

(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2014-501098
(P2014-501098A)

(43) 公表日 平成26年1月20日(2014.1.20)

(51) Int.Cl.	F I	テーマコード (参考)
C12Q 1/68 (2006.01)	C12Q 1/68 ZNAA	4B024
C12N 15/09 (2006.01)	C12N 15/00 A	4B029
C12M 1/00 (2006.01)	C12N 15/00 F	4B063
GO1N 33/53 (2006.01)	C12M 1/00 A	
GO1N 37/00 (2006.01)	GO1N 33/53 D	

審査請求 未請求 予備審査請求 未請求 (全 114 頁) 最終頁に続く

(21) 出願番号 特願2013-536914 (P2013-536914)
 (86) (22) 出願日 平成23年12月22日 (2011.12.22)
 (85) 翻訳文提出日 平成25年6月26日 (2013.6.26)
 (86) 国際出願番号 PCT/US2011/058669
 (87) 国際公開番号 WO2012/058689
 (87) 国際公開日 平成24年5月3日 (2012.5.3)
 (31) 優先権主張番号 61/498,965
 (32) 優先日 平成23年6月20日 (2011.6.20)
 (33) 優先権主張国 米国 (US)

(71) 出願人 513105292
 ビービーディー バイオダイアグノスティックス, エルエルシー
 PBD BIODIAGNOSTICS, LLC
 アメリカ合衆国 メリーランド州 21035 デイヴィッドソンビル アッシュランドドライブ 3539
 3539 Ashland Drive
 Davidsonville, MD 21035 United States of America
 (74) 代理人 100126572
 弁理士 村越 智史

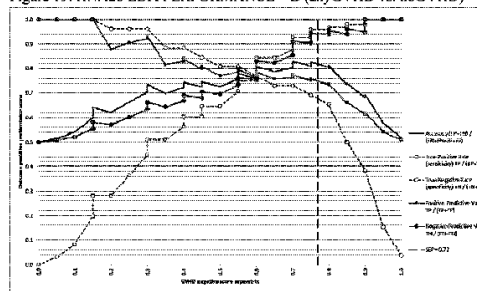
最終頁に続く

(54) 【発明の名称】 造血細胞移植のドナーをスクリーニング、予想、特定する方法、キット及びアレイ、並びに、造血細胞移植 (HCT) が移植片対宿主病 (GVHD) を引き起こすリスクを予測する方法、キット

(57) 【要約】

本発明は、ドナーからの造血細胞移植 (HCT) が HCT レシピエントにおいて移植片対宿主病 (GVHD) を引き起こすリスクを予測又は決定すること、HCT レシピエントにおいてGVHDを引き起こすリスクに応じてドナー候補からのHCTを分類すること、及び一又は複数のドナー候補のHCTがHCT レシピエントにおいてGVHDを引き起こすリスクに応じて分類又は採点されている組織的構成 (例えば、データベース) 及び組織的構成 (例えば、データベース) の生成方法に関する。本発明はまた、ドナー候補からのHCTがHCT レシピエントにおいてGVHDを引き起こすリスクを予測又は決定するため、及び、HCT レシピエントにおいてGVHDを引き起こすリスクに応じてドナーを分類又は採点するために有用なキット及びアレイに関する。

Figure 19: RNA20 LDA PERFORMANCE - B (anyGVHD vs. noGVHD)



【特許請求の範囲】

【請求項 1】

ドナー候補からの造血細胞移植（HCT）がHCTレシピエントにおいて移植片対宿主病（GVHD）を引き起こす又は引き起こさないリスクを予測又は決定する方法であって、

a) ドナー候補のCD4陽性T細胞又はCD8陽性T細胞において、表1（RNA1538）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）もしくは表18（SG64）から選択された一又は複数の陽性又は陰性GVHD予測遺伝子又は陽性及び／又は陰性GVHD予測遺伝子の組み合わせ又はこれらの多型の発現を測定する工程と、

b) 工程a)において測定された発現に基づいて、陽性又は陰性GVHD予測遺伝子の発現量を得るか、又は、工程a)において測定された発現に基づいて、陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合を得る工程と、

c) 前記陽性又は陰性GVHD予測遺伝子の発現量と陽性又は陰性GVHD予測遺伝子の既定の基準発現量とを比較するか、又は、前記陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合と陽性及び／又は陰性GVHD予測遺伝子の線形結合又は非線型結合の既定の基準値とを比較する工程と、

d) 前記ドナー候補からの前記HCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いか低いかを示す発現量、陽性又は陰性GVHD予測遺伝子の総数又は同一性、又は前記陽性及び／又は陰性GVHD予測遺伝子の組み合わせの評価に基づいて、前記ドナー候補からの前記HCTがHCTレシピエントにおいてGVHDを引き起こすか又は引き起こさないリスクを予測し又は決定する工程と、

を備え、

陽性GVHD予測遺伝子の発現量が陽性GVHD予測遺伝子の既定の基準発現量よりも大きい小さいかということが、前記ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示し、又は、

陰性GVHD予測遺伝子の発現量が陰性GVHD予測遺伝子の基準発現量よりも大きい小さいかということが、前記ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示し、又は、

陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きい小さいかということが、前記ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示し、又は、

陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きい小さいかということが、前記ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示す、

方法。

【請求項 2】

ドナー候補からの（HCTがHCTレシピエントにおいて移植片対宿主病（GVHD）を引き起こす又は引き起こさないリスクを予測又は決定する方法であって、

a) ドナー候補のCD4陽性T細胞もしくはCD8陽性T細胞、又は、CD4陽性T細胞又はCD8陽性T細胞によって発現された核酸又はタンパク質を、表1（RNA1538）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）又は表18（SG64）に挙げられた一又は複数の陽性又は陰性GVHD予測遺伝子又はこれらの多型の発現を検出する検体と接触させる工程と、

b) CD4陽性T細胞又はCD8陽性T細胞における前記一又は複数の陽性又は陰性GVHD予測遺伝子の発現を測定して陽性又は陰性GVHD予測遺伝子の発現量を得るか、又は、前記陽性及び／又は陰性予測遺伝子の組み合わせの発現を測定して陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合を得る工程と、

10

20

30

40

50

c) 前記陽性又は陰性 G V H D 予測遺伝子の発現量と陽性又は陰性 G V H D 予測遺伝子の既定の基準発現量とを比較するか、又は、前記陽性及び / 又は陰性 G V H D 予測遺伝子の組み合わせの発現量の線形結合又は非線型結合と前記陽性及び / 又は陰性 G V H D 予測遺伝子の組み合わせの発現量の線形結合又は非線型結合の既定の基準値とを比較する工程と

d) 前記ドナー候補からの前記 H C T が H C T レシピエントにおいて G V H D を引き起こすリスクが高いか低いかを示す発現量、陽性又は陰性 G V H D 予測遺伝子の総数又は同一性、又は陽性及び / 又は陰性 G V H D 予測遺伝子の組み合わせの評価に基づいて、前記ドナー候補からの H C T が H C T レシピエントにおいて G V H D を引き起こすリスクを予測又は決定する工程と、

を備え、

陽性 G V H D 予測遺伝子の発現量が陽性 G V H D 予測遺伝子の既定の基準発現量よりも大きい小さいかということが、前記ドナー候補からの H C T が H C T レシピエントにおいて G V H D を引き起こすリスクが高いこと又は低いことを示し、

陰性 G V H D 予測遺伝子の発現量が陰性 G V H D 予測遺伝子の基準発現量よりも大きい小さいかということが、前記ドナー候補からの H C T が H C T レシピエントにおいて G V H D を引き起こすリスクが低いこと又は高いことをそれぞれ示し、又は、

陽性及び / 又は陰性 G V H D 予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きい小さいかということが、前記ドナー候補からの H C T が H C T レシピエントにおいて G V H D を引き起こすリスクが高いこと又は低いことを示し

陽性及び / 又は陰性 G V H D 予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きい小さいかということが、前記ドナー候補からの H C T が H C T レシピエントにおいて G V H D を引き起こすリスクが低いこと又は高いことをそれぞれ示す、

方法。

【請求項 3】

ドナー候補からの造血細胞移植 (H C T) が H C T レシピエントにおいて移植片対宿主病 (G V H D) を引き起こすリスクを分類する方法であって、

a) 前記 H C T ドナー候補の C D 4 陽性 T 細胞又は C D 8 陽性 T 細胞において、表 1 (R N A 1 5 3 8)、表 2、表 2 A (R N A 1 4 3)、表 2 B (R N A 1 9 2)、表 3、表 1 3 (S G 1 7 5)、表 1 5 (S G 1 2 8) 又は表 1 8 (S G 6 4) に挙げられた遺伝子から選択された複数の陽性又は陰性 G V H D 予測遺伝子又はこれらの多型の発現を測定する工程と、

b) 工程 a) で測定された前記発現に基づいて、陽性又は陰性 G V H D 予測遺伝子の発現量を得るか、又は、工程 a) で測定された前記発現に基づいて、陽性及び / 又は陰性 G V H D 予測遺伝子の組み合わせの発現量の線形結合又は非線型結合を得る工程と、

c) 前記陽性又は陰性 G V H D 予測遺伝子の発現量と陽性又は陰性 G V H D 予測遺伝子の既定の基準発現量とを比較するか、又は、前記陽性及び / 又は陰性 G V H D 予測遺伝子の組み合わせの発現量の線形結合又は非線型結合と陽性及び / 又は陰性 G V H D 予測遺伝子の線形結合又は非線型結合の既定の基準値とを比較する工程と、

d) 前記ドナー候補からの前記 H C T が H C T レシピエントにおいて G V H D を引き起こすリスクが高いか低いかを示す発現量、陽性又は陰性 G V H D 予測遺伝子の総数又は同一性、又は陽性及び / 又は陰性 G V H D 予測遺伝子の組み合わせの評価に基づいて、前記 H C T ドナー候補が G V H D を引き起こすか又は引き起こさないリスクを分類する工程と、

陽性 G V H D 予測遺伝子の発現量が陽性 G V H D 予測遺伝子の既定の基準発現量よりも大きい小さいかということが、前記ドナー候補からの H C T が H C T レシピエントにおいて G V H D を引き起こすリスクが高いこと又は低いことを示し、又は

陰性 G V H D 予測遺伝子の発現量が陰性 G V H D 予測遺伝子の基準発現量よりも大きい小さいかということが、前記ドナー候補からの H C T が H C T レシピエントにおいて G

10

20

30

40

50

VHDを引き起こすリスクが低いこと又は高いことをそれぞれ示し、又は

陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きい小さいかということが、前記ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示し、又は

陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きい小さいかということが、前記ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示す、

方法。

10

【請求項4】

実際のドナー又はドナー候補のHCTがHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こす又は引き起こさない可能性又はリスクに基づいて各々にスコアが割り当てられた複数の実際のHCTドナー又はHCTドナー候補を含むデータベース又は組織的構成の生成方法であって、

a)実際のドナー又はドナー候補のCD4陽性T細胞又はCD8陽性T細胞において、表1(RNA1538)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)又は表18(SG64)に挙げられた一又は複数の陽性又は陰性GVHD予測遺伝子又はこれらの多型の発現を測定する工程と、

b)工程a)で測定された前記発現に基づいて陽性又は陰性GVHD予測遺伝子の発現量を得るか、又は、工程a)で測定された前記発現に基づいて、陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合を得る工程と、

20

c)前記陽性又は陰性GVHD予測遺伝子の発現量と陽性又は陰性GVHD予測遺伝子の既定の基準発現量とを比較するか、又は、前記陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合と陽性及び/又は陰性GVHD予測遺伝子の線形結合又は非線型結合の既定の基準値とを比較する工程と、

を備え、

陽性GVHD予測遺伝子の発現量が陽性GVHD予測遺伝子の既定の基準発現量よりも大きい小さいかということが、前記ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示し、又は

30

陰性GVHD予測遺伝子の発現量が陰性GVHD予測遺伝子の基準発現量よりも大きい小さいかということが、前記ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示し、又は

陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きい小さいかということが、前記ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示し、又は

陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きい小さいかということが、前記ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示し、さらに、

40

d)前記ドナー候補からの前記HCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いか低いかを示す発現量、陽性又は陰性GVHD予測遺伝子の総数又は同一性、又は陽性及び/又は陰性GVHD予測遺伝子の組み合わせの評価に基づいて、実際のドナー又はドナー候補のHCTがHCTレシピエントにおいてGVHDを引き起こす可能性又はリスクを示すスコアを割り当てる工程と、

e)前記実際のHCTドナー又はHCTドナー候補のスコアを記録又は記憶する工程と、

f)一又は複数の追加的な実際のHCTドナー又はHCTドナー候補について、工程a)、工程b)、工程c)、工程d)及び工程e)を繰り返す工程と、

を備えることにより、前記実際のドナー又はドナー候補のHCTがHCTレシピエント

50

において移植片対宿主病（GVHD）を引き起こす又は引き起こさない可能性又はリスクに基づいて各々にスコアが割り当てられた実際のHCTドナー又はHCTドナー候補を含むデータベース又は組織的構成を生成する方法。

【請求項5】

複数の陽性又は陰性予測遺伝子の発現を測定して前記複数の陽性又は陰性予測遺伝子の発現量を得る工程と、前記陽性又は陰性予測遺伝子の前記発現量と当該対応する陽性又は陰性予測遺伝子の既定の基準発現量とを比較する工程と、を含む請求項1、2又は4に記載の方法。

【請求項6】

前記陽性又は陰性予測遺伝子が、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、又は表18（SG64）に挙げられた一又は複数の陽性又は陰性予測遺伝子又はこれらの多型から選択される、請求項1ないし請求項4のいずれか1項に記載の方法。

10

【請求項7】

測定された前記複数の陽性又は陰性予測遺伝子が、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、又はこれ以上の陽性又は陰性予測遺伝子。である請求項6に記載の方法。

【請求項8】

前記陽性予測遺伝子の前記既定の基準発現量よりも大きい前記陽性予測遺伝子の発現量が、GVHDを引き起こすことが分かっている単数又は複数のHCTドナーにおける前記陽性予測遺伝子の発現と相関がある、請求項1ないし請求項4のいずれか1項に記載の方法。

20

【請求項9】

前記陰性予測遺伝子の前記既定の基準発現量よりも大きい前記陰性予測遺伝子の発現量が、GVHDを引き起こさないことが分かっている単数又は複数のHCTドナーにおける前記陰性予測遺伝子の発現と相関がある、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項10】

前記陽性予測遺伝子の前記既定の基準発現量が、GVHDを引き起こす2以上のHCTドナー及びGVHDを引き起こさない2以上のHCTドナーの前記陽性予測遺伝子の発現レベルの平均値又は中央値の間にある、請求項1ないし請求項4のいずれか1項に記載の方法。

30

【請求項11】

前記陰性予測遺伝子の前記既定の基準発現量が、GVHDを引き起こす2以上のHCTドナー及びGVHDを引き起こさない2以上のHCTドナーの前記陰性予測遺伝子の発現レベルの平均値又は中央値の間にある、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項12】

前記陽性又は陰性予測遺伝子の前記既定の基準発現量が、GVHDを引き起こす5以上のHCTドナー及びGVHDを引き起こさない5以上のHCTドナーの前記陽性又は陰性予測遺伝子の発現レベルの平均値又は中央値の間にある、請求項1ないし請求項4のいずれか1項に記載の方法。

40

【請求項13】

前記陽性予測遺伝子の前記既定の基準発現量が、GVHDを引き起こすことが知られている複数のHCTドナーの遺伝子発現の中央値又は平均値と、GVHDを引き起こさないことが知られている複数のHCTドナーの遺伝子発現の中央値又は平均値との間にある、請求項1ないし請求項4のいずれか1項に記載の方法

【請求項14】

50

前記陰性予測遺伝子の前記既定の基準発現量が、GVHDを引き起こすことが知られている複数のHCTドナーの遺伝子発現の中央値又は平均値と、GVHDを引き起こさないことが知られている複数のHCTドナーの遺伝子発現の中央値又は平均値との間にある、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項15】

前記陽性予測遺伝子の前記既定の基準発現量が、GVHDを引き起こす一又は複数のドナーの前記陽性予測遺伝子の発現レベルと、GVHDを引き起こさない一又は複数のドナーの前記陽性予測遺伝子の発現レベルとの間にある中間値であり、

前記中間値よりも大きい前記陽性予測遺伝子の発現量が、前記ドナー候補からの前記HCTが、移植片対宿主病(GVHD)を引き起こすリスクが高いことを示す。請求項1ないし請求項4のいずれか1項に記載の方法。

10

【請求項16】

前記陰性予測遺伝子の前記既定の基準発現量は、GVHDを引き起こさない一又は複数のドナーの前記陰性予測遺伝子の発現レベルとGVHDを引き起こす一又は複数のドナーの前記陰性予測遺伝子の発現レベルとの間にある中間値であり、

前記中間値よりも大きい前記陰性予測遺伝子の発現量 indicates that 前記ドナー候補からの前記HCTが、移植片対宿主病(GVHD)を引き起こすリスクが低いことを示す、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項17】

前記中間値には0.5の値が割り当てられ、0.5よりも大きい前記一又は複数の陰性予測遺伝子の発現量が、前記ドナー候補からの前記HCTが、移植片対宿主病(GVHD)を引き起こすリスクが低いことを示す、請求項10ないし請求項16のいずれか1項に記載の方法。

20

【請求項18】

前記中間値には0.5の値が割り当てられ、0.55以上の前記一又は複数の陰性予測遺伝子の発現量が、前記ドナー候補からの前記HCTが、移植片対宿主病(GVHD)を引き起こすリスクが引くことを示す、請求項10ないし請求項16のいずれか1項に記載の方法。

【請求項19】

前記中間値には0.5の値が割り当てられ、0.60以上の前記一又は複数の陰性予測遺伝子の発現量が、前記ドナー候補からの前記HCTが、移植片対宿主病(GVHD)を引き起こすリスクが低いことを示す、請求項10ないし請求項16のいずれか1項に記載の方法。

30

【請求項20】

前記中間値には0.5の値が割り当てられ、0.5より大きい前記一又は複数の陽性予測遺伝子の発現量が、前記ドナー候補からの前記HCTが、移植片対宿主病(GVHD)を引き起こすリスクが高いことを示す、請求項10ないし請求項16のいずれか1項に記載の方法。

【請求項21】

前記中間値には0.5の値が割り当てられ、0.55以上の前記一又は複数の陽性予測遺伝子の発現量が、前記ドナー候補からの前記HCTが、移植片対宿主病(GVHD)を引き起こすリスクが高いことを示す、請求項10ないし請求項16のいずれか1項に記載の方法。

40

【請求項22】

前記中間値には0.5の値が割り当てられ、0.60以上の前記一又は複数の陽性予測遺伝子の発現量が、前記ドナー候補からの前記HCTが、移植片対宿主病(GVHD)を引き起こすリスクが高いことを示す、請求項10ないし請求項16のいずれか1項に記載の方法。

【請求項23】

前記陽性又は陰性予測遺伝子の前記既定の基準発現量が、GVHDを引き起こすことが

50

分かっている HCT ドナー 及び GVHD を引き起こさないことが分かっている HCT ドナーにおける遺伝子発現の判別分析によって決定される値である、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

【請求項 2 4】

前記陽性又は陰性予測遺伝子について得られた発現量は、前記陽性又は陰性予測遺伝子の前記発現量を前記陽性又は陰性予測遺伝子の前記既定の基準発現量と比較する前に、一又は複数の基準遺伝子の発現で調整又は正規化される、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

【請求項 2 5】

前記発現量は、前記陽性又は陰性予測遺伝子の遺伝子発現の 一又は複数の基準遺伝子に対する比で表され、レシオメトリック遺伝子対 (RGP) と呼ばれる、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

10

【請求項 2 6】

前記発現量は、前記陽性又は陰性予測遺伝子の遺伝子発現の基準遺伝子に対する比で表されてレシオメトリック遺伝子対 (RGP) と呼ばれ、また、式「N/D」で表され、ここで「N」は前記陽性又は陰性予測遺伝子の発現レベルであり、「D」は一又は複数の基準遺伝子の発現レベルである、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

【請求項 2 7】

前記分子値N又は分母値Dは、それぞれ、一又は複数の陽性又は陰性予測遺伝子の発現の平均値又は中央値、又は一又は複数の基準遺伝子である、請求項 2 6 の方法。

20

【請求項 2 8】

前記発現量は、前記陽性又は陰性予測遺伝子の遺伝子発現の 基準遺伝子に対する比で表されてレシオメトリック遺伝子対 (RGP) と呼ばれ、また、前記陽性又は陰性予測遺伝子の発現が \log_n で表されるときには、前記発現量は、式 $(\log_n X - \log_n Y)$ で表される請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

ここで、「X」は、前記陽性又は陰性予測遺伝子の前記発現レベル、「Y」は前記基準遺伝子の前記発現レベル、及び「n」は、2、10、e (自然対数の底) 又は任意の正の整数。

【請求項 2 9】

レシオメトリック遺伝子対 (RGP) を含む遺伝子の少なくとも 1 つが、表 2 A (RNA 143)、表 2 B (RNA 192)、表 3、表 13 (SG175)、表 1 5 (SG128)、又は表 1 8 (SG64) に記載されている遺伝子又はこれらの多型である請求項 2 5、2 6 又は 2 8 のいずれか 1 項に記載の方法。

30

【請求項 3 0】

前記基準遺伝子は、遺伝子発現の前記比を得るために用いられた前記陽性又は陰性予測遺伝子と異なる陽性又は陰性予測遺伝子を含む、請求項 2 4、2 5、2 6 又は 2 8 のいずれか 1 項に記載の方法。

【請求項 3 1】

前記レシオメトリック遺伝子対 (RGP) の少なくとも 1 つは、表 1 4 (RGP348) に記載された RGP から選択される、請求項 2 5、2 6 又は 2 8 のいずれか 1 項に記載の方法。

40

【請求項 3 2】

測定された遺伝子発現比の数は、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、又はそれ以上の遺伝子発現比である、請求項 2 5、2 6 又は 2 8 のいずれか 1 項に記載の方法。

【請求項 3 3】

前記比を含む前記陽性又は陰性予測遺伝子の少なくとも 1 つは、表 1 3 (SG175)、表 1 5 (SG128) 又は表 1 8 (SG64) に記載された from 一又は複数の単一遺伝子 (SG) から選択されるか、表 1 4 (RGP348) に記載されたレシオメトリ

50

ック遺伝子対 (R G P) から選択される、請求項 2 5、2 6 又は 2 8 のいずれか 1 項に記載の方法。

【請求項 3 4】

前記基準遺伝子は、その発現が構成的であり C D 4 陽性 T 細胞又は C D 8 陽性 T 細胞において比較的一定のレベルで発現する遺伝子、ハウスキーピング遺伝子、又は、その発現が前記数値を決定するために用いられない陽性又は陰性予測遺伝子である、請求項 2 4、2 5、2 6 又は 2 8 のいずれか 1 項に記載の方法。

【請求項 3 5】

前記ハウスキーピング遺伝子は、ベータアクチン (A C T B)、アルドラーゼ A (A L D O A)、乳酸脱水素酵素 A (L D H A)、ホスホグリセリン酸キナーゼ 1 (P G K 1)、トランスフェリン受容体 (T F R C)、ベータチューブリン (T U B B)、ベータチューブリン 2 A (T U B B 2 A)、チオレドキシン (T X N)、ユビキチン C (U B C)、ユビキチン活性化酵素 E 1 (U B E 1)、表 2 B (R N A 1 9 2、denoted HSK、配列番号：1 6 9 0 - 1 7 3 8) に記載された配列、又は、表 1 2 (H S K l i s t) に記載された配列から選択される、請求項 3 4 に記載の方法。

10

【請求項 3 6】

前記陽性又は陰性予測遺伝子は、表 1 3 (S G 1 7 5)、表 1 5 (S G 1 2 8) 又は表 1 8 (S G 6 4) から選択される、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

【請求項 3 7】

表 1 3 (S G 1 7 5)、表 1 5 (S G 1 2 8) 又は表 1 8 (S G 6 4) から選択される複数の陽性又は陰性予測遺伝子、又は、表 1 4 (R G P 3 4 8) に記載された R G P から選択される複数のレシオメトリック遺伝子対 (R G P) を含む請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

20

【請求項 3 8】

表 1 3 (S G 1 7 5)、表 1 5 (S G 1 2 8) 又は表 1 8 (S G 6 4) から選択される複数の陽性及び陰性の予測遺伝子、及び、表 1 4 (R G P 3 4 8) から選択される複数のレシオメトリック遺伝子対 (R G P) を備える、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

【請求項 3 9】

前記陽性又は陰性予測遺伝子の前記発現量に基づいてスコアを割り当てる工程をさらに含み、前記スコアは、前記 H C T ドナー候補 H C T レシピエントにおいて移植片対宿主病 (G V H D) を引き起こす又は引き起こさない可能性又はリスクを反映している、請求項 1 ないし 3 のいずれか 1 項に記載の方法。

