解決手段】、多能性幹細胞株の機能性及び適切性を予測するための、エピジェネティックなプロファイリング、分化アッセイ、及び遺伝子発現アッセイから選択される少なくとも2つの幹細胞アッセイを用いることを含むスコアカードの作成法、システム及びキットの提供。スコアカード参照データを多能性幹細胞株データと比較して、所定の応用に関する多能性幹細胞の有用性を有効かつ正確に予測することができ、たとえば治療的使用、薬物スクリーニング、及び毒性アッセイ、及び細胞系列への分化についてのその適切性を決定するための多能性幹細胞株の特定の特徴を同定することができる。

【選択図】 図 2 5



【特許請求の範囲】

【請求項1】

以下の段階を含む、多能性幹細胞株を選択する方法：

- a. 多能性幹細胞株における標的遺伝子セットのDNAメチル化を測定して、DNAメチル化データと、同じ標的遺伝子の参照DNAメチル化データとの比較を行う段階；
- b. 複数の系列マーカー遺伝子の遺伝子発現および/またはDNAメチル化を測定することによって、多能性幹細胞の非定方向性または定方向性分化による多能性幹細胞株の分化能を測定する段階と；遺伝子発現および/またはDNAメチル化分化を、同じ系列マーカー遺伝子の参照遺伝子発現および/またはDNAメチル化分化と比較する段階；ならびに
- c. 参照DNAメチル化レベルと比較して標的遺伝子のDNAメチル化に統計学的に有意な量で差がなく、かつ参照分化能と比較して中胚葉、外胚葉、および内胚葉系列に沿って分化する指向性に統計学的に有意な量で差がない多能性幹細胞株を選択する段階；または、参照DNAメチル化レベルと比較して標的遺伝子のDNAメチル化に統計学的に有意な量で差があり、かつ参照分化能と比較して中胚葉、外胚葉、および内胚葉系列に沿って分化する指向性に統計学的に有意な量で差がある多能性幹細胞株を廃棄する段階。

10

【請求項2】

DNAメチル化が、少なくとも1つの多能性幹細胞に、DNAにおけるエピジェネティックな修飾に差次的に結合する剤を接触させることによって測定される、請求項1記載の方法。

【請求項3】

少なくとも1つの多能性幹細胞に、メチル化DNAおよび非メチル化DNAに差次的に結合する剤を接触させて、DNAメチル化データと、同じ標的遺伝子の参照DNAメチル化データとの比較を行うことによって、DNAメチル化を測定することができる、請求項2記載の方法。

20

【請求項4】

DNAメチル化を、濃縮に基づく方法（たとえば、MeDIP、MBD-seq、およびMethylCap）、バイサルファイトシークエンシングおよびバイサルファイトに基づく方法（たとえば、RRBS、バイサルファイトシークエンシング、Infinium、GoldenGate、COBRA、MSP、Methylight）、ならびに制限消化法（たとえば、MRE-seq）、または同じ標的遺伝子の参照DNAメチル化データと比較した多能性幹細胞のDNAメチル化標的遺伝子の差次的変換、差次的制限、差次的重量からなる群より選択されるいずれか1つによって測定することができる、請求項2記載の方法。

30

【請求項5】

以下の段階をさらに含む、請求項1~4のいずれか一項記載の方法：

- a. 多能性幹細胞株における第二の標的遺伝子セットの遺伝子発現を測定して、該遺伝子発現データと、同じ標的遺伝子の参照遺伝子発現レベルとの比較を行う段階；
- b. 参照遺伝子発現レベルと比較して標的遺伝子の遺伝子発現レベルに統計学的に有意な量で差がない多能性幹細胞株を選択する段階；または、参照遺伝子発現レベルと比較して標的遺伝子の発現レベルに統計学的に有意な量で差がある多能性幹細胞株を廃棄する段階。

【請求項6】

参照DNAメチル化レベルが、前記DNAメチル化標的遺伝子のメチル化の正常な変動の範囲である、請求項1~5のいずれか一項記載の方法。

40

【請求項7】

参照DNAメチル化レベルが、前記DNAメチル化標的遺伝子のDNAメチル化の平均値 ± 任意で1標準偏差であり、平均値が複数の多能性幹細胞株における前記標的遺伝子のDNAメチル化から算出される、請求項1~6のいずれか一項記載の方法。

【請求項8】

複数の多能性幹細胞株が少なくとも5個またはそれより多くの多能性幹細胞株である、請求項7記載の方法。

【請求項9】

多能性細胞株および/または参照DNAメチル化がバイサルファイトアッセイによって決

50

定される、請求項1~8のいずれか一項記載の方法。

【請求項10】

多能性細胞株および/または参照DNAメチル化が、全ゲノムバイサルファイトアッセイによって決定される、請求項1~9のいずれか一項記載の方法。

【請求項11】

多能性細胞株および/または参照DNAメチル化が、簡約表示(reduced representation)バイサルファイトシーケンシング(RBBS)アッセイによって決定される、請求項1~10のいずれか一項記載の方法。

【請求項12】

参照遺伝子発現レベルが前記標的遺伝子の正常な変動の範囲である、請求項5記載の方法。 10

【請求項13】

参照遺伝子発現レベルが前記標的遺伝子の発現レベルの平均値であり、平均値が、複数の多能性幹細胞株における該標的遺伝子の発現レベルから算出される、請求項5~12のいずれか一項記載の方法。

【請求項14】

複数の多能性幹細胞株が、少なくとも5個またはそれより多くの異なる多能性幹細胞株である、請求項13記載の方法。

【請求項15】

多能性細胞株および/または参照遺伝子発現がマイクロアレイアッセイによって決定される、請求項5~14のいずれか一項記載の方法。 20

【請求項16】

多能性細胞株の分化能が、定量的分化アッセイによって決定される、請求項1~15のいずれか一項記載の方法。

【請求項17】

参照分化能が、中胚葉、内胚葉、外胚葉、ニューロン、造血系列、およびこれらの任意の組み合わせからなる群より選択される系列への分化能である、請求項1~16のいずれか一項記載の方法。

【請求項18】

参照分化能データが、複数の多能性幹細胞株から作成される、請求項1~17のいずれか一項記載の方法。 30

【請求項19】

複数の多能性幹細胞株が少なくとも5個の異なる多能性幹細胞株である、請求項18記載の方法。

【請求項20】

多能性細胞株のDNAメチル化標的遺伝子および/または参照DNAメチル化標的遺伝子が、癌遺伝子、腫瘍遺伝子、腫瘍抑制遺伝子、発達遺伝子、系列マーカー遺伝子、およびこれらの任意の組み合わせからなる群より選択される、請求項1~19のいずれか一項記載の方法。

【請求項21】

多能性細胞株のDNAメチル化標的遺伝子および/または参照DNAメチル化標的遺伝子が、表12Aまたは表13Aまたは表14に記載される群、およびこれらの任意の組み合わせから選択される、請求項1~19のいずれか一項記載の方法。 40

【請求項22】

腫瘍遺伝子の遺伝子がc-Sis、上皮細胞増殖因子受容体、血小板由来増殖因子受容体、血管内皮増殖因子受容体、HER2/new、チロシンキナーゼのSrcファミリー、チロシンキナーゼのSyk-Zap-70ファミリー、チロシンキナーゼのBTKファミリーチロシンキナーゼ、Rafキナーゼ、サイクリン依存的キナーゼ、Rasタンパク質、およびmyc遺伝子から選択される、請求項20記載の方法。

【請求項23】

腫瘍抑制遺伝子が、TP53、PTEN、APC、CD95、ST5、ST7、およびST14遺伝子から選択される、請求項20記載の方法。

【請求項24】

発達遺伝子が、表7または表13Aまたは表14に記載される遺伝子の任意の組み合わせから選択される、請求項20記載の方法。

【請求項25】

系列マーカー遺伝子が、VEGF受容体II (KDR)、アクチン-2平滑筋 (ACTA2)、ネスチン、チューブリン-3、 α -フェトプロテイン (AFP)、シンデカン-4、CD64IFcyRI、Oct-4、 β -HCG、 β -LH、oct-3、Brachyury T、Fgf-5、nodal、GATA-4、flk-1、Nkx-2.5、EKLF、およびMsx3から選択される、請求項20記載の方法。

10

【請求項26】

多能性細胞株のDNAメチル化標的遺伝子および/または参照DNAメチル化標的遺伝子が、BMP4、CAT、CD14、CXCL5、DAZL、DNMT3B、GATA6、GAPDH、LEFTY2、MEG3、PAX6、S100A6、SOX2、SNAI1、TF、およびこれらの任意の組み合わせからなる群より選択される、請求項1~26のいずれか一項記載の方法。

【請求項27】

統計学的な差が、参照レベルからの少なくとも1標準偏差、少なくとも2標準偏差、または少なくとも3標準偏差の差である、請求項1~25のいずれか一項記載の方法。

【請求項28】

多能性細胞株の遺伝子発現標的遺伝子および/または参照遺伝子発現標的遺伝子が、表12Bまたは表13Aまたは表14において記載される群およびこれらの任意の組み合わせから選択される、請求項1~27のいずれか一項記載の方法。

20

【請求項29】

表12Aまたは表13Aまたは表14のリストにおける遺伝子の任意の組み合わせから選択される少なくとも約200個の標的遺伝子のDNAメチル化が、多能性細胞株において測定されて、同じ少なくとも200個の標的遺伝子セットの参照DNAメチル化レベルと比較される、請求項1~28のいずれか一項記載の方法。

【請求項30】

表12Aまたは表13Aまたは表14のリストにおける遺伝子の任意の組み合わせから選択される少なくとも約200個の標的遺伝子のDNAメチル化が、表12Aまたは表13Aまたは表14に記載される1~500番の遺伝子の任意の組み合わせから選択される、請求項1~29のいずれか一項記載の方法。

30

【請求項31】

少なくとも約200個の標的遺伝子のDNAメチル化が、表12Aまたは表13Aまたは表14に記載される1~200番から選択される、請求項1~30のいずれか一項記載の方法。

【請求項32】

表12Aまたは表13Aまたは表14のリストにおける遺伝子の任意の組み合わせから選択される少なくとも約500個の標的遺伝子のDNAメチル化が、多能性細胞株において測定されて、同じ少なくとも500個の標的遺伝子セットの参照DNAメチル化レベルと比較される、請求項1~31のいずれか一項記載の方法。

40

【請求項33】

表12Aまたは表13Aまたは表14のリストにおける遺伝子の任意の組み合わせから選択される少なくとも約500個の標的遺伝子のDNAメチル化が、表12Aまたは表13Aまたは表14に記載される1~1000番の遺伝子の任意の組み合わせから選択される、請求項1~32のいずれか一項記載の方法。

【請求項34】

少なくとも約500個の標的遺伝子のDNAメチル化が、表12Aまたは表13Aまたは表14に記載される1~500番から選択される、請求項1~33のいずれか一項記載の方法。

【請求項35】

表12Aまたは表13Aまたは表14のリストにおける遺伝子の任意の組み合わせから選択され

50

る少なくとも約1000個の標的遺伝子のDNAメチル化が、多能性細胞株において測定され、同じ少なくとも1000個の標的遺伝子セットの参照DNAメチル化レベルと比較される、請求項1~29のいずれか一項記載の方法。

【請求項36】

少なくとも約1000個の標的遺伝子のDNAメチル化が、表12Aまたは表13Aまたは表14に記載される1~2000番から選択される、請求項1~35のいずれか一項記載の方法。

【請求項37】

表12Bまたは表13Aまたは表14のリストにおける遺伝子の任意の組み合わせから選択される少なくとも約200個の標的遺伝子の遺伝子発現が、多能性細胞株において測定され、同じ少なくとも200個の標的遺伝子セットの参照遺伝子発現レベルと比較される、請求項1~36のいずれか一項記載の方法。

10

【請求項38】

少なくとも約200個の標的遺伝子の遺伝子発現が、表12Bまたは表13Aまたは表14に記載される1~500番から選択される、請求項1~37のいずれか一項記載の方法。

【請求項39】

表12Bまたは表13Aまたは表14のリストにおける遺伝子の任意の組み合わせから選択される少なくとも約500個の標的遺伝子の遺伝子発現が、多能性細胞株において測定され、同じ少なくとも500個の標的遺伝子セットの参照遺伝子発現レベルと比較される、請求項1~38のいずれか一項記載の方法。

【請求項40】

少なくとも約500個の標的遺伝子の遺伝子発現が、表12Bまたは表13Aまたは表14に記載される1~1000番から選択される、請求項1~39のいずれか一項記載の方法。

20

【請求項41】

表12Bまたは表13Aまたは表14のリストにおける遺伝子の任意の組み合わせから選択される少なくとも約1000個の標的遺伝子の遺伝子発現が、多能性細胞株において測定され、同じ少なくとも1000個の標的遺伝子セットの参照遺伝子発現レベルと比較される、請求項1~40のいずれか一項記載の方法。

【請求項42】

少なくとも約1000個の標的遺伝子の遺伝子発現が、表12Bまたは表13Aまたは表14に記載される1~2000番から選択される、請求項1~41のいずれか一項記載の方法。

30

【請求項43】

参照遺伝子と比較してメチル化に統計学的に有意な差を有する多能性幹細胞株におけるDNAメチル化遺伝子の数が、10、9、8、7、6、5、4、3、2、1、または0個である、請求項1~42のいずれか一項記載の方法。

【請求項44】

参照遺伝子と比較して遺伝子発現レベルに統計学的に有意な差を有する多能性幹細胞株における遺伝子の数が、10、9、8、7、6、5、4、3、2、1、または0個である、請求項1~43のいずれか一項記載の方法。

【請求項45】

多能性幹細胞が哺乳動物の多能性幹細胞である、請求項1~44のいずれか一項記載の方法。

40

【請求項46】

多能性幹細胞がヒト多能性幹細胞である、請求項1~45のいずれか一項記載の方法。

【請求項47】

生物活性に関して化合物をスクリーニングするための多能性幹細胞の使用であって、該多能性細胞が請求項1~46のいずれか一項記載の方法によって選択される、使用。

【請求項48】

スクリーニングが以下の段階を含む、請求項47記載の使用：

(iv) 任意で、多能性幹細胞を特定系列に沿って分化させるか、または分化することを可能にする段階；

50

(v) 該細胞に試験化合物を接触させる段階；および

(vi) 該細胞に及ぼす該化合物の任意の効果を決定する段階。

【請求項 49】

試験化合物が、有機低分子、無機低分子、多糖類、ペプチド、タンパク質、核酸、細菌、植物、真菌、動物細胞、動物組織などの生物材料から作製された抽出物、およびその任意の混合物からなる群より選択される、請求項47～48のいずれか一項記載の使用。

【請求項 50】

試験化合物が、約0.01 nMから約1000 mMの範囲の濃度で試験される、請求項47～49のいずれか一項記載の使用。

【請求項 51】

方法が、ハイスループットスクリーニング法である、請求項47～50のいずれか一項記載の使用。

【請求項 52】

生物活性が、生物アッセイにおける刺激、阻害、調節、毒性、または致死反応の誘発である、請求項47～51のいずれか一項記載の使用。

【請求項 53】

生物活性が、酵素活性の調節、受容体の不活化、受容体の刺激、1つまたは複数の遺伝子の発現レベルの調節、細胞増殖の調節、細胞分裂の調節、細胞形態の調節、およびこれらの任意の組み合わせからなる群より選択される、請求項47～52のいずれか一項記載の使用。

【請求項 54】

特定系列が、疾患の遺伝子型または表現型である、請求項47～53のいずれか一項記載の使用。

【請求項 55】

特定系列が、器官、組織、またはその一部の遺伝子型または表現型である、請求項47～54のいずれか一項記載の使用。

【請求項 56】

請求項1～46のいずれか一項記載の方法によって選択される多能性幹細胞を対象に投与する段階によって対象を処置するための、多能性幹細胞の使用。

【請求項 57】

対象が哺乳動物である、請求項56記載の使用。

【請求項 58】

対象がマウスである、請求項56～57のいずれか一項記載の使用。

【請求項 59】

対象がヒトである、請求項56～57のいずれか一項記載の使用。

【請求項 60】

対象が、癌、糖尿病、心不全、筋損傷、セリアック病、神経障害、神経変性障害、リソソーム蓄積症、およびこれらの任意の組み合わせからなる群より選択される疾患または状態を患っているかまたは有すると診断されている、請求項56～59のいずれか一項記載の使用。

【請求項 61】

投与が局所である、請求項56～60のいずれか一項記載の使用。

【請求項 62】

投与が多能性幹細胞の対象への移植である、請求項56～61のいずれか一項記載の使用。

【請求項 63】

多能性幹細胞またはその分化した子孫を対象に投与する前に、該多能性幹細胞を分化させる段階をさらに含む、請求項56～62のいずれか一項記載の使用。

【請求項 64】

多能性幹細胞が、中胚葉、内胚葉、外胚葉、ニューロン、造血系列、およびこれらの任意の組み合わせからなる群より選択される系列に沿って分化する、請求項63記載の使用。

10

20

30

40

50

【請求項 65】

多能性幹細胞が、インスリン産生細胞（膵細胞、細胞等）、神経細胞、筋細胞、皮膚細胞、心筋細胞、肝細胞、血液細胞、適応免疫細胞、および先天免疫細胞などへと分化する、請求項63～64のいずれか一項記載の使用。

【請求項 66】

請求項1～26のいずれか一項記載の方法によって選択される多能性幹細胞を含む、キット。

【請求項 67】

使用説明書をさらに含む、請求項66記載のキット。

【請求項 68】

多能性幹細胞が、請求項47～55のいずれか一項記載の使用にとって有用である、請求項66～67のいずれか一項記載のキット。

【請求項 69】

多能性幹細胞が、請求項56～65のいずれか一項記載の使用にとって有用である、請求項66～67のいずれか一項記載のキット。

【請求項 70】

以下の少なくとも2つを含む、多能性細胞の複数の特性の特徴決定を行うためのアッセイ：

- a. DNAメチル化アッセイ；
- b. 遺伝子発現アッセイ；および
- c. 分化アッセイ。

【請求項 71】

DNAメチル化アッセイがパイサルファイトシークエンシングアッセイである、請求項70記載のアッセイ。

【請求項 72】

DNAメチル化アッセイが、全ゲノムパイサルファイトシークエンシングアッセイである、請求項70～71のいずれか一項記載のアッセイ。

【請求項 73】

DNAメチル化アッセイが、濃縮に基づく方法（たとえば、MeDIP、MBD-seq、およびMethylCap）、パイサルファイトシークエンシングおよびパイサルファイトに基づく方法（たとえば、RRBS、パイサルファイトシークエンシング、Infinium、GoldenGate、COBRA、MSP、MethylLight）ならびに制限消化法（たとえば、MRE-seq）からなる群より選択される、請求項70～72のいずれか一項記載のアッセイ。

【請求項 74】

遺伝子発現アッセイがマイクロアレイアッセイである、請求項70～73のいずれか一項記載のアッセイ。

【請求項 75】

分化アッセイが、定量的分化アッセイである、請求項70～74のいずれか一項記載のアッセイ。

【請求項 76】

分化アッセイが、以下の系列、すなわち中胚葉、内胚葉、外胚葉、ニューロン、または造血系列のうち少なくとも1つへの多能性細胞の分化能を評価する、請求項70～75のいずれか一項記載のアッセイ。

【請求項 77】

以下の系列、すなわち中胚葉、内胚葉、および外胚葉のうち少なくとも1つへの多能性細胞の分化能が、中胚葉、内胚葉、および外胚葉系列の少なくとも1つのマーカーに対する抗体を用いる免疫染色またはFACソーティングによって決定される、請求項70～76のいずれか一項記載のアッセイ。

【請求項 78】

以下の系列、すなわち中胚葉、内胚葉、および外胚葉のうち少なくとも1つへの多能性

10

20

30

40

50

細胞の分化能が、EBにおいて少なくとも約7日後に多能性幹細胞を免疫染色することによって決定される、請求項70～77のいずれか一項記載のアッセイ。

【請求項79】

中胚葉系列に沿った多能性細胞の分化能が、VEGF受容体II（KDR）またはアクチン-2平滑筋（ACTA2）に関する免疫染色陽性によって決定される、請求項70～78のいずれか一項記載のアッセイ。

【請求項80】

外胚葉系列に沿った多能性細胞の分化能が、ネスチンまたはチューブリン-3に関する免疫染色陽性によって決定される、請求項70～79のいずれか一項記載のアッセイ。

【請求項81】

内胚葉系列に沿った多能性細胞の分化能が、 α -フェトプロテイン（AFP）に関する免疫染色陽性によって決定される、請求項70～80のいずれか一項記載のアッセイ。

【請求項82】

複数の異なる多能性幹細胞をアッセイするためのハイスループットアッセイである、請求項70～81のいずれか一項記載のアッセイ。

【請求項83】

ハイスループットアッセイが、対象由来の複数の異なる人工多能性幹細胞を評価する、請求項81記載のアッセイ。

【請求項84】

対象が哺乳動物である、請求項83記載のアッセイ。

【請求項85】

対象がヒト対象である、請求項83記載のアッセイ。

【請求項86】

DNAメチル化遺伝子が、癌遺伝子、腫瘍遺伝子、腫瘍抑制遺伝子、発達遺伝子、系列マーカー遺伝子、およびこれらの任意の組み合わせからなる群より選択される、請求項70～85のいずれか一項記載のアッセイ。

【請求項87】

DNAメチル化遺伝子が、BMP4、CAT、CD14、CXCL5、DAZL、DNMT3B、GATA6、GAPDH、LEFTY2、MEG3、PAX6、S100A6、SOX2、SNAI1、TF、およびこれらの任意の組み合わせからなる群より選択される、請求項70～86のいずれか一項記載のアッセイ。

【請求項88】

遺伝子発現アッセイが、表7または表13Aまたは表14に記載される遺伝子の任意の組み合わせから選択される遺伝子の発現を決定する、請求項70～86のいずれか一項記載のアッセイ。

【請求項89】

DNAメチル化アッセイが、表12Aまたは表13Aまたは表14に記載される群から選択される複数の標的遺伝子の任意の組み合わせのDNAメチル化レベルを決定する、請求項70～88のいずれか一項記載のアッセイ。

【請求項90】

DNAメチル化アッセイが、表12Aまたは表13Aまたは表14に記載される少なくとも200個の遺伝子の任意の組み合わせのDNAメチル化レベルを決定する、請求項70～89のいずれか一項記載のアッセイ。

【請求項91】

DNAメチル化アッセイが、表12Aまたは表13Aまたは表14に記載される1～500番の遺伝子のうち少なくとも200個の遺伝子の任意の組み合わせのDNAメチル化レベルを決定する、請求項70～89のいずれか一項記載のアッセイ。

【請求項92】

DNAメチル化アッセイが、表12Aまたは表13Aまたは表14に記載される少なくとも500個の遺伝子の任意の組み合わせのDNAメチル化レベルを決定する、請求項70～91のいずれか一項記載のアッセイ。

10

20

30

40

50

【請求項 9 3】

DNAメチル化アッセイが、表12Aに記載される1～1000番の遺伝子のうち少なくとも500個の遺伝子の任意の組み合わせのDNAメチル化レベルを決定する、請求項70～92のいずれか一項記載のアッセイ。

【請求項 9 4】

DNAメチル化アッセイが、表12Aまたは表13Aまたは表14に記載される少なくとも1000個の遺伝子の任意の組み合わせのDNAメチル化レベルを決定する、請求項70～93のいずれか一項記載のアッセイ。

【請求項 9 5】

DNAメチル化アッセイが、表12Aまたは表13Aまたは表14に記載される1～2000番の遺伝子のうち少なくとも1000個の遺伝子の任意の組み合わせのDNAメチル化レベルを決定する、請求項70～92のいずれか一項記載のアッセイ。

10

【請求項 9 6】

遺伝子発現アッセイが、表12Bに記載される群から選択される複数の標的遺伝子の任意の組み合わせの遺伝子発現レベルを決定する、請求項70～95のいずれか一項記載のアッセイ。

【請求項 9 7】

遺伝子発現アッセイが、表12Bまたは表13Aまたは表14に記載される少なくとも200個の遺伝子の任意の組み合わせの遺伝子発現レベルを決定する、請求項70～96のいずれか一項記載のアッセイ。

20

【請求項 9 8】

遺伝子発現アッセイが、表12Bまたは表13Aまたは表14に記載される1～500番の遺伝子のうち少なくとも200個の遺伝子の任意の組み合わせの遺伝子発現レベルを決定する、請求項70～97のいずれか一項記載のアッセイ。

【請求項 9 9】

遺伝子発現アッセイが、表12Bまたは表13Aまたは表14に記載される少なくとも500個の遺伝子の任意の組み合わせの遺伝子発現レベルを決定する、請求項70～96のいずれか一項記載のアッセイ。

【請求項 1 0 0】

遺伝子発現アッセイが、表12Bまたは表13Aまたは表14に記載される1～1000番の遺伝子のうち少なくとも500個の遺伝子の任意の組み合わせの遺伝子発現レベルを決定する、請求項70～97のいずれか一項記載のアッセイ。

30

【請求項 1 0 1】

遺伝子発現アッセイが、表12Bまたは表13Aまたは表14に記載される少なくとも1000個の遺伝子の任意の組み合わせの遺伝子発現レベルを決定する、請求項70～96のいずれか一項記載のアッセイ。

【請求項 1 0 2】

遺伝子発現アッセイが、表12Bまたは表13Aまたは表14に記載される1～2000番の遺伝子のうち少なくとも1000個の遺伝子の任意の組み合わせの遺伝子発現レベルを決定する、請求項70～97のいずれか一項記載のアッセイ。

40

【請求項 1 0 3】

少なくとも1つまたは複数の多能性幹細胞株からスコアカードを作成するための、請求項70～102のいずれか一項記載のアッセイの使用。

【請求項 1 0 4】

(i) 複数の多能性幹細胞株における第一の標的遺伝子セットにおけるDNAメチル化を測定する段階；

(ii) 複数の多能性幹細胞株における第二の標的遺伝子セットにおける遺伝子発現を測定する段階；および

(iii) 複数の多能性幹細胞株の分化能を測定する段階を含む、多能性幹細胞スコアカードを作成する方法。

50

【請求項105】

(i) 第一の標的遺伝子セットにおける各標的遺伝子の平均メチル化レベルを算出する段階；

(ii) 第二の標的遺伝子セットにおける各標的遺伝子の平均遺伝子発現レベルを算出する段階

をさらに含む、請求項104記載の方法。

【請求項106】

分化能が、中胚葉、内胚葉、外胚葉、ニューロン、造血系列、およびこれらの任意の組み合わせからなる群より選択される系列への分化能である、請求項104～105のいずれか一項記載の方法。

10

【請求項107】

複数の多能性幹細胞株が少なくとも5個の多能性幹細胞株である、請求項104～106のいずれか一項記載の方法。

【請求項108】

DNAメチル化が、パイサルファイトシークエンシングアッセイによって測定される、請求項104～107のいずれか一項記載の方法。

【請求項109】

DNAメチル化が全ゲノムパイサルファイトシークエンシングアッセイによって測定される、請求項104～108のいずれか一項記載の方法。

20

【請求項110】

DNAメチル化が、濃縮に基づく方法（たとえば、MeDIP、MBD-seq、およびMethylCap）、パイサルファイトシークエンシングおよびパイサルファイトに基づく方法（たとえば、RRBS、パイサルファイトシークエンシング、Infinium、GoldenGate、COBRA、MSP、MethylLight）、ならびに制限消化法（たとえば、MRE-seq）の群から選択される方法のいずれか1つによって測定される、請求項104～109のいずれか一項記載の方法。

【請求項111】

遺伝子発現がマイクロアレイアッセイによって測定される、請求項104～110のいずれか一項記載の方法。

【請求項112】

分化能が定量的分化アッセイによって測定される、請求項104～111のいずれか一項記載のアッセイ。

30

【請求項113】

以下の系列、すなわち中胚葉、内胚葉、および外胚葉のうち少なくとも1つへの多能性細胞の分化能が、中胚葉、内胚葉、および外胚葉系列に関する少なくとも1つのマーカーに対する抗体を用いる免疫染色またはFACソーティングによって決定される、請求項104～112のいずれか一項記載の方法。

【請求項114】

以下の系列、すなわち中胚葉、内胚葉、および外胚葉のうち少なくとも1つへの多能性細胞の分化能が、EBにおいて少なくとも約7日後に多能性幹細胞を免疫染色することによって決定される、請求項104～113のいずれか一項記載の方法。

40

【請求項115】

中胚葉系列に沿った多能性細胞の分化能が、VEGF受容体II（KDR）またはアクチン-2平滑筋（ACTA2）に関する免疫染色陽性によって決定される、請求項104～114のいずれか一項記載の方法。

【請求項116】

外胚葉系列に沿った多能性細胞の分化能が、ネスチンまたはチューブリン-3に関する免疫染色陽性によって決定される、請求項104～115のいずれか一項記載の方法。

【請求項117】

内胚葉系列に沿った多能性細胞の分化能が、 α -フェトプロテイン（AFP）に関する免疫染色陽性によって決定される、請求項104～116のいずれか一項記載の方法。

50

【請求項 1 1 8】

第一の遺伝子セットが、癌遺伝子、腫瘍遺伝子、腫瘍抑制遺伝子、発達遺伝子、系列マーカー遺伝子、およびこれらの任意の組み合わせからなる群より選択される、請求項104～117のいずれか一項記載の方法。

【請求項 1 1 9】

第一の遺伝子セットが、BMP4、CAT、CD14、CXCL5、DAZL、DNMT3B、GATA6、GAPDH、LEFT Y2、MEG3、PAX6、S100A6、SOX2、SNAI1、TF、およびこれらの任意の組み合わせからなる群より選択される少なくとも1つの遺伝子を含む、請求項104～118のいずれか一項記載の方法。

【請求項 1 2 0】

第一のDNAメチル化遺伝子セットが、表12Aまたは表13Aまたは表14に記載される群から選択される複数の標的遺伝子の任意の組み合わせを含む、請求項104～119のいずれか一項記載の方法。

【請求項 1 2 1】

第一のDNAメチル化遺伝子セットが、表12Aまたは表13Aまたは表14に記載される少なくとも200個の遺伝子の任意の組み合わせを含む、請求項104～120のいずれか一項記載の方法。

【請求項 1 2 2】

第一のDNAメチル化遺伝子セットが、表12Aまたは表13Aまたは表14に記載される1～500番の遺伝子のうち少なくとも200個の遺伝子の任意の組み合わせを含む、請求項104～121のいずれか一項記載の方法。

【請求項 1 2 3】

第一のDNAメチル化遺伝子セットが、表12Aまたは表13Aまたは表14に記載される少なくとも500個の遺伝子の任意の組み合わせを含む、請求項104～122のいずれか一項記載の方法。

【請求項 1 2 4】

第一のDNAメチル化遺伝子セットが、表12Aまたは表13Aまたは表14に記載される1～1000番の遺伝子のうち少なくとも500個の遺伝子の任意の組み合わせを含む、請求項104～123のいずれか一項記載の方法。

【請求項 1 2 5】

第一のDNAメチル化遺伝子セットが、表12Aまたは表13Aまたは表14に記載される少なくとも1000個の遺伝子の任意の組み合わせを含む、請求項104～124のいずれか一項記載の方法。

【請求項 1 2 6】

第一のDNAメチル化遺伝子セットが、表12Aまたは表13Aまたは表14に記載される1～2000番の遺伝子のうち少なくとも1000個の遺伝子の任意の組み合わせを含む、請求項104～125のいずれか一項記載の方法。

【請求項 1 2 7】

第二の遺伝子発現遺伝子セットが、表12Bまたは表13Aまたは表14に記載される群から選択される複数の標的遺伝子の任意の組み合わせを含む、請求項104～126のいずれか一項記載の方法。

【請求項 1 2 8】

第二の遺伝子発現遺伝子セットが、表12Bまたは表13Aまたは表14に記載される少なくとも200個の遺伝子の任意の組み合わせを含む、請求項104～127のいずれか一項記載の方法。

【請求項 1 2 9】

第二の遺伝子発現遺伝子セットが、表12Bまたは表13Aまたは表14に記載される1～500番の遺伝子のうち少なくとも200個の遺伝子の任意の組み合わせを含む、請求項104～128のいずれか一項記載の方法。

【請求項 1 3 0】

10

20

30

40

50

第二の遺伝子発現遺伝子セットが、表12Bまたは表13Aまたは表14に記載される少なくとも500個の遺伝子の任意の組み合わせを含む、請求項104～129のいずれか一項記載の方法。

【請求項131】

第二の遺伝子発現遺伝子セットが、表12Bまたは表13Aまたは表14に記載される1～1000番の遺伝子のうち少なくとも500個の遺伝子の任意の組み合わせを含む、請求項104～130のいずれか一項記載の方法。

【請求項132】

第二の遺伝子発現遺伝子セットが、表12Bに記載される少なくとも1000個の遺伝子の任意の組み合わせを含む、請求項104～131のいずれか一項記載の方法。

10

【請求項133】

第二の遺伝子発現遺伝子セットが、表12Bまたは表13Aまたは表14に記載される1～2000番の遺伝子のうち少なくとも1000個の遺伝子の任意の組み合わせを含む、請求項104～132のいずれか一項記載の方法。

【請求項134】

(i) 複数の多能性幹細胞株由来の複数のDNAメチル化標的遺伝子のDNAメチル化レベルを含む第一のデータセット；

(ii) 複数の多能性幹細胞株由来の複数の遺伝子発現標的遺伝子の遺伝子発現レベルを含む第二のデータセット；ならびに

(iii) 複数の多能性幹細胞株由来の外胚葉、中胚葉、および内胚葉系列への分化に関する分化指向性レベルを含む第三のデータセット

20

を含む、多能性幹細胞の性能パラメータのスコアカード。

【請求項135】

複数の参照DNAメチル化遺伝子が、少なくとも約500個、少なくとも約1000個、少なくとも約1500個、または少なくとも約200個の参照DNAメチル化遺伝子である、請求項134記載のスコアカード。

【請求項136】

複数の参照DNAメチル化遺伝子が、表12Aまたは表13Aまたは表14に記載される遺伝子の任意の組み合わせから選択される、請求項134または135記載のスコアカード。

【請求項137】

複数の参照DNAメチル化遺伝子が、表12Aまたは表13Aまたは表14に記載される遺伝子の任意の組み合わせから選択される、請求項134または136記載のスコアカード。

30

【請求項138】

複数の参照DNAメチル化遺伝子が、表12Aまたは表13Aまたは表14に記載される少なくとも200個の遺伝子の任意の組み合わせから選択される、請求項134～137のいずれか一項記載のスコアカード。

【請求項139】

複数の参照DNAメチル化遺伝子が、表12Aまたは表13Aまたは表14に記載される1～500番の遺伝子のうち少なくとも200個の遺伝子の任意の組み合わせから選択される、請求項134～138のいずれか一項記載のスコアカード。

40

【請求項140】

複数の参照DNAメチル化遺伝子が、表12Aまたは表13Aまたは表14に記載される少なくとも500個の遺伝子の任意の組み合わせから選択される、請求項134～139のいずれか一項記載のスコアカード。

【請求項141】

複数の参照DNAメチル化遺伝子が、表12Aまたは表13Aまたは表14に記載される1～1000番の遺伝子のうち少なくとも500個の遺伝子の任意の組み合わせから選択される、請求項134～140のいずれか一項記載のスコアカード。

【請求項142】

複数の参照DNAメチル化遺伝子が、表12Aまたは表13Aまたは表14に記載される少なくと

50

も1000個の遺伝子の任意の組み合わせから選択される、請求項134～141のいずれか一項記載のスコアカード。

【請求項143】

複数の参照DNAメチル化遺伝子が、表12Aまたは表13Aまたは表14に記載される1～2000番の遺伝子のうち少なくとも1000個の遺伝子の任意の組み合わせから選択される、請求項134～142のいずれか一項記載のスコアカード。

【請求項144】

複数の参照DNAメチル化遺伝子が全ゲノムのDNAメチル化状態である、請求項134～143のいずれか一項記載のスコアカード。

【請求項145】

複数の参照DNAメチル化遺伝子が、癌遺伝子、腫瘍遺伝子、腫瘍抑制遺伝子、発達遺伝子、および系列マーカー遺伝子を含む、請求項134～144のいずれか一項記載のスコアカード。

【請求項146】

複数の参照DNAメチル化遺伝子が、BMP4、CAT、CD14、CXCL5、DAZL、DNMT3B、GATA6、GAPDH、LEFTY2、MEG3、PAX6、S100A6、SOX2、SNAI1、TF、およびこれらの任意の組み合わせからなる群より選択される少なくとも1つの遺伝子を含む、請求項134～145のいずれか一項記載のスコアカード。

【請求項147】

少なくとも第一および/または第二のデータセットがデータ保存デバイスに接続される、請求項134～146のいずれか一項記載のスコアカード。

【請求項148】

少なくとも第一および/または第二のデータセットが、データ保存デバイスに接続され、データ保存デバイスがコンピュータデバイス上に存在するデータベースである、請求項134～147のいずれか一項記載のスコアカード。

【請求項149】

複数の幹細胞株が、少なくとも5個、少なくとも10個、少なくとも15個、または少なくとも20個の多能性幹細胞株である、請求項134～148のいずれか一項記載のスコアカード。

【請求項150】

複数の幹細胞株が、HUES64、HUES3、HUES8、HUES53、HUES28、HUES49、HUES9、HUES48、HUES45、HUES1、HUES44、HUES6、H1、HUES62、HUES65、H7、HUES13、HUES63、HUES66、およびこれらの任意の組み合わせからなる群より選択される少なくとも1つの多能性幹細胞株を含む、請求項134～149のいずれか一項記載のスコアカード。

【請求項151】

複数の幹細胞株が、HUES64、HUES3、HUES8、HUES53、HUES28、HUES49、HUES9、HUES48、HUES45、HUES1、HUES44、HUES6、H1、HUES62、HUES65、H7、HUES13、HUES63、HUES66からなる群より独立して選択される少なくとも5個の多能性幹細胞株を含む、請求項134～149のいずれか一項記載のスコアカード。

【請求項152】

複数の多能性幹細胞株が、少なくとも1つの哺乳動物多能性幹細胞株を含む、請求項134～151のいずれか一項記載のスコアカード。

【請求項153】

複数の多能性幹細胞株の全ての多能性幹細胞株が、哺乳動物多能性幹細胞株である、請求項134～152のいずれか一項記載のスコアカード。

【請求項154】

複数の多能性幹細胞株が、少なくともヒト多能性幹細胞株を含む、請求項134～153のいずれか一項記載のスコアカード。

【請求項155】

複数の多能性幹細胞株の全ての多能性幹細胞株が、ヒト多能性幹細胞株である、請求項134～154のいずれか一項記載のスコアカード。

10

20

30

40

50

- 【請求項 1 5 6】
多能性幹細胞が哺乳動物多能性幹細胞である、請求項134～155のいずれか一項記載のスコアカード。
- 【請求項 1 5 7】
多能性幹細胞がヒト多能性幹細胞である、請求項134～156のいずれか一項記載のスコアカード。
- 【請求項 1 5 8】
多能性幹細胞が人工多能性幹（iPS）細胞である、請求項134～157のいずれか一項記載のスコアカード。
- 【請求項 1 5 9】 10
多能性幹細胞が胚幹細胞である、請求項134～158のいずれか一項記載のスコアカード。
- 【請求項 1 6 0】
多能性幹細胞が成体幹細胞である、請求項134～159のいずれか一項記載のスコアカード。
- 【請求項 1 6 1】
多能性幹細胞が自己幹細胞である、請求項134～160のいずれか一項記載のスコアカード。
- 【請求項 1 6 2】
請求項134～161のいずれか一項記載のスコアカードを含む、キット。
- 【請求項 1 6 3】 20
使用説明書をさらに含む、請求項162記載のキット。
- 【請求項 1 6 4】
人工多能性幹細胞を胚幹細胞株と区別するための、請求項134～161のいずれか一項記載のスコアカードの使用。
- 【請求項 1 6 5】
（iii）DNAメチル化状態を測定するための試薬；および
（iv）多能性幹細胞の分化指向性を測定するための試薬
を含む、請求項1～46のいずれか一項記載の方法を行うためのキット。
- 【請求項 1 6 6】 30
標的遺伝子発現遺伝子の遺伝子発現レベルを測定するための試薬をさらに含む、請求項165記載のキット。
- 【請求項 1 6 7】
使用説明書をさらに含む、請求項165～166のいずれか一項記載のキット。
- 【請求項 1 6 8】
請求項134～161のいずれか一項記載のスコアカードをさらに含む、請求項165～166のいずれか一項記載のキット。
- 【請求項 1 6 9】 40
（c）（i）多能性幹細胞株におけるDNAメチル化標的遺伝子セットのDNAメチル化データを受信して、該DNAメチル化データと、同じ標的遺伝子の参照DNAメチル化レベルとの比較を行う段階；
（ii）多能性幹細胞株の分化能データを受信して、該分化能データを参照分化能データと比較する段階；
（iii）参照DNAメチル化パラメータと比較したDNAメチル化データの比較、および参照分化データと比較した分化指向性の比較に基づいて、品質保証スコアカードを作成する段階
を含む少なくとも1つのプログラムを含む、少なくとも1つのメモリ；ならびに
（d）プログラムを作動させるためのプロセッサ
を含む、多能性幹細胞の品質保証スコアカードを作成するためのコンピュータシステム。
- 【請求項 1 7 0】 50
プログラムが以下の段階をさらに含む、請求項169記載のシステム：

(i) 多能性幹細胞株における第二の標的遺伝子セットの遺伝子発現データを受信して、該発現データを、同じ第二の標的遺伝子セットの参照遺伝子発現レベルと比較する段階；

(ii) 参照DNAメチル化パラメータと比較したDNAメチル化データの比較、および参照分化データと比較した分化指向性の比較、および参照遺伝子発現レベルと比較した遺伝子発現データの比較に基づいて、品質保証スコアカードを作成する段階。

【請求項 171】

DNAメチル化標的遺伝子が多様なメチル化を有する、請求項169～170のいずれか一項記載のシステム。

【請求項 172】

DNAメチル化標的遺伝子が、癌遺伝子、腫瘍遺伝子、腫瘍抑制遺伝子、発達遺伝子、系列マーカー遺伝子、およびこれらの任意の組み合わせから選択される、請求項169～171のいずれか一項記載のシステム。

【請求項 173】

DNAメチル化標的遺伝子が、BMP4、CAT、CD14、CXCL5、DAZL、DNMT3B、GATA6、GAPDH、LEF1、MEG3、PAX6、S100A6、SOX2、SNAI1、TF、およびこれらの任意の組み合わせからなる群より選択される、請求項169～172のいずれか一項記載のシステム。

【請求項 174】

参照DNAメチル化レベルが、腫瘍遺伝子のエピジェネティックなサイレンシングに関しては高レベルのメチル化、ならびに腫瘍抑制遺伝子および発達遺伝子の活発な転写に関しては低レベルのメチル化である、請求項169～173のいずれか一項記載のシステム。

【請求項 175】

DNAメチル化標的遺伝子が表12Aに記載される遺伝子の任意の組み合わせから選択される、請求項167～174のいずれか一項記載のシステム。

【請求項 176】

DNAメチル化標的遺伝子が、表12Aに記載される少なくとも200個の遺伝子から選択される、請求項167～175のいずれか一項記載のシステム。

【請求項 177】

DNAメチル化標的遺伝子が、表12Aまたは表13Aまたは表14に記載される1～500番の遺伝子のうち少なくとも200個の遺伝子の任意の組み合わせから選択される、請求項167～176のいずれか一項記載のシステム。

【請求項 178】

DNAメチル化標的遺伝子が、表12Aに記載される少なくとも500個の遺伝子から選択される、請求項167～177のいずれか一項記載のシステム。

【請求項 179】

DNAメチル化標的遺伝子が、表12Aまたは表13Aまたは表14に記載される1～1000番の遺伝子のうち少なくとも500個の遺伝子の任意の組み合わせから選択される、請求項167～178のいずれか一項記載のシステム。

【請求項 180】

DNAメチル化標的遺伝子が、表12Aに記載される少なくとも1000個の遺伝子から選択される、請求項167～179のいずれか一項記載のシステム。

【請求項 181】

DNAメチル化標的遺伝子が、表12Aまたは表13Aまたは表14に記載される1～3000番の遺伝子のうち少なくとも1000個の遺伝子の任意の組み合わせから選択される、請求項167～180のいずれか一項記載のシステム。

【請求項 182】

多能性幹細胞株の品質に基づいて幹細胞スコアカードレポートを作成するレポート作成モジュールをさらに含む、請求項167～181のいずれか一項記載のシステム。

【請求項 183】

メモリがデータベースをさらに含む、請求項167～182のいずれか一項記載のシステム。

10

20

30

40

50

- 【請求項184】
データベースがDNAメチル化遺伝子セットを階層的様式で配置する、請求項167～183のいずれか一項記載のシステム。
- 【請求項185】
データベースが、異なる系列への分化指向性を階層的様式で配置する、請求項167～184のいずれか一項記載のシステム。
- 【請求項186】
データベースが、遺伝子発現レベルデータセットを階層的様式で配置する、請求項167～185のいずれか一項記載のシステム。
- 【請求項187】 10
メモリが、ネットワークを介して第一のコンピュータに接続される、請求項167～186のいずれか一項記載のシステム。
- 【請求項188】
ネットワークが広域ネットワークを含む、請求項187記載のシステム。
- 【請求項189】
スコアカードが、多能性幹細胞の適切な用途または適用の指標を提供する、請求項167～188のいずれか一項記載のシステム。
- 【請求項190】 20
参照DNAメチル化レベルが、前記DNAメチル化標的遺伝子に関するメチル化の正常な変動の範囲である、請求項167～189のいずれか一項記載のシステム。
- 【請求項191】
参照DNAメチル化レベルが、前記DNAメチル化標的遺伝子のDNAメチル化の平均値であり、平均値が、複数の多能性幹細胞株における前記標的遺伝子のDNAメチル化から算出される、請求項167～190のいずれか一項記載のシステム。
- 【請求項192】
多能性細胞株の分化能が、定量的分化アッセイによって決定される、請求項167～191のいずれか一項記載のシステム。
- 【請求項193】 30
参照分化能が、中胚葉、内胚葉、外胚葉、ニューロン、造血系列、およびこれらの任意の組み合わせからなる群より選択される系列への分化能である、請求項167～192のいずれか一項記載のシステム。
- 【請求項194】
参照遺伝子発現レベルが、前記遺伝子発現標的遺伝子の遺伝子発現の正常な変動の範囲である、請求項167～193のいずれか一項記載のシステム。
- 【請求項195】
参照遺伝子発現レベルが、前記標的遺伝子の遺伝子発現レベルの平均値であり、平均値が、複数の多能性幹細胞株における前記標的遺伝子の発現レベルから算出される、請求項111～128のいずれか一項記載の方法。
- 【請求項196】 40
参照DNAメチル化、参照分化能データ、および参照遺伝子発現レベルデータが、複数の多能性幹細胞株から作成される、請求項167～194のいずれか一項記載のシステム。
- 【請求項197】
複数の多能性幹細胞株が、少なくとも5個、少なくとも10個、少なくとも15個、または少なくとも20個の多能性幹細胞株である、請求項196記載のシステム。
- 【請求項198】
DNAメチル化標的遺伝子が、少なくとも1つまたは複数の遺伝子発現標的遺伝子を含む、請求項167～197のいずれか一項記載のシステム。
- 【請求項199】 50
遺伝子発現標的遺伝子が少なくとも1つまたは複数のDNAメチル化標的遺伝子を含む、請求項167～198のいずれか一項記載のシステム。

【請求項200】

以下を含む、多能性幹細胞株の品質保証スコアカードの作成に関する命令を含む、コンピュータ読み取り可能な媒体：

(i) 多能性幹細胞株におけるDNAメチル化標的遺伝子セットのDNAメチル化データを受信して、該DNAメチル化データと、同じ標的遺伝子の参照DNAメチル化レベルとの比較を行うこと；

(ii) 多能性幹細胞株の分化能データを受信して、該分化能データを参照分化能データと比較すること；ならびに

(iii) 参照DNAメチル化パラメータと比較したDNAメチル化データの比較、および参照分化データと比較した分化指向性の比較に基づいて、品質保証スコアカードを作成すること。

10

【請求項201】

以下に関する命令をさらに含む、請求項200記載のコンピュータ読み取り可能な媒体：

a. 多能性幹細胞株における第二の標的遺伝子セットの遺伝子発現データを受信して、該発現データを同じ第二の標的遺伝子セットの参照遺伝子発現レベルと比較すること；ならびに

b. 参照DNAメチル化パラメータと比較したDNAメチル化データの比較、および参照分化データと比較した分化指向性の比較、および参照遺伝子発現レベルと比較した遺伝子発現データの比較に基づいて、品質保証スコアカードを作成すること。

【請求項202】

以下の少なくとも2つを含む、多能性幹細胞株の品質を決定するためのキット：

a. 複数のDNAメチル化遺伝子のメチル化状態を測定するための試薬；

b. 複数の遺伝子の遺伝子発現レベルを測定するための試薬；ならびに

c. 多能性幹細胞の外胚葉、中胚葉、および内胚葉系列への分化指向性を測定するための試薬。

20

【請求項203】

使用説明書をさらに含む、請求項202記載のキット。

【請求項204】

少なくとも1つの多能性幹細胞株をさらに含む、請求項202～203のいずれか一項記載のキット。

30

【請求項205】

請求項134～161のいずれか一項記載のスコアカードをさらに含む、請求項202～204のいずれか一項記載のキット。

【請求項206】

a. (i) DNAメチル化標的遺伝子セットのDNAメチル化データを得ること、および関心対象の少なくとも1つの多能性幹細胞株における遺伝子発現遺伝子セットの遺伝子発現データを得ること、ならびに

(ii) DNAメチル化標的遺伝子セットのDNAメチル化データを得ること、および少なくとも1つの参照多能性幹細胞株における遺伝子発現遺伝子セットの遺伝子発現データを得ること、

40

(iii) 要素(i)および(ii)で得られた遺伝子発現データのデータ正規化を行うこと、

(iv) 要素(i)および(ii)で得られたDNAメチル化データおよび遺伝子発現データの遺伝子マッピングを行うこと、

(v) 要素(i)および(iii)で得られた関心対象の多能性幹細胞株由来のDNAメチル化データおよび正規化遺伝子発現データを、要素(ii)および(iii)で得られた参照多能性幹細胞株由来の標準化DNAメチル化データおよび正規化遺伝子発現データと比較して、参照多能性幹細胞株のDNAメチル化レベルまたは遺伝子発現レベルの正常な範囲から統計学的に有意な量で外れているDNAメチル化レベルまたは正規化遺伝子発現レベルを有する多能性幹細胞株における遺伝子を同定すること、

50

(vi) 要素(v)で同定された遺伝子の関連性フィルターを適用して、参照多能性幹細胞株の参照DNAメチル化レベルまたは遺伝子発現レベルと比較して、15%を超えるDNAメチル化の差または1.5倍を超える遺伝子発現の変化を有する遺伝子を同定すること、

(vii) DNAメチル化標的遺伝子および遺伝子発現標的遺伝子および系列マーカーの遺伝子セットを得ること

のうち1つまたは複数を実施するのに適合させた1つまたは複数のプログラムを実行するための関連メモリおよびプロセッサを、コンピュータに設ける段階；ならびに

b. 少なくとも1つの参照多能性幹細胞株と比較した関心対象の多能性幹細胞株におけるDNAメチル化および/または遺伝子発現の偏差を有する、要素(vi)で同定された遺伝子の数および/または遺伝子の数の百分率を含む多能性スコアカードレポートを作成する段階

を含む、関心対象の幹細胞株の多能性を同定するためのスコアカードを作成する方法。

【請求項207】

段階(v)で同定された遺伝子が、参照多能性幹細胞株の正常なDNAメチル化範囲または遺伝子発現範囲の四分位数間範囲の少なくとも1.2倍で中心四分位数の外側となるDNAメチル化レベルまたは正規化遺伝子発現レベルを有する、請求項206記載の方法。

【請求項208】

段階(vi)で同定された遺伝子が、参照多能性幹細胞株の参照DNAメチル化レベルまたは遺伝子発現レベルと比較して、20%を超えるDNAメチル化の差または2倍を超える遺伝子発現の変化を有する、請求項206記載の方法。

【請求項209】

レポートスコアカードが、少なくとも1つの参照多能性幹細胞株と比較して関心対象の多能性幹細胞株におけるDNAメチル化および/または遺伝子発現から外れている、影響を受けた遺伝子の名称をさらに含む、請求項206記載の方法。

【請求項210】

DNAメチル化データが、全ゲノムDNAメチル化、または簡約表示バイサルファイトシーケンシング(RRBS)によって得られる、請求項206記載の方法。

【請求項211】

遺伝子発現データが、マイクロアレイデータまたは定量的PCR(qPCR)によって得られる、請求項206記載の方法。

【請求項212】

DNAメチル化標的遺伝子の遺伝子セットにおいて、遺伝子発現標的遺伝子および系列マーカーが、表7、表12A、表12B、表12C、表13A、表13B、または表14から選択される群から選択される表に記載される、請求項206記載の方法。

【請求項213】

コンピュータ上で行われる、請求項206~212のいずれか一項記載の方法。

【請求項214】

コンピュータシステムである、請求項206~213のいずれか一項記載の方法。

【請求項215】

1つまたは複数のプログラムが、コンピュータ読み取り可能な媒体上のスコアカードソフトウェアプログラムによって行われる、請求項206~214のいずれか一項記載の方法。

【請求項216】

a. (i) 関心対象の少なくとも1つの多能性幹細胞株の胚様体(EB)における標的系列マーカー遺伝子セットのDNAメチル化データおよび遺伝子発現データを得ること、ならびに

(ii) 少なくとも1つの参照多能性幹細胞株の胚様体(EB)における標的系列マーカー遺伝子セットのDNAメチル化データおよび遺伝子発現データを得ること、

(iii) 任意で、要素(i)および(ii)で得られたDNAメチル化データおよび遺伝子発現データを陽性対照を用いてサイズ変更することによって、アッセイの標準化を行うこと

、

10

20

30

40

50

(iv) 任意で、レプリケート実験全体にわたって要素(i)および(ii)で得られたDNAメチル化データおよび遺伝子発現データの試料の標準化および変動の安定化を行うこと、

(v) 要素(i)で得られた関心対象の多能性幹細胞株由来の系列マーカー遺伝子のDNAメチル化データおよび遺伝子発現データを、要素(ii)で得られた参照多能性幹細胞株由来の系列マーカー遺伝子のDNAメチル化データおよび遺伝子発現データと比較し、参照多能性幹細胞株のDNAメチル化レベルまたは遺伝子発現レベルの正常な範囲と比較して統計学的に有意な量で増加または減少するDNAメチル化レベルまたは標準化遺伝子発現レベルを有する多能性幹細胞株における系列遺伝子を同定し、それによって各個々の系列マーカー遺伝子に関する分散値を作成すること、

(vi) 関心対象の特徴的な細胞系列または胚葉に関する系列マーカー遺伝子の遺伝子セットを得ること、

(vii) 要素(vi)で得られた系列マーカー遺伝子セットに記載される各系列マーカーに関する個々の変動値(要素(v)で得られた)から変動の平均値を算出することによって、濃縮分析を行うこと

のうち1つまたは複数を実施するのに適合させた1つまたは複数のプログラムを実行するための関連メモリおよびプロセッサを、コンピュータに設ける段階;ならびに

b. 少なくとも1つの参照多能性幹細胞株と比較して多能性幹細胞株の系列マーカー遺伝子セットにおける全ての遺伝子に関する変動の平均値を含む系列スコアカードレポートを作成する段階

をシステムが含む、関心対象の多能性幹細胞株の分化指向性を同定するための系列スコアカードを作成するための方法。

【請求項217】

多能性幹細胞株が、請求項206記載のスコアカードを特徴としている、請求項216記載の方法。

【請求項218】

DNAメチル化データおよび遺伝子発現データに関する標的系列遺伝子マーカーセットが、表7、表13A、表13B、または表14から選択される群から選択される表に記載される、請求項216~217のいずれか一項記載の方法。

【請求項219】

要素(v)における参照との比較が、参照多能性幹細胞株のDNAメチル化または遺伝子発現と比較してDNAメチル化または遺伝子発現の統計学的に有意な増加または減少を有する系列マーカー遺伝子を同定するために、調整(moderated)t-検定を用いる、請求項216~218のいずれか一項記載の方法。

【請求項220】

調整t-検定を用いる参照との比較が、Bioconductors Limmaパッケージを用いて行われる、請求項216~219のいずれか一項記載の方法。

【請求項221】

系列マーカー遺伝子セットを、遺伝子オントロジー、MoSigDBプログラムまたはキュレーションによって得ることができる、請求項216~220のいずれか一項記載の方法。

【請求項222】

要素(vii)の濃縮分析が、各系列マーカーに関する個々のt-スコアからt-スコア平均値を算出する、請求項216~221のいずれか一項記載の方法。

【請求項223】

要素(iv)の試料標準化が、Bioconductor VSNパッケージによって行われる、請求項216記載の方法。

【請求項224】

要素(vi)における系列マーカー遺伝子セットが、外胚葉、中胚葉、内胚葉、神経系列遺伝子セット、造血系列遺伝子セット、多能性細胞シグネチャー遺伝子セット、表皮系列遺伝子セット、間葉幹細胞系列遺伝子セット、骨系列遺伝子セット、軟骨系列遺伝子セット、脂肪系列遺伝子セット、筋系列遺伝子セット、血管系列遺伝子セット、心臓系列遺伝

10

20

30

40

50

子セット、リンパ球系列遺伝子セット、骨髄細胞系列遺伝子セット、肝臓系列遺伝子セット、膵臓系列遺伝子セット、上皮系列遺伝子セット、運動ニューロン系列遺伝子セット、単球-マクロファージ系列遺伝子セット、iSCi系列遺伝子セット、または表7もしくは13Aおよび13Bおよび表14に記載される遺伝子の任意の選択物の群から選択される遺伝子セットである、請求項216~223のいずれか一項記載の方法。

【請求項225】

コンピュータ上で行われる、請求項216~224のいずれか一項記載の方法。

【請求項226】

システムがコンピュータシステムである、請求項216~225のいずれか一項記載の方法。

【請求項227】

1つまたは複数のプログラムが、コンピュータ読み取り可能な媒体上のスコアカードソフトウェアプログラムによって行われる、請求項216~226のいずれか一項記載の方法。

【請求項228】

a. 関心対象の多能性幹細胞株における、DNAメチル化標的遺伝子のDNAメチル化レベルおよび/または遺伝子発現標的遺伝子の遺伝子発現レベルを測定するための決定モジュール；

b. (i) 決定モジュールによって測定されたDNAメチル化レベルおよび遺伝子発現レベルを保存するための、ならびに1つまたは複数の参照多能性幹細胞株のDNAメチル化標的遺伝子の参照DNAメチル化レベルおよび遺伝子発現標的遺伝子の参照遺伝子発現レベルを保存するための、保存モジュール；

(ii) 決定モジュールによって測定された遺伝子発現レベルを正規化するための正規化モジュール；

(iii) 多能性幹細胞株において測定したDNAメチル化標的遺伝子のDNAメチル化レベルを、1つもしくは複数の参照多能性幹細胞株のDNAメチル化標的遺伝子のDNAメチル化レベルとマッチさせるための、および/または多能性幹細胞株において測定された遺伝子発現標的遺伝子の遺伝子発現レベルを、1つもしくは複数の参照多能性幹細胞株の遺伝子発現標的遺伝子の遺伝子発現レベルとマッチさせるための、遺伝子マッピングモジュール；

(iv) (i) 関心対象の多能性幹細胞株由来のDNAメチル化標的遺伝子のDNAメチル化レベルを、1つもしくは複数の参照多能性幹細胞株由来の同じDNAメチル化標的遺伝子のDNAメチル化レベルと比較するための、および/または(ii) 関心対象の多能性幹細胞株の遺伝子発現標的遺伝子の遺伝子発現レベルを、1つもしくは複数の参照多能性幹細胞株由来の同じ遺伝子発現標的遺伝子の遺伝子発現レベルと比較して、参照多能性幹細胞株のDNAメチル化レベルまたは遺伝子発現レベルの正常範囲から統計学的に有意な量で外れているDNAメチル化レベルまたは正規化遺伝子発現レベルを有する多能性幹細胞株における遺伝子を同定するための、比較モジュール；

(v) 参照多能性幹細胞株の参照DNAメチル化レベルまたは遺伝子発現レベルと比較して、少なくとも15%を超えるDNAメチル化の差または少なくとも1.5倍を超える遺伝子発現の変化を有する比較モジュールによって同定された遺伝子を選択するための関連性フィルターモジュール；

(vi) 関心対象の比較モジュールおよび/または関連性フィルターモジュールによって同定された遺伝子を選択するための遺伝子セットモジュール

のうち1つまたは複数を含む、プロセッサおよび関連メモリを含むコンピュータモジュール；

c. 少なくとも1つの参照多能性幹細胞株と比較した関心対象の多能性幹細胞株におけるDNAメチル化および/または遺伝子発現の偏差を有する、比較モジュールおよび/または関連性フィルターモジュールおよび/または遺伝子セットモジュールによって同定された遺伝子の数および/または遺伝子の数の百分率を含むスコアカードレポートを表示するための、ディスプレイモジュール

のうち少なくとも1つまたは複数を含む、関心対象の幹細胞株の多能性を同定するためのスコアカードを作成するシステム。

10

20

30

40

50

【請求項 2 2 9】

決定モジュールが、1つまたは複数の参照多能性幹細胞株におけるDNAメチル化標的遺伝子のDNAメチル化レベルおよび/または遺伝子発現遺伝子もしくは系列マーカー遺伝子の遺伝子発現レベルを測定することができる、請求項228記載のシステム。

【請求項 2 3 0】

保存モジュールが、1つまたは複数の参照多能性幹細胞株におけるDNAメチル化標的遺伝子のDNAメチル化レベルおよび/または遺伝子発現遺伝子もしくは系列マーカー遺伝子の遺伝子発現レベルの測定値を保存することができる、請求項228記載のシステム。

【請求項 2 3 1】

1つまたは複数のモジュールを組み合わせることで1つのモジュールにすることができる、請求項228記載のシステム。

10

【請求項 2 3 2】

a. 関心対象の多能性幹細胞株の胚様体 (EB) における複数の系列マーカー遺伝子の系列遺伝子発現レベルを測定するための決定モジュール；

b. (i) 決定モジュールによって測定された系列遺伝子発現レベルを保存するための、および1つまたは複数の参照多能性幹細胞株の胚様体 (EB) における系列マーカー遺伝子の参照系列遺伝子発現レベルを保存するための、保存モジュール；

(ii) 遺伝子発現陽性対照に基づいて遺伝子発現レベルを標準化するためのアッセイ標準化モジュール；

(iii) 関心対象の同じ多能性幹細胞株由来の胚様体 (EB) における同じ系列マーカー遺伝子のレプリケート遺伝子発現レベル測定値全体にわたり系列マーカー遺伝子の遺伝子発現レベルを標準化および変動安定化するための試料標準化モジュール；

20

(iv) 関心対象の多能性幹細胞株由来の胚様体 (EB) 由来の系列マーカー遺伝子の遺伝子発現レベルを、1つまたは複数の参照多能性幹細胞株由来の胚様体 (EB) からの同じ系列マーカー遺伝子の遺伝子発現レベルと比較して、各系列マーカー遺伝子に関する参照多能性幹細胞株の系列遺伝子発現レベルと比較した多能性幹細胞株における系列遺伝子発現レベルの差の統計学的差を算出するための、比較モジュール；

(v) 関心対象の特定の細胞系列に特徴的な系列マーカー遺伝子のサブセットを選択するための遺伝子セットモジュール；

(vi) 遺伝子セットモジュールによって選択された系列マーカー遺伝子のサブセットの遺伝子の比較モジュールによって算出された統計学的差の平均値を算出するための濃縮分析モジュール

30

のうち1つまたは複数を含む、プロセッサおよび関連メモリを含むコンピュータモジュール；

c. 少なくとも1つの参照多能性幹細胞株と比較した多能性幹細胞株の系列マーカー遺伝子セットの各サブセットにおける系列マーカー遺伝子に関する系列遺伝子発現の統計学的差の平均値を含む系列スコアカードレポートを表示するための、ディスプレイモジュールのうち少なくとも1つまたは複数を含む、関心対象の幹細胞株の分化指向性を同定するための系列スコアカードを作成するためのシステム。

【請求項 2 3 3】

1つまたは複数のモジュールを組み合わせることで1つのモジュールにすることができる、請求項232記載のシステム。

40

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

関連出願の相互参照

本出願は、35 U.S.C. 119(e)の下で、その全内容が参照により本明細書に組み入れられる、2010年9月17日に提出された米国特許仮出願第61/384,030号および2011年1月5日に提出された米国特許仮出願第61/429,965号の優先権を主張する。

【0 0 0 2】

50

発明の分野

本発明は、ハイスループット法による特徴決定などの、幹細胞を特徴決定するための方法、ならびに疾患のモデリング、幹細胞集団の研究、および疾患の治療的処置のためのその使用のための多能性細胞株の選択を標準化および最適化するための方法および組成物に関する。

【0003】

政府の支援

本発明は一部、米国国立衛生研究所によって与えられたエピゲノミクスに関するNIHロードマップイニシアティブ助成金番号U01ES017155により米国政府の支援を受けて行われた。米国政府は、本発明において一定の権利を有する。

10

【0004】

表の参照

本出願は、当初にファイルされた主題の一部として、「コピー1」および「コピー2」および「コピー3」と表示される3枚のコンパクトディスクを含み、各々のディスクは11個のテキストファイルを含む。コンパクトディスク（「コピー1」、「コピー2」、および「コピー3」）の各々は、10個の個別の長い表に関する11個のテキストファイルを含み、ファイル名は以下のとおりである：「002806-067741-P2_TABLE 3.txt」（9,919 KB、1/7/2011作成）、「002806-067741-P2_TABLE 4.txt」（19,381 KB、1/7/2011作成）、「002806-067741-P2_TABLE 5.txt」（10,006 KB、1/7/2011作成）、「002806-067741-P2_TABLE 8.txt」（98 KB、1/7/2011作成）、「002806-067741-P2_TABLE 10.txt」（180 KB、1/7/2011作成）、「002806-067741-P2_TABLE 12A.txt」（160 KB、1/7/2011作成）、「002806-067741-P2_TABLE 12B.txt」（160 KB、1/7/2011作成）、「002806-067741-P2_TABLE 12C.txt」（31 KB、1/7/2011作成）、「002806-067741-P2_TABLE 13A.txt」（25 KB、1/7/2011作成）、「002806-067741-P2_TABLE 13B.txt」（28 KB、1/7/2011作成）、「002806-067741-P2_TABLE 14.txt」（10KB、1/7/2011作成）。各コンパクトディスク（「コピー1」、「コピー2」、および「コピー3」）の機器フォーマットは、IBM-PCであり、各コンパクトディスクのオペレーティングシステムは、MS-ウィンドウズである。「コピー1」、「コピー2」、および「コピー3」と表示されるコンパクトディスクの内容は、その全体が参照により本明細書に組み入れられる。

20

【0005】

長い表

本明細書は、11個の長い表、すなわち表3、表4、表5、表8、表10、表12A、表12B、表12C、表13A、表13Bおよび表14を含む。長い表3は、Ensembl遺伝子およびプロモーター領域（Ensembl注釈付き転写開始部位周囲の-5 kbから+1 kbとして定義される）に関するDNAメチル化および遺伝子発現をまとめたデータであり、本明細書においてCD上のファイル「002806-067741-P2_TABLE 3.txt」として電子フォーマットで提供される。長い表4は、全てのES細胞株においてエピジェネティックな変動の多い順に選別した、35個の細胞株および31,929個のEnsembl遺伝子プロモーター領域に関するDNAメチル化データであり（BFの縦列）、本明細書においてCD上のファイル「002806-067741-P2_TABLE 4.txt」として電子フォーマットで提供される。長い表5は、全てのES細胞株における転写の変動の多い順に選別した、35個の細胞株および15,079個のEnsembl遺伝子の遺伝子発現データであり（BGの縦列）、本明細書においてCD上のファイル「002806-067741-P2_TABLE 5.txt」として電子フォーマットで提供される。長い表8は、系列スコアカード予測に参与する個々の測定の詳細に関する表であり、本明細書においてCD上のファイル「002806-067741-P2_TABLE 8.txt」として電子フォーマットで提供される。長い表10は、系列スコアカードの構築および検証のために用いられる遺伝子発現データの表であり、本明細書においてCD上のファイル「002806-067741-P2_TABLE 10.txt」として電子フォーマットで提供される。長い表12A、表12B、および12Cは、スコアカードまたはアッセイ、および方法において用いるための標的遺伝子のリストであり、表12Aは、ヒト多能性細胞株における参照DNAメチル化変動セットの変動性に基づいて同定されている、優先順位の高い順に記載される遺伝子を示し、表

30

40

50

12Bは、ヒト多能性細胞株における参照遺伝子発現変動セットの変動性に基づいて同定されている、優先順位の高い順に記載される遺伝子を示し、および表12Cは、優先順位の高い順に記載され、統計的順位づけおよび情報検索スキームを用いて文献から検索されている遺伝子を示し、表12Aおよび/または表12Bおよび/または表12Cの遺伝子は、スコアカードを決定するために用いることができ、それぞれ本明細書において、CD上のファイル「002806-067741-P2_TABLE 12A.txt」、「002806-067741-P2_TABLE 12B.txt」および「002806-067741-P2_TABLE 12C.txt」として電子フォーマットで提供される。長い表13Aおよび13Bは、スコアカードおよび系列スコアカードを決定するためのDNAメチル化および遺伝子発現測定のために用いることができる「含まれる遺伝子」として記載される標的遺伝子のもう1つのリストであり、本明細書においてそれぞれCD上のファイル「002806-067741-P2_ 10
TABLE 13A.txt」および「002806-067741-P2_TABLE 13B.txt」として電子フォーマットで提供される。長い表14は、スコアカードおよび系列スコアカードを決定するためのDNAメチル化および遺伝子発現測定のために用いることができる表13Aの遺伝子の亜群である標的遺伝子のもう1つのリストであり、CD上のファイル「002806-067741-P2_TABLE 14.txt」として電子フォーマットで提供される。本明細書において、CD上でファイル「002806-067741-P2_TABLE 3.txt」；「002806-067741-P2_TABLE 4.txt」；「002806-067741-P2_TABLE 5.txt」；「002806-067741-P2_TABLE 8.txt」；「002806-067741-P2_TABLE 10.txt」；「002806-067741-P2_TABLE 12A.txt」；「002806-067741-P2_TABLE 12B.txt」、「002806-067741-P2_TABLE 12C.txt」、「002806-067741-P2_TABLE 13A.txt」、「002806-067741-P2_ 20
TABLE 13B.txt」および「002806-067741-P2_TABLE 14.txt」としてそれぞれ、電子フォーマットで提供される表3、表4、表5、表8、表10および表12A~12Cは、その全内容が参照により本明細書に組み入れられる。アクセスの説明に関しては本明細書末尾を参照されたい。

【背景技術】

【0006】

発明の背景

再生医療の1つの目標は、組織修復および再生のために、多能性細胞を他の細胞型に変換できることである。ヒト多能性細胞株は、初期胚と類似である発達可塑性レベルを示し、インビトロで3つ全ての胚性胚葉へと分化することができる (Rossant, 2008; Thomson et al., 1998)。同時に、これらの多能性細胞株は、何代もの継代のあいだ未分化状態で維持することが可能である (Adewumi et al., 2007)。これらの独自の特徴により、ヒト胚幹細胞 (ES) およびヒト人工多能性幹 (iPS) 細胞は、生物医学研究のための有望なツールとなっている (Colman and Dreesen, 2009)。ES細胞株は既に、単一遺伝子性ヒト疾患の細胞基本原理を分析するためのモデル系として確立されている。たとえば、脆弱X症候群を引き起こす変異を有するES細胞は、インビトロで分化させると、この疾患の表現型局面を再現することが示されている (Eiges et al., 2007)。さらに、ヒトES細胞由来運動ニューロンは、薬物スクリーニングに用いるために適合する家族性筋萎縮性側索硬化症 (ALS) のインビトロモデルを開発するために用いられている (Di Giorgio et al., 2008)。明確なリプログラミング法の発見 (Takahashi and Yamanaka, 2006) および患者特異的 iPS細胞株の誘導におけるその使用 (Dimos et al., 2008; Park et al., 2008) により、単一遺伝子性疾患モデリングのための多能性細胞の有用性はさらに拡大され、脊髄筋萎縮 (Ebert et al., 2009) および家族性自律神経障害 (Lee et al., 2009) のインビトロでの研究が可能となっている。

【0007】

最近まで、生物医学研究のために広く利用することができるヒト多能性細胞株はごく少数であった。この理由から、研究者らは、たいいていの場合、これらの容易に入手可能で良好に特徴決定された細胞株 (たとえば、Thomson、bresigen、およびHUES 1-17細胞株) に依存している。さらに、米国におけるES細胞研究に対する資金上の制限により、広く用いられる細胞株の数はさらに制限された。その結果、研究者らは、自身の関心対象の応用にとって利用可能である細胞株を用いるしかなく、所定のアッセイにおいて細胞株がどのよ

10

20

30

40

50

うな挙動を示すかを予測することができる診断の必要性はほとんどなかった。

【0008】

胚幹細胞は、培養において有意な期間、多能性を維持することができるという点において独自であり、それにより、胚幹細胞は細胞治療において用いるための第1候補となっている。胚幹(ES)細胞の分化は、系列特異的遺伝子発現パターンを制御するためのエピソード的な機序を伴う。ES細胞に基づく治療は、現在難治性の多くの遺伝障害、外傷障害、および変性障害を処置するために非常に有望である。しかし、これらの治療戦略は、所望の前駆細胞(たとえば、体幹細胞等)を提供するためにエクスピボで維持され、操作され、および/または分化しているヒト細胞の導入を必然的に伴い、そのため、異常な細胞(たとえば、そのような操作および分化プロトコールの際に生じうる癌細胞または癌

10

【0009】

しかし、最近のいくつかの進展により、多能性ヒト細胞株の挙動を予測することができる診断の必要性は大きく増加している。第一に、多くの研究室によってヒトES細胞株が絶えず誘導されていること、および米国における資金制限の解除により、研究者らが選択しうるES細胞株の数は実質的に増加している。さらに、必ずしも全てのヒトES細胞株があらゆる目的に等しく適しているわけではないことも明らかとなっている(Osafune et al., 2008)。このことは、いかなる新規研究プロジェクトも、関心対象の応用にとって最も適格である細胞株を、熟考して情報を得た上で選択すべきであることを示唆している。

20

【0010】

患者からの体細胞をiPS細胞へとリプログラムする因子が発見されたこともまた、研究者らが利用可能なおよび研究者らが使用する多能性細胞株の数をさらに増加させている。研究者らが既存の細胞株を収集する際、または関心対象の応用のために新しい株を誘導する際に、最も使用に適している細胞株を選択する方法に関しては、ほとんど情報または手引きがない。

【0011】

ヒト多能性幹細胞株のさらなる応用には、個人の遺伝子型とその環境とのあいだの複雑な相互作用の結果として生じる共通の疾患の研究が含まれる可能性が高いと考えられている(Colman and Dreesen, 2009)。さらに、多能性細胞は、最終的に移植医療のための細胞および組織の双方の再生可能な起源として役立つであろう(Daley, 2010)。多能性幹細胞に関して提唱されるこれらの応用はいずれも、疾患に関連する細胞型へと信頼可能に、再現可能に、効率よく、および安定的に分化する細胞株の選択を必要とするであろう。しかし、様々なヒトES細胞株が3つの胚性胚葉の異なる派生物へと分化する効率には有意な量の変動があることが報告されている(Di Giorgio et al., 2008; Osafune et al., 2008)。多能性幹細胞株間の変動が機能に及ぼす結末に関する懸念は、iPS細胞株の研究によってさらに高まっている。具体的に、iPS細胞は全体的に、何百もの遺伝子の発現(Chin et al., 2009)、そのゲノム全体のDNAメチル化パターン(Doi et al., 2009)、およびその運動ニューロン系列への分化能(Hu et al., 2010)の点で、ES細胞から逸脱していることが報告されている。対照的に、いくつかの文脈において、iPS細胞株は、ES細胞と同程度に効率よく分化することができること(Boland et al., 2009; Miura et al., 2009; Zhao et al., 2009)、およびiPS細胞の公表された遺伝子発現シグネチャーが、再現可能ではない可能性があることも同様に報告されている(Stadtfeld et al., 2010)。ヒトES細胞およびiPS細胞株を疾患モデリングまたは移植治療のいずれかのツールとして広く利用できるようになる前に、これらの相違を解決しなければならない。特に、細胞株毎の変動を同定することができるベースラインを提供するため、および多能性細胞のクラス(たとえば、ES細胞対iPS細胞株、特異的変異を有するiPS細胞株対変異を有しないiPS細胞、異なるリプログラミングプロトコールによって誘導されたiPS細胞)間の系統的な比較を行うことができるようになるためには、高品質の多能性細胞株における正常な変動に関する参照値を確立する必要がある。

30

40

50

【0012】

それゆえ当技術分野において、多能性幹細胞のモニタリングおよび検証のための、および正常な変動のスペクトルの中で他の多能性幹細胞と比較して1つの多能性幹細胞株を決定するための新規で、有効かつ効率的な方法が必要であり、ならびにたとえば異常な細胞（たとえば、癌細胞または癌に対する素因を有する細胞）の投与を除外するために治療的投与においてそれを用いる前に、または疾患モデリング、薬物の開発およびスクリーニング、および毒性アッセイにおいて用いる前に、多能性幹細胞集団の安全性プロファイルおよび分化指向性を決定するための有効かつ効率的な方法が必要である。

【発明の概要】

【0013】

本発明は、幹細胞を、その全般的品質および分化能ならびに起こりうる悪性成長の指向性に関して迅速かつ比較的安価にスクリーニングするためのシステムおよび方法に向けられる。本発明のシステムおよび方法は、細胞の迅速な同定および選択を可能にする、いくつかの例において将来的に使用するために適した細胞または特定の有用性を有する特定細胞の自動的選択を行うことができるハイスループットスクリーニングシステムを可能にする。本発明は、人工多能性幹細胞（iPSC）を含む多能性幹細胞を特徴決定する方法に関し、ここで、自然の分化指向性の分析によって、特定細胞株が定方向性分化の体制および枠組みの中でどのような挙動を示すかを高度に予測する。

【0014】

現在、既存の方法は、多能性幹細胞株が所定の定方向性分化の枠組みにおいてどのような挙動を示すかを予測することができない。本明細書において開示される方法およびシステムは、用いるのが面倒で、時間を浪費し、および非常に費用が高く、このため幹細胞の大規模特徴決定において有用とされない奇形腫形成などの現行の広く用いられるシステムと比較して、多能性幹細胞特徴決定に関するはるかに優れたシステムを提供する。たとえば、奇形腫形成またはリプログラミング因子のサイレンシング分析のみを用いると、細胞株が定方向性分化においてどのような挙動を示すかを予測することができず、しかもこれらの方法は最適以下の幹細胞株を同定することができない。本発明の方法およびシステムは、迅速で、安価で、自動化にとって適しているのみならず、現行の標準的な方法（たとえば、奇形腫形成を用いる）と比較して、適切なまたは適切でない幹細胞およびクローンを同定するために有意に感度がよく、ならびに最適な多能性幹細胞を同定するためおよび適切に分化することができない幹細胞株（たとえば、非効率的に分化するまたは性能が不良な多能性幹細胞である幹細胞）を同定するために用いることができる、強力な多能性幹細胞の特徴決定を提供する。したがって、本明細書において開示される方法、システム、およびキットは、従来の方法と比較して細胞の分化能の予測において非常に有用であり、悪性細胞株になる高い素因などの理由のために不適切でありうる幹細胞株を同定することができる、多能性幹細胞株を特徴決定するための迅速で、安価で、しかも定量的なアプローチを提供する。

【0015】

このように、本明細書において開示される方法およびシステムにより、分析される多能性幹細胞株の分化効率を予報することができる。たとえば、方法およびシステムは、特定の系列、たとえば運動ニューロン系列などのニューロン系列に沿った多能性幹細胞株の分化を高度に予測することが証明されている。本明細書において開示される方法およびシステムは幅広い有用性を有し、所定の多能性幹細胞が将来的に任意の所望の系列、たとえば造血系列、内胚葉系列、および膵臓系列などに沿ってどのように良好に分化するかを予測するために用いることができる。

【0016】

本明細書において開示される方法およびシステムは、選択された幹細胞特徴に関するハイスループット様式でのスクリーニングを可能にする既定の遺伝子セットの遺伝子発現に基づく新規システムの開発に基づいている。さらに、新規システムはまた、既定の遺伝子セットのDNAメチル化の決定にも基づいている。遺伝子発現およびDNAメチル化のための遺

10

20

30

40

50

伝子セットは、本明細書において開示される任意の既定の遺伝子セットでありえて、たとえば系列マーカー遺伝子ならびに腫瘍遺伝子および腫瘍抑制遺伝子などを含むがこれらに限定されるわけではない。方法およびシステムによりさらに、得られたデータを自動的にまとめて、適切な細胞またはクローンを選択することが可能である。具体的に、システムは、翻訳後修飾、遺伝子発現データ、DNAメチル化、およびエピジェネティックな修飾および分化マーカーなどの機能的ゲノムデータの決定に依存しており、それによってDNAメチル化、エピジェネティックな修飾、翻訳後修飾、および分化マーカー発現パターンを含む機能的ゲノムデータの正常範囲から逸脱している細胞を除外することができ、および正常範囲内に入る細胞を今後使用するために選択することができる。統計学的分析法を用いてシステムを自動化する。いくつかの態様において、機能的ゲノムデータは、DNAメチル化である。代替の態様において、機能的ゲノムデータは、たとえばヒストンおよび非ヒストンタンパク質（標準的なタンパク質およびその変種を含む）のメチル化、ユビキチン化、リン酸化、グリコシル化、スモイル化、アセチル化、S-ニトロシル化またはニトロシル化、シトルリン化または脱イミノ化、ネジル化（neddylation）、OClcNAc、ADP-リボシル化、ヒドロキシル化、ファテニル化（fattenylation）、ウフミル化（ufmylation）、プレニル化、ミリストイル化、S-パルミトイル化、チロシン硫酸化、ホルミル化、およびカルボキシル化などの翻訳後修飾のいずれか、またはその組み合わせである。いくつかの態様において、機能的ゲノムデータ、たとえばメチル化および/または翻訳後修飾は、遺伝子配列においてならびに低分子非コードRNAおよびクロマチンの非共有的構造修飾（たとえば、凝縮および脱凝縮）に関して決定される。

10

20

【0017】

メチル化の差などのエピジェネティックな修飾および機能的ゲノム修飾は、たとえば悪性の細胞成長に関連する。本発明は、外れ値でありかつ従ってたとえば悪性成長能を有する細胞を、本発明のシステムがスクリーニングして除外することができるような、メチル化パターンの正常範囲を提供する。

【0018】

所望の細胞分化マーカーセットのスクリーニングにより、所望の組織への発達能を有するクローンを選択することができる。たとえば、中胚葉、内胚葉、および外胚葉系列への発達に関するマーカーに関してスクリーニングすることができる。幹細胞が、適切なマーカーセットを発現する多能性細胞に関する既定のパラメータ内に適合しない場合、これを廃棄することができる。

30

【0019】

ヒト多能性幹細胞の長期間増殖能および分化能から、それらが、疾患モデリングおよび移植治療のための様々な細胞型を大量に産生できることが示唆されている。しかし、胚幹（ES）細胞または人工多能性幹（iPS）細胞を治療応用または疾患モデリングにおいて確信を持って用いることができるためには、または薬物スクリーニングもしくは毒性アッセイにおいて用いることができるためには、ヒト多能性細胞株間の変動の程度を理解しなければならない。そのような変動に関する包括的な見方を得るために、本発明者らは、31個のヒトES細胞株およびiPS細胞株を全ゲノムDNAメチル化および転写分析に供すると共に、そのインビトロ分化指向性を定量した。

40

【0020】

多能性幹細胞株間に存在する変動の性質および程度を確固として確立するために、本発明者らは、19個のES細胞株、12個のiPS細胞株、および6個の初代培養線維芽細胞株に対して3つのゲノム規模のアッセイを行った。3つのアッセイは、ゲノム規模のバイサルファイトシーケンシングによるDNAメチル化マッピング（Gu et al., 2010; Meissner et al., 2008）、ハイスループットマイクロアレイを用いる遺伝子発現プロファイリング、および胚様体における500個の遺伝子の転写物計数を使用する定量的分化アッセイを含んだ。

【0021】

本発明者らは、ヒトiPS細胞株およびES細胞株の大きいコホートにおけるDNAメチル化および遺伝子転写プロファイルの全ゲノム分析の使用を実例により示し、多能性幹細胞株間

50

の一般的変動に関して新たに発見された参照値を提供する。本発明者らは、任意の多能性細胞株の分化指向性および有用性を予測するために用いることができる「系列スコアカード」を提供するために、DNAメチル化および遺伝子転写の全ゲノム分析を用いる。本発明者らはまた、ヒトES細胞が変動を示すこと、およびiPS細胞が類似の遺伝子座で変動を示すことも証明する。本発明者らは、ヒトES細胞とヒトiPS細胞とを正確に区別することができる1つの遺伝子座を検出することができなかつた。それゆえ、意図される目的にとって有用である幹細胞をスクリーニングするためには、多数のマーカーのパターンに依存するシステムの開発が重要である。

【0022】

特に、本発明者らは、個々の多能性幹細胞集団、たとえば幹細胞株の挙動を予測するために用いることができる、多様な異なる多能性細胞株におけるDNAメチル化レベルおよび/または遺伝子発現レベルの正常な変動の参照レベルを提供する、ならびに異なるクラスの多能性幹細胞（たとえば、ES細胞対iPS細胞、またはiPS細胞対部分的人工iPS細胞など）を系統的に比較するための土台を提供する、複数の多能性幹細胞集団からのデータを獲得する方法を証明する。

10

【0023】

いくつかの態様において、本発明者らは、多能性幹細胞株がたとえば運動ニューロンへと最適に分化するか否かを予測することによって、およびES細胞とiPS細胞株との定量的比較を行うことによって、本発明の方法およびシステムの有用性を証明する。この比較は、iPS細胞株とES細胞株とを区別するために普遍的に用いることができるDNAメチル化または転写の特異的变化がないことを証明する。したがって、本発明者らは、本明細書において「スコアカード」と呼ばれるデータセットおよびバイオインフォマティクスデータツールを用いることにより、ゲノムアッセイを用いたiPS細胞株および胚細胞株などのヒト多能性細胞株のハイスループットな特徴決定が可能となることを証明する。

20

【0024】

したがって、本発明者らは、異なる多能性細胞集団間の変動性を決定するために、その治療的有用性および安全性プロファイルを予測するために（たとえば、多能性幹細胞を治療的使用のために移植する際に重要である、多能性幹細胞集団が持続的自己再生に対して素因を有するか否か、および高い悪性形質転換能を有するか否かを判定するために）、多能性幹細胞集団を検証するために用いることができる、ならびに同様に多能性幹細胞株が効率よく分化する系列および発達経路への多能性幹細胞集団の分化能を予測することができる、効率的かつ有効な方法、システム、およびキットを発見した。そのため、本明細書において開示される方法、システム、およびキットにより、望ましい特徴を有する多能性幹細胞を選択することができ、たとえば他の多能性幹細胞と類似の特徴を有する多能性幹細胞、または所望の細胞型へと、もしくは特定細胞系列に沿って最適に分化する素因を有する多能性幹細胞に関して正の選択を行うことができ、または代替的に、本方法は、望ましくない特徴を有する多能性幹細胞、たとえば癌細胞へと発達する素因を有する細胞に対して負の選択を行う（例えば同定して廃棄する）ことができる。

30

【0025】

したがって、本発明は、有効かつ効率的な多能性幹細胞および/または前駆細胞のモニタリングおよび検証のための、ならびに特定応用、たとえば新規治療法に適した、または特定系列に沿って分化するのに適した多能性幹細胞を同定するための、方法、システム、およびキットに関し、方法は、異常な細胞の導入を排除するために（たとえば、癌細胞になる、または特定の所望の系列に沿って分化する可能性が低い細胞になると提唱される多能性幹細胞株の投与を回避するために）、治療的に投与する前に多能性幹細胞をモニターおよび/または検証する段階を含む。

40

【0026】

具体的に、本発明のいくつかの局面に従って、出願人は、多能性幹細胞の特徴を同定することができるように、およびどの多能性幹細胞株が幹細胞起源の癌に寄与する可能性があるのかを予測するために、(i) 特定遺伝子、たとえば腫瘍遺伝子、腫瘍抑制遺伝子お

50

よび発達遺伝子のプロモーターメチル化による特定遺伝子のエピジェネティックなサイレンシングの同定、(ii)たとえば発達遺伝子および系列マーカー遺伝子の遺伝子発現の同定、ならびに(iii)異なる系列に沿って分化する分化指向性から選択される少なくとも2つのデータセットに関して多能性幹細胞をモニターすることができることを示す。たとえば、可逆的な遺伝子発現が永続的なサイレンシングに取って代わっている癌特異的プロモーターDNA高度メチル化、永久的自己再生状態への細胞の固定化、およびそれによってその後の悪性の形質転換に細胞を罹りやすくする素因を有する細胞を、選択して除外することができる。

【0027】

1つの態様において、本発明は、一般的に、所望の使用のために多能性幹細胞株の機能性および適切性を予測するための方法および複数のアッセイに関する。いくつかの態様において、所望の使用に関する多能性幹細胞株の機能性および適切性を予測するために、少なくとも1つ、または少なくとも2つまたは少なくとも3つの幹細胞アッセイが単独または組み合わせて用いられる。いくつかの態様において、1つのアッセイは、エピジェネティックなプロファイリングであり、これはたとえば多能性幹細胞株において活性化される遺伝子を決定するための定義された特定の遺伝子セットの遺伝子メチル化の評価である。1つの態様において、第二のアッセイは、多能性幹細胞株が特定系列に沿って分化する指向性を決定するための分化アッセイである。いくつかの態様において、アッセイは、遺伝子発現アッセイ、たとえば、細胞分化関連遺伝子の遺伝子発現パターンを決定するための全ゲノム遺伝子発現アッセイである。

10

20

【0028】

いくつかの態様において、エピジェネティックなプロファイリングを最初に行い、分化に関する遺伝子発現分析を次に行う。いくつかの態様において、分化関連遺伝子の遺伝子発現分析を先に行い、エピジェネティックなマーカープロファイリングを次に行う。いくつかの態様において、アッセイを行う効率を増加させて費用を低減させるために、第一のスクリーニングを用いて正常なパラメータ内であると決定された細胞に限って第二のスクリーニングを行う。

【0029】

もう1つの局面は、本明細書において「スコアカード」と呼ばれ、本発明の3つの複合アッセイからの異なる多数の多能性幹細胞株の結果の平均値データまたはそうでなければ集合データを意味する参照データセットに関する。当業者は、「スコアカード」を構成する参照データを用いて、たとえばコンピュータアルゴリズムまたはソフトウェアを用いて、関心対象の多能性幹細胞株を正常な良好に機能する幹細胞と比較することができる。参照「スコアカード」との比較を用いて、所定の応用に関する多能性幹細胞の有用性、ならびに関心対象の多能性幹細胞株、たとえばES細胞またはiPS細胞株の任意の特異的特徴を有効かつ正確に予測することができる。したがって、本明細書において開示される方法、アッセイおよびスコアカードは、治療的使用、薬物スクリーニング、および毒性アッセイ、ならびに所望の細胞系列への分化についてのその適切性などの下流の応用に関するその適切性を決定するための幹細胞の特定の特徴を同定するために用いることができる。

30

40

【0030】

特定の態様は、以下の段階を含む、好ましい多能性幹細胞を同定、スクリーニング、選択、または濃縮する方法を提供する：多能性幹細胞において(i)高度メチル化DNAプロモーターを有する遺伝子の有無を同定する段階、または正常な変動と比較して特異的メチル化標的遺伝子のメチル化状態に統計学的有意差を有する(増加または減少する)遺伝子を同定する段階；および(ii)特定の標的遺伝子、たとえば発達遺伝子および/または系列マーカー遺伝子の遺伝子発現レベルを同定する段階；および(iii)望ましい特徴を有する多能性幹細胞株を同定するために異なる系列に沿って分化する分化指向性を同定する段階。

【0031】

本発明のもう1つの局面は、多能性幹細胞株が悪性の形質転換を受ける可能性があるか

50

否か、および所望のまたは特定の発達経路に沿ったおよび特定細胞系列への分化能を有するか否かを予測するために、以下の少なくとも1つ、すなわち標的メチル化遺伝子のDNAメチル化状態、標的遺伝子の発現レベル、ならびに外胚葉、中胚葉、および内胚葉への分化指向性、の少なくとも1つをスクリーニングまたはモニターする段階を含む、幹細胞、たとえば多能性、多分化能、単能性、または体性幹細胞、または最終的に分化した細胞集団、たとえばこれらに限定されるわけではないが、前駆細胞、胚幹 (ES) 細胞、体性幹細胞、癌幹細胞、前駆細胞、人工多能性幹 (iPS)、部分的な人工多能性 (piPS) 細胞、リプログラム細胞、直接リプログラム細胞等を検証および/またはモニターする方法を提供する。

【0032】

本発明の1つの態様は、以下の段階を含む、特定の適応に関して多能性幹細胞株または前駆細胞集団を検証および選択する方法を提供する：(i) 本明細書において開示される定量的分化アッセイを用いて多能性幹細胞集団の分化能を測定する段階、および(ii) 所望の細胞系列に沿った、または所望の細胞型への分化の効率が中程度であるかまたは高い多能性幹細胞集団を選択する段階、(iii) 多能性幹細胞集団におけるDNAメチル化標的遺伝子セットのDNAメチル化を測定して、該DNAメチル化データと、同じ標的遺伝子の参照DNAメチル化レベルとの比較を行う段階；ならびに(iv) 参照DNAメチル化レベルと比較して標的遺伝子のメチル化に統計学的に有意な量で差がない多能性幹細胞株を選択する段階、ならびに任意で段階(v) および段階(vi) を行う段階で、段階(v) は、多能性幹細胞株における標的遺伝子の発現レベルを測定して、該遺伝子発現レベルデータと、同じ標的遺伝子の参照遺伝子発現レベルとの比較を行う段階を含み；かつ、段階(vi) は、参照遺伝子発現レベルと比較して標的遺伝子の遺伝子発現レベルに統計学的に有意な量で差がない多能性幹細胞株を選択する段階を含む。いくつかの態様において、多能性幹細胞を最初に、所望の細胞系列に沿ったまたは所望の細胞型への分化に基づいて選択して、次に、多能性幹細胞における遺伝子のDNAメチル化または発現レベルのいずれかに基づいて選択して、望ましくない特徴を有する多能性幹細胞、たとえば腫瘍遺伝子および/または腫瘍抑制遺伝子の異常な(増加または減少した)発現を有する多能性幹細胞の負の選択を行う(たとえば、廃棄する)。例に過ぎないが、腫瘍遺伝子の低いメチル化または腫瘍抑制遺伝子の高い発現を有する細胞を廃棄することができ、および/または腫瘍抑制遺伝子の高いメチル化または腫瘍抑制遺伝子の高い遺伝子発現を有する細胞を廃棄することができる。代替の態様において、多能性幹細胞が分化する予定である所望の細胞において通常発現される発達遺伝子および/または系列マーカー遺伝子の高いメチル化を有する細胞を廃棄することができる。

【0033】

本発明の1つの局面は、(i) 少なくとも5個の多能性幹細胞集団からの複数のDNAメチル化標的遺伝子に関するDNAメチル化レベルを含む第一のデータセット；(ii) 少なくとも5個の多能性幹細胞集団からの複数の標的遺伝子に関する遺伝子発現レベルを含む第二のデータセット；ならびに(iii) 少なくとも5個の多能性幹細胞集団からの外胚葉、中胚葉、および内胚葉系列への分化に関する分化指向性レベルを含む第三のデータセット、を含む、多能性幹細胞の性能パラメータのスコアカードに関する。いくつかの態様において、複数の参照DNAメチル化遺伝子は、少なくとも約1000個の参照DNAメチル化遺伝子であるか、または少なくとも約2000個の参照DNAメチル化遺伝子であり、またはいくつかの態様において、全ゲノムのDNAメチル化状態である。いくつかの態様において、参照DNAメチル化遺伝子は、癌遺伝子、腫瘍遺伝子、および腫瘍抑制遺伝子、系列マーカー遺伝子、および発達遺伝子を含む群から選択される任意の遺伝子である。

【0034】

いくつかの態様において、DNAメチル化標的遺伝子は、BMP4、CAT、CD14、CXCL5、DAZL、DNMT3B、GATA6、GAPDH、LEFTY2、MEG3、PAX6、S100A6、SOX2、SNAI1、TFからなる群より選択される任意の遺伝子または任意の遺伝子の組み合わせである。

