

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2016-531580

(P2016-531580A)

(43) 公表日 平成28年10月13日(2016.10.13)

(51) Int. Cl.	F I	テーマコード (参考)
C 1 2 Q 1/68 (2006.01)	C 1 2 Q 1/68 A	4 B 0 2 9
C 1 2 N 15/09 (2006.01)	C 1 2 N 15/00 A	4 B 0 6 3
C 1 2 M 1/00 (2006.01)	C 1 2 M 1/00 A	4 C 0 8 4
G 0 1 N 33/53 (2006.01)	G 0 1 N 33/53 M	
A 6 1 K 45/00 (2006.01)	A 6 1 K 45/00	

審査請求 未請求 予備審査請求 未請求 (全 111 頁) 最終頁に続く

(21) 出願番号 特願2016-540430 (P2016-540430)
 (86) (22) 出願日 平成26年9月5日 (2014.9.5)
 (85) 翻訳文提出日 平成28年4月28日 (2016.4.28)
 (86) 国際出願番号 PCT/US2014/054342
 (87) 国際公開番号 W02015/035203
 (87) 国際公開日 平成27年3月12日 (2015.3.12)
 (31) 優先権主張番号 61/874, 970
 (32) 優先日 平成25年9月6日 (2013.9.6)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 61/987, 342
 (32) 優先日 平成26年5月1日 (2014.5.1)
 (33) 優先権主張国 米国 (US)

(71) 出願人 514116143
 イムコア ジーティーアイ ダイアグノス
 ティクス インコーポレイテッド
 アメリカ合衆国 ジョージア州 3007
 1, ノークロス, ゲートウェイドライブ
 3130
 (74) 代理人 100091096
 弁理士 平木 祐輔
 (74) 代理人 100118773
 弁理士 藤田 節
 (74) 代理人 100122389
 弁理士 新井 栄一
 (74) 代理人 100111741
 弁理士 田中 夏夫

最終頁に続く

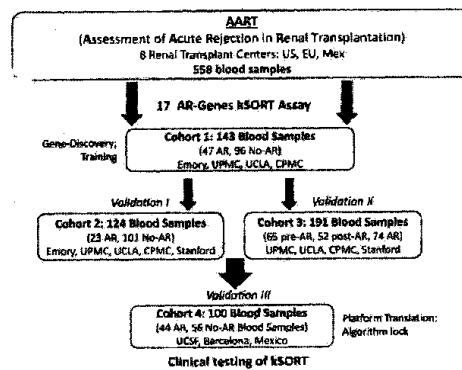
(54) 【発明の名称】 腎臓移植における急性拒絶を評価するための組成物および方法

(57) 【要約】

本明細書では、分類子遺伝子セットの遺伝子発現プロファイルを用いて腎臓移植の急性拒絶を診断するための方法、組成物、およびキットが提供される。このような方法および組成物は、レシピエントの年齢、移植センター、RNAソース、アッセイ、末期腎疾患の原因、併存症、免疫抑制使用などの外的交絡因子に依存しない。

【選択図】 図 1

Figure 1



【特許請求の範囲】

【請求項 1】

腎臓同種移植を受けた個人において (AR) の診断に使用するため、非ARの診断に使用するため、またはARを発症するリスクの診断に使用するための方法であって、

a) 前記個人からの生体サンプルにおいてCEACAM4、ならびにCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37から選択される6~16の間の他の遺伝子のレベルを測定して遺伝子発現結果を得ること；ならびに

b) 各遺伝子に対するARサンプル由来の単一の参照発現ベクターと各遺伝子に対する非ARサンプル由来の単一の参照発現ベクターとを含んでなる参照標準を使用し、診断のために前記遺伝子発現結果が前記参照標準と比較されることを含んでなる、方法。

10

【請求項 2】

前記個人が23歳以上の成人である、請求項1に記載の方法。

【請求項 3】

前記個人が小児または23歳未満の若年成人である、請求項1に記載の方法。

【請求項 4】

6~16の間の他の遺伝子がCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37を含んでなる、請求項1~3のいずれか一項に記載の方法。

20

【請求項 5】

前記測定工程が前記サンプルの遺伝子発現結果をマイクロアレイチップでアッセイすることまたは前記サンプルの遺伝子発現結果をqPCRを用いてアッセイすることを含んでなる、請求項1~4のいずれか一項に記載の方法。

【請求項 6】

前記測定工程が前記サンプルの遺伝子発現結果をビーズでアッセイすることを含んでなる、請求項1~5のいずれか一項に記載の方法。

【請求項 7】

前記測定工程が前記サンプルの遺伝子発現結果をナノ粒子でアッセイすることを含んでなる、請求項1~6のいずれか一項に記載の方法。

30

【請求項 8】

前記生体サンプルが血液サンプルである、請求項1~7のいずれか一項に記載の方法。

【請求項 9】

前記血液サンプルが末梢血白血球または末梢血単核サンプルである、請求項8に記載の方法。

【請求項 10】

前記血液サンプルが全血である、請求項8に記載の方法。

【請求項 11】

前記遺伝子発現結果と前記参照標準の比較が、70%を超える感度でのARの予測を含んでなる、請求項1~10のいずれか一項に記載の方法。

40

【請求項 12】

前記遺伝子発現結果と前記参照標準の比較が、70%を超える特異度でのARの予測を含んでなる、請求項1~11のいずれか一項に記載の方法。

【請求項 13】

前記遺伝子発現結果と前記参照標準の比較が、70%を超える陽性的中率 (ppv) でのARの予測を含んでなる、請求項1~12のいずれか一項に記載の方法。

【請求項 14】

前記遺伝子発現結果と前記参照標準の比較が、70%を超える陰性的中率 (npv) でのARの予測を含んでなる、請求項1~13のいずれか一項に記載の方法。

50

【請求項 15】

腎臓移植の急性拒絶（AR）の処置のための個人の同定において使用する方法であって

a) 前記個人からの生体サンプルにおいてCEACAM4、ならびにCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37から選択される6～16の間の他の遺伝子のレベルを測定して遺伝子発現結果を得ること；および

b) 各遺伝子に対するARサンプル由来の単一の参照発現ベクターと各遺伝子に対する非ARサンプル由来の単一の参照発現ベクターとを含んでなる参照標準を使用し、同定のために前記遺伝子発現結果が前記参照標準と比較されることを含んでなる、方法。

10

【請求項 16】

前記個人が23歳以上の成人である、請求項15に記載の方法。

【請求項 17】

前記個人が小児または23歳未満の若年成人である、請求項15に記載の方法。

【請求項 18】

6～16の間の他の遺伝子がCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37を含んでなる、請求項15～17のいずれか一項に記載の方法。

20

【請求項 19】

前記測定工程が前記サンプルの遺伝子発現結果をマイクロアレイチップでアッセイすることまたは前記サンプルの遺伝子発現結果をqPCRを用いてアッセイすることを含んでなる、請求項15～18のいずれか一項に記載の方法。

【請求項 20】

前記測定工程が前記サンプルの遺伝子発現結果をビーズでアッセイすることを含んでなる、請求項15～19のいずれか一項に記載の方法。

【請求項 21】

前記測定工程が前記サンプルの遺伝子発現結果をナノ粒子でアッセイすることを含んでなる、請求項15～20のいずれか一項に記載の方法。

30

【請求項 22】

前記生体サンプルが血液サンプルである、請求項15～21のいずれか一項に記載の方法。

【請求項 23】

前記血液サンプルが末梢血白血球または末梢血単核細胞である、請求項22に記載の方法。

【請求項 24】

前記血液サンプルが全血である、請求項22に記載の方法。

【請求項 25】

前記遺伝子発現結果と前記参照標準の比較が、70%を超える感度でのARの予測を含んでなる、請求項15～24のいずれか一項に記載の方法。

40

【請求項 26】

前記遺伝子発現結果と前記参照標準の比較が、70%を超える特異度でのARの予測を含んでなる、請求項15～25のいずれか一項に記載の方法。

【請求項 27】

前記比較工程が、70%を超える陽性的中率（ppv）でのARの予測を含んでなる、請求項15～26のいずれか一項に記載の方法。

【請求項 28】

前記遺伝子発現結果と前記参照標準の比較が、70%を超える陰性的中率（npv）でのARの予測を含んでなる、請求項15～27のいずれか一項に記載の方法。

50

【請求項 29】

腎臓同種移植を受けた個人において急性拒絶（AR）の診断に使用するためのシステムであって、

a) 前記個人由来の生体サンプルにおいてCEACAM4、ならびにCF CFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37から選択される6～16の他の遺伝子のレベルを測定して遺伝子発現結果を得るための遺伝子発現評価素子；ならびに

b) 診断のために前記遺伝子発現結果と前記参照標準を比較するための、単一の腎臓移植センターにおける各遺伝子に対するARサンプル由来の単一の参照発現ベクターと、単一の腎臓移植センターにおける各遺伝子に対する非ARサンプル由来の単一の参照発現ベクターとを含んでなる参照標準素子とを含んでなる、システム。

10

【請求項 30】

前記遺伝子発現評価素子がマイクロアレイチップまたはqPCR装置を含んでなる、請求項29に記載のシステム。

【請求項 31】

前記遺伝子発現評価素子がビーズを含んでなる、請求項30に記載のシステム。

【請求項 32】

前記遺伝子発現評価素子がナノ粒子を含んでなる、請求項29～31のいずれか一項に記載のシステム。

20

【請求項 33】

前記参照標準素子がコンピューター生成される、請求項29～32のいずれか一項に記載のシステム。

【請求項 34】

前記遺伝子発現結果と前記参照標準の比較がコンピューターまたは個人により行われる、請求項29～33のいずれか一項に記載のシステム。

【請求項 35】

前記個人が23歳以上の成人である、請求項29～34のいずれか一項に記載のシステム。

30

【請求項 36】

前記個人が小児または23歳未満の若年成人である、請求項29～34のいずれか一項に記載のシステム。

【請求項 37】

6～16の他の遺伝子がCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37を含んでなる請求項29～36のいずれか一項に記載のシステム。

【請求項 38】

前記生体サンプルが血液サンプルである、請求項29～37のいずれか一項に記載のシステム。

40

【請求項 39】

前記血液サンプルが末梢血白血球または末梢血単核細胞である、請求項38に記載のシステム。

【請求項 40】

前記血液サンプルが全血である、請求項38に記載のシステム。

【請求項 41】

前記遺伝子発現結果と前記参照標準の比較が70%を超える感度でARを予測する、請求項29～40のいずれか一項に記載のシステム。

【請求項 42】

50

前記遺伝子発現結果と前記参照標準の比較が70%を超える特異度でARを予測する、請求項29～41のいずれか一項に記載のシステム。

【請求項43】

前記遺伝子発現結果と前記参照標準の比較が70%を超える陽性的中率（ppv）でARを予測する、請求項29～42のいずれか一項に記載のシステム。

【請求項44】

前記遺伝子発現結果と前記参照標準の比較が70%を超える陰性的中率（npv）でARを予測する、請求項29～43のいずれか一項に記載のシステム。

【請求項45】

腎臓同種移植を受けた個人において急性拒絶（AR）の診断に使用するためのキットであって、

10

a) 前記個人由来の生体サンプルにおいてCEACAM4、ならびにCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXTA、およびSLC25A37から選択される6～16の他の遺伝子のレベルを評価して遺伝子発現結果を得るための遺伝子発現評価素子；

b) 単一の腎臓移植センターにおける各遺伝子に対するARサンプル由来の単一の参照発現ベクターと、単一の腎臓移植センターにおける各遺伝子に対する非ARサンプル由来の単一の参照発現ベクターとを含んでなる参照標準素子；および

c) 前記遺伝子発現結果と前記参照標準の比較を含んでなるARを診断するための説明書一式

20

を含んでなる、キット。

【請求項46】

前記個人が23歳以上の成人である、請求項45に記載のキット。

【請求項47】

前記個人が23歳未満の小児または若年成人である、請求項45に記載のキット。

【請求項48】

6～16の他の遺伝子がCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37を含んでなる、請求項45～47のいずれか一項に記載のキット。

30

【請求項49】

前記遺伝子発現評価素子がマイクロアレイチップで前記サンプルの遺伝子発現結果をアッセイすることを含んでなる、請求項45～48のいずれか一項に記載のキット。

【請求項50】

前記遺伝子発現評価素子がビーズで前記サンプルの遺伝子発現結果をアッセイすることを含んでなる、請求項45～49のいずれか一項に記載のキット。

【請求項51】

前記遺伝子発現評価素子がナノ粒子で前記サンプルの遺伝子発現結果をアッセイすることを含んでなる、請求項45～50のいずれか一項に記載のキット。

40

【請求項52】

前記生体サンプルが血液サンプルである、請求項45～50のいずれか一項に記載のキット。

【請求項53】

前記血液サンプルが末梢血白血球または末梢血単核細胞である、請求項52に記載のキット。

【請求項54】

前記血液サンプルが全血である、請求項52に記載のキット。

【請求項55】

前記遺伝子発現結果と前記参照標準の比較が70%を超える感度でARを予測する、請

50

求項 45 ~ 54 のいずれか一項に記載のキット。

【請求項 56】

前記遺伝子発現結果と前記参照標準の比較が70%を超える特異度でARを予測する、請求項 45 ~ 55 のいずれか一項に記載のキット。

【請求項 57】

前記遺伝子発現結果と前記参照標準の比較が70%を超える陽性的中率 (p p v) でARを予測する、請求項 45 ~ 56 のいずれか一項に記載のキット。

【請求項 58】

前記遺伝子発現結果と前記参照標準の比較が70%を超える陰性的中率 (n p v) でARを予測する、請求項 45 ~ 57 のいずれか一項に記載のキット。

10

【請求項 59】

前記遺伝子発現結果と前記参照標準の比較がコンピューターまたは個人により行われる、請求項 45 ~ 58 のいずれか一項に記載のキット。

【請求項 60】

腎臓同種移植を受けた個人由来の生体サンプルにおいてCEACAM4、ならびにCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37から選択される6~16の他の遺伝子のレベルを測定することにより得られる遺伝子発現結果と比較するための参照標準を含んでなり、単一の腎臓移植センターにおける各遺伝子に対するARサンプル由来の単一の参照発現ベクターと、単一の腎臓移植センターにおける各遺伝子に対する非ARサンプル由来の単一の参照発現ベクターとを含んでなり、前記遺伝子発現と前記参照標準の比較が、前記個人において急性拒絶 (AR) の診断に使用するため、非ARの診断に使用するため、またはARを発症するリスクの診断に使用するためのものである、製品。

20

【請求項 61】

前記個人が23歳以上の成人である、請求項 60 に記載の製品。

【請求項 62】

前記個人が小児または23歳未満の若年成人である、請求項 60 に記載の製品。

【請求項 63】

6~16の他の遺伝子がCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37を含んでなる、請求項 60 ~ 62 のいずれか一項に記載の製品。

30

【請求項 64】

CEACAM4および6~16の他の遺伝子のレベルを測定することが、前記サンプルの遺伝子発現結果をマイクロアレイチップでアッセイすること、または前記サンプルの遺伝子発現結果をqPCRを用いてアッセイすることを含んでなる、請求項 60 ~ 63 のいずれか一項に記載の製品。

【請求項 65】

CEACAM4および6~16の他の遺伝子のレベルを測定することが、前記サンプルの遺伝子発現結果をビーズでアッセイすることを含んでなる、請求項 60 ~ 64 のいずれか一項に記載の製品。

40

【請求項 66】

CEACAM4および6~16の他の遺伝子のレベルを測定することが、前記サンプルの遺伝子発現結果をナノ粒子でアッセイすることを含んでなる、請求項 60 ~ 65 のいずれか一項に記載の製品。

【請求項 67】

前記生体サンプルが血液サンプルである、請求項 60 ~ 66 のいずれか一項に記載の製品。

【請求項 68】

50

前記血液サンプルが末梢血白血球または末梢血単核細胞である、請求項 67 に記載の製品。

【請求項 69】

前記血液サンプルが全血である、請求項 67 に記載の製品。

【請求項 70】

前記遺伝子発現と前記参照標準の比較が、70%を超える感度でのARの予測を含んでなる、請求項 60 ~ 69 のいずれか一項に記載の製品。

【請求項 71】

前記遺伝子発現と前記参照標準の比較が、70%を超える特異度でのARの予測を含んでなる、請求項 60 ~ 70 のいずれか一項に記載の製品。

10

【請求項 72】

前記遺伝子発現と前記参照標準の比較が、70%を超える陽性的中率 (ppv) でのARの予測を含んでなる、請求項 60 ~ 71 のいずれか一項に記載の製品。

【請求項 73】

前記遺伝子発現と前記参照標準の比較が、70%を超える陰性的中率 (npv) でのARの予測を含んでなる、請求項 60 ~ 72 のいずれか一項に記載の製品。

【請求項 74】

腎臓移植患者のための処置の方法であって、

a) 前記個人からの生体サンプルにおいてCEACAM4、ならびにCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37から選択される6~16の間の他の遺伝子のレベルを測定して遺伝子発現結果を得ること；

20

b) 各遺伝子に対するARサンプル由来の単一の参照発現ベクターと各遺伝子に対する非ARサンプル由来の単一の参照発現ベクターとを含んでなる参照標準を使用し、前記遺伝子発現結果が前記参照標準と比較され、それにより、被験者を腎臓移植のARを有するまたは腎臓移植のARを有さないとして同定すること；

c) 腎臓移植のARを有する被験者では治療上有効な量の1種類以上の治療薬の投与を増やし、腎臓移植のARを有さない被験者では治療上有効な量の1種類以上の治療薬の投与を維持し、または腎臓移植のARを有さない被験者では治療上有効な量の1種類以上の治療薬の投与を減らすことを含んでなる試験を指示することを含んでなる、方法。

30

【請求項 75】

前記個人が23歳以上の成人である、請求項 74 に記載の方法。

【請求項 76】

前記個人が小児または23歳未満の若年成人である、請求項 74 に記載の方法。

【請求項 77】

6~16の他の遺伝子がCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37を含んでなる、請求項 74 ~ 76 のいずれか一項に記載の方法。

40

【請求項 78】

前記測定工程が前記サンプルの遺伝子発現結果をマイクロアレイチップでアッセイすること、または前記サンプルの遺伝子発現結果をqPCRを用いてアッセイすることを含んでなる、請求項 74 ~ 77 のいずれか一項に記載の方法。

【請求項 79】

前記測定工程が前記サンプルの遺伝子発現結果をビーズでアッセイすることを含んでなる、請求項 74 ~ 78 のいずれか一項に記載の方法。

【請求項 80】

前記測定工程が前記サンプルの遺伝子発現結果をナノ粒子でアッセイすることを含んでなる、請求項 74 ~ 79 のいずれか一項に記載の方法。

50

【請求項 8 1】

前記生体サンプルが血液サンプルである、請求項 7 4 ~ 8 0 のいずれか一項に記載の方法。

【請求項 8 2】

前記血液サンプルが末梢血白血球または末梢血単核細胞である、請求項 8 1 に記載の方法。

【請求項 8 3】

前記血液サンプルが全血である、請求項 8 1 に記載の方法。

【請求項 8 4】

前記遺伝子発現結果と前記参照標準の比較が、70%を超える感度でのARの予測を含んでなる、請求項 7 4 ~ 8 3 のいずれか一項に記載の方法。

10

【請求項 8 5】

前記遺伝子発現結果と前記参照標準の比較が、70%を超える特異度でのARの予測を含んでなる、請求項 7 4 ~ 8 4 のいずれか一項に記載の方法。

【請求項 8 6】

前記比較工程が70%を超える陽性的中率 (ppv) でのARの予測を含んでなる、請求項 7 4 ~ 8 5 のいずれか一項に記載の方法。

【請求項 8 7】

前記遺伝子発現結果と前記参照標準の比較が、70%を超える陰性的中率 (npv) でのARの予測を含んでなる、請求項 7 4 ~ 8 6 のいずれか一項に記載の方法。

20

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、分類子遺伝子セットの遺伝子発現プロファイルを用いて腎臓移植の急性拒絶を評価するための方法、組成物、およびキットに関する。記載の方法および組成物は、レシピエントの年齢、移植センター、RNAソース、アッセイ、末期腎疾患の原因、併存症、免疫抑制使用などの外的交絡因子に依存しない。

【背景技術】

【0002】

ドナーから宿主レシピエントへの臓器移植は、特定の医療行為および治療計画の一要素である。移植後には、レシピエントによる移植片拒絶を回避する必要がある。ドナー臓器の生存力を維持するためには、一般に免疫抑制療法が用いられる。しかしながらやはり、固形臓器移植拒絶はなお起こり得る。

30

【0003】

臓器移植拒絶は、超急性、急性、境界急性、準臨床急性、または慢性として分類される。腎臓を含むほとんどの臓器では、臓器拒絶は、その臓器の生検を行うことでしか明白に診断できない。しかしながら、実用上の理由のため、生検は急性拒絶が疑われる場合にいつも行われるわけではない。さらに、生検にはサンプリングおよび解釈によりバイアスがかかる場合があり(Furness, P.N. et al. Transplantation 2003, 76, 969-973; Furness, P.N. Transplantation 2001, 71, SS31-36)、それらは予測されない。適時に傷害を検出することは、同種移植の健康および長期生存を保証するために重要である。

40

【0004】

臓器移植レシピエントは直面する主要な臨床問題の1つは、患者の同種免疫閾値および急性移植片拒絶のリスクを連続的にモニタリングするために使用可能な高感度、特異的、かつ、非侵襲的なアッセイが無いことである。極めて冗長かつ非特異的な機能性マーカーの上昇(例えば、移植片の機能不全を示す手段としての血清クレアチニンの上昇)が急性拒絶を示唆し得る。しかしながら、腎臓移植では、監視生検時の予期されない診断(Racusen, L. C. et al. Kidney International 1999, 55, 713-723; Solez, K. et al. Am. J. Transplant. 2008, 8, 753-760; Naesens, M. et al. Am. J. Transplant. 2012, 12, 2730-2743)まで、血清クレアチニンのドリフトにより、損傷が検出されずに持続する(準

50

臨床急性拒絶)という認識が高まっている(Lerut, E. et al. Transplantation 2007, 83, 1416-1422; Sigdel, T. K. et al. J. Am. Soc. Nephrol. 2012, 23, 750-763; Moreso, F. et al. Am. J. Transplant. 2006, 6, 747-752; Moreso, F. et al. Transplantation 2012, 93, 41-46; Heilman, R. L. et al. Am. J. Transplant. 2010, 10, 563-570)。

【0005】

急性移植片拒絶(AR)の検出を高い特異度(ARのリスクの低い患者において侵襲的プロトコルの生検を少なくするため)かつ高感度(ARのリスクの高い患者臨床監視を高めるため)で、現在可能なものよりも早く可能とするアッセイシリーズは、ARを緩和するとともに寛解期患者および安定疾患患者に対する免疫抑制プロトコルを軽減するための適時な臨床的介入をもたらす。多くのアッセイは、レシピエントの年齢、併存症、移植センター、免疫抑制使用、および/または末期腎疾患の原因など依存すると思われる。本明細書では、これらの変数に依存しないアッセイの開発を通してこの問題に対する解決策を記載する。

10

【0006】

本明細書に引用されている全ての特許、特許出願、刊行物、文書、および文献は、特に断りのない限り、引用することによりそれらの全内容が本明細書の一部とされる。

【先行技術文献】

【非特許文献】

【0007】

- 【非特許文献1】Furness, P.N. et al. Transplantation 2003, 76, 969-973
- 【非特許文献2】Furness, P.N. Transplantation 2001, 71, SS31-36
- 【非特許文献3】Racusen, L. C. et al. Kidney International 1999, 55, 713-723
- 【非特許文献4】Solez, K. et al. Am. J. Transplant. 2008, 8, 753-760
- 【非特許文献5】Naesens, M. et al. Am. J. Transplant. 2012, 12, 2730-2743
- 【非特許文献6】Lerut, E. et al. Transplantation 2007, 83, 1416-1422
- 【非特許文献7】Sigdel, T. K. et al. J. Am. Soc. Nephrol. 2012, 23, 750-763
- 【非特許文献8】Moreso, F. et al. Am. J. Transplant. 2006, 6, 747-752
- 【非特許文献9】Moreso, F. et al. Transplantation 2012, 93, 41-46
- 【非特許文献10】Heilman, R. L. et al. Am. J. Transplant. 2010, 10, 563-570

20

30

【発明の概要】

【課題を解決するための手段】

【0008】

発明の簡単な概要

本明細書では、ある個人を腎臓移植の急性拒絶(AR)のリスクが高いとして、および/または急性拒絶のリスクが低いもしくはリスクが無い(非AR)として分類するための組成物および方法が開示される。これらの組成物および方法は、分類子遺伝子セットの遺伝子発現レベルを含んでなり、小児患者および成人患者の両方でこのような分類に使用できる。

【0009】

よって、一態様において、本発明は、腎臓同種移植を受けた個人において急性拒絶(AR)の診断において使用する、非ARの診断に使用するため、またはARを発症するリスクの診断に使用するための方法であって、a)前記個人からの生体サンプルにおいてCFLLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37から選択される6~16の間の他の遺伝子のレベルを測定して遺伝子発現結果を得ること;およびb)各遺伝子に対するARサンプル由来の単一の参照発現ベクターと各遺伝子に対する非ARサンプル由来の単一の参照発現ベクターとを含んでなる参照標準を使用し、前記遺伝子発現結果が前記参照標準と相関されることを含んでなる方法を提供する。本明細書の実施形態のいずれにおいても、個人は23歳以上の

40

50

成人であり得る。本明細書の実施形態のいずれにおいても、個人は小児または23歳未満の若年成人であり得る。本明細書の実施形態のいずれにおいても、6~16の他の遺伝子は、CFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37を含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果をマイクロアレイチップでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果をqPCRを用いてアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果をビーズでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果をナノ粒子でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果を、有孔質または非孔質であり得、かつ、サイズに幅があり得る固相表面でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、全血サンプルであり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、血液サンプルであり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血白血球であり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血単核細胞であり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える感度でのARの予測を含んでなり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える特異度でのARの予測を含んでなり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える陽性的中率(ppv)でのARの予測を含んでなり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える陰性的中率(npv)でのARの予測を含んでなり得る。

【0010】

別の態様において、本発明は、腎臓移植の急性拒絶(AR)の処置のための個人の同定において使用する方法であって、a)前記個人からの生体サンプルにおいてCEACAM4、ならびにCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37から選択される6~16の間の他の遺伝子のレベルを測定して遺伝子発現結果を得ること；およびb)各遺伝子に対するARサンプル由来の単一の参照発現ベクターと各遺伝子に対する非ARサンプル由来の単一の参照発現ベクターとを含んでなる参照標準を使用し、同定のために前記遺伝子発現結果が前記参照標準と相関されることを含んでなる方法を提供する。本明細書の実施形態のいずれにおいても、個人は、23歳以上の成人であり得る。本明細書の実施形態のいずれにおいても、個人は、小児または23歳未満の若年成人であり得る。本明細書の実施形態のいずれにおいても、6~16の他の遺伝子は、CFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37を含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果をマイクロアレイチップでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果をビーズでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果をナノ粒子でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果を、有孔質または非孔質であり得、かつ、サイズに幅があり得る固相表面でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、血液サンプルであり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血白血球であり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血

10

20

30

40

50

単核細胞であり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、全血サンプルであり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える感度でのARの予測を含んでなり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える特異度でのARの予測を含んでなり得る。本明細書の実施形態のいずれにおいても、前記比較工程は、70%を超える陽性的中率（ ppv ）でのARの予測を含んでなり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える陰性的中率（ npv ）でのARの予測を含んでなり得る。

【0011】

別の態様において、本発明は、腎臓同種移植を受けた個人において急性拒絶（AR）の診断に使用するためのシステムであって、a）前記個人由来の生体サンプルにおいてCEACAM4、ならびにCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37から選択される6~16の他の遺伝子のレベルを測定して遺伝子発現結果を得るための遺伝子発現評価素子；およびb）診断のために前記遺伝子発現結果と前記参照標準を相関させるための、各遺伝子に対するARサンプル由来の単一の参照発現ベクターと、各遺伝子に対する非ARサンプル由来の単一の参照発現ベクターとを含んでなる参照標準素子とを含んでなるシステムを提供する。本明細書の実施形態のいずれにおいても、遺伝子発現評価素子は、マイクロアレイチップを含んでなり得る。本明細書の実施形態のいずれにおいても、遺伝子発現評価素子は、ビーズを含んでなり得る。本明細書の実施形態のいずれにおいても、遺伝子発現評価素子は、ナノ粒子を含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果を、有孔質または非孔質であり得、かつ、サイズに幅があり得る固相表面でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、参照標準素子は、コンピューター生成され得る。本明細書の実施形態のいずれにおいても、前記参照標準に対する前記遺伝子発現結果は、コンピューターまたは個人により行うことができる。本明細書の実施形態のいずれにおいても、個人は、23歳以上の成人であり得る。本明細書の実施形態のいずれにおいても、個人は、小児または23歳未満の若年成人であり得る。本明細書の実施形態のいずれにおいても、6~16の他の遺伝子は、CFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37を含んでなり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、血液サンプルであり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血白血球であり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血単核細胞であり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、全血サンプルであり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える感度でARを予測し得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える特異度でARを予測し得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える陽性的中率（ ppv ）でARを予測し得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える陰性的中率（ npv ）でARを予測し得る。

【0012】

別の態様において、本発明は、腎臓同種移植を受けた個人において急性拒絶（AR）の診断に使用するためのキットであって、a）前記個人由来の生体サンプルにおいてCEACAM4、ならびにCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37から選択される6~16の他の遺伝子のレベルを評価して遺伝子発現結果を得るための遺伝子発現評価素子；b）各遺伝子

に対するARサンプル由来の単一の参照発現ベクターと、各移植センターに対する非ARサンプル由来の単一の参照発現ベクターとを含んでなる参照標準素子；およびc)前記遺伝子発現結果と前記参照標準の相関を含んでなるARを診断するための説明書一式を含んでなるキットを提供する。本明細書の実施形態のいずれにおいても、個人は、23歳以上の成人であり得る。本明細書の実施形態のいずれにおいても、個人は、小児または23歳未満の若年成人であり得る。本明細書の実施形態のいずれにおいても、6～16の他の遺伝子は、CFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37を含んでなり得る。本明細書の実施形態のいずれにおいても、遺伝子発現評価素子は、前記サンプルの遺伝子発現結果をマイクロアレイチップでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、遺伝子発現評価素子は、前記サンプルの遺伝子発現結果をビーズでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、遺伝子発現評価素子は、前記サンプルの遺伝子発現結果をナノ粒子でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果を、有孔質または非孔質であり得、かつ、サイズに幅があり得る固相表面でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、血液サンプルであり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、全血サンプルであり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血白血球であり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血単核細胞であり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える感度でARを予測し得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える特異度でARを予測し得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える陽性的中率(ppv)でARを予測し得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える陰性的中率(npv)でARを予測し得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、コンピューターまたは個人により行うことができる。

【0013】

別の態様において、本発明は、腎臓同種移植を受けた個人由来の生体サンプルにおいてCEACAM4、ならびにCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37から選択される6～16の他の遺伝子のレベルを測定することにより得られる遺伝子発現結果と比較するための参照標準を含んでなり、単一の腎臓移植センターにおける各遺伝子に対するARサンプル由来の単一の参照発現ベクターと、各遺伝子に対する非ARサンプル由来の単一の参照発現ベクターとを含んでなり、前記遺伝子発現と前記参照標準の相関が、前記個人において急性拒絶(AR)の診断、非ARの診断、またはARを発症するリスクの診断に使用するためのものである、製品を提供する。本明細書の実施形態のいずれにおいても、個人は、23歳以上の成人であり得る。本明細書の実施形態のいずれにおいても、個人は、小児または23歳未満の若年成人であり得る。本明細書の実施形態のいずれにおいても、6～16の他の遺伝子は、CFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37を含んでなり得る。本明細書の実施形態のいずれにおいても、CEACAM4および6～16の他の遺伝子のレベルを測定することは、前記サンプルの遺伝子発現結果をマイクロアレイチップでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、CEACAM4および6～16の他の遺伝子のレベルを測定することは、前記サンプルの遺伝子発現結果をビーズでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、CEACAM

4 および 6 ~ 16 の他の遺伝子のレベルを測定することは、前記サンプルの遺伝子発現結果をナノ粒子でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果を、有孔質または非孔質であり得、かつ、サイズに幅があり得る固相表面でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、血液サンプルである。本明細書の実施形態のいずれにおいても、生体サンプルは、全血サンプルである。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血白血球であり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血単核細胞であり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現と前記参照標準の比較は、70%を超える感度でのARの予測を含んでなり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現と前記参照標準の比較は、70%を超える特異度でのARの予測を含んでなり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現と前記参照標準の比較は、70%を超える陽性的中率 (p p v) でのARの予測を含んでなり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現と前記参照標準の比較は、70%を超える陰性的中率 (n p v) でのARの予測を含んでなり得る。

10

20

30

40

50

【0014】

別の態様において、本発明は、腎臓移植患者のための処置の方法であって、a) 前記個人からの生体サンプルにおいてCEACAM4、ならびにCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37から選択される6~16の間の他の遺伝子のレベルを測定して遺伝子発現結果を得ること；b) 各遺伝子に対するARサンプル由来の単一の参照発現ベクターと各遺伝子に対する非ARサンプル由来の単一の参照発現ベクターとを含んでなる参照標準を使用し、前記遺伝子発現結果が前記参照標準と比較され、それにより、被験者を腎臓移植のARを有するまたは腎臓移植のARを有さないとして同定すること；およびc) 腎臓移植のARを有する被験者では治療上有効な量の1種類以上の治療薬の投与を増やし、腎臓移植のARを有さない被験者では治療上有効な量の1種類以上の治療薬の投与を維持し、または腎臓移植のARを有さない被験者では治療上有効な量の1種類以上の治療薬の投与を減らすことを含んでなる試験を指示することを含んでなる方法を提供する。本明細書の実施形態のいずれにおいても、個人は、23歳以上の成人であり得る。本明細書の実施形態のいずれにおいても、個人は、小児または23歳未満の若年成人であり得る。本明細書の実施形態のいずれにおいても、6~16の他の遺伝子は、CFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37を含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果をマイクロアレイチップでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果をビーズでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果をナノ粒子でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果を、有孔質または非孔質であり得、かつ、サイズに幅があり得る固相表面でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、血液サンプルであり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血白血球であり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血単核細胞であり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、全血サンプルであり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える感度でのARの予測を含んでなり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える特異度でのARの予測を含んでなり得る。本明細書の実施形態のいずれにおいても、前記比較工程は、70%を超える陽性的中率 (p p v) でのARの予測を含んでな

り得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える陰性的中率（ npv ）でのARの予測を含んでなり得る。

【0015】

別の態様において、本発明は、腎臓同種移植を受けた個人において非急性拒絶（非AR）の診断における使用方法であって、a)前記個人由来の生体サンプルにおいてCEACAM4、ならびにCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37から選択される6~16の他の遺伝子のレベルを測定して遺伝子発現結果を得ること；およびb)各遺伝子に対するARサンプル由来の単一の参照発現ベクターと各遺伝子に対する非ARサンプル由来の単一の参照発現ベクターとを含んでなる参照標準を使用し、前記遺伝子発現結果が前記参照標準と相関されることを含んでなる方法を提供する。本明細書の実施形態のいずれにおいても、個人は、23歳以上の成人であり得る。本明細書の実施形態のいずれにおいても、個人は、小児または23歳未満の若年成人であり得る。本明細書の実施形態のいずれにおいても、6~16の他の遺伝子は、CFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37を含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果をマイクロアレイチップでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果をビーズでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果をナノ粒子でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果を、有孔質または非孔質であり得、かつ、サイズに幅があり得る固相表面でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、全血サンプルであり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、血液サンプルであり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血白血球であり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血単核細胞であり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える感度での非ARの予測を含んでなり得る。本明細書の実施形態のい

10

20

30

【0016】

別の態様において、本発明は、腎臓移植の非急性拒絶（非AR）の処置のための個人の同定において使用する方法であって、a)前記個人由来の生体サンプルにおいてCEACAM4、ならびにCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37から選択される~16の他の遺伝子のレベルを測定して遺伝子発現結果を得ること；およびb)単一の腎臓移植センターにおける各遺伝子に対するARサンプル由来の単一の参照発現ベクターと、単一の腎臓移植センターにおける各遺伝子に対する非ARサンプル由来の単一の参照発現ベクターとを含んでなる参照標準を使用し、前記遺伝子発現結果が同定のために前記参照標準と相関されることを含んでなる方法を提供する。本明細書の実施形態のいずれにおいても、個人は、23歳以上の成人であり得る。本明細書の実施形態のいずれにおいても、個人は、小児または23歳未満の若年成人であり得る。本明細書の実施形態のいずれにおいても、6~16の他の遺伝子は、CFLAR、DUSP1、ITGAX、NAMPT、NKTR、PS

40

50

EN1、EPOR、GZMK、RARA、RHEB、およびSLC25A37を含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果をマイクロアレイチップでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果をビーズでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果をナノ粒子でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果を、有孔質または非孔質であり得、かつ、サイズに幅があり得る固相表面でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、全血サンプルであり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、血液サンプルであり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血白血球であり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血単核細胞であり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える感度での非ARの予測を含んでなり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える特異度での非ARの予測を含んでなり得る。本明細書の実施形態のいずれにおいても、前記比較工程は、70%を超える陽性的中率 (ppv) での非ARの予測を含んでなり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える陰性的中率 (npv) での非ARの予測を含んでなり得る。

10

20

30

40

50

【0017】

別の態様において、本発明は、腎臓同種移植を受けた個人における非急性拒絶 (非AR) の診断に使用するためのシステムであって、a) 前記個人由来の生体サンプルにおいてCEACAM4、ならびにCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37から選択される6~16の他の遺伝子のレベルを測定して遺伝子発現結果を得るための遺伝子発現評価素子；およびb) 診断のために前記遺伝子発現結果と前記参照標準を相関させるための、単一の腎臓移植センターにおける各遺伝子に対するARサンプル由来の単一の参照発現ベクターと、単一の腎臓移植センターにおける各遺伝子に対する非ARサンプル由来の単一の参照発現ベクターとを含んでなる参照標準素子とを含んでなるシステムを提供する。本明細書の実施形態のいずれにおいても、遺伝子発現評価素子は、マイクロアレイチップを含んでなり得る。本明細書の実施形態のいずれにおいても、遺伝子発現評価素子は、ビーズを含んでなり得る。本明細書の実施形態のいずれにおいても、遺伝子発現評価素子は、ナノ粒子を含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果を、有孔質または非孔質であり得、かつ、サイズに幅があり得る固相表面でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、参照標準素子は、コンピューター生成され得る。本明細書の実施形態のいずれにおいても、前記参照標準に対する前記遺伝子発現結果は、コンピューターまたは個人により行われ得る。本明細書の実施形態のいずれにおいても、個人は、23歳以上の成人であり得る。本明細書の実施形態のいずれにおいても、個人は、小児または23歳未満の若年成人であり得る。本明細書の実施形態のいずれにおいても、6~16の他の遺伝子は、CFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37を含んでなり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、全血サンプルであり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、血液サンプルであり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血白血球であり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血単核細胞であり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える感度で非ARを予測し得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超