30

【請求項 4 0】

陰性又は陽性予測遺伝子の複数の発現量が決定され、前記遺伝子の前記発現量がドナー候補又は実際のドナーが G V H D を引き起こすリスクを示すか又は引き起こさないリスクを示すかによって、各陰性又は陽性予測遺伝子に票が割り当てられ、G V H D を引き起こすこと又は引き起こさないことを示す又は示さない票の総数に基づいて前記ドナー候補又は実際のドナーにスコアが割り当てられ、前記スコアは、ドナー候補又は実際のドナーからの造血細胞移植 (H C T) が H C T レシピエントにおいて G V H D を引き起こすか又は引き起こさないリスクを反映している、請求項 1 ないし 3 のいずれか 1 項に記載の方法。

40

【請求項 4 1】

前記票の 5 0 % 以上が G V H D を引き起こすことを示している場合、前記スコアは、前記ドナー候補又は実際のドナーからの造血細胞移植 (H C T) が H C T レシピエントにおいて G V H D を引き起こすリスクが高いことを反映している、請求項 4 0 に記載の方法。

【請求項 4 2】

前記票の 5 0 % 以上が G V H D を引き起こさないことを示している場合、前記スコアは、前記ドナー候補又は実際のドナーからの造血細胞移植 (H C T) が H C T レシピエントにおいて G V H D を引き起こすリスクが低いことを反映している、請求項 4 0 に記載の方法。

50

【請求項 4 3】

前記票の少なくとも 55%、60%、65%、70%、75%、80%、85%、90%、95%、又はそれ以上が G V H D を引き起こすことを示している場合、前記スコアは、前記ドナー候補又は実際のドナーからの造血細胞移植 (HCT) が HCT レシピエントにおいて G V H D を引き起こすリスクが高いことを反映しており、又は、前記票の少なくとも 55%、60%、65%、70%、75%、80%、85%、90%、95%、又はそれ以上が G V H D を引き起こさないことを示している場合、前記スコアは、前記ドナー候補又は実際のドナーからの造血細胞移植 (HCT) が HCT レシピエントにおいて G V H D を引き起こすリスクが低いことを反映している、請求項 4 0 に記載の方法。

【請求項 4 4】

前記ドナー候補からの前記 HCT が G V H D を引き起こすリスクが高いことを示す陽性又は陰性予測遺伝子又は前記陽性及び / 又は陰性 G V H D 予測遺伝子の組み合わせの数が、前記ドナー候補からの前記 HCT が HCT レシピエントにおいて G V H D を引き起こすリスクが小さいことを示す陽性又は陰性予測遺伝子又は前記陽性及び / 又は陰性 G V H D 予測遺伝子の組み合わせの数よりも大きく、ドナー候補からの前記 HCT が HCT レシピエントにおいて G V H D を引き起こすリスクが高いことを予測又は決定する、請求項 1 ないし 3 のいずれか 1 項に記載の方法。

【請求項 4 5】

前記ドナー候補からの前記 HCT が G V H D を引き起こすリスクが低いことを示す陽性又は陰性予測遺伝子又は前記陽性及び / 又は陰性 G V H D 予測遺伝子の組み合わせの数が、前記ドナー候補からの前記 HCT が HCT レシピエントにおいて G V H D を引き起こすリスクが高いことを示す陽性又は陰性予測遺伝子又は前記陽性及び / 又は陰性 G V H D 予測遺伝子の組み合わせの数よりも大きく、ドナー候補からの前記 HCT が HCT レシピエントにおいて G V H D を引き起こすリスクが低いことを予測又は決定する、請求項 1 ないし 3 のいずれか 1 項に記載の方法。

【請求項 4 6】

ドナー候補からの造血細胞移植 (HCT) が HCT レシピエントにおいて移植片対宿主病 (G V H D) を引き起こすか又は引き起こさないリスクを予測又は決定するために用いられる前記陰性及び陽性予測遺伝子は、表 1 8 (VmodSG64) に記載された一又は複数の遺伝子を含む、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

【請求項 4 7】

前記陰性及び陽性予測遺伝子は、表 1 (RNA1538)、表 2、表 2 A (RNA143)、表 2 B (RNA192)、表 3、表 1 3 (SG175)、表 1 5 (SG128) 又は表 1 8 (SG64) に記載された 2 以上の遺伝子の複数のレシオメトリック遺伝子対 (RGP) を含む、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

【請求項 4 8】

ドナー候補からの造血細胞移植 (HCT) が HCT レシピエントにおいて移植片対宿主病 (G V H D) を引き起こすか又は引き起こさないリスクを予測又は決定するために用いられる前記レシオメトリック遺伝子対 (RGP) は、表 1 7 (VmodRGP100) に記載された一又は複数の遺伝子対 (RGP) を含む、請求項 4 5 に記載の方法。

【請求項 4 9】

前記陰性及び陽性予測遺伝子は、ドナー候補からの造血細胞移植 (HCT) が HCT レシピエントにおいて移植片対宿主病 (G V H D) を引き起こすか否かのリスクを予測又は決定するために用いられる単一遺伝子 (SG) 及びレシオメトリック遺伝子対 (RGP) の組み合わせを含み、前記組み合わせは、表 1 8 (VmodSG64) に記載された単一遺伝子 (SG) から選択される複数の遺伝子及び表 1 7 (VmodRGP100) に記載された R G P から選択される複数のレシオメトリック遺伝子対 (RGP) を含む、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

【請求項 5 0】

単一遺伝子 (SG) 及びレシオメトリック遺伝子対 (RGP) の前記組み合わせは、SG43RGP

10

20

30

40

50

46-GPperformance、SG42RGP21-GPminimalist、SG43RGP37-GPconnectivity、SG43RGP51-PRGPminranksort、SG43RGP55-PRGPmedranksort、SG43RGP36-RGPgreedysearch、又はSG21RGP28-RGPmaxgreedysearchに記載されたものであり、これらの組み合わせの各々は、表17及び表18において「x」で示されたSG及びRGPを含む、請求項49に記載の方法。

【請求項51】

前記ドナー候補及びHCTレシピエントは、10/10又は9/10のヒト白血球抗原(HLA)マーカー座適合を有する、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項52】

前記ドナー候補及びHCTレシピエントは、HLA-A、HLA-B、HLA-C、HLA-DRB1及びHLA-DQB1座の全てのHLAマーカー座適合を有するか、又は、HLA-A、HLA-B、HLA-C、HLA-DRB1又はHLA-DQB1座のうちの任意の4つのHLAマーカー座適合を有する、請求項1ないし請求項4のいずれか1項に記載の方法。

10

【請求項53】

前記HLAマーカー座適合は、血清学的に又はHLA遺伝子の配列分析によって決定される、請求項51又は52に記載の方法。

【請求項54】

前記ドナー候補及びHCTレシピエントは、兄弟ではないか、又は、血縁関係にない、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項55】

前記ドナー候補及びHCTレシピエントは、兄弟か又は血縁関係にある、請求項1ないし請求項4のいずれか1項に記載の方法。

20

【請求項56】

前記方法は、HCTレシピエントと10/10又は9/10HLAマーカー座適合を有することに基づいてGVHD陰性ドナーを特定することよりも優れている、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項57】

前記方法は、少なくとも60%の正確性でGVHD陰性HCTドナーを予測する、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項58】

前記方法は、少なくとも70%の正確性でGVHD陰性HCTドナーを予測する、請求項1ないし請求項4のいずれか1項に記載の方法。

30

【請求項59】

前記方法は、少なくとも80%の正確性でGVHD陰性HCTドナーを予測する、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項60】

GVHD陰性ドナーを予測する正確性は、HCTレシピエントとの10/10HLAマーカー座適合によって陰性に分類されたHCT候補ドナーのグループの中で、GVHD陰性ドナーを正しく特定する可能性又はリスクである、請求項57ないし請求項59のいずれか1項に記載の方法。

40

【請求項61】

前記方法は、GVHD陽性HCTドナーを少なくとも60%の正確性で予測する、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項62】

前記方法は、GVHD陽性HCTドナーを少なくとも70%の正確性で予測する、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項63】

前記方法は、GVHD陽性HCTドナーを少なくとも80%の正確性で予測する、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項64】

50

前記陽性又は陰性予測遺伝子の閾値は、前記HCTドナー候補がHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすリスクが高いことを予測又は決定するために、HCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすリスクが高いことを示さなければならない、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項65】

前記陽性又は陰性予測遺伝子の閾値は、前記HCTドナー候補がHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすリスクが低いことを予測又は決定するために、HCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすリスクが低いことを示さなければならない、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項66】

HCTドナーがHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすリスクが高いこと又は低いことを示す陽性又は陰性予測遺伝子の数に基づいて、スコアを割り当てる工程をさらに備え、前記スコアは、前記HCTドナー候補がHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こす可能性又はリスクを反映している、請求項64又は65に記載の方法。

【請求項67】

前記HCTドナー候補がHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすリスクが高いことを予測又は決定するために、前記陽性又は陰性予測遺伝子の過半数が、HCTレシピエントにおいて移植片対宿主病(GVHD)を引き越すリスクが高いことを示していなければならない、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項68】

前記HCTドナー候補がHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすリスクが低いことを予測又は決定するために、前記陽性又は陰性予測遺伝子の過半数が、HCTレシピエントにおいて移植片対宿主病(GVHD)を引き越すリスクが低いことを示していなければならない、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項69】

前記HCTドナー候補がHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすリスクが高いことを予測又は決定するために、前記陽性又は陰性予測遺伝子の少なくとも66%が、HCTレシピエントにおいて移植片対宿主病(GVHD)を引き越すリスクが高いことを示していなければならない、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項70】

前記HCTドナー候補がHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすリスクが低いことを予測又は決定するために、前記陽性又は陰性予測遺伝子の少なくとも66%が、HCTレシピエントにおいて移植片対宿主病(GVHD)を引き越すリスクが低いことを示していなければならない、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項71】

前記HCTドナー候補がHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすリスクが低いことを予測又は決定するために、前記陽性又は陰性予測遺伝子の少なくとも75%が、HCTレシピエントにおいて移植片対宿主病(GVHD)を引き越すリスクが低いことを示していなければならない、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項72】

前記検体が、プライマー対、オリゴ-又はポリ-ヌクレオチドプローブ、又は抗体もしくはその抗原結合断片を含む請求項2に記載の方法。

【請求項73】

前記測定する工程が、前記陽性又は陰性予測遺伝子もしくはこれらの多型のうちの1つ

10

20

30

40

50

から生成されるRNA 転写産物、又は、前記陽性又は陰性予測遺伝子もしくはこれらの多型の前記RNA 転写産物に由来するcDNAへの、オリゴ - 又はポリ - ヌクレオチドプローブを用いたハイブリダイゼーションを含む、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

【請求項 7 4】

前記測定する工程が、表 1 (RNA 1538)、表 2、表 2 A (RNA 143)、表 2 B (RNA 192)、表 3、表 1 3 (SG175)、表 1 5 (SG128) 又は表 1 8 (SG64) に記載されている遺伝子の転写産物とハイブリダイズするオリゴ - 又はポリ - ヌクレオチドプローブ又はプライマーを用いたハイブリダイゼーションを含む、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

10

【請求項 7 5】

前記測定する工程は、表 1 (RNA 1538)、表 2、表 2 A (RNA 143)、表 2 B (RNA 192)、表 3、表 1 3 (SG175)、表 1 5 (SG128) 又は表 1 8 (SG64) に記載されているオリゴ - 又はポリ - ヌクレオチドプローブ又はプライマーを用いたハイブリダイゼーションを含む、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

【請求項 7 6】

前記測定する工程は、プライマー対のハイブリダイゼーション、及び、その後に行われ、前記陽性又は陰性予測遺伝子又はこれらの多型から生成される前記陽性又は陰性予測遺伝子のRNA 転写産物に由来するcDNAの増幅を含む、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

20

【請求項 7 7】

前記プライマー対は、表 2 B (RNA 192) に記載された配列における対、又は、表 1 (RNA 1538)、表 2、表 2 A (RNA 143)、表 2 B (RNA 192)、表 3、表 1 3 (SG175)、表 1 5 (SG128) 又は表 1 8 (SG64) に記載された遺伝子の転写産物とハイブリダイズするプライマー対である、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

【請求項 7 8】

前記測定する工程は、cDNAを生成して一又は複数の陽性又は陰性予測遺伝子の発現レベルを決定するためにRNA 転写産物の逆転写を含む、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

30

【請求項 7 9】

前記 C D 4 陽性 T 細胞又は C D 8 陽性 T 細胞は、ドナー候補又は実際のドナーの血液からのもの、又は、当該血液に存在するものである、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

【請求項 8 0】

前記 G V H D は、グループ 1、2、3、4、5、又は 6 のいずれかに分類される、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

【請求項 8 1】

前記 G V H D は、慢性 G V H D を伴うもしくは伴わないグレード I、II、III 又は IV の急性 G V H D、又は、急性 G V H D を伴わない慢性 G V H D に分類される、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

40

【請求項 8 2】

H C T レシピエントについて移植片対宿主病 (G V H D) を引き起こすリスクが低い H C T ドナーを選択する工程をさらに含む、請求項 1 ないし 3 のいずれか 1 項に記載の方法。

【請求項 8 3】

H C T ドナー候補又は実際の H C T ドナーの前記遺伝子発現プロファイル、又は、G V H D を引き起こすか否かについてのスコアもしくはリスクプロファイルは、コンピュータ読み取り可能な媒体、電子記憶媒体、又はデータベースもしくはそれ以外組織的構成に記録又は記憶される、請求項 1 ないし請求項 4 のいずれか 1 項に記載の方法。

【請求項 8 4】

50

移植片対宿主病（GVHD）を引き起こす又は引き起こさないリスクが低い又は高いHCT候補ドナーが特定される、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項85】

ドナー候補又は実際のドナーからのHCTがHCTレシピエントにおいてGVHDを引き起こすか否かのリスク又はスコアは、又はコンピュータ読み取り可能な媒体に記録又は記憶される、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項86】

前記ドナー候補からの前記HCTがHCTレシピエントにおいて移植片対宿主病（GVHD）を引き起こすか否かについてのリスク又はスコアのレポートを作成する工程をさらに含む、請求項1ないし請求項4のいずれか1項に記載の方法。

10

【請求項87】

前記陽性又は陰性予測遺伝子又はハウスキーピング遺伝子の発現は、RT-PCRによって決定される、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項88】

2以上のプライマー対を含むキットであって、
各プライマー対は互いに逆方向に配向しており、
前記プライマー対の第1の部分は、前記陽性又は陰性予測遺伝子の1つから生成されるRNA又はcDNAにハイブリダイズし、第2の部分は、表1（RNA1538）、表2、表2A、表2B（RNA192）、表3及び/又は表12に記載されたハウスキーピング遺伝子にハイブリダイズする、
キット。

20

【請求項89】

互いに逆方向に配向した5以上のプライマー対を含み、当該5つのプライマー対の各々は、表1（RNA1538）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）又は表18（SG64）に記載された陽性又は陰性予測遺伝子又はこれらの多型のRNA又はcDNAにハイブリダイズする、請求項88に記載のキット。

【請求項90】

互いに逆方向に配向した10以上のプライマー対を含み、当該5つのプライマー対の各々は、表1（RNA1538）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）又は表18（SG64）に記載された陽性又は陰性予測遺伝子又はこれらの多型のRNA又はcDNAにハイブリダイズする、請求項88に記載のキット。

30

【請求項91】

互いに逆方向に配向した20以上のプライマー対を含み、当該5つのプライマー対の各々は、表1（RNA1538）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）又は表18（SG64）に記載された陽性又は陰性予測遺伝子又はこれらの多型のRNA又はcDNAにハイブリダイズする、請求項88に記載のキット。

【請求項92】

前記プライマー対の1つによって増幅された核酸配列にハイブリダイズするプローブを含む、請求項88に記載のキット。

40

【請求項93】

前記プライマー対の各々は、支持体又は基板に固定されていない、請求項88に記載のキット。

【請求項94】

一又は複数の核酸プローブを含み、
前記一又は複数のプローブは、表1（RNA1538）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）又は表18（SG64）に記載された前記陽性又は陰性予測遺伝子の一又は複数又はこ

50

これらの多型のRNA又はcDNAにハイブリダイズする、
キット。

【請求項95】

表1(RNA 1538)、表2、表2A(RNA 143)、表2B(RNA 192)、表3、表13(SG175)、表15(SG128)又は表18(SG64)に記載された前記陽性又は陰性予測遺伝子のうちの5つ又はこれらの多型のRNA又はcDNAにハイブリダイズする5以上のプローブを含む、請求項94に記載のキット。

【請求項96】

表1(RNA 1538)、表2、表2A(RNA 143)、表2B(RNA 192)、表3、表13(SG175)、表15(SG128)又は表18(SG64)に記載された前記陽性又は陰性予測遺伝子のうちの10個又はこれらの多型のRNA又はcDNAにハイブリダイズする10以上のプローブを含む、請求項94に記載のキット。

10

【請求項97】

表1(RNA 1538)、表2、表2A(RNA 143)、表2B(RNA 192)、表3、表13(SG175)、表15(SG128)又は表18(SG64)に記載された前記陽性又は陰性予測遺伝子のうちの20個又はこれらの多型のRNA又はcDNAにハイブリダイズする20以上のプローブを含む、請求項94に記載のキット。

【請求項98】

複数の実際のHCTドナー又はHCTドナー候補の2以上の陽性又は陰性予測遺伝子の遺伝子発現プロファイルを含むデータベース又は組織的構成であって、

20

前記2以上の陽性又は陰性予測遺伝子は、表1(RNA 1538)、表2、表2A(RNA 143)、表2B(RNA 192)、表3、表13(SG175)、表15(SG128)又は表18(SG64)に記載されたもの又はこれらの多型であり、

前記データベース又は組織的構成は、前記遺伝子発現プロファイルを、前記実際のHCTドナー又はHCTドナー候補の各々に関連づける、

データベース又は組織的構成。

【請求項99】

前記データベースは、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、又はこれ以上の陽性又は陰性予測遺伝子の発現プロファイルを含み、及び、任意に一又は複数のハウスキーピング遺伝子の発現プロファイルを含む、請求項98に記載のデータベース又は組織的構成。

30

【請求項100】

前記実際のドナー又はドナー候補からのHCTがHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすリスクが高いか低いか特定される、請求項98に記載のデータベース又は組織的構成。

【請求項101】

前記陽性又は陰性予測遺伝子の発現は、実際のドナー又はドナー候補のCD4陽性T細胞又はCD8陽性T細胞を含む生体サンプルからのものである、請求項98に記載のデータベース又は組織的構成。

40

【請求項102】

前記データベースは、プロセッサと動作可能に接続されており、前記プロセッサは、データ入力モジュール又はデータクエリモジュールを備える、請求項98に記載のデータベース又は組織的構成。

【請求項103】

実際のHCTドナー又はHCTドナー候補の一又は複数には、各々のHCTがHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすか否かの可能性又はリスクに基づいてスコアが割り当てられている、請求項98に記載のデータベース又は組織的構成。

50

【請求項104】

2以上のプライマー対を含むプライマーのアレイであって、各プライマー対は、互いに逆方向に配向しており、前記プライマー対の各々は、表1(RNA 1538)、表2、表2A(RNA 143)、表2B(RNA 192)、表3、表13(SG175)、表15(SG128)又は表18(SG64)に記載されている陽性又は陰性予測遺伝子の1つ又はこれらの多型から生成されるRNA又はcDNAにハイブリダイズし、各プライマー対は、支持体又は基板に付着又は固定されている、アレイ。

【請求項105】

プローブのアレイであって、各プローブは、表1(RNA 1538)、表2、表2A(RNA 143)、表2B(RNA 192)、表3、表13(SG175)、表15(SG128)又は表18(SG64)に記載された陽性又は陰性予測遺伝子又はこれらの多型のRNA又はcDNAにハイブリダイズし、各プローブは支持体又は基板に付着又は固定されている、アレイ。

10

【請求項106】

ハウスキーピング遺伝子から生成されたRNA又はcDNAにハイブリダイズするプライマー対又はプローブをさらに含む、請求項104又は請求項105に記載のアレイ。

【請求項107】

各プライマー対又はプローブは、前記支持体又は基板上に、既知の位置又はアドレスを有する、請求項104又は請求項105に記載のアレイ。

【請求項108】

前記プライマー対又はプローブの全てが、表1(RNA 1538)、表2、表2A(RNA 143)、表2B(RNA 192)、表3、表13(SG175)、表15(SG128)又は表18(SG64)に記載された陽性又は陰性予測遺伝子又はこれらの多型のRNA又はcDNAにハイブリダイズする、請求項104又は請求項105に記載のアレイ。

20

【請求項109】

表1(RNA 1538)、表2、表2A、表2B(RNA 192)及び/又は表3に記載された5個、10個、20個、30個またはそれ以上の陽性又は陰性予測遺伝子又はこれらの多型のRNA又はcDNAにハイブリダイズするプライマー対又はプローブを含む、請求項104又は請求項105に記載のアレイ。

30

【請求項110】

前記総プライマー対又はプローブは、20,000未満、15,000未満、10,000未満、5,000未満、2,500未満、2,000未満、1,500未満、1,000未満、500未満、400未満、300未満、200未満、100未満、50未満、又は25未満のプライマー対又はプローブを含む、請求項104又は請求項105に記載のアレイ。

【請求項111】

前記支持体又は基板は、マルチウェルフォーマットを有する、請求項104又は請求項105に記載のアレイ。

【請求項112】

前記支持体又は基板は、マルチウェルプレートを備える、請求項104又は請求項105に記載のアレイ。

40

【請求項113】

前記プライマー対の1つによって増幅される核酸配列にハイブリダイズするプローブをさらに含む、請求項104に記載のアレイ。

【請求項114】

前記CD4陽性T細胞又はCD8陽性T細胞が生体サンプルを含む、請求項1ないし請求項4のいずれか1項に記載の方法。

【請求項115】

前記オリゴ-又はポリ-ヌクレオチドプローブ又はプライマーが、約5~10、10~20、20~30、30~40、40~50、50~60、60~70、70~80、9

50

0 ~ 100、100 ~ 150、150 ~ 200、200 ~ 250、250 ~ 300、300 ~ 400、400 ~ 500、500 ~ 000、又は1000 ~ 2000のヌクレオチドの長さを有する、請求項72ないし請求項77のいずれか1項に記載の方法、請求項88から請求項94のいずれか1項に記載のキット、又は請求項104又は105に記載のアレイ。

【発明の詳細な説明】

【技術分野】

【0001】

本出願は、2011年6月20日に出願された特許出願第61/498,965号及び2010年10月29日に出願された特許出願第61/408,491号に基づく優先権を主張し、これらの出願の内容は全体として参照により本明細書に明示的に組み込まれる。

10

【0002】

本発明は、ドナーからの造血細胞移植（HCT）が、HCTレシピエントに移植片対宿主病（GVHD）を引き起こすリスクの予測又は決定に関する。本発明はまた、ドナー候補のHCTを、HCTレシピエントにおいてGVHDを引き起こすリスクに従って分類することに関する。本発明はさらに、HCTレシピエントにおいてGVHDを引き起こすリスクに従って一又は複数のドナー候補のHCTを分類又は採点した組織的構成（例えば、データベース）及び組織的構成（例えば、データベース）の生成方法に関する。本発明はさらにまた、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクを予測又は決定するために有用なキット及びアレイ、及び、HCTレシピエントにおいてGVHDを引き起こすリスクに従ってかかるドナーを分類又は採点するために有用なに関する。

20

【背景技術】

【0003】

造血細胞移植（HCT：Hematopoietic cell transplantation又はHematopoietic cell transplant）（比較的新しい用語である）又は骨髄移植（BMT）（より一般的な用語である）は、急性リンパ性白血病、急性骨髄性白血病、慢性骨髄性白血病、慢性リンパ球性白血病、髄異形成症候群（それぞれALL、AML、CML、CLL、及びMDS）等の様々な種類の血液がん及び血液疾患にとって、しばしば延命効果があるし、根治療法となることもある。HCTをより幅広く思い通りに利用することを阻害する主な要因は、HCTレシピエントにおけるGVHD（移植片対宿主病）のリスクである。

30

【0004】

米国で年間10000例のHCTのうち（これは保守的な数字であり、より正確には年間12000以上に近い数である。）大部分（約75%）は、HCTレシピエントの非血縁ドナーを用いて実施されている。HCTの医療においては、平均で4人に1人のHCT候補がドナーとして適した兄弟やそれ以外の血縁者を持たず、それゆえに米国や欧州の多くで行われているHCTの約4件に3件は患者と親族関係にないドナーが関わるものとなっていることがはっきりしている。これらの約7500件の非血縁ドナー移植のうち、約5600件（約75%）は、いわゆる10/10適合（10/10 major alleles）（正式名称はHLA-A、HLA-B、HLA-C、HLA-DRB1、HLA-DQB1）のHLA適合のドナー・レシピエントペアを用いている。

40

【0005】

移植片対宿主病（GVHD）は、ドナーの移植された造血幹細胞（移植片）から生成される免疫系T細胞によるHCTレシピエントの組織及び臓器（宿主）の深刻で致命的な拒絶反応となることがある（Bhushan & Collins.2003; Ferrara, et al.,2005）。10/10適合（HLA-A、HLA-B、HLA-C、HLA-DRB1、HLA-DQB1）でHCTドナーとHCTレシピエントとの間のHLA（ヒト白血球抗原）適合度が高い場合であっても、移植レシピエントの50%から60%にGVHDが起こる。これは、血縁ドナーであっても非血縁ドナーであっても変わらない。したがって、ドナーからの造血細胞移植（HCT）がHCTレシピエントにおいてGVHDを引き起こすリスクを予測し決定することが必要とされている