【0035】

10

20

30

40

50

いくつかの態様において、スコアカードの第一および第二のデータセットは、コンピュータデバイス上に存在するデータベースであるデータ保存デバイスなどのデータ保存デバイスに接続される。

【0036】

いくつかの態様において、少なくとも15個の多能性幹細胞株を用いて、スコアカードのための第一、または第二、または第三のデータセットを作成する。いくつかの態様において、第一、第二、または第三のデータセットは、HUES64、HUES3、HUES8、HUES53、HUES28、HUES49、HUES9、HUES48、HUES45、HUES1、HUES44、HUES6、H1、HUES62、HUES 65、H7、HUES13、HUES63、HUES66の群から選択される多能性幹細胞株の少なくとも5個またはそれより多く、または少なくとも6個、または少なくとも7個、または少なくとも8個、または少なくとも9個、または少なくとも10個、または少なくとも11個、または少なくとも12個、または少なくとも13個、または少なくとも14個、または少なくとも15個、または少なくとも16個、または少なくとも17個、または少なくとも18個、または19個全てから得られる。

10

【0037】

いくつかの態様において、スコアカードのデータセットを生成するために用いられる多能性幹細胞集団は、ヒト多能性幹細胞集団または人工多能性幹(iPS)細胞集団、または胚幹細胞集団、または成体幹細胞集団、または自己幹細胞集団、または胚幹(ES)細胞集団などの哺乳動物多能性幹細胞集団である。

【0038】

いくつかの態様において、本明細書において開示されるスコアカードを、関心対象の多能性幹細胞集団のDNAメチル化レベル、遺伝子発現レベル、および分化指向性レベルと比較することができ、特定系列に沿った最適な分化および/または望ましくない特徴を有する指向性を予測することによって、たとえば癌細胞へと発達する素因を有する多能性幹細胞集団を予測することによって、多能性幹細胞集団の挙動を検証および/または予測するために用いることができる。このように、いくつかの態様において、スコアカードを、望ましい特徴(たとえば、特定系列に沿った高い分化能)を有する関心対象の多能性幹細胞集団を選択する、たとえば正の選択を行う方法において、および/または望ましくない特徴を有する細胞、たとえば癌細胞へと発達する素因を有する細胞の負の選択を行う方法において用いることができる。

20

30

【0039】

本発明のもう1つの局面は、(i)複数の多能性幹細胞集団における標的遺伝子セットにおけるDNAメチル化を測定する段階；(ii)複数の多能性幹細胞株における第二の標的遺伝子セットにおける遺伝子発現を測定する段階、および(iii)複数の多能性幹細胞株の分化能を測定する段階を含む、多能性幹細胞スコアカードを作成する方法に関する。いくつかの態様において、多能性幹細胞スコアカードを作成する方法を用いて、複数の多能性幹細胞株、たとえば、少なくとも5個、または少なくとも6個、または少なくとも7個、または少なくとも8個、または少なくとも9個、または少なくとも10個、または少なくとも15個、または少なくとも20個、または少なくとも30個、または少なくとも40個、または40個より多くの異なる多能性幹細胞集団からのDNAメチル化の正常な変動、DNA遺伝子発現の正常な変動、および正常な分化指向性の値を含むスコアカードを作成することができる。

40

【0040】

本発明のもう1つの局面は、(i)多能性幹細胞集団におけるDNAメチル化標的遺伝子セットのDNAメチル化を測定して、該DNAメチル化データと、同じ標的遺伝子の参照DNAメチル化レベルとの比較を行う段階；(ii)多能性幹細胞集団の分化能を測定して、分化能データを参照分化能データと比較する段階；ならびに(iii)参照DNAメチル化レベルと比較して標的遺伝子のメチル化に統計学的に有意な量で差がなく、かつ参照分化能と比較して、中胚葉、外胚葉、および内胚葉系列に沿って分化する指向性に統計学的に有意な量で差がない多能性幹細胞株を選択する段階を含む、多能性幹細胞集団を選択する方法に関する。

50

【0041】

いくつかの態様において、多能性幹細胞集団を選択する方法はさらに、(i)多能性幹細胞株における第二の標的遺伝子セットの遺伝子発現レベルを測定して、該遺伝子発現レベルデータと、同じ標的遺伝子の参照遺伝子発現レベルとの比較を行う段階；および(ii)参照遺伝子発現レベルと比較して標的遺伝子の遺伝子発現レベルに統計学的に有意な量で差がない多能性幹細胞株を選択する段階を含む。

【0042】

本発明の1つの局面は、(a)(i)多能性幹細胞株におけるDNAメチル化標的遺伝子セットのDNAメチル化データを受信して、該DNAメチル化データと、同じ標的遺伝子の参照DNAメチル化レベルとの比較を行う段階；(ii)多能性幹細胞株の分化能データを受信して、該分化能データを参照分化能データと比較する段階；(iii)参照DNAメチル化パラメータと比較したDNAメチル化データの比較、および参照分化データと比較した分化指向性の比較に基づいて品質保証スコアカードを作成する段階を含む少なくとも1つのプログラムを含む、少なくとも1つのメモリ；ならびに(b)該プログラムを作動させるためのプロセッサを含む、多能性幹細胞の品質保証スコアカードを作成するためのコンピュータシステムに関する。

【0043】

いくつかの態様において、システムのプログラムは、(i)多能性幹細胞株における第二の標的遺伝子セットの遺伝子発現データを受信して、該発現データを同じ第二の標的遺伝子セットの参照遺伝子発現レベルと比較する段階；(ii)参照DNAメチル化パラメータと比較したDNAメチル化データの比較、および参照分化データと比較した分化指向性の比較、および参照遺伝子発現レベルと比較した遺伝子発現データの比較に基づいて、品質保証スコアカードを作成する段階をさらに含む。

【0044】

本発明の全ての局面のいくつかの態様において、DNAメチル化標的遺伝子は、多様なメチル化を有し、いくつかの態様において、DNAメチル化標的遺伝子は、癌遺伝子、腫瘍遺伝子、腫瘍抑制遺伝子、発達遺伝子、系列マーカー遺伝子のいずれかおよびその全ての組み合わせから選択される。いくつかの態様において、DNAメチル化標的遺伝子は、BMP4、CAT、CD14、CXCL5、DAZL、DNMT3B、GATA6、GAPDH、LEFTY2、MEG3、PAX6、S100A6、SOX2、SNAI1、TFからなる群より選択される。

【0045】

本発明の全ての局面のいくつかの態様において、参照DNAメチル化レベルは、参照多能性幹細胞集団におけるDNAメチル化標的遺伝子のメチル化の正常な変動のレベルである。いくつかの態様において、参照DNAメチル化レベル(たとえば、DNAメチル化標的遺伝子のメチル化の正常な変動レベル)は、複数の異なる多能性幹細胞集団、たとえば少なくとも2個、または少なくとも3個、または少なくとも4個、または少なくとも5個、または少なくとも6個、または少なくとも10個の異なる多能性幹細胞集団からの標的DNAメチル化遺伝子に関するメチル化レベルの変動から作成される。いくつかの態様において、関心対象の多能性幹細胞のDNAメチル化標的遺伝子のメチル化レベルが、参照DNAメチル化レベルから外れる場合、例えば参照DNAメチル化レベルと比較して統計学的に有意な量で増加または減少したメチル化レベルである場合、標的DNAメチル化遺伝子のエピジェネティックなサイレンシングの増加または減少をそれぞれ示すことができる。

【0046】

いくつかの態様において、DNAメチル化標的遺伝子が腫瘍遺伝子である場合、その腫瘍遺伝子に関する参照DNAメチル化レベルと比較してメチル化が統計学的に有意なレベルで減少していれば、腫瘍遺伝子のエピジェネティックなサイレンシングの減少および抑制の欠如を示すことができ、多能性幹細胞が癌細胞への悪性の形質転換に関する素因を有することを示すことができる。または、DNAメチル化標的遺伝子が腫瘍抑制遺伝子であるいくつかの態様において、その腫瘍抑制遺伝子に関する参照DNAメチル化レベルと比較してメチル化が統計学的に有意なレベルで増加していれば、腫瘍抑制遺伝子の発現のエピジェネ

ティックなサイレンシングおよび抑制の増加を示すことができ、多能性幹細胞が癌細胞への悪性の形質転換に関する素因を有することを示すことができる。

【0047】

いくつかの態様において、DNAメチル化標的遺伝子が発達遺伝子または系列マーカー遺伝子である場合、その発達遺伝子または系列マーカー遺伝子に関する参照DNAメチル化レベルと比較してメチル化が統計学的に有意なレベルで増加していれば、発達遺伝子または系列マーカー遺伝子のエピジェネティックなサイレンシングおよび発現の抑制の増加を示すことができ、多能性幹細胞が、発達遺伝子が正常に発現される発達経路に沿って分化する効率が低いか、または系列マーカーを発現する細胞型に分化する効率が低いと予測することができる。逆に、DNAメチル化標的遺伝子が発達遺伝子または系列マーカー遺伝子である態様において、その発達遺伝子または系列マーカー遺伝子に関する参照DNAメチル化レベルと比較してメチル化が統計学的に有意なレベルで減少していれば、発達遺伝子または系列マーカー遺伝子のエピジェネティックなサイレンシングの減少および発現の抑制の減少を示すことができ、関心対象の多能性幹細胞が、発達遺伝子が正常に発現されている発達経路に沿って分化する効率が低いもしくは最適であるか、または系列マーカーを発現する細胞型に分化する効率が低いことを予測するために用いることができる。

10

【0048】

いくつかの態様において、システムはさらに、多能性幹細胞集団の品質に基づいて幹細胞スコアカードレポートを作成するためのレポート作成モジュールを含む。いくつかの態様において、システムはメモリを含み、メモリはさらにデータベースを含む。いくつかの態様において、データベースは、DNAメチル化遺伝子セットを階層的様式で配置して、たとえばデータベースは、関心対象の多能性幹細胞の異なる系列への分化指向性を階層的様式で配置することができる。いくつかの態様において、データベースは遺伝子発現データを階層的様式で配置することができる。いくつかの態様において、システムのメモリは、ネットワーク、たとえば広域ネットワーク、またはワールドワイドネットワークを介して第一のコンピュータに接続される。

20

【0049】

いくつかの態様において、スコアカードレポートは、多能性幹細胞集団の適切な用途または応用の指標を提供し、または代替の態様において、多能性幹細胞株が適していない用途または応用の指標を提供する。

30

【0050】

いくつかの態様において、参照DNAメチル化レベルは、複数の多能性幹細胞における前記DNAメチル化標的遺伝子のメチル化の正常な変動の範囲である。いくつかの態様において、参照遺伝子発現レベルは、複数の多能性幹細胞におけるその標的遺伝子に関する遺伝子発現レベルの正常な変動の範囲内である。いくつかの態様において、DNAメチル化標的遺伝子は、遺伝子発現標的遺伝子と同じであり、いくつかの態様において、DNAメチル化標的遺伝子は、遺伝子発現標的遺伝子の少なくとも1つまたは複数を含み、いくつかの態様において、遺伝子発現標的遺伝子は、DNAメチル化標的遺伝子の少なくとも1つまたは複数を含む。

40

【0051】

本発明のもう1つの局面は、(i)多能性幹細胞株におけるDNAメチル化標的遺伝子セットのDNAメチル化データを受信して、該DNAメチル化データと、同じ標的遺伝子の参照DNAメチル化レベルとの比較を行うこと；(ii)多能性幹細胞株の分化能データを受信して、該分化能データを参照分化能データと比較すること；(iii)参照DNAメチル化パラメータと比較したDNAメチル化データの比較、および参照分化データと比較した分化指向性の比較に基づいて品質保証スコアカードを作成することを含む、多能性幹細胞株の品質保証スコアカードの作成に関する命令を含む、コンピュータ読み取り可能な媒体に関する。いくつかの態様において、コンピュータ読み取り可能な媒体は、(i)多能性幹細胞株における第二の標的遺伝子セットの遺伝子発現データを受信して、該発現データを同じ第二の標的遺伝子セットの参照遺伝子発現レベルと比較すること；(ii)参照DNAメチル化パラメ

50

ータと比較したDNAメチル化データの比較、および参照分化データと比較した分化指向性の比較、および参照遺伝子発現レベルと比較した遺伝子発現データの比較に基づいて品質保証スコアカードを作成することに関する命令をさらに含む。

【0052】

本発明のもう1つの局面は、(i) DNAメチル化アッセイ；(ii) 遺伝子発現アッセイ；および(iii) 分化アッセイの少なくとも2つを含む、多能性細胞の複数の特性の特徴決定を行うためのアッセイに関する。いくつかの態様において、DNAメチル化アッセイは、パイサルファイトシーケンシングアッセイまたは全ゲノムシーケンシングアッセイ、たとえば簡約表示(reduced representation)パイサルファイトシーケンシング(RRBS)である。いくつかの態様において、遺伝子発現アッセイはマイクロアレイアッセイである。

10

【0053】

いくつかの態様において、分化アッセイは、定量的分化アッセイ、たとえば多能性細胞の以下の系列、すなわち中胚葉、内胚葉、および外胚葉、ニューロン、造血系列の少なくとも1つへの分化能を評価することができる分化アッセイである。いくつかの態様において、多能性細胞の以下の系列、すなわち中胚葉、内胚葉、および外胚葉の少なくとも1つへの分化能は、中胚葉、内胚葉、および外胚葉系列の少なくとも1つのマーカに対する抗体を用いる免疫染色またはFACソーティングによって決定される。いくつかの態様において、多能性細胞の以下の系列、すなわち中胚葉、内胚葉、および外胚葉の少なくとも1つへの分化能は、EBにおいて少なくとも約0日目後に多能性幹細胞の免疫染色によって決定される。いくつかの態様において、多能性細胞の以下の系列、すなわち中胚葉、内胚葉、および外胚葉の少なくとも1つへの分化能は、EBにおいて0日、EBにおいて0~32日、たとえば、少なくとも1日、または少なくとも2日、または少なくとも約3日、または少なくとも約4日、または少なくとも約5日、または少なくとも約6日、または少なくとも約7日、またはEBにおいて約7日超、たとえばEBにおいて5から7日、またはEBにおいて約7~10日、またはEBにおいて約10~14日、またはEBにおいて約14~21日、またはEBにおいて約21~32日、またはEBにおいて32日超のいずれかの日に決定される。いくつかの態様において、多能性幹細胞の分化能は、5~10日目のEB、たとえばEBにおいて約7日目に決定される。中胚葉、内胚葉、および外胚葉系列に関する系列マーカの例は、当業者に周知であり、中胚葉系列マーカ-VEGF受容体II(KDR)、またはアクチン-2平滑筋(ACTA2)、外胚葉系列マーカ-ネスチンまたはチューブリン-3および内胚葉系列マーカ-フェトプロテイン(AFP)が挙げられるがこれらに限定されるわけではない。いくつかの態様において、当業者は、分化に関して結果が得られるまでの時間を増加させるために、ならびに多能性幹細胞が中胚葉、内胚葉、および外胚葉系列に沿って分化する指向性を決定する際のシグナル対ノイズ比および変動性を低減させるために、化学刺激または他の刺激、たとえば増殖因子等を用いることができる。

20

30

【0054】

いくつかの態様において、アッセイは、複数の異なる多能性幹細胞をアッセイするためのハイスループットアッセイであり、それによってたとえば同じまたは異なる対象、たとえば哺乳動物対象またはヒト対象から得られた体細胞のリプログラミングに由来する異なる複数の人工多能性幹細胞を評価することができる。

40

【0055】

いくつかの態様において、本明細書において開示されるアッセイは、少なくとも1つまたは複数の多能性幹細胞集団から本明細書において開示されるスコアカードを作成するために用いることができる。

【0056】

本明細書において開示される全ての局面のいくつかの態様において、参照DNAメチル化レベルは、多能性幹細胞集団における前記DNAメチル化標的遺伝子に関するメチル化の正常な変動の範囲である。

【0057】

50

本明細書において開示される全ての局面のいくつかの態様において、参照遺伝子発現レベルは、多能性幹細胞集団におけるその標的遺伝子に関する遺伝子発現レベルの正常な変動の範囲である。

【0058】

本発明のもう1つの局面は、以下を含む、多能性幹細胞株の品質を決定するためのキットに関する：(i)複数のDNAメチル化遺伝子のメチル化状態を測定するための試薬；(ii)複数の遺伝子の遺伝子発現レベルを測定するための試薬；ならびに(iii)多能性幹細胞の外胚葉、中胚葉、および内胚葉系列への分化指向性を測定するための試薬。いくつかの態様において、キットは、本明細書において開示されるスコアカードをさらに含む。いくつかの態様において、キットは、使用説明書をさらに含む。

10

【0059】

本発明者らは本明細書において、患者の試料から、完全にリプログラムされたiPS細胞まで、すなわち疾患モデリングのために妥当な規模で用いることができる選択され取り扱い可能な多能性iPS細胞株セットまで研究者が進行するよう案内することができる明確な経路を提供している。特に、多能性幹細胞株に存在する変動の性質および程度を確固として確立するために、3つのゲノム規模のアッセイを19個のES細胞株、12個のiPS細胞株、および6個の初代培養線維芽細胞株に適用した。これらのアッセイは、ゲノム規模のバイサルファイトシーケンシングによるDNAメチル化マッピング(Gu et al., 2010; Meissner et al., 2008)、ハイスループットマイクロアレイを用いる遺伝子発現プロファイリング、および胚様体における500個の遺伝子の転写物計数を使用する定量的分化アッセイを含んだ。

20

【0060】

全体として、本発明者らは、本明細書において開示されるシステムおよび方法を用いて、ヒト多能性細胞株におけるDNAメチル化レベルおよび遺伝子発現レベルの正常な変動に関する参照レベルを含む少なくとも1つのスコアカードを提供するための3つのアッセイの少なくとも2つから、データを作成する。大部分の遺伝子に関して、本発明者らは、DNAメチル化および転写レベルの変動をほとんど観察できなかった。しかし、本発明者らは、個々の多能性細胞株間でかなり多様なDNAメチル化または転写を示す注目すべき遺伝子クラスが存在することを発見した。驚くべきことに本発明者らは、この変動を理解することが重要であり、かつこれによって所定の多能性幹細胞株の挙動の予測が可能になることを証明する。さらに、定量的分化アッセイを用いて、本発明者らは、多能性幹細胞の特定系列への最適な分化の予測が正確であること、および同様に、各多能性細胞株が、所定の発達経路への分化に関して特異的で再現可能な指向性をそれ自体が有することを証明した。重要なことに、本発明者らはまた、分化指向性の知識を用いて、Boultingおよび共同研究者によって独立して行われた定方向性分化実験において各細胞株が行う効率を正確に予測することができることを証明する。要約すると、本発明者らは、これらの3つのアッセイ(DNAメチル化、遺伝子発現プロファイリング、および定量的分化アッセイ)の結果をまとめて、所定の応用に関する特定のES細胞またはiPS細胞株の有用性を予測するために誰でも用いることができる、「系列スコアカード」を作成した。

30

【0061】

本明細書において開示される「要約スコアカード」は、ヒト多能性細胞株における正常な変動に関する参照値を提供する「偏差スコアカード」および「系列スコアカード」を含む。分析される遺伝子の大部分に関する偏差スコアカードにおいて、本発明者らは、DNAメチル化および転写レベルに関してほとんど変動を観察しなかった。しかし、本発明者らは、個々の細胞株において高度に変動するDNAメチル化または転写のいずれかを示す注目すべき遺伝子サブセットまたはクラスを発見した。本明細書において、本発明者らは、この変動を理解することが重要であり、これを所定の多能性幹細胞株の挙動を予測するために使用できることを証明する。

40

【0062】

たとえば、本発明の局面は、多能性幹細胞株を特徴決定するための2つのスコアカード

50

に関する方法および作成に関し、「偏差スコアカード」または「多能性スコアカード」と呼ぶことができる第一のスコアカードは、関心対象の多能性幹細胞株が、既に確立されたまたは対照の多能性幹細胞株にどの程度匹敵するかという情報を提供するために有用であり、1つの参照多能性幹細胞株および/または複数の参照多能性幹細胞株と比較して、DNAメチル化または遺伝子発現に関して外れている遺伝子の数または%を同定するために用いることができる。そのようなスコアカードは、関心対象の幹細胞株の多能性を同定するために有用であると共に、関心対象の幹細胞株を後の時点で異常増殖および癌形成に罹りやすくする可能性がある癌遺伝子の異型の遺伝子発現またはDNAメチル化を関心対象の幹細胞株が有するか否かを同定するために、有用である。本明細書において「系列スコアカード」と呼ばれる第二のスコアカードは、関心対象の多能性幹細胞の分化能の定量として有用であり、関心対象の多能性幹細胞株が、既に確立されたまたは対照の多能性幹細胞株と比較して、どれほど効率的に関心対象の特定の系列へと分化するかに関する情報を提供する。

10

【0063】

要約すると、単独でまたは任意の組み合わせで用いられ、3つ全てのアッセイの複合結果を含む本明細書記載の3つのアッセイを用いて、当業者は、多能性幹細胞を検証するため、および特定の多能性幹細胞、たとえば所定の応用に関するES細胞またはiPS細胞株の有用性を予測するために用いることができる「要約スコアカード」（たとえば、偏差スコアカードおよび/または系列スコアカードを含む）を作成することができる。

20

【0064】

本明細書において開示されるアッセイは、たとえばハイスループット施設において100個レベルおよび1000個レベルの幹細胞株を特徴決定することが望ましい場合のように、一度に何百または何千ものES細胞株および/またはiPS細胞株の特徴決定を可能にするであろう偏差スコアカードおよび系列スコアカードを作成するため、たとえば治療的に用いるための薬物スクリーニングにおける有用性に関して幹細胞株を決定するために、たとえば多重qPCRおよびハイスループット試料処理を用いてハイスループットとなるように構成することができる。本明細書において開示される方法およびスコアカードを用いることにより、従来の奇形腫による特徴決定方法を用いると費用が高く、実際的ではない、多数の幹細胞株の迅速かつ安価な特徴決定が可能となる。または、本明細書において開示されるアッセイ、方法、システム、およびスコアカードは、研究を加速させるために個別に用いることができ、および関心対象の研究課題に取り組むために研究において用いることができ、たとえば本明細書において開示されるアッセイ、方法、システム、およびスコアカードを用いて、関心対象の研究課題に取り組むためのさらなる分析にとって最も適切な多能性幹細胞株を同定するために、多能性幹細胞株を特徴決定することができる。

30

【図面の簡単な説明】

【0065】

本特許または出願ファイルは、カラーで描かれる少なくとも1つの図面を含む。カラーの図面と共に、本特許または特許出願の刊行物のコピーは、要求に応じて、必要な料金を支払うことにより、担当局によって提供されるであろう。

【図1A】ヒトES細胞株の参照マップが、多能性細胞株における正常な変動の範囲帯にわたることを示す図である。この図は、19個のヒトES細胞株と6個の初代培養線維芽細胞株のジョイント階層クラスタリングを示す。DNAメチル化レベルを、各Ensembl注釈付き転写開始部位周囲の-5 kbから+1 kbの範囲のプロモーター領域において平均した。遺伝子発現レベルを、マイクロアレイにおける全ての関連するプローブについて平均することにより、各Ensembl遺伝子に関して算出した。階層クラスタリングの前に、2つのデータセットを、平均ゼロおよび分散1へと個々に正規化した。DNAメチル化および遺伝子発現の両方に関してユークリッド距離行列を算出して、2つの距離行列を平均した。平均連鎖を用いて階層クラスタリングを行い、ヒートマップは、250個の遺伝子の代表的選択を示す。明るい色は、より高レベルのDNAメチル化（赤色）または遺伝子発現（緑色）を示し、暗い色は、より低いレベルを示す、DNAメチル化と遺伝子発現の複合データを表3に示す。全ての

40

50

遺伝子およびプロモーター領域をそのエピジェネティックな転写の変動レベルにより並べたリストを、表4および5に示す。

【図1B】ヒトES細胞株の参照マップが、多能性細胞株における正常な変動の範囲帯にわたることを示す図である。この図は、選択された4つの遺伝子でのDNAメチル化および遺伝子発現測定の高解像度視野を示す。Ensembl注釈付き転写開始部位周囲の-5 kbから+1 kbの範囲のプロモーター領域に関するDNAメチル化パターンを示す。左側の各々の四角は、プロモーター領域内に位置する1つのCpGジヌクレオチドを表す（暗い赤色：高いメチル化、明るい赤色：部分的メチル化、白色：完全なメチル化）。右側の1つの四角は、各遺伝子の標準化された発現レベルを可視化する（暗い緑色：高い発現、明るい赤色：中程度の発現、白色：発現なし）。測定を4つの代表的なES細胞株および1つの代表的な線維芽細胞株に関して示す。DNAメチル化パターンは縮尺どおりに書かれていないことに注意されたい。全ての高解像度データは、<http://scorecard.computational-epigenetics.org/>の補助ウェブサイトによりゲノムブラウザトラックとして利用可能である。

【図1C】ヒトES細胞株の参照マップが、多能性細胞株における正常な変動の範囲帯にわたることを示す図である。この図は、19個の低継代ヒトES細胞株における遺伝子特異的DNAメチル化（左）および遺伝子発現（右）の箱ひげ図を示し、エピジェネティックな転写の参照範囲帯の概念を描写する。多くのES細胞株の複合データは、ヒト多能性幹細胞株において観察された変動を定量化して参照値を提供し、それに対して1つの細胞株を比較することができる。範囲帯は、全体で31,929個のプロモーター領域（DNAメチル化）および15,079個の遺伝子（発現）にわたり、このダイアグラムは、広範囲の異なる変動レベルをカバーする選択された15個の遺伝子を中心に説明する。箱ひげ図の箱は、中央値を黒色の線で印す中心四分位数に対応し、ひげ部分は、箱からの四分位数間範囲の1.5倍以下の極値データポイントまで伸びる。完全なES細胞の参照範囲帯は、その全体が参照により本明細書に組み入れられる、ウェブサイト<http://scorecard.computational-epigenetics.org/>から入手可能である（データは示していない）。

【図2A】エピジェネティックな転写の変動が特定遺伝子を標的として、細胞の分化に影響を及ぼすことを示す図である。この図は、19個のES細胞株において平均したES細胞参照からの細胞株特異的偏差の分布を示し、エピジェネティックな転写の変動に対する感受性の、遺伝子特異的測定を提供する。ヒストグラムは、平均偏差レベル（x-軸）の各区間内に入る遺伝子数（y-軸）を示す。各ヒストグラム内の選択された遺伝子の位置を上部に強調する。DNAメチル化ヒストグラム（左）は、極めて偏っていることに注意されたい：よりわかりやすく表示するために、ダイアグラムの右半分に関してx-軸は5倍圧縮されており、このためヒストグラムの中心で偽のピークが生じている。遺伝子発現ヒストグラム（右）では、ゼロで強いピークが存在するが、これは全てのES細胞株においてゼロ発現（およびこのようにゼロ偏差）を示す遺伝子が多数存在するためである。

【図2B】エピジェネティックな転写の変動が特定遺伝子を標的として、細胞の分化に影響を及ぼすことを示す図である。この図は、DNAメチル化（左上）または遺伝子発現（左下）に関して最も変動性の高い1,000個の遺伝子の染色体分布を示し、エピジェネティックに変動するが、転写的には変動しない遺伝子がヒト性染色体XおよびYに主に位置することを示している。変動性は、19個のES細胞株において平均したES細胞参照からの細胞株特異的偏差として測定された。ダイアグラムはまた、十分なDNAメチル化（右上）または遺伝子発現データ（右下）を有する全ての遺伝子の染色体分布を示し、最も変動性の高い遺伝子のゲノム位置の差が、偏ったシーケンシングカバー率の副作用ではないことを強調している。

【図2C】エピジェネティックな転写の変動が特定遺伝子を標的として、細胞の分化に影響を及ぼすことを示す図である。この図は、DNAメチル化（上）および遺伝子発現（下）に関して最も変動性の高い1,000個の遺伝子の比較を示す。性染色体の偏りがこの分析に影響を及ぼすのを防止するために、全てのX-連鎖およびY-連鎖遺伝子を除外した。重なり部分の有意性は、フィッシャー正確確率検定を用いて確立した。

【図2D】エピジェネティックな転写の変動が特定遺伝子を標的として、細胞の分化に影響

10

20

30

40

50

響を及ぼすことを示す図である。この図は、DNAメチル化（上）および遺伝子発現（下）に関して最も変動性の高い遺伝子1,000個の構造的および機能的特徴を示す。機能的アノテーションクラスタリングを、DAVIDソフトウェアによって分析して（Huang et al., 2007）、プロモーターの特徴を、EpiGRAPHウェブサービス（Bock et al., 2009）によって分析した。このパネルは、結果の要約を提供する：完全な結果を表3および5に示す。性染色体の偏りがこの分析に影響を及ぼすのを防止するために、全てのX-連鎖およびY-連鎖遺伝子を除外した。

【図2E】エピジェネティックな転写の変動が特定遺伝子を標的として、細胞の分化に影響を及ぼすことを示す図である。この図は、非定方向性EB分化の際の2つのES細胞株におけるDNAメチル化（左、中央）および遺伝子発現（右）の差の散布図を示し、これは、ES細胞状態（左）のDNAメチル化の差が16日目のEB（中央）において維持され、EBにおける遺伝子発現（右）と負の相関を有することを示す。多能性状態の2つのES細胞株において差次的にメチル化される（閾値：20パーセントポイント）（左）それらの遺伝子を、3つ全てのダイアグラムにおいて強調する（オレンジ色：HUES6における高度メチル化、青色：HUES8における高度メチル化）。マクロファージ/顆粒球特異的マーカー遺伝子CD14の位置を矢印で示し、16日目のEBにおけるその細胞株特異的な差次的メチル化を維持し、かつそのプロモーターでDNAメチル化が存在しない場合に限ってアップレギュレートされる遺伝子の例を提供する。

【図2F】エピジェネティックな転写の変動が特定遺伝子を標的として、細胞の分化に影響を及ぼすことを示す図である。この図は、定義された造血分化プロトコールに供した2つのES細胞株（HUES6およびHUES8）のエピジェネティックな転写の差を示す。DNAメチル化レベルを、分化プロトコールの0日および18日目でのクロールバイサルファイトシーケンシングによって測定した。白色のピーズは、非メチル化CpGに対応し、黒色のピーズは、メチル化CpGに対応する。列は、個々のクローンに対応し、行は、CD14のプロモーター領域における特異的CpGに対応する。同様に、CD14およびさらに2つのマクロファージマーカー遺伝子（CD33およびCD64）の遺伝子発現も、分化プロトコールの0日目および18日目に、独立した2つの実験においてqPCRによって測定した（3つのテクニカルレプリケートを示す）。

【図2G】エピジェネティックな転写の変動が特定遺伝子を標的として、細胞の分化に影響を及ぼすことを示す図である。この図は、造血（TFCP2、LY6H）および神経プロセス（COMT、CAT）において公知の役割を有する4つの遺伝子における細胞株特異的DNAメチル化および遺伝子発現レベルを示す。各データポイントは、ES細胞株（「ES」）または対応する16日胚様体（「EB」）のDNAメチル化（x-軸）および遺伝子発現（y-軸）の複合レベルを指す。

【図3A】iPS細胞株においてより高い変動性の方向に向かう傾向を検出するが、iPS特異的欠損は検出されないゲノムマップを示す図である。この図は、11個のiPS細胞株（「hiPSx」）、19個のES細胞株（「HUESx」または「Hx」）および6個の初代培養線維芽細胞株（「hFibx」）のジョイント階層クラスタリングを示し、全てのiPS細胞株がES細胞株とクラスタを形成して、多能性細胞株におけるサブクラスタへの明確な分離が存在しないことを示す。クラスタリングは、図1Aと同じように行った。ヒートマップおよびMEG3発現状態を伴う拡大版は図9Bから入手可能である。

【図3B】iPS細胞株においてより高い変動性の方向に向かう傾向を検出するが、iPS特異的欠損は検出されないゲノムマップを示す図である。この図は、19個のES細胞株の細胞株特異的偏差（x-軸）を、11個のiPS細胞株の細胞株特異的偏差（y-軸）と比較する散布図を示し、いずれの場合においてもES細胞参照と比較して測定して、関連する細胞株について平均した。細胞株を自身と比較するのを防止するために、各ES細胞株を、参照に対して採点する際にES細胞参照から一時的に除去した。選択された遺伝子はオレンジ色で強調してあり、挿入されたベン図は、全てのES細胞株においておよび全てのiPS細胞株において平均した最も外れている2,000個の遺伝子の重なり部分を可視化している。トランスジェンサイレンシングがiPS細胞株において偽性の高度メチル化を生じることから、リプログ

10

20

30

40

50

ラミング因子OCT4、SOX2、およびKLF4を分析から除外した(図9C)。全ての遺伝子およびプロモーター領域と共にES細胞およびiPS細胞株におけるその細胞株特異的偏差の平均値のリストを表4および5に示す。

【図3C】iPS細胞株においてより高い変動性の方向に向かう傾向を検出するが、iPS特異的欠損は検出されないゲノムマップを示す図である。この図は、ES細胞参照値と比較して測定し、全ての遺伝子において平均した、19個のES細胞株、11個のiPS細胞株、および6個の初代培養線維芽細胞株の細胞株特異的偏差の箱ひげ図を示す。19個のES細胞株における細胞株特異的偏差の分布を、平均ゼロおよび分散1へと正規化して、他の2つの分布をそれに合わせて変倍した。(この正規化は同じ変倍パラメータを用いたことから、3つの分布間の比較に影響を及ぼさない)。

【図3D】iPS細胞株においてより高い変動性の方向に向かう傾向を検出するが、iPS特異的欠損は検出されないゲノムマップを示す図である。この図は、ES細胞株とiPS細胞株とを区別するための、既に公表された3つのiPS細胞シグネチャーと新たに誘導された3つの分類子の予測力を要約する性能表を示す。比較するために、表はさらに、ES細胞株と線維芽細胞(陽性対照)とを区別する3つの新しく誘導された分類子の性能および3つのトリビアルな分類子(陰性対照)の性能も記載する。iPS細胞株として間違っ予測される(偽陽性、FP)またはES細胞株として間違っ予測される(偽陰性、FN)細胞株の数を最小限にしながら、ES細胞(真の陰性、TN)の中でiPS細胞株(真の陽性、TP)を同定するための予測精度、感度、および特異性を示す。結果の頑健性を増加させるために、全ての値を、交差検証を100回無作為に繰り返して平均した。表中の軽微な数値の不一致は、全ての値の端数を切り捨てて整数にしたことによる。交差検証した分類子および公表されたシグネチャーの性能推定値は、同じタイプの新しいデータ(同じ培養条件、同じアッセイ等)について再現可能でありうる試験独自の精度であると見なすべきである。

【図4A】ES細胞参照との統計学的比較がES/iPS細胞株特異的偏差を同定することを示す図である。この図は、箱ひげ図で示されるES細胞参照範囲帯と比較した19個のES細胞株および11個のiPS細胞株におけるDNAメチル化(左)および遺伝子発現(右)の分布を示す(詳細に関しては図1Cを参照されたい)。ES細胞参照から20%ポイントより多く、および0.1%未満のFDR(DNAメチル化)で外れている、または1より上の絶対対数倍率変化で、および10%未満のFDR(遺伝子発現)で外れているESまたはiPS細胞株を、カラーの三角印で強調する。細胞株をそれ自身と比較することを防止するために、各ES細胞株を参照に対して採点する際には、それをES細胞参照から一時的に除去した。差次的にメチル化および発現された遺伝子の完全なリストは、ウェブサイト"<http://scorecard.computational-epigenetics.org/>"から入手可能であり、本明細書において開示される表4および5において利用可能である。

【図4B】ES細胞参照との統計学的比較がES/iPS細胞株特異的偏差を同定することを示す図である。この図は、DNAメチル化(左)および遺伝子発現(右)に関して、ES細胞参照と比較した細胞株特異的外れ値の数を要約する偏差スコアカードを示す。細胞株の品質のさらなる指標として、スコアカードは、図2DにおけるCD14に関して示されるように、ある特定の経路に沿った細胞株の分化指向性を弱める能力を有する、影響を受けた系列マーカー遺伝子の数を記載する。表はまた、20個の低継代ES細胞株(下の行)において外れている遺伝子の平均数を示し、どれほどの数が低継代ES細胞株において同様に観察される範囲内にあるかという指標を提供する。全てのES細胞株のデータを含み、影響を受ける全ての遺伝子を記載するこのスコアカードのよりわかりやすいバージョンを、表6に示す。FDRが10%未満である差は有意であると見なされたが、絶対差が20%ポイントを超えた場合(DNAメチル化)、または絶対対数倍率変化が1を超えた場合(遺伝子発現)に限られた。細胞株選択のためにスコアカードを用いる場合、関心対象の応用を妨害しうる遺伝子特異的偏差の形跡がないかどうか、これらのデータを注意深く再検討すべきである。

【図5A】細胞株特異的分化指向性を、定量的EBアッセイによって測定することができることを示す図である。この図は、細胞株特異的分化指向性を定量するためのアッセイの概略図を示す。このアッセイの主な結果は、図5Bおよび5Dに示される系列スコアカードであ

10

20

30

40

50

る。

【図5B】細胞株特異的分化指向性を、定量的EBアッセイによって測定することができることを示す図である。この図は、低継代ヒトES細胞株セットの細胞株特異的分化指向性を要約する系列スコアカードを示す。数値は、線形目盛での相対的濃縮（正の値）または枯渇（負の値）を示す。数値は、所与のES細胞株に関する全ての生物学的レプリケートをES細胞参照（他の全てのES細胞株の生物学的レプリケートからなる）と比較する調整（moderated）t-検定を行った後、関心対象の細胞系列または胚葉に関して関連性を有するマーカー遺伝子セットの遺伝子セット濃縮分析によって算出した（表7）。ES細胞株は、アッセイを較正するために用いられる他の全てのES細胞株の平均とちょうど同じように分化すると、分化指向性ゼロを示すことから、縦列は全てゼロに集まる。値は、互いに対して解釈されるべきであり、高い数値は、より高い分化指向性を示し、より低い値は、より低い分化指向性を示すが、絶対値は、測定単位および直接的な生物学的解釈を有しない。代表的なEBの写真を図10Aに示し；予測のサブセットを検証する免疫染色を図10Bに示し；全てのマーカー遺伝子のリストは表7から利用可能であり；スコアカード作成の元となる遺伝子発現データは、表10において利用可能であり；かつ、1つの遺伝子発現レベルと系列スコアカード分化指向性との関連に関する報告を表8に示す。

10

【図5C】細胞株特異的分化指向性を、定量的EBアッセイによって測定することができることを示す図である。この図は、ES細胞株およびiPS細胞株、ES由来およびiPS由来EB、ならびに初代培養線維芽細胞株の転写類似性に関する二次元の多次元尺度構成法マップを示す。500個の系列マーカー遺伝子の遺伝子発現を、nCounterシステムを用いて測定し、正規化したデータを、点の互いの距離が、遺伝子発現レベルの500次元空間におけるその距離を表すように、平面に投射した。各点は、1つの生物学的レプリケートに対応し、投射は多次元尺度構成法を用いて行われた。2つのiPS細胞株は、正常なEBの形成能が有意に損なわれ（hiPS 15b、hiPS 29e、矢印で強調して、「壊れたEB」と表示される）、および1個のiPS細胞株は、正常なEBを全く形成できなかったが（hiPS27e、矢印で強調して、「不良EB」と表示される）、16日EB分化後であっても多能性細胞を想起させる遺伝子発現プロファイルを維持した。これらの3つの細胞株の生物学的レプリケートを全て矢印で強調し、3つ全ての細胞株はまた、系列スコアカードに従って有意に低減された分化指向性を示す（図5D）。

20

【図5D】細胞株特異的分化指向性を、定量的EBアッセイによって測定することができることを示す図である。この図は、ヒトiPS細胞株セットの系列特異的分化指向性を要約する系列スコアカードを示す。スコアカードを、図5Bに記述したように誘導して、ES細胞参照に対して標準化した。スコアを、各細胞株に関して利用可能な全ての生物学的レプリケートに対して算出した。代表的なEBの写真を図10Cに示す。系列スコアカードの特定局面を検証するFACS分析を図10Dに示す。

30

【図6A】系列スコアカードが、運動ニューロン分化の細胞株特異的な差を予測することを示す図である。この図は、インビトロで運動ニューロンを作製する効率の、細胞株特異的な差を測定するための技法の概要を示す。13個のiPS細胞株（表1を参照されたい）を、32日の神経分化プロトコルに供して、分化効率を、運動ニューロンマーカーISL1およびHB9に関して染色陽性である細胞の自動計数によって定量した（Boulting et al., 同時提出）。実験は全て少なくとも生物学的に3通りで行った。

40

【図6B】系列スコアカードが、運動ニューロン分化の細胞株特異的な差を予測することを示す図である。この図は、神経系列分化に関する系列スコアカード推定値とインビトロで運動ニューロンを作製する細胞株特異的効率との相関を示す（ r_p 、ピアソン相関係数； r_s 、スピアマン相関係数）。運動ニューロンの効率を、32日神経分化プロトコルの終了点で、ISL1-陽性（左）およびHB9陽性細胞（右）の百分率によって測定した。生物学的レプリケートおよび標準誤差を含むさらなる詳細を表9に示す。

【図6C】系列スコアカードが、運動ニューロン分化の細胞株特異的な差を予測することを示す図である。この図は、3つの胚葉の系列スコアカード推定値とインビトロで運動ニューロンを作製する細胞株特異的効率との相関を示す（ r_p 、ピアソン相関係数； r_s 、スピアマン相関係数）。

50

アマン相関係数)。運動ニューロン効率を、32日神経分化プロトコルの終了点でのISL1陽性細胞の百分率によって測定した。HB9陽性細胞の百分率との類似の比較を図11Aに示す。生物学的レプリケートおよび標準誤差を含むさらなる詳細を表9に示す。

【図7A】スコアカードの小さい変更により、ヒトiPS細胞株のハイスループットな特徴決定が可能となることを示す図である。この図は、多次元に沿ってES/iPS細胞株の品質および有用性を定量するためのスコアカードの1つの態様の要約を示す。この表は、図4Bおよび図5Dからのデータをまとめて、(i) 遺伝子特異的DNAメチル化のES細胞参照からの偏差、(ii) ES細胞参照と比較してアップまたはダウンレギュレートされた遺伝子、および(iii) 3つの胚葉に関する定量的分化指向性、に関する概要を提供する。

【図7B】スコアカードの小さい変更により、ヒトiPS細胞株のハイスループットな特徴決定が可能となることを示す図である。この図は、スコアカードの異なる次元間のペアワイズ相関を示し、エピジェネティックな転写の偏差を示す遺伝子数ならびに分化指向性の推定値が、ES/iPS細胞株の品質および有用性に関する余分の情報ではなく補足情報を提供することを示す。

【図7C】スコアカードの小さい変更により、ヒトiPS細胞株のハイスループットな特徴決定が可能となることを示す図である。この図は、DNAメチル化アッセイのゲノムカバー率が低減された、スコアカード性能のシミュレーションを示す。19個全てのES細胞株(または大きさ10、5、および1個の無作為なサブセット)のデータに基づいて、全ての遺伝子を、ES細胞参照からの平均偏差に従って順位づけした。次に、上位1%、5%、10%、90%までの変動性の高いES細胞遺伝子を選択し、これらの遺伝子のみを偏差に関してモニターする場合に検出されるであろうiPS細胞株特異的偏差の百分率に関して評価した。これらのデータは、20%の最も感受性が高いプロモーター領域に焦点を当てることによって、iPS細胞株特異的偏差の90%を検出することが可能であることを示している。図12は、最も転写が変動する遺伝子に同様に焦点を当てることにより、DNAメチル化の場合より、遺伝子発現の細胞株特異的偏差の検出能が極めて強く低減されることを示す。

【図7D】スコアカードの小さい変更により、ヒトiPS細胞株のハイスループットな特徴決定が可能となることを示す図である。この図は、EB分化がないスコアカード性能のシミュレーションを示す。遺伝子発現プロファイルを、nCounterシステムを用いてES細胞株およびiPS細胞株に関して得て、16日目のEBからの遺伝子発現プロファイルと同じように処理し、通常の成長条件下で維持されたES/iPS細胞株の遺伝子発現プロファイルのみに基づく系列スコアカードを作成した。散布図は、16日目のEBから算出された系列スコアカード推定値(x-軸)と、多能性状態から算出された系列スコアカード推定値(y-軸)との相関を可視化し、両者のあいだに良好な一致を示したが、後者では実質的に低減されたダイナミックレンジを示した。

【図7E】スコアカードの小さい変更により、ヒトiPS細胞株のハイスループットな特徴決定が可能となることを示す図である。この図は、ヒト多能性細胞株のハイスループットな特徴決定に関するワークフローの概要に関する概略図を示す。細胞株の特徴決定は、おそらく最も情報に富む定量的分化アッセイから始め、関心対象の応用にとって有用であると系列スコアカードが特定する細胞株に限ってさらなる特徴決定を行うことを繰り返し行う。必ずしもあらゆる細胞株が全ての応用にとって等しく適しているわけではないことに注意されたい。現在の試験からのデータは、ES等級のiPS細胞株が存在することを明らかに示している。

【図8A】本試験に含まれるES細胞株の代表的な画像および免疫染色を示す。

【図8B】RRBSによって得られたDNAメチル化データのゲノムカバー率を示す(要約)。遺伝子プロモーター、CpGアイランドおよび推定のエンハンサーにおけるRRBSカバー率を示す円グラフ。カバー率は、所与のタイプの各領域内でCpGでの個々の観察数(すなわち、高品質なシーケンシングのリード)として測定される。代表的なヒトES細胞株(H1)に関するデータを示す。

【図8C】RRBSによって得られたDNAメチル化データのゲノムカバー率を示す(特異的遺伝子座)。SNA11遺伝子座でのRRBSカバー率を示すUCSCゲノムブラウザスクリーンショット

10

20

30

40

50

ト。SNAI1のプロモーター領域（紫色）は、最高密度のCpG（黒色）を示し、同様に最高のRRBSカバー率（青色）を示す。さらなるRRBSのカバーは、下流のCpGアイランド（緑色）および上流の調節エレメント（オレンジ色）に集中している。ほとんどのCpGリッチ領域はメチル化されていないが（明るい青色）、CpGプア領域は、メチル化される傾向がある（暗い青色）。各青色のドットは、RRBSによってカバーされる1つのCpGに対応する。H1とH7のあいだでいくつかのエピジェネティックな変動を認めることができるが、全体としてプロモーター領域は、示された全てのES細胞株においてメチル化されない。

【図8D】19個の異なるES細胞株に対するプロモーターDNAのメチル化の全体的な比較を示す。19個のES細胞株に対するプロモーターDNAのメチル化レベルの平均値を比較するペアワイズ散布図。全てのペアワイズ比較に関して高い類似性が観察された。しかし、ES細胞株の対のあいだには、このダイアグラムから見ることもできる2つのタイプの差が存在した：(i) XまたはY軸に近い左下に位置する小さいが密度の高い点の集まり：これらは広範囲のX不活化を有する雌性のES細胞株を雄性のES細胞株と区別するX-染色体関連の差である、(ii) ダイアグラム全体に散在する対角線から外れた点：これらの差のほとんどは、常染色体に存在し、ES細胞株間のエピジェネティックな差を構成する。

【図9A】11個のiPS細胞株および6個の初代培養線維芽細胞株におけるプロモーターDNAのメチル化の全体的な比較を示す。11個のiPS細胞株と6個の初代培養線維芽細胞株に対するプロモーターDNAのメチル化レベルの平均値を比較するペアワイズ散布図。iPS細胞株において高い類似性が観察されたが、実質的な差により、iPS細胞株は線維芽細胞株と区別される。

【図9B】DNAのメチル化と遺伝子発現データとのジョイントクラスタリング分析の結果の一例を示す。ヒトES細胞株、iPS細胞株、および線維芽細胞の合同階層クラスタリングおよびヒートマップ。クラスタリングを、図1の説明文に記述されるように行った。「MEG3」の列では、MEG3非コードRNAの発現状態を示す：「+」は、それぞれの細胞株において発現されるMEG3を表し（MEG3発現レベル > 1）、「-」は、MEG3が発現されないことを示す（MEG3発現レベル < 1）。

【図9C】トランスジーンサイレンシングによるKLF4のコード領域における偽性高度メチル化を示す。トランスジーンサイレンシングがリプログラミング因子の内因性の遺伝子座でどのように偽性高度メチル化を生じるかを説明するUCSCゲノムブラウザスクリーンショット。RRBSのリードをゲノムに整列させる様式により、ほとんどのウイルストランスジーンリードは、OCT4、SOX2、およびKLF4の内因性の遺伝子座に置かれる。この現象をKLF4に関して説明する。ES細胞において、KLF4遺伝子は主としてメチル化されていない（緑色）のに対し、iPS細胞では部分的にメチル化されているように思われるが、トランスジーンの一部であるエキソンのみでメチル化されているのであり（赤色）、トランスジーンの一部ではないイントロンではメチル化されない（青色）。さらに、hiPS27e（黄色）における不完全なトランスジーンサイレンシングは、トランスジェニックKLF4における実質的により低いDNAメチル化レベルと相関する。

【図9D】MEG3発現が、ES細胞参照からのエピジェネティックなまたは転写の偏差の強い予測因子ではないことを示す。以下の細胞株に関する、全ての遺伝子において平均したES細胞参照からの細胞株特異的偏差の箱ひげ図：(i) MEG3非コードRNAが発現されるES細胞株（図9Bを参照されたい）、(ii) MEG3が発現されない細胞株（HUES1、HUES3、HUES13、HUES44、HUES45、HUES53、HUES66、H1およびH7）、ならびに(iii) 6個の初代培養線維芽細胞株。

【図10A】スコアカードが、ヒト多能性細胞株の迅速かつ包括的な特徴決定を可能にすることを示す図である。この図は、3つのES細胞株（HUES1、継代28および29；HUES8、継代29および30；H1、継代37および38）の生物学的レプリケート間のDNAメチル化を比較するペアワイズ相関係数および散布図を示す。さらに、DNAメチル化比較は、ウィスコンシン大学（継代25）およびCellular Dynamics社（継代32）でそれぞれ成長させたH1の2つの生物学的レプリケートを含む。全てのペアワイズ比較に関して高い類似性が観察された。しかし、ES細胞株の対には2つのタイプの差があることを、これらのダイアグラムから

10

20

30

40

50

見ることができる：(i) x-軸またはy-軸（DNAメチル化のみ）に近い左下に位置する小さい密度の濃い点の集まり。これらの点は、広範囲のX不活化を有する雌性ES細胞株を雄性ES細胞株と区別するX染色体関連の差に対応する。(ii) ダイアグラム全体に散在する対角線を外れた点。これらの差の多くは常染色体に存在し、ES細胞株間のエピジェネティックなまたは転写の差を構成する。

【図10B】スコアカードが、ヒト多能性細胞株の迅速かつ包括的な特徴決定を可能にすることを示す図である。この図は、3つのES細胞株（HUES1、継代28および29；HUES8、継代29および30；H1、継代37および38）の生物学的レプリケート間の遺伝子発現を比較するペアワイズ相関係数および散布図を示す。

【図10C】スコアカードが、ヒト多能性細胞株の迅速かつ包括的な特徴決定を可能にすることを示す図である。この図は、不均一な細胞集団におけるDNAメチル化の差に関する最小閾値の図を示す。細胞株における小さいDNAメチル化の差でさえも、変動が低ければ統計学的に高度に有意でありうる。しかし、これは必ずしも生物学的有意性を暗示するわけではない。それゆえ、関連すると見なされるためには、10%の偽発見率（FDR）の統計学的有意性閾値のほかに、2つの細胞株間（または1つの細胞株とES細胞参照とのあいだ）での20パーセントポイントを超えるDNAメチル化の差が必要である。ほとんどの細胞株が何らかの程度の不均一性を示すことを考慮に入れると、細胞株が、ES細胞参照から20パーセントポイントより多く外れることができるいくつかの場合がある：(i) 全ての細胞が20パーセントポイント増加した（減少した）DNAメチル化レベルを示す；(ii) 全ての細胞の20%のサブセットが、100パーセントポイント増加する（減少する）DNAメチル化レベルを示すが、残りの80%はいかなる差も示さない；(iii) 図面に示される任意の組み合わせ。

【図10D】スコアカードが、ヒト多能性細胞株の迅速かつ包括的な特徴決定を可能にすることを示す図である。この図は、エピジェネティックな転写の空間におけるES細胞株とiPS細胞株の類似性の模式図を示す。左側の密度プロットは、ヒトES細胞において観察された変動を描写する。2つの交叉点は、全てのES細胞株およびiPS細胞株の（理論的）平均値を示し、本試験は、この値を20個のヒトES細胞株および12個のヒトiPS細胞株のプロファイリングによって近似する。右側の散布図は、その中程度に増加した変動（図3C）ならびに微量のiPS細胞株はES細胞株と区別できないという知見（図3D）を考慮に入れて、多数のヒトiPS細胞株の分布をシミュレートする。ガウスを用いて、インシリコでES細胞およびiPS細胞の分布をシミュレートした。

【図11A】ゲノム全域にわたるDNAメチル化および/または遺伝子発現データに基づいて偏差スコアカードを算出するためのアルゴリズムの概要を示す。

【図11B】分化途中のEBにおけるマーカー遺伝子発現に基づいて系列スコアカードを算出するためのアルゴリズムの概要を示す。

【図12A】ES細胞由来EBの代表的な画像の例を示す。低継代ヒトES細胞株に由来する16日目の胚様体の画像を用いて、系列スコアカードの参照データセットを確立した。

【図12B】選択された系列マーカー遺伝子の免疫染色の画像を示す。選択された系列スコアカード推定値の免疫染色による検証は、5つのマーカー遺伝子に関する系列スコアカードの分化特性、mRNAレベル、およびタンパク質染色のあいだの良好な定性的一致を示す。非定方向性EB分化を4つの代表的なES細胞株について行った。2日後、EBをマトリゲルにおいて平板培養して、さらに5日間分化させた。EB分化の7日後、3つの胚葉のマーカー遺伝子に関する免疫染色を行った。図面は、未分化ES細胞、7日目のEB、および免疫染色の代表的な写真を示す。nCounterシステムを用いて、16日目のEBについて遺伝子発現レベルを得た（表10）。

【図12C】iPS細胞株および誘導EBの画像を示す。系列スコアカードに関するiPS細胞株および誘導EBの画像。

【図12D】内胚葉マーカー遺伝子AFPのFACS分析を示す。hiPS17およびhiPS27eの16日目のEBにおけるFACSによって決定したAFP陽性細胞数とmRNA発現レベルの比較。

【図12E】外胚葉分化（青色）および中胚葉分化（赤色）にとって都合がよい条件下で

分化させた4つのES細胞株 (HUES1、HUES8、H1、H9) の系列スコアカードの平均値を示す。

【図13A】運動ニューロン効率 (HB9+細胞) と胚葉に関する系列スコアカード指向性との相関を示す図である。この図は、外胚葉分化への細胞株特異的分化指向性の系列スコアカード推定値と、運動ニューロンへの定方向性分化効率との相関を示す散布図である。

【図13B】運動ニューロン効率 (HB9+細胞) と胚葉に関する系列スコアカード指向性との相関を示す図である。この図は、中胚葉分化への細胞株特異的分化指向性の系列スコアカード推定値と、運動ニューロンへの定方向性分化効率との相関を示す散布図である。

【図13C】運動ニューロン効率 (HB9+細胞) と胚葉に関する系列スコアカード指向性との相関を示す図である。この図は、内胚葉分化への細胞株特異的分化指向性の系列スコアカード推定値と、運動ニューロンへの定方向性分化効率との相関を示す散布図である。各細胞株に関して運動ニューロン効率を、32日間の運動ニューロン分化プロトコールの終了点でのHB9陽性細胞の百分率の自動計数によって測定した。HB9は、ほとんどの他の神経細胞型において発現されない運動ニューロンの非常に特異的なマーカーである。

【図14A】最も転写的に変動性の遺伝子のカバー率 (遺伝子発現) が低いスコアカード (図7Cのような) 性能により、DNAメチル化の場合よりも、遺伝子発現における細胞株特異的偏差の検出能が極めて強く低減されることを示す。飽和曲線は、ES細胞株においてES細胞参照からの最高の平均絶対偏差を示す上位Xパーセントの遺伝子だけに焦点を当てた場合に検出されるES細胞参照と比較したiPS細胞株特異的偏差の数を示す。

【図14B】ゲノムカバー率が低いDNAメチル化アッセイに関するスコアカード性能を推定する飽和プロットを示す。飽和プロットは、20個全てのES細胞株 (またはサイズ10、5、および1個の無作為サブセット) のデータに基づき、全ての遺伝子を、ES細胞参照からの平均偏差に従って順位づけした。次に、上位1%、5%、10%、または90%までの変動性の高いES細胞遺伝子を選択して、これらの遺伝子のみを偏差に関してモニターした場合に検出されるiPS細胞株特異的偏差の百分率を算出した。

【図14C】ゲノムカバー率が低い遺伝子発現アッセイに関するスコアカード性能を推定する飽和プロットを示す。飽和プロットは、20個全てのES細胞株 (またはサイズ10、5、および1個の無作為サブセット) のデータに基づき、全ての遺伝子を、ES細胞参照からの平均偏差に従って順位づけした。次に、上位1%、5%、10%、または90%までの変動性の高いES細胞遺伝子を選択して、これらの遺伝子のみを偏差に関してモニターした場合に検出されるiPS細胞株特異的偏差の百分率を算出した。

【図15】ヒト多能性細胞株の品質評価のために現在用いられている方法のいくつかを示す。安価で単純なアッセイは特異性を欠き、最も厳格なアッセイは、ヒトでは利用できない。奇形腫はヒトに関して最も標準的であると見なされているが、奇形腫は労力および費用がかかり、高い動物試験負荷を課して、資格を有する病理学者の評価に大きく依存し、このため定量することが難しい。

【図16】異なるヒストンメチル化マークに関するChIP-seqアプローチを用いてヒストンメチル化プロファイリングを行った1つの態様を示す。ChIP-seq法のこの態様を用いると、全てのES/iPS細胞において良好な定性的一致が存在することが認められるが、ChIP-seq法は異なる定量をもたらし、かつ多数の細胞を必要とする。したがって、DNAメチル化を決定するために代替法を用いる可能性がある。

【図17】異常なDNAメチル化遺伝子を有するiPS細胞株を選択する概略図を示す。パイサルファイトDNAメチル化シーケンシングを用いた多くのES細胞株におけるDNAメチル化マッピングを用いて、正常な変動を確立する。次に、関心対象細胞における異なる遺伝子のDNAメチル化レベルを、それらの遺伝子に関する正常なDNAメチル化レベルと比較して、正常範囲外に外れるメチル化レベルを有する遺伝子は外れ値であると見なす。

【図18】本試験において用いられる多様な異なるES細胞株およびiPS細胞株において増加または減少したメチル化レベルを有する遺伝子の数を示す一例を示す。

【図19】ES、iPS、および線維芽細胞における高度メチル化 (図19A) および低メチル化 (図19B) 遺伝子数のベン図を示す。図19Aは、116個の遺伝子がES細胞およびiPS細胞の両

10

20

30

40

50

方において高度メチル化された態様を示し、この態様において、11個はES細胞および線維芽細胞の両方において高度メチル化され、65個はiPS細胞と線維芽細胞の両方において高度メチル化された。この態様のこの例において、6個のみの遺伝子が3つ全てのタイプの細胞において高度メチル化された。図19Bは、ES細胞とiPS細胞の両方において低メチル化される116個の遺伝子が存在し；かつ、83個はES細胞と線維芽細胞の両方において高度メチル化され、および217個がiPS細胞と線維芽細胞の両方において高度メチル化される1つの態様を示す。本態様のこの例において、58個のみの遺伝子が3つ全ての細胞において高度メチル化された。

【図20】多様な異なるES細胞およびiPS細胞におけるメチル化レベルの正常な変動と比較して増加または減少したメチル化を有する遺伝子の数、ならびに参照メチル化レベルの正常な変動と比較して増加または減少したメチル化レベルを有する癌遺伝子の数を示すスコアカードの1つの態様を示す。少数の高度メチル化および/または低メチル化癌遺伝子を有する多能性細胞株は、エピジェネティックに「安全な」ESまたはiPS細胞と呼ばれ、より多数の高度メチル化および/または低メチル化癌遺伝子を有する細胞を、エピジェネティックな外れ値と呼び、これらは治療および/または他の応用において用いるためにおそらく安全ではない。

【図21】ヒトiPS株セットの分化の偏りまたは指向性を決定するための細胞株分化アッセイを要約する系列スコアカード作成の模式図を示す。この態様において、16日目の胚体(EB)分化プロトコルを用いてスコアカードを誘導したが、より短い分化プロトコル、たとえばEB0(0日目のEB)からEB32(21日目のEB)またはそれより長い任意の期間を用いることができる。500個の「系列遺伝子発現遺伝子」の遺伝子発現プロファイリングを用いて、異なる細胞型および系列に沿って分化する多能性幹細胞株の指向性を定量し、バイオインフォマティクス分析を用いて、濃縮対枯渇遺伝子セットを決定し、複数の他の多能性細胞株(たとえば、ES細胞株およびiPS細胞株)と比較して、系列スコアカードを作成した。

【図22】図22Aは、ヒトiPS株の運動ニューロンへの定方向性分化における系列スコアカードの実験的検証を示す。全てのiPS細胞株が運動ニューロンへと分化した。図22Bは、運動ニューロンへの分化効率を示す系列スコアカードの1つの態様を示し、これはIsl1に関する染色(>60,000個の細胞による2~3回の独立した反復)によって測定された。トランスクリプト発現をqPCRによってアッセイした。そのような系列スコアカードを、500個の「系列遺伝子発現遺伝子」の遺伝子発現プロファイリングによって作成して、異なる細胞型および系列に沿って分化する多能性幹細胞株の指向性を定量し、バイオインフォマティクス分析を用いて、濃縮対枯渇遺伝子セットを決定して、複数の他の多能性幹細胞株(たとえば、ES細胞株およびiPS細胞株)と比較して、系列スコアカードを作成した。

【図23-01】関心対象の多能性幹細胞株に関する偏差スコアカード作成用のコンピュータプログラムの命令の態様のフローチャートを示す。データは、プロセッサおよび関連メモリまたは保存デバイス、および遺伝子マッピングモジュール、参照比較モジュール、正規化モジュール、関連性フィルターモジュール、遺伝子セットモジュール、および偏差スコアカードを表示するためのスコアカードディスプレイモジュールを含むコンピュータに入力される。

【図23-02】図23-01の続きを示す。

【図24】関心対象の多能性幹細胞株に関する系列スコアカード作成用のコンピュータプログラムの命令の1つの態様のフローチャートを示す。偏差スコアカードを作成するために得られたデータ(たとえば、関心対象の多能性幹細胞株に関するDNAメチル化データおよび/または遺伝子発現データ)を用いることができるが、この態様において、入力されるデータは、関心対象の多能性幹細胞株の遺伝子発現データである。データは、プロセッサおよび関連メモリおよび/または保存デバイス、およびアッセイ標準化モジュール、試料標準化モジュール、参照比較モジュール、遺伝子セットモジュール、濃縮分析モジュール、および系列スコアカードを表示するためのスコアカードディスプレイモジュールを含むコンピュータに入力される。

10

20

30

40

50

【図25】関心対象の多能性幹細胞を特徴決定して、偏差および/または系列スコアカードを作成するためのハイスループットシステムに関する本発明の1つの態様の単純化ブロックダイアグラムを示す。決定モジュールは、遺伝子発現および/またはDNAメチル化を測定するための任意の装置または機器でありうる。

【図26】DNAメチル化アッセイおよび遺伝子発現アッセイからのデータが、任意の場所のコンピュータシステムによって処理され、かつユーザーインターフェースを通してアクセスすることができるように構成されることを可能にする、本発明の1つの態様の単純化ブロックダイアグラムを示す。この態様において、各多能性幹細胞のデータはデータベースに保存されている。

【図27】図23および24において概要した命令を実行するように構成することができるコンピュータシステムの例示的なブロックダイアグラムを示す。

【発明を実施するための形態】

【0066】

発明の詳細な説明

本発明は全般的に、多能性幹細胞に関する参照データセットまたは「スコアカード」、ならびに所望の使用に関する多能性幹細胞株の機能性および適切性を予測するためのスコアカードを作成するための方法、システム、およびキットに関する。「スコアカード」は、幹細胞におけるメチル化などの少なくとも1つの正常な翻訳後修飾に関する参照値範囲、および任意で幹細胞における分化関連遺伝子の正常な発現パターンに関する参照値範囲、および任意でさらに神経幹細胞、造血幹細胞、膀胱幹細胞、および他のより限られた幹細胞マーカーなどの系列特異的マーカーの正常範囲を提供する。いくつかの局面において、スコアカードは、DNAメチル化参照セットなどの翻訳後修飾参照セット、分化指向性参照セット、および遺伝子発現データセットから選択される少なくとも2つの参照データセットを含む。いくつかの態様において、スコアカードはさらに、関心対象の多能性幹細胞が、正常な多能性幹細胞変動の正常なパラメータ内に入るか否かを判定するためのガイドラインを提供する。そのようなガイドラインは、好ましくはコンピュータ実行可能なフォーマットで存在する。

【0067】

いくつかの態様において、スコアカードは、エピジェネティックなまたは翻訳後修飾、例えば本明細書において記載した19個の異なるES細胞株のデータから編集された、DNAメチル化参照セット、分化指向性参照セット、および遺伝子発現データセットから選択される、少なくとも2つの参照データセットを含む。代替の態様において、スコアカードは、望ましい特徴を有する多能性幹細胞、たとえば、膀胱系などの内胚葉系列に分化する分化指向性、たとえば外胚葉または中胚葉分化マーカーを有する多能性幹細胞のデータから編集されたスコアカードである。

【0068】

本発明のもう1つの局面は、所望の使用に関する多能性幹細胞株の機能性および適切性を予測するために、エピジェネティックなプロファイリング、分化アッセイ、および遺伝子発現アッセイから選択される少なくとも2つの幹細胞アッセイを用いる段階を含む、スコアカードを作成する方法に関する。いくつかの態様において、所定の応用に関する多能性幹細胞の有用性を有効かつ正確に予測するために、ならびにたとえば治療的使用、薬物スクリーニング、および毒性アッセイ、ならびに所望の細胞系列への分化についてのその適切性などの下流の応用に関するその適切性を決定するための多能性幹細胞株の特定の特徴を同定するために、スコアカード参照データを多能性幹細胞データと比較することができる。

【0069】

いくつかの態様において、DNAメチル化参照セットは、第一の参照遺伝子セットのメチル化レベルに関し、DNAメチル化参照遺伝子は癌遺伝子および/または発達遺伝子でありえて、表12Aに開示される。いくつかの態様において、第一の参照DNAメチル化遺伝子セットにおいて用いられる遺伝子は、表12Aおよび/または表12Cおよび/または表13A、13B、

10

20

30

40

50

または表14における遺伝子のリストから選択され、任意に組み合わせた、少なくとも約200個、または少なくとも約300個、または少なくとも約400個、または少なくとも約500個、または少なくとも約600個、または少なくとも約800個、または少なくとも約1000個、または少なくとも約1500個、または少なくとも約2000個、または少なくとも約3000個、または少なくとも約4000個、または少なくとも約5000個の遺伝子である。いくつかの態様において、遺伝子は、表12A、表12C、表13A、表13B、または表14のいずれかに記載される遺伝子の1~200番、または1~500番、または1~1000番から選択される遺伝子セットの任意の組み合わせである。

【0070】

したがって、本発明の1つの局面は、所望の使用に関する多能性幹細胞株の機能性および適切性を予測するための方法および複数のアッセイに関する。いくつかの態様において、所望の使用に関する多能性幹細胞株の機能性および適切性を予測するために、少なくとも1つ、または少なくとも2つ、または少なくとも3つの幹細胞アッセイを単独または任意の組み合わせで用いることができる。いくつかの態様において、第一のアッセイは、エピジェネティックなプロファイリング、たとえば、多能性幹細胞株において活性化された遺伝子を決定するための定義された特定の遺伝子セットの遺伝子メチル化の評価である。いくつかの態様において、第二のアッセイは、多能性幹細胞株が特定の系列に沿って分化する指向性を決定するための分化アッセイである。いくつかの態様において、アッセイは、遺伝子発現アッセイ、たとえばを決定するための全ゲノム遺伝子発現アッセイである。

【0071】

もう1つの局面は、本明細書において「スコアカード」と呼ばれる参照データセットに関し、これは、本発明の3つの複合アッセイに由来する多数の異なる多能性幹細胞株の結果からの平均データであり、これにより、関心対象のその多能性幹細胞株と比較するために当業者が用いることができる「スコアカード」を構成する参照データが提供され、ここで、参照「スコアカード」との比較を用いて、所定の応用に関する多能性幹細胞の有用性を、ならびに関心対象の多能性幹細胞株たとえばES細胞またはiPS細胞株の任意の特異的特徴を有効かつ正確に予測することができる。したがって、本明細書において開示される方法、アッセイ、およびスコアカードは、たとえば治療的使用、薬物スクリーニングおよび毒性アッセイ、ならびに所望の細胞系列への分化についてのその適切性などの下流の応用に関するその適切性を決定するための幹細胞の特定の特徴を同定するために用いることができる。

【0072】

いくつかの態様において、本明細書において開示されるアッセイは、たとえば胚幹細胞、自己成体幹細胞、iPS細胞、およびリプログラム細胞、直接リプログラム細胞、または部分リプログラム細胞などの他の多能性幹細胞株などの、しかしこれらに限定されるわけではない多様な多能性幹細胞株の品質を特徴決定および決定するために用いることができる。いくつかの態様において、幹細胞株はヒト幹細胞株である。いくつかの態様において、多能性幹細胞株は、遺伝子改変多能性幹細胞株である。いくつかの態様において、多能性幹細胞株が、治療的に用いるためのまたは対象に移植するための細胞株である場合、多能性幹細胞株は、自己多能性幹細胞株、たとえば幹細胞集団が移植によって戻される対象に由来する幹細胞株であり、代替りの態様において、多能性幹細胞株は同種異系多能性幹細胞株である。

【0073】

定義

便宜上、本明細書、実施例、および添付の特許請求の範囲において用いられる一定の用語を本章に集める。それ以外であることを明記している場合を除き、または文脈上暗黙に意味している場合を除き、以下の用語および句は、以下に提供される意味を含む。それ以外であると明白に述べている場合を除き、または本文から明らかである場合を除き、以下の用語および句は、本発明が属する技術分野においてその用語または句が持つ意味を除外しない。定義は、特定の態様の記述を助けるために提供され、本発明の範囲は特許請求の

10

20

30

40

50

範囲によってのみ制限されることから、特許請求される本発明を制限すると意図されない。それ以外であると定義している場合を除き、本明細書において用いられる全ての科学技術用語は、本発明が属する当業者によって一般的に理解される意味と同じ意味を有する。

【0074】

本明細書において開示される「スコアカード」という用語は、参照多能性幹細胞株と比較して、関心対象の1つまたは複数の多能性幹細胞株における選択された遺伝子のDNAメチル化および/または遺伝子発現の差の要約のリストを意味し、多能性幹細胞の予測される性能、たとえば分化能および/または多能性能、および/または癌様細胞株になる素因の記録として機能する。スコアカードは、任意の形状で存在することができ、たとえばデータベース、書面、電子型などで存在することができ、注釈付きデータベースに電子的すなわちデジタル式に記録および保存することができる。いくつかの態様において、スコアカードは、参照多能性細胞株または複数の細胞株と比較した多能性幹細胞の能力（たとえば、分化能、多能性等）の予測のグラフ表示でありうる。したがって、本明細書において開示されるスコアカードは、多能性幹細胞株の特徴および潜在力の指標またはリストとして役立ち、特定の使用のためにおよび/または特定の目的に達するために特定の多能性幹細胞株の迅速かつ効率的な選択を助けるために用いることができる。

10

【0075】

本明細書において用いられる「リプログラミング」という用語は、分化した細胞（たとえば、体細胞）の分化状態を変化させるまたは逆転させるプロセスを意味する。別の言い方をすれば、リプログラミングは、より未分化のまたはより始原的な細胞型へと細胞が戻る方向への分化を促進するプロセスを意味する。完全なリプログラミングは、接合体が成体へと発達する細胞分化の際に起こる、核酸修飾（たとえば、メチル化）、クロマチンの凝縮、エピジェネティックな変化、ゲノムインプリンティング等の遺伝可能なパターンの少なくともいくつかの完全な逆転を伴う。リプログラミングは、既に多能性である細胞の既に存在する未分化状態を単に維持することは異なり、または既に多分化能細胞である細胞（たとえば、造血幹細胞）の既に存在する完全ではない分化状態を維持することとも異なる。リプログラミングはまた、既に多能性または多分化能である細胞の自己再生または増殖を促進することとも異なるが、本発明の組成物および方法はまた、そのような目的のためにも有用でありうる。

20

【0076】

本明細書において用いられる「安定にリプログラムされた細胞」という用語は、分化した細胞（たとえば、体細胞）の部分的または不完全なリプログラミングから生じる細胞を意味する。安定にリプログラムされた細胞は、本明細書において「piPSC」と互換的に用いられる。安定にリプログラムされた細胞は、完全なリプログラミングを受けておらず、したがって細胞のエピゲノムの全体的なリモデリングを行っていない。安定にリプログラムされた細胞は多能性幹細胞であり、iPSC（本明細書において定義される用語である）へとさらにリプログラムすることができ、または異なる系列に沿って分化させることができる。いくつかの態様において、部分的にリプログラムされた細胞は、3つ全ての胚性胚葉（すなわち、内胚葉、中胚葉、または外胚葉の3つ全ての胚葉）由来のマーカーを発現する。マウスにおいて、内胚葉生殖細胞のマーカーは、Gata4、FoxA2、PDX1、Nodal、Sox7、およびSox17を含む。マウスにおいて、中胚葉生殖細胞のマーカーは、Brachycury、GSC、LEF1、Mox1およびTie1を含む。マウスにおいて、外胚葉生殖細胞のマーカーは、cripto1、EN1、GFAP、Islet 1、LIM1、およびネスチンを含む。いくつかの態様において、部分的にリプログラムされた細胞は、未分化細胞である。ヒト内胚葉生殖細胞、外胚葉生殖細胞、および中胚葉生殖細胞のマーカーは、本明細書において表7において開示され、たとえば外胚葉生殖細胞のマーカーには、NCAM1、EN1、FGFR2、GATA2、GATA3、HAND1、MNX1、NEFL、NES、NOG、OTX2、PAX3、PAX6、PAX7、SNAI2、SOX10、SOX9、TDGF1、APOE、PDGFRA、MCAM、FUT4、NGFR、ITGB1、CD44、ITGA4、ITGA6、ICAM1、THY1、FAS、ABCG2、CRABP2、MAP2、CDH2、NES、NEUROG3、NOG、NOTCH1、SOX2、SYP、MAPT、THが挙げられるがこれらに限定されるわけではない。ヒト内胚葉生殖細胞のマーカーには、APOE、CDX2、FOXA2、GAT

30

40

50

A4、GATA6、GCG、ISL1、NKX2-5、PAX6、PDX1、SLC2A2、SST、ITGB1、CD44、ITGA6、THY1、CDX2、GATA4、HNF1A、HNF1B、CDH2、NEUROG3、CTNNB1、SYPが挙げられるがこれらに限定されるわけではなく、中胚葉生殖細胞のマーカには、CD34、DLL1、HHEX、INHBA、LEF1、SRF、T、TWIST1、ADIPOQ、MME、KIT、ITGAL、ITGAM、ITGAX、TNFRSF1A、ANPEP、SDC1、CDH5、MCAM、FUT4、NGFR、ITGB1、PECAM1、CDH1、CDH2、CD36、CD4、CD44、ITGA4、ITGA6、ITGAV、ICAM1、NCAM1、ITGB3、CEACAM1、THY1、ABCG2、KDR、GATA3、GATA4、MYOD1、MYOG、NES、NOTCH1、SPI1、STAT3が挙げられるがこれらに限定されるわけではない。

【0077】

「人工多能性幹細胞」または「iPSC」または「iPS細胞」という用語は、分化した細胞（たとえば、体細胞）の分化状態の完全な逆転またはリプログラミングに由来する細胞を意味する。本明細書において用いられるiPSCは、完全にリプログラムされて、完全なエピジェネティックなリプログラミングを受けた細胞である。本明細書において用いられるiPSCは、これ以上リプログラムすることができない細胞である（たとえば、iPSCは最終的にリプログラムされている）。

10

【0078】

「エピゲノムのリモデリング」という用語は、細胞においてゲノム配列または遺伝子の塩基対配列を変化させないが発現を変化させる、ゲノムの化学的修飾を意味する。

【0079】

「エピゲノムの全体的なリモデリング」という用語は、リプログラム細胞またはiPSCが由来する分化細胞からのこれまでの遺伝子発現の記憶が存在しないようなゲノムの化学修飾が起こることを意味する。

20

【0080】

「エピゲノムの不完全なリモデリング」という用語は、安定なリプログラム細胞またはpiPSCが由来する分化細胞からのこれまでの遺伝子発現の記憶が存在するようなゲノムの化学修飾が起こることを意味する。

【0081】

本明細書において用いられる「エピジェネティックなリプログラミング」という用語は、細胞におけるゲノム配列または遺伝子の塩基対配列を変化させない化学修飾を介した、細胞における遺伝子発現パターンの変化を意味する。

【0082】

本明細書において用いられる「エピジェネティックな」という用語は、「ゲノム上の」を意味する。遺伝子の配列を変化させないが、遺伝子発現に影響を及ぼし、同様に遺伝されうるDNAの化学修飾。エピジェネティックな修飾はまた、いくつかの例において、遺伝子DNAまたは核酸の配列を変化させないが、たとえばインプリンティングおよび細胞リプログラミングにおいて重要であるDNAに対する変化である翻訳後修飾または「PTM」を含むことができる。翻訳後修飾には、たとえばDNAメチル化、ユビキチン化、リン酸化、グリコシル化、スモイル化、アセチル化、S-ニトロシル化、またはニトロシル化、シトルリン化、または脱イミノ化、ネディレーション (neddylation)、OClcNAc、ADP-リボシル化、ヒドロキシル化、ファテニル化、ウフミル化、プレニル化、ミリストイル化、S-パルミトイル化、チロシン硫酸化、ホルミル化、およびカルボキシル化が挙げられる。

30

40

【0083】

本明細書において用いられる「メチル化」という用語は、遺伝子調節領域のCpGジヌクレオチド内のヌクレオチド塩基シトシンのC5位でのメチル基の共有結合を意味する。「メチル化状態」または「メチル化状態 (status)」という用語は、DNA配列内での1つまたは複数のCpGジヌクレオチドにおける5-メチル-シトシン (「5-mCyt」) の存在または非存在を意味する。本明細書において用いられる「メチル化状態 (status)」および「メチル化状態」は互換的に用いられる。メチル化部位は、配列特異的メチラーゼによって認識されてメチル化される連続的に連結したヌクレオチドの配列である。メチラーゼは、メチル化部位の1つまたは複数のヌクレオチドをメチル化する (すなわち、メチル基を共有結合させる) 酵素である。

50

【0084】

「メチル化レベル」という用語は、たとえば全てのゲノム領域およびいくつかの非ゲノム領域における標的DNAメチル化遺伝子のDNA配列上に存在するメチル化の量を意味する。いくつかの態様において、メチル化レベルは、標的遺伝子のプロモーター領域において決定される。

【0085】

本明細書において用いられる「CpGアイランド」という用語は、CpGジヌクレオチドに富む短いDNA配列であり、全てのヒト遺伝子の約半数の5'領域において見いだすことができる。「CpG部位」という用語は、CpGアイランド内のCpGジヌクレオチドを意味する。CpGアイランドは、典型的に長さが約0.2から約1 kbであるが、必ずしもこの範囲ではない。

10

【0086】

本明細書において用いられる「遺伝子プロファイル」という用語は、多能性幹細胞試料における遺伝子または遺伝子セットの遺伝子発現レベルを意味すると意図される。本発明の1つの態様において、「遺伝子プロファイル」という用語は、表12Bおよび/または表12Cに記載される遺伝子もしくは遺伝子セット、または本明細書において記述される表12B、または表12C、表13A、表13B、または表14の遺伝子の任意の選択物を意味する。

【0087】

本発明の文脈における「差次的発現」という用語は、遺伝子が、多能性幹細胞における発現のその正常な変動と比較してアップレギュレートまたはダウンレギュレートされることを意味する。遺伝子の差次的発現を算出するための統計学的方法については、本明細書の他所で述べる。

20

【0088】

「表12Bの遺伝子」とは、本明細書において「表12Bに記載される遺伝子」と互換的に用いられ、表12Bにおいて「遺伝子の名称」として記載される遺伝子の遺伝子産物を意味する。「遺伝子産物」とは、天然または人工的手段によって産生されるか否かによらず、遺伝子の転写または翻訳の任意の産物を意味する。本発明のいくつかの態様において、本明細書において言及される遺伝子は、2番目の列において「遺伝子の名称」として定義される表12Aおよび12Bおよび12Cに記載される遺伝子である。遺伝子はまた、表12A、表12C、表13A、表13B、または表14においても記載される。

【0089】

本明細書において用いられる「多能性」という用語は、異なる条件で3つ全ての胚葉（内胚葉、中胚葉、および外胚葉）の細胞型特徴へと分化する能力を有する細胞を意味する。多能性細胞は、たとえばヌードマウス奇形腫形成アッセイを用いる3つ全ての胚葉への分化能によって主に特徴決定される。多能性はまた、胚幹（ES）細胞マーカーの発現によっても証明されるが、多能性の好ましい試験は、3つの胚葉の各々の細胞へと分化できることの証明である。いくつかの態様において、多能性細胞は未分化細胞である。

30

【0090】

本明細書において用いられる「多能性」または「多能性状態」という用語は、3つ全ての胚性胚葉、すなわち内胚葉（腸管組織）、中胚葉（血液、筋肉、および血管を含む）、および外胚葉（皮膚および神経などの）への分化能を有する細胞を意味し、典型的にインビトロで長期間、たとえば1年より長くまたは30回継代を超えて分裂能を有する。

40

【0091】

「多分化能細胞」に関して用いる場合の「多分化能」という用語は、3つ全ての胚葉に由来する細胞の全てではないがいくつかへと分化することができる細胞を意味する。このように、多分化能細胞は部分的に分化した細胞である。多分化能細胞は当技術分野において周知であり、多分化能細胞の例は、たとえば造血幹細胞および神経幹細胞などの成体幹細胞を含む。多分化能とは、幹細胞が所定の系列における多くの細胞型を形成しうるが、他の系列の細胞を形成しないことを意味する。たとえば、多分化能血液幹細胞は多くの異なるタイプの血液細胞（赤血球、白血球、血小板等）を形成することができるが、ニューロンを形成することはできない。

50

【0092】

「多分化能」という用語は、全能性および多能性より低い程度の発達上の多様性を有する細胞を意味する。

【0093】

「全能性」という用語は、成体ならびに胎盤を含む胚外組織における細胞の全てを作る能力を説明する分化の程度を有する細胞を意味する。受精卵（接合体）は初期分割球（割球）であることから全能性である。

【0094】

「分化した細胞」という用語は、その天然形態において（本明細書において定義される用語である）多能性ではない、任意の初代細胞を意味する。「分化した細胞」という用語はまた、多分化能細胞などの部分的に分化した細胞、または安定な非多能性の部分的リプログラム細胞も包含する。多くの初代細胞を培養すると、完全に分化した特徴の一部が喪失されうることには注意すべきである。このように、単なる培養中のそのような細胞は、分化した細胞という用語には含まれず、これらの細胞は非分化細胞（たとえば未分化細胞）または多能性細胞とはならない。分化した細胞の多能性への移行は、培養中の分化特徴の部分的喪失に至る刺激を超えるリプログラミング刺激を必要とする。リプログラムされた細胞はまた、一般的に培養において限られた回数 of 分裂能を有する初代親細胞と比較して、成長能を失うことなく、長期間継代することができるという特徴を有する。いくつかの態様において、「分化した細胞」という用語はまた、細胞分化プロセスを受けている、より特殊でない細胞型の細胞に由来する（たとえば、未分化細胞またはリプログラム細胞に由来する）より特殊化された細胞型の細胞を意味する。

10

20

【0095】

本明細書において用いられる「体細胞」という用語は、生殖細胞以外、着床前胚に存在するもしくは着床前胚から得られる細胞以外、およびインビトロでそのような細胞の増殖に起因する細胞以外の、任意の細胞を意味する。別の言い方をすれば、体細胞は、生殖系列細胞に対して、生物の体を形成する任意の細胞を意味する。哺乳動物において、生殖系列細胞（「接合子」としても知られる）は精子と卵子であり、これらは受精の際に融合して接合体と呼ばれる細胞を産生し、そこから完全な哺乳動物胚が発達する。哺乳動物の体におけるあらゆる他の細胞型は、それらが作製される細胞（生殖母細胞）である精子および卵子ならびに未分化幹細胞を別として、体細胞であり、内臓、皮膚、骨、血液、および結合組織は全て体細胞で構成される。いくつかの態様において、体細胞は、「非胚体細胞」であり、これは胚に存在せずまたは胚から得られず、インビトロでそのような細胞の増殖に起因しない体細胞を意味する。いくつかの態様において、体細胞は「成体体細胞」であり、これは胚または胎児以外の生物に存在するもしくは生物から得られる、またはインビトロでそのような細胞の増殖に起因する細胞を意味する。それ以外であると明記している場合を除き、分化した細胞をリプログラムする方法は、インビボおよびインビトロの両方で行うことができる（インビボは、分化した細胞が対象に存在する場合に実行され、インビトロは、培養で維持される単離された分化した細胞を用いて実行される）。いくつかの態様において、分化した細胞または分化した細胞集団をインビトロで培養する場合、分化した細胞は、その全内容が参照により本明細書に組み入れられる、meneghel- Rozzo et al., (2004), Cell Tissue Res, 316(3);295-303において記述される培養などの器官切片培養において培養することができる。

30

40

【0096】

本明細書において用いられる「成体細胞」という用語は、胚の発達後に体の至るところで見いだされる細胞を意味する。

【0097】

細胞の個体発生の文脈において、「分化する」または「分化している」という用語は相対的な用語であり、「分化した細胞」は、その前駆細胞より発達経路のさらに下へと進行した細胞を意味する。したがって、いくつかの態様において、本明細書において定義される用語であるリプログラム細胞は、系列限定前駆細胞（中胚葉幹細胞などの）へと分化す

50

ることができ、次に、経路のさらに下の他のタイプの前駆細胞（組織特異的前駆細胞、たとえば心筋細胞前駆細胞などの）へと分化することができ、次に最終段階分化細胞へと分化することができ、この細胞はある組織型において特徴的な役割を果たし、さらなる増殖能を保持してもよく保持しなくてもよい。

【0098】

「胚幹細胞」という用語は、胚盤胞の内細胞塊の多能性幹細胞を意味するために用いられる（参照により本明細書に組み入れられる、米国特許第5,843,780号、第6,200,806号を参照されたい）。そのような細胞はまた、体細胞核移入に由来する胚盤胞の内細胞塊からも同様に得ることができる（たとえば、参照により本明細書に組み入れられる、米国特許第5,945,577号、第5,994,619号、第6,235,970号を参照されたい）。胚幹細胞の注目すべき特徴によって、胚幹細胞表現型が定義される。したがって、細胞は、該細胞を他の細胞と区別することができる胚幹細胞の独自の特徴の1つまたは複数を保有するならば、胚幹細胞の表現型を有する。例示的な注目すべき胚幹細胞特徴には、遺伝子発現プロファイル、増殖能、分化能、核型、および特定の培養条件に対する反応性などが挙げられるがこれらに限定されるわけではない。

10

【0099】

「表現型」という用語は、実際の遺伝子型によらず、環境条件と因子の特定のセットの下で細胞または生物を定義する、1つまたは多数の全体的な生物学的特徴を意味する。

【0100】

「発現」という用語は、RNAおよびタンパク質の産生、ならびに適当であればタンパク質の分泌、例えば、これらに限定されるわけではないが、適用可能である場合には転写、翻訳、フォールディング、修飾、およびプロセッシングなどに関係する、細胞プロセスを意味する。「発現産物」は、遺伝子から転写されたRNA、および、遺伝子から転写されたmRNAの翻訳によって得られたポリペプチドを含む。

20

【0101】

「外因性」という用語は、その本来の起源以外の細胞に存在する物質を意味する。本明細書において用いられる場合の「外因性」という用語は、それが通常見いだされないまたはより低い量で見いだされない細胞または生物などの生物系にヒトの手を伴うプロセスによって導入されている核酸（たとえば、sox2転写因子をコードする核酸）またはタンパク質（たとえば、sox2ポリペプチド）を意味する。物質（たとえば、sox2転写因子をコードする核酸またはタンパク質、たとえばsox2ポリペプチド）が細胞または該物質を遺伝する細胞の祖先に導入されたら、これは外因性であると見なされるであろう。これに対し、「内因性」という用語は、生物系または細胞（たとえば、分化した細胞）にとって生得的である物質を意味する。

30

【0102】

本明細書において用いられる「単離された」または「部分的に精製された」という用語は、核酸またはポリペプチドの場合には、その天然の起源において見いだされる核酸またはポリペプチドと共に存在する、および/または細胞によって発現される場合または分泌型ポリペプチドのように分泌される場合に、核酸またはポリペプチドと共に存在するであろう少なくとも1つの他の成分（たとえば、核酸またはポリペプチド）から分離された核酸またはポリペプチドを意味する。化学合成された核酸もしくはポリペプチド、またはインビトロ転写/翻訳を用いて合成された核酸もしくはポリペプチドは、「単離された」と見なされる。

40

【0103】

本明細書において用いられる「単離された細胞」という用語は、当初見いだされる生物またはそのような細胞の子孫から取り出された細胞を意味する。その後細胞は任意で、インビトロで、たとえば他の細胞の存在下で培養されている。任意で、細胞は後に、第二の生物に導入されるか、または該細胞（またはその子孫である細胞）が単離された生物に再導入される。

【0104】

50

本明細書において用いられる単離細胞集団に関して「単離された集団」という用語は、混合または不均一な細胞集団から取り出されて分離されている細胞集団を意味する。いくつかの態様において、単離集団は、そこから細胞が単離または濃縮される不均一な集団と比較して実質的に純粋な細胞集団である。いくつかの態様において単離集団は、リプログラム細胞およびリプログラム細胞が由来する細胞を含む不均一な細胞集団と比較して、リプログラム細胞の実質的に純粋な集団である単離されたリプログラム細胞集団である。

【0105】

特定の細胞集団に関する「実質的に純粋な」という用語は、総細胞集団を構成する細胞に関して少なくとも約75%、好ましくは少なくとも約85%、より好ましくは少なくとも約90%、および最も好ましくは少なくとも約95%純粋である細胞集団を意味する。言い換えれば、「実質的に純粋な」または「本質的に精製された」という用語は、リプログラム細胞集団に関して、本明細書における用語によって定義されるリプログラム細胞またはその子孫ではない細胞を約20%より少なく、より好ましくは約15%、10%、8%、7%より少なく、最も好ましくは約5%、4%、3%、2%、1%より少なく、または1%未満を含む細胞集団を意味する。いくつかの態様において、本発明は、リプログラム細胞集団を拡大するための方法を包含し、拡大されたリプログラム細胞集団は、リプログラム細胞の実質的に純粋な集団である。

10

【0106】

本明細書において用いられる「増殖する」および「増殖」は、細胞分裂によって集団中の細胞数が増加すること（成長）を意味する。細胞増殖は一般的に、増殖因子および他のマイトゲンを含む環境に反応した多数のシグナル伝達経路の協調的活性化に起因すると理解される。細胞増殖はまた、細胞内または細胞外シグナルの作用および細胞増殖を遮断するまたは負の影響を及ぼす機構からの解放によっても促進されうる。

20

【0107】

「濃縮する」または「濃縮された」という用語は、本明細書において互換的に用いられ、1つの種類の細胞の収量（分画）が、開始培養物または調製物における該種類の細胞の分画に対して少なくとも10%増加することを意味する。

【0108】

「再生」または「自己再生」または「増殖」という用語は、本明細書において互換的に用いられ、自身の細胞のより多くのコピー（たとえば、複製）を作製する細胞のプロセスを意味する。いくつかの態様において、リプログラム細胞は、長期間、および/または多数の年月のあいだ、同じ未分化細胞（たとえば、多能性または非特殊化細胞型）へと分裂することによって、自身を再生することができる。いくつかの例において、増殖は、2つの同一の娘細胞への1つの細胞の繰り返し分裂によるリプログラム細胞の拡大を意味する。

30

【0109】

本明細書において言及される「細胞培養培地」（同様に本明細書において「培養培地」または「培地」とも呼ばれる）は、細胞生存率を維持して増殖を支持する栄養を含む細胞を培養するための培地である。細胞培養培地は以下のいずれかを適切な組み合わせで含む：塩、緩衝剤、アミノ酸、グルコースまたは他の糖、抗生物質、血清または血清代替物、およびペプチド増殖因子などの他の成分等。特定の細胞型のために通常用いられる細胞培養培地は、当業者に公知である。

40

【0110】

「細胞株」という用語は、典型的に、1つの祖先細胞に由来している、または定義されたおよび/または実質的に同一の祖先細胞集団に由来している大部分が同一のまたは実質的に同一の細胞集団を意味する。細胞株は、長期間（たとえば、数ヶ月、数年、制限されない期間）培養において維持されていてもよく、維持されることが可能であってもよい。細胞株は、細胞に対して無限の培養寿命を付与する自発的なまたは誘導された形質転換プロセスを受けていてもよい。細胞株は、当技術分野においてそのように認識される全ての細胞株を含む。細胞は、細胞株の個々の細胞の少なくともいくつかの特性が互いに関して

50

異なりうるように、時間と共に変異およびおそらくエピジェネティックな変化を獲得すると認識されるであろう。

【0111】

本明細書において用いられる「系列」という用語は、共通の祖先を有する細胞または共通の発達の運命を有する細胞を記述する。一例であるに過ぎないが、内胚葉起源の細胞または「内胚葉系列」である細胞は、内胚葉細胞に由来して、最終的な内胚葉細胞を生じる1つまたは複数の発達系列経路などの、内胚葉系列限定経路に沿って分化することができ、次に肝細胞、胸腺、膵臓、肺、および腸管へと分化することができる細胞を意味する。

【0112】

「減少」、「低減された」、「低減」、「減少する」、または「阻害する」という用語は全て、統計学的に有意な量の減少を意味するために本明細書において一般的に用いられる。しかし、紛らわしさを回避するために、「低減された」、「低減」、または「減少する」または「阻害する」は、参照レベルと比較して少なくとも10%の減少、たとえば少なくとも約20%、または少なくとも約30%、または少なくとも約40%、または少なくとも約50%、または少なくとも約60%、または少なくとも約70%、または少なくとも約80%、または少なくとも約90%、または100%までおよび100%を含む減少（たとえば、参照試料と比較して存在しないレベル）、または参照レベルと比較して10~100%のあいだの任意の減少を意味する。

10

【0113】

「増加した」、「増加する」、または「増強する」、または「活性化する」という用語は全て、統計学的に有意な量の増加を意味するために本明細書において一般的に用いられる；いかなる紛らわしさも回避するために、「増加した」、「増加する」、または「増強する」、または「活性化する」という用語は、参照レベルと比較して少なくとも10%の増加、たとえば少なくとも約20%、または少なくとも約30%、または少なくとも約40%、または少なくとも約50%、または少なくとも約60%、または少なくとも約70%、または少なくとも約80%、または少なくとも約90%、または100%までおよび100%を含む増加、または参照レベルと比較して10~100%のあいだの任意の増加、または参照レベルと比較して少なくとも約2倍、または少なくとも約3倍、または少なくとも約4倍、または少なくとも約5倍、または少なくとも約10倍の増加、または2倍から10倍のあいだ、またはそれ以上の任意の増加を意味する。

20

30

【0114】

「統計学的に有意な」または「有意に」という用語は、統計学的有意性を意味し、一般的に、マーカーの通常濃度またはそれ未満の濃度を2標準偏差（2SD）下回ることを意味する。この用語は、差が存在する統計学的証拠を意味する。これは、帰無仮説が実際に真実である場合に帰無仮説を棄却する決定を下す確率として定義される。決定はしばしば、p-値を用いて行われる。

【0115】

本明細書において用いられる「DNA」という用語は、デオキシリボ核酸として定義される。

【0116】

本明細書において用いられる「分化」という用語は、始原段階からより成熟した（すなわちより始原的でない）細胞に向かう細胞の細胞発達を意味する。

40

【0117】

本明細書において用いられる「定方向性分化」という用語は、遺伝子および/または環境操作を介して、未分化（たとえば、より始原的な細胞）からより成熟した細胞型（すなわちより始原的でない細胞）へと細胞を強制的に分化させることを意味する。いくつかの態様において、本明細書において開示されるリプログラム細胞は、ニューロン細胞型、および筋細胞型などの特定細胞型への定方向性分化に供される。