える特異度で非ARを予測し得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える陽性的中率(ppv)で非ARを予測し得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える陰性的中率(npv)で非ARを予測し得る。

【0018】

別の態様において、本発明は、腎臓同種移植を受けた個人における非拒絶(非AR)の診断に使用するためのキットであって、a)前記個人由来の生体サンプルにおいてCEACAM4、ならびにCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37から選択される6~16の他の遺伝子のレベルを測定して遺伝子発現結果を得るための遺伝子発現評価素子；b)単一の腎臓移植センターにおける各遺伝子に対するARサンプル由来の単一の参照発現ベクターと、単一の腎臓移植センターにおける各遺伝子に対する非ARサンプル由来の単一の参照発現ベクターとを含んでなる参照標準素子；およびc)ARを診断するための説明書一式を含んでなり、前記遺伝子発現結果と参照標準の相関を含んでなるキットを提供する。本明細書の実施形態のいずれにおいても、個人は、23歳以上の成人であり得る。本明細書の実施形態のいずれにおいても、個人は、小児または23歳未満の若年成人であり得る。本明細書の実施形態のいずれにおいても、6~16の他の遺伝子は、CFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37を含んでなり得る。本明細書の実施形態のいずれにおいても、遺伝子発現評価素子は、前記サンプルの遺伝子発現結果をマイクロアレイチップでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、遺伝子発現評価素子は、前記サンプルの遺伝子発現結果をビーズでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、遺伝子発現評価素子は、前記サンプルの遺伝子発現結果をナノ粒子でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果を、有孔質または非孔質であり得、かつ、サイズに幅があり得る固相表面でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、全血サンプルであり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、血液サンプルであり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血白血球であり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血単核細胞であり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える感度で非ARを予測し得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える特異度で非ARを予測し得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える陽性的中率(ppv)で非ARを予測し得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、70%を超える陰性的中率(npv)で非ARを予測し得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現結果と前記参照標準の比較は、コンピューターまたは個人により行うことができる。

【0019】

別の態様において、本発明は、腎臓同種移植を受けた個人由来の生体サンプルにおいてCEACAM4、ならびにCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37から選択される6~16の他の遺伝子のレベルを測定することにより得られる遺伝子発現結果と比較するための参照標準を含んでなり、単一の腎臓移植センターにおける各遺伝子に対するARサンプル由来の単一の参照発現ベクターと、単一の腎臓移植センターにおける各遺伝子に対する非ARサンプル由来の単一の参照発現ベクターとを含んでなり、前記遺伝子発現と前記参照標準の相関が、前記個人において非急性拒絶(非AR)の診断に使用するためのものである製

10

20

30

40

50

品を提供する。本明細書の実施形態のいずれにおいても、個人は、23歳以上の成人であり得る。本明細書の実施形態のいずれにおいても、個人は、小児または23歳未満の若年成人であり得る。本明細書の実施形態のいずれにおいても、6~16の他の遺伝子は、CFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37を含んでなり得る。本明細書の実施形態のいずれにおいても、CEACAM4および6~16の他の遺伝子のレベルを測定することは、前記サンプルの遺伝子発現結果をマイクロアレイチップでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、CEACAM4および6~16の他の遺伝子のレベルを測定することは、前記サンプルの遺伝子発現結果をビーズでアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、CEACAM4および6~16の他の遺伝子のレベルを測定することは、前記サンプルの遺伝子発現結果をナノ粒子でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、前記測定工程は、前記サンプルの遺伝子発現結果を、有孔質または非孔質であり得、かつ、サイズに幅があり得る固相表面でアッセイすることを含んでなり得る。本明細書の実施形態のいずれにおいても、生体サンプルは、全血サンプルである。本明細書の実施形態のいずれにおいても、生体サンプルは、血液サンプルである。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血白血球であり得る。本明細書の実施形態のいずれにおいても、血液サンプルは、末梢血単核細胞であり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現と前記参照標準の比較は、70%を超える感度での非ARの予測を含んでなり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現と前記参照標準の比較は、70%を超える特異度での非ARの予測を含んでなり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現と前記参照標準の比較は、70%を超える陽性的中率(ppv)での非ARの予測を含んでなり得る。本明細書の実施形態のいずれにおいても、前記遺伝子発現と前記参照標準の比較は、70%を超える陰性的中率(npv)での非ARの予測を含んでなり得る。

【図面の簡単な説明】

【0020】

【図1】図1は、世界8か所の移植センターからの438名のユニークな成人/小児腎臓移植患者での腎臓移植における急性拒絶の評価(Assessment of Acute Rejection in Renal Transplantation)(AART)研究計画を記載する。

【図2A】図2Aは、ペナルティ付きロジスティック回帰による、15の遺伝子を用いた4センターからの192名の患者における急性拒絶(AR)の予測を示すグラフである。

【図2B】図2Bは、ペナルティ付きロジスティック回帰による、15の遺伝子を用いた4センターからの192名の患者における急性拒絶(AR)の予測を示すグラフである。

【図3】図3Aは、15の遺伝子がペナルティ付きロジスティック回帰により細胞性拒絶および体液性拒絶を検出することを示すグラフである。図3Bは、ペナルティ付きロジスティック回帰による15の遺伝子を用いたARおよび非ARの検出は移植後の時間と交絡がないことを示す。

【図4】図4A~Bは、生検がARエピソードを証明する2年~0か月前、または生検がARエピソードを証明した0~16か月後に採取された156の小児および成人サンプルに関するARの予測確率を示す。図4Aは、成人サンプル集団における15の遺伝子の発現が、ペナルティ付きロジスティック回帰によりARに関して、生検最大33か月前、生検1か月後までARを指示することを示す。図4Bは、10のうち5つの遺伝子の発現が、ロジスティック回帰により、AR生検の最大3か月前、生検後に成人サンプル集団においてARを予測することを示す。

【図5】図5は、修正系統的プロファイラー(kSAS)のワークフローを示す。図5Aは、サンプルがバッチ効果補正の必要なく、AR参照およびSTA参照との全体的類似性に基づいて分類できることを示す。図5Bは、kSAS(修正系統的プロファイラー)をqPCRデータからAR相対リスクモデルまでのワークフローにどのように適合させるか

を示す。

【図6】図6A～Bは、部分最小二乗判別分析(partial least square Discriminant analysis) (plsDA)による17の遺伝子を用いた143の成人サンプルにおけるARと非ARの分類を示す。17の遺伝子を用い、plsDAにより、4か所からの143の成人血液サンプル(コホート1)においてARを予測した。6Aは、各採取場所におけるARと非ARの平均[%]予測確率が、その場所でもARの方が有意に高いこと($p < 0.0001$)、および非ARサンプルではAR予測の閾値(予測確率AR = 50%)に達しなかったことを示す。6Bは、トレーニングセットにおけるARの受信者動作特性(receiver operating characteristic) (ROC) AUCは0.94 (95% CI 0.91 ~ 0.98)であったことを示す。

10

【図7】図7A～Cは、17の遺伝子を用いた124の成人および小児サンプルにおけるARと非ARの分類を示す。フリータイムでの固定plsDA 17遺伝子モデルを用いた、124の成人および小児ARおよび非AR血液サンプル(コホート2)における独立バリデーション。22/23のARは正確にARと分類され、100/101の非ARが正確に非ARと分類された。7A:各サンプルで、表現型(ARと非AR)および患者の年齢(成人;小児)により分離した[%]予測AR確率が示される。7B:全サンプルの平均予測AR確率は、非ARに対してARで有意に高かった($p < 0.0001$)。7C:17遺伝子ARモデルに関するROC分析は、AR予測に関して高い感度および特異度を示した(AUC = 0.95 [95% CI 0.88 ~ 1.0])。

【図8】図8は、17の遺伝子を用いた191の成人および小児サンプルにおけるARの予測を示す。191の一連の血液サンプル(コホート3)を、ARを確認した生検前6か月以内(AR前)または生検後(AR後)にプロファイリングした。74のARサンプル、および117のAR生検前および生検後サンプル、ならびに216の非AR/安定サンプルを含む各群において、ARおよび非ARの平均出現率を示す。柱内にアッセイによって計算されたARの平均予測確率スコアを示す。17遺伝子腎臓AR予測モデルは、AR前3か月以内に採取されたサンプルの62.9%でARを予測し、極めて高い平均ARスコアであった($96.4\% \pm 0.08$)。ARスコアは、AR後3か月以内に採取されたサンプルの51.6%で持続し、この場合にも極めて高い平均予測ARスコアであり($94.6\% \pm 0.14$)；非ARサンプルの83.8%は、常に非ARと予測された(平均予測AR確率 = $8.2\% \pm 0.12$)。平均ARスコアは、AR前サンプル(0～3か月)と非AR/安定サンプルの間で有意に異なっていた($p = 3.72E - 47$)。

20

30

【図9】図9A～Cは、17の遺伝子を用いたkSASアルゴリズムの開発を示す。個々のサンプルのARリスクスコアとARリスクカテゴリーを提供するためのkSASを開発した。図9Aは、未知のサンプルにおける17遺伝子腎臓AR予測アッセイモデルの発現値が、ピアソン相関により対応するAR参照値および非AR参照値と相関が見出されたことを示し；図9Bは、100サンプルからの17遺伝子ARアッセイ生成QPCRデータがトレーニングセット($n = 32$)と独立バリデーションセット($n = 68$)に分類されたことを示し；17遺伝子腎臓AR予測アッセイモデルからの13の12遺伝子モデルは、各サンプルに関する数的統合ARリスクスコアを生成し、それらを高リスクAR(統合ARリスクスコア ≥ 9)、低リスクAR(統合ARリスクスコア ≤ -9)および中間(統合ARリスクスコア $-9 < \text{スコア} < 9$)カテゴリー9Cの3群に類別した。

40

【図10】図10A～Cは、kSASを用いた100サンプルにおける17遺伝子AR予測アッセイの成績を示す。図10Aは、予測統合ARリスクスコアが各サンプルに関して計算され、AR予測アッセイは、4か所の異なるサンプル採取場所、および成人/小児レシピエント年齢で、36/39のARを高リスクARとして(92.3% ; リスクスコア ≥ 9)、43/46の非ARを低リスクARとして(93.5% 、リスクスコア ≤ -9)正確に分類し；残りの11サンプルは中間に分類された(リスクスコア $-9 < \text{スコア} < 9$)ことを示す。図10Bは、統合AR - リスクスコア[%]は非ARに対してARで有意に高かった($p < 0.0001$)ことを示す。図10Cは、ROC分析がAR予測アッセイの高い感度および特異度を示したことを示す；AUC = 0.93 (95% CI 0.86 ~ 0

50

. 9)。

【図 1 1 A】図 1 1 A ~ D は、フリータイム Q P C R データにおける交絡因子分析およびデータの正規化を示す。1 4 3 の A R および非 A R 成人サンプル (コホート 1) 由来の Q P C R データの、4 3 の拒絶遺伝子に関する主成分分析 (P C A) は、表現型 (図 1 1 B) によるよりもむしろサンプル採取場所 (図 1 1 A) によるサンプル分離を明らかにした。混合 A N O V A による Q P C R データの正規化は、遺伝子発現に対するサンプル採取場所の主要な影響を補正し (図 1 1 C)、A R および非 A R へのサンプルの分離をもたらした (図 1 1 D)。P C A は、4 3 の遺伝子の相対的遺伝子発現値 (d C t 1 8 S) を用いて行った。混合 A N O V A モデルは、ランダムカテゴリー因子としてのサンプル採取場所、R N A ソースおよびチップと、カテゴリー因子としての表現型を用いて構築した。各球は 1 サンプルを表し; 記号はサンプル採取場所を表し (* = U P M C ; = U C L A ; X = C P M C ; # = E M O R Y); この図はまた、生検診断に基づく患者の表現型 (A R ; 非 A R) も表す。

10

【図 1 1 B】図 1 1 A ~ D は、フリータイム Q P C R データにおける交絡因子分析およびデータの正規化を示す。1 4 3 の A R および非 A R 成人サンプル (コホート 1) 由来の Q P C R データの、4 3 の拒絶遺伝子に関する主成分分析 (P C A) は、表現型 (図 1 1 B) によるよりもむしろサンプル採取場所 (図 1 1 A) によるサンプル分離を明らかにした。混合 A N O V A による Q P C R データの正規化は、遺伝子発現に対するサンプル採取場所の主要な影響を補正し (図 1 1 C)、A R および非 A R へのサンプルの分離をもたらした (図 1 1 D)。P C A は、4 3 の遺伝子の相対的遺伝子発現値 (d C t 1 8 S) を用いて行った。混合 A N O V A モデルは、ランダムカテゴリー因子としてのサンプル採取場所、R N A ソースおよびチップと、カテゴリー因子としての表現型を用いて構築した。各球は 1 サンプルを表し; 記号はサンプル採取場所を表し (* = U P M C ; = U C L A ; X = C P M C ; # = E M O R Y); この図はまた、生検診断に基づく患者の表現型 (A R ; 非 A R) も表す。

20

【図 1 1 C】図 1 1 A ~ D は、フリータイム Q P C R データにおける交絡因子分析およびデータの正規化を示す。1 4 3 の A R および非 A R 成人サンプル (コホート 1) 由来の Q P C R データの、4 3 の拒絶遺伝子に関する主成分分析 (P C A) は、表現型 (図 1 1 B) によるよりもむしろサンプル採取場所 (図 1 1 A) によるサンプル分離を明らかにした。混合 A N O V A による Q P C R データの正規化は、遺伝子発現に対するサンプル採取場所の主要な影響を補正し (図 1 1 C)、A R および非 A R へのサンプルの分離をもたらした (図 1 1 D)。P C A は、4 3 の遺伝子の相対的遺伝子発現値 (d C t 1 8 S) を用いて行った。混合 A N O V A モデルは、ランダムカテゴリー因子としてのサンプル採取場所、R N A ソースおよびチップと、カテゴリー因子としての表現型を用いて構築した。各球は 1 サンプルを表し; 記号はサンプル採取場所を表し (* = U P M C ; = U C L A ; X = C P M C ; # = E M O R Y); この図はまた、生検診断に基づく患者の表現型 (A R ; 非 A R) も表す。

30

【図 1 1 D】図 1 1 A ~ D は、フリータイム Q P C R データにおける交絡因子分析およびデータの正規化を示す。1 4 3 の A R および非 A R 成人サンプル (コホート 1) 由来の Q P C R データの、4 3 の拒絶遺伝子に関する主成分分析 (P C A) は、表現型 (図 1 1 B) によるよりもむしろサンプル採取場所 (図 1 1 A) によるサンプル分離を明らかにした。混合 A N O V A による Q P C R データの正規化は、遺伝子発現に対するサンプル採取場所の主要な影響を補正し (図 1 1 C)、A R および非 A R へのサンプルの分離をもたらした (図 1 1 D)。P C A は、4 3 の遺伝子の相対的遺伝子発現値 (d C t 1 8 S) を用いて行った。混合 A N O V A モデルは、ランダムカテゴリー因子としてのサンプル採取場所、R N A ソースおよびチップと、カテゴリー因子としての表現型を用いて構築した。各球は 1 サンプルを表し; 記号はサンプル採取場所を表し (* = U P M C ; = U C L A ; X = C P M C ; # = E M O R Y); この図はまた、生検診断に基づく患者の表現型 (A R ; 非 A R) も表す。

40

【図 1 2】2 6 7 の成人および小児サンプルにおける A R および非 A R 特異的遺伝子の同

50

定のための方法を示す。AR予測のための最終的な17の腎臓AR遺伝子の発見は、合計43の遺伝子、すなわち、本発明者らがこれまでに同定した10の小児AR遺伝子；成人および小児移植拒絶における新規な発見に関する33の候補遺伝子、について行ったマイクロ流体ハイスループットフリーダムQPCRからの267の成人および小児血液サンプル（コホート1、コホート2）由来の遺伝子発現データにて行った。143のARおよび非ARサンプルの成人セットにおいて小児の10遺伝子を確認すると、ARが87.4%で正確に予測された。267のARおよび非ARサンプル（コホート1、コホート2）の成人および小児のデータセットを合わせたもので、新規な発見およびバリデーションを行った。スチューデントのT検定、ANOVAおよびペナルティ付きロジスティック回帰の結果、10の拒絶セットと合わせてAR予測のための17の遺伝子の最終選択を定義した7つのさらなる遺伝子が定義される。等しい事前確率の部分最小二乗判別分析によれば、17の遺伝子は、143サンプルのトレーニングセット（コホート1；AUC = 0.944）ならびに事前の分析には含まれなかった（AUC = 0.948）124サンプルの独立バリデーションセット（コホート2）において、高い感度および特異度でARを予測した。この分析に用いた遺伝子発現データは、サンプル採取場所、RNAソース、および混合ANOVAモデルを用いた実行に関してさらに正規化したフリーダムQPCRプラットフォームから18Sに対するdCt値を表した。

【図13】図13A～Dは、17の遺伝子を用いた、各参加センターにおけるARと非ARの個々の分類を示す。ROC分析は、異なるサンプル採取場所での腎臓AR予測アッセイの成績を評価するために、AART試験に含まれる各移植センターに関して行った。エモリー大学で採取されたARと非ARについて計算されたAUCは0.8765（95%CI 0.7538～0.9993）（図13A；n = 42）；UPMCで採取されたARと非ARについては0.9825（95%CI 0.9608～1.0）（13B；n = 81）、UCLAで採取されたARと非ARについては0.9360（95%CI 0.8648～1.0）（13C、n = 44）、およびCPMCで採取されたARと非ARについては1.0（95%CI 1.0～1.0）であった（図13D、n = 35）。最後のものは、ARサンプルが2つだけというアンバランスなデータセットであり、腎臓AR予測アッセイ成績が過剰適合した可能性がある。各ROC曲線の隣の表は、評価された各センターでのサンプルの分布を示す。

【図14】図14A～Bは、17の遺伝子はp1sDAにより抗体媒介性および細胞媒介性ARを検出し、ARと非ARの分類は移植後の時間に依存しないことを示す。図14Aは、明確な抗体媒介性拒絶（AMR、C4D陽性生検染色、DSA+）だけを有する19名の患者のサブセットにおける固定された17遺伝子腎臓AR予測アッセイモデルによるARの予測確率が、明確な細胞媒介性拒絶（ACR、C4d-およびDSA-）を有する51名の患者のサブセットと比較され；この固定された17遺伝子モデルは、体液性ARと細胞性ARを等しく検出する（14A、p1sDA、p = 0.9906；平均ACR = 80.84% ± 4.4；平均AMR = 80.75% ± 6.6）ことを示す。図14Bは、同様に、17固定遺伝子p1sDAモデルは、非AR患者中のARに関しては一貫して低い予測確率で、また、AR患者群中、一貫して高いAR予測確率で、移植後の時間に依存せずにARを予測したことを示す（図14Bは、p1us SEMの平均予測確率を示す）。3つの移植後時間カテゴリー（0～6か月、6か月～1年、>1年）のうち1つに入るサンプルに関して平均AR予測確率を計算し、スチューデントのT検定により比較し；p値は有意性に達しなかった（p > 0.05）。

【図15】図15A～Cは、17遺伝子の生物学的基礎を示す。経路解析およびネットワーク解析は、QPCRによるARサンプルおよび非ARサンプルにおける遺伝子発現に見られる相関を裏付ける遺伝子の強い生物学的相関を示した。図15Aは、17の遺伝子と有意に（p < 0.05）関連があったのがアポトーシスの調節、免疫表現型および細胞表面タンパク質であったことを示し；図15Bは、インジェニイティ・パスウェイ・アナリシス（Ingenuity Pathway Analyses）（IPA、Qiagen、レッドウッド・シティ、CA）が、17のうち11の遺伝子の、癌、細胞死および細胞生存における一般的役割を

10

20

30

40

50

さらに実証した ($p < 0.05$) ことを示す。図 15C は、さらなるネットワーク解析が 17 のうち 7 つの遺伝子が調節相互作用の単一のネットワークを形成したことを示したことを示す。

【図 16】図 16 は、複数の異なるタイプの臓器移植拒絶にわたる共通の拒絶モジュールに相当する、臓器移植拒絶において過剰発現されることが判明した 12 の遺伝子を示す。

【発明を実施するための形態】

【0021】

本発明者らは、腎臓移植を受けた個人が移植臓器の急性拒絶 (AR) を受けているかどうか、またはその後受けるかどうかを決定することができる遺伝子発現プロファイルを見つけた。これらの遺伝子発現プロファイルは、レシピエントの年齢、移植センター、RNA ソース、アッセイ、末期腎疾患の原因、併存症、免疫抑制使用などに依存しない。本明細書に記載の発明は、腎臓同種移植を受けた個人において AR または非 AR を評価するための方法、ならびに腎臓移植における AR の処置のための個人を同定する方法を提供する。本発明はまた、これらのシステムの成分としてのマイクロアレイチップの使用を含む、腎臓同種移植において AR を評価するためのシステムを記載する。本発明はさらに、腎臓同種移植を受けた個人において AR および AR の確率を評価するための、これらのシステムに基づくキットを提供する。

【0022】

定義

本明細書を説明する目的で、以下の定義が当てはまり、適切な場合、単数で用いられる用語はその複数も含み、逆もそうである。以下に示される定義が、引用することにより本明細書の一部とされる文書と矛盾する場合には、以下に示される定義が優位とすべきである。

【0023】

「急性拒絶」「急性同種移植片拒絶」または「AR」は、移植組織が免疫学的に外来となる場合の、組織/臓器移植レシピエントの免疫系による拒絶である。AR は、レシピエントの免疫細胞による移植組織への浸潤を特徴とし、免疫細胞はそれらのエフェクター機能を遂行し、移植組織を破壊する。AR はまた、抗体媒介性拒絶 (AMR) と呼ばれる診断であるドナー特異的抗体の生成も特徴とし得る。AR は、超急性、急性、境界急性、または準臨床 AR としてさらに分類することができる。超急性拒絶の誘導は一般に急速で、人では一般に移植手術後数分から数時間内に起こる。AR の誘導は人では一般に、移植手術後数ヶ月以内、多くの場合にはおよそ 6 ~ 12 か月で起こる。境界急性および準臨床 AR は、軽度炎症性アロ応答の結果である。一般に、AR は、ラパマイシン、シクロスポリン A、抗 CD40L モノクローナル抗体などの免疫抑制薬で処置、阻害、または抑制することができる。

【0024】

「非急性拒絶」または「非 AR」または「安定」または「STA」は、本明細書では互換的に使用される。非 AR / STA は、移植後 AR が低リスクまたはリスク無しの患者を表す。非 AR は、組織移植レシピエントにとって免疫学的に外来である移植組織の長期移植片生着を特徴とし得る。

【0025】

用語「腎臓同種移植」は、ある個人から別の個人への腎臓移植を意味する。

【0026】

本明細書で使用する場合、「遺伝子」は、エキソン配列および (場合により) イントロン配列を含むポリペプチドをコードするオープンリーディングフレームを含んでなる核酸を意味する。用語「イントロンとは、タンパク質に翻訳されず、DNA 分子中のエキソン間に一般に見られる、所与の遺伝子中に存在する DNA 配列を意味する。さらに、遺伝子は、場合によりその天然プロモーター (すなわち、非組換え細胞において遺伝子のエキソンおよびイントロンが作動可能なように連結されているプロモーター)、および関連の調節配列を含んでよく、かつ、AUG 開始部位の上流配列、非翻訳リーダー配列、シグナル

配列、下流非翻訳配列、転写開始配列および終結配列、ポリアデニル化シグナル、翻訳開始および終結配列、リボソーム結合部位などを含んでも含まなくてもよい。

【0027】

用語「参照」は、観測値が比較できる既知値または既知値のセットを意味する。一実施形態では、参照は、移植片生着表現型における遺伝子の遺伝子発現の値（またはレベル）を意味する。別の実施形態では、参照は、移植片欠損表現型における遺伝子の遺伝子発現の値（またはレベル）である。

【0028】

本明細書で使用する場合、「参照発現ベクター」は、参照標準を意味する。一実施形態では、参照発現ベクターは、ARサンプルに関して、所与の移植センターにおいて各発現遺伝子に対して作出された参照標準である。別の実施形態では、参照発現ベクターは、非ARサンプルに関して、所与の移植センターにおいて各発現遺伝子に対して作出された参照標準である。別の実施形態では、参照発現ベクターは、ARサンプルに関して、移植センター間の各発現遺伝子に対して作出された参照標準である。別の実施形態では、参照発現ベクターは、非ARサンプルに関して、移植センター間の各発現遺伝子に対して作出された参照標準である。

10

【0029】

「個人」または「被験者」は、「患者」であり得る。「患者」は、治療医の医療行為の下にある「個人」を意味する。患者は男性であっても女性であってもよい。一実施形態では、患者は、腎臓移植を受けている。別の実施形態では、患者は腎臓移植を受け、臓器拒絶を受けている。さらに別の実施形態では、患者は腎臓移植を受け、ARを受けている。

20

【0030】

「患者部分集団」およびその文法的変形形態は、本明細書で使用する場合、ある患者のサブセットを他から、それが属するより広い疾病カテゴリーにおいて識別する1以上の明瞭な測定可能なおよび/または識別可能な特徴を有することを特徴とする患者のサブセットを意味する。

【0031】

用語「サンプル」は、本明細書で使用する場合、ゲノム情報を含む、個人から得られるまたは個人に由来する組成物を意味する。一実施形態では、サンプルは全血である。一実施形態では、サンプルは血液である。別の実施形態では、サンプルは末梢血白血球である。別の実施形態では、サンプルは末梢血単核細胞である。別の実施形態では、サンプルは組織生検である。別の実施形態では、サンプルは、移植臓器由来の組織生検である。別の実施形態では、サンプルは、レシピエントにおける移植前の臓器由来の組織生検である。

30

【0032】

本明細書で使用する場合、「マイクロアレイ」は、集中的場所におけるヌクレオチド配列のコレクションの配置を意味する。アレイは、ガラス、プラスチック、またはケイ素から構成される表面などの固相基板であり得る。これらのヌクレオチド配列は、DNA、RNA、またはその任意の順列であり得る。ヌクレオチド配列はまた、遺伝子由来の部分配列、プライマー、全遺伝子配列、非コード配列、コード配列、公開配列、既知配列、または新規配列であり得る。

40

【0033】

「予測する」および「予測」とは、本明細書で使用する場合、その結果が100%の確実性で起こることを意味するものではなく、その結果が起こらないことより起こる可能性が高いことを意味するものとする。「予測する」または「予測を行う」ために採られる行為は、結果が起こらないことより起こる可能性が高い確率の決定を含み得る。このような決定または予測を行うためには、本明細書に記載の複数の因子の評価を使用することができる。

【0034】

「比較する」または「比較」とは、第1の分析の結果を第2および/または第3の分析の結果と何らかの方法で関連させることを意味する。例えば、1の分析の結果を用い、そ

50

の結果が第3の結果よりも第2の結果に類似性が高いと分類してもよい。ある個人由来の生体サンプルのAR評価の実施形態に関して、それらの結果を用いて、その個人がAR応答を受けているかどうか決定してもよい。ある個人由来の生体サンプルの非AR評価の実施形態に関して、それらの結果を用いて、その個人が非AR応答を受けているかどうか決定してもよい。

【0035】

用語「評価」および「決定」は、任意の測定形態を意味して互換的に使用され、定量的および定性的両方の測定を含む。例えば、「評価」は相対であっても絶対的であってもよい。

【0036】

用語「診断」は、本明細書では、分子状態または病的状態、疾患の同定または分類を意味して使用される。例えば、「診断」は、臓器拒絶の同定を意味する場合がある。「診断」はまた、ARなどの臓器拒絶の特定のサブタイプの分類を意味する場合もある。

【0037】

本明細書で使用する場合、「処置」は、処置される個人の自然経過を変化させる試みでの臨床的介入を意味する。処置の望ましい効果には、疾患またはその病態もしくは症状の発生または再発の予防、疾患の病態または症状の緩和、疾患の任意の直接的または間接的な病理学的帰結の軽減、疾患進行速度の減速、病的状態の改善または消散、および予後改善の達成が含まれる。特定の実施形態では、処置は、個人におけるARの疾患進行速度の減速、病的状態の改善または消散、および予後改善の達成を意味する。いくつかの実施形態では、処置は、1以上の治療薬の治療計画の投与を修正または変更する臨床的介入を意味する。

【0038】

「約」という場合、本明細書の値またはパラメータは、その値またはパラメータそれ自体に対する具体例を含む（および記載する）。例えば、「約X」という記載は、「X」の記載を含む。用語「約」は、所与の値が所望の結果に影響を及ぼさずに終点より「少し上」または「少し下」であってもよいことを示すことにより、数値範囲終点に柔軟性を持たせるために使用される。濃度、量、および他の数値データは、本明細書では範囲形式で表現または提示される場合がある。このような範囲形式は単に便宜上および簡潔さのために使用されるものであり、範囲の限界として明示された数値を含むだけでなく、各数値および部分範囲が明示されているかのうように個々の数値またはその範囲内に包含される部分範囲の全ても含むと柔軟に解釈されるべきであると理解されるべきである。

【0039】

本明細書に記載の本発明の態様および実施形態は、態様および実施形態「を含んでなる」、「からなる」および「から本質的になる」を含むと理解される。本明細書に記載の全ての組成物および本明細書に記載の組成物を用いた全ての方法に関して、これらの組成物は、挙げられている成分もしくは工程を含んでなり得るか、または挙げられている成分もしくは工程「から本質的になり得る」かのいずれかである。組成物が挙げられている成分「から本質的になる」と記載されている場合、本組成物は挙げられている成分を含み、処置される病態に実質的に影響を及ぼさない他の成分を含み得るが、明示的に挙げられている成分以外に、治療される病態に実質的に影響しない他のいずれの成分も含まず；または本組成物治療される病態に実質的に影響を及ぼす、挙げられているもの以外の付加的成分を含む場合には、本組成物は処置される病態に実質的に影響を及ぼすに十分な濃度もしくは量の付加的成分を含まない。方法が挙げられている工程「から本質的になる」と記載されている場合、その方法は挙げられている工程を含み、かつ、処置される病態に実質的に影響を及ぼさない他の工程を含んでよいが、その方法は明示的に挙げられている工程以外に、処置される病態に実質的に影響を及ぼす他のいずれの工程も含まない。限定されない具体例として、組成物がある成分「から本質的になる」と記載されている場合、その組成物は、任意の量の薬学上許容可能な担体、ビヒクル、または希釈剤、および処置される病態に実質的に影響を及ぼさない他のこのような成分をさらに含んでよい。

10

20

30

40

50

【 0 0 4 0 】

本明細書および添付の特許請求の範囲で使用する場合、単数形「a」、「an」および「the」は、そうではないことが明示されない限り、複数の指示対象を含む。

【 0 0 4 1 】

一般技術

そうではないことが定義されない限り、本明細書で使用される技術用語および科学用語は、本発明が属する技術分野の熟練者により共通に理解されているものと同じ意味を有する。

【 0 0 4 2 】

本発明の実施は、特に断りのない限り、タンパク質生物学、タンパク質化学、分子生物学（組換え技術を含む）、微生物学、細胞生物学、生化学、および免疫学の従来技術を用い、これらは当業者の技術の範囲内である。このような技術は、“Molecular Cloning: A Laboratory Manual”，第2版（Sambrook et al., 1989）；“Current Protocols in Molecular Biology”（Ausubel et al.編，1987，定期更新）；“PCR: The Polymerase Chain Reaction”，（Mullis et al.編，1994）；およびSingleton et al., Dictionary of Microbiology and Molecular Biology, 第2版, J. Wiley & Sons (New York, N.Y. 1994)などの文献に十分に説明されている。

10

【 0 0 4 3 】

腎臓同種移植レシピエント

腎臓同種移植レシピエントは、いずれの年齢であってもよい。いくつかの実施形態では、個人は小児である。一実施形態では、小児は乳児である。別の実施形態では、小児は幼児である。他の実施形態では、個人は23歳未満の若年成人である。いくつかの実施形態では、個人はおよそ1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、または22歳である。さらなる実施形態では、個人は23歳を超える成人である。いくつかの実施形態では、個人は、およそ23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、67、68、69、60、61、62、63、64、65、66、67、68、69、70、71、72、73、74、75、76、77、78、79、80、81、82、83、84、85、86、87、88、89、90、91、92、93、94、95、96、97、98、99、または100歳である。一実施形態では、腎臓同種移植レシピエントは女性である。別の実施形態では、腎臓同種移植レシピエントは男性である。

20

30

【 0 0 4 4 】

腎臓移植手術/手術は、特殊設計の治療施設または移植センターで行うことができる。移植センターは世界のどこに位置していてもよい。一実施形態では、移植センターはアメリカ合衆国にある。いくつかの実施形態では、移植センターはエモリー大学（アトランタ、ジョージア州）、カルフォルニア大学ロサンゼルス校（ロサンゼルス、CA）、ピッツバーグ大学（ピッツバーグ、PA）、カリフォルニア・パシフィック・メディカルセンター（サンフランシスコ、CA）、またはカルフォルニア大学サンフランシスコ校（サンフランシスコ、CA）である。他の実施形態では、移植センターはヨーロッパにある。一実施形態では、移植センターは、大学病院（パルセロナ、スペイン）である。さらなる実施形態では、移植センターはメキシコにある。一実施形態では、移植センターは、Laboratorio de Investigacion en Nefrologia、Hospital Infantil de Mexico（メキシコシティ、メキシコ）である。

40

【 0 0 4 5 】

腎臓同種移植レシピエントからの生体サンプルの採取

生体サンプルは、腎臓同種移植を受けた個人から採取する。いくつかの実施形態では、腎臓同種移植レシピエントは、ARの外面的症状はない。他の実施形態では、腎臓同種移

50

植レシピエントは、A Rの症状を示す。限定されるものではないが、全血、血液、血清、血漿、尿、粘液、唾液、脳脊髄液、組織、生検およびそれらの組合せを含むいずれのタイプの生体サンプルを採取してもよい。一実施形態では、生体サンプルは全血である。一実施形態では、生体サンプルは血液である。いくつかの実施形態では、血液サンプルは末梢血である。別の実施形態では、生体サンプルは末梢血単核細胞である。いくつかの実施形態では、生体サンプルは末梢血リンパ球である。いくつかの実施形態では、生体サンプルは組織生検である。

【0046】

腎臓同種移植レシピエント由来の生体サンプルの採取は、臓器移植後のいずれの時点で行ってもよい。いくつかの実施形態では、生体サンプルは、P A X g e n e (商標)チューブ(Q i a g e nから入手可能)中に採取することができる。他の実施形態では、生体サンプルは、R N A分解を防ぐためにR Nアーゼ阻害剤の入ったコレクションチューブ中に採取することができる。いくつかの実施形態では、生体サンプルは、慣例のプロトコール監視検査(protocol surveillance examination)の際に採取する。他の実施形態では、生体サンプルは、治療臨床医にその個人がA R応答を受けていると疑う理由が合った場合に採取される。

10

【0047】

腎臓同種移植レシピエントから採取される生体サンプルは、A Rサンプルは非A Rサンプルの参照を作成する場合、同じ患者由来の同時的腎臓同種移植生検と対となるようにすればよい。一般に、腎臓同種移植生検は、生体サンプル採取の48時間以内にレシピエントから採取する。いくつかの実施形態では、生検は、移植時に採取する。他の実施形態では、生検は、移植後24か月以内に採取する。一実施形態では、生検は、移植後約3か月；移植後約6か月；移植後約12か月；移植後約18か月；または移植後約24か月で採取してよい。生検および/または生体サンプルは移植後のいずれの時点で採取してもよいので、これらの時点は限定として見なされるべきでない。むしろこれらの時点は、慣例の監視が大多数の腎臓同種移植レシピエントにおいて行われる可能性の最も高い移植後の期間を示すために提供される。さらに、これらの時点は、A R応答が生じる可能性の最も高い移植後の期間を示す。

20

【0048】

採取された各腎臓同種移植生検は、B a n f f分類システム(Solez, K. et al. Am. J. Transplant., 2008, 8, 753-760; Mengel, M. et al. Am. J. Transplant. 2012, 12, 563-570)に従ってスコア化してもよい。このシステムは、観察された腎臓生検サンプルの病状を正常病歴、超急性拒絶、境界性変化、急性拒絶、慢性同種移植腎症、およびその他の変化として分類する。このB a n f f分類は腎臓移植病状の標準を設定し、新規の抗拒絶薬の国際的臨床試験において広く使用されている。本明細書に記載されるように、「急性拒絶」(A R)は、B a n f f尿細管炎スコア(t)が1以下かつ間質性浸潤物スコアが0以下の生検サンプルに関して定義され；「安定」(「S T A」)/「非A R」は、A Rの不在(非A R)または他の任意の実質的病状の不在を示す生検サンプルに関して定義され；また「その他」は、B a n f fグレード分類のA Rの不在を示すが、慢性同種移植損傷、慢性カルシニユリン阻害剤毒性、B Kウイルス感染、または他の移植片損傷のB a n f f判定基準のいずれかを満たすサンプルに関して定義される。

30

40

【0049】

生体サンプルにおける遺伝子発現の評価

腎臓同種移植レシピエントから採取された生体サンプルは、A R応答を受けている個人で差次的に発現される遺伝子のレベルを評価するために使用することができる。遺伝子発現を測定する種々の技術は当業者に公知である。1つの非限定的方法は、採取した生体サンプルからR N Aを抽出し、c D N Aを合成することである。c D N Aは、標的遺伝子(すなわち、A R応答を受けている個人において差次的に発現される遺伝子)に特異的なプライマーまた標識プライマーを用いて増幅し、次に、定量的ポリメラーゼ連鎖反応(q P C R)を用いて分析することができる。B i o M a r k (フリーダタイム、サザンサンフ