50

。また、HCTレシピエントにおいてGVHDを引き起こすリスクが低くGVHDを減少させることができるドナーを特定することが必要とされている。本発明は、このニーズを満たし、追加的な効果も提供する。

【発明の概要】

【0006】

本発明は、米国の57の移植センターにおいて、レシピエント（すなわち患者）と血縁関係にないドナーを用いて行われた180例のHCTから得られたサンプルの分析に、少なくとも部分的に基づくものである。遺伝子発現分析は、末梢血液に由来する移植前ドナーのCD4陽性T細胞における分子RNAマーカープロファイルを明らかにした。この分子RNAマーカープロファイルは、HCTレシピエントにおける急性の又は慢性のGVHD転帰を高度に予測するものである。概して、当該データは、様々な遺伝子マーカーの組み合わせを用い、急性及び慢性のGVHDの異なる程度の転帰予測を含む（表20参照）複数の多遺伝子予測モデルを明らかにする。この多遺伝子予測モデルは、（1）陰性予測値（Negative Predictive Value）：GVHD陰性として予測されるHCTの比で正しく予測されるもの。全てのGNOS（GVHD陰性転帰スコア）閾値にわたって平均で82%、GNOS閾値0.50については平均で78%（2）特異度（Specificities）：真陰性率、すなわち正しくGVHD陰性と予測されたGVHD陰性HCTの割合。全てのGNOS閾値にわたって平均で50%、GNOS閾値0.50については平均で78%、及び（3）感度（Sensitivities）：真陽性率、すなわち正しくGVHD陽性と予測されたGVHD陽性HCTの割合。全てのGNOS閾値にわたって平均で88%、GNOS閾値0.50については平均で78%。具体的には、最も性能がよい多遺伝子予測モデルとしてSG43RGP36-RGPgreedysearch、Gneg対Gag3分割（「no GVHD」対「グレードIII又はIVの急性GVHD」）、GNOS閾値を0.55としたとき、観測された陰性予測値は92%、特異度は80%、及び感度は94%である（表20参照）。正確で、ドナーベースの、移植前GVHD転帰予測は、移植センターサンプル出所の変化、血液病疾患転帰の医師による分類、及びHCTドナーが骨髄とPBMC（末梢血単核細胞）のいずれの形式かに関してロバスト性が高い。HLA適合に対する随意的追加的な処置として、HCTレシピエントと血縁の及び非血縁のドナーにおけるドナーT細胞のRNA発現測定に基づいてGVHDを信頼性高く予測でき、GVHDリスクが低いドナーからのHCTを選択できるので、HCTレシピエントにおけるGVHDの発生を大きく減少させ、また、HCTレシピエントにおけるGVHDの強度/重症度を減少させることができる。

【0007】

このように、本発明によれば、実際のドナー又はドナー候補からの造血細胞移植（HCT）がHCTレシピエントにおいて移植片対宿主病（GVHD）を引き起こす（又は引き起こさない）リスクを予測又は決定する方法が提供される。一実施形態における方法は、ドナー候補者のCD4陽性T細胞又はCD8陽性T細胞において、表1（RNA 1538）、表2、表2A（RNA 143）、表2B（RNA 192）、表3、表13（SG175）、表15（SG128）又は表18（SG64）から選択される一又は複数の陽性又は陰性のGVHD予測遺伝子（GVHD predictor遺伝子）又は陽性及び/もしくは陰性GVHD予測遺伝子の組み合わせ、又はこれらの多型の発現を測定することを含む。遺伝子発現レベルに基づいて、陽性又は陰性GVHD予測遺伝子の発現量が得られる。これとは代替的又は追加的に、測定された発現レベルに基づく陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線形結合が得られる。陽性又は陰性GVHD予測遺伝子の発現量と陽性又は陰性GVHD予測遺伝子の発現量の既定の基準値との比較、又は、陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線形結合と陽性及び/又は陰性GVHD予測遺伝子の線形結合又は非線形結合の既定の基準値との比較が実行される。この比較に基づいて、（1）陽性GVHD予測遺伝子の発現量が陽性GVHD予測遺伝子の既定の基準発現量よりも大きい小さいかということが、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示し、（2）陰性GVHD予測遺伝子の発現量が陰性GVHD予測遺伝子の既定の基準発現量よりも大きい小さいかということが、ドナー候補からのHCTがHCTレシピエントにおいてGV

10

20

30

40

50

H Dを引き起こすリスクが低いこと又は高いことをそれぞれ示し、(3)陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きい小さいかということがドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示し、(4)陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きい小さいかということが、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示す。ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示す発現量の比較の評価、陽性又は陰性GVHD予測遺伝子総数又は同一性、又は、陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合の比較に基づいて、ドナー候補からのHCTがHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすか否かのリスク又は可能性が予測及び/又は決定される。

10

20

30

40

50

【0008】

本発明によれば、実際のドナー又はドナー候補からのHCTがHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こす(又は引き起こさない)リスクを予測又は決定する方法が提供される。一実施形態における方法は、ドナー候補のCD4陽性T細胞又はCD8陽性T細胞又はCD4陽性T細胞又はCD8陽性T細胞によって発現された核酸もしくはタンパク質を、表1(RNA 1538)、表2、表2A(RNA 143)、表2B(RNA 192)、表3、表13(SG175)、表15(SG128)又は表18(SG64)に挙げられている一又は複数の陽性又は陰性GVHD予測遺伝子又はこれらの多型の発現を検出する検体と接触させること、及び、CD4陽性T細胞又はCD8陽性T細胞中の一又は複数の陽性又は陰性GVHD予測遺伝子の発現を測定して陽性又は陰性GVHD予測遺伝子の発現量を得ること、又は、陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現を測定して当該陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合(linear or non-linear combinations)を得ることを含む。当該陽性又は陰性GVHD予測遺伝子の発現量と当該陽性又は陰性GVHD予測遺伝子の既定の基準発現量との比較、又は、陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合と、陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合の既定の基準値との比較が行われる。この比較に基づいて、(1)陽性GVHD予測遺伝子の既定の基準発現量よりも大きい又は小さい陽性GVHD予測遺伝子の発現量が、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことをそれぞれ示し、(2)陰性GVHD予測遺伝子の既定の基準発現量よりも大きい又は小さい当該陰性GVHD予測遺伝子の発現量が、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示し、(3)陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きい小さいかということが、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことをそれぞれ示し、(4)陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きい小さいかということが、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示す。発現量の比較、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示す陽性又は陰性GVHD予測遺伝子の総数もしくは同一性、又は、陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合の比較の評価に基づいて、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクを予測し決定することができる。

【0009】

本発明によれば、HCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こす(又は引き起こさない)リスクに関して、実際のドナー又はドナー候補からの造血細胞移

植（HCT）を分類する方法がさらに提供される。一実施形態における方法は、HCTドナー候補のCD4陽性T細胞又はCD8陽性T細胞において、表1（RNA 1538）、表2、表2A（RNA 143）、表2B（RNA 192）、表3、表13（SG175）、表15（SG128）又は表18（SG64）に挙げられた遺伝子又はこれらの多型から選択された複数の陽性又は陰性GVHD予測遺伝子の発現を測定すること、及び、測定された発現に基づいて陽性又は陰性GVHD予測遺伝子の発現量を得ること、又は、測定された発現に基づいて陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合を得ることを含む。陽性又は陰性GVHD予測遺伝子の発現量と陽性又は陰性GVHD予測遺伝子の既定の基準発現量との比較、又は、陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合と陽性及び／又は陰性GVHD予測遺伝子の線形結合又は非線型結合の既定の基準値との比較が行われる。この比較に基づいて、（1）陽性GVHD予測遺伝子の発現量が陽性GVHD予測遺伝子の既定の基準発現量よりも大きい小さいかということが、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示し、（2）陰性GVHD予測遺伝子の発現量が陰性GVHD予測遺伝子の基準発現量よりも大きい小さいかということが、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示し、（3）陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きい小さいかということが、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示し、（4）陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きい小さいかということが、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示す。実際のHCTドナー又はドナー候補は、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いか低いかを示す発現量、陽性又は陰性GVHD予測遺伝子の総数又は同一性、又は陽性及び／又は陰性GVHD予測遺伝子の組み合わせの評価に基づいて、移植片対宿主病（GVHD）を引き起こすか否かのリスクに関して分類される。

【0010】

本発明によれば、実際のドナー又はドナー候補のHCTがHCTレシピエントにおいて移植片対宿主病（GVHD）を引き起こす又は引き起こさない可能性又はリスクの程度に基づいてそれぞれスコアが割り当てられた（又は分類された）複数の実際のHCTドナー又はHCTドナー候補を含むデータベース又は組織的構成の生成方法がさらに提供される。一実施形態における方法は、実際のドナー又はドナー候補のCD4陽性T細胞又はCD8陽性T細胞において、表1（RNA 1538）、表2、表2A（RNA 143）、表2B（RNA 192）、表3、表13（SG175）、表15（SG128）又は表18（SG64）に挙げられた一又は複数の陽性又は陰性GVHD予測遺伝子又はこれらの多型の発現を測定すること、及び、測定された発現に基づいて陽性又は陰性GVHD予測遺伝子の発現量を得ること、又は、測定された発現に基づいて陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合を得ることを含む。陽性又は陰性GVHD予測遺伝子の発現量と陽性又は陰性GVHD予測遺伝子の既定の基準発現量との比較、又は、陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合と陽性及び／又は陰性GVHD予測遺伝子の線形結合又は非線型結合の既定の基準値との比較が実行される。この比較に基づいて、（1）陽性GVHD予測遺伝子の発現量が陽性GVHD予測遺伝子の既定の基準発現量よりも大きい小さいかということが、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示し、（2）陰性GVHD予測遺伝子の発現量が陰性GVHD予測遺伝子の基準発現量よりも大きい小さいかということが、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示し、（3）陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きい小さいかということが、ドナー候補からのHCTがHCTレシピエント

10

20

30

40

50

においてGVHDを引き起こすリスクが高いこと又は低いことを示し、(4)陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きいか小さいかということが、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示す。実際のドナー又はドナー候補のHCTは、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示す発現量の比較、陽性又は陰性GVHD予測遺伝子の総数又は同一性、又は陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合の比較の評価に基づいて、スコアが割り当てられ又は分類される。ここで、割り当てられるスコアは、実際のドナー又はドナー候補のHCTがHCTレシピエントにおいてGVHDを引き起こす可能性又はリスクを示す。このスコアは、続いて、記録され又は記憶され得る。続いて、一又は複数の追加的な実際のHCTドナー又はHCTドナー候補について、上記の行程が繰り返され、これにより、実際のドナー又はドナー候補のHCTがHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こす又は引き起こさない可能性又はリスクに基づいて各々にスコアが割り当てられた実際のHCTドナー又はHCTドナー候補を含むデータベース又は組織的構成を生成することができる。

10

【0011】

測定のための例示的な陽性及び陰性の「GVHD」予測遺伝子及び例示的なハウスキーピング(「HSK」)遺伝子は、表1(RNA1538、配列番号:1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表12(HSK6)、表13(SG175)、表15(SG128)及び表18(SG64)に列挙されており、これらの中から選択され得る。1546、192、175、128及び64の例示的な陽性及び陰性のGVHD予測遺伝子及びHSK(ハウスキーピング遺伝子)の配列は、配列表に記載されている(配列番号:1~1738)。陽性及び陰性の予測遺伝子の発現を検出、測定又は分析するために用いられるハイブリダイゼーション(検出)及び/又はRT-PCRのための例示的なプローブ及びプライマーは、例えば、表1(RNA1538、配列番号:1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表12(HSK6)、表13(SG175)、表15(SG128)及び表18(SG64)の配列から又はこれらの配列に基づいて得られる。

20

【0012】

本発明によれば、データベース及び組織的構成がさらに提供される。一実施形態において、データベース又は組織的構成は、2以上の陽性又は陰性GVHD予測遺伝子の遺伝子発現プロファイル、陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合、又は、複数の実際のHCTドナー又はHCTドナー候補がGVHDを引き起こす又は引き起こさないスコア又はリスク確率を含む。ここで、当該2以上の陽性又は陰性GVHD予測遺伝子は、表1(RNA1538、配列番号:1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表12(HSK6)、表13(SG175)、表15(SG128)及び表18(SG64)に挙げられた遺伝子又はこれらの多型の任意の組み合わせである。または、当該スコア又はリスク確率は、表1(RNA1538、配列番号:1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表12(HSK6)、表13(SG175)、表15(SG128)及び表18(SG64)に挙げられた一又は複数の陽性又は陰性GVHD予測遺伝子の発現に基づく。また、当該データベース又は組織的構成は、当該遺伝子発現プロファイル、GVHDを引き起こす又は引き起こさないスコア又はリスク確率を、実際のHCTドナー又はHCTドナー候補の各々と関連づける。

30

40

【0013】

本発明によれば、キットがさらに提供される。一実施形態において、キットは、一又は複数の陽性及び/又は陰性GVHD予測遺伝子を検出、測定又は分析するための一又は複数の検体を含む。具体的には、キットは、2以上のプライマーペアを含み、この2以上の

50

プライマーペアの各々は、互いに逆向きに配向しており、また、この2以上のプライマーペアの各々は、表1 (RNA 1538、配列番号：1～1546)、表2、表2A (RNA 143)、表2B (RNA 192)、表3、表12 (HSK6)、表13 (SG175)、表15 (SG128)又は表18 (SG64)に挙げられた陽性又は陰性予測遺伝子又はこれらの多型のうちの一つから生成されるRNA又はcDNAとハイブリダイズする。他の具体的な態様においては、キットは、一又は複数の核酸プローブを含み、当該一又は複数のプローブの少なくとも一つが、表1 (RNA 1538、配列番号：1～1546)、表2、表2A (RNA 143)、表2B (RNA 192)、表3、表12 (HSK6)、表13 (SG175)、表15 (SG128)又は表18 (SG64)に挙げられた陽性又は陰性GVHD予測遺伝子又はこれらの多型の一又は複数のRNA又はcDNAとハイブリダイズする。

10

【0014】

本発明によれば、さらにアレイが提供される。一実施形態において、アレイは、一又は複数の陽性及び/又は陰性GVHD予測遺伝子を検出、測定又は分析するための一又は複数の検体を含む。具体的には、アレイは2以上のプライマーペアを含み、この2以上のプライマーペアの各々は、互いに逆向きに配向している。また、この2以上のプライマーペアの各々は、表1 (RNA 1538、配列番号：1～1546)、表2、表2A (RNA143)、表2B (RNA192)、表3、表12 (HSK6)、表13 (SG175)、表15 (SG128)又は表18 (SG64)に挙げられた陽性又は陰性GVHD予測遺伝子又はこれらの多型のうちの一つから生成されるRNA又はcDNAとハイブリダイズする。また、当該2以上のプライマーペアの各々は、支持体又は基板に貼り付けられ又は固定されている。他の具体的な態様においては、アレイは、一又は複数のプローブを含み、当該プローブの少なくとも一つが表1 (RNA 1538、配列番号：1～1546)、表2、表2A (RNA 143)、表2B (RNA 192)、表3、表12 (HSK6)、表13 (SG175)、表15 (SG128)又は表18 (SG64)に挙げられた陽性又は陰性GVHD予測遺伝子又はこれらの多型から生成されるRNA又はcDNAとハイブリダイズする。また、プローブの各々は、支持体又は基板に付着され又は固定される。

20

【図面の簡単な説明】

【0015】

【図1】48803個のプローブ×48個のサンプルのイルミナ社のシグナル値の典型的な未平滑化ヒストグラムである (Plot 5,1)。

30

【0016】

【図2】イルミナ社の計測値による $\log_{10}(\text{bead_stderr})$ 対 $\log_{10}(\text{プラスのシグナル})$ の散布図である。

【0017】

【図3】イルミナ社の未加工のシグナルデータについて、数万のレベルにおいて多くのシグナルが発生しているものの、大部分は約1500未満のレベルにおいて発生することを経験的に示している。98%のシグナルについては、標準偏差又は分散のシグナルレベルに対する明確かつ顕著な依存が依然として存在する。そして図3に表されているようなデータは、各サンプルについて別個にVSTデータ依存のパラメータ c_1 及び c_2 を計算する際に採用される。

40

【0018】

【図4】48803×48個のサンプル集合のすべてのシグナル値の、当該集合がVSTに変換された後のヒストグラムである。横軸に沿ってより見やすくするために、最大の5%のデータは除外されている。

【0019】

【図5】6つのGVHD転帰クラスの各々について、全122サンプルについて登り順でプロットされたRNA発現測定値を示す。

【0020】

【図6】6つのGVHD転帰クラスの各々について、全122サンプルについて登り順でプ

50

ロットされたRNA発現測定値を示す。

【0021】

【図7】6つのGVHD転帰クラスの各々について、全122サンプルについて登り順でプロットされたRNA発現測定値を示す。

【0022】

【図8】GVHDグループの番号とともに、安定的に、一本調子に増加するGVHDグループ平均の系列。

【0023】

【図9】GVHDグループの番号とともにGVHDグループ平均の安定的な下降トレンドを示す。

【0024】

【図10】20個の性能のよい個別のLDA遺伝子からのGVHD陰性票の相対スコアのプロット(RNA20 GROUPS)を示し、また、GVHD重症度が増加するにつれてGVHDグループ平均スコアの安定的な下降トレンドを描くことを示す。

【0025】

【図11】個別のRNA発現マーカーに対応するLDAモデルに関し、「anyGVGD」対「no GVHD」についてサンプル特異的GVHD転帰予測を示す。CTCF LDAサンプルは、区分線の下にあるときにGVHD陰性に分類される。

【0026】

【図12】個別のRNA発現マーカー-BLVRAに対応するLDAモデルに関し、「anyGVGD」対「no GVHD」についてサンプル特異的なGVHD転帰予測を示す。BLVRA LDAサンプルは区分線の上にあるときにGVHD陰性に分類される。

【0027】

【図13】RNAマーカー投票モデルに対応するLDAモデルRNA20 LDA - Aに関し、「anyGVGD」対「no GVHD」についてサンプル特異的GVHD転帰予測を示す。RNA 20 LDAサンプルは、区分線よりも上方にあるときにGVHD陰性と分類される。

【0028】

【図14】慢性GVHD(単独又は任意の急性GVHDとの組み合わせ)を「no GVHD」転帰(cGVHD対noGVHD)と区別する際に、わずか2例の偽陰性分類が報告されたことを示す。(RNA20 LDA - B) (Negative Predictive Value =0.95)

【0029】

【図15】急性GVHDの任意の態様(単独又は慢性GVHDとの組み合わせ)を「no GVHD」転帰(aGVHD対noGVHD)と区別する際に、わずか3例の偽陰性分類が報告されたことを示す。(RNA20 LDA - C) (Negative Predictive Value =0.94)

【0030】

【図16】急性GVHD(任意の態様)と組み合わせられた慢性GVHDを「no GVHD」転帰(「a&cGVHD」対「noGVHD」)と区別する際に、わずか1例の偽陰性分類が報告されたことを示す。(RNA20 LDA - D) (Negative Predictive Value= 0.96)

【0031】

【図17】最も重要なグレード3又は4の急性GVHD(単独又は慢性GVHDとの組み合わせ)を「no GVHD転帰」(「a34GVHD」対「noGVHD」)と区別する際に、偽陰性分類が全く報告されなかったことを示す。(RNA20 LDA - E) (Negative Predictive Value = 1.00)

【0032】

【図18】相対的に高い真陰性及び高い真陰性率を維持しながら偽陰性を最小化し陰性予測値を最大化するために、閾値0.77を選択したことを示す。(RNA20 LDA PERFORMANCE - A、任意のGVHD対「no GVHD」)

【0033】

【図19】5つのLDA正確性測定の詳細な挙動を示し、陽性予測値(PPV)及び真陽性率(TPR、感度)を含む。RNA20 LDA PERFORMANCE - B、「any GVHD」対「no GVHD」

10

20

30

40

50

【発明を実施するための形態】

【0034】

本発明は、A H C T（同種造血細胞移植）ドナーのC D 4 陽性T細胞の遺伝子発現プロファイルに関する。かかるドナーは、H C Tレシピエントにおいて、G V H Dを引き起こすこと、及び、引き起こさないことが知られている。本明細書において説明する研究は、H C TドナーのC D 4 陽性中における多数の遺伝子を同定する。この特定される遺伝子は、H C TレシピエントにおいてG V H Dを引き起こさなかったH C Tドナーにおいて発現が増加したものであり、陰性予測遺伝子と称される。本明細書において説明される研究は、H C TドナーのC D 4 陽性T細胞における多数の遺伝子であって、H C TレシピエントにおいてG V H Dを引き起こしたH C Tドナーにおいて発現が増加した遺伝子も同定する。このような遺伝子は、陽性予測遺伝子と呼ばれる。一又は複数の「G V H D」予測遺伝子の発現の測定は、あるドナー候補からのH C TがあるH C TレシピエントにおいてG V H Dを引き起こすリスクを解明又は予測するために用いられ得る。例えば、随意にH C TレシピエントにH L A適合するH C Tドナー候補のC D 4 陽性T細胞における一又は複数のかかる遺伝子の発現は測定が可能である。H T CがH C TレシピエントにおいてG V H Dを引き起こすにつれて増加することが知られている一又は複数の遺伝子の発現の増加は、H C TドナーがH C TレシピエントにおいてG V H Dを引き起こしやすいか否かに関する情報を提供することができる。同様に、H C TがH C TレシピエントにおいてG V H Dを引き起こさないことにつれて増加することが知られている一又は複数の遺伝子の発現の増加は、ドナーがH C TレシピエントにおいてG V H Dを引き起こしにくい否かに関する情報を提供することができる。一又は複数のかかる陽性又は陰性G V H D予測遺伝子もしくはかかる陽性又は陰性G V H D予測遺伝子の組み合わせ、複数の陽性及び陰性のG V H D予測遺伝子、又は、特にかかる陽性及び/又は陰性G V H D予測遺伝子の比率の測定は、適度の、高い、又は非常に高い信頼性をもって、任意のH C TドナーがH C TレシピエントにおいてG V H Dを引き起こすリスクを予測又は決定するために用いられる。

10

20

【0035】

したがって、本発明は、ドナー候補からの造血細胞移植（H C T）がH C Tレシピエントにおいて移植片対宿主病（G V H D）を引き起こすか又は引き起こさないリスクを予測及び/又は決定する方法を提供する。一実施形態における方法は、ドナー候補のC D 4 陽性T細胞又はC D 8 陽性T細胞において、表1（R N A 1 5 3 8）、表2、表2 A（R N A 1 4 3）、表2 B（R N A 1 9 2）、表3、表1 3（S G 1 7 5）、表1 5（S G 1 2 8）又は表1 8（S G 6 4）から選択された一又は複数の陽性又は陰性G V H D予測遺伝子又は陽性及び/もしくは陰性G V H D予測遺伝子の組み合わせ又はこれらの多型の発現を測定することを含む。測定された遺伝子発現レベルに基づいて、陽性又は陰性G V H D予測遺伝子の発現量が得られる。または、測定された発現レベルに基づいて、陽性及び/又は陰性G V H D予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が得られる。陽性又は陰性G V H D予測遺伝子の発現量と陽性又は陰性G V H D予測遺伝子の既定の基準発現量との比較、又は、陽性及び/又は陰性G V H D予測遺伝子の組み合わせの発現量の線形結合又は非線型結合と陽性及び/又は陰性G V H D予測遺伝子の線形結合又は非線型結合の既定の基準値との比較が実行される。この比較において、（1）陽性G V H D予測遺伝子の発現量が陽性G V H D予測遺伝子の既定の基準発現量よりも大きい小さいかということが、ドナー候補からのH C TがH C TレシピエントにおいてG V H Dを引き起こすリスクが高いこと又は低いことを示し、（2）陰性G V H D予測遺伝子の発現量が陰性G V H D予測遺伝子の既定の基準発現量よりも大きい小さいかということがドナー候補からのH C TがH C TレシピエントにおいてG V H Dを引き起こすリスクが高いこと又は高いことをそれぞれ示し、（3）陽性及び/又は陰性G V H D予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きい小さいかということが、ドナー候補からのH C TがH C TレシピエントにおいてG V H Dを引き起こすリスクが高いこと又は低いことを示し、（4）陽性及び/又は陰性G V H D予測遺伝子の

30

40

50

組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きいか小さいかということが、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示す。当該ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いか低いかを示す発現量の比較、陽性又は陰性GVHD予測遺伝子の総数又は同一性、又は陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合の比較の評価に基づいて、当該ドナー候補からのHCTがHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすか引き起こさないかのリスク又は可能性が予測及び/又は決定される。