【0118】

本明細書において用いられる「機能的アッセイ」という用語は、ある状況下での細胞の

50

成長または生存能を評価することによって、細胞の遺伝子発現または発達状態などの細胞の特性を評価する試験である。いくつかの態様において、リプログラム細胞は、リプログラム細胞が本明細書において開示される多能性状態であるか否かを判定するために機能的アッセイによって同定することができる。

【0119】

本明細書において用いられる「疾患モデリング」という用語は、ヒト疾患または病気に関する新規情報を得るために実験的細胞培養または動物研究を用いることを意味する。いくつかの態様において、本明細書において開示される方法によって産生されるリプログラム細胞は、疾患モデリング実験において用いることができる。

【0120】

本明細書において用いられる「薬物スクリーニング」という用語は、特定機能を有する薬物を同定するために研究室において細胞および組織を用いることを意味する。いくつかの態様において、本発明は、分化した細胞をリプログラム細胞（たとえば、本明細書において開示される、多能性状態であるリプログラム細胞、または安定な中間体であるリプログラム細胞、部分的リプログラム細胞）へとリプログラムする化合物または薬物を同定するための、分化した細胞の薬物スクリーニング法を提供する。いくつかの態様において、本発明は、分化した細胞を完全にリプログラムされた細胞（たとえば、多能性状態にあるリプログラム細胞）へとリプログラムする化合物または薬物を同定するための、安定な中間的部分的リプログラム細胞の薬物スクリーニング法を提供する。代替の態様において、本発明は、疾患または病気（たとえば、ヒトの疾患または病気）のための治療として有用である化合物または薬物を同定するためのリプログラム細胞（たとえば、ヒトリプログラム細胞）に対する薬物スクリーニングを提供する。

【0121】

本明細書において用いられる「マーカー」は、細胞の特徴および/または表現型を記述するために用いられる。マーカーは、関心対象の特徴を含む細胞の選択のために用いることができる。マーカーは、特定細胞型と共に変化するであろう。マーカーは、特徴、すなわち特定の細胞型の細胞またはその細胞型によって発現される分子の形態学的、機能的、または生化学的（酵素的）特徴である。好ましくは、そのようなマーカーはタンパク質であり、およびより好ましくは当技術分野において入手可能な抗体のエピトープまたは他の結合分子を保有する。しかし、マーカーは、タンパク質（ペプチドおよびポリペプチド）、脂質、多糖類、核酸およびステロイドを含むがこれらに限定されるわけではない、細胞において見いだされる任意の分子からなりうる。形態学的特徴または形質の例には、形状、大きさ、および核対細胞質比が挙げられるがこれらに限定されるわけではない。機能的特徴または形質の例には、特定の基質に対する接着能、特定の色素の取り込み能または排除能、特定の条件下での遊走能、および特定の系列に沿った分化能が挙げられるがこれらに限定されるわけではない。マーカーは、当業者に入手可能な任意の方法によって検出されうる。マーカーはまた、形態学的特徴が存在しないこと、またはタンパク質、脂質等が存在しないことでありうる。マーカーは、ポリペプチドおよび他の形態学的特徴の有無に関する独自の特徴のパネルの組み合わせでありうる。

【0122】

「選択可能なマーカー」という用語は、発現されると、細胞障害剤または細胞抑制剤に対する耐性（たとえば、抗生物質耐性）、栄養原栄養性、またはタンパク質を発現する細胞と発現しない細胞とを区別するための根拠として用いることができる特定のタンパク質の発現などの選択可能な表現型を細胞に付与する遺伝子、RNA、またはタンパク質を意味する。蛍光または発光タンパク質などの、その発現を容易に検出することができるタンパク質、または基質に作用して着色、蛍光、または発光物質を生じる酵素（「検出可能マーカー」）は、選択可能マーカーのサブセットを構成する。多能性幹細胞において通常、選択的にまたは独占的に発現される遺伝子にとって生得的である発現制御エレメントに連結させた選択可能マーカーが存在することにより、多能性状態へとリプログラムされている体細胞を同定および選択することができる。ネオマイシン耐性遺伝子（neo）、ピューロ

10

20

30

40

50

マイシン耐性遺伝子 (puro)、グアニンホスホリボシルトランスフェラーゼ (gpt)、ジヒドロ葉酸レダクターゼ (DHFR)、アデノシンデアミナーゼ (ada)、ピューロマイシン-N-アセチルトランスフェラーゼ (PAC)、ヒグロマイシン耐性遺伝子 (hyg)、多剤耐性遺伝子 (mdr)、チミジンキナーゼ (TK)、ヒポキサンチン-グアニンホスホリボシルトランスフェラーゼ (HPRT)、および hisD 遺伝子などの多様な選択可能マーカー遺伝子を用いることができる。検出可能なマーカーには、緑色蛍光タンパク質 (GFP) ブルー、サファイア、イエロー、レッド、オレンジ、およびシアン蛍光タンパク質、およびこれらの任意の変種が挙げられる。ルシフェラーゼ (たとえば、ホタルまたはウミシイタケルシフェラーゼ) などの発光タンパク質も同様に有用である。当業者に明らかであるように、本明細書において用いられる「選択可能マーカー」という用語は、遺伝子または遺伝子の発現産物、たとえばコードされたタンパク質を意味することができる。

10

【0123】

いくつかの態様において、選択可能マーカーは、それを発現しない細胞または有意により低いレベルで発現する細胞と比較して、それを発現する細胞に増殖および/または生存上の利点を付与する。そのような増殖および/または生存上の利点は典型的に、細胞がある条件下、たとえば「選択的条件」下で維持される場合に起こる。有効な選択を確実にするために、マーカーを発現しない細胞が増殖および/または生存せず、集団から消失するか、またはその数が集団の非常に小さい分画にまで低減される条件下で十分な期間、細胞集団を維持することができる。マーカーを発現しない細胞をほぼまたは完全に消失させるために、選択的条件下で細胞集団を維持することによって、増殖および/または生存上の利点を付与するマーカーを発現する細胞を選択するプロセスは、本明細書において「正の選択」と呼ばれ、マーカーは「正の選択にとって有用である」と言われる。負の選択および負の選択にとって有用なマーカーも同様に、本明細書において記述されるある種の方法において重要である。そのようなマーカーの発現は、マーカーを発現しないまたは有意に低いレベルでマーカーを発現する細胞と比較して、マーカーを発現する細胞に対して増殖および/または生存を不利にする (または別の考え方をすれば、マーカーを発現しない細胞は、マーカーを発現する細胞と比較して増殖および/または生存上の利点を有する)。それゆえ、マーカーを発現する細胞を、選択的条件下で十分な期間維持すると、細胞集団からほぼまたは完全に消失させることができる。

20

【0124】

本明細書において用いられる「処置する」および「処置」という用語は、対象において疾患の少なくとも1つの症状の低減、または疾患の改善、たとえば有益なまたは所望の臨床結果が得られるような組成物の有効量を、対象に投与することを意味する。本発明の目的に関して、有益なまたは所望の臨床結果は、検出可能であるか検出不能であるかによらず、1つまたは複数の症状の軽減、疾患の程度の縮小、疾患の安定化 (たとえば、悪化していない) 状態、疾患進行の遅延または鈍化、疾患状態の改善または緩和、および寛解 (部分的または完全によらず) を含むがこれらに限定されるわけではない。いくつかの態様において、処置するとは、処置を受けない場合に予想される生存と比較した、生存の延長を意味することができる。このように、当業者は、処置により疾患状態が改善されうるが疾患が完全に治癒するわけではないことを認識する。本明細書において用いられる「処置」という用語は、予防を含む。または処置は、疾患の進行が低減されるかまたは停止する場合に「有効」である。いくつかの態様において、「処置」という用語はまた、処置を受けない場合に予想される生存と比較した、生存の延長を意味することができる。処置を必要とする対象は、疾患または状態を有すると既に診断された対象、ならびに遺伝的感受性、または疾患もしくは状態に関与する他の要因により疾患または状態を発症する可能性がある対象を含み、非制限的な例として、対象の体重、食事、および健康は、糖尿病を発症する可能性がある対象に寄与しうる要因である。処置を必要とする対象は、内科的または外科的な注意、ケア、または管理を必要とする対象を含む。対象は通常、病気であるかもしくは損傷を有し、または集団の平均的なメンバーと比較して病気に罹るリスクが増加しており、かつそのような注意、ケア、または管理を必要とする。

30

40

50

【0125】

本明細書において用いられる「投与する」、「導入する」、および「移植する」という用語は、リプログラム細胞またはその分化した子孫の所望の部位での少なくとも部分的局在が起こる方法または経路によって、本明細書において開示されるリプログラム細胞またはその分化した子孫を対象に留置するという文脈において互換的に用いられる。リプログラム細胞またはその分化子孫は、関心対象組織に直接投与することができ、またはリプログラム細胞もしくはその子孫の少なくとも一部もしくは細胞の成分が生存したままで対象における所望の位置への送達を得られる任意の適切な経路によって、投与することができる。対象に投与した後のリプログラム細胞の生存期間は、数時間、たとえば24時間程度の短さから数日、数年程度の長さでありうる。

10

【0126】

本明細書において用いられる「移植」という用語は、宿主（すなわち、移植レシピエントまたは移植対象）への新しい細胞（たとえば、リプログラム細胞）、組織（リプログラム細胞から生じた分化した細胞などの）、または器官の導入を意味する。

【0127】

「コンピュータ」という用語は、構造化入力を受けることができ、定法に従って構造化入力を処理することができ、および処理の結果を出力として生じることができ、任意の非ヒト装置を意味し得る。コンピュータの例には：コンピュータ；汎用コンピュータ；スーパーコンピュータ；メインフレーム；スーパーミニコンピュータ；ミニコンピュータ；ワークステーション；マイクロコンピュータ；サーバー；インタラクティブテレビ；コンピュータとインタラクティブテレビのハイブリッド組み合わせ；ならびに、コンピュータおよび/またはソフトウェアを模倣するためのアプリケーション特異的ハードウェアが挙げられる。コンピュータは、1つのプロセッサを、または同時に作動することができ、かつ/もしくは同時に作動しない多数のプロセッサを有することができる。コンピュータはまた、コンピュータ間で情報を伝達または受信するためのネットワークを介して共に接続された2つまたはそれより多くのコンピュータを意味する。そのようなコンピュータの例には、ネットワークによって連結されたコンピュータを介して情報を処理するための分散型コンピュータシステムが挙げられる。

20

【0128】

「コンピュータ読み取り可能な媒体」という用語は、コンピュータによってアクセス可能なデータを保存するために用いられる任意の保存デバイスならびにコンピュータによるデータへのアクセスを提供するための他の任意の手段を意味しうる。保存デバイス型のコンピュータ読み取り可能な媒体の例には：磁気ハードディスク；フロッピーディスク；CD-ROMおよびDVDなどの光学ディスク；磁気テープ；メモリーチップが挙げられる。

30

【0129】

「ソフトウェア」という用語は、本明細書において「プログラム」と互換的に用いられ、コンピュータを操作するための定法を意味する。ソフトウェアの例には：ソフトウェア；コードセグメント；命令；コンピュータプログラム；およびプログラムロジックが挙げられる。

【0130】

「コンピュータシステム」という用語は、コンピュータを操作するためにソフトウェアを具現化するコンピュータ読み取り可能な媒体を含むコンピュータを有するシステムを意味しうる。

40

【0131】

「プロテオミクス」という用語は、細胞内のタンパク質が作用して互いに相互作用する様式を含み、遺伝子発現のゲノム分析とは異なる情報を提供する、細胞内のタンパク質の発現、構造、および機能に関する研究を意味しうる。

【0132】

本明細書において用いられる「含む (comprising)」または「含む (comprises)」という用語は、本発明にとって必須であるものの、（本質的であるか否かによらず）明記さ

50

れていない要素を含めることを容認する、組成物、方法、およびその各々の成分に関して用いられる。

【0133】

本明細書において用いられる「から本質的になる」という用語は、所定の態様にとって必要な要素を意味する。この用語は、本発明のその態様の基本的小よび新規または機能的な特徴に著しい影響を及ぼさない追加の要素の存在を許容する。

【0134】

「からなる」という用語は、態様のその説明において引用されていないいかなる要素も除外する、本明細書において記述される組成物、方法、およびその各々の成分を意味する。

10

【0135】

本明細書および添付の特許請求の範囲において用いられる単数形「ある」、「1つの」、および「その」は、文脈上特記される場合を除き、複数形を含む。このように、たとえば、「その方法」という言及は、本明細書において記述されるおよび/または本開示を読むことにより当業者に明らかとなるであろうタイプの1つまたは複数の方法、および/または段階を含む。

【0136】

実施例または特記される場合を除き、本明細書において用いられた成分または反応条件の量を表現する全ての数値は、全ての例において「約」という用語によって修飾されると理解すべきである。百分率に関連して用いる場合の「約」という用語は、 $\pm 1\%$ を意味する。本発明はさらに、実施例を含み以下に詳細に説明するが、本発明の範囲はそれらに限定されるべきではない。

20

【0137】

前述の詳細な説明および以下の実施例は、単なる説明であり、本発明の範囲に対する制限と解釈されるべきではないと理解される。開示の態様に対する様々な変化および修飾が当業者に明らかとなるであろうが、それらも本発明の趣旨および範囲に含まれる。さらに、同定された全ての特許、特許出願、および刊行物は、たとえば本発明に関連して用いられるそのような刊行物において記述される方法論を記述および開示する目的のために明白に参照により本明細書に組み入れられる。これらの刊行物は、単にその開示が本出願の提出日以前になされたために提供される。この点においていかなるものも、本発明者らが、先行発明に基づいてまたは任意の他の理由からそのような開示に先行する権利がないのを自認したと解釈すべきではない。これらの文書の日付または内容に関する表現に関する全ての声明は、出願人が入手可能な情報に基づいており、これらの文書の日付または内容の正確性に関していかなる承認も構成するものではない。

30

【0138】

全般

本発明の1つの局面は、多能性幹細胞株を特徴決定するための2つのスコアカードを作成するための方法、システム、およびアッセイに関し、第一のスコアカードは、「偏差スコアカード」または「多能性スコアカード」と呼ぶことができ、これは関心対象の多能性幹細胞株が既に確立されたまたは対照の多能性幹細胞株にどれほど匹敵するかに関する情報を提供するために有用であり、1つの参照多能性幹細胞株および/または複数の参照多能性幹細胞株と比較して、DNAメチル化または遺伝子発現に関して外れている遺伝子の数または%を同定するために用いることができる。そのようなスコアカードは、関心対象の幹細胞株の多能性を同定するために有用であると共に、関心対象の幹細胞株を後の時点で異常増殖および癌形成に罹りやすくする可能性がある癌遺伝子の異型の遺伝子発現またはDNAメチル化を関心対象の幹細胞株が有するか否かを同定するために、有用である。第二のスコアカードは、本明細書において「系列スコアカード」と呼ばれ、これは関心対象の多能性幹細胞の分化能の定量として有用であり、関心対象の多能性幹細胞株が、既に確立されたまたは対照の多能性幹細胞株と比較して関心対象の特定の系列へとどのように効率的に分化するかに関する情報を提供する。「要約スコアカード」は、関心対象の1つまたは

40

50

複数の多能性幹細胞株の偏差スコアカードと系列スコアカードとを含むことができる。

【0139】

したがって、本発明のさらなる局面は、多能性幹細胞の特徴を同定することができるように、およびどの多能性幹細胞株が幹細胞起源の癌に寄与する可能性があるかを予測するために、(i) 特定遺伝子、たとえば腫瘍遺伝子、腫瘍抑制遺伝子および発達遺伝子のプロモーターメチル化による特定遺伝子のエピジェネティックなサイレンシングの同定、(ii) たとえば発達遺伝子および系列マーカー遺伝子の遺伝子発現の同定、ならびに(iii) 異なる系列に沿って分化する分化指向性から選択される少なくとも2つのデータセットをモニターすることによって多能性幹細胞株のスコアカードを作成する段階を含む、多能性幹細胞集団を検証および/またはモニターする方法を提供する。

10

【0140】

いくつかの態様において、たとえば、所定の細胞株に関する分化指向性(系列マーカー遺伝子の差次的に修飾されるメチル化および/または差次的遺伝子発現を用いて)を決定した後、参照または「標準的な」多能性幹細胞株と比較した、標的遺伝子のDNAメチル化(たとえば、表12Aおよび/または表12C、表13A、表13B、または表14のいずれかに記載される遺伝子のいくつかまたは組み合わせ)の変化の決定、および/または標的遺伝子の遺伝子発現レベル(たとえば、表12Bおよび/または表12Cのいずれかに記載される、または表13A、表13B、または表14から選択される遺伝子のいくつかまたは組み合わせ)の変化の決定についての品質の判定を行うことができる。

【0141】

本明細書において述べるように、スコアカードは、以下のいくつかの成分:(i) 参照多能性細胞株における標的遺伝子のDNAメチル化の正常な変動と比較した多能性細胞におけるDNAメチル化遺伝子外れ値の同定、(ii) 参照多能性細胞株における標的遺伝子のDNA発現レベルの正常な変動と比較した多能性細胞株における遺伝子発現外れ値の同定、(iii) (i) および(ii) からのDNAメチル化および/もしくは遺伝子発現データならびに/または分化誘導されている多能性細胞株由来の遺伝子発現/DNAメチル化データに基づく、細胞分化の偏りの予測、を含む。

20

【0142】

本発明は、様々なタイプの多能性幹細胞および前駆細胞(たとえば、ES細胞、体性幹細胞、造血幹細胞、白血病幹細胞、皮膚幹細胞、腸管幹細胞、性腺幹細胞、脳幹細胞、筋幹細胞(筋芽細胞等)、乳腺幹細胞、神経幹細胞(たとえば、小脳顆粒ニューロン前駆細胞等)等)について、ならびに様々な幹細胞または前駆細胞(たとえば、参照により本明細書に組み入れられるSparmann & Lohuizen, Nature 6, 2006 (Nature Reviews Cancer, November 2006)の表1に記述される細胞など)について、ならびにインビトロおよびインビボ由来幹細胞、例えば人工多能性幹細胞(iPSC)、ならびに最終的に分化した細胞の品質および有用性の決定に関して、実質的有用性を有する。

30

【0143】

本発明のいくつかの局面において、本発明は、多能性幹細胞の特徴の同定を可能にするためおよびどの多能性幹細胞株が幹細胞起源の癌に寄与する可能性があるかを予測するために、(i) 特定遺伝子、たとえば腫瘍遺伝子、腫瘍抑制遺伝子、および発達遺伝子のプロモーターメチル化による特定遺伝子のエピジェネティックなサイレンシングの同定、(ii) たとえば発達遺伝子および系列マーカー遺伝子の遺伝子発現の同定、ならびに(iii) 異なる系列に沿って分化する分化指向性から選択される少なくとも2つのデータセットをモニターすることによる、多能性幹細胞株を検証およびモニターするためのならびに多能性幹細胞株の全般的品質管理として役立つ、多能性幹細胞株のスコアカードの作成に関する。

40

【0144】

いくつかの態様において、本発明は、以下の段階を含む、多能性幹細胞株を選択するための方法を提供する:(i) 少なくとも1つの多能性幹細胞に、DNAにおけるエピジェネティックな修飾に差次的に結合する剤を接触させることによって、多能性幹細胞株における

50

標的遺伝子セットのエピジェネティックな修飾を測定して、該エピジェネティックな修飾データと、同じ標的遺伝子の参照のエピジェネティックな修飾データとの比較を行う段階；(ii)多能性幹細胞の非定方向性または定方向性分化により多能性幹細胞株の分化能を測定して、複数の系列マーカー遺伝子の遺伝子発現レベルを検出することができるように転写物を標識する段階；および分化能データを参照分化能データと比較する段階；ならびに(iii)参照のエピジェネティックな修飾レベルと比較した標的遺伝子のDNAのエピジェネティックな修飾に統計学的に有意な量で差がなく、かつ参照分化能と比較して中胚葉、外胚葉、および内胚葉系列に沿って分化する指向性に統計学的に有意な量で差がない多能性幹細胞株を選択する段階；または、参照のエピジェネティックな修飾レベルと比較して標的遺伝子のエピジェネティックな修飾に統計学的に有意な量で差があり、かつ参照分化能と比較して中胚葉、外胚葉、および内胚葉系列に沿って分化する指向性に統計学的に有意な量で差がある多能性幹細胞株を廃棄する段階。

10

【0145】

いくつかの態様において、エピジェネティックな修飾は、多能性幹細胞株における標的遺伝子セットにおいてエピジェネティックな修飾を測定する段階を含み、たとえばエピジェネティックな修飾は以下からなる群より選択される以下のいずれか1つによって測定することができる：濃縮に基づく方法（たとえば、MeDIP、MBD-seqおよびMethylCap）、バイサルファイトシーケンシングおよびバイサルファイトに基づく方法（たとえば、RRBS、バイサルファイトシーケンシング、Infinium、GoldenGate、COBRA、MSP、MethylLight）、ならびに制限消化法（たとえば、MRE-seq）、または同じ標的遺伝子の参照DNAメチル化データと比較した多能性幹細胞のDNAメチル化標的遺伝子の差次的変換、差次的制限、差次的重量。

20

【0146】

いくつかの態様において、方法はさらに、(iv)多能性幹細胞株における第二の標的遺伝子セットの遺伝子発現を測定して、該遺伝子発現データと、同じ標的遺伝子の参照遺伝子発現レベルとの比較を行う段階；ならびに(v)参照遺伝子発現レベルと比較して標的遺伝子の遺伝子発現レベルに統計学的に有意な量で差がない多能性幹細胞株を選択する段階、または参照遺伝子発現レベルと比較して標的遺伝子の発現レベルに統計学的に有意な量で差がある多能性幹細胞株を廃棄する段階を含む。

30

【0147】

いくつかの態様において、参照DNAメチル化レベルは、前記DNAメチル化標的遺伝子のメチル化の正常な変動範囲であり、いくつかの例において前記DNAメチル化標的遺伝子に関するDNAメチル化の平均値±任意で1標準偏差でありえて、平均値は、複数の多能性幹細胞株における、たとえば少なくとも5個またはそれより多くの多能性幹細胞株におけるその標的遺伝子のDNAメチル化から算出される。

【0148】

いくつかの態様において、参照遺伝子発現レベルはその標的遺伝子に関する正常な変動の範囲であり、いくつかの態様において、これはその標的遺伝子の発現レベルの平均値であり、平均値は、複数の多能性幹細胞株、たとえば少なくとも5個またはそれより多くの異なる多能性幹細胞株におけるその標的遺伝子の発現レベルから算出される。

40

【0149】

いくつかの態様において、遺伝子発現は、定量的分化アッセイなどのマイクロアレイアッセイによって決定される。

【0150】

いくつかの態様において、参照分化能は、中胚葉、内胚葉、外胚葉、ニューロン、造血系列、およびこれらの任意の組み合わせからなる群より選択される系列への分化能であり、参照分化能データは、複数の多能性幹細胞株、たとえば少なくとも5個の異なる多能性幹細胞株から作成される。いくつかの態様において、試験多能性幹細胞および/または参照多能性幹細胞の分化能は、多能性幹細胞を分化させて（既定の期間での定方向性分化または自発的分化のいずれか）、DNAメチル化および/または遺伝子発現の差を決定するこ

50

とによって決定される。

【0151】

本発明の全ての局面のいくつかの態様において、DNAメチル化標的遺伝子および/または参照DNAメチル化標的遺伝子は、癌遺伝子、腫瘍遺伝子、腫瘍抑制遺伝子、発達遺伝子、系列マーカー遺伝子、およびこれらの任意の組み合わせからなる群より選択され、表12Aに記載される群から選択される、または表13A、表13B、または表14およびこれらの任意の組み合わせから選択されるDNAメチル化標的遺伝子および/または参照DNAメチル化標的遺伝子を含む。いくつかの態様において、腫瘍抑制遺伝子は、c-Sis、上皮細胞増殖因子受容体、血小板由来増殖因子受容体、血管内皮増殖因子受容体、HER2/new、チロシンキナーゼのSrcファミリー、チロシンキナーゼのSyk-Zap-70ファミリー、チロシンキナーゼのBTKファミリー、Rafキナーゼ、サイクリン依存的キナーゼ、Rasタンパク質、およびmyc遺伝子から選択される。いくつかの態様において、腫瘍抑制遺伝子は、TP53、PTEN、APC、CD95、ST5、ST7、およびST14遺伝子から選択される。いくつかの態様において、発達遺伝子は、表7に記載される遺伝子の任意の組み合わせから選択される。いくつかの態様において、系列マーカー遺伝子は、VEGF受容体II (KDR)、アクチン-2平滑筋 (ACTA2)、ネスチン、チューブリン 3、 α -フェトプロテイン (AFP)、シンデカン-4、CD64IFc γ RI、Oct-4、 β -HCG、 β -LH、oct-3、Brachyury T、Fgf-5、nodal、GATA-4、flk-1、Nkx-2.5、EKLF、およびMsx3から選択される。いくつかの態様において、DNAメチル化標的遺伝子および/または参照DNAメチル化標的遺伝子は、BMP4、CAT、CD14、CXCL5、DAZL、DNMT3B、GATA6、GAPDH、LEFTY2、MEG3、PAX6、S100A6、SOX2、SNAI1、TF、およびこれらの任意の組み合わせからなる群より選択される。いくつかの態様において、表12Aのリストにおける遺伝子の任意の組み合わせから選択される、または表13A、表13B、もしくは表14から選択される少なくとも約200個の標的遺伝子のDNAメチル化が多能性細胞株において測定され、少なくとも200個の同じ標的遺伝子セットの参照DNAメチル化レベルと比較されるが、少なくとも200個の標的遺伝子は、表12Aのリストにおける遺伝子の任意の組み合わせから選択される、または表13A~13B、もしくは表14から選択される少なくとも約200個の標的遺伝子でありえて、または表12Aに記載される、もしくは表13A、表13B、もしくは表14から選択される1~500番の遺伝子の任意の組み合わせから選択される少なくとも約200個の標的遺伝子でありえて、または表12Aに記載される、もしくは表13A、表13B、もしくは表14から選択される1~200番から選択される少なくとも約200個の標的遺伝子でありうる。いくつかの態様において、表12Aのリストにおける遺伝子の任意の組み合わせから選択される少なくとも約500個の標的遺伝子のDNAメチル化を、多能性細胞株において測定して、同じ少なくとも500個の標的遺伝子セットの参照DNAメチル化レベルと比較する。いくつかの態様において、表12Aのリストにおける遺伝子の任意の組み合わせから選択される、または表13A、表13Bもしくは表14から選択される少なくとも約500個の標的遺伝子のDNAメチル化は、表12Aに記載される、または表13A、表13B、もしくは表14から選択される1~1000番の遺伝子の任意の組み合わせから選択される。

【0152】

本発明の全ての局面のいくつかの態様において、遺伝子発現標的遺伝子および/または参照遺伝子発現標的遺伝子は、表12Bに記載される群から、または表13A、表13Bもしくは表14、およびこれらの任意の組み合わせから選択され、たとえば少なくとも約200個または少なくとも約500個の標的遺伝子は、表12Aに記載される1~500番から選択され、または表12Aのリストにおける遺伝子の任意の組み合わせから選択される、もしくは表13A、表13Bもしくは表14から選択される少なくとも約1000個の標的遺伝子から選択され、または少なくとも約1000個の標的遺伝子は、表13A、表13B、または表14Aに記載される、またはそれらから選択される1~2000番から選択される。

【0153】

いくつかの態様において、参照遺伝子と比較してメチル化に統計学的有意差を有する多能性幹細胞株における多数のDNAメチル化遺伝子は、10、9、8、7、6、5、4、3、2、1、または0個である。いくつかの態様において、参照遺伝子と比較して遺伝子発現レベルに統

計学的有意差を有する多能性幹細胞株における多数の遺伝子は、10、9、8、7、6、5、4、3、2、1、または0個である。

【0154】

いくつかの態様において、多能性幹細胞は、ヒト多能性幹細胞などの哺乳動物多能性幹細胞である。

【0155】

本発明のもう1つの局面は、生物活性に関して化合物をスクリーニングするための多能性幹細胞株の使用に関する。たとえば、そのような態様は、(i) 任意で特定系列に沿って多能性幹細胞を分化させるか、または分化することを可能にすること；(ii) 細胞に試験化合物を接触させること；および(iii) 該細胞に及ぼす該化合物の任意の効果を決定

10

【0156】

いくつかの態様において、化合物は、有機低分子、無機低分子、多糖類、ペプチド、タンパク質、核酸、細菌、植物、真菌、動物細胞、動物組織などの生物材料から作製された抽出物、およびその任意の混合物からなる群より選択され、約0.01 nMから約1000 mMの範囲の濃度で用いることができる。いくつかの態様において、スクリーニングは、ハイスループットスクリーニング法である。いくつかの態様において、生物活性は、生物アッセイにおける刺激、阻害、調節、毒性、電気刺激、または致死反応の誘発である。いくつかの態様において、生物活性は、酵素活性の調節、受容体の不活化、受容体の刺激、1つまたは複数の遺伝子の発現レベルの調節、細胞増殖の調節、細胞分裂の調節、細胞形態の調節

20

【0157】

本発明のもう1つの局面は、再生医療および細胞補充/増強治療などのために、哺乳動物対象、たとえばマウスもしくは齧歯類動物モデルまたはヒト対象を処置するために、対象に多能性幹細胞を投与することによって、対象を処置するために本明細書において開示される方法およびスコアカードを用いて検証および特徴決定される多能性幹細胞の使用に関する。いくつかの態様において、対象は、癌、糖尿病、心不全、筋障害、セリアック病、神経障害、神経変性障害、リソソーム蓄積症、およびこれらの任意の組み合わせからなる群より選択される疾患または状態を患っているかまたは有すると診断されている。いくつかの態様において、多能性幹細胞は局所に投与される、または投与は対象への多能性幹細胞の移植である。

30

【0158】

いくつかの態様において、多能性幹細胞を、対象に該多能性幹細胞またはその分化した子孫を投与する前に分化させる、たとえば、中胚葉、内胚葉、外胚葉、ニューロン、造血系列およびこれらの任意の組み合わせからなる群より選択される系列に沿って分化させるか、またはインスリン産生細胞(膵細胞、細胞等)、ニューロン細胞、筋細胞、皮膚細胞、心筋細胞、肝細胞、血液細胞、適応免疫細胞、および先天免疫細胞などへと分化させる。

40

【0159】

本発明のもう1つの局面は、本明細書において開示される方法、アッセイおよびスコアカードを用いることによって選択される多能性幹細胞を含むキットに関する。キットはさらに使用説明書を含む。

【0160】

本発明のもう1つの局面は、以下の少なくとも2つを含む、多能性細胞の複数の特性の特徴決定を行うためのアッセイに関する：(i) DNAメチル化アッセイ；(ii) 遺伝子発現アッセイ；および(iii) 分化アッセイ。いくつかの態様において、アッセイはキットの形状でありうる。いくつかの態様において、アッセイは研究者またはサービス提供者によって行われる。いくつかの態様において、アッセイは、本明細書において開示される方法に

50

従って多能性幹細胞株を検証および/または特徴決定するためにスコアカードの形式でレポートを提供する。

【0161】

いくつかの態様において、アッセイは、パイサルファイトシーケンシングアッセイもしくは全ゲノムパイサルファイトシーケンシングアッセイであるDNAメチル化アッセイを含み、または濃縮に基づく方法（たとえば、MeDIP、MBD-seqおよびMethylCap）、パイサルファイトシーケンシングおよびパイサルファイトに基づく方法（たとえば、RRBS、パイサルファイトシーケンシング、Infinium、GoldenGate、COBRA、MSP、MethylLight）、ならびに制限消化法（たとえば、MRE-seq）からなる群より選択される任意のDNAメチル化アッセイでありうる。

10

【0162】

いくつかの態様において、アッセイは、マイクロアレイアッセイ、たとえば定量的分化アッセイである遺伝子発現アッセイを含む。いくつかの態様において、アッセイは、以下の系列、すなわち中胚葉、内胚葉、外胚葉、ニューロン、または造血系列のうち少なくとも1つへの多能性細胞の分化能を評価する分化アッセイを含み、特定の系列への多能性細胞の分化能は、本明細書において開示されるDNAメチル化アッセイおよび/または遺伝子発現アッセイによって決定され、または中胚葉、内胚葉、および外胚葉系列に関する少なくとも1つのマーカーに対する抗体を用いる免疫染色もしくはFACSソーティングによって決定される。いくつかの態様において、多能性細胞の特定系列への分化能は、EB培養の少なくとも約0日後、たとえば約0~3日、または約3~7日、または約7~10日、または約10~14日、または14日より後に決定される。

20

【0163】

いくつかの態様において、分化アッセイは、VEGF受容体II（KDR）もしくはアクチン - 2平滑筋（ACTA2）に関する免疫染色陽性によって決定される中胚葉系列に沿った多能性細胞の分化能を評価するか、またはネスチンもしくはチューブリン 3に関する免疫染色陽性によって決定される外胚葉系列に沿った多能性細胞の分化能を評価することができ、または - フェトプロテイン（AFP）に関する免疫染色陽性によって決定される内胚葉系列に沿った多能性細胞の分化能を評価することができる。

【0164】

いくつかの態様において、アッセイはヒトまたは他の哺乳動物対象などの対象由来の複数の異なる人工多能性幹細胞を含む、複数の異なる多能性幹細胞をアッセイするためのハイスループットアッセイである。

30

【0165】

本発明のもう1つの局面は、少なくとも1つまたは複数の多能性幹細胞株からのスコアカードを作成するための、本明細書において開示されるアッセイの使用に関する。

【0166】

本発明のもう1つの局面は、以下の段階を含む、多能性幹細胞スコアカードを作成する方法に関する：(i) 複数の多能性幹細胞株における第一の標的遺伝子セットにおけるDNAメチル化を測定する段階；(ii) 複数の多能性幹細胞株における第二の標的遺伝子セットにおける遺伝子発現を測定する段階；および(iii) 複数の多能性幹細胞株の分化能を測定する段階。いくつかの態様において、方法はさらに、(iv) 第一の標的遺伝子セットにおける各標的遺伝子の平均メチル化レベルを算出する段階；および(v) 第二の標的遺伝子セットにおける各標的遺伝子の平均遺伝子発現レベルを算出する段階を含む。

40

【0167】

本発明のもう1つの局面は、以下を含む、多能性幹細胞の性能パラメータのスコアカードに関する：(i) 複数の多能性幹細胞株由来の複数のDNAメチル化標的遺伝子のDNAメチル化レベルを含む第一のデータセット；(ii) 複数の多能性幹細胞株由来の複数の遺伝子発現標的遺伝子の遺伝子発現レベルを含む第二のデータセット；ならびに(iii) 複数の多能性幹細胞株由来の外胚葉、中胚葉、および内胚葉系列への分化に関する分化指向性レベルを含む第三のデータセット。

50

【0168】

いくつかの態様において、スコアカードは、表12Aもしくは12Cに記載される、または表13A、表13Bもしくは表14から選択される遺伝子の任意の組み合わせからの任意のDNAメチル化遺伝子などの、少なくとも約500個、少なくとも約1000個、少なくとも約1500個、または少なくとも約200個の参照DNAメチル化遺伝子のDNAメチル化レベルを測定することに由来する。

【0169】

いくつかの態様において、スコアカードは、表12Bもしくは12Cに記載される、または表13A、表13Bもしくは表14から選択される遺伝子の任意の組み合わせからの任意のDNAメチル化遺伝子などの、少なくとも約500個、少なくとも約1000個、少なくとも約1500個、または少なくとも約200個の参照DNAメチル化遺伝子の遺伝子発現レベルを測定することに由来する。

10

【0170】

いくつかの態様において、少なくとも第一および/または第二のデータセットは、データ保存デバイス、たとえばコンピュータデバイス上に位置するデータベースであるデータ保存デバイスに接続される。

【0171】

いくつかの態様において、本明細書において開示されるスコアカードは、少なくとも5個、少なくとも10個、少なくとも15個、または少なくとも20個の多能性幹細胞株などの複数の幹細胞株から決定される。いくつかの態様において、本明細書において開示されるスコアカードは1つの幹細胞株から決定され、ここで、各アッセイは3通りでまたはそれより多くで行われる。いくつかの態様において、「参照スコアカード」が所望である場合、スコアカードを作成するための複数の幹細胞株は、HUES64、HUES3、HUES8、HUES53、HUES28、HUES49、HUES9、HUES48、HUES45、HUES1、HUES44、HUES6、H1、HUES62、HUES65、H7、HUES13、HUES63、HUES66、およびこれらの任意の組み合わせからなる群より選択される少なくとも1つの多能性幹細胞株を含む。

20

【0172】

いくつかの態様において、スコアカードを作成するための幹細胞株は、哺乳動物多能性幹細胞株、たとえば、胚幹細胞および/もしくは人工多能性幹(iPS)細胞株を含むヒト多能性幹細胞株、ならびに/または、成体幹細胞、または体性幹細胞、または自己幹細胞である。

30

【0173】

本発明のもう1つの局面は、人工多能性幹細胞を胚幹細胞株と区別するための、本明細書において開示されるスコアカードの使用に関する。

【0174】

本発明のもう1つの局面は、以下を含む、本明細書において開示される方法を実行するためのキットに関する：(i) DNAメチル化状態を測定するための試薬；および(ii) 多能性幹細胞の分化指向性を測定するための試薬。

【0175】

本発明のもう1つの局面は、(i) (a) 多能性幹細胞株におけるDNAメチル化標的遺伝子セットのDNAメチル化データを受信して、該DNAメチル化データと、同じ標的遺伝子の参照DNAメチル化レベルとの比較を行う段階；(b) 多能性幹細胞株の分化能データを受信して、該分化能データを参照分化能データと比較する段階；(c) 参照DNAメチル化パラメータと比較したDNAメチル化データの比較、および参照分化データと比較した分化指向性の比較に基づいて、品質保証スコアカードを作成する段階を含む、少なくとも1つのプログラムを含む、少なくとも1つのメモリ；ならびに(ii) 該プログラムを作動させるためのプロセッサを含む、多能性幹細胞の品質保証スコアカードを作成するためのコンピュータシステムに関する。いくつかの態様において、システムのプログラムは、(d) 多能性幹細胞株における第二の標的遺伝子セットの遺伝子発現データを受信して、該発現データを同じ第二の標的遺伝子セットの参照遺伝子発現レベルと比較する段階；(e) 参照DNAメチル

40

50

化パラメータと比較したDNAメチル化データの比較、および参照分化データと比較した分化指向性の比較、および参照遺伝子発現レベルと比較した遺伝子発現データの比較に基づいて、品質保証スコアカードを作成する段階をさらに含む。いくつかの態様において、システムは、多能性幹細胞株の品質に基づいて幹細胞スコアカードレポートを作成するレポート作成モジュールをさらに含む。いくつかの態様において、システムは、データベースを含むメモリを含む。いくつかの態様において、データベースは、DNAメチル化遺伝子セットを階層的様式で配置し、たとえばDNAメチル化遺伝子を表12Aもしくは12B、または表13A、表13B、もしくは表14の順に並べて、および遺伝子発現遺伝子を表12Bもしくは表12Cの順に並べる。いくつかの態様において、データベースは、異なる系列への分化指向性を階層的様式で配置する。いくつかの態様において、メモリは、ネットワークへのアクセスが安全なサイトまたはパスワードアクセスを介して行われるネットワークを介して、たとえばインターネットなどのローカルネットワーク（LAN）または広域ネットワークを介して第一のコンピュータに接続される。

【0176】

いくつかの態様において、本明細書において開示されるシステムは、試験される多能性幹細胞株の適切な使用、有用性、または応用の指標を提供するスコアカードを提供する。

【0177】

本発明のもう1つの局面は、以下を含む、多能性幹細胞株の品質保証スコアカードの作成に関する命令を含む、コンピュータ読み取り可能な媒体に関する：(i) 多能性幹細胞株におけるDNAメチル化標的遺伝子セットのDNAメチル化データを受信して、該DNAメチル化データと、同じ標的遺伝子の参照DNAメチル化レベルとの比較を行うこと；(ii) 多能性幹細胞株の分化能データを受信して、該分化能データを参照分化能データと比較すること；ならびに(iii) 参照DNAメチル化パラメータと比較したDNAメチル化データの比較、および参照分化データと比較した分化指向性の比較に基づいて、品質保証スコアカードを作成すること。いくつかの態様において、コンピュータ読み取り可能な媒体は、(iv) 多能性幹細胞株における第二の標的遺伝子セットの遺伝子発現データを受信して、該発現データを同じ第二の標的遺伝子セットの参照遺伝子発現レベルと比較すること；ならびに(v) 参照DNAメチル化パラメータと比較したDNAメチル化データの比較、および参照分化データと比較した分化指向性の比較、および参照遺伝子発現レベルと比較した遺伝子発現データの比較に基づいて、品質保証スコアカードを作成することに関する命令をさらに含む。

【0178】

本発明のもう1つの局面は、以下の少なくとも2つを含む、多能性幹細胞株の品質を決定するためのキットに関する：(i) 複数のDNAメチル化遺伝子のメチル化状態を測定するための試薬、(ii) 複数の遺伝子の遺伝子発現レベルを測定するための試薬；ならびに(iii) 多能性幹細胞の外胚葉、中胚葉、および内胚葉系列への分化指向性を測定するための試薬。

【0179】

スコアカード

本発明の1つの局面は、以下を含む、多能性幹細胞の性能パラメータのスコアカードに関する：(i) 少なくとも5個の多能性幹細胞集団からの複数のDNAメチル化標的遺伝子のDNAメチル化レベルを含む第一のデータセット；(ii) 少なくとも5個の多能性幹細胞集団からの複数の遺伝子発現標的遺伝子の遺伝子発現レベルを含む第二のデータセット；ならびに(iii) 少なくとも5個の多能性幹細胞集団からの外胚葉、中胚葉、および内胚葉系列への分化に関する分化指向性レベルを含む第三のデータセット。いくつかの態様において、複数の参照DNAメチル化遺伝子は、少なくとも約1000個の参照DNAメチル化遺伝子、または少なくとも約2000個の参照DNAメチル化遺伝子であり、またはいくつかの態様において、全ゲノムのDNAメチル化状態である。いくつかの態様において、参照DNAメチル化遺伝子は、癌遺伝子、腫瘍遺伝子、および腫瘍抑制遺伝子、系列マーカー遺伝子、および発達遺伝子を含む群から選択される任意の遺伝子である。

【0180】

いくつかの態様において、DNAメチル化標的遺伝子は、BMP4、CAT、CD14、CXCL5、DAZL、DNMT3B、GATA6、GAPDH、LEFTY2、MEG3、PAX6、S100A6、SOX2、SNAI1、TFからなる群より選択される任意の遺伝子および遺伝子の組み合わせである。いくつかの態様において、DNAメチル化標的遺伝子は、表12Aもしくは表12Cから選択されるまたは表13A、表13B、もしくは表14から選択される遺伝子の任意の組み合わせである。いくつかの態様において、DNAメチル化は、表12Aおよび表12Cに記載される標的遺伝子のプロモーター領域において決定されるが、本発明は、表13A、表13B、または表14に記載される遺伝子のプロモーター領域を含む、全てのゲノム領域（ならびに非ゲノム領域）におけるDNAメチル化を決定することを包含する。いくつかの態様において、DNAメチル化は、プロモーター、エンハンサー、隔離エレメント、CpGアイランド、CpGアイランドショア等などの任意のゲノム領域、またはゲノム領域の特定タイプにおいて決定される。さらに、DNAメチル化を、非コード遺伝子ならびに非コード転写物、たとえば天然のアンチセンス転写物（NAT）、マイクロRNA（miRNA）遺伝子、および他の全てのタイプの核酸および/またはRNA転写物において決定することができる。いくつかの態様において、非常に多様である領域を直接導き出すためにDNAメチル化データを用いることができ、エピジェネティックな変化を受けやすいゲノム領域を予測するためにDNA配列データを用いることができる。さらに、いくつかの態様において、癌、正常および異常な発達ならびに疾患に関係する遺伝子およびゲノム領域に関するこれまでの知識を候補として用いることができる。いくつかの態様において、DNAメチル化標的遺伝子は、表12Aおよび/または表12Cにおける遺伝子のリストから選択される、または表13A、表13B、もしくは表14から選択される、任意の組み合わせの、少なくとも約200個、または少なくとも約300個、または少なくとも約400個、または少なくとも約500個、または少なくとも約600個、または少なくとも約800個、または少なくとも約1000個、または少なくとも約1500個、または少なくとも約2000個、または少なくとも約3000個、または少なくとも約4000個、または少なくとも約5000個の遺伝子である。いくつかの態様において、遺伝子は、表12Aもしくは12Cに記載される、または表13A、表13Bもしくは表14から選択される遺伝子の1~200番、または1~500番、または1~1000番から選択される遺伝子セットの任意の組み合わせである。

【0181】

いくつかの態様において、スコアカードの第一および第二のデータセットは、コンピュータデバイス上に存在するデータベースであるデータ保存デバイスなどのデータ保存デバイスに接続される。

【0182】

いくつかの態様において、少なくとも15個の多能性幹細胞株を用いて、スコアカードのための第一または第二または第三のデータセットを作成する。いくつかの態様において、第一、第二、または第三のデータセットは、以下の群から選択される以下の多能性幹細胞株の少なくとも5個もしくはそれより多く、または少なくとも6個、または少なくとも7個、または少なくとも8個、または少なくとも9個、または少なくとも10個、または少なくとも11個、または少なくとも12個、または少なくとも13個、または少なくとも14個、または少なくとも15個、または少なくとも16個、または少なくとも17個、または少なくとも18個、または19個全てから得られる：HUES64、HUES3、HUES8、HUES53、HUES28、HUES49、HUES9、HUES48、HUES45、HUES1、HUES44、HUES6、H1、HUES62、HUES65、H7、HUES13、HUES63、HUES66。

【0183】

いくつかの態様において、スコアカードのためのデータセットを作成するために用いられる多能性幹細胞集団は、ヒト多能性幹細胞集団または人工多能性幹（iPS）細胞集団または胚幹細胞集団または成体幹細胞集団または自己幹細胞集団、または胚幹（ES）細胞集団などの哺乳動物多能性幹細胞集団である。

【0184】

いくつかの態様において、本明細書において開示されるスコアカードを、関心対象の多

能性幹細胞集団のDNAメチル化レベル、遺伝子発現レベル、および分化指向性レベルと比較することができ、これを用いて特定系列に沿った最適な分化および/または望ましくない特徴を有する指向性、たとえば癌細胞へと発達する素因を有する多能性幹細胞集団を予測することによって、多能性幹細胞集団の挙動を検証および/または予測することができる。このように、いくつかの態様において、スコアカードは、たとえば望ましい特徴（たとえば、特定系列に沿った高い分化能）を有する関心対象の多能性幹細胞集団に対して正の選択を行うための、および/または、負の選択を行うための、たとえば望ましくない特徴を有する細胞、たとえば癌細胞へと発達する素因を有する細胞を同定して廃棄するための方法などの選択方法において用いることができる。

【0185】

いくつかの態様において、多能性幹細胞におけるその遺伝子（たとえば、正常な参照値）のDNAメチル化の正常な変動と比較して、DNAメチル化レベルの統計学的に有意な（FDR < 5%）および/または20%ポイントを超える絶対差を有する標的遺伝子のDNAメチル化レベルを有する多能性幹細胞株は、エピジェネティックな外れ値のDNAメチル化遺伝子であると見なされるであろう。参照多能性幹細胞と比較して、多数の、たとえば少なくとも約5個、または少なくとも約6個、または少なくとも約7個、または少なくとも約8個、または少なくとも約5~10個、または少なくとも約10~15個、または少なくとも約10~50個、または少なくとも約50~100個、または少なくとも約100~150個、または少なくとも約150~200個または200個より多くの完全なエピジェネティックな外れ値のDNAメチル化遺伝子を有する多能性幹細胞は、外れ値の多能性幹細胞であると見なされるであろう。したがって、そのような多能性幹細胞を用いて、負の選択を行う、たとえば、望ましくない特徴を有する細胞を単離して廃棄することができる。

【0186】

いくつかの態様において、多能性幹細胞においてその標的癌遺伝子のDNAメチル化の正常な変動（たとえば、癌遺伝子に関する正常な参照DNAメチル化レベル）と比較してDNAメチル化レベルの統計学的に有意な（FDR < 5%）および/または20%ポイントを超える絶対差を有する標的癌遺伝子のDNAメチル化レベルを有する多能性幹細胞株は、エピジェネティックな外れ値のDNAメチル化癌遺伝子であると見なされるであろう。参照多能性幹細胞と比較して、多数の、たとえば少なくとも約5個、または少なくとも約6個、または少なくとも約7個、または少なくとも約8個、または少なくとも約5~10個、または少なくとも約10~15個、または少なくとも約10~50個、50個より多くの完全なエピジェネティックな外れ値のDNAメチル化癌遺伝子を有する多能性幹細胞は、外れ値の多能性幹細胞であると見なされるであろう。したがって、そのような多能性幹細胞を用いて、負の選択を行う、たとえば癌遺伝子のDNAメチル化の増加または減少などの望ましくない特徴を有する細胞を単離して廃棄することができる。

【0187】

いくつかの態様において、多能性幹細胞においてその遺伝子の遺伝子発現の正常な変動（たとえば、正常な参照値）と比較して遺伝子発現レベルの統計学的に有意な（FDR < 10%）および/または1 log-2倍を超える変化の絶対差を有する標的遺伝子の遺伝子発現レベルを有する多能性幹細胞株は、外れ値の遺伝子発現遺伝子であると見なされるであろう。参照多能性幹細胞と比較して、多数の、たとえば少なくとも約5個、または少なくとも約6個、または少なくとも約7個、または少なくとも約8個、または少なくとも約5~10個、または少なくとも約10~15個、または少なくとも約10~50個、または少なくとも約50~100個、またはそれより多くの完全な外れ値の遺伝子発現遺伝子を有する多能性幹細胞は、外れ値の多能性幹細胞であると見なされるであろう。したがって、そのような多能性幹細胞を用いて、負の選択を行う、たとえば望ましくない特徴を有する細胞を単離して廃棄することができる。

【0188】

いくつかの態様において、多能性幹細胞においてその系列遺伝子の遺伝子発現の正常な変動（たとえば、正常な参照値）と比較して系列遺伝子発現レベルの統計学的に有意な（

10

20

30

40

50

FDR < 5%) および / または 1 log-2倍を超える変化の絶対差を有する系列遺伝子の遺伝子発現レベルを有する多能性幹細胞株は、外れ値の分化遺伝子であると見なされるであろう。参照多能性幹細胞と比較して、多数の、たとえば少なくとも約5個、または少なくとも約6個、または少なくとも約7個、または少なくとも約8個、または少なくとも約5~10個、または少なくとも約10~15個、または少なくとも約10~50個、または少なくとも約50~100個、またはそれより多くの完全な外れ値の系列遺伝子発現化遺伝子を有する多能性幹細胞は、参照多能性幹細胞株と同じ系列に沿って分化しない可能性がある外れ値の多能性幹細胞であると見なされるであろう。したがって、そのような多能性幹細胞を用いて、負の選択を行う、たとえば望ましくない特徴、たとえば特定の系列に沿って分化しない可能性がある細胞を単離して廃棄することができる。

10

【0189】

好ましい多能性幹細胞のスコアカードを作成する方法

本発明のもう1つの局面は、以下の段階を含む多能性幹細胞スコアカードを作成する方法に関する：(i) 複数の多能性幹細胞集団における標的遺伝子セットにおけるDNAメチル化を測定する段階；(ii) 複数の多能性幹細胞株における第二の標的遺伝子セットにおける遺伝子発現を測定する段階；および(iii) 複数の多能性幹細胞株の分化能を測定する段階。いくつかの態様において、多能性幹細胞スコアカードを作成する方法を用いて、複数の多能性幹細胞株、たとえば、少なくとも5個、または少なくとも6個、または少なくとも7個、または少なくとも8個、または少なくとも9個、または少なくとも10個、または少なくとも15個、または少なくとも20個、または少なくとも30個、または少なくとも40個または40個より多くの異なる多能性幹細胞集団からのDNAメチル化の正常な変動、DNA遺伝子発現の正常な変動、および正常な分化指向性の値を含むスコアカードを作成することができる。

20

【0190】

アッセイ

本発明のもう1つの局面は、以下の少なくとも2つを含む、多能性細胞の複数の特性の特徴決定を行うためのアッセイに関する：(i) DNAメチル化アッセイ；(ii) 遺伝子発現アッセイ；および(iii) 分化アッセイ。

【0191】

いくつかの態様において、DNAメチル化アッセイは、バイサルファイトシーケンシングアッセイ、または全ゲノムシーケンシングアッセイ、たとえば、簡約表示バイサルファイトシーケンシング(RRBS)である。いくつかの態様において、DNAメチル化アッセイは、濃縮に基づくDNAメチル化アッセイ(たとえば、MeDIP)、または制限酵素に基づくDNAメチル化アッセイ(たとえば、CHARMまたはHELP)、または本明細書においておよび実施例において開示されるDNAメチル化アッセイの他の手段である。いくつかの態様において、DNAメチル化アッセイは、Illuminaメチル化アッセイである。いくつかの態様において、遺伝子発現アッセイはマイクロアレイアッセイである。

30

【0192】

いくつかの態様において、分化指向性アッセイは、定量的分化アッセイ、たとえば多能性細胞の以下の系列：中胚葉、内胚葉、および外胚葉、ニューロン造血系列の少なくとも1つへの分化能を評価することができる分化アッセイである。いくつかの態様において、多能性細胞の以下の系列：中胚葉、内胚葉、および外胚葉の少なくとも1つへの分化能は、分化指向性を評価するためのバイオインフォマティクスアルゴリズムと組み合わせた胚様体(EB)における遺伝子発現プロファイリングによって決定され、ここで、本明細書において表7に開示される系列遺伝子の遺伝子発現レベルが決定され、系列マーカー遺伝子の遺伝子発現レベルの統計学的に有意な差(FDR < 5%)の変化、および / または遺伝子発現レベルの1 log-2倍を超える変化により、参照多能性幹細胞株と比較して異なる系列に沿って分化する指向性が示される。代替の態様において、以下の系列、すなわち中胚葉、内胚葉、および外胚葉のうち少なくとも1つへの多能性細胞の分化能は、中胚葉、内胚葉、および外胚葉系列の少なくとも1つのマーカーに対する抗体を用いる免疫染色またはFAC

40

50

ソーティングによって決定される。いくつかの態様において、以下の系列、すなわち中胚葉、内胚葉、および外胚葉のうち少なくとも1つへの多能性細胞の分化能は、EBにおいて少なくとも約7日後に多能性幹細胞を免疫染色することによって決定される。中胚葉、内胚葉、および外胚葉系列の系列マーカーの例は、当業者に周知であり、中胚葉系列マーカー-VEGF受容体II (KDR) またはアクチン α -2平滑筋 (ACTA2)、外胚葉系列マーカー-ネスチンまたはチューブリン β 3、および内胚葉系列マーカー α -フェトプロテイン (AFP) が挙げられるがこれらに限定されるわけではない。

【0193】

いくつかの態様において、アッセイは、たとえば、同じまたは異なる対象、たとえば哺乳動物対象またはヒト対象から得られた体細胞のリプログラミングに由来する複数の異なる人工多能性幹細胞を評価することができる、複数の異なる多能性幹細胞をアッセイするためのハイスループットアッセイである。

10

【0194】

いくつかの態様において、本明細書において開示されるアッセイを用いて、少なくとも1つまたは複数の多能性幹細胞集団から本明細書において開示されるスコアカードを作成することができる。

【0195】

エピジェネティックなマッピング

理論に拘束されたくはないが、エピジェネティックな事象は、遺伝子の発現において有意な役割を果たし、癌の発生および進行において重要である。DNAメチル化などのエピジェネティックな変化は、哺乳動物の正常な発達において遺伝子発現を調節するように作用する。プロモーター高度メチル化もまた、腫瘍抑制遺伝子などの重要な増殖調節因子の転写サイレンシングを通して癌において主要な役割を果たしている。腫瘍抑制遺伝子などの遺伝子の機能の喪失は、DNAメチル化などのエピジェネティックな変化を通して起こりうる。「エピジェネティック」という用語は、遺伝子のヌクレオチド配列における変化に起因しない、遺伝子発現における遺伝可能な変化を意味する。たとえば、転写が開始される遺伝子のプロモーター領域でDNAがメチル化されると、遺伝子は不活化されてサイレンシングされる。エピジェネティックな修飾には、たとえば、DNAメチル化、クロマチンの翻訳後修飾、低分子非コードRNA、ならびにクロマチンの凝縮および脱凝縮などのクロマチンに対する非共有構造修飾が挙げられるがこれらに限定されるわけではない。いくつかの例において、エピジェネティックな修飾はまた、DNAメチル化、ユビキチン化、リン酸化、グリコシル化、スモイル化、アセチル化、S-ニトロシル化、またはニトロシル化、シトルリン化、または脱イミノ化、neddylation、O⁶-methylguanine (O⁶-MeG), ADP-リボシル化、ヒドロキシル化、ファテニル化、ウフミル化、プレニル化、ミリストイル化、S-パルミトイル化、チロシン硫酸化、ホルミル化、およびカルボキシル化を含むタンパク質の翻訳後修飾 (PTM) の形でありうる。

20

30

【0196】

本発明の方法、システム、およびキットのいくつかの態様において、エピジェネティックな修飾のレベルは、関心対象の多能性幹細胞株において決定される。いくつかの態様において、エピジェネティックな修飾はDNAメチル化である。いくつかの態様において、DNAメチル化標的遺伝子のメチル化が決定される。したがって、いくつかの態様において、DNAメチル化標的遺伝子は、遺伝子の発現の抑制 (たとえば、エピジェネティックなサイレンシング) を決定することが望ましい任意の遺伝子である。いくつかの態様において、DNAメチル化標的遺伝子は癌遺伝子、たとえば、腫瘍遺伝子または腫瘍抑制遺伝子である。いくつかの態様において、DNAメチル化標的遺伝子は発達遺伝子であり、いくつかの態様において、DNAメチル化標的遺伝子は系列マーカー遺伝子である。

40

【0197】

いくつかの態様において、DNAメチル化は、BMP4、CAT、CD14、CXCL5、DAZL、DNMT3B、GATA6、GAPDH、LEFTY2、MEG3、PAX6、S100A6、SOX2、SNAI1、TFの群から選択される任意の遺伝子において決定または測定される。いくつかの態様において、DNAメチル化は、DAZL

50

、LEFTY2、CXCL5、MEG3、S100A6、CAT、TF、CD14などの多様なDNAメチル化レベルを有する遺伝子である。いくつかの態様において、DNAメチル化は、PAX6、DNMT3B、GATA6、GAPDH、SOX2、SNAI1、BMP4などの低いDNAメチル化変動性を有する遺伝子である。

【0198】

いくつかの態様において、DNAメチル化は、参照DNAメチル化標的遺伝子セットにおいて決定または測定され、ここで、参照DNAメチル化遺伝子は癌遺伝子および/または発達遺伝子でありえて、表12Aに開示される。いくつかの態様において、第一の参照DNAメチル化遺伝子セットにおいて用いられる遺伝子は、表12Aおよび/または表12Cにおける遺伝子のリストから選択される、または表13A、表13B、もしくは表14から選択される、任意の組み合わせの少なくとも約200個、または少なくとも約300個、または少なくとも約400個、または少なくとも約500個、または少なくとも約600個、または少なくとも約800個、または少なくとも約1000個、または少なくとも約1500個、または少なくとも約2000個、または少なくとも約3000個、または少なくとも約4000個、または少なくとも約5000個の遺伝子である。いくつかの態様において、遺伝子は、表12Aもしくは表12Cに記載される、または表13A、表13B、もしくは表14から選択される遺伝子の1~200番、または1~500番、または1~1000番から選択される遺伝子セットの任意の組み合わせである。

10

【0199】

いくつかの態様において、DNAメチル化は、以下の140個の遺伝子セットを任意に組み合わせた少なくとも50個の遺伝子または少なくとも100個の遺伝子において測定される：

20

PON3; CD14; PEG3AS; CRCT1, LCE5A;
 HIST1; H2BB; HIST1; H3C, CRCT1, LCE5A, PTK2B, TF, CAT, SLC38A11, ZNF528, CALCB, ERAS, INGX, TMRSS12, ZNF248, ZNF876P, SLC17A3, TDRD5, LCE3A, ASB3, GPR75, ZNF354C, PEG3AS, KAAG1, PCDHA2, HPDL, ZNF737, AGBL2, COMT, TXNRD2, SLC30A8, H2AFZP1, CTSE, ZNF833, S100A5, S100A6, PRDM9, CYP2E1, ZNF177, CR1L, ZNF572, MOS, FAM70A, GP5, PAPOLB, ZDHHC15, HSF5, CDX4, GOLGA8B, KLF8; ARMCX5; CBLN4, POU3F4, LYNX1, DENND2D, CYP2E1, ZNF562, PPYR1, KLHL34, ZNF562, TMLHE, CCDC11, GYG2P, TCEAL2, ZNF454, ZNF667, TRIM4, FAM24B, ZNF397OS, PAQR6, DENND2D, LYNX1, BHMT2, DMGDH, PF4, LTF, NAP1L6, ALOX15B, CES1, PPP1R13L, COMT, TXNRD2, LYNX1, DNAJC15, ARMCX1, TRPM2, GOLGA8A, ZPBP, ZNF630, BHMT2, DMGDH, SLC7A3, SLFN13, PLEK2, DYNLT3, SLC2A14, SPATS1, SLC01A2, TCEAL6, SLC2A14, TAF9B, KIAA1210, CNTD2, PLD6, CFLAR, PHF8, TBPL2, RWDD2B, DEFB124, REM1, TCEAL6, CD14, BCL2L10, ZNF630, DCDC2, CRYGD, ZNF440, RFPL2, MYCL2, TRPM2, MEG3, TEK4, FAM104B, EDNRB, OSGIN1, NKAP, NR0B1, SPIN3, NDUFA1, RNF113A, ZNF726, ZNF502 および C3orf62

30

【0200】

現在では多くの遺伝子の機能が公知であることから、癌のリスクの増加または減少、特定細胞型および系列への分化能の差、薬物に対する耐性、ならびに疾患モデリング、薬物スクリーニング、および再生治療に関する全般的有用性などの推定の効果を、癌遺伝子の差次的発現および/またはDNAメチル化に対して割付けすることができる。

40

【0201】

癌細胞は、プロモーターCpGアイランドDNA高度メチル化、ならびにヒストン修飾およびクロマチン構造における関連する変化を含む、広範囲に及ぶ異常なエピジェネティックな変化を含む。癌における腫瘍抑制遺伝子の異常なエピジェネティックなサイレンシングは、遺伝子発現、クロマチン構造、ヒストン修飾、およびシトシン-5-DNAメチル化の変化を伴う。

【0202】

50

したがって、いくつかの態様において、DNAメチル化標的遺伝子は、癌遺伝子、たとえば腫瘍遺伝子および腫瘍抑制遺伝子、および発達遺伝子ならびに系列マーカー遺伝子を含む。例として、腫瘍遺伝子のプロモーターの高度メチル化の存在が検出される場合、エピジェネティックなサイレンシングが起こっていること、および腫瘍遺伝子が抑制または永続的にサイレンシングされることが示され、これは、望ましい特徴でありうる。しかし、メチル化レベルが減少していることは、エピジェネティックなサイレンシングが存在しないこと、および腫瘍遺伝子が発現されうることを示しており、このことは、多能性幹細胞が自己再生および高い悪性形質転換能の素因を有することを示しうる。同様に、癌遺伝子が腫瘍抑制遺伝子である場合、高度メチル化プロモーターの存在、または該腫瘍抑制遺伝子に関するメチル化の正常な変動と比較して統計学的に有意に高いレベルのメチル化は、

10

エピジェネティックなサイレンシングを示し、かつ、腫瘍抑制因子の発現が永続的に抑制されていることを示し、このことは、多能性幹細胞が持続的な自己再生および高い悪性形質転換能の素因を有することを示している。したがって、腫瘍遺伝子および/または腫瘍抑制遺伝子のメチル化状態を用いて、多能性幹細胞が持続的な自己再生および高い悪性形質転換能に対して素因を有するか否かを予測することができる。さらに、いくつかの態様において、DNAメチル化レベルを癌遺伝子セット、たとえば腫瘍遺伝子および腫瘍抑制遺伝子において測定および決定することにより、多能性幹細胞が持続的な自己再生および高い悪性形質転換能に対して素因を有するか否かを予測することができる。

【0203】

代わりの態様において、DNAメチル化レベルを系列特異的遺伝子セット（たとえば、系列マーカー遺伝子）または発達特異的遺伝子セットにおいて測定および決定することにより、多能性幹細胞が、特定の発達経路に沿ってまたは系列マーカーを発現する細胞型へと分化できるか否かを予測することができる。

20

【0204】

重要なことに、本明細書において開示される分化指向性アッセイおよび方法において、系列特異的遺伝子セット（たとえば、系列マーカー遺伝子）または発達特異的遺伝子セットにおけるDNAメチル化レベルを、多能性幹細胞株を培養して、既定の期間自発的に分化させた後に決定すると、系列マーカー遺伝子セットのDNAメチル化アッセイの結果から、多能性幹細胞株の系列分化の偏りを予測することができる。分化指向性アッセイのいくつかの態様において、系列マーカー遺伝子セットのDNAメチル化アッセイは、特定の系列に沿った定方向性分化後の多能性幹細胞株において行われる。

30

【0205】

メチル化標的遺伝子が発達遺伝子または系列マーカー遺伝子である例において、遺伝子プロモーターの高度メチル化の存在、または、該発達遺伝子もしくは系列マーカー遺伝子のDNAメチル化の正常な変動と比較して統計学的に有意に高いレベルのDNAメチル化は、エピジェネティックなサイレンシングを示し、かつ、該発達遺伝子または系列マーカーの発現が永続的に抑制されることを示し、このことは、多能性幹細胞が該発達遺伝子および/または系列マーカーを発現しない素因を有し、かつしたがって発達遺伝子の発達経路に沿って分化しないこと、または系列マーカーを発現する細胞型へと分化しないと予測されることを示す。代替的状況において、多能性幹細胞における発達遺伝子または系列マーカー

40

遺伝子のメチル化レベルが該遺伝子のメチル化レベルの正常な変動内である場合、これを用いて、該多能性幹細胞は該発達遺伝子の発達経路に沿って分化するように、または系列マーカーを発現する細胞型へと分化するように進行できると予測することができる。したがって、発達遺伝子および/または系列マーカーのメチル化状態を用いて、多能性幹細胞が特定の発達経路に沿ってまたは該系列マーカーを発現する細胞型へと、分化できるか否かを予測することができる。

【0206】

上記のDNAメチル化の測定は、主に1つの遺伝子の効果に焦点をあてているが、いくつかの態様において、スコアカードは、たとえば細胞株の品質（たとえば、癌様株へと発達する可能性）および有用性（たとえば、関心対象の特定系列に沿って分化する可能性、また

50

は分化しない可能性)を予測するために、多数の遺伝子、たとえば「癌遺伝子」セットにおける多数の遺伝子または「系列マーカー遺伝子」セットにおける多数の遺伝子のデータの組み合わせでDNAメチル化を測定する。したがって、特定の所望のまたは望ましくない特徴を同定するために多能性幹細胞株を感度よく正確に特徴決定するための「カスタムスコアカード」を開発するために、特定のDNAメチル化標的遺伝子セットを選択することができる。これは、特定の多能性幹細胞株の品質および有用性を決定するために本明細書において開示されるスコアカードを用いる重要な利点の1つである。

【0207】

本発明のいくつかの態様において、DNAメチル化状態は、PRC2遺伝子ならびにDlx、Irx、Lhx、およびPax遺伝子ファミリー（神経発生、造血、および体軸パターン形成に関係する）またはFox、Sox、Gata、およびTbxファミリー（発達プロセスに関係する）の他の転写因子において同定される。

10

【0208】

本明細書において述べるように、いくつかの態様において、多能性幹細胞においてその遺伝子に関するDNAメチル化の正常な変動（たとえば、正常な参照値）と比較してDNAメチル化レベルの統計学的に有意な（FDR < 5%）および/または20パーセントポイントを超える絶対差を有する標的遺伝子のDNAメチル化レベルを有する多能性幹細胞株は、エピジェネティックな外れ値のDNAメチル化遺伝子であると見なされるであろう。参照多能性幹細胞と比較して、多数の、たとえば少なくとも約5個、または少なくとも約6個、または少なくとも約7個、または少なくとも約8個、または少なくとも約5~10個、または少なくとも約10~15個、または少なくとも約10~50個、または少なくとも約50~100個、または少なくとも約100~150個、または少なくとも約150~200個または200個より多くの完全なエピジェネティックな外れ値のDNAメチル化遺伝子を有する多能性幹細胞は、外れ値の多能性幹細胞であると見なされるであろう。したがって、そのような多能性幹細胞を用いて、負の選択を行う、たとえば、望ましくない特徴を有する細胞を単離して廃棄することができる。

20

【0209】

いくつかの態様において、多能性幹細胞においてその標的癌遺伝子のDNAメチル化の正常な変動（たとえば、癌遺伝子に関する正常な参照DNAメチル化レベル）と比較してDNAメチル化レベルの統計学的に有意な（FDR < 5%）および/または20%ポイントを超える絶対差を有する標的癌遺伝子のDNAメチル化レベルを有する多能性幹細胞株は、エピジェネティックな外れ値のDNAメチル化癌遺伝子であると見なされるであろう。参照多能性幹細胞と比較して、多数の、たとえば少なくとも約5個、または少なくとも約6個、または少なくとも約7個、または少なくとも約8個、または少なくとも約5~10個、または少なくとも約10~15個、または少なくとも約10~50個、50個より多くの完全なエピジェネティックな外れ値のDNAメチル化癌遺伝子を有する多能性幹細胞は、外れ値の多能性幹細胞であると見なされるであろう。したがって、そのような多能性幹細胞を用いて、負の選択を行う、たとえば癌遺伝子のDNAメチル化の増加または減少などの望ましくない特徴を有する細胞を単離して廃棄することができる。

30

【0210】

DNAメチル化法およびアッセイ

DNAメチル化を測定するために、濃縮に基づく方法（たとえば、MeDIP、MBD-seqおよびMethylCap）、バイサルファイトに基づく方法（たとえば、RRBS、バイサルファイトシーケンシング、Infinium、GoldenGate、COBRA、MSP、MethylLight）、ならびに制限消化法（たとえば、MRE-seq）を含むがこれらに限定されるわけではない当業者に一般的に公知である任意の方法を用いることができる。1つの態様において、エピジェネティックなプロファイリングおよびエピジェネティックなマッピングのための方法は、全ゲノムエピジェネティックなマッピングである。当業者に公知の多能性幹細胞株のエピジェネティックなマッピングに関する任意の方法を用いることができ、これは、たとえば簡約表示バイサルファイトシーケンシング（RRBS）ならびにその全内容が参照により本明細書に組み入れ

40

50

られる米国特許出願第US2010/0172880号に開示される方法を含む。他のDNAメチル化アッセイは、その全内容が参照により本明細書に組み入れられる、米国特許出願US2008/0213789号および第US2010/0075331号、ならびに米国特許第6,960,434号および第7,425,415号に開示されている。多能性幹細胞のDNAメチル化を測定する方法はまた、その全内容が参照により本明細書に組み入れられる、"Genome-wide mapping of DNA methylation: a quantitative technology comparison" by Bock et al.,においても記述され、ここで発明者らは、多能性幹細胞の正確なDNAメチル化データを生じる多様なDNAメチル化法 (MeDIP-seq: メチル化DNA免疫沈降、MethylCap-seq: アフィニティ精製によるメチル化DNA捕捉、RRBS: 簡約表示パイサルファイトシークエンシング、およびInfiniumヒトメチル化アッセイ) を評価した。

10

【0211】

いくつかの態様において、DNAメチル化アッセイは、種特異的であり、このためマウス胚線維芽細胞をヒト多能性幹細胞のフィーダー層として用いても、エピジェネティックな分析を妨害しないであろう。

【0212】

ゲノム規模でのDNAメチル化プロファイリングを可能にするためにいくつかの方法が開発されている。これらの方法の多くは、マイクロアレイまたはハイスループットシークエンシングによるDNA分析を、DNAメチル化パターンをDNA配列情報またはライブラリ濃縮へと翻訳する4つの方法の1つと組み合わせる: (i) メチル化DNA免疫沈降 (MeDIP) は、5-メチル-シトシンに対して特異的な抗体を用いて、超音波処理したDNA11からのメチル化断片を回収する、(ii) アフィニティ精製によるメチル化DNA捕捉 (MethylCap) はメチル結合ドメインタンパク質を使用して、類似のメチル化レベルを有するDNA分画を得る; (iii) パイサルファイトに基づく方法は非メチル化 (メチル化は変換しない) シトシンをウラシルに選択的に変換する化学反応を利用して、このように、メチル化特異的ヌクレオチド多形をDNA配列に導入する; (iv) メチル化特異的消化は、原核細胞制限酵素を用いてDNAをメチル化特異的に分画する。

20

【0213】

MeDIP-seq、MethylCap-seq、RRBS、およびInfiniumヒトメチル化アッセイを含む4つの一般的な方法は、生物医学研究およびバイオマーカー開発に関するその実際的な有用性を特に強調して、本発明者らによって既に評価された ("Genome-wide mapping of DNA methylation: a quantitative technology comparison" by Bock et al.,を参照されたい)。これらの方法は、以下の検討に基づいて本発明の方法、システム、およびアッセイにおいて有用である: (i) 4つ全ての方法は、詳細なプロトコルが公表されておりおよび/または市販のキットが使用できることから、比較的容易にセットアップされる、(ii) RRBSは、その試料あたりの費用が他の方法と同等であり、大規模な試料にとって現実的であることから、他の全ゲノムパイサルファイトシークエンシングに対して利点を有する、(iii) Infiniumヒトメチル化アッセイは、その広い用途、および既存の遺伝子タイピングパイプラインに容易に組み込めること、および同様にマイクロアレイに基づく方法であることから、本明細書において開示される方法、システム、およびアッセイにおいて有用である。いくつかの態様において、マイクロアレイおよび/またはメチル化特異的消化を利用する他のDNAメチル化法は、既にベンチマークされていることから、これらを本明細書において開示される方法、システム、およびアッセイにおいて用いることができる。現在のこれらのアッセイおよび分析を行う方法は、本明細書において実施例において、サブタイトル「他のDNAメチル化マッピング法」と題する方法の章で開示される。

30

40

【0214】

多数の異なるエピジェネティックなプロファイリング技術が開発されている (たとえば、全てが参照により本明細書に組み入れられるLaird, P.W. Hum Mol Genet 14, R65-R76, 2005; Laird, P.W. Nat Rev Cancer 3, 253-66, 2003; Squazzo, S.L. et al. Genome Res 16, 890-900, 2006; and Lieb, J. D. et al. Cytogenet Genome Res 114, 1-15, 2006)。これらは、特定クロマチン成分またはヒストン修飾に向けられる抗体によるクロマ

50

チンの免疫沈降に主に依存するクロマチン照合技術と、DNAメチル化分析技術とに大きく分けることができる。クロマチンの免疫沈降を、高密度ゲノムタイリングマイクロアレイ (ChIP-Chip) とのハイブリダイゼーションと組み合わせて、包括的ゲノムデータを得ることができる。しかし、クロマチンの免疫沈降は、小さい割合の細胞においてエピジェネティックな異常を検出することができないが、DNAメチル化分析は、癌患者の血流中の腫瘍由来の遊離のDNAの高感度検出に応用されて成功している (Laird, P.W. *Nat Rev Cancer* 3, 253-66, 2003)。好ましくは、非メチル化分子の10,000倍過剰量中で異常にメチル化された分子を検出することができる (Eads, CA. et al., *Nucleic Acids Res* 28, E32, 2000) 感度のよい、正確な、蛍光に基づくメチル化特異的PCRアッセイ (たとえば、METHYLIGHT (商標))、または非メチル化分子の非常に大量または過剰量の中で1つの異常にメチル化されたDNA分子を検出することができるさらにより感度の高いMETHYLIGHT (商標) の変化形が用いられる。特定の局面において、METHYLIGHT (商標) 分析は、本出願人によって既に記述されたように行われる (たとえば、Weisenberger, DJ. et al. *Nat Genet* 38:787-793, 2006; Weisenberger et al., *Nucleic Acids Res* 33:6823-6836, 2005; Siegmund et al., *Bioinformatics* 25, 25, 2004; Eads et al., *Nucleic Acids Res* 28, E32, 2000; Virmani et al., *Cancer Epidemiol Biomarkers Prev* 11 :291-297, 2002; Uhlmann et al., *Int J Cancer* 106:52-9, 2003; Ehrlich et al., *Oncogene* 25:2636-2645, 2006; Eads et al., *Cancer Res* 61 :3410-3418, 2001 ; Ehrlich et al., *Oncogene* 21 :6694-6702, 2002; Marjoram et al., *BMC Bioinformatics* 7, 361, 2006; Eads et al., *Cancer Res* 60:5021-5026, 2000; Marchevsky et al., *J Mol Diagn* 6:28-36, 2004; Sarter et al., *Hum Genet* 117:402-403, 2005; Trinh et al., *Methods* 25:456-462, 2001 ; Ogino et al., *Gut* 55: 1000-1006, 2006; Ogino et al., *J Mol Diagn* 8:209-217, 2006, and Woodson, K. et al. *Cancer Epidemiol Biomarkers Prev* 14: 1219-1223, 2005)。

10

20

30

40

50

【0215】

ハイスループットIlluminaプラットフォームは、たとえばヒトES細胞DNA試料の大きな集合 (または他の誘導体および/または前駆細胞集団) における異常なDNAメチル化に関してPRC2標的 (または他の標的) をスクリーニングするために用いることができ、次にMETHYLIGHT (商標) およびMETHYLIGHT (商標) の変化形を用いて、限られた数の遺伝子座で (たとえば、細胞培養および分化の際の特定の数の細胞株において) 異常なDNAメチル化を感度よく検出することができる。

【0216】

Illumina DNAメチル化プロファイリング。Illumina, Inc. (San Diego) は、最近、そのGOLDENGATE (商標) プラットフォームに基づいてフレキシブルDNAメチル化分析技術を開発しており、これは1枚のプレート上で96個の異なる試料に関して1,536個の異なる遺伝子座を照合することができる。最近、Illumina社は、このプラットフォームを用いて、ヒト胚幹細胞における独自のエピジェネティックなシグネチャーを同定できることを報告した (Bibikova, M. et al. *Genome Res* 16: 1075-83, 2006)。それゆえ、Illumina分析プラットフォームが好ましくは用いられる。ハイスループットIlluminaプラットフォームは、たとえばヒトES細胞DNA試料の大きな集合 (または他の誘導体および/または前駆細胞集団) における異常なDNAメチル化に関してPRC2標的 (または他の標的) をスクリーニングするために用いることができ、次にMethyLightおよびMethyLightの変化形を用いて、限られた数の遺伝子座で (たとえば、細胞培養および分化の際の特定の数の細胞株において) 異常なDNAメチル化を感度よく検出することができる。

【0217】

好ましく用いることができるDNAメチル化データの分析およびクラスタリングならびにDNAメチル化マーカー選択には十分な経験がある (たとえば、その全てが参照により本明細書に組み入れられる、Weisenberger, DJ. et al. *Nat Genet* 38:787-793, 2006; Siegmund et al., *Bioinformatics* 25, 25, 2004; Virmani et al. *Cancer Epidemiol Biomarkers Prev* 11 :291-297, 2002; Marjoram et al., *Bioinformatics* 7, 361, 2006); Siegmund

d et al., *Cancer Epidemiol Biomarkers Prev* 15, :567-572, 2006); and Siegmund & Laird, *Methods* 27: 170-178, 2002)。たとえば、ES細胞における腫瘍遺伝子エピジェネティックなサイレンシングの標的であるDNAメチル化マーカーを提供するために、段階的戦略（たとえば、本明細書に組み入れられるWeisenberger et al., *Nat Genet* 38:787-793, 2006）を、本明細書において例示される方法によって教示されるように用いる。

【0218】

単なる例に過ぎないが、メチル化アッセイは、サービス提供者、たとえばエピゲノミクス（Berlin）および他のサービスの提供者が行うことができる。簡単に説明すると、品質管理を試料に行った後、ゲノムDNAを亜硫酸水素ナトリウムによって処置する。明記された遺伝子における関心対象領域に関してPCRプライマーを設計した。選択された関心対象遺伝子、たとえば、表12Aおよび/または表12Cに記載される遺伝子、または表13A、表13B、もしくは表14から選択される任意の遺伝子などのDNAメチル化標的遺伝子を評価する。たとえば、評価される1つのDNAメチル化標的遺伝子が、POU5F1（注釈つきOCT4オルトログヒト遺伝子）およびNANOG遺伝子である場合、POU5F1遺伝子（参照配列：NM.sub.--002701）AMP1000122は、TSSの150 bp上流の注釈つきEnsembl転写物POU5F1_HUMAN（ENST00000259915）の59 UTRに存在する。NANOG遺伝子（参照配列：NM.sub.--024865）AMP 1000123は、TSSの25 bp上流の注釈つきEnsembl転写物NANOG_HUMAN（ENST00000229307）の59 UTRに位置する。以下のパイサルファイトプライマーをPCRおよびシーケンシングのために用いることができる：

POU5F1 5'-ATGGTGTGGTGGGAAGGGG-AA-3' (SEQ ID NO: 1) および 5'-

TCCAAACAACCTAAAATATACAAAACCT-3' (SEQ ID NO: 2); NANOG 5'-

TAATATGAGGTAATTAGTTTGTAGTTTGTAGT-3' (SEQ ID NO: 3) および 5'-

TAATTTCAAACCTCTAACTTCAAATAAT-3' (SEQ ID NO: 4)

【0219】

遺伝子発現プロファイリング

いくつかの態様において、アッセイ、システム、および方法は、マイクロアレイまたはその他などの定量的遺伝子プロファイリングアッセイを含む。当業者に一般的に公知の遺伝子発現レベルを測定するための任意の方法が、本明細書において開示される方法、システム、およびアッセイにおいて用いるために包含され、これらは、Affymetrixマイクロアレイ法、およびDNAまたは転写物発現を測定するための他の方法を含む。いくつかの態様において、遺伝子発現は、cDNAおよびRNAシーケンシング、NanoStringなどのイメージングに基づく方法、およびPCRと共にqPCRを用いる広範囲の方法を用いて測定される。これらの方法の正規化は広く記述されている。本発明者らは、Affymetrixマイクロアレイデータを正規化するためにgcRMAアルゴリズムを用いている。

【0220】

いくつかの態様において、遺伝子発現レベルは遺伝子発現標的遺伝子セットにおいて測定され、遺伝子発現標的遺伝子は癌遺伝子および/または発達遺伝子でありえて、表12Bに開示される。いくつかの態様において、本発明の方法、システム、およびアッセイにおいて測定される遺伝子は、表12Bおよび/または表12Cにおける遺伝子のリストから選択される、または表13A、表13B、もしくは表14に記載される遺伝子のリストから選択される、任意の組み合わせの少なくとも約200個、または少なくとも約300個、または少なくとも約400個、または少なくとも約500個、または少なくとも約600個、または少なくとも約800個、または少なくとも約1000個、または少なくとも約1500個、または少なくとも約2000個、または少なくとも約3000個、または少なくとも約4000個、または少なくとも約5000個の遺伝子である。いくつかの態様において、遺伝子は、表12Bもしくは表12Cに記載される、または表13A、表13B、もしくは表14に記載される遺伝子のリストから選択される遺伝子の1~200番、または1~500番、または1~1000番から選択される遺伝子セットの任意の組み合わせである。

10

20

30

40

50

【 0 2 2 1 】

いくつかの態様において、DNAメチル化は、以下の134個の遺伝子セットを任意に組み合わせた少なくとも50個の遺伝子または少なくとも100個の遺伝子において測定される：
PON3, CD14, PEG3AS, CRCT1, LCE5A,

HIST1, H2BB, HIST1, H3C, CRCT1, LCE5A, PTK2B, TF, CAT, SLC38A11, ZNF528, CALCB, ERAS, INGX, TMPRSS12, ZNF248, ZNF876P, SLC17A3, TDRD5, LCE3A, ASB3, GPR75, ZNF354C, PEG3AS, KAAG1, PCDHA2, HPDL, ZNF737, AGBL2, COMT, TXNRD2, SLC30A8, H2AFZP1, CTSF, ZNF833, S100A5, S100A6, PRDM9, CYP2E1, ZNF177, CR1L, ZNF572, MOS, FAM70A, GP5, PAPOLB, ZDHHC15, HSF5, CDX4, GOLGA8B, KLF8, ARMCX5, CBLN4, POU3F4, LYNX1, DENND2D, CYP2E1, ZNF562, PPYR1, KLHL34, ZNF562, TMLHE, CCDC11, GYG2P, TCEAL2, ZNF454, TRIM4, FAM24B, ZNF397OS, PAQR6, DENND2D, LYNX1, BHMT2, DMGDH, PF4, LTF, NAP1L6, ALOX15B, CES1, PPP1R13L, COMT, TXNRD2, LYNX1, DNAJC15, ARMCX1, TRPM2, GOLGA8A, ZBPB, ZNF630, BHMT2, DMGDH, SLC7A3, SLFN13, PLEK2, DYNLT3, SLC2A14, SPATS1, SLC01A2, TCEAL6, SLC2A14, TAF9B, KIAA1210, CNTD2, PLD6, CFLAR, PHF8, TBPL2, RWDD2B, DEFB124, REM1, TCEAL6, BCL2L10, ZNF630, DCDC2, CRYGD, ZNF440, RFPL2, MYCL2, TRPM2, MEG3, TEKT4, FAM104B, EDNRB, OSGIN1, NKAP, NR0B1, SPIN3, SPIN3, NDUFA1, RNF113A, ZNF726

10

20

【 0 2 2 2 】

代わりの態様において、遺伝子発現は、系列特異的遺伝子セット（たとえば、系列マーカー遺伝子）または発達特異的遺伝子セットにおいて測定および決定され、それによって、多能性幹細胞が特定の発達経路に沿ってまたは該系列マーカーを発現する細胞型へと分化できるか否かを予測することができる。

【 0 2 2 3 】

重要なことに、本明細書において開示される分化指向性アッセイおよび方法において、系列特異的遺伝子セット（たとえば、系列マーカー遺伝子）または発達特異的遺伝子セットの遺伝子発現レベルは、多能性幹細胞株を既定の期間培養して自発的に分化させた後に決定され、系列マーカー遺伝子セットの遺伝子発現アッセイからの結果により、多能性幹細胞株の系列分化の偏りを予測することができる。分化指向性アッセイのいくつかの態様において、系列マーカー遺伝子セットの遺伝子発現アッセイは、特定の系列に沿った定方向性分化後の多能性幹細胞株について行われる。

30

【 0 2 2 4 】

遺伝子発現標的遺伝子が発達遺伝子または系列マーカー遺伝子である場合、その発達遺伝子または系列マーカーの遺伝子発現レベルの正常な変動と比較して高レベルの発現および/または統計学的に有意に高レベルのDNAメチル化は、発達遺伝子または系列マーカーの発現が増加していることを示し、かつ、多能性幹細胞が、発達遺伝子の発達経路に沿って分化するかまたは系列マーカーを発現する細胞型へと分化する素因を有することを示している。同様に、多能性幹細胞における発達遺伝子または系列マーカー遺伝子の遺伝子発現レベルが、該遺伝子の遺伝子発現レベルの正常な変動内である状況では、その情報を用いて、多能性幹細胞が発達遺伝子の発達経路に沿って分化するかまたは系列マーカーを発現する細胞型へと分化するように進行できると予測することができる。したがって、発達遺伝子および/または系列マーカーの遺伝子発現レベルを用いて、多能性幹細胞が、特定の発達経路に沿ってまたは該系列マーカーを発現する細胞型へと分化できるか否かを予測することができる。

40

【 0 2 2 5 】

上記の遺伝子発現の測定は、主に1つの遺伝子の効果に焦点を当てているが、いくつか

50

の態様において、スコアカードは、たとえば細胞株の品質（たとえば、癌様株へと発達する可能性）および有用性（たとえば、関心対象の特定系列に沿って分化するまたは分化しない可能性）を予測するために、遺伝子発現標的遺伝子の組み合わせ（たとえば、表12Aおよび/または表12Cに記載される遺伝子の任意の組み合わせ）、たとえば「癌遺伝子」セットにおける多数の遺伝子または「系列マーカー遺伝子」セットにおける多数の遺伝子の遺伝子発現を測定する。したがって、特定の所望のまたは望ましくない特徴を同定するために、特定の遺伝子発現標的遺伝子セットを選択して、多能性幹細胞株を感度よく正確に特徴決定するための「カスタムスコアカード」を作成することができる。このことは、特定の多能性幹細胞株の品質および有用性を決定するために本明細書において開示されるスコアカードを用いる重要な利点の1つである。

10

【0226】

本明細書において述べるように、いくつかの態様において、多能性幹細胞においてその遺伝子の遺伝子発現の正常な変動（たとえば、正常な参照値）と比較して遺伝子発現レベルの統計学的に有意な（ $FDR < 10\%$ ）および/または1 log-2倍を超える変化の絶対差を有する標的遺伝子の遺伝子発現レベルを有する多能性幹細胞株は、外れ値の遺伝子発現遺伝子であると見なされるであろう。参照多能性幹細胞と比較して、多数の、たとえば少なくとも約5個、または少なくとも約6個、または少なくとも約7個、または少なくとも約8個、または少なくとも約5~10個、または少なくとも約10~15個、または少なくとも約10~50個、または少なくとも約50~100個、またはそれより多くの完全な外れ値の遺伝子発現遺伝子を有する多能性幹細胞は、外れ値の多能性幹細胞であると見なされるであろう。したがって、そのような多能性幹細胞を用いて、負の選択を行う、たとえば望ましくない特徴を有する細胞を単離して廃棄することができる。

20

【0227】

遺伝子発現アッセイ

いくつかの態様において、遺伝子発現は任意の遺伝子レベルで、たとえば非コード遺伝子ならびに非コード転写物、たとえば天然のアンチセンス転写物（NAT）、マイクロRNA（miRNA）遺伝子、ならびに多能性細胞および分化した細胞に正常または異常に存在する全ての他のタイプの核酸および/またはRNA転写物の発現に関して決定される。

【0228】

いくつかの態様において、測定される遺伝子発現レベルが、測定される遺伝子転写物発現レベルである場合、タンパク質発現遺伝子の転写物発現を、メッセンジャーRNA（mRNA）レベルで測定することができる。いくつかの態様において、検出は、たとえばDNA、RNA、PNA、偽相補的DNA（pcDNA）、ロックド（locked）核酸、ならびにその変種および相同体を含むがこれらに限定されるわけではない核酸または核酸アナログを用いる。いくつかの態様において、遺伝子転写物発現は、当業者に一般的に公知の方法によって、逆転写ポリメラーゼ連鎖反応（RT-PCR）または定量的RT-PCRによって評価することができる。

30

【0229】

核酸およびリボ核酸（RNA）分子は、当技術分野において周知である多くの技法のいずれかを用いて特定の生物試料から単離することができ、選択される特定の単離技法は、特定の生物試料にとって適切である。たとえば、凍結融解およびアルカリ溶解技法は、固体材料から核酸分子を得るために有用でありうる；熱アルカリ溶解技法は、尿から核酸を得るために有用でありうる；かつ、プロテイナーゼK抽出は、血液から核酸を得るために用いることができる（Roiff, A et al. PCR: Clinical Diagnostics and Research, Springer (1994)）。

40

【0230】

一般的に、PCR技法は、(i)核酸試料またはライブラリにおける特定遺伝子に対するプライマーの配列特異的ハイブリダイゼーション、(ii)DNAポリメラーゼを用いたアニーリング、伸長、および変性の多数回ラウンドを伴うその後の増幅、ならびに(iii)正確な大きさのバンドに関するPCR産物のスクリーニングを含む、遺伝子増幅法を記述する。用いられるプライマーは、重合化の開始を提供するために十分な長さを有しかつ適切であ

50

る配列のオリゴヌクレオチドであり、すなわち各プライマーは、増幅されるゲノム遺伝子座の各々の鎖に対して相補的となるように特異的に設計される。

【0231】

代わりの態様において、遺伝子発現標的遺伝子は、逆転写(RT)PCRによって、および定量的RT-PCR(QRT-PCR)またはリアルタイムPCR法によって決定することができる。RT-PCRおよびQRT-PCR法は当技術分野において周知であり、以下により詳細に記述される。

【0232】

リアルタイムPCRは、mRNA発現レベルを決定するために用いることができる増幅技術である(たとえば、Gibson et al., Genome Research 6:995-1001, 1996; Heid et al., Genome Research 6:986-994, 1996を参照されたい)。リアルタイムPCRは、増幅の際のPCR産物の蓄積レベルを評価する。この技術は、多数の試料におけるmRNAレベルの定量的評価を可能にする。mRNAレベルに関して、mRNAは、生物試料、たとえば腫瘍および正常組織から抽出され、cDNAは、標準的な技術を用いて調製される。リアルタイムPCRは、たとえばPerkin Elmer/ Applied Biosystems (Foster City, Calif.) 7700 Prism機器を用いて行うことができる。たとえばPerkin Elmer/ Applied Biosystems (Foster City, Calif.) によって提供されるプライマー発現プログラムを用いて、関心対象遺伝子に関してマッチするプライマーおよび蛍光プローブを設計することができる。プライマーおよびプローブの最適な濃度は、当業者が最初に決定することができ、対照(たとえば、アクチン)プライマーおよびプローブは、たとえばPerkin Elmer/ Applied Biosystems (Foster City, Calif.) から購入することにより得ることができる。試料中の特定の関心対象核酸の量を定量するために、対照を用いて検量線を作成する。検量線は、リアルタイムPCRにおいて決定されたCt値を用いて作成することができ、これはアッセイにおいて用いられる関心対象核酸の初回濃度に関連する。関心対象遺伝子の10~106コピーの範囲にわたる標準的な希釈物が一般的に十分である。さらに、対照配列に関して検量線を作成する。これにより、比較目的のために対照の量に対する組織試料中の関心対象核酸の初回含有量の標準化が可能となる。

【0233】

TaqMan(登録商標)プローブを用いるリアルタイム定量的PCR法は、当技術分野において周知である。リアルタイム定量的PCRに関する詳細なプロトコールは、たとえばRNAに関してGibson et al., 1996, A novel method for real time quantitative RT-PCR. Genome Res., 10:995-1001において、およびDNAに関してHeid et al., 1996, Real time quantitative PCR. Genome Res., 10:986-994に提供される。

【0234】

TaqManに基づくアッセイは、5'蛍光色素および3'消光剤を含む蛍光発生オリゴヌクレオチドプローブを用いる。プローブは、PCR産物にハイブリダイズするが、それ自身は3'末端でのブロック剤により伸長することができない。PCR産物をその後のサイクルにおいて増幅すると、ポリメラーゼ、たとえば、AmpliTaq(登録商標)の5'ヌクレアーゼ活性により、TaqManプローブの切断が起こる。この切断は5'蛍光色素と3'消光剤とを分離して、それによって増幅の関数として蛍光の増加が起こる(たとえば、ワールドワイドウェブサイト: perkin-elmer-dot-comを参照されたい)。

【0235】

もう1つの態様において、RNA転写物の検出は、ノザンプロットティングによって達成することができ、この場合、RNAの調製を変性アガロースゲルにおいて行って、活性化セルロース、ニトロセルロースまたはガラスもしくはナイロンメンブレンなどの適切な支持体に転写する。標識された(たとえば、放射標識された)cDNAまたはRNAを調製物にハイブリダイズさせて、洗浄し、オートラジオグラフィーなどの方法によって分析する。

【0236】

RNA転写物の検出はさらに、公知の増幅法を用いて達成することができる。たとえば、mRNAをcDNAに逆転写させた後、ポリメラーゼ連鎖反応(RT-PCR)を行うこと、または米国特許第5,322,770号に記述されるように両方の段階のために1つの酵素を用いること、また

10

20

30

40

50

はR. L. Marshall, et al., PCR Methods and Applications 4: 80-84 (1994)によって記述されるように、mRNAをcDNAに逆転写させた後、対称ギャブリガーゼ連鎖反応(RT-AGL CR)を行うことは本発明の範囲内である。酵素mRNA転写物を検出するための1つの適切な方法は、その全内容が参照により本明細書に組み入れられる、参考文献Pabic et. al. Hepatology, 37(5): 1056-1066, 2003において記述される。