50

ランシスコ、CA)またはABI via7(Life Technologies、フォスターシティ、CA)などのqPCRプラットフォームが使用可能である。

【0050】

いくつかの実施形態では、遺伝子特異的プライマーまたはdNTP、好ましくはdNTPのいずれか一方を、合成されるcDNAが標識されるように標識する。標識とは、それらの存在が、シグナル生成系のメンバーを含んでなり、従って、直接的に、またはシグナル生成系の1以上の付加的メンバーとの組合せ作用を介して検出可能である。直接検出可能な標識の例としては、ヌクレオチドモノマー単位、例えば、dNTPまたはプライマーのモノマー単位に組み込まれた、通常には共有結合された同位元素部分および蛍光部分が含まれる。対象とする同位元素部分または標識としては、 ^{32}P 、 ^{33}P 、 ^{35}S 、 ^{125}I などが含まれる。対象とする蛍光部分または標識としては、クマリンおよびその誘導体、例えば、7-アミノ-4-メチルクマリン、アミノクマリン、ボディパイFLなどのボディパイ色素、カスケードブルー、フルオレセインおよびその誘導体、例えば、フルオレセインイソチシアネート、オレゴングリーン、ローダミン色素、例えば、テキサスレッド、テトラメチルローダミン、エオジンおよびエリスロシン、シアニン色素、例えば、Cy3およびCy5、大ランタニドイオンの環状キレート、例えば、クワンタム(quantum)色素(商標)、蛍光エネルギー移動色素、例えば、チアゾールオレンジ-エチジウムヘテロ二量体、TOTABなどが含まれる。標識はまた、同じ系の1以上の付加的メンバーと協調して作用して検出可能なシグナルをもたらすシグナル生成系のメンバーであってもよい。このような標識の例は、リガンド、例えば、ビオチン、フルオレセイン、ジゴキシゲニン、抗原、多価陽イオン、キレーター基などのような特異的結合対のメンバーであり、これらのメンバーはシグナル生成系の付加的メンバーと特異的に結合し、その付加的メンバーは直接的または間接的のいずれかで検出可能なシグナル、例えば、蛍光部分または基質を発色生成物に変換することができる酵素部分とコンジュゲートさせた抗体、例えば、アルカリ性ホスファターゼコンジュゲート抗体などを提供する。また、標識核酸は標識プライマーの存在下でPCRを行うことによっても作製することができる。米国特許第5,994,076号は、単に修飾プライマーおよびそのdNTPのその教示のために引用することにより本明細書の一部とされる。

10

20

【0051】

AR応答を受けている腎臓同種移植レシピエントにおいて差次的に発現される遺伝子の例を表1に示す。一実施形態では、差次的発現遺伝子は、p値が0.05以下、または偽陽性率が5%以下により示され、かつ、有意であると見なすことができ、予測モデルの構築に使用することができる。別の実施形態では、絶対的変化倍率が1.5以上かつp値が0.05以下であるか、または偽陽性率が5%以下である遺伝子が有意であると見なされ、予測モデルの構築に使用することができる。種々のタイプのソフトウェアが統計分析に使用可能である。このようなソフトウェアの一例は、Partek Genomics Suiteである。これらの遺伝子を統計分析にかけてARの検出および/または予測のためのロバストなモデルを選択することができる。ペナルティ付きロジスティック回帰、サポートベクターマシン、および等しい事前確率の部分最小二乗判別分析などの種々の分類モデルが使用可能である。実施例にさらに詳細に示すように、主成分分析を用いて生のqPCRデータを可視化することができ、ANOVAおよび学生t検定で有意に差次的に発現される遺伝子を検出することができ、また、Shrinking Centroidsを適用してARサンプルと非ARサンプル間を識別する遺伝子を特定することができる。

30

40

【0052】

表1に挙げられている遺伝子から、患者の年齢、移植センター、RNAソース、アッセイ、末期腎疾患の原因、併存症、および/または免疫抑制使用とは無関係に患者をARまたは非ARとして分類することができる17の遺伝子のサブセットが同定された。この17遺伝子のセットは、小児患者においてARの指標となることが従前に示されていた10遺伝子の組合せ(CFLAR、DUSP1、ITGAX、RNF130、PSEN1、N

50

KTR、RYBP、NAMPT、MAPK9、およびIFNGR1)、新たに定義された成人患者においてARの指標となる6つの遺伝子(CEACAM4、RHEB、GZMK、RARA、SLC25A37、およびEPOR)、およびレチノイドX受容体(RXRA)から構成される。これらの遺伝子の配列は付表AおよびBに示されている。本明細書に開示される遺伝子は、腎臓同種移植を受けた個人においてARを診断する種々方法のため、処置のための患者を選択するため、ならびに本明細書に記載の他の使用のために使用することができる。

【0053】

【表1】

遺伝子記号	Ensembl ID	Entrez ID	定義	TaqMan アッセイ ID
RYBP	ENSG00000163602	23429	RING1 および YY1 結合タンパク質	Hs00171928_m1
RNF130	ENSG00000113269	55819	リングフィンガータンパク質 130	Hs00218335_m1
PSEN1	ENSG00000080815	5663	プレセニン 1	Hs00997789_m1
NKTR	ENSG00000114857	4820	ナチュラルキラー腫瘍認識配列	Hs00234637_m1
NAMPT	ENSG00000105835	10135	ニコチンアミドホスホリボシルトランスフェラーゼ	Hs00237184_m1
MAPK9	ENSG00000050748	5601	マイトジェン活性化タンパク質キナーゼ 9	Hs00177102_m1
ITGAX	ENSG00000140678	3687	インテグリン, αX (補体成分 3 受容体 4 サブユニット)	Hs00174217_m1
IFNGR1	ENSG00000027697 LRG_66	3459	インターフェロン γ 受容体 1	Hs00166223_m1
DUSP1	ENSG00000120129	1843	二重特異的ホスファターゼ 1	Hs00610256_g1
CFLAR	ENSG00000003402	8837	CASH および FADD 様アポトーシスレギュレーター	Hs00236002_m1
SLC25A37	ENSG00000147454	51312	溶質担体ファミリー-25, メンバー-37	Hs00249769_m1
RXRA	ENSG00000186350	6256	レチノイド X 受容体, α	Hs01067640_m1
RHEB	ENSG00000106615	6009	脳に豊富な Ras ホモログ	Hs02858186_m1
RARA	ENSG00000131759	5914	レチノイン酸受容体, α	Hs00940446_m1
GZMK	ENSG00000113088	3003	グランザイム K(グランザイム 3; トリプターゼ II)	Hs00157875_m1
EPOR	ENSG00000187266	2057	エリスロポエチン受容体	Hs00959427_m1
CEACAM4	ENSG00000105352	1089	癌胎児性抗原関連細胞接着分子 4	Hs00156509_m1

10

20

30

40

NFE2	ENSG00000123405	4778	核因子(赤血球由来 2), 45kDa	Hs00232351_m1
MPP1	ENSG00000130830	4354	膜タンパク質, パルミト イル化 1, 55kDa	Hs00609971_m1
MAP2K3	ENSG00000034152	5606	ミトジェン活性化タ ンパク質キナーゼキナ ーゼ 3	Hs00177127_m1
IL2RB	ENSG00000100385	3560	インターロイキン 2 受 容体, β	Hs01081697_m1
FOXP3	ENSG00000049768 LRG_62	50943	フォーヘッドボックス P3	Hs00203958_m1
CXCL10	ENSG00000169245	3627	ケモカイン(C-X-Cモチ ーフ)リガンド 10	Hs00171042_m1
C1orf38	ENSG00000130775	9473	染色体 1 オープンリー ディングフレーム 38	Hs00985482_m1
GZMB	ENSG00000100453	3002	グランザイム B	Hs00188051_m1
ABTB1	ENSG00000114626	80325	アンキリンリピートお よび BTB (P02) ドメイ ン含有 1	Hs00261395_m1
IL7R	ENSG00000168685 LRG_74	3575	インターロイキン 7 受 容体	Hs00233682_m1
STAT3	ENSG00000168610	6774	シグナル伝達兼転写活 性化因子 3(急性期応答 因子)	Hs01047580_m1
YPEL3	ENSG00000090238	83719	イッピー様 3(ショウジ ョウバエ)	Hs00368883_m1
PFN1	ENSG00000108518	5216	プロフィリン 1	Hs00748915_s1
IL7	ENSG00000104432	3574	インターロイキン 7	Hs00174202_m1
PCTP	ENSG00000141179	58488	ホスファチジルコリン 転移タンパク質	Hs00221886_m1
GBP2	ENSG00000162645	2634	グアニル酸結合タンパ ク質 2, インターフェロ ン誘導性	Hs00894837_m1
GBP1	ENSG00000117228	2633	グアニル酸結合タンパ ク質 1, インターフェロ ン誘導性, 67kDa	Hs00977005_m1
ANK1	ENSG00000029534	286	アンキリン 1, 赤血球	Hs00986657_m1
INPP5D	ENSG00000168918	3635	イノシトールポリリン 酸-5-ホスファターゼ, 145kDa	Hs00183290_m1

10

20

30

40

CHST11	ENSG00000171310	50515	炭水化物(コンドロイチン4)スルホトランスフェラーゼ 11	Hs00218229_ml
TNFRSF1A	ENSG00000067182 LRG_193	7132	腫瘍壊死因子受容体スーパーファミリー, メンバー1A	Hs01042313_ml
LYST	ENSG00000143669	1130	リソソーム輸送レギュレーター	Hs00915897_ml
ADAMS	ENSG00000151651	101	ADAM メタロペプチダーゼドメイン 8	Hs00923282_g1
RUNX3	ENSG00000020633	864	runt 関連転写因子 3	Hs00231709_ml
PSMB9	ENSG00000240065 ENSG00000239836 ENSG00000243958 ENSG00000243594 ENSG00000243067 ENSG00000243067 ENSG00000242711 ENSG00000240508 ENSG00000240118	5698	プロテアソーム(プロソーム, マクロペイ)サブユニット, β 型, 9 (高分子量多機能性ペプチダーゼ 2)	Hs00544762_ml
ISG20	ENSG00000172183	3669	インターフェロン刺激エキソヌクレアーゼ遺伝子 20kDa	Hs00158122 ml

10

20

【 0 0 5 4 】

遺伝子発現を測定する別の非限定的方法はノーザンブロット法である。タンパク質をコードする遺伝子の遺伝子発現レベルはまた、ウエスタンブロット法などのタンパク質量法を用いて決定することもできる。差次的発現遺伝子のレベルを測定するためのプロテオミクスアッセイの使用も本明細書に包含される。当業者は、遺伝子発現のレベルを測定するために標準的なプロテオミクスアッセイをどのように使用するかを知っている。

30

【 0 0 5 5 】

参照発現ベクター

本発明は、年齢、移植センター、RNAソース、アッセイ、末期腎疾患の原因、併存症、および/または免疫抑制使用に依存しない参照発現ベクターの作出を提供する。これらの参照発現ベクターの使用は、Partekなどの市販のソフトウェアパッケージまたはRなどのオープンソースソフトウェアによって一般に必要とされるバッチ効果の除を必要としない。

【 0 0 5 6 】

データに対する有意な変量効果は移植センターの違いにより推定される。これらの変量効果は、種々の移植センターでの生体サンプル採取プロトコルおよび免疫抑制計画の違いから生じる。よって、個々の移植センター特異的AR予測モデルは、全移植センターの単一AR予測モデルよりも正確である。

40

【 0 0 5 7 】

実施例および付表に例示されるように、所与の移植センターに関して、AR予測モデルは、各遺伝子に対する、その移植センターで採取されたARサンプルの第1の参照発現ベクター、および各遺伝子に対する、同じ移植センターで採取された非ARサンプルの第2の参照発現ベクターを作出することによって開発することができる。参照発現ベクターを作出するために使用されるサンプルは、同種移植生検を用いて分類してもよい。次に、同じ移植センターで腎臓同種移植レシピエントから採取された生体サンプル(すなわち、「未知の」サンプル)から得られた差次的発現遺伝子の発現レベルをARサンプルおよび非ARサンプルの2つの参照発現ベクターと比較することができる。LineagePro

50

filerの修正バージョンであるkSASなどのコンピュータプログラムを用い、ARのリスクまたは非ARのリスクを示す各評価遺伝子にカテゴリー的値またはスコアおよび/または数的値またはスコア(付表Cに示されるソースコード)を割り付けることができる。複数の遺伝子セットモデルが使用可能である。複数の遺伝子セットモデルを使用する利点は、明瞭に異なる値またはスコアを各遺伝子セットに割り付けられ、従って、単一の遺伝子モデルに基づくバイアスのリスクを最小限にすることである。

【0058】

一実施形態では、ARの診断のために2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、または17の参照発現ベクターが存在する。関連の実施形態では、非ARの診断のために2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、または17の参照発現ベクターが存在する。1つの特定の実施形態では、ARの診断のための17の参照発現ベクターと非ARの診断のための17の参照発現ベクターが存在する。別の特定の実施形態では、ARの診断のための16の参照発現ベクターと非ARの診断のための16の参照発現ベクターが存在する。別の特定の実施形態では、ARの診断のための15の参照発現ベクターと非ARの診断のための15の参照発現ベクターが存在する。別の特定の実施形態では、ARの診断のための12の参照発現ベクターと非ARの診断のための12の参照発現ベクターが存在する。

10

【0059】

一実施形態では、参照発現ベクターを作出るために、生体サンプルを採取し、未知のサンプルの分析の前に12遺伝子モデルセットを用いてプロファイリングを行う。12遺伝子モデルの例を表2に示す。別の実施形態では、生体サンプルを採取し、未知のサンプルの分析の前にBASP1、CD6、CD7、CXCL10、CXCL9、INPP5D、ISG20、LCK、NKG7、PSMB9、RUNX3、およびTAP1を含んでなる12遺伝子モデルセットを用いてプロファイリングを行う。一実施形態では、12遺伝子セットは、CFLAR、PSEN1、CEACAM4、NAMPT、RHEB、GZMK、NKTR、DUSP1、RARA、ITGAX、SLC25A37、およびEPORから構成される。一実施形態では、12遺伝子セットは、CFLAR、PSEN1、CEACAM4、NAMPT、RHEB、GZMK、NKTR、DUSP1、ITGAX、SLC25A37、RXRA、およびEPORから構成される。一実施形態では、12遺伝子セットは、CFLAR、PSEN1、CEACAM4、RHEB、GZMK、NKTR、DUSP1、RARA、ITGAX、SLC25A37、RXRA、およびEPORから構成される。一実施形態では、12遺伝子セットは、CFLAR、PSEN1、CEACAM4、NAMPT、GZMK、NKTR、DUSP1、ITGAX、SLC25A37、RYBP、RXRA、およびEPORから構成される。一実施形態では、12遺伝子セットは、CFLAR、PSEN1、CEACAM4、NAMPT、GZMK、NKTR、DUSP1、RARA、ITGAX、SLC25A37、RXRA、およびEPORから構成される。一実施形態では、12遺伝子セットは、CFLAR、PSEN1、CEACAM4、NAMPT、GZMK、NKTR、DUSP1、RARA、ITGAX、SLC25A37、RYBP、RXRA、およびEPORから構成される。一実施形態では、12遺伝子セットは、CFLAR、PSEN1、CEACAM4、NAMPT、GZMK、NKTR、DUSP1、RARA、ITGAX、SLC25A37、RXRA、およびEPORから構成される。一実施形態では、12遺伝子セットは、CFLAR、MAPK9、PSEN1、CEACAM4、GZMK、NKTR、DUSP1、RARA、ITGAX、SLC25A37、RYBP、RXRA、およびEPORから構成される。一実施形態では、12遺伝子セットは、CFLAR、MAPK9、PSEN1、CEACAM4、GZMK、NKTR、D

20

30

40

50

USP1、RARA、ITGAX、SLC25A37、RYBP、およびEPORから構成される。一実施形態では、12遺伝子セットは、CFLAR、MAPK9、PSEN1、CEACAM4、NAMPT、GZMK、NKTR、ITGAX、SLC25A37、RYBP、RXRA、およびEPORから構成される。一実施形態では、12遺伝子セットは、CFLAR、MAPK9、PSEN1、CEACAM4、NAMPT、GZMK、NKTR、DUSP1、RARA、ITGAX、SLC25A37、およびEPORから構成される。一実施形態では、12遺伝子セットは、BASP1、CD6、CD7、CXCL10、CXCL9、INPP5D、ISG20、LCK、NKG7、PSMB9、RUNX3、およびTAP1から構成される。

【0060】

【表2】

表2: 腎臓AR予測アッセイ成績—選択された17遺伝子から選択された14の12遺伝子モデル

モデル	成人トレーニングセット (n=32)	小児+成人テストセット (n=68)	記載
1	90.63%	88.24%	CFLAR, PSEN1, CEACAM4, NAMPT, RHEB, GZMK, NKTR, DUSP1, RARA, ITGAX, SLC25A37, EPOR
2	90.63%	86.27%	CFLAR, PSEN1, CEACAM4, NAMPT, RHEB, GZMK, NKTR, DUSP1, ITGAX, SLC25A37, RXRA, EPOR
3	90.63%	86.27%	CFLAR, PSEN1, CEACAM4, RHEB, GZMK, NKTR, DUSP1, RARA, ITGAX, SLC25A37, RXRA, EPOR
4	90.63%	84.31%	CFLAR, PSEN1, CEACAM4, NAMPT, GZMK, NKTR, DUSP1, ITGAX, SLC25A37, RYBP, RXRA, EPOR
5	90.63%	82.35%	CFLAR, MAPK9, PSEN1, CEACAM4, GZMK, NKTR, DUSP1, RARA, SLC25A37, RYBP, RXRA, EPOR
6	90.63%	80.39%	CFLAR, PSEN1, CEACAM4, GZMK, NKTR, DUSP1, RARA, ITGAX, SLC25A37, RYBP, RXRA, EPOR
7	90.63%	80.39%	CFLAR, MAPK9, PSEN1, CEACAM4, NAMPT, GZMK, NKTR, DUSP1, ITGAX, SLC25A37, RXRA, EPOR
8	90.63%	80.39%	CFLAR, PSEN1, CEACAM4, NAMPT, GZMK, NKTR, DUSP1, RARA, ITGAX, SLC25A37, RYBP, EPOR
9	90.63%	80.39%	CFLAR, PSEN1, CEACAM4, NAMPT, GZMK, NKTR, DUSP1, RARA, ITGAX, SLC25A37, RXRA, EPOR
10	90.63%	78.43%	CFLAR, MAPK9, PSEN1, CEACAM4, GZMK, NKTR, DUSP1, RARA, ITGAX, SLC25A37, RXRA, EPOR
11	90.63%	78.43%	CFLAR, MAPK9, PSEN1, CEACAM4, GZMK, NKTR, DUSP1, RARA, ITGAX, SLC25A37, RYBP, EPOR
12	90.63%	78.43%	CFLAR, MAPK9, PSEN1, CEACAM4, NAMPT, GZMK, NKTR, ITGAX, SLC25A37, RYBP, RXRA, EPOR
13	90.63%	78.43%	CFLAR, MAPK9, PSEN1, CEACAM4, NAMPT, GZMK, NKTR, DUSP1, RARA, ITGAX, SLC25A37, EPOR
14	N/A	N/A	BASP1, CD6, CD7, CXCL10, CXCL9, INPP5D, ISG20, LCK, NKG7, PSMB9, RUNX3, and TAP1

【0061】

これらのサンプルのqPCRプロファイルを得た後に、全ARサンプルおよび非ARサンプルの平均発現を別に採り、アッセイした全遺伝子の2列参照を作成する。あるいは、

10

20

30

40

50

個々のサンプルの代わりにプールしたRNA参照を使用しても十分であり得る。データを、第1列が遺伝子識別を含み、第2列がAR参照を含み、第3列が非AR参照を含む3列参照ファイルとして保存する。この参照に用いた元のサンプルを再分析すれば、例えば、ARサンプルと非ARサンプル間の不十分な分類スコアのためにこれらの参照サンプル間に有意な変動が存在するかどうかを判定することができる。

【0062】

別の実施形態では、参照発現ベクターを作出するために、生体サンプルを採取し、未知のサンプルの分析の前に、CEACAM4、CFLAR、DUSP1、ITGAX、RNFI30、PSEN1、NKTR、RYBP、NAMPT、MAPK9、IFNGR1、RHEB、GZMK、RARA、SLC25A37、EPOR、およびRXRAを含んでなる17遺伝子モデルセットを用いてプロファイリングを行う。いくつかの態様では、生体サンプルを採取し、BASP1、CD6、CD7、CXCL10、CXCL9、INPP5D、ISG20、LCK、NKG7、PSMB9、RUNX3、およびTAP1を含んでなる表2の12遺伝子モデルセットを用いてプロファイリングを行う。これらのサンプルは、移植センター特異的参照として役立つ。これらのサンプルのqPCRプロファイルを得た後に、全ARサンプルおよび非ARサンプルの平均発現を別に採り、アッセイした全遺伝子の2列参照を作成する。あるいは、個々のサンプルの代わりにプールしたRNA参照を使用しても十分であり得る。データを、第1列が遺伝子識別を含み、第2列がAR参照を含み、第3列が非AR参照を含む3列参照ファイルとして保存する。この参照に用いた元のサンプルを再分析すれば、例えば、ARサンプルと非ARサンプル間の不十分な分類スコアのためにこれらの参照サンプル間に有意な変動が存在するかどうかを判定することができる。

10

20

【0063】

「未知の」サンプルをARまたは非ARとして分類するために、「未知の」サンプルの発現プロファイル参照ARプロファイルおよび参照非ARプロファイルと直接比較する。サンプル発現プロファイルが参照非AR発現プロファイルよりも参照AR発現プロファイルに密接に適合すれば、そのサンプルはARとして分類される。zスコアを精度の指標として計算することができる(実施例2参照)。発現プロファイルは、mRNAの発現を評価することにより評価することができ、mRNAの発現は、mRNAから逆転写されたcDNAを評価得ることにより評価することができる。

30

【0064】

腎臓同種移植レシピエントにおけるAR/非ARを評定するための遺伝子発現の使用法

本明細書に記載の差次的発現遺伝子は、ARを受けているかもしくはARを受けるであろう個人の診断のためまたは診断を助けるために使用することができる。発現された遺伝子はまた、ARの進行の監視、ARの緩解の監視、ARに関して処置されるべきもしくはARの処置を継続すべき患者の同定、ARの処置の有効性の評定、ARに関して監視されるべき患者の同定、および/またはARのリスクの無い個人の同定のために使用することもできる。本明細書に記載の差次的発現遺伝子は、ARを受けていない個人の診断のためまたは診断を助けるため、ARを受けていない個人の診断のためまたは診断を助けるため、個人がARを受けるまたはARを受けないリスクの予測の診断のためまたは診断を助けるために使用することができる。

40

【0065】

診断アレイは、腎臓同種移植レシピエントから採取された生体サンプル中に存在する差次的発現遺伝子を定量するために使用することができる。このアレイは、基板上に複数の不連続の既知領域を含んでなるDNAコーティング基板を含み得る。これらのアレイは、粒子、ナノ粒子、ビーズ、ナノビーズ、または有孔質または非孔質であり得、かつ、サイズに幅があり得る他の固相表面を含んでなり得る。一実施形態では、アレイはマイクロアレイチップである。別の実施形態では、診断アレイはビーズを含んでなる。さらなる実施形態では、診断アレイはナノ粒子を含んでなる。さらなる実施形態では、診断アレイは微小流体技術を含んでなる。

50

【0066】

本明細書で開示される差次的発現遺伝子を使用する1つの利益は、ARの決定を高い精度で行うことができるということである。精度は感度（AR患者が正確に同定される精度）および特異度（非AR患者が正確に同定される精度）；それぞれ陽性的中率（PPV）および陰性的中率（NPV）によって表すことができる。

【0067】

本明細書に示される実施形態では、差次的発現遺伝子を用いたARの決定は、ARの検出または予測に関して高精度である。本明細書に示される実施形態では、本方法は、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または100%の精度を提供する。さらに、本明細書に示される実施形態では、本方法は、ARの検出または予測に関して少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または100%の精度を提供する。

10

【0068】

本明細書に示される実施形態では、差次的発現遺伝子を用いたARの決定は、ARの検出または予測に関して高感度である。本明細書に示される実施形態では、本方法は、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または100%の感度を提供する。さらに、本明細書に示される実施形態では、本方法は、ARの検出または予測に関して少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または100%の感度を提供する。

20

【0069】

さらに、本明細書に示される実施形態では、差次的発現遺伝子を用いたARの分析は、ARの検出または予測に関して高い特異度である。本明細書に示される実施形態では、本方法は、少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または100%の特異度を提供する。さらに、本明細書に示される実施形態では、本方法は、ARの検出または予測に関して少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または100%の特異度を提供する。

30

【0070】

さらに、本明細書に示される実施形態では、差次的発現遺伝子を用いたARの分析は、ARの検出または予測に関してある陽性的中率（PPV；真の陽性である陽性検査結果の割合/正確な診断）を有する。本明細書に示される実施形態では、本方法は、ARの検出または予測に関して少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または100%のPPVを提供する。また、本明細書に示される実施形態では、差次的発現遺伝子を用いたARの分析は、ARの検出または予測に関してある陰性的中率（NPV；正確に診断される陰性検査結果を有する被験者の割合）を有する。本明細書に示される実施形態では、本方法は、ARの検出または予測に関して少なくとも70%、少なくとも75%、少なくとも80%、少なくとも85%、少なくとも90%、少なくとも95%、少なくとも96%、少なくとも97%、少なくとも98%、少なくとも99%、または100%のNPVを提供する。

40

【0071】

腎臓同種移植レシピエント由来の生体サンプルの分析は、本明細書で開示される1以上

50

、 2 以上、 3 以上、 4 以上、 5 以上、 6 以上、 7 以上、 8 以上、 9 以上、 10 以上、 11 以上、 12 以上、 13 以上、 14 以上、 15 以上、 16 以上、 17 以上、 18 以上、 19 以上、 20 以上、 21 以上、 22 以上、 23 以上、 24 以上、 25 以上、 26 以上、 27 以上、 28 以上、 29 以上、 30 以上、 31 以上、 32 以上、 33 以上、 34 以上、 35 以上、 36 以上、 37 以上、 38 以上、 39 以上、 40 以上、 41 以上、 42 以上、 43 以上、 44 以上、 45 以上、 46 以上、 47 以上、 48 以上、 49 以上、 50 以上、 51 以上、 52 以上、 53 以上、 54 以上、 55 以上、 56 以上、 57 以上、 58 以上、 59 以上、 60 以上、 61 以上、 62 以上、 63 以上、 64 以上、 65 以上、 66 以上、 67 以上、 68 以上、 69 以上、 70 以上、 71 以上、 72 以上、 73 以上、 74 以上、 75 以上、 76 以上、 77 以上、 78 以上、 79 以上、 80 以上、 81 以上、 82 以上、 83 以上、 84 以上、 85 以上、 86 以上、 87 以上、 88 以上、 89 以上、 90 以上、 91 以上、 92 以上、 93 以上、 94 以上、 95 以上、 96 以上、 97 以上、 98 以上、 99 以上、 100 以上、 101 以上、 またはさらに 102 の本明細書差次的発現遺伝子の組合せの評価を含む。いくつかの実施形態では、明示された表 1 の範囲内の遺伝子の数の整数の全ての反復を含む約 1 ~ 約 43 の遺伝子が、本明細書に記載の方法により腎臓同種移植レシピエント由来の生体サンプルから測定される。いくつかの実施形態では、明示された表 2 の範囲内の遺伝子の数の整数の全ての反復を含む約 1 ~ 約 12 の遺伝子が、本明細書に記載の方法により腎臓同種移植レシピエント由来の生体サンプルから測定される。いくつかの実施形態では、明示された表 3 の範囲内の遺伝子の数の整数の全ての反復を含む約 1 ~ 約 102 の遺伝子が、本明細書に記載の方法により腎臓同種移植レシピエント由来の生体サンプルから測定される。

【 0 0 7 2 】

一実施形態では、腎臓同種移植レシピエント由来の差次的発現遺伝子の分析は、CEACAM4、ならびにCFLAR、DUSP1、ITGAX、RNF130、PSEN1、NKTR、RYBP、NAMPT、MAPK9、IFNGR1、RHEB、GZMK、RARA、SLC25A37、EPOR、およびRXRAからなる群から選択される6つの遺伝子のレベルを測定することを含んでなる。別の実施形態では、腎臓同種移植レシピエント由来の差次的発現遺伝子の分析は、CEACAM4、ならびにCFLAR、DUSP1、ITGAX、RNF130、PSEN1、NKTR、RYBP、NAMPT、MAPK9、IFNGR1、RHEB、GZMK、RARA、SLC25A37、EPOR、およびRXRAからなる群から選択される7つの遺伝子のレベルを測定することを含んでなる。さらなる実施形態では、腎臓同種移植レシピエント由来の差次的発現遺伝子の分析は、CEACAM4と、CFLAR、DUSP1、ITGAX、RNF130、PSEN1、NKTR、RYBP、NAMPT、MAPK9、IFNGR1、RHEB、GZMK、RARA、SLC25A37、EPOR、およびRXRAからなる群から選択される8つの遺伝子のレベルを測定することを含んでなる。別の実施形態では、腎臓同種移植レシピエント由来の差次的発現遺伝子の分析は、CEACAM4、ならびにCFLAR、DUSP1、ITGAX、RNF130、PSEN1、NKTR、RYBP、NAMPT、MAPK9、IFNGR1、RHEB、GZMK、RARA、SLC25A37、EPOR、およびRXRAからなる群から選択される9つの遺伝子のレベルを測定することを含んでなる。さらに別の実施形態では、腎臓同種移植レシピエント由来の差次的発現遺伝子の分析は、CEACAM4、ならびにCFLAR、DUSP1、ITGAX、RNF130、PSEN1、NKTR、RYBP、NAMPT、MAPK9、IFNGR1、RHEB、GZMK、RARA、SLC25A37、EPOR、およびRXRAからなる群から選択される10の遺伝子のレベルを測定することを含んでなる。さらなる実施形態では、腎臓同種移植レシピエント由来の差次的発現遺伝子の分析は、CEACAM4、ならびにCFLAR、DUSP1、ITGAX、RNF130、PSEN1、NKTR、RYBP、NAMPT、MAPK9、IFNGR1、RHEB、GZMK、RARA、SLC25A37、EPOR、およびRXRAからなる群から選択される11の遺伝子のレベルを測定することを含んでなる。別の実施形態では、腎臓同種移植レシピエント由来の差次的発現

遺伝子の分析は、CEACAM4、ならびにCFLAR、DUSP1、ITGAX、RNF130、PSEN1、NKTR、RYBP、NAMPT、MAPK9、IFNGR1、RHEB、GZMK、RARA、SLC25A37、EPOR、およびRXRAからなる群から選択される12の遺伝子のレベルを測定することを含んでなる。さらなる実施形態では、腎臓同種移植レシピエント由来の差次的発現遺伝子の分析は、CEACAM4、ならびにCFLAR、DUSP1、ITGAX、RNF130、PSEN1、NKTR、RYBP、NAMPT、MAPK9、IFNGR1、RHEB、GZMK、RARA、SLC25A37、EPOR、およびRXRAからなる群から選択される13の遺伝子のレベルを測定することを含んでなる。別の実施形態では、腎臓同種移植レシピエント由来の差次的発現遺伝子の分析は、CEACAM4、ならびにCFLAR、DUSP1、ITGAX、RNF130、PSEN1、NKTR、RYBP、NAMPT、MAPK9、IFNGR1、RHEB、GZMK、RARA、SLC25A37、EPOR、およびRXRAからなる群から選択される14の遺伝子のレベルを測定することを含んでなる。さらなる実施形態では、腎臓同種移植レシピエント由来の差次的発現遺伝子の分析は、CEACAM4、ならびにCFLAR、DUSP1、ITGAX、RNF130、PSEN1、NKTR、RYBP、NAMPT、MAPK9、IFNGR1、RHEB、GZMK、RARA、SLC25A37、EPOR、およびRXRAからなる群から選択される15の遺伝子のレベルを測定することを含んでなる。別の実施形態では、腎臓同種移植レシピエント由来の差次的発現遺伝子の分析は、BASP1、CD6、CD7、CXCL10、CXCL9、INPP5D、ISG20、LCK、NKG7、PSMB9、RUNX3、およびTAP1のレベルを測定することを含んでなる。

【0073】

さらなる実施形態では、腎臓同種移植レシピエント由来の差次的発現遺伝子の分析は、表2から選択される12の遺伝子の組合せのレベルを測定することを含んでなる。

【0074】

さらなる実施形態では、腎臓同種移植レシピエント由来の差次的発現遺伝子の分析は、遺伝子CEACAM4、CFLAR、DUSP1、ITGAX、RNF130、PSEN1、NKTR、RYBP、NAMPT、MAPK9、IFNGR1、RHEB、GZMK、RARA、SLC25A37、EPOR、およびRXRAのレベルを測定することを含んでなる。補正されたこの17遺伝子セットは、サンプルの88%をARとして、サンプルの95%を非ARとして予測する。いくつかの実施形態では、合計17遺伝子の発現レベルが測定される。

【0075】

さらなる実施形態では、腎臓同種移植レシピエント由来の差次的発現遺伝子の分析は、遺伝子BASP1、CD6、CD7、CXCL10、CXCL9、INPP5D、ISG20、LCK、NKG7、PSMB9、RUNX3、およびTAP1のレベルを測定することを含んでなる。

【0076】

別の実施形態では、腎臓同種移植レシピエント由来の差次的発現遺伝子の分析は、遺伝子CEACAM4、CFLAR、DUSP1、ITGAX、NAMPT、NKTR、PSEN1、EPOR、GZMK、RARA、RHEB、およびSLC25A37のレベルを測定することを含んでなる。この遺伝子セットは、86%感度および90%特異度でARを分類する。

【0077】

別の実施形態では、本明細書に記載の差次的発現遺伝子の分析は、腎臓同種移植に対する慢性損傷を予測するために有用である。慢性損傷は一般に、最も多くはドナー臓器に対して生じた持続的免疫応答による移植臓器における長期機能欠失として表される。一態様において、差次的発現遺伝子は、被験者由来の組織生検サンプルにおいて評定される。別の態様において、組織生検における差次的発現遺伝子の測定は、免疫組織化学的技術、本明細書に記載の核酸法、またはタンパク質検出法（例えば、ウエスタンブロット法）また

は当技術分野で公知の他の一般的な遺伝子発現法によって行うこともできる。別の態様において、CEACAM4、ならびにCFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、RYBP、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37から選択される6~16の他の遺伝子のレベルが、ARの評価のために、腎臓同種移植を受けた個人に由来する組織生検において測定される。別の態様において、CEACAM4、ならびにBASP1、CD6、CD7、CXCL10、CXCL9、INPP5D、ISG20、LCK、NKG7、PSMB9、RUNX3、およびTAP1から選択される6~16の他の遺伝子のレベルが、ARの評価のために、腎臓同種移植を受けた個人に由来する組織生検において測定される。別の態様において、表1からの、明示された範囲内の遺伝子の数の整数の全ての反復を含む約1~約43の遺伝子のレベルが、ARの評価のために、腎臓同種移植を受けた個人に由来する組織生検において測定される。別の実施形態では、表3からの、明示された範囲内の遺伝子の数の整数の全ての反復を含む約1~約102の遺伝子が、ARの評価のために、腎臓同種移植を受けた個人に由来する組織生検において測定される。

10

20

30

40

50

【0078】

いくつかの実施形態では、統合遺伝子モデルが用いられる。すなわち、上記のような複数の遺伝子セットが使用され、各遺伝子セットはカテゴリー的値もしくはスコアおよび/または数的値またはスコアを提供する。このように、統合モデルは、単一の遺伝子セットに偏りを持たない。ARのリスクが高い患者のうち91%が正確にARと分類された。ARのリスクが低い患者のうち92%が正確に非ARと分類された。

【0079】

本発明の差次的発現遺伝子はまた、ARの処置のための個人を同定するために使用することができる。いくつかの実施形態では、この個人はAR症状の進行または緩解に関して監視される。いくつかの実施形態では、この個人はAR症状の発症前または発症時にARに関して処置される。いくつかの実施形態では、この処置は、コルチコステロイド療法である。他の実施形態では、この処置は、ムロモナブ-CD3(Orthoclone OKT3)などの抗T細胞抗体の投与である。さらなる実施形態では、この処置は、血漿交換と抗CD20抗体の投与の組合せである。場合によっては、処置が継続されるべきかどうかを決定するためまたは処置が有効であるかどうかをみるために監視が行われる。

【0080】

本明細書に記載の方法のいくつかの実施形態では、これらの方法はAR応答の予測に用途を持つ。これらの方法では、被験者はまず、対象の方法に従ってARに関して監視され、次に、監視結果に少なくとも部分的に基づいて決定されたプロトコルを用いて処置される。一実施形態では、被験者は、本明細書に記載の方法の1つに従って急性拒絶の有無に関して監視される。被験者は次に、監視工程の結果を用いてその適合性が決定されるプロトコルを用いて処置することができる。例えば、被験者が次の1~6か月以内に急性拒絶応答を示すと予測される場合、急性拒絶の治療/予防に関して当技術分野で公知のように、免疫抑制療法を調整、例えば、増強または薬物の変更を行うことができる。同様に、被験者がその時および短期的に急性拒絶を受けていないと予測される場合、薬物毒性の可能性を軽減するために免疫抑制療法を低減することができる。本明細書に記載の方法のいくつかの実施形態では、被験者は、移植片の受容または移植後に急性拒絶に関して監視される。被験者は、移植重要後1回または連続的に、例えば、毎週、毎月、隔月、半年ごと、毎年などにスクリーニングすることができる。いくつかの実施形態では、被験者は、急性拒絶エピソードの発生前に監視される。他の実施形態では、被験者は、急性拒絶エピソードの発生後に監視される。

【0081】

本明細書に記載の方法のいくつかの実施形態では、本方法は、ARの治療を必要とする被験者の治療パラダイムまたは治療計画の変更または変化に用途を持つ。非限定的免疫抑制療法またはそれを必要とする被験者の治療に有用な治療薬の例としては、ステロイド(例

例えば、プレドニゾン（デルタゾン）、プレドニゾロン、メチル・プレドニゾロン（メドロール、ソルメドロール）、抗体（例えば、ムロモナブ - CD3（Orthoclone - OKT3）、抗胸腺細胞免疫グロブリン（ATGAM、サイモグロブリン）、ダクリズマブ（ゼナバックス）、パシリキシマブ（シムレクト）、リツキシマブ、サイトメガロウイルス免疫グロブリン（サイトガム）、免疫グロブリン（ポリガム）、カルシニューリン阻害剤（例えば、シクロスポリン（サンディミュン）、タクロリムス（プロGRAF））、細胞増殖阻害薬（例えば、ミコフェノール酸モフェチル（セルセプト）、アザチオプリン（イムラン）、TOR阻害剤（例えば、ラパマイシン（ラパミュン、シロリムス）、エベロリムス（セルチカン））、またはそれらの併用療法を含んでなる。