【0036】

他の実施形態において、ドナー候補からの造血細胞移植(HCT)がHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすか引き起こさないかのリスクを予測及び/又は決定する方法は、ドナー候補のCD4陽性T細胞又はCD8陽性T細胞又はCD4陽性T細胞又はCD8陽性T細胞によって発現された核酸又はタンパク質を、表1(RNA1538)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)又は表18(SG64)に記載された一又は複数の陽性又は陰性GVHD予測遺伝子又はこれらの多型の発現を検出する検体に接触させること、及び、CD4陽性T細胞又はCD8陽性T細胞内の一又は複数の陽性又は陰性GVHD予測遺伝子の発現を測定して陽性又は陰性GVHD予測遺伝子の発現量を得ること、又は、当該陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現を測定して陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合を得ることを含む。陽性又は陰性GVHD予測遺伝子の発現量と陽性又は陰性GVHD予測遺伝子の既定の基準発現量との比較、又は、陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合と陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合の既定の基準値との比較が実行される。この比較に基づいて、(1)陽性GVHD予測遺伝子の発現量が陽性GVHD予測遺伝子の既定の基準発現量よりも大きいか小さいかということが、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示し、(2)陰性GVHD予測遺伝子の発現量が陰性GVHD予測遺伝子の基準発現量よりも大きいか小さいかということが、当該ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示し、(3)陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きいか小さいかということが、当該ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示し、(4)陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きいか小さいかということが、当該ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示す。当該ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いか低いかを示す発現量の比較、陽性又は陰性GVHD予測遺伝子の総数又は同一性、又は陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合の比較の評価に基づいて、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクを予測又は決定することができる。

【0037】

本発明はまた、HCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすか引き起こさないかのリスク又は可能性に応じて、造血細胞移植(HCT)ドナー候補を分類又は類型化する方法を提供する。一実施形態における方法は、HCTドナー候補のCD4陽性T細胞又はCD8陽性T細胞において、表1(RNA1538)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)又は表18(SG64)に記載されている遺伝子から選択された複数の陽性又は陰性GVHD予測遺伝子又はこれらの多型の発現を測定すること、及び、測定された発現に基づいて、陽性又は陰性GVHD予測遺伝子の発現量を得ること、又は、測定された発

10

20

30

40

50

現に基づいて、陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合を得ることを含む。陽性又は陰性GVHD予測遺伝子の発現量と陽性又は陰性GVHD予測遺伝子の既定の基準発現量との比較、又は、陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合と陽性及び／又は陰性GVHD予測遺伝子の線形結合又は非線型結合の既定の基準値との比較が実行される。この比較に基づいて、(1)陽性GVHD予測遺伝子の発現量が陽性GVHD予測遺伝子の既定の基準発現量よりも大きいか小さいかということが、当該ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示し、(2)陰性GVHD予測遺伝子の発現量が陰性GVHD予測遺伝子の基準発現量よりも大きいか小さいかということが、当該ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示し、(3)陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きいか小さいかということが、当該ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示し、(4)陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きいか小さいかということが、当該ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示す。実際のHCTドナー又はHCTドナー候補は、当該ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いか低いかを示す発現量の比較、陽性又は陰性GVHD予測遺伝子の総数又は同一性、又は陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合の比較の評価に基づいて、移植片対宿主病(GVHD)を引き起こす又は引き起こさないリスクに関して分類される。

10

20

30

40

50

【0038】

本発明はさらに、データベース及び組織的構成を生成又は作成する方法を提供する。このデータベース又は組織的構成は、実際のドナー又はドナー候補からのHCTがHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こす又は引き起こさない可能性又はリスクに基づいて、随意に分類され、類型化されもしくはスコアが割り当てられ又は特定された複数の実際のHCTドナー及び／又はHCTドナー候補を含む。一実施形態における方法は、実際のドナー又はドナー候補のCD4陽性T細胞又はCD8陽性T細胞において、表1(RNA1538)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)又は表18(SG64)に記載された一又は複数の陽性又は陰性GVHD予測遺伝子又はこれらの多型の発現を測定すること、及び、測定された発現に基づいて陽性又は陰性GVHD予測遺伝子の発現量を得ること、又は、測定された発現に基づいて陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合を得ることを含む。陽性又は陰性GVHD予測遺伝子の発現量と陽性又は陰性GVHD予測遺伝子の既定の基準発現量との比較、又は、陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合と陽性及び／又は陰性GVHD予測遺伝子の線形結合又は非線型結合の既定の基準値との比較が実行される。この比較に基づいて、(1)陽性GVHD予測遺伝子の発現量が陽性GVHD予測遺伝子の既定の基準発現量よりも大きいか小さいかということが、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示し、(2)陰性GVHD予測遺伝子の発現量が陰性GVHD予測遺伝子の基準発現量よりも大きいか小さいかということが、当該ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示し、(3)陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きいか小さいかということが、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことを示し、(4)陽性及び／又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合が既定の基準値よりも大きいか小さいかということが、当該ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いこ

とをそれぞれ示す。実際のHCTドナー又はHCTドナー候補は、当該ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いか低いかを示す発現量の比較、陽性又は陰性GVHD予測遺伝子の総数又は同一性、又は陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合の比較に基づいてスコアを割り当てられ又は分類される。このスコアは、実際のドナー又はHCTドナー候補のHCTレシピエントにおいてGVHDを引き起こす可能性又はリスクを反映したものである。このスコアは、その後、記録又は記憶され、一又は複数の追加的な実際のHCTドナー又はHCTドナー候補について上記の工程が随意に繰り返され、実際のドナー又はドナー候補のHCTがHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こす又は引き起こさない可能性又はリスクに基づいてスコアが割り当てられた実際のHCTドナー又はHCTドナー候補を含むデータベース又は組織的構成が生成される。

10

【0039】

本発明の方法のさらに他の具体的な態様において、HCTドナー候補の陽性又は陰性遺伝子発現プロファイル、HCTドナー候補の陽性又は陰性GVHD予測遺伝子の発現量、当該発現量とHCTドナー候補の陽性又は陰性予測遺伝子の対応する既定の基準発現量との比較、又は陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合の比較の一又は複数が、例えば、電子媒体、電子フォーマット又は電子書式に記録又は記憶されてもよい。これらは、コンピュータ読み取り可能又はコンピュータアクセス可能であってもよい。

20

【0040】

追加的な実施形態において、本発明の方法は、遺伝子に特異的にハイブリダイズする一又は複数のプローブ又はプライマーを用いて実行されてもよい。この一又は複数のプローブ又はプライマーは、表1(RNA1538、配列番号：1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表12(HSK6)、表13(SG175)、表15(SG128)又は表18(SG64)に記載された配列から選択され、これらの配列に由来し、又はこれらの配列に基づくプローブ又はプライマーから選択される。例えば、ドナー候補からの造血細胞移植(HCT)がHCTレシピエントにおいてGVHDを引き起こす又は引き起こさないリスクを予測及び/又は決定する方法においては、一又は複数の陽性又は陰性GVHD予測遺伝子の発現は、表1(RNA1538、配列番号：1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表12(HSK6)、表13(SG175)、表15(SG128)又は表18(SG64)に記載された配列から選択され、これらの配列に由来し、又はこれらの配列に基づく一又は複数のプローブ又はプライマーを用いる。かかるプローブ及びプライマーは、表1(RNA1538、配列番号：1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表12(HSK6)、表13(SG175)、表15(SG128)及び表18(SG64)に記載された対応する遺伝子とハイブリダイズすると仮定される。したがって、他のプローブ及びプライマーは、当該遺伝子の核酸配列に基づいて、本明細書で説明されるように当該遺伝子の発現を測定又は分析するためにデザインされる。しかしながら、当該プローブ又はプライマーが異なる遺伝子とハイブリダイズする場合には、本発明の方法は、表1(RNA1538、配列番号：1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表12(HSK6)、表13(SG175)、表15(SG128)又は表18(SG64)に記載された特定のプローブ(又は類似配列及び/又は長さのプローブ)又はプライマーは、陰性又は陽性GVHD予測遺伝子に特異的なので、当該プローブ又はプライマーが表に記載された特定の遺伝子とハイブリダイズしないとしても、当該プローブ又はプライマーの一又は複数を用いて実行できる。

30

40

【0041】

その発現が、HCTドナーがHCTレシピエントにおいてGVHDを引き起こすリスクの減少と相関を有する特定の遺伝子は、表1(RNA1538、配列番号：1~1546)、表

50

2、表2A(RNA143)、表2B(RNA192)、表3、表12(HSK6)、表13(SG175)、表15(SG128)又は表18(SG64)において特定されている。このような遺伝子は、陰性予測遺伝子と呼ばれる。本発明における陰性予測遺伝子は、それゆえ、ドナー候補のCD4陽性T細胞又はCD8陽性T細胞における発現の増加がHCTレシピエントにおいてGVHDを引き起こすリスクの減少と相関する遺伝子である。例示的な陰性予測遺伝子は、表1(RNA1538、配列番号：1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、及び表13(SG175)において、記号「N」で示されている。また、所定の陰性予測遺伝子については、ドナーCD4陽性T細胞又はCD8陽性T細胞における当該陰性予測遺伝子の発現が多くなるほど、HCTドナーがHCTレシピエントにおいてGVHDを引き起こすリスク又は可能性が小さくなる。

10

【0042】

本明細書において説明されているとおり、CD4陽性T細胞における陰性予測遺伝子の発現の増加は、GVHDを引き起こさないHCTと相関があり、HCTドナーがHCTレシピエントにおいてGVHDを引き起こすリスク又は可能性が減少することを示す。したがって、陰性予測遺伝子の発現の減少は、HCTドナーがHCTレシピエントにおいてGVHDを引き起こすリスク又は可能性の増加と相関し、かかる増加を示すものである。

【0043】

その発現が、HCTドナーがHCTレシピエントにおいてGVHDを引き起こすリスクの増加と相関を有する特定の遺伝子は、表1(RNA1538、配列番号：1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)又は表18(SG64)において特定されている。このような遺伝子は、陽性予測遺伝子と呼ばれる。本発明における陽性予測遺伝子は、それゆえ、ドナー候補のCD4陽性T細胞又はCD8陽性T細胞における発現の減少がHCTレシピエントにおいてGVHDを引き起こすリスクの増加と相関する遺伝子である。例示的な陽性予測遺伝子は、表1(RNA1538、配列番号：1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、及び表13(SG175)において、記号「P」で示されている。また、所定の陽性予測遺伝子については、ドナーCD4陽性T細胞における当該陽性予測遺伝子の発現が多くなるほど、HCTドナーがHCTレシピエントにおいてGVHDを引き起こすリスク又は可能性が大きくなる。

20

【0044】

本明細書において説明されているとおり、CD4陽性T細胞における陽性予測遺伝子の発現の増加は、GVHDを引き起こすHCTと相関があり、HCTドナーがHCTレシピエントにおいてGVHDを引き起こすリスク又は可能性が増加することを示す。したがって、陽性予測遺伝子の発現の減少は、HCTドナーがHCTレシピエントにおいてGVHDを引き起こすリスク又は可能性の減少と相関し、かかる減少を示すものである。

30

【0045】

本発明又は本発明に係る任意の他の方法における陰性及び陽性GVHD予測遺伝子は、表1(RNA1538、配列番号：1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表12(HSK6)、表13(SG175)、表15(SG128)及び表18(SG64)に記載されており、HCTドナー候補(又は実際のHCTドナー)がレシピエントにおいてGVHDを引き起こし又は引き起こさないリスクを予測又は決定するために、個別に測定又は評価され得るし、または、複数のかかる遺伝子は、HCTドナー候補(又は実際のHCTドナー)のCD4陽性T細胞又はCD8陽性T細胞において測定又は分析され得る。このように、これらの表に記載されている陰性及び陽性予測遺伝子のグループ化は、説明の便宜のために過ぎず、如何なる意味においても、表に記載されている全ての遺伝子が分析されなければならないことや、表中の陰性及び/又は陽性予測遺伝子の最小限の数が分析されなければならないこと等を意味するものではない。むしろ、本明細書のガイダンスに照らして、表1(RNA1538、配列番号：1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表12(HSK6)、表13(SG175)、表15(SG128)及び表18(SG64)中

40

50

の陰性及び／又は陽性GVHD予測遺伝子の任意の組み合わせが、本発明の方法を実行するために測定又は分析され、または、本発明キットやアレイを生成するために用いられ得る。このように、非限定的な例示として、表2B(RNA192)から選択された一又は複数の陰性及び／又は陽性GVHD予測遺伝子は、表12(HSK6)、表13(SG175)、表15(SG128)又は表18(SG64)のいずれかに記載された任意の遺伝子と組み合わせることができ、表13(SG175)から選択された一又は複数の陰性及び／又は陽性GVHD予測遺伝子は、表2B(RNA192)、表12(HSK6)、表15(SG128)又は表18(SG64)のいずれかに記載されている任意の遺伝子と組み合わせることが可能である。

【0046】

本発明によれば、測定又は分析される遺伝子の数は、単一の遺伝子であってもよく(表1(RNA1538、配列番号:1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表12(HSK6)、表13(SG175)、表15(SG128)又は表18(SG64)に記載された任意の配列)、または、表1(RNA1538、配列番号:1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表12(HSK6)、表13(SG175)、表15(SG128)又は表18(SG64)に記載された全ての遺伝子以下の任意の数の陰性及び／又は陽性GVHD予測遺伝子であってもよく、制限はない。また、任意の特定の陰性又は陽性予測遺伝子が分析されなければならないということもない。同様に、遺伝子の比率及び陽性及び／又は陰性GVHD予測遺伝子の組み合わせの分析は、表1(RNA1538、配列番号:1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表12(HSK6)、表13(SG175)、表15(SG128)又は表18(SG64)に記載された配列に基づいて実行され得る。繰り返しになるが、本明細書における遺伝子の表は、説明の便宜のためのものに過ぎず、特定の遺伝子又は遺伝子の特定の組み合わせに限定されない。例えば、表3は、20の遺伝子モデル(RNA20モデルとしても知られる)を示しており、そこでのドナーのCD4陽性T細胞におけるかかる遺伝子の分析又は測定は、ドナーとレシピエントとの間の標準の10分の10HLA適合(10 out of 10 HLA matches)を用いる場合よりも、HCTドナーがHCTレシピエントにおいてGVHDを引き起こすリスクを予測又は決定するより高い能力を提供する。同様に、表1(RNA1538、配列番号:1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)及び表18(SG64)は、HCT候補ドナーのCD4陽性T細胞において分析又は測定されると、ドナーとレシピエントとの間の標準の10分の10HLA適合を用いるよりも、HCTドナーがHCTレシピエントにおいてGVHDを引き起こすリスクを予測又は決定するより高い能力を示す遺伝子を示している。HCTドナーがHCTレシピエントにおいてGVHDを引き起こすリスクを予測又は決定する他の好適なモデルは、表1(RNA1538、配列番号:1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表12(HSK6)、表13(SG175)、表15(SG128)及び表18(SG64)に記載された陰性及び陽性予測遺伝子の任意の組み合わせ並びに本明細書の教示に基づいて容易に得られる。したがって、本発明の方法は、表1(RNA1538、配列番号:1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表12(HSK6)、表13(SG175)、表15(SG128)又は表18(SG64)に記載された陰性及び／又は陽性予測遺伝子の任意のもの又はこれらの表に記載された陰性及び／又は陽性予測遺伝子の任意の組み合わせを測定又は分析することを含む。同様に、キット、アレイ及びデータベース等の本発明の合成物は、表1(RNA1538、配列番号:1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表12(HSK6)、表13(SG175)、表15(SG128)及び表18(SG64)に記載された陰性及び／又は陽性GVHD予測遺伝子の任意のもの又はこれらの遺伝子の任意の組み合わせの分析用又は測定用プライマー及び／又はプローブ、又は、これらの発現プロファイルを有するデータベースを含むがこれらには限定されない。

10

20

30

40

50

【 0 0 4 7 】

本明細書での用法のとおり、遺伝子発現プロファイル又は「発現プロファイル」は、HCT候補ドナーのCD4陽性T細胞又はCD8陽性T細胞における一又は複数の陽性及び/又は陰性GVHD予測遺伝子の発現レベルを意味し、GVHDの予後の予測又は決定に関連する。このようなプロファイルは、遺伝子比率及び陽性及び/又は陰性GVHD予測遺伝子の発現量の組み合わせを含んでもよい。プロファイルは特定のドナー候補と対応するので、HCTレシピエントのドナーとしてHCTの安定性を採点、特定又は文書かする方法を提供する。

【 0 0 4 8 】

HCT候補ドナー又はHCTレシピエントの遺伝子発現レベル、プロファイル、スコア、及びこれら以外の指標は、本明細書で説明される方法（例えば、比較及び評価）で用いるのに適した任意のデータ形式で表現され得る。レベル、プロファイル、及びスコアは、物理的表現（例えば、グラフ等の紙）、コンピュータ（例えば、スクリーン上）又は電子媒体もしくはコンピュータ読み取り可能な媒体に記憶されるデジタル表現として提示され得る。かかるデータは、例えば、HCTレシピエントにおいてGVHDを引き起こすリスク又は可能性が低いHCTドナー候補を特定するために、ユーザによってアクセスされる。

10

本明細書での用法のとおり、表1（RNA1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表12（HSK6）、表13（SG175）、表15（SG128）及び表18（SG64）に記載されている陰性及び陽性GVHD予測遺伝子の多型が含まれる。多型は、RNA又はゲノムDNA配列レベルでの遺伝的異形である。かかる多型は、典型的には、自然に発生する配列変異であり、単一の又は複数のヌクレオチド変化である。多型は、機能に影響を与えずコード化タンパク質のアミノ酸残基を変更するという点でサイレントであってもよく、または、遺伝子、mRNA又はコード化タンパク質の活性、発現、半減期等に影響を与えてもよい。しかしながら、このような多型は、サイレントでは無く、機能に影響を与え、コード化タンパク質のアミノ酸残基を変化させ、又は、遺伝子、mRNA又はコード化タンパク質の活性、発現、半減期等に影響を与えても良い。表1ないし表3に記載された陰性及び陽性予測遺伝子の特定の多型は、当業者に知られているものであり、本明細書において説明されているように、又はそれ以外の方法を用いて、測定又は分析され得る。

20

30

【 0 0 4 9 】

本明細書での用法のとおり、「複数の」という語は、二以上を意味する。本明細書での用法のとおり、複数の陽性及び/又は陰性予測遺伝子が測定又は分析され得る。このように、表1（RNA1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表12（HSK6）、表13（SG175）、表15（SG128）又は表18（SG64）の二以上の遺伝子が、本発明の方法において測定又は分析され得る。特定の実施形態において、測定又は分析される陰性及び/又は陽性予測遺伝子の数は、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20又はそれ以上であり、例えば、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50等であり、表1（RNA1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表12（HSK6）、表13（SG175）、表15（SG128）及び表18（SG64）に記載された全ての遺伝子の数以下であればよい。

40

同様に、本発明のキット及び/又はアレイにおける複数の検体（例えば、プライマー、プローブ又は抗体）は、表1（RNA1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表12（HSK6）、表13（SG175）、表15（SG128）又は表18（SG64）に記載されている陽性又は陰

50

性予測遺伝子又はこれらの遺伝子でコード化された発現産物（タンパク質）と結合又はハイブリダイズし、当該陽性又は陰性GVHD予測遺伝子の発現量を得ることができ、当該陽性又は陰性予測遺伝子の発現量を既定の基準発現量と比較することができる。このように、本発明のキット及び/又はアレイにおける検体（例えば、プライマー、プローブ又は抗体）は、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20又はそれ以上、例えば、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50等の遺伝子で、表1（RNA1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表12（HSK6）、表13（SG175）、表15（SG128）及び表18（SG64）に記載された全ての遺伝子の数以下のもの又はこのような遺伝子によってコード化された発現産物（タンパク質）と結合又はハイブリダイズできるものを含むことができる。

複数のこのような遺伝子について、HCTドナー候補がHCTレシピエントにおいてGVHDを引き起こすリスク又は可能性に応じたGVHD転帰の予測や決定、又は、分類、類型化、採点又は特定は、前記複数の陽性又は陰性予測遺伝子の発現量とこれらに対応する既定の基準発現量との比較の総計に基づく。遺伝子発現プロファイル又はより単純に発現プロファイルは、所定のHCTドナー候補又は二以上のHCT候補ドナーの複数の陰性及び/又は陽性予測遺伝子の発現を意味し、又は、前記複数の陽性又は陰性予測遺伝子の発現量のデータセットであり、又は、陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合の比較のデータセットであり、随意に対応する既定の基準発現量と比較される。このように、GVHDの転帰、スコア等のリスクの決定又は予測を提供するために、十分な複数の陰性及び/又は陽性予測遺伝子が発現のために測定され、発現量又は発現量の組み合わせが各々について特定される。

当然、表1（RNA1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表12（HSK6）、表13（SG175）、表15（SG128）及び表18（SG64）に記載されていない追加的な遺伝子及びこのような遺伝子によってコード化された発現産物（タンパク質）も、測定又は分析され、又は本発明の方法に含められる。また、本発明のキット及びアレイにおける検体（例えば、プライマー、プローブ又は抗体）は、表1（RNA1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表12（HSK6）、表13（SG175）、表15（SG128）又は表18（SG64）に記載されていない一又は複数の遺伝子と結合又はハイブリダイズすることができる。しかしながら、ドナー候補からのHCTがGVHDを引き起こす又は引き起こさないリスク又は可能性の程度を予測又は決定するためには、発現が測定又は分析される遺伝子は、表1（RNA1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表12（HSK6）、表13（SG175）、表15（SG128）又は表18（SG64）に記載された遺伝子から選択された一又は複数の遺伝子又はこれらの遺伝子によってコード化された発現産物（タンパク質）である。

【0050】

本発明に係る方法において、GVHD転帰の予測又は決定は、一又は複数の陽性又は陰性予測遺伝子の発現レベルと当該特定の陽性又は陰性予測遺伝子の既定の又は所定の基準発現量との比較による。GVHDの高リスクと相関する値に近いHCT候補ドナーの遺伝子発現は、当該特定の遺伝子がGVHDを引き起こすリスクが高いことを示すと考えられることを意味する。一方、GVHDの低リスクと相関する値に近いHCT候補ドナーの遺伝子発現は、GVHDを引き起こすリスクが低いことを示すと考えられることを意味する。具体的には、陽性予測遺伝子については、当該特定の陽性予測遺伝子の既定の又は所定の基準発現量よりも高いレベルの発現は、GVHDを引き起こすことが知られている一又は複数のHCTドナーにおける陽性予測遺伝子の発現と相関するので、HCTがレシピエ

10

20

30

40

50

ントにおいてGVHDを引き起こすリスク又は可能性が高いことを示す。したがって、既定の又は所定の基準発現量よりも陽性予測遺伝子の発現量が多いことは、当該ドナー候補からのHCTが移植片対宿主病(GVHD)を引き起こすリスクが高いことを示す。陰性予測遺伝子については、当該特定の陰性予測遺伝子の既定の又は所定の基準発現量よりも高いレベルの発現は、GVHDを引き起こさないことが知られている一又は複数のHCTドナーにおける陰性予測遺伝子の発現と相関するので、HCTがレシピエントにおいてGVHDを引き起こすリスク又は可能性が低いことを示す。したがって、既定の又は所定の基準発現量よりも陰性予測遺伝子の発現量が多いことは、当該ドナー候補からのHCTが移植片対宿主病(GVHD)を引き起こすリスクが低いことを示す。

10

【0051】

陽性及び陰性のGVHD予測遺伝子の既定の又は所定の基準発現量は、HCTレシピエントにおいて少なくともある程度GVHDを引き起こすことが知られているHCTドナー及びHCTレシピエントにおいてGVHDを引き起こさないことが知られているHCTドナーの発現分析により決定又は設定される値である。したがって、陽性及び陰性のGVHD予測遺伝子の既定の又は所定の基準発現量(又は、同様に、陽性及び/又は陰性GVHD予測遺伝子の組み合わせの発現量の線形結合又は非線型結合は既定の又は所定の基準値を有する)は、高いレベルの発現が、あるドナー候補のHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いこと又は低いことをそれぞれ示すと考えられるように設定される値である。当然、対応する陽性又は陰性予測遺伝子の既定の又は所定の基準発現量よりも少ない陽性又は陰性GVHD予測遺伝子の発現は、あるドナー候補のHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いこと又は高いことをそれぞれ示すと考えられる。したがって、既定の又は所定の基準発現量は、HCTレシピエントにおけるHCTドナー候補のGVHD転帰の高いリスク又は可能性と低いリスク又は可能性とを区別する境界値(すなわち、区分線)と考えられる。

20

【0052】

既定の又は所定の基準発現量は、判別分析によって決定される。この分析は、統計的に有意でGVHDを引き起こす高リスクと低リスクとの間でGVHD転帰の予測又は決定を区別する陽性又は陰性予測遺伝子発現の量を決定する。例えば、線形判別分析(LDA)又は二次判別分析(QDA)等の判別分析は、HCTレシピエントにおいてGVHDを引き起こすこと又はGVHDを引き起こさないことが分かっているHCTドナー候補の遺伝子発現量を判別する基礎を提供する。

30

【0053】

既定の又は所定の基準発現量はユーザによって設定され得る。例えば、所定の陽性又は陰性予測遺伝子の既定の又は所定の基準発現量は、HCTレシピエントにおいてGVHDを引き起こすことが分かっているHCTドナーのCD4陽性T細胞又はCD8陽性T細胞における陽性又は陰性予測遺伝子の発現と、GVHDを引き起こさないことが分かっているHCTドナーのCD4陽性T細胞又はCD8陽性T細胞における当該陽性又は陰性予測遺伝子の発現との概ね又は正確に中間(midway)で設定され得る。したがって、前記中間値よりも大きい陽性予測遺伝子の発現量は、ドナー候補からのHCTが移植片対宿主病(GVHD)を引き起こすリスクが高いことを示し、前記中間値よりも小さい陰性予測遺伝子の発現量は、ドナー候補からのHCTが移植片対宿主病(GVHD)を引き起こすリスクが低いことを示し、前記中間値よりも小さい陽性予測遺伝子の発現量は、ドナー候補からのHCTが移植片対宿主病(GVHD)を引き起こすリスクが低いことを示し、前記中間値よりも小さい陰性予測遺伝子の発現量は、ドナー候補からのHCTが移植片対宿主病(GVHD)を引き起こすリスクが高いことを示す。