【0237】

本明細書において使用することができる他の公知の増幅法には、PNAS USA 87: 1874-1878 (1990) および同様にNature 350 (No. 6313): 91-92 (1991)に記述されるいわゆる「NASBA」または「3SR」技術、公表された欧州特許出願(EPA)第4544610号に記述されるQ-増幅、鎖置換増幅(G. T. Walker et al., Clin. Chem. 42: 9-13 (1996)および欧州特許出願第684315号に記述される)、およびPCR公開番号WO 9322461に記述される標的媒介増幅が挙げられるがこれらに限定されるわけではない。

10

【0238】

インサイチュ-ハイブリダイゼーションの可視化も同様に使用することができ、この場合、放射活性標識アンチセンスRNAプローブを生検標本の薄片にハイブリダイズさせて、洗浄し、RNアーゼによって切断して、オートラジオグラフィーのために感受性乳剤に曝露する。試料の組織学組成を証明するために、試料をヘマトキシリンによって染色することができ、適切な光フィルターによる暗視野イメージングにより、現像された乳剤を示す。ジゴキシゲニンなどの非放射活性標識も同様に用いることができる。

【0239】

または、mRNA発現をDNAアレイ、チップ、またはマイクロアレイにおいて検出することができる。そのような態様において、プローブを「遺伝子チップ」として用いるために表面に付着させることができる。そのような遺伝子チップを用いて当業者に公知の多数の技術によって遺伝子変動を検出することができる。1つの技術において、オリゴヌクレオチドを、米国特許第6,025,136号および第6,018,041号に概要されるアプローチなどのハイブリダイゼーションアプローチによるシーケンシングによりDNA配列を決定するために遺伝子チップ上に配置する。本発明のプローブはまた、遺伝子配列の蛍光検出のためにも用いることができる。そのような技術は、たとえば米国特許第5,968,740号および第5,858,659号に記述されている。プローブはまた、Kayyem et al., 米国特許第5,952,172号およびKelley, S.O. et al. (1999) Nucleic Acids Res. 27:4830-4837によって記述される配列などの核酸配列の電気化学的検出のために電極表面に付着させることができる。

20

30

【0240】

遺伝子発現標的遺伝子に対応するオリゴヌクレオチドを、チップ上に固定して、次にこれを患者から得られた試験試料の標識核酸とハイブリダイズさせる。遺伝子発現標的遺伝子mRNA転写物を含む試料について陽性ハイブリダイゼーションシグナルが得られる。DNAアレイを調製する方法およびその用途は、当技術分野において周知である(たとえば、その全内容が参照により本明細書に組み入れられる、米国特許第6,618,679号;第6,379,897号;第6,664,377号;第6,451,536号;第548,257号;U.S.20030157485およびSchena et al. 1995 Science 20:467-470; Gerhold et al. 1999 Trends in Biochem. Sci. 24, 168-173; and Lennon et al. 2000 Drug discovery Today 5: 59-65を参照されたい)。遺伝子発現連鎖解析(SAGE)も同様に、行うことができる(たとえば、米国特許出願第2003021585号を参照されたい)。

40

【0241】

マイクロアレイ

マイクロアレイは、個別領域、典型的には核酸のアレイであり、互いから離れており、典型的には約100/cm²から1000/cm²のあいだの密度で配置されるが、より高い密度、例えば10000/cm²で配置してもよい。マイクロアレイ実験の原理は、所定の細胞株または組織からのmRNAを用いて標識試料、典型的に「標的」と呼ばれる標識cDNAを作製し、これを、整列配置で固体表面上に固定された多数の核酸配列、典型的にはDNA配列に同時にハイブリダイズさせることである。

50

【0242】

10000個の転写物種を同時に検出および定量することができる。多くの異なるマイクロアレイシステムが開発されているが、今日最も一般的に用いられるシステムは、配置される材料に従って2つの群、すなわち相補的DNA (cDNA) マイクロアレイとオリゴヌクレオチドマイクロアレイに分けられる。アレイ材料は一般的に、それがノザンプロット分析において用いられるプローブと同等であることからプローブと呼ばれている。cDNAアレイのプローブは通常、ベクター特異的または遺伝子特異的なプライマーのいずれかを用いてcDNAライブラリまたはクローン集合体から作製されたポリメラーゼ連鎖反応 (PCR) の産物であり、スライドガラスまたはナイロンメンブレン上に既定の位置でスポットとしてプリントされる。スポットは典型的に、大きさが10~300 μmであり、ほぼ同じ距離離れて存在する。この技術を用いて、30,000個より多くのcDNAからなるアレイを従来の顕微鏡スライドガラスの表面上に適合させることができる。オリゴヌクレオチドアレイに関して、短い20~25量体を、シリコンウェーハへのフォトリソグラフィ (Affymetrixの高密度オリゴヌクレオチドアレイ) またはインクジェット技術 (Rosetta Inpharmaticsが開発して、Agilent Technologiesに販売認可されている) のいずれかによってインサイチューで合成する。

10

【0243】

または、予め合成されたオリゴヌクレオチドをスライドガラス上にプリントすることができる。合成オリゴヌクレオチドに基づく方法は、配置されるDNAを作製するのに配列情報のみで十分であるので、時間のかかるcDNAリソースの取り扱いが不要であるという利点を提供する。また、所定の転写物の最も特異的な部分を表すようにプローブを設計することができ、これにより、密接に関連する遺伝子またはスプライス変種の検出が可能となる。短いオリゴヌクレオチドは、ハイブリダイゼーション特異性の低下および感度の低下をもたらす。最近、これらの短所を克服するために、より長い予め合成されたオリゴヌクレオチド (50~100量体) の配置が開発されている。

20

【0244】

このように、多能性幹細胞における標的遺伝子発現遺伝子の遺伝子発現レベルを確認するためマイクロアレイを行う場合、以下の段階を行うことができる：多能性幹細胞を含む試料からmRNAを得て、核酸標的を調製する段階、アレイを、典型的にはマイクロアレイの製造元によって提唱される条件 (好適には3×SSC、0.1% SDS、50 などのストリンジェントなハイブリダイゼーション条件) 下で接触させてアレイ上の対応するプローブに結合させる段階、必要であれば未結合の核酸標的を除去するために洗浄する段階、および結果を分析する段階。

30

【0245】

mRNAは、プライマー特異的cDNA合成などの当技術分野において公知の方法によって、本明細書において記述される遺伝子プロファイルに存在する配列などの関心対象配列に関して濃縮されうると認識されるであろう。集団はさらに、たとえばPCR技術を用いることによって増幅されうる。標的またはプローブは、標的分子のマイクロアレイに対するハイブリダイゼーションを検出できるように標識される。適切な標識は、プローブに組み入れることができる同位元素または蛍光標識を含む。

40

【0246】

Affymetrix HG-U133. Plus2.0遺伝子チップを、Affymetrix社の標準プロトコールに従って用いてハイブリダイズさせ、洗浄して、走査することができる。いくつかのRNAをアレイ上で複製させることができ、その後の分析のために利用可能なハイブリダイゼーションの総数は96となる。

【0247】

mRNAレベルをモニターするために、たとえば、mRNAを、試験される多能性幹細胞を含む試料から抽出して、逆転写させて、蛍光標識cDNAプローブを作製する。次に、遺伝子発現標的cDNAにハイブリダイズすることができるマイクロアレイを標識cDNAプローブによってプローブして、スライドガラスを走査して、蛍光強度を測定する。この強度は、ハイブリ

50

ダイゼーション強度および発現レベルと相関する。

【0248】

「定量的」増幅法は、当業者に周知である。たとえば、定量的PCRは、同じプライマーを用いて対照配列の公知の品質を同時に増幅する段階を伴う。これは、PCR反応を較正するために用いることができる内部標準を提供する。定量的PCRのための詳細なプロトコールは、たとえばInnis et al. (1990) PCR Protocols, A Guide to Methods and Applications, Academic Press, Inc. N.Y.に提供される。

【0249】

Affymetrix社によって記述される同じ技法およびハードウェアを、本発明に関連して用いることができるが、他の代替法も同様に利用可能である。マイクロアレイを作製する方法およびアッセイを行う方法を詳述する多くの論評が記述されている（たとえば、Bowtell, Nature Genetics Suppl. 27:25-32 (1999); Constantine, et al., Life Sci News 7: 11-13 (1998); Ramsay, Nature Biotechnol. 16:40-44 (1998)を参照されたい）。さらに、マイクロアレイプレート、スライド、および関連機器を作製する技術（その全内容が参照により本明細書に組み入れられる、米国特許第6,902,702号；第6,594,432号；第5,622,826号）、およびアッセイを行う技術（その全内容が参照により本明細書に組み入れられる、米国特許第6,902,900号；第6,759,197号）を記述する特許が発行されている。プレートまたはスライドを作製する2つの主な技術は、ポリリソグラフィ法（その全内容が参照により本明細書に組み入れられる、米国特許第5,445,934号；第5,744,305号を参照されたい）、またはロボットスポット法（その全内容が参照により本明細書に組み入れられる、米国特許第5,807,522号）のいずれかを伴う。他の技法は、インクジェットプリンティングまたはキャピラリースポットティング（たとえば、その全内容が参照により本明細書に組み入れられるWO 98/29736またはWO 00/01859を参照されたい）を伴う。

【0250】

マイクロアレイプレートまたはスライドのために用いられる基板は、プラスチック、白金などの金属、およびガラスを含む、オリゴヌクレオチドに結合して固定することができる任意の材料でありうる。好ましい基板は、ポリリジンなどのオリゴヌクレオチド結合を促進する材料によってコーティングされたガラスである（Chena, et al., Science 270:467-470 (1995)を参照されたい）。オリゴヌクレオチドを共有結合させる多くのスキームが記述されており、本発明に関連して用いるために適している（たとえば、その全内容が参照により本明細書に組み入れられる、米国特許第6,594,432号を参照されたい）。固定されたオリゴヌクレオチドは、長さが少なくとも20塩基でなければならず、かつハイブリダイゼーションの標的とされる遺伝子中のセグメントに正確に対応する配列を有しなければならない。

【0251】

分化指向性アッセイ

本明細書において開示されるように、スコアカードを作成するために本明細書において開示される方法、システム、およびアッセイは、任意で分化指向性アッセイを含むことができる。いくつかの態様において、たとえばDNAメチル化アッセイおよび遺伝子発現アッセイを、分化指向性アッセイの後に行うことができる。いくつかの態様において、分化指向性アッセイは、所望の細胞系列に沿って分化することをユーザーが既に知っている多能性幹細胞株の品質（たとえば、安全性）を決定することが関心対象である場合には省略することができる。

【0252】

一般的に、分化指向性アッセイは、多能性幹細胞株を既定の期間、異なる系列に沿って自発的に分化させて、分化した細胞からの核酸材料を収集して、これを本明細書において述べるようにDNAメチル化アッセイおよび/または遺伝子発現アッセイの開始材料として用いる。代替の態様において、分化指向性アッセイはまた、特定系列（たとえば、ニューロン系列、膵臓系列、心臓系列等）に沿った既定の期間の多能性幹細胞株の直接の分化、その後分化した細胞からの核酸材料を収集してDNAメチル化アッセイおよび/または遺

伝子発現アッセイの開始材料として用いることを包含する。いくつかの態様において、分化指向性アッセイは、分化した細胞を、本明細書において開示されるDNAメチル化アッセイおよび/または遺伝子発現アッセイにおいて処理する前に、少なくとも0日間、または約1日間、または約2日間、または約3日間、または約4日間、または約5日間、または約6日間、または約7日間、または約8日間、または約8~10日間、または約10~12日間、または約12~14日間、または約14~16日間、または約16~20日間、または20日間超の多能性幹細胞株の自発的または直接の分化を包含する。

【0253】

分化指向性アッセイにおいて、DNAメチル化アッセイおよび/または遺伝子発現アッセイはそれぞれ、本明細書において開示される多様な系列マーカー遺伝子および/または発達遺伝子におけるDNAメチル化および遺伝子発現の測定に関して行われる。いくつかの態様において、DNAメチル化アッセイおよび/または遺伝子発現は、表7に記載される複数の系列マーカー遺伝子および/または発達遺伝子において測定される。

10

【0254】

本明細書において述べるように、いくつかの態様において、多能性幹細胞においてその系列遺伝子の遺伝子発現の正常な変動（たとえば、正常な参照値）と比較して系列遺伝子発現レベルの統計学的に有意な（FDR < 5%）および/または1 log-2倍を超える変化の絶対差を有する系列遺伝子の遺伝子発現レベルを有する多能性幹細胞株は、外れ値の分化遺伝子であると見なされるであろう。参照多能性幹細胞と比較して、多数の、たとえば少なくとも約5個、または少なくとも約6個、または少なくとも約7個、または少なくとも約8個、または少なくとも約5~10個、または少なくとも約10~15個、または少なくとも約10~50個、または少なくとも約50~100個、またはそれより多くの完全な外れ値の系列遺伝子発現遺伝子を有する多能性幹細胞は、参照多能性幹細胞株と同じ系列に沿って分化しない可能性がある外れ値の多能性幹細胞であると見なされるであろう。したがって、そのような多能性幹細胞を用いて、負の選択を行う、たとえば望ましくない特徴を有する細胞、たとえば特定の系列に沿って分化しない細胞を単離して廃棄することができる。

20

【0255】

いくつかの態様において、本発明の方法において用いるために自発的に分化させるために培養される多能性幹細胞は、たとえば形態学に関しておよび培地交換のために、毎日モニターすることができる。さらなる分析および検証を、ルーチンで5回継代毎のアルカリホスファターゼ、OCT4、NANOG、TRA-160、TRA-181、SEAA-4、CD30、および10~15回継代毎のG-バンド分染法による核型を含む幹細胞マーカーについて任意で行うと、多能性幹細胞が多能性幹細胞から分化しているか否かが同定されるであろう。

30

【0256】

さらなる局面において、多能性幹細胞を、異なる分化プロトコルの条件下で培養して、多能性幹細胞に異常なエピジェネティックな変化を獲得させやすくするその傾向に関して分析する。たとえば、フィーダー層を交換せずに4から7週間の高密度での培養などの最適以下の培養条件で維持することによる非定方向性分化を、そのような傾向を有する例示的な状態として分析する。このおよび/または他の培養条件および/またはプロトコルに関して、たとえばDNA試料を、同時に行われた分化培養から定期的な間隔で採取して、異常なエピジェネティックな変化の進行を調べる。同様に、神経系列32'33、膵臓系列（Segev et al., J. Stem Cells 22:265-274, 2004; and Xu, X. et al. Cloning Stem Cells 8:96-107, 2006、参照により本明細書に組み入れられる）および/または心筋細胞（Yoon, B. S. et al. Differentiation 74: 149-159, 2006; and Beqqali et al., Stem Cells 24: 1956-1967, 2006、参照により本明細書に組み入れられる）への分化などの定方向性分化プロトコルを、ES細胞が異常なエピジェネティックな変化を獲得しやすい傾向があるか否かに関して分析することができる。

40

【0257】

いくつかの態様において、多能性幹細胞株は、1つまたは複数の異なる系列に沿って分化するように指示される。いくつかの態様において、多能性幹細胞株の分化は、本明細書

50

において開示されるDNAメチル化および/または遺伝子発現アッセイによって評価することができる。代替りの態様において、多能性幹細胞株の分化は、当業者に一般的に公知の免疫染色およびイムノアッセイによって評価することができる。例示的なイムノアッセイには、酵素免疫測定法 (ELISA)、ラジオイムノアッセイ (RIA)、免疫放射測定アッセイ (IRMA)、ウェスタンブロッティング、免疫細胞化学、または免疫組織化学が挙げられ、その各々は以下により詳細に記述される。極めて迅速でありうるELISAまたはRIAなどのイムノアッセイがより一般的に好ましい。抗体アレイまたはタンパク質チップも同様に使用することができ、たとえば、その全内容が参照により本明細書に組み入れられる、米国特許出願第20030013208 AI号；第20020155493A1号；第20030017515号および米国特許第6,329,209号；第6,365,418号を参照されたい。

10

【0258】

イムノアッセイ：最も一般的な酵素イムノアッセイは、「酵素免疫測定法 (ELISA)」である。ELISAは、標識 (たとえば、酵素結合) 型の抗体を用いて抗原の濃度を検出および測定する技術である。ELISAには異なる型があり、それらは当業者に周知である。ELISAに関して当技術分野において公知の標準的な技術は、"Methods in Immunodiagnosis", 2nd Edition, Rose and Bigazzi, eds. John Wiley & Sons, 1980; Campbell et al., "Methods and Immunology", W. A. Benjamin, Inc., 1964; and Oellerich, M. 1984, J. Clin. Chem. Clin. Biochem., 22:895-904において記述されている。「サンドイッチELISA」では、抗体 (たとえば、抗酵素) を固相 (すなわち、マイクロタイタープレート) に連結させて、抗原 (たとえば、酵素) を含む生物試料に曝露する。次に固相を洗浄して、未結合抗原を除去する。標識抗体 (たとえば、酵素結合) を結合抗原 (存在すれば) に結合させて、抗体-抗原-抗体サンドイッチを形成する。抗体に連結させることができる酵素の例は、アルカリホスファターゼ、西洋ワサビペルオキシダーゼ、ルシフェラーゼ、ウレアーゼ、および α -ガラクトシダーゼである。酵素結合抗体は、基質と反応して、測定することができる着色反応産物を生じる。

20

【0259】

「競合的ELISA」において、抗体は抗原 (すなわち、酵素) を含む試料と共にインキュベートされる。次に、抗原-抗体混合物を、抗原 (すなわち、酵素) をコーティングした固相 (たとえば、マイクロタイタープレート) に接触させる。試料中により多くの抗原が存在すれば、固相に結合するために利用できる遊離の抗体はより少なくなるであろう。標識された (たとえば、酵素結合) 二次抗体を固相に添加して、固相に結合した一次抗体の量を決定する。

30

【0260】

「免疫組織化学アッセイ」において、アッセイされるタンパク質に対して特異的である抗体に組織を曝露することによって、組織の切片を特定タンパク質に関して調べる。次に、抗体を多数の方法のいずれかによって可視化して、タンパク質の有無および存在するタンパク質の量を決定する。抗体を可視化するために用いられる方法の例は、抗体に結合した酵素 (たとえば、ルシフェラーゼ、アルカリホスファターゼ、西洋ワサビペルオキシダーゼ、または α -ガラクトシダーゼ) による方法、または化学法 (たとえば、DAB/基質色原体) である。次に、試料を顕微鏡によって、最も好ましくは可視スペクトルで検出される染料によって染色した試料を、そのような多様な染色法のいずれかおよび当業者に公知の試薬を用いて光学顕微鏡により分析する。

40

【0261】

または、「ラジオイムノアッセイ」を使用することができる。ラジオイムノアッセイは、標識 (たとえば、放射活性または蛍光標識) 型の抗原を用いて抗原の濃度を検出および測定する技術である。抗原に関する放射活性標識の例には、 ^3H 、 ^{14}C 、および ^{125}I が挙げられる。生物試料中の抗原酵素の濃度は、生物試料中の抗原を、抗原に対する抗体との結合に関して標識 (たとえば、放射活性) 抗原と競合させることによって測定される。標識抗原と非標識抗原のあいだの競合的結合を確保するために、標識抗原は、抗体の結合部位を飽和するために十分な濃度で存在する。試料中の抗原濃度がより高ければ、抗体に結合

50

する標識抗原の濃度はより低くなるであろう。

【0262】

ラジオイムノアッセイにおいて、抗体に結合した標識抗原の濃度を決定するためには、抗原-抗体複合体を、遊離の抗原から分離しなければならない。抗原-抗体複合体を遊離の抗原から分離する1つの方法は、抗アイソタイプ抗血清によって抗原-抗体複合体を沈殿させることによる。抗原-抗体複合体を遊離の抗原から分離するもう1つの方法は、ホルマリノ殺菌黄色ブドウ球菌による抗原-抗体複合体の沈殿による。抗原-抗体複合体を遊離の抗原から分離するなもう1つの方法は、抗体をセファロースビーズ、ポリスチレンウェル、ポリ塩化ビニルウェル、またはマイクロタイターウェルに結合させる（たとえば共有結合により）「固相ラジオイムノアッセイ」を行うことによる。抗体に結合した標識抗原の濃度を、抗原の公知の濃度を有する試料に基づく検量線と比較することにより、生物試料中の抗原の濃度を決定することができる。

10

【0263】

「免疫放射測定アッセイ」(IRMA)は、抗体試薬が放射活性標識されるイムノアッセイである。IRMAは、タンパク質、たとえばウサギ血清アルブミン(RSA)とのコンジュゲーションなどの技術による多価抗原コンジュゲートの産生を必要とする。多価抗原コンジュゲートは、1分子あたり、少なくとも2個の抗原残基を有しなければならない、抗原残基は、少なくとも2つの抗体が抗原に結合できるように十分な距離離れていなければならない。たとえば、IRMAにおいて多価抗原コンジュゲートを、プラスチック球体などの固相表面に結合させることができる。非標識「試料」抗原および抗原に対する放射活性標識抗体を、多価抗原コンジュゲートをコーティングした球体を含む試験管に加える。試料中の抗原は、抗原抗体結合部位に対して多価抗原コンジュゲートと競合する。適切なインキュベーション期間の後、未結合反応体を洗浄によって除去して、固相上の放射活性量を決定する。結合した放射活性抗体の量は、試料中の抗原濃度と反比例する。

20

【0264】

分化した多能性幹細胞集団によって発現される系列マーカーレベルを検出するために用いることができる他の技術を、実施者の好みに従って行うことができる。そのような技術の1つは、ウェスタンブロッティング(Towbin et al., Proc. Nat. Acad. Sci. 76:4350 (1979))であり、この場合、適切に処置された試料をSDS-PAGEゲルにおいて泳動させた後、ニトロセルロースフィルターなどの固相支持体に転写する。次に、検出可能に標識された抗体またはタンパク質結合分子を用いて、発現された系列マーカーレベルを評価することができ、検出可能な標識からのシグナル強度は、発現された系列マーカーの量に対応する。存在する発現された系列マーカーの量のレベルも同様に、たとえば密度測定によって定量することができる。

30

【0265】

1つの態様において、生物試料中の発現された系列マーカーレベルは、MALDI/TOF(飛行時間型)、SELDI/TOF、液体クロマトグラフィー-質量分析(LC-MS)、ガスクロマトグラフィー-質量分析(GC-MS)、高速液体クロマトグラフィー-質量分析(HPLC-MS)、キャピラリー電気泳動-質量分析、核磁気共鳴分光法、またはタンデム質量分析(たとえば、MS/MS、MS/MS/MS、ESI-MS/MS、等)などの質量分析によって決定することができる。たとえば、参照により本明細書に組み入れられる、米国特許出願第20030199001号、第20030134304号、第20030077616号を参照されたい。特定の態様において、これらの方法論を機器、コンピュータシステムおよび媒体と組み合わせ、多能性幹細胞集団において発現された発現系列マーカーレベルを測定するための自動システムを作製することができ、および生物試料中のタンパク質発現レベルを同定する印刷可能なレポートを作成するための分析を行うことができる。

40

【0266】

スコアカードの作成に用いるための、またはスコアカードとの比較により機能性を決定するための多能性幹細胞

本明細書において開示される方法、キット、システム、およびスコアカードを用いて、

50

任意の種、たとえばヒトなどの哺乳動物種由来の任意の多能性幹細胞を検証およびモニターすることができる。

【0267】

一般的に、方法、アッセイ、システム、キットにおいて用いるためのおよびスコアカードを作成するための多能性幹細胞は、任意の入手可能な起源から得るまたはそれらに由来することができる。したがって、多能性細胞は、脊椎動物または無脊椎動物から得るまたはそれらに由来することができる。いくつかの態様において、多能性幹細胞は、哺乳動物多能性幹細胞である。本明細書において開示される全ての局面において、方法、アッセイにおいて用いるための、およびスコアカードの作成、または本明細書において開示される既存のスコアカードとの比較に用いるための多能性幹細胞は、任意の多能性幹細胞でありうる。たとえば、多能性幹細胞は、脊椎動物または無脊椎動物から得るまたはそれらに由来することができる。本発明の局面のいくつかの態様において、多能性幹細胞は哺乳動物多能性幹細胞である。

10

【0268】

本発明の局面のいくつかの態様において、多能性幹細胞は、霊長類または齧歯類多能性幹細胞である。本発明の局面のいくつかの態様において、多能性幹細胞は、チンパンジー、カニクイザル、クモザル、マカク（たとえば、アカゲザル）、マウス、ラット、ウッドチャック、フェレット、ウサギ、ハムスター、ウシ、ウマ、ブタ、シカ、バイソン、バッファロー、ネコ（たとえば、ネコ）、イヌ（たとえば、イヌ、キツネ、およびオオカミ）、トリ（たとえば、ニワトリ、エミュ、およびダチョウ）、および魚（たとえば、マス、ナマズ、およびサケ）の多能性幹細胞からなる群より選択される。

20

【0269】

本発明の局面のいくつかの態様において、多能性幹細胞はヒト多能性幹細胞である。いくつかの態様において、多能性幹細胞は、当業者に公知のヒト幹細胞株である。いくつかの態様において、多能性幹細胞は、人工多能性幹（iPS）細胞であるか、または、中間の多能性幹細胞でありiPS細胞へとさらにリプログラムすることができる、安定にリプログラムされた細胞、たとえば部分的人工多能性幹細胞（「piPS細胞」とも呼ばれる）である。いくつかの態様において、多能性幹細胞、iPSCまたはpiPSCは、遺伝子改変多能性幹細胞である。

【0270】

いくつかの態様において、本発明において用いられる多能性幹細胞の多能性状態を、様々な方法によって確認することができる。たとえば、細胞を特徴的なES細胞マーカーの有無に関して試験することができる。ヒトES細胞の場合、そのようなマーカーの例は前記で明らかにされ、例えばSSEA-4、SSEA-3、TRA-1-60、TRA-1-81、およびOCT 4であり、これらは当技術分野において公知である。

30

【0271】

同様に多能性は、細胞を適切な動物、たとえばSCIDマウスに注射して、分化した細胞および組織の産生を観察することによって確認することができる。多能性を確認するなおもう1つの方法は、キメラ動物を作製するために本発明の多能性細胞を用いること、および異なる細胞型に対する導入された細胞の関与を観察することである。キメラ動物を作製する方法は、当技術分野において周知であり、参照により本明細書に組み入れられる米国特許第6,642,433号において記述される。

40

【0272】

多能性を確認するなおもう1つの方法は、分化に都合がよい条件下（たとえば、線維芽細胞フィーダー層の除去）で培養した場合の、ES細胞の胚様体および他の分化した細胞型への分化を観察することである。この方法は利用されており、本発明の多能性幹細胞が組織培養において胚様体および異なる分化した細胞型を生じることを確認している。

【0273】

得られた多能性細胞および細胞株、好ましくは完全に雌性起源のDNAに由来するヒト多能性細胞および細胞株は、多数の治療応用および診断応用を有する。そのような多能性細

50

胞は、多数の疾患状態の処置において細胞移植治療または遺伝子治療（遺伝子改変される場合）のために用いられる。

【0274】

この点において、いくつかのマウス胚幹（ES）細胞が、他の細胞型と比較してより高い効率でいくつかの細胞型へと分化する指向性を有することは公知である。同様に、ヒト多能性（ES）細胞は、類似の選択的分化能を保有する。したがって、本発明を用いて、多能性幹細胞の所望の使用のために所望の特徴および分化指向性を有する多能性幹細胞を同定および選択することができる。たとえば、多能性細胞株が本発明の方法に従ってスクリーニングされている場合、特定の細胞株に沿って分化する効率の増加（ならびに、腫瘍遺伝子のエピジェネティックなサイレンシング、腫瘍抑制遺伝子および/または特定の発達遺伝子の低メチル化などの他の望ましい特徴）によって、多能性幹細胞を選択することができ、公知の方法に従って所望の細胞型を得るために分化誘導することができる。たとえば、ヒト多能性幹細胞、たとえばES細胞またはiPS細胞を、当業者に公知の方法に従って、細胞の分化を提供する分化培地および条件下で培養することによって、造血幹細胞、筋細胞、心筋細胞、肝細胞、島細胞、網膜細胞、軟骨細胞、上皮細胞、尿管細胞等へと分化誘導することができる。ES細胞の分化が起こる培地および方法は、適切な培養条件と同様に当技術分野において公知である。

10

【0275】

いくつかの態様において、多能性幹細胞は、人工多能性幹細胞（たとえば、iPS細胞）または安定な部分的リプログラム細胞、たとえばpiPSCである。いくつかの態様において、本明細書において開示される安定なリプログラム細胞は、体細胞の不完全なリプログラミングから産生することができる。いくつかの態様において、体細胞はヒト細胞であり、たとえば病態を有する対象から得られたまたは疾患もしくは障害を有する遺伝的素因を有するもしくはそれらのリスクがある対象から得られた、疾患を有する体細胞でありうる。

20

【0276】

たとえば、その内容の全体が参照により本明細書に組み入れられる、国際特許出願WO2007/069666；WO2008/118820；WO2008/124133；WO2008/151058；WO2009/006997；および米国特許出願US2010/0062533；US2009/0227032；US2009/0068742；US2009/0047263；US2010/015705；US2009/0081784；US2008/0233610；US7615374；米国特許出願第12/595,041号、EP2145000、CA2683056、AU8236629、12/602,184、EP2164951、CA2688539、US2010/0105100；US2009/0324559；US2009/0304646、US2009/0299763、US2009/0191159に開示されるように、体細胞をiPS細胞またはpiPSC細胞へとリプログラムするために任意の方法を用いることができる。いくつかの態様において、方法、アッセイにおいて用いるための、およびスコアカードの作成、または本明細書において開示される既存のスコアカードとの比較に用いるためのiPS細胞は、その内容の全体が参照により本明細書に組み入れられる、EP1970446、US2009/0047263、US2009/0068742、および2009/0227032に開示されるように、細胞をリプログラムするための当技術分野において公知の任意の方法によって、たとえばリプログラム細胞のウイルス誘導または化学誘導による生成によって、産生することができる。

30

【0277】

いくつかの態様において、方法、アッセイにおいて用いるための、およびスコアカードの作成、または本明細書において開示される既存のスコアカードとの比較に用いるためのiPS細胞は、その内容の全体が参照により本明細書に組み入れられる、WO2010/033906に開示される方法などの、化学的リプログラミングによる体細胞の不完全なリプログラミングから産生することができる。代替の態様において、本明細書において開示される安定なリプログラム細胞は、その全内容が参照により本明細書に組み入れられる、WO2010/048567に開示される方法などの非ウイルス手段による体細胞の不完全なリプログラミングから産生することができる。

40

【0278】

方法、アッセイにおいて用いるための、およびスコアカードの作成、または本明細書において開示される既存のスコアカードとの比較に用いるための他の多能性幹細胞は、当業

50

者に公知の任意の多能性幹細胞でありうる。例示的な幹細胞には、胚幹細胞、成体幹細胞、多能性幹細胞、神経幹細胞、肝幹細胞、筋幹細胞、筋前駆幹細胞、内皮前駆細胞、骨髄幹細胞、軟骨形成幹細胞、リンパ幹細胞、間葉幹細胞、造血幹細胞、中枢神経系幹細胞、および末梢神経系幹細胞などが挙げられる。幹細胞を単離および培養する方法を含む、幹細胞に関する記述は、中でも、Embryonic Stem Cells, Methods and Protocols, Turksen, ed., Humana Press, 2002; Weisman et al., Annu. Rev. Cell. Dev. Biol. 17:387 403; Pittinger et al., Science, 284: 143 47, 1999; Animal Cell Culture, Masters, ed., Oxford University Press, 2000; Jackson et al., PNAS 96(25): 14482 86, 1999; Zuk et al., Tissue Engineering, 7:211 228, 2001 (「Zuk et al.」); Atala et al., particularly Chapters 33 41 ;ならびに米国特許第5,559,022号、第5,672,346号、および第5,827,735号において見いだされうる。間質細胞を単離する方法を含む、間質細胞に関する記述は、中でもProckop, Science, 276:71 74, 1997; Theise et al., Hepatology, 31 :235 40, 2000; Current Protocols in Cell Biology, Bonifacino et al., eds., John Wiley & Sons, 2000 (2002年3月までの改訂版を含む) ; および米国特許第4,963,489号に見いだされうる。当業者は、混合SVF細胞またはSVF-マトリクス構築物(たとえば、本明細書において開示される構築物および方法に従って組織または細胞移植物を封入するための)と共に移植物に含めるために選択される幹細胞および/または間質細胞が、典型的にその構築物の意図される用途にとって適切であることを理解するであろう。

10

【0279】

方法、アッセイにおいて用いるための、およびスコアカードの作成、または本明細書において開示される既存のスコアカードとの比較に用いるためのさらなる多能性幹細胞は、任意の種類(たとえば胎児もしくは前胎児組織などの胚組織または成体組織)に由来する任意の細胞であり、幹細胞は適切な条件下で3つの胚葉(内胚葉、中胚葉、および外胚葉)の全ての派生物である異なる細胞型の子孫を産生することができるという特徴を有する。これらの細胞型は、確立された細胞株の形で提供されてもよく、またはそれらを初代培養胚組織から直接得て、分化のために直ちに用いてもよい。NIHヒト胚幹細胞レジストリに記載される細胞、たとえばhESBGN-01、hESBGN-02、hESBGN-03、hESBGN-04 (BresaGen, Inc.) ; HES-1、HES-2、HES-3、HES-4、HES-5、HES-6 (ES Cell International) ; Miz-hES1 (MizMedi Hospital-Seoul National University) ; HSF-1、HSF-6 (University of California at San Francisco) ; およびH1、H7、H9、H13、H14 (Wisconsin Alumni Research Foundation (WiCell Research Institute)) が含まれる。いくつかの態様において、方法、アッセイ、システムにおいて用いるための、およびスコアカードの作成、または本明細書において開示される既存のスコアカードとの比較に用いるための多能性幹細胞を得る際に、胚は破壊されていない。

20

30

【0280】

もう1つの態様において、幹細胞、たとえば成体幹細胞または胚幹細胞は、幹細胞源として文献においてまだ同定されていない固体組織(固体組織の例外は、血液、血漿、および骨髄を含む全血である)を含む組織から単離することができる。いくつかの態様において、組織は心臓または心組織である。他の態様において、組織はたとえば、臍帯血、胎盤、骨髄、または軟骨絨毛であるが、これらに限定されるわけではない。

40

【0281】

方法、アッセイ、システムにおいて用いるための、およびスコアカードの作成、または本明細書において開示される既存のスコアカードとの比較に用いるための関心対象の幹細胞はまた、Thomson et al. (1998) Science 282: 1145によって記述されるヒト胚幹(hES)細胞; アカゲザル幹細胞(Thomson et al. (1995) Proc. Natl. Acad. Sci USA 92:7844)、マーモセット幹細胞(Thomson et al. (1996) Biol. Reprod. 55:254)、およびヒト胚性生殖(hEG)細胞(Shambloft et al., Proc. Natl. Acad. Sci. USA 95: 13726, 1998)などの他の霊長類からの胚幹細胞によって例証される様々なタイプの胚細胞を含む。同様に、中胚葉幹細胞および他の初期心臓性細胞などの系列限定幹細胞も重要である(Reyes et al. (2001) Blood 98:2615-2625; Eisenberg & Bader (1996) Circ Res. 78(2)

50

:205-16; etcを参照されたい)。いくつかの態様において、多能性幹細胞は、任意の哺乳動物種、たとえば、ヒト、ウマ、ウシ、ブタ、イヌ、ネコ、齧歯類、たとえばマウス、ラット、ハムスター、霊長類等から得られうる。いくつかの態様において、多能性幹細胞がヒト多能性幹細胞である場合、方法、アッセイ、システムにおいて用いるための、およびスコアカードの作成、または本明細書において開示される既存のスコアカードとの比較に用いるための多能性幹細胞を得る際に、胚は破壊されていない。

【0282】

ES細胞は、それらが特定の分化系列に限定されていない場合に未分化であると見なされるという事情がある。そのような細胞は、胚または成体起源の分化した細胞からそれらを区別する形態学的特徴を示す。未分化ES細胞は、当業者によって容易に認識され、顕微鏡視野の平面において、典型的に高い核/細胞質比および明確な核小体を有する細胞のコロニーで出現する。未分化ES細胞は、未分化細胞の存在を検出するためのマーカーとして用いられうる遺伝子を発現し、そのポリペプチド産物は、負の選択のためのマーカーとして用いられうる。たとえば、その各々が参照により本明細書に組み入れられる、米国特許出願第2003/0224411 AI号；Bhattacharya (2004) *Blood* 103(8):2956-64; およびThomson (1998)、前記を参照されたい。ヒトES細胞株は、ステージ特異的胚抗原 (SSEA) -3、SSEA-4、TRA-1-60、TRA-1-81、およびアルカリホスファターゼを含む、未分化非ヒト霊長類ES細胞およびヒトEC細胞を特徴決定する細胞表面マーカーを発現する。SSEA-4エピトープを有するグロボ系糖脂質GL7は、SSEA-3エピトープを有するグロボ系糖脂質Gb5にシアル酸を付加することによって形成される。このように、GL7は、SSEA-3およびSSEA-4の両方に対する抗体と反応する。未分化ヒトES細胞株は、SSEA-1に関して染色されなかったが、分化した細胞は、SSEA-1に関して強く染色された。未分化型hES細胞を増殖させるための方法は、その全内容が参照により本明細書に組み入れられる、WO 99/20741、WO 01/51616、およびWO 03/020920に記述される。

【0283】

いくつかの態様において、方法、アッセイ、システムにおいて用いるための、およびスコアカードの作成、または本明細書において開示される既存のスコアカードとの比較に用いるための多能性幹細胞は、ヒト臍帯血細胞である。ヒト臍帯血細胞 (HUCBC) は、最近、造血および間葉前駆細胞に富む起源として認められている (Broxmeyer et al., 1992 *Proc. Natl. Acad. Sci. USA* 89:4109-4113)。これまで、臍帯血および胎盤血は、乳児の出生時に通常廃棄される廃棄物と見なされていた。臍帯血細胞は、移植可能な幹細胞および前駆細胞源として、ならびに悪性疾患 (すなわち、急性リンパ球性白血病、急性骨髄性白血病、慢性骨髄性白血病、骨髄異形成症候群、および神経芽腫)、およびファンコーニ貧血および再生不良性貧血などの非悪性疾患 (Kohli-Kumar et al., 1993 *Br. J. Haematol.* 85:419-422; Wagner et al., 1992 *Blood* 79:1874-1881; Lu et al., 1996 *Crit. Rev. Oncol. Hematol.* 22:61-78; Lu et al., 1995 *Cell Transplantation* 4:493-503) を処置するための骨髄再構成細胞源として用いられる。HUCBCの明確な利点は、胎児細胞と非常に類似のこれらの細胞の未成熟免疫であり、これは宿主による拒絶のリスクを有意に低減させる (Taylor & Bryson, 1985 *J. Immunol.* 134: 1493-1497)。

【0284】

ヒト臍帯血は、組織培養において拡大させることができる間葉および造血前駆細胞、ならびに内皮細胞前駆体を含む (Broxmeyer et al., 1992 *Proc. Natl. Acad. Sci. USA* 89:4109-4113; Kohli-Kumar et al., 1993 *Br. J. Haematol.* 85:419-422; Wagner et al., 1992 *Blood* 79:1874-1881; Lu et al., 1996 *Crit. Rev. Oncol. Hematol.* 22:61-78; Lu et al., 1995 *Cell Transplantation* 4:493-503; Taylor & Bryson, 1985 *J. Immunol.* 134: 1493-1497 Broxmeyer, 1995 *Transfusion* 35:694-702; Chen et al., 2001 *Stroke* 32:2682-2688; Nieda et al., 1997 *Br. J. Haematology* 98:775-777; Erices et al., 2000 *Br. J. Haematology* 109:235-242)。臍帯血における造血前駆細胞の総含有量は、骨髄に等しいかまたは骨髄を上回り、かつさらに、増殖性の高い造血細胞は、HUCBCでは骨髄より8倍高く、CD14、CD34、およびCD45などの造血マーカーを発現する (Sanchez-Ram

10

20

30

40

50

os et al., 2001 Exp. Neur. 171 : 109-115; Bicknese et al., 2002 Cell Transplantation 11 :261-264; Lu et al., 1993 J. Exp Med. 178:2089-2096)。1つの細胞源は、循環中の末梢血などの、好ましくは哺乳動物の末梢血の単核球分画、臍帯血、骨髓、胎児肝臓、または卵黄嚢からの造血微小環境である。いくつかの態様において、多能性幹細胞、特に神経幹細胞もまた、髄膜を含む中枢神経系に由来しうる。

【0285】

コンピュータシステム

本発明の1つの局面は、アッセイデータを処理するための、および、多能性幹細胞の1つまたは複数の品質保証スコアカードなどの1つまたは複数の標的細胞の測定または評価を作成するための、コンピュータ化されたシステムに関する。コンピュータシステムは、(a) (i) DNAメチル化データ、たとえば関心対象の多能性幹細胞株におけるDNAメチル化標的遺伝子セットのメチル化レベルを受信して、DNAメチル化データと、1つの対照多能性幹細胞株または複数の参照多能性幹細胞株における同じ標的遺伝子の参照DNAメチル化レベルとの比較を行う段階；(ii) 多能性幹細胞株の分化能データを受信して、該分化能データを参照分化能データと比較する段階；(iii) 参照DNAメチル化データパラメータと比較したDNAメチル化データの比較に基づいて偏差スコアカードを作成して、参照分化データと比較した関心対象の幹細胞株の分化指向性の比較に基づいて系列スコアカードを作成する段階を含む方法を実行するためのコンピュータシステムの操作を制御するように適合させた少なくとも1つのコンピュータプログラムを含む、少なくとも1つのメモリ、ならびに (b) コンピュータプログラムを実行するための少なくとも1つのプロセッサを含むことができる。

【0286】

いくつかの態様において、コンピュータシステムは、(a) (i) DNAメチル化データ、たとえば関心対象の多能性幹細胞株におけるDNAメチル化標的遺伝子セットのメチル化レベルを受信して、1つの対照多能性幹細胞株または複数の参照多能性幹細胞株における同じDNAメチル化標的遺伝子のDNAメチル化データ(たとえば、DNAメチル化レベル)との比較を行う段階；(ii) 遺伝子発現データ、たとえば関心対象の多能性幹細胞株における系列マーカー遺伝子セットの遺伝子発現レベルを受信して、1つの対照多能性幹細胞株または複数の参照多能性幹細胞株における同じ系列マーカー遺伝子の遺伝子発現データ(たとえば、遺伝子発現レベル)の比較を行う段階；(iii) 参照DNAメチル化パラメータと比較したDNAメチル化データの比較に基づいて偏差スコアカードを作成して、その遺伝子の系列マーカーの参照遺伝子発現レベルと比較して、関心対象の多能性幹細胞株における系列マーカー遺伝子の遺伝子発現レベルの比較に基づいて系列スコアカードを作成する段階を含む方法を実行するためのコンピュータシステムの操作を制御するように適合させた少なくとも1つのコンピュータプログラムを含む、少なくとも1つのメモリ、ならびに (b) コンピュータプログラムを実行するための少なくとも1つのプロセッサを含むことができる。

【0287】

いくつかの態様において、コンピュータプログラムは、(i) 関心対象の多能性幹細胞株における第二の標的遺伝子セットの遺伝子発現データ(たとえば、遺伝子発現レベル)を受信して、遺伝子発現データ(たとえば、遺伝子発現レベル)を参照遺伝子発現データ(たとえば、対照多能性幹細胞株または複数の多能性幹細胞株における同じ第二の標的遺伝子セットの遺伝子発現レベル)と比較する段階；(ii) 参照遺伝子発現データ(たとえば、参照多能性幹細胞株における参照遺伝子発現レベル)と比較した遺伝子発現データ(たとえば、遺伝子発現レベル)の比較に基づいて偏差スコアカードを作成する段階をさらに含む方法を実行するためのコンピュータシステムの操作を制御するように適合させる。

【0288】

本発明のもう1つの局面は、(i) DNAメチル化データ、たとえば関心対象の多能性幹細胞株におけるDNAメチル化標的遺伝子セットのメチル化レベルを受信して、1つの対照多能性幹細胞株または複数の参照多能性幹細胞株における同じDNAメチル化標的遺伝子のDNAメ

チル化データ（たとえば、DNAメチル化レベル）との比較を行う段階；（ii）遺伝子発現データ、たとえば関心対象の多能性幹細胞株における系列マーカー遺伝子セットの遺伝子発現レベルを受信して、1つの対照多能性幹細胞株または複数の参照多能性幹細胞株における同じ系列マーカー遺伝子の遺伝子発現データ（たとえば、遺伝子発現レベル）の比較を行う段階；（iii）参照DNAメチル化パラメータと比較したDNAメチル化データの比較に基づいて偏差スコアカードを作成する段階、およびその遺伝子の系列マーカーの参照遺伝子発現レベルと比較した関心対象の多能性幹細胞における系列マーカー遺伝子の遺伝子発現レベルの比較に基づいて系列スコアカードを作成する段階を含む、アッセイデータを処理するための、および多能性幹細胞株の1つまたは複数の品質保証スコアカードを作成するためのコンピュータシステムの制御に関するコンピュータプログラムおよびソフトウェアなどの命令を含む、コンピュータ読み取り可能な媒体に関する。いくつかの態様において、コンピュータ読み取り可能な媒体は、（i）関心対象の多能性幹細胞株における第二の標的遺伝子セットの遺伝子発現データ（たとえば、遺伝子発現レベル）を受信して、遺伝子発現データ（たとえば、遺伝子発現レベル）と、対照多能性幹細胞株または複数の多能性幹細胞株における同じ第二の標的遺伝子セットの参照遺伝子発現データ（たとえば、参照遺伝子発現レベル）とを比較すること；（ii）参照遺伝子発現データ（たとえば、参照多能性幹細胞株における参照遺伝子発現レベル）と比較した遺伝子発現データ（たとえば、遺伝子発現レベル）の比較に基づいて偏差スコアカードを作成することに関する命令をさらに含む。

10

20

30

40

50

【0289】

コンピュータシステムは、1つまたは複数の汎用または専用プロセッサならびに揮発性および非揮発性メモリデバイスを含む関連メモリを含むことができる。コンピュータシステムメモリは、本発明に従う専用システムを作成するために、または本発明に従う方法を行うためのシステムを実行するために、コンピュータシステムの操作を制御するためのソフトウェアまたはコンピュータプログラムを保存することができる。コンピュータシステムは、IntelまたはAMD x86に基づくシングルまたはマルチコア中央演算処理装置（CPU）、ARMプロセッサ、またはデータを処理するための類似のコンピュータプロセッサを含むことができる。CPUまたはマイクロプロセッサは、Intel Pentiumプロセッサ、Intel 8051プロセッサ、RISCもしくはMIPSプロセッサ、Power PCプロセッサ、またはALPHAプロセッサなどの任意の通常の汎用のシングルまたはマルチチップマイクロプロセッサでありうる。さらに、マイクロプロセッサは、デジタルシグナルプロセッサまたはグラフィクスプロセッサなどの任意の通常のまたは専用マイクロプロセッサでありうる。マイクロプロセッサは典型的に、通常のアドレスライン、通常データライン、および1つまたは複数の通常の制御ラインを有する。以下に記述するように、本発明に従うソフトウェアは、専用システム上で、またはDOS、CPM、Windows、Unix、Linux、もしくは他のオペレーティングシステムを有する汎用目的コンピュータ上で実行することができる。システムは、コンピュータプログラム、ソフトウェア、およびデータを保存するためのディスクメモリおよびソリッドステートメモリなどの非揮発性メモリ、ならびにプログラムおよびソフトウェアを実行するための高速ramなどの揮発性メモリを含むことができる。

【0290】

本発明の様々な態様において有用なコンピュータ読み取り可能な物理的保存媒体は、任意のコンピュータ読み取り可能な物理的保存媒体、たとえば、ソリッドステートメモリ（フラッシュメモリなどの）、磁気および光学コンピュータ読み取り可能保存媒体およびデバイス、ならびに他の永続的保存技術を用いるメモリを含むことができる。いくつかの態様において、コンピュータ読み取り可能な媒体は、コンピュータプログラムおよびデータにコンピュータがアクセスすることができる任意の有形媒体でありうる。コンピュータ読み取り可能な媒体は、コンピュータ読み取り可能な命令、プログラムモジュール、プログラム、データ、データ構造、およびデータベース情報などの情報を保存することができる任意の方法または技術において実行される、揮発性および非揮発性、リムーバブルおよび非リムーバブル有形媒体を含むことができる。本発明のいくつかの態様において、コンピ

ユーザ読み取り可能な媒体には、RAM（ランダムアクセスメモリ）、ROM（読み出し専用メモリ）、EPROM（消去可能書き換え可能読み出し専用メモリ）、EEPROM（電子的消去可能書き換え可能読み出し専用メモリ）、フラッシュメモリ、または他のメモリ技術、CD-ROM（コンパクトディスク読み出し専用メモリ）、DVD（デジタルバーサタイルディスク）、または他の光学保存媒体、磁気カセット、磁気テープ、磁気ディスク保存または他の磁気保存媒体、他のタイプの揮発性および非揮発性メモリ、ならびに情報を保存するために用いることができ、前述の任意の適切な組み合わせを含むコンピュータによって読み取り可能な他の任意の有形媒体が挙げられるがこれらに限定されるわけではない。

【0291】

本発明は、スタンドアロンコンピュータ上で、またはネットワークコンピュータシステムの一部として実行することができる。スタンドアロンコンピュータでは、ソフトウェアおよびデータは全て、ローカルメモリデバイスに存在することができる、たとえば光学ディスクまたはフラッシュメモリデバイスを用いて、本発明を実行するためのコンピュータソフトウェアならびにデータを保存することができる。代替りの態様において、遠隔デバイスとのネットワーク接続を通して、ソフトウェアまたはデータまたはその両方にアクセスすることができる。1つのネットワークコンピュータシステムの態様において、本発明は、遠隔および/または中央に位置する場所で保存されるデータおよびリソースに接続するために、インターネットなどの公共のネットワークまたはプライベートネットワークに対してクライアント-サーバー環境を用いる。この態様において、ウェブサーバーを含むサーバーは、本発明に従って提供される情報に対するアクセス、オープンアクセス、現金払いまたは会費に基づくアクセスを提供することができる。クライアントサーバー環境において、ウェブブラウザなどのクライアントソフトウェアまたはプログラムを実行するクライアントコンピュータは、ネットワークを介してサーバーに接続する。クライアントソフトウェアまたはウェブブラウザは、本発明のユーザーがデータおよび情報を入力してデータおよび情報へのアクセスを受けるためのユーザーインターフェースを提供する。クライアントソフトウェアは、ローカルコンピュータディスプレイまたは他の出力デバイス上で見ることができ、ユーザーは、コンピュータキーボード、マウス、または他の入力デバイスなどを用いることによって情報を入力することができる。サーバーは、クライアントソフトウェアが、本発明に従うデータを入力、データを処理、およびユーザーにデータを出力することができ、かつローカルおよび遠隔コンピュータリソースへのアクセスを提供することができる、1つまたは複数のコンピュータプログラムを実行する。たとえば、ユーザーインターフェースは、アッセイからのデータ、たとえば参照多能性幹細胞集団および/または関心対象の多能性幹細胞集団の標的遺伝子のDNAメチル化データレベルまたはDNA遺伝子発現レベルの登録を可能にするテキストボックスなどのアクセス要素、ならびにスコアカードとの比較結果、またはコンピュータ読み取り可能な媒体においてコードされた命令の実行後にプロセッサに伝達されるまたはプロセッサによって利用可能となるデータセットのグラフィック読み出しを提供することができるディスプレイ要素を含むグラフィカルユーザーインターフェースを含むことができる。

【0292】

本発明の態様はまた、本明細書において開示される方法に従って多能性幹細胞集団の品質保証を決定する方法を行うためのシステム（およびコンピュータシステムを引き起こさせるためのコンピュータ読み取り可能な媒体）を提供する。

【0293】

本発明のいくつかの態様において、コンピュータシステムソフトウェアは1つまたは複数の機能モジュールを含むことができ、これは、コンピュータ読み取り可能な媒体に記録されたコンピュータ実行可能な命令によって定義されることができ、かつ、実行されると、本発明に従う方法をコンピュータに行わせる。モジュールは、明快にするために機能ごとに分離することができるが、モジュールは、コードの個別のブロックに対応する必要はなく、記述の機能は、様々な媒体に保存され、様々な時間で実行される様々なソフトウェアコード部分の実行によって行うことができると理解すべきである。さらに、モジュール

は、他の機能を行うことができ、このようにモジュールは、任意の特定の機能または機能セットを有することに限定されるわけではないと認識すべきである。いくつかの態様において、偏差スコアカードを作成するための機能モジュールは、たとえば保存モジュール、遺伝子マッピングモジュール、参照比較モジュール、正規化モジュール、関連性フィルターモジュール、遺伝子セットモジュール、および偏差スコアカードを表示するためのスコアカードディスプレイモジュールであるが、これらに限定されるわけではない。系列スコアカードを作成するための機能モジュールは、たとえば保存デバイス、アッセイ標準化モジュール、試料標準化モジュール、参照比較モジュール、遺伝子セットモジュール、濃縮分析モジュール、および系列スコアカードを表示するためのスコアカードディスプレイモジュールであるが、これらに限定されるわけではない。機能モジュールは、1つまたは複数のコンピュータを用いて、および1つまたは複数のコンピュータネットワークを用いて実行することができる。

【0294】

1つまたは複数のコンピュータ読み取り可能な媒体上で具現化される情報は、データ、コンピュータソフトウェアまたはプログラムを含むことができ、かつ、コンピュータによって実行された結果として、該コンピュータを特殊用途マシンに変換し、本明細書において記述される機能の1つまたは複数、該コンピュータに行わせることができる、プログラム命令を含むことができる。そのような命令は本来、複数のプログラミング言語、たとえばJava、J#、Visual Basic、C、C#、C++、Fortran、Pascal、Eiffel、Basic、およびCOBOLアセンブリ言語など、またはその任意の多様な組み合わせのいずれかで書くことができる。そのような命令が具現化されるコンピュータ読み取り可能な媒体は、本発明に従うコンピュータシステムまたはコンピュータシステムのネットワークの成分の1つまたは複数に存在することができる。

【0295】

いくつかの態様において、コンピュータ読み取り可能な媒体は、本明細書において述べる本発明の局面を実行するために、保存される命令を任意のコンピュータリソースにロードすることができるように、移動可能でありうる。さらに、コンピュータ読み取り可能な媒体に保存される命令は、ホストコンピュータ上で作動するアプリケーションプログラムの一部として具現化される命令に限定されないと認識すべきである。むしろ、命令は、本発明の局面を実行するためにコンピュータをプログラムするために用いることができる任意のタイプのコンピュータコード（たとえば、オブジェクトコード、ソフトウェア、またはマイクロコード）として具現化されうる。コンピュータ実行可能な命令は、適切なコンピュータ言語またはいくつかの言語の組み合わせで書かれてもよい。基本的な計算生物学の方法は、当業者に公知であり、たとえば、Setubal and Meidanis et al., Introduction to Computational Biology Methods (PWS Publishing Company, Boston, 1997); Salzberg, Searles, Kasif, (Ed.), Computational Methods in Molecular Biology, (Elsevier, Amsterdam, 1998); Rashidi and Buehler, Bioinformatics Basics: Application in Biological Science and Medicine (CRC Press, London, 2000) and Ouelette and Bzevanis Bioinformatics: A Practical Guide for Analysis of Gene and Proteins (Wiley & Sons, Inc., 2nd ed., 2001)に記述される。

【0296】

いくつかの態様において、本明細書において開示されるシステムは、自動遺伝子発現分析システム、たとえばMALDI-TOF、またはマトリクス支援レーザー脱離イオン化飛行時間型システムを含む質量分析システム；SELDI-TOF-MSプロテインチップアレイプロファイリングシステム、たとえばCIPHERGEN Protein Biology System II（商標）ソフトウェアを有する機器；遺伝子発現データを分析するためのシステム（たとえば、U.S. 2003/0194711を参照されたい）；アレイに基づく発現分析のためのシステム、たとえばHTアレイシステムおよびAffymetrix（Santa Clara, CA 95051）から入手可能なカートリッジアレイシステム、オートローダー、Complete GeneChip（登録商標）機器システム、フルイディクスステーション450、ハイブリダイゼーションオープン645、QCツールボックスソフトウェ

10

20

30

40

50

アキット、スキャナ3000 7G、スキャナ3000 7Gプラス標的化遺伝子タイピングシステム、スキャナ3000 7G全ゲノムアソシエーションシステム、GeneTitan (商標) 機器、GeneChip (登録商標) アレイステーション、HTアレイ; 自動ELISAシステム (たとえば、Dynax, Chantilly, VA のDSX (登録商標) またはDS2 (登録商標) またはENEASYSTEM III (登録商標)、Triturus (登録商標)、The Mago (登録商標) プラス); 密度計 (たとえば、X-Rite-508-Spectro Densitometer (登録商標)、The HYRYS (商標) 2密度計); 自動蛍光インサイチューハイブリダイゼーションシステム (たとえば、米国特許第6,136,540号を参照されたい); 2-Dイメージングソフトウェアと連結させた2Dゲルイメージングシステム; マイクロプレートリーダー; 蛍光活性化セルソーター (FACS) (たとえば、フローサイトメーター、FACS Vantage SE, Becton Dickinson); 放射性同位元素分析装置 (たとえば、シンチレーションカウンター) が挙げられるがこれらに限定されるわけではない自動タンパク質発現分析から遺伝子発現レベルデータを受信することができる。

10

【0297】

本発明のいくつかの態様において、参照データを、電子的またはデジタル式に記録して、注釈をつけて、genome、ESTs、SNPS、Traces、Celera、Ventor Reads、Watson reads、HGTS等などのGenBank (NCBI) タンパク質およびDNAデータベース; ENZYME、PROSITE、SWISS-2DPAGE、Swiss-Prot、およびTrEMBLデータベースなどのSwiss Institute of Bioinformaticsデータベース; MelanieソフトウェアパッケージまたはExPASy WWWサーバー等、SWISS-MODEL、Swiss-Shop、および他のネットワークに基づく計算ツール; 包括的微生物リソースデータベース (The institute of Genomic Research) が挙げられるがこれらに限定されるわけではないデータベースから検索することができる。得られた情報を関連するデータベースに保存することができ、これを用いて、ゲノム内およびゲノムにおける参照データまたは遺伝子またはタンパク質のあいだの相同性を決定してもよい。

20

【0298】

いくつかの態様において、多能性幹細胞における標的遺伝子の遺伝子発現レベルはメモリ、保存デバイス、またはデータベースから受信することができる。メモリ、保存デバイス、またはデータベースを、データを検索するコンピュータシステムに直接接続することができ、または有線もしくは無線接続技術を通してコンピュータに接続して、有線または無線接続を通して遠隔デバイスまたはシステムから検索することができる。さらに、メモリ、保存デバイス、またはデータベースは、検索されるコンピュータシステムから離れた場所に存在することができる。

30

【0299】

本発明と共に用いるのに適した接続技術の例には、たとえばパラレルインターフェース (たとえば、PATA)、シリアルインターフェース (たとえば、SATA、USB、ファイアワイア)、ローカルエリアネットワーク (LAN)、広域ネットワーク (WAN)、インターネット、イントラネット、およびエクストラネット、および無線 (たとえば、Blue Tooth、Zigbee、WiFi、WiMAX、3G、4G) 通信技術が挙げられる。

【0300】

保存デバイスはまた、当技術分野において一般的に、様々な態様において有用である「コンピュータ読み取り可能な物理的保存媒体」と呼ばれ、これは任意のコンピュータ読み取り可能な物理的保存媒体、中でもたとえば磁気および光学コンピュータ読み取り可能な保存媒体を含むことができる。搬送波および他のシグナルに基づく保存または伝達媒体は、この用語に包含されて本発明にとって有用である保存デバイスまたはコンピュータ読み取り可能な物理的保存媒体の範囲内に含まれない。保存デバイスは、サイトカインレベル情報を記録するように適応または構成される。そのような情報は、たとえばインターネットを介して、ディスク上で、USB (汎用シリアルバス) を介して、または他の任意の適切な通信様式を介して電子的に伝達および読み取りが可能なデジタル形式で、提供することができる。

40

【0301】

本明細書において用いられる「保存される」とは、情報、たとえばデータ、プログラム

50

および命令を、後に読み返すことができる保存デバイス上に記録するプロセスを意味する。当業者は、参照スコアカードデータ情報を与える、たとえば本明細書の方法において開示される多能性幹細胞のDNAメチル化レベルおよび/または遺伝子発現レベルおよび/または分化指向性データを公知の媒体に記録するための、現在公知の任意の方法を容易に採用することができる。

【0302】

スコアカードデータおよび情報を保存デバイスに保存するために、多様なソフトウェアプログラムおよびフォーマットを用いることができる。任意の数のデータプロセッサ構造化フォーマット（たとえば、テキストファイルまたはデータベース）を使用して、スコアカードが記録された媒体を得るまたは作成することができる。

10

【0303】

1つの態様において、参照スコアカードデータを、電子的またはデジタル的に記録して、Yaleタンパク質発現データベース（YPED）ならびにgenome、ESTs、SNPS、Traces、Celera、Ventor Reads、Watson reads、およびHGTSなどのGenBank（NCBI）タンパク質およびDNAデータベース；ENZYME、PROSITE、SWISS-2DPAGE、Swiss-Prot、およびTrEMBLデータベースなどのSwiss Institute of Bioinformaticsデータベース；MelanieソフトウェアパッケージまたはExpASY WWWサーバーなど；SWISS-MODEL、およびSwiss-Shopなどのネットワークに基づく計算ツール；包括的微生物リソースデータベース（The Institute of Genomic Researchから入手可能）などの、当技術分野において一般的に公知のタンパク質発現データベースが挙げられるがこれらに限定されるわけではないデータベースから注釈をつけることができる。多能性幹細胞株のDNAメチル化レベルおよび/または遺伝子発現レベルおよび/または分化指向性データに関して得られた情報をリレーショナルデータベースに保存することができ、該データベースは、異なる多能性幹細胞集団と比較した、または、異なる多能性幹細胞集団の間で、たとえばES細胞、iPS細胞およびpiPS細胞、ならびに体性幹細胞のあいだで、または、異なるゲノム、種、および異なる個体集団由来の同じタイプの多能性幹細胞（たとえば、iPS細胞）の中で、参照DNAメチル化レベル、参照遺伝子発現レベル、および参照指向性分化データと比較した差を決定するために用いられる。

20

【0304】

いくつかの態様において、システムは、1つまたは複数のプログラムを作動させるためのプロセッサを有し、たとえばプログラムはオペレーティングシステム（たとえば、UNIX、Windows）、リレーショナルデータベース管理システム、アプリケーションプログラム、およびワールドワイドウェブサーバープログラムを含むことができる。アプリケーションプログラムは、データベース言語ステートメント（たとえば、構造化問い合わせ言語（SQL）ステートメント）の作成にとって必要な実行可能なコードを含むワールドワイドウェブアプリケーションでありうる。実行可能なファイルは、埋め込みSQLステートメントを含むことができる。さらに、ワールドワイドウェブアプリケーションは、ワールドワイドウェブサーバー機能と、ユーザーの要求に応えるようにアクセス可能である様々な外部および内部データベースとを提供する様々なソフトウェアエンティティに対するポインターおよびアドレスを含むコンフィギュレーションファイルを含むことができる。コンフィギュレーションファイルはまた、2つまたはそれより多くの別個のコンピュータにサーバーが分散している場合に必要であるような、適切なハードウェアデバイスに対するサーバーリソースの要求を指示することができる。1つの態様において、ワールドワイドウェブサーバーは、TCP/IPプロトコルをサポートする。このようなローカルネットワークは、時に「イントラネット」と呼ばれる。そのようなイントラネットの利点は、それらが、ワールドワイドウェブに存在する公共ドメインデータベース（たとえば、GenBankまたはSwiss Proワールドワイドウェブサイト）と容易に通信できる点である。このように、本発明の特に好ましい態様において、ユーザーは、ウェブブラウザおよびウェブサーバーによって提供されるHTMLインターフェースを用いてインターネットデータベース上に存在するデータに（たとえばハイパーテキストリンクを介して）直接アクセスすることができる。

30

40

【0305】

50

1つの態様において、本明細書において開示されるシステムを用いて、DNAメチル化データ（たとえば、複数のDNAメチル化標的遺伝子のDNAメチル化プロファイルまたはDNAメチル化レベル）および/または遺伝子発現プロファイル（たとえば、複数の遺伝子発現標的遺伝子の遺伝子発現プロファイルまたは遺伝子発現レベル）を比較することができる。たとえば、システムは、そのメモリ上に試験多能性幹細胞株の遺伝子発現プロファイルまたはデータを受信して、これを1つまたは複数の保存された遺伝子発現プロファイル（たとえば、1つまたは複数の参照多能性幹細胞株における遺伝子発現の正常な変動）と比較する、または初期の時点で既に分析された多能性幹細胞株由来の1つまたは複数の遺伝子発現プロファイルと比較することができる。いくつかの態様において、プローブセットからのシグナル強度に基づいて1つの遺伝子または複数の遺伝子の相対量を分析するために、Affymetrix Microarray Suiteソフトウェアバージョン5.0 (MAS 5.0) (Affymetrix, Santa Clara, Californiaから入手可能)を用いて遺伝子発現プロファイルを得て、MAS 5.0データファイルをデータベースに転送し、マイクロソフトExcelおよびGeneSpring 6.0ソフトウェア (Agilent Technologies, Santa Clara, Californiaから入手可能)によって分析することができる。いくつかの態様において、MAS 5.0ソフトウェアの比較アルゴリズムを用いて、どれほど多くの転写物が所定の試料中に検出され、2つまたはそれより多くのマイクロアレイデータセットの比較分析を可能にするかに関する包括的概要を得ることができる。

【0306】

本発明のこの局面および他の全ての局面のいくつかの態様において、システムは、決定モジュールにおいて決定された配列情報を参照データと比較するように作用する比較のために、利用できる多様なソフトウェアプログラムおよびフォーマットを用いることができる「比較モジュール」においてデータを比較することができる。1つの態様において、比較モジュールは、1つまたは複数の登録からの配列情報を1つまたは複数の参照データパターンと比較するためにパターン認識技術を用いるように構成される。パターンを比較するために、既存の市販のまたは無料のソフトウェアを用いて比較モジュールを構成してもよく、行われる特定のデータ比較のために最適化してもよい。比較モジュールはまた、たとえば、DNA配列におけるCpGメチル化部位の有無の検出、メチル化レベルの決定、試料中の配列の濃度の決定（たとえば、アミノ酸配列/タンパク質発現レベル、またはヌクレオチド (RNAもしくはDNA) 発現レベル)、または遺伝子発現プロファイルの決定を含むことができる配列情報に関連するコンピュータ読み取り可能な情報を提供することができる。

【0307】

本発明のいくつかの態様において、システムは、関心対象の多能性幹細胞に関するDNAメチル化データ、または関心対象の多能性幹細胞に関する遺伝子発現レベルデータが、本明細書において開示される参照DNAメチル化レベル（たとえば、DNAメチル化の正常な変動）または参照遺伝子発現レベルの範囲から外れるか否か、たとえば複数の多能性幹細胞に関する標的遺伝子の遺伝子発現レベルの正常な変動の範囲から外れるか否かを判定するために用いられる比較ソフトウェアを含む。1つの態様において、関心対象の多能性幹細胞のDNAメチル化レベルが、参照DNAメチル化レベルより統計学的に有意な量で高ければ、そのDNAメチル化標的遺伝子のエピジェネティックなサイレンシングおよび抑制の可能性を示している。DNAメチル化標的遺伝子が腫瘍抑制遺伝子である例では、これは、多能性幹細胞が癌細胞になる素因を有することを示す。DNAメチル化標的遺伝子が発達遺伝子および/または系列マーカー遺伝子である例では、多能性幹細胞株が、低い分化効率を有するか、または特定の発達経路に沿って分化しないか、または系列マーカー遺伝子を発現する細胞へと分化しないことを示すまたはシグナルを発するように、ソフトウェアを構成することができる。

【0308】

同様に、関心対象の多能性幹細胞の遺伝子発現レベルがその遺伝子に関する参照遺伝子発現レベルより統計学的に有意な量で高い場合、標的遺伝子が発現される可能性を示しており、DNA標的遺伝子が発達または系列特異的マーカーである場合、その細胞系列に沿っ

て最適に分化する可能性のシグナルを送る（またはそうでなければ示す）ようにソフトウェアを構成することができる。DNAメチル化標的遺伝子が腫瘍遺伝子である例では、関心対象の多能性幹細胞株が癌細胞になる素因を有する、または制御されない増殖を有する可能性があるシグナルを送るように、ソフトウェアを構成することができる。

【0309】

DNAメチル化データおよび/または遺伝子発現レベルデータをコンピュータ読み取り可能な形状で提供することによって、多能性幹細胞に関するDNAメチル化データおよび/または遺伝子発現レベルデータを用いて、保存デバイス内の他の多能性幹細胞の参照DNAメチル化レベルおよび参照遺伝子発現レベルと比較することができる。たとえば、検索プログラムを用いて、関心対象の多能性幹細胞の同じ標的遺伝子のDNAメチル化レベルとマッチする関連する参照データ（すなわち、標的遺伝子の参照DNAメチル化レベル）を同定することができる。コンピュータ読み取り可能な形式で行われる比較は、多様な手段によって処理することができるコンピュータ読み取り可能な内容を提供する。内容、検索内容を比較モジュールから検索することができる。

10

【0310】

いくつかの態様において、比較モジュールは、既定の基準またはユーザーによって定義される基準によってコンピュータ読み取り可能な形式で処理することができるコンピュータ読み取り可能な比較結果を提供し、ディスプレイモジュールを用いてユーザーの要求により保存および出力される該比較結果に部分的に基づく内容を含むレポートを提供する。いくつかの態様において、ディスプレイモジュールは、比較結果に部分的に基づく内容をユーザーに表示することを可能にし、該内容は、関心対象の多能性幹細胞とスコアカードとの比較結果を示す、または、多能性幹細胞の有用性、たとえば特定の癌（たとえば、腫瘍遺伝子および腫瘍抑制遺伝子）のメチル化状態ならびに特定の発達遺伝子および/もしくは系列マーカー遺伝子のメチル化状態を示す、レポートである。

20

【0311】

いくつかの態様において、ディスプレイモジュールは、比較結果に部分的に基づくレポートまたは内容を末端ユーザーに表示することを可能にし、該内容は、関心対象の多能性幹細胞とスコアカードとの比較結果を示す、または、多能性幹細胞の有用性、たとえば特定の癌（たとえば、腫瘍遺伝子および腫瘍抑制遺伝子）のメチル化状態ならびに特定の発達遺伝子および/もしくは系列マーカー遺伝子のメチル化状態を示す、レポートである。

30

【0312】

本発明のこの局面および他の全ての局面のいくつかの態様において、比較モジュールまたは本発明の他の任意のモジュールは、その上で関連するデータベース管理システムを動作させるオペレーティングシステム（たとえば、UNIX、Windows）、ワールドワイドウェブアプリケーション、およびワールドワイドウェブサーバーを含むことができる。ワールドワイドウェブアプリケーションは、データベース言語ステートメント[たとえば標準化問い合わせ言語（SQL）ステートメント]の作成にとって必要な実行可能なコードを含むことができる。実行可能なファイルは、埋め込みSQLステートメントを含むことができる。さらに、ワールドワイドウェブアプリケーションは、サーバーと、ユーザーの要求に応えるためにアクセスする必要がある様々な外部および内部データベースとを含む様々なソフトウェアエンティティに対するポインターおよびアドレスを含むコンフィギュレーションファイルを含むことができる。コンフィギュレーションファイルはまた、2つまたはそれより多くの別個のコンピュータにサーバーが分散している場合に必要でありうるような、適切なハードウェアデバイスに対するサーバーリソースの要求を指示する。1つの態様において、ワールドワイドウェブサーバーは、TCP/IPプロトコルをサポートする。このようなローカルネットワークは時に「イントラネット」と呼ばれる。そのようなイントラネットの利点は、それらが、ワールドワイドウェブに存在する公共ドメインデータベース（たとえば、GenBankまたはSwiss Proワールドワイドウェブサイト）と容易に通信できる点である。このように、本発明の特に好ましい態様において、ユーザーは、ウェブブラウザおよびウェブサーバーによって提供されるHTMLインターフェースを用いてインターネットデ

40

50

ータベース上に存在するデータに（たとえばハイパーテキストリンクを介して）直接アクセスすることができる。本発明の他の態様において、HTTP、FTP、SSH、およびVPNに基づくインターフェースなどの他のインターフェースを用いてインターネットデータベースに接続することができる。

【0313】

本発明のこの局面および他の全ての局面のいくつかの態様において、コンピュータ読み取り可能な媒体は、コンピュータプログラムおよびソフトウェアなどのそれ自体に保存された命令を、本明細書において述べる本発明の局面を実行するための任意のコンピュータリソースにロードすることができるように、移動可能でありうる。さらに、上記のようにコンピュータ読み取り可能な媒体に保存される命令は、ホストコンピュータ上で作動するアプリケーションプログラムの一部として具現化される命令に限定されないと認識されるべきである。むしろ、命令は、本発明の局面を実行するためにプロセッサをプログラムするために使用することができる任意のタイプのコンピュータコード（たとえば、ソフトウェアまたはマイクロコード）として具現化されうる。コンピュータ実行可能な命令は、適切なコンピュータ言語またはいくつかの言語の組み合わせで書くことができる。基本的な計算生物学の方法は、たとえば、Setubal and Meidanis et al., Introduction to Computational Biology Methods (PWS Publishing Company, Boston, 1997); Salzberg, Searles, Kasif, (Ed.), Computational Methods in Molecular Biology, (Elsevier, Amsterdam, 1998); Rashidi and Buehler, Bioinformatics Basics: Application in Biological Science and Medicine (CRC Press, London, 2000)およびOuelette and Bzevanis Bioinformatics: A Practical Guide for Analysis of Gene and Proteins (Wiley & Sons, Inc., 2nd ed., 2001)に記述される。

10

20

【0314】

コンピュータ命令は、ソフトウェア、ファームウェア、またはハードウェアにおいて実行することができる、情報処理システムのモジュールが行ういかなるタイプのプログラムされた工程も含むことができる。コンピュータシステムは、ローカルエリアネットワーク（LAN）または広域ネットワーク（WAN）に接続することができる。ローカルエリアネットワークの一例は、データ処理システムを含むコンピュータおよびコンピューティングデバイスが接続される、インターネットへのアクセスを含むコーポレートコンピューティングネットワークでありうる。1つの態様において、LANは、通信のために産業用標準伝送制御プロトコル/インターネットプロトコル（TCP/IP）ネットワークプロトコルを用いる。伝送制御プロトコル（TCP）は、コンピュータシステムにおいて高信頼性の接続指向トランスポート層リンクを提供するためのトランスポート層プロトコルとして用いることができる。ネットワーク層は、トランスポート層に対してサービスを提供する。双方向ハンドシェイクスキームを用いて、TCPは、コンピュータシステム間のロジック接続を確立、維持、および終了させるための機構を提供する。TCPトランスポート層は、そのネットワーク層プロトコルとしてIPを用いる。さらに、TCPは、各メッセージと共に目的地およびソースポート数を含めることによって1つのデバイスについて実行される多数のプログラムを区別するためにプロトコルポートを提供する。TCPは、バイトストリーム転送、データフロー定義、データ承認、失われたまたは壊れたデータの再転送、および1つのネットワーク接続を通じた多数の接続の多重化などの機能を行う。最後に、TCPは、情報のデータグラム構造への隠蔽の原因である。代替の態様において、LANは、International Standards Organizationのオープンシステム相互接続、IBMのSNA、NovellのNetware、およびBanyan VINESが挙げられるがこれらに限定されるわけではない他のネットワーク標準に従うことができる。

30

40

【0315】

いくつかの態様において、本明細書において記述されるコンピュータシステムは、例として以下のネットワーク：インターネット、イントラネット、ローカルエリアネットワーク（LAN）、または広域ネットワーク（WAN）を含む電子的に接続された任意のタイプのコンピュータ群を含むことができる。さらに、ネットワークとの接続性は、たとえばリモ-

50

トモデム、イーサネット（IEEE802.3）、トークンリング（IEEE 802.5）、光ファイバー分散データリンクインターフェース（FDDI）、または非同期転送モード（ATM）でありうる。コンピューティングデバイスは、デスクトップデバイス、サーバー、ポータブルコンピュータ、携帯型コンピューティングデバイス、スマートフォン、セットトップデバイス、または他の任意の所望のタイプもしくは構造でありうる。本明細書において用いられるように、ネットワークは、公共のインターネット、プライベートインターネット、セキュアインターネット、プライベートネットワーク、公共ネットワーク、付加価値ネットワーク、イントラネット、エクストラネット、および前述の組み合わせを含む、1つまたは複数を含む。

【0316】

本発明の1つの態様において、コンピュータシステムはパターン比較ソフトウェアを含むことができ、これは、関心対象の多能性幹細胞株におけるDNAメチル化レベルまたは遺伝子発現レベルのパターンが、該細胞株が外れ値であることを示すか否か、および、参照多能性幹細胞株の正常な特徴外で機能する幹細胞株を予測するか否か、または多能性幹細胞株が関心対象の特定の細胞株に沿って分化する効率が低いまたは癌様の特性（たとえば制御されない増殖に対する素因）を保持する可能性を予測するか否かを判定するために用いることができる。この態様において、パターン比較ソフトウェアは、関心対象の多能性幹細胞のデータの少なくともいくつか（たとえば、DNAメチル化レベルおよび/または遺伝子発現レベル）を、参照多能性幹細胞株のDNAメチル化レベルおよび（DNAメチル化標的遺伝子、および/または遺伝子発現標的遺伝子、および/または系列マーカー標的遺伝子の）遺伝子発現レベルの既定のパターンと比較して、それらがどれほど厳密にマッチするかを決定することができる。マッチングは、パターンの全てまたは一部がマッチする程度を示す部分または程度で評価および報告することができる。

【0317】

本発明のこの局面および他の全ての局面のいくつかの態様において、比較モジュールは、既定の基準またはユーザーによって定義される基準によってコンピュータ読み取り可能な形式で処理することができるコンピュータ読み取り可能なデータを提供し、ディスプレイモジュールを用いてユーザーの要求により保存および出力されうる検索内容を提供する。

【0318】

ディスプレイモジュール

本発明のいくつかの態様において、コンピュータ化システムは、コンピュータモニター、タッチスクリーンまたはビデオディスプレイシステムなどのディスプレイモジュールを含むことができ、またはディスプレイモジュールに機能的に接続することができる。ディスプレイモジュールにより、ユーザー命令がシステムのユーザーに提示され、システムへの入力が可視化され、かつ、システムが、ユーザーインターフェースの一部として結果をユーザーに表示することができる。任意で、コンピュータ化システムは、システムによる情報出力の印刷されたコピーを作成するための印刷デバイスを含むことができ、または印刷デバイスに機能的に接続することができる。

【0319】

いくつかの態様において、関心対象の多能性幹細胞の品質および/または有用性、たとえば癌細胞へと発達する可能性のリスクが低いことに基づく特定の治療的使用に関する有用性、ならびに/または、発達遺伝子および系列特異的マーカーのDNAメチル化および/もしくは遺伝子発現からのデータ、ならびに分化指向性データに基づいて、ある細胞株系列に沿って分化する可能性に基づく特定の目的に関する有用性を示すために、結果をディスプレイモジュールに表示することができる、または、レポート、たとえばスコアカードレポートに印刷することができる。

【0320】

いくつかの態様において、スコアカードレポートは、プリンタから印刷されたハードコピーである。代替の態様において、コンピュータ化システムは、スコアカードを報告す

10

20

30

40

50

るために、たとえば関心対象の多能性幹細胞株の品質および有用性を示すために、光または音を用いることができる。たとえば、本発明の全ての局面において、本明細書において開示される方法、アッセイ、システムによって生じた、およびキットに存在するスコアカードは、1つまたは複数の参照多能性幹細胞株（たとえば、本明細書において試験される標準的なヒトES細胞株およびiPS細胞株）と比較して、または研究者が選択する別の「最も」標準的な多能性幹細胞株と比較して、関心対象の多能性幹細胞の品質に関するシグナルを送るまたは品質を示すように色分けされたレポートを含むことができる。

【0321】

たとえば、赤色または他の既定のシグナルは、多能性幹細胞株が外れ値多能性幹細胞株であり、かつ、DNAメチル化レベルおよび/または遺伝子発現レベルが1つまたは複数の参照多能性幹細胞株におけるレベルと比較して統計学的に有意な量変化する1つまたは複数の遺伝子を有することを示すことができ、このように、多能性幹細胞株が参照多能性幹細胞株とは異なる特徴を有する、たとえば癌細胞株へと分化する素因を有しうる、および/または特定の細胞系列への低い分化効率を有しうるというシグナルが送られる。別の態様において、黄色またはオレンジ色または他の既定のシグナルは、多能性幹細胞株が1つまたは複数の参照多能性幹細胞株のレベルと比較してDNAメチル化レベルおよび/または遺伝子発現レベルが統計学的に有意な量変化する1つの遺伝子を有しうることを示すことができ、したがって、多能性幹細胞株が参照多能性幹細胞株とはわずかに異なる特徴を有するが、差は機能にとっては重要ではない可能性があり、たとえば関心対象の多能性幹細胞株は用いられる特徴的品質の細胞株であるものの、癌細胞株へと分化する素因等を有しないというシグナルを送ることができる。別の態様において、緑色または他の既定のシグナルは、多能性幹細胞株が高い品質の細胞株であり、遺伝子の大多数のDNAメチル化レベルおよび/または遺伝子発現レベルが、1つまたは複数の参照多能性幹細胞株におけるレベルと比較して統計学的に有意な量変化しないことを示すことができ、このように多能性幹細胞株が高い品質の細胞株であり、参照多能性幹細胞株と類似の特徴を有する可能性があるというシグナルを送ることができる。いくつかの態様において、「ヒートマップ」または勾配のある配色設計をレポート、たとえば多能性幹細胞株の品質に関するシグナルを送るためのスコアカードレポートに用いることができ、ここで勾配は赤色から黄色を経て緑色への勾配であり、赤色シグナルは、1つまたは複数の参照多能性幹細胞株と比較して、劣ったおよび/または不良な品質のシグナルであり、黄色のシグナルは良好な品質を示し、および緑色のシグナルは、関心対象の高い品質の多能性幹細胞株を示すであろう。赤色と黄色および黄色と緑色のあいだの色は、赤色-黄色-緑色スケールに関する多能性幹細胞株の特徴に関するシグナルを送るであろう。レポートにおける他の配色設計および勾配設計も同様に包含される。

【0322】

いくつかの態様において、レポート、たとえばスコアカードは、DNAメチル化の正常な変動と比較してDNAメチル化レベルに差がある遺伝子の総%および/または絶対総数を表示することができる。同様に、レポート、たとえばスコアカードは、遺伝子発現の正常な変動と比較して差次的遺伝子発現レベルを有する遺伝子の総%および/または絶対総数を表示することができる。1つの事例であるに過ぎないが、スコアカードは、試験多能性幹細胞が、評価した差次的メチル化された21%の遺伝子および/または1057個の遺伝子を有することを示すことができ、同様に差次的メチル化された遺伝子に関する正常な変動（たとえば、複数の参照多能性幹細胞株において）が、14.6~15.7%および/または731~785個の遺伝子であることを示すことができる。この例は、たとえば表12Aに示したように約5000個の遺伝子のDNAメチル化分析に基づくことに注意されたい。

【0323】

いくつかの態様において、レポート、たとえばスコアカードは、参照多能性幹細胞株（たとえば、研究者が選択する選択された「最も」標準的な株）または参照多能性幹細胞株における正常な変動に対して標準化された、試験多能性幹細胞株の標準化された値を示すことができる。したがって、スコアカードは、DNAメチル化の正常な変動と比較して変化

10

20

30

40

50

したDNAメチル化レベルを有する遺伝子の%差および/または絶対数の変化を表示することができる。同様に、レポート、たとえばスコアカードは、遺伝子発現レベルの正常な変動と比較して差次的に発現される遺伝子の%差および/または絶対数の変化を表示することができる。1つの実例であるに過ぎないが、スコアカードは、試験多能性幹細胞が、差次的メチル化遺伝子（たとえば、複数の参照多能性幹細胞株）の正常な変動と比較して、差次的にメチル化される遺伝子の34%増加および/または272個の遺伝子の増加を有することを示すことができる。

【0324】

いくつかの態様において、レポート、たとえばスコアカードは、DNAメチル化遺伝子の結果および遺伝子発現の結果を癌遺伝子および/または発達遺伝子に細分することができる。たとえば、スコアカードは、DNAメチル化レベルの正常な変動と比較して異なるDNAメチル化レベルを有する癌遺伝子および/または系列マーカー遺伝子の%（総%または%変化）および/または絶対数（総数または数の変化）を表示することができ、ならびに遺伝子発現の正常な変動レベルと比較して差次的に発現される癌遺伝子および/または系列マーカー遺伝子の%（総%または%変化）および/または絶対数（総数または数の変化）を表示することができる。

10

【0325】

いくつかの態様において、レポートを色分けすることができ、たとえば差次的DNAメチル化遺伝子または差次的発現遺伝子の%または絶対数が、ある既定の閾値レベルより上であれば、その%値または絶対数の値の色を、多能性幹細胞株がいくつかの望ましくない特徴を有しうること、および問題のある品質を有する（たとえば、癌を形成する素因を有する可能性）および/または限られた有用性を有する細胞株である可能性があることをこの値が示すことを容易に同定できるように、明るい色（たとえば、赤色）もしくはそれ以外の印を付すか（たとえば、*によって）、または強調することができる。

20

【0326】

いくつかの態様において、スコアカードはまた、試験される多能性幹細胞株の値と比較するために用いることができる参照多能性幹細胞株における差次的メチル化遺伝子の正常な数の参照値（%または絶対数のいずれかで）を表示することができる。同様に、いくつかの態様において、スコアカードはまた、試験される多能性幹細胞株の値と比較するために用いることができる参照多能性幹細胞株において差次的に発現される遺伝子の正常な数の参照値（%または絶対数のいずれかで）を表示することができる。

30

【0327】

代わりの態様において、レポート、たとえばスコアカードは、特定系列、たとえばニューロン、内胚葉、外胚葉、中胚葉、膵臓、心臓系列等に沿って分化する%または相対的分化指向性を表示することができる。

【0328】

いくつかの態様において、レポート、たとえばスコアカードは、多能性細胞株が適切である応用および/または有用性に関する推奨を与える、および/または多能性細胞株が適切でない応用および/または有用性に関する推奨を与える、口頭または書面での文章を提示することができる。

40

【0329】

本発明のこの局面および他の全ての局面に関するいくつかの態様において、レポートデータ、たとえば比較モジュールからのスコアカードを、印刷されたレポート、たとえばスコアカードの1ページまたは複数のページとして、コンピュータモニター上に表示することができる。本発明の1つの態様において、検索された内容のページを、印刷可能な媒体を通して表示することができる。ディスプレイモジュールは、コンピュータ読み取り可能な情報をユーザーに表示するように適合させた任意のデバイスまたはシステムでありうる。ディスプレイモジュールは、スピーカー、ブラウン管（CRT）、プラズマディスプレイ、発光ダイオード（LED）ディスプレイ、液晶ディスプレイ（LCD）、プリンタ、真空蛍光ディスプレイ（VFD）、表面伝導型電界放出ディスプレイ（SED）、電界放出ディスプレイ

50

(FED)等を含むことができる。

【0330】

本発明のいくつかの態様において、情報を入力するため、要求を行うため、および検索した内容を表示するためにユーザーがシステムと対話できるようなユーザーインターフェースを提供するために、ワールドワイドウェブブラウザを用いることができる。さらに、システムの様々な機能モジュールを、ユーザーインターフェースを提供するためにウェブブラウザが用いられるように適合させることができる。ウェブブラウザを用いて、ユーザーは、データベースなどのデータ源からデータを検索するための要求を行うことができ、比較およびパターンマッチングを行うために比較モジュールと対話することができる。ユーザーは、システムと対話するためにグラフィカルユーザーインターフェースにおいて従来使用されているボタン、プルダウンメニュー、スクロールバーなどのユーザーインターフェース要素にポインタを合わせてクリックし、該システムに本発明の方法を行わせることができる。ユーザーのウェブブラウザによって公式化された要求は、該要求を処理またはフォーマットできるウェブアプリケーションにネットワークを通して伝達され、ウェブアプリケーションは、DNAメチル化レベルおよび遺伝子発現レベル、検索された内容に関連する適切な情報を提供するために用いることができる1つまたは複数のデータベースの問い合わせを作成し、この情報を処理して、結果、たとえば以下のいずれかの少なくとも1つを出力することができる：(i)参照DNAメチル化レベル(たとえば、参照多能性幹細胞株の)と比較してDNAメチル化レベルの変動を有するDNAメチル化標的遺伝子の有無(%および/または絶対数)の指標の表示；(ii)参照遺伝子発現レベル(たとえば、参照多能性幹細胞株の)と比較して遺伝子発現レベルの変動を有する遺伝子発現標的遺伝子の有無(%および/または絶対数)の表示；(iii)参照系列マーカー遺伝子発現レベル(たとえば、参照多能性幹細胞株の)と比較して遺伝子発現レベルの変動を有する系列マーカー標的遺伝子の有無(%および/または絶対数)の表示。1つの態様において、1つまたは複数の参照多能性幹細胞株のDNAメチル化レベルまたは遺伝子発現レベルまたは系列マーカー遺伝子の遺伝子発現レベルも同様に、表示することができる。

10

20

【0331】

本明細書において記述されるアッセイ、方法、システム、およびキットは、DNAメチル化に関連するが、他のエピジェネティックなマーカーも同様に本発明のアッセイ、方法、システムおよびキットにおいて用いることができると理解される。たとえば、DNAメチル化および/または遺伝子発現レベルの代わりにまたはそれに加えて、ヒストン修飾または翻訳後修飾のパターンおよびレベルを用いることができる。あるポリペプチドにおける翻訳後変化のパターンは、アルツハイマー病および癌などのある疾患と相関することが知られている。たとえば、国際特許出願WO/2010/044892の表3を参照されたい。本明細書において用いられる「翻訳後修飾」または「PTM」という用語は、化学部分がタンパク質に共有的に付加される反応を意味する。多くのタンパク質は、ポリペプチド鎖の初回合成(すなわち、翻訳)後、化学部分の共有的付加(本明細書において同様に「修飾」部分とも呼ばれる)を通して翻訳後修飾されることができる。そのような化学部分は、通常、酵素によってポリペプチド鎖のアミノ酸側鎖に、またはカルボキシルもしくはアミノ末端に付加され、別の酵素によって切断されうる。同じまたは異なる化学部分のいずれかである1つまたは複数の化学部分を、1つのタンパク質分子に付加することができる。タンパク質のPTMは、その酵素活性、他のタンパク質とのその結合または他のタンパク質の活性化、またはその代謝回転などのその生物機能を変化させることができ、細胞シグナル伝達事象、生物および疾患の発生において重要である。PTMの例には、ユビキチン化、リン酸化、グリコシル化、スモイル化、アセチル化、S-ニトロシル化またはニトロシル化、シトルリン化、または脱イミノ化、neddylation、OClcNAc、ADP-リボシル化、メチル化、ヒドロキシル化、ファテニル化、ウフミル化、プレニル化、ミリストイル化、S-パルミトイル化、チロシン硫酸化、ホルミル化、およびカルボキシル化が挙げられるがこれらに限定されるわけではない。翻訳後修飾を決定およびマッピングするためのアッセイは、当業者に公知である。たとえば、その全内容が参照により本明細書に組み入れられる、米国特許第6,465,199

30

40

50

号および第6,495,664号；ならびに米国特許出願公開第2006/0078998号、第2006/0210978号、および第2008/007025号を参照されたい。

【0332】

キット

本発明のもう1つの局面は、以下を含む、多能性幹細胞株の品質を決定するためのキットに関する：(i)複数のDNAメチル化遺伝子のメチル化状態を測定するための試薬、(ii)複数の遺伝子発現遺伝子の遺伝子発現レベルを測定するための試薬；ならびに(iii)多能性幹細胞の外胚葉、中胚葉、および内胚葉系列への分化指向性を測定するための試薬。いくつかの態様において、キットは、本明細書において開示されるスコアカードをさらに含む。いくつかの態様において、キットは使用説明書をさらに含む。

10

【0333】

1つの局面において、本発明は、スコアカードを含むキットを提供する。いくつかの態様において、キットは、体細胞または分化した細胞を人工多能性幹細胞(iPSC)へとリプログラムするための試薬をさらに含み、同様に生成されたiPS細胞株の品質を評価するための試薬を含む。体細胞を人工多能性幹(iPS)細胞へとリプログラムするために用いられる試薬の例は、当業者に周知であり、本明細書において述べる試薬を含むが、たとえば、その内容の全体が参照により本明細書に組み入れられる、国際特許出願WO2007/069666；WO2008/118820；WO2008/124133；WO2008/151058；WO2009/006997；および米国特許出願US2010/0062533；US2009/0227032；US2009/0068742；US2009/0047263；US2010/0015705；US2009/0081784；US2008/0233610；US7615374；米国特許出願第12/595,041号、EP2145000、CA2683056、AU8236629、12/602,184、EP2164951、CA2688539、US2010/0105100；US2009/0324559、US2009/0304646、US2009/0299763、US2009/0191159に開示される、体細胞をiPS細胞またはpiPS細胞へとリプログラムするための方法およびキットに限定されるわけではない。いくつかの態様において、キットは、その全体が参照により本明細書に組み入れられる、EP1970446、US2009/0047263、US2009/0068742、および2009/0227032に開示されるように、リプログラム細胞、たとえばiPS細胞をウイルス法または化学法によって生成するための試薬を含む。

20

【0334】

いくつかの態様において、本明細書において開示されるキットはまた、多くの細胞株の中から所望の多能性幹細胞株を選択するための少なくとも1つの試薬、たとえば細胞株の意図される用途のために1つまたは複数の適切な多能性幹細胞株を選択するための試薬を含む。そのような剤は、当技術分野において周知であり、細胞特異的系列マーカーに関して選択するための標識抗体などを含むがこれらに限定されるわけではない。いくつかの態様において、標識抗体は、蛍光標識される、または磁気ビーズなどによって標識される。いくつかの態様において、本明細書において開示されるキットは、本明細書において開示される方法に従って、既存のES細胞および/またはiPS細胞バンクをハイスループットでプロファイルおよび注釈付けするための少なくとも1つまたは複数の試薬を含む。

30

【0335】

1つの局面において、本発明は、本発明のアッセイ、方法、またはシステムによって選択される多能性幹細胞株を含むキットを提供する。上記の成分に加えて、キットはまた情報提供材料を含むことができる。情報提供材料は、本明細書において記述される方法、および/または本明細書において記述されるアッセイ、方法、およびシステムの成分の使用に関する説明、指示、販売、または他の材料でありうる。たとえば、情報提供材料は、多能性幹細胞を選択するための、多能性細胞の複数の特性の特徴決定を行うための、または本発明に従うスコアカードを作成するための方法を記述してもよい。キットが対象に投与するのに適した材料を含む場合、キットは任意で送達デバイスを含むことができるが、これらに限定されるわけではない。

40

【0336】

いくつかの態様において、各々の試料が多能性幹細胞株または異なる多能性幹細胞集団を含む1つまたは複数の試料(たとえば、試料のアレイ)を研究者が有することができる

50

場合、サービス提供者が運営する診断研究所において本明細書において開示される方法、キットおよびシステムを用いて評価するために、本明細書において開示される方法、システム、キット、およびデバイスをサービス提供者が行うことができる。そのような態様において、開示の本発明のアッセイ、方法、およびシステムを行った後、サービス提供者は、分析を行って、分析した各多能性幹細胞株の特徴に関するレポート、たとえばスコアカードを研究者に提供することができる。代替りの態様において、サービス提供者は、研究者にアッセイの未加工データを提供して、分析を研究者に任せることができる。いくつかの態様において、レポートは、電子的手段により研究者に通信または送付され、たとえば安全なウェブサイト上にアップロードされるかまたはe-メールもしくは他の電子通信手段により送付される。いくつかの態様において、研究者は、試料を任意の手段、たとえば郵便、速達郵便等によりサービス提供者に送ることができ、またはサービス提供者は、研究者から試料を収集して、それらをサービス提供者の診断研究所に送るというサービスを提供することができる。いくつかの態様において、研究者は、分析される試料を、サービス提供者の診断研究所の所在地に委託することができる。代替りの態様において、サービス提供者は、研究者の研究所に人員を送って、同様に研究者の研究所において研究者の多能性幹細胞株について、本明細書において開示される本発明のアッセイ、方法、およびシステムを行うためのキット、装置、および試薬を提供して、結果を分析して、分析した各多能性幹細胞株または複数の多能性幹細胞株の特徴に関するレポートを研究者に提供するという訪問サービスを提供する。