【0082】

被験者が本発明に記載の方法を用いてARを持たないと同定されるいくつかの実施形態では、被験者は、抗増殖薬（例えば、ミコフェノール酸モフェチルおよび/またはアザチオプリン）、カルシニューリン阻害剤（例えば、シクロスポリンおよび/またはタクロリムス）、ステロイド（例えば、プレドニゾン、プレドニゾロン、および/またはメチルプレドニゾロン）またはそれらの組合せの投与を含んでなる免疫抑制標準治療維持療法に留まることができる。例えば、本発明に記載の方法を用いてARを持たないと同定された被験者は、低用量のプレドニゾン（例えば、約 $0.1\text{ mg} \cdot \text{kg}^{-1} \cdot \text{d}^{-1}$ ~ 約 $1\text{ mg} \cdot \text{kg}^{-1} \cdot \text{d}^{-1}$ ）、低用量のシクロスポリン（例えば、約 $4\text{ mg} \cdot \text{kg}^{-1} \cdot \text{d}^{-1}$ ~ 約 $8\text{ mg} \cdot \text{kg}^{-1} \cdot \text{d}^{-1}$ ）、および低用量のミコフェノラート（例えば、約 $1 \sim 1.5\text{ g}$ 1日2回）の投与を含んでなる維持療法におくことができる。別の例では、本発明に記載の方法を用いてARを持たないと同定された被験者は、ステロイド療法から離脱させ、低用量のシクロスポリン（例えば、約 $4\text{ mg} \cdot \text{kg}^{-1} \cdot \text{d}^{-1}$ ~ 約 $8\text{ mg} \cdot \text{kg}^{-1} \cdot \text{d}^{-1}$ ）、および低用量のミコフェノラート（例えば、約 $1 \sim 1.5\text{ g}$ 1日2回）の投与を含んでなる維持療法におくことができる。別の例では、本発明に記載の方法を用いてARを持たないと同定された被験者は、本明細書に記載の全ての免疫抑制療法から離脱させることができる。

【0083】

被験者が本発明に記載の方法を用いてARを有すると同定されるいくつかの実施形態では、被験者は、救済療法または高用量のステロイド（例えば、プレドニゾン、プレドニゾロン、および/またはメチルプレドニゾロン）、高用量のポリクローナルまたはモノクローナル抗体（例えば、ムロモナブ - CD3（OKT3）、抗胸腺細胞免疫グロブリン、ダクリズマブ、リツキシマブ、パシリキシマブ、サイトメガロウイルス免疫グロブリン、および/または免疫グロブリン）、高用量の抗増殖薬（例えば、ミコフェノール酸モフェチルおよび/またはアザチオプリン）、またはそれらの組合せの投与を含んでなる免疫抑制薬の増強におくことができる。

【0084】

いくつかの実施形態では、被験者が本発明に記載の方法を用いてARを持たないと同定されるか、またはARを持つと同定される場合の療法の経過は、移植後の時間および拒絶の重篤度、治療医、および移植センターに依存する。

【0085】

よって、本発明の差次的発現遺伝子および本明細書に記載の方法論を用い、当業者は、腎臓同種移植レシピエントにおいてARを診断すること、腎臓同種移植レシピエントにおいて非ARを診断すること、ARの診断を助けること、ARのリスクに関する診断を助けること、ARの進行を監視すること、ARの緩解を監視すること、ARに関して処置すべきまたはARに関する処置を継続すべき個人を同定すること、ARの処置の有効性を評定すること、および/またはAR症状に関して監視すべき個人を同定することが可能である。

【0086】

いくつかの実施形態では、本発明の差次的発現遺伝子および本明細書に記載の方法論は、抗体媒介性ARの層別化または同定に使用することができる。他の実施形態では、本発

10

20

30

40

50

明の差次的発現遺伝子および本明細書に記載の方法論は、T細胞媒介性ARの層別化または同定に使用することができる。本明細書で提供される遺伝子は、いくつかの態様では、それらがB細胞で発現されるか、またはT細胞で発現されるかもしくは活性化T細胞の既知のマーカであるので、B細胞またはT細胞媒介性ARの同定に有用である。

【0087】

ARの診断、検出、または予測のためのキット

本発明はさらに、ARの診断、検出、および予測のアッセイキットを提供する。本キットは、腎臓同種移植を受けた個人由来の生体サンプルにおいてARに関連する差次的発現遺伝子のレベルを測定するための遺伝子発現評価素子を含んでなる。いくつかの実施形態では、本キットは、生体サンプルにおいて対象とする差次的発現遺伝子のレベルを測定するための試薬を含んでなる。いくつかの実施形態では、本キットは、生体サンプルにおける差次的発現遺伝子の測定のための1以上の固相表面を含んでなる組成物を含んでなる。一実施形態では、固相表面は、マイクロアレイチップを含んでなる。別の実施形態では、固相表面は、ビーズを含んでなる。さらなる実施形態では、固相表面は、ナノ粒子を含んでなる。一実施形態では、本キットは、CEACAM4、ならびにCFLAR、DUSP1、ITGAX、RNF130、PSEN1、NKTR、RYBP、NAMPT、MAPK9、IFNGR1、RHEB、GZMK、RARA、SLC25A37、EPOR、およびRXRAから選択される少なくとも6、7、8、9、10、または11の他の遺伝子の測定のための1以上の固相表面を含んでなる組成物を含んでなる。いくつかの実施形態では、合計17の遺伝子の発現レベルが測定される。

10

20

【0088】

本キットはさらに、腎臓同種移植を受けた個人のARの診断において使用するための参照標準素子を含んでなる。いくつかの実施形態では、参照標準素子は、単一の移植センターまたは複数の移植センターの腎臓同種移植レシピエントから得られた各差次的発現遺伝子に対するARサンプル由来の単一の参照発現ベクターを含んでなる。いくつかの実施形態では、参照標準素子は、単一の移植センターまたは複数の移植センターの腎臓同種移植レシピエントから得られた各差次的発現遺伝子に対する非ARサンプル由来の単一の参照発現ベクターを含んでなる。参照標準素子は、ARを有するレシピエントを診断するために腎臓同種移植レシピエントからの遺伝子発現と比較するために使用される。

30

【0089】

いくつかの実施形態では、比較はコンピューターにより行われる。他の実施形態では、比較は、個人により行われる。一実施形態では、比較は、医師により行われる。各移植センターの参照標準は上記のようにして作成することができる。

【0090】

いくつかの実施形態では、コンピューターは、AR応答の誘導の予測、AR応答の診断、および被験者におけるAR応答の特性決定のうちの少なくとも1つをユーザーに出力するように構成されており、この出力は、7、8、9、10、11、12、13、14、15、16、または17の遺伝子の遺伝子発現結果を対照参照発現プロファイルと比較することにより決定される。

40

【0091】

本キットはまた、アッセイの使用に関する説明書を含んでなる。

【0092】

ARの診断、検出、または予測のためのシステム

本発明はさらに、ARの診断、検出、および予測のためのシステムを提供する。本システムは、腎臓同種移植を受けた個人由来の生体サンプルにおいてARに関連する差次的発現遺伝子のレベルを測定するための遺伝子発現評価素子を含んでなる。一実施形態では、本システムは、マイクロアレイチップを含んでなる。別の実施形態では、システムは、ビーズを含んでなる。さらなる実施形態では、システムは、ナノ粒子を含んでなる。種々の実施形態では、システムは、CEACAM4、ならびにCFLAR、DUSP1、ITGAX、RNF130、PSEN1、NKTR、RYBP、NAMPT、MAPK9、IF

50

NGR1、RHEB、GZMK、RARA、SLC25A37、EPOR、およびRXRAから選択される少なくとも6、7、8、9、10、11、12、13、14、15、または16の他の遺伝子の測定のための遺伝子発現評価素子を含んでなる。いくつかの実施形態では、合計17の遺伝子の発現レベルが測定される。

【0093】

特定の実施形態では、遺伝子発現評価素子は、遺伝子発現結果を生成するために、CEACAM4、ならびにCFLAR、DUSP1、ITGAX、RNF130、PSEN1、NKTR、RYBP、NAMPT、MAPK9、IFNGR1、RHEB、GZMK、RARA、SLC25A37、EPOR、およびRXRAから選択される少なくとも6、7、8、9、10、11、12、13、14、15、または16の他の遺伝子を選択的に増幅するように設計された標識遺伝子プライマーまたは標識プローブを含んでなる。いくつかの実施形態では、標識は、非天然である。他の実施形態では、遺伝子プライマーまたはプローブは、標識を含むように共有結合的に修飾される。関連の実施形態では、標識は、蛍光団または放射性標識からなる群から選択することができる。

10

【0094】

本システムは、腎臓同種移植を受けた個人においてARを評定するための参照標準素子をさらに含んでなる。いくつかの実施形態では、参照標準素子は、単一の移植センターの腎臓同種移植レシピエントから得られた各差次的発現遺伝子に対するARサンプル由来の単一の参照発現ベクターを含んでなる。いくつかの実施形態では、参照標準素子は、単一の移植センターの腎臓同種移植レシピエントから得られた各差次的発現遺伝子に対する非ARサンプル由来の単一の参照発現ベクターを含んでなる。参照標準素子は、ARを有するレシピエントを診断するために、腎臓同種移植レシピエントからの遺伝子発現と比較するために使用される。いくつかの実施形態では、比較は、コンピューターにより行われる。他の実施形態では、比較は、個人により行われる。一実施形態では、比較は、医師により行われる。各移植センターの参照標準は上記のようにして作成することができる。

20

【0095】

ARの診断、検出、または予測のための組成物

本発明は、腎臓同種移植を受けた個人由来の生体サンプルにおいてARに関連する差次的発現遺伝子のレベルを測定するための1以上の固相表面を含んでなる組成物を提供する。いくつかの実施形態では、組成物は製品である。一実施形態では、製品は、腎臓同種移植を受けた個人由来の生体サンプルにおいて差次的発現遺伝子のレベルを測定するための参照標準を含んでなる。いくつかの実施形態では、固相表面は、差次的発現遺伝子のcDNAの付着を提供する。他の実施形態では、固相表面は、差次的発現遺伝子の増幅のためのプライマーまたは標識プライマーの付着を提供する。特定の実施形態では、固相表面は、本明細書に開示される少なくとも1、1以上、2以上、3以上、4以上、5以上、6以上、7以上、8以上、9以上、10以上、11以上、12以上、13以上、14以上、15以上、16以上、17以上、18以上、19以上、20以上、21以上、22以上、23以上、24以上、25以上、26以上、27以上、28以上、29以上、30以上、31以上、32以上、33以上、34以上、35以上、36以上、37以上、38以上、39以上、40以上、41以上、42以上、43以上、44以上、45以上、46以上、47以上、48以上、49以上、50以上、51以上、52以上、53以上、54以上、55以上、56以上、57以上、58以上、59以上、60以上、61以上、62以上、63以上、64以上、65以上、66以上、67以上、68以上、69以上、70以上、71以上、72以上、73以上、74以上、75以上、76以上、77以上、78以上、79以上、80以上、81以上、82以上、83以上、84以上、85以上、86以上、87以上、88以上、89以上、90以上、91以上、92以上、93以上、94以上、95以上、96以上、97以上、98以上、99以上、100以上、101以上、またはさらには102の遺伝子の測定を可能とする。一実施形態では、表1からの、明示された範囲内の遺伝子の数の整数の全ての反復を含む約1～約43の遺伝子が、ARの評価のために、腎臓同種移植を受けた個人由来の生体サンプルにおいて測定される。別の実施形態で

30

40

50

は、表 3 からの、明示された範囲内の遺伝子の数の整数の全ての反復を含む約 1 ~ 約 1 0 2 の遺伝子が、A R の評価のために、腎臓同種移植を受けた個人由来の生体サンプルにおいて測定される。別の実施形態では、最低 7 つの遺伝子が、A R の評価のために測定される。別の実施形態では、最大 1 7 の遺伝子が、A R の評価のために測定される。

【 0 0 9 6 】

1 つの特定の実施形態では、本発明は、C E A C A M 4、ならびに C F L A R、D U S P 1、I T G A X、R N F 1 3 0、P S E N 1、N K T R、R Y B P、N A M P T、M A P K 9、I F N G R 1、R H E B、G Z M K、R A R A、S L C 2 5 A 3 7、E P O R、および R X R A からなる群から選択される 6 つの遺伝子の遺伝子発現レベルの測定のための 1 以上の固相表面を含む組成物を提供する。別の実施形態では、組成物は、C E A C A M 4、ならびに C F L A R、D U S P 1、I T G A X、R N F 1 3 0、P S E N 1、N K T R、R Y B P、N A M P T、M A P K 9、I F N G R 1、R H E B、G Z M K、R A R A、S L C 2 5 A 3 7、E P O R、および R X R A からなる群から選択される 7 つの遺伝子の遺伝子発現レベルを測定するための 1 以上の固相表面を含む。さらなる実施形態では、組成物は、C E A C A M 4、ならびに C F L A R、D U S P 1、I T G A X、R N F 1 3 0、P S E N 1、N K T R、R Y B P、N A M P T、M A P K 9、I F N G R 1、R H E B、G Z M K、R A R A、S L C 2 5 A 3 7、E P O R、および R X R A からなる群から選択される 8 つの遺伝子の遺伝子発現レベルを測定するための 1 以上の固相表面を含む。別の実施形態では、組成物は、C E A C A M 4、ならびに C F L A R、D U S P 1、I T G A X、R N F 1 3 0、P S E N 1、N K T R、R Y B P、N A M P T、M A P K 9、I F N G R 1、R H E B、G Z M K、R A R A、S L C 2 5 A 3 7、E P O R、および R X R A からなる群から選択される 9 つの遺伝子の遺伝子発現レベルを評価するための 1 以上の固相表面を含む。さらに別の実施形態では、組成物は、C E A C A M 4、ならびに C F L A R、D U S P 1、I T G A X、R N F 1 3 0、P S E N 1、N K T R、R Y B P、N A M P T、M A P K 9、I F N G R 1、R H E B、G Z M K、R A R A、S L C 2 5 A 3 7、E P O R、および R X R A からなる群から選択される 1 0 の遺伝子の遺伝子発現レベルを測定するための 1 以上の固相表面を含む。さらなる実施形態では、組成物は、C E A C A M 4、ならびに C F L A R、D U S P 1、I T G A X、R N F 1 3 0、P S E N 1、N K T R、R Y B P、N A M P T、M A P K 9、I F N G R 1、R H E B、G Z M K、R A R A、S L C 2 5 A 3 7、E P O R、および R X R A からなる群から選択される 1 1 の遺伝子の遺伝子発現レベルを測定するための 1 以上の固相表面を含む。別の実施形態では、組成物は、C E A C A M 4、ならびに C F L A R、D U S P 1、I T G A X、R N F 1 3 0、P S E N 1、N K T R、R Y B P、N A M P T、M A P K 9、I F N G R 1、R H E B、G Z M K、R A R A、S L C 2 5 A 3 7、E P O R、および R X R A からなる群から選択される 1 2 の遺伝子の遺伝子発現レベルを測定するための 1 以上の固相表面を含む。さらなる実施形態では、組成物は、C E A C A M 4、ならびに C F L A R、D U S P 1、I T G A X、R N F 1 3 0、P S E N 1、N K T R、R Y B P、N A M P T、M A P K 9、I F N G R 1、R H E B、G Z M K、R A R A、S L C 2 5 A 3 7、E P O R、および R X R A からなる群から選択される 1 3 の遺伝子の遺伝子発現レベルを測定するための 1 以上の固相表面を含む。別の実施形態では、組成物は、C E A C A M 4、ならびに C F L A R、D U S P 1、I T G A X、R N F 1 3 0、P S E N 1、N K T R、R Y B P、N A M P T、M A P K 9、I F N G R 1、R H E B、G Z M K、R A R A、S L C 2 5 A 3 7、E P O R、および R X R A からなる群から選択される 1 4 の遺伝子の遺伝子発現レベルを測定するための 1 以上の固相表面を含む。さらなる実施形態では、組成物は、C E A C A M 4、ならびに C F L A R、D U S P 1、I T G A X、R N F 1 3 0、P S E N 1、N K T R、R Y B P、N A M P T、M A P K 9、I F N G R 1、R H E B、G Z M K、R A R A、S L C 2 5 A 3 7、E P O R、および R X R A からなる群から選択される 1 5 の遺伝子の遺伝子発現レベルを測定するための 1 以上の固相表面を含む。さらなる実施形態では、組成物は、C E A C A M 4、ならびに C F L A R、D U S P 1、I T G A X、R N F 1 3 0、P S E N 1、N K T R、R Y B P、N A M P T、M A P K 9、I F N G R 1

10

20

30

40

50

、RHEB、GZMK、RARA、SLC25A37、EPOR、およびRXRAの遺伝子発現レベルを測定するための1以上の固相表面を含む。いくつかの実施形態では、合計17の遺伝子の発現レベルが測定される。別の実施形態では、組成物は、CEACAM4、CFLAR、DUSP1、ITGAX、NAMPT、NKTR、PSEN1、EPOR、GZMK、RARA、RHEB、およびSLC25A37の遺伝子発現レベルを測定するための1以上の固相表面を含む。

【0097】

【表3】

遺伝子記号	Ensembl ID	Entrez ID	定義
AIF1	ENSG00000206428	199	同種移植炎症因子1
F13A1	ENSG00000124491	2162	凝固因子 XIII, A1 ポリペプチド
MMP9	ENSG00000100985	4318	マトリックスメタロペプチダーゼ9(ゼラチナーゼ B, 92kDa ゼラチナーゼ, 92kDa IV型コラゲナーゼ)
NELL2	ENSG00000184613	4753	NEL様2(ニワトリ)
CD53	ENSG00000143119	963	CD53分子
CXCL10	ENSG00000169245	3627	ケモカイン(C-X-Cモチーフ)リガンド10
ISG20	ENSG00000172183	3669	インターフェロン刺激エキソヌクレアーゼ遺伝子 20kDa
CD48	ENSG00000117091	962	CD48分子
BATF	ENSG00000156127	10538	塩基性ロイシンジッパー転写因子, ATF様
DDX11	ENSG00000013573	1663	DEAD/H (Asp-Glu-Ala-Asp/His) ボックスポリペプチド11 (CHL1様ヘリカーゼホモログ, S. セレビシエ(S. cerevisiae))
CD44	ENSG00000026508	960	CD44分子(インディアン血液型)
DDX23	ENSG00000174243	9416	DEAD (Asp-Glu-Ala-Asp) ボックスポリペプチド23
IL15RA	ENSG00000134470	3601	インターロイキン15受容体, α
ADAM8	ENSG00000151651	101	ADAMメタロペプチダーゼドメイン8
RAB27A	ENSG00000069974	5873	RAB27A, メンバー-RAS 癌遺伝子ファミリー
CD3D	ENSG00000167286	915	CD3d分子, δ (CD3-TCR複合体)
HLA-A	ENSG00000206503	3105	主要組織適合性遺伝子複合体, クラス I, A
TNFRSF14	ENSG00000157873	8764	腫瘍壊死因子受容体スーパーファミリー, メンバー-14(ヘルペスウイルスエントリーメディエーター)
BASP1	ENSG00000176788	10409	脳に豊富な, 膜結合シグナルタンパク質1

10

20

30

40

HLA-E	ENSG00000204592	3133	主要組織適合性遺伝子複合体, クラス I, E	
HLA-G	ENSG00000204632	3135	主要組織適合性遺伝子複合体, クラス I, G	
HLA-F	ENSG00000206509	3134	主要組織適合性遺伝子複合体, クラス I, F	
ARPC1B	ENSG00000130429	10095	アクチン関連タンパク質 2/3 複合体, サブユニット 1B, 41kDa; アクチン関連タンパク質 2/3 複合体サブユニット 1B (ARP2/3 複合体 41kDa サブユニット) (p41-ARC) に類似	10
KRT17	ENSG00000186831	729682	ケラチン 17; ケラチン 17 偽遺伝子 3	
ADAMTS3	ENSG00000156140	9508	トロンボスポンジン 1 型モチーフ, 3 を有する ADAM メタロペプチダーゼ	
BTN3A2	ENSG00000186470	11118	ブチロフィリン, サブファミリー 3, メンバー A2	20
TNFAIP2	ENSG00000185215	7127	腫瘍壊死因子, α 誘導タンパク質 2	
GBP2	ENSG00000162645	2634	グアニル酸結合タンパク質 2, インターフェロン誘導性	
IFITM3	ENSG00000142089	10410	インターフェロン誘導膜貫通タンパク質 3 (1-8U)	
STK10	ENSG00000072786	6793	セリン/トレオニンキナーゼ 10	
MAP4K1	ENSG00000104814	11184	マイトジェン活性化タンパク質キナーゼキナーゼキナーゼキナーゼ 1	30
ITGB2	ENSG00000160255	3689	インテグリン, β 2 (補体成分 3 受容体 3 および 4 サブユニット)	
PTPRCAP	ENSG00000213402	5790	タンパク質チロシンホスファターゼ, 受容体型, C 結合タンパク質	
MDK	ENSG00000110492	4192	ミドカイン (神経突起成長促進因子 2)	
SERPINH1	ENSG00000149257	871	セルピンペプチダーゼ阻害剤, クレイド H (熱ショックタンパク質 47), メンバー 1, (コラーゲン結合タンパク質 1)	40

ITGB7	ENSG00000139626	3695	インテグリン, $\beta 7$
ZAP70	ENSG00000115085	7535	ζ 鎖(TCR)結合タンパク質キナーゼ 70kDa
FCER1G	ENSG00000158869	2207	IgE の Fc フラグメント, 高親和性 I, 受容体; γ ポリペプチド
RUNX3	ENSG0000020633	864	runt 関連転写因子 3
ARHGDIB	ENSG00000111348	397	Rho GDP 解離阻害剤(GDI) β
CDC20	ENSG00000117399	991	細胞分裂周期 20 ホモログ(S. セレビシエ)
AIM2	ENSG00000163568	9447	黒色腫に不在 2
LGALS9	ENSG00000168961	3965	レクチン, ガラクトシド結合, 可溶性, 9
PSMB9	ENSG00000206232	5698	プロテアソーム(プロソーム, マクロペイン)サブユニット, β 型, 9 (高分子量多機能性ペプチダーゼ 2)
ATF5	ENSG00000169136	22809	活性化転写因子 5
IKBKE	ENSG00000143466	9641	B 細胞における κ 軽鎖ポリペプチド遺伝子エンハンサーの阻害剤, キナーゼ ϵ
CORO1A	ENSG00000102879	11151	コロニン, アクチン結合タンパク質, 1A
BBC3	ENSG00000105327	27113	BCL2 結合成分 3
UCP2	ENSG00000175567	7351	脱共役タンパク質 2 (ミトコンドリア, プロトン担体)
MARCKS	ENSG00000155130	4082	ミリストイル化アラニンリッチタンパク質キナーゼ C 基質
NKG7	ENSG00000105374	4818	ナチュラルキラー細胞グループ 7 配列
NNMT	ENSG00000166741	4837	ニコチンアミド N-メチルトランスフェラーゼ
CD8A	ENSG00000153563	925	CD8a 分子
HLA-DMA	ENSG00000204257	3108	主要組織適合性遺伝子複合体, クラス II, DM α
ARHGAP4	ENSG00000089820	393	Rho GTP アーゼ活性化タンパク質 4
CASP4	ENSG00000196954	837	カスパーゼ 4, アポトーシス関連システインペプチダーゼ
HCP5	ENSG00000206337	10866	HLA 複合体 P5
MAN2B1	ENSG00000104774	4125	マンノシダーゼ, α , クラス 2B,

10

20

30

40

			メンバー1
PLCB2	ENSG00000137841	5330	ホスホリパーゼ C, $\beta 2$
GZMA	ENSG00000145649	3001	グランザイム A (グランザイム 1, 細胞傷害性 Tリンパ球結合セリン エステラーゼ 3)
LEF1	ENSG00000138795	51176	リンパ系エンハンサー結合因子 1
MCM5	ENSG00000100297	4174	ミニ染色体維持複合体成分 5
SH2D2A	ENSG00000027869	9047	SH2 ドメインタンパク質 2A
PRKD2	ENSG00000105287	25865	タンパク質キナーゼ D2
TNFRSF9	ENSG00000049249	3604	腫瘍壊死因子受容体スーパーファミリー, メンバー9
GZMK	ENSG00000113088	3003	グランザイム K (グランザイム 3; トリプターゼ II)
PSME1	ENSG00000092010	5720	プロテアソーム (プロソーム, マクロペイン)アクチベーターサブユニット 1 (PA28 α)
LCK	ENSG00000182866	3932	リンパ球特異的タンパク質チロシンキナーゼ
DDB2	ENSG00000134574	1643	損傷特異的 DNA 結合タンパク質 2, 48kDa
ADAM19	ENSG00000135074	8728	ADAM メタロペプチダーゼドメイン 19 (メルトリン β)
MAP3K11	ENSG00000173327	4296	マイトジェン活性化タンパク質キナーゼキナーゼキナーゼ 11
HLA-DQB1	ENSG00000179344	3119	主要組織適合性遺伝子複合体, クラス II, DQ $\beta 1$;主要組織適合性遺伝子, クラス II, DQ $\beta 1$ に類似
TNF	ENSG00000206439	7124	腫瘍壊死因子 (TNF スーパーファミリー, メンバー2)
FOXM1	ENSG00000111206	2305	フォークヘッドボックス M1
PML	ENSG00000140464	652346	前骨髄球性白血病;前骨髄球性白血病タンパク質アイソフォーム 1 に類似
CXCL9	ENSG00000138755	4283	ケモカイン (C-X-C モチーフ)リガンド 9
POLR2A	ENSG00000181222	5430	ポリメラーゼ (RNA) II (DNA 依存性)ポリペプチド A, 220kDa
TNFRSF1A	ENSG00000067182	7132	腫瘍壊死因子受容体スーパーファミリー, メンバー1A

10

20

30

40

RGS10	ENSG00000148908	6001	Gタンパク質シグナル伝達10のレギュレーター
TNFRSF1B	ENSG00000028137	7133	腫瘍壊死因子受容体スーパーファミリー, メンバー1B
NUP210	ENSG00000132182	23225	ヌクレオポリン210kDa
IL10RA	ENSG00000110324	3587	インターロイキン10受容体, α
TAP1	ENSG00000206233	6890	輸送体1, ATP結合カセット, サブファミリーB (MDR/TAP)
CD2	ENSG00000116824	914	CD2分子
INPP5D	ENSG00000168918	3635	イノシトールポリリン酸-5-ホスファターゼ, 145kDa
CD6	ENSG00000013725	923	CD6分子
CD7	ENSG00000173762	924	CD7分子
PTPRC	ENSG00000081237	5788	タンパク質チロシンホスファターゼ, 受容体タイプ, C
IL2RB	ENSG00000100385	3560	インターロイキン2受容体, β
PLEK	ENSG00000115956	5341	プレクストリン
BIRC5	ENSG00000089685	332	バキュロウイルスIAPリピート含有5
FZD2	ENSG00000180340	2535	frizzledホモログ2(ショウジョウバエ)
STAT1	ENSG00000115415	6772	シグナル伝達兼転写活性化因子1, 91kDa
CCL13	ENSG00000181374	6357	ケモカイン(C-Cモチーフ)リガンド13
IRF5	ENSG00000128604	3663	インターフェロン調節因子5
STAB1	ENSG00000010327	23166	サビリン
IRF1	ENSG00000125347	3659	インターフェロン調節因子1
IRF3	ENSG00000126456	3661	インターフェロン調節因子3
IRF4	ENSG00000137265	3662	インターフェロン調節因子4
CD14	ENSG00000170458	929	CD14分子
CLORF38			染色体1オープンリーディングフレーム38
VAMP5	ENSG00000168899	10791	小胞結合膜タンパク質5

10

20

30

50

【0098】

ARと非ARの分類のための相関に基づくアルゴリズム用のソフトウェア

本明細書に記載の相関に基づく分析は、AltAnalyzeバージョン2.0.8以降を用いて行うことができる。LineageProfilerは、オープンソースソフトウェアAltAnalyze (<http://code.google.com/p/altanalyze/downloads>, version 2.0.8 or higher)のグラフィック・ユーザー・インターフェースから、また、スタンドアロンpythonスクリプトとして (<https://github.com/nsalomonis/LineageProfilerIterate>)入手可能である。AltAnalyzeは、<http://www.altanalyze.org>からダウンロードし、ハードドライブに抽出し、初回起動の後にプロンプト(現在、Ensmart65)が表示された際に最新のヒットデータベースとともにインストールすることができる。あるいは、LineageProfiler・ファンクションは、このソフトウェアのコマンドラインバージョンを、<https://github.com/nsalomonis/LineageProfilerIterate>で入手可能なgene model discoveryのオプションとともに用いて実行することもできる。LineageProfilerのスタンドアロン・グラフィック・ユーザー・インターフェース・バージョンおよびコマンド・ライン・バージョンの実行命令は、<http://>

/code.google.com/p/altanalyze/wiki/SampleClassificationに記載されている。LineageProfilerのソースコードは、本明細書に記載の実施形態で使用するために修正され、LineageProfiler Iterateとなっている。本明細書で使用する場合、修正版LineageProfilerであるLineageProfiler IterateおよびkSASは互換的に使用される。kSASのソースコードは、付表Cに示されている。このソフトウェアを用いて、所与のサンプルセットに関する定量的発現値を、特定の疾患クラス、表現型、または処置カテゴリーに属すとして分類することができる。簡単に述べると、このアルゴリズムはこれを、所与のサンプルに関する発現値の入力セットを2以上参照条件と関連させることにより行う。このサンプルを参照と直接関連させるといよりも、既知のクラスに属するサンプルを用いて高い程度の予測成功をもたらすことが従前に確認されているモデルファイルから遺伝子のサブセットを選択することができる。このアルゴリズムはまた、別のまたは新たな遺伝子モデルを発見するために新たなデータに適用することもできる。

10

【0099】

以下の実施例は例示目的で示される。これらは本発明の特定の態様および実施形態を示すことを意図し、本発明を何ら限定するものではない。

【実施例】

【0100】

実施例1：腎臓移植における急性拒絶を評定するための組成物および方法の開発のための研究計画

20

異なる年齢、異なる免疫抑制、および移植センター特異的プロトコルを受けたレシピエントにおける急性拒絶（AR）診断および予測のための簡単な血液QPCR検査を開発するために、腎臓移植における急性拒絶（AART）の評定試験を、世界8か所の腎臓移植センターでの共同努力において計画し、438名の成人および小児腎臓移植患者由来の558の末梢血（PB）サンプルを使用した。

【0101】

図1は、世界8か所の移植センターからの438名のユニークな成人/小児腎臓移植患者における腎臓移植における急性拒絶（AART）の評定研究計画を示す：エモリー、UCLA、UPMC、CPMC、UCSF、およびバルセロナが寄与した成人サンプル、メキシコ、およびスタンフォード小児サンプル。AR QPCR分析については、サンプルを4つのコホート：遺伝子モデル化のためのコホート1 n = 143成人サンプル；独立ARバリデーションのためのコホート2 n = 124成人/小児サンプル；AR予測のためのコホート3 n = 191成人/小児サンプル；最終的なARアッセイ固定および臨床解釈のためのコホート4：n = 100成人/小児サンプルに分けた。

30

【0102】

臨床経過観察受診の際に移植レシピエントから横断的に血液サンプルを採取し、同時腎臓同種移植生検と照合させた。本AART試験に参加したセンターは、スタンフォード大学（スタンフォード；n = 162小児サンプル）；Laboratorio de Investigacion en Nefrologia, Hospital Infantil de Mexico（Mex；n = 23小児サンプル）；エモリー大学、アトランタ、ジョージア州（エモリー、n = 43成人サンプル）；カリフォルニア大学ロサンゼルス校、ロサンゼルス、CA（UCLA、n = 105成人サンプル）；ピッツバーグ大学、ピッツバーグ、PA（UPMC、n = 132成人サンプル）；カリフォルニア・パシフィック・メディカルセンター、サンフランシスコ、CA（CPMC、n = 37成人サンプル）；カリフォルニア大学サンフランシスコ校、サンフランシスコ、CA（UCSF、n = 40成人サンプル）；ベルピッジ大学病院、バルセロナ、スペイン（バルセロナ、n = 16サンプル）であった。サンプルを、遺伝子選択とモデルトレーニングのための143のARおよび非AR成人サンプル（コホートI）のトレーニングセット、AR検出用の遺伝子のバリデーションのための124のARおよび非AR成人（>21歳）および小児（<21歳）サンプル（コホート2）の第1バリデーションセット、およびAR予測の評

40

50

価のための、拒絶生検の前と後の6か月以内に一連の採取を行った191の成人および小児サンプル(コホート3)の第2の予測的バリデーションセットに分けた。これらの3コホートを構成する血液サンプルを同時にマイクロ流体ハイスループットフリーダイムQPCRプラットフォーム(Biomark、Fluidigm Inc.、サンフランシスコ、CA)で、合計43の遺伝子に関して測定した。ARの非侵襲的検出のための17遺伝子の最終的な腎臓AR予測アッセイは、新規な数学アルゴリズム(kSAS)の開発を伴うABI QPCRプラットフォームにて、100の成人および小児サンプル(コホート4)の独立バリデーションセットに固定された(図1-研究計画、および表4、表5、患者の人口統計)。

【0103】

【表4】

表4: 438名のユニークな患者の統計

パラメータ	スタンフォード (P)	メキシコ (P)	UCLA (A)	エモリー (A)	UPMC (A)	CMPC (A)	UCSF (A)	バルセロナ (A)
全サンプル (Bxで確認済みのAR)	102 (54)	23 (15)	59 (37)	43 (18)	120 (48)	35 (4)	40 (14)	16 (8)
ドナーの年齢 (歳, 平均±SD)	33.3±14.6	34.3±6	41.4±12.4	36.2±12	49.8±12.8	48±13	43±12.8	45.2±11.3
ドナーの性別 (%男性)	47	44	42.7	38	37.6	46	45	61.7
移植片タイプ (縮小%)	37.4	28	42.7	70.4	17	43%	60	91.4
レシピエントの年齢 (歳, 平均±SD)	12.5±5	14.7±6	44.9±12.2	46±15	47±18	51.6±16	53.4±11	47±10.8
レシピエントの性別 (%男性)	62	65	74.8	62	61.4	60.7%	56.5%	55%
HLA 不適合 (x/6 平均±SD)	3.1±1.5	3.3±1.1	3.8±1.5	3.7±1.5	3±1.7	3.1±2	3.4±1.8	3.9±0.75
PRA >20% (%患者)	3	0	10	16.7	11.4	22.7%	8	6.3
誘導療法	Dac	Thymo	Thymo	Thymo	抗CD52	Thymo	Thymo	Thymo
一次IS	CNI/MMF/±CS	CNI/MMF/±CS	CNI/MMF/CS	CNI/MMF/CS	CNI/MMF	CNI/MMF/CS	CNI/MMF/CS	CNI/MMF/CS
採血法	PAXgene	PAXgene	PBMC	PBMC	PBL	PAXgene	PBMC	PAXgene
集中SOP	有	無	無	無	無	無	無	無
集中RNA抽出	有	無	無	有	無	有	有	無

HLA=ヒト白血球抗原;PRA=パネル反応性抗体;P=小児;A=成人;UPMC=ピッツバーグ大学メディカルセンター;UCLA=カリフォルニア大学ロサンゼルス校;CPMC=カリフォルニア・パシフィック・メディカルセンター,スタンフォードU=スタンフォード大学,エモリーU=エモリー大学;CNI=カルシニューリン阻害剤,DAC=ダクリズマブ,Thymo=サイモグロブリン,MMF=ミコフェノール酸モフェチル;CS=コルチコステロイド。

【0104】

【表 5】

表 5: 遺伝子セットのバリデーションに用いた 659 のユニークな小児 (n=293) および成人 (n=366) サンプルの患者およびサンプルの統計

パラメーター	スタン フォード (Peds)	SNS (Peds)	メキシ コ (Peds)	UCLA (成 人)	エモリ ー (成人)	UPMC (成 人)	CMPC (成 人)	UCSF (成 人)	バルセ ロナ (成人)
全サンプル(生 検で確認済みの AR)	93 (22)	177 (22)	23 (15)	109 (32)	60 (17)	92 (28)	53 (4)	36 (14)	16 (8)
ドナーの年齢 (歳, 平均±SD)	33.3±1 4.6	29.3± 10.4		41.4 ±12. 4	36.2± 12	49.8 ±12. 8			45.2± 11.3
ドナーの性別 (%男性)	47	61		42.7	38	37.6			61.7
移植片タイプ (縮小%)	37.4	53		42.7	70.4	9			91.4
レシピエントの 年齢 (歳, 平均±SD)	12.5±5	11.6± 6		44.9 ±12. 2	46±15	47±1 8			47±10 .8
レシピエントの 性別 (%男性)	62	57		74.8	62	61.4			
HLA 不適合 (x/6, 平均±SD)	3.1±1. 5	3.9±1 .5		3.8± 1.5	3.7±1 .5	3±1. 7			3.9±0 .75
誘導療法	Dac ^a	Dac ^a	Thymo ^b	Thymo ^b	Thymo ^b	抗 CD52	Thymo ^b	Thymo ^b	Thymo ^b
一次免疫抑制薬	CNI ^c , MMF ^d , +/- CS ^e	CNI ^c , MMF ^d , +/- CS ^e	CNI ^c , MMF ^d	CNI ^c , MMF ^d , CS ^e	CNI ^c , MMF ^d , CS ^e	CNI ^c , MMF ^d	CNI ^c , MMF ^d , CS ^e	CNI ^c , MMF ^d , CS ^e	CNI ^c , MMF ^d , CS ^e
採血法	PAXgene	PAXgen e	PAXgen e	PBMC	PBMC	PBL	PAXge ne	PBMC	PAXgen e
中央研究室採血 /保存/処理 SOP (対して、セン ター特異的 SOP)	有	有	無	無	無	無	無	無	無
中央研究室での RNA 単離	有	有	無	無	有	無	有	有	無

^a Dac=ダクリズマブ; ^b Thymo=サイモグロブリン; ^c CNI=カルシニューリン阻害剤;

^d MMF=ミコフェノール酸モフェチル; ^e CS=コルチコステロイド

実施例 2 : 血液サンプル

ユニークな小児 (移植時のレシピエントの年齢 = 0.8 ~ 21.9 歳; n = 200) および成人 (移植時のレシピエントの年齢 = 23 ~ 78 歳; n = 315) 腎臓移植レシピエントに由来する末梢血サンプル (n = 518) を、生検により確認された急性腎臓同種移植片拒絶の非侵襲的診断のための一般末梢血遺伝子パネルの開発に用いた。200 サンプルの小児コホートのうち 177 サンプルは、12 の米国移植センターから組織学的にグレード分類された AR を有する患者と有さない患者の両方が組み入れられた前向き多施設共同 NIH / NIAID 助成臨床試験の一環として従前に得られたものであった (SNS 0