40

【0054】

一般に、より信頼性の高い既定の又は所定の基準発現量は、HCTレシピエントにおいてGVHDを引き起こす複数のHCTドナーのCD4陽性T細胞又はCD8陽性T細胞における陽性又は陰性GVHD予測遺伝子の発現の平均値又は中央値に基づくものである可能

50

性があり、また、HCTレシピエントにおいてGVHDを引き起こさない複数のHCTドナーのCD4陽性T細胞又はCD8陽性T細胞におけるthe陽性又は陰性予測遺伝子の発現の平均値又は中央値に基づくものである可能性がある。したがって、一実施形態において、陽性予測遺伝子の既定の又は所定の基準発現量は、GVHDを引き起こす2以上のHCTドナー及びGVHDを引き起こさない2以上のHCTドナーの陽性予測遺伝子の発現の平均値又は中央値レベルの概ね又は正確に中間に設定される。他の実施形態において、陰性予測遺伝子の既定の又は所定の基準発現量は、GVHDを引き起こす2以上のHCTドナーの陰性予測遺伝子の発現の平均値又は中央値レベルとGVHDを引き起こさない2以上のHCTドナーの陰性予測遺伝子の発現の平均値又は中央値レベルとの、概ね又は正確に中間に設定される。さらに他の実施形態において、陽性又は陰性予測遺伝子の既定の基準発現量は、GVHDを引き起こす少なくとも2、3、4、5又はそれ以上のHCTドナー（例えば、6、7、8、9、10、11、12、13、14、15、16、17、18、19、又はそれ以上のHCTドナー、例えば、20、21、22、23、24、25等又はそれ以上）の陽性又は陰性予測遺伝子の発現の平均値又は中央値レベルと、GVHDを引き起こさない少なくとも2、3、4、5又はそれ以上のHCTドナー（例えば、6、7、8、9、10、11、12、13、14、15、16、17、18、19、又はそれ以上のHCTドナー、例えば、20、21、22、23、24、25等、又はそれ以上）の陽性又は陰性予測遺伝子の発現の平均値又は中央値レベルと、概ね又は正確に中間に設定される。

10

【0055】

陽性又は陰性GVHD予測遺伝子の既定の又は所定の基準発現量は、陽性又は陰性予測遺伝子について測定された発現量との比較を容易にするために、数値を割り当てられても良い。当該数値よりも大きい発現は、HCTレシピエントにおいてGVHDを引き起こすHCTドナーのリスクが高いこと又は低いことを示すものとして理解される。具体的な実施形態において、既定の又は所定の基準発現量（例えば、中間値）は、0.5の数値を割り当てられ、0.5よりも大きい一又は複数の陰性予測遺伝子の発現量は、ドナー候補からのHCTが移植片対宿主病（GVHD）を引き起こすリスクが低いことを示す。他の具体的な実施形態において、既定の又は所定の基準発現量（例えば、中間値）は、0.5の数値を割り当てられ、0.5よりも大きい一又は複数の陽性予測遺伝子の発現量は、ドナー候補からのHCTが移植片対宿主病（GVHD）を引き起こすリスクが高いことを示す。当然、GVHD転帰の予測又は決定により高い信頼性が望まれる場合には、既定の又は所定の基準発現（数）値を超えることが要求される発現量は、増加され得る。このように、例えば、ドナー候補からのHCTが移植片対宿主病（GVHD）を引き起こすリスクが低いことを示すためには、陰性予測遺伝子は、0.55以上（例えば、0.60、0.65、0.70、0.75、又は0.80）の発現量を有していなければならない。他の例において、陽性予測遺伝子は、ドナー候補からのHCTが移植片対宿主病（GVHD）を引き起こすリスクが高いことを示すためには、0.55以上（例えば、0.60、0.65、0.70、0.75、又は0.80）の発現量を有していなければならない。

20

30

【0056】

基準発現量（又は既定の基準値）は、より高い又はより低い閾値に設定されてもよい。したがって、このような基準発現量は、信頼性、正確性、及び再現性を向上させるため、及び、統計誤差等の変数を考慮してGVHDの決定又は予測のロバスト性を改善するために、調整可能であってもよい。一般に、HCTドナー候補がHCTレシピエントにおいてGVHDを引き起こすリスク又は可能性を減少又は最小化するために（すなわち、偽陰性を減少させるため、すなわち、レシピエントにおいてGVHDを引き起こす高いリスクを有するドナー候補を正しく予測するため）、ユーザは、HCTドナー候補のCD4陽性T細胞又はCD8陽性T細胞の遺伝子発現プロファイルにおいて、基準発現量を高く設定することにより、陰性予測遺伝子のより高い発現を選択することができ、及び/又は、基準発現量を低く設定することにより、陽性予測遺伝子のより低い発現を選択することができる。

40

50

【0057】

陽性又は陰性 G V H D 予測遺伝子について得られる発現量は、陽性又は陰性予測遺伝子の発現量を当該陽性又は陰性予測遺伝子の既定の基準発現量と比較する前に一又は複数の基準遺伝子の発現で調整又は正規化され得る。遺伝子発現のレベルを正規化する方法は、当業者に知られている。例えば、陽性又は陰性予測遺伝子の発現は、サンプル量、抽出効率、抽出量、又は測定用化学物質 (measurement chemistry) 又は器具性能 (instrumentation performance) の変化が遺伝子発現量又はレベルを測定するときに減少するように、基準遺伝子の mRNA レベルに対する当該遺伝子の mRNA レベルの相対比率に基づいて正規化され得る。基準遺伝子は、その発現が構成的であり C D 4 陽性 T 細胞又は C D 8 陽性 T 細胞において比較的一定のレベルで発現する遺伝子、又は、その発現が発現量を決定するために使用されない陽性又は陰性予測遺伝子である。特定の実施形態において、基準遺伝子は、(例えば、表 1 2 又は表 1 3 内の) ハウスキーピング遺伝子である。本明細書での用法のとおり、「ハウスキーピング遺伝子」は、その発現がサンプルごと組織ごと実質的に同じ遺伝子であり、又は、外部刺激に対して相対的に変化しにくい遺伝子である。ハウスキーピング遺伝子は、発現量が決定される陽性又は陰性予測遺伝子以外のサンプル RNA の正規化を可能にする任意の遺伝子であってもよく、各反応において加えられる総 RNA 量を正規化するために用いられるそれ以外の任意のマーカであってもよい。ハウスキーピング遺伝子の非限定的な例は、表 1 (RNA 1538)、表 2 A、表 2 B (RNA 192)、表 1 2 及び表 1 3 において、記号「H S K」で示されているものを含み、より具体的には、真核生物翻訳開始因子 4 H (E I F 4 H) 転写産物 1、3 ベータアクチン (A C T B)、アルドラーゼ A (A L D O A)、乳酸脱水素酵素 A (L D H A)、ホスホグリセリン酸キナーゼ 1 (P G K 1)、トランスフェリン受容体 (T F R C)、ベータチューブリン (T U B B)、ベータチューブリン 2 A (T U B B 2 A)、チオレドキシン (T X N)、ユビキチン C (U B C)、又はユビキチン活性化酵素 E 1 (U B E 1) を含む。一又は複数の G V H D 予測遺伝子に関して「組み合わせ」という語が用いられる場合には、2 つの予測遺伝子の最小限の組み合わせを意味し、より多くの予測遺伝子、例えば 3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、57、48、49、50、又はこれ以上の「n」個以下の陽性及び / 又は陰性 G V H D 予測遺伝子を意味してもよい。ここで、「n」は自然数である。このように、あくまで例示に過ぎず本発明を限定するものではないが、組み合わせは、二以上の陽性 G V H D 予測遺伝子の組み合わせ、二以上の陰性 G V H D 予測遺伝子の組み合わせ、又は二以上の陽性及び / 又は陰性 G V H D 予測遺伝子の組み合わせであってもよい。陽性及び / 又は陰性予測遺伝子のこのような組み合わせの数は、2 つの遺伝子の組み合わせについては 2^2 であり (P - P、P - N、N - P、N - N)、3 つの遺伝子の組み合わせについては 2^3 であり (P - P - P、P - P - N、P - N - P、P - N - N、N - P - P、N - P - N、N - N - P、N - N - N)、「n」個の遺伝子のより高次の組み合わせについては 2^4 、 2^5 、又は 2^n である。組み合わせの文脈において、例えば本発明の方法における比較工程において用いられる「既定の基準値」は、単一の発現量ではなく、発現量の結合を意味する。

【0058】

発現量の「結合」に関連して「線形結合」が用いられる場合には、最小限では、2 つの発現量の差 $X - Y$ 、又は 2 つの発現量の対数の差 $\log X - \log Y$ 、又は 2 つの発現量の和 $X + Y$ 、又は 2 つの発現量の対数の和 $\log X + \log Y$ 、又は 2 以上の発現量の組み合わせられた差及び / 又は和を意味する。ここで、任意の遺伝子の発現量又は発現量の対数には、因子「c」が乗じられる。「c」は実数であり、「c」の値は遺伝子の各々で異なってもよい。また、定数項「d」が任意の遺伝子の発現量に加算され又は減算される。ここで、「d」は実数であり、「d」の値は遺伝子の各々で異なってもよい。発現量に関連して「非線形結合」が用いられる場合には、最小限では、2 つの発現量の比 X / Y 、又は 2 つの発現量の対数の

10

20

30

40

50

比 $\log X / \log Y$ 、又は 2 つの発現量の積 $X \cdot Y$ 、又は 2 つの発現量の対数の積 $\log X \cdot \log Y$ 、又は 2 以上の発現量の組み合わせられた比及び / 又は積を意味する。which the of 任意の遺伝子の発現量又は発現量の対数は指数「b」で累乗される。ここで「b」は実数であり、「b」の値は遺伝子の各々で異なってもよい。また、任意の遺伝子の発現量又は発現量の対数には、因子「c」が乗じられる。ここで「c」は実数であり、「c」の値は遺伝子の各々で異なってもよい。また、定数項「d」が任意の遺伝子の発現量又は発現量の対数に加算又は減算される。ここで、「d」は実数であり、「d」の値は遺伝子の各々で異なってもよい。

【0059】

遺伝子発現の正規化は、競合関係にある予測遺伝子対、すなわち予測遺伝子対（レシオメトリック遺伝子対又はRGPと称する）における遺伝子2に対する遺伝子1の比が関連する予測モデルに関して簡単な方法で実行され、追加的な基準遺伝子の必要性がなくなる（実施例参照）。分離されたハウスキーピング遺伝子及び / 又は基準サンプルに対する陽性又は陰性予測遺伝子のレベルを報告する代わりに、予測遺伝子2に対する予測遺伝子1のレベル（その比）が、情報量が多い相対的な発現測定の前を提供する。

【0060】

したがって、陽性又は陰性GVHD予測遺伝子の発現量は、レシオメトリック遺伝子対（RGP）におけるのと同様に、比でも表され得る。遺伝子発現データの比は、様々な方法で表され得る。一実施形態において、発現量は、レシオメトリック遺伝子対（RGP）で示され、陽性又は陰性GVHD予測遺伝子の遺伝子発現の一又は複数の基準遺伝子に対する比で表される。より具体的な実施形態において、発現量は、レシオメトリック遺伝子対（RGP）で示され、陽性又は陰性予測GVHD遺伝子の遺伝子発現の基準遺伝子に対する比で表され、式「N/D」（分子/分母）で表される。ここで分子の「N」は陽性又は陰性GVHD予測遺伝子の発現レベルであり、分母の「D」は一又は複数の基準遺伝子の発現レベルである。N及びDの値は、一又は複数の陽性又は陰性GVHD予測遺伝子又は一又は複数の基準遺伝子の対応する発現の平均値又は中央値を反映していてもよいし、複数のサンプルにおける発現を反映していてもよい。このようなRGPは、陽性及び陰性のGVHD予測遺伝子（N-P及びP-N）の組み合わせ、陽性GVHD予測遺伝子（P-P）の組み合わせ、及び陰性GVHD予測遺伝子（N-N）の組み合わせを含む。

【0061】

このような発現量の決定、発現の正規化及び発現比の決定（例えば、RGP）に関し、基準遺伝子は、ハウスキーピング（HSK）遺伝子であってもよく、遺伝子発現の比を得るために用いられる陽性又は陰性予測遺伝子とは異なる陽性又は陰性GVHD予測遺伝子であってもよく、又は、ユーザによって選択される任意のこれら以外の遺伝子であってもよい。

【0062】

本発明によれば、GVHDに関して発現が測定される陽性及び陰性のGVHD予測遺伝子は、単一遺伝子の発現であれ、2つ（又はそれ以上の）遺伝子（RGP、遺伝子対等）の比であれ、遺伝子の組み合わせであれ、表1（RNA 1538）、表2A（RNA 143）、表2B（RNA 192）、表3、表13（SG175）、表15（SG128）、又は表18（SG64）に記載されこれらの表から選択されたもの、又はこれらの多型である。一実施形態において、その発現が測定される陽性又は陰性GVHD予測遺伝子の少なくとも一つは、表1（RNA 1538）、表2A（RNA143）、表2B（RNA 192）、表3、表13（SG175）、表15（SG128）又は表18（SG64）に記載された一又は複数の単一遺伝子（SG）から選択され、又は、表1（RNA 1538）、表2A（RNA143）、表2B（RNA 192）、表3、表13（SG175）、表15（SG128）又は表18（SG64）に記載されたレシオメトリック遺伝子対（RGP）又は単一遺伝子（SG）から選択される。例示的な非限定のレシオメトリック遺伝子対（RGP）は、表14（RGP348）及び表17（VmodRGP100）に記載されたものでありこれらから選択され、「遺伝子対の対」等の比における例示的な複数の遺伝子の非限定の例は、表16（「PRGP348」）に記載されこの表から選択され得る。したがって、表1（RNA 1538）、表2A（RNA

10

20

30

40

50

143)、表 2 B (RNA 192)、単一遺伝子の発現、遺伝子の比(例えば、RGP)、及び表 3、表 1 3 (SG175)、表 1 5 (SG128)、表 1 7 (VmodRGP100)及び表 1 8 (VmodSG64)のうちの任意のものにおける陰性 G V H D 予測遺伝子、陽性 G V H D 予測遺伝子及び/又は陰性及び陽性 G V H D 予測遺伝子の混合の多重遺伝子比(multi-gene ratios)を含む遺伝子の組み合わせは、任意に組み合わせ、本発明を実行するために行われる。

【0063】

本発明のより具体的な実施形態において、ドナー候補からの造血細胞移植(HCT)がHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こす又は引き起こさないリスクを予測又は決定するために用いられる陰性及び/又は陽性GVHD予測遺伝子は、表 1 8 (VmodSG64)に記載された一又は複数の遺伝子から選択される。他のより具体的な本発明の実施形態において、ドナー候補からの造血細胞移植(HCT)がHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こす又は引き起こさないリスクを予測又は決定するために用いられる陰性及び/又は陽性GVHD予測遺伝子は、表 1 (RNA 1538)、表 2、表 2 A (RNA 143)、表 2 B (RNA 192)、表 3、表 1 3 (SG175)、表 1 5 (SG128)又は表 1 8 (SG64)に記載されている遺伝子から選択される 2 以上の遺伝子の複数のレシオメトリック遺伝子対(RGP)である。さらに具体的な本発明の実施形態において、ドナー候補からの造血細胞移植(HCT)がHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こす又は引き起こさないリスクを予測又は決定するために用いられるレシオメトリック遺伝子対(RGP)は、表 1 7 (VmodRGP100)に記載された遺伝子から選択される一又は複数の遺伝子対(RGP)であり。本発明の具体的な追加実施形態において、陰性及び/又は陽性GVHD予測遺伝子は、単一遺伝子(SG)及びレシオメトリック遺伝子対(RGP)の組み合わせを含み、ドナー候補からの造血細胞移植(HCT)がHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすか又は引き起こさないリスクを予測又は決定するために、表 1 8 (VmodSG64)に記載されている単一遺伝子(SG)から選択される複数の遺伝子及び表 1 7 (VmodRGP100)に記載されているRGPから選択されるレシオメトリック遺伝子対(RGP)である。

【0064】

本発明によれば、複数の陽性及び/又は陰性GVHD予測遺伝子が発現のために測定又は分析される場合には、HCTドナー候補がHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすリスクが高いか低いかを予測又は決定するために、典型的には、遺伝子の数、又は発現レベルもしくは発現量又は評価される遺伝子の種類の閾値(例えば、最小値)が存在する。評価は、一又は複数の基準に基づく分析を意味する。この基準には、閾値発現レベルよりも大きい又は小さい遺伝子発現、又は、ユーザによって設定可能な閾値より多い又は少ない陽性及び/又は陰性GVHD予測遺伝子の数、又はその発現がGVHD転帰と高い相関を有する傾向がある特定の遺伝子のGVHDの予測方向が含まれるがこれらには限られない。このような基準の全ては、ユーザによって設定可能であり、信頼性又は正確性の所望の程度に基づくものであってもよい。非限定的な例として、発現のために測定又は分析される単一遺伝子(SG)、遺伝子発現比(例えば、RGP)又は多重遺伝子比(例えば、表 1 6 等のPRGP)の数は、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、57、48、49、50、又はそれ以上の遺伝子及び/又は遺伝子発現比である。

【0065】

遺伝子の数又は発現レベル、又は発現量の線形結合又は非線型結合は、発現が測定された遺伝子の総数の百分率によって表すことが可能で、例えば、陽性及び/又は陰性予測遺伝子の総数の少なくとも30%、40%、50%、60%、70%、80%又はそれ以上である。このように、10個の予測遺伝子の全体の発現が測定される場合には、閾値は、遺伝子の3、4、5、6、7、8又はそれ以上であり、HCTがGVHDを引き起こすリスクが高いか低いかを予測又は決定するために、HCTのリスクが低いこと又は高いことを

示さなければならない。特定の実施形態において、HCTドナー候補がHCTレシピエントにおいて移植片対宿主病（GVHD）を引き起こすリスクが高いと予測又は決定するには、陽性又は陰性GVHD予測遺伝子の大多数がHCTレシピエントにおいて移植片対宿主病（GVHD）を引き起こすリスクが高いことを示さなければならない。又は、HCTドナー候補がHCTレシピエントにおいて移植片対宿主病（GVHD）が引き起こすリスクが低いと予測又は決定するためには、陽性又は陰性GVHD予測遺伝子の大部分が、HCTレシピエントにおいて移植片対宿主病（GVHD）を引き起こすリスクが低いことを示さなければならない。特定の実施形態において、ドナー候補からのHCTがGVHDを引き起こすリスクが高いことを示す陽性又は陰性GVHD予測遺伝子又は陽性及び/又は陰性GVHD予測遺伝子の組み合わせの数が、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いことを示す陽性又は陰性予測遺伝子又は陽性及び/又は陰性GVHD予測遺伝子の組み合わせの数よりも多いとき、ドナー候補のHCTがHCTレシピエントにおいてGVHDを引き起こすリスクは高いと予測又は決定される。さらに他の実施形態において、ドナー候補からのHCTがGVHDを引き起こすリスクが低いことを示す陽性又は陰性GVHD予測遺伝子又は陽性及び/又は陰性GVHD予測遺伝子の組み合わせの数が、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが高いことを示す陽性又は陰性予測遺伝子又は陽性及び/又は陰性GVHD予測遺伝子の組み合わせの数よりも多い場合、ドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすリスクが低いと予測又は決定される。

10

20

【0066】

さらに具体的な実施形態において、HCTドナー候補がHCTレシピエントにおいて移植片対宿主病（GVHD）を引き起こすリスクが高いと予測又は決定するためには、陽性又は陰性予測遺伝子の少なくとも66%が、HCTレシピエントにおいて移植片対宿主病（GVHD）を引き起こすリスクが高いことを示さなければならない。または、HCTドナー候補がHCTレシピエントにおいて移植片対宿主病（GVHD）を引き起こすリスクが低いと予測又は決定するためには、陽性又は陰性予測遺伝子の少なくとも66%がHCTレシピエントにおいて移植片対宿主病（GVHD）を引き起こすリスクが低いことを示さなければならない。追加的な具体的実施形態において、HCTドナー候補がHCTレシピエントにおいて移植片対宿主病（GVHD）を引き起こすリスクが低いと予測又は決定するためには、陽性又は陰性予測遺伝子の少なくとも75%がHCTレシピエントにおいて移植片対宿主病（GVHD）を引き起こすリスクが低いことを示さなければならない。

30

【0067】

例示として、レシピエントにおけるGVHDのリスクを調べる非限定的なモデルの1つは、GVHDを引き起こす又は引き起こさないリスクを調べるために、その発現が分析される陽性及び/又は陰性GVHD予測遺伝子の各々に「票」を割り当てることである。この票は、測定された各遺伝子から得られた発現量又は発現量の組み合わせがGVHDのリスクが高いこと又は低いことを示すかに応じて、一覧表にされる。例えば、合計10の陽性及び/又は陰性予測遺伝子の発現が測定された場合、過半数（すなわち、10のうち6つ）はGVHDのリスクが低いことを示し、10のうち4つはGVHDのリスクが高いことを示す。このように、6つの遺伝子がGVHDのリスクが低いことを表明する場合には、4つの遺伝子がGVHDのリスクが高いことを示す。遺伝子及びGVHDのリスクを正確に予測する遺伝子の能力によっては、10の遺伝子のような過半数の投票は、GVHDを引き起こすリスクが低いと結論付けるのに十分なことがある。予測においてより高い信頼性が望まれる場合には、特定のGVHD転帰を予測するのに求められる遺伝子「票」の閾値の値を増加させることができる。例えば、10の遺伝子の6から7へ、又は10の遺伝子の6から8へ、又はそれ以上である。

40

【0068】

本発明によれば、例示的なモデルの1つは、その発現が測定される各遺伝子に「票」を割り当て、票が割り当てられた各遺伝子から得られた発現量に重み、また票の合計に基づい

50

て、GVHDを引き起こす又は引き起こさないリスクが決定又は予測される。一実施形態において、陰性又は陽性GVHD予測遺伝子の複数の発現量が決定され、遺伝子の発現量がドナー候補又は実際のドナーがGVHDを引き起こすリスクを示すか又は引き起こさないリスクを示すかに応じて、各陰性又は陽性予測遺伝子に票が割り当てられる。続いて、HCTレシピエントにおけるGVHDを引き起こすこと又は引き起こさないことを示す票又は示さない票の総数に基づいて、ドナー候補又は実際のドナーにスコアが割り当てられる。具体的な態様において、50%以上の票がGVHDを引き起こすことを示す場合には、スコアは、ドナー候補又は実際のドナーからの造血細胞移植(HCT)がHCTレシピエントにおいてGVHDを引き起こすリスクが高いことを反映したものとなる。または、50%以上の票がGVHDを引き起こさないことを示す場合には、スコアは、ドナー候補又は実際のドナーからの造血細胞移植(HCT)がHCTレシピエントにおいてGVHDを引き起こすリスクが低いことを反映したものとなる。追加的な態様において、少なくとも55%、60%、65%、70%、75%、80%、85%、90%、95%、又はそれ以上の票がGVHDを引き起こすことを示す場合には、スコアは、ドナー候補又は実際のドナーからの造血細胞移植(HCT)がHCTレシピエントにおいてGVHDを引き起こすリスクが高いことを反映したものとなる。または、少なくとも55%、60%、65%、70%、75%、80%、85%、90%、95%、又は又はそれ以上の票がGVHDを引き起こさないことを示す場合にはスコアは、ドナー候補又は実際のドナーからの造血細胞移植(HCT)がHCTレシピエントにおいてGVHDを引き起こすリスクが低いことを反映したものとなる。

10

20

【0069】

GVHDを引き起こす及び引き起こさないリスクを予測又は決定する多数の非限定的で代表的な投票モデル(Vmods)が本明細書に開示される。このような非限定的な投票モデルの例には、SG43RGP46-GPperformance、SG42RGP21-GPminimalist、SG43RGP37-GPconnectivity、SG43RGP51-PRGPminranksort、SG43RGP55-PRGPmedranksort、SG43RGP36-RGPgreedysearch、又はSG21RGP28-RGPmaxgreedysearchにおいて説明される単一遺伝子(SG)及びレシオメトリック遺伝子対(RGP)の組み合わせが含まれる。これらの組み合わせの各々は、SG及びRGPを含む。投票モデル(Vmods)の各々に組み込まれ、その発現が測定されるSG及びRGPは、表17及び表18において「x」で示されている。