10

20

【0337】

偏差または系列スコアカードを作成するためのハイスループット試料処理のワークフローの例

制限的ではない実例として、スコアカードワークフローを、以下のケーススタディによって説明する：大企業（または基金）が米国集団のX%のために、10,000個のiPS細胞株を必要とするHLA-マッチiPS細胞株を提供する幹細胞バンクを設立することを計画している。細胞株は全て販売されており、研究者および企業にとってリソースを最も貴重にするために、各細胞株に関するスコアカード特徴を公表することが計画されている。自動化を促進するために、全てのiPS細胞株を96ウェルプレートまたは384ウェルプレートにおいて成長させる。ほとんどの試料処理はロボット化され、細胞株は全てバーコード処理されて、中央LIMSによって追跡される。スコアカード特徴決定は以下のように行う：

30

【0338】

(1) 偏差スコアカード / 多能性の確認：研究者は、液体取り扱いロボットを以下のようにロードする：(i) ウェル毎に1つのiPS細胞株を入れた1枚の96ウェルプレート、(ii) 96ウェルRNA抽出キット、(iii) 96個のマーカ-遺伝子および対照に関するプライマーを予めスポットしたカスタムqPCRプレート（96ウェルまたは384ウェル）。

【0339】

(2) ロボットは、プレート全体のRNA抽出を行って、各ウェルからのRNAを個別のqPCRプレート（96ウェルqPCRプレートを用いる場合）またはプレートの4分の1（384ウェルqPCRプレートを用いる場合）にピペットで移す。同じプレートにおいて逆転写を行って、バーコード処理したCt表をLIMSに転送する。

40

【0340】

(3) 系列スコアカード / 分化能の定量：ウェル当たり1つのiPS細胞株を有する96ウェルプレートから開始して、研究者らは、各ウェルから細胞を採取して、それらを新しい3枚の96ウェルプレートで平板培養して、胚様体（EB）分化のための3つの生物学的レプリケートを生じる。分化誘導培地を加えて、培地交換をせずにプレートをインキュベータの中にN日間放置する。

【0341】

(4) EB分化の一定期間（たとえば、n日）後、プレートを液体取り扱いロボットにロードして、分化特異的マーカ-遺伝子を有するカスタムqPCRプレートを用いることを唯一の例外として、qPCR分析を段階1および2に記述されるように行う。

50

【0342】

(5) 実験の終了時、研究者は、未加工Ct値をカスタムスコアカードソフトウェアにロードする。このソフトウェアは、一般的な任意のqPCR機器からの出力データフォーマットをインポートして、多数のハウスキーピング遺伝子を用いて相対的標準化を行い、スコアカード予測を算出する。

【0343】

(6) 遺伝子セットの選択。本明細書において開示されるように、スコアカードは、2つの独立した、しかし相補的な部分を含む：(i) 偏差スコアカード、および(ii) 系列スコアカード。いくつかの態様において、偏差スコアカードのデータを作成するためのアッセイは、所定の細胞株が多能性であるとして分類されるか否かを判定するために最も関連する遺伝子を有する1枚の96ウェルqPCRプレート（またはいくつかの態様において、384ウェルqPCRプレート上の4つの試料）からなりうる。いくつかの態様において、系列スコアカードのためのデータを作成するためのアッセイは、所定の細胞株の分化指向性を定量するための最も関連する遺伝子を有する、2枚の96ウェルプレート（またはいくつかの態様において、384ウェルqPCRプレート上の2つの試料）からなりうる。

10

【0344】

いくつかの態様において、多重qPCRアッセイを用いて両方のスコアカードに関する両方のアッセイのための最適な遺伝子選択をさらに検証および最適化することができる。さらに、いくつかの態様において、系列スコアカードアッセイの前に偏差アッセイを行って、関心対象の幹細胞株の多能性状態を決定して、系列スコアカードアッセイのためのEB分化アッセイの必要性をおそらく不要にすることができる。したがって、いくつかの態様において、偏差スコアカードアッセイと系列スコアカードアッセイの両方のために設計された1枚の384ウェルqPCRプレートを用いて検証相を行うことができる。いくつかの態様において、各細胞株のアッセイに関して、関心対象の各生物学的幹細胞株レプリケート用のプレート、その多能性状態の幹細胞株用のプレート、およびそのEB状態の幹細胞株用のプレートを含む、複数のプレートを用いる。いくつかの態様において、そのような384ウェルqPCRプレート（「技術由来プレート」）に含まれる遺伝子は、以下の遺伝子セット選択を用いて選択することができる：

20

【0345】

1. 標準化：各プレートは、6個の標準化遺伝子の技術的デュプリケート、3個の陽性対照、および1個の陰性対照を含む。

30

【0346】

2. 支持される細胞型 / 系列：qPCRに基づくスコアカードに関してNanoStringに基づくプロトタイプと同じである系列マーカー遺伝子を選択することができる（外胚葉、中胚葉、および内胚葉ならびに神経および造血系列、または表7、または表13Aおよび13Bおよび表14に記載される遺伝子の任意の選択物）。さらに、いくつかの態様において、系列マーカー遺伝子は、多能性細胞シグネチャー、表皮、間葉幹細胞、骨、軟骨、脂肪、筋肉、血管、心臓、リンパ細胞、骨髄細胞、肝臓、膵臓、上皮、運動ニューロン、単球-マクロファージが挙げられるがこれらに限定されるわけではないさらなる遺伝子セットカテゴリを含むことができる（表13Aおよび13Bおよび表14を参照されたい）。

40

【0347】

3. 追加の特色：いくつかの態様において、偏差および系列スコアカードアッセイのためのqPCRプレートはまた、以下を含むことができる：(i) 体細胞をiPSCにリプログラムするために一般的に用いられる4つのリプログラムウイルスに関するqPCRプライマー（たとえば、リプログラミング遺伝子Sox2、Oct4、c-myc、Klf4等のいずれかに対するプライマー）；ならびに(ii) 可能性がある試料の混合を検出するために雌雄分類のための5遺伝子シグネチャー（表14を参照されたい）；ならびに(iii) 広範なアポトーシスを検出するための1遺伝子シグネチャー。いくつかの態様において、偏差および系列スコアカードアッセイのためのqPCRプレートはまた、本発明者らが本明細書において同定した、ES細胞株およびiPS細胞株における最も転写的および / またはエピジェネティックに変動性の遺

50

伝子のサブセットを含むことができる。

【0348】

検証：いくつかの態様において、偏差スコアカードおよび系列スコアカードのためのデータを生成するためのアッセイに関してqPCRプレートを検証することができる。検証は3つの相で行うことができる。初回検証相では、NanoStringアッセイと類似の精度および予測力を提供するか否かを判定するために、qPCRプレートを評価するであろう。より多くの多能性幹細胞株に関するqPCRに基づくスコアカードの予測性、および関心対象の多様な異なる系列への分化指向性を評価および確認する第二の生物学的検証相を行うことができる。より初期の全てのデータとの技術的一貫性に関してqPCRプレートを最適にする最後のアッセイ検証を行うことができる。より具体的に、いくつかの態様において、検証相は、以下のように行われるであろう：

10

【0349】

1. 技術的qPCRアッセイ検証。NanoStringに基づくスコアカードからの結果を、qPCRに基づくスコアカードと直接比較することができ、NanoStringとqPCRプラットフォームのあいだの各遺伝子の精度、感度、および頑健性を比較することができる。さらに、qPCRに基づくスコアカードが定方向性の運動ニューロン分化効率の細胞特異的差を予測することができることを確認することができる。

【0350】

2. 生物学的qPCRアッセイの検証および範囲の拡大。本発明者らは、EBに基づくプロトコルを用いて運動ニューロン分化を予測するための系列スコアカードを広く検証した。類似のEBに基づくプロトコルを用いて造血細胞の分化に関する系列スコアカードの類似の検証を行うことができる。したがって、様々な異なる系列への分化効率を定量的に決定するためにいくつかの異なるさらなる分化プロトコルを用いて系列スコアカード予測性を検証することができる。さらに、偏差スコアカードを校正するために、たとえば、ヒト多能性細胞株、部分的リプログラム細胞株、胎児性癌細胞株等から選択されるがこれらに限定されるわけではない少なくとも約100個またはそれより多くの多能性幹細胞株を用いてqPCRアッセイを検証することができる。そのような検証を用いて、大規模産生のためであり、特定の幹細胞株または系列の好みに合うように調整される、qPCRに基づくスコアカードアッセイを最適化して再設計することができる。

20

【0351】

3. 技術的検証。いくつかの態様において、qPCRアッセイのソフトウェアおよびアッセイ取り扱い、たとえばプレートの安定性、およびqPCRプレートからの出力の解読の容易さなどを検証するために、さらなる検証が望ましい可能性がある。そのような検証および最適化は、一般的に当業者に公知である。

30

【0352】

スコアカードの使用

いくつかの態様において、本明細書において開示される方法、システム、キット、およびスコアカードは、多様な臨床および研究応用において用いることができる。例として、本明細書において開示される方法、システム、キット、およびスコアカードは、薬物に反応した多能性幹細胞株におけるエピジェネティックなおよび機能的ゲノム変化を同定するために有用であり、または薬物スクリーニングにおいて用いられる複数の多能性幹細胞株が同じ特性を有するように選択するために有用であり、このことは薬物スクリーニングの品質を確実にするために、および可能性がある任意のヒットが、異なる多能性幹細胞株の変動によるのではなくむしろ薬物の効果であることを確実にするために有用である。いくつかの態様において、本明細書において開示される方法、システム、キット、およびスコアカードは、移植された幹細胞株が癌細胞へと分化する素因を有しないことを確実にするための、治療的使用、たとえば幹細胞治療または他の再生医療にとって適している多能性幹細胞株を同定および選択するために有用である。同様に、本明細書において開示される方法、システム、キット、およびスコアカードは、iPSCが品質を保有して、他の多能性幹細胞と比較できることを確実にするための、哺乳動物、たとえばヒトから生成されたiPSC

40

50

を特徴決定および検証するために有用である。

【0353】

いくつかの態様において、本明細書において開示される方法、システム、キット、およびスコアカードは、特定の多能性幹細胞株の臨床での安全性および有用性を決定するために診療所で用いることができる。

【0354】

いくつかの態様において、本明細書において開示される方法、システム、キット、およびスコアカードは、たとえば有意なエピジェネティックなまたは機能的なゲノムの変化が時間と共に（たとえば、何回もの継代により、および凍結保存後に）起こらないことを確実にするために、異なる継代に対しておよび/または凍結保存技法の前および後で多能性幹細胞の特徴をモニターするための品質管理として用いることができる。たとえば、本明細書において開示される方法、システム、キット、およびスコアカードは、幹細胞バンクにおける全ての幹細胞を特徴決定するために、そのバンクに預けられた各幹細胞株のカatalogを作成するために、および幹細胞が、融解後にそれらが凍結前と同じ特性を確実に有することを確実にするために用いることができる。

10

【0355】

いくつかの態様において、各多能性幹細胞株に関する未加工データ（たとえば、DNAメチル化および/または遺伝子発現データ）および/またはスコアカードデータを、中央データベースに保存することができ、ここでデータおよび/またはスコアカードを用いて特定の用途または有用性に関して多能性幹細胞株を選択することができる。したがって、本発明の1つの局面は、複数の多能性幹細胞株に関するDNAメチル化データ、遺伝子発現データ、およびスコアカードの少なくとも1つを含むデータベースに関し、いくつかの態様において、データベースは、幹細胞バンクにおける複数の多能性幹細胞株に関するDNAメチル化データ、遺伝子発現データ、および/またはスコアカードを含む。

20

【0356】

いくつかの態様において、本明細書において開示される方法、システム、キット、およびスコアカードは、多能性幹細胞が異なる系列に分化する際の機能的なゲノムの変化をモニターするために研究において用いることができる。いくつかの態様において、本明細書において開示される方法、システム、キット、およびスコアカードは、特定の疾患からの多能性幹細胞の特徴をモニターおよび決定するために用いることができ、たとえば遺伝子欠損または特定の遺伝的多形を有する対象、および/または特定の疾患を有する対象由来の多能性幹細胞をモニターすることができ、たとえば健康な兄弟などの健康な対象由来の正常なiPSC細胞と比較して、ALSなどの神経変性疾患を有する対象に由来するiPSC細胞の機能的なゲノムの差をモニターおよび決定することができる。同様に、iPS細胞が、ES細胞または他の多能性幹細胞と比較して機能的ゲノミクスおよび分化指向性に関して同等であるか否かを判定することができる。さらに、本明細書において開示される方法、システム、キット、およびスコアカードは、奇形腫アッセイおよび/またはキメラマウスを生成する必要がなく、幹細胞株の多能性を十分に特徴決定することができ、それゆえ、多能性幹細胞株をハイスループットで特徴決定する能力を有意に増加させる。

30

【0357】

いくつかの態様において、スコアカードは、患者特異的iPS細胞株を作製および検証するための「全て含まれる」キットの中に入れることができる。たとえば、そのような態様において、キットは、(i) 試料収集デバイス、たとえば患者の体細胞または分化した細胞を収集するために必要な針またはチューブ、およびいくつかの態様において、患者の同意書、(ii) 患者の収集した体細胞または分化した細胞をiPS細胞へとリプログラムするための試薬、たとえばここで該キットは、本明細書において記述されるウイルス/DNA/RNA/タンパク質などの任意の数または組み合わせのリプログラミング因子、およびES細胞培地を含む、ならびに(iii) 本明細書において開示されるスコアカードを作成するためのアッセイ、たとえばDNAメチル化アッセイを行うための試薬、遺伝子発現アッセイを行うための試薬、およびiPS細胞株分化能の確認を行うための試薬を含むことができる。い

40

50

くつかの態様において、キットは、キットの品質管理としての、陽性対照（または陰性対象、たとえば多能性幹細胞株が望ましくない特徴を有すると同定されている場合）として用いることができる1つまたは複数の参照多能性幹細胞株を含むことができる。いくつかの態様において、キットはまた、たとえば評価される患者のiPS細胞との比較目的のために用いられる参照多能性幹細胞のスコアカードを含むことができる。いくつかの態様において、「全て含まれる」キットは、品質管理からの結果に基づく患者のiPS細胞株の有用性の予測（たとえば、本明細書において開示されるバイオインフォマティクスによる判定によって決定される）のために用いることができる。いくつかの態様において、「全て含まれる」キットはまた、関心対象の特定の細胞型（たとえば、心筋細胞、細胞、肝細胞、毛包幹細胞、軟骨、および造血細胞など）への新しく生成された本発明のiPS細胞株の定方向性分化のための材料、試薬、およびプロトコルをさらに含むことができる。

10

20

30

40

50

【0358】

いくつかの態様において、本明細書において開示されるスコアカード、方法、キット、およびアッセイを用いて、たとえば診療所もしくは臨床診断研究所において直接行うことができる、または専用の施設によって行われる「メールイン」サービスとして行うことができる、「細胞品質保証付（cell-to-quality assured）多能性幹細胞株」サービスなどのサービスを提供することができる。たとえば、そのようなサービスは、研究者または患者が体細胞（たとえば、分化した細胞）をサービス提供者に提出して、それによってサービス提供者は、本明細書において開示される一般的に公知の方法を用いて、体細胞からiPS細胞株を生成して、サービス提供者は、生成された多能性iPS細胞株に対して本明細書において開示される方法およびアッセイを行い、たとえばサービス提供者は（i）分化指向性アッセイ、（ii）DNAメチル化アッセイ、および任意で（iii）遺伝子発現アッセイを行って、その後分析を行って、分析された各個々のiPS細胞に関するスコアカードを作成する、という形で運営されるであろう。サービス提供者はまた、任意で特定の用途のために1つまたは複数の選択されたiPS細胞株の適格性を示唆し、たとえばサービス提供者は、運動ニューロン分化経路に沿って分化する効率が高いと同定された「iPS細胞株1」がニューロン分化にとって適しているであろうと示唆することができ、または同様に、肝臓系列に沿って分化する効率が高いと同定された「iPS細胞株2」が肝細胞再生医療において用いるために肝細胞への分化にとって適していることを示唆することができる。同様に、サービス提供者は、外れ値DNAメチル化遺伝子、および/または特定遺伝子の外れ値遺伝子発現レベル、たとえば癌遺伝子の外れ値DNAメチル化または遺伝子発現であると同定された「iPS細胞株6」が、癌形成リスクの可能性により再生医療において治療的に用いるために適していない可能性があることを示唆することができる。いくつかの態様において、サービス提供者は、サービス提供者が生成して分析した各iPS細胞株に関して推奨を行うことはできず、スコアカードの報告を行うことができる。いくつかの態様において、サービス提供者は、レポートスコアカードのコピーをつけて研究者または患者にiPS細胞株を返却する。

【0359】

いくつかの態様において、本明細書において開示されるスコアカード、方法、キット、およびアッセイは、データベースを作成するために用いることができ、ここでそのようなデータベースは、研究者が意図する用途について特定の多能性幹細胞株を特異的に選択するためにバンクにおける各多能性幹細胞株に関する各スコアカードのデータを含むデータベースを用いることができるように、品質管理され有用性が予測された多数の多能性幹細胞を含む多能性幹細胞保管庫、たとえば中央保管庫（たとえば、組織および/または細胞バンク）を構築してカタログ作成するために有用であろう。たとえば、データベースのユーザーは、データベースにリンクしたウェブサイト上の「私の応用にとって最善の細胞株を提案」というボタンをクリックして、研究者の特定の用途にとって有用な多数の細胞株の情報および同一性を得ることができる。いくつかの態様において、そのようなデータベースの使用は、ユーザーが、関心対象の特定の細胞型に関するマイクロアレイデータ（たとえば、DNAメチル化データおよび/または遺伝子発現データ）をアップロードすることができるように容易に拡張することができ、このマイクロアレイデータを、スコアカード

アルゴリズムを通して実行させて、結果を多能性幹細胞バンクのデータベーススコアカード結果と比較することができる。単純な類推において、データベースは、Googleの「類似サイトの検索」と類似のように機能することができ、それによってデータベースは、新規および/または混合された組織型に関して有用な細胞株を選択するための、または所望の分化した幹細胞株への分化能を有する細胞バンク中の多能性幹細胞株を同定するための、効率的な方法として用いることができるであろう。

【0360】

いくつかの態様において、本明細書において開示されるスコアカード、方法、キット、およびアッセイは、大量生産のために所望の多能性幹細胞株を同定および選択するために用いることができ、たとえば大量に、たとえば大きなバッチ培養またはバイオリクターにおいて良好におよび/または効率的に成長する多能性幹細胞株の品質を同定および特徴決定および検証するために、およびバルク培養において明記された細胞型へと効率的に分化することができる多能性幹細胞株を選択するために、本明細書において開示される方法、アッセイ、およびスコアカードを用いることができる。

10

【0361】

別の態様において、本明細書において開示されるスコアカード、方法、キット、およびアッセイは、多能性の頑健性の特性に基づいて多能性幹細胞株を選択するために用いることができ、たとえば本明細書において開示される方法、アッセイ、およびスコアカードは、インピットで容易に培養される（たとえば、ほとんど注意を必要としない、および/または容易に自発的に分化しない、および/または多能性特性を維持する）多能性幹細胞株を同定するために用いることができる。たとえば、いくつかの態様において、多能性幹細胞株を、培養前に、その後培養中および培養後の異なる時点で、ならびに異なる培養条件および培地条件で、方法、アッセイ、およびスコアカードを用いて評価して、短期間および長期間培養条件でその最初の品質を維持する1つまたは複数の多能性幹細胞株を同定することができる。

20

【0362】

別の態様において、本明細書において開示されるスコアカード、方法、キット、およびアッセイは、薬物の反応性に関して多能性幹細胞株を選択するために用いることができ、たとえば多能性幹細胞株は、薬物または他の剤または刺激との接触の前、最中、および後に、本明細書において開示される方法、アッセイ、およびスコアカードを用いて評価して、多能性幹細胞株の薬物代謝および/または薬物ゲノムシグネチャーを生成することができ、これは、たとえば薬物毒性アッセイを含む薬物スクリーニングおよび新薬開発にとって特に有用でありうる多能性幹細胞株を同定するために用いることができる。

30

【0363】

別の態様において、本明細書において開示されるスコアカード、方法、キット、およびアッセイは、その安全性プロファイルに基づいて多能性幹細胞株を選択するために用いることができ、たとえば、多能性幹細胞株を、本明細書において開示される方法、アッセイ、およびスコアカードを用いて評価して、それが癌細胞へと形質導入する可能性、または転移の可能性、もしくは特定の細胞型へと分化する可能性、または脱分化する可能性を同定することができ、これは、細胞補充治療および再生医療などの臨床応用において多能性幹細胞株またはその分化した子孫の安全性を検証するために特に有用である。

40

【0364】

別の態様において、本明細書において開示されるスコアカード、方法、キット、およびアッセイは、効能に関して多能性幹細胞株を選択するために用いることができる。たとえば、多能性細胞株に由来する分化した細胞が特定の所望の細胞系列に沿って分化し続けるか否かおよび/またはどのように良好に分化し続けるか、および/またはそれらが対象、たとえばヒト患者もしくは動物モデル（たとえば、ラットまたはマウス疾患モデル等）に移植された後に増殖するか否かを予測するために、特定の多能性幹細胞株のスコアカード予測を用いることができる。より一般的に、いくつかの態様において、スコアカードを用いて、多能性細胞株の挙動のみならず、多能性細胞株に直接または間接的に由来する分化

50

した細胞の挙動を予測することができる。

【0365】

別の態様において、本明細書において開示されるスコアカード、方法、キット、およびアッセイは、インビボで多能性幹細胞の同じまたは非常に類似した特徴を有する多能性幹細胞株を選択するために（たとえば、インビボ環境における細胞を忠実に表す多能性幹細胞を選択するために）用いることができる。たとえば、インビボでその対応する細胞に極めて類似した多能性幹細胞株を用いることが重要であることから、疾患モデリングに適した多能性幹細胞株を同定するために、本明細書において開示される方法、アッセイ、およびスコアカードを用いて多能性幹細胞株を評価することができる。したがって、当業者は、本明細書において開示されるスコアカードを容易に用いて、インビボでの細胞の挙動と比較して参照の純粋ES細胞株集団の変動を最小限にするために、たとえば多能性幹細胞株の特性（スコアカードに記載される）を、対象（たとえば、動物モデルまたは齧歯類疾患モデルなどの疾患モデル）から採取した対応する細胞の特性と比較することによって、どの多能性細胞株がインビボでその対応する細胞に類似するかを予測することができる。

10

【0366】

別の態様において、本明細書において開示されるスコアカード、方法、キット、およびアッセイは、たとえばインビトロで分化させて細胞型を作製するために有用であるが、ヒトES細胞株（たとえば、多能性幹細胞へとさらにリプログラムすることができる、ヒト基底状態ES細胞および部分的リプログラム細胞株、たとえば部分的人工多能性幹（piPS）細胞）の通常の見解に当てはまらない多能性幹細胞株に関する情報を提供するために、異なるまたは新しい多能性または多分化能状態の多能性幹細胞株の選択、および/または品質管理、および/または検証のために用いることができる。

20

【0367】

インビトロ培養および継代を継続すると、iPS細胞株の品質が改善することが示されている（Polo et al., Nat Biotechnol. 2010 Aug;28(8):848-55, and Nat Rev Mol Cell Biol. 2010 Sep;11(9):601, and Nat Rev Genet. 2010 Sep;11(9):593を参照されたい）。一方で、持続的な継代は費用がかさむ。したがって、いくつかの態様において、本明細書において開示されるスコアカード、方法、キット、およびアッセイは、多能性幹細胞株の品質を改善するために何回継代すれば十分であるかを測定するために用いることができる。

30

【0368】

さらなる態様において、本明細書において開示されるスコアカード、方法、キット、およびアッセイは、多能性幹細胞を特徴決定およびモニターおよび検証するために多様な異なる研究および臨床での用途において用いることができ、たとえば典型的な応用は、（i）疾患の機序に関心を有する研究所および/または企業（たとえば、疾患モデリングおよび小規模薬物スクリーニングのためにiPS細胞株ならびに分化した細胞を作製する複雑さを低減させるために、本明細書において開示されるキットまたはサービスを用いて）、（ii）疾患の所定の標的に関する低分子および/または生物学的製剤を同定しようと試みる研究所および/または企業（たとえば、薬物スクリーニングのために高度に標準化された細胞を多数産生することができるように、本明細書において開示されるキットおよび/またはサービスを用いて）、（iii）ヒトまたは動物に移植するための細胞を産生することが関心対象である、多能性幹細胞株を品質管理および検証するための臨床および前臨床研究グループ（たとえば、規制当局の承認、たとえばFDAの承認を得るために十分である精度のレベルで品質管理を可能にするために本明細書において開示されるキットおよび/またはサービスを用いて）、（iv）顧客に、提供時に、アドバイス、ならびに多能性幹細胞株の性能および品質および有用性に関するデータを含む情報を提供しようと考えている組織バンク（たとえば、本明細書において開示されるキットおよび/またはサービスを用いて、多数の多能性幹細胞株の品質および/または有用性に関する偏りのない評価を、たとえば安価にハイスループット様式で提供する、たとえば最終的に、細胞バンクに保存された全細胞株集団をカバーするために100,000個の多能性幹細胞株に対してアッセイを行う

40

50

)、(v)たとえば将来の再生医療目的のためにあるタイプの健康保険政策として、消費者自身および/またはその子供またはその子孫の体の分化した細胞から生成された少なくとも1つまたは複数の多能性細胞株、たとえばiPS細胞株(またはpiPS細胞株)を作製してバンクに預けることを望む個人消費者、などの、しかしこれらに限定されるわけではない領域での応用を含む。

【0369】

治療的使用

癌、糖尿病、心不全、筋障害、セリアック病、神経障害、神経変性障害、およびリソソーム蓄積症、ならびにALS、パーキンソン病、単一遺伝子疾患、およびメンデル遺伝病、加齢、ヒトの体の全身消耗、リウマチ性関節炎、および他の炎症疾患、出生時欠損等の疾患のいずれかなどの様々な疾患および障害が、幹細胞治療の可能性がある標的として示唆されている。したがって、本発明のアッセイ、方法、システム、およびキットは、処置のために対象に投与するための多能性幹細胞を選択するために用いることができる。

10

【0370】

それゆえ、1つの局面において、本発明は、対象に多能性幹細胞(たとえば、多能性細胞、多能性細胞に由来する分化した細胞、およびリプログラミング(たとえば、トランス分化)を伴う他の方法によって得られた分化した細胞)を投与する段階を含む、対象における疾患または障害を処置、予防、または改善する方法であって、多能性幹細胞が本発明のアッセイ、キット、方法、またはシステムによって選択される方法を提供する。多能性幹細胞は、対象に投与される前に特定系列に沿って分化するように処置することができるが、これらに限定されるわけではない。

20

【0371】

本発明の方法に適した投与経路は、局所投与および全身投与の両方を含む。一般的に、局所投与によって、対象の全身と比較して特定位置への細胞の送達が行われるが、全身投与では対象の本質的に全身への細胞の送達が行われる。例示的な投与様式には、注射、注入、点滴注入、吸入、または摂取が挙げられるがこれらに限定されるわけではない。「注射」には、静脈内、筋肉内、動脈内、髄腔内、心室内、包内、眼窩内、心臓内、皮内、腹腔内、経気管、皮下、表皮下、関節内、嚢下、クモ膜下、脊髄内、脳脊髄内、および胸骨内の注射および注入が挙げられるがこれらに限定されるわけではない。1つの局所投与法は筋肉内注射である。

30

【0372】

1つの好ましい投与法は、そのような多能性細胞、または多能性幹細胞に由来する分化した子孫の対象への移植である。「移植」という用語は、たとえば自家移植(患者内のある位置から細胞を取り出して、同じ患者の同じまたは別の位置へと移入)、同種異系移植(同じ種のメンバー間での移植)および異種移植(異なる種のメンバー間での移植)を含む。当業者は、本発明に従う様々な疾患を処置するために、細胞を植え込むまたは移植する方法を承知しているであろう。

【0373】

対象に投与するために、多能性幹細胞を、薬学的に許容される組成物で提供することができる。これらの薬学的に許容される組成物は、1つまたは複数の薬学的に許容される担体(添加剤)および/または希釈剤と共に調合された1つまたは複数の多能性細胞を含む。以下に詳細に記述するが、本発明の薬学的組成物は、以下のために適合させた剤形を含む、固体または液体剤形で投与するために特に調合することができる:(1)経口投与、たとえば水薬(水性または非水性溶液または懸濁液)、胃管栄養、ロゼンジ、糖衣錠、カプセル剤、丸剤、錠剤(たとえば、口腔内、舌下、および全身吸収を標的とした錠剤)、ポーラス、粉剤、顆粒剤、舌に適用するためのパスタ剤;(2)非経口投与、たとえば、滅菌水溶液または懸濁液として、または徐放性製剤としての、たとえば皮下、筋肉内、静脈内、または硬膜外注射;(3)たとえば、皮膚に適用するためのクリーム、軟膏、または徐放性パッチまたはスプレーなどの表面適用;(4)腔内または直腸内、たとえばベッサリー、クリームまたはフォームとして;(5)舌下;(6)眼;(7)経皮;(8)経粘膜

40

50

；または(9)経鼻。さらに、細胞を、対象に植え込むことができ、またはドラッグデリバリーシステムを用いて注射することができる。たとえば、その全ての内容が参照により本明細書に組み入れられる、Urquhart, et al., Ann. Rev. Pharmacol. Toxicol. 24: 199-236 (1984); Lewis, ed. "Controlled Release of Pesticides and Pharmaceuticals" (Plenum Press, New York, 1981)；米国特許第3,773,919号；および米国特許第353,270,960号を参照されたい。

【0374】

本明細書において用いられる「薬学的に許容される」という用語は、健全な医学的判断の範囲内で、過度の毒性、刺激、アレルギー反応、または他の問題または合併症を生じることなく、妥当な利益/リスク比に釣り合っ、ヒトおよび動物の組織に接触して用いるのに適した化合物、材料、組成物、および/または投与剤形を意味する。

10

【0375】

本明細書において用いられる「薬学的に許容される担体」という用語は、液体または固体増量剤、希釈剤、賦形剤、製造補助剤(たとえば、潤滑剤、タルク、ステアリン酸マグネシウム、ステアリン酸カルシウム、またはステアリン酸亜鉛、またはステアリン酸(steric acid))、または身体の器官の1つまたはその一部から身体の別の器官またはその一部への本発明の化合物の運搬または輸送に関係する溶媒封入材料などの薬学的に許容される材料、組成物、または媒体を意味する。各担体は、製剤の他の成分と適合性であり、患者にとって有害ではないという意味において「許容可能」でなければならない。薬学的に許容される担体として役立つことができる材料のいくつかの例は：(1)乳糖、グルコースおよびスクロースなどの糖；(2)コーンスターチおよびジャガイモデンプンなどのデンプン；(3)セルロース、ならびにカルボキシメチルセルロース、メチルセルロース、エチルセルロース、結晶セルロースおよび酢酸セルロースなどのその誘導体；(4)粉末トラガカント；(5)麦芽；(6)ゼラチン；(7)ステアリン酸マグネシウム、ラウリル硫酸ナトリウムおよびタルクなどの潤滑剤；(8)カカオバターおよび坐剤用ロウなどの賦形剤；(9)落花生油、綿実油、サフラワー油、ゴマ油、オリーブ油、コーン油、および大豆油などの油；(10)プロピレングリコールなどのグリコール；(11)グリセリン、ソルビトール、マンニトールおよびポリエチレングリコール(PEG)などのポリオール；(12)オレイン酸エチル、およびラウリン酸エチルなどのエステル；(13)寒天；(14)水酸化マグネシウムおよび水酸化アルミニウムなどの緩衝剤；(15)アルギン酸；(16)発熱物質を含まない水；(17)等張食塩水；(18)リンゲル液；(19)エチルアルコール；(20)pH緩衝液；(21)ポリエステル、ポリカーボネート、および/またはポリアンヒドリド；(22)ポリペプチドおよびアミノ酸などの増量剤；(23)血清アルブミン、HDL、およびLDLなどの血清成分；(22)エタノールなどのC₂~C₁₂アルコール；ならびに(23)薬学的製剤において用いられる他の非毒性の適合性の物質が挙げられる。湿潤剤、着色剤、放出剤、コーティング剤、甘味料、香味料、香料、保存剤、および抗酸化剤も同様に製剤に存在することができる。「賦形剤」、「担体」、「薬学的に許容される担体」またはその他などの用語は、本明細書において互換的に用いられる。

20

30

【0376】

多能性幹細胞を投与するという文脈において、「投与する」という用語はまた、対象におけるそのような細胞の移植を含む。本明細書において用いられる「移植」という用語は、対象に少なくとも1つの細胞を植え込むまたは移入するプロセスを意味する。「移植」という用語は、たとえば自家移植(患者内のある位置から細胞を取り出して、同じ患者の同じまたは別の位置へと移入)、同種異系移植(同じ種のメンバー間での移植)および異種移植(異なる種のメンバー間での移植)を含む。

40

【0377】

多能性幹細胞は、薬学的活性物質と併用して対象に投与することができる。本明細書において用いられる「薬学的活性物質」という用語は、インピボで放出された場合に、所望の生物活性、たとえばインピボで治療的、診断的、および/または予防的特性を保有する剤を意味する。この用語は、安定化されたおよび/または長期間放出されるように調合さ

50

れた薬学的活性物質を含むと理解される。例示的な薬学的活性物質は、その全ての完全な内容が参照により本明細書に組み入れられる、Harrison's Principles of Internal Medicine, 13th Edition, Eds. T.R. Harrison et al. McGraw-Hill N.Y., NY; Physicians Desk Reference, 50th Edition, 1997, Oradell New Jersey, Medical Economics Co.; Pharmacological Basis of Therapeutics, 8th Edition, Goodman and Gilman, 1990; United States Pharmacopeia, The National Formulary, USP XII NF XVII, 1990; current edition of Goodman and Gilman's The Pharmacological Basis of Therapeutics; and current edition of The Merck Indexにおいて見いだされるものを含むがこれらに限定されるわけではない。

【0378】

10

本明細書において用いられる「対象」は、ヒトまたは動物を意味する。通常、動物は霊長類、齧歯類、家畜動物、または競技用動物などの脊椎動物である。霊長類には、チンパンジー、カニクイザル、クモザル、およびマカク、たとえばアカゲザルが挙げられる。齧歯類には、マウス、ラット、ウッドチャック、フェレット、ウサギ、およびハムスターが挙げられる。家畜および競技用動物には、ウシ、ウマ、ブタ、シカ、バイソン、水牛、ネコ種、たとえばネコ、イヌ種、たとえばイヌ、キツネ、オオカミ、トリ種、たとえばニワトリ、エミュー、ダチョウ、および魚、たとえばマス、ナマズ、およびサケが挙げられる。患者または対象は、前述の任意のサブセット、たとえば上記の全てを含むが、ヒト、霊長類、または齧歯類などの1つまたは複数の群または種を除外する。本明細書において記述される局面のある態様において、対象は哺乳動物、たとえば霊長類、たとえばヒトである。「患者」および「対象」という用語は本明細書において互換的に用いられる。「患者」および「対象」という用語は、本明細書において互換的に用いられる。対象は雄性または雌性でありうる。

20

【0379】

好ましくは、対象は哺乳動物である。哺乳動物はヒト、非ヒト霊長類、マウス、ラット、イヌ、ネコ、ウマ、またはウシでありうるが、これらの例に限定されるわけではない。ヒト以外の哺乳動物は、自己免疫疾患または炎症に関連する障害の動物モデルを表す対象として都合よく用いることができる。さらに、本明細書において記述される方法および組成物は、家畜動物および/またはペットを処置するために用いることができる。

【0380】

30

対象は、幹細胞に基づく治療が有用である疾患を特徴とする障害を患っているまたは有すると既に診断されているかまたは同定されている対象でありうる。

【0381】

対象は、幹細胞に基づく治療によって現在処置されていない対象でありうる。

【0382】

本明細書において記述される局面のいくつかの態様において、方法は、幹細胞に基づく治療により恩典が得られる疾患を有する対象を選択する段階をさらに含む。

【0383】

本明細書において用いられる、「神経変性疾患または障害」という用語は、特にアルツハイマー病、パーキンソン病、ハンチントン病、筋萎縮性側索硬化症、てんかん、および筋ジストロフィーなどのニューロンレベルでの中枢神経系(CNS)の変性または変化を特徴とする疾患または状態を含む。この用語はさらに、白質脳症および白質ジストロフィーなどの神経-炎症および脱髄性の状態または疾患を含む。例示的な神経変性障害には、AIDS認知症複合症、副腎脳白質ジストロフィー、アレキサンダー病、アルパーズ病、アルツハイマー病、筋萎縮性側索硬化症、毛細血管拡張性運動失調、バッテン病、ウシ海綿状脳症、キャナヴァン病、皮質基底核変性症、クロイツフェルト-ヤコブ病、レヴィー小体認知症、致死性家族性不眠症、前頭側頭葉変性、ハンチントン病、乳児レフサム病、ケネディ病、クラッペ病、ライム病、マチャド・ジョゼフ病、多発性硬化症、多系統萎縮、神経有棘赤血球増加症、ニーマン・ピック病、パーキンソン病、ピック病、原発性側索硬化症、進行性核上麻痺、レフサム病、ザントホフ病、びまん性ミエリン破壊性硬化症、脊髄小

40

50

脳運動失調、脊髄の亜急性複合変性、脊髄ろう、テイ・サックス病、毒性脳症、および感染性海綿状脳症が挙げられるがこれらに限定されるわけではない。

【0384】

本明細書において用いられる「癌」という用語は、調節されないまたは制御されない細胞の成長を特徴とする悪性疾患、たとえば癌腫、肉腫、白血病、およびリンパ腫を含む。

「癌」という用語は、原発性の悪性腫瘍（たとえば、その細胞が当初の腫瘍部位以外の対象の体の部位に移動していない腫瘍）、および二次性悪性腫瘍（たとえば、転移、すなわち当初の腫瘍部位とは異なる二次的な部位への腫瘍細胞の移動により生じる腫瘍）を含む。

【0385】

「癌腫」という用語は、呼吸器系の癌腫、消化管の癌腫、尿性器系の癌腫、精巣の癌腫、乳房の癌腫、前立腺の癌腫、内分泌系の癌腫、黒色腫、絨毛上皮腫、ならびに子宮頸部、肺、頭頸部、結腸および卵巣の癌腫を含む、上皮または内分泌組織の悪性疾患を含む。

「癌腫」という用語はまた、癌腫様および肉腫様組織で構成される悪性腫瘍を含む癌肉腫を含む。「腺癌」は、腫瘍細胞が認識可能な腺様構造を形成する腺様組織または腫瘍に由来する癌腫を意味する。

【0386】

「肉腫」という用語は、中胚葉結合組織の悪性腫瘍、たとえば骨、脂肪、および軟骨の腫瘍を含む。

【0387】

「白血病」および「リンパ腫」という用語は、骨髄の造血細胞の悪性疾患を含む。白血病は単細胞として増殖する傾向があるが、リンパ腫は固形腫瘍塊として増殖する傾向がある。白血病の例には、急性骨髄性白血病（AML）、急性前骨髄球性白血病、慢性骨髄性白血病、混合型白血病、急性単芽球性白血病、急性リンパ芽球性白血病、急性非リンパ芽球性白血病、芽球性マントル細胞白血病、骨髄異形成症候群、T細胞白血病、B細胞白血病、および慢性リンパ球性白血病が挙げられる。リンパ腫の例には、ホジキン病、非ホジキンリンパ腫、B細胞リンパ腫、上皮親和性リンパ腫、混合リンパ腫、未分化大細胞リンパ腫、胃粘膜および非胃粘膜関連リンパ組織リンパ腫、リンパ増殖疾患、T細胞リンパ腫、パーキットリンパ腫、マントル細胞リンパ腫、びまん性大細胞型リンパ腫、リンパ形質細胞性リンパ腫、および多発性骨髄腫が挙げられる。

【0388】

たとえば、本発明のアッセイ、キット、方法、およびシステムによって選択される多能性細胞は、乏突起神経膠腫、星状細胞腫、多形グリア芽細胞腫、子宮頸癌、子宮類内膜癌、漿液性内膜癌、卵巣類内膜癌、卵巣ブレンナー腫瘍、ムチン性卵巣癌、漿液性卵巣癌、子宮癌肉腫、乳房小葉癌、乳管癌、乳房髓質癌、ムチン性乳癌、乳腺管状癌、甲状腺腺癌、甲状腺濾胞癌、甲状腺髓様癌、甲状腺乳頭癌、副甲状腺腺癌、副腎腺腫、副腎癌、褐色細胞腫、軽度異形成大腸腺腫、中程度異形成大腸腺腫、高度異形成大腸腺腫、大腸腺癌、食道腺癌、肝細胞癌、口癌、胆嚢腺癌、膵臓腺癌、小腸腺癌、びまん性胃腺癌、前立腺（ホルモン不応性）、前立腺（無処置）、嫌色素性腎癌、腎明細胞癌、腎好酸性細胞腫、腎乳頭癌、精巣非セミノーマ癌、精巣セミノーマ、移行上皮膀胱癌、肺腺癌、肺大細胞癌、肺小細胞癌、肺扁平上皮癌、ホジキンリンパ腫、MALTリンパ腫、びまん性大B細胞性非ホジキンリンパ腫（NHL）、NHL、胸腺腫、皮膚悪性黒色腫、皮膚基底細胞腫、皮膚扁平上皮癌、皮膚メルケル細胞癌、皮膚良性母斑、脂肪腫、および異常な細胞成長を伴う脂肪肉腫などの多くの種類の癌を処置するために用いることができる。

【0389】

薬物スクリーニング

本発明の方法、アッセイ、システム、およびキットは、良好に定義されたヒト細胞に基づくインビトロアッセイを開発するために用いることができる。薬物スクリーニング/試験および毒性試験に関する既存のアッセイは、それらが、動物起源、不死化細胞株、または死体由来の細胞であることからいくつかの欠点を有する。これらの代替物は正常なヒト

10

20

30

40

50

細胞の生理学を十分に反映していないことが多いので、幹細胞由来のアッセイ（たとえば、心臓および肝細胞の均一な集団）を将来的に確立することができれば、これらの目的にとって重要な役割を果たしうる。たとえば、本発明の方法、アッセイ、システム、およびキットは、疾患の表現型である系列に沿って分化することができる多能性幹細胞を同定および/もしくは検証するために用いることができる。さらにもしくはまたは、本発明の方法、アッセイ、システム、およびキットは、器官、および/または組織系列、またはその一部へと分化することができる多能性幹細胞を同定および/または検証するために用いることができる。次に、そのような同定された多能性細胞を、試験化合物をスクリーニングするために用いることができる。

【0390】

さらに、ヒト疾患に関連して分子および細胞レベルで目下入手可能となった相次ぐ新しい情報（たとえば、マイクロアレイデータ）により、病原性の相互関係に関する仮説を立てて試験することが重要となっている。全ての発達段階由来の特定細胞型およびさらに着床前遺伝子診断に基づいて病態を有すると思われる胚盤胞由来の特定細胞型の実験的利用は、ヒト疾患のモデリングおよび理解の局面において有用である可能性がある。したがって、そのような細胞株はまた、薬物の試験にとっても貴重であろう。

【0391】

したがって、本発明は、以下の段階を含む、生物活性に関して試験化合物をスクリーニングする方法を提供する：(a) 特定系列に沿った分化に関して同定および検証される多能性幹細胞を得る段階；(b) 任意で、多能性幹細胞を特定系列に分化させるか、または分化することを可能にする段階；(c) 細胞に試験化合物を接触させる段階；ならびに(d) 該細胞に及ぼす該化合物の任意の効果を決定する段階。細胞に及ぼす効果は、直接観察することができるか、またはレポーター分子を用いることによって間接的に観察することができる効果でありうる。

【0392】

本明細書において用いられる「生物活性」または「生物活性 (bioactivity)」という用語は、試験化合物が生物試料に影響を及ぼす能力を意味する。生物活性は、生物アッセイにおける刺激、阻害、調節、毒性、または致死反応の誘発を含むことができるがこれらに限定されるわけではない。たとえば、生物活性は、化合物が酵素の効果を調整する、受容体を遮断する、受容体を刺激する、1つまたは複数の遺伝子の発現レベルを調整する、細胞増殖を調整する、細胞分裂を調整する、細胞形態を調整する、またはその組み合わせを行うことができることを意味することができる。いくつかの例において、生物活性は、生物試料において毒性効果を生じる試験化合物の能力を意味することができる。

【0393】

先に述べたように、特定系列は、疾患の表現型および/または遺伝子型である系列でありうる。または、特定系列は、器官および/または組織またはその一部の表現型および/または遺伝子型である系列でありうる。

【0394】

本明細書において用いられる「試験化合物」という用語は、細胞に対して効果を有するか否かに関してスクリーニングされる化合物の集合体を意味する。試験化合物は、化学物質、化学物質の混合物、たとえば、多糖類、有機または無機低分子（たとえば、分子量200ダルトン未満、1000ダルトン未満、1500ダルトン未満、1000ダルトン未満、または500ダルトン未満を有する分子）、生体高分子、たとえばペプチド、タンパク質、ペプチドアナログ、およびそのアナログおよび誘導體、ペプチド模倣体、核酸、核酸アナログおよび誘導體、細菌、植物、真菌、または動物の細胞または組織などの生物材料から作製された抽出物、天然または合成の組成物を含む多様な異なる化合物を含みうる。

【0395】

実践される特定の態様に応じて、試験化合物は、溶液中で遊離状態で提供されうる、または担体、もしくは固相支持体、たとえばビーズに結合されうる。多くの適切な固相支持体が、試験化合物の固定のために用いられうる。適切な固相支持体の例には、アガロース

10

20

30

40

50

、セルロース、デキストラン（すなわちSephadex、Sepharoseとして市販される）、カルボキシメチルセルロース、ポリスチレン、ポリエチレングリコール（PEG）、濾紙、ニトロセルロース、イオン交換樹脂、プラスチックフィルム、ポリアミンメチルビニルエーテルマレイン酸コポリマー、ガラスビーズ、アミノ酸コポリマー、エチレン-マレイン酸コポリマー、ナイロン、絹等が挙げられる。さらに、本明細書において記述される方法に関して、試験化合物は、個別にまたはグループでスクリーニングされうる。グループスクリーニングは、有効な試験化合物に関するヒット率が、所定の群に関して1つより多くの陽性結果が予想されないほど低いと予想される場合には、特に有用である。

【0396】

多数の低分子ライブラリが当技術分野において公知であり、市販されている。これらの低分子ライブラリは、本明細書において記述されるスクリーニング法を用いて炎症の阻害に関してスクリーニングすることができる。たとえば、NIH Roadmap, Molecular Libraries Screening Centers Network (MLSCN)からの10,000個の化合物および86,000個の化合物からのライブラリなどの、Vitas-M LabのライブラリおよびBiomol International, Inc.の化学物質ライブラリをスクリーニングすることができる。化合物ライブラリの包括的リストは、http://www.broad.harvard.edu/chembio/platform/screening/compound_libraries/index.htmで見いだされうる。化学物質ライブラリまたは化合物ライブラリは、ハイスループットスクリーニングまたは工業的製造において通常、最終的に用いられる保存された化学物質の集合体である。化学物質ライブラリは、単純な意味で、一連の保存された化学物質からなりうる。各化学物質は、化合物の化学構造、純度、量、および物理化学特徴などの情報と共に、何らかの種類のデータベースにおいて保存された関連情報を有する。

10

20

【0397】

化合物は、対照と比較して適切な期間にわたって細胞に対して効果を発揮することができる任意の濃度で試験することができるが、これらに限定されるわけではない。いくつかの態様において、化合物は、約0.01 nMから約1000 mM、約0.1 nMから約500 μ M、約0.1 μ Mから約20 μ M、約0.1 μ Mから約10 μ M、または約0.1 μ Mから約5 μ Mの範囲の濃度で試験される。

【0398】

化合物のスクリーニングアッセイは、ハイスループットスクリーニングで用いられうる。ハイスループットスクリーニングは、所定の活性に関して化合物のライブラリが試験されるプロセスである。ハイスループットスクリーニングは、多数の化合物を迅速かつ同時にスクリーニングすることを求める。たとえば、マイクロタイタープレートおよび自動化アッセイ機器を用いて、製薬企業は、1日あたり100,000回程度の多数のアッセイを同時に行いうる。

30

【0399】

本発明の化合物スクリーニングアッセイは、観察可能なレポーター機能の2回以上の測定を伴いうる。多数回の測定は、試験化合物のインキュベーション時間のあいだの生物活性の追跡を可能にしうる。1つの態様において、レポーター機能を複数回測定することにより、異なるインキュベーション時間での試験化合物の効果のモニタリングが可能となる。

40

【0400】

同定された試験化合物が意図される用途にとって望ましい特性を有するか否かをさらに同定するために、スクリーニングアッセイの後に、引き続きアッセイを行ってもよい。たとえば、スクリーニングアッセイの後に、生物学的利用率、毒性、または薬物動態のいずれかの測定からなる群より選択されるがこれらの方法に限定されるわけではない第二のアッセイを行ってもよい。

【0401】

多能性幹細胞株のスコアカードを作成するためのバイオインフォマティクス分析のアルゴリズムおよび方法

本明細書において述べるように、スコアカードはいくつかの成分を含む：(i) エピジ

50

エネティックな修飾を同定するための、たとえば、正常なエピジェネティックな変動と比較した、たとえば参照多能性細胞株における標的遺伝子セットに関するDNAメチル化の正常な変動と比較した多能性細胞におけるDNAメチル化遺伝子の外れ値を同定するための、DNAメチル化アッセイの使用、(ii)参照多能性細胞株における標的遺伝子セットに関するDNA発現レベルの正常な変動と比較して遺伝子発現レベルが多能性細胞株において外れ値である遺伝子を同定するための、遺伝子発現アッセイの使用、(iii)(i)および(ii)からのエピジェネティックな修飾(たとえば、DNAメチル化)および/または遺伝子発現データを用いて、および/または分化誘導されている、たとえば定方向性分化誘導されている多能性細胞株由来の遺伝子発現/DNAメチル化データを用いて細胞の分化の偏りを予測するための、分化アッセイの使用。

10

【0402】

これらの3つの応用またはアッセイの各々は、多能性細胞株の品質および有用性についての実際的に有用な指標を得るために、異なるバイオインフォマティクス法を必要とする。

【0403】

いくつかの態様において、および本明細書において述べるように、任意のDNAメチル化法を用いることができ、たとえばDNAメチル化分析を、濃縮に基づく方法(たとえば、MeDIP、MBD-seqおよびMethylCap)、パイサルファイトに基づく方法(たとえば、RRBS、パイサルファイトシークエンシング、Infinium、GoldenGate、COBRA、MSP、MethyLight)、ならびに制限消化法(たとえば、MRE-seq)が挙げられるがこれらに限定されるわけではない多数の方法によって行うことができる。これらのDNAメチル化法の各々は、スコアカード分析にとって有用なデータを作成するために、データの前処理および標準化のための特殊なバイオインフォマティクス法を必要とする。これらは、たとえばGCおよびCpGパイアスの補正、ゲノムDNA配列に対するパイサルファイト特異的アラインメント等を含む。

20

【0404】

DNAメチル化データを適切に標準化した後、多能性細胞株またはその子孫の意図される使用を促進しうるまたは妨害しうる、変化したDNAメチル化レベルを示す任意の遺伝子および/またはゲノム領域を同定する。いくつかの態様において、本発明者らは、関心対象の多能性細胞株のDNAメチル化プロファイルを、1つまたは複数の参照多能性幹細胞株、たとえば既に特徴決定された良好な、または既に特徴決定された不良な多能性細胞株と比較することによってそのようなゲノム領域を同定する統計学的アルゴリズムを開発した。技術的には、これは、所定の候補遺伝子座のセットの各々に統計学的検定(たとえば、t-検定、フィッシャーの正確確率検定、ANOVA)を適用することによって行われる。頑健性を改善するために、偽発見率および細胞株と参照多能性幹細胞株とのDNAメチル化の絶対差に対して閾値を用いることができ、参照多能性幹細胞株の変動性を考慮に入れることができる。

30

【0405】

実施例において開示されるように、本明細書において開示されるスコアカードは、関心対象の1つまたは複数の多能性幹細胞株が、ES細胞の参照細胞株から外れているか否かを要約する。本明細書において用いられるように、ES細胞の参照株は、関心対象の任意の数のES細胞でありうる。代わりの態様において、ES細胞の参照株は、多数のiPSCおよび/またはES細胞、たとえば本明細書において実施例において用いられる少なくとも約10個または少なくとも約20個の低継代ES細胞株に関するDNAメチル化および遺伝子発現の正常範囲を構成することができる。

40

【0406】

偏差スコアカードを算出するためのアルゴリズム(図11Aに概要)は、DNAメチル化および遺伝子発現データに関するアルゴリズムと同じであり、唯一の例外は、マイクロアレイデータがさらなる正規化段階を必要とする点である。

【0407】

いくつかの態様において、遺伝子発現またはDNAメチル化スコアカードを決定するため

50

のアルゴリズムは、以下の段階を含む：

【0408】

(i) データインポート：関心対象の多能性幹細胞、および参照の高品質多能性幹細胞対照株として用いられる少なくとも1個または少なくとも約10個またはそれより多くの参照多能性幹細胞株由来の遺伝子発現および/またはDNAメチル化データをインポートする。いくつかの態様において、遺伝子発現データはマイクロアレイデータであり、いくつかの態様において、DNAメチル化データは全ゲノムDNAメチル化またはRRBS（簡約表示バイサルファイトシーケンシング）である。

【0409】

(ii) 任意のデータ正規化段階（遺伝子発現のみにとって必要）：マイクロアレイデータのgcRNAによる正規化などの、遺伝子発現データの正規化を行い、全ての遺伝子発現値を、0から10までの標的区間範囲の縮尺にする。いくつかの態様において、標的間隔参照範囲は、0から100または0から1000、または0から約500、または任意の好ましい標的区間範囲へと正規化される。

【0410】

(iii) 遺伝子マッピング：各遺伝子に関してDNAメチル化レベル（プロモーター領域における全てのCpGについての平均）および遺伝子発現レベル（別の転写物についての平均）を決定するために遺伝子マッピングを行う。いくつかの態様において、Ensembl遺伝子注釈付けは、各遺伝子に関してDNAメチル化レベルと遺伝子発現レベルをマッチさせるために有用である。いくつかの態様において、重み付けスキームは、試料間の異なるシーケンシングカバー率を補正する。別の言い方をすれば、「参照範囲帯」または「参照DNAメチル化レベル」または「参照遺伝子発現レベル」は、参照の高品質ES細胞における任意の遺伝子に関するDNAメチル化および遺伝子発現転写物レベルの予想されるレベルまたは範囲の値の範囲を提供する。

【0411】

(iv) 参照との比較：各遺伝子の標準化DNAメチル化値および正規化遺伝子発現値を、参照多能性幹細胞株に関する標準化DNAメチル化値および正規化遺伝子発現値と比較する。DNAメチル化または遺伝子発現に関するその値が四分位数間範囲（たとえば、テューキーの外れ値フィルターを用いて）の約1.2倍を超えてまたは1.5倍を超えて中心四分位数の外側となる場合、多能性幹細胞株を「外れ値」細胞株であると同定する。別の言い方をすれば、DNAメチル化レベルまたは遺伝子発現レベルが、「参照範囲帯」の外側または「参照DNAメチル化範囲」もしくは「参照遺伝子発現範囲」から外れる場合（例示的な例として図1Cを参照されたい）、多能性幹細胞株は「外れ値」幹細胞株であると見なされる。

【0412】

(v) 関連性フィルター：関連性フィルターを適用して、約15%または約20パーセントポイント（20%）を超えるDNAメチル化の差または少なくとも約1.5倍または少なくとも約2倍の発現の変化を有する多能性幹細胞を「外れ値」幹細胞株として同定して、多能性幹細胞の外れ値幹細胞株を、使用または今後の分析から無視する。

【0413】

(vi) 遺伝子セット：表12A、12B、12C、13A、13B、および14に記載される遺伝子、ならびに系列マーカー遺伝子（たとえば表7、表13A～13B、および表14に記載される遺伝子）、ならびに癌遺伝子（たとえば、表6Aおよび6Bに記載される遺伝子などの）などの関心対象の応用にとって適切な遺伝子を含む遺伝子セットをロードする。

【0414】

(v) レポートの要約：関心対象の各多能性幹細胞株に関する偏差の数のリスト。たとえば、レポートは、正常からの偏差の%、正常からの偏差の絶対数、および任意で影響を受けた遺伝子の名称を提供することができる（たとえば、4B、および表6A、6B、9Aを参照されたい）。

【0415】

いくつかの態様において、偏差スコアカードは、テューキーの外れ値フィルターを用い

10

20

30

40

50

るノンパラメトリック外れ値検出に基づく (Tukey, 1977)。関心対象の細胞株のDNAメチル化または遺伝子発現値が、四分位数間範囲の1.5倍を超えて中心四分位数の外側となる全ての遺伝子は、外れ値の疑いがあると見なされ、そのように目印をつけられる。

【0416】

次に、変化の規模を考慮し、ES細胞参照からの偏差が生物学的に意味があると十分見なされるほどの大きさである遺伝子のみを最終的に外れ値として報告する。現在の試験に関して、本発明者らは、DNAメチル化に関して少なくとも20パーセントポイントおよび遺伝子発現に関して少なくとも2倍という閾値を用いたが、これらの閾値はこれまでの研究 (Bock et al., 2010) と一貫しており、図10Cにおいてさらに正当化された。どの遺伝子が影響を受けるかに応じて偏差が多かれ少なかれ関係するという事実を説明するために、いくつかの態様において、遺伝子の多数のリスト、たとえばDNAメチル化欠損に関して特に詳しくモニターする必要がある遺伝子、すなわち系列マーカー遺伝子および癌遺伝子の2つまたはそれより多くのリストを作成することができる。これらの遺伝子の偏差を、偏差スコアカードの拡大版において具体的に強調する (表12A、表12B、および表12C)。最後に、いくつかの態様において、本発明者らは、外れ値多能性幹細胞株を同定するまたは目印をつけるために、たとえば調整t-検定に基づくパラメトリックアプローチを含む代替りの戦略を用いることができる。いくつかの態様において、テューキーの外れ値フィルターを、外れ値多能性幹細胞株を同定するために用いることができ、これは、「参照範囲帯」箱ひげ図によって直観的に可視化することができるというさらなる利点を有する (図10および4Aを参照されたい)。

10

20

【0417】

系列スコアカード計算

本明細書において開示される系列スコアカードは、1つまたは複数の参照多能性幹細胞株、たとえば本明細書の実施例において用いられる19個の低継代ES細胞株に関する参照値などの、高品質および/または低継代多能性幹細胞株と比較した関心対象の細胞株の分化指向性を定量する。系列スコアカードを算出するためのアルゴリズム (図11Bに概要) は、調整t-検定 (Smyth, 2004) とt-スコアについて行われる遺伝子セット濃縮分析 (Nam and Kim, 2008; Subramanian et al., 2005) との組み合わせを用いる。

【0418】

系列特異的分化指向性を定量に関する生物学的な根拠を提供するために、本発明者らは、3つの胚葉 (外胚葉、中胚葉、内胚葉) ならびに神経および造血系列の各々に関するいくつかのマーカー遺伝子セットを作製した (図7および13Aを参照されたい)。次に、Bioc onductorのlimmaパッケージを用いて、関心対象の細胞株について得られたEBにおける遺伝子発現をES細胞参照について得られたEBと比較する調整t-検定を行い、関連する遺伝子セットに寄与する全ての遺伝子において平均t-スコアを算出した。高い平均t-スコアは、試験されるEBにおける遺伝子セットの遺伝子の発現増加を示しており、対応する系列に関する高い分化指向性を示すと見なされる。対照的に、低い平均t-スコアは、関連遺伝子の発現減少を示し、対応する系列に関する低い分化指向性を示すと見なされる。分析の頑健性を増加させるために、平均t-スコアを、所定の系列に割付される全ての遺伝子セットについて平均した。系列スコアカードダイアグラム (図5Bおよび5D) は、これらの「遺伝子セット平均t-スコアの平均値」を、細胞株特異的分化指向性の定量的指標として記載する。系列スコアカード分析および検証は、カスタムRスクリプト (<http://www.r-project.org/>) を用いて行った。

30

40

【0419】

本明細書の実施例の章で示されるように、特異的な細胞分化の効率を、特定の細胞系列への多能性幹細胞株の分化能を予測するための信頼可能な頑健な試験として用いることができる。たとえば、本明細書において実施例において示されるように、Boulting et al. によって実験的に誘導された運動ニューロン分化効率は、系列スコアカードの予測力を決定するための真の試験セットを提供した。系列スコアカードのバイオインフォマティクスアルゴリズムは、2つのデータセット間の第1の比較を行う前に既に確定されており、スコ

50

アカードの局面は、適合を改善するために遡及的に最適化されなかった。

【0420】

系列スコアカードを算出するためのアルゴリズム（図11Bに概要）は、以下の段階を含む。

【0421】

(i) データのインポート：(i) 関心対象の多能性幹細胞の胚様体 (EB)、および (ii) 参照多能性幹細胞株（たとえば、高品質の参照多能性幹細胞対照細胞株として用いられる多能性幹細胞株）からの少なくとも1個、または少なくとも約5個、または少なくとも約10個またはそれより多くの胚様体 (EB) からの、少なくとも200、または少なくとも約300、または少なくとも約400、または少なくとも約500、またはそれより多くのマーカー遺伝子の遺伝子発現および/またはDNAメチル化データをインポートする。いくつかの態様において、遺伝子発現データはマイクロアレイデータであり、いくつかの態様において、DNAメチル化データは、全ゲノムDNAのメチル化またはRRBS（簡約表示バイサルファイトシーケンシング）である。

10

【0422】

(ii) 任意のアッセイ標準化段階：陽性spike-inコントロールを用いてアッセイ標準化因子を計算してそれに従ってデータのサイズを変更する。いくつかの態様において、spike-inによる標準化は、各実験またはレプリケート実験において必要である。

【0423】

(iii) 試料の標準化：全ての実験にわたって変動の安定化および標準化を行う。いくつかの態様において、変動の安定化および標準化は、BioconductorのVSNパッケージなどの、当業者が容易に入手可能なソフトウェアによって行うことができる。

20

【0424】

(iv) 参照値との比較：関心対象の各多能性幹細胞株由来のEBの、各系列マーカー遺伝子（たとえば、表7、13A～13B、および14に記載される）に関する標準化DNAメチル化値および標準化遺伝子発現値を、参照多能性幹細胞株のEBの、同じ系列マーカー遺伝子に関する標準化DNAメチル化値および標準化遺伝子発現値と比較する。いくつかの態様において、比較のために統計分析を用い、たとえば関心対象の多能性幹細胞株のEBレプリケートを、参照の高品質EBについて得られた参照値セットと比較するために、各マーカー遺伝子に関して調整t-検定を用いる。いくつかの態様において、任意の統計学パッケージを用いることができ、たとえばBioconductorのlimmaパッケージまたはその他を用いることができる。

30

【0425】

(v) 遺伝子セット：関心対象の細胞系列または胚葉に関して特徴的である関連遺伝子を含む系列マーカー遺伝子セットをロードする。任意の遺伝子のリストを用いることができ、それらを当業者は、遺伝子オンロジー、MolSigDB、またはマニュアルキュレーションエフォートを用いて容易に編集することができる。そのような遺伝子のリストの例は、本明細書において表7、13A、13B、および表14に開示される。

【0426】

(vi) 濃縮分析：各遺伝子セット（DNAのメチル化および/または遺伝子転写物発現レベルが決定される）に関して、各セットに属する全てのマーカー遺伝子に関する平均t-スコアを算出する。

40

【0427】

(vii) 系列スコアカードレポート：関心対象の各多能性幹細胞株に関して、多能性幹細胞が分化する系列に関するスコアカード推定値を提供するために、全ての関連する遺伝子セットに関するt-スコアの平均値をリストにする（たとえば、図5Aおよび5Bを参照されたい）。

【0428】

バイオインフォマティクス分析およびデータアクセス

方法特異的なデータ標準化およびスコアカードの計算（上記の）のほかに、データセッ

50

トのバイオインフォマティクス分析を以下のように行うことができる。

【0429】

(i) 階層的クラスタリング

本明細書の実施例の章(図1、3、8および9を参照されたい)に開示されるように、DNAメチル化レベル(たとえば、Ensembl注釈付き転写物のプロモーター領域における全てのCpGについてのカバー率重み付け平均値のもの)ならびに遺伝子発現レベル(たとえば、マイクロアレイ上の全ての関連プローブについて平均することにより各Ensembl遺伝子に関する)の階層的クラスタリングを行うことができる。階層的クラスタリングの前に、両方のデータセットに対して等しい重みを与えるために、2つのデータセットの各々を平均ゼロおよび分散1へと個々に個別に正規化することができる。図1、3、8、および9に示されるヒートマップは、250個の遺伝子の代表的な選択物である。

10

【0430】

(ii) 注釈付きクラスタリングおよびプロモーターの特徴(図2D)

たとえばDAVID(Huang et al., 2007)およびEpiGRAPH(Bock et al., 2009)などの一般的に入手可能なソフトウェアパッケージを用いて、デフォルトパラメータを用いて、Ensembl遺伝子注釈(プロモーターは、転写開始部位周囲の-5 kbから+1 kbの配列ウィンドウとして定義される)に基づいて最も変動性の遺伝子における一般的な特徴を同定することができる。

【0431】

(iii) ES対iPS細胞株の分類(図3D)

所与のシグネチャーにおける全ての遺伝子についての平均のDNAメチル化または発現レベルを用いて、ES遺伝子シグネチャーおよびiPS遺伝子シグネチャーを容易に検証することができる。ロジスティック回帰分析を用いて、判別閾値を選択することができ、各シグネチャーの予測性を、leave-one-out交差検証によって評価することができる。新しい分類子を誘導するために、サポートベクターマシンをDNAメチル化データ、遺伝子発現データ、または両方のデータセットの組み合わせに関して訓練することができる。本明細書の実施例の章に開示されるように、7500個の無作為選択属性に関して、各分類を行うことができ、これは1回の分析で分析するために容易で、計算上実現可能な属性の最大数である。いくつかの態様において、全ての分類子の予測性を、leave-one-out交差検証によって評価し、無作為な属性セット(図3Dに示されるように)による100個の分類に対して性能を平均することができる。いくつかの態様において、教師ありまたは教師なしの特色選択を用いて予測精度を増加させることができる。いくつかの態様において、予測は、容易に入手可能なソフトウェア、たとえばWekaソフトウェア(Frank et al., 2004)を用いて行うことができる。

20

30

【0432】

(iv) エピジェネティックな記憶の線形モデル

同様に、DNAのメチル化および/または遺伝子発現レベルの線形モデルを作成することができる。たとえば、本明細書において開示されるように、DNAのメチル化および遺伝子発現の両方に関して2つの代替線形モデルを構築することができる。1つのモデルは、ES細胞特異的平均DNAメチル化(または遺伝子発現)レベルに対して各遺伝子のiPS細胞特異的平均DNAメチル化(または遺伝子発現)レベルを回帰推定するために用いることができる。第二のモデルは、ES細胞特異的および線維芽細胞特異的平均DNAメチル化(または遺伝子発現)レベルに対して各遺伝子のiPS細胞特異的平均DNAメチル化(または遺伝子発現)レベルを回帰推定する。

40

【0433】

差次的メチル化領域(DMR)の同定

古典的なピーク検出(その全内容が参照により本明細書に組み入れられる、Bock, C. et al., *Bioinformatics* 24, 1 (2008) および (Park, P. J., *Nat. Rev. Genet.* 10, 669 (2009)において考察されている)などの一般的に公知の方法を用いて、差次的メチル化ゲノム領域、たとえば差次的メチル化遺伝子を同定することができる。しかし、古典的な

50

ピーク検出は、境界線のピークが1つの試料に検出されるが他の試料には検出されない場合に生じる偽性ヒット数が多いために (C. Bock、未発表知見)、差次的メチル化領域 (DMR) の同定にとってあまり適していない可能性がある。

【0434】

その代わりに、いくつかの態様において、2つの試料を互いに直接比較するために統計学的検定を用いて差次的メチル化領域を同定することができる。RRBSデータを有する所とのゲノム領域に関して、両方の試料におけるメチル化 対 非メチル化CpGの数を計数し、フィッシャーの正確確率検定を行って、DMRである領域の可能性を示すp値を得ることができる。同様に、MeDIPおよびMethylCapに関して、両方の試料に関して領域内部で整列するリード数を計数することができ、フィッシャーの正確確率検定を行って、これらの値をゲノムの他所で整列するリード総数と対比させることができる。たとえば、Infiniumアッセイを用いてメチル化を測定する場合、領域内部で全てのInfiniumプローブに関する2つの試料の値を比較するために、対応試料t-検定 (paired-samples t-test) を用いることができる。これらの検定は、同時に多数のゲノム領域について行われ (たとえば、全てのCpGアイランドについて)、q-値法 (Storey, et al., PNAS 100, 9440 (2003)) を用いてp-値を多数の検定に関して補正する。0.1未満のq-値を有するゲノム領域は、高度メチル化または低メチル化であるとしてフラグが立てられるが (差の方向性に依じて)、これは、DNAメチル化の絶対差が20%を超える場合 (RRBSおよびInfiniumに関して) またはリード数に少なくとも2倍の差がある場合 (MeDIPおよびMethylCapに関して) に限られた。これらの閾値は、異なる細胞型間の多数の比較におけるその実際的な有用性によって本発明者らによって選択され、これ以上の正当な理由はない。いくつかの態様において、シーケンシングのカバー率が不十分であるゲノム領域も同様にマークすることができるが、差次的メチル化領域 (DMR) 分析から除外しない。いくつかの態様において、MeDIPおよびMethylCapアッセイを用いてメチル化を測定する場合、より高い解読カバー率を有する試料に関して総解読数1000万回当たり少なくとも10回の解読を有することが推奨され、RRBSを用いてメチル化を測定する場合には、両方の試料におけるそれぞれ少なくとも5回の解読について少なくとも5個のCpGを有することが推奨される。

【0435】

いくつかの態様において、差次的メチル化領域 (DMR) 同定に対するこの統計学的アプローチでは、分析が行われるゲノム領域セットまたは一連のゲノム領域セットを定義する必要があり、たとえば、表12Aおよび/または12Cに記載される遺伝子セットまたは一連の遺伝子セットを選択することができる。いくつかの態様において、多能性幹細胞における興味深いDMRを発見する機会を最大限にするために二方向の戦略を追求することができる。いくつかの態様において、ゲノム領域セットまたは一連のゲノム領域セットを選択した後、エピジェネティックな調節に関する最も重要な候補である、CpGアイランドおよびゲノムプロモーターに対して分析を特異的に集中させることができる。このアプローチは、CpGアイランドおよび遺伝子プロモーターの数が比較的少ないことにより、ゲノム全域にわたる場合と比較して多数の試験補正を行う負担を軽減することから、周知の機能的役割を有する領域に関して増加した統計力を提供するため、有用である。代替の態様において、任意の候補領域の外側に位置するDMRを検出するためにゲノムの1キロベース (または他の既定のゲノムサイズ) のタイリングを用いることができる。いくつかの態様において、およびより広い網を打つために、CpGアイランドおよび遺伝子プロモーターのみならず、CpGアイランドショア (Irizarry, R. A. et al., Nat. Genet. 41, 178 (2009))、エンハンサー (Heintzman, N. D. et al., Nature 459, 108 (2009))、進化的保存領域および他のタイプのゲノム領域を含む包括的な13タイプのゲノム領域セットを収集することができる。いくつかの態様において、これらの領域セットの全てに関する差次的メチル化領域 (DMR) データを、PythonおよびRスクリプトのセットを用いて算出することができ、これらはオンラインで入手可能である (ワールドワイドウェブ: <http://meth-benchmark.computational-epigenetics.org/>)。

【0436】

エピジェネティックな修飾、たとえば異なるレベルのDNAメチル化を決定するための候補遺伝子座は、全てのゲノム領域、またはプロモーター、エンハンサー、インスレーター配列、CpGアイランド、CpGアイランドショア等などの特定タイプのゲノム領域を含むことができる。いくつかの態様において、同様にDNAメチル化データを用いて高度に変動性である領域を直接導き出すことができ、DNA配列データを用いてエピジェネティックな変化に対して感受性があるゲノム領域を予測することができる。さらに、いくつかの態様において、癌、正常および異常な発達ならびに疾患に関係する遺伝子およびゲノム領域に関するこれまでの知識を候補として用いることができる。

【0437】

さらに、当業者は、公共の利用可能な情報および全ての種類の機能的ゲノミクスデータセットに基づいて遺伝子およびゲノム領域の優先順位をつけるために、テキストマイニング、情報の検索、統計的学習および順位法のいずれか1つ、またはそれらの組み合わせを用いることができる。本発明者らはこれらの方法を用いて、遺伝子セット、ネットワーク、および経路を定義した。

10

【0438】

いくつかの態様において、DNAメチル化の代替として、またはそれに加えて、ヒストン修飾などの、しかしこれらに限定されるわけではない他のエピジェネティックな修飾を評価することができる。DNAメチル化および他のエピジェネティックな修飾は高度に相関しており、そのためDNAメチル化データから得ることができる情報はまた、ヒストンメチル化およびアセチル化等などの他のエピジェネティックな修飾からも得ることができることは直ちに明白である。

20

【0439】

遺伝子発現分析はまた、DNAメチル化分析方法より広く用いられる多数の方法によって行うことができる。典型的な例には、遺伝子発現マイクロアレイ、cDNAおよびRNAシーケンシング、NanoStringなどのイメージングに基づく方法、ならびにPCRおよびqPCRを用いる広範囲の方法が挙げられるがこれらに限定されるわけではない。これらの方法の正規化は広く記述されている。本明細書において、本発明者らは、Affymetrixマイクロアレイデータを正規化するためにgcRMAアルゴリズムを用いている。

【0440】

いくつかの態様において、NanoStringデータを用いることができ、本発明者らは、本明細書において、このデータに基づいて多数のアルゴリズムを系統的に評価している。これらの結果に基づき、本発明者らは、VSNアルゴリズムがNanoStringデータを標準化するために最も適していることを発見した。

30

【0441】

いくつかの態様において、遺伝子発現、たとえば非コード遺伝子、マイクロRNA遺伝子、ならびに多能性細胞および分化した細胞に正常または異常に存在する他の全てのタイプのRNA転写物の発現を、任意の遺伝子レベルで決定する。

【0442】

遺伝子発現データを標準化した後、細胞株の品質および有用性に関して適切な遺伝子を、試料および/または試料群のあいだの差次的遺伝子発現を検出するための標準的な方法を用いて同定する。例には、t-検定およびその変法、t-検定のノンパラメトリック代替法、およびANOVAが挙げられる。本明細書の実施例において本発明者らは、調整t-統計学を実行するlimmaパッケージを用いた。

40

【0443】

多くの遺伝子の機能が現在では公知であることを考慮すると、増加または減少した癌のリスク、特定の細胞型および系列への分化能の差、薬物に対する耐性、疾患モデリング、薬物スクリーニング、および再生医療に関する全般的有用性などの推定の効果を、差次的発現および/またはDNAメチル化に対して割付することが可能である。

【0444】

先に記述したDNAメチル化および遺伝子発現アッセイは、たいてい1つの遺伝子の効果に

50

集中しているが、いくつかの態様において、系列スコアカードは、細胞株の品質および有用性を予測するために多数の遺伝子に関するデータの組み合わせを用いる。これは、系列スコアカードの作成のために最も重要かつバイオインフォマティクスにおいて複雑な段階である。

【0445】

多数の遺伝子からの情報は現在、平均値および標準偏差の計算によって集計されているが、サポートベクターマシン、線形およびロジスティック回帰、階層モデル、およびBayesianアルゴリズムなどの統計学的学習法の使用は、集計の効果を低減させる可能性がある。多能性幹細胞の品質および有用性の局面を記述する数値またはカテゴリ値を生じるために、遺伝子発現および/またはDNAメチル化に関する候補遺伝子またはゲノム領域の多数の測定を考慮に入れる任意の数学的関数を、本明細書において開示されるスコアカードの予測因子および要素であると見なすことができる。

10

【0446】

重要なことに、これらの数学的関数は、多くの場合において、過去の生物学的知識を考慮に入れるであろう。特に、本発明者らは、これらの予測因子に情報を与えるために、文献から、公共のデータベースから、および機能的ゲノムクスデータからの遺伝子セットの実質的な数を管理している。スコアカードの1つの態様において、各遺伝子およびゲノム領域に対して差次的メチル化/発現スコアを割付するために、多能性細胞またはその分化子孫からのDNAメチル化および/または遺伝子発現データを用いることができ、次に、得られたt-スコアを用いて、3つの胚葉ならびに他の興味深い細胞型、細胞経路、およびネットワークを表す遺伝子セット、ならびに他の機能的もしくはそれ以外により定義された遺伝子セットの、(パラメトリックまたはノンパラメトリックな)遺伝子セット濃縮分析を行うことができる。

20

【0447】

先に記述したバイオインフォマティクス法は、本明細書において記述される実施例において応用されたが、それらはまた、多能性細胞のDNAメチル化、遺伝子発現、ならびに他のエピジェネティックなおよび機能的ゲノムクスデータにも直接適用することができ、同様に、品質および有用性のある局面がより明らかとなるように多能性細胞株を分化誘導することが可能である。これは、広範囲の化学物質、ペプチド、およびタンパク質処置(しばしば組み合わせられる)にわたる、単純な増殖因子除去および物理的操作(本明細書において非定方向性の胚様体分化のために用いられる)から、専用の表面上での平板培養および特定遺伝子の発現誘導に至るまでの、広範囲の摂動を用いて行うことができる。

30

【0448】

たとえば、Harr et al., Nucleic acid research, 2006; 34(2): e8, "Comparison of algorithms for the analysis of Affymetrix microarray data as evaluated by co-expression of genes in known operons", および"Methods in microarray normalization"と題する文献; Edited By Phillip Stafford, Drug Discovery Series/10, CRC Press 発行(その全体が参照により本明細書に組み入れられる)に開示される多様な方法を用いて遺伝子発現データを分析することができる。cgRMAアルゴリズム(GC[GC含有量]頑健性マルチチップ分析(RMA))は、RMAアルゴリズムの分位点標準化およびミディアムポリッシュ要約法(medium polish summarization)の両方を用いる。確率的モードを用いて、アレイ上の各プローブ対に関して観察されたPMおよびMMプローブシグナルを記述する。特に、モデルは、

40

$$PM_{\mu_i} = 0_{n_i} + N_{1n_i} + S_{n_i}$$

$$NM_{n_i} = 0_{n_i} + N_{2n_i}$$

である。

【0449】

0_{n_i} は、光学ノイズを表し、 N_1 および N_2 は、非特異的結合を表し、 S_{n_j} は試料中のRNA発現に比例した量である。さらに、モデルは、 0 が正規分布 $N(\mu_0, \sigma_0^2)$ に従うこと、ならびに $\log_2(N_{1n_i})$ および $\log_2(N_{2n_i})$ が、プローブ対全体にわたって一定の等分散 σ_N^2 および

50

相関係数0.7で二変量正規分布に従うことを仮定している。非特異的結合項に関する分布の平均値は、プローブ配列に依存する。光学ノイズおよび非特異的結合項は独立していると仮定される。

【0450】

gcRNAにプローブ配列に関する情報を含める方法は、位置依存的塩基親和性の合計に基づいて親和性を比較することである。特に、プローブの親和性は以下の式で与えられ：

$$A = \sum_{k=1}^{25} \sum_{b \in (A,C,G,T)} \mu_b(k) 1_{\beta_k=j}$$

式中、 $\mu_b(k)$ は、自由度5のスプライン関数としてモデルされる。実際に、1つのマイクロアレイ（たとえば、U113Aマイクロアレイチップ）に関する $\mu_b(k)$ は、実験において全てのチップに関して観察されたデータを用いるか、またはgcRMAの作製者によって実行される特異的NSB実験からのいくつかのハードコード化された推定値に基づいて推定される。このことは、gcRMAモデルにおける N_1 および N_2 無作為変数が、プローブ親和性のなめらかな関数 h を用いてモデルとされることを意味する。

【0451】

光学ノイズパラメータ μ_0 、 σ_0^2 は、次のように推定される：光学ノイズによる変動は非特異的結合による変動よりかなり小さく、このように有効に一定である。単純にするために、これを0に設定する。平均値は、負の値を避けるために相関係数を用いて、アレイ上の最低のPMまたはMMプローブ強度を用いて推定される。次に、この定数 μ_0 を差し引くことによって、全てのプローブ強度を相関させる。 $h(A_{n_i})$ を推定するために、全てのMMプローブ親和性に対する補正 $\log(\text{MM})$ 強度に関する散布図に対してloess曲線を適合させる。このloessプロットからの負の余りを用いて σ_N^2 を推定する。最後に、gcRMAのバックグラウンド調節技法は、観察されたPM、MMおよびモデルパラメータを考慮して、Sの期待値を算出することである。gcRMAは、RMAのミディアムポリッシュ要約を用いるが、このようにして生成された発現推定値はそれ以外の点では満足であっても、品質評価を行いたい場合には、PLM要約アプローチをその場で用いてはならないことに注意されたい。

【0452】

いくつかの態様において、たとえばPhillip Staffordによる"method for microarray normalization"において詳細に説明されるMAS5.0アルゴリズム（Microarray suite 5.0）、RMAアルゴリズム（頑健なマルチチップ分析）を用いる、遺伝子発現正規化のための他の方法を用いることができる。

【0453】

統計法

統計的クラスタリングのための方法およびそのためのソフトウェアを以下に述べる。たとえば、遺伝子の差次的発現を定量するために用いられる1つのパラメータは、変化倍率であり、これは、2つの別個の実験条件での遺伝子のmRNA発現レベルを比較するための測定である。その算術的定義は、研究者によって異なる。しかし、変化倍率が大きければ、関連遺伝子の差次的発現が適切に分離する可能性はより高くなり、どのカテゴリに患者が入るかをより容易に決定することができるであろう。

【0454】

アップレギュレートされた遺伝子の変化倍率は、たとえば少なくとも1.4、少なくとも1.5、少なくとも1.6、少なくとも1.7、少なくとも1.8、少なくとも1.9、または少なくとも2.0またはそれより多くの \log_2 変化でありうる。ある態様において、発現レベルがPCRを用いて測定される場合、変化倍率は少なくとも2.0である。

【0455】

ダウンレギュレートされた遺伝子の変化倍率は、0.6もしくは0.6未満であってもよく、たとえば0.5もしくは0.5未満、0.4もしくは0.4未満、0.3もしくは0.3未満、0.2もしくは0.2未満、または0.1もしくは0.1未満の \log_2 変化でありうる。したがって、変化倍率0.1は

10

20

30

40

50

、遺伝子の発現が10倍ダウンレギュレートされることを示す。変化倍率2.0は、遺伝子の発現が2倍アップレギュレートされることを示す。

【0456】

たとえば、多能性幹細胞における遺伝子発現標的遺伝子の変化倍率が、 $=2.0$ である場合（その遺伝子の遺伝子発現の正常な変動と比較）、その遺伝子は「外れ値」遺伝子であることを示している。同様に、多能性幹細胞における遺伝子発現標的遺伝子の変化倍率が $=0.5$ である場合（その遺伝子の遺伝子発現の正常な変動と比較）（遺伝子 $=0.5$ ）、その遺伝子は外れ値遺伝子であることを示している。試験多能性幹細胞株において、「外れ値」遺伝子である遺伝子発現遺伝子の数が多ければ、多能性幹細胞株が、望ましくない特徴、たとえば品質を有するおよび/または特定の有用性にとって適していないことを示している。たとえば、試験多能性幹細胞が、少なくとも約50個、または少なくとも約100個、または100個より多くの外れ値遺伝子発現遺伝子を有する場合、多能性幹細胞株は、外れ値多能性幹細胞株であると同定され、標準的な多能性幹細胞株と比較して異なる、おそらく望ましくない特徴を有する、例として不良な品質（たとえば、癌様細胞株へのトランスデューサーへの高い傾向）、および/または特定の系列に沿って分化する効率が低い細胞である可能性がある。

10

【0457】

差次的発現を定量するために用いられるもう1つのパラメータは、「p」値である。P値がより低ければ、遺伝子はより差次的に発現される可能性があり、このことは遺伝子が多能性幹細胞における遺伝子発現の正常な変動と比較して外れ値遺伝子であることを示していると考えられる。P値はたとえば、0.1またはそれ未満、たとえば0.05またはそれ未満など、特に0.01またはそれ未満を含みうる。本明細書において用いられるp値は、補正「P」値、および/または同様に非補正「P」値を含む。

20

【0458】

本発明は、以下の番号付きパラグラフのいずれかにおいて定義されうる：

1. 以下の段階を含む、多能性幹細胞株を選択する方法：

- a. 多能性幹細胞株における標的遺伝子セットのDNAメチル化を測定して、DNAメチル化データと、同じ標的遺伝子の参照DNAメチル化データとの比較を行う段階；
- b. 複数の系列マーカー遺伝子の遺伝子発現および/またはDNAメチル化を測定することによって、多能性幹細胞の非定方向性または定方向性分化による多能性幹細胞株の分化能を測定する段階と；遺伝子発現および/またはDNAメチル化分化を、同じ系列マーカー遺伝子の参照遺伝子発現および/またはDNAメチル化分化と比較する段階；ならびに
- c. 参照DNAメチル化レベルと比較して標的遺伝子のDNAメチル化に統計学的に有意な量で差がなく、かつ参照分化能と比較して中胚葉、外胚葉、および内胚葉系列に沿って分化する指向性に統計学的に有意な量で差がない多能性幹細胞株を選択する段階；または、参照DNAメチル化レベルと比較して標的遺伝子のDNAメチル化に統計学的に有意な量で差があり、かつ参照分化能と比較して中胚葉、外胚葉、および内胚葉系列に沿って分化する指向性に統計学的に有意な量で差がある多能性幹細胞株を廃棄する段階。

30

2. DNAメチル化が、少なくとも1つの多能性幹細胞に、DNAにおけるエピジェネティックな修飾に差次的に結合する剤を接触させることによって測定される、パラグラフ1の方法。

40

3. 少なくとも1つの多能性幹細胞に、メチル化DNAおよび非メチル化DNAに差次的に結合する剤を接触させて、DNAメチル化データと、同じ標的遺伝子の参照DNAメチル化データとの比較を行うことによって、DNAメチル化を測定することができる、パラグラフ2の方法。

4. DNAメチル化を、濃縮に基づく方法（たとえば、MeDIP、MBD-seq、およびMethylCap）、パイサルファイトシーケンシングおよびパイサルファイトに基づく方法（たとえば、RRBS、パイサルファイトシーケンシング、Infinium、GoldenGate、COBRA、MSP、MethylLight）、ならびに制限消化法（たとえば、MRE-seq）、または同じ標的遺伝子の参照DNAメチル化データと比較した多能性幹細胞のDNAメチル化標的遺伝子の差次的変換、差次的

50

制限、差次的重量からなる群より選択されるいずれか1つによって測定することができる、パラグラフ2の方法。

5. 以下の段階をさらに含む、パラグラフ1~4のいずれかの方法：

a. 多能性幹細胞株における第二の標的遺伝子セットの遺伝子発現を測定して、該遺伝子発現データと、同じ標的遺伝子の参照遺伝子発現レベルとの比較を行う段階；

b. 参照遺伝子発現レベルと比較して標的遺伝子の遺伝子発現レベルに統計学的に有意な量で差がない多能性幹細胞株を選択する段階；または、参照遺伝子発現レベルと比較して標的遺伝子の発現レベルに統計学的に有意な量で差がある多能性幹細胞株を廃棄する段階。

6. 参照DNAメチル化レベルが、前記DNAメチル化標的遺伝子のメチル化の正常な変動の範囲である、パラグラフ1~5のいずれかの方法。

7. 参照DNAメチル化レベルが、前記DNAメチル化標的遺伝子のDNAメチル化の平均値 ± 任意で1標準偏差であり、平均値が複数の多能性幹細胞株における前記標的遺伝子のDNAメチル化から算出される、パラグラフ1~6のいずれかの方法。

8. 複数の多能性幹細胞株が少なくとも5個またはそれより多くの多能性幹細胞株である、パラグラフ7の方法。

9. 多能性細胞株および/または参照DNAメチル化がバイサルファイトアッセイによって決定される、パラグラフ1~8のいずれかの方法。

10. 多能性細胞株および/または参照DNAメチル化が、全ゲノムバイサルファイトアッセイによって決定される、パラグラフ1~9のいずれかの方法。

11. 多能性細胞株および/または参照DNAメチル化が、簡約表示(reduced representation)バイサルファイトシーケンシング(RBBS)アッセイによって決定される、パラグラフ1~10のいずれかの方法。

12. 参照遺伝子発現レベルが前記標的遺伝子の正常な変動の範囲である、パラグラフ5の方法。

13. 参照遺伝子発現レベルが前記標的遺伝子の発現レベルの平均値であり、平均値が、複数の多能性幹細胞株における該標的遺伝子の発現レベルから算出される、パラグラフ5~12のいずれかの方法。

14. 複数の多能性幹細胞株が、少なくとも5個またはそれより多くの異なる多能性幹細胞株である、パラグラフ13の方法。

15. 多能性細胞株および/または参照遺伝子発現がマイクロアレイアッセイによって決定される、パラグラフ5~14のいずれかの方法。

16. 多能性細胞株の分化能が、定量的分化アッセイによって決定される、パラグラフ1~15のいずれかの方法。

17. 参照分化能が、中胚葉、内胚葉、外胚葉、ニューロン、造血系列、およびこれらの任意の組み合わせからなる群より選択される系列への分化能である、パラグラフ1~16のいずれかの方法。

18. 参照分化能データが、複数の多能性幹細胞株から作成される、パラグラフ1~17のいずれかの方法。

19. 複数の多能性幹細胞株が少なくとも5個の異なる多能性幹細胞株である、パラグラフ18の方法。

20. 多能性細胞株のDNAメチル化標的遺伝子および/または参照DNAメチル化標的遺伝子が、癌遺伝子、腫瘍遺伝子、腫瘍抑制遺伝子、発達遺伝子、系列マーカー遺伝子、およびこれらの任意の組み合わせからなる群より選択される、パラグラフ1~19のいずれかの方法。

21. 多能性細胞株のDNAメチル化標的遺伝子および/または参照DNAメチル化標的遺伝子が、表12Aまたは表13Aまたは表14に記載される群、およびこれらの任意の組み合わせから選択される、パラグラフ1~19のいずれかの方法。

22. 腫瘍遺伝子の遺伝子がc-Sis、上皮細胞増殖因子受容体、血小板由来増殖因子受容体、血管内皮増殖因子受容体、HER2/new、チロシンキナーゼのSrcファミリー、チロシ

10

20

30

40

50

ンキナーゼのSyk-Zap-70ファミリー、チロシンキナーゼのBTKファミリーチロシンキナーゼ、Rafキナーゼ、サイクリン依存的キナーゼ、Rasタンパク質、およびmyc遺伝子から選択される、パラグラフ20の方法。

23. 腫瘍抑制遺伝子が、TP53、PTEN、APC、CD95、ST5、ST7、およびST14遺伝子から選択される、パラグラフ20の方法。

24. 発達遺伝子が、表7または表13Aまたは表14に記載される遺伝子の任意の組み合わせから選択される、パラグラフ20の方法。

25. 系列マーカー遺伝子が、VEGF受容体II (KDR)、アクチン-2平滑筋 (ACTA2)、ネスチン、チューブリン-3、 α -フェトプロテイン (AFP)、シンデカン-4、CD64IFcyRI、Oct-4、 β -HCG、 β -LH、oct-3、Brachyury T、Fgf-5、nodal、GATA-4、flk-1、Nkx-2.5、EKLF、およびMsx3から選択される、パラグラフ20の方法。

26. 多能性細胞株のDNAメチル化標的遺伝子および/または参照DNAメチル化標的遺伝子が、BMP4、CAT、CD14、CXCL5、DAZL、DNMT3B、GATA6、GAPDH、LEFTY2、MEG3、PAX6、S100A6、SOX2、SNAI1、TF、およびこれらの任意の組み合わせからなる群より選択される、パラグラフ1~26のいずれかの方法。

27. 統計学的な差が、参照レベルからの少なくとも1標準偏差、少なくとも2標準偏差、または少なくとも3標準偏差の差である、パラグラフ1~25のいずれかの方法。

28. 多能性細胞株の遺伝子発現標的遺伝子および/または参照遺伝子発現標的遺伝子が、表12Bまたは表13Aまたは表14において記載される群およびこれらの任意の組み合わせから選択される、パラグラフ1~27のいずれかの方法。

29. 表12Aまたは表13Aまたは表14のリストにおける遺伝子の任意の組み合わせから選択される少なくとも約200個の標的遺伝子のDNAメチル化が、多能性細胞株において測定されて、同じ少なくとも200個の標的遺伝子セットの参照DNAメチル化レベルと比較される、パラグラフ1~28のいずれかの方法。

30. 表12Aまたは表13Aまたは表14のリストにおける遺伝子の任意の組み合わせから選択される少なくとも約200個の標的遺伝子のDNAメチル化が、表12Aまたは表13Aまたは表14に記載される1~500番の遺伝子の任意の組み合わせから選択される、パラグラフ1~29のいずれかの方法。

31. 少なくとも約200個の標的遺伝子のDNAメチル化が、表12Aまたは表13Aまたは表14に記載される1~200番から選択される、パラグラフ1~30のいずれかの方法。

32. 表12Aまたは表13Aまたは表14のリストにおける遺伝子の任意の組み合わせから選択される少なくとも約500個の標的遺伝子のDNAメチル化が、多能性細胞株において測定されて、同じ少なくとも500個の標的遺伝子セットの参照DNAメチル化レベルと比較される、パラグラフ1~31のいずれかの方法。

33. 表12Aまたは表13Aまたは表14のリストにおける遺伝子の任意の組み合わせから選択される少なくとも約500個の標的遺伝子のDNAメチル化が、表12Aまたは表13Aまたは表14に記載される1~1000番の遺伝子の任意の組み合わせから選択される、パラグラフ1~32のいずれかの方法。

34. 少なくとも約500個の標的遺伝子のDNAメチル化が、表12Aまたは表13Aまたは表14に記載される1~500番から選択される、パラグラフ1~33のいずれかの方法。

35. 表12Aまたは表13Aまたは表14のリストにおける遺伝子の任意の組み合わせから選択される少なくとも約1000個の標的遺伝子のDNAメチル化が、多能性細胞株において測定され、同じ少なくとも1000個の標的遺伝子セットの参照DNAメチル化レベルと比較される、パラグラフ1~29のいずれかの方法。

36. 少なくとも約1000個の標的遺伝子のDNAメチル化が、表12Aまたは表13Aまたは表14に記載される1~2000番から選択される、パラグラフ1~35のいずれかの方法。

37. 表12Bまたは表13Aまたは表14のリストにおける遺伝子の任意の組み合わせから選択される少なくとも約200個の標的遺伝子の遺伝子発現が、多能性細胞株において測定され、同じ少なくとも200個の標的遺伝子セットの参照遺伝子発現レベルと比較される、パラグラフ1~36のいずれかの方法。