10

20

30

40

50

1 ; NCT00141037 ; www.ClinicalTrials.gov ; Li, L., et al. Am. J. Transplant. 2012, 12, 2710-2718)。残りの23サンプルは、本試験のために the Laboratorio de Investigacion en Nefrologia、Hospital Infantil de Mexicoからのみ得られたものであった。315サンプルの成人コホートのうち、米国およびヨーロッパの6か所の移植センターからサンプルが得られた (n = 48 : エモリー大学、アトランタ、ジョージア州、外科 (エモリー) ; n = 97 : カリフォルニア大学ロサンゼルス校、ロサンゼルス、CA、免疫遺伝学センター (UCLA) ; n = 92 : ピッツバーグ大学、ピッツバーグ、PA、E. Starzl 移植センター (ピッツバーグ) ; n = 39 : カリフォルニア・パシフィック・メディカルセンター、サンフランシスコ、CA (CPMC) ; n = 23 : カリフォルニア大学サンフランシスコ校、サンフランシスコ、CA、腎臓科 (UCSF) ; n = 16 : ベルビッジェ大学病院、腎臓移植ユニット バルセロナ、スペイン (バルセロナ))。本試験は全ての現地IRB委員会により承認され、全患者がインフォームド・コンセントにより参加に同意した。

【0105】

本試験の各末梢血サンプルは、同じ患者からの同時 (48時間以内) 腎臓同種移植生検と対となるようにした。監視生検は全ての患者から、移植時、移植後3、6、12、および24か月の時点、ならびに移植片の機能不全が疑われた時点で得た。各生検が決定的な表現型診断を有する限り、同じ患者からの複数の末梢血 - 生検対を使用した。各生検は、Banff分類 (Solez, K. et al. Am. J. Transplant., 2008, 8, 753-760; Mengel, M. et al. Am. J. Transplant. 2012, 12, 563-570) に従い、各登録臨床サイトに関して、中央の病理学者によりスコア化された。末梢血 - 生検対は、「急性拒絶」(AR ; n = 130)、「安定」(非AR) または「その他」診断 (その他) として類別された。「急性拒絶」は、Banff尿細管炎スコア (t) が1であり、間質性浸潤物スコアが0であるサンプルに対して定義された。「安定」は、ARまたは他の任意の実質的病態の不在を示すサンプルに関して定義された。「その他」は、サンプルに関して定義された。Banffによりグレード分類されたARの不在を示すが、慢性同種移植損傷 (CAI ; サンプルのIFTAグレードは1である ; n = 51)、慢性カルシニユリン阻害剤毒性 (CNIT ; n = 19)、BKウイルス感染 (BKV ; n = 3)、または他の移植片損傷 (OGI ; n = 153) のBanff判定基準のいずれかを満たすサンプルに関して定義される。

【0106】

実施例3 : 患者

成人および小児セットI

表5は成人および小児セットIを示す。

【0107】

一例では、小児サンプルと成人サンプルを合わせたものを、テスト (n = 236 ; 143名の成人、93名の小児) およびバリデーション (n = 292 ; 208名の成人、84名の小児、表5) のための2群に分けた。

【0108】

成人および小児セットII

別の例では、小児サンプルと成人サンプルを合わせたものをトレーニングおよびテスト (n = 143名の成人)、バリデーション (n = 124 ; 59名の成人、65名の小児)、および独立予測 (n = 191 ; 130名の成人、61名の小児、表4) のための3群に分けた。

【0109】

成人および小児セットIII

別の例では、小児サンプルと成人サンプルを合わせたものをバリデーションのために100サンプルに分けた (77名の成人、23名の小児、表4)。

【0110】

10

20

30

40

50

実施例 4：血液サンプルの採取およびRNAの処理**血液サンプルの採取**

血液は、2.5 mLのPAXgene（商標）血液RNAチューブ（PreAnalytiX、Qiagen、パレンシア、CA）または末梢血リンパ球（PBL）単離用のフィコールチューブに採取した。PBLサンプルは、Affymetrixシステムでのマイクロアレイ・ディスカバリーにのみ使用した。PAXgene（商標）チューブにはPreAnalytiXの列に基づく方法キット（Qiagen）またはPBLサンプルにはRNeasy（Qiagen）を製造者のプロトコルに従って用いて全RNAを抽出した。

【0111】**RNA抽出**

全RNAを、2100バイオアナライザーにてRNA 6000 Nano LabChip（登録商標）キット（両方ともAgilent Technologies、サンタクララ、CA）を用いてRNAの完全性に関して評価した。なお、好適なRNAはRNA integrity number（RIN）が7を超える場合と定義される（Fleige, S. and Pfaffl, M. W. Mol. Aspects. Med. 2006, 27, 126-139.; Schroeder, A. et al. BMC Mol. Biol. 2006, 7, 3）。

【0112】**cDNAの合成**

cDNAの合成は、SuperScript（登録商標）II第1鎖cDNA合成キット（Invitrogen、カールスバッド、CA）を製造者のプロトコルに従って用い、末梢血サンプルからの抽出品質のmRNA 250 ngを用いて行った。

【0113】**実施例 5：QPCR****マイクロ流体QPCRのための全RNAサンプルの調製**

マイクロ流体qPCR分析向けのサンプルは、特異的標的増幅のためのcDNA合成からの1.52 ng（相対量）の全RNAとサンプル希釈溶液を用い、フリューダインプロトコル（Fluidigm、サウスサンフランシスコ、CA）に従い、18Sを除く各検討遺伝子に対する個々のABI Taqmanアッセイをプールしたものをを用いて調製した。簡単に述べれば、1.52 ngのcDNAを用い、プールしたTaqmanアッセイをTaqman PreAmp Master mix（ABI）と複合化して、最終容量5 µLで特異的標的増幅を行った。増幅は、サーマルサイクラー（Eppendorf Vapo-Protect、ハンブルク、ドイツ）にて18サイクルの後に達成された。次に、サンプルを無菌水（Gibco、Invitrogen、カールズバッド、CA）で1：5希釈した。

【0114】**QPCR**

マイクロ流体qPCRは、96.96 Dynamic Array（Fluidigm）にて、製造者のプロトコルに概説されているように、各mRNAに対するTaqman Assays（Applied Biosystems、フォスターシティ、CA）、Taqman Universal master mix（Applied Biosystems）、およびローディング試薬（Fluidigm）とともに、特異的標的増幅から得られた2.25 µLの希釈サンプルを用いて行った。このチップをブライムし、HX IFC Controller（Fluidigm）によってロードし、製造者のプロトコル（Fluidigm）に示されているように、BioMark（Fluidigm）にて遺伝子発現データ採取用のデフォルトパラメーターを用い、qPCRを行った。標準比較Ct値を使用し、内部対照参照としての18Sと外部比較参照としてのUniversal Human Total RNAを用い、遺伝子発現の相対的変化倍率値を求めた（Qiagen、Venlo、Limburg）。

【0115】

10

20

30

40

50

ABI QPCR

標準プロトコールは、ABI 7900 Sequence Detection SystemまたはViiA7 (Applied Biosystems)で、標準条件(95 で10分、95 で15秒、60 で30秒の40サイクル)および遺伝子発現アッセイ(Applied Biosystems)を使用するqPCR反応に従った。RNA発現の相対量は比較Ct法を用いて算出した。発現値は、リボソームRNA内部参照およびUniversal Human Total RNA (Qiagen.)を用いて18Sに対して正規化した。

【0116】

実施例6：データの前処理および正規化

マイクロ流体QPCRデータ/前処理および正規化

フリューダ임ハイスルーブットマイクロ流体qPCR技術を用いて43の遺伝子のより大きな遺伝子パネルの発現を測定するために236の成人および小児サンプルからのRNAで行った42, 792のqPCR反応から得られた生のCt値を、6つの96.96マイクロ流体チップ(フリューダ임)から採取した。フリューダ임リアルタイムPCR分析ソフトウェアによりCt値を抽出し、エクセル(Microsoft Office 2012, Microsoft, CA)にアップロードした。テクニカルレプリケートに関する平均Ct値は、反復に関する標準偏差が<0.5であった場合に計算した。Ct値は、Ct(dCt)法(Livak, K. J. and Schmittgen, T. D. Methods 2001, 25, 402-408)を用い、内在対照遺伝子に対して正規化した。4つの異なる対照遺伝子、リボソーム18S、アクチン(ACTB)、グリセルアルデヒドリン酸デヒドロゲナーゼ(GAPDH)、およびβ2ミクログロブリン(B2M)を検定した。18Sが全てのサンプルで最小の変動を示したことから、それをdCt値の計算のために選択した。欠損値は5個の近傍を用いるK近傍法(K nearest neighbor) (KNN; Troyanskaya, O. et al. Bioinformatics 2001, 17, 520-525)により入力した。生のqPCRデータの可視化は、Partek Genomics Suite v. 6.6 (Partek Inc., USA)にて、教師なしの主成分分析(Principal Component Analysis) (PCA)および階層的クラスタリングを用いて行った。遺伝子発現の交絡因子はPCAおよび分散分析(ANOVA)により特定し、Partek (カテゴリー因子と数値因子を組み合わせた混合モデルANOVA)でのバッチ効果除去のより、また、SVARパッケージ内のコンバットファンクション(combat function)とともに経験ベイズ法を使用することにより補正した。この方法は小規模なサンプルのアウトライアーに対してロバストなものである(Chapelle, O., Haffner, P., and Vapnik, V. N. IEEE Trans. Neural Netw. 1999, 10, 1055-1064)。次に、43遺伝子のより大きなパネルのための正規化発現データを、以下に概略を示すように、ARと非ARの間の差次的発現遺伝子の同定のため、また、異なる年齢群のARの機序のより良い理解のため、また、ARの最高の予測力、感度、および特異度を有する遺伝子の選択ために用いた。

【0117】

Partekにおけるバッチ効果除去を用いたマイクロ流体QPCRデータにおける交絡因子の補正

143のARおよび非ARサンプルの成人データセットにおいて、サンプル間の差次的遺伝子発現に対するRNAソース、PCRプレートのテクニカル効果、および移植センターの外的効果を混合ANOVAモデルで評価した。RNAソース、PCRプレート、および移植センターはランダムカテゴリー因子として含まれ、表現型(AR、非AR)はカテゴリー因子として含まれていた。各因子に関してp値を計算し、<0.05のp値は、特定の遺伝子の差次的発現がANOVAに含まれた因子のいずれか1つに関連していたことを示した。ANOVAモデルに基づくPartekのバッチ効果除去の特徴は、マイクロアレイチップが異なるバッチでハイブリダイズされた場合にマイクロアレイデータにおける差次的遺伝子発現の効果を除去するように初期設計されていた。次に、この特徴を用いて、RNAソース、PCRプレート、および移植センターのp値が1となるようにデータ

10

20

30

40

50

を調整する混合4元配置ANOVAモデルを構築することにより、RNAソース、PCRプレート、および移植センターの望ましくないランダム因子に関して補正を行う。この方法では、これらの因子による遺伝子発現の違いは無く、表現型のp値は最大となった(図11A~D)。

【0118】

43の拒絶遺伝子に関する143のARおよび非AR成人サンプル(コホート1)からのQPCRデータの主成分分析(PCA)により、表現型(図11B)によるよりもむしろサンプル採取場所によるサンプル分離が明らかとなった(図11A)。混合ANOVAによるQPCRデータの正規化は、遺伝子発現に対するサンプル採取場所の主要な効果を補正し(図11C)、ARおよび非ARへのサンプルの分離をもたらした(図11D)。PCAは、43の遺伝子の相対的遺伝子発現値(dCt_{18S})を用いて行った。混合ANOVAモデルは、ランダムカテゴリー因子としてサンプル採取場所、RNAソースおよびチップ、ならびにカテゴリー因子として表現型を用いて構築した。各球は1サンプルを表し;記号はサンプル採取場所を表し(* = UPMC; = UCL A; X = CPMC; # = エモリー);この図はまた生検診断に基づく患者表現型(AR;非AR)も表す。

10

【0119】

Rにおける経験ベイズ法を用いたマイクロ流体QPCRデータの交絡因子の補正

143ARおよび非ARサンプルの成人データセットにおける変数選択の前に、バッチ効果を除去するためにSVA Rパッケージのコンバットファンクションとともに経験ベイズ法を用い、43遺伝子の発現を正規化した。この方法は小規模サンプルのアウトライアーに対してロバストなものである。

20

【0120】

Abi QPCRデータの処理および正規化

17遺伝子の発現を測定するための100の成人および小児サンプルに由来するRNAで行ったABI QPCR反応から得られた生のCt値を、384ウェルプレートから採取した。ABI v i i a 7 PCR分析ソフトウェアによりCt値を抽出し、エクセル(Microsoft Office 2012、Microsoft、CA)にアップロードした。テクニカルレプリケートに関する平均Ct値は、反復に関する標準偏差が<0.5であった場合に計算した。Ct値は、ここに記載の方法(Livak, K. J. and Schmittgen, T. D. Methods 2001, 25, 402-408)を用い、t値(dCt)の計算のために内部対照遺伝子としてのリボゾーム18S RNAに対して、さらにまた、Ct値(ddCt)の計算のためにヒトUniversal RNA(Qiagen)に対して正規化した。

30

【0121】

実施例7: ARおよび非AR特異的遺伝子の選択法

合計43の遺伝子を、ARおよび非AR特異的遺伝子の選択に用いた。遺伝子は、小児および成人移植拒絶における従前のマイクロアレイ試験(Li, L. et al. Am. J. Transplant. 2012, 12, 2710-2718; Naesens, M. et al. Kidney Int. 2011, 80, 1364-1376; Sarwal, M. et al. N. Engl. J. Med. 2003, 349, 125-138)に基づき、安定な同種移植(表2)に比べて差次的に変更されかつARに関連していることが確認されたものである。これら合計43の遺伝子のうち、10(CFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130およびRYBP)は、小児腎臓移植においてARの予測モデルの開発に焦点を当てた従前の研究で同定されたものである(Li, L. et al. Am. J. Transplant. 2012, 12, 2710-2718)。残りの33遺伝子は、種々のタイプの実質臓器移植の安定同種移植に比べてARにおいて、マイクロアレイデータのメタ分析で決定した際に差次的に変更されていた(Khatri et al. JEM, 2013, 出版者受理済み)。

40

【0122】

実施例8: ARと非ARの間の差次的発現遺伝子の同定方法

ARと非ARの間で有意に差次的に発現される遺伝子を検出するため、および年齢群間

50

のARの機序理解を助けるために、一元配置および多元配置ANOVA、有意に異なる分散の場合にはウェルチ補正を伴う対応のない学生t検定、および多重比較に関して補正するための偽陽性率(FDR)の算出を用いた； < 0.05 のp値または $FDR < 5\%$ を有意と見なした(図12)。

【0123】

実施例9：ARおよび非ARを識別する遺伝子の同定方法

143の成人サンプルにおけるこれまでに公開されている遺伝子の評価

より大きな43遺伝子のパネルに由来するこれまでに公開されている10遺伝子パネル(CFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130およびRYBP；Li, L. et al., Am. J. Transplant 2012, 12, 2710-2718)を、143サンプルの成人検定データセットにおいてAR表現型の識別および予測のために用いた(図12)。

【0124】

236の成人および小児サンプルにおける新規な遺伝子の同定

年齢、移植センター、およびRNAソースに依存せず、最小数の遺伝子を用いてARの高い予測力を有する新規な遺伝子パネルを定義するために、236サンプルの成人データおよび小児データを合わせたセットに、Shrinking Centroids (Tibshirani, R. et al. Proc. Natl. Acad. Sci. USA 2002, 99, 6567-6572; Storey, J. D. and Tibshirani, R. Proc. Natl. Acad. Sci. USA 2003, 100, 9440-9445)、前進的および後退的選択、および遺伝的アルゴリズム(Zhu, Z. et al., IEEE Trans. Syst. Man, Cybern. B Cybern. 2007, 37, 70-76)を適用した。加えて、徹底的な検索を適用し、236のサンプルにおいて分析した43の遺伝子から5遺伝子のあり得る全ての組合せを検索することによって上位ランク遺伝子を定義した。Shrinking Centroidsアプローチにおいて、43の遺伝子からあり得る全ての遺伝子組合せを、最小数の遺伝子セットを5、最大数の遺伝子セットを20として、1遺伝子刻みで一度に検定した。遺伝子をそれらのAR予測確率に関して交差検証(1-LOCV)によって検定した。合計1872のモデルを、117の異なるアルゴリズムを用いて検定した。結果を、遺伝子の数、および5遺伝子から43遺伝子の範囲の同じ遺伝子組合せのAUCに従ってランク付けした。結果として最高の平均AUCを有する組合せを選択した。前進的選択(ステップアップ)および遺伝的アルゴリズムにおいて、遺伝子パネルの増加数を検定した： $n = 5, 7, 10, 12, 13, 15, 17$ 、および20をそれぞれ上記のような117の異なるアルゴリズムを用いて検定した。結果を比較し、これらのモデルの少なくとも50%で選択された最終遺伝子を選んだ。この遺伝的アルゴリズムでは、遺伝子パネルが無作為に選択された初期集団を、各遺伝子が少なくとも50回検定されるように定義した。以下の集団を検定した：下式：

【数1】

$$N = \frac{43 \times 50}{n}$$

【0125】

(式中、Nは初期集団サイズであり、nは遺伝子パネルのサイズであり、50はある遺伝子が初期集団に現れなければならない回数を表し、43は引き出される遺伝子の合計数である)(図12)に従い、 $n = 5, 7, 10, 12, 13, 17$ 、および20の遺伝子パネルについてそれぞれ430、308、215、180、166、127、および108。

【0126】

143の成人サンプルにおける新規な遺伝子の同定

変数選択の前に、バッチ効果を除去するために、経験ベイズ法をSVARパッケージのコンパクトファンクションとともに用いることによって遺伝子を正規化した。この方法は小規模なサンプルのアウトライアーに対してロバストなものである。このデータは、1

43の観測値および43の遺伝子と、疎であることから、本発明者らは、Rパッケージを用いて患者サンプルを分類するためにペナルティ付きロジスティック回帰を使用した。このアプローチにより、回帰係数の正確な推定値だけでなく、各患者の確率推定値も得られる。本発明者らは、評価のためのCoordinate Descentによる一般線形モデルに関する正側化パスを用いた(図12)。

【0127】

実施例10：ARおよび非ARを判別するための遺伝子の評価方法

遺伝子選択を、ARの判別および予測に関して、等しいおよび比例する事前確率を有する判別分析(Discriminant Analyses)(DA)、サポートベクターマシン(Support Vector Machine)(SVM)、ロジスティック回帰(logistic regression)(LR)および等しい事前確率の部分最小二乗DA(Chapelle, O. et al., IEEE Trans. Neural Netw. 1999, 10, 1055-1064; Brown, M. P. et al., Proc. Nat. Acad. Sci. 2000, 97, 262-267)の、核関数ラジアル基底関数(radial basis function)(rbf)、部分最小二乗(partial least square)(pls)DA(Perez-Enciso, M. and Tenenhaus, M. Hum. Genet. 2003, 112, 581-592; Gottfries, J. et al., Dementia 1995, 6, 83-88)との併用によって評価した。SVM分類では、経験分類誤差の最小化と幾何学幅の最大化を同時に行うことによってARと非ARの最も広い分離を与えた決定面を定義する最良の一般非線形ベクター(「サポートベクター」)を見出すために、正側化パスラジアル基底関数(rbf)を使用する。SVMは、特徴(遺伝子)およびサンプル数の希薄なデータセットに対して十分な性能を示す(Nouretdinov, I. et al., Neuroimage 2011, 56, 809-813)。第一種過誤を最小にするために、データセットを分離したトレーニングセットとテストセットに分けるのではなく、10分の1レベル一個抜き交差検証(a ten-fold one level leave one-out cross validation)(1-LLCV)を行った。成人データセットと小児データセットを合わせたものでこれらの遺伝子による予測力、感度、および特異度を評定するために、各分類法にARの曲線下面積(AUC)および事後確率を与えた。各遺伝子選択アプローチから最高のAR予測力、最高の感度、および最高の特異度を有する遺伝子を、ABIプラットフォーム(Abi v i i a 7、L i f e T e c h n o l o g i e s、フォスターシティ、CA)でのqPCR用の17遺伝子の最終選択のために比較した。必要な場合には、ARと非ARを比較する学生t検定およびANOVAからのp値およびFDR値を使用した。最終遺伝子選択のワークフローを図12に示す。

【0128】

実施例11：フリータイムQPCRデータにおけるARと非ARの分類のためのアルゴリズムの開発方法

17遺伝子を用いた236の小児および成人サンプルにおけるアルゴリズムの同定

合計122の分類アルゴリズムを、選択された遺伝子(合計17)を用い、2レベル一個抜きネステッド交差検証(two level-leave one out nested cross validation)(2-LLCV)および5アウター、5インナーのデータ分割で検定した。「インナー」交差検証(cross validation)(CV)は、予測因子変数と最適モデルパラメーター選択のために行い、「アウター」CVは、分類子に関して総合的精度推定値を得るために用いた。「インナー」CVは、テストセットの提供に適用される最適モデルを選択するために、テストデータとして提供されなかったトレーニングデータで行った。236サンプルで検定された分類モデルは、判別関数および等しいまたは比例する事前確率、ユークリッド距離測度、平均ユークリッド距離測度またはコサイン非類似性距離測度および5個の近傍を用いるKNN、等しいまたは比例する事前確率、LR、およびSVMを用いる最近傍重心を含んだ。

【0129】

17遺伝子を用いた143の成人サンプルにおけるアルゴリズムの同定

合計122の分類アルゴリズムを、143の成人サンプルにおいて、選択された遺伝子(合計17)を用い、2レベル一個抜きネステッド交差検証(2-LLCV)および10アウター、10インナーのデータ分割で検定した。「インナー」交差検証(CV)は、

予測因子変数と最適モデルパラメータ選択のために行い、「アウター」CVは、分類子に関して総合的精度推定値を得るために用いた。「インナー」CVは、テストセットの提供に適用される最適モデルを選択するために、テストデータとして提供されなかったトレーニングデータで行った。143サンプルで検定された分類モデルは、等しいまたは比例する事前確率を用いる部分最小二乗判別分析および線形判別分析、サポートベクターマシン、ユークリッド距離測度、平均ユークリッド距離測度またはコサイン非類似性距離測度および5個の近傍を用いるKNN、等しいまたは比例する事前確率、およびLRを用いる最近傍重心を含んだ。上位モデルを143サンプルにおいて、1-LLOCVを用いて評価した。精度の尺度は、正解率、感度、特異度、NPV、PPV、および受信者動作者曲線下面積(AUC)であった。

10

【0130】

15遺伝子を用いた143の成人サンプルにおけるアルゴリズムの同定

本発明者らは、ARと非ARを分類するために、ブートストラップサンプル(29テスト、114トレーニング、置換によるサンプリング)を用い、100のElastic Netロジスティック回帰モデルを43の遺伝子に当てはめた。各ブートストラップに関して、ネステッド交差検証ループは、逸脱度に従っての最良値を推定した。Elastic-Netのパラメータは推奨値の.95に固定された。これらの遺伝子のランク付けを行うために、本発明者らは、各遺伝子が100ブートストラップにおいてElastic-Netにより選択された回数を数えた。ブートストラップサンプルのそれぞれについて、Elastic-Netは43の遺伝子のサブセットを非ゼロ係数に当てはめる。100のブートストラップモデルの実行の後、本発明者らは最大数の非ゼロ係数を有するK遺伝子を選択した。第2工程において、予測成績(分類率、感度、特異度、PPV、NPV)のバイアスのない推定を得るために、本発明者らは、に対するネステッド交差検証を用い、この場合には工程1で選択されたK遺伝子のセットのみを用いて、別のセットの100ブートストラップElastic-Net分類を復興した。本発明者らは分類率、感度、特異度、陽性的中率(PPV)および陰性的中率(NPV)を報告する。

20

【0131】

実施例12：フリータイムおよびABI QPCRデータにおけるARおよび非ARの判別および予測のためのアルゴリズムの開発方法

相関に基づくARおよび非AR分類の開発

各患者サンプルについてピアソンの相関係数()を計算するために、 Ct値をクエリーサンプルとして用い、AR分類サンプルまたは非AR分類サンプルのいずれかに関する平均遺伝子 Ct値と比較した

30

【0132】

各患者サンプルについてピアソンの相関係数()を計算するために、 Ct値をクエリーサンプルとして用い、AR分類サンプルまたは非AR分類サンプルのいずれかに関する平均遺伝子 Ct値と比較した。

【0133】

Zスコアは、次のように、各サンプル に関して、全サンプル比較からの全 値の平均(μ)および標準偏差()に関して計算する。

40

【数2】

$$Z = \frac{x - \mu_{\text{グローバル}}}{\sigma_{\text{グローバル}}}$$

【0134】

サンプルを、サンプルARおよび非AR zスコアの比較(ARまたは非ARにおいてはより大きなz)に基づきARまたは非ARとして分類した。これらの関数はAltAnalyzeのLineageProfilerIterate.pyモジュールに見出すことができる。

50

【0135】

相関分析は、適用可能であれば、4、5、6、7、8、9、10、11および12遺伝子セットの、全てのあり得る組合せに関して行った。ABI分析に関する最良レポートモデルを、サイズの異なる遺伝子セットを比較する場合には、全サンプルのうち正確に分類された患者サンプルのパーセンテージに基づきスコア化した。

【0136】

ARと非ARの分類のための相関に基づくアルゴリズムとしてのLineage Profilerの開発

Lineage Profiler (LP)の名称の新たな相関に基づくオープンソースのアルゴリズムを用い、さらなるqPCR評価のために最適な遺伝子モデルの発見のためにさらなる修正を行った。LPの入力はCt正規化患者サンプルqPCR値および2つの参照qPCRプロファイル(AR参照プロファイルおよび非AR参照プロファイル)である。この分析は5つの工程からなる:工程1:評価した遺伝子のパネルに対してRNA発現値のマトリックスをインポートする;工程2:各遺伝子について、全ARサンプルからの単一の参照発現ベクター(平均)および全非ARサンプルに対する単一の参照発現ベクターを作出し、保存する;工程3:各qPCRセット(遺伝子セット)に関して分析した遺伝子のあり得る全ての組合せを同定する;工程4:患者サンプルを分類するために、各遺伝子セットについて各患者RNAプロファイルを参照ARプロファイルおよび参照非ARプロファイルと直接比較する(LPを使用);および工程5:関連の参照プロファイルに関して上位予後リストを同定するために、既知のARおよび非AR状態に基づく遺伝子セットをランク付けする。17遺伝子由来の、4~12の範囲の数種類の長さの遺伝子セットを、各独特な測定プラットフォーム(フリータイムまたはABI)のため、およびあり得る全ての組合せのために作出した。フリータイム分析については、全遺伝子で開始して上位スコアモデルを反復同定し、さらにその後の全てのモデルの導出を分析する最適化関数が書かれていた。最も成績のよい遺伝子セットを同定した後、これらの遺伝子セットを固定し、別個のパリレーションデータセットに適用した。対応する参照発現プロファイルを用いた既存のまたは新たなデータセットの分析は、オープンソースソフトウェアAltAnalyzeバージョン2.0.8(<http://www.altanalyze.org>)にてLP関数を用いて行うことができる(図4A~B)。

【0137】

ARと非ARの分類のための相関に基づくアルゴリズムとしてのkSASの開発

サンプルのARまたは非ARとしてのロバストなリスク層別化のため、kSASという名称の新規な相関に基づくアルゴリズムを開発した。大規模なデータセットが評価される発見および交差検証における好適なアプローチである経験ベイズ法およびANOVAなどの方法により外的交絡因子を補正するのではなく、数、サンプル採取場所に依存しない、従って通常の臨床現場により適用度の高いサンプルの正確な前向き予測を可能とする、17遺伝子パネルに固定されたARおよび非AR QPCR参照プロファイルを適用するためにkSASを開発した。kSASでは、患者サンプル、および2つの参照QPCRプロファイル(1つは既知のAR用、1つは既知の非AR用)においてQPCR dCt(18S)値を使用する。kSAS分析は、トレーニングおよびテストのために5つの主要工程を含んでなる:1)全サンプルに対して17遺伝子dCt(18S)発現マトリックスをインポートする、2)各遺伝子に対する既知のAR発現ベクターおよび非AR発現ベクターを定義する;3)全遺伝子で開始して上位スコアモデルを反復同定する最適化関数を用いてあり得る全ての遺伝子の組合せを同定する;4)各患者について得られた全てのモデルを参照ARおよび非ARプロファイルと比較して、相関度(ピアソン相関係数)に基づき患者サンプルを分類する;5)上位予後リストを同定するために、相関により遺伝子セットをランク付けする。各患者サンプルについてピアソンの相関係数()を計算するために、本発明者らは、クエリーサンプルの各遺伝子のdCt(18S)値を、ARまたは非ARいずれかの参照の同じ遺伝子の平均dCt(18S)値と比較する。得られた各遺伝子モデルについて、AR - 非AR $\times 10$ を計算することによりリスクスコア

を算出した。得られた全てのモデルリスクスコアを合計し、各サンプルに関する統合 A R リスクスコアを得た。サンプルは、サンプル A R および非 A R リスクスコア (A R または非 A R においてより大きな相関) の比較に基づき A R または非 A R として分類された。相関分析は、適用可能であれば、4、5、6、7、8、9、10、11 および 12 遺伝子セットのあり得る全ての組合せについて行った。最良レポートモデルを、サイズの異なる遺伝子セットを比較する場合には、全てのサンプルのうち正確に分類された患者サンプルのパーセンテージに基づきスコア化した。遺伝子セットの例は表 2 にある。A R および非 A R プロファイルにおける採取場所に関連する変動に取り組むためには、各採取場所に対する個別の A R 参照および非 A R 参照は、各モデルについての相関により導かれたリスクスコアを算定する場合には、各個人サンプル比較のための最も相関の高い場所参照対を選択するために単一の表に示した (図 9 A ~ C) 。

10

【 0 1 3 8 】

相関に基づく A R と非 A R の分類のための新規参照データの作出

新規移植センターに関して参照を使用するためには、未知のサンプルの分析の前に、同じ方法で未知のサンプルとして採取され A R または非 A R として分類される血液を採取し、推奨される 1 2 遺伝子モデルセット (下記参照) を用いてプロファイリングを行うべきである。機械およびサンプル採取センターのバイアスがこれまでに見られていたことから、これらのサンプルは移植センター特異的参照として役立つ。十分な数のサンプルに対して q P C R プロファイルを得た後に、全ての A R サンプルおよび非 A R サンプルの平均発現を個別に採り、アッセイした全ての遺伝子について 2 列参照を作成した。あるいは、個々のサンプルの代わりにプールした R N A 参照の使用でも十分なはずである。このデータは、第 1 列が遺伝子 I D を含み、第 2 および第 3 列がそれぞれ A R 参照および非 A R 参照を含む、3 列タブ区切りテキストファイルとして保存する。まず、これらの参照サンプル間に有意な変動性が存在するかどうか (例えば、A R サンプルと非 A R サンプル間の不十分な分類スコア) を決定するためのこの参照に用いたオリジナルサンプルの再分析が奨励される。

20

【 0 1 3 9 】

実施例 1 3 : フリュウダインおよび A B I Q P C R データにおける A R および非 A R の判別および予測のための相関に基づくアルゴリズムの評価方法

非移植データにおける k S A S の評価

30

k S A S を A R 患者および非 A R 患者のデータに適用する前に、本発明者らは、従前に記載されている、G E I C A M / 9 9 0 6 臨床試験から得られた 8 1 4 サンプルに適用された 5 0 の乳癌予後マーカー遺伝子の Q P C R 分析に対してこのアプローチを評価した。k S A S は、無作為に選択した患者テストセット (2 7 2 の患者サンプル) を、残りのサンプルで参照作成 (トレーニング) を行った後に、5 つの明瞭に異なる予後乳癌群に上手く分類することができ、5 0 全てのマーカー遺伝子を用いて成功率は > 8 5 % であった。2 4 および 2 5 遺伝子のより小さな予後遺伝子モデルもまた、トレーニングセットにおいてより高いパーセンテージ (9 0 . 0 % 対 8 5 . 6 %) およびテストセットにおいて同等の精度 (8 3 . 1 ~ 8 3 . 8 %) で患者を分類することができた。

40

【 0 1 4 0 】

1 4 3 の成人フルウダイン Q P C R データにおける k S A S の評価

本発明者らは、1 4 3 の成人サンプル (コホート 1) の同じ正規化データセットで k S A S を評価した。参照 A R プロファイルおよび参照非 A R プロファイルをコホート 1 のランダム 2 / 3 トレーニングサンプルセットから、4 3 全ての遺伝子について得た。次に、このトレーニングセットを、上位スコア遺伝子モデルを同定するために、サイズの等しい 1 0 の A R / 非 A R 2 / 3 および 1 / 3 セットにプログラムによりさらに再分割した。このトレーニングセットからの最高スコアモデルを、トレーニングセット A R 参照プロファイルおよび非 A R 参照プロファイルを用いて、オリジナル 1 / 3 トレーニングセットで評価した。

【 0 1 4 1 】

50

100の成人および小児ABI v i i a Q P C Rデータによるにおけるk S A Sの評価

本発明者らは、100の成人および小児サンプルにおいて、k S A Sにより定義された13全ての12遺伝子モデルの、単一の遺伝子モデルに基づくものではないが13全ての12遺伝子モデルを含む、各患者についての単一の信頼スコアを提供する組合せ能を評価した。本発明者らは、100のARおよび非ARサンプル(26のAR、42の非AR)の合わせたデータセットに関する統合ARリスクスコアを算出した。統合ARリスク分析は、13の12遺伝子モデルによって患者が非ARと予測された回数を、同じ患者がARと予測された回数から差し引くことによって、各患者に関する数的ARリスクスコア(-13~13)を生成した。この統合リスクスコアに基づき、患者は高リスクAR、低リスクARまたは中リスクとして類別することができる。高リスクARのカットオフは統合スコア 9であり、低リスクARは統合スコア -9であった。統合スコア -7および7の患者は中リスクと見なされた(図9C)。

【0142】

実施例14：ARと非ARの分類のための相関に基づくアルゴリズムのためのソフトウェアの開発方法

本明細書に記載の相関に基づく分析は、AltAnalyzeバージョン2.0.8以降を用いて行うことができる。LineageProfilerは、オープンソースソフトウェアAltAnalyze(<http://code.google.com/p/altanalyze/downloads>, version 2.0.8 or higher)のグラフィック・ユーザー・インターフェースから、また、スタンドアロンpythonスクリプトとして(<https://github.com/nsalomonis/LineageProfilerIterate>)入手可能である。AltAnalyzeは、<http://www.altanalyze.org>からダウンロードし、ハードドライブに抽出し、初回起動の後にプロンプト(現在、EnsMart65)が表示された際に最新のヒトデータベースとともにインストールすることができる。あるいは、LineageProfiler・ファンクションは、このソフトウェアのコマンドラインバージョンを、<https://github.com/nsalomonis/LineageProfilerIterate>で入手可能なgene model discoveryのオプションとともに用いて実行することもできる。LineageProfilerのスタンドアロン・グラフィック・ユーザー・インターフェース・バージョンおよびコマンド・ライン・バージョンの実行命令は、<http://code.google.com/p/altanalyze/wiki/SampleClassification>に記載されている。LineageProfilerのソースコードは、本明細書に記載の実施形態で使用するために修正され、LineageProfiler Iterateとなっている。本明細書で使用する場合、修正版LineageProfilerであるLineageProfiler Iterateおよびk S A Sは互換的に使用される。k S A Sのソースコードは、付表Cに示されている。このソフトウェアを用いて、所与のサンプルセットに関する定量的発現値を、特定の疾患クラス、表現型、または処置カテゴリーに属すとして分類することができる。簡単に述べると、このアルゴリズムはこれを、所与のサンプルに関する発現値の入力セットを2以上参照条件と相関させることにより行う。このサンプルを参照と直接相関させるというよりも、既知のクラスに属するサンプルを用いて高い程度の予測成功をもたらすことが従前に確認されているモデルファイルから遺伝子のサブセットを選択することができる。このアルゴリズムはまた、別のまたは新たな遺伝子モデルを発見するために新たなデータに適用することもできる。

【0143】

Ct値(dCt)を用いたk S A SにおけるARおよび非AR分類のための発現ファイルの開発

AR分類は、ARおよび非AR判別遺伝子のパネルに関するqPCRにより誘導される発現値を対照18S遺伝子とともに用いて実行される。ABI v i i a 7プラットフォームにてqPCRから生成したCt値(18Sと比較)がこのアルゴリズムの未知のサンプル入力として使用される。さらに、参照ARプロファイルおよび参照非ARプロファイル(dCt)を含む参照ファイルもこのソフトウェアに供給される。

【0144】

Ct値(dCt)を用いたkSASにおけるARおよび非AR分類のための発現ファイルの開発

AR分類は、ARおよび非AR判別遺伝子のパネルに関するQPCRにより誘導される発現値を、対照18S遺伝子とともに用いて実行される。ABI viia7プラットフォームにてQPCRから生成した、18SおよびユニバーサルヒトRNAとの比較によるCt値がこのアルゴリズムの未知のサンプル入力として使用される。さらに、参照ARプロファイルおよび参照非ARプロファイル(ddCt)を含む参照ファイルもQPCRデータから導かれる。

【0145】

Ct値を用いたkSASにおけるAR分類のための発現ファイルの作成

発現ファイルは、file extension .txt.を有するタブ区切りテキストファイル形式の正規化発現値(qPCR Ct値)からなる。このファイルの第1列は、参照ファイルの第1列に一致するID(遺伝子記号)を含み、第1行はサンプル名を含み、残りのデータは正規化発現値(すなわち、Ct値)からなる。

【0146】

参照ファイルは、発現ファイルに見られるものと同じ値の範囲でのAR qPCR Ct値と非AR qPCR Ct値の複合体である。このファイルの全ての遺伝子記号は、発現ファイルに存在するものと一致すべきである。ソフトウェアの実行の際に、参照ファイルと発現ファイルの値が全体的に低い相関を示す場合(<90%)には警告が与えられる。理想的には、報告される相関係数の範囲は0.92~0.96またはそれを超えるべきである。そうではない場合には、試験を繰り返すかまたはさらなる品質対照に関して評価する必要のある場合がある。