30

【0070】

本発明の方法は、典型的には、HCTドナーとHCTレシピエントとの10/10HLAマーカー座適合(10 out of 10 HLA マーカー-loci matches)を有するGVHD陰性HCTドナーの特定よりも優れている。特定の実施形態においては、HCTレシピエントにおいてGVHDを引き起こす又は引き起こさないHCTドナーを予測する方法は、少なくとも60%、少なくとも70%、少なくとも80%、又は少なくとも90%の正確性を有する。他の具体的な実施形態において、GVHD陰性ドナーの予測の正確性は、HCTレシピエントとの10/10HLAマーカー座適合(10 out of 10 HLA マーカー-loci matches)によって陰性に分類されたHCT候補ドナーのグループ内において、GVHD陰性ドナーを正しく特定する可能性又はリスクである。

40

【0071】

本明細書での用法のとおり、「測定」又は「分析」という語は、発現の決定又は遺伝子発現の定量化の文脈において、絶対的な又は相対的な定量化を意味する。遺伝子発現の文脈において、測定は、CD4陽性T細胞又はCD8陽性T細胞等のサンプル又は標本を用いて実施される一又は複数の分離、精製、プロセッシング(processing)、操作、抽出、又は決定の工程を含む検査法を意味し、一又は複数の遺伝子の発現量は、いかなる心理的な工程とも区別される。絶対定量化は、未知のものと既知の標的核酸又は発現産物(例えば、標準曲線の生成を通じて)のハイブリダイゼーション又は結合強度を参照し、一又は複数の標的核酸又は発現産物の既知の濃縮を含めることによって実現しても良い。または、相対定量化は、2以上の遺伝子又は2以上のサンプルの間で信号を比較して、信号の変化、及び含意的に転写産物又は発現産物及び遺伝子発現量を定量化することにより実現される

50

。

【 0 0 7 2 】

比較は、目視検査によって又はコンピュータアルゴリズムを用いて実現され得る。アルゴリズムの例は、include 線形又は非線形回帰アルゴリズム、線形又は非線形分類アルゴリズム、ANOVA（分散分析）、コンピュータによるニューラル・ネットワークアルゴリズム、コンピュータによる遺伝的アルゴリズム、サポートベクターマシナアルゴリズム、階層分析又はクラスタリングアルゴリズム、決定木を用いた階層アルゴリズム、カーネルベースマシナアルゴリズム、ルックアップテーブルアルゴリズム、部分最小二乗法アルゴリズム等の判別アルゴリズム、マッチング追跡アルゴリズム、重判別分析アルゴリズム、主成分分析アルゴリズム、特異値分解アルゴリズム、ベイズ推定機能アルゴリズム、マルコフブランケットアルゴリズム、隠れマルコフアルゴリズム、決定的最適化（deterministic optimization）アルゴリズム、確率的探索最適化（stochastic search optimization）又は焼きなまし法アルゴリズム、再帰的将来除去（recursive feature elimination）又はエントロピーに基づく再帰的将来除去（entropy-based recursive feature elimination）アルゴリズム、組み合わせられたアルゴリズム、コミッティーネットワークにおいて準備された複数のアルゴリズム、及び前方浮動探索（forward floating search）又は後方浮動探索（backward floating search）アルゴリズムを含む。一又は複数の単一遺伝子又はレシオメトリック遺伝子対（RGP）、遺伝子対の対（PRGP）等を用いてGVHD転帰を決定又は予測する値を得るためのさらなる方法は、実施例20において説明される。

10

【 0 0 7 3 】

HCTドナー候補及び実際のHCTドナー並びにHCTレシピエントは動物を含み、典型的にはヒト等のほ乳類の動物（ほ乳類）を含む。ヒトには、HCTレシピエント候補と遺伝的に関連がある家族が含まれるが、これには限られない。ヒトには、HCTレシピエント候補と遺伝的に関連がない家族以外の者も含まれ、HCTレシピエント候補とHLA適合を有する非家族の実際のHCTドナー又はHCTドナー候補を含む。より具体的には、実際のHTCドナー又はHCTドナー候補及びHCTレシピエントは、10/10（10 out of 10）又は9/10（9 out of 10）ヒト白血球抗原（HLA）マーカー座適合、例えば、HLA-A、HLA-B、HLA-C、HLA-DRB1及びHLA-DQB1座のHLAマーカー座適合、又はHLA-A座適合、HLA-B座適合、HLA-C座適合、HLA-DRB1又はHLA-DQB1座適合の4つの任意の組み合わせを有する。このようなHLAマーカー座適合は、血清学的に又はHLA遺伝子の配列分析のいずれかによって決定され得る。分析に適した動物は、他の動物、例えばHCT・GVHDの動物モデルのHCTレシピエントのためのHCTドナーとなりえる動物を含む。

20

30

【 0 0 7 4 】

GVHDを引き起こす実際のドナー又はHCTドナー候補を定義するために、HCTレシピエントがドナーからの移植後にGVHDの症状を示した場合には、当該HCTドナーはGVHDを引き起こしたと考える。GVHDを引き起こさない実際のドナー又はHCTドナー候補を定義するために、HCTレシピエントがドナーからの移植後にGVHDの症状を示さなかった場合には、当該HCTドナーはGVHDを引き起こさないと考える。または、当該ドナーからの移植後に、レシピエントが最小限の重篤な急性GVHD、すなわち急性グレードIのGVHDの症状しか示さず、HCT後にそれ以外の急性又は慢性のGVHDの症状を示さない場合、実際のHCTドナー又はHCTドナー候補は、GVHDを引き起こさないHCTドナーと定義され得る。

40

【 0 0 7 5 】

GVHDは、症状の重症度及び期間に応じて分類又はグループ化され、本明細書ではグループ1からグループ6に分類される。このグループは、概して、重症度の違いを反映する。例示的な分類は、急性及び慢性のGVHDをいずれも示さないグループ1で始まり、グレード3又はグレード4の重篤な急性GVHD及び広範囲な慢性GVHDを示すグループ6で終わる。グループ5も、グレード3又はグレード4GVHDを示すが、慢性GVHDは示さない。グループ4及びグループ3は、グレード1又はグレード2急性GVHDを示

50

し、グループ4は慢性GVHDがあり、グループ3は慢性GVHDがない。グループ2は、慢性GVHDのみを示し、急性GVHDは示さない。グレード3又はグレード4の急性GVHDは、最も激しく及び生命を脅かすGVHDの態様を特徴付けるが、グレード1又はグレード2の急性GVHDは、重症度ははるかに軽微で軽度と考えられることもある。急性GVHDのグレードの分類は、GVHDの重症度を分類する医師にとって医療実務において確立された複数の症状の診断的な分類である。グループの定義は、医学的に意味のあるGVHDの重症度のグループである。これ以外の分類も可能である。例えば、グループ1～グループ6は本明細書で定義された用語であるが、急性GVHD、慢性GVHD、グレード0～グレード4は、確立され、受け入れられ、医学的に定義された用語である。

【0076】

本発明の方法は、さらに実際のHTCドナー又はHCTドナー候補にスコアを割り当てること、又は、実際のHCTドナー又はHCTドナー候補を特定することを含む。このようなスコア又は特定は、HCTドナー遺伝子発現プロファイル、HCTドナーの陽性及び/又は陰性予測遺伝子の発現量、HCTドナー候補に関する情報の総計、又はHLAマーカー座プロファイルに基づくものである。スコア又は特定は、リスク予測又は決定に基づいて、実際のドナー又はドナー候補のHCTがHCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こす又は引き起こさない可能性又はリスクを反映することができる。スコア又は特定は、起こることが予想され又は決定されたGVHDの分類又はグループを反映することもできる。このGVHDの分類又はグループは、GVHD転帰又は重症度(例えば、本明細書においてはグループ1～6で定義され、又は慢性GVHDを伴う又は伴わないグレードI、II、III又はIVの急性GVHDによって定義され、又は急性GVHDを伴わない慢性GVHDで定義される)を示す。

【0077】

本明細書での用法のとおり、本発明は、CD4陽性T細胞における、陰性及び/又は陽性GVHD予測遺伝子また基準遺伝子(例えば、HSK遺伝子)を含む遺伝子の発現レベルの分析によって例示される。本発明の方法は、他の種類のT細胞を用いることもできる。例えば、本発明の方法は、CD8陽性T細胞における、陰性及び/又は陽性GVHD予測遺伝子また基準遺伝子(例えば、HSK遺伝子)の発現レベルを解明することができる。したがって、本発明は、様々なT細胞を用いて実施することができる。このT細胞は、CD4陽性T細胞、CD8陽性T細胞、制御性T細胞、及びこれらの及び他のT細胞亜群の混合物を含むがこれらには限定されない。

【0078】

生体サンプルは、生体物質を有し得る任意のサンプルを含む。生体サンプルには、HCTドナー候補の細胞物質を含む任意の生体物質が含まれる。典型的には、このようなサンプルには、免疫細胞、例えば、CD4陽性T細胞及び/又はCD8陽性T細胞が含まれる。したがって、生体サンプルには、GVHD転帰の予測又は決定のためのHCTドナー候補の一又は複数の陽性及び/又は陰性予測遺伝子の発現の測定又は分析に適したDNA、RNA又はポリペプチド(タンパク質)等の核酸を含む生体物質又は生体液又は任意の物質が含まれる。生体サンプルは、したがって、一又は複数の陽性及び/又は陰性予測遺伝子の発現の測定又は分析に適したものでありさえすればよく、GVHD転帰と相関のある核酸及び/又はタンパク質を含む。典型的には、生体サンプルには、CD4陽性T細胞、CD8陽性T細胞、又は細胞物質が含まれる。非限定的な例には、血液、血液細胞(例えば、末梢血単核細胞)、血清、血漿、骨髄、唾液、排泄物、脳脊髄液、又は尿が含まれる。

【0079】

生体サンプルは、例えば、遺伝子発現又は発現産物の量又はレベル又は機能の存在を決定し、これらの測定又は分析を行うために、形質転換され、プロセスされ、又は操作され得る。典型的には、生体サンプルは、一又は複数の陽性及び/又は陰性GVHD予測遺伝子の発現及び/又は量もしくはレベルを直接的又は間接的に示す核酸(例えば、全核酸又はmRNA)又は遺伝子発現産物(例えば、タンパク質又は断片)を精製又は分離するために、形質転換又はプロセスされる。このように、サンプルには、HCTドナー候補のCD4陽

10

20

30

40

50

性T細胞又はCD8陽性T細胞から精製され、分離され、由来し、抽出され、又は取得された核酸及びタンパク質も含まれる。

【0080】

陰性及び/又は陽性GVHD予測遺伝子発現レベルは、HCTドナー候補のCD4陽性T細胞又はCD8陽性T細胞を含むサンプルのmRNA(又は当該mRNAから逆転写されたcDNA)を測定することにより決定され得る。陰性又は陽性GVHD予測遺伝子は、タンパク質をコード化することができる。したがって、遺伝子発現レベルは、ポリペプチド又はタンパク質等の発現産物を測定することにより決定され得る。表1(RNA1538、配列番号:1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)、又は表18(SG64)に記載された陰性及び/又は陽性予測遺伝子によってコード化された転写産物及び/又はタンパク質の発現は、当業者に知られた任意の様々な方法によって、測定及び/又は分析され得る。

10

【0081】

検出、測定又は分析に適した核酸サンプルには、関連する転写産物(すなわち、HCTレシピエントにおけるGVHDを引き起こすHCTの陽性及び/又は陰性予測遺伝子由来の転写産物、RNA、プリプロセスRNA(preprocessed RNA)、又はmRNA)。このように、RNA発現(例えば、mRNA)を測定又は分析するときには、かかるRNAは直接測定され得る。

【0082】

スクリーニングのために適した核酸サンプルもまた、関連する転写産物(例えば、陽性及び/又は陰性予測遺伝子由来のmRNAから転写されたcDNA)に由来する核酸を含む。転写産物に由来する核酸は、その合成においてmRNA転写産物又はその部分配列が(究極的には)鋳型として働く核酸を意味する。このような核酸の例は、転写産物から逆転写されるcDNA、cDNAから転写されるRNA、cDNAから増幅されるDNA、増幅されたDNAから転写されるRNA等を含み、これらは全て転写産物に由来する。このような由来産物の測定は、陽性及び/又は陰性遺伝子発現の存在及び/又は量を示す。例えば、cDNAの量は発現したRNAの量と相関するので、陽性又は陰性予測遺伝子のRNAは、cDNAに転写され(相補DNA)、その後測定される。

20

【0083】

一般に、サンプル中の核酸(例えば、DNA又はRNA)は、遺伝子配列又は発現又は量を測定又は検出する任意の好適な方法又は技術により検出することができる。遺伝子(例えば、核酸発現)を測定する非限定的で例示的な方法には、ポリメラーゼ連鎖反応(PCR)、逆転写ポリメラーゼ連鎖反応(RT-PCR)、in situ PCR、定量PCR(q-PCR)、in situハイブリダイゼーション、サザンブロット、ノーザンブロット、配列分析、マイクロアレイ分析、レポーター遺伝子検出、又はこれら以外の核酸ハイブリダイゼーションプラットフォームが含まれるがこれらには限られない。RNAの発現を測定する方法には、本明細書において説明する陰性及び/又は陽性予測遺伝子の一又は複数の全部又は一部の転写産物とハイブリダイズする標識プローブを用いた細胞性mRNA抽出及びノーザンブロットイング、特異的プライマーを用いて陰性及び/又は陽性予測遺伝子の一又は複数から発現したmRNAの増幅、ポリメラーゼ連鎖反応(PCR)、定量PCR(q-PCR)、及び逆転写ポリメラーゼ連鎖反応(RT-PCR)(これらの産物は、定量検出及び細胞からの全RNA抽出が行われ、その後プロセスされ(例えば逆転写又は増幅され)、陰性及び/又は陽性予測遺伝子の全部又は一部をコード化するcDNAs又はオリゴヌクレオチドを標識化し調査するために用いられる)、及びin situハイブリダイゼーションを含むがこれらには限られない。陽性及び陰性のGVHD予測遺伝子及びハウスキーピング遺伝子に対応するRT-PCR用のプライマーは、例えば、表2B(RNA192)に記載されており、商業的に利用可能なABI Assay ID番号に従って記述されている。他のプライマー及びプローブは、表1(RNA1538、配列番号:1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)、及び表18(SG64)に記載された遺伝子配列に由来し、又は、これらの遺伝子配列に基づくものであってもよい。

30

40

【0084】

50

全RNA又はmRNA等のRNAを分離する方法は、当業者に知られている。非限定的な例には、例えば、生体サンプルから全核酸を得るための酸性グアニジン・フェノール・クロロホルム抽出、及び、オリゴdTカラムクロマトグラフィー又は(dT)_n磁性ビーズを用いたmRNAの分離が含まれる(例えば、Sambrook et al., *Molecular Cloning: A Laboratory Manual* (2nd ed.), Vols. 1-3, Cold Spring Harbor 研究所, (1989), 又はCurrent Protocols in Molecular Biology, F. Ausubel et al., ed. Greene Publishing and Wiley-Interscience, New York (1987)参照)。

【0085】

核酸が増幅される実施形態においては、任意の増幅方法を用いることができるが、遺伝子発現量又はレベルを反映する結果が望まれる場合には、定量増幅を実現するために、増幅された核酸の相対頻度を維持又は制御する方法が使用される。「定量」増幅の様々な方法が当業者に知られている。例えば、定量PCRは、同一のプライマーを用いて既知量の対照配列を同時に増幅することを含む。これにより、PCR反応を調整するために用いることができる内部標準が提供される。このように、内部標準に特異的なプライマー及び/又はプローブは、増幅された核酸を定量化するために用いることができる。これ以外の好適な増幅方法には、ポリメラーゼ連鎖反応(PCR、Innis, et al., *PCR Protocols. A Guide to Methods and Application*. Academic Press, Inc. San Diego, (1990))、リガーゼ連鎖反応(LCR、Wu and Wallace, *Genomics*, 4:560; Landegren et al., *Science*, 241: 1077; and Barringer, et al., *Gene*, 89:117))、転写産物増幅(Kwoh et al., *Proc. Natl. Acad. Sci. USA*, 86:1173)、及び自家持続配列複製法(Guatelli et al., *Proc. Nat. Acad. Sci. USA*, 87:1874)が含まれるが、これらには限られない。

【0086】

したがって、遺伝子発現レベルは、概して、細胞のmRNA(又はそのcDNA)等のRNAを検出し、及び/又は、ポリペプチド又はタンパク質等の遺伝子発現産物を検出することにより、測定又は分析することができる。表1(RNA 1538、配列番号: 1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)、及び表18(SG64)に記載された陽性及び/又は陰性予測遺伝子によってコード化された転写産物及び/又はタンパク質の発現は、当業者に既知の様々な任意の方法によって測定され得る。したがって、本発明に係る検体は、核酸配列を含む。

【0087】

本明細書での用法のとおり、「核酸」及び「ポリヌクレオチド」等の用語は、リン酸エステル結合又は同等の共有結合によって結合している少なくとも2以上のリボ核酸塩基又はデオキシリボ核酸塩基(ヌクレオチド)を意味する。核酸には、ポリヌクレオチドが含まれる。核酸には、一本鎖、二本鎖又は三本鎖、環状又は線状、核酸分子が含まれる。核酸には、センス配列及びアンチセンス配列が含まれ、例えば、表1(RNA 1538)、表2、表2A、表2B(RNA 192)、及び/又は、表3に記載された任意の配列、又は、表1(RNA 1538)、表2、表2A、表2B(RNA 192)及び/又は表3に記載された任意の配列の相補的配列の全部又は一部に結合するセンス配列及びアンチセンス配列が含まれる。例示的な核酸には、全RNA、mRNA、DNA、cDNA、ゲノム核酸、天然型及び非天然型核酸例えば合成核酸が含まれるが、これらには限られない。

【0088】

核酸は、様々な長さを取りえる。核酸の長さは、典型的には、約10ヌクレオチドから20Kbであり、又は、かかる長さ内の又はかかる長さを含む任意の数値であり、例えば、10ヌクレオチドから250Kb、1~15Kb又はそれ以下、1000~約5000ヌクレオチド又はそれ以下、500~1000ヌクレオチドの長さである。核酸は、もっと短くともよい。例えば、100から約500ヌクレオチド、又は約10~25、25~50、50~100、100~250、又は約250~500ヌクレオチドの長さであってもよく、又は、かかる長さ内の又はかかる長さを含む任意の数値である。具体的な態様において、核酸配列は、約5~10、10~20、20~30、30~40、40~50、50~60、60~70、70~80、80~90、90~100、100~150、1

10

20

30

40

50

50～200、200～250、250～300、300～400、400～500、500～1000、1000～2000ヌクレオチドの長さを有し、又は、かかる長さ内の又はかかる長さを含む任意の数値である。より短いポリヌクレオチドは、一般に「オリゴヌクレオチド」又は一本鎖DNA又は二本鎖DNAの、「プローブ」又は「プライマー」とよばれ、典型的には約10～20、20～30、30～50、50～100ヌクレオチドの長さを有する。しかしながら、このようなオリゴヌクレオチドの長さには上限はない。

【0089】

核酸には、例えば、陰性及び/又は陽性予測遺伝子配列（又は転写産物、RNA又はそのcDNA）、例えば、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、又は表18（SG64）に記載された任意の配列の全部又は一部、又は、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）に記載された任意の配列の全部又は一部に相補的な配列、とハイブリダイズするポリヌクレオチド及びオリゴヌクレオチド（プライマー及びプローブ）が含まれる。このようなハイブリダイズした核酸は、標的配列、転写産物、又は相補的な又は増幅された配列の検出を可能とし、HCTがHCTレシピエントにおいてGVHDを引き起こす又は引き起こさないリスクを予測又は決定する本発明の方法、キット及びアレイにおいて用いられ得る。

【0090】

陰性及び/又は陽性予測遺伝子の発現を検出又は測定するために、核酸（例えば、オリゴ-又はポリヌクレオチドプローブ又はプライマー）は、対応する陰性及び/又は陽性予測遺伝子配列（又はRNA 転写産物又はそのcDNA）又は相補的配列の全部又は一部、すなわち、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、表18（SG64）に記載された任意の配列の全部又は一部、又は、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、表18（SG64）に記載された任意の配列の全部又は一部に相補的な配列と「ハイブリダイズ」する。これは、2以上の核酸配列の結合を意味する。「十分に相補的な」配列は、核酸配列の標的配列（陰性及び/又は陽性予測遺伝子配列、又は 転写産物、RNAもしくはそのcDNA）、例えば、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）に記載された任意の配列の全部又は一部、又は、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）に記載された任意の配列の全部又は一部と相補的な配列、との安定的なハイブリダイゼーションを可能にする。したがって、2つの配列が完全に相補的でなくとも検出が可能となる。検出は、直接的であってもよく（すなわち、配列と直接ハイブリダイズしたプローブから得られる）、間接的であってもよい（すなわち、プローブを標的配列に結びつける中間分子構造とハイブリダイズしたプローブから得られる）。

【0091】

ハイブリダイズした配列は、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、又は表18（SG64）に記載された任意の配列の全部又は一部、又は、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、又は表18（SG64）に記載された任意の配列の全部又は一部と相補的な配列と、概して約50%以上相補的となる。典型的には、ハイブリダイズした配列は、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、又は表18（SG64）に記載された任意の配列の全部又は一部、又は、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、又は表1

8 (SG64) に記載された任意の配列の全部又は一部と相補的な配列と、60%、70%、80%、85%、90%、又は95%又はそれ以上相補的である。ハイブリダイズした配列間のハイブリダイゼーション領域は、典型的には、少なくとも約5~10、10~15ヌクレオチド、15~20ヌクレオチド、20~30ヌクレオチド、30~50ヌクレオチド、50~75ヌクレオチド、75~100ヌクレオチド、100~200ヌクレオチド、300~400ヌクレオチド、400~500ヌクレオチド又はそれ以上、この長さ内の又はこの長さを含む任意の数値又は領域である。

【0092】

二本鎖分子を形成するための二本鎖の核酸の相補領域間のハイブリダイゼーションは、ハイブリダイゼーション方法の性質並びにハイブリダイズした核酸配列の構成及び長さによって変わる。一般に、ハイブリダイゼーションの温度及びハイブリダイゼーションバッファのイオン強度(例えばNaイオン濃度)が、will determine the of ハイブリダイゼーションの厳密さ(stringency)を決定する(特定の厳密さを達成するためのハイブリダイゼーション条件はSambrook et al., (1989)Molecular Cloning, second edition, Cold Spring Harbor 研究所, Plainview, N.Y.で報告されている)。

【0093】

以下に、例示的な非限定的なハイブリダイゼーション条件を示す。非常に高い厳密さ(90%の同一性を有する配列を検出) - ハイブリダイゼーション: 5 X S S C、65°Cで16時間、洗浄: 2 X S S C中で室温で2度(RT)、各々15分間、洗浄: 0.5 X S S C中で65°Cで2度、各々20分間

高い厳密さ(80%以上の同一性を有する配列を検出) - ハイブリダイゼーション: 5~6 X S S C、65°C~70°Cで16~20時間、洗浄: 2 X S S C中で2度、各々5~20分間(RT)、洗浄: 1 X S S C中で55°C~70°Cで2度、各々30分間。

低い厳密さ(50%以上の同一性を有する配列を検出) - ハイブリダイゼーション: 6 X S S C、室温~55°Cで16~20時間、洗浄: 3 X S S C中で室温~55°Cで少なくとも2度、各々20~30分間。

【0094】

したがって、様々な実施形態において、ハイブリダイゼーションのためのポリヌクレオチド及びオリゴヌクレオチド(プライマー及びプローブ)は、陽性又は陰性予測遺伝子又はこれらの多型から生成されるRNA転写産物に対する(例えば、接触)オリゴ-又はポリ-ヌクレオチドプローブ、又は、陽性又は陰性予測遺伝子又はこれらの多型のRNA転写産物に由来するcDNAとのオリゴ-又はポリ-ヌクレオチドプローブのハイブリダイゼーションを含む。具体的な実施形態において、ハイブリダイゼーションのためのポリヌクレオチド及びオリゴヌクレオチド(プライマー及びプローブ)は、陽性又は陰性GVHD予測遺伝子配列又はその断片に結合する(例えば、表1(RNA 1538、配列番号: 1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)、又は表18(SG64)に記載された遺伝子の全部又は一部、又は、表1(RNA 1538、配列番号: 1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)、又は表18(SG64)に記載された任意の配列の全部又は一部と相補的な配列に結合する)(例えば、接触)オリゴ-又はポリ-ヌクレオチドプローブを含む。したがって、このような配列は、表1(RNA 1538、配列番号: 1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)、及び表18(SG64)に記載された配列、及び、表1(RNA 1538、配列番号: 1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)、及び表18(SG64)に記載された任意の配列の全部又は一部と50%、60%、70%、80%、85%、90%、又は95%同一の配列、又は、表1(RNA 1538、配列番号: 1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)、又は表18(SG64)に記載された任意の配列の全部又は一部と相補的な配列の断片を含む。

【0095】

本発明の方法、アレイ及びキットにおいては、複数のポリヌクレオチドを用いることができる。複数のポリヌクレオチド（例えば、プローブ又はプライマー対）は、陽性及び/又は陰性予測遺伝子、例えば、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）に記載された任意の配列又はこれらの多型の発現を検出、測定又は分析するために用いることができる。