10

20

30

40

50

38. 少なくとも約200個の標的遺伝子の遺伝子発現が、表12Bまたは表13Aまたは表14に記載される1~500番から選択される、パラグラフ1~37のいずれかの方法。
39. 表12Bまたは表13Aまたは表14のリストにおける遺伝子の任意の組み合わせから選択される少なくとも約500個の標的遺伝子の遺伝子発現が、多能性細胞株において測定され、同じ少なくとも500個の標的遺伝子セットの参照遺伝子発現レベルと比較される、パラグラフ1~38のいずれかの方法。
40. 少なくとも約500個の標的遺伝子の遺伝子発現が、表12Bまたは表13Aまたは表14に記載される1~1000番から選択される、パラグラフ1~39のいずれかの方法。
41. 表12Bまたは表13Aまたは表14のリストにおける遺伝子の任意の組み合わせから選択される少なくとも約1000個の標的遺伝子の遺伝子発現が、多能性細胞株において測定され、同じ少なくとも1000個の標的遺伝子セットの参照遺伝子発現レベルと比較される、パラグラフ1~40のいずれかの方法。 10
42. 少なくとも約1000個の標的遺伝子の遺伝子発現が、表12Bまたは表13Aまたは表14に記載される1~2000番から選択される、パラグラフ1~41のいずれかの方法。
43. 参照遺伝子と比較してメチル化に統計学的に有意な差を有する多能性幹細胞株におけるDNAメチル化遺伝子の数が、10、9、8、7、6、5、4、3、2、1、または0個である、パラグラフ1~42のいずれかの方法。
44. 参照遺伝子と比較して遺伝子発現レベルに統計学的に有意な差を有する多能性幹細胞株における遺伝子の数が、10、9、8、7、6、5、4、3、2、1、または0個である、パラグラフ1~43のいずれかの方法。 20
45. 多能性幹細胞が哺乳動物の多能性幹細胞である、パラグラフ1~44のいずれかの方法。
46. 多能性幹細胞がヒト多能性幹細胞である、パラグラフ1~45のいずれかの方法。
47. 生物活性に関して化合物をスクリーニングするための多能性幹細胞の使用であって、該多能性細胞がパラグラフ1~46のいずれかの方法によって選択される、使用。
48. スクリーニングが以下の段階を含む、パラグラフ47の使用：
(iv) 任意で、多能性幹細胞を特定系列に沿って分化させるか、または分化することを可能にする段階；
(v) 該細胞に試験化合物を接触させる段階；および
(vi) 該細胞に及ぼす該化合物の任意の効果を決定する段階。 30
49. 試験化合物が、有機低分子、無機低分子、多糖類、ペプチド、タンパク質、核酸、細菌、植物、真菌、動物細胞、動物組織などの生物材料から作製された抽出物、およびその任意の混合物からなる群より選択される、パラグラフ47~48のいずれかの使用。
50. 試験化合物が、約0.01 nMから約1000 mMの範囲の濃度で試験される、パラグラフ47~49のいずれかの使用。
51. 方法が、ハイスループットスクリーニング法である、パラグラフ47~50のいずれかの使用。
52. 生物活性が、生物アッセイにおける刺激、阻害、調節、毒性、または致死反応の誘発である、パラグラフ47~51のいずれかの使用。
53. 生物活性が、酵素活性の調節、受容体の不活化、受容体の刺激、1つまたは複数の遺伝子の発現レベルの調節、細胞増殖の調節、細胞分裂の調節、細胞形態の調節、およびこれらの任意の組み合わせからなる群より選択される、パラグラフ47~52のいずれかの使用。 40
54. 特定系列が、疾患の遺伝子型または表現型である、パラグラフ47~53のいずれかの使用。
55. 特定系列が、器官、組織、またはその一部の遺伝子型または表現型である、パラグラフ47~54のいずれかの使用。
56. パラグラフ1~46のいずれかの方法によって選択される多能性幹細胞を対象に投与する段階によって対象を処置するための、多能性幹細胞の使用。
57. 対象が哺乳動物である、パラグラフ56の使用。 50

58. 対象がマウスである、パラグラフ56~57のいずれかの使用。
59. 対象がヒトである、パラグラフ56~57のいずれかの使用。
60. 対象が、癌、糖尿病、心不全、筋損傷、セリアック病、神経障害、神経変性障害、リソソーム蓄積症、およびこれらの任意の組み合わせからなる群より選択される疾患または状態を患っているかまたは有すると診断されている、パラグラフ56~59のいずれかの使用。
61. 投与が局所である、パラグラフ56~60のいずれかの使用。
62. 投与が多能性幹細胞の対象への移植である、パラグラフ56~61のいずれかの使用。
63. 多能性幹細胞またはその分化した子孫を対象に投与する前に、該多能性幹細胞を分化させる段階をさらに含む、パラグラフ56~62のいずれかの使用。 10
64. 多能性幹細胞が、中胚葉、内胚葉、外胚葉、ニューロン、造血系列、およびこれらの任意の組み合わせからなる群より選択される系列に沿って分化する、パラグラフ63の使用。
65. 多能性幹細胞が、インスリン産生細胞（膵細胞、細胞等）、神経細胞、筋細胞、皮膚細胞、心筋細胞、肝細胞、血液細胞、適応免疫細胞、および先天免疫細胞などへと分化する、パラグラフ63~64のいずれかの使用。
66. パラグラフ1~26のいずれかの方法によって選択される多能性幹細胞を含む、キット。
67. 使用説明書をさらに含む、パラグラフ66のキット。 20
68. 多能性幹細胞が、パラグラフ47~55のいずれかの使用にとって有用である、パラグラフ66~67のいずれかのキット。
69. 多能性幹細胞が、パラグラフ56~65のいずれかの使用にとって有用である、パラグラフ66~67のいずれかのキット。
70. 以下の少なくとも2つを含む、多能性細胞の複数の特性の特徴決定を行うためのアッセイ：
- a. DNAメチル化アッセイ；
- b. 遺伝子発現アッセイ；および
- c. 分化アッセイ。
71. DNAメチル化アッセイがパイサルファイトシークエンシングアッセイである、パラグラフ70のアッセイ。 30
72. DNAメチル化アッセイが、全ゲノムパイサルファイトシークエンシングアッセイである、パラグラフ70~71のいずれかのアッセイ。
73. DNAメチル化アッセイが、濃縮に基づく方法（たとえば、MeDIP、MBD-seq、およびMethylCap）、パイサルファイトシークエンシングおよびパイサルファイトに基づく方法（たとえば、RRBS、パイサルファイトシークエンシング、Infinium、GoldenGate、COBRA、MSP、MethyLight）ならびに制限消化法（たとえば、MRE-seq）からなる群より選択される、パラグラフ70~72のいずれかのアッセイ。
74. 遺伝子発現アッセイがマイクロアレイアッセイである、パラグラフ70~73のいずれかのアッセイ。 40
75. 分化アッセイが、定量的分化アッセイである、パラグラフ70~74のいずれかのアッセイ。
76. 分化アッセイが、以下の系列、すなわち中胚葉、内胚葉、外胚葉、ニューロン、または造血系列のうち少なくとも1つへの多能性細胞の分化能を評価する、パラグラフ70~75のいずれかのアッセイ。
77. 以下の系列、すなわち中胚葉、内胚葉、および外胚葉のうち少なくとも1つへの多能性細胞の分化能が、中胚葉、内胚葉、および外胚葉系列の少なくとも1つのマーカーに対する抗体を用いる免疫染色またはFACソーティングによって決定される、パラグラフ70~76のいずれかのアッセイ。
78. 以下の系列、すなわち中胚葉、内胚葉、および外胚葉のうち少なくとも1つへの 50

多能性細胞の分化能が、EBにおいて少なくとも約7日後に多能性幹細胞を免疫染色することによって決定される、パラグラフ70~77のいずれかのアッセイ。

79. 中胚葉系列に沿った多能性細胞の分化能が、VEGF受容体II (KDR) またはアクチン-2平滑筋 (ACTA2) に関する免疫染色陽性によって決定される、パラグラフ70~78のいずれかのアッセイ。

80. 外胚葉系列に沿った多能性細胞の分化能が、ネスチンまたはチューブリン 3に関する免疫染色陽性によって決定される、パラグラフ70~79のいずれかのアッセイ。

81. 内胚葉系列に沿った多能性細胞の分化能が、 α -フェトプロテイン (AFP) に関する免疫染色陽性によって決定される、パラグラフ70~80のいずれかのアッセイ。

82. 複数の異なる多能性幹細胞をアッセイするためのハイスループットアッセイである、パラグラフ70~81のいずれかのアッセイ。

83. ハイスループットアッセイが、対象由来の複数の異なる人工多能性幹細胞を評価する、パラグラフ81のアッセイ。

84. 対象が哺乳動物である、パラグラフ83のアッセイ。

85. 対象がヒト対象である、パラグラフ83のアッセイ。

86. DNAメチル化遺伝子が、癌遺伝子、腫瘍遺伝子、腫瘍抑制遺伝子、発達遺伝子、系列マーカー遺伝子、およびこれらの任意の組み合わせからなる群より選択される、パラグラフ70~85のいずれかのアッセイ。

87. DNAメチル化遺伝子が、BMP4、CAT、CD14、CXCL5、DAZL、DNMT3B、GATA6、GAPDH、LEFTY2、MEG3、PAX6、S100A6、SOX2、SNAI1、TF、およびこれらの任意の組み合わせからなる群より選択される、パラグラフ70~86のいずれかのアッセイ。

88. 遺伝子発現アッセイが、表7または表13Aまたは表14に記載される遺伝子の任意の組み合わせから選択される遺伝子の発現を決定する、パラグラフ70~86のいずれかのアッセイ。

89. DNAメチル化アッセイが、表12Aまたは表13Aまたは表14に記載される群から選択される複数の標的遺伝子の任意の組み合わせのDNAメチル化レベルを決定する、パラグラフ70~88のいずれかのアッセイ。

90. DNAメチル化アッセイが、表12Aまたは表13Aまたは表14に記載される少なくとも200個の遺伝子の任意の組み合わせのDNAメチル化レベルを決定する、パラグラフ70~89のいずれかのアッセイ。

91. DNAメチル化アッセイが、表12Aまたは表13Aまたは表14に記載される1~500番の遺伝子のうち少なくとも200個の遺伝子の任意の組み合わせのDNAメチル化レベルを決定する、パラグラフ70~89のいずれかのアッセイ。

92. DNAメチル化アッセイが、表12Aまたは表13Aまたは表14に記載される少なくとも500個の遺伝子の任意の組み合わせのDNAメチル化レベルを決定する、パラグラフ70~91のいずれかのアッセイ。

93. DNAメチル化アッセイが、表12Aに記載される1~1000番の遺伝子のうち少なくとも500個の遺伝子の任意の組み合わせのDNAメチル化レベルを決定する、パラグラフ70~92のいずれかのアッセイ。

94. DNAメチル化アッセイが、表12Aまたは表13Aまたは表14に記載される少なくとも1000個の遺伝子の任意の組み合わせのDNAメチル化レベルを決定する、パラグラフ70~93のいずれかのアッセイ。

95. DNAメチル化アッセイが、表12Aまたは表13Aまたは表14に記載される1~2000番の遺伝子のうち少なくとも1000個の遺伝子の任意の組み合わせのDNAメチル化レベルを決定する、パラグラフ70~92のいずれかのアッセイ。

96. 遺伝子発現アッセイが、表12Bに記載される群から選択される複数の標的遺伝子の任意の組み合わせの遺伝子発現レベルを決定する、パラグラフ70~95のいずれかのアッセイ。

97. 遺伝子発現アッセイが、表12Bまたは表13Aまたは表14に記載される少なくとも200個の遺伝子の任意の組み合わせの遺伝子発現レベルを決定する、パラグラフ70~96のい

10

20

30

40

50

ずれかのアッセイ。

98. 遺伝子発現アッセイが、表12Bまたは表13Aまたは表14に記載される1~500番の遺伝子のうち少なくとも200個の遺伝子の任意の組み合わせの遺伝子発現レベルを決定する、パラグラフ70~97のいずれかのアッセイ。

99. 遺伝子発現アッセイが、表12Bまたは表13Aまたは表14に記載される少なくとも500個の遺伝子の任意の組み合わせの遺伝子発現レベルを決定する、パラグラフ70~96のいずれかのアッセイ。

100. 遺伝子発現アッセイが、表12Bまたは表13Aまたは表14に記載される1~1000番の遺伝子のうち少なくとも500個の遺伝子の任意の組み合わせの遺伝子発現レベルを決定する、パラグラフ70~97のいずれかのアッセイ。

101. 遺伝子発現アッセイが、表12Bまたは表13Aまたは表14に記載される少なくとも1000個の遺伝子の任意の組み合わせの遺伝子発現レベルを決定する、パラグラフ70~96のいずれかのアッセイ。

102. 遺伝子発現アッセイが、表12Bまたは表13Aまたは表14に記載される1~2000番の遺伝子のうち少なくとも1000個の遺伝子の任意の組み合わせの遺伝子発現レベルを決定する、パラグラフ70~97のいずれかのアッセイ。

103. 少なくとも1つまたは複数の多能性幹細胞株からスコアカードを作成するための、パラグラフ70~102のいずれかのアッセイの使用。

104. (i) 複数の多能性幹細胞株における第一の標的遺伝子セットにおけるDNAメチル化を測定する段階；

(ii) 複数の多能性幹細胞株における第二の標的遺伝子セットにおける遺伝子発現を測定する段階；および

(iii) 複数の多能性幹細胞株の分化能を測定する段階

を含む、多能性幹細胞スコアカードを作成する方法。

105. (i) 第一の標的遺伝子セットにおける各標的遺伝子の平均メチル化レベルを算出する段階；

(ii) 第二の標的遺伝子セットにおける各標的遺伝子の平均遺伝子発現レベルを算出する段階

をさらに含む、パラグラフ104の方法。

106. 分化能が、中胚葉、内胚葉、外胚葉、ニューロン、造血系列、およびこれらの任意の組み合わせからなる群より選択される系列への分化能である、パラグラフ104~105のいずれかの方法。

107. 複数の多能性幹細胞株が少なくとも5個の多能性幹細胞株である、パラグラフ104~106のいずれかの方法。

108. DNAメチル化が、パイサルファイトシーケンシングアッセイによって測定される、パラグラフ104~107のいずれかの方法。

109. DNAメチル化が全ゲノムパイサルファイトシーケンシングアッセイによって測定される、パラグラフ104~108のいずれかの方法。

110. DNAメチル化が、濃縮に基づく方法（たとえば、MeDIP、MBD-seq、およびMethyICap）、パイサルファイトシーケンシングおよびパイサルファイトに基づく方法（たとえば、RRBS、パイサルファイトシーケンシング、Infinium、GoldenGate、COBRA、MSP、MethyLight）、ならびに制限消化法（たとえば、MRE-seq）の群から選択される方法のいずれか1つによって測定される、パラグラフ104~109のいずれかの方法。

111. 遺伝子発現がマイクロアレイアッセイによって測定される、パラグラフ104~110のいずれかの方法。

112. 分化能が定量的分化アッセイによって測定される、パラグラフ104~111のいずれかのアッセイ。

113. 以下の系列、すなわち中胚葉、内胚葉、および外胚葉のうち少なくとも1つへの多能性細胞の分化能が、中胚葉、内胚葉、および外胚葉系列に関する少なくとも1つのマーカーに対する抗体を用いる免疫染色またはFACソーティングによって決定される、パ

10

20

30

40

50

ラグラフ104~112のいずれかの方法。

114. 以下の系列、すなわち中胚葉、内胚葉、および外胚葉のうち少なくとも1つへの多能性細胞の分化能が、EBにおいて少なくとも約7日後に多能性幹細胞を免疫染色することによって決定される、パラグラフ104~113のいずれかの方法。

115. 中胚葉系列に沿った多能性細胞の分化能が、VEGF受容体II (KDR) またはアクチン-2平滑筋 (ACTA2) に関する免疫染色陽性によって決定される、パラグラフ104~114のいずれかの方法。

116. 外胚葉系列に沿った多能性細胞の分化能が、ネスチンまたはチューブリン-3に関する免疫染色陽性によって決定される、パラグラフ104~115のいずれかの方法。

117. 内胚葉系列に沿った多能性細胞の分化能が、 α -フェトプロテイン (AFP) に関する免疫染色陽性によって決定される、パラグラフ104~116のいずれかの方法。

118. 第一の遺伝子セットが、癌遺伝子、腫瘍遺伝子、腫瘍抑制遺伝子、発達遺伝子、系列マーカー遺伝子、およびこれらの任意の組み合わせからなる群より選択される、パラグラフ104~117のいずれかの方法。

119. 第一の遺伝子セットが、BMP4、CAT、CD14、CXCL5、DAZL、DNMT3B、GATA6、GAPDH、LEFTY2、MEG3、PAX6、S100A6、SOX2、SNAI1、TF、およびこれらの任意の組み合わせからなる群より選択される少なくとも1つの遺伝子を含む、パラグラフ104~118のいずれかの方法。

120. 第一のDNAメチル化遺伝子セットが、表12Aまたは表13Aまたは表14に記載される群から選択される複数の標的遺伝子の任意の組み合わせを含む、パラグラフ104~119のいずれかの方法。

121. 第一のDNAメチル化遺伝子セットが、表12Aまたは表13Aまたは表14に記載される少なくとも200個の遺伝子の任意の組み合わせを含む、パラグラフ104~120のいずれかの方法。

122. 第一のDNAメチル化遺伝子セットが、表12Aまたは表13Aまたは表14に記載される1~500番の遺伝子のうち少なくとも200個の遺伝子の任意の組み合わせを含む、パラグラフ104~121のいずれかの方法。

123. 第一のDNAメチル化遺伝子セットが、表12Aまたは表13Aまたは表14に記載される少なくとも500個の遺伝子の任意の組み合わせを含む、パラグラフ104~122のいずれかの方法。

124. 第一のDNAメチル化遺伝子セットが、表12Aまたは表13Aまたは表14に記載される1~1000番の遺伝子のうち少なくとも500個の遺伝子の任意の組み合わせを含む、パラグラフ104~123のいずれかの方法。

125. 第一のDNAメチル化遺伝子セットが、表12Aまたは表13Aまたは表14に記載される少なくとも1000個の遺伝子の任意の組み合わせを含む、パラグラフ104~124のいずれかの方法。

126. 第一のDNAメチル化遺伝子セットが、表12Aまたは表13Aまたは表14に記載される1~2000番の遺伝子のうち少なくとも1000個の遺伝子の任意の組み合わせを含む、パラグラフ104~125のいずれかの方法。

127. 第二の遺伝子発現遺伝子セットが、表12Bまたは表13Aまたは表14に記載される群から選択される複数の標的遺伝子の任意の組み合わせを含む、パラグラフ104~126のいずれかの方法。

128. 第二の遺伝子発現遺伝子セットが、表12Bまたは表13Aまたは表14に記載される少なくとも200個の遺伝子の任意の組み合わせを含む、パラグラフ104~127のいずれかの方法。

129. 第二の遺伝子発現遺伝子セットが、表12Bまたは表13Aまたは表14に記載される1~500番の遺伝子のうち少なくとも200個の遺伝子の任意の組み合わせを含む、パラグラフ104~128のいずれかの方法。

130. 第二の遺伝子発現遺伝子セットが、表12Bまたは表13Aまたは表14に記載される少なくとも500個の遺伝子の任意の組み合わせを含む、パラグラフ104~129のいずれかの

10

20

30

40

50

方法。

131. 第二の遺伝子発現遺伝子セットが、表12Bまたは表13Aまたは表14に記載される1~1000番の遺伝子のうち少なくとも500個の遺伝子の任意の組み合わせを含む、パラグラフ104~130のいずれかの方法。

132. 第二の遺伝子発現遺伝子セットが、表12Bに記載される少なくとも1000個の遺伝子の任意の組み合わせを含む、パラグラフ104~131のいずれかの方法。

133. 第二の遺伝子発現遺伝子セットが、表12Bまたは表13Aまたは表14に記載される1~2000番の遺伝子のうち少なくとも1000個の遺伝子の任意の組み合わせを含む、パラグラフ104~132のいずれかの方法。

134. (i) 複数の多能性幹細胞株由来の複数のDNAメチル化標的遺伝子のDNAメチル化レベルを含む第一のデータセット；

(ii) 複数の多能性幹細胞株由来の複数の遺伝子発現標的遺伝子の遺伝子発現レベルを含む第二のデータセット；ならびに

(iii) 複数の多能性幹細胞株由来の外胚葉、中胚葉、および内胚葉系列への分化に関する分化指向性レベルを含む第三のデータセット

を含む、多能性幹細胞の性能パラメータのスコアカード。

135. 複数の参照DNAメチル化遺伝子が、少なくとも約500個、少なくとも約1000個、少なくとも約1500個、または少なくとも約200個の参照DNAメチル化遺伝子である、パラグラフ134のスコアカード。

136. 複数の参照DNAメチル化遺伝子が、表12Aまたは表13Aまたは表14に記載される遺伝子の任意の組み合わせから選択される、パラグラフ134または135のスコアカード。

137. 複数の参照DNAメチル化遺伝子が、表12Aまたは表13Aまたは表14に記載される遺伝子の任意の組み合わせから選択される、パラグラフ134または136のスコアカード。

138. 複数の参照DNAメチル化遺伝子が、表12Aまたは表13Aまたは表14に記載される少なくとも200個の遺伝子の任意の組み合わせから選択される、パラグラフ134~137のいずれかのスコアカード。

139. 複数の参照DNAメチル化遺伝子が、表12Aまたは表13Aまたは表14に記載される1~500番の遺伝子のうち少なくとも200個の遺伝子の任意の組み合わせから選択される、パラグラフ134~138のいずれかのスコアカード。

140. 複数の参照DNAメチル化遺伝子が、表12Aまたは表13Aまたは表14に記載される少なくとも500個の遺伝子の任意の組み合わせから選択される、パラグラフ134~139のいずれかのスコアカード。

141. 複数の参照DNAメチル化遺伝子が、表12Aまたは表13Aまたは表14に記載される1~1000番の遺伝子のうち少なくとも500個の遺伝子の任意の組み合わせから選択される、パラグラフ134~140のいずれかのスコアカード。

142. 複数の参照DNAメチル化遺伝子が、表12Aまたは表13Aまたは表14に記載される少なくとも1000個の遺伝子の任意の組み合わせから選択される、パラグラフ134~141のいずれかのスコアカード。

143. 複数の参照DNAメチル化遺伝子が、表12Aまたは表13Aまたは表14に記載される1~2000番の遺伝子のうち少なくとも1000個の遺伝子の任意の組み合わせから選択される、パラグラフ134~142のいずれかのスコアカード。

144. 複数の参照DNAメチル化遺伝子が全ゲノムのDNAメチル化状態である、パラグラフ134~143のいずれかのスコアカード。

145. 複数の参照DNAメチル化遺伝子が、癌遺伝子、腫瘍遺伝子、腫瘍抑制遺伝子、発達遺伝子、および系列マーカー遺伝子を含む、パラグラフ134~144のいずれかのスコアカード。

146. 複数の参照DNAメチル化遺伝子が、BMP4、CAT、CD14、CXCL5、DAZL、DNMT3B、GATA6、GAPDH、LEFTY2、MEG3、PAX6、S100A6、SOX2、SNAI1、TF、およびこれらの任意の組み合わせからなる群より選択される少なくとも1つの遺伝子を含む、パラグラフ134~145のいずれかのスコアカード。

10

20

30

40

50

147. 少なくとも第一および/または第二のデータセットがデータ保存デバイスに接続される、パラグラフ134~146のいずれかのスコアカード。
148. 少なくとも第一および/または第二のデータセットが、データ保存デバイスに接続され、データ保存デバイスがコンピュータデバイス上に存在するデータベースである、パラグラフ134~147のいずれかのスコアカード。
149. 複数の幹細胞株が、少なくとも5個、少なくとも10個、少なくとも15個、または少なくとも20個の多能性幹細胞株である、パラグラフ134~148のいずれかのスコアカード。
150. 複数の幹細胞株が、HUES64、HUES3、HUES8、HUES53、HUES28、HUES49、HUES9、HUES48、HUES45、HUES1、HUES44、HUES6、H1、HUES62、HUES65、H7、HUES13、HUES63、HUES66、およびこれらの任意の組み合わせからなる群より選択される少なくとも1つの多能性幹細胞株を含む、パラグラフ134~149のいずれかのスコアカード。 10
151. 複数の幹細胞株が、HUES64、HUES3、HUES8、HUES53、HUES28、HUES49、HUES9、HUES48、HUES45、HUES1、HUES44、HUES6、H1、HUES62、HUES65、H7、HUES13、HUES63、HUES66からなる群より独立して選択される少なくとも5個の多能性幹細胞株を含む、パラグラフ134~140のいずれかのスコアカード。
152. 複数の多能性幹細胞株が、少なくとも1つの哺乳動物多能性幹細胞株を含む、パラグラフ134~151のいずれかのスコアカード。
153. 複数の多能性幹細胞株の全ての多能性幹細胞株が、哺乳動物多能性幹細胞株である、パラグラフ134~152のいずれかのスコアカード。 20
154. 複数の多能性幹細胞株が、少なくともヒト多能性幹細胞株を含む、パラグラフ134~153のいずれかのスコアカード。
155. 複数の多能性幹細胞株の全ての多能性幹細胞株が、ヒト多能性幹細胞株である、パラグラフ134~154のいずれかのスコアカード。
156. 多能性幹細胞が哺乳動物多能性幹細胞である、パラグラフ134~155のいずれかのスコアカード。
157. 多能性幹細胞がヒト多能性幹細胞である、パラグラフ134~156のいずれかのスコアカード。
158. 多能性幹細胞が人工多能性幹(iPS)細胞である、パラグラフ134~157のいずれかのスコアカード。 30
159. 多能性幹細胞が胚幹細胞である、パラグラフ134~158のいずれかのスコアカード。
160. 多能性幹細胞が成体幹細胞である、パラグラフ134~159のいずれかのスコアカード。
161. 多能性幹細胞が自己幹細胞である、パラグラフ134~160のいずれかのスコアカード。
162. パラグラフ134~161のいずれかのスコアカードを含む、キット。
163. 使用説明書をさらに含む、パラグラフ162のキット。
164. 人工多能性幹細胞を胚幹細胞株と区別するための、パラグラフ134~161のいずれかのスコアカードの使用。 40
165. (iii) DNAメチル化状態を測定するための試薬; および
(iv) 多能性幹細胞の分化指向性を測定するための試薬
を含む、パラグラフ1~46のいずれかの方法を行うためのキット。
166. 標的遺伝子発現遺伝子の遺伝子発現レベルを測定するための試薬をさらに含む、パラグラフ165のキット。
167. 使用説明書をさらに含む、パラグラフ165~166のいずれかのキット。
168. パラグラフ134~161のいずれかのスコアカードをさらに含む、パラグラフ165~166のいずれかのキット。
169. (c)(i) 多能性幹細胞株におけるDNAメチル化標的遺伝子セットのDNAメチル化データを受信して、該DNAメチル化データと、同じ標的遺伝子の参照DNAメチル化レベル 50

との比較を行う段階；

(ii) 多能性幹細胞株の分化能データを受信して、該分化能データを参照分化能データと比較する段階；

(iii) 参照DNAメチル化パラメータと比較したDNAメチル化データの比較、および参照分化データと比較した分化指向性の比較に基づいて、品質保証スコアカードを作成する段階

を含む少なくとも1つのプログラムを含む、少なくとも1つのメモリ；ならびに

(d) プログラムを作動させるためのプロセッサ

を含む、多能性幹細胞の品質保証スコアカードを作成するためのコンピュータシステム。

170. プログラムが以下の段階をさらに含む、パラグラフ169のシステム：

(i) 多能性幹細胞株における第二の標的遺伝子セットの遺伝子発現データを受信して、該発現データを、同じ第二の標的遺伝子セットの参照遺伝子発現レベルと比較する段階；

(ii) 参照DNAメチル化パラメータと比較したDNAメチル化データの比較、および参照分化データと比較した分化指向性の比較、および参照遺伝子発現レベルと比較した遺伝子発現データの比較に基づいて、品質保証スコアカードを作成する段階。

171. DNAメチル化標的遺伝子が多様なメチル化を有する、パラグラフ169~170のいずれかのシステム。

172. DNAメチル化標的遺伝子が、癌遺伝子、腫瘍遺伝子、腫瘍抑制遺伝子、発達遺伝子、系列マーカー遺伝子、およびこれらの任意の組み合わせから選択される、パラグラフ169~171のいずれかのシステム。

173. DNAメチル化標的遺伝子が、BMP4、CAT、CD14、CXCL5、DAZL、DNMT3B、GATA6、GAPDH、LEFTY2、MEG3、PAX6、S100A6、SOX2、SNAI1、TF、およびこれらの任意の組み合わせからなる群より選択される、パラグラフ169~172のいずれかのシステム。

174. 参照DNAメチル化レベルが、腫瘍遺伝子のエピジェネティックなサイレンシングに関しては高レベルのメチル化、ならびに腫瘍抑制遺伝子および発達遺伝子の活発な転写に関しては低レベルのメチル化である、パラグラフ169~173のいずれかのシステム。

175. DNAメチル化標的遺伝子が表12Aに記載される遺伝子の任意の組み合わせから選択される、パラグラフ167~174のいずれかのシステム。

176. DNAメチル化標的遺伝子が、表12Aに記載される少なくとも200個の遺伝子から選択される、パラグラフ167~175のいずれかのシステム。

177. DNAメチル化標的遺伝子が、表12Aまたは表13Aまたは表14に記載される1~500番の遺伝子のうち少なくとも200個の遺伝子の任意の組み合わせから選択される、パラグラフ167~176のいずれかのシステム。

178. DNAメチル化標的遺伝子が、表12Aに記載される少なくとも500個の遺伝子から選択される、パラグラフ167~177のいずれかのシステム。

179. DNAメチル化標的遺伝子が、表12Aまたは表13Aまたは表14に記載される1~1000番の遺伝子のうち少なくとも500個の遺伝子の任意の組み合わせから選択される、パラグラフ167~178のいずれかのシステム。

180. DNAメチル化標的遺伝子が、表12Aに記載される少なくとも1000個の遺伝子から選択される、パラグラフ167~179のいずれかのシステム。

181. DNAメチル化標的遺伝子が、表12Aまたは表13Aまたは表14に記載される1~3000番の遺伝子のうち少なくとも1000個の遺伝子の任意の組み合わせから選択される、パラグラフ167~180のいずれかのシステム。

182. 多能性幹細胞株の品質に基づいて幹細胞スコアカードレポートを作成するレポート作成モジュールをさらに含む、パラグラフ167~181のいずれかのシステム。

183. メモリがデータベースをさらに含む、パラグラフ167~182のいずれかのシステム。

184. データベースがDNAメチル化遺伝子セットを階層的様式で配置する、パラグラフ167~183のいずれかのシステム。

10

20

30

40

50

185. データベースが、異なる系列への分化指向性を階層的様式で配置する、パラグラフ167~184のいずれかのシステム。
186. データベースが、遺伝子発現レベルデータセットを階層的様式で配置する、パラグラフ167~185のいずれかのシステム。
187. メモリが、ネットワークを介して第一のコンピュータに接続される、パラグラフ167~186のいずれかのシステム。
188. ネットワークが広域ネットワークを含む、パラグラフ187のシステム。
189. スコアカードが、多能性幹細胞の適切な用途または適用の指標を提供する、パラグラフ167~188のいずれかのシステム。
190. 参照DNAメチル化レベルが、前記DNAメチル化標的遺伝子に関するメチル化の正常な変動の範囲である、パラグラフ167~189のいずれかのシステム。 10
191. 参照DNAメチル化レベルが、前記DNAメチル化標的遺伝子のDNAメチル化の平均値であり、平均値が、複数の多能性幹細胞株における前記標的遺伝子のDNAメチル化から算出される、パラグラフ167~190のいずれかのシステム。
192. 多能性細胞株の分化能が、定量的分化アッセイによって決定される、パラグラフ167~191のいずれかのシステム。
193. 参照分化能が、中胚葉、内胚葉、外胚葉、ニューロン、造血系列、およびこれらの任意の組み合わせからなる群より選択される系列への分化能である、パラグラフ167~192のいずれかのシステム。
194. 参照遺伝子発現レベルが、前記遺伝子発現標的遺伝子の遺伝子発現の正常な変動の範囲である、パラグラフ167~193のいずれかのシステム。 20
195. 参照遺伝子発現レベルが、前記標的遺伝子の遺伝子発現レベルの平均値であり、平均値が、複数の多能性幹細胞株における前記標的遺伝子の発現レベルから算出される、パラグラフ111~128のいずれかの方法。
196. 参照DNAメチル化、参照分化能データ、および参照遺伝子発現レベルデータが、複数の多能性幹細胞株から作成される、パラグラフ167~194のいずれかのシステム。
197. 複数の多能性幹細胞株が、少なくとも5個、少なくとも10個、少なくとも15個、または少なくとも20個の多能性幹細胞株である、パラグラフ196のシステム。
198. DNAメチル化標的遺伝子が、少なくとも1つまたは複数の遺伝子発現標的遺伝子を含む、パラグラフ167~197のいずれかのシステム。 30
199. 遺伝子発現標的遺伝子が少なくとも1つまたは複数のDNAメチル化標的遺伝子を含む、パラグラフ167~198のいずれかのシステム。
200. 以下を含む、多能性幹細胞株の品質保証スコアカードの作成に関する命令を含む、コンピュータ読み取り可能な媒体：
- (i) 多能性幹細胞株におけるDNAメチル化標的遺伝子セットのDNAメチル化データを受信して、該DNAメチル化データと、同じ標的遺伝子の参照DNAメチル化レベルとの比較を行うこと；
- (ii) 多能性幹細胞株の分化能データを受信して、該分化能データを参照分化能データと比較すること；ならびに
- (iii) 参照DNAメチル化パラメータと比較したDNAメチル化データの比較、および参照分化データと比較した分化指向性の比較に基づいて、品質保証スコアカードを作成すること。 40
201. 以下に関する命令をさらに含む、パラグラフ200のコンピュータ読み取り可能な媒体：
- a. 多能性幹細胞株における第二の標的遺伝子セットの遺伝子発現データを受信して、該発現データを同じ第二の標的遺伝子セットの参照遺伝子発現レベルと比較すること；ならびに
- b. 参照DNAメチル化パラメータと比較したDNAメチル化データの比較、および参照分化データと比較した分化指向性の比較、および参照遺伝子発現レベルと比較した遺伝子発現データの比較に基づいて、品質保証スコアカードを作成すること。 50

202. 以下の少なくとも2つを含む、多能性幹細胞株の品質を決定するためのキット
:

- a. 複数のDNAメチル化遺伝子のメチル化状態を測定するための試薬;
- b. 複数の遺伝子の遺伝子発現レベルを測定するための試薬; ならびに
- c. 多能性幹細胞の外胚葉、中胚葉、および内胚葉系列への分化指向性を測定するための試薬。

203. 使用説明書をさらに含む、パラグラフ202のキット。

204. 少なくとも1つの多能性幹細胞株をさらに含む、パラグラフ202~203のいずれかのキット。

205. パラグラフ134~161のいずれかのスコアカードをさらに含む、パラグラフ202~204のいずれかのキット。

206. a. (i) DNAメチル化標的遺伝子セットのDNAメチル化データを得ること、および関心対象の少なくとも1つの多能性幹細胞株における遺伝子発現遺伝子セットの遺伝子発現データを得ること、ならびに

(ii) DNAメチル化標的遺伝子セットのDNAメチル化データを得ること、および少なくとも1つの参照多能性幹細胞株における遺伝子発現遺伝子セットの遺伝子発現データを得ること、

(iii) 要素(i)および(ii)で得られた遺伝子発現データのデータ正規化を行うこと

(iv) 要素(i)および(ii)で得られたDNAメチル化データおよび遺伝子発現データの遺伝子マッピングを行うこと、

(v) 要素(i)および(iii)で得られた関心対象の多能性幹細胞株由来のDNAメチル化データおよび正規化遺伝子発現データを、要素(ii)および(iii)で得られた参照多能性幹細胞株由来の標準化DNAメチル化データおよび正規化遺伝子発現データと比較して、参照多能性幹細胞株のDNAメチル化レベルまたは遺伝子発現レベルの正常な範囲から統計学的に有意な量で外れているDNAメチル化レベルまたは正規化遺伝子発現レベルを有する多能性幹細胞株における遺伝子を同定すること、

(vi) 要素(v)で同定された遺伝子の関連性フィルターを適用して、参照多能性幹細胞株の参照DNAメチル化レベルまたは遺伝子発現レベルと比較して、15%を超えるDNAメチル化の差または1.5倍を超える遺伝子発現の変化を有する遺伝子を同定すること、

(vii) DNAメチル化標的遺伝子および遺伝子発現標的遺伝子および系列マーカーの遺伝子セットを得ること

のうち1つまたは複数を実施するのに適合させた1つまたは複数のプログラムを実行するための関連メモリおよびプロセッサを、コンピュータに設ける段階; ならびに

b. 少なくとも1つの参照多能性幹細胞株と比較した関心対象の多能性幹細胞株におけるDNAメチル化および/または遺伝子発現の偏差を有する、要素(vi)で同定された遺伝子の数および/または遺伝子の数の百分率を含む多能性スコアカードレポートを作成する段階

を含む、関心対象の幹細胞株の多能性を同定するためのスコアカードを作成する方法。

207. 段階(v)で同定された遺伝子が、参照多能性幹細胞株の正常なDNAメチル化範囲または遺伝子発現範囲の四分位数間範囲の少なくとも1.2倍で中心四分位数の外側となるDNAメチル化レベルまたは正規化遺伝子発現レベルを有する、パラグラフ206の方法。

208. 段階(vi)で同定された遺伝子が、参照多能性幹細胞株の参照DNAメチル化レベルまたは遺伝子発現レベルと比較して、20%を超えるDNAメチル化の差または2倍を超える遺伝子発現の変化を有する、パラグラフ206の方法。

209. レポートスコアカードが、少なくとも1つの参照多能性幹細胞株と比較して関心対象の多能性幹細胞株におけるDNAメチル化および/または遺伝子発現から外れている、影響を受けた遺伝子の名称をさらに含む、パラグラフ206の方法。

210. DNAメチル化データが、全ゲノムDNAメチル化、または簡約表示バイサルファイトシーケンシング(RRBS)によって得られる、パラグラフ206の方法。

10

20

30

40

50

2 1 1 . 遺伝子発現データが、マイクロアレイデータまたは定量的PCR (qPCR) によって得られる、パラグラフ206の方法。

2 1 2 . DNAメチル化標的遺伝子の遺伝子セットにおいて、遺伝子発現標的遺伝子および系列マーカーが、表7、表12A、表12B、表12C、表13A、表13B、または表14から選択される群から選択される表に記載される、パラグラフ206の方法。

2 1 3 . コンピュータ上で行われる、パラグラフ206~212のいずれかの方法。

2 1 4 . コンピュータシステムである、パラグラフ206~213のいずれかの方法。

2 1 5 . 1つまたは複数のプログラムが、コンピュータ読み取り可能な媒体上のスコアカードソフトウェアプログラムによって行われる、パラグラフ206~214のいずれかの方法。

2 1 6 . a . (i) 関心対象の少なくとも1つの多能性幹細胞株の胚様体 (EB) における標的系列マーカー遺伝子セットのDNAメチル化データおよび遺伝子発現データを得ること、ならびに

(ii) 少なくとも1つの参照多能性幹細胞株の胚様体 (EB) における標的系列マーカー遺伝子セットのDNAメチル化データおよび遺伝子発現データを得ること、

(iii) 任意で、要素 (i) および (ii) で得られたDNAメチル化データおよび遺伝子発現データを陽性対照を用いてサイズ変更することによって、アッセイの標準化を行うこと、

(iv) 任意で、レプリケート実験全体にわたって要素 (i) および (ii) で得られたDNAメチル化データおよび遺伝子発現データの試料の標準化および変動の安定化を行うこと、

(v) 要素 (i) で得られた関心対象の多能性幹細胞株由来の系列マーカー遺伝子のDNAメチル化データおよび遺伝子発現データを、要素 (ii) で得られた参照多能性幹細胞株由来の系列マーカー遺伝子のDNAメチル化データおよび遺伝子発現データと比較し、参照多能性幹細胞株のDNAメチル化レベルまたは遺伝子発現レベルの正常な範囲と比較して統計学的に有意な量で増加または減少するDNAメチル化レベルまたは標準化遺伝子発現レベルを有する多能性幹細胞株における系列遺伝子を同定し、それによって各個々の系列マーカー遺伝子に関する分散値を作成すること、

(vi) 関心対象の特徴的な細胞系列または胚葉に関する系列マーカー遺伝子の遺伝子セットを得ること、

(vii) 要素 (vi) で得られた系列マーカー遺伝子セットに記載される各系列マーカーに関する個々の変動値 (要素 (v) で得られた) から変動の平均値を算出することによって、濃縮分析を行うこと

のうち1つまたは複数を実施するのに適合させた1つまたは複数のプログラムを実行するための関連メモリおよびプロセッサを、コンピュータに設ける段階；ならびに

b . 少なくとも1つの参照多能性幹細胞株と比較して多能性幹細胞株の系列マーカー遺伝子セットにおける全ての遺伝子に関する変動の平均値を含む系列スコアカードレポートを作成する段階

をシステムが含む、関心対象の多能性幹細胞株の分化指向性を同定するための系列スコアカードを作成するための方法。

2 1 7 . 多能性幹細胞株が、パラグラフ206のスコアカードを特徴としている、パラグラフ216の方法。

2 1 8 . DNAメチル化データおよび遺伝子発現データに関する標的系列遺伝子マーカーセットが、表7、表13A、表13B、または表14から選択される群から選択される表に記載される、パラグラフ216~217のいずれかの方法。

2 1 9 . 要素 (v) における参照との比較が、参照多能性幹細胞株のDNAメチル化または遺伝子発現と比較してDNAメチル化または遺伝子発現の統計学的に有意な増加または減少を有する系列マーカー遺伝子を同定するために、調整 (moderated) t-検定を用いる、パラグラフ216~218のいずれかの方法。

2 2 0 . 調整t-検定を用いる参照との比較が、Bioconductors Limmaパッケージを用いて行われる、パラグラフ216~219のいずれかの方法。

10

20

30

40

50

221. 系列マーカー遺伝子セットを、遺伝子オントロジー、MoSigDBプログラムまたはキュレーションによって得ることができる、パラグラフ216~220のいずれかの方法。

222. 要素(vii)の濃縮分析が、各系列マーカーに関する個々のt-スコアからt-スコア平均値を算出する、パラグラフ216~221のいずれかの方法。

223. 要素(iv)の試料標準化が、Bioconductor VSNパッケージによって行われる、パラグラフ216の方法。

224. 要素(vi)における系列マーカー遺伝子セットが、外胚葉、中胚葉、内胚葉、神経系遺伝子セット、造血系遺伝子セット、多能性細胞シグネチャー遺伝子セット、表皮系遺伝子セット、間葉幹細胞系遺伝子セット、骨系遺伝子セット、軟骨系遺伝子セット、脂肪系遺伝子セット、筋系遺伝子セット、血管系遺伝子セット、心臓系遺伝子セット、リンパ球系遺伝子セット、骨髄細胞系遺伝子セット、肝臓系遺伝子セット、膵臓系遺伝子セット、上皮系遺伝子セット、運動ニューロン系遺伝子セット、単球-マクロファージ系遺伝子セット、iSCi系遺伝子セット、または表7もしくは13Aおよび13Bおよび表14に記載される遺伝子の任意の選択物の群から選択される遺伝子セットである、パラグラフ216~223のいずれかの方法。

225. コンピュータ上で行われる、パラグラフ216~224のいずれかの方法。

226. システムがコンピュータシステムである、パラグラフ216~225のいずれかの方法。

227. 1つまたは複数のプログラムが、コンピュータ読み取り可能な媒体上のスコアカードソフトウェアプログラムによって行われる、パラグラフ216~226のいずれかの方法。

228. a. 関心対象の多能性幹細胞株における、DNAメチル化標的遺伝子のDNAメチル化レベルおよび/または遺伝子発現標的遺伝子の遺伝子発現レベルを測定するための決定モジュール；

b. (i) 決定モジュールによって測定されたDNAメチル化レベルおよび遺伝子発現レベルを保存するための、ならびに1つまたは複数の参照多能性幹細胞株のDNAメチル化標的遺伝子の参照DNAメチル化レベルおよび遺伝子発現標的遺伝子の参照遺伝子発現レベルを保存するための、保存モジュール；

(ii) 決定モジュールによって測定された遺伝子発現レベルを正規化するための正規化モジュール；

(iii) 多能性幹細胞株において測定したDNAメチル化標的遺伝子のDNAメチル化レベルを、1つもしくは複数の参照多能性幹細胞株のDNAメチル化標的遺伝子のDNAメチル化レベルとマッチさせるための、および/または多能性幹細胞株において測定された遺伝子発現標的遺伝子の遺伝子発現レベルを、1つもしくは複数の参照多能性幹細胞株の遺伝子発現標的遺伝子の遺伝子発現レベルとマッチさせるための、遺伝子マッピングモジュール；

(iv) (i) 関心対象の多能性幹細胞株由来のDNAメチル化標的遺伝子のDNAメチル化レベルを、1つもしくは複数の参照多能性幹細胞株由来の同じDNAメチル化標的遺伝子のDNAメチル化レベルと比較するための、および/または(ii) 関心対象の多能性幹細胞株の遺伝子発現標的遺伝子の遺伝子発現レベルを、1つもしくは複数の参照多能性幹細胞株由来の同じ遺伝子発現標的遺伝子の遺伝子発現レベルと比較して、参照多能性幹細胞株のDNAメチル化レベルまたは遺伝子発現レベルの正常範囲から統計学的に有意な量で外れているDNAメチル化レベルまたは正規化遺伝子発現レベルを有する多能性幹細胞株における遺伝子を同定するための、比較モジュール；

(v) 参照多能性幹細胞株の参照DNAメチル化レベルまたは遺伝子発現レベルと比較して、少なくとも15%を超えるDNAメチル化の差または少なくとも1.5倍を超える遺伝子発現の変化を有する比較モジュールによって同定された遺伝子を選択するための関連性フィルターモジュール；

(vi) 関心対象の比較モジュールおよび/または関連性フィルターモジュールによって同定された遺伝子を選択するための遺伝子セットモジュール

のうち1つまたは複数を含む、プロセッサおよび関連メモリを含むコンピュータモジュ

10

20

30

40

50

ール；

c. 少なくとも1つの参照多能性幹細胞株と比較した関心対象の多能性幹細胞株におけるDNAメチル化および/または遺伝子発現の偏差を有する、比較モジュールおよび/または関連性フィルターモジュールおよび/または遺伝子セットモジュールによって同定された遺伝子の数および/または遺伝子の数の百分率を含むスコアカードレポートを表示するための、ディスプレイモジュール

のうち少なくとも1つまたは複数を含む、関心対象の幹細胞株の多能性を同定するためのスコアカードを作成するシステム。

229. 決定モジュールが、1つまたは複数の参照多能性幹細胞株におけるDNAメチル化標的遺伝子のDNAメチル化レベルおよび/または遺伝子発現遺伝子もしくは系列マーカー遺伝子の遺伝子発現レベルを測定することができる、パラグラフ228のシステム。

10

230. 保存モジュールが、1つまたは複数の参照多能性幹細胞株におけるDNAメチル化標的遺伝子のDNAメチル化レベルおよび/または遺伝子発現遺伝子もしくは系列マーカー遺伝子の遺伝子発現レベルの測定値を保存することができる、パラグラフ228のシステム

231. 1つまたは複数のモジュールを組み合わせて1つのモジュールにすることができる、パラグラフ228のシステム。

232. a. 関心対象の多能性幹細胞株の胚様体(EB)における複数の系列マーカー遺伝子の系列遺伝子発現レベルを測定するための決定モジュール；

20

b. (i) 決定モジュールによって測定された系列遺伝子発現レベルを保存するための、および1つまたは複数の参照多能性幹細胞株の胚様体(EB)における系列マーカー遺伝子の参照系列遺伝子発現レベルを保存するための、保存モジュール；

(ii) 遺伝子発現陽性対照に基づいて遺伝子発現レベルを標準化するためのアッセイ標準化モジュール；

(iii) 関心対象の同じ多能性幹細胞株由来の胚様体(EB)における同じ系列マーカー遺伝子のレプリケート遺伝子発現レベル測定値全体にわたり系列マーカー遺伝子の遺伝子発現レベルを標準化および変動安定化するための試料標準化モジュール；

(iv) 関心対象の多能性幹細胞株由来の胚様体(EB)由来の系列マーカー遺伝子の遺伝子発現レベルを、1つまたは複数の参照多能性幹細胞株由来の胚様体(EB)からの同じ系列マーカー遺伝子の遺伝子発現レベルと比較して、各系列マーカー遺伝子に関する参照多能性幹細胞株の系列遺伝子発現レベルと比較した多能性幹細胞株における系列遺伝子発現レベルの差の統計学的差を算出するための、比較モジュール；

30

(v) 関心対象の特定の細胞系列に特徴的な系列マーカー遺伝子のサブセットを選択するための遺伝子セットモジュール；

(vi) 遺伝子セットモジュールによって選択された系列マーカー遺伝子のサブセットの遺伝子の比較モジュールによって算出された統計学的差の平均値を算出するための濃縮分析モジュール

のうち1つまたは複数を含む、プロセッサおよび関連メモリを含むコンピュータモジュール；

c. 少なくとも1つの参照多能性幹細胞株と比較した多能性幹細胞株の系列マーカー遺伝子セットの各サブセットにおける系列マーカー遺伝子に関する系列遺伝子発現の統計学的差の平均値を含む系列スコアカードレポートを表示するための、ディスプレイモジュールのうち少なくとも1つまたは複数を含む、関心対象の幹細胞株の分化指向性を同定するための系列スコアカードを作成するためのシステム。

40

233. 1つまたは複数のモジュールを組み合わせて1つのモジュールにすることができる、パラグラフ232のシステム。

【実施例】

【0459】

本明細書を通して、様々な刊行物を参照する。刊行物およびそれらの刊行物の中で引用された参考文献の全てに関する開示は、その全内容が、本発明が属する当技術分野の技術

50

の現状をより完全に記述するために、本出願に参照により本明細書に組み入れられる。以下の実施例は、本発明の特許請求の範囲の範囲を限定することは意図されず、むしろ、ある態様の例示であることが意図される。例示される方法において当業者に想起されるいかなる変更も、本発明の範囲内に入ることが意図される。

【0460】

ヒト多能性幹細胞の発達能は、それらが、生物医学研究のための疾患に関連する細胞型をもたらすことを示唆している。しかし、多能性細胞株には、その有用性および臨床での安全性に影響を及ぼしうる実質的な変動があることが報告されている。胚幹 (ES) 細胞または人工多能性幹 (iPS) 細胞を翻訳研究において確信を持って用いることができるためには、そのような細胞株特異的な差をよりよく理解しなければならない。この目標に向かつて、本発明者らは、既に誘導された20個のヒトES細胞株および12個のヒトiPS細胞株に関してDNAメチル化および遺伝子発現に関する全ゲノム参照マップを確立して、これらの細胞株のインビトロ分化指向性を測定した。このリソースにより、本発明者らは、ES細胞およびiPS細胞のエピジェネティックな転写の類似性を評価して個々の細胞株の分化効率を予測することができる。アッセイの組み合わせにより、多能性細胞株を迅速かつ包括的に特徴決定するためのスコアカードが得られる。

10

【0461】

多能性細胞株は、疾患モデリング、薬物スクリーニング、および再生医療のための貴重なツールである。しかし、ヒト多能性細胞株に関する現行の検証アッセイは、面倒で、必ずしも正確ではなく、研究を遅らせる傾向があり、ヒトiPS細胞の効能に関して何らかの混乱を引き起こしている。これらの問題に系統的に取り組むために、本発明者らは、31個の低継代のES細胞株およびiPS細胞株に焦点を当てて、本明細書において多能性のメチロームおよびトランスクリプトームに関する「スコアカード」と呼ばれる参照マップを確立した。さらに、本発明者らはまた、定量的分化アッセイを開発して、これらの細胞株の分化指向性を測定した。このデータセットを用いて、本発明者らは、ES細胞参照からの各ESまたはiPS細胞株の変動を定量して、細胞株の品質および有用性に関するスコアカードを作成した。本発明者らは、(i) CD14陽性細胞への分化を防止するDNAメチル化欠損を検出する、(ii) 運動ニューロンの作製効率の細胞株特異的な差を正確に予測することを示すことによって、このスコアカードを検証した。本発明者らはまた、ヒトES細胞株およびiPS細胞株を、そのDNAメチル化、遺伝子発現、および分化指向性に関して比較したところ、iPS細胞株に関してより高い変動を観察したが、ES細胞株とiPS細胞株とを正確に区別することができる1つの遺伝子座または遺伝子シグネチャーは観察されなかった。要約すると、本発明者らのデータセットは、ゲノムアッセイを用いて、ヒト多能性細胞株のハイスループットな特徴決定のための参照を提供する。

20

30

【0462】

方法

ES細胞株およびiPSC細胞株および培養条件

全体で20個のヒトES細胞株、13個のヒトiPS細胞株、および6個の初代培養線維芽細胞株を本試験に含めた (表1)。ES細胞株は、Human Embryonic Stem Cell Facility of the Harvard Stem Cell Institute (17個のES細胞株)、およびWiCell (3個のES細胞株) から得た。iPS細胞株は、皮膚線維芽細胞において、OCT4、SOX2、およびKLF4のレトロウイルス形質導入によって誘導した。線維芽細胞は、各々のそれぞれのドナーの前腕の皮膚穿刺に由来し、既に記述されたように成長させた (Dimos et al., 2009)。多能性細胞株は全て、従来の方法によって特徴決定されており (Chen et al., 2009; Cowan et al., 2004, Boulting et al., 提出)、確立された標準に従って (Maherali and Hochedlinger, 2008)、それらが多能性として適格であることを確認した。多能性幹細胞を、KO-DMEM (Invitrogen)、10% KOSR (Invitrogen)、10% プラスマネート (Talecris)、1% glutamaxまたはL-グルタミン、非必須アミノ酸、ペニシリン/ストレプトマイシン、0.1% 2-メルカプトエタノール、および10~20 ng/ml bFGFからなるヒトES培地において成長させた。培養物を、放射線照射CF1-MEFs (GlobalStem) の単層上で成長させて、トリプシン (0.05%

40

50

）またはディスペーゼ（Invitrogen）を用いて継代した。分析のためにDNAおよびRNAを収集する前に、ES細胞およびiPS細胞をトリプシン（0.05%）もしくはディスペーゼ処置によって単離するか、またはマトリゲル（BD Biosciences）上で1継代、平板培養して、CF1-MEF中で条件付けしたヒトES培地を24時間与えた。

【0463】

分化プロトコール

全体で5つのES/iPS細胞分化プロトコールを本研究に用いた。

【0464】

(i) 非定方向性EB分化

未分化細胞をディスペーゼまたはトリプシンを用いて採取して、bFGFおよびプラスマネートを含まないヒトES細胞培養培地の存在下で低接着プレートにおいて浮遊培養させた。細胞凝集体（EB）を、48時間毎に培地を交換しながら全体で16日間成長させた。

10

【0465】

(ii) 単球/マクロファージ分化

未分化細胞を、造血細胞分化に関して公表されたプロトコールに従って多数の組み換え型タンパク質によって処置した（Grigoriadis et al., 2010）。簡単に説明すると、フィーダー枯渇多能性細胞を、6ウェル低接着プレート（Corning）中で、ペニシリン/ストレプトマイシン、グルタミン（2 mM）、モノチオグリセロール（0.0004 M）、アスコルビン酸（50 μg/ml）（Sigma-Aldrich）およびBMP4（10 ng/ml）（R & D Systems）を含むStem Pro-34培地（Invitrogen）中で、浮遊させて小さい凝集体として24時間成長させた。原条/中胚葉形成を誘導するために、EBを洗浄して、ヒト組み換え型bFGF（5 ng/ml）（Millipore）を添加したStemPro-34分化培地においてさらに3日間さらに培養した。4日目に、EBを再度採取して、hVEGF（10 ng/ml）（PeproTech）、hbFGF（1 ng/ml）、hIL-6（10 ng/ml）（PeproTech）、hIL-3（40 ng/ml）（PeproTech）、hIL-11（5 ng/ml）（PeproTech）、およびヒト組み換え型SCF（100 ng/ml）（PeproTech）をさらに含む上記の分化培地においてさらに4日間培養して、造血性の特異化を誘導した。8日目以降、造血細胞の成熟および拡大を促進するために、hVEGF（10 ng/ml）、ヒトエリスロポエチン（4 U/ml）（Cell Sciences）、ヒトトロポエチン（50 ng/ml）（Cell Sciences）、ならびにヒト幹細胞因子、hIL-6、hIL-11、およびhIL-3を含むStemPro-34培地においてさらに培養した。

20

【0466】

(iii) 中胚葉分化

未分化細胞を、中胚葉分化を促進する公表されたプロトコール（Laflamme et al., 2007）に従って、アクチビンAおよびBMP4によって処置した。簡単に説明すると、コラゲナーゼIV（Invitrogen）とのインキュベーションにより細胞を採取して、マトリゲルコーティング細胞培養皿上で平板培養した。中胚葉分化を誘導するために、細胞を、ヒト組み換え型アクチビンA（100 ng/ml）（R & D Systems）を添加したRPMI-B27培地（Invitrogen）において24時間培養した。ヒト組み換え型BMP4（10 ng/ml）を培地に4日間添加した後、添加物を含まないRPMI-B27培地を細胞にさらに与えた。

30

【0467】

(iv) 外胚葉分化

未分化細胞をコラゲナーゼIV（Invitrogen）とのインキュベーションによって採取して、マトリゲルコーティング細胞培養皿上で平板培養した。細胞を、Noggin（500 ng/ml）（R&D Systems）およびSB431542（10 μM）（Tocris）を添加したノックアウト血清補充物（Invitrogen）を含むKO-DMEM（Invitrogen）培地において成長させた。

40

【0468】

(v) 運動ニューロン分化

未分化細胞を、Boulting et al.（提出）によってより詳細に記述される公表されたプロトコール（DiGiorgio et al., 2008）に従って分化させた。

【0469】

DNAメチル化マッピング

50

簡約表示バイサルファイトシーケンシング (RRBS)

RRBS (Cowan, C. A. et al., *N. Engl. J. Med.* 350, 1353 (2004)) を、既に公表されたプロトコール (Smith, et al., *Methods* 48, 226 (2009)) に従って、臨床試料および入力された少量のDNAに関する何らかの最適化を行って (Gu, H. et al., *Nat. Methods* 7, 133 (2010)) 行った。主なステップは以下のとおりであった: (i) 全体で50 ng (ES細胞) または1 μg (結腸試料) のゲノムDNAを5 Uから20 UのMspI (New England Biolabs, NEB) によって16時間まで消化した。(ii) 消化したDNAの末端修復およびアデニル化を、10 Uクレノウ断片 (3' 5' エキソ、NEB)、プレミクスヌクレオチド三リン酸 (1 mM dGTP、10 mM dATP、1 mM 5'メチル化dCTP) 2 μl からなる20 μl 反応物中で行った。反応物を30分で30分間インキュベートした後、37 °C でさらに30分間インキュベートした。(iii) プレアニルした5-メチルシトシン含有Illuminaアダプターを、濃縮T4リガーゼ (NEB) 1 μl、15 μMアダプター1~2 μlを含む20 μl 反応物において16 °C で16~20時間、アデニル化DNA断片にライゲートした。(iv) インサートサイズ40~120塩基対および120~220塩基対を有する断片の、ゲルに基づく選択を、既に記述されたように行った (Gu, H. et al., *Nat. Methods* 7, 133 (2010))。(v) EpiTectバイサルファイトキット (Qiagen) によるバイサルファイト処置を、ホルマリン固定およびパラフィン包埋組織から単離されたDNAに関して設計されたプロトコールに従って行った。バイサルファイト変換率を最大限にするために、2回の変換ラウンドを行った。最終的なバイサルファイト変換DNAを、予め加温した (65 °C) EB緩衝液2 × 20 μl によって溶出した。(vi) 最終的なライブラリの濃縮のためのPCRサイクルの最少数を決定するために、バイサルファイト処置DNA 0.5 μl、各0.2 μM のIllumina PCRプライマー-LPX1.1および2.1、ならびに0.5 U PfuTurbo Cx Hotstart DNAポリメラーゼ (Stratagene) を含む分析的 (10 μl) PCR反応をセットアップした。サーモサイクラー条件は: 95 °C で5分、95 °C で20秒、65 °C で30秒、72 °C で30秒で多様なサイクル数 (10~20回) の後、72 °C で7分間であった。PCR産物を4~20% ポリアクリルアミドCriterion TBEゲル (Bio-Rad) において泳動させることによって可視化して、SYBRグリーンによって染色した。最終的なライブラリは、各々がバイサルファイト変換鑄型2~3 μl、1.25 U PfuTurbo Cx Hotstartポリメラーゼ、ならびに各0.2 μMのIllumina LPX 1.1および2.1 PCRプライマーを含む25 μl PCR反応の8によって生成された。ライブラリをPCR増幅して、既に記述されているように (Gu, H. et al., *Nat. Methods* 7, 133 (2010))、IlluminaゲノムアナライザIIにおいてシーケンシングした。シーケンシングのリードを、RRBSデータのために開発されたカスタムアラインメントソフトウェア (Meissner, A. et al., *Nature* 454, 766 (2008)) を用いてヒトゲノムのNCBI36 (hg18) アセンブリと整列させた。

【0470】

いくつかの態様において、RRBSは、既に公表されたプロトコール (Smith et al., 2009) に、細胞数が少ないことに対する何らかの最適化を行って (Gu et al., 2010) 行った。未加工のシーケンシングのリードをMaqのバイサルファイトアラインメントモード (Li et al., 2008) を用いて整列させて、カスタムソフトウェア (Gu et al., 2010) を用いてDNAメチル化の呼び出しを行った。所定の細胞株が全てのヒトES細胞株の参照から外れる遺伝子プロモーターを同定するために、本発明者らは、関心対象の細胞株と、試験に含まれる (しかし、試験される細胞株を除外する) 全てのヒトES細胞株の参照とのあいだで所定の遺伝子プロモーターにおける各CpGのDNAメチル化状態を比較する重み付けt-検定を行って、フィッシャーの複合確率検定の重み付けバージョンを用いて、対応するp-値を1つの領域特異的p-値へとまとめた。遺伝子プロモーターは、Ensembl注釈付き遺伝子の注釈付き転写開始部位周囲の-5 kbから+ 1 kbの配列ウィンドウとして定義された (Hubbard et al., 2009)。各CpGでのシーケンシングカバー率に従って重み付けを行った。最後に、q-値法を用いて多数の検定を説明して (Storey and Tibshirani, 2003)、偽発見率 (FDR) 5%未満で統計学的に有意であり、かつDNAメチル化の絶対差が、図8Eにおいても理由が説明される一般的に用いられる20パーセントポイントの閾値 (Bibikova et al., 2009) を超える場合、ゲノム領域を差次的にメチル化されたと呼んだ。試料間でのシーク

エンシングの深度およびカバー率の差は、この検定の統計力に影響を及ぼす可能性があるが、低メチル化または高度メチル化のいずれかへと試験を偏らせることはないことに注意されたい。統計分析は全て、R統計値パッケージ（ワールドワイドウェブ：r-project.org/）を用いて行われ、ソースコードは著者からの要求により入手可能である。

【0471】

クローナルバイサルファイトシーケンシング

ゲノムDNAを、PureLinkゲノムDNAミニキット（Invitrogen）を用いて単離し、DNAを、EpiTectキット（Qiagen）を用いてバイサルファイト変換し、バイサルファイト変換DNA 50 ngをPCR増幅した。プライマー配列は、CD14フォワード

5'-AGTTGTGGTTGAGGTTTAGGTT-3' (SEQ ID NO: 5)

およびリバース

5'-ACCACAAAACCTTACACTTTCCA-3' (SEQ ID NO: 6)

であった。アンプリコンをゲル精製して、TOPO TAクローニングキット（Invitrogen）を用いてサブクローニングした。クローンを、シーケンシングのために無作為に選択して、シーケンシングデータを、BiQAアナライザソフトウェア（Bock et al., 2005）を用いて処理した。

【0472】

他のDNAメチル化マッピング法

メチル-DNA免疫沈降（MeDIP）

MeDIP（Down, et al., Nat. Biotechnol. 26, 779 (2008)）を、EZ DNAメチル化キット（Zymo Research）を用いて行った。全体で1試料あたりDNA 300 ngを、Bioruptor（Diagenode）を用いて間隔を8回開けて10分間（30秒オン、30秒オフ）超音波処理し、150塩基対の平均断片サイズを得た。超音波処理したDNAを末端修復して、既に記述されたように（Down, et al., Nat. Biotechnol. 26, 779 (2008)）シーケンシングアダプターにライゲートした。100から200塩基対の断片サイズの、ゲルに基づく選択の後に、製造元のプロトコールに従ってメチル化DNA免疫沈降を行った。全体で、5-メチルシトシンに対するモノクローナル抗体（EZ DNAメチル化キットに含まれる）1 μgを、免疫沈降のために用いた。免疫沈降したDNAをPCR増幅して、濃縮の特異性を、既に記述されたように（Rakyan, V. K. et al, Genome Res. 18, 1518 (2008)）選択された遺伝子座に関するqPCRによって確認した。2レーンの36塩基対シングルエンドシーケンシングを、IlluminaゲノムアナライザIIにおいて製造元の標準的なプロトコールに従って行った。デフォルトパラメータによるMqを用いてシーケンシングのリードを、ヒトゲノムのNCBI36（hg18）アセンブリ（Li, H., Ruan, J., and Durbin, R., Genome Res. 18, 1851 (2008)）に整列させた。

【0473】

メチル化DNA捕捉（MethyCap）

MethyCap（Brinkman, A. B. et al., Methods (2010)）は、SX-8G / IP-Star（Diagenode）を用いてロボット技法で行った。His6-GST-MBD（Diagenode）2 μgを、200 mM NaClを含む結合緩衝液（BB、20 mM Tris-HCl、pH 8.5、0.1% Triton X-100）200 μl中で、超音波処理したDNA 1 μgと混合した。この溶液を4℃で2時間インキュベートした。磁気GST-ビーズは、十分に混合したMagneGSTグルタチオン粒子浮遊液（Promega）35 μlを、結合緩衝液に200 mM NaClを加えた溶液200 μlによって4℃で洗浄することにより調製した。洗浄を1回繰り返して、上清を除去した。GST-MBD-DNA溶液を、洗浄して収集したビーズに加えて、この浮遊液を4℃でさらに1時間回転させた。上清（これはフロースルーである）を除去した後、ビーズ-GST-MBD-DNA複合体を、洗浄により溶出した。異なる濃度のNaClを有する結合緩衝液200 μlを加えて、浮遊液を4℃で10分間回転させた。磁石を用いてビーズを捕捉して、上清を収集した。溶出技法は1×300 mM（洗浄）、2×400 mM（洗浄）、1×500 mM（「低い」溶出）、1×600 mM（「中程度の」溶出）、1×800 mM NaCl（「高い」溶出）からなった。収集された溶出液を、QIAquick PCR精製スピンカラム（Qiagen）を用いて精製し、溶出緩衝液100 μlによって溶出し、既に記述されたように（Brinkman, A. B. et al., Methods (2010)）シーケンシングのために調製した。Illuminaゲノムアナライ

10

20

30

40

50

ザIIにおいて行った1レーンの36塩基対のシングルエンドシーケンシングをそれぞれ、低、中程度、および高い溶出物に関して行った。シーケンシングのリードを、デフォルトパラメータを用いて、Illuminaの分析パイプライン(ELAND)を用いてヒトゲノムのNCBI36(hg18)アセンブリに整列させた。3つの溶出物の各々に関するレーンを図2に個別に示し、さらなる情報を考慮に入れることによって、Infiniumアッセイと比較して精度が改善されるか否かを判定するために試験した。しかし、3つのレーンの個別のリード数に基づく線形モデルは、3つのレーンの総和に基づくモデルより優れた性能を示さなかった。

【0474】

マイクロアレイに基づくエピゲノタイピング(Infinium)

Infinium(Bibikova, M. et al., Epigenomics 1, 177 (2009))分析は、Broad InstituteのGenetic Analysis Platformによって行われた。1試料あたり全体で1 μ gのゲノムDNAを、製造元のプロトコールに従ってパイサルファイト処置して、Infinium Human Methylationビーズアレイ(Illumina)にハイブリダイズした。本発明者らは、技術的レプリケートにおけるほぼ完全な一致を以前に観察しており(ピアソン係数>0.98)、これは各試料に関して1回のみハイブリダイゼーションを行った理由である。

【0475】

データ作成および品質管理

MeDIPおよびMethylCapに関して、整列させたリードを、超音波処理の際に得られた平均断片長に伸長させ、デュプリケートのリード(すなわち、同じ染色体上の厳密に同じ開始位置に整列させたリード)の各群から、下流の分析に及ぼすPCRバイアスの影響を最小限にするために、1つを除く全てのリードを廃棄した。RRBSに関して、整列させたリードを参照ゲノムと比較して、DNAメチル化状態を、既に記述されたように(Gu, H. et al., Nat. Methods 7, 133 (2010))カスタムソフトウェアを用いて決定した。Infinium Human Methylation27データを、標準化のためにデフォルトバックグラウンドサブトラクション法を用いてIllumina BeadStudio 3.2ソフトウェアによって処理した。UCSCゲノムブラウザトラックは、Pythonプログラミング言語で実行されるカスタムスクリプトによって構築された(<http://www.python.org/>)。

【0476】

絶対的なDNAメチル化レベルの定量

本発明者らは、線形回帰モデルを用いて、MeDIPおよびMethylCapのリード数からの絶対的なDNAメチル化レベルを推定した。多数の異なる特色選択実験に基づいて、本発明者らは、以下の変数の組み合わせが、DNAメチル化レベルを頑健に予測することを発見した：(i)所定の領域におけるMeDIPまたはMethylCapのリードの総数の平方根、(ii)領域内の全細胞抽出物(WCE)のリード総数の平方根(ChIP-seqデータ標準化のために、本発明者らがルーチンで用いている組織にまたがるWCEトラックに基づく)、(iii)領域内でのCpG頻度のロジット、(iv)領域の相対的GC含有量、(v)CpGに対するCsの比率、および(vi)RepeatMasker(<http://www.repeatmasker.org>)によって決定された領域の相対的反復配列含有量。MeDIPおよびMethylCapの両方に関して、本発明者らは、リード頻度がInfiniumデータに従って絶対的なメチル化レベルと強い正の関連があるが、反復配列含有量は中程度に正の関連があることを発見した。対照的に、CpG頻度のロジットは、DNAメチル化と高度に負の関連があり、他の全ての変数ならびにモデルの切片は、中程度に負の関連を示した。モデルの適合および性能評価のために、現在のデータセットを等しい大きさの訓練セットと試験セットとに分けた。モデルの適合は全て、R統計値パッケージ(<http://www.r-project.org/>)を用いて行った。

【0477】

差次的メチル化領域の同定

本発明者らの経験において、古典的なピーク検出(Park, P. J., Nat. Rev. Genet. 10, 669 (2009) and Storey, et al, PNAS 100, 9440 (2003))は、境界線ピークが1つの試料において検出されるが他の試料では検出されない場合に生じる多数の偽性ヒットのため

10

20

30

40

50

に (C. Bock、未発表知見)、DMR同定には適していない。代わりに、本発明者らは、2つの試料を互いに直接比較するために統計検定を用いた。RRBSデータを有する所定の領域に関して、本発明者らは、両方の試料におけるメチル化 対 非メチル化CpGの数を計数し、フィッシャーの正確確率検定を行って、領域がDMRである可能性を示すp-値を得た。同様に、MeDIPおよびMethylCapに関して、本発明者らは、両方の試料に関して領域内部で整列するリード数を計数し、フィッシャーの正確確率検定を用いてこれらの値をゲノムの他所で整列する総リード数と対比させた。Infiniumアッセイに関して、本発明者らは、領域内で全てのInfiniumプローブの2つの試料の値を比較するために対応試料t-検定を用いた。これらの検定を、同時に多数のゲノム領域 (たとえば、全てのCpGアイランド) について行い、q-値法 (Storey, et al, PNAS 100, 9440 (2003)) を用いて、p-値を多数の検定に関して補正した。0.1未満のq-値を有するゲノム領域は、高度メチル化または低メチル化であるとフラグを立てられるが (差の方向性に依じて)、これはDNAメチル化の絶対差が20%を超える場合 (RRBSおよびInfiniumの場合)、またはリード数に少なくとも2倍の差がある場合 (MeDIPおよびMethylCapの場合) に限られた。これらの閾値は、異なる細胞型のあいだでの多数の比較におけるその実際的な有用性によって選択され、これ以上の正当な理由はない。本発明者らはまた、不十分なシーケンシングカバー率を有するゲノム領域をマークするが、それらをDMR分析から除外しない。MeDIPおよびMethylCapに関して、本発明者らは、より高い解読カバー率を有する試料に関して1000万回の総解読数当たり少なくとも10回の解読を推奨し、RRBSに関して、本発明者らは両方の試料における少なくとも5回の解読について少なくとも5個のCpGを用いることを推奨した。

10

20

30

40

50

【0478】

DMR同定に対するこの統計学的アプローチでは、本発明者らは分析が行われるゲノム領域セットを定義する必要がある。本発明者らは、興味深いDMRを発見する機会を最大限にするために二元戦略を追求した。一方で、本発明者らは、エピジェネティックな調節の重要な候補であるCpGアイランドおよび遺伝子プロモーターに関して特に焦点を当てた。このアプローチは、全ゲノムの場合と比較してCpGアイランドおよび遺伝子プロモーターが比較的少ない数であることにより、多数の試験の補正をするという負担が低減されることから、周知の機能的役割を有する領域に関する統計力の増加をもたらす。他方で、本発明者らは、任意の候補領域の外側に位置するDMRを検出するためにゲノムの1キロベースタイリングを用いた。さらにより広い網を打つために、本発明者らは、CpGアイランドおよび遺伝子プロモーターのみならず、CpGアイランドショア³⁰、エンハンサー⁶⁰、進化的に保存された領域および他のタイプのゲノム領域を含む、包括的な13タイプのゲノム領域セットを収集した。これらの領域セットの全てに関するDMRデータは、PythonおよびRスクリプトセットを用いて算出され、オンラインで利用可能である (<http://meth-benchmark.computational-epigenetics.org/>)。

【0479】

実験的検証

2つの異なるES細胞株において差次的にメチル化されたとして検出されたCpGアイランドに基づき、本発明者らは、実験的検証のために8個の方法特異的DMRを手動で選択した。この目的に関して、1つの方法によって (しかし、他の2つの方法によっては同定されない) 統計学的に有意なDMRとして同定されたCpGアイランドを、UCSCゲノムブラウザにおいて肉眼で検分して、データが方法特異的DMRとしてのその分類を完全に支持している場合に限って、領域を検証のために選択した。特に、第二の方法が、第一の方法と同じ方向の示唆に富むが有意でない傾向を既にピックアップしている場合には、または第一の方法のデータが、DMRが偽陽性ヒットであることを既に示唆している場合には (たとえば、DMRの近傍で矛盾する傾向のために)、領域を選択しなかった。実験的検証は、確立されたプロトコール⁶¹に従ってクローナルバイサルファイトシーケンシングによって行われた。MethPrimer⁶²を用いて、アンプリコンが、本発明者らの当初のデータに従って差次的メチル化の最高レベルを示すCpGと重なり合うように、プライマーを設計した。バイサルファイトシーケンシングに関して調製するために、DNA 1 μ gを、EpiTectキット (Qiagen) を用い

てバイサルファイト変換して、バイサルファイト変換したDNA 50 ngをPCR増幅し、精製したアンプリコンをTOPO TAクローニングキット (Invitrogen) を用いてクローニングした。各領域に関して平均で11個のクローンを、シーケンシングのために無作為に選択した。シーケンシングデータは全て、BiQAアナライザソフトウェア (Bock, C. et al., *Bioinformatics* 21, 4067 (2005)) を用いて処理した。

【0480】

反復配列DNAの分析

反復配列を、RepBase Updateのバージョン14.07データベースから得たが (Jurka, J., *Trends Genet.* 16, 418 (2000))、これはオンラインで (<http://www.girinst.org/server/RepBase/index.php>) 公共に入手可能である。全体で11,670個のプロトタイプ反復配列から、本発明者らは、分類学ツリーにおいてヒトまたはその先祖のいずれかに対して注釈をつけられた、1,267個の配列を選択し、これらのプロトタイプ反復配列を仮のゲノムファイルにまとめた。デフォルトパラメータによるMaqを用いてMeDIP、MethylCap、RRBS、ChIP-seq (H3K4me3)、および全細胞抽出物 (WCE) シーケンシングのリード (Li, H., Ruan, J., and Durbin, R., *Genome Res.* 18, 1851 (2008)) をこの仮ゲノムに対して整列させた。RRBSの場合、リードおよび参照ゲノムをいずれも、アラインメントの前にインシリコでバイサルファイト変換した。各プロトタイプの反復配列のエピジェネティックな状態を、以下のように定量した：(i) MeDIP、MethylCap、およびChIP-seqに関して、本発明者らは、WCEデータに対するオッズ比を算出した。(ii) RRBSに関して、本発明者らは、整列させたリードとプロトタイプ反復配列との比較に基づいて、メチル化CpGの数、CpG測定の総数およびDNAメチル化の百分率を算出した。

【0481】

本発明者らは、WCEカバー率が整列させたリード100個未満である、または25個未満のCpG測定によるRRBSカバー率を有する、まれな反復配列を廃棄し、それによって今後の分析のために用いられる553個のプロトタイプ反復配列を得た。これらは、LINEクラス配列が97個 (それらのうち92個は、L1ファミリーに由来する)、SINEが51個 (それらの48個はAluファミリーに由来する)、SVAが6個、DNA反復配列が62個、サテライト反復配列が15個、LTRが315個、低複雑度反復配列が1個、およびRNA反復配列が6個であった。MeDIPおよびMethylCap試料の対のあいだの差次的メチル化を定量するために、本発明者らは、各プロトタイプ反復配列に関して解読カバー率のペアワイズオッズ比を算出したが、RRBSの場合では、DNAメチル化の絶対差を用いた。差の有意性を、非反復ゲノムの場合と同じように (先に記述したように)、フィッシャーの正確確率検定を用いて評価した。

【0482】

遺伝子発現プロファイリング

Broad Instituteのマイクロアレイコア施設によってマイクロアレイ分析を行った。全体を通してAffymetrix GeneChip HT HG-U133Aマイクロアレイを用いた。マイクロアレイの強度データを、BioconductorのgcRMAパッケージ (Gentleman et al., 2004) を用いて正規化し、およびアレイ品質測定 (Kauffmann et al., 2009) を用いて品質を管理した。所定の細胞株が全てのヒトES細胞株試料の参照から外れる遺伝子を同定するために、本発明者らは、limmaパッケージ (Smyth, 2005) において実行される調整t-検定を行って、関心対象の細胞株を、この試験に含まれる (しかし試験される細胞株を除外する) 全てのヒトES細胞株の参照と比較した。本発明者らは、発現レベルが、その遺伝子の参照遺伝子発現と比較して、10%未満のFDRで統計学的に有意であるか、および/または少なくとも2倍もしくは1 log-2倍を超えてアップレギュレートもしくはダウンレギュレートされた発現レベルを有する場合に、遺伝子を差次的に発現されたと呼んだ。統計分析は全て、R統計値パッケージ (ワールドワイドウェブ: r-project.org/) を用いて行い、ソースコードは著者からの要求により入手可能である。

【0483】

定量的RT-PCR分析

総RNAを、RNeasyキット (Qiagen) を用いて、製造元の推奨に従って単離した後、標準

10

20

30

40

50

的なプロトコルを用いるcDNA合成を行った。簡単に説明すると、cDNAをSuperscript II 逆転写酵素 (Invitrogen) およびランダム6量体 (Invitrogen) を用いて、総RNA投入量500 ngで合成した。SYBR Green PCRマスターミクス (Applied Biosystems) をqPCR分析のために用い、これはStepOnePlusリアルタイムPCRシステム (Applied Biosystems) において行われた。PCR条件は以下のとおりであった：94 °Cで5分間の初回変性、94 °Cで15秒、60 °Cで15秒、72 °Cで30秒を40サイクル、および72 °Cで10分。プライマー配列は、CD14フォワード

5'-ACGCCAGAACCTTGTGAGC-3' (SEQ ID NO: 7)

およびリバース

5'-GCATGGATCTCCACCTCTACTG-3' (SEQ ID NO: 8)

; CD33フォワード

5'-TCTTCTCCTGGTTGTCAGCT-3' (SEQ ID NO: 9)

およびリバース

5'-GAGGCAGAGACAAAGAGCG-3' (SEQ ID NO: 10)

(Garnache-Ottou et al., 2005) ; CD64フォワード

5'-GTGTCATGCGTGGAAGGATA-3' (SEQ ID NO: 11)

およびリバース

5'-GCACTGGAGCTGGAAATAGC-3' (SEQ ID NO: 12)

(Li et al., 2010) ; ならびにGAPDHフォワード

5'-ACCCACTCCTCCACCTTTGAC-3' (SEQ ID NO: 13)

およびリバース

5'-ACCCTGTTGCTGTAGCCAAATT-3' (SEQ ID NO: 14)

であった。比較閾値サイクル (Ct) 法を用いて相対的定量を算出した。

【0484】

定量的胚様体アッセイおよび系列スコアカード

胚様体の分化に関して、ES/iPS細胞をディスペルゼまたはトリプシンによって処置し、bFGFおよびプラスマネートを含まないヒトES培養培地の存在下で低接着プレートにおいて浮遊培養した。48時間毎に培地を交換しながら、細胞凝集体または胚様体を全体で16日間成長させた。16日目に、細胞を溶解して、Trizol (Invitrogen) を用いて総RNAを抽出した後、RNeasyキット (Qiagen) を用いてカラムを洗浄した。次いで、RNA 300から500 ngを、製造元の説明書に従ってNanoString nCounterシステムでの分析のために用いた。nCounterコードセットは、細胞状態、多能性、および分化をモニターできることに基づいてコンピュータにより選択された500個の遺伝子を含んだ。nCounterシステムは、ごく最近導入されたことから、発現値を標準化するための最善の実践は存在しない。本発明者らは、いくつかの異なる技法を試して、陽性対照を用いるspike-inによる標準化とVSNアルゴリズム (Huber et al., 2002) の組み合わせが、最善の結果を生じることを見いだした。データ分析を、マイクロアレイデータと同じ様式で行った。具体的には、本発明者らは、関心対象の細胞株に関する胚様体における遺伝子発現を、試験に含まれる (しかし、試験される細胞株を除外する) 全てのES細胞由来胚様体の参照と比較するために、調整t-検定を用いた。遺伝子セットの試験を準備するために、本発明者らは、全ての遺伝子に対するt-スコアの平均値および標準偏差を算出した。次に、本発明者らは、予め定義された全ての遺伝子セットに関して平均t-スコアを個別に算出して、本発明者らは、既に記述されたように全ての遺伝子に対する平均値に対してパラメトリック検定を行った (Kim 2005)。系列スコアカードダイアグラムに関して、本発明者らは、遺伝子試験の平均値と、全ての寄与する遺伝子セットについて平均した有意性とは無関係なt-スコアの全体的な平均値とのあいだの符号のある差をプロットした。

【0485】

免疫細胞化学およびFACS分析

免疫染色は以下の一次抗体を用いて行った：AFP (Dako)、NESTIN (Chemicon)、OCT4

10

20

30

40

50

(Santa Cruz Biotechnology)、 α -SMA (Sigma)、SSEA3 (Biolegend)、SSEA4 (Chemicon)、TRA-1-60 (Chemicon)、TRA-1-81 (Chemicon)、IIIチューブリン (Abcam)、VEGFR II (Abcam)。FACS分析に関して、EBをトリプシンにより単細胞へと解離し、PBSによって洗浄し、4%パラホルムアルデヒドによって終夜固定し、0.5%PBS-Tweenによって20分から1時間透過性にした。細胞(約500個)を、10%ロバ血清を添加した0.1%PBS-Tween中で1時間ブロックし、一次抗体(AFP:1:300倍希釈、DakoCtomatic)と共に終夜インキュベートし、二次抗体と共に1時間インキュベートした後、洗浄して、0.1%ロバ血清を含むPBS 1 ml中に浮遊させた。試料をBD Biosystems LSR IIアナライザを用いて分析した。FACS分析に関して、EBをトリプシンによって単細胞へと解離させ、PBSによって洗浄し、4%パラホルムアルデヒドによって終夜固定し、0.5%PBS-Tweenによって20分から1時間透過性にした。細胞(約500個)を、10%ロバ血清を添加した0.1%PBS-Tween中で1時間ブロックし、一次抗体(AFP:1:300倍希釈、DakoCtomatic)と共に終夜インキュベートし、二次抗体と共に1時間インキュベートした後、洗浄し、0.1%ロバ血清を含むPBS 1 ml中に浮遊させた。試料をBD Biosystems LSR IIアナライザを用いて分析した。

10

20

30

40

50

【0486】

偏差スコアカードの計算

偏差スコアカードは、関心対象の細胞株においてどの遺伝子が、およびどれほど多くの遺伝子が、ES細胞参照から外れるかを要約する。参照は、20個の低継代ES細胞株、または通常は参照の一部である1つの細胞株に関する偏差スコアカードを算出する場合には残りの19個のES細胞株で構成される。偏差スコアカードを算出するためのアルゴリズム(図11Aに概要)は、DNAメチル化および遺伝子発現データと同じであるが、唯一の例外はマイクロアレイデータがさらなる正規化段階を必要とする点である。統計学的見地から、偏差スコアカードは、チューキーの外れ値フィルター(Tukey, 1977)を用いるノンパラメトリック外れ値検出に基づく。関心対象の細胞株のDNAメチル化または遺伝子発現値が、四分位数間範囲の1.5倍を超えて中心四分位数の外側となる全ての遺伝子を、疑外れ値と見なし、そのようにフラグを立てる。次に、変化の程度を検討して、生物学的に意味があると見なされるほどES細胞参照からの偏差が十分に大きい遺伝子のみ、最終的に外れ値として報告する。DNAメチル化に関して少なくとも20パーセントポイント、および遺伝子発現に関して少なくとも2倍の閾値を、本明細書において用いるが、この閾値はこれまでの研究と一貫しており(Bock et al., 2010)、図10Cにおいてさらに理由が説明される。偏差が多かれ少なかれ、どの遺伝子が影響を受けるかに関係する可能性があるという事実を説明するために、DNAメチル化欠損に関して特に厳密にモニターすることが推奨される遺伝子、すなわち系列マーカー遺伝子と癌遺伝子(たとえば、腫瘍抑制遺伝子および腫瘍遺伝子)の2つのリストを作成した。これらの遺伝子の偏差は、偏差スコアカードの拡張版(表6)において具体的に強調される。最後に、本発明者らはまた、調整t-検定に基づくパラメトリックアプローチを含む、外れ値にフラグを立てるための代替りの戦略を評価している。全体として、チューキーの外れ値フィルターは、最も関連する結果を与えると判断され、「参照範囲帯」の箱ひげ図によって直観的に可視化することができるというさらなる利点を有する(図1Cおよび4A)。

【0487】

系列スコアカード計算

系列スコアカードは、19個の低継代ES細胞株によって構成される参照に対して関心対象の細胞株の分化指向性を定量する。系列スコアカードを算出するためのアルゴリズム(図11Bに概要)は、調整t-検定(Smyth, 2004)とt-スコアについて行われた遺伝子セット濃縮分析(Nam and Kim, 2008; Subramanian et al., 2005)との組み合わせを用いる。系列特異的分化指向性の定量に関する生物学的根拠を提供するために、3つの胚葉(外胚葉、中胚葉、内胚葉)ならびに神経および造血系列の各々のマーカー遺伝子のいくつかのセットを収集した(表7、表13A、および表14)。次に、関心対象の細胞株に関して得られたEBにおける遺伝子発現と、ES細胞参照から得られたEBとを比較する調整t-検定をBioconductorのlimmaパッケージを用いて行い、関連遺伝子セットに寄与する全ての遺伝子に対し

て平均 t - スコアを算出した。高い平均 t - スコアは、試験されるEBにおける遺伝子セット遺伝子の発現の増加を示しており、対応する系列に関する高い分化指向性を示すと見なされる。対照的に、低い平均 t - スコアは、関連する遺伝子の発現の減少を示し、対応する系列に関する低い分化指向性を示すと見なされる。分析の頑健性を増加させるために、平均 t - スコアを、所定の系列に割付された全ての遺伝子セットについて平均した。系列スコアカードダイアグラム (図5BおよびD) は、これらの「遺伝子セット平均 t - スコアの平均値」を、細胞株特異的分化指向性の定量的指標として記載する。系列スコアカード分析および検証は、カスタムRスクリプト (ワールドワイドウェブ : r-project.org/ から入手可能) を用いて行った。最後に、Boulting et al.、によって実験的に誘導された運動ニューロン分化効率、系列スコアカードの予測力を決定するための、細胞株の真の試験セットを提供する。さらに、系列スコアカードのバイオインフォマティクスアルゴリズムは、2つのデータセット間の第一の比較の前に既に確定されており、スコアカードの局面は、適合を改善するために適時的に最適化されなかった。

【 0 4 8 8 】

バイオインフォマティクス分析およびデータアクセス

方法特異的なデータ標準化およびスコアカードの計算 (上記の) に加えて、バイオインフォマティクス分析を以下のように行った。

【 0 4 8 9 】

(i) 階層的クラスタリング (図1、3、8、および9)

DNAメチル化レベルを、Ensembl注釈付き転写物のプロモーター領域における全てのCpGに対するカバー率重み付け平均として算出した：遺伝子発現レベルを、マイクロアレイ上の全ての関連するプローブについて平均することによって、各Ensembl遺伝子に関して算出した。階層的クラスタリングの前に、2つのデータセットを、両方のデータセットに対して等しい重みを与えるために、平均値ゼロおよび分散1へと個々に正規化した。ヒートマップは、250個の遺伝子の代表的な選択物を示す。階層的クラスタリングは、ユークリッド距離関数および平均連鎖法を用いて、R (ワールドワイドウェブ : r-project.org/ から入手可能) において行った。

【 0 4 9 0 】

(ii) 注釈付きクラスタリングおよびプロモーター特徴 (図2D)

最も変動性の遺伝子における一般的な特徴の同定は、DAVID (Huang et al., 2007) およびEpiGRAPH (Bock et al., 2009) を用いて、デフォルトパラメータを用いてEnsembl遺伝子注釈 (プロモーターは、転写開始部位周囲の -5 kb から + 1 kb の配列ウィンドウとして定義される) に基づいて行った。

【 0 4 9 1 】

(iii) ES対iPS細胞株の分類 (図3D)

これまでに報告されたiPS遺伝子シグネチャーを検証するために、所定のシグネチャーにおける全ての遺伝子に対する平均DNAメチル化または遺伝子発現レベルを、現在のデータセットから算出した。最も識別的な閾値を選択するために、ロジスティック回帰を用い、各シグネチャーの予測性をleave-one-out交差検証によって評価した。新しい分類子を誘導するために、サポートベクターマシンをDNAメチル化データ、遺伝子発現データ、または両方のデータセットの組み合わせに関して訓練した。

【 0 4 9 2 】

各分類は無作為に選択された7500個の属性に基づき、これは、1回の分析において計算上実現可能である属性の最大数であった。全ての分類子の予測性を、leave-one-out交差検証によって評価し、無作為な属性セットを有する100個の分類子についての平均性能を図3Dに報告する。これらの分類はいずれも、特色の選択を用いなかったことに注意されたい。教師ありまたは教師なしの特色選択は、予測精度を増加させる可能性があるが、第二の検証データセットがない場合には、そのような改善が、予測性の真の増加を反映するのかが、または現行のデータセットに対する過剰適合を反映するかは不明である。予測は全て、Wekaソフトウェア (Frank et al., 2004) を用いて行われた。

10

20

30

40

50

【 0 4 9 3 】

(iv) エピジェネティックな記憶の線形モデル

2つの代わりの線形モデルをDNAメチル化および遺伝子発現の両方に関して構築した。第一のモデルは、各遺伝子のiPS細胞特異的平均DNAメチル化（または遺伝子発現）レベルを、ES細胞特異的平均DNAメチル化（または遺伝子発現）レベルに回帰推定する。第二のモデルは、各遺伝子のiPS細胞特異的平均DNAメチル化（または遺伝子発現）レベルを、ES細胞特異的および線維芽細胞特異的平均DNAメチル化（または遺伝子発現）レベルに回帰推定する。両方のモデルを、分散分析（ANOVA）によって比較した。計算は全てR（ワールドワイドウェブ：r-project.org/から入手可能）において行った。

【 0 4 9 4 】

実施例1

hES細胞株間のDNAメチル化および転写の変動

そのDNAメチル化、転写、または分化指向性に影響を及ぼしうる、所与のES細胞株の多くの特性が存在する。これらは、細胞株の遺伝的背景、細胞株を培養する方法、長期間のインビトロ成長によって与えられる選択圧、または説明されない確率的ノイズを含みうる。多能性幹細胞株の挙動における変動の潜在的な根本原因を調べようと試みる前に、細胞株の実質的なコホート内に存在する変動の性質および程度の両方を最初に決定することが重要である。

【 0 4 9 5 】

多能性幹細胞集団または株における細胞株間変動を調べるために、本発明者らは、低継代数の19個のヒトES細胞株（p15~25）を得て、それらを標準化条件で数継代培養した後、DNAメチル化の分析のためにDNAおよび転写プロファイリングのためにRNAを収集した（表1、図8A）。別の細胞型との比較を行うために、RNAおよびDNAをいずれも、遺伝的に無関係なドナーの上腕から得られた6個の低継代ヒト皮膚線維芽細胞株から分析した。

【 0 4 9 6 】

（表1）ハイスループット実験において用いられた細胞株の要約

* RRBS、マイクロアレイ、および/またはNanoStringデータにおける、chrYの有無およびX-染色体の不活化の証拠によって確認した。

10

20

細胞株	参考文献	ドナーの年齢	ドナーの性別*	同胞対(ES)/ドナー(iPS)	RRBSに関する継代数	マイクロアレイに関する継代数	系列スコアカードに関する継代数
HUES1	Cowan et al. 2004	NA	雌性		22	26	26,26
HUES3	Cowan et al. 2004	NA	雄性		27	27	27,28
HUES6	Cowan et al. 2004	NA	雌性		23	23	19,21
HUES8	Cowan et al. 2004	NA	雄性		27	27	25,26
HUES9	Cowan et al. 2004	NA	雌性		21	21	19,18
HUES13	Cowan et al. 2004	NA	雄性		47	47	NA
HUES28	Chen et al. 2009	NA	雌性		17	17	13,15
HUES44	Chen et al. 2009	NA	雌性		18	18	15,16
HUES45	Chen et al. 2009	NA	雌性		20	20	17,19
HUES48	Chen et al. 2009	NA	雌性		19	19	16,17
HUES49	Chen et al. 2009	NA	雌性		17	17	14,14
HUES53	Chen et al. 2009	NA	雄性	A	17	18	17,18
HUES62	Chen et al. 2009	NA	雌性	B	14	17	15,16,16,18
HUES63	Chen et al. 2009	NA	雄性	B	19	14	19,17
HUES64	Chen et al. 2009	NA	雄性	B	19	19	18,20
HUES65	Chen et al. 2009	NA	雄性		19	19	16,17
HUES66	Chen et al. 2009	NA	雌性	A	20	20	15,15
H1	Thomson et al. 1998	NA	雄性		34	34	33,34
H7	Thomson et al. 1998	NA	雌性		48	48	NA
H9	Thomson et al. 1998	NA	雌性		NA	58	57,58
hiPS 11a	Boulting et al.	36	雄性	11	22	22	14,18,27,29
hiPS 11b	Boulting et al.	36	雄性	11	13	13	15,18,25,31
hiPS 15b	Boulting et al.	48	雌性	15	27	16	29,30,41,44
hiPS 17a	Boulting et al.	71	雌性	17	14	12	10,16,17,19
hiPS 17b	Boulting et al.	71	雌性	17	32	32	18,20,38
hiPS 18a	Boulting et al.	48	雌性	18	30	30	31,32,46
hiPS 18b	Boulting et al.	48	雌性	18	27	27	20,37
hiPS 18c	Boulting et al.	48	雌性	18	36	27	30,32
hiPS 20b	Boulting et al.	55	雄性	20	43	43	26,31,46,50
hiPS 27b	Boulting et al.	29	雌性	27	31	31	27,28
hiPS 27e	Boulting et al.	29	雌性	27	32	30	30,31,32,32,35
hiPS 29d	Boulting et al.	82	雌性	29	NA	NA	14,15
hiPS 29e	Boulting et al.	82	雌性	29	NA	NA	25,27
hFib_11	Boulting et al.	36	雄性	11	8	8	7,8
hFib_15	Boulting et al.	48	雌性	15	7	7	6,7
hFib_17	Boulting et al.	71	雌性	17	7	7	6,7
hFib_18	Boulting et al.	48	雌性	18	7	7	6,7
hFib_20	Boulting et al.	55	雄性	20	7	7	6,7
hFib_27	Boulting et al.	29	雌性	27	7	7	6,7

* RRBS、マイクロアレイ、および/またはNanoStringデータにおける、chrYの有無およびX-染色体の不活化の証拠によって確認した。

【0497】

本発明者らは、いくつかの理由から他のクロマチン修飾よりむしろES細胞におけるDNAメチル化を試験することを選択した。プロモーター領域におけるCpGジヌクレオチドのメチル化は、長期間の有糸分裂遺伝可能な遺伝子サイレンシングに関連する(Bird, 2002; Reik, 2007)。それゆえ、細胞株のあいだの差次的なDNAメチル化によって、分化の際に多様な遺伝子発現が起こり、場合により発達能に影響を及ぼす可能性がある。DNAメチル

10

20

30

40

50

化を研究するもう1つの合理的根拠は、これが、非常に定量的なアッセイによって、すなわちDNAのバイサルファイト修飾後にDNAのシーケンシングによって測定することができる (Laird, 2010) という点である。DNAメチル化の全ゲノムレベルを測定するための確立された方法の系統的比較を行った後 (Bock et al., 提出)、本発明者らは、この試験に用いるために簡約表示バイサルファイトシーケンシング (RRBS) を選択した (Gu et al., 2010; Meissner et al., 2008)。