【0147】

Ct値を用いたkSASにおけるAR分類のための発現ファイルの作成

発現ファイルは、file extension .txt.を有するタブ区切りテキストファイル形式の正規化発現値(qPCR Ct値)からなる。このファイルの第1列は、参照ファイルの第1列に一致するID(遺伝子記号)を含み、第1行はサンプル名を含み、残りのデータは正規化発現値(すなわち、Ct値)からなる。

【0148】

参照ファイルは、発現ファイルに見られるものと同じ値の範囲でのAR qPCR Ct値と非AR qPCR Ct値の複合体である。このファイルの全ての遺伝子記号は、発現ファイルに存在するものと一致すべきである。ソフトウェアの実行の際に、参照ファイルと発現ファイルの値が全体的に低い相関を示す場合(<90%)には警告が与えられる。理想的には、報告される相関係数の範囲は0.92~0.96またはそれを超えるべきである。そうではない場合には、試験を繰り返すかまたはさらなる品質対照に関して評価する必要のある場合がある。

【0149】

グラフィック・ユーザー・インターフェースを介したkSASにおけるARおよび非AR分類のためのkSASの使用

このアルゴリズムもまた、オープンソース分析パッケージAltAnalyzeで利用可能であり、別にインストールする必要はない。AltAnalyzeは、いくつかの異なる分析機能を含む大きなトランスクリプトーム分析ツールキットである。AltAnalyzeは大規模データベースのインストールを必要とするので、スクリプトのコマンドラインバージョンの使用が勧められる。

【0150】

AltAnalyzeの最新バージョンをインストールするためには、以下の5工程に従えばよい: 1) <http://code.google.com/p/altanalyze/downloads>へ行く; 2) 所与のオペレーティングシステムに適切な最新バージョンを探し当て、ダウンロードリンクに従う; 3) ハードドライブおよびアクセス可能な場所に.zipまたは.dmgファイルを抽出す

10

20

30

40

50

る；4) AltAnalyzeプログラムフォルダを開き、実行可能なAltAnalyze.exe (ウィンドウズ(登録商標))または等価なものをダブルクリックする；5) 小規模データベース(例えば、トウモロコシ(Zea mays))をダウンロードする手順を行い、“Download/update all gene-set analysis databases”のオプションの選択を解除する(提供されている遺伝子注釈はサンプルの分類には必要とされない)。

【0151】

入力ファイルは、未知のサンプル用の発現ファイルからなる。参照ファイルは、参照ARサンプルおよび参照非ARサンプル用の発現ファイルからなる。

【0152】

モデルファイルは、参照入力ファイルおよび発現入力ファイルの両方で一致するが、遺伝子セットのサブセットに相当しない遺伝子記号からなる。標準的なAR分類パネルは、13の12遺伝子モデルからなる。このファイルは毎回の分析に再使用することができる。

10

【0153】

kSASの出力は、全ての参照プロファイルに関連するスコアを含むタブ区切りテキストファイルである。この結果ファイルは、トレーニングセットサンプルの分析のために作成された。

【0154】

コマンドラインオプションを介したARおよび非AR分類のためのkSASの使用

LineageProfilerIterate/kSASスクリプトがダウンロードされれば、それを容易にアクセス可能な場所に移動するべきである。次に、ターミナルウィンドウを開かなければならない(PCでのコマンドプロンプトとも呼ばれる)。ターミナルまたはコマンドプロンプトウィンドウを開くための所与のオペレーティングシステムに対する命令は、オンラインで容易に見つけることができる。次に、ターミナルウィンドウでは、ディレクトリからLineageProfilerIterate/kSASスクリプトを含むフォルダにアクセスすべきである。

20

【0155】

入力ファイル、参照ファイル、およびモデルファイルの3つのファイルを作成する。

【0156】

Ct発現値を分析するために、入力ファイルおよび参照ファイル用の、Ct値を含む3ファイルの記憶位置をLineageProfilerIterate/kSASに提供する。コマンド - - i は、サンプル Ct発現値用である。コマンド - - r は、参照発現ファイル用である。コマンド - - m は、提供された13の12遺伝子モデル用である。このコマンドを実行した後、種々の印刷物が見られる。この時に結果を指定の結果ディレクトリに保存する。

30

【0157】

Ct発現値を分析するために、入力ファイルおよび参照ファイル用の、Ct値を含む3ファイルの記憶位置をLineageProfilerIterate/kSASに提供する。コマンド - - i は、サンプル Ct発現値用である。コマンド - - r は、参照発現ファイル用である。コマンド - - m は、提供された13の12遺伝子モデル用である。このコマンドを実行した後、種々の印刷物が見られる。この時に結果を指定の結果ディレクトリに保存する。

40

【0158】

AltAnalyzeのkSASの実行

上記の手順を用いてAltAnalyzeをインストールした後、入力データの分析を実行することができる。このために、適当な発現、参照、およびモデルファイルが必要である。

【0159】

Ct値を用いてkSASを実行するためには、以下の6工程に従えばよい：1) AltAnalyzeを開き、“Begin Analysis”を選択する；2) プラットフォーム分析メ

50

ニューの“Continue”を選択する；3) “Additional Analyses”を選択し、継続する；
4) “Lineage Analysis”を選択し、継続する；4) 発現ファイル(d C t)、参照ファイル(d C t)およびモデルファイルを提供し、継続する；5) 分類分析の進行が印刷される；および6) 完了したら、継続を選択し、結果フォルダが発現ファイルの記憶場所に提示される。

【0160】

C t値を用いてk S A Sを実行するためには、以下の6工程に従えばよい：1) A l t A n a l y z eを開き、“Begin Analysis”を選択する；2) プラットフォーム分析メニューの“Continue”を選択する；3) “Additional Analyses”を選択し、継続する；4) “Lineage Analysis”を選択し、継続する；4) 発現ファイル(d d C t)、参照ファイル(d d C t)およびモデルファイルを提供し、継続する；5) 分類分析の進行が印刷される；および6) 完了したら、継続を選択し、結果フォルダが発現ファイルの記憶場所に提示される。

10

【0161】

k S A Sにおいて生成された結果の解釈

複数のフィールドがフォルダサンプル分類内の結果ファイルに提示される。タブ区切りテキストファイルはエクセルで開くことができる。データは次のように提示される：

A列：サンプル - サンプル名を示す

B列：A R予測ヒット - A Rが予測されるモデルの数を示す

C列：非A R予測ヒット - 非A Rが予測されるモデルの数を示す

D列：コンポジット予後スコア - B ~ C列のスコアを合わせたもの

E列：Zスコアの差の中央値 - G ~ S列のZスコアの中央値

F列：予後リスク - 総合的予測リスク評定

G ~ S列：A R予測ヒット - 各サンプルおよびモデルについての個々のスコア

20

【0162】

予後リスク(F列)は、サンプルを「高リスクA R」、「中リスクA R」および「低リスクA R」として表示する。「低リスクA R」は、履歴により証明された安定移植片を有する個人に最もよく類似し、「高リスクA R」は、生検により証明されたA R移植片に最もよく類似すると見なされる。中リスクは、予後評価において13のモデル間でいずれかの不一致があるサンプルに割り当てられる。

30

【0163】

U C S Fからの40サンプルにおいて、生検により証明されたA R - サンプルは、13の全遺伝子セットのうち、A Rとして8遺伝子セットの予測および非A Rとして5遺伝子セットの予測を示し、各遺伝子セットは12遺伝子から構成された。従って、このサンプルは不確定リスクとして見なされた。

【0164】

実施例15：成人および小児A Rと非A Rの間で差次的に発現される遺伝子

236の成人および小児A Rサンプルと非A Rサンプルの間で差次的に発現される遺伝子

成人と小児の両方でA R患者を非A R患者から識別し、A Rの非侵襲的検出のためのロバストなバイオマーカーを提供した遺伝子を同定するために、マイクロ流体ハイスループットq P C Rプラットフォームフリーダイム(B i o m a r k , F l u i d i g m I n c .)にて、成人および小児患者由来の236の血液サンプルで、上記で定義された43遺伝子(42遺伝子とハウスキーピング遺伝子であるリボソームR N A 1 8 S)の発現の同時測定を行った。教師なしのP C AおよびA N O V Aによって評価した場合、患者が同種移植を受けた特定の移植センター(「センター」)は、拒絶状態に関する患者の分離を説明する最大の変数であることが判明した。教師なしのP C Aによれば、移植センターにより分離したサンプルは、未補正のデータセットにおいて表現型(A Rと非A R)により推論される遺伝子発現の違いを無効にすることが判明した。除去されるランダムカテゴリー因子として移植センター、R N A ソース、およびq P C Rチップ、残すカテゴリー因子として表現型(A R、非A R)が含まれた混合A N O V Aモデルを用いてデータを補正

40

50

すると、移植センターによりサンプルを分離することなく、むしろ表現型によりサンプルを分離した遺伝子発現が得られた。この正規化セットの分析により、これらの遺伝子の大規模サブセットがARと非ARの間で差次的に発現されたことが示された（スチューデントのT検定： $n = 32$ 、 $p < 0.05$ ）。

【0165】

267の成人および小児ARサンプルと非ARサンプルの間で差次的に発現される遺伝子合計31の遺伝子が、267の成人および小児両方のARと非ARの間で差次的に発現された（コホート1、 $n = 143$ ；コホート2、 $n = 124$ ； $FDR < 5\%$ 、ボンフェローニ事後検定を伴うANOVA）。興味深いことに、8/10遺伝子の小児パネルは、成人サンプルにおいて有意に異なっていた（ $p < 0.05$ ）。

【0166】

実施例16：10遺伝子を用いたARサンプルと非ARサンプルの分類

サポートベクターマシンのによる10遺伝子を用いた成人ARサンプルと非ARサンプルの分類

異なる採取センター、性別、血液RNAサンプルソース、およびレシピエント年齢でのAR分類に関するこれらの遺伝子セットの潜在的有効性を評価するために、PartekおよびRにおいて利用可能な2つの異なる分類アプローチを利用した。Partekでは、SVMアルゴリズム（コストパラメーター $c = 701$ 、核関数 = ラジアル基底関数 $\exp(-\gamma ||x-y||^2)$ ここで $\gamma = 3$ ）、Rでは、Elastic-Netを用いるペナルティ付きロジスティック回帰を、サンプルのARまたは非ARへの分類に適用した。両分類アルゴリズムとも二元分類子であり、SVMは2つのクラスを分離する境界を最大化するように設計され、従って、トレーニングされたモデルは、所見のデータに対して、データを過剰適合することなく、十分な一般化を行う。しかしながら、SVMは非確率的分類子であり、個々の予測精度スコアを提供しない。ロジスティック回帰は、各サンプルに対して予測的確率スコアを提供する。これらの方法は、生検読み取りと臨床機能が一致した143の成人AR（ $n = 47$ ）および非AR（ $n = 96$ ）サンプルのテストセットにおいて、小児および若年成人血液サンプルでのAR検出に関して従前にバリデートされた（AR検出に関して92%精度、91%感度、および94%特異度）、従前に公表された10遺伝子小児モデル（CFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NAMPT、NKTR、PSEN1、RNF130、およびRYBP）を用いて、センターに関して正規化されたデータに適用した。上記のSVMを用い、同じ10遺伝子は、143サンプルの成人データセットに適用した場合に、87%精度、70%感度、および96%特異度で成人におけるARを検出した。

【0167】

ロジスティック回帰による10遺伝子を用いたAR生検の最大6か月前および1か月後の小児サンプルにおけるARの検出

生検により証明されたARを有する40名の患者について、AR生検の最大7か月前（ $n = 27$ ）および6か月後まで（ $n = 30$ ）に採取した小児同種移植レシピエントからの一連のサンプルが利用可能であった。小児5遺伝子発現モデル（DUSP1、NAMPT、PSEN1、MAPK9、およびNKTR）により、ARの生検の最大6か月前（平均スコア0~3か月 = 88%；平均スコア3~6か月 = 58%）および1か月後まで（平均スコア = 63%）、高いAR予測スコアが明らかになった。40の適合ARサンプルの平均スコアは91%であった。AR生検後1か月以降に採取されたサンプルでは、ARの平均予測スコアは3か月で42%、>3か月後で48%であった（図3B）。

【0168】

実施例17：15遺伝子を用いたARサンプルと非ARサンプルの分類

ペナルティ付きロジスティック回帰による15遺伝子を用いた143の成人ARサンプルと非ARサンプルの分類

精度および感度を改善するために、さらに32遺伝子の成人検定セットに対する影響を、年齢非依存的AR予測アルゴリズムを開発するために含めることができるさらなる遺伝

10

20

30

40

50

子の選択のためのペナルティ付きロジスティック回帰を用いて調べた。結果として、32 遺伝子から15のさらなる遺伝子が選択された (CEACAM4、SLC25A37、RARAA、CXCL10、GZB、IL2RB、RHEB、C1orf38、EPOR、GZMK、ABTB1、NFE2、FOXP3、MPP1、およびMAP2K3)。これらのさらなる遺伝子の使用により、ペナルティ付きロジスティック回帰による成人データセットにおけるARの予測に改善がもたらされた (92%精度、86%感度、および94%特異度)。5サンプル (非ARが2、ARが3) だけが不正確な分類であった。このペナルティ付きロジスティック回帰モデルにおけるAR予測の精度は50%であり、これらのサンプルの分類が、サンプルをARと表示するために > 50%の確率スコアで、また、サンプルを非ARと表示するために < 50%の確率スコアで達成されたことを示した (図1A)。

10

【0169】

ペナルティ付きロジスティック回帰による15遺伝子を用いた49の成人ARサンプルと非ARサンプルの分類

49サンプルの独立セットでの成人15遺伝子モデルの成績を検定した。サンプルは、採血時に生検により確認された病状レポートを有する8名のAR患者と6名の非AR患者を含んだ。残りの20サンプルは、血液サンプル採取時に適合した生検を持たなかった患者 (N.A., n = 22)、またはBK腎症患者 (BK, n = 2) に加え、急性尿細管腎炎 (ATN, n = 3)、急性薬物毒性 (CNIT, n = 4) を含む、もしくは生検時に慢性同種移植損傷を示していた (IF/TA, n = 4)、他の形態の移植片機能不全を受けていた患者 (n = 13) から採取された。未知の表現型の患者に由来するこれら22サンプルに、サンプル採取の前または後に生検により証明された拒絶を持っていたものは無かった。遺伝子発現を用いたところ、全ての非ARサンプルが正確に非ARと予測され、8のうち5のARサンプルが正確にARと分類された。ARと他起源の移植片機能不全の間の予測スコアは、ARの方が有意に高かった ($p = 0.0162$)。未知の表現型 (N.A.) を有する患者由来の全てのサンプルは非ARと予測された (図1B)。

20

【0170】

ペナルティ付きロジスティック回帰による15遺伝子を用いた成人における抗体媒介性急性拒絶および細胞媒介性急性拒絶の均等検出

大部分のARサンプルは、細胞性拒絶と体液性拒絶の混合型である。ドナー特異的抗体 (donor specific antibody) (DSA) データは、全ての場合で生検時には得られなかった。抗体媒介性拒絶 (antibody-mediated rejection) (AMR、C4D陽性生検染色、DSA+) のみを示した患者5名のサブセットを用い、予測スコアを、明らかな細胞媒介性拒絶 (ACR、C4d-、DSA-; n = 33) のみを有する患者と比較した。純粋な抗体媒介性拒絶エピソードおよび細胞媒介性拒絶エピソードの数は比較的少なかったものの、2つのARサブグループにおける平均AR予測スコアの比較により、このモデルはAMRおよびACRを、高い予測確率 (平均スコアAMR = $82.9\% \pm 0.16$; 平均スコアACR = $89.5\% \pm 0.12$; $p = 0.413$) で均等に検出したことが明らかになった (図2)。図2Aは、143のARおよび非AR成人患者におけるARの予測確率を示す。図2Bは、49の独立した患者 (AR8、非AR6、移植片不全13、および未知22) におけるARの予測確率を示す。

30

40

【0171】

ペナルティ付きロジスティック回帰による15遺伝子を用いた成人腎臓レシピエント由来の血液におけるAR生検3か月前および1か月後のARの予測

生検により証明されたARを有する患者のサブセット (n = 59) に関して、AR生検の最大2年前 (n = 23) および1.5年後 (n = 19) に採取した一連の血液サンプルが利用可能であった。遺伝子発現により、ARは、この成人集団においてARに関する生検の最大3か月前および1か月後まで示された (平均AR確率0~3か月前 = 43%; 平均AR確率0~1か月後 = 50%)。AR生検の3か月前または1か月後を超えて採取された血液サンプルでは、本遺伝子発現モデルを用いてARを検出する確率がそれぞれ24

50

%および24%の確率に落ちた。17の適合ARサンプルの平均スコアは82%であった(図3A)。

【0172】

実施例18: サポートベクターマシンによる17遺伝子を用いたARと非ARの分類

レシピエント年齢に依存せずにARを検出するために、93の末梢血サンプル(AR22、非AR71)からなる小児および若年成人患者の独立したサブセットからのqPCRデータ(Li, L. et al. Am. J. Transplant. 2012, 12, 2710-2718)を、成人移植レシピエントからの143のサンプル(AR47、非AR96)と合わせた。Shrinking Centroidsを用い、患者をARまたは非ARに分類した17遺伝子のセットを、SVMアルゴリズムを分類のためのコストパラメーター=701、カーネル=rbf、および $\gamma=3$ とともに用いて同定した。この17遺伝子モデルは、SVMを用い、236の小児、若年成人、および成人患者の複合データセットで、94%精度、88%感度、および95%特異度でARを検出した。この17遺伝子セットは10の小児遺伝子(CFLAR、DUSP1、ITGAX、RNF130、PSEN1、NKTR、RYBP、NAMPT、MAPK9、およびIFNGR1)の組合せ、新たに定義された15の成人遺伝子のうち6(CEACAM4、RHEB、GZMK、RARA、SLC25A37、およびEPOR)、ならびにレチノイドX受容体(RXRA)を使用した。これらの17遺伝子を用いたところ、69のARサンプルのうち8サンプルだけが不正確に非ARと予測され、169の非ARサンプルのうち8サンプルだけが不正確にARと予測された。明らかに、成人特異的遺伝子と小児特異的遺伝子の組合せは、高い精度、感度、および特異度を持った年齢非依存性のARの予測の開発に必要である。

10

20

【0173】

実施例19: 等しい事前確率の部分最小二乗判別分析による17遺伝子を用いたARと非ARの分類

等しい事前確率の部分最小二乗判別分析による17遺伝子を用いた143の成人ARサンプルと非ARサンプルの分類

腎臓AR予測アッセイを定義するための最終的な17遺伝子は、小児10遺伝子パネル(DUSP1、CFLAR、ITGAX、NAMPT、MAPK9、RNF130、IFNGR1、PSEN1、RYBP、NKTR)と成人拒絶に有益なさらなる7遺伝子(SLC25A37、CEACAM4、RARA、RXRA、EPOR、GZMK、RHEB)からなった(図12); これらの17遺伝子は、レシピエント異年齢間でARを識別するために最適な成績を示し: 143の成人サンプルのトレーニングセット(コホート1)では、17遺伝子は39/47のサンプルを正確にAR予測し、87/96のサンプルを正確に非ARと予測し、等しい事前確率の部分最小二乗判別分析において感度83%、および特異度91%が得られた(p LSDA; 図6A~B)。平均予測AR確率は、各センターにおいてARと非ARを比較すると高い有意性で異なっていた(CPMC: $p < 0.0001$; エモリー: $p = 0.002$; UPMC: $p < 0.0001$; UCLA: $p < 0.0001$) (図6A)。17遺伝子の受信者動作特性(ROC)曲線下の総面積は、 p LSDAによればAUC = 0.94 (95%CI 0.91~0.98; $p < 0.0001$)であった(図6B)。

30

40

【0174】

等しい事前確率の部分最小二乗判別分析による17遺伝子を用いた独立した124の成人および小児レシピエントの分類

成人および小児の両レシピエントにおいて非AR表現型からARを識別するための17遺伝子腎臓AR予測アッセイモデルを独立にバリデートするために、本発明者らは、フリーダムプラットフォームでも行った、成人($n = 59$)および小児($n = 65$)を合わせた124の独立したサンプルのセット(コホート2; 後向きバリデーション)におけるその成績を検定した。17遺伝子腎臓AR予測アッセイモデルは、21/23のサンプルをARとして正確に予測し、BKウイルス腎炎患者4名を含む100/101のサンプルを非ARと正確に予測し(図7A)、アッセイ感度91.3%および特異度99.01

50

%が得られた。2つの誤分類ARサンプルのうちの1つは、重度の慢性傷害を有し（IF/TAGグレードIII）、拒絶時に生検サンプルにおけるグローバルな退縮は>33%であった。トレーニングセット（コホート1）で見られるように、ARの平均予測確率は、バリデーションセット（コホート2）のARサンプル（80.55%）と非ARサンプル（9.2%）の間でも有意に異なっており（ $p < 0.0001$ ；図7B）；BKV群のARの平均予測確率は12.76%と低かった。124サンプルのROC分析の結果はAUC = 0.9479（95%CI 0.88~1.0）であった（図7C）。各サンプル採取場所における17遺伝子腎臓AR予測アッセイモデルの成績を評価するために、本発明者らは、コホート1およびコホート2からのエモリー（ $n = 42$ ）、UPMC（ $n = 81$ ）、UCLA（ $n = 44$ ）およびCPMC（ $n = 35$ ）における予測のためのROC AUCを計算した。移植センターによるアッセイの成績は、個々のROC AUCは4か所全てのセンターについて>0.8であることを示した（図13A~D）。

10

【0175】

等しい事前確率の部分最小二乗判別分析による17遺伝子を用いた抗体媒介性急性拒絶と細胞媒介性急性拒絶の均等検出

フリーダタイムプラットフォームで分析された大部分のARサンプルは、いくらかの細胞性拒絶と体液性拒絶の混合状態または関連の慢性変化を示した。固定された17遺伝子モデルによって評定した場合、明らかな抗体媒介拒絶（AMR、C4D+生検染色、DSA+）のみを有する19名の患者と明らかな細胞媒介性拒絶（ACR、C4d-およびDSA-、およびBanff t-およびiスコア>1）を有する51名の患者の間でAR予測スコアに違いは見られなかった（pISA； $p = 0.9906$ ；平均ACR = 80.84% ± 4.4；平均AMR = 80.75% ± 6.6；図14A）。

20

【0176】

等しい事前確率の部分最小二乗判別分析による17遺伝子を用いたARの分類は移植後時間に依存しない

移植後拒絶の時間が17遺伝子の予測精度に影響をしたかどうかを評価するために、移植後0~6か月、6~12か月の間、および>1年に採取されたARサンプルおよび非ARサンプルにおける予測AR確率を評価したところ、移植後の時間による影響を認めなかった（図14B）。

【0177】

17遺伝子は等しい事前確率の部分最小二乗判別分析により191サンプルにおいて臨床移植片機能不全前に生検により確認されるARを予測する

17遺伝子腎臓AR予測アッセイモデルの予測性を評価するために、生検適合ARエピソード（ $n = 74$ ）の前（0.2~6.8か月、 $n = 65$ ）または後（0.2~7か月； $n = 52$ ）のいずれかに採取した191の血液サンプル（コホート3；前向きバリデーション）を分析した。AR生検の0~3か月前の血液サンプルを有する患者（ $n = 35$ ）のうち、安定な移植片機能の時点で、62.9%（22/35）が極めて高いAR予測スコア（96.4% ± 0.8）を有しており（図8）、経過観察時に安定な移植片機能を有しかつARを示さない患者のスコアよりも有意に高かった（19.4% ± 0.3； $p < 0.0001$ ）。AR処置の0~3か月後に採取された血液サンプルを有する患者（ $n = 31$ ）のうち、51.6%（16/31）はなお高いARスコアを有しており（86% ± 0.17）；15/31サンプルはAR処置の0~3か月後にAR閾値より低いARスコアを示した（6.59% ± 0.13%）。低いAR予測スコアを有する患者のクレアチニンレベルが1.8 ± 0.4 mg/dLであるのに対して、高いAR予測スコアを有する患者の血清クレアチニンレベルは2.04 ± 0.4 mg/dLであったことから、前者はAR処置に奏効した患者n相当すると思われる（図8）。

30

40

【0178】

実施例20：kSASによるARと非ARの分類

標準的QPCRのためのABI viia7 QPCRプラットフォームの選択

フリーダタイムプラットフォームなどのハイスループットQPCRプラットフォームは

50

、診断用バイオマーカーパネルの発見および初期開発のために極めて好適であるが、費用効果がある性能を得るためには大きなサンプルサイズおよび遺伝子数が必要である。従って、変動のあるより少ないサンプル数のための、カスタマイズ可能な形式と費用効果のある性能を有する臨床適用可能なアッセイを開発するために、17遺伝子モデルを44名のAR患者および56名の非AR患者から採取した100サンプルを用い、標準的qPCR (ABI v i i a 7、Life Technologies、フォスターシティ、CA)により分析した。これらの遺伝子セットを臨床分析(スケーラビリティ、コスト、マシンアベイラビリティ、プロトコルの簡素性)に関して最適化するため、下流での発見およびバリデーションのためにABI qPCRプラットフォームを用いた。

【0179】

k S A Sによる10遺伝子を用いた成人および小児のARと非ARの分類

発見のため、k S A S分析は2か所の成人センター(UCSFおよびピッツバーグ)と1か所の小児センター(SNS)に制限した。この分析により、両成人センターおよび小児センターで89%の割合でAR状態を分類することができた7遺伝子モデル(CFLAR、DUSP1、IFNGR1、ITGAX、MAPK9、NKTR、およびRYBP)が得られた。もう1つの3~10遺伝子のモデルサイズは、最終的な7遺伝子のセットよりも全体的な成績が低かった。成人に関して合わせた分類率は、16のARサンプルおよび16の非ARサンプル(感度=88%、特異度=75%)に基づけば81%精度となり、小児セットでは、22のARサンプルおよび155の非ARサンプル(感度=91%、感度=90%)に基づけば、90%精度となった。

【0180】

k S A Sによる17遺伝子を用いたARと非ARの分類

従前に発見された10の小児遺伝子に加え、フリーダイン分析から同定された7つの成人分類子遺伝子(CEACAM4、EPOR、GZMK、RARA、RHEB、RXRA、およびSLC25A37)をABI遺伝子パネルに加えた。これら17遺伝子の配列をゲノムDNA配列として付表Aに示す。これらの遺伝子のほとんど全てが、143の患者フリーダインqPCRデータセットを、k S A Sを用いて再分析した際にも高い値の予後マーカーとして同定された。この大きなABI遺伝子セット分析は、最初は、ARまたは非AR状態が確認された成人患者の血液サンプルに限られた。これら17の小児および成人遺伝子から性能の改良された遺伝子サブセットを同定するために、まず、2か所の

【0181】

これらの遺伝子モデルの独立検証として、このモデルセットを成人および小児患者(それぞれバルセロナおよびメキシコ)の新たなセットならびに1か所のトレーニングセンター(UCSF)からの独立した患者の第2セットに対して検定した。発見分析からのARおよび非AR発現参照(以前のUCSFサンプル)を用いたUCSF患者の分析により、76~86%の範囲のバリデーション率が得られた。全ての新規サンプルからの結果を統合した場合、上位モデル分類率は88%(86%感度、90%特異度)となり、成人サンプルと小児サンプルの間の分類率は同等であった。この上位12遺伝子モデル(CFLAR、PSEN1、CEACAM4、NAMPT、RHEB、GZMK、NKTR、DUSP1、RARA、ITGAX、SLC25A37、およびEPOR)は、小児分類セット(CFLAR、PSEN1、NAMPT、NKTR、およびDUSP1)からの5遺伝子を含み、年齢、人口統計、誘導、維持免疫抑制、併存症、または交絡する移植片病状とは無関係にAR状態を分類する。実験的に予測される相互作用に関して評価したところ、これらの遺伝子の過半数が直接的または間接的に関連していた。

【0182】

10

20

30

40

50

k S A Sによる17遺伝子を用いた100の成人および小児サンプルに関するARリスクスコアの計算

複数のモデルアプローチが各遺伝子セットに対して異なるスコアを与えたことから、各患者の単一の遺伝子モデルを偏らせない信頼スコアを提供するこれらのモデルの組み合わせ能力を評価した。この統合ARリスク分析は各患者についてスコア(-13~13)を生成し、ARのリスクを示す(13=高リスク、-13=極めて低リスク)。「高リスクのAR」の患者のうち、91%(34名中31)が正確にARと分類され、「極めて低リスクのAR」の患者については、92%(38名中35)が正確に非ARと分類された。残りの患者(n=15)は不確定リスクを有すると予測された(図10A)。

【0183】

平均算出ARリスクスコアは、k S A Sを用いれば、ARでは非ARに比べて有意に高かった($p < 0.0001$) (図10B)。

【0184】

計算されたAUCは、明確なk S A S判定(高リスクAR、低リスクAR、n=85)に関して0.93(95%CI 0.86~0.99)であった(図10C)。

【0185】

示されたアッセイの強度は、末梢血サンプルにおいてARを検出する、その高いPPV(92.3%)である。現在、移植に利用可能な唯一の診断検査は、中等度/重度の急性細胞性心臓拒絶(ISHLT 3A)の不在を検出するが、ARの存在の検出に関する性能は低い(PPV=6.8%)(Deng et al., 2006, Am J. Transplant 6:150-160)。同様に、閉塞性冠動脈疾患を評価するための血液遺伝子発現検査(Corus(登録商標)Cad、CARDIODX(登録商標)、パロ・アルト、CA)は、多施設共同バリデーション試験において46%のPPVを示した(Rosenberg et al, 2010, Ann Intern Med 153:425-434)。拒絶の時点でARを検出するための高感度のアッセイ(現行のゴールドスタンダードにより診断)に加え、このアッセイはまた、12症例で準臨床拒絶を検出し、移植片機能不全および組織学的ARの最大3か月前に採取されたサンプルの>60%において臨床ARを予測し;準臨床および臨床ARは慢性拒絶および移植片欠損の前兆であるので、これは拒絶検査の重要な能力である(Nasesens et al., 2012, Am J Transplant 12: 2730-2743)。適応同種免疫応答を評価するための、循環中のドナー特異的抗体または記憶エフェクターT細胞のいずれかを評価する現行の免疫監視ツールは、ARの潜在リスクを予測するためのそれらの有用性を示したが(Loupy et al, 2013, N Engl J Med 369: 1215-1216;およびBestard et al, 2013 Kidney Int 84: 1226-1236)、それらの検出は必ずしも進行中の免疫介在性同種移植損傷に置き換えられず、さらに、これらのエフェクター機構は生検により証明されるARの前またはその時点で常に検出されるわけではない。さらに、ほとんどのセンターは目下のところ準臨床ARを検出する手段としてプロトコール生検を行っておらず、従って、これらは多くが検出されないままである。本明細書で提供されるアッセイによる慣例の移植後監視はARを予測し、適時の介入で組織損傷を制限し、コスト集約的な透析へ戻る患者の数を最小にすることにより医療システムへの金融負担を低減することができる。

【0186】

実施例21: ARと非AR分類のための17遺伝子の生物学

経験的に予測される相互作用に関して評価した場合、過半数の遺伝子は共通の分子経路、特に、アポトーシスの調節、免疫表現型および細胞表面により直接的または間接的に互いに関連付けられた(図15a~15c)。小児ARに関して末梢バイオマーカーとして従前に評価された、大部分が単球系譜の末梢血細胞で発現されることが知られている10遺伝子に加え、さらなる末梢7AR遺伝子のうち6遺伝子はまた、末梢循環中の活性化単球(RXRA、RARA、CEACAM4)、内皮細胞(EPOR、SLC25A37)およびT細胞(GZMK)によっても発現された。17のうち11の遺伝子は細胞死、および細胞生存ネットワークに共通の役割を果たしていた(フィッシャーの正確確率検定、 $p < 0.05$; IPA; 図15c)。

10

20

30

40

50

【 0 1 8 7 】

実施例 2 1 : 一臓器抜き分析 (leave-one-organ-out analysis) を用いた共通拒絶モジュール (common rejection module) (CRM) の同定

共通拒絶モジュールは、腎臓、肺、心臓、および肝臓移植患者由来の 2 3 6 の独立した生検サンプルからの全ゲノム発現データを分析することにより同定した。各データセットは g c R M A 正規化を行った (Irizarry, E. et al. Nucleic Acids Res. 2003, 31, e15 参照)。移植データベースを、サイズ効果を組み合わせる、また、20% の F D R で 1 0 2 の遺伝子 (表 3 に列挙) を同定する p 値を組み合わせるメタ分析法により分析した。異なる臓器の反復組合せが得られる、一度に一臓器を抜き取る反復をそれぞれメタ分析により分析したところ、全ての臓器で過剰発現される、B A S P 1、C D 6、C D 7、C X C L 1 0、C X C L 9、I N P P 5 D、I S G 2 0、L C K、N K G 7、P S M B 9、R U N X 3、および T A P 1 を含んでなる 1 2 遺伝子が明らかになった (図 1 6)。

10

【 0 1 8 8 】

付表 C : L i n e a g e P r o f i l e r I t e r a t e ソースコード

```
### Based on code from AltAnalyze's LineageProfiler (http://altanalyze.org)
#Author Nathan Salomonis - nsalomonis@gmail.com
```

```
#Permission is hereby granted, free of charge, to any person obtaining a copy
#of this software and associated documentation files (the "Software"), to deal
#in the Software without restriction, including without limitation the rights
#to use, copy, modify, merge, publish, distribute, sublicense, and/or sell
#copies of the Software, and to permit persons to whom the Software is furnished
#to do so, subject to the following conditions:
```

20

```
#THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND,
EXPRESS OR IMPLIED,
#INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY,
FITNESS FOR A
#PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE
AUTHORS OR COPYRIGHT
#HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY,
WHETHER IN AN ACTION
#OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN
CONNECTION WITH THE
#SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.
```

30

【 0 1 8 9 】

```
"""
```

このスクリプトは、1 以上の遺伝子モデルが与えられた場合に、2 つの参照のうち 1 つと比較してサンプルタイプを同定するために L i n e a g e P r o f i l e r アルゴリズム (相関に基づく分類法) を反復する。主関数は r u n L i n e a g e P r o f i l e r である。

40

【 0 1 9 0 】

このプログラムは以下のアクションを実行する :

- 1) 3 列 (I D、生物学的群 1、群 2) およびヘッダー行 (生物学的群の名称) を含むタブ区切り参照発現ファイルをインポートする。
- 2) 遺伝子 I D (列 1)、サンプル名 (行 1) および正規化発現値 (例えば C T 値) を含むタブ区切り発現ファイルをインポートする。
- 3) (任意選択 - 既存モデルのインポート) カンマ区切りの分析用遺伝子モデルを含むタブ区切りファイルをインポートする。

50

- 4) (任意選択 - 新たなモデルの発見) 供給されたモデルサイズ変数 (例えば - - s 7) に対して遺伝子モデルのあり得る全ての組合せを同定する。
- 5) 供給されたまたは同定された遺伝子モデルを反復して、新規なまたは既知のサンプルタイプに関して予測を得る。
- 6) 全ての分析モデルに関する予測結果をフォルダSampleClassificationにエクスポートする。
- 7) (任意選択) サイズ - - s のあり得る全てのモデル組合せに関する上位 20 スコアおよびモデルを印刷する。

```
"""
```

10

```
import sys, string
import math
import os.path
import copy
import time
import getopt
try: import scipy
except Exception: pass
```

```
try: import unique ### Not required (used in AltAnalyze)
except Exception: None
try: import export ### Not required (used in AltAnalyze)
except Exception: None
```

20

```
#import salstat_stats; reload(salstat_stats)
try:
    from scipy import stats
    use_scipy = True
except Exception:
    use_scipy = False ### scipy is not required but is used as a faster implementation of Fisher
Exact Test when present
```

30

```
def filepath(filename):
    try: fn = unique.filepath(filename)
    except Exception: fn = filename
    return fn
```

```
def exportFile(filename):
    try: export_data = export.ExportFile(filename)
    except Exception: export_data = open(filename,'w')
    return export_data
```

40

```
def makeUnique(item):
    db1={}; list1=[]; k=0
    for i in item:
        try: db1[i]=[]
        except TypeError: db1[tuple(i)]=[]; k=1
    for i in db1:
        if k==0: list1.append(i)
        else: list1.append(list(i))
    list1.sort()
```

50

```
return list1
```

```
def cleanUpLine(line):
```

```
    line = string.replace(line, '\n', '')
    line = string.replace(line, '\c', '')
    data = string.replace(line, '\r', '')
    data = string.replace(data, '""', '')
    return data
```

```
def returnLargeGlobalVars():
```

10

```
    """ Prints all large global variables retained in memory (taking up space)
    all = [var for var in globals() if (var[:2], var[-2:]) != ("__", "__")]
    for var in all:
        try:
            if len(globals()[var])>1:
                print var, len(globals()[var])
        except Exception: null=[]
```

```
def clearObjectsFromMemory(db_to_clear):
```

20

```
    db_keys={ }
    for key in db_to_clear: db_keys[key]=[]
    for key in db_keys:
        try: del db_to_clear[key]
        except Exception:
            try:
                for i in key: del i """ For lists of tuples
            except Exception: del key """ For plain lists
```

```
def int_check(value):
```

30

```
    val_float = float(value)
    val_int = int(value)
    if val_float == val_int:
        integer_check = 'yes'
    if val_float != val_int:
        integer_check = 'no'
    return integer_check
```

```
def IQR(array):
```

40

```
    k1 = 75
    k2 = 25
    array.sort()
    n = len(array)
    value1 = float((n*k1)/100)
    value2 = float((n*k2)/100)
```

```

if int_check(value1) == 'no':
    k1_val = int(value1) + 1
if int_check(value1) == 'yes':
    k1_val = int(value1)
if int_check(value2) == 'no':
    k2_val = int(value2) + 1
if int_check(value2) == 'yes':
    k2_val = int(value2)
try: median_val = scipy.median(array)
except Exception: median_val = Median(array)
upper75th = array[k1_val]
lower25th = array[k2_val]

int_qrt_range = upper75th - lower25th
T1 = lower25th - (1.5 * int_qrt_range)
T2 = upper75th + (1.5 * int_qrt_range)
return lower25th, median_val, upper75th, int_qrt_range, T1, T2

class IQRData:
    def __init__(self, maxz, minz, medz, iq1, iq3):
        self.maxz = maxz; self.minz = minz
        self.medz = medz; self.iq1 = iq1
        self.iq3 = iq3
    def Max(self): return self.maxz
    def Min(self): return self.minz
    def Medium(self): return self.medz
    def IQ1(self): return self.iq1
    def IQ3(self): return self.iq3
    def SummaryValues(self):
        vals =
string.join([str(self.IQ1()), str(self.Min()), str(self.Medium()), str(self.Max()), str(self.IQ3())], '\t')
        return vals

def importGeneModels(geneModels):
    fn = filepath(geneModels); x = 0
    geneModels = []
    for line in open(fn, 'rU').xreadlines():
        genes = cleanUpLine(line)
        genes = string.replace(genes, '"', '')
        genes = string.replace(genes, ';;')
        genes = string.split(genes, ';')
        models = []
        for gene in genes:
            if len(gene) > 0:

```

```

        models.append(gene)
    if len(models)>0:
        geneModels.append(models)
    return geneModels

##### Below code deals is specific to this module #####
def
runLineageProfiler(species,array_type,exp_input,exp_output,codingtype,compendium_platform,
modelSize=None,customMarkers=False,geneModels=False,permute=False,useMulti=False):
    """ This code differs from LineageProfiler.py in that it is able to iterate through the
    LineageProfiler functions with distinct geneModels
    that are either supplied by the user or discovered from all possible combinations. """
    10

    global exp_output_file; exp_output_file = exp_output; global targetPlatform
    global tissues; global sample_headers
    global analysis_type; global coding_type; coding_type = codingtype
    global tissue_to_gene; tissue_to_gene = { }; global platform; global cutoff
    global customMarkerFile; global delim; global keyed_by; global pearson_list
    global Permute; Permute=permute; global useMultiRef; useMultiRef = useMulti
    pearson_list=[]
    #global tissue_specific_db
    20

    customMarkerFile = customMarkers
    if geneModels == False: geneModels = []
    else:
        geneModels = importGeneModels(geneModels)

    if '\n' in exp_input: delim = '\n'
    elif '/' in exp_input: delim = '/'
    else: delim = "/"
    30

    print "\nRunning LineageProfiler analysis on',string.split(exp_input,delim)[-1][:4]

    global correlate_by_order; correlate_by_order = 'no'
    global rho_threshold; rho_threshold = -1
    global correlate_to_tissue_specific; correlate_to_tissue_specific = 'no'
    platform = array_type
    cutoff = 0.01
    global value_type

    if 'stats.' in exp_input:
        value_type = 'calls'
    else:
        value_type = 'expression'
    40

```

```

tissue_specific_db={ }; sample_headers=[]; tissues=[]
if len(array_type)==2:
    ### When a user-supplied expression is provided (no ExpressionOutput files provided -
importGeneIDTranslations)
    vendor, array_type = array_type
    platform = array_type
    else: vendor = 'Not needed'

if 'RawSplice' in exp_input or 'FullDatasets' in exp_input or coding_type == 'AltExon':
    analysis_type = 'AltExon'
    if platform != compendium_platform: ### If the input IDs are not Affymetrix Exon 1.0 ST
probesets, then translate to the appropriate system
        translate_to_genearray = 'no'
        targetPlatform = compendium_platform
        translation_db = importExonIDTranslations(array_type,species,translate_to_genearray)
        keyed_by = 'translation'
    else: translation_db=[]; keyed_by = 'primaryID'; targetPlatform = compendium_platform
elif array_type == "3'array" or array_type == 'AltMouse':
    ### Get arrayID to Ensembl associations
    if vendor != 'Not needed':
        ### When no ExpressionOutput files provided (user supplied matrix)
        translation_db = importVendorToEnsemblTranslations(species,vendor,exp_input)
    else:
        translation_db = importGeneIDTranslations(exp_output)
        keyed_by = 'translation'
        targetPlatform = compendium_platform
        analysis_type = 'geneLevel'
else:
    translation_db=[]; keyed_by = 'primaryID'; targetPlatform = compendium_platform;
analysis_type = 'geneLevel'

targetPlatform = compendium_platform ### Overrides above
try: importTissueSpecificProfiles(species,tissue_specific_db)
except Exception:
    try:
        try:
            targetPlatform = 'exon'
            importTissueSpecificProfiles(species,tissue_specific_db)
        except Exception:
            try:
                targetPlatform = 'gene'
                importTissueSpecificProfiles(species,tissue_specific_db)
            except Exception:

```

```

        targetPlatform = "3'array"
        importTissueSpecificProfiles(species,tissue_specific_db)
    except Exception:
        print 'No compatible compendiums present...'
        print e
        forceError

all_marker_genes=[]
for gene in tissue_specific_db:
    all_marker_genes.append(gene)
10

if len(geneModels)>0:
    allPossibleClassifiers = geneModels
elif modelSize == None or modelSize == 'optimize':
    allPossibleClassifiers = [all_marker_genes]
else:
    ### A specific model size has been specified (e.g., find all 10-gene models)
    allPossibleClassifiers = getRandomSets(all_marker_genes,modelSize)

num=1
all_models=[]
20
if len(allPossibleClassifiers)<16:
    print 'Using:'
    for model in allPossibleClassifiers:
        print 'model',num,model
        num+=1
        all_models+=model

#all_models = unique.unique(all_models)
#print len(all_models);sys.exit()
30

### This is the main analysis function
print 'Number of references to compare to:',len(tissues)
if len(tissues)<16:
    print tissues

if modelSize != 'optimize':
    hit_list, hits, fails, prognostic_class_db,sample_diff_z, evaluate_size, prognostic_class1_db,
    prognostic_class2_db = iterateLineageProfiler(exp_input, tissue_specific_db,
    allPossibleClassifiers,translation_db,compendium_platform,modelSize)
else:
40
    summary_hit_list=[]
    evaluate_size = len(allPossibleClassifiers[0])

```

```

hit_list, hits, fails, prognostic_class_db, sample_diff_z, evaluate_size, prognostic_class1_db,
prognostic_class2_db = iterateLineageProfiler(exp_input, tissue_specific_db,
allPossibleClassifiers, translation_db, compendium_platform, None)
while evaluate_size > 4:
    hit_list.sort()
    top_model = hit_list[-1][-1]
    top_model_score = hit_list[-1][0]
    """
    try: ### Used for evaluation only - gives the same top models
        second_model = hit_list[-2][-1]
        second_model_score = hit_list[-2][0]
        if second_model_score == top_model_score:
            top_model = second_model_score ### Try this
            print 'selecting secondary'
    except Exception: None
    """
    allPossibleClassifiers = [hit_list[-1][-1]]

    hit_list, hits, fails, prognostic_class_db, sample_diff_z, evaluate_size,
prognostic_class1_db, prognostic_class2_db = iterateLineageProfiler(exp_input,
tissue_specific_db, allPossibleClassifiers, translation_db, compendium_platform, modelSize)
    summary_hit_list += hit_list
    hit_list = summary_hit_list

    exp_output_file = string.replace(exp_output_file, '\\', '/')
    root_dir = string.join(string.split(exp_output_file, '/')[:-1], '/') + '/'
    dataset_name = string.replace(string.split(exp_input, '/')[-1][:-4], 'exp.', '')
    output_classification_file = root_dir + 'SampleClassification/' + dataset_name + '-
SampleClassification.txt'
    try: os.mkdir(root_dir + 'SampleClassification')
    except Exception: None
    export_summary = exportFile(output_classification_file)
    models = []
    for i in allPossibleClassifiers:
        i = string.replace(str(i), '"', '')[1:-1]
        models.append(i)

    class_headers = map(lambda x: x + ' Predicted Hits', tissues)
    headers = string.join(['Samples'] + class_headers + ['Composite Prognostic Score', 'Median Z-
score Difference', 'Prognostic Risk'] + models, '\t') + '\n'
    export_summary.write(headers)
    sorted_results = [] ### sort the results
    for sample in prognostic_class_db:
        if len(tissues) == 2:

```

```

class1_score = prognostic_class1_db[sample]
class2_score = prognostic_class2_db[sample]
zscore_distribution = map(str,sample_diff_z[sample])
dist_list=[]
for i in zscore_distribution:
    try: dist_list.append(float(i))
    except Exception: None ### Occurs for 'NA's
try: median_score = scipy.median(dist_list)
except Exception: median_score = Median(dist_list)

```

10

```

class_db = prognostic_class_db[sample]
class_scores=[]; class_scores_str=[]; class_scores_refs=[]
for tissue in tissues:
    class_scores_str.append(str(class_db[tissue]))
    class_scores.append(class_db[tissue])
    class_scores_refs.append((class_db[tissue],tissue))
overall_prog_score = str(max(class_scores)-min(class_scores))
if len(tissues)==2:
    class_scores_str = [str(class1_score),str(class2_score)] ### range of positive and negative
scores for a two-class test
    if class2_score == 0:
        call = 'High Risk '+ tissues[0]
    elif class1_score == 0:
        call = 'Low Risk '+ tissues[0]
    else:
        call = 'Intermediate Risk '+ tissues[0]
    overall_prog_score = str(class1_score-class2_score)
else:
    class_scores_refs.sort()
    call=class_scores_refs[-1][1]
    if ':' in call:
        call = string.split(call,':')[0]
    if 'non' in call:
        overall_prog_score = str(float(overall_prog_score)*-1)
        median_score = median_score*-1
values = [sample]+class_scores_str+[overall_prog_score,str(median_score),call]
values = string.join(values+zscore_distribution,'\t')+'\n'
sorted_results.append([float(overall_prog_score),median_score,values])
sample_diff_z[sample] = dist_list

```

20

30

```

sorted_results.sort()
sorted_results.reverse()
for i in sorted_results:
    export_summary.write(i[-1])

```

40

```

export_summary.close()
print 'Results file written to:',root_dir+'SampleClassification/'+dataset_name+'-
SampleClassification.txt,'\n'

hit_list.sort(); hit_list.reverse()
top_hit_list=[]
top_hit_db={}
hits_db={}; fails_db={}

avg_pearson_rho = Average(pearson_list) 10

for i in sample_diff_z:
    zscore_distribution = sample_diff_z[i]
    maxx = max(zscore_distribution); minz = min(zscore_distribution)
    sample_diff_z[i] = string.join(map(str,zscore_distribution),'\t')
    try:
        lower25th,medz,upper75th,int_qrt_range,T1,T2 = IQR(zscore_distribution)
        if float(maxx)>float(T2): maxx = T2
        if float(minz) < float(T1): minz = T1
        #iqr = IQRData(maxx,minz,medz,lower25th,upper75th) 20
        #sample_diff_z[i] = iqr
    except Exception:
        pass
for i in hits:
    try: hits_db[i]+=1
    except Exception: hits_db[i]=1
for i in fails:
    try: fails_db[i]+=1
    except Exception: fails_db[i]=1
for i in fails_db:
    if i not in hits: 30
        try:
            #print i+'\t'+'\t'+str(fails_db[i])+'\t'+ sample_diff_z[i]
            None
        except Exception:
            #print i
            None

if modelSize != False:
    print 'Returning all model overall scores'
    hits=[] 40
    for i in hits_db:
        hits.append([hits_db[i],i])

```

```

hits.sort()
hits.reverse()
for i in hits:
    if i[1] in fails_db: fail = fails_db[i[1]]
    else: fail = 0
    try:
        #print i[1]+'\\t'+str(i[0]+'\\t'+str(fail)+'\\t'+sample_diff_z[i[1]])
        None
    except Exception:
        #print i[1]
        None

```

10

```

for i in hit_list:
    if i[0]>0:
        top_hit_list.append(i[-1])
        top_hit_db[tuple(i[-1])]=i[0]

if len(geneModels) > 0:
    for i in hit_list:
        print i[:5],i[-1],i[-2] ### print all
else:
    """
    print 'Returning all over 90'
    for i in hit_list:
        if i[0]>85:
            print i[:5],i[-1],i[-2] ### print all

    sys.exit()"""
    #print 'Top hits'
    for i in hit_list[:500]:
        print i[:5],i[-1],i[-2]
    """

```

20

```

try:
    if hit_list[0][0] == hit_list[20][0]:
        for i in hit_list[20:]:
            if hit_list[0][0] == i[0]:
                print i[:5],i[-1],i[-2]
            else: sys.exit()
except Exception: None ### Occurs if less than 20 entries here
"""

```

30

```

print 'Average Pearson correlation coefficient:', avg_pearson_rho
if avg_pearson_rho<0.9:
    print '\\n\\nWARNING!!!!!!!!!!'

```

40

```

    print '\tThe average Pearson correlation coefficient for all example models is less than 0.9.'
    print '\tYour data may not be comparable to the provided reference (quality control may be
needed).\n\n'
    else:
        print 'No unusual warning.\n'
    return top_hit_db

```

```
def
```

```
iterateLineageProfiler(exp_input,tissue_specific_db,allPossibleClassifiers,translation_db,compen
dium_platform,modelSize):
```

10

```

    hit_list=[]
    ### Iterate through LineageProfiler for all gene models (allPossibleClassifiers)
    times = 1; k=1000; l=1000; hits=[]; fails=[]; f=0; s=0; sample_diff_z={ };
    prognostic_class1_db={ }; prognostic_class2_db={ }
    prognostic_class_db={ }
    begin_time = time.time()

    evaluate_size=len(allPossibleClassifiers[0]) ### Number of reference markers to evaluate

    if modelSize=='optimize':
        evaluate_size -= 1
        allPossibleClassifiers = getRandomSets(allPossibleClassifiers[0],evaluate_size)

    for classifiers in allPossibleClassifiers:
        tissue_to_gene={ }; expression_subset=[]; sample_headers=[]; classifier_specific_db={ }
        for gene in classifiers:
            try: classifier_specific_db[gene] = tissue_specific_db[gene]
            except Exception: None
        expression_subset, sampleHeaders =
importGeneExpressionValues(exp_input,classifier_specific_db,translation_db,expression_subset)
    ### If the incorrect gene system was indicated re-run with generic parameters
    if len(expression_subset)==0:
        translation_db=[]; keyed_by = 'primaryID'; targetPlatform = compendium_platform;
analysis_type = 'geneLevel'
        tissue_specific_db={ }
        importTissueSpecificProfiles(species,tissue_specific_db)
        expression_subset, sampleHeaders =
importGeneExpressionValues(exp_input,tissue_specific_db,translation_db,expression_subset)
    if len(sample_diff_z)==0: ### Do this for the first model examine only
        for h in sampleHeaders:
            sample_diff_z[h]=[] ### Create this before any data is added, since some models will
exclude data for some samples (missing dCT values)
            if len(expression_subset)!=len(classifiers): f+=1
            #if modelSize=='optimize': print len(expression_subset), len(classifiers);sys.exit()

```

30

40

```

if len(expression_subset)==len(classifiers): ### Sometimes a gene or two are missing from
one set
    s+=1
    #print classifiers,'\t',
    zscore_output_dir,tissue_scores =
analyzeTissueSpecificExpressionPatterns(tissue_specific_db,expression_subset)
    #except Exception: print len(classifier_specific_db), classifiers; error
    headers = list(tissue_scores['headers']); del tissue_scores['headers']
    if times == k:
        end_time = time.time()
        print int(end_time-begin_time),'seconds'
        k+=1
    times+=1; index=0; positive=0; positive_score_diff=0
    sample_number = (len(headers)-1)
    population1_denom=0; population1_pos=0; population2_pos=0; population2_denom=0
    diff_positive=[]; diff_negative=[]
    while index < sample_number:
        scores = map(lambda x: tissue_scores[x][index], tissue_scores)
        scores_copy = list(scores); scores_copy.sort()
        diff_z = scores_copy[-1]-scores_copy[-2] ### Diff between the top two scores
        j=0
        for tissue in tissue_scores:
            if scores[j] == max(scores):
                hit_score = 1
            else: hit_score = 0
            if len(tissues)>2:
                if tissue+'.' in headers[index+1] and hit_score==1:
                    positive+=1
            try:
                class_db = prognostic_class_db[headers[index+1]]
                try: class_db[tissue]+=hit_score
                except Exception: class_db[tissue]=hit_score
            except Exception:
                class_db={ }
                class_db[tissue]=hit_score
                prognostic_class_db[headers[index+1]] = class_db
            j+=1
        if len(tissues)==2:
            diff_z = tissue_scores[tissues[0]][index]-tissue_scores[tissues[-1]][index]
            if headers[index+1] not in prognostic_class1_db:
                prognostic_class1_db[headers[index+1]]=0 ### Create a default value for each
sample
            if headers[index+1] not in prognostic_class2_db:

```

```

    prognostic_class2_db[headers[index+1]]=0 ### Create a default value for each
sample
if diff_z>0:
    prognostic_class1_db[headers[index+1]]+=1
if diff_z<0:
    prognostic_class2_db[headers[index+1]]+=1
if diff_z>0 and (tissues[0]+'.' in headers[index+1]):
    positive+=1; positive_score_diff+=abs(diff_z)
    population1_pos+=1; diff_positive.append(abs(diff_z))
    hits.append(headers[index+1]) ### see which are correctly classified 10
elif diff_z<0 and (tissues[-1]+'.' in headers[index+1]):
    positive+=1; positive_score_diff+=abs(diff_z)
    population2_pos+=1; diff_positive.append(abs(diff_z))
    hits.append(headers[index+1]) ### see which are correctly classified
elif diff_z>0 and (tissues[-1]+'.' in headers[index+1]): ### Incorrectly classified
    diff_negative.append(abs(diff_z))
    fails.append(headers[index+1])
elif diff_z<0 and (tissues[0]+'.' in headers[index+1]): ### Incorrectly classified
    #print headers[index+1]
    diff_negative.append(abs(diff_z))
    fails.append(headers[index+1]) 20
if (tissues[0]+'.' in headers[index+1]):
    population1_denom+=1
else:
    population2_denom+=1
sample_diff_z[headers[index+1]].append(diff_z)
index+=1
percent_positive = (float(positive)/float(index))*100
if len(tissues)==2:
    hit_list.append([percent_positive,population1_pos,
population1_denom,population2_pos,population2_denom,[Average(diff_positive),Average(diff_
negative)],positive_score_diff,len(classifiers),classifiers]) 30
else:
    hit_list.append([percent_positive,len(classifiers),classifiers])
for sample in sample_diff_z:
    if len(sample_diff_z[sample]) != (times-1): ### Occurs when there is missing data for
a sample from the analyzed model
        sample_diff_z[sample].append('NA') ### add a null result
return hit_list, hits, fails, prognostic_class_db, sample_diff_z, evaluate_size,
prognostic_class1_db, prognostic_class2_db

def factorial(n):
    ### Code from http://docs.python.org/lib/module-doctest.html

```

```

if not n >= 0:
    raise ValueError("n must be >= 0")
if math.floor(n) != n:
    raise ValueError("n must be exact integer")
if n+1 == n: # catch a value like 1e300
    raise OverflowError("n too large")
result = 1
factor = 2
while factor <= n:
    result *= factor
    factor += 1
return result

```

10

```

def choose(n,x):
    """Equation represents the number of ways in which x objects can be selected from a total of n
objects without regard to order."""
    #(n x) = n!/(x!(n-x)!)
    f = factorial
    result = f(n)/(f(x)*f(n-x))
    return result

```

20

```

def getRandomSets(a,size):
    #a = ['a','b','c','d','e','f','g','h','i','j','k','l','m','n','o','p','q','r','s','t','u','v','w','x','y','z']
    #size = 4

select_set={'ENSG00000140678':'ITGAX','ENSG00000105835':'NAMPT','ENSG00000027697'
:'IFNGR1','ENSG00000120129':'DUSP1','ENSG00000003402':'CFLAR','ENSG00000113269':'R
NF130'}
select_set={}

select_set2={'ENSG00000163602':'RYBP'}
negative_select = {'ENSG00000105352':'CEACAM4'}
negative_select={}

import random
possible_sets = choose(len(a),size)
print 'Possible',size,'gene combinations to test',possible_sets
permute_ls = []; done = 0; permute_db={}
while done == 0:
    b = list(tuple(a)); random.shuffle(b)
    bx_set={}
    i = 0
    while i < len(b):
        try:

```

30

40

```

        bx = b[i:i+size]; bx.sort()
        if len(bx)==size: permute_db[tuple(bx)]=None
        else: break
    except Exception: break
    i+=1
if len(permute_db) == possible_sets:
    done=1; break
for i in permute_db:
    add=0; required=0; exclude=0
    for l in i:
        if len(select_set)>0:
            if l in select_set: add+=1
            #if l in select_set2: required+=1
            #if l in negative_select: exclude+=1
        else: add = 1000
    if add>2 and exclude==0:# and required==1:
        permute_ls.append(i)
#print len(permute_ls)
return permute_ls

```

10

```

def importVendorToEnsemblTranslations(species,vendor,exp_input):
    translation_db={ }
    """
    ### Faster method but possibly not as good
    uid_db = simpleUIDImport(exp_input)
    import gene_associations
    ### Use the same annotation method that is used to create the ExpressionOutput annotations
    array_to_ens = gene_associations.filterGeneToUID(species,'Ensembl',vendor,associated_IDs)
    for arrayid in array_to_ens:
        ensembl_list = array_to_ens[arrayid]
        try: translation_db[arrayid] = ensembl_list[0] ### This first Ensembl is ranked as the most
        likely valid based on various metrics in getArrayAnnotationsFromGOElite
        except Exception: None
    """
    translation_db={ }
    import BuildAffymetrixAssociations

    ### Use the same annotation method that is used to create the ExpressionOutput annotations
    use_go = 'yes'
    conventional_array_db={ }
    conventional_array_db =
    BuildAffymetrixAssociations.getArrayAnnotationsFromGOElite(conventional_array_db,species,
    vendor,use_go)
    for arrayid in conventional_array_db:

```

20

30

40

```

    ca = conventional_array_db[arrayid]
    ens = ca.Ensembl()
    try: translation_db[arrayid] = ens[0] ### This first Ensembl is ranked as the most likely
valid based on various metrics in getArrayAnnotationsFromGOElite
    except Exception: None

    return translation_db

def importTissueSpecificProfiles(species,tissue_specific_db):
    if analysis_type == 'AltExon':
        filename = 'AltDatabase/ensembl/'+species+'/'+species+'_'+targetPlatform+'_tissue-
specific_AltExon_protein_coding.txt'
    else:
        filename = 'AltDatabase/ensembl/'+species+'/'+species+'_'+targetPlatform+'_tissue-
specific_'+coding_type+'.txt'
    if customMarkerFile != False:
        filename = customMarkerFile

    if value_type == 'calls':
        filename = string.replace(filename, '.txt', '_stats.txt')
        fn=filepath(filename); x=0
        tissues_added={}
        for line in open(fn,'rU').xreadlines():
            data = cleanUpLine(line)
            t = string.split(data,'\t')

            if x==0:
                print 'Importing the tissue compedium database:',string.split(filename,delim)[-1][:-4]
                headers = t; x=1; index=0
                for i in headers:
                    if 'UID' == i: ens_index = index; uid_index = index
                    if analysis_type == 'AltExon': ens_index = ens_index ### Assigned above when
analyzing probesets
                    elif 'Ensembl' in i: ens_index = index
                    if 'marker-in' in i: tissue_index = index+1; marker_in = index
                    index+=1
                try:
                    for i in t[tissue_index:]: tissues.append(i)
                except Exception:
                    for i in t[1:]: tissues.append(i)
                if keyed_by == 'primaryID':
                    try: ens_index = uid_index
                    except Exception: None
            else:

```

```

try:
    gene = t[0]
    tissue_exp = map(float, t[1:])
    tissue_specific_db[gene]=x,tissue_exp ### Use this to only grab relevant gene
expression profiles from the input dataset
except Exception:
    gene = string.split(t[ens_index],')[0] ### Only consider the first listed gene - this gene
is the best option based on ExpressionBuilder rankings
    #if 'Pluripotent Stem Cells' in t[marker_in] or 'Heart' in t[marker_in]:
    #if t[marker_in] not in tissues_added: ### Only add the first instance of a gene for that
tissue - used more for testing to quickly run the analysis
    tissue_exp = map(float, t[tissue_index:])
    if value_type == 'calls':
        tissue_exp = produceDetectionCalls(tissue_exp,platform) ### 0 or 1 calls
        tissue_specific_db[gene]=x,tissue_exp ### Use this to only grab relevant gene
expression profiles from the input dataset
        tissues_added[t[marker_in]]=[]
    x+=1
print len(tissue_specific_db), 'genes in the tissue compendium database'

if correlate_to_tissue_specific == 'yes':
try: import TissueCorrelations(filename)
except Exception:
    null=[]
    #print '\nNo tissue-specific correlations file present. Skipping analysis.'; kill

    useMultiRef
return tissue_specific_db

def importTissueCorrelations(filename):
    filename = string.replace(filename,'specific','specific_correlations')
    fn=filepath(filename); x=0
    for line in open(fn,'rU').xreadlines():
        data = cleanUpLine(line)
        if x==0: x=1 ### Ignore header line
        else:
            uid,symbol,rho,tissue = string.split(data,'\t')
            if float(rho)>rho_threshold: ### Variable used for testing different thresholds internally
                try: tissue_to_gene[tissue].append(uid)
                except Exception: tissue_to_gene[tissue] = [uid]

def simpleUIDImport(filename):
    """Import the UIDs in the gene expression file"""
    uid_db={}

```



```

    if value_type == 'calls': ### Hence, this is a DABG or RNA-Seq expression
        exp_vals = produceDetectionCalls(exp_vals,targetPlatform) ### 0 or 1 calls
        gene_expression_db[gene] = [index,exp_vals]
    #print len(gene_expression_db), 'matching genes in the dataset and tissue compendium
    database'

    for gene in genes_added:
        if genes_added[gene]>1: del gene_expression_db[gene] ### delete entries that are present
in the input set multiple times (not trustworthy)
        else: expression_subset.append(gene_expression_db[gene]) ### These contain the rank order
and expression
    #print len(expression_subset);sys.exit()
    expression_subset.sort() ### This order now matches that of
    gene_expression_db=[]
    return expression_subset, sample_headers

def produceDetectionCalls(values,Platform):
    # Platform can be the compendium platform (targetPlatform) or analyzed data platform
(platform or array_type)
    new=[]
    for value in values:
        if Platform == 'RNASeq':
            if value>1:
                new.append(1) ### expressed
            else:
                new.append(0)
        else:
            if value<cutoff: new.append(1)
            else: new.append(0)
    return new

def importGeneIDTranslations(filename):
    ### Import ExpressionOutput/DATASET file to obtain Ensembl associations (typically for
Affymetrix 3' arrays)
    fn=filepath(filename); x=0; translation_db={ }
    for line in open(fn,'rU').xreadlines():
        data = cleanUpLine(line)
        t = string.split(data,'\t')
        if x==0:
            headers = t; x=1; index=0
            for i in headers:
                if 'Ensembl' in i: ens_index = index; break
                index+=1
            else:

```

```

uid = t[0]
ens_geneids = t[ens_index]
ens_geneid = string.split(ens_geneids,')[0] #### In v.2.0.5, the first ID is the best protein
coding candidate
if len(ens_geneid)>0:
    translation_db[uid] = ens_geneid
return translation_db

```

```

def remoteImportExonIDTranslations(array_type,species,translate_to_genearray,targetplatform):
    global targetPlatform; targetPlatform = targetplatform
    translation_db = importExonIDTranslations(array_type,species,translate_to_genearray)
    return translation_db

```

10

```

def importExonIDTranslations(array_type,species,translate_to_genearray):
    gene_translation_db={ }; gene_translation_db2={ }
    if targetPlatform == 'gene' and translate_to_genearray == 'no':
        #### Get gene array to exon array probeset associations
        gene_translation_db = importExonIDTranslations('gene',species,'yes')
        for geneid in gene_translation_db:
            exonid = gene_translation_db[geneid]
            gene_translation_db2[exonid] = geneid
            #print exonid, geneid
        translation_db = gene_translation_db2
    else:

```

20

```

        filename = 'AltDatabase/'+species+'/'+array_type+'/'+species+'_'+array_type+'-
exon_probesets.txt'

```

```

        #### Import exon array to target platform translations (built for DomainGraph visualization)
        fn=filepath(filename); x=0; translation_db={}
        print 'Importing the translation file',string.split(fn,delim)[-1][:-4]

```

```

        for line in open(fn,'rU').xreadlines():

```

30

```

            data = cleanUpLine(line)

```

```

            t = string.split(data,'\t')

```

```

            if x==0: x=1

```

```

            else:

```

```

                platform_id,exon_id = t

```

```

                if targetPlatform == 'gene' and translate_to_genearray == 'no':

```

```

                    try:
                        translation_db[platform_id] = gene_translation_db[exon_id] #### return RNA-Seq
to gene array probeset ID

```

```

                        #print platform_id, exon_id, gene_translation_db[exon_id];sys.exit()

```

40

```

                    except Exception: null=[]

```

```

                else:

```

```

                    translation_db[platform_id] = exon_id

```

```

del gene_translation_db; del gene_translation_db2
return translation_db

def analyzeTissueSpecificExpressionPatterns(tissue_specific_db,expression_subset):
    tissue_specific_sorted = []; genes_present={}; tissue_exp_db={}; gene_order_db={};
    gene_order=[]
    gene_list=[]
    for (index,vals) in expression_subset: genes_present[index]=[]
    for gene in tissue_specific_db:
        gene_list.append(gene)
        tissue_specific_sorted.append(tissue_specific_db[gene])
        gene_order_db[tissue_specific_db[gene][0]] = gene ### index order (this index was created
before filtering)
        tissue_specific_sorted.sort()

    new_index=0
    for (index,tissue_exp) in tissue_specific_sorted:
        try:
            null=genes_present[index]
            i=0
            gene_order.append([new_index,gene_order_db[index]]); new_index+=1
            for f in tissue_exp:
                ### The order of the tissue specific expression profiles is based on the import gene
order
                try: tissue_exp_db[tissues[i]].append(f)
                except Exception: tissue_exp_db[tissues[i]] = [f]
                i+=1

        except Exception: null=[] ### Gene is not present in the input dataset

    ### Organize sample expression, with the same gene order as the tissue expression set
sample_exp_db={}
    for (index,exp_vals) in expression_subset:
        i=0
        for f in exp_vals:
            ### The order of the tissue specific expression profiles is based on the import gene order
            try: sample_exp_db[sample_headers[i]].append(f)
            except Exception: sample_exp_db[sample_headers[i]] = [f]
            i+=1

    if correlate_by_order == 'yes':
        ### Rather than correlate to the absolute expression order, correlate to the order of
expression (lowest to highest)
        sample_exp_db = replaceExpressionWithOrder(sample_exp_db)

```

```
tissue_exp_db = replaceExpressionWithOrder(tissue_exp_db)
```

```
global tissue_comparison_scores; tissue_comparison_scores={ }
```

```
if correlate_to_tissue_specific == 'yes':
```

```
    ### Create a gene_index that reflects the current position of each gene
```

```
    gene_index={ }
```

```
    for (i, gene) in gene_order: gene_index[gene] = i
```

```
    ### Create a tissue to gene-index from the gene_index
```

```
    tissue_to_index={ }
```

10

```
    for tissue in tissue_to_gene:
```

```
        for gene in tissue_to_gene[tissue]:
```

```
            if gene in gene_index: ### Some are not in both tissue and sample datasets
```

```
                index = gene_index[gene] ### Store by index, since the tissue and expression lists
```

```
are sorted by index
```

```
                try: tissue_to_index[tissue].append(index)
```

```
                except Exception: tissue_to_index[tissue] = [index]
```

```
            tissue_to_index[tissue].sort()
```

```
    sample_exp_db, tissue_exp_db =
```

```
    return TissueSpecificExpressionProfiles(sample_exp_db, tissue_exp_db, tissue_to_index)
```

20

```
distributionNull = True
```

```
if Permute:
```

```
    import copy
```

```
    sample_exp_db_original = copy.deepcopy(sample_exp_db)
```

```
    tissue_exp_db_original = copy.deepcopy(tissue_exp_db)
```

```
    group_list=[]; group_db={ }
```

```
    for sample in sample_exp_db:
```

```
        group = string.split(sample, ':')[0]
```

```
        try: group_db[group].append(sample)
```

```
        except Exception: group_db[group] = [sample]
```

30

```
import random
```

```
if distributionNull:
```

```
    group_lengths=[]
```

```
    for group in group_db:
```

```
        group_lengths.append(len(group_db[group]))
```

```
    group_db={ }
```

```
    for sample in sample_exp_db:
```

```
        group = 'null1'
```

```
        try: group_db[group].append(sample)
```

```
        except Exception: group_db[group] = [sample]
```

40

```
    group_db['null2'] = group_db['null1']
```

```

    choice = random.sample
    tissue_groups = ['null1','null2']
else:
    choice = random.choice
    tissue_groups = tuple(tissues)

permute_groups=[]
groups=[]
gn=0
for group in group_db:
    samples = group_db[group]
    permute_db={}; x=0
    while x<200:
        if distributionNull:
            size = group_lengths[gn]
            psamples = choice(samples,size)
        else: psamples = [choice(samples) for _ in xrange(len(samples))] ### works for
random.sample or choice (with replacement)
        permute_db[tuple(psamples)]=None
        x+=1
    permute_groups.append(permute_db)
    groups.append(group); gn+=1 ### for group sizes

groups.sort()
permute_group1 = permute_groups[0]
permute_group2 = permute_groups[1]

permute_group1_list=[]
permute_group2_list=[]
for psamples in permute_group1:
    permute_group1_list.append(psamples)
for psamples in permute_group2:
    permute_group2_list.append(psamples)

i=0; diff_list=[]
group_zdiff_means={}
sample_diff_zscores=[]
for psamples1 in permute_group1_list:
    psamples2 = permute_group2_list[i] #this is the second group to compare to
    x=0; permute_sample_exp_db={}
    for sample in psamples1:
        if distributionNull:
            nsample = 'null1:'+string.split(sample,')[1] ### reassign group ID
            new_sampleID=nsample+str(x)

```

```

else: new_sampleID=sample+str(x)
try: permute_sample_exp_db[new_sampleID]=sample_exp_db[sample]
except Exception: print sample, new_sampleID, sample_exp_db[sample];sys.exit()
x+=1
for sample in psamples2:
    if distributionNull:
        nsample = 'null2:'+string.split(sample,':')[1] ### reassign group ID
        new_sampleID=nsample+str(x)
    else: new_sampleID=sample+str(x)
    permute_sample_exp_db[new_sampleID]=sample_exp_db[sample]
    x+=1
i+=1

new_tissue_exp_db={ }
### Create a new reference from the permuted data
for sample in permute_sample_exp_db:
    group = string.split(sample,':')[0]
    try: new_tissue_exp_db[group].append(permute_sample_exp_db[sample])
    except Exception: new_tissue_exp_db[group] = [permute_sample_exp_db[sample]]
for group in new_tissue_exp_db:
    k = new_tissue_exp_db[group]
    new_tissue_exp_db[group] = [Average(value) for value in zip(*k)] ### create new
reference from all same group sample values

PearsonCorrelationAnalysis(permute_sample_exp_db,new_tissue_exp_db)

zscore_output_dir,tissue_scores = exportCorrelationResults()
tissue_comparison_scores={ }

headers = list(tissue_scores['headers']); del tissue_scores['headers']
index=0; positive=0; positive_score_diff=0
sample_number = (len(headers)-1)
diff_z_list=[]
population1_denom=0; population1_pos=0; population2_pos=0; population2_denom=0

group_diff_z_scores={ } ### Keep track of the differences between the z-scores between
the two groups
while index < sample_number:
    j=0
    #ref1 = tissue_groups[0]+'; ref2 = tissue_groups[-1]+'
    sample = headers[index+1]
    diff_z = tissue_scores[tissue_groups[0]][index]-tissue_scores[tissue_groups[-
1]][index]
    diff_list.append((diff_z,sample))

```

```

group = string.split(sample,':')[0]
try: group_diff_z_scores[group].append(diff_z)
except Exception: group_diff_z_scores[group] = [diff_z]
sample_diff_zscores.append(diff_z)
index+=1

for group in group_diff_z_scores:
    avg_group_zdiff = Average(group_diff_z_scores[group])
    try: group_zdiff_means[group].append(avg_group_zdiff)
    except Exception: group_zdiff_means[group] = [avg_group_zdiff]
10

diff_list.sort()

all_group_zdiffs=[]
for group in group_zdiff_means:
    all_group_zdiffs += group_zdiff_means[group]
all_group_zdiffs.sort()

print sample_diff_zscores;sys.exit()

#for i in diff_list: print i
#sys.exit()

i=1
groups.reverse()
group1,group2 = groups[:2]
group1+=': '; group2+=': '
scores=[]
print max(diff_list), min(diff_list);sys.exit()
while i < len(diff_list):
    g1_hits=0; g2_hits=0
    list1 = diff_list[:i]
    list2 = diff_list[i:]
    for (z,s) in list1:
        if group1 in s: g1_hits+=1
    for (z,s) in list2:
        if group2 in s: g2_hits+=1
    sensitivity = float(g1_hits)/len(list1)
    specificity = float(g2_hits)/len(list2)
    accuracy = sensitivity+specificity
    #accuracy = g1_hits+g2_hits
30

#print g1_hits, len(list1)
#print g2_hits, len(list2)
40

```

```

    #print sensitivity, specificity;sys.exit()
    z_cutoff = Average([list1[-1][0],list2[0][0]])
    scores.append([accuracy,z_cutoff])
    i+=1

scores.sort(); scores.reverse()
print scores[0][0],'\t',scores[0][1]

sample_exp_db = sample_exp_db_original
tissue_exp_db = tissue_exp_db_original

```

10

```

PearsonCorrelationAnalysis(sample_exp_db,tissue_exp_db)
sample_exp_db=[]; tissue_exp_db=[]
zscore_output_dir,tissue_scores = exportCorrelationResults()
return zscore_output_dir, tissue_scores

```

```

def returnTissueSpecificExpressionProfiles(sample_exp_db,tissue_exp_db,tissue_to_index):
    tissue_exp_db_abbreviated={}
    sample_exp_db_abbreviated={} ### This db is designed differently than the non-tissue specific
    (keyed by known tissues)

```

20

```

    ### Build the tissue specific expression profiles
    for tissue in tissue_exp_db:
        tissue_exp_db_abbreviated[tissue] = []
        for index in tissue_to_index[tissue]:
            tissue_exp_db_abbreviated[tissue].append(tissue_exp_db[tissue][index]) ### populate
with just marker expression profiles

    ### Build the sample specific expression profiles
    for sample in sample_exp_db:
        sample_tissue_exp_db={}
        sample_exp_db[sample]
        for tissue in tissue_to_index:
            sample_tissue_exp_db[tissue] = []
            for index in tissue_to_index[tissue]:
                sample_tissue_exp_db[tissue].append(sample_exp_db[sample][index])
        sample_exp_db_abbreviated[sample] = sample_tissue_exp_db
    return sample_exp_db_abbreviated, tissue_exp_db_abbreviated