【0096】

「相補的」又は「アンチセンス」という語は、特定のDNA又はRNA配列、例えば、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）に記載された任意の配列の全部又は一部、又は、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）に記載された任意の配列の全部又は一部と相補的な配列、に結合することができるポリヌクレオチド又はペプチド核酸（PNA）を意味する。アンチセンスは、RNA転写産物又はDNAを結合する一本鎖、二本鎖、三本鎖又はこれ以上の鎖のRNA及びDNAポリヌクレオチド及びペプチド核酸（PNA）を含む。具体例は、センスRNAに結合するRNA及びDNAアンチセンスを含む。例えば、一本鎖核酸は、陰性及び/又は陽性予測遺伝子の転写産物を標的にすることができる。アンチセンス/センス分子は、典型的にはセンス/アンチセンス鎖と100%相補的であるが、部分的に相補的であってもよい。このとき、ヌクレオチドの一部のみがセンス/アンチセンス分子と結合する（100%未満相補的、例えば、95%、90%、80%、70%でありそれ未満でもよい）。または、このパーセント値以内の又はその値に含まれる任意の数値であってもよい。

10

20

30

40

50

【0097】

本発明の方法、アレイ及びキットでプライマー及びプローブとして使用できるポリヌクレオチドは、典型的には、所定のサンプル（例えば、CD4陽性T細胞又はCD8陽性T細胞を含むサンプル）における、遺伝子（又はその部分/断片）の同定、検出、測定又は分析のためのハイブリダイゼーションプローブ又はプライマーとして用いるのに好適な遺伝子（センス又はアンチセンス）の部分/断片である。典型的には、プライマーは、（例えば、PCRによって）核酸配列にハイブリダイズ又は核酸配列を増幅できるように、逆方向に（すなわち、第1のプライマーは5'の位置で第2のプライマーは3'の位置）配向している。

したがって、他の実施形態において、測定には、プライマー対（逆方向に配向している）のハイブリダイゼーション及び陽性又は陰性予測遺伝子又はこれらの多型（例えば、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）のいずれかに記載された遺伝子）から産出された陽性又は陰性GVHD予測遺伝子のRNA転写産物に由来するcDNAの増幅が含まれる。他の実施形態において、測定には、cDNAを生成して、一又は複数の陽性又は陰性GVHD予測遺伝子（例えば、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）のいずれかに記載された遺伝子）の発現レベルを決定する、RNA転写産物の逆転写（例えば、逆向きに配向したプライマー対を用いる）が含まれる。

[0101] 核酸配列は、can includeヌクレオチド及びヌクレオシド代替物、添加物、欠失、誘導体及び融合配列/キメラ配列（例えば、コード化組み換えポリペプチド）、及びこれらの変異体（例えば、代替物、添加物、挿入及び欠失）。このような変異体の具体例には、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）に記載された任意の配列、又は、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）に記載された任意の配列と相補的な配列の多型及び断片が含まれる。[0102]

「同一性」及びこの文法的変異は、2以上の言及対象が同一であることを意味する。この

ように、2つの配列が同一の場合には、それらは同じアミノ酸配列を有する。「同一のエリア、領域又はドメイン」は、2以上の言及対象の一部が同一であることを意味する。このように、2つの配列が一又は複数の配列領域と同一又は相同である場合には、この2つの配列は、これらの領域において同一性を有する。

【0098】

「同一性」及び「相同性」の程度は、配列における各位置を比較することによって決定される。同一性又は相同性の程度は、当該配列によって共有されている位置における、同一の又は一致する位置（例えば、一致するヌクレオチド又はアミノ酸残基）の数の関数である。「同一性」及び「相同性」の具体例には、include（例えば、1~3、3~5、5~10、10~20、20~30、又はそれ以上の）配列の残基が含まれる。配列は、基準配列、表1（RNA 1538、配列番号：1~1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）に記載された任意の配列の全部又は一部、又は、表1（RNA 1538、配列番号：1~1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）に記載された任意の配列の全部又は一部と相補的な配列に対して、50%、60%、70%、75%、80%、85%、90%、95%、96%、97%、98%、99%、又はそれ以上の同一性又は相同性を有することができる。本明細書での用法のとおり、配列間の同一性又は相同性の所定の割合は、整列した配列において配列の同一性の程度を示すこともできる。2つの配列間の同一性の程度は、コンピュータプログラム及び数学アルゴリズムを用いて説明することができる。配列の同一性（相同性）の割合を計算するアルゴリズムは、概して、比較領域における配列ギャップ及び不一致を考慮する。例えば、BLAST（例えば、BLAST 2.0）検索アルゴリズム（例えば、Altschul et al., J. Mol. Biol. 215:403 (1990), publicly available through the National Center for Biotechnology Information, NCBI参照）は、Mismatch -2; gap open 5; gap extension 2といった例示的な検索パラメータを有する。BLASTアルゴリズムは、データベース配列中の同一長のワードと比べるときに、ポジティブ閾値スコア（positive-valued threshold score）Tと一致するか又はこのTを満足するクエリ配列中のショートワードの長さWを特定することにより、高スコアの配列ペア（HSP）を特定することを含む。Tは、近傍ワードスコア閾値と称される。初期近傍ワードヒットは、より長いHSPを見つけるためのサーチ開始のためのシードとして働く。累積アラインメントスコアが増加し得る限り、ワードヒットを各配列に沿って両方向に伸ばされる。各方向におけるワードヒットの伸長は、累積アラインメントスコアがその最大到達値から量Xだけ低下する場合、1個以上のネガティブスコアリング残基アラインメントの累積のために累積スコアがゼロ以下になる場合、またはいずれかの配列が末端に到達する場合に停止される。BLASTアルゴリズムパラメーターW、TおよびXは、アラインメントの感度および速度を決定する。BLASTプログラムは、デフォルト値としてワード長(W)11、BLOSUM62スコアリングマトリックス(Henikoff and Henikoff, 1992, Proc. Natl. Acad. Sci. USA 89: 10915-10919参照)、アラインメント50、期待値(E)10（又は1又は0.1又は0.01又は0.001又は0.0001）を、M=5、N=-4、および両鎖の比較を用いる。BLASTアルゴリズムを用いた2つの配列間の統計的類似性の尺度の一つは最小合計確率(P(N))であり、これは2つのヌクレオチド配列間または2つのアミノ酸配列間の一致が偶然発生する確率の指標を提供するものである。

【0099】

ポリペプチド配列の比較に関し、BLASTPアルゴリズムは、典型的には、PAM100、PAM 250、BLOSUM 62又はBLOSUM 50等のスコアリングマトリックスと組み合わせて用いられる。FASTA（例えば、FASTA2及びFASTA3）及びSSEARCH 配列比較プログラムは、同一性の程度を定量化するために用いられる（Pearson et al., Proc. Natl. Acad. Sci. USA 85:2444 (1988)、Pearson, Methods Mol Biol. 132:185 (2000)、及びSmith et al., J. Mol. Biol. 147:195 (1981)）。DeLaunay-based位相マッピングを用いたタンパク質構造類似性を定量化するプログラム開発されている（Bostick et al., Biochem Biophys Res Commun. 304:320 (2003)）。

10

20

30

40

50

【0100】

核酸は、様々な標準的クローニング技術及び化学合成技術を用いて生成される。このような技術には、抗体コード化配列にアニーリングすることができるプライマー（例えば縮重プライマー混合物）を用いた、ゲノムDNA又はcDNA標的との核酸増幅、例えば、ポリメラーゼ連鎖反応（PCR）を含むがこれには限られない。核酸は、化学合成（例えば、固相アミド垂りん酸エステル合成）又は遺伝子からの転写によっても生成できる。生成された配列は、その後、in vitroで翻訳され、又は、プラスミドにクローン化され、増殖された後、細胞（例えば、動物又は植物内における真核生物細胞又はほ乳類細胞等の宿主細胞、酵母又はバクテリア）内で発現される。

【0101】

本明細書において説明されているように、遺伝子発現は、発現産物の検出によって測定及び/又は分析できる。本明細書での用法のとおり、「発現産物」という語は、遺伝子でコード化されたアミノ酸配列、タンパク質、ポリペプチド、又はペプチドである。具体的には、発現産物は、例えば、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、又は表18（SG64）の配列に示されている陰性又は陽性GVHD予測遺伝子の全部又は一部によってコード化される。本発明の方法、キット及びアレイは、例えば、in 表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、又は表18（SG64）に記載されている一又は複数の陰性又は陽性GVHD予測遺伝子によってコード化された発現産物の検出、測定又は分析を含む。

【0102】

したがって、検体は、発現産物に結合する、すなわち、bind to encoded by陰性又は陽性GVHD予測遺伝子（例えば、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA143）、表2B（RNA192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）のいずれかに記載されている配列）の全部又は一部によってコード化されたアミノ酸配列、タンパク質、ポリペプチド、又はペプチドに結合する分子をさらに含む。本明細書での用法のとおり、「アミノ酸配列」、「タンパク質」、「ポリペプチド」及び「ペプチド」という用語は、are used interchangeably to refer toアミド結合又はそれと同等のものによって共有結合する2以上のアミノ酸又は「残基」を指し示すために、相互に交換可能に用いられる。このようなアミノ酸配列の例示的な長さは、アミノ酸残基の長さで、約5～10、10～20、20～25、25～50、50～100、100～150、150～200、又は200～300、400～500、500～1000、又はそれ以上である。

【0103】

したがって、本発明にかかる検体は、陽性又は陰性GVHD予測遺伝子によってコード化されるタンパク質又は断片（ペプチド、ポリペプチド等）に結合する抗体及びその部分配列を含む。「抗体」という語は、重鎖可変ドメイン V_H 及び/又は軽鎖可変ドメイン V_L を介して他の分子（抗原）と結合するタンパク質を意味する。「抗体」は、任意のモノクローナル免疫グロブリン分子又はポリクローナル免疫グロブリン分子、例えばIgG、IgA、IgD、IgE、IgM、及びこれらの任意のサブクラス（例えば、IgG₁、IgG₂、IgG₃又はIgG₄）を意味する。抗体には、2つの重鎖配列及び2つの軽鎖配列を含む全長抗体が含まれる。抗体は、カッパ重鎖配列又はラムダ重鎖配列（いずれが天然抗体と同様に全長であってもよい）、これらの混合物（すなわち、カッパ鎖配列とラムダ鎖配列との融合）、及びこれらの部分配列又は断片を有していてもよい。天然抗体分子は、2つのカッパ軽鎖又は2つのラムダ軽鎖を有する。

【0104】

抗体及びその部分配列には、哺乳類抗体、霊長類抗体、ヒト抗体及び完全ヒト抗体及びその部分配列が含まれる。抗体及びその部分配列には、形質転換細胞又はハイブリドーマ、又はB細胞によって又はこれらにおいて生成され又は発現されたもの、又は、合成的に生成され又は他の有機体（植物、昆虫、バクテリア等）によって生成されたものが含まれる

10

20

30

40

50

。

【0105】

抗体には、ポリクローナル抗体及びモノクローナル抗体が含まれる。「モノクローナル」抗体は、真核クローン、原核クローン、又はファージクローンを含む単クローンに基づき、この単クローンから得られ、この単クローンから誘導される抗体を意味する。このように、「モノクローナル」抗体は、構造的に定義され、生成方法によっては定義されない。

【0106】

抗体には、部分配列が含まれる。非限定的で代表的な抗体の部分配列の例には、Fab、Fab'、F(ab')₂、Fv、Fd、単鎖Fv(scFv)、ジスルフィド連結Fvs(sdFv)、V_L、V_H、Camel Ig、V-NAR、VHH、trispesific(Fab₃)、二方特異的抗体(Fab₂)、二重特異性抗体((V_L-V_H)₂又は(V_H-V_L)₂)、三重特異性抗体(trivalent)、四重特異性抗体(tetavalent)、小型抗体((scFv-C_{H3})₂)、二方特異的単鎖Fv(Bis-scFv)、IgGdeltaCH2、scFv-Fc、(scFv)₂-Fc、アフィボディ、アプタマー、アビマーもしくはナノボディ、又はこれら以外の完全型免疫グロブリンの抗原結合の部分配列が含まれるがこれらには限られない。抗体には、1以上のエピトープ(例えば、二重特異性抗体)と結合できる抗体、又は、一又は複数の異なる抗原(例えば、二重又は多重特異性抗体)と結合できる抗体が含まれる。

10

【0107】

抗体及びその部分配列は、生成することができ、商業的に利用可能であり、それ以外の提供元から入手することもできる。例えば、表1(RNA 1538)、表2、表2A、表2B(RNA 192)及び/又は表3内の任意の配列の全部又は一部をコード化して得られる発現産物又は断片に結合する抗体は、当業者に既知の標準的な免疫学的方法を用いて生成できる。

20

【0108】

哺乳類抗体は、トランスジェニック又は非トランスジェニックの哺乳類、又は、哺乳類抗体を生成するように操作された哺乳類以外の生命体、例えば、非哺乳類細胞(バクテリア細胞、酵母細胞、昆虫細胞)、動物又は植物によって生成される抗体である。「ヒト」抗体は、抗体のアミノ酸配列が完全ヒト型、すなわち、ヒト重鎖可変領域及びヒト軽鎖可変領域並びに及びヒト不変領域であることを意味する。このように、アミノ酸の全てがヒト型であるか、又は、ヒト抗体に存在する。「ヒト化」抗体は、抗体のアミノ酸配列が、アクセプターヒト免疫グロブリン分子中で所望の抗原と特異的に結合する一又は複数の相補性決定領域(CDR)の非ヒトアミノ酸残基(例えば、マウス、ラット、ヤギ、ウサギ等)、及びFvフレームワーク領域(FR)中にCDRにフランキングする一又は複数のヒトアミノ酸残基を有することを意味する。

30

【0109】

陰性及び/又は陽性予測遺伝子でコード化された発現産物の量を測定する方法は、当業者に既知である。タンパク質検出、測定及び分析方法の非限定的な例には、ウェスタンブロット、免疫ブロット、酵素結合免疫測定法(ELISA)、放射免疫測定法(RIA)、免疫沈降、表面プラズモン共鳴、化学発光、吸収、発光、蛍光分極、リン光、免疫組織化学的分析、マトリックス支援レーザー脱離イオン化飛行時間(MALDI-TOF)、質量分析、マイクロサイトメトリー、マイクロアレイ、顕微鏡法、蛍光活性化細胞分類(FACS)及びフローサイトメトリーが含まれる。陰性及び陽性予測遺伝子によってコード化された発現産物の量は、タンパク質の機能、例えば、酵素もしくは触媒機能、DNA結合機能、リガンド又はレセプター結合、シグナル伝達等に基づく機能的解析も含む。

40

【0110】

「結合」又は「結合する」という語は、検体に関して用いられる場合には、核酸配列又は遺伝子発現産物(例えば、タンパク質)全部又は一部と分子レベルで相互作用する結合部分を意味する。特異的な結合は、配列又は発現産物に対して選択的である。特異的及び選択的結合は、当業者に既知の解析法(例えば、免疫沈降、ELISA、フローサイトメトリー、免疫組織化学法、ウェスタンブロットティング、核酸ハイブリダイゼーション等)を用い

50

る非特異的結合とは区別され得る。

【0111】

検体は、検出を可能とするために、標識化又はタグ付けされる。検出可能な標識、マーカ及びタグには、遺伝子発現又は発現産物の検出、測定、分析及び/又は定量化に好適な標識が含まれ、また、酵素、生化学、分光、光化学、免疫化学、同位体、電気、光学、化学又はこれら以外の方法によって検出可能な任意の構成物を含む。検出可能な標識は、検体に取り付けられ(例えば、結合又は接合され)、検体を構成する一又は複数の原子内にあるか、当該原子である。検体の構成は、炭素、水素、窒素、酸素、硫黄、リン等の一又は複数のを含むことができるので、炭素、水素、窒素、酸素、硫黄、リン等の任意の放射性同位体は、分離可能に標識化された検体内に含められ得る。

10

【0112】

非限定的で例示的な分離可能な標識には、放射性同位体、金属又は金属酸化物等の放射性物質も含まれる。放射性同位体には、アルファ線、ベータ線又はガンマ線を放出する放射性核種を含む。具体的な実施形態において、放射性同位体は、C、N、O、H、S、Cu、Fe、Ga、Ti、Sr、Y、Tc、In、Pm、Gd、Sm、Ho、Lu、Re、At、Bi又はAcの一又は複数であってもよい。追加的な実施形態において、放射性同位体は、 ^3H 、 ^{11}C 、 ^{14}C 、 ^{13}N 、 ^{18}O 、 ^{15}O 、 ^{32}P 、 ^{33}P 、 ^{35}S 、 ^{125}I 又は ^{131}I の一又は複数であってもよい。

【0113】

さらに非限定的で例示的な分離可能な標識には、造影剤(例えば、ガドリニウム、マンガン、バリウム、硫黄、ヨード剤もしくは非ヨード剤、イオン剤又は非イオン性剤)、磁化剤及び常磁性剤(例えば、酸化鉄キレート)、ナノ粒子、酵素(西洋ワサビペルオキシダーゼ、アルカリホスファターゼ、ベータガラクトシダーゼ、又はアセチルコリンエステラーゼ)、補欠分子族(例えば、ストレプトアビジン/ビオチン及びアビジン/ビオチン)、コロイド金又は着色ガラス又はプラスチック(例えば、ポリスチレン、ポリプロピレン、ラテックス等)等の比色分析標識、ビーズ、蛍光物質又は蛍光色素(例えば、ウンベリフェロン、フルオレセイン、フルオレセインイソチオシアネート、ローダミン、ジクロロトリアジニルアミンフルオレセン、塩化ダンシル、テキサスレッド、ローダミン)、発光物質(例えば、ルミノール)、又は生物発光物質(例えば、緑色蛍光タンパク質、ルシフェラーゼ、ルシフェリン、エクオリン)を含む。標識は、遺伝子発現又は発現産物の検出、測定、分析及び/又は定量化のために用いられる任意の造影剤である(例えば、コンピュータ制御による体軸断層撮影レントゲン写真術(CAT又はCT)、蛍光透視法、単光子放出コンピュータ断層撮影(SPECT)イメージング、光学イメージング、陽電子放出型断層撮影法(PET)、磁気共鳴映像法(MRI)、ガンマ線イメージングのためのもの)。

20

30

【0114】

分離可能な標識は、検体に結合又は接合されてもよい(例えば共有結合)。様々な実施形態において、放射性核種又は金属もしくは金属酸化物等の分離可能な標識は、直接又は間接に検体に結合又は接合される。リンカー又は中間官能基は、検体を分離可能な標識に結合するために用いられ得る。

【0115】

「融合」又は「キメラ」又は「コンジュゲート」という語及びこれらの文法的変異体は、分子に関して用いられる場合、典型的には天然には共には存在しない互いに区別される2つの異なる分子の実体に由来し、このような分子の実体から取得もしくは分離され、又はこれらの分子の実体を基礎もしくは原型とする部分又は部位を含む分子を意味する。すなわち、例えば、検体の融合又はコンジュゲートの部分は、遺伝子産物(陽性又は陰性予測遺伝子によってコード化されたもの)に結合する部分(例えば、抗体)及び第2の部分を含み又はこのような部分及び第2の部分から成る。第2の部分は、分離可能な部分又は物質を含み又はこのような部分又は物質から成る。第1及び第2の部分は構造的に区別される。

40

【0116】

融合、キメラ及びコンジュゲートは、共有結合又は非共有結合により、直接又は間接に結

50

合される。共有結合の非限定的な例は、アミド結合、非天然非アミド化学結合であり、例えば、グルタルアルデヒド、N-ヒドロキシスクシンイミドエステル、二官能性マレイミド、N、N'-ジシクロヘキシルカルボジイミド (DCC) 又はN,N'-ジイソプロピルカルボジイミド (DIC) を含む。アミド結合以外の連結基には、例えば、ケトメチレン (例えば、-C(=O)-CH₂- for -C(=O)-NH-)、アミノメチレン (CH₂-NH₂)、エチレン、オレフィン (CH=CH)、エーテル (CH₂-O)、チオエーテル (CH₂-S)、テトラゾール (CN₄-)、チアゾール、レトロアミド、チオアミド、又はエステル (例えば、Spatola (1983) in Chemistry and Biochemistry of Amino Acids, Peptides and Proteins, Vol. 7, pp 267-357, "Peptide and Backbone Modifications," Marcel Decker, NY参照) が含まれる。本発明の構成物及び方法は、may be contacted又はprovided in vitro、ex vivo又はin vivoで接触される。「接触」及びこれらの文法的変異体は、2以上の実体 (例えば、検体及び核酸又は発現産物) の間での物理的相互作用 (直接又は間接) を許容する状態を意味する。一例において、接触は、検体 (例えば、ポリヌクレオチド、プローブ、プライマー、抗体又は断片等) と生体サンプル、例えば、CD4陽性T細胞、CD8陽性T細胞、又は核酸、タンパク質等の生体サンプルに由来する細胞性の又はそれ以外の物質との相互作用 (例えば、結合) を意味する。

【0117】

本発明における発現の検出、測定又は分析方法に関し、本明細書における接触には、溶液中、固相中、in situ、in vitro、ex vivo、細胞中におけるものを含む。この細胞は、例えば、in vivo CD4陽性細胞又はCD8陽性T細胞、invitro CD4陽性細胞又はCD8陽性T細胞、初代細胞分離株、通過細胞、培養細胞、又はex vivo細胞を含むサンプルなどである。このように、本発明の方法は、陽性及び/又は陰性予測遺伝子発現の量又はレベルを示す他の実体と検体が結合できる条件下での接触を含む。

【0118】

検体 (すなわち、核酸、タンパク質、抗体又はその断片) は、自由状態、溶液中、固相中のいずれにあってもよく、例えば、基板又は支持体 (例えば、固体のもの) に固定されている。基板及び支持体の例には、マルチウェルプレート、ビーズもしくは球体、チューブ又はバイアル、マイクロアレイ又はこれら以外の任意の好適な基板又は支持体を含む。固定は、基板又は支持体と検体とのパッシブ吸着 (非共有結合) もしくは共有結合、又は、検体を試薬に付着させその後試薬を基板又は支持体に取り付ける間接的なもの (例えば、リガンドレセプター機構。分子が検体及び基板又は支持体に固定された対応するレセプターに接合され、ピオチン-ストレプトアビジン機構が例である) であってもよい。

【0119】

「結合 (bind)」又は「結合する (binding)」という語は、分子レベルでの (直接又は間接) 物理的相互作用を意味する。典型的には、結合は、標的に対して特異的又は選択的である。すなわち、アッセイに関するバックグラウンド結合又は対照結合よりも統計的に有意に高いことを意味する。「特異的に結合」という語は、優先的に又は選択的に標的と結合する能力を意味する。例えば、ポリヌクレオチド、プライマー、プローブ、又は抗体等の検体が核酸又は遺伝子発現産物と結合 (又はハイブリダイズ) する。特異的及び選択的な結合は、公知の解析法 (例えば、核酸検出に関しては、ポリメラーゼ連鎖反応、DNA転写、ノーザンブロットティング及びサザンブロットティング等、及び/又はタンパク質検出、免疫沈降、ELISA、フローサイトメトリー、及びウェスタンブロットティング) を用いて、非特異的な結合と区別される。例えば、免疫測定を行うときには、対照には、典型的には抗体又は抗原結合断片単体 (すなわち、タンパク質サンプルが存在しない) を含む反応ウェル又は反応チューブが含まれる。このとき、反応量 (例えば、ウェルとの非特異的結合) は、タンパク質サンプルが存在しない抗体又はその抗原結合断片によって、バックグラウンドと考えられる。

【0120】

本発明は、データベース及び組織的構成をさらに提供する。「データベース」又は「組織的構成」は、情報の収集に関連する。データベース又は組織的構成には、典型的には、一

10

20

30

40

50

又は複数の実際のHCTドナー及び/又はHCTドナー候補の遺伝子発現プロファイル、又は、実際のドナー及び/又はドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こす又は引き起こさないリスク又は可能性を示すスコア又はそれ以外の指標が含まれる。一実施形態において、データベース又は組織的構成には、実際のHCTドナー又はHCTドナー候補の複数の陽性及び/又は陰性予測遺伝子の遺伝子発現プロファイル、又は、実際のドナー又はドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすかどうかのリスク又は可能性を示すスコアが含まれる。他の実施形態において、データベース又は組織的構成には、複数の実際のドナー又はHCTドナー候補の複数の陽性及び/又は陰性予測遺伝子の遺伝子発現プロファイル、又は、複数の実際のドナー又はドナー候補からのHCTがHCTレシピエントにおいてGVHDを引き起こすかどうかのリスク又は可能性を示すスコアが含まれる。

10

【0121】

所定の実際のドナー又はドナー候補のHCTのリスク(GVHDが含まれる)は、GVHDがHCTレシピエントにおいて引き起こされるか否か、及び、それはどの程度か(例えば、重症度)を予測するために用いられる。例えば、所定のHCTレシピエントについて相性のあうHCTドナーが限られている場合には、いくらかのGVHDを引き起こすリスクを有するHCTドナーがHCTレシピエントとして選択される。GVHDはレシピエントへの移植後に起こり得るので、レシピエントは、レシピエントへのHCTの実施前又は後に、効果的な量の拒絶反応抑制剤で治療を受けることができる。GVHDを引き起こすリスクに応じて、レシピエントは、予期されるリスクに基づいて、より積極的に又は消極的に治療を受けることができ、又は、レシピエントに標準的プロトコルに従って治療を施すか否かを決定することができる。このようにして治療を受けたHCTレシピエントは、移植に関連するGVHD等の合併症を減少させ又は防止することができる。したがって、本発明は、HCTレシピエントにおけるGVHDのリスクを予測することができ、レシピエントへのHCTの実施前又は後に、レシピエントが抗GVHD拒絶反応改善治療を受けることができるようになる方法を提供することができる。