【0498】

RRBSを用いて、本発明者らは、各細胞株に関して400万個を上回る個々のCpGジヌクレオチドのメチル化状態を定量した。このゲノムスケールのカバー率により、本発明者らは、全ての遺伝子プロモーターの4分の3、CpGアイランドの大部分、および他の多くのゲノム要素におけるメチル化状態を決定することが可能となった (図8Bおよび8C、ならびにデータは示していない)。本発明者らは、各細胞株におけるおよそ400万個のCpGでの平均15~20回のDNAメチル化測定により、細胞株間のDNAメチル化の小さい定量的な差の検出が可能となったと判断した。

【0499】

この規模の試験に関する一般的な実践と同様に (Adewumi et al., 2007; ENCODE Project Consortium, 2007; Meissner et al., 2008; Miiller et al., 2008; Narva et al., 2010)、本発明者らは、ほとんどの細胞株に関して1つのレプリケートのみを分析した。しかし、細胞株のサブセットに関して ($n=4$)、本発明者らは測定の一貫性を評価するためにさらなるレプリケートを行った。本発明者らは、RRBSおよびマイクロアレイプロファイリングの双方に関して優れた技術的再現性があること (ピアソン係数 $r > 0.99$) を証明した。生物学的再現性も同様に高く (ピアソン係数 $r > 0.95$)、2から7継代離して同じ細胞株から収集した生物学的レプリケートも同様に、他のES細胞株よりも互いに対して類似であった。本発明者らは、同じ株からの高継代 (継代 > 45) および低継代 (継代 < 30) の細胞を比較すると、強い相関 (ピアソン係数 $r > 0.95$) があることを証明したが、これらの試料は異なるES細胞株由来の試料の場合ほど互いに決して類似ではなかった (データは示していない)。長期間の培養は、DNAメチル化および転写にさらなる変動を誘導することから、本発明者らは、その後の分析を19個の低継代試料のみに集中した (表1を参照されたい)。

【0500】

転写およびDNAメチル化の集約的な全体的パターンが、ES細胞株を異なる機能的特性を有するサブクラスに分離するために十分であるか否かを判定するために、本発明者らは、データセットに対してジョイント階層的クラスタリングを行った (図1A)。対照として、本発明者らは、6個の非多能性線維芽細胞株由来の類似のデータセットを分析に含めた。予想されたように、2つの良好に分離された細胞株のクラスタが出現した。1つのクラスタは、ES細胞株の全てを含み、他は線維芽細胞対照細胞株の全てを含んだ。重要なことに、ヒトES細胞株のクラスタにおいて、さらなるサブクラスタリングの形跡はほとんどまたは全くなかった。サブクラスタリングがこのようにならないことは、その後の分析を歪めうる全体的なメチル化および転写シグネチャーを有する外れ値ES細胞株が存在しなかったことを示唆している。さらに、別個のES細胞サブクラスが存在しないことも、19個のES細胞株全てが転写およびDNAメチル化の類似する全体的なパターンを有することを強く示唆した。

【0501】

メチル化および転写の全体的なパターンは、各々のES細胞株において良好に保存されたが、多数の遺伝子座が株間で変動を示した (図1A)。その遺伝子発現およびDNAメチル化パターンに基づいて、本発明者らは、ほとんどの遺伝子座を、4つの異なるカテゴリのうち1つへと分類することができることを明らかにした。図1Bは、各クラスの代表的な例を示す。SOX2などの多くの本質的な遺伝子が、DNAメチル化または転写のいずれにおいても株間で変動を示さなかった。対照的に、CD14などのいくつかの遺伝子は、株間で多様なメチル化を有したが、GATA6などの他の遺伝子は、明確なレベルの転写を示し、一方DNAメチル化の変動は示さなかった。最後に、S100A6を含むさらに小さいクラスの遺伝子は、転

10

20

30

40

50

写およびメチル化の両方において変動を示した（図1B）。

【0502】

株間でのDNAメチル化または転写における変動が、部分的に細胞株の挙動の差が原因であるか否かを判定するために、本発明者らは、多様な特性を有する遺伝子の各々を同定した後、任意の所与の株の分化指向性を予測することができる、その変動の程度を明らかにした。それゆえ、本発明者らは、19個のES細胞株における各々の遺伝子座に関するメチル化および転写レベルの平均値、ならびにこれらの測定における変動の量を算出した（表3~5）。これらの結果は、任意の遺伝子、たとえば標的とするDNAメチル化遺伝子および標的とする遺伝子発現遺伝子に関する、ES細胞におけるDNAメチル化または転写レベルの予想されたレベルおよび範囲の値の範囲を提供するために、「参照範囲帯」または「参照DNAメチル化レベル」、または「参照遺伝子発現レベル」を包含する。これを図1Cに示し、いくつかの選択された遺伝子に関するDNAメチル化または転写レベルの平均値および範囲を示すために箱ひげ図を用いて「参照範囲帯」の概念を表示する（図1C）。これらのプロットは、「ES細胞参照の範囲内」であると見なされる各遺伝子座に関するDNAメチル化および発現レベルに対して、より上のおよびより下の閾値を課す。本発明者らはまた、「範囲帯」から外れる19個の株からの全ての測定値に対して重要な偏差スコアを割付した（図8Dおよび8Eは、DNAメチル化データおよび細胞株間で有意な違いを識別するために用いられる閾値を例証する）。この参照が手元があれば、当業者は、厳格な統計検定を行うことによって、任意の多能性細胞株における範囲帯からの偏差の数および同一性を決定することができる。さらに、細胞株間の変動に関するこの「参照マップ」を用いて、本発明者らは、この変動の性質および可能性がある起源の両方を調べて、遺伝子発現および/またはDNAメチル化が幹細胞挙動にどのように影響を及ぼすかを決定することができる。

10

20

【0503】

実施例2

ヒトES細胞株におけるエピジェネティックな転写の変動の原因および結果

ES細胞株間での転写およびメチル化における変動の原因および結果を理解し始めるために、本発明者らは、各遺伝子座に関してこれらの測定における変動レベルを定量するための「参照マップ」を用いた（表4および5）。この定量により、本発明者らは、変動する遺伝子の割合および最少のまたは実質的な変動のいずれかを有する遺伝子の同一性を決定することが可能となった。得られた分布は非常に偏っており、全ての遺伝子の16%のみがDNAメチル化変動の50%を占め、全ての遺伝子のわずか28%が遺伝子発現変動の50%を占めた（図2A）。このように、細胞株間のほとんどの変動は遺伝子座のサブセットのみに限定され、これらの2つのクラスにおける遺伝子の同一性が、それらの変動する理由に対して洞察を提供しうること、およびその変動が所定の株の特性に対して何らかの関係を有するか否かを示唆する。

30

【0504】

本発明者らは、次に、細胞株のコホート内で高度に変動性および非変動性の遺伝子座の両方の同一性を注目することに進んだ（図2A、表4および5）。予想されたように、GAPDHなどのハウスキーピング遺伝子は、幹細胞株において最も変動性の少ない遺伝子であった。同様に、本発明者らは、その機能が多能性状態に関連するSOX2およびDNMT3Bなどの遺伝子においてごく低い変動から中程度の変動が観察されたことを証明した（図2A）。対照的に、本発明者らは、GATA6、LEFTY2およびPAX6を含む胚発達を調節するいくつかの遺伝子に関するエピジェネティックなまたは転写の変動が中程度から高レベルであることを驚くべきことに発見した。最後に、株間でDNAメチル化の非常に変動するレベルを示す少数の遺伝子座が存在した。これらのゲノム要素に関して、DNAメチル化レベルは、いくつかの細胞株におけるほぼ0%のメチル化から、他の細胞株におけるほぼ100%まで変動した。これらのまれな、しかし非常に変動性の高い遺伝子は、トランスフェリンコード遺伝子TF、カタラーゼコード遺伝子CAT、およびマクロファージ/顆粒球特異的マーカー遺伝子CD14を含んだ。

40

【0505】

50

本発明者らは次に、変種遺伝子の同一性が、それらの特性が細胞株において変動する理由に洞察を提供するか否かを評価した。本発明者らは最初に、エピジェネティックな転写の変動の最高レベルを有する遺伝子にそれぞれ焦点を当てた。驚くべきことに、本発明者らは、最も変動性の遺伝子の実質的な割合が性染色体上に位置することを証明した(図2B)。この発見は、おそらく雄性および雌性細胞株の両方を含めたためである。Yに関連するメチル化および転写は、その染色体が雌性株では存在しないことから、細胞株間で変動すると予想されるであろう。X染色体の不活化における実質的な変動も同様に、別個の雌性ES細胞株において報告されており、X-連鎖遺伝子における高い程度のメチル化および転写変動に関して可能性がある説明を提供する(図2B)(Hanna et al., 2010; Lengner et al., 2010)。性染色体連鎖遺伝子は、そのような有意な変動源であることから、本発明者らは、それらが、その転写またはエピジェネティックな変動に対しよりわずかな影響を及ぼす遺伝子特色の同定能を制限しうることを心配した。それゆえ、その後の分析において、本発明者らはXおよびY染色体に連鎖した遺伝子座を除外した。

10

20

30

40

50

【0506】

本発明者らが常染色体の遺伝子座のみに焦点を当てた場合、本発明者らは、最大の構成的および転写の変動性をそれぞれ示した遺伝子セット間に明確で有意な重なりが存在することを証明した($p < 10^{-11}$ 、フィッシャーの正確確率検定、図2C)。この相関は、DNAメチル化が、最も転写的に変動性の遺伝子のサブセットに関する調節機構でありうることを証明している。遺伝子機能およびプロモーター特徴の分析は、変動するおよび変動しない遺伝子間の関連する差を強調した(図2D)。本発明者らは、転写の変動を有する遺伝子座が、細胞シグナル伝達および外部刺激に対する反応に関連する遺伝子オントロジーカテゴリに関して高度に濃縮されることを証明した。

【0507】

対照的に、多様なメチル化レベルを有する遺伝子は、任意の特定の機能に関して濃縮の証拠をほとんど示さなかった。そのかわりに、本発明者らは、これらの遺伝子のプロモーターが共通の構造特徴を有することを証明した。最も注目すべきことには、これらのプロモーターでは、DNAメチル化における変動に対して感受性があるゲノム領域の公知の特徴であるCpGジヌクレオチドが、相対的に枯渇した(Bock et al., 2006; Keshet et al., 2006; Meissner et al., 2008)。

【0508】

ヒトES細胞株における変動の機能的結果を調べるために、本発明者らは次に、ES細胞株において非常に変動性のDNAメチル化レベルを示すが、ES細胞株において常にサイレントである遺伝子をより詳細に調べた(図1B)。本発明者らは、これらの遺伝子でのエピジェネティックな欠損が、転写に対して遅れた効果を有するか否か、影響を受けた遺伝子が関連する経路に沿った分化を損なうか否かを評価した。これを証明するために、本発明者らは、強いDNAメチル化の差を有する2つのES細胞株(HUES6およびHUES8)の偏りのない胚様体(EB)分化を行って、16日目のEBにおいてDNAのメチル化ならびに遺伝子発現を測定した(図2D)。データは、2つの細胞株のあいだのDNAメチル化の差の大部分が16日目のEBにおいて保持されること($p < 10^{-16}$ 、フィッシャーの正確確率検定)、およびこれらのDNAメチル化の差がしばしば2つの細胞株間の差次的遺伝子発現に関連することを証明した($p < 10^{-5}$ 、フィッシャーの正確確率検定)。CD14は、両方のES細胞株においてサイレントであるが、HUES8に限って高度メチル化される遺伝子の例である。EB分化の際に、CD14は、HUES6に限ってアップレギュレートされる；HUES8におけるその高度メチル化遺伝子プロモーターは、それが、分化の際にそのES細胞株において活性化できないことと相関する。マクロファージおよび好中球顆粒球の正規の表面マーカーとしてのCD14の役割を考慮して、本発明者らは、定方向性の分化によりこれらの細胞の多数を生成しようとする人は、HUES8のこの特定の株を避けるべきであると判断した。より一般的に、未分化ES細胞における転写レベルでは検出できない分化における制限、または可能性がある偏りを予測するためのマーカーとしてDNAメチル化をモニターする適切性を強調する。

【0509】

実施例3

hES細胞およびhiPS細胞における、DNAメチル化および転写の全体的なパターンの類似性

ヒトES細胞株の変動に関する本発明者らの「参照マップ」により、本発明者らは、ES細胞「参照範囲帯」との統計的比較を通して、任意の新しい細胞株において正常から外れる遺伝子の数および同一性を決定することができる。様々な応用のためにヒトiPS細胞株を産生するために定義されたリプログラミング因子を用いる場合には (Park et al., 2008b; Takahashi et al., 2007; Yu et al., 2007)、所定の目的に関して最も適切なiPS細胞株を選択する方法を決定することがますます必要である。iPS細胞株に対してDNAメチル化および転写の変動をマッピングすることにより、当業者は、リプログラム細胞とそのES細胞相対物とのあいだで系統的に異なる遺伝子座が存在するか否かを判定することができるであろう。これはさらに、本明細書においてES細胞に関して記述される内容と類似する高品質iPS細胞株の選択の手引きとするために役立つであろう。

10

【0510】

それゆえ、本発明者らは、OCT4、SOX2、およびKLF4のレトロウイルス形質導入による6個の別個のドナーに由来する11個のiPS細胞株 (表1を参照されたい) におけるDNAのメチル化および遺伝子発現をマップした。これらのiPS細胞株は、十分に特徴決定されており (Boulting et al., 同時提出)、19個の参照ES細胞株と類似する培養条件で維持され、同等の継代回数でDNAおよびRNAを採取した。これらのiPS細胞株のDNAのメチル化および転写のプロファイリングを、ES細胞株の場合と同様に行い、この場合も非常に再現可能なデータが得られた (図9A)。

20

【0511】

本発明者らは、最初に、iPS細胞株がES細胞とは異なる全体的な転写およびDNAメチル化パターンを有するか否かを問うた。本発明者らは、19個のES細胞株および11個のiPS細胞株由来の完全なデータセットを用いてジョイント階層的クラスタリングを行った。対照として、本発明者らはまた、クラスタリング分析のために用いられる6個の線維芽細胞株由来のデータセットを含めた (図1A)。これまでの分析と同様に、2つの良好に分離されるクラスタが出現した。1つのクラスタは、線維芽細胞株を含み、他のクラスタは全てのES細胞株およびiPS細胞株を含んだ (図3Aおよび図9B)。重要なことに、多能性細胞株におけるサブクラスタリングは本発明者により同定されず、ES細胞とiPS細胞のあいだに何らかの系統的な差があっても、それらはこの型の分析において表れるほど強くないことを証明した。

30

【0512】

これらの2つの多能性細胞型間でより定量的な比較を行うために、本発明者らは、30個全ての細胞株由来のデータから始めて、データセットにおける各遺伝子に関するES細胞の「参照範囲帯」からの変動の程度の平均値を算出した (表4および5)。参照由来の19個のES細胞株の変動と、参照由来の11個のiPS細胞株の変動とのあいだで観察された一致は高く、ピアソン相関係数は、DNAメチル化および遺伝子発現の両方に関して $r = 0.89$ であり、iPS細胞において変動を示すほとんどの遺伝子が、ES細胞株においても非常に変動性であることを示した (図3B)。たとえば、ES細胞株においてDNAメチル化の最も変動性であるレベルを示したTF、CATおよびCD14などの遺伝子は、同様に、iPS細胞株においても最大の変動を示した。同様に予想されるように、GAPDHは、ES細胞株またはiPS細胞株において変化しなかった (図3B)。ES細胞およびiPS細胞における変種遺伝子の性質の相関は高かったが、これらの遺伝子に関するES細胞参照からのエピジェネティックな転写の偏差の定量的程度は、iPS細胞株のほうがわずかに高かった (図3C)。結論すると、メチル化および転写の変動しないおよび変動するレベルを有する遺伝子のリストは、本明細書におけるES細胞およびiPS細胞の試料採取においてほぼ完全に重なっている。

40

【0513】

実施例4

個々の遺伝子の差次的なメチル化または転写は、ES細胞とiPS細胞とを正確に区別することができない

50

全体的な類似性にもかかわらず、本発明者らは、少数の遺伝子がiPS細胞株におけるメチル化および転写の「参照」レベルから実質的に増加した偏差を示すことを証明した。プロテアーゼHTRA4 (iPS細胞株11個中9個)、ニューロン特異的RNA結合タンパク質NOVA1 (iPS細胞株11個中2個)、およびリラキシンホルモンRLN1/2 (RLN1:iPS細胞株11個中8個、RLN2:iPS細胞株11個中5個)などのいくつかの遺伝子は、iPS細胞株のサブセットにおいて高度メチル化された。他は、リゾホスホリパーゼCLC (iPS細胞株11個中3個)、およびクリスタリンCRYBB1 (iPS細胞株11個中3個)などのように、iPS細胞株においてより高レベルで転写された(図3B)。

【0514】

HTRA4のプロモーター領域は、iPS細胞株11個中9個において、および線維芽細胞株6個中6個において、高度メチル化されたが、全てのES細胞株においてはメチル化されない(n=19)。ES細胞およびiPS細胞間の、DNAメチル化パターンのそのような偏差は、分化状態の不完全なリプログラミングおよびエピジェネティックな「記憶」の証拠として解釈することができる。そのような「記憶」により、ある遺伝子座でのiPS細胞と体細胞とのあいだのDNAメチル化レベルの鏡像反映が起こると予想されるであろう。iPS細胞における体細胞のエピジェネティックな状態の有意な記憶が存在するか否かを直接および定量的に試験するために、本発明者らは、ES細胞株における変動性の交絡因子を制御しながら、遺伝子特異的な体細胞記憶の予測性に関して試験する統計学的モデルを構築した。具体的には、本発明者らは、ES細胞参照の平均値および変動、またはES細胞参照の平均値および変動、ならびにES細胞参照から線維芽細胞が逸脱する方向および程度のいずれかに基づいてES細胞参照からのiPS細胞偏差の方向および程度を予測する線形モデルを誘導した。本発明者らがこれら2つのモデルを統計学的に比較すると、本発明者らは、「エピジェネティックな記憶」を考慮に入れる後者のモデルが、iPS細胞株におけるエピジェネティックな偏差のレベルを、前者よりごくわずかに良好に説明することを証明した(0.5%の追加の変動を説明した)。エピジェネティックな記憶によって説明される変動を中程度に低減させよう、本発明者らが制御できなかった他の交絡因子が存在する可能性があるが、本発明者らのデータは、エピジェネティックな記憶が、ヒトES細胞とiPS細胞とのあいだのDNAメチル化レベルにおける変動の、有意な決定因子ではないことを明らかに証明している。

【0515】

注目すべきもう1つの遺伝子であるMEG3は、マウスのES細胞およびiPS細胞において差次的に発現されるが、四倍体胚補完法によりマウスを生成することができないと報告されている(Liu et al., 2010; Stadtfeld et al., 2010b)。MEG3は、ヒト12番染色体においてインプリンティングされたDLK1/D103ドメインにおいて見いだされる、インプリンティング遺伝子であり、様々な組織において発生段階で調節される発現パターンを示す。MEG3の発現は、ヒトES細胞株19個中10個において非常に変動性であり、残りの9個においてサイレントであった。ES細胞株におけるその変動性の発現とは対照的に、MEG3転写はiPS細胞株のいずれにおいても検出されず、iPS細胞株が由来する6個の線維芽細胞株中1個のみで中程度に発現されたに過ぎなかった(図9B)。

【0516】

本発明者らは、MEG3のサイレンシングをiPS特異的現象としてみなすべきではないことを発見した。本発明者らは、MEG3が多くの皮膚線維芽細胞株においても同様にサイレントであることを証明し、このことは、リプログラミングの際の不適切なサイレンシングの何らかの形態が、ヒトiPS細胞株において観察された低レベルのMEG3に達するのに必要であるわけではないことを暗示した。さらに、多くのヒトES細胞株がMEG3を発現せず、その発現がヒト多能性にとって必要ではないことを証明した。しかし、差次的なMEG3発現によって引き起こされたわずかな影響をヒト多能性細胞株の状況において検出することは、この影響が四倍体胚補完法によってマウスに限って観察できることを考慮すると、難しい可能性がある(Stadtfeld et al., 2010b)。より実際的な見通しから、MEG3を発現する細胞株と発現しない細胞株はいずれも広く豊富に用いられていることは安心材料である。最後の可能性として、本発明者らは、MEG3発現の変動が、ESまたはiPS細胞株のエピジェネテ

10

20

30

40

50

ィックなおよび/または転写の変動の全体的なレベルの有用なマーカーおよび指標として役立つか否かを評価した。しかし、本発明者らは、このことが当てはまらなかったことを発見した(図9D)。

【0517】

実施例5

DNAメチル化および転写の変動の統計学的モデリングはiPS細胞とES細胞とを区別する力が限定的である

本発明者らのiPS細胞とES細胞とのあいだの差を調べるアプローチは、階層的クラスタリング、ならびにHTRA4およびMEG3などの個々の精選された候補の非常に全体的なアプローチまたは系統的ベンチマーキングのいずれかを使用した。これらのアプローチはいずれもES細胞株とiPS細胞株の全体的な差異を正確に記述することができない。もう1つのアプローチは、ES細胞株とiPS細胞株を区別するために多数の遺伝子に依存する転写シグネチャーを用いることである(Chin et al., 2009)。さらに、複数のゲノム領域でのDNAメチル化レベルを合わせたものは、細胞がES細胞またはiPS細胞であるか否かを予測する(Doi et al., 2009)。したがって、本発明者らは、データセットにおける転写およびDNAメチル化シグネチャーの両方を評価して、ESまたはiPSとして細胞株を分類するが遺伝子セットそのものは分類しない閾値を再度最適化した。遺伝子発現シグネチャーに関して、本発明者らは、精度67%を証明し、これは偶然のみによって予想されるより良好であった。しかし、既に報告されたDNAメチル化シグネチャー(Doi et al., 2009)は、本発明者らの研究においてiPS細胞株をいずれも正確に同定することができなかった(図3D)。

【0518】

次に、本発明者らは、データセットからのメチル化または転写シグネチャーを調べた(表2)。既に報告された遺伝子発現シグネチャー(Chin et al., 2009)を用いて、本発明者らは、遺伝子を分類する(ES対iPS)頑健な3.4倍の濃縮が、両方の研究において同じ方向性の効果を示すことを証明したが、厳格な統計学的検定に合格した遺伝子はわずか5個に過ぎなかった。それゆえ、ES細胞株とiPS細胞株の平均遺伝子発現プロファイルの差は、本研究およびこれまでの研究(Chin et al., 2009)において保存されているが、この差は、ESまたはiPSのいずれかとして細胞株を正確に同定するには弱すぎる。

【0519】

DNAメチル化シグネチャーに関して、十分なデータを有するiPS特異的な差次的メチル化領域(Doi et al., 2009)の3分の1は、前記データセットにおいて同様に差次的にメチル化されたが、領域12個中7個は、既に報告された傾向とは反対の傾向を示した。重要なことに、同じ研究からの線維芽細胞とiPS細胞との差の98%が、この研究において同じ方向性により確認され、このことは、iPS特異的な差次的メチル化領域に関する一致の欠如が、DNAメチル化マッピングのために用いた異なる方法の副作用ではないことを示している(Doi et al., 2009)。それゆえ、本発明者らは、Doi et al.によるこれまでの研究は、真のiPS特異的なDNAメチル化欠損よりむしろ偶然に差次的にメチル化された高度に変動性のゲノム領域をおそらくは取り上げたと判断した。

【0520】

表2. 既に報告されたiPS特異的なDNAメチル化および遺伝子発現の検証
DNAメチル化データ

既に公表された、ES細胞とiPS細胞とを区別する遺伝子/ゲノム領域の検証

表11A~11Cは、DNAメチル化データである(Doi et al. 2009 Nature Genetics, <http://www.ncbi.nlm.nih.gov/pubmed/19881528>に基づく)。表11D~11Fは、遺伝子発現データである(Chin et al. 2009 Cell Stem Cell, ワールドワイドウェブサイト: "[ncbi.nlm.nih.gov/pubmed/19570518](http://www.ncbi.nlm.nih.gov/pubmed/19570518)"に基づく)。

【0521】

(表2A) DNAメチル化データ

有意な変化 ($FDR < 0.1$)		Doi et al.	
		ES細胞において上昇	iPS細胞において上昇
現在のデータセット	ES細胞において上昇	0	0
	iPS細胞において上昇	7	5
	p値	1.00	
	オッズ比	0.00	
わずかな変化 ($p < 1$)		Doi et al.	
		ES細胞において上昇	iPS細胞において上昇
現在のデータセット	ES細胞において上昇	6	5
	iPS細胞において上昇	13	11
	p値	1.00	
	オッズ比	1.02	
線維芽細胞 ($FDR < 0.1$)		Doi et al.	
		線維芽細胞において上昇	iPS細胞において上昇
現在のデータセット	線維芽細胞において上昇	572	1
	iPS細胞において上昇	20	300
	p値	$< 2.2e-16$	
	オッズ比	7792.74	

10

20

【0522】

(表2B) 遺伝子発現データ

有意な変化 ($FDR < 0.1$)		Chin et al.	
		ES細胞において上昇	iPS細胞において上昇
現在のデータセット	ES細胞において上昇	3	1
	iPS細胞において上昇	1	2
	p値	0.486	
	オッズ比	4.45	
わずかな変化 ($p < 1$)		Chin et al.	
		ES細胞において上昇	iPS細胞において上昇
現在のデータセット	ES細胞において上昇	122	92
	iPS細胞において上昇	45	114
	p値	$3.61E-08$	
	オッズ比	3.35	

30

【0523】

最後に、本発明者らは、ES細胞株とiPS細胞株とをそのDNAメチル化および/または遺伝子発現プロファイルに基づいて区別するための新規かつより正確な方法を開発するために、19個のES細胞株および11個のiPS細胞株のデータセットを用いることができるか否かを評価した。訓練データを過剰適合させる、または分類子の予測精度を過剰推定するリスクを最小限にするために、本発明者らは、厳格な統計学習アプローチを用いた (Hastie et al., 2001)。本発明者らは、任意の手動のパラメータ最適化または教師ありの特色選択を控えた (これらは正しく用いられないと予測精度を膨張させることから悪名高い)。具体的には、本発明者らは、(i) DNAメチル化データ、(ii) 遺伝子発現データ、および (iii) 両方の組み合わせに関して、ロジスティック回帰モデルならびにサポートベクターマシンを訓練した後、訓練データセットに含まれないテストケースにおける訓練した分類子の性能を評価した。サポートベクターマシンは、90%の精度に達したが (これは無作為に予想される50%または63.3%より実質的に高い)、分類子はいずれもES細胞株とiPS細胞株とを完全に識別することはできなかった (図3D)。

40

50

【 0 5 2 4 】

実施例6

ヒト多能性細胞株の品質評価のためのスコアカード

本発明者らの結果はこれまで、DNAメチル化と転写における変動が、ヒトES細胞株とiPS細胞株とのあいだに存在すること(図1)、この変動は、遺伝子のサブセットに限定されること、および所定の細胞株における遺伝子座の変動に関する知識が部分的にその挙動を予測すること(図2)を示している。しかし、ヒトES細胞とiPS細胞とを頑健に区別することができる遺伝子シグネチャーは存在しないように思われる(図3)。これらのデータからの1つの結論は、iPS細胞株が集合的に集団レベルでES細胞株の鏡像反映であり、それゆえiPS細胞が、全体に類似する程度までヒト多能性幹細胞の特徴を示すという点である。それにもかかわらず、限られた数のES細胞株および/またはiPS細胞株について研究する個々の研究者のレベルでは、これらの群のいずれかにおける、確実な遺伝子変動が実験転帰に影響を及ぼす程度を決定することは、重要である。

10

【 0 5 2 5 】

所定の応用に関する細胞株を選択するための単純かつ効率的なアプローチを開発するために、本発明者らは、特定細胞系列におけるエピジェネティックな転写の偏差を、その挙動を予測する「スコアカード」へとまとめ上げるために統計学的検定を用いた(図4A、4B、および表6)。これを行うために、本発明者らは、細胞株の、それを正常と区別する特徴に焦点を当てた。これらの選択基準はまた、一定の株を除外するための基準として用いることができる。

20

【 0 5 2 6 】

例示的な例は、マクロファージ分化が関心対象である研究者が、CD14プロモーターが高度メチル化される細胞株を使用しないように、スコアカードが役立つであろうという点であろう(図2E)。しかし、「参照」データセットの転写およびメチル化の変動から予測することができない、細胞株の多くの特徴が存在しうる。これらは、各細胞株の個々の遺伝子構成、DNAメチル化のモニタリングでは説明することができないエピジェネティックな変動、または本発明者らがまだ認識していない他の要因を含む。これらの制限を克服するために、本発明者らは、所定の分化模範例において良好に作用するその可能性に基づいて細胞株を選択するための手段を提供しうる測定を「スコアカード」に追加することを試みた。

30

【 0 5 2 7 】

表6：各ES/iPS細胞株に関するES細胞参照マップからの変動の要約

表6Aは、各ES/iPS細胞株に関するDNAメチル化変動データである。表6Bは、各ES/iPS細胞株に関する遺伝子発現変動データである。各欄の省略語の説明を表6Bの末尾に示す。

【 0 5 2 8 】

(表6A) DNAメチル化

細胞株							
試料の名称	変動	増加数	減少数	系列数	癌の数	系列マーカー	癌遺伝子
hES_HUES1	108.0%	289	19	6	13	CHRDL1+, CHRDL1+, CHRDL1+, EDA+, EDA+, ZIC3+	ARHGEF6+, FGF13+, FOXO4+, FOXO4+, FOXO4+, LCK+, LCK+, LCK+, PAK3+, PAK3+, PIM2+, RUNX1T1+, STK3+
hES_HUES3	92.2%	50	27	3	1	CD14+, CD14+, CDX4-	BCL2L10+
hES_HUES6	124.3%	66	65	1	2	SP7-	ERN2+, RARB-
hES_HUES8	73.0%	23	19	2	0	CD14+, CD14+	<なし>
hES_HUES9	73.6%	62	21	1	1	ERAS+	ERAS+

40

細胞株							
hES_HUES13	117.1%	212	168	9	12	AMN+, CAMK2A-, CAMK2A-, CD14+, CD14+, CDX4+, POU5F1+, WNT16+, ZFP42+	BCL2L10+, CAMK2A-, CAMK2A-, CFLAR+, CFLAR+, GNA14+, MX1+, NCR1-, POU5F1+, PRKCZ-, WNT16+, ZNF266+
hES_HUES28	96.0%	47	146	2	1	CD14-, GCNT2-	ALOX15B+
hES_HUES44	90.7%	318	2	10	7	AMN+, CD14+, CD14+, ERAS+, RENBP+, RENBP+, RENBP+, SYP+, SYP+, ZIC3+	ERAS+, FAM123B+, FGF13+, MAOA+, PAK3+, PAK3+, STK3+
hES_HUES45	80.3%	49	20	3	1	CD14+, CD14+, ERAS+	ERAS+
hES_HUES48	88.4%	48	3	2	0	CD14-, DDX3X+	<なし>
hES_HUES49	98.5%	248	4	13	10	CITED1+, CITED1+, EDA+, EDA+, HTATSF1+, MTM1+, MTM1+, RENBP+, RENBP+, RENBP+, SYN1+, SYP+, SYP+	AR+, ARAF+, ARAF+, FAM123B+, FAM123B+, PIM2+, SEPT6+, SEPT6+, SEPT6+, SFN+
hES_HUES53	104.4%	41	176	6	2	ANGPTL2+, ANGPTL2+, CDX4-, DPPA3-, ERAS-, SP7-	ERAS-, FGF17+
hES_HUES62	114.1%	327	44	12	20	ABCB7+, ABCB7+, CAMK2A-, CD14-, CD40+, CD40+, DES+, ERAS+, LAMP2+, RBPJ+, SYN1+, ZFP42+	ALOX15B+, CAMK2A-, CD40+, CD40+, CD74+, CD74+, CFLAR+, CFLAR+, ELK1+, ELK1+, ERAS+, ERN2+, SEPT9-, SRC+, SRC+, TCL1A+, TNFRSF25+, XIAP+, XIAP+, XIAP+
hES_HUES63	98.3%	59	21	0	0	<なし>	<なし>
hES_HUES64	87.3%	126	13	3	3	DES+, ERAS+, RBPJ+	ALOX15B+, ERAS+, SRC+
hES_HUES65	114.3%	18	293	6	7	ANGPTL2+, ANGPTL2+, CDX4-, CSF1R-, DES-, ERAS-	ALOX15B-, ERAS-, FGF17+, PSEN1-, TCL1A-, WHSC1-, ZFP37-
hES_HUES66	112.3%	32	278	2	5	CD14-, DPPA3-	BCL2L10-, ELN-, ELN-, ELN-, ZFP37-
hES_H1	95.2%	138	69	13	10	CD14+, CD14+, CD14+, CDX4-, CEACAM5+, CEACAM5+, DES-, ERAS-, GRM1+, GRM1+, ITGB2-, ITGB2-, ITGB2-	ALOX15B-, BCL2L10+, CEACAM5+, CEACAM5+, ERAS-, ERN2-, LGALS1+, SEPT9+, TCL1A-, ZFP37-
hES_H7	132.1%	428	144	10	29	ALX1+, AMN+, CDX4+, ERAS+, GCNT2+, GRM1+, GRM1+, LAMP2+, PCSK9+, ZFP42+	ALOX15B-, BCL2L10+, CACNA1B+, CACNA1B+, CACNA1B+, CACNA1B+, CASC5-, CASC5-, CFLAR+, CFLAR+, DCTN1+, DCTN1+, ERAS+, ERN2+, GOPC+, LGALS1-, NOS3-, PCSK9+, PIK3R5-, RAC2-, RAC2-, RAC2-, RAC2-, SEPT9-, SFN-, SRD5A2+, SRD5A2+, ZNF443+

10

20

30

40

細胞株							
hiPS_11a	119.9%	128	40	10	2	AMN+, ANGPTL2+, ANGPTL2+, CD14+, CD14+, CD14+, CD8A+, CD8A+, KLF4+, POU5F1+	KLF4+, POU5F1+
hiPS_11b	104.2%	56	106	10	10	CAMK2A-, CAMK2A-, CD72+, CD72+, CD93-, ELAVL4-, ERAS-, IGHD-, POU5F1+, SOX2+	CAMK2A-, CAMK2A-, CD74- , CD74-, ERAS-, GLI2-, POU5F1+, TCL1A-, ZFP37+, ZNF471+
hiPS_15b	92.5%	75	52	10	6	ANGPTL2+, ANGPTL2+, ERAS- , KLF4+, POU5F1+, RENBP+, RENBP+, RENBP+, SOX2+, SP7-	ERAS-, KLF4+, POU2AF1-, POU5F1+, TCL1A-, ZNF471+
hiPS_17a	144.4%	472	7	17	19	ARX+, CD14+, CD14+, CD72+, CD72+, CDX4+, DES+, ERAS+, POU5F1+, RENBP+, RENBP+, RENBP+, SIPA1+, SOX2+, WNT16+, ZFP42+, ZIC3+	BCL2L10+, CFLAR+, CFLAR+, ELF4+, ELF4+, ERAS+, ERN2+, GNA14+, PDE4DIP+, PDE4DIP+, PDE4DIP+, PDE4DIP+, PDE4DIP+, POU5F1+, RPL31+, SRC+, STK3+, WNT16+, ZNF471+
hiPS_17b	120.7%	511	3	19	20	CD14+, CD14+, CD14+, CD8A+, CD8A+, CD8A+, CDX4+, DES+, ERAS+, GCNT2+, HFE2+, LAMP2+, PCSK9+, PLAGL1-, POU5F1+, RBPJ+, SIPA1+, WNT16+, ZFP42+	ALOX15B+, BCL2L10+, CFLAR+, CFLAR+, DCTN1+, ERAS+, ERN2+, GNA14+, GOPC+, MAOA+, PCSK9+, PLAGL1-, POU5F1+, RUNX1T1+, SFN+, SRC+, TNFRSF25+, TNFRSF25+, WNT16+, ZNF471+
hiPS_18a	95.5%	168	44	5	8	CDX4+, DES+, KLF4+, POU5F1+, SOX2+	CFLAR+, CFLAR+, FGF13+, KLF4+, POU5F1+, RPS4X+, RPS4X+, STK3+
hiPS_18b	107.7%	287	49	16	12	ABCB7+, ABCB7+, CD40+, CD40+, CDX4+, CHRDL1+, CHRDL1+, CHRDL1+, ERAS+, FOXG1+, FOXG1+, LAMP2+, POU5F1+, SIPA1+, SOX2+, ZFP42+	CD40+, CD40+, CFLAR+, CFLAR+, ERAS+, LSM5+, MAOA+, PAK3+, PAK3+, POU5F1+, RPS4X+, RPS4X+
hiPS_18c	93.4%	377	23	19	18	CDX4+, CITED1+, CITED1+, DDX3X+, EDA+, EDA+, GPC3+, GPC3+, GPC3+, HTATSF1+, KLF4+, MTM1+, MTM1+, OSR1+, POU5F1+, SIPA1+, SOX2+, SYN1+, ZIC3+	ARHGEF6+, ARHGEF6+, ELF4+, ELF4+, ELK1+, ELK1+, FGF13+, GPC3+, GPC3+, GPC3+, KLF4+, POU5F1+, RPS4X+, RPS4X+, STK3+, XIAP+, XIAP+, XIAP+

10

20

30

40

細胞株							
hiPS_20b	119.7%	432	26	11	17	CDX4+, CFDP1+, DES+, ERAS+, GCNT2+, ID2+, PCSK9+, POU5F1+, RBPI+, WNT16+, ZFP42+	ALOX15B+, DCTN1+, ERAS+, ERN2+, GNA14+, GNAS+, GNAS+, GNAS+, GOPC+, PCSK9+, POU5F1+, SEPT9-, SRC+, TFCEP2+, TNFRSF25+, WNT16+, ZNF471+
hiPS_27b	107.5%	291	32	10	16	CDX4+, DES+, ERAS+, HFE2+, ID2+, PAX4+, PAX4+, POU5F1+, TNFRSF8+, ZFP42+	CFLAR+, CFLAR+, ERAS+, ERN2+, FGF13+, GOPC+, PDE4DIP+, PDE4DIP+, PDE4DIP+, POU5F1+, RPS4X+, RPS4X+, STK3+, TNFRSF8+
hiPS_27e	169.1%	59	504	16	12	ANGPTL2+, ANGPTL2+, CD14-, CDX4-, CSF3-, CSF3-, CSF3-, ELAVL4-, ERAS-, GCNT2-, ITGB2-, ITGB2-, ITGB2-, PLAGL1-, POU5F1+, TNNT3-	ALOX15B-, CD74+, CD74-, ERAS-, ERN2-, FZD10+, LGALS1-, PLAGL1-, POU2AF1-, POU5F1+, TCL1A-, TPM4-
hES_最小	73.0%	18	2	0	0	N/A	N/A
hES_第1四分位	89.6%	48	19	2	1	N/A	N/A
hES_平均	100.0%	136	81	5	7	N/A	N/A
hES_第3四分位	113.2%	230	145	10	10	N/A	N/A
hES_最大	132.1%	428	293	13	29	N/A	N/A
hiPS_最小	92.5%	56	3	5	2	N/A	N/A
hiPS_第1四分位	99.8%	102	25	10	9	N/A	N/A
hiPS_平均	115.9%	260	81	13	13	N/A	N/A
hiPS_第3四分位	120.3%	405	51	17	18	N/A	N/A
hiPS_最大	169.1%	511	504	19	20	N/A	N/A

10

20

30

【 0 5 2 9 】

(表 6 B) 遺 伝 子 発 現

細胞株							
試料の名称	変動	増加数	減少数	系列数	癌の数	系列マーカー	癌遺伝子
hES_HUES1	74.6%	7	1	1	0	LHX2+	<なし>
hES_HUES3	81.6%	5	2	1	0	CD151-	<なし>
hES_HUES6	88.5%	18	2	1	0	HLA-DRA+	<なし>
hES_HUES8	82.7%	6	1	0	1	<なし>	MSN+
hES_HUES9	72.0%	5	0	0	0	<なし>	<なし>
hES_HUES13	215.3%	847	500	100	131	ABCB1+, ACTA2+, AGR2+, ALB+, ALDH1A1+, ALPL-, ARID3B-, ASCL1+, BGN+, BMI1+, BMPR2+, BSG-, CAPN1-, CD55-, CD9-, CDCP1-, CDH1-, CDH3-, CEACAM6+, CLDN6-, COL1A1+, COL1A2+, COL2A1+, COL3A1+, COL4A2+, CSPG5+, CST3+, CTNND2+, DCN+, DCX+, DPPA4-, DZIP1+, ELAVL4+	ABCB1+, AGR2+, AKT3+, ALB+, ALPL-, ARNT2+, ASXL1+, BCL11A-, BCL2+, BCL7A+, BIK-, BMI1+, BNIP3L+, BOP1-, BRAF-, CANT1+, CAPN2+, CARD8+, CASP9-, CCL2+, CCND2+, CCNE1-, CDCP1-, CDH1-, CDH11+, CDKN2D+, CHEK2-, COL1A1+, COL4A1+, COL4A2+, COL4A6+, COPZ2+, CRTC3
hES_HUES28	112.8%	34	17	1	3	UTF1+	CHN1-, HRK+, MLH1-
hES_HUES44	92.0%	5	2	0	2	<なし>	CREB5+, DPF1+
hES_HUES45	72.0%	1	0	0	1	<なし>	LMO1+
hES_HUES48	104.6%	15	4	0	0	<なし>	<なし>
hES_HUES49	75.6%	5	0	0	0	<なし>	<なし>
hES_HUES53	80.6%	20	0	2	0	CGB+, FABP1+	<なし>
hES_HUES62	117.8%	40	7	2	6	CITED1+, PPARGC1A+	ARC+, FGF3+, HOXA2+, NAIP+, VLDLR-, WNT4+
hES_HUES63	92.3%	6	1	0	1	<なし>	BCL6+
hES_HUES64	84.0%	0	2	0	0	<なし>	<なし>
hES_HUES65	110.6%	43	2	7	6	DPP4+, FOXA2+, GATA4+, LHX1+, SST+, TBX3+, 注釈なし+	GATA4+, IL6+, LAMC3+, LIFR+, SST+, TBX3+
hES_HUES66	108.8%	21	21	2	6	BST2-, FGF8+	EIF4A3+, FGF8+, GGPS1-, GRB2+, HRAS+, PHB+
hES_H1	126.5%	58	55	5	9	BMP4-, ETV1+, FAM65B+, GABRA1+, NEFH-	BMP4-, CCND2+, DHCR7+, EIF5B-, ETV1+, FANCF+, LAMB1-, PSMC3-, RHOH+
hES_H7	107.5%	28	8	2	2	LLGL1+, NGFR+	NGFR+, SEPT9+
hiPS_11a	154.1%	161	255	10	29	CLDN6+, CST3+, IFNGR1-, ITGA6-, PUM2-, ROCK1-, SOX12+, TNNT2+, UTF1+, ZMYM2-	CCNA1+, CD74+, CDK2-, CHEK2-, CHN1-, CREB1-, CRK-, DHX9-, DPF1+, EIF4EBP1+, EML4-, ERC1-, FOXO4+, HRAS+, ITGA6-, MSH6-, NONO-, PAFAH1B2+, PIK3CA-, PMS1-, PSEN1-, PTK2-, PTPN11-, SFRS1-, TFCP2-, TNFAIP8-, TOP2A-, TSC1-, ZMYM2-

10

20

30

40

細胞株								
hiPS_11b	195.3%	390	129	38	40	AGR2+, ALB+, ALDH1A1+, BMI1+, BMPR2+, COL2A1+, DCN+, DLX2+, DPPA4-, ELAVL4+, EPYC+, GDF10+, GREM2+, HOXA5+, HOXC4+, ISL1+, LEF1+, LHX2+, LMO2+, LPL+, MAP2+, MEF2C+, MEIS1+, MEOX1+, MSX1+, NEFL+, NEFM+, NR2F1+, PDGFC+, PLAGL1+, SLC2A1+, SOX9+, SST+, TACSTD	AGR2+, ALB+, ASXL1+, BAX-, BCL11B+, BMI1+, BNIP3L+, BTG1+, CCNE1-, COL4A6+, COPZ2+, CTBP1+, DAP+, DDB2-, EGLN1+, FGF9+, FZD1+, GDF10+, GLUT25D2+, HTATIP2-, LEF1+, LMO2+, MITF+, MLLT3+, NR2F1+, PDGFC+, PDGFD+, PGF+, PIK3CD-, PIK3R1+, PLAGL1+, PRRX1+, RALGDS	10
hiPS_15b	122.8%	43	39	4	4	CD46-, DGCR6+, IFITM3-, ZMYM2-	CCNL1-, ORM2+, RNF7+, ZMYM2-	
hiPS_17a	146.9%	132	208	15	25	CD81-, COL1A1-, COL1A2-, COL4A2-, DGCR6+, IFITM3-, ITGAE+, LAMP1-, LXN+, MKI67-, NCSTN-, NES-, NOTCH1-, NOTCH2-, SMARCA4-	ACSL3-, BAX-, BCL6-, BID+, COL1A1-, COL4A1-, COL4A2-, CRADD+, JUP-, LAMA5-, LASP1-, LMO1+, LSM5+, MEN1-, MYH9-, NOTCH1-, NOTCH2-, NR3C1+, RNF7+, SMARCA4-, SOCS2+, TPR-, TRAF6+, TSC2-, VLDLR-	20
hiPS_17b	83.4%	0	3	0	0	<なし>	<なし>	
hiPS_18a	85.0%	3	2	0	0	<なし>	<なし>	
hiPS_18b	102.3%	32	3	0	5	<なし>	CREB5+, DDB2+, FOXL2+, IL1A+, LAMC2+	30
hiPS_18c	121.3%	57	103	2	11	CD46-, LHX1+	AXIN1+, BCL6-, ELP4-, EML4-, FANCG-, NUDT2-, PALB2-, PJA2-, SS18L1-, TNFAIP8-, TRAF5-	
hiPS_20b	172.2%	338	361	16	55	AHCTF1-, BST2-, CD46-, CNN1+, CNN2+, CSPG5+, DGCR6+, ITGA6-, ITGAE+, KLF6-, MKI67-, ROCK1-, SDC1+, TCF4-, TNNT2+, ZMYM2-	ACSL3-, ARHGGEF6-, ATM-, BAK1+, BID+, BRCA2-, C16orf5+, CASP6+, CCNL1-, CHIC2+, CIAPIN1+, CLTC-, DDB2+, DEK-, DICER1-, EIF4EBP1+, EIF5B-, ERC1-, FUS-, GNA14+, GPX1+, HRAS+, HSP90B1-, IL1A+, ITGA6-, KLF6-, KTN1-, LAMB1-, MLL-, NRAS+, OPA1-, PCM1-, PEA15+, P	40
hiPS_27b	97.5%	21	0	1	5	FZD9+	ARC+, CEP110+, FZD9+, JUNB+, PROC+	
hiPS_27e	101.9%	27	1	1	5	PPP1R13B+	EIF2S2+, ELF4+, MX1+, PPP1R13B+, TFE3+	
hES_最小	72.0%	0	0	0	0	N/A	N/A	
hES_第1四分位	81.1%	5	1	0	0	N/A	N/A	

細胞株							
hES_平均	100.0%	61	33	7	9	N/A	N/A
hES_第3四分位	109.7%	31	8	2	5	N/A	N/A
hES_最大	215.3%	847	500	100	131	N/A	N/A
hiPS_最小	83.4%	0	0	0	0	N/A	N/A
hiPS_第1四分位	99.7%	24	3	1	5	N/A	N/A
hiPS_平均	125.7%	109	100	8	16	N/A	N/A
hiPS_第3四分位	150.5%	147	169	13	27	N/A	N/A
hiPS_最大	195.3%	390	361	38	55	N/A	N/A

10

【 0 5 3 0 】

表6Aおよび6Bの説明

変動	全てのES細胞株と比較したパーセント値に標準化した、全ての遺伝子における変動 (DNAのメチル化または遺伝子発現)の平均値 例: 100% → ES細胞株平均値と同じ量の変動
増加数	全てのES細胞の参照値と比較して有意に増加したDNAメチル化/遺伝子発現レベルを有する遺伝子の数
減少数	全てのES細胞の参照値と比較して有意に減少したDNAメチル化/遺伝子発現レベルを有する遺伝子の数
系列数	有意に増加または減少した系列マーカー遺伝子の数
癌の数	有意に増加または減少した系列マーカー遺伝子の数
系列マーカー	DNAメチル化/遺伝子発現レベルが有意に増加した (+) または減少した (-) 系列マーカー遺伝子 (*)
癌遺伝子	DNAメチル化/遺伝子発現レベルが有意に増加した (+) または減少した (-) 癌遺伝子 (*)
(*) 2回の実験 (duplicate) は同じ遺伝子の別のプロモーターによる	

20

【 0 5 3 1 】

細胞株の陽性選択のための任意の適切な方法は、短時間で行うために単純であるべきであり、安価であるべきであり、および可能な限り多くの異なる系列への分化に応用するために予測的であるべきである。本発明者らは、所定の細胞株の分化が比較的偏りのない様式で開始されるか否か、次にその自然の分化指向性が、定方向性の分化プロトコールにおけるその働きを予測するか否かを評価した。言い換えれば、本発明者らは、外胚葉または神経系列の細胞を形成する自然の指向性を有する細胞株が、たとえば運動ニューロンに向かう分化において最適に機能するか否かを評価した。これを評価するために、本発明者らは、多能性細胞株の分化特性に関する単純で、迅速かつ安価なアッセイを設計し、定方向性の分化の下でそれが細胞株の挙動を予測するか否かを判定した (図5A)。

30

【 0 5 3 2 】

分化指向性を測定するために、本発明者らは、ESまたはiPS細胞株を酵素を使用して継代することによって最初に分化を開始させた後、それらをbFGFおよびプラスマネートを含まないヒトES培養培地の存在下で浮遊培養に供した。EBをこの環境で全体で16日間培養した後、総RNAの単離のために収集した。RNAを、3つの胚性胚葉ならびに神経および造血系列などの特定の体細胞系列を表す500個の系列特異的遺伝子を含めるように設計したシグネチャー遺伝子を用いてNanoString nCounterシステムを用いて分析した (表7)。標準的なマイクロアレイに対するnCounterシステムの利点は、測定の高い感受性、大きいダイナミックレンジ (Geiss et al., 2008)、および試料当たりの低い費用と共に容易で迅速な取り扱いである。データ収集後、本発明者らは、2つの生物学的レプリケートの遺伝子発現プロファイルを、対照EBからの「参照」の測定セットのプロファイルと統計学的に比較した (表10)。最後に、本発明者らは、対照の「参照」のEBと比較して細胞株特異的の分化指向性を定量するために、差次的発現t-スコアに関する遺伝子セット濃縮分析 (Nam

40

50

and Kim, 2008; Subramanian et al., 2005) を行った。

【 0 5 3 3 】

(表7) 系列スコアカードの構築のために用いられる遺伝子セットの注釈

神経系列		
リスト1__外胚葉	NCAM1, EN1, FGFR2, GATA2, GATA3, HAND1, MNX1, NEFL, NES, NOG, OTX2, PAX3, PAX6, PAX7, SNAI2, SOX10, SOX9, TDGF1	
リスト2__外胚葉	APOE, PDGFRA, MCAM, FUT4, NGFR, ITGB1, CD44, ITGA4, ITGA6, ICAM1, NCAM1, THY1, FAS, ABCG2, CRABP2, MAP2, CDH2, NES, NEUROG3, NOG, NOTCH1, SOX2, SYP, MAPT, TH	
リスト3__神経_幹_細胞	ABCG2, BMP2, CAMK2A, DLX5, EOMES, FGF2, FGFR3, FOXD3, ISL1, ITGA4, LMX1A, MAP2, MNX1, MSI1, NES, NEUROG1, NGFR, NOTCH1, NR2E1, OLIG2, PAX3, SHH, SNAI2, SOX1, SOX4, SOX9, TCF3, TCF4	10
リスト4__ニューロン_マーカー	CAMK2A, CD34, CEACAM1, CEACAM5, DLX5, EOMES, EPHB4, ISL1, ITGAM, ITGB1, MAP2, MNX1, MSI1, NCAM1, NEFL, NES, NEUROG1, NR2E1, OLIG2, PAX6, POU5F1, SDC1, SNAI2, SOX10, SOX2, SOX4, THY1, TWIST1	
リスト5__神経_幹_細胞	ABCG2, BMP2, CAMK2A, DLL1, DLX5, EOMES, FGF2, FGFR3, FOXD3, ISL1, ITGA4, LMX1A, MAP2, MNX1, MSI1, NES, NEUROG1, NGFR, NOTCH1, NR2E1, OLIG2, PAX3, SHH, SNAI2, SOX1, SOX4, SOX9, TCF3, TCF4	
リスト6__神経_幹_細胞	MCAM, FUT4, NGFR, ITGB1, ITGA6, ICAM1, FAS, ABCG2, NES, NOG, NOTCH1, SOX2	
リスト7__ニューロン	APOE, NGFR, NCAM1, THY1, MAP2, CDH2, NES, SYP, MAPT, TH	
造血系列		
リスト1__中胚葉	CD34, DLL1, HHEX, INHBA, LEF1, SRF, T, TWIST1	
リスト2__中胚葉	CD34, HHEX, INHBA, LEF1, SRF, T, TWIST1	
リスト3__中胚葉	ADIPOQ, MME, KIT, ITGAL, ITGAM, ITGAX, TNFRSF1A, ANPEP, SDC1, CDH5, MCAM, FUT4, NGFR, ITGB1, PECAM1, CDH1, CDH2, CD34, CD36, CD4, CD44, ITGA4, ITGA6, ITGAV, ICAM1, NCAM1, ITGB3, CEACAM1, THY1, ABCG2, KDR, GATA3, GATA4, MYOD1, MYOG, NES, NOTCH1, SPI1, STAT3	
リスト4__造血_前駆細胞	ABCG2, ANPEP, BMI1, BMPR1A, CD22, CD28, CD34, CD36, CD3E, CD4, CD40, CD44, CDH2, CEACAM1, DLL1, EBF1, EPHB4, ERG, ETV2, FAS, FASLG, FUT4, GATA1, ICAM1, IFNGR1, ITGA6, ITGAL, ITGAM, ITGAV, ITGAX, ITGB3, JMJD6, KDR, KIT, MME, MPL, NCAM1, NOTCH1, PECAM1, PODXL, RUNX1, SDC1, SPEN, T, TAL1, THY1, ZBTB16, ZFX	30
リスト5__血液	ANPEP, CD36, ITGAV, PECAM1, THPO	
リスト6__適応_免疫	CD22, CD28, NCAM1, CD3E, CD4, CD40, CEACAM1, CEACAM5, FASLG, GATA3, ICAM1, MME, THY1	
リスト7__先天_免疫	FAS, FASLG, IFNGR1, IRF6, JMJD6, TNFRSF1A	
リスト8__造血_前駆細胞	ABCG2, ANPEP, BMI1, BMPR1A, CD22, CD28, CD34, CD36, CD3E, CD4, CD40, CD44, CDH2, CEACAM1, CEACAM5, DLL1, EBF1, EPHB4, ERG, ETV2, FAS, FASLG, FUT4, GATA1, ICAM1, IFNGR1, ITGA6, ITGAL, ITGAM, ITGAV, ITGAX, ITGB3, JMJD6, KDR, KIT, MME, MPL, NCAM1, NOTCH1, PECAM1, PODXL, RUNX1, SDC1, SPEN, T, TAL1, THY1, ZBTB16, ZFX	40
外胚葉		
リスト1__外胚葉	NCAM1, EN1, FGFR2, GATA2, GATA3, HAND1, MNX1, NEFL, NES, NOG, OTX2, PAX3, PAX6, PAX7, SNAI2, SOX10, SOX9, TDGF1	

リスト2__外胚葉	APOE, PDGFRA, MCAM, FUT4, NGFR, ITGB1, CD44, ITGA4, ITGA6, ICAM1, NCAM1, THY1, FAS, ABCG2, CRABP2, MAP2, CDH2, NES, NEUROG3, NOG, NOTCH1, SOX2, SYP, MAPT, TH
中胚葉	
リスト1__中胚葉	CD34, DLL1, HHEX, INHBA, LEF1, SRF, T, TWIST1
リスト2__中胚葉	CD34, HHEX, INHBA, LEF1, SRF, T, TWIST1
リスト3__中胚葉	ADIPOQ, MME, KIT, ITGAL, ITGAM, ITGAX, TNFRSF1A, ANPEP, SDC1, CDH5, MCAM, FUT4, NGFR, ITGB1, PECAM1, CDH1, CDH2, CD34, CD36, CD4, CD44, ITGA4, ITGA6, ITGAV, ICAM1, NCAM1, ITGB3, CEACAM1, THY1, ABCG2, KDR, GATA3, GATA4, MYOD1, MYOG, NES, NOTCH1, SPI1, STAT3
内胚葉	
リスト1__内胚葉	APOE, CDX2, FOXA2, GATA4, GATA6, GCG, ISL1, NKX2-5, PAX6, PDX1, SLC2A2, SST
リスト2__内胚葉	APOE, ITGB1, CD44, ITGA6, THY1, CDX2, GATA4, HNF1A, HNF1B, CDH2, NEUROG3, CTNNB1, SYP

10

20

30

40

50

【0534】

多能性細胞に関する「スコアカード」のこの新しい陽性成分を評価および較正するために、本発明者らは最初に、このレポートにおいて他の分析のために用いられる19個の低継代ES細胞株における遺伝子発現をモニターするためにスコアカードを用いた（図5B、図10B、および表8）。この実験の結果は、各細胞株が3つの胚葉の各々への分化のその指向性に定量的な差を示すことを証明した。たとえば、HUES8は、内胚葉分化に対して最大の指向性を示し、この細胞株が定方向性の内胚葉分化において良好に機能するというこれまでの報告を裏付ける（Osafune et al., 2008）。この結果はまた、HUES8が定方向性の内胚葉分化に関係する細胞株に関して頻りに用いられる細胞株である理由を証明する（Borowiak et al., 2009）。

【0535】

対照的に、H1およびH9は、神経系列分化に関して高い「スコア」を有し（図5B）、それらが神経変性の研究または処置での応用にとって優れた選択でありうることを証明した。実際に、これらの細胞株が、運動ニューロンに向かう分化のアッセイにおいて良好に機能することが既に報告されている（Hu et al., 2010）。しかし、本明細書において開示されるスコアカードの本発明者らの最初の使用は過去の有用性の予測に有用であったが、本発明者らは、系列スコアカードの再現性をさらに検証した。この目的のために、本発明者らは、特定の系列の産生において比較的良好にまたは比較的不良に機能する細胞株を「スコアカード」に基づいて選択し、これらの指向性が再現性があるか否かおよびそれらが独立したアッセイによって検証できるか否かを評価した。本発明者らは、いくつかの細胞株に関するさらなる独立したラウンドのEB分化を行い、個別の系列に限って発現される5個の遺伝子（NES、TUBB3、KDR、ACTA2、AFP）のmRNAレベルを測定すると、本発明者らは、各遺伝子のRNAレベルと、本明細書において開示される「スコアカード」によって予測される分化指向性とのあいだに良好な一致を観察した（図11B）。さらに、これらの分化実験の、より定性的な評価を、接着条件でEBを平板培養した後、3つ全ての胚葉を表す様々な分化細胞型に対して特異的な抗体によって免疫染色することによって行った。この場合も、本発明者らは、所定の細胞株の分化挙動に関して良好な予測を提供した（図19および20）。

【0536】

本発明者らの最初の結果は、単純な転写アッセイが、所与のES細胞株の再現可能な挙動を予測することができることを証明した。本発明者らは次に、この同じ系列「スコアカード」を用いてiPS細胞の挙動を予測することができるか否かを評価した。この目的に関して、本発明者らは、いくつかの良好に特徴決定されたiPS細胞株を選択して（Boulting et al., 同時提出）、標準的なEB分化を行い、RNAを収集してそれらをNanoStringを用いて分析し、得られたデータを「参照」のES細胞由来EBに対して標準化した。結果は、選択されたiPS細胞株の挙動に関する系列「スコアカード」であった（図5Cおよび5D、および図10C

)。表9は、所定の多能性幹細胞株、たとえばES細胞株またはiPS細胞株の再現可能な挙動を予測するための系列スコアカードを明示する。

【0537】

表9：系列スコアカード予測（表9A）および運動ニューロンへの分化効率（表9B）。

【0538】

（表9A）系列スコアカード予測

細胞株	hiPS 11a	hiPS 11b	hiPS 15b	hiPS 17a	hiPS 17b	hiPS 18a	hiPS 18b	hiPS 18c	hiPS 20b	hiPS 27b	hiPS 27e	hiPS 29d	hiPS 29e
レプリケートの数	4	4	4	5	3	3	2	2	4	2	5	2	2
神経系列 (平均値)	-0.41	-0.73	0.34	0.14	0.02	0.24	0.74	0.84	-0.12	0.49	-1.11	0.10	0.96
造血系列 (平均値)	-0.12	-0.43	0.56	-0.11	0.39	0.44	0.54	0.55	-0.39	0.49	-0.81	0.20	0.76
外胚葉 (平均値)	-0.28	-0.68	0.50	0.17	0.01	0.21	0.75	0.89	-0.13	0.56	-1.50	0.03	1.19
中胚葉 (平均値)	-0.43	-1.01	0.84	-0.18	0.65	0.57	0.46	0.35	-0.83	0.63	-1.35	0.33	1.31
内胚葉 (平均値)	0.23	-0.05	1.90	0.41	0.11	0.11	0.08	0.08	0.06	0.57	-2.20	0.45	1.31
神経系列 (標準偏差)	0.25	0.61	0.63	0.31	0.40	0.45	0.01	0.08	0.38	0.13	0.11	0.20	0.55
造血系列 (標準偏差)	0.10	0.52	0.29	0.17	0.19	0.22	0.01	0.12	0.19	0.20	0.19	0.17	0.06
外胚葉 (標準偏差)	0.16	0.75	0.83	0.29	0.44	0.50	0.06	0.02	0.44	0.23	0.18	0.21	0.58
中胚葉 (標準偏差)	0.18	0.82	0.71	0.28	0.52	0.50	0.30	0.49	0.53	0.08	0.44	0.21	0.22
内胚葉 (標準偏差)	0.19	0.89	0.80	0.33	0.21	0.45	0.30	0.09	0.69	0.08	0.21	0.15	0.22
神経系列 (標準誤差)	0.12	0.30	0.31	0.14	0.23	0.26	0.01	0.06	0.19	0.09	0.05	0.14	0.39
造血系列 (標準誤差)	0.05	0.26	0.14	0.08	0.11	0.13	0.01	0.09	0.10	0.14	0.09	0.12	0.05
外胚葉 (標準誤差)	0.08	0.38	0.41	0.13	0.25	0.29	0.04	0.02	0.22	0.17	0.08	0.15	0.41
中胚葉 (標準誤差)	0.09	0.41	0.36	0.12	0.30	0.29	0.22	0.35	0.26	0.06	0.20	0.15	0.16
内胚葉 (標準誤差)	0.09	0.45	0.40	0.15	0.12	0.26	0.22	0.07	0.34	0.06	0.09	0.11	0.16
神経系列 (階級)	10	11	9	5	6	4	2	1	8	3	13	7	12
造血系列 (階級)	2	6	11	1	5	7	9	10	4	8	13	3	12
外胚葉 (階級)	9	11	10	5	7	4	2	1	8	3	13	6	12
中胚葉 (階級)	4	11	10	1	8	6	5	3	9	7	13	2	12
内胚葉 (階級)	3	5	12	2	8	9	6	7	4	10	13	1	11

【0539】

（表9B）運動ニューロンへの分化効率（ISL1陽性細胞の百分率）

細胞株	hiPS 11a	hiPS 11b	hiPS 15b	hiPS 17a	hiPS 17b	hiPS 18a	hiPS 18b	hiPS 18c	hiPS 20b	hiPS 27b	hiPS 27e	hiPS 29d	hiPS 29e
実験数 (各3個の レプリケート)	5	4	1	1	2	6	5	6	1	5	6	6	3
効率(平均)	6.23	0.00	13.29	8.32	7.17	10.73	13.04	15.26	7.61	11.27	0.00	9.87	0.00
効率(標準偏差)	1.67	0.00	2.63	2.63	0.29	3.02	3.86	4.34	2.63	5.03	0.00	4.37	0.00
効率(標準誤差)	0.75	0.00	2.63	2.63	0.21	1.23	1.73	1.77	2.63	2.25	0.00	1.78	0.00
効率(階級)	10	11	2	7	9	5	3	1	8	4	11	6	11

【 0 5 4 0 】

別のアッセイによる分化「スコアカード」を独立して検証するために、本発明者らは、いくつかのiPS細胞株の分化を繰り返し、フローサイトメトリーを用いて、内胚葉に対して特異的な遺伝子（AFP）を発現する細胞の百分率を分析した（図10D）。この場合も、スコアカードは、内胚葉分化に対して指向性を有する細胞株を正確に予測することができた（図10D）。

【 0 5 4 1 】

iPS細胞株の挙動を予測するためスコアカードの頑健性および再現性をさらに確認するために、本発明者らは、5回までの独立した時間で各iPS細胞株を分化させた後、単純な転写アッセイを用いて、採取したRNAを分析した（表11A、および表11B）。重要なことに、本発明者らは、所与の細胞株由来の各レプリケートによって生成されたスコアカード予測に優れた全体的相関があることを観察した（ピアソン係数 $r = 0.82$ ）。

10

【 0 5 4 2 】

表11：系列スコアカードアッセイの一貫性および再現性

【 0 5 4 3 】

（表 1 1 A）系列スコアカードアッセイの一貫性および再現性

生物学的 レプリケート	神経系列	造血系列	外胚葉	中胚葉	内胚葉	生物学的 レプリケート間 の相関
hEB16d_11a_p14	-0.68	-0.24	-0.44	-0.33	0.36	0.81
hEB16d_11a_p18	-0.13	-0.03	-0.16	-0.24	0.12	0.91
hEB16d_11a_p27	-0.53	-0.04	-0.39	-0.56	0.03	0.81
hEB16d_11a_p29	-0.28	-0.16	-0.12	-0.60	0.42	
hEB16d_11b_p18	-1.56	-1.14	-1.72	-2.09	-1.38	0.73
hEB16d_11b_p25	-0.50	-0.41	-0.49	-1.12	0.21	0.76
hEB16d_11b_p15	-0.13	-0.27	0.08	-0.19	0.48	0.55
hEB16d_11b_p31	-0.73	0.11	-0.58	-0.62	0.48	
hEB16d_15b_p29	0.57	-0.17	0.71	0.22	-0.72	0.72
hEB16d_15b_p30	-0.66	-0.62	-1.01	-1.06	-2.48	0.97
hEB16d_15b_p41	-0.44	-0.57	-0.67	-1.19	-2.27	1.00
hEB16d_15b_p44	-0.83	-0.87	-1.04	-1.31	-2.13	
hEB16d_17a_p17	-0.16	0.04	-0.02	-0.12	0.91	0.81
hEB16d_17a_p10	-0.16	-0.32	-0.17	-0.57	0.21	0.90
hEB16d_17a_p19	0.26	-0.15	0.36	-0.23	0.48	0.69
hEB16d_17a_p16	0.56	-0.20	0.56	-0.17	0.05	0.69
hEB16d_17a_p12	0.18	0.09	0.10	0.20	0.38	
hEB16d_17b_p18	0.49	-0.17	0.51	-0.11	0.03	0.81

20

30

生物学的 レプリケート	神経系列	造血系列	外胚葉	中胚葉	内胚葉	生物学的 レプリケート間 の相関
hEB16d_17b_p20	-0.23	-0.49	-0.27	-0.71	-0.35	0.92
hEB16d_17b_p38	-0.19	-0.52	-0.22	-1.14	0.00	0.66
hEB16d_18a_p31	0.36	-0.54	0.33	-0.65	-0.28	0.93
hEB16d_18a_p32	0.61	-0.18	0.63	-0.03	0.40	0.78
hEB16d_18a_p46	-0.26	-0.59	-0.34	-1.02	-0.45	
hEB16d_18b_p20	0.73	-0.54	0.79	-0.24	0.14	0.95
hEB16d_18b_p37	0.74	-0.53	0.71	-0.67	-0.29	1.00
hEB16d_18c_p30	0.89	-0.63	0.90	-0.69	-0.14	0.94
hEB16d_18c_p32	0.78	-0.46	0.87	0.00	-0.01	
hEB16d_20b_p31	-0.02	-0.21	0.04	-0.43	0.40	0.96
hEB16d_20b_p26	0.36	-0.27	0.39	-0.33	0.79	0.72
hEB16d_20b_p50	-0.50	-0.46	-0.59	-1.24	-0.18	0.66
hEB16d_20b_p46	-0.32	-0.63	-0.37	-1.33	-0.78	0.78
hEB16d_27b_p27	0.58	-0.63	0.72	-0.69	-0.62	0.99
hEB16d_27b_p28	0.40	-0.35	0.39	-0.57	-0.51	
hEB16d_27e_p30	-1.01	-0.51	-1.28	-0.70	-1.85	0.99
hEB16d_27e_p32	-1.26	-0.79	-1.73	-1.13	-2.33	0.92
hEB16d_27e_p31	-1.00	-0.83	-1.51	-1.47	-2.36	0.97
hEB16d_27e_p32	-1.11	-0.90	-1.39	-1.72	-2.28	0.99
hEB16d_27e_p35	-1.17	-1.03	-1.60	-1.74	-2.20	
hEB16d_29d_p15	0.04	-0.32	0.17	-0.47	0.34	0.61
hEB16d_29d_p14	-0.24	-0.08	-0.12	-0.18	0.55	
hEB16d_29e_p25	-1.35	-0.80	-1.60	-1.46	-1.46	0.40
hEB16d_29e_p27	-0.57	-0.71	-0.78	-1.15	-1.15	
hFib_11_p7	-1.35	0.14	-1.03	-0.51	-2.16	0.89
hFib_11_p8	-1.58	0.36	-1.51	-0.81	-1.65	
hFib_15_p6	-1.85	0.26	-1.87	-0.64	-2.08	0.95
hFib_15_p7	-2.15	0.10	-2.11	-0.92	-1.63	
hFib_17_p6	-1.60	0.17	-1.56	-0.71	-2.46	0.83
hFib_17_p7	-1.74	0.30	-1.76	-0.51	-1.28	
hFib_18_p6	-1.61	0.60	-1.58	-0.25	-2.37	0.96
hFib_18_p7	-1.32	0.39	-1.25	-0.86	-2.04	
hFib_20_p6	-2.12	0.22	-2.17	-0.74	-2.30	0.98
hFib_20_p7	-1.95	0.16	-1.94	-0.82	-1.68	
hFib_27_p6	-1.75	0.88	-1.81	0.70	-2.57	1.00
hFib_27_p7	-1.74	0.95	-1.87	0.59	-2.68	
hMN_11a_p21	-0.95	-0.49	-1.29	-1.45	-1.58	
hMN_15b_p27	-0.60	-0.84	-1.34	-1.93	-1.36	
hMN_17a_p9	-0.92	-0.49	-1.48	-1.33	-1.80	
hMN_17b_p31	-0.92	-0.82	-1.42	-1.90	-1.53	
hMN_18a_p28	-0.30	-0.78	-0.55	-1.42	-1.50	
hMN_18b_p25	-0.51	-0.71	-0.94	-1.48	-1.39	
hMN_18c_p34	-0.07	-0.57	-0.37	-1.27	-1.28	

10

20

30

40

生物学的 レプリケート	神経系列	造血系列	外胚葉	中胚葉	内胚葉	生物学的 レプリケート間 の相関
hMN_20b_p33	0.08	-0.56	-0.36	-0.28	-1.28	
hMN_27b_p34	-0.92	-0.72	-1.03	-2.16	-1.05	
hES_HUES1_p26	-0.15	-0.31	-0.53	-0.26	-1.59	1.00
hES_HUES1_p26	-0.10	-0.25	-0.49	-0.27	-1.51	
hES_HUES3_p27	-0.69	-0.42	-1.25	-0.59	-1.80	0.91
hES_HUES3_p28	-0.70	-0.44	-1.33	-0.72	-1.26	
hES_HUES6_p19	-0.80	-0.46	-1.27	-0.83	-1.43	0.97
hES_HUES6_p21	-0.58	-0.14	-1.20	-0.52	-1.84	
hES_HUES8_p25	-0.50	0.02	-1.14	-0.22	-0.69	0.88
hES_HUES8_p26	-0.61	0.29	-1.25	0.19	-1.51	
hES_HUES9_p19	-0.94	-0.11	-1.66	-0.38	-1.95	0.93
hES_HUES9_p18	-0.64	-0.47	-1.22	-0.71	-1.19	
hES_HUES28_p13	-0.69	-0.30	-1.49	-0.17	-1.64	0.98
hES_HUES28_p15	-0.53	-0.23	-1.21	-0.13	-1.67	
hES_HUES44_p15	-0.67	-0.34	-1.36	-0.66	-1.41	1.00
hES_HUES44_p16	-0.60	-0.23	-1.31	-0.57	-1.25	
hES_HUES45_p17	-0.06	-0.20	-0.49	-0.24	-0.82	0.99
hES_HUES45_p19	-0.06	-0.28	-0.51	-0.31	-0.83	
hES_HUES48_p16	-0.11	0.56	-0.69	0.42	-1.04	0.99
hES_HUES48_p17	-0.11	0.45	-0.64	0.36	-1.27	
hES_HUES49_p14	-0.67	-0.12	-1.36	-0.37	-1.46	1.00
hES_HUES49_p14	-0.72	-0.17	-1.40	-0.51	-1.43	
hES_HUES53_p17	-0.80	-0.35	-1.20	-0.43	-0.87	0.97
hES_HUES53_p18	-0.57	-0.35	-0.92	-0.35	-0.78	
hES_HUES62_p16	-0.08	0.45	-0.54	0.39	-0.62	0.92
hES_HUES62_p15	-0.57	-0.37	-1.21	-0.58	-1.59	0.66
hES_HUES62_p16	0.72	0.03	0.42	0.28	-1.03	1.00
hES_HUES62_p16	0.78	0.03	0.50	0.28	-0.96	1.00
hES_HUES62_p18	0.70	0.01	0.41	0.28	-0.91	
hES_HUES63_p19	-0.51	-0.15	-1.24	-0.43	-1.54	0.97
hES_HUES63_p17	-0.67	-0.26	-1.43	-0.20	-1.65	
hES_HUES64_p18	-0.09	0.41	-0.56	0.37	-0.61	0.98
hES_HUES64_p20	-0.15	0.54	-0.73	0.38	-1.15	
hES_HUES65_p16	-0.21	0.09	-0.67	0.25	-0.56	0.27
hES_HUES65_p17	0.71	-0.02	0.46	0.30	-1.04	
hES_HUES66_p15	-0.84	-0.32	-1.56	-0.68	-1.58	0.97
hES_HUES66_p15	-0.49	-0.13	-1.21	-0.41	-1.58	
hES_H1_p33	-0.43	-0.22	-0.92	-0.30	-2.29	1.00
hES_H1_p34	-0.57	-0.39	-1.07	-0.52	-2.76	
hES_H9_p57	0.33	-0.01	-0.05	0.45	-1.07	0.99
hES_H9_p58	0.30	0.06	0.00	0.59	-0.98	
hiPS_11a_p14	-0.89	0.32	-1.27	0.41	-2.10	0.77
hiPS_11a_p18	-1.11	-0.24	-1.68	-0.77	-1.25	

10

20

30

40

生物学的 レプリケート	神経系列	造血系列	外胚葉	中胚葉	内胚葉	生物学的 レプリケート間 の相関
hiPS_11b_p15	-0.73	-0.16	-1.19	-0.33	-0.99	0.83
hiPS_11b_p18	-0.92	-0.22	-1.38	-0.66	-2.16	
hiPS_15b_p29	-1.33	-0.55	-1.83	-1.17	-2.89	0.99
hiPS_15b_p30	-1.40	-0.55	-1.92	-1.11	-2.57	
hiPS_17a_p16	-0.65	-0.28	-1.07	-0.27	-1.68	0.74
hiPS_17a_p16	-0.37	0.07	-0.84	0.34	-0.48	
hiPS_17b_p18	-0.78	-0.18	-1.15	-0.20	-1.57	0.92
hiPS_17b_p20	-0.55	-0.42	-0.96	-0.40	-1.85	0.77
hiPS_17b_p38	-0.80	-0.20	-1.37	-0.44	-1.27	
hiPS_18a_p31	-0.40	-0.23	-0.72	-0.35	-1.85	0.29
hiPS_18a_p32	-1.02	-0.49	-1.45	-0.44	-0.89	
hiPS_18b_p20	-1.12	-0.54	-1.56	-0.78	-1.97	0.86
hiPS_18b_p37	-0.17	-0.18	-0.44	0.17	-1.51	
hiPS_18c_p30	-0.18	-0.28	-0.30	-0.28	-1.79	0.78
hiPS_18c_p32	-0.68	-0.04	-1.04	-0.03	-1.70	
hiPS_20b_p31	-0.37	-0.33	-0.62	-0.25	-1.05	0.32
hiPS_20b_p26	-1.19	-0.60	-1.65	-0.69	-0.97	
hiPS_27b_p27	-0.66	-0.16	-1.10	-0.29	-1.62	1.00
hiPS_27b_p28	-0.93	-0.32	-1.35	-0.47	-1.96	
hiPS_27e_p30	-1.04	-0.33	-1.73	-0.51	-2.21	0.98
hiPS_27e_p32	-1.48	-0.46	-2.03	-1.08	-2.71	
hiPS_29d_p15	-0.49	-0.28	-0.75	-0.40	-1.12	0.70
hiPS_29d_p14	-0.58	-0.15	-1.06	-0.45	-0.73	
hiPS_29e_p25	-1.57	-0.90	-2.13	-1.59	-1.74	0.91
hiPS_29e_p27	-1.55	-0.92	-2.08	-1.46	-1.31	

10

20

【 0 5 4 4 】

(表 1 1 B)

30

試料の タイプ	説明	レプリケート間の 相関の平均値
hEB16d	16日目の胚様体	0.82
hFib	ヒト線維芽細胞	0.93
hES	ヒトES細胞株	0.92
hiPS	ヒトiPS細胞株	0.78

【 0 5 4 5 】

多能性細胞の分化指向性に関する本発明者らの「スコアカード」の有用性は、定方向性の分化アッセイにおいて所与の細胞株がどのように機能するかを予測することができれば実質的に増加するであろう。本発明者らは、所定の系列に向かう分化に関して自然の指向性を有する細胞株が、その系列から特定の細胞型を産生することをねらいとした定方向性の分化戦略において良好に機能するか否かを評価した。本発明者らはこれを、ヒトESまたはiPS細胞が定方向性の分化のために用いられる任意の応用のための細胞株選択に、本明細書において開示される「スコアカード」が広い有用性を有するか否かを判定するために評価した。これを評価するために、本発明者らは、頑健な定方向性の分化プロトコールに供した場合に、iPS細胞株の大きいコホートからの各細胞株が運動ニューロンを産生する効率をスコアカードが予測することができるか否かを評価した (Wichterle et al., 2002) (Di Giorgio et al., 2008) (Boulting et al., 同時提出)。

40

50

【0546】

簡単に説明すると、各iPS細胞株を運動ニューロンの定方向性の分化に供し、運動ニューロンの産生効率を、運動ニューロン特異的転写因子ISL1/2およびHB9に対して免疫反応性である細胞の自動定量によってモニターした（同時提出されたBoulting et al.の図6A）。これらの定方向性の分化のデータは、この状況において「スコアカード」の予測力を決定するための真の試験セットを提供した。その発現が単純な転写アッセイによってモニターされる遺伝子の同一性は、2つのデータセット間の第一の比較を行う前に既に確定されており、「スコアカード」のパラメータは、適合を改善するために遡及的に最適化されなかった。「スコアカード」によって作製された所定の細胞株の神経系列分化指向性に関する推定値と、各細胞株が運動ニューロンを産生する実際の効率とを本発明者らが比較したところ、本発明者らは、顕著に高い相関を観察した（図6B）（ISL1に関してピアソン係数 $r = 0.85$ 、HB9に関して $r = 0.86$ ）。この最初の結果は、所与の細胞株の分化指向性の測定を用いて、定方向性の分化のプロトコールにおける多能性幹細胞の挙動を予測することができることを証明する。しかし、「スコアカード」が、任意の種類の細胞への分化に対する細胞株の全体的な反抗性または柔軟性を予測するために限って有用である場合、スコアカードは、その株が運動ニューロンを生成する効率によって決定することができる。

10

【0547】

所与の系列に関するスコアカード予測の特異性を決定するために、本発明者らは、運動ニューロン分化効率を、3つの胚性胚葉の各々への分化指向性に関するスコアカード予測と相関させた（図6Cおよび図11A）。本発明者らは、外胚葉分化指向性の推定と、運動ニューロン産生とのあいだに優れた相関があることを証明した（LSI1に関してピアソン係数 $r = 0.83$ 、HB9に関して $r = 0.82$ ）。対照的に、細胞株が運動ニューロンを産生する効率と、その予想される中胚葉分化指向性（ISL1に関してピアソン係数 $r = 0.48$ 、HB9に関して $R = 0.44$ ）または内胚葉分化指向性（ISL1に関してピアソン係数 $r = 0.23$ 、HB9に関して $R = 0.26$ ）との相関はかなり不良であった。要約すると、本発明者らは、所定の応用に関してiPSまたはES細胞株を最適に選択するために、任意の研究室の当業者が行うことができる迅速なアッセイを明らかに証明した。

20

【0548】

実施例7

多能性細胞の品質および有用性のハイスループット評価に向けて

30

本発明者らは、ヒトES細胞株およびiPS細胞株の品質評価のために用いることができる3つのゲノムアッセイを記述し、低継代ヒトES細胞株において各測定に存在する変動の「参照マップ」を確立することによって、これらのアッセイを校正した。本発明者らは、任意の多能性幹細胞の分化指向性を予測することができることを示す最初の「スコアカード」を設計するために、本明細書において開示されるアッセイを用いることを実例により示した。図7Aに示されるスコアカード出力は、任意の新規ESまたはiPS細胞株におけるエピジェネティックな転写の偏差の数および同一性を要約し、同様に、細胞株の分化指向性の系統的な推定値を提供する。多能性幹細胞株の有用性を増加させ、かつ多能性幹細胞株の特徴を任意の当業者の研究者に手が届くようにするために、本発明者らは、最初のスコアカードの重要な成分を再検討し、アッセイを単純化する機会の特定およびさらなる費用削減を試みた。

40

【0549】

最初に、本発明者らは、3つ全てのアッセイが厳密に必要であるか否かを評価し、またはDNAメチル化、遺伝子発現、もしくは定量的分化アッセイを、スコアカードの精度を損なうことなく省略できるか否かを評価した。本発明者らのデータは明らかに3つのアッセイの重要性を指摘している：いかなる1つのアッセイも、異なるiPS細胞株のその順位がもう1つのアッセイの結果と完全に相関するという意味において余分ではなかった（図7B）。それにも関わらず、非常に感受性の高い少数の遺伝子に向かうDNAメチル化欠損の偏りを利用することによって、DNAメチル化アッセイの費用および複雑度を低減させることが可能であるように思われる（図2A）。本発明者らのデータセットに基づくと、本発明者ら

50

は、ES細胞における最も変動性の上位10%の遺伝子のみをモニターすることによって、iPS細胞株におけるDNAメチル化偏差の80%を検出するのであろう(図7C)。上位約3000個(加えて、まれな欠損に関してもモニターすべきである、さらに約1000個の手動で選択された遺伝子)の最も変動性の遺伝子に集中することにより、プロモーター領域の数は、マイクロアレイコア施設を通して広く入手可能である市販のエピゲノタイピングアッセイ(Bibikova et al., 2009)の範囲内となる。

【0550】

対照的に、遺伝子発現の場合、iPS特異的偏差の完全な範囲をなおもとらえながら少数のES細胞変動遺伝子に集中することは不可能である(図12)。しかし、本発明者らは、実際には制限がないことを証明した。転写をモニターするための市販のマイクロアレイは、当業者に広く利用可能であり、容易に用いることができ、比較的費用効果が高い。

10

【0551】

さらなる測定として、本発明者らは、定量的分化アッセイを行うために必要な時間を低減させることをねらいとした。したがって、アッセイ時間の短縮は、結果を得るまでの時間を短縮させ、同様にインキュベータ空間および培地交換の必要性に関する論理的費用を最小限にすることから利点を有する。本発明者らは、最も多くは、それ以外では自己再生性である培養物において低レベルの細胞分化を検出することにより、未分化多能性幹細胞株から直接単離したRNAを用いて分化指向性を推定するために十分な感度となるよう定量的分化アッセイを最適にした。

【0552】

定量的分化アッセイの期間を短縮する効果を評価するために、本発明者らは、自己再生条件下で各ES細胞株およびiPS細胞株から総RNAを精製して、NanoStringを用いて転写分析を行い、これらのES細胞株およびiPS細胞株に関する新規「スコアカード」を作製した(図7D)。興味深いことに、この新規ES/iPSスコアカードと当初のEBスコアカードとのあいだにはいくつかの限定的な相関が存在し(「r」は0.59から0.82の範囲であった)(図7D)、多能性細胞株自身から発現されたRNAを用いていくつかの妥当な予測を行うことができることを証明している。驚くべきことに、未分化細胞について行った予測のダイナミックレンジは、16日間の分化に供したEBからのRNAを用いて作成されたスコアカードのダイナミックレンジより実質的に低かった。それゆえ、多能性幹細胞株由来のRNAの分析を行うことはできるが、アッセイの頑健性を低減させる可能性がある。代わりに、本発明者らは、EBアッセイの期間を16日から7日まで低減させることができるか否かを評価した。この場合、本発明者らは、4つの代表的なiPS細胞株に関して2つのアッセイ間で優れた一致があることを証明し(ピアソン係数 $r > 0.9$)、その精度を危険にさらすことなく、分化アッセイの期間を低減させることが可能であることを証明した。

20

30

【0553】

実施例8

本発明者らはまた、複数の継代における同じ多能性幹細胞株および無関係な研究室間での同じ多能性幹細胞株を本発明者らが比較した場合に、「スコアカード」からの結果がどれほど頑健で再現可能であるかを調べた。DNAメチル化および転写を分析するための本発明者らの方法は再現可能であることが示されていることから(Gu et al., 2010; Irizarry et al., 2005)、および本発明者らはこれらの測定が継代によってどのように変化するかを既に調べていることから(データは示していない)、本発明者らは、定量的分化アッセイの再現性に焦点を当てた。EBにおけるES細胞の分化は、おそらく物理的取り扱い、培地の交換、およびプラスチックウェアなどのパラメータの差に対して感受性がある可能性があることから、本発明者らは、分化アッセイからの結果が別の研究室での、または別個の研究者による細胞株の挙動をどのように予測するかを評価した。

40

【0554】

それゆえ、本発明者らは、1つの細胞株(hiPS 17b)を、同様にEBアッセイを個別におよび独立して行う異なる2つの研究室の異なる2人の研究者によって2継代培養する系統的比較を行った。系列スコアカード予測の相関は、アッセイを同じ研究者が同じ研究室にお

50

いて行った場合に先に観察された $r=0.82$ より低かった。しかし、本発明者らは、相関が再現可能であると見なされることを証明した($r=0.59$)。それゆえ、最適な細胞株を選択するために、各研究室は、その自身の培養条件下でその自身の株のスコアカードを作成するために、本明細書において記述される複合アッセイを用いるべきであることを本発明者らは推奨する。分化指向性の正確な推定値を維持するために、本発明者らは、核型分析に関する一般的な実践と同様に、細胞株を新たにサブクローニングする際または実質的な継代に供する際にはスコアカードアッセイを繰り返すことを推奨する。

【0555】

実施例9

本明細書において記述される試験において、本発明者らは、多能性細胞株の大きいコホートにおいて観察した変動を調べるためにいくつかのゲノムアッセイを使用して、既存のまたは新しく誘導された株(ES細胞およびiPS細胞)の分類、およびその分化指向性の予測に適用することができるスコアカードを開発した。一般的に観察される変動に関する本発明者らの「参照レベル」および本明細書において開示される「スコアカード」の開発は、ヒト幹細胞の分野におけるいくつかの発展により特に有意義である。

【0556】

最近まで、生物医学研究において広く利用されていたのはごく少数のヒト多能性細胞株であった。この理由から、研究者はたいてい、これらの容易にアクセス可能かつ良好に特徴決定された細胞株に依存してきた(Cowan et al., 2004; Mitalipova et al., 2003; Thomson et al., 1998)。米国におけるヒトES細胞研究に対する資金の制限は、利用可能な細胞株の選択をさらに制限した。その結果、研究者は所定のアッセイにおいて所与の細胞株がどのように良好に挙動を示すかを予測することができる診断の必要がほとんどない、関心対象のその応用のために用いることが可能な任意の株を単に用いていた。

【0557】

しかし、多くの研究室(Chen et al., 2009)によるヒトES細胞株の持続的な誘導および米国における資金制限の解除により、研究者らが選択しうるES細胞株の数は実質的に増加している。さらに、必ずしも全てのヒトES細胞株があらゆる目的にとって等しく適しているわけではないことが明らかとなっている(Osafune et al., 2008)。このことは、いかなる新規研究プロジェクトも、関心対象の応用にとって最も適格である細胞株を十分に考慮して情報を得た上で選択すべきであることを示唆している。

【0558】

患者からの体細胞をiPS細胞へとリプログラムする因子の発見により、研究者らが利用でき、かつ研究者らによって必要とされる多能性細胞株の数にさらに変化が起こっている。研究者らが、既存の細胞株を共に集める際、または関心対象の応用のために新しい細胞株を誘導する際に、最も適切である細胞株をどのように選択するかに関しては情報または手引きはほとんどない。本発明者らは、本明細書において、研究者らが、患者の試料から、完全にリプログラムされたiPS細胞へと、疾患モデリングのために妥当な規模で用いることができる選択された管理可能な細胞株セットへと実験を進めるための手引きとなる、明確な進路を提供する。

【0559】

本明細書において、本発明者らは、それによって研究者らがその所定の応用において最適に機能する株を選択することができる、ヒト多能性細胞株の指向性を正確に予測する方法を明らかにする。重要なことに、多能性幹細胞株の品質および有用性のために本明細書において開示される「スコアカード」を用いることは、任意の数の多能性細胞株、たとえば約5個程度の少ない多能性幹細胞株から10個レベルおよび100個レベルの多能性幹細胞株を特徴決定するために容易にスケール変更することができる。

【0560】

まとめると、本明細書において開示されるスコアカードは、任意の特定の応用において多能性幹細胞を用いるために有意な時間およびリソースをつぎ込む前に、研究者らが理解したいと望む所定の多能性幹細胞株の状態および挙動に関する多くの異なる特徴を報告す

10

20

30

40

50

る。例として、本明細書において開示されるスコアカードは、多能性幹細胞株に関する遺伝子発現プロファイルを組み入れており、それにより研究者らは、研究者らが選択する細胞株が、多能性細胞において通常発現される遺伝子の適切なレベルを転写することを確信することができる(図1)。いくつかの態様において、これらの遺伝子発現プロファイルはまた、関心対象の細胞株が誤って取り扱われていないことを確実にするため、ならびにいくつかの細胞が多能性細胞および分化した細胞の双方の混合集団となるように分化していることを確実にするため、体細胞遺伝子発現シグネチャーを測定するために用いることができる。

【0561】

細胞治療の開発に関心を有する研究者らにとって、臨床開発へと進められる多能性幹細胞株が、調製物毎の「標準的な」基準に適合して、腫瘍抑制遺伝子または腫瘍遺伝子の異常なレベルを発現しないことを証明することは極めて重要でありうる。したがって、本明細書において開示される「スコアカード」を本発明者らが作成して使用することは、治療的使用において対象に多能性幹細胞またはその子孫を投与する前にこれらの重要な安全性測定を行うのに有用である。

【0562】

いくつかの態様において、本発明者らの「スコアカード」はまた、未分化細胞の転写プロファイルにおいて反映されないエピジェネティックな変動を検出するために、DNAメチル化レベルのプロファイリングを含む(図1および2)。本明細書において、本発明者らは、関心対象の所定の株におけるDNAメチル化の特異的測定と共に、この変動を全般的に理解することが、そのエピジェネティックなプロファイルが関心対象の系列への分化を妨害しうる(図2E)細胞株を回避するかもしれないか、もしくはそれらを負の選択により除くために用いられ得ること、または多能性幹細胞株が腫瘍抑制遺伝子もしくは腫瘍遺伝子のいずれかの異常なレベルを発現しないことを示すことを証明した。

【0563】

多能性細胞株の指向性に関する情報をスコアカードに与えるアッセイの1つは、新規な定量的分化アッセイである。この定量的分化アッセイは、不均一なEBにおける各系列からの細胞型の発生率を定量するために計数デバイスとして特定系列において発現された遺伝子の転写測定を用いる。

【0564】

ヒトiPS細胞株およびES細胞株の両方に関して用いられる「スコアカード」を包括的に較正および検証するために、本発明者らは、少なくとも19個のES細胞株および11個のiPS細胞株の転写およびDNAメチル化の全ゲノムレベルに関する「参照」マップを確立した。1つのスコアカードがヒトES細胞およびiPS細胞の両方に対して適切であることを確実にするために、本発明者らは、これらの2つの多能性細胞型において両方の測定の包括的な統計的比較を行った。これらの比較の結果は、本発明者らの「スコアカード」が両方の細胞型に対して高度に適切であることを裏付ける。重要なことに、これらの統計学的結果はまた、定方向性の内胚葉分化アッセイにおける多数のヒトES細胞株の過去の挙動を予測するために、ならびにiPS細胞株の11個が運動ニューロンへと分化することができる効率を高い精度で予測するために「スコアカード」を実行することによって、機能的に確認された(図6および7)。

【0565】

本題から外れるが、本発明者らのデータセットおよび細胞株間でなされた統計的比較により、本発明者らはまた、ES細胞およびiPS細胞株が互いに異なるか否かを評価することが可能となった。これまでの報告(Doi et al., 2009; Stadtfeld et al., 2010b)とは異なり、本明細書において本発明者らが分析した30個の細胞株は、この疑問に対する統計学的情報に富む解答に至るために十分な「数の力」をデータセットに提供した。頑健な統計学的学習アプローチを用いて、本発明者らは、既に公表されたiPS特異的シグネチャーを評価し、無作為より高い精度でこの試験において用いられるES細胞株とiPS細胞株とを区別することができる分類子を誘導した(図3D)。1つの遺伝子座または遺伝子シグネチ

10

20

30

40

50

ヤーが、全てのES細胞株およびiPS細胞株を正確に区別できるわけではないことは本発明者らの分析から明らかであった。言い換えれば、エピジェネティックな転写の差は、平均的なES細胞株を平均的なiPS細胞株と区別することができるが、これらの差は、検討される任意の1つのESまたはiPS細胞株の特徴に関して結論を引き出すには不十分である。言い換えれば、本発明者らは、いくつかのES細胞株が所定の応用に関して他より適しているが、同じことがiPS細胞株に当てはまることを突き止めた。これらの研究の結果、本発明者らは、現在のリプログラミング法が驚くべきことに頑健であると判断した。

【0566】

本発明者らはまた、必要であるように思われるのは、全ての必要性および応用にとって最適なES細胞株または完全なリプログラミングプロトコルを発見しようと試みるのではなく、適切な細胞株を所定の応用にマッチさせることができる迅速なアッセイであることを明らかにした。したがって、本明細書において開示される方法、システム、および「スコアカード」は、ヒト多能性細胞株の指向性を決定および予測するために有用であり、それにより所望の指向性を有する適切な多能性幹細胞を、特定の下流の応用において用いるためにマッチさせて選択することができるであろう。

10

【0567】

本発明者らは、多能性細胞に関する「スコアカード」を本明細書において実証したが、本発明者らはまた、多能性幹細胞のエピジェネティックな転写の調節に対する貴重な生物学的洞察源を提供する多能性のエピゲノムおよびトランスクリプトームの「参照マップ」を実証している。たとえば、本発明者らは、ES細胞株におけるエピジェネティックな変動が、ゲノム領域をDNAメチル化に対して感受性にすることが既に示されているDNA配列モチーフと高度に相関することを証明した (Bock et al., 2006; Keshet et al., 2006; Meissner et al., 2008)。

20

【0568】

驚くべきことに、本発明者らは、最も転写的に変動性の遺伝子のクラスにおいて細胞シグナル伝達において機能する遺伝子の遺伝子発現が顕著に濃縮されたことを証明した。これは、各多能性細胞株が、インビトロ培養の選択圧に対して多様な様式で適合しうることを証明した。したがって、このデータに基づき、ES細胞株はさらに、細胞の競合および成長条件に対するエピジェネティックな適応に関する問題を調べるためのモデルシステムを提供するために有用である。最後に、本発明者らはまた、いくつかの多能性幹細胞株が、CD14プロモーターで多様なメチル化レベルを有することを証明し、プロモーターの高度メチル化が多能性幹細胞株に存在する発達経路における重要な遺伝子をサイレンシングさせる手段であり、ヒト胚発生の際の「ゲートキーパー遺伝子」(Hemberger et al., 2009)のエピジェネティックな調節に対するさらなる洞察を決定するために発達研究にとって有用であることを証明している。

30

【0569】

要約すると、本発明者らは、多くのヒト多能性細胞株においてDNAメチル化、転写、および分化指向性を分析および測定し、この結果、任意の新規なまたは既存の多能性細胞株の挙動を予測するための「スコアカード」を作成するために当業者の任意の研究者が利用することができる単純なシステム、方法およびアッセイを開発した(図7E)。現在、本発明がなければ、研究者らは、既存の多能性幹細胞株を得た後または新しい株を生成した後、特異的抗原による免疫染色および奇形腫生成を含む時間のかかる、労力のかかる、および費用がかかる多数のアッセイを行うであろう。これらのアッセイは、所定の細胞株が多能性であるという何らかの確証を与える可能性があるが、それらは多能性細胞株が所定の応用にとって十分に適しているか否かを予測することはできない。対照的に、本明細書において開示される本発明の方法、キット、システム、アッセイ、およびスコアカードは、多能性幹細胞の挙動を、迅速に、効率的に、および有効に予測するために有用であり、時間または労力がかからず、比較的安価である。

40

【0570】

したがって、本明細書において開示される方法、キット、システム、アッセイ、および

50

スコアカードを用いて、たとえば筋萎縮性側索硬化症（ALS）の疾患モデリングが関心対象である研究者らは、関心対象の多能性幹細胞を分析して、本明細書において開示される定量的分化アッセイを行うことができるであろう（図5D）。次に、研究者らは、将来の研究のために、神経系列に関して正常～高い分化指向性を示す多能性幹細胞株を選択することができる。次に選択された多能性細胞株を、DNAメチル化分析および/または転写プロファイリングに供することができる。したがって、本明細書において開示される方法、システム、およびスコアカードを用いることによって、研究者はその特定の所望の応用における多能性幹細胞株の有用性を最も良好に予測するパラメータの変動に関して細胞株を調べることができる（図7E）。

【0571】

本明細書において開示される本発明者らの方法、アッセイ、スコアカードおよびキットは、選択された多能性細胞株が運動ニューロンまたは関心対象の他の細胞へと高い効率で分化する可能性があり、かつ他の重篤な制限を示さない（たとえば、腫瘍遺伝子の発現または腫瘍抑制遺伝子の抑制等）ことを「スコアカード」が予測する場合に限り、研究者らが、最も時間のかかり費用のかかるアッセイである奇形腫形成を特定の多能性幹細胞株において開始するのを遅らせることができる。いずれ、本明細書において開示される方法、アッセイ、スコアカードが奇形腫形成能を有する多能性幹細胞株を正確に予測するために用いられれば、本明細書において開示される方法、アッセイ、スコアカード、およびキットを用いることにより、奇形腫形成アッセイを完全に消失させることができるであろう。

【0572】

結論すると、ヒト多能性細胞の発見および選択された患者集団からヒトiPS細胞を産生するリプログラミング法は、ヒト疾患の研究および処置に関する研究者らの考え方に対して革命をもたらした。しかし、ヒト多能性幹細胞およびiPS細胞の使用が、研究において、ならびに患者の生命を改善するために細胞治療および治療的使用において効率よく有効に用いられることである場合、新薬開発および毒性アッセイに関して試験する際の、ならびに特定の治療的応用の際の、または所定の適応もしくは疾患を処置する際の多能性細胞株の選択を能率的にし、標準化し、および最適化するために、本明細書において開示される方法、アッセイ、システム、および「スコアカード」などの品質評価および検証法を確立することが緊急に必要である。

【0573】

参考文献

参考文献は、その全文が参照により本明細書に組み入れられる。

10

20

30

Adewumi, O., Aflatoonian, B., Ahrlund-Richter, L., Amit, M., Andrews, P.W., Beighton, G., Bello, P.A., Benvenisty, N., Berry, L.S., Bevan, S., *et al.* (2007). Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. *Nat Biotechnol* 25, 803-816

Allison, D.B., Cui, X., Page, G.P., and Sabripour, M. (2006). Microarray data analysis: from disarray to consolidation and consensus. *Nat Rev Genet* 7, 55-65.