```

30

```

def replaceExpressionWithOrder(sample_exp_db):
    for sample in sample_exp_db:
        sample_exp_sorted=[]; i=0
        for exp_val in sample_exp_db[sample]: sample_exp_sorted.append([exp_val,i]); i+=1
        sample_exp_sorted.sort(); sample_exp_resort = []; order = 0

```

40

```

for (exp_val,i) in sample_exp_sorted: sample_exp_resort.append([i,order]); order+=1
sample_exp_resort.sort(); sample_exp_sorted=[] ### Order lowest expression to highest
for (i,o) in sample_exp_resort: sample_exp_sorted.append(o) ### The expression order
replaces the expression, in the original order
sample_exp_db[sample] = sample_exp_sorted ### Replace exp with order
return sample_exp_db

```

```

def PearsonCorrelationAnalysis(sample_exp_db,tissue_exp_db):
    #print "Beginning LineageProfiler analysis"
    k=0
    original_increment = int(len(tissue_exp_db)/15.00); increment = original_increment
    p = 1 ### Default value if not calculated
    for tissue in tissue_exp_db:
        #print k,"of",len(tissue_exp_db),"classifier tissue/cell-types"
        if k == increment: increment+=original_increment; #print '*',
        k+=1
        tissue_expression_list = tissue_exp_db[tissue]
        for sample in sample_exp_db:
            if correlate_to_tissue_specific == 'yes':
                ### Keyed by tissue specific sample profiles
                sample_expression_list = sample_exp_db[sample][tissue] ### dictionary as the value
for sample_exp_db[sample]
                #print tissue, sample_expression_list
                #print tissue_expression_list; sys.exit()
            else: sample_expression_list = sample_exp_db[sample]
            try:
                ### p-value is likely useful to report (not supremely accurate but likely sufficient)
                rho,p = stats.pearsonr(tissue_expression_list,sample_expression_list)
                pearson_list.append(rho)
                try: tissue_comparison_scores[tissue].append([rho,p,sample])
                except Exception: tissue_comparison_scores[tissue] = [[rho,p,sample]]
            except Exception:
                ### simple pure python implementation - no scipy required (not as fast though and no
p-value)
                try:
                    rho = pearson(tissue_expression_list,sample_expression_list); p=0
                    try: tissue_comparison_scores[tissue].append([rho,p,sample])
                    except Exception: tissue_comparison_scores[tissue] = [[rho,p,sample]]
                    pearson_list.append(rho)
                except Exception: None ### Occurs when an invalid string is present - ignore and
move onto the next model
                #tst = salstat_stats.TwoSampleTests(tissue_expression_list,sample_expression_list)
                #pp,pr = tst.PearsonsCorrelation()
                #sp,sr = tst.SpearmanCorrelation()

```

```

        #print tissue, sample
        #if rho>.5: print [rho, pr, sr],[pp,sp];sys.exit()
        #if rho<.5: print [rho, pr, sr],[pp,sp];sys.exit()
sample_exp_db=[]; tissue_exp_db=[]
#print 'Correlation analysis finished'

def pearson(array1,array2):
    item = 0; sum_a = 0; sum_b = 0; sum_c = 0
    while item < len(array1):
        a = (array1[item] - Average(array1))*(array2[item] - Average(array2))
        b = math.pow((array1[item] - Average(array1)),2)
        c = math.pow((array2[item] - Average(array2)),2)
        sum_a = sum_a + a
        sum_b = sum_b + b
        sum_c = sum_c + c
        item = item + 1
    r = sum_a/math.sqrt(sum_b*sum_c)
    return r
10

def Median(array):
    array.sort()
    len_float = float(len(array))
    len_int = int(len(array))
    if (len_float/2) == (len_int/2):
        try: median_val = avg([array[(len_int/2)-1],array[(len_int/2)]])
        except IndexError: median_val = "
    else:
        try: median_val = array[len_int/2]
        except IndexError: median_val = "
    return median_val
20
30

def Average(array):
    try: return sum(array)/len(array)
    except Exception: return 0

def adjustPValues():
    """ Can be applied to calculate an FDR p-value on the p-value reported by scipy.
    Currently this method is not employed since the p-values are not sufficiently
    stringent or appropriate for this type of analysis """
40

import statistics
all_sample_data={}
for tissue in tissue_comparison_scores:

```

```

    for (r,p,sample) in tissue_comparison_scores[tissue]:
        all_sample_data[sample] = db = { } ### populate this dictionary and create sub-
dictionaries
        break

```

```

for tissue in tissue_comparison_scores:
    for (r,p,sample) in tissue_comparison_scores[tissue]:
        gs = statistics.GroupStats(", ",p)
        all_sample_data[sample][tissue] = gs
for sample in all_sample_data:
    statistics.adjustPermuteStats(all_sample_data[sample])

```

10

```

for tissue in tissue_comparison_scores:
    scores = []
    for (r,p,sample) in tissue_comparison_scores[tissue]:
        p = all_sample_data[sample][tissue].AdjP()
        scores.append([r,p,sample])
    tissue_comparison_scores[tissue] = scores

```

```

def stdev(array):
    sum_dev = 0
    try: x_bar = scipy.average(array)
    except Exception: x_bar=Average(array)
    n = float(len(array))
    for x in array:
        x = float(x)
        sq_deviation = math.pow((x-x_bar),2)
        sum_dev += sq_deviation

```

20

```

try:
    s_sqr = (1.0/(n-1.0))*sum_dev #s squared is the variance
    s = math.sqrt(s_sqr)
except Exception:
    s = 'null'
return s

```

30

```

def replacePearsonPvalueWithZscore():
    adjust_rho=True
    all_sample_data={ }
    for tissue in tissue_comparison_scores:
        for (r,p,sample) in tissue_comparison_scores[tissue]:
            all_sample_data[sample] = [] ### populate this dictionary and create sub-dictionaries
        break

```

40

```

for tissue in tissue_comparison_scores:
    for (r,p,sample) in tissue_comparison_scores[tissue]:
        if adjust_rho:
            try: r = 0.5*math.log(((1+r)/(1-r)))
            except Exception: print tissue, sample, r, p; sys.exit()
        all_sample_data[sample].append(r)
        #print tissue, sample, r

sample_stats={}
all_dataset_rho_values=[]
### Get average and standard deviation for all sample rho's
for sample in all_sample_data:
    all_dataset_rho_values+=all_sample_data[sample]
    try: avg=scipy.average(all_sample_data[sample])
    except Exception: avg=Average(all_sample_data[sample])
    st_dev=stdev(all_sample_data[sample])
    sample_stats[sample]=avg,st_dev

try: global_rho_avg = scipy.average(all_dataset_rho_values)
except Exception: global_rho_avg=Average(all_sample_data[sample])
global_rho_stdev = stdev(all_dataset_rho_values)

### Replace the p-value for each rho
for tissue in tissue_comparison_scores:
    scores = []
    for (r,p,sample) in tissue_comparison_scores[tissue]:
        if adjust_rho:
            try: r = 0.5*math.log(((1+r)/(1-r)))
            except Exception: print tissue, sample, r, p; sys.exit()
        #u,s=sample_stats[sample]
        #z = (r-u)/s
        z = (r-global_rho_avg)/global_rho_stdev ### Instead of doing this for the sample
background, do it relative to all analyzed samples
        #z_alt = (r-global_rho_avg)/global_rho_stdev
        scores.append([r,z,sample])
        #print sample, r, global_rho_avg, global_rho_stdev, z
    tissue_comparison_scores[tissue] = scores

def exportCorrelationResults():
    corr_output_file = string.replace(exp_output_file,'DATASET','LineageCorrelations')
    corr_output_file = string.replace(corr_output_file,'.txt','-'+coding_type+'.txt')
    if analysis_type == 'AltExon':
        corr_output_file = string.replace(corr_output_file,coding_type,'AltExon')
    filename = string.split(corr_output_file,delim)[-1][:4]

```

```

#score_data = exportFile(corr_output_file)

zscore_output_dir = string.replace(corr_output_file,'.txt','-zscores.txt')
#probability_data = exportFile(zscore_output_dir) .
#adjustPValues()
replacePearsonPvalueWithZscore()

### Make title row
headers=['Sample_name']
for tissue in tissue_comparison_scores:
    for (r,z,sample) in tissue_comparison_scores[tissue]: headers.append(sample)
    break
#title_row = string.join(headers,'\t')+'\n'
#score_data.write(title_row)
#if use_scipy: probability_data.write(title_row)
### Export correlation data
tissue_scores = { }; tissue_probabilities={ }; tissue_score_list = [] ### store and rank tissues
according to max(score)
for tissue in tissue_comparison_scores:
    scores=[]
    probabilities=[]
    for (r,z,sample) in tissue_comparison_scores[tissue]:
        scores.append(r)
        probabilities.append(z)
    tissue_score_list.append((max(scores),tissue))
    tissue_scores[tissue] = probabilities ### These are actually z-scores
    #tissue_scores[tissue] = string.join(map(str,[tissue]+scores),'\t')+'\n' ### export line
    if use_scipy:
        tissue_probabilities[tissue] = string.join(map(str,[tissue]+probabilities),'\t')+'\n'

tissue_score_list.sort()
tissue_score_list.reverse()
#for (score,tissue) in tissue_score_list:
#score_data.write(tissue_scores[tissue])
#if use_scipy: probability_data.write(tissue_probabilities[tissue])
#score_data.close()
#if use_scipy: probability_data.close()
#print filename,'exported...'
tissue_scores['headers'] = headers
return zscore_output_dir, tissue_scores

def visualizeLineageZscores(zscore_output_dir,grouped_lineage_zscore_dir,graphic_links):
import clustering
### Perform hierarchical clustering on the LineageProfiler Zscores

```

```

graphic_links = clustering.runHCOOnly(zscore_output_dir,graphic_links)
return graphic_links

if __name__ == '__main__':

    ##### Default Variables #####
    species = 'Hs'
    platform = "exon"
    vendor = 'Affymetrix'
    compendium_platform = "exon"
    codingtype = 'protein_coding'
    platform = vendor, platform
    exp_output = None
    geneModels = False
    modelSize = None
    permute = False
    useMulti = False

    """ This script iterates the LineageProfiler algorithm (correlation based classification method)
    to identify sample types relative
    two one of two references given one or more gene models. The program '
    """

    #python LineageProfilerIterate.py --i
    "/Users/nsalomonis/Desktop/dataAnalysis/qPCR/ExpressionInput/exp.ABI_Pediatric.txt" --r
    "/Users/nsalomonis/Desktop/dataAnalysis/qPCR/ExpressionOutput/MarkerFinder/MarkerFinder
    -ABI_Pediatric.txt" --m
    "/Users/nsalomonis/Desktop/dataAnalysis/qPCR/ExpressionInput/7GeneModels.txt"
    #python LineageProfilerIterate.py --i
    "/Users/nsalomonis/Desktop/dataAnalysis/qPCR/deltaCT/LabMeeting/ExpressionInput/exp.ABI
    _PediatricSNS.txt" --r
    "/Users/nsalomonis/Desktop/dataAnalysis/qPCR/ExpressionOutput/MarkerFinder/MarkerFinder
    -ABI_PediatricSNS.txt" --s 4

    ##### Comand-line arguments #####
    if len(sys.argv[1:])<=1: ### Indicates that there are insufficient number of command-line
    arguments
        print "Warning! Please designate a tab-delimited input expression file in the command-line"
        print 'Example: python LineageProfilerIterate.py --i "/Users/me/qPCR.txt" --r
        "/Users/me/reference.txt" --m "/Users/me/models.txt"'
    else:
        try:
            options, remainder = getopt.getopt(sys.argv[1:],
            ['i=', 'species=', 'o=', 'platform=', 'codingtype=',

```

```

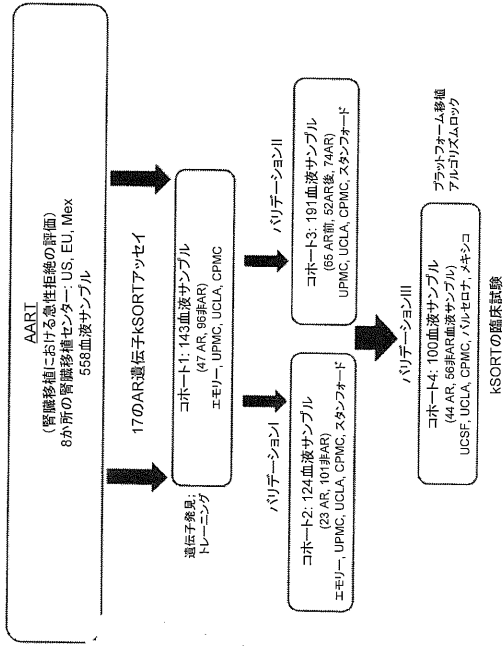
'compendium_platform=', 'r=', 'm=', 'v=', 's=', 'permute=', 'useMulti='])
except Exception, e:
    print ""
for opt, arg in options:
    if opt == '--i': exp_input=arg
    elif opt == '--o': exp_output=arg
    elif opt == '--platform': platform=arg
    elif opt == '--codingtype': codingtype=arg
    elif opt == '--compendium_platform': compendium_platform=arg
    elif opt == '--r': customMarkers=arg
    elif opt == '--m': geneModels=arg
    elif opt == '--v': vendor=arg
    elif opt == '--permute': permute=True
    elif opt == '--useMulti': useMulti=True
    elif opt == '--s':
        try: modelSize = int(arg)
        except Exception:
            modelSize = arg
            if modelSize != 'optimize':
                print 'Please specify a modelSize (e.g., 7-gene model search) as a single integer
(e.g., 7)'
                sys.exit()
    else:
        print "Warning! Command-line argument: %s not recognized. Exiting..." % opt;
sys.exit()

if exp_output == None: exp_output = exp_input

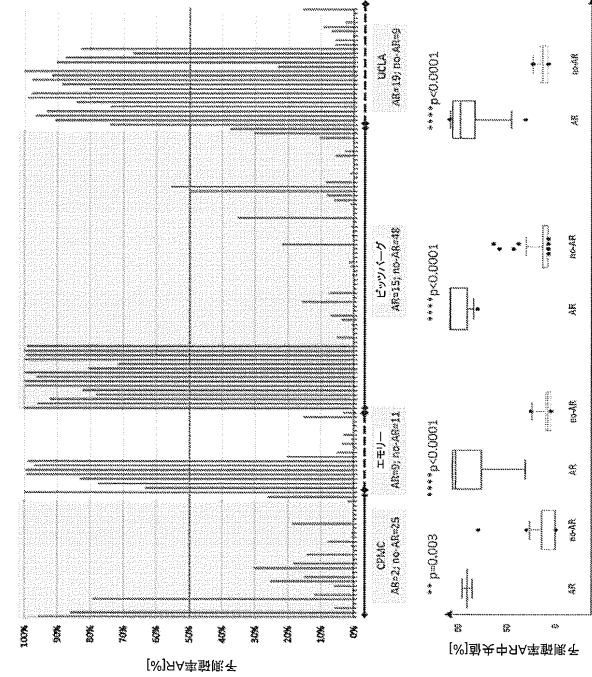
runLineageProfiler(species, platform, exp_input, exp_output, codingtype, compendium_platform, modelSize=modelSize, customMarkers=customMarkers, geneModels=geneModels, permute=permute, useMulti=useMulti)

```

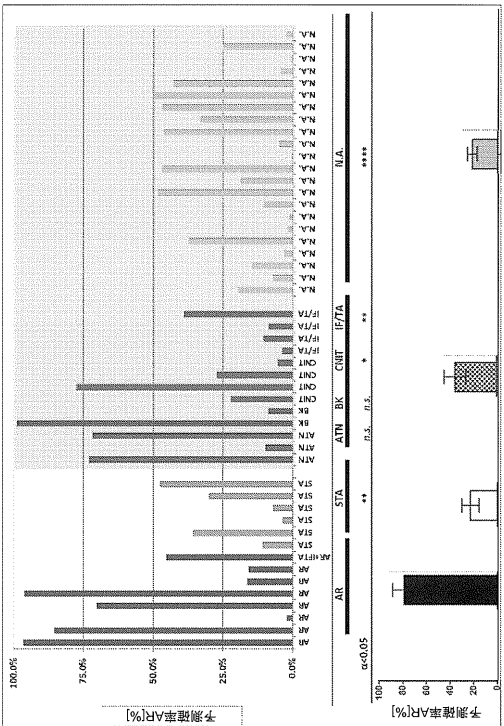
【 図 1 】



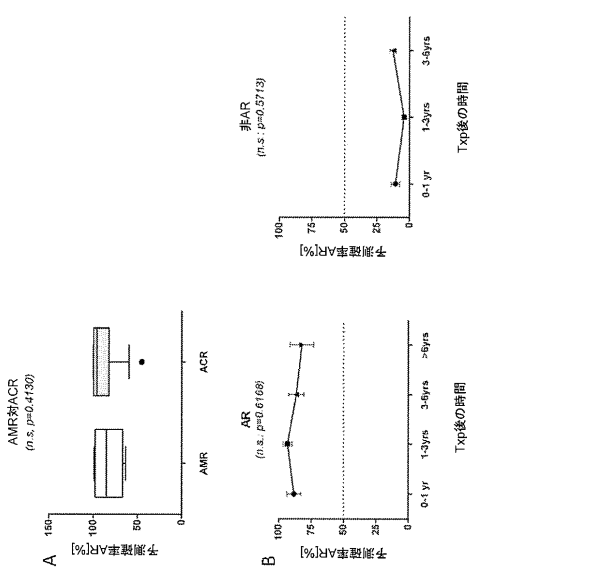
【 図 2 A 】



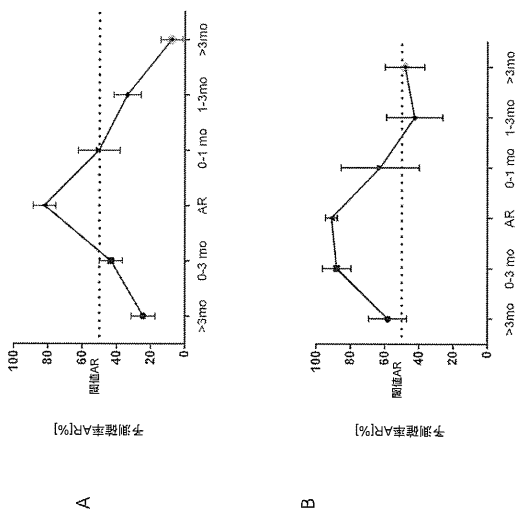
【 図 2 B 】



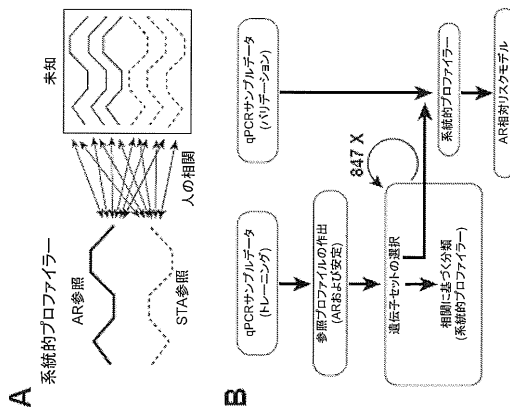
【 図 3 】



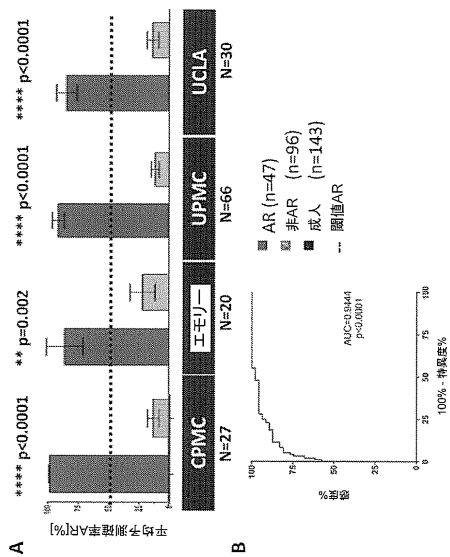
【 図 4 】



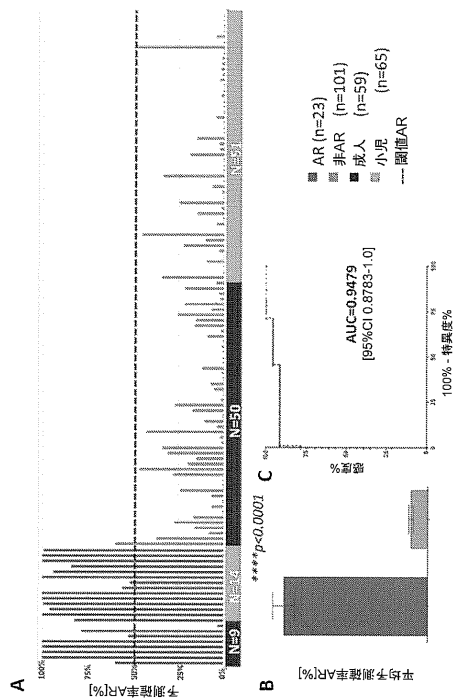
【 図 5 】



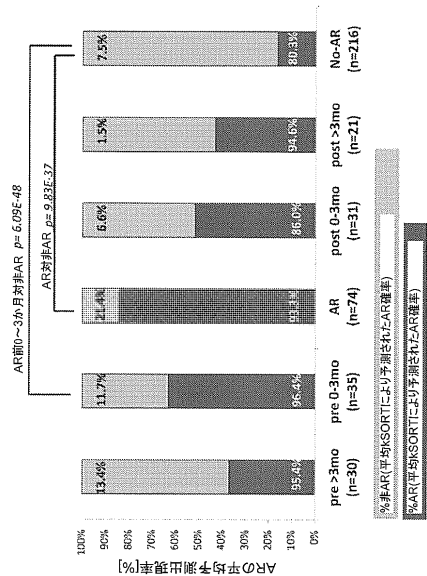
【 図 6 】



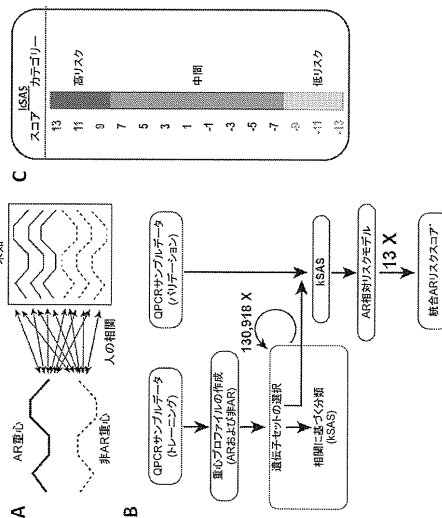
【 図 7 】



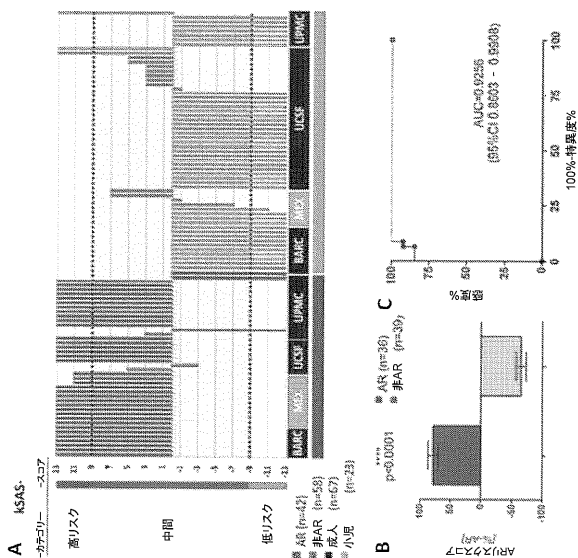
【 8 】



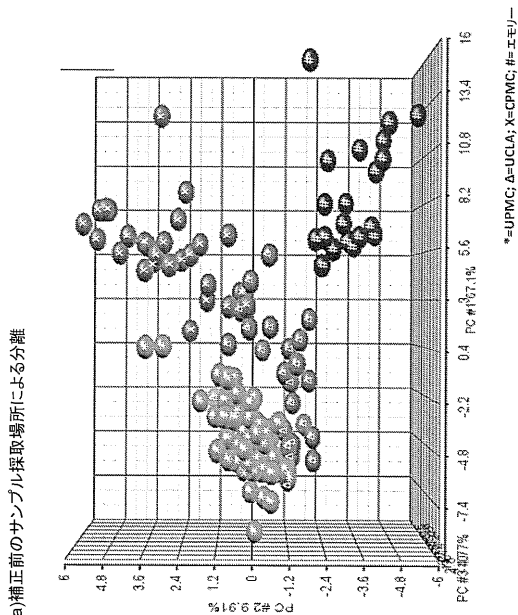
【 9 】



【 10 】

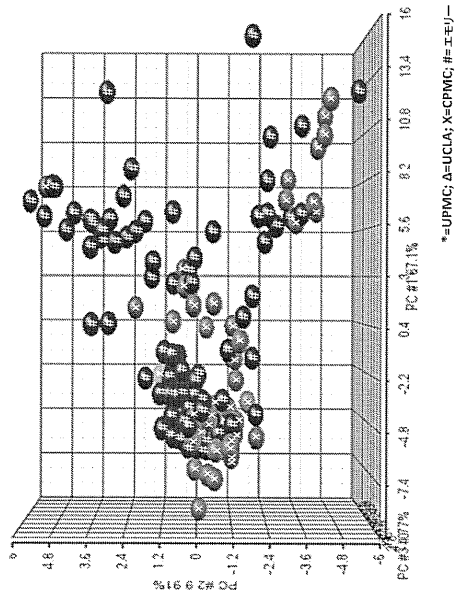


【 11 A 】



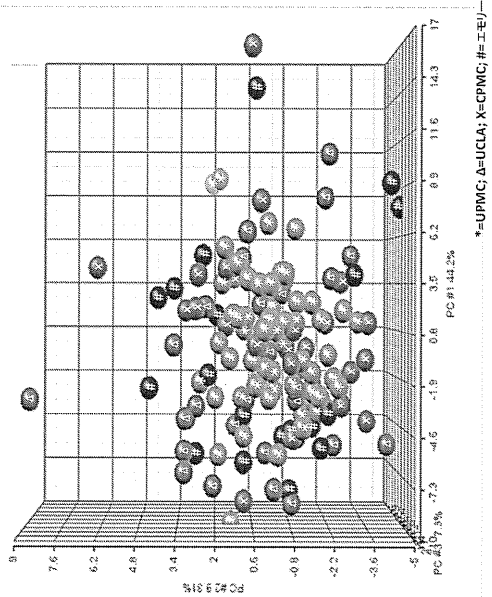
【 図 1 1 B 】

b) 補正前の表現型による分離なし



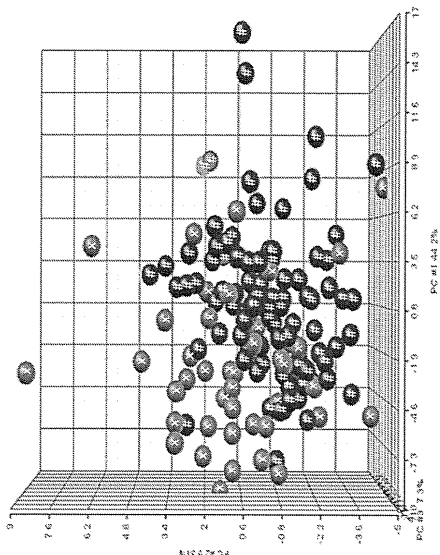
【 図 1 1 C 】

c) 補正後のサンプル採取場所による分離なし

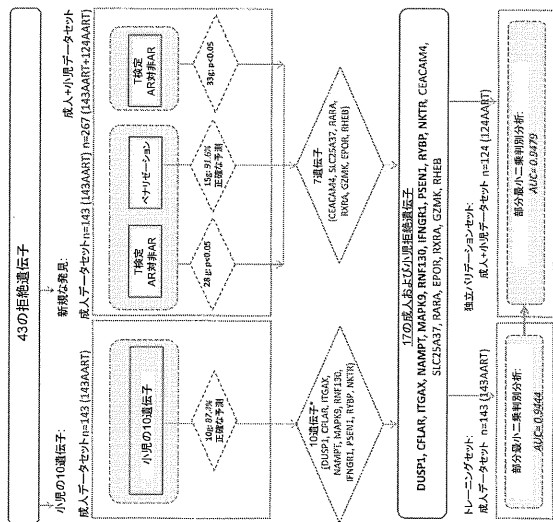


【 図 1 1 D 】

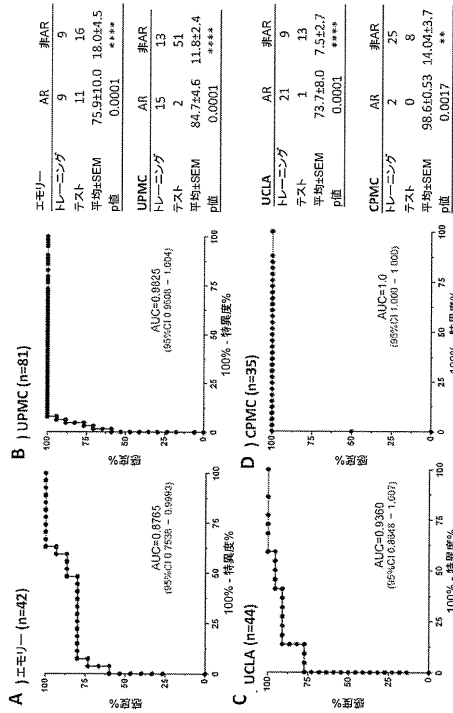
d) 補正後の表現型による分離



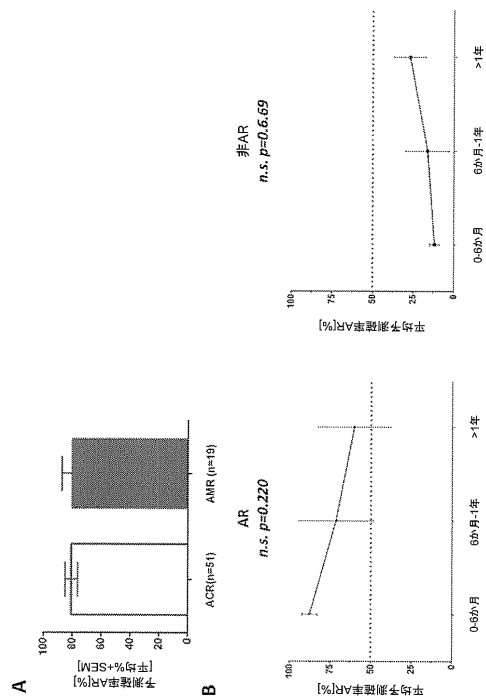
【 図 1 2 】



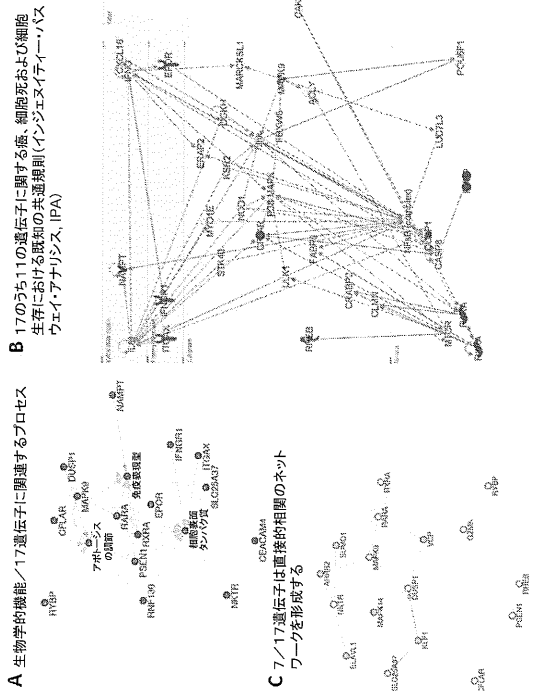
【 図 1 3 】



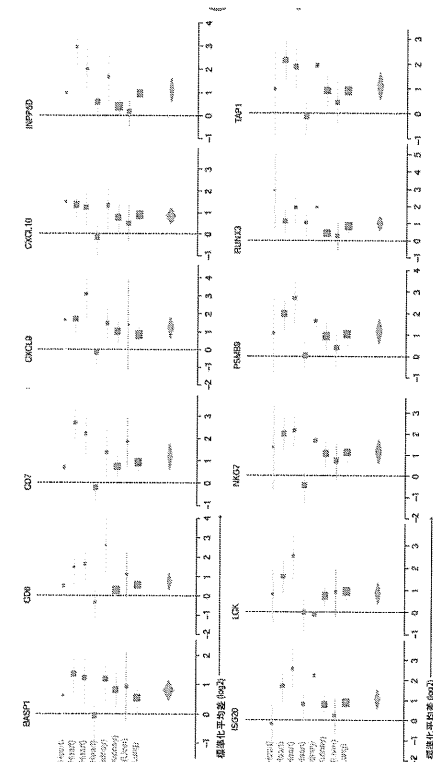
【 図 1 4 】





【 図 1 5 】



【 図 1 6 】



【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US2014/054342
A. CLASSIFICATION OF SUBJECT MATTER C12Q 1/68(2006.01)i		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) C12Q 1/68; G01N 33/68		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Korean utility models and applications for utility models Japanese utility models and applications for utility models		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) eKOMPASS(KIPO internal) & Keywords: acute rejection, renal allograft, biomarker, CEACAM4		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2013-0143755 A1 (SARWAL et al.) 06 June 2013 See abstract and claims 1, 15-16.	29-32, 45-48, 60-63
A	LI et al., 'A peripheral blood diagnostic test for acute rejection in renal transplantation' American Journal of Transplantation, Vol.12, No.10, pp.2710-2718 (2012) See abstract and discussion.	29-32, 45-48, 60-63
A	HEIDT et al., 'Peripheral blood sampling for the detection of allograft rejection: biomarker identification and validation' Transplantation, Vol.92, No.1, pp.1-9 (2011) See the whole document.	29-32, 45-48, 60-63
A	HARTONO et al., 'Noninvasive diagnosis of acute rejection of renal allografts' Current Opinion in Organ Transplantation, Vol.15, No.1, pp.35-41 (2010) See the whole document.	29-32, 45-48, 60-63
A	NAESENS et al., 'Progressive histological damage in renal allografts is associated with expression of innate and adaptive immunity genes' Kidney International, Vol.80, No.12, pp.1364-1376 (2011) See the whole document.	29-32, 45-48, 60-63
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed		"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family
Date of the actual completion of the international search 28 November 2014 (28.11.2014)		Date of mailing of the international search report 01 December 2014 (01.12.2014)
Name and mailing address of the ISA/KR  International Application Division Korean Intellectual Property Office 189 Cheongsa-ro, Seo-gu, Daejeon Metropolitan City, 302-701, Republic of Korea Facsimile No. +82-42-472-7140		Authorized officer KIM, Seung Beom Telephone No. +82-42-481-3371 

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2014/054342

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

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

1. Claims Nos.: 1-28,74-87
because they relate to subject matter not required to be searched by this Authority, namely:
Claims 1-28 and 74-87 pertain to methods for treatment of the human body by therapy, as well as diagnostic methods, and thus relate to a subject matter which this International Searching Authority is not required, under PCT Article 17(2)(a)(i) and PCT Rule 39.1(iv), to search.
2. Claims Nos.: 9-10,23-24,39-40,53-54,68-69,82-83
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
Claims 9,10,23,24,39,40,53,54,68,69,82 and 83 are unclear since they refer to multiple dependent claims which do not comply with PCT Rule 6.4(a).
3. Claims Nos.: 5-8,11-14,19-22,25-28,33-38,41-44,49-52,55-59,64-67,70-73,78-81,84-87
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

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

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

Remark on Protest

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

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2014/054342

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2013-0143755 A1	06/06/2013	CA 2794255 A1 EP 2550371 A1 EP 2550371 A4 WO 2011-119980 A1	29/09/2011 30/01/2013 20/11/2013 29/09/2011

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 P 13/12 (2006.01) A 6 1 P 13/12

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

(74)代理人 100169971
 弁理士 菊田 尚子

(74)代理人 100144794
 弁理士 大木 信人

(72)発明者 サーワル, ミニー エム.
 アメリカ合衆国 9 4 0 4 3 カリフォルニア州, マウンテン ビュー, スイート 3 0 7, ベイ
 ショア パークウェイ 2 6 8 0

F ターム(参考) 4B029 AA07 AA23 BB11 CC01 FA12 FA15 GA08 GB10
 4B063 QA01 QA05 QA19 QQ03 QQ08 QQ52 QR32 QR35 QR55 QR59
 QR62 QS25 QS28 QS34 QX02
 4C084 AA17 NA14 ZA812

专利名称(译)	用于评估肾移植中急性排斥的组合物和方法		
公开(公告)号	JP2016531580A	公开(公告)日	2016-10-13
申请号	JP2016540430	申请日	2014-09-05
[标]申请(专利权)人(译)	IMMUCOR GTI诊断		
申请(专利权)人(译)	Imukoa啧啧茶眼诊断公司		
[标]发明人	サーワルミニーエム		
发明人	サーワル, ミニー エム.		
IPC分类号	C12Q1/68 C12N15/09 C12M1/00 G01N33/53 A61K45/00 A61P13/12		
CPC分类号	A61P13/12 C12Q1/6883 C12Q2600/118 C12Q2600/158		
FI分类号	C12Q1/68.A C12N15/00.A C12M1/00.A G01N33/53.M A61K45/00 A61P13/12		
F-TERM分类号	4B029/AA07 4B029/AA23 4B029/BB11 4B029/CC01 4B029/FA12 4B029/FA15 4B029/GA08 4B029/GB10 4B063/QA01 4B063/QA05 4B063/QA19 4B063/QQ03 4B063/QQ08 4B063/QQ52 4B063/QR32 4B063/QR35 4B063/QR55 4B063/QR59 4B063/QR62 4B063/QS25 4B063/QS28 4B063/QS34 4B063/QX02 4C084/AA17 4C084/NA14 4C084/ZA812		
代理人(译)	荒井英一 Nobuto 滤纸冲		
优先权	61/874970 2013-09-06 US 61/987342 2014-05-01 US		
其他公开文献	JP2016531580A5		
外部链接	Espacenet		

摘要(译)

本文提供了使用分类器基因集的基因表达谱诊断肾移植急性排斥反应的方法，组合物和试剂盒。这样的方法和组合物与外部混杂因素无关，例如受体年龄，移植中心，RNA来源，测定法，终末期肾脏疾病的病因，合并症，免疫抑制用途。 [选型图]图1