Bibikova, M., Le, J., Barnes, B., Saedinia-Melnyk, S., Zhou, L., Shen, R., and Gunderson, K.L. (2009). Genome-wide DNA methylation profiling using Infinium assay. *Epigenomics* 1, 177-200.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev* 16, 6-21.

Bock, C., Halachev, K., Büch, J., and Lengauer, T. (2009). EpiGRAPH: User-friendly software for statistical analysis and prediction of (epi-) genomic data. *Genome Biol* 10, R14.

Bock, C., Paulsen, M., Tierling, S., Mikeska, T., Lengauer, T., and Walter, J. (2006). CpG island methylation in human lymphocytes is highly correlated with DNA sequence, repeats, and predicted DNA structure. *PLoS Genet* 2, e26.

Borowiak, M., Maehr, R., Chen, S., Chen, A.E., Tang, W., Fox, J.L., Schreiber, S.L., and Melton, D.A. (2009). Small molecules efficiently direct endodermal differentiation of mouse and human embryonic stem cells. *Cell Stem Cell* 4, 348-358.

Carvajal-Vergara, X., Sevilla, A., D'Souza, S.L., Ang, Y.S., Schaniel, C., Lee, D.F., Yang, L., Kaplan, A.D., Adler, E.D., Rozov, R., *et al.* (2010). Patient-specific induced pluripotent stem-cell-derived models of LEOPARD syndrome. *Nature* 465, 808-812.

Chen, A.E., Egli, D., Niakan, K., Deng, J., Akutsu, H., Yamaki, M., Cowan, C., Fitz-Gerald, C., Zhang, K., Melton, D.A., *et al.* (2009). Optimal timing of inner cell mass isolation increases the efficiency of human embryonic stem cell derivation and allows generation of sibling cell lines. *Cell stem cell* 4, 103-106.

10

20

30

Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., Ambartsumyan, G., Aimiwu, O., Richter, L., Zhang, J., *et al.* (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* 5, 111-123.

Colman, A., and Dreesen, O. (2009). Pluripotent stem cells and disease modeling. *Cell Stem Cell* 5, 244-247. Cowan, C.A., Klimanskaya, I., McMahon, J., Atienza, J., Witmyer, J., Zucker, J.P., Wang, S., Morton, C.C., McMahon, A.P., Powers, D., *et al.* (2004). Derivation of embryonic stem-cell lines from human blastocysts. *N Engl J Med* 350, 1353-1356.

Daley, G. (2010). Straight talk with...George Daley. Interview by Elie Dolgin. *Nat Med* 16, 624. 10

Di Giorgio, F.P., Boulting, G.L., Bobrowicz, S., and Eggan, K.C. (2008). Human embryonic stem cell-derived motor neurons are sensitive to the toxic effect of glial cells carrying an ALS-causing mutation. *Cell Stem Cell* 3, 637-648.

Dimos, J.T., Rodolfa, K.T., Niakan, K.K., Weisenthal, L.M., Mitsumoto, H., Chung, W., Croft, G.F., Saphier, G., Leibel, R., Golland, R., *et al.* (2008). Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. *Science* 321, 1218-1221. 20

Doi, A., Park, I.H., Wen, B., Murakami, P., Aryee, M.J., Irizarry, R., Herb, B., Ladd-Acosta, C., Rho, J., Loewer, S., *et al.* (2009). Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat Genet.*

Ebert, A.D., Yu, J., Rose, F.F., Jr., Mattis, V.B., Lorson, C.L., Thomson, J.A., and Svendsen, C.N. (2009). Induced pluripotent stem cells from a spinal muscular atrophy patient. *Nature* 457, 277-280.

Eiges, R., Urbach, A., Malcov, M., Frumkin, T., Schwartz, T., Amit, A., Yaron, Y., Eden, A., Yanuka, O., Benvenisty, N., *et al.* (2007). Developmental study of fragile X syndrome using human embryonic stem cells derived from preimplantation genetically diagnosed embryos. *Cell Stem Cell* 1, 568-577. 30

ENCODE Project Consortium (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447, 799-816.

Geiss, G.K., Bumgarner, R.E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D.L., Fell, H.P., Ferree, S., George, R.D., Grogan, T., *et al.* (2008). Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nature Biotechnology* 26, 317-325. 40

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., *et al.* (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5, R80.

Gu, H., Bock, C., Mikkelsen, T.S., Jager, N., Smith, Z.D., Tomazou, E., Gnirke, A., Lander, E.S., and Meissner, (2010). Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. *Nat Methods* 7, 133-136.

Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., Cassady, J.P., Muffat, J., Carey, B.W., and Jaenisch, R. (2010). Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proc Natl Acad Sci U S A* *107*, 9222-9227.

Hastie, T., Tibshirani, R., and Friedman, J.H. (2001). *The elements of statistical learning : data mining, inference, and prediction* (New York, Springer).

Hawkins, R.D., Hon, G.C., Lee, L.K., Ngo, Q., Lister, R., Pelizzola, M., Edsall, L.E., Kuan, S., Luu, Y., Klugman, S., *et al.* (2010). Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* *6*, 479-491. 10

Hemberger, M., Dean, W., and Reik, W. (2009). Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. *Nature Reviews Molecular Cell Biology* *10*, 526-537.

Hu, B.Y., Weick, J.P., Yu, J., Ma, L.X., Zhang, X.Q., Thomson, J.A., and Zhang, S.C. (2010). Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. *Proc Natl Acad Sci U S A* *107*, 4335-4340.

Huang, D.W., Sherman, B.T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M.W., Lane, H.C., *et al.* (2007). DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res* *35*, W169-175. 20

Hubbard, T.J., Aken, B.L., Ayling, S., Ballester, B., Beal, K., Bragin, E., Brent, S., Chen, Y., Clapham, P., Clarke, L., *et al.* (2009). Ensembl 2009. *Nucleic Acids Res* *37*, D690-697.

Huber, W., von Heydebreck, A., Sultmann, H., Poustka, A., and Vingron, M. (2002). Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* *18 Suppl 1*, S96-104.

Irizarry, R.A., Warren, D., Spencer, F., Kim, I.F., Biswal, S., Frank, B.C., Gabrielson, E., Garcia, J.G., Geoghegan, J., Germino, G., *et al.* (2005). Multiple-laboratory comparison of microarray platforms. *Nature Methods* *2*, 345-350. 30

Kauffmann, A., Gentleman, R., and Huber, W. (2009). arrayQualityMetrics--a bioconductor package for quality assessment of microarray data. *Bioinformatics* *25*, 415-416.

Keshet, I., Schlesinger, Y., Farkash, S., Rand, E., Hecht, M., Segal, E., Pikarski, E., Young, R.A., Niveleau, A., Cedar, H., *et al.* (2006). Evidence for an instructive mechanism of de novo methylation in cancer cells. *Nat Genet* *38*, 149-153.

Laird, P.W. (2010). Principles and challenges of genome-wide DNA methylation analysis. *Nat Rev Genet* *11*, 191-203. 40

Lee, G., Papapetrou, E.P., Kim, H., Chambers, S.M., Tomishima, M.J., Fasano, C.A., Ganat, Y.M., Menon, J., Shimizu, F., Viale, A., *et al.* (2009). Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. *Nature*.

Lengner, C.J., Gimelbrant, A.A., Erwin, J.A., Cheng, A.W., Guenther, M.G., Welstead, G.G., Alagappan, R., Frampton, G.M., Xu, P., Muffat, J., *et al.* (2010). Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. *Cell* *141*, 872-883.

Li, H., Ruan, J., and Durbin, R. (2008). Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res* 18, 1851-1858.

Lister, R., Pelizzola, M., Downen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.M., *et al.* (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462, 315-322.

Liu, L., Luo, G.Z., Yang, W., Zhao, X., Zheng, Q., Lv, Z., Li, W., Wu, H.J., Wang, L., Wang, X.J., *et al.* (2010). Activation of the imprinted Dlk1-Dio3 region correlates with pluripotency levels of mouse stem cells. *J Biol Chem* 285, 19483-19490. 10

Lu, R., Markowitz, F., Unwin, R.D., Leek, J.T., Airoidi, E.M., MacArthur, B.D., Lachmann, A., Rozov, R., Ma'ayan, A., Boyer, L.A., *et al.* (2009). Systems-level dynamic analyses of fate change in murine embryonic stem cells. *Nature* 462, 358-362.

Maherali, N., and Hochedlinger, K. (2008). Guidelines and techniques for the generation of induced pluripotent stem cells. *Cell Stem Cell* 3, 595-605.

Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., *et al.* (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454, 766-770. 20

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. *Nature* 454, 49-55.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., *et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448, 553-560.

Mitalipova, M., Calhoun, J., Shin, S., Winingar, D., Schulz, T., Noggle, S., Venable, A., Lyons, I., Robins, A., and Stice, S. (2003). Human embryonic stem cell lines derived from discarded embryos. *Stem Cells* 21, 521-526. 30

Müller, F.J., Laurent, L.C., Kostka, D., Ulitsky, I., Williams, R., Lu, C., Park, I.H., Rao, M.S., Shamir, R., Schwartz, P.H., *et al.* (2008). Regulatory networks define phenotypic classes of human stem cell lines. *Nature* 455, 401-405.

Nam, D., and Kim, S.Y. (2008). Gene-set approach for expression pattern analysis. *Briefings in Bioinformatics* 9, 189-197. 40

Narva, E., Autio, R., Rahkonen, N., Kong, L., Harrison, N., Kitsberg, D., Borghese, L., Itskovitz-Eldor, J., Rasool, O., Dvorak, P., *et al.* (2010). High-resolution DNA analysis of human embryonic stem cell lines reveals culture-induced copy number changes and loss of heterozygosity. *Nat Biotechnol.*

Osafune, K., Caron, L., Borowiak, M., Martinez, R.J., Fitz-Gerald, C.S., Sato, Y., Cowan, C.A., Chien, K.R., and Melton, D.A. (2008). Marked differences in differentiation propensity among human embryonic stem cell lines. *Nat Biotechnol* 26, 313-315.

Park, I.H., Arora, N., Huo, H., Maherali, N., Ahfeldt, T., Shimamura, A., Lensch, M.W., Cowan, C., Hochedlinger, K., and Daley, G.Q. (2008a). Disease-specific induced pluripotent stem cells. *Cell* 134, 877-886.

Park, I.H., Zhao, R., West, J.A., Yabuuchi, A., Huo, H., Ince, T.A., Lerou, P.H., Lensch, M.W., and Daley, G.Q. (2008b). Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* 451, 141-146.

Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* 447, 425-432.

10

Rossant, J. (2008). Stem cells and early lineage development. *Cell* 132, 527-531.

Smith, Z.D., Gu, H., Bock, C., Gnirke, A., and Meissner, A. (2009). High-throughput bisulfite sequencing in mammalian genomes. *Methods* 48, 226-232.

Smyth, G.K. (2005). Limma: linear models for microarray data. In *Bioinformatics and Computational Biology Solutions using R and Bioconductor*, R. Gentleman, V. Carey, S. Dudoit, R. Irizarry, and W. Huber, eds. (New York, Springer), pp. 397-420.

Stadtfeld, M., Apostolou, E., Akutsu, H., Fukuda, A., Follett, P., Natesan, S., Kono, T., Shioda, T., and Hochedlinger, K. (2010a). Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature*.

20

Stadtfeld, M., Apostolou, E., Akutsu, H., Fukuda, A., Follett, P., Natesan, S., Kono, T., Shioda, T., and Hochedlinger, K. (2010b). Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* 465, 175-181.

Storey, J.D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A* 100, 9440-9445.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., *et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* 102, 15545-15550.

30

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861-872.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663-676.

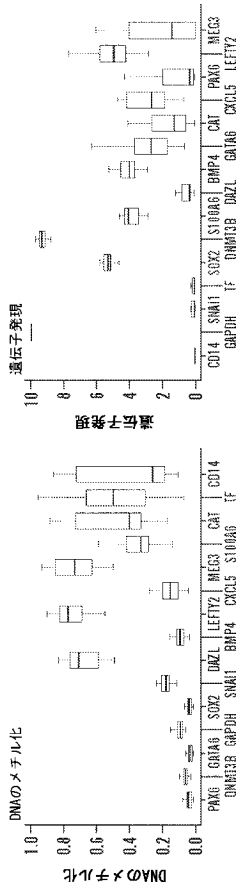
Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science* 282, 1145-1147.

40

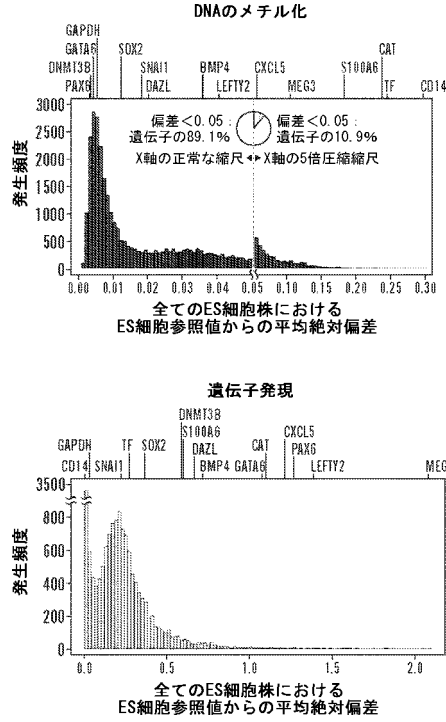
Wichterle, H., Lieberam, I., Porter, J.A., and Jessell, T.M. (2002). Directed differentiation of embryonic stem cells into motor neurons. *Cell* 110, 385-397.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., *et al.* (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318, 1917-1920.

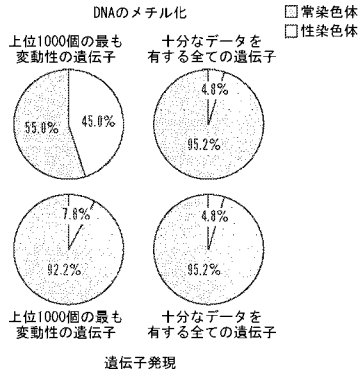
【図 1 C】



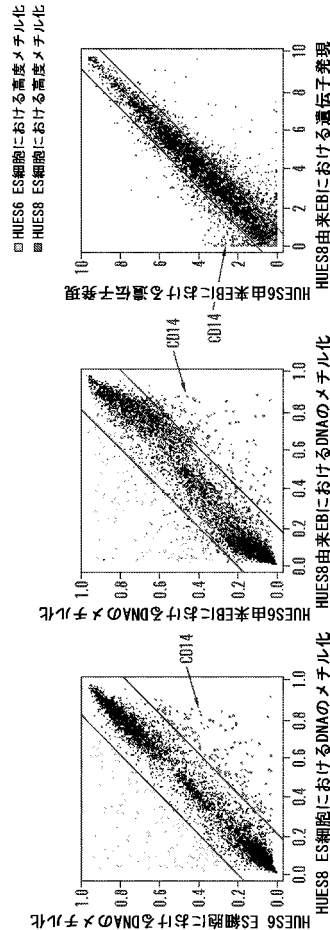
【図 2 A】



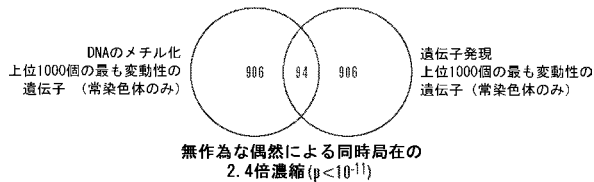
【図 2 B】



【図 2 E】



【図 2 C】

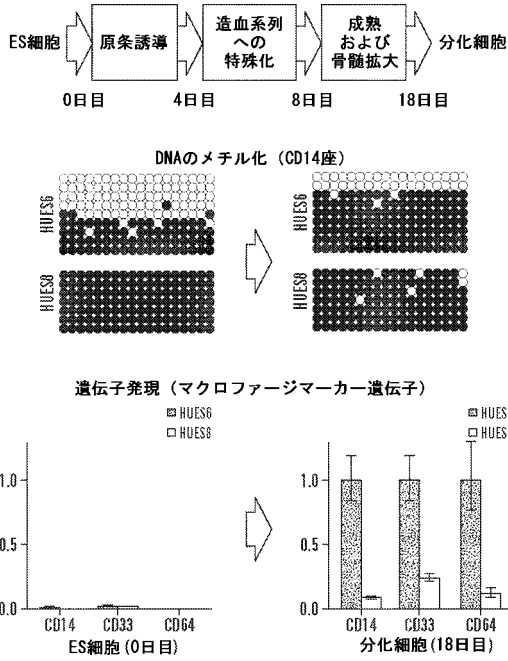


【図 2 D】

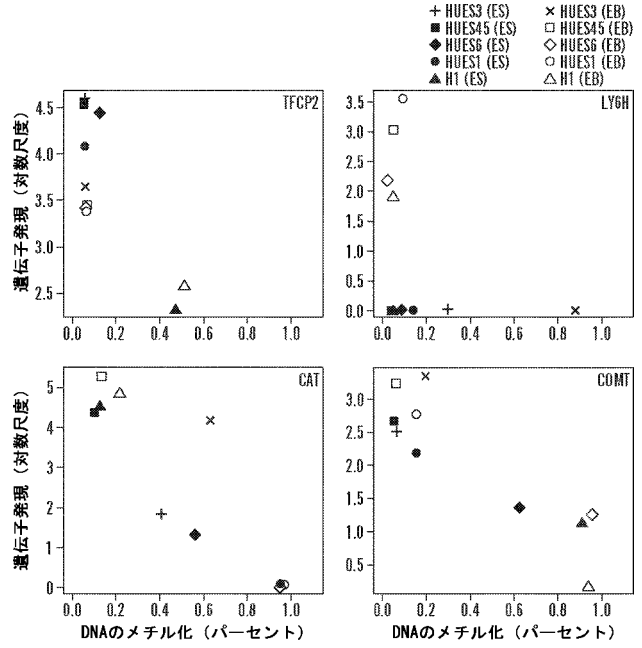
上位1000個の最も変動性の遺伝子 (常染色体のみ)

	DNAのメチル化		遺伝子発現
機能的アテネーションクラスターリング (DAVID)	強い	なし	シグナル伝達、糖タンパク質、細胞外マトリクス
	中程度	ジンクフィンガータンパク質	発達、刺激に対する反応
プロモーターの特徴 (epiGRAPH)	濃縮	CpGプリアリピート (L1, L1R)	抑制性クロマチンマーク (H3K27me3, H3K9me3)
	減少	CpGリッチ配列パターン、CpGアイランド、CpGリッチリピート	活性クロマチンマーク (H3K4me3, POUH 結合)

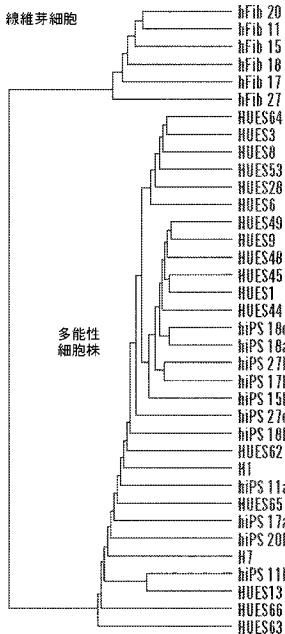
【図 2 F】



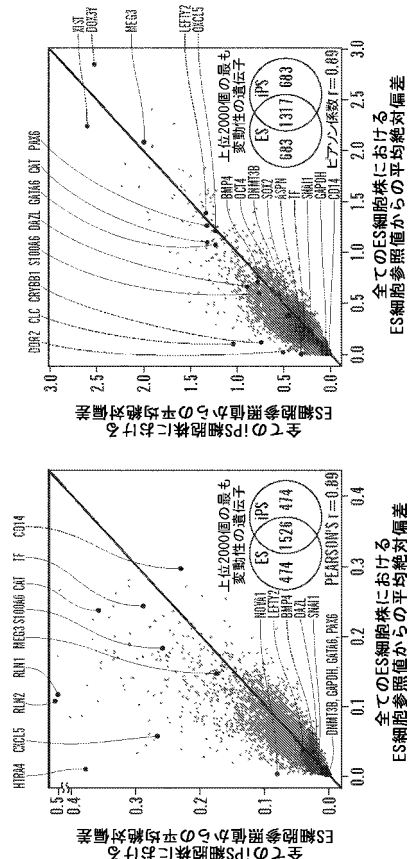
【図 2 G】



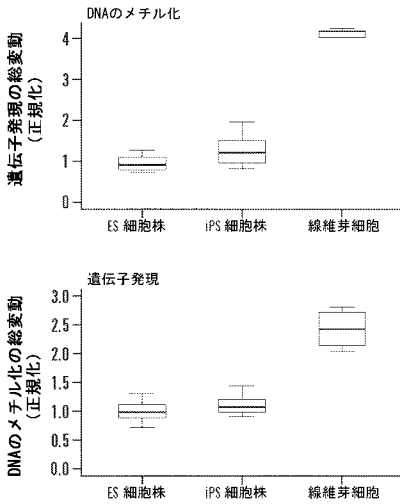
【図 3 A】



【図 3 B】



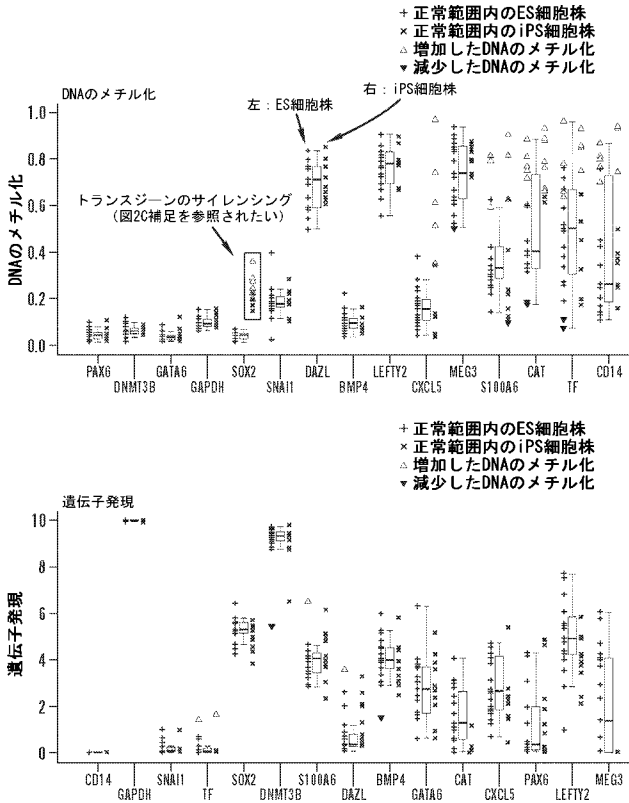
【 図 3 C 】



【 図 3 D 】

IPSC細胞分類子	精度	感度	特異性	TN (ES)	FN	FP	IP
公認されたiPS細胞シグネチャー Chin2009 遺伝子発現シグネチャー Doi2009 DNAメチル化シグネチャー Stadler0210 単遺伝子シグネチャー (MEG3)	63%	0%	100%	20	12	0	0
	63%	0%	100%	20	12	0	0
	72%	100%	55%	11	0	9	12
	78%	67%	86%	17	4	3	8
交差検証されたES対iPS分類子 ES細胞株対iPS細胞株 (DNAメチル化) ES細胞株対iPS細胞株 (遺伝子発現) ES細胞株対iPS細胞株 (両方のデータタイプ)	69%	48%	81%	16	6	4	6
	81%	64%	91%	18	4	2	8
	100%	100%	100%	20	0	0	6
陰性対照 (ES対繊維芽細胞分類子) ES細胞株対繊維芽細胞 (DNAメチル化) ES細胞株対繊維芽細胞 (遺伝子発現) ES細胞株対繊維芽細胞 (両方のデータタイプ)	94%	73%	100%	20	2	0	4
	99%	95%	100%	20	0	0	6
	38%	100%	0%	0	0	20	12
陰性対照 (トリリアルな分類子) あらゆる事象をES細胞株として予測する いかなる事象もES細胞株として予測しない コインをばく (ES細胞株として予測する可能性は50%)	63%	0%	100%	20	12	0	0
	50%	50%	50%	10	6	10	6

【 図 4 A 】



【 図 4 B 】

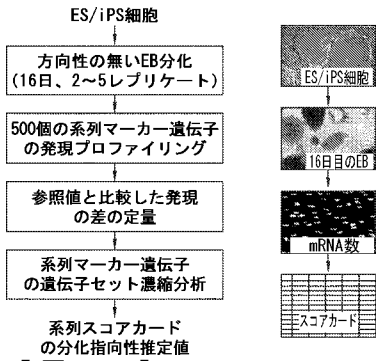
細胞株	DNAのメチル化		影響を受けた系列マーカー遺伝子の数
	メチル化が増加した遺伝子の数	メチル化が減少した遺伝子の数	
hiPS 11a	39	60	5
hiPS 11b	63	142	6
hiPS 11c	80	36	7
hiPS 15b	51	41	2
hiPS 17a	260	14	11
hiPS 17b	368	11	19
hiPS 18a	75	50	0
hiPS 18b	158	69	10
hiPS 18c	73	35	2
hiPS 20b	384	35	7
hiPS 27b	219	50	9
hiPS 27e	274	62	9
12個のiPS細胞株の平均値	181	54	7
20個のES細胞株の平均値	61	70	4

□高い □中程度 ■低い

細胞株	DNAのメチル化		影響を受けた系列マーカー遺伝子の数
	メチル化が増加した遺伝子の数	メチル化が減少した遺伝子の数	
hiPS 11a	468	452	40
hiPS 11b	970	596	123
hiPS 11c	280	418	36
hiPS 15b	236	211	22
hiPS 17a	310	457	44
hiPS 17b	279	57	11
hiPS 18a	74	71	9
hiPS 18b	288	171	38
hiPS 18c	168	340	17
hiPS 20b	164	170	26
hiPS 27b	249	97	28
hiPS 27e	184	149	20
12個のiPS細胞株の平均値	293	257	35
20個のES細胞株の平均値	244	143	28

□高い □中程度 ■低い

【図5A】

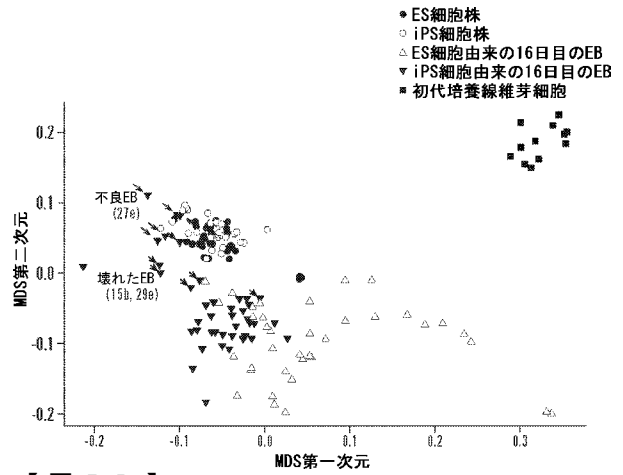


【図5B】

細胞株	神経系列	造血系列	外胚葉	中胚葉	内胚葉
HUES1	1.04	-0.30	-1.56	0.06	-0.99
HUES3	-0.29	-0.01	-0.23	-0.07	0.08
HUES6	-0.78	-0.26	-0.51	-0.05	-0.47
HUES8	-0.15	0.69	-0.17	0.68	
HUES9	-0.89	0.31	-0.75	0.51	0.37
HUES28	1.33	-0.11		1.03	-0.07
HUES44	0.70	-0.27	0.52	-0.48	-0.45
HUES45	-0.46	-0.26	-0.49	-0.02	0.65
HUES48	0.83	0.18	0.70	0.24	-0.55
HUES49	0.19	0.07	0.03	-0.56	-0.26
HUES53		0.65		-0.22	-0.20
HUES62	0.25	-0.15	0.15	-0.50	0.24
HUES63	0.62	0.39	0.72	0.34	0.61
HUES64		-0.07		-0.56	-0.61
HUES65	0.19	0.02	0.22	0.19	-0.15
HUES66	0.59	-0.67	0.36		-0.37
H1	1.88	-0.29		0.07	-0.56
H9		0.01		0.55	-0.16

分化指向性: ■高い □中程度 ▨低い

【図5C】

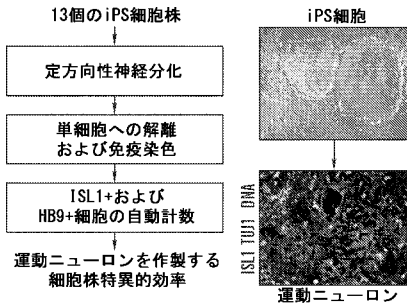


【図5D】

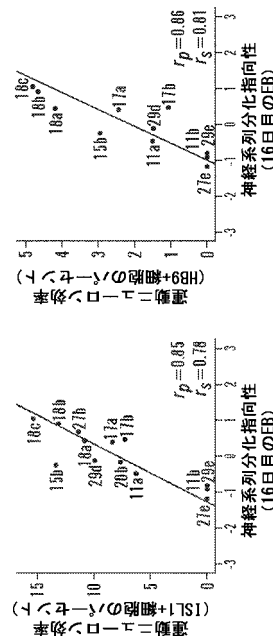
細胞株	神経系列	造血系列	外胚葉	中胚葉	内胚葉
hiPS 11a	-0.69	0.18	-0.37	-0.23	0.63
hiPS 11b		-0.23	-0.66	-1.03	0.47
hiPS 11c	-0.22	0.40	-0.03	-0.16	0.37
hiPS 15b	-0.48	-0.78	-0.63	-1.11	-2.48
hiPS 17a	0.19	0.05	0.33	0.00	1.16
hiPS 17b	-0.07	-0.48	-0.02	-0.63	0.20
hiPS 18a	0.28	-0.52	0.31	-0.67	0.20
hiPS 18b	0.80	-0.72	0.84	-0.62	0.15
hiPS 18c	0.93	-0.65	1.05	-0.41	0.10
hiPS 20b	-0.37	-0.47	-0.30		0.56
hiPS 27b	0.52	-0.50	0.68	-0.71	-0.42
hiPS 27e	-1.61	-1.04	-2.12	-1.62	-3.27
hiPS 29d	-0.25	-0.04	0.00	-0.11	0.83
hiPS 29e	-0.99	-0.60			-1.04

分化指向性: ■高い □中程度 ▨低い

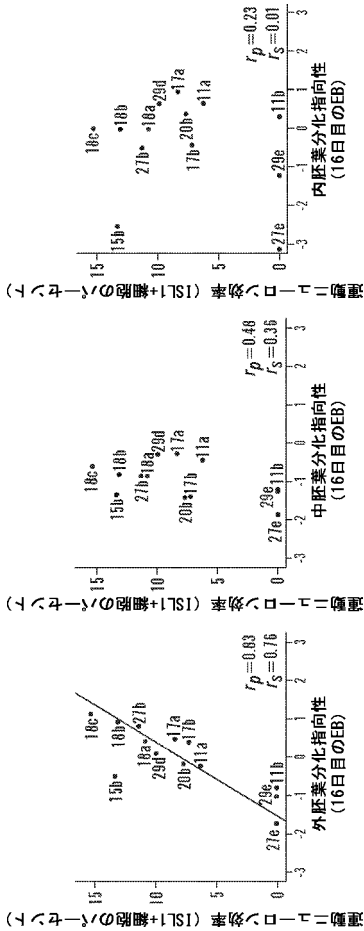
【図6A】



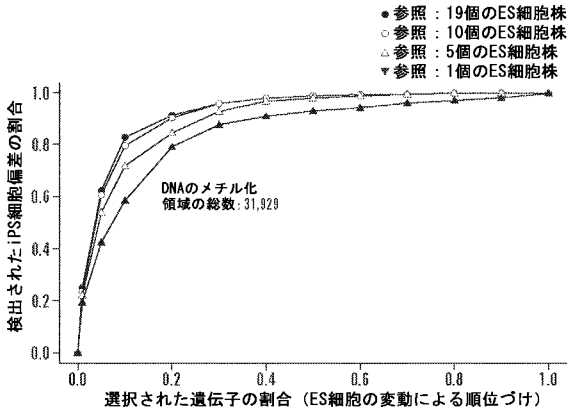
【図6B】



【図6C】



【図7C】



【図7A】

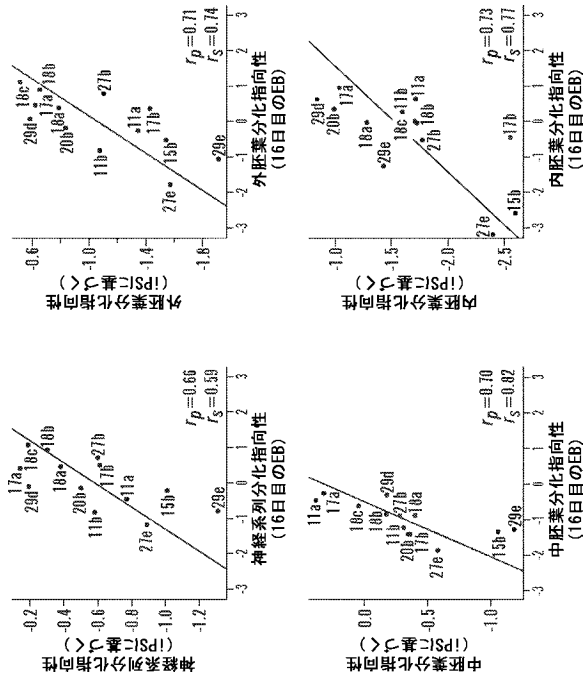
細胞株	外れている遺伝子の数 (メチル化)	外れている遺伝子の数 (発現)	外胚葉指向性	中胚葉指向性	内胚葉指向性	運動ニューロン効率
hiPS 11a	169	416	-0.24	-0.44	0.62	6.23
hiPS 11b	162	519	-0.79	-1.20	0.28	0.00
hiPS 15b	127	82	-0.49	-1.32	0.55	13.29
hiPS 17a	479	340	0.47	-0.26	0.93	8.32
hiPS 17b	514	3	0.38	-1.38	-0.43	7.17
hiPS 18a	212	3	0.40	-0.87	-0.01	10.73
hiPS 18b	336	38	0.91	-0.82	-0.03	13.04
hiPS 18c	400	160	1.12	-0.59	-0.02	15.26
hiPS 20b	458	699	-0.17	-1.40	0.37	7.61
hiPS 27b	323	21	0.80	-0.85	-0.52	11.27
hiPS 27e	563	28	-1.74	-1.05	0.15	0.00

【図7B】

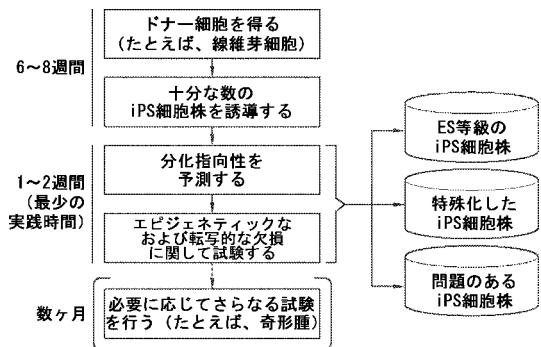
ペアワイズ相関	外れている遺伝子の数 (メチル化)	外れている遺伝子の数 (発現)	外胚葉指向性	中胚葉指向性	内胚葉指向性	運動ニューロン効率
外れている遺伝子の数 (メチル化)	1.00	-1.10	0.01	-0.26	-0.10	-0.17
外れている遺伝子の数 (発現)	-0.10	1.00	-0.22	0.10	0.51	-0.36
外胚葉指向性	0.01	-0.22	1.00	0.68	0.50	0.80
中胚葉指向性	-0.26	0.10	0.68	1.00	0.72	0.47
内胚葉指向性	-0.10	0.51	0.58	0.72	1.00	0.14
運動ニューロン効率	-0.17	-0.36	0.80	0.47	0.14	1.00

ピアソン相関係数: □高い □中程度 □低い

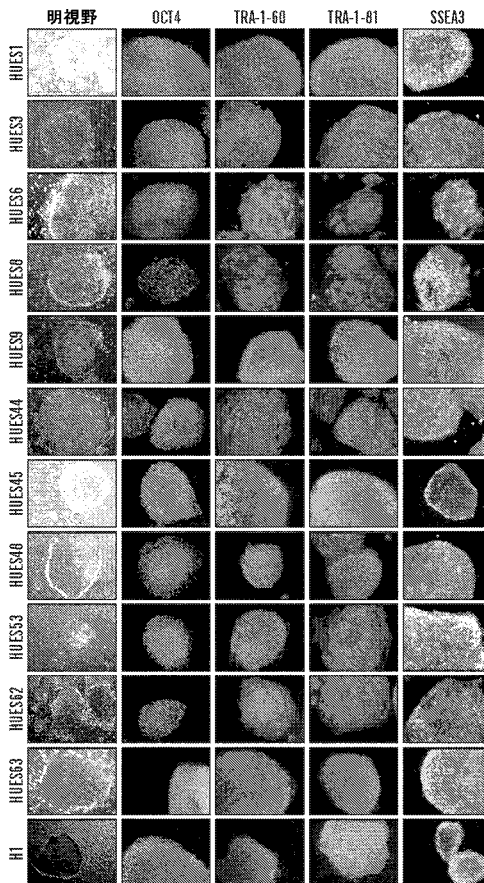
【図7D】



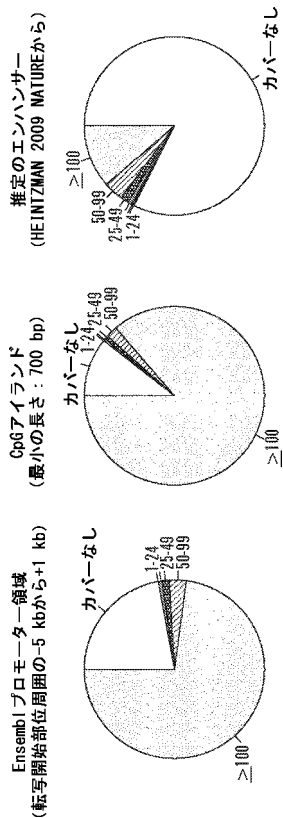
【 図 7 E 】



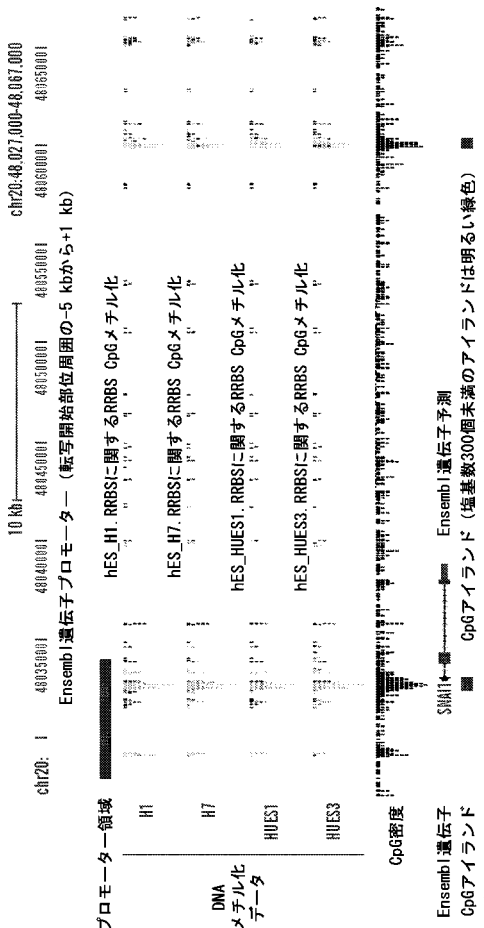
【 図 8 A 】



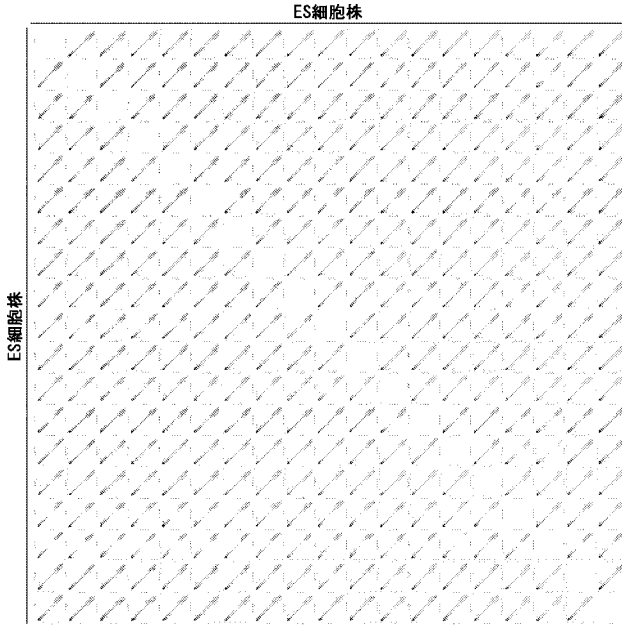
【 図 8 B 】



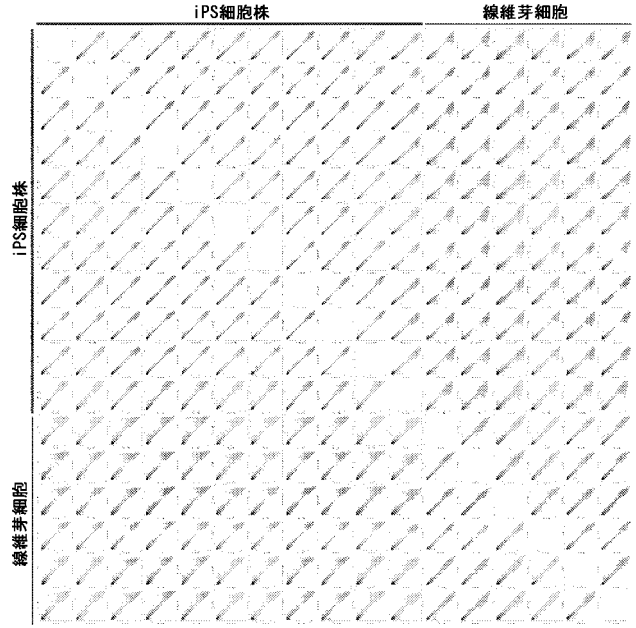
【 図 8 C 】



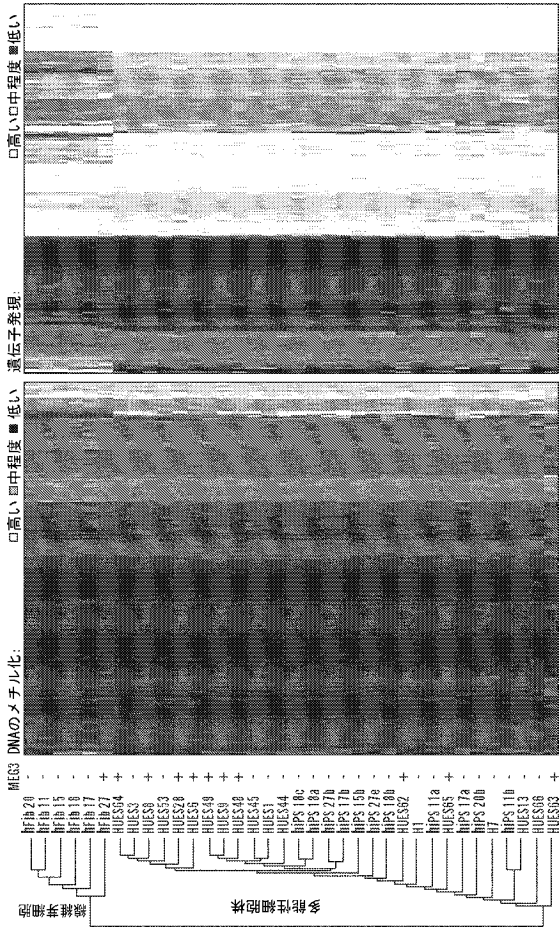
【 図 8 D 】



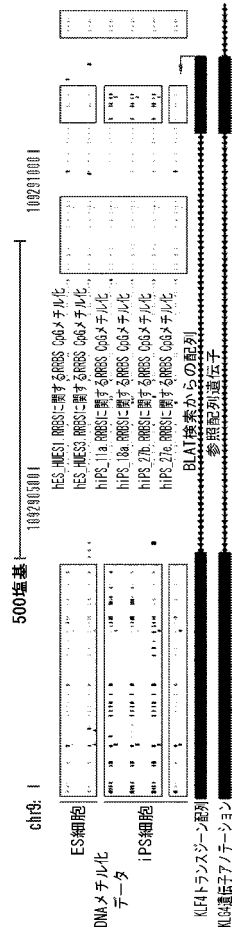
【 図 9 A 】



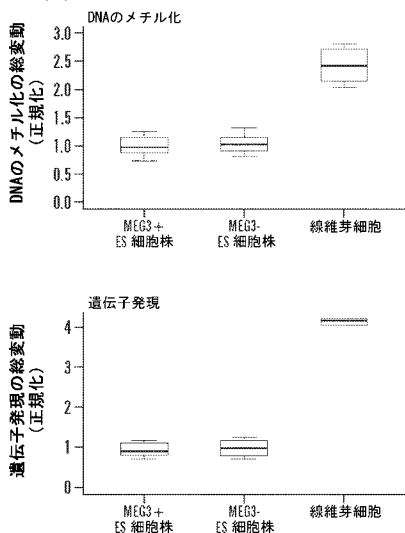
【 図 9 B 】



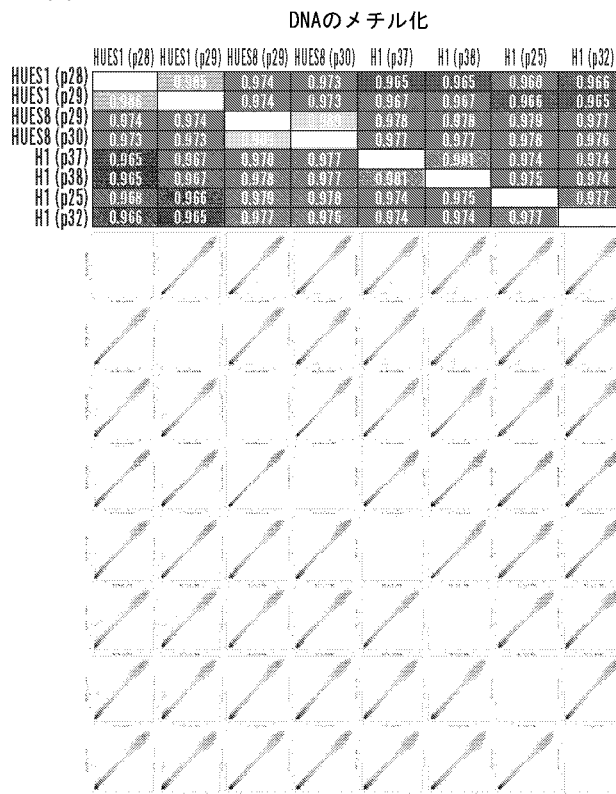
【 図 9 C 】



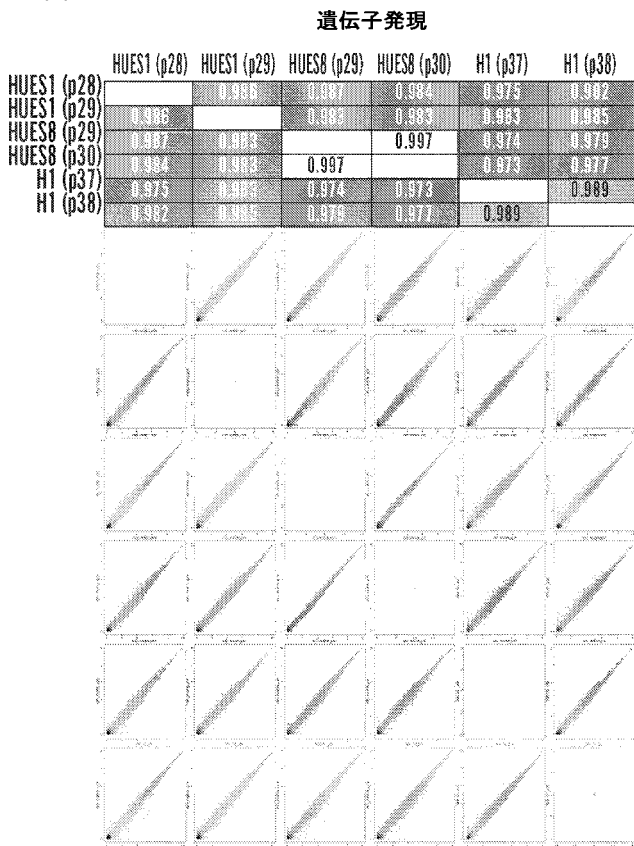
【 図 9 D 】



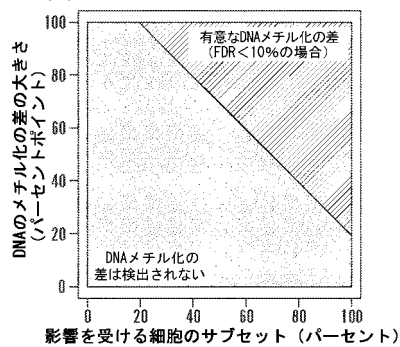
【 図 1 0 A 】



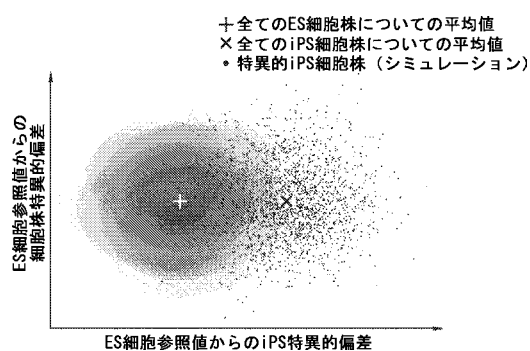
【 図 1 0 B 】



【 図 1 0 C 】



【 図 1 0 D 】



【 図 1 1 A 】

偏差スコアカードを計算するためのアルゴリズム

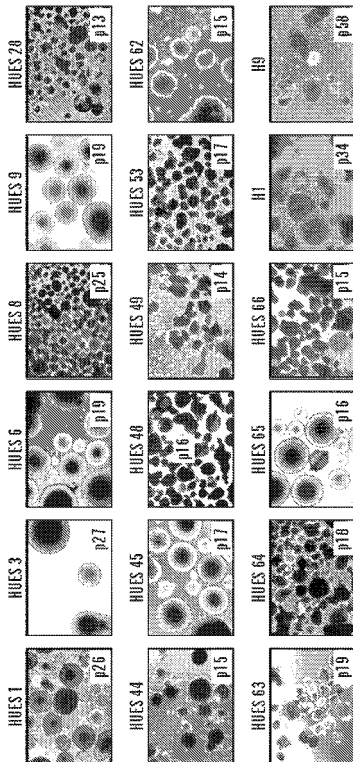
- データのインポート: (i) 関心対象の細胞株、および (ii) 高品質多能性細胞株の参照セットに関する DNA メチル化データ (RRBS によって測定) および遺伝子発現データ (Affymetrix マイクロアレイによって測定) をロードする。
- データの正規化: Affymetrix マイクロアレイの gcRMA による正規化を行い、全ての遺伝子発現値を 0 から 10 の範囲の標的区間の縮尺にする。
- 遺伝子マッピング: Ensembl 遺伝子アノテーションをロードして、各遺伝子に関して、DNA メチル化レベル (プロモーター領域における全ての CpG について平均する) および遺伝子発現レベル (別の転写物について平均する) を計算する。重み付けスキームにより、試料間の異なるシーケンシングカバー率を補正する。
- 参照との比較: 高品質多能性細胞株の参照セットに基づいて、DNA のメチル化および発現値の予想される範囲を計算し、特異的遺伝子に関する値が四分位数間範囲の 1.5 倍より大きい値だけ中心四分位数の外側に外れる場合、その関心対象の細胞株を外れ値として示す (テューキーの外れ値フィルター)。
- 関連性フィルター: 20 パーセントポイント未満の DNA のメチル化の差または 2 倍未満の発現の変化を有する軽微な外れ値を破棄する。
- 遺伝子セット: 系列マーカー遺伝子または癌遺伝子などの関心対象の応用のために関連遺伝子を含む遺伝子セットをロードする。
- 外れ値の要約: 影響を受けた遺伝子のリストと共に、各細胞株に関する偏差の数をリストにする (図 4B および補足表 5)。

【 図 1 1 B 】

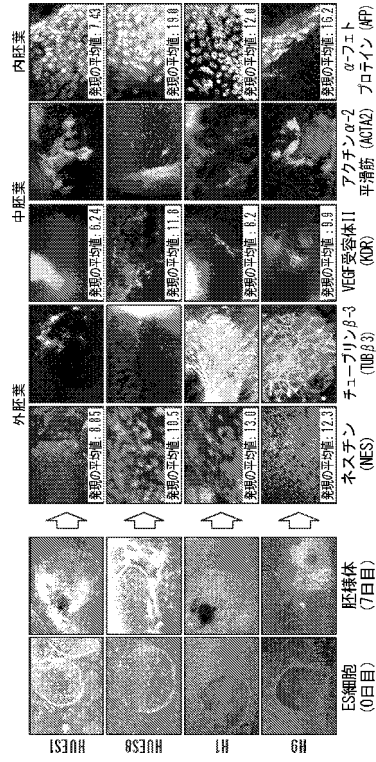
系列スコアカードを計算するためのアルゴリズム

- データのインポート: 以下の試料における 500 個のマーカー遺伝子に関する転写物の絶対数 (nCounter システムによって測定) をロードする: (i) 関心対象の細胞株に由来する胚様体 (EB)、および (ii) 高品質多能性細胞株に由来する EB の参照セット。
- アッセイの標準化: 陽性 Spike-In コントロールを用いて、アッセイ標準化因子を計算し、それによってデータのサイズを変更する (各実験に関して個別に)。
- 試料の標準化: 全ての実験にわたって分散の安定化および標準化を行う (BioConductor の VSN パッケージを用いる)。
- 参照値との比較: 各マーカー遺伝子に関して調整 t-検定を行い、高品質 EB の参照セットに対して、関心対象の細胞株に関する全ての EB レプリケートと比較する (BioConductor の LIMMA パッケージを用いる)。
- 遺伝子セット: 関心対象の細胞系列または胚葉に関して特異的である遺伝子セットをロードする (遺伝子セットは、遺伝子オントロジー、MoSigDB、またはマニュアルキュレーション作業により得ることができる)。
- 濃縮分析: 各遺伝子セットに関して、遺伝子セットに属する全てのマーカー遺伝子の t-スコアの平均値を計算する。
- スコアカードの計算: 各細胞株に関して、全ての関連する遺伝子セットに関する t-スコアの平均値が、遺伝子セットが表す系列に関するスコアカード推定値を構成する (図 5B, D)。

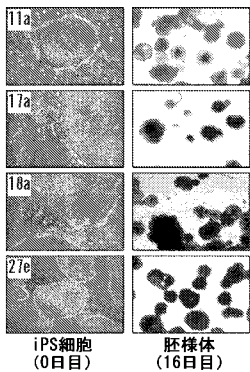
【 図 1 2 A 】



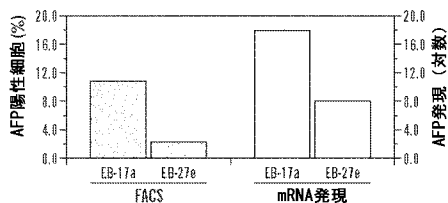
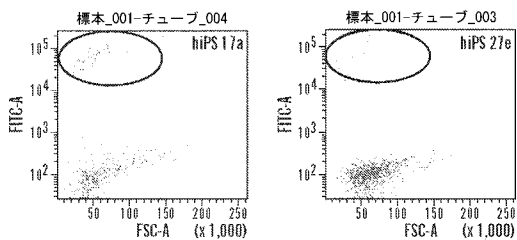
【 図 1 2 B 】



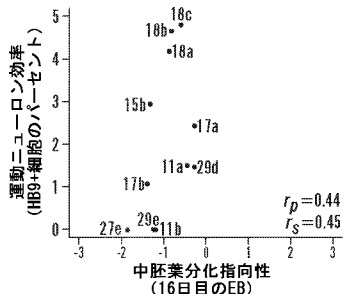
【 図 1 2 C 】



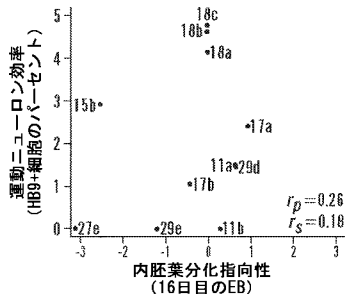
【 図 1 2 D 】



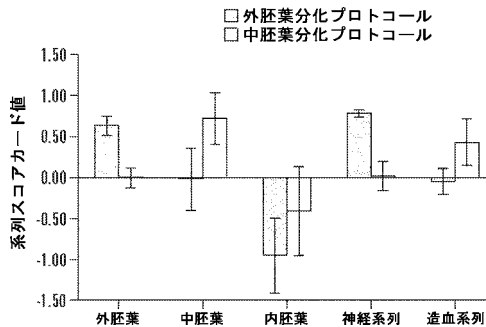
【 図 1 3 B 】



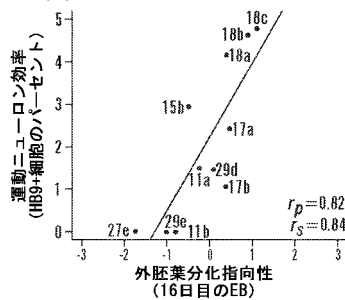
【 図 1 3 C 】



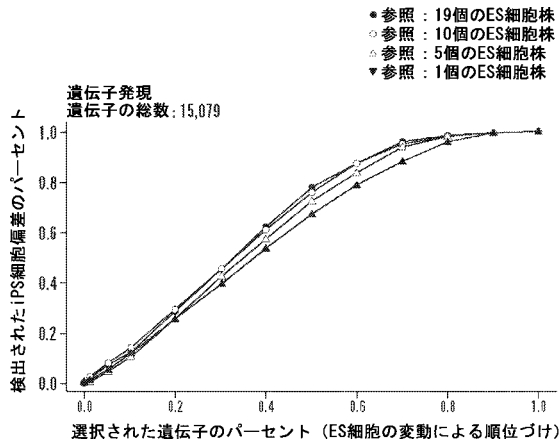
【 図 1 2 E 】



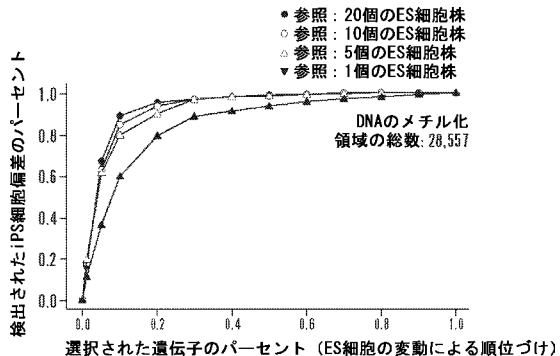
【 図 1 3 A 】



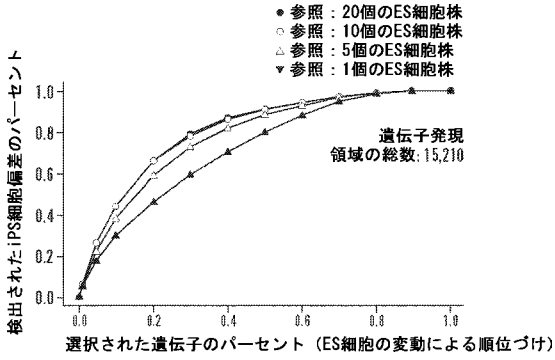
【 図 1 4 A 】



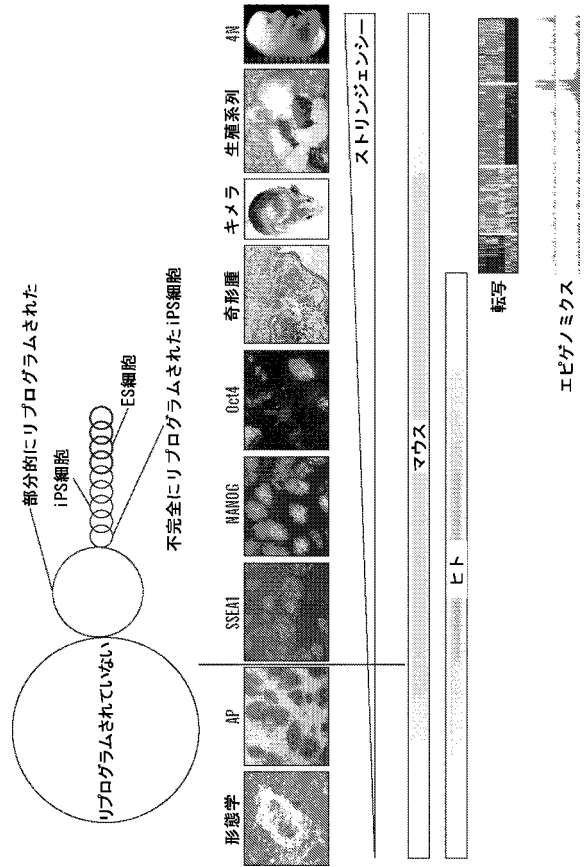
【 図 1 4 B 】



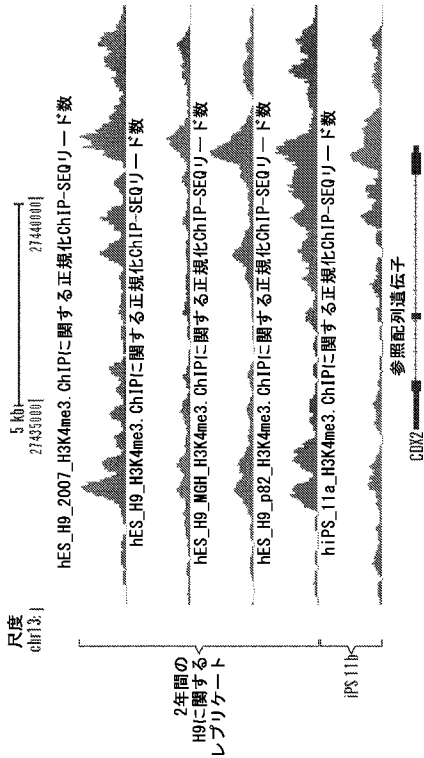
【 図 1 4 C 】



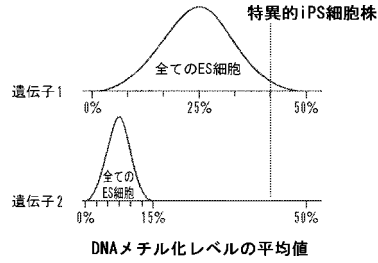
【 図 1 5 】



【 図 1 6 】



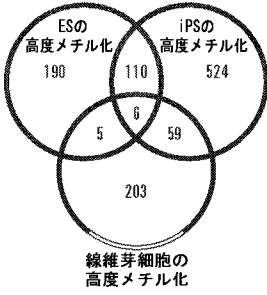
【 図 1 7 】



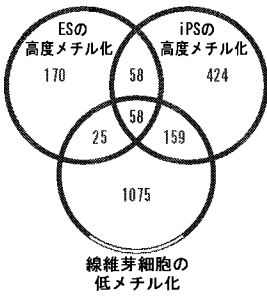
【 図 1 8 】

	増加したメチル化	減少したメチル化
ES	235	456
ES	344	262
ES	47	293
ES	134	126
ES	296	210
ES	177	182
iPS 11a	462	45
iPS 11b	98	169
iPS 15b	241	126
iPS 17a	592	3
iPS 17b	102	72
iPS 18a	86	28
iPS 18b	234	39
iPS 18c	314	18
iPS 20b	650	172

【 図 19 】

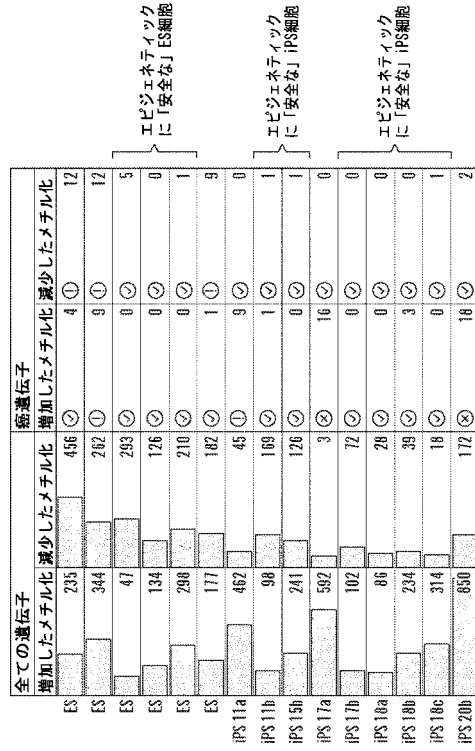


A

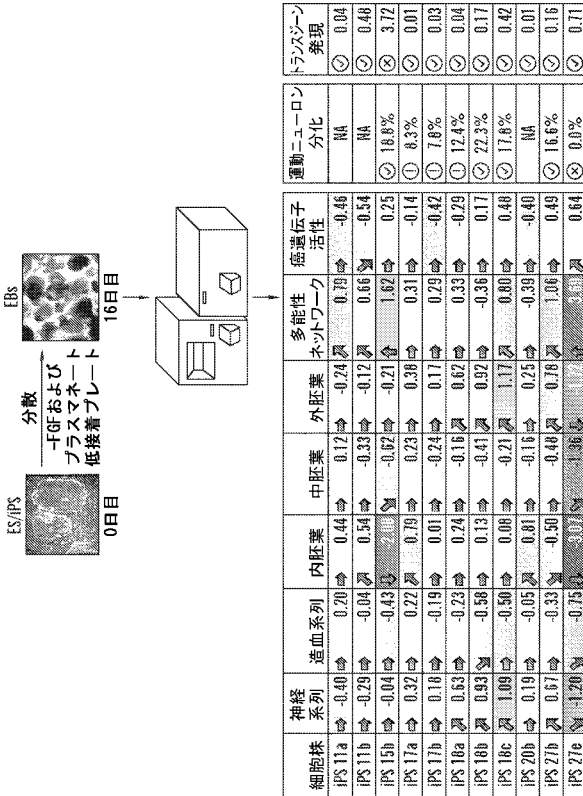


B

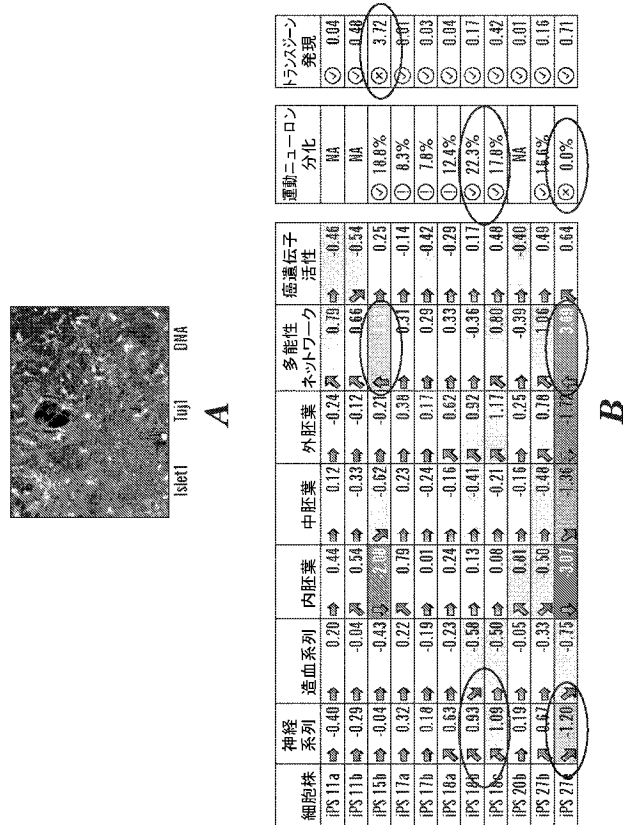
【 図 20 】



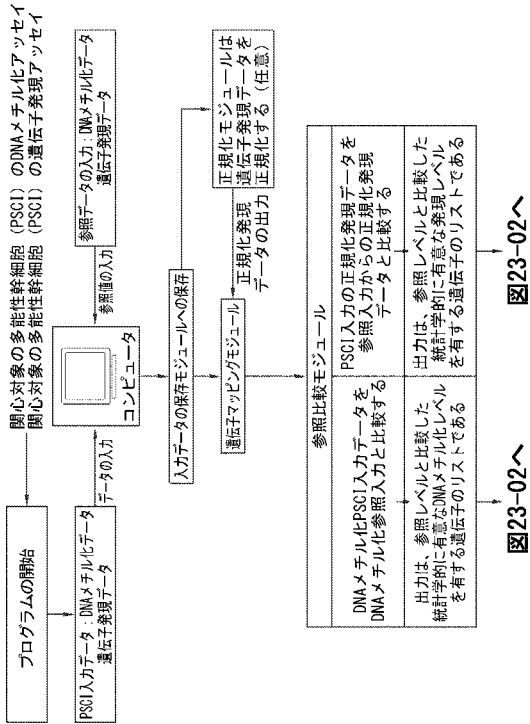
【 図 21 】



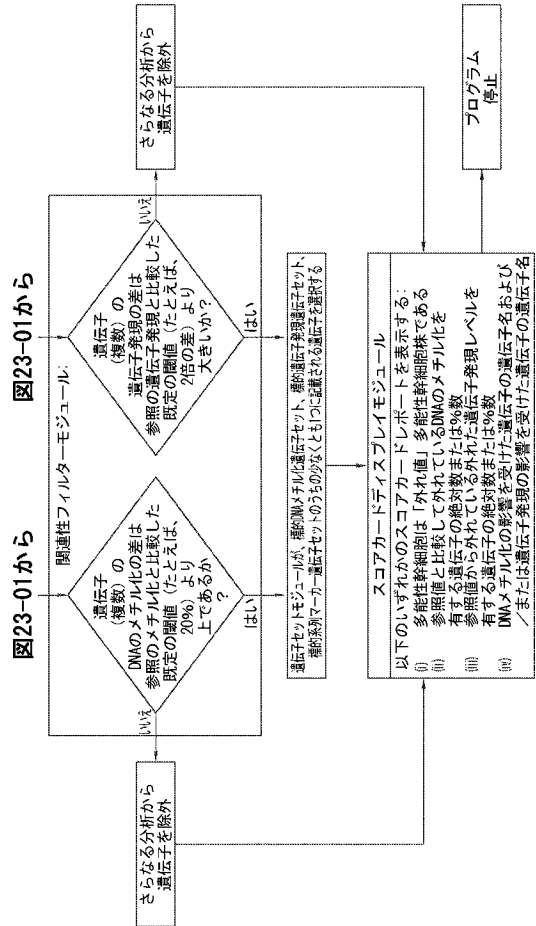
【 図 22 】



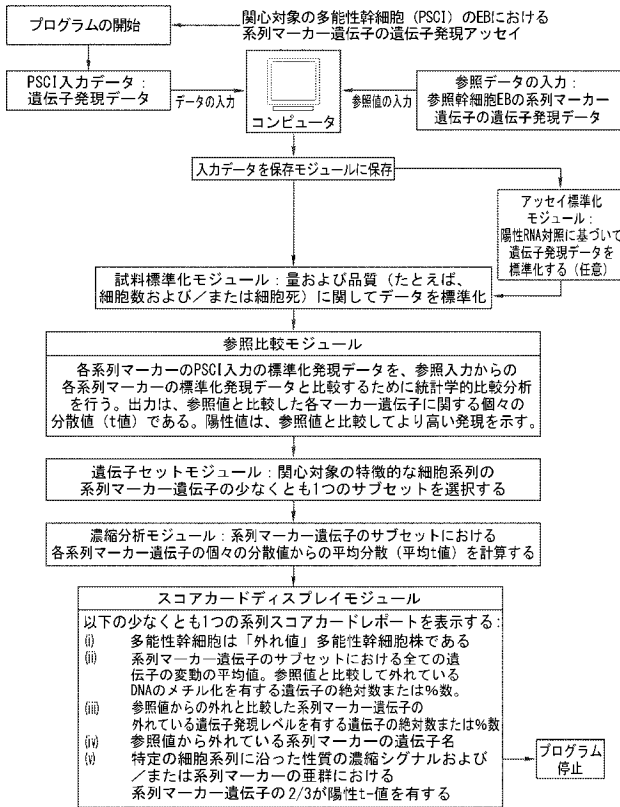
【 図 2 3 - 0 1 】



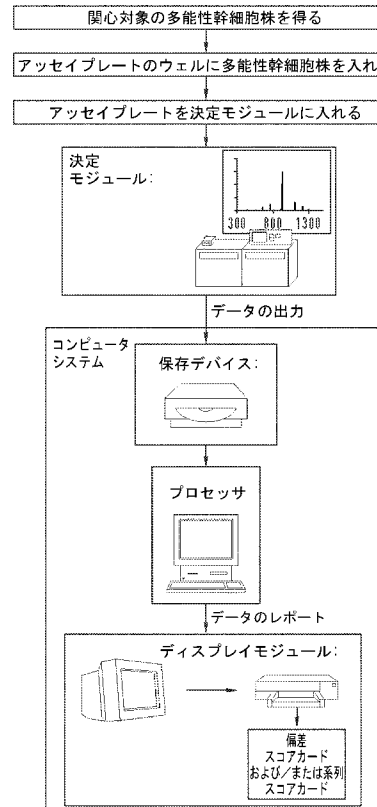
【 図 2 3 - 0 2 】



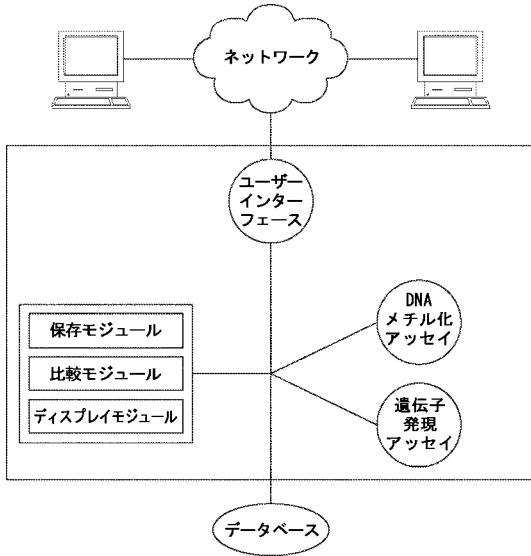
【 図 2 4 】



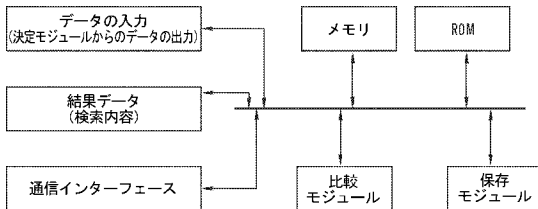
【 図 2 5 】



【図 2 6】



【図 2 7】



【配列表】

2017104105000001.app

【手続補正書】

【提出日】平成28年12月15日(2016.12.15)

【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

系列マーカのセットのmRNAまたはcDNAを増幅するオリゴヌクレオチドまたはオリゴヌクレオチド対を含む、細胞を特徴決定するための組成物であって、該系列マーカのセットは少なくとも一つの中胚葉系列マーカ、少なくとも一つの内胚葉系列マーカ、および少なくとも一つの外胚葉系列マーカを含み、

(i) 中胚葉系列マーカがCD34、DLL1、HHEX、INHBA、LEF1、SRF、T、TWIST1、ADIPOQ、MME、KIT、ITGAL、ITGAM、ITGAX、TNFRSF1A、ANPEP、SDC1、CDH5、MCAM、FUT4、NGFR、ITGB1、PECAM1、CDH1、CDH2、CD36、CD4、CD44、ITGA4、ITGA6、ITGAV、ICAM1、NCAM1、ITGB3、CEACAM1、THY1、ABCG2、KDR、GATA3、GATA4、MYOD1、MYOG、NES、NOTCH1、SPI1、またはSTAT3から選択されるものであり；

(ii) 外胚葉系列マーカがNCAM1、EN1、FGFR2、GATA2、GATA3、HAND1、MNX1、NEFL、NES、NOG、OTX2、PAX3、PAX6、PAX7、SNAI2、SOX10、SOX9、TDGF、APOE、PDGFRA、MCAM、FUT4、NGFR、ITGB1、CD44、ITGA4、ITGA6、ICAM1、NCAM1、THY1、FAS、ABCG2、CRABP2、MAP2、CDH2、NES、NEUROG3、NOG、NOTCH1、SOX2、SYP、MAPT、またはTHから選択されるものであり；ならびに

(iii) 内胚葉系列マーカーがAPOE、CDX2、FOXA2、GATA4、GATA6、GCG、ISL1、NKX2-5、PDX1、SLC2A2、SST、ITGB1、CD44、ITGA6、THY1、CDX2、GATA4、HNF1A、HNF1B、CDH2、NEUROG3、CTNNB1、またはSYPから選択されるものである、組成物。

【請求項2】

中胚葉系列マーカーがCD34、DLL1、HHEX、INHBA、LEF1、SRF、T、TWIST1、ADIPOQ、MME、KIT、ITGAL、ITGAM、ITGAX、TNFRSF1A、ANPEP、SDC1、CDH5、MCAM、PECAM1、CDH1、CDH2、CD36、CD4、ITGAV、ITGB3、CEACAM1、ABCG2、KDR、MYOD1、MYOG、NES、NOTCH1、SP11、またはSTAT3から選択されるものであり；

外胚葉系列マーカーがNCAM1、EN1、FGFR2、GATA2、HAND1、MNX1、NEFL、NES、NOG、OTX2、PAX3、PAX6、PAX7、SNAI2、SOX10、SOX9、TDGF、PDGFRA、MCAM、FAS、ABCG2、CRABP2、MAP2、NES、NOG、NOTCH1、SOX2、MAPT、またはTHから選択されるものであり；ならびに内胚葉系列マーカーがCDX2、FOXA2、GATA4、GATA6、GCG、ISL1、NKX2-5、PDX1、SLC2A2、SST、CDX2、GATA4、HNF1A、HNF1B、またはCTNNB1から選択されるものである、

請求項1記載の組成物。

【請求項3】

多能性マーカーのセットのmRNAまたはcDNAを増幅するオリゴヌクレオチドまたはオリゴヌクレオチド対をさらに含む、請求項1または2記載の組成物。

【請求項4】

多能性マーカーのセットが、CXCL5、NANOG、POU5F1、SOX2から選択される少なくとも1つの多能性マーカーを含む、請求項3記載の組成物。

【請求項5】

少なくとも1つの制御遺伝子に対応する配列を増幅する少なくとも1つのオリゴヌクレオチドまたは少なくとも1つのオリゴヌクレオチド対をさらに含む、請求項1～4のいずれか一項記載の組成物。

【請求項6】

オリゴヌクレオチドが、固相支持体の表面に固定または付着される、請求項1～5のいずれか一項記載の組成物。

【請求項7】

固相支持体がマイクロタイタープレートである、請求項6記載の組成物。

【請求項8】

マイクロタイタープレートが、96または384以下のウェルを有し、各ウェルが、系列マーカーのためのオリゴヌクレオチドまたはオリゴヌクレオチド対を有する、請求項7記載の組成物。

【請求項9】

オリゴヌクレオチドまたはオリゴヌクレオチド対が蛍光標識されている、請求項1～8のいずれか一項記載の組成物。

【請求項10】

細胞系の分化能を決定するための方法における請求項1記載の組成物の使用であって、該方法が、

細胞由来の核酸を用いて増幅を行う段階；

細胞系における系列マーカーのセットの発現を検出する段階であって、該系列マーカーのセットは少なくとも一つの中胚葉系列マーカー、少なくとも一つの内胚葉系列マーカー、および少なくとも一つの外胚葉系列マーカーを含む、段階；

該系列マーカーのセットの発現を、少なくとも一つまたは複数の参照多能性幹細胞試料における同じ系列マーカーのセットの発現と比較する段階、およびこの比較に基づき；

中胚葉系列、内胚葉系列、または外胚葉系列へ分化するための細胞系の分化能を決定する段階

を含む、使用。

【請求項11】

アレイ増幅を行った後に、データをウェブサーバー上に位置するソフトウェアを用いて解析する、請求項10記載の使用。

【請求項12】

ソフトウェアが、中胚葉系列、内胚葉系列、および外胚葉系列から選択される系列に分化する可能性があることを示すシグナルを出力する、請求項11記載の使用。

【請求項13】

ソフトウェアが、多能性幹細胞の多能性を示すシグナルを出力する、請求項12記載の使用。

【請求項14】

増幅が、ポリメラーゼ連鎖反応(PCR)；逆転写ポリメラーゼ連鎖反応(RT-PCR)、定量的RT-PCRを含む方法によって行われる、請求項10記載の使用。

【請求項15】

細胞の分化能を特徴決定することによって、中胚葉系列、内胚葉系列、または外胚葉系列から選択される系列へ分化する細胞系を選択するためのアッセイ法であって、

i. 系列遺伝子マーカーのセットの発現レベルを測定する段階であって、該系列遺伝子マーカーのセットは少なくとも一つの中胚葉系列マーカー、少なくとも一つの内胚葉系列マーカー、および少なくとも一つの外胚葉系列マーカーを含み、

中胚葉系列マーカーがCD34、DLL1、HHEX、INHBA、LEF1、SRF、T、TWIST1、ADIPOQ、MME、KIT、ITGAL、ITGAM、ITGAX、TNFRSF1A、ANPEP、SDC1、CDH5、MCAM、FUT4、NGFR、ITGB1、PECAM1、CDH1、CDH2、CD36、CD4、CD44、ITGA4、ITGA6、ITGAV、ICAM1、NCAM1、ITGB3、CEACAM1、THY1、ABCG2、KDR、GATA3、GATA4、MYOD1、MYOG、NES、NOTCH1、SPI1、またはSTAT3から選択されるものであり；

外胚葉系列マーカーがNCAM1、EN1、FGFR2、GATA2、GATA3、HAND1、MNX1、NEFL、NES、NOG、OTX2、PAX3、PAX6、PAX7、SNAI2、SOX10、SOX9、TDGF、APOE、PDGFRA、MCAM、FUT4、NGFR、ITGB1、CD44、ITGA4、ITGA6、ICAM1、NCAM1、THY1、FAS、ABCG2、CRABP2、MAP2、CDH2、NES、NEUROG3、NOG、NOTCH1、SOX2、SYP、MAPT、またはTHから選択されるものであり、；ならびに

内胚葉系列マーカーがAPOE、CDX2、FOXA2、GATA4、GATA6、GCG、ISL1、NKX2-5、PDX1、SLC2A2、SST、ITGB1、CD44、ITGA6、THY1、CDX2、GATA4、HNF1A、HNF1B、CDH2、NEUROG3、CTNNB1、またはSYPから選択されるものである、段階；

ii. 細胞における該系列マーカーのセットの遺伝子発現レベルを、同じ系列マーカーのセットの参照遺伝子発現レベルと比較する段階；ならびに

iii. 測定された系列マーカーのうち少なくとも一つの遺伝子発現レベルが、同じ系列マーカーの参照遺伝子発現レベルと比較して統計学的に有意な差がないことに基づき、細胞系を選択する段階；または

測定された系列マーカーのうち少なくとも1つの発現レベルが、同じ系列マーカーの参照発現レベルと比較して統計学的に有意な差があることに基づき、細胞系を選択する段階を含む、アッセイ法。

【請求項16】

中胚葉系列マーカーがCD34、DLL1、HHEX、INHBA、LEF1、SRF、T、TWIST1、ADIPOQ、MME、KIT、ITGAL、ITGAM、ITGAX、TNFRSF1A、ANPEP、SDC1、CDH5、MCAM、PECAM1、CDH1、CDH2、CD36、CD4、ITGAV、ITGB3、CEACAM1、ABCG2、KDR、MYOD1、MYOG、NES、NOTCH1、SPI1、またはSTAT3から選択されるものであり；

外胚葉系列マーカーがNCAM1、EN1、FGFR2、GATA2、HAND1、MNX1、NEFL、NES、NOG、OTX2、PAX3、PAX6、PAX7、SNAI2、SOX10、SOX9、TDGF、PDGFRA、MCAM、FAS、ABCG2、CRABP2、MAP2、NES、NOG、NOTCH1、SOX2、MAPT、またはTHから選択されるものであり；ならびに

内胚葉系列マーカーがCDX2、FOXA2、GATA4、GATA6、GCG、ISL1、NKX2-5、PDX1、SLC2A2、SST、CDX2、GATA4、HNF1A、HNF1B、またはCTNNB1から選択されるものである、請求項15記載のアッセイ法。

フロントページの続き

(51)Int.Cl.		F I		テーマコード(参考)
A 6 1 K	35/545 (2015.01)	C 1 2 M	1/00	A
A 6 1 P	35/00 (2006.01)	A 6 1 K	35/545	
A 6 1 P	3/10 (2006.01)	A 6 1 P	35/00	
A 6 1 P	9/04 (2006.01)	A 6 1 P	3/10	
A 6 1 P	21/00 (2006.01)	A 6 1 P	9/04	
A 6 1 P	1/04 (2006.01)	A 6 1 P	21/00	
A 6 1 P	3/08 (2006.01)	A 6 1 P	1/04	
A 6 1 P	25/00 (2006.01)	A 6 1 P	3/08	
A 6 1 P	25/28 (2006.01)	A 6 1 P	25/00	
A 6 1 P	43/00 (2006.01)	A 6 1 P	25/28	
G 0 1 N	33/48 (2006.01)	A 6 1 P	43/00	1 1 1
G 0 1 N	33/50 (2006.01)	G 0 1 N	33/48	M
G 0 1 N	37/00 (2006.01)	G 0 1 N	33/50	P
G 0 1 N	33/15 (2006.01)	G 0 1 N	37/00	1 0 2
G 0 1 N	33/53 (2006.01)	G 0 1 N	33/50	Z
C 1 2 N	15/09 (2006.01)	G 0 1 N	33/15	Z
C 1 2 N	5/10 (2006.01)	G 0 1 N	33/53	Y
A 6 1 K	35/54 (2015.01)	C 1 2 N	15/00	A
A 6 1 K	35/28 (2015.01)	C 1 2 N	5/10	
A 6 1 K	35/15 (2015.01)	A 6 1 K	35/54	
A 6 1 K	35/39 (2015.01)	A 6 1 K	35/28	
A 6 1 K	35/30 (2015.01)	A 6 1 K	35/15	Z
A 6 1 K	35/36 (2015.01)	A 6 1 K	35/39	
A 6 1 K	35/34 (2015.01)	A 6 1 K	35/30	
A 6 1 K	35/407 (2015.01)	A 6 1 K	35/36	
A 6 1 K	35/14 (2015.01)	A 6 1 K	35/34	
A 6 1 K	35/17 (2015.01)	A 6 1 K	35/407	
		A 6 1 K	35/14	Z
		A 6 1 K	35/17	Z

4 . V I S U A L B A S I C

5 . Z I G B E E

6 . イ ー サ ネ ッ ト

7 . B L U E T O O T H

- (74)代理人 100142929
 弁理士 井上 隆一
- (74)代理人 100148699
 弁理士 佐藤 利光
- (74)代理人 100128048
 弁理士 新見 浩一
- (74)代理人 100129506
 弁理士 小林 智彦
- (74)代理人 100114340
 弁理士 大関 雅人
- (74)代理人 100121072
 弁理士 川本 和弥

- (72)発明者 エッグガン ケビン シー .
アメリカ合衆国 マサチューセッツ州 ボストン エセックス ストリート 181 #イー40
2
- (72)発明者 マイスナー アレクサンダー
アメリカ合衆国 マサチューセッツ州 ケンブリッジ フォース ストリート プレイス 6 ア
パートメント 3
- (72)発明者 ボック クリストフ
アメリカ合衆国 マサチューセッツ州 ケンブリッジ グリーン ストリート 515
- (72)発明者 キスキニス エバンゲロス
アメリカ合衆国 マサチューセッツ州 ボストン ビーコン ストリート 504 #56
- (72)発明者 ベルスタッペン グリエト アニー フランス
ベルギー王国 モルツェル リダー ヴァン ランストレイ 16 / 1

F ターム(参考) 2G045 AA25 DA13

4B029 AA07 BB20 FA04 FA12 FA15

4B063 QA01 QA13 QA20 QQ42 QQ53 QR32 QR55 QR62 QR66 QR82

QS33 QS34 QX02

4B065 AA90X AB01 AC20 BA02 BA30 CA46

4C087 AA01 AA02 BB34 BB37 BB44 BB45 BB47 BB48 BB51 BB52

BB57 BB64 CA04 NA14 ZA01 ZA15 ZA37 ZA66 ZA94 ZB26

ZC35 ZC41

专利名称(译)	功能基因组学分析，以表征多能干细胞的实用性和安全性		
公开(公告)号	JP2017104105A	公开(公告)日	2017-06-15
申请号	JP2016233779	申请日	2016-12-01
[标]申请(专利权)人(译)	哈佛大学校长及研究员协会		
申请(专利权)人(译)	总裁和研究员哈佛学院		
[标]发明人	エッガンケビンシー マイスナーアレクサンダー ボッククリストフ キスキニスエバンゲロス ベルスタッペングリエトアニーフランス		
发明人	エッガン ケビン シー マイスナー アレクサンダー ボック クリストフ キスキニス エバンゲロス ベルスタッペン グリエト アニー フランス		
IPC分类号	C12N5/074 C12N5/0735 C12Q1/02 C12Q1/68 C12M1/00 A61K35/545 A61P35/00 A61P3/10 A61P9/04 A61P21/00 A61P1/04 A61P3/08 A61P25/00 A61P25/28 A61P43/00 G01N33/48 G01N33/50 G01N37 /00 G01N33/15 G01N33/53 C12N15/09 C12N5/10 A61K35/54 A61K35/28 A61K35/15 A61K35/39 A61K35/30 A61K35/36 A61K35/34 A61K35/407 A61K35/14 A61K35/17		
CPC分类号	A61P1/04 A61P21/00 A61P25/00 A61P25/28 C12Q1/6881 C12Q2600/154 C12Q2600/158 C12N15 /1072		
FI分类号	C12N5/074.ZNA C12N5/0735 C12Q1/02 C12Q1/68.A C12Q1/68.Z C12M1/00.A A61K35/545 A61P35 /00 A61P3/10 A61P9/04 A61P21/00 A61P1/04 A61P3/08 A61P25/00 A61P25/28 A61P43/00.111 G01N33/48.M G01N33/50.P G01N37/00.102 G01N33/50.Z G01N33/15.Z G01N33/53.Y C12N15/00.A C12N5/10 A61K35/54 A61K35/28 A61K35/15.Z A61K35/39 A61K35/30 A61K35/36 A61K35/34 A61K35 /407 A61K35/14.Z A61K35/17.Z C12Q1/6806.Z C12Q1/6834.Z C12Q1/686.Z		
F-TERM分类号	2G045/AA25 2G045/DA13 4B029/AA07 4B029/BB20 4B029/FA04 4B029/FA12 4B029/FA15 4B063 /QA01 4B063/QA13 4B063/QA20 4B063/QQ42 4B063/QQ53 4B063/QR32 4B063/QR55 4B063/QR62 4B063/QR66 4B063/QR82 4B063/QS33 4B063/QS34 4B063/QX02 4B065/AA90X 4B065/AB01 4B065 /AC20 4B065/BA02 4B065/BA30 4B065/CA46 4C087/AA01 4C087/AA02 4C087/BB34 4C087/BB37 4C087/BB44 4C087/BB45 4C087/BB47 4C087/BB48 4C087/BB51 4C087/BB52 4C087/BB57 4C087 /BB64 4C087/CA04 4C087/NA14 4C087/ZA01 4C087/ZA15 4C087/ZA37 4C087/ZA66 4C087/ZA94 4C087/ZB26 4C087/ZC35 4C087/ZC41		
代理人(译)	清水初衷 井上隆一 佐藤俊光 小林智彦 正人大关		
优先权	61/429965 2011-01-05 US 61/384030 2010-09-17 US		
外部链接	Espacenet		
摘要(译)			

要解决的问题：提供多能干细胞的参考数据集，用于预测多能干细胞系的功能和适用性，用于制备记分卡的方法，系统和试剂盒。记分卡，包括使用至少两种选自表观遗传分析，分化分析和基因表达分析的干细胞分析来预测多能干细胞系的功能和适用性，创建方法和系统津市的门的报价。记分卡参考数据可以与多能干细胞系数数据进行比较，以有效和准确地预测多能干细胞在特定应用中的效用，并且可以用于例如治疗用途，药物筛选和毒性测定，并且它适合细胞谱系分化