20

【0122】

本発明はキットを提供する。このキットには、例えば、好適なパッケージ材にパッケージされた検体、核酸配列、プライマー、プローブ、抗体及びアレイが含まれる。キットの成分は、陽性及び/又は陰性GVHD予測遺伝子(例えば、表1(RNA 1538)、表2、表2A、表2B(RNA 192)及び/又は表3内のもの)、例えば、関連する陽性又は陰性予測遺伝子(例えば、核酸又は発現産物)に特的に結合するプローブ、プライマー対又は抗体(例えば、その発現レベルが、HCTドナーがGVHDを引き起こすリスクと相関のある遺伝子)の発現を検出、測定又は分析のために用いられる。したがって、一実施形態において、キットには、例えば、表1(RNA 1538)、表2、表2A、表2B(RNA 192)及び/又は表3に記載された一又は複数の陽性及び/又は陰性GVHD予測遺伝子又はこれらの任意の配列によってコード化された発現産物の発現の検出、測定又は分析を可能とする検体、核酸配列、プライマー、プローブ、抗体又はアレイが含まれる。

30

【0123】

「パッケージ材」という語は、キットの一又は複数の構成要素を収容する物理的構造を意味する。パッケージ材は、構成要素を殺菌包装することができ、この種の目的に常用される材料(例えば、紙、段ボール素材、ガラス、プラスチック、箔、アンプル、バイアル、チューブ等)から成る。キットは、複数の構成要素、例えば、2以上の検体を、単独で又は組み合わせて含む。

40

【0124】

キットには、構成要素に関する説明(種類、量、用量等)を含むラベル又は添付文書や、固相中、液体中、in vitro、in situ、又はin vivo及びこれ以外の任意の状態での使用説明書が含まれてもよい。ラベル又は添付文書には、本明細書で説明する方法の任意のものを実施する際の説明、例えば、実際のドナー又はドナー候補のHCTがHCTレシピエントにおいてGVHDを引き起こすか否かのリスクを決定又は予測するために遺伝子発現を

50

測定及び／又は分析するための説明が含まれる。この説明は、追加的に、既定の基準発現量（例えば、標準量又は対照量に対するもの）よりも大きい遺伝子発現レベルが、HCTドナーがHCTレシピエントにおいてGVHDを引き起こすリスクが高い又は低いことを示すことを示してもよい。

【0125】

ラベル又は添付文書には、製造者、ロット番号、製造場所及び製造日、使用期限を特定する情報が含まれてもよい。ラベル又は添付文書には、紙又は厚紙等の「印刷物」が含まれてもよい。この印刷物は、キットやパッキング材（例えば、箱）と別であってもよく、キットもしくはパッキング材又はキットの構成要素を含むアンプル、チューブ又はバイアルに添付されていてもよい。ラベル又は添付文書には、追加的に、コンピュータ読み取り可能な媒体、例えば、バーコードが印刷されたラベル、ディスク、CD-ROM、CD-RAM、DVD-ROM、DVD-RAM、DVD等の光学ディスク、MP3、磁気テープ、又はRAM及びROM等の電子記憶媒体、又は磁気光学記憶媒体等のこれらの組み合わせ、FLASH媒体又はメモリカード、が含まれる。

10

【0126】

本発明のキットには、検体（例えば、例えば、表1（RNA 1538）、表2、表2A、表2B（RNA 192）及び／又は表3に記載された陽性又は陰性GVHD予測遺伝子の発現又はこれらの配列によってコード化された発現産物の検出、測定又は分析を可能にする核酸配列、プライマー、プローブ又は抗体）を含む剤形中に、追加的に、緩衝剤、保存剤、安定剤が含まれてもよい。キットの各構成要素は、個別の容器に入れられ、この容器の全てを1つのパッケージに入れるようにしてもよい。

20

【0127】

本発明のキットには、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA 143）、表2B（RNA 192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）に記載された任意の遺伝子配列の全部又は一部と100%同一の又はこれらの100%相補的な核酸（例えば、オリゴヌクレオチド、プライマー、又はプローブ）、及び、表1（RNA 1538）、表2、表2A、表2B（RNA 192）及び／又は表3に記載された遺伝子配列の全部又は一部と100%未満の同一性を有し又は100%未満の相補性を有する核酸（例えば、オリゴヌクレオチド、プライマー、又はプローブ）（例えば、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA 143）、表2B（RNA 192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）に記載された任意の遺伝子配列の全部又は一部と、60%、70%、80%、85%、90%、又は95%の同一性又は相補性）が含まれてもよい。したがって、キットには、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA 143）、表2B（RNA 192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）に記載された陽性又は陰性GVHD予測遺伝子配列の全部又は一部（又はこれらの多型）、又は、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA 143）、表2B（RNA 192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）に記載された遺伝子の全部又は一部配列の相補的な配列、にハイブリダイズするセンス及び／又はアンチセンス核酸配列を含んでもよい。このような核酸は、表1（RNA 1538、配列番号：1～1546）、表2、表2A（RNA 143）、表2B（RNA 192）、表3、表13（SG175）、表15（SG128）、及び表18（SG64）に記載された核酸配列の全部又は一部と同一であってもよく相補的であってもよい。

30

40

【0128】

一実施形態において、キットには、2以上（例えば、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20又はそれ以上、例えば、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、4

50

7、48、49、50又はそれ以上)のプライマー対が含まれる。各プライマー対は、互いに逆方向に配向しており、当該プライマー対は、表1(RNA 1538、配列番号：1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)、又は表18(SG64)に記載された陽性又は陰性GVHD予測遺伝子の1つ、又は、表1(RNA 1538、配列番号：1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)、又は表18(SG64)に記載された陽性又は陰性GVHD予測遺伝子配列(又はこれらの多型)の全部又は一部、又は、表1(RNA 1538、配列番号：1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)、又は表18(SG64)に記載された遺伝子の全部又は一部配列の相補的な配列、から生成されるRNA又はcDNAにハイブリダイズする。

10

【0129】

本発明のキットには、上記以外の検体が含まれる。一実施形態において、キットには、プライマー対の1つによって増幅される核酸配列にハイブリダイズするプローブを含むことができる。このプライマー対は、表1(RNA 1538、配列番号：1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)、又は表18(SG64)に記載された陽性又は陰性GVHD予測遺伝子配列(又はこれらの多型)の全部又は一部、又は、表1(RNA 1538、配列番号：1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)、又は表18(SG64)に記載された遺伝子の全部又は一部配列の相補的な配列、例えば、表1(RNA 1538、配列番号：1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)、又は表18(SG64)に記載された陽性又は陰性GVHD予測遺伝子(又はこれらの多型)の一又は複数のRNA又はcDNAにハイブリダイズする。具体的な態様において、複数のプローブ(例えば、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20又はそれ以上、例えば、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50又はそれ以上)は、表1(RNA 1538、配列番号：1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)、又は表18(SG64)に記載された陽性又は陰性GVHD予測遺伝子配列(又はこれらの多型)の全部又は一部、又は、このような遺伝子配列の全部又は一部の相補的な配列、例えば、表1(RNA 1538、配列番号：1~1546)、表2、表2A(RNA143)、表2B(RNA192)、表3、表13(SG175)、表15(SG128)、又は表18(SG64)に記載された陽性又は陰性予測遺伝子のRNA又はcDNAにハイブリダイズする。

20

30

【0130】

検体を含む本発明のキットにおいては、当該検体の全部又は一部が、支持体や基板に付着又は固定されている必要がない。一実施形態において、キットにはプライマー対又はプローブが含まれ、当該プライマー対及び/又はプローブは支持体又は基板に付着されたり固定されたりしない。

【0131】

本発明のキットには、核酸の発現レベルを評価する際に有用な上記以外の試薬(例えば、PCR反応を行うためやプローブと核酸との結合を検出するための緩衝剤等の試薬)を含んでもよい。例えば、キットには、追加的な有用な物質及び材料、例えば基準物質(例えば、発現の結果が比較される既知量の核酸を含むサンプル)を含んでもよい。キットには、追加的に include a コンピュータ読み取り可能媒体(例えば、データ分析プログラム、基準遺伝子発現プロファイル等を含むもの)、対照サンプル、及びサンプル及び分析を取得及び/又は処理し、取得された遺伝子発現データを分析するための前記以外の試薬を含むことができる。

40

【0132】

本発明はアレイを提供する。このアレイには、例えば、支持体又は基板(例えば、マルチウェルフォーマットもしくはマルチウェルプレート又はディッシュ)に付着され又は固定

50

される一又は複数の検体、核酸配列、ポリヌクレオチド、オリゴヌクレオチド、プライマー、プローブ又は抗体を含むことができる。「アレイ」又は「マイクロアレイ」は、「バイオチップ」とも呼ばれ、ポリヌクレオチド、オリゴヌクレオチド、プライマー、プローブ又は抗体等の（例えばハイブリダイズ可能な）検体を、基板に結合する構成を意味する。このようなアレイは、遺伝子発現レベルにおける変異を定量化するために好適であり、したがって本明細書において説明している方法、例えば、一又は複数の陽性及び/又は陰性予測遺伝子の発現を検出、測定又は分析する方法にとって有用である。

【0133】

典型的には、アレイにおいて、既知の遺伝子（例えば、陽性又は陰性予測遺伝子単鎖、センス又はアンチセンス）の一部であり、又は、遺伝子発現産物（例えば、陽性又は陰性予測遺伝子の遺伝子発現産物）に結合する検体（例えば、核酸配列、オリゴヌクレオチド、プローブ又は抗体）は、基板又は支持体上で所定の又は既知のアドレス又は位置を占める。したがって、核酸配列、ポリヌクレオチド、オリゴヌクレオチド、プライマー、プローブ又は抗体等の検体は、支持体又は基板で、所定の又は既知の場所、位置又はアドレスを有する。

10

【0134】

検体は、典型的には、アレイの2以上の次元に配置される。アレイは、様々な形状をとれる。例えば、アレイは、規則的であってもよく（例えば、均一な行及び列に配置される）又は不規則であってもよい。このように、整然としたアレイでは、サンプルの位置又は場所は、サンプルがアレイに付着されるときに当該サンプルに割り当てられ、キーによって、各々の位置又は場所を適当な標的と関連づけられる。整然としたアレイは、シンメトリックなグリッド状に配置されてもよいが、サンプルはこれ以外のパターン（例えば、半径方向に延びる線状、らせん状、又は整然としたクラスター状）で配置されてもよい。アレイは、通常、アレイの特定のアドレスをその位置におけるサンプルのID（例えば、ハイブリダイゼーション又は例えばシグナル強度を含む結合データ）と関連づけるようにコンピュータがプログラムされ得るという点で、コンピュータ読み取り可能である。コンピュータ読み取り可能なフォーマットの非限定的な例において、アレイにおける個別のサンプルは、規則的に、例えばデカルト格子パターンに配置され、コンピュータによってアドレス情報と関連づけられる。

20

【0135】

アレイ「フォーマット」には、検体が支持体又は基板に付着され又は固定される任意のフォーマット、例えば、マイクロタイタープレートもしくはマルチウェルプレート又はディッシュ、テストチューブ、無機系シート、ディップスティック等が含まれる。特定のフォーマットが重要ということはない。必要なことは、吸着された検体の機能的挙動に影響を与えずに検体が支持体又は基板に付着され又は固定されることだけである。

30

【0136】

支持体又は基板は、ガラス又はプラスチック等の不活性物質であってもよい。このような材料の1つは、ポリプロピレン等の有機高分子である。ポリプロピレンは、化学的に不活性で疎水性であり、様々な有機酸、溶剤、塩、酸化剤、及び鉱酸に対する優れた耐薬品性を有している。追加的な非限定的の例には、ポリエチレン、ポリブチレン、ポリイソブチレン、ポリブタジエン、ポリイソプレン、ポリビニルピロリジン、ポリテトラフルオロエチレン、ポリフッ化ビニリデン、ポリフッ化エチレン・プロピレン、ポリエチレンビニルアルコール、ポリメチルペンテン、ポリクロロトリフルオロエチレン、ポリスルホン、水酸化二軸延伸ポリプロピレン、アミノ化二軸延伸ポリプロピレン、チオール化二軸延伸ポリプロピレン、エチレンアクリル酸、エチレンメタクリル酸、ナイロン類、及びこれらの混合物又は共重合体（例えば、ポリプロピレン、ポリエチレン、ポリブチレン、ポリイソブチレン等の混合物、交互ブロック、又は交互材料）が含まれる。

40

【0137】

一実施形態において、アレイには、2以上のプライマー対が含まれる。各プライマー対は、互いに逆方向に配列している。プライマー対の各々は、表1（RNA 1538、配列番号：1

50

～ 1 5 4 6)、表 2、表 2 A (RNA143)、表 2 B (RNA192)、表 3、表 1 3 (SG175)、表 1 5 (SG128)、又は表 1 8 (SG64)に記載された任意の遺伝子配列(又はこれらの多型)の全部又は一部、又は、表 1 (RNA 1538、配列番号: 1 ~ 1 5 4 6)、表 2、表 2 A (RNA143)、表 2 B (RNA192)、表 3、表 1 3 (SG175)、表 1 5 (SG128)、又は表 1 8 (SG64)に記載された任意の遺伝子配列の全部又は一部の相補的な配列、にハイブリダイズする。また、各プライマー対は、支持体又は基板に付着され又は固定される。具体的な態様において、プライマー対の一又は複数のプライマーは、表 1 (RNA 1538、配列番号: 1 ~ 1 5 4 6)、表 2、表 2 A (RNA143)、表 2 B (RNA192)、表 3、表 1 3 (SG175)、表 1 5 (SG128)、又は表 1 8 (SG64)に記載された任意の遺伝子配列の全部又は一部と 1 0 0 % 同一又は 1 0 0 % 相補的である。または、プライマー対の一又は複数のプライマーは、表 1 (RNA 1538、配列番号: 1 ~ 1 5 4 6)、表 2、表 2 A (RNA143)、表 2 B (RNA192)、表 3、表 1 3 (SG175)、表 1 5 (SG128)、又は表 1 8 (SG64)に記載された任意の遺伝子配列の全部又は一部と 1 0 0 % 未満の同一性又は 1 0 0 % 未満の相補性を有する(例えば、表 1 (RNA 1538、配列番号: 1 ~ 1 5 4 6)、表 2、表 2 A (RNA143)、表 2 B (RNA192)、表 3、表 1 3 (SG175)、表 1 5 (SG128)、又は表 1 8 (SG64)に記載された任意の遺伝子配列の全部又は一部と 60%、70%、80%、85%、90%、又は 95%の同一性又は相補性を有する)。さらに具体的な態様において、アレイには、プライマー対(例えば、表 1 (RNA 1538、配列番号: 1 ~ 1 5 4 6)、表 2、表 2 A (RNA143)、表 2 B (RNA192)、表 3、表 1 3 (SG175)、表 1 5 (SG128)、又は表 1 8 (SG64)に記載された任意の配列の全部又は一部、又は、表 1 (RNA 1538、配列番号: 1 ~ 1 5 4 6)、表 2、表 2 A (RNA143)、表 2 B (RNA192)、表 3、表 1 3 (SG175)、表 1 5 (SG128)、又は表 1 8 (SG64)に記載された任意の配列の全部又は一部と相補的な配列の全部又は一部)の 1 つによって増幅された核酸配列にハイブリダイズするプローブ(又は複数のプローブ)がさらに含まれる。

【 0 1 3 8 】

他の実施形態において、アレイには、2 以上のプローブが含まれる。各プローブは、表 1 (RNA 1538、配列番号: 1 ~ 1 5 4 6)、表 2、表 2 A (RNA143)、表 2 B (RNA192)、表 3、表 1 3 (SG175)、表 1 5 (SG128)、又は表 1 8 (SG64)に記載された遺伝子配列(又はこれらの多型)の全部又は一部、又は、表 1 (RNA 1538、配列番号: 1 ~ 1 5 4 6)、表 2、表 2 A (RNA143)、表 2 B (RNA192)、表 3、表 1 3 (SG175)、表 1 5 (SG128)、又は表 1 8 (SG64)に記載された遺伝子の全部又は一部配列の相補的な配列にハイブリダイズする。各プローブは、支持体又は基板に付着され又は固定される。具体的な態様において、一又は複数のプローブは、表 1 (RNA 1538、配列番号: 1 ~ 1 5 4 6)、表 2、表 2 A (RNA143)、表 2 B (RNA192)、表 3、表 1 3 (SG175)、表 1 5 (SG128)、又は表 1 8 (SG64)に記載された遺伝子配列の全部又は一部と 1 0 0 % の同一性又は 1 0 0 % の相補性を有し、又は、表 1 (RNA 1538、配列番号: 1 ~ 1 5 4 6)、表 2、表 2 A (RNA143)、表 2 B (RNA192)、表 3、表 1 3 (SG175)、表 1 5 (SG128)、又は表 1 8 (SG64)に記載された遺伝子配列の全部又は一部と 1 0 0 % 未満の同一性又は 1 0 0 % 未満の相補性を有する(例えば、表 1 (RNA 1538)、表 2、表 2 A、表 2 B (RNA 192)及び/又は表 3に記載された遺伝子配列の全部又は一部と、60%、70%、80%、85%、90%、又は 95%の同一性又は相補性を有する)。

【 0 1 3 9 】

本明細書において説明されるハイブリダイズプローブ及び/又はプライマー配列及び陽性及び陰性の予測遺伝子の配列は、表 1 (RNA 1538、配列番号: 1 ~ 1 5 4 6)、表 2、表 2 A (RNA143)、表 2 B (RNA192)、表 3、表 1 3 (SG175)、表 1 5 (SG128)、又は表 1 8 (SG64)に記載されている配列内に提供されている。このように、本明細書で説明されている遺伝子の配列及び同一性が分かっているので、核酸及びその他の検体のアレイは、基板上にデノボ合成することにより、又は、基板の特定の位置に核酸配列を配置又は輸送することにより作成できる。例えば、生体物質、例えば C D 4 陽性 T 細胞又は C D 8 陽性 T 細胞を含むサンプルから精製及び/又は分離された核酸は、オリゴヌクレオチド

又はプローブのアレイとハイブリダイズされ、その後、アレイ中で各オリゴヌクレオチド又はプローブにハイブリダイズする標的核酸の量が決定される。It is noted that all of the本明細書において説明される遺伝子の全ては、かかる遺伝子の検出、測定及び分析に好適なオリゴヌクレオチドが生成できるように、少なくとも部分的に、既に配列が決定されている。

【0140】

さらに他の実施形態において、アレイには、表1 (RNA 1538、配列番号：1～1546)、表2、表2A (RNA143)、表2B (RNA192)、表3、表13 (SG175)、表15 (SG128)、又は表18 (SG64)に記載された陽性又は陰性予測遺伝子(又はこれらの多型)又はこれらの遺伝子の全部又は一部の相補的な配列のうちの5個、10個、20個、30個又はそれ以上とハイブリダイズするプライマー及び/又はプローブが含まれる。さらに他の実施形態において、アレイには、表1 (RNA 1538、配列番号：1～1546)、表2、表2A (RNA143)、表2B (RNA192)、表3、表13 (SG175)、表15 (SG128)、又は表18 (SG64)に記載された遺伝子(又はこれらの多型)の全部又は一部、又は、表1 (RNA 1538、配列番号：1～1546)、表2、表2A (RNA143)、表2B (RNA192)、表3、表13 (SG175)、表15 (SG128)、又は表18 (SG64)に記載された遺伝子配列の全部又は一部の相補的な配列、に全てがハイブリダイズするプライマー及び/又はプローブを含むことができる。

10

【0141】

さらに他の実施形態において、アレイには、プライマー対及び/又はプローブの総数が、30000未満、20000未満、15000未満、10000未満、5000未満、2500未満、2000未満、1500未満、1000未満、500未満、400未満、300未満、200未満、100未満、50未満、又は25未満のプライマー対及び/又はプローブが含まれる。

20

【0142】

例示として、核酸、ポリヌクレオチド、オリゴヌクレオチド、プライマー又はプローブのアレイは、マイクロチップ上に固定され、核酸サンプルへのハイブリダイゼーションに好適である。蛍光標識されたcDNAプローブ(例えば、CD4陽性T細胞又はCD8陽性T細胞から抽出されたRNAの逆転写による蛍光ヌクレオチドの導入によって生成される)は、アレイに結合又は付着され、当該アレイ上の核酸の各々のスポットに特異的にハイブリダイズできるようになる。非特異的に結合されたcDNAプローブを洗浄して除去した後、アレイは、検出方法(例えば、共焦点レーザー顕微鏡検査又はCCDカメラ)によってスキャンされる。各アレイ要素のハイブリダイゼーションの定量化は、mRNAの存在量の評価を可能にする。二色蛍光法を用いることにより、2つのRNAのソースから生成され別に標識化されたcDNAプローブは、アレイに対でハイブリダイズされる。このようにして、各々の特定の遺伝子に対応するこの2つのソースからの転写産物の相対的な存在量が決定される。このような方法は、各細胞で僅かに発現する稀な転写産物を検出するために要求される感度を有しており、発現レベルにおける少なくとも約2倍の差を再現可能な方法で検出することが分かっている(Schena et al., Proc. Natl. Acad. Sci. USA 93:106-149 (1996))。

30

【0143】

アレイは、様々な方法で準備される。一例において、オリゴヌクレオチド又はタンパク質配列は、個別に合成され、その後固体の支持体に固定されてもよい(米国特許第6,013,789号参照)。他の例において、配列は、支持体の上に直接合成されて、所望のアレイを提供する(米国特許第5,554,501号参照)。オリゴヌクレオチド及びタンパク質を固体の支持体に共有結合させ及びオリゴヌクレオチド又はタンパク質を当該支持体に直接合成する好適な方法が知られている(好適な方法のサマリーはMatson et al., Anal. Biochem. 217:306-10 (1994)に記載されている)。一例において、オリゴヌクレオチドは、オリゴヌクレオチドを固体の支持体上に準備する従来の化学技術(WO85/01051、WO89/10977、及び米国特許第5,554,501号)を用いて、支持体上に合成される。

40

【0144】

50

本発明は、データベース及び組織的構成を提供する。データベース及び/又は組織的構成は、プロセッサに動作可能に接続される。このようなプロセッサには、データ入力モジュール又はクエリモジュールが含まれる。

【0145】

一実施形態において、データベース又は組織的構成には、複数の実際のHCTドナー又はHCTドナー候補について、表1(RNA 1538、配列番号:1~1546)、表2、表2A(RNA 143)、表2B(RNA 192)、表3、表13(SG175)、表15(SG128)、又は表18(SG64)に記載された2以上の陽性及び/又は陰性予測遺伝子(例えば、CD4陽性T細胞又はCD8陽性T細胞の生体サンプルから得られるもの)又はこれらの多型の遺伝子発現プロファイルが含まれる。また、遺伝子発現プロファイルは、データベース又は組織的構成内の実際のHCTドナー又はHCTドナー候補の各々に関連づけられる。他の実施形態において、データベース又は組織的構成には、実際のHCTドナー又はHCTドナー候補がHCTレシピエントにおいてGVHDを引き起こすか否かの可能性又はリスクに基づいて割り当てられるスコアが含まれる。このスコアの各々は、実際のHCTドナー又はHCTドナー候補の各々について、表1(RNA 1538)、表2、表2A、表2B(RNA 192)及び/又は表3に記載された2以上の陽性及び/又は陰性予測遺伝子又はこれらの多型の遺伝子発現プロファイルに基づくものである。また、各スコアは、データベース又は組織的構成内の実際のHCTドナー又はHCTドナー候補の各々に関連づけられる。具体的な態様において、データベース又は組織的構成には、実際のHCTドナー又はHCTドナー候補の各々について、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20又はそれ以上、例えば、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50又はそれ以上の陽性又は陰性予測遺伝子又はこれらの多型に関する発現情報が含まれる。この陽性又は陰性予測遺伝子は、表1(RNA 1538、配列番号:1~1546)、表2、表2A(RNA 143)、表2B(RNA 192)、表3、表13(SG175)、表15(SG128)、又は表18(SG64)に記載されたものである。さらに具体的な態様においては、実際のドナー又はドナー候補からのHCTについて、HCTレシピエントにおいて移植片対宿主病(GVHD)を引き起こすリスクが高いか低いが特定される。

(表1)

(表2)

(表2B)

(表3)

【0146】

別途定義されない限り、本明細書において用いられる全ての技術用語及び科学用語は、本発明の技術分野における当業者によって一般に理解されている意味と同じ意味で用いられる。本明細書で説明されている方法及び物質と類似の又は同等のものが本発明の実施や試験において用いられるが、本明細書においては、好適な方法及び物質が説明される。

【0147】

全ての出願、公報、論文及び他の文献、本明細書において引用されているジェンバンク(GenBank)の引用及びATCCの引用は、全体として、参照により本明細書に組み込まれる。食い違いがある場合には、定義を含め本明細書が優先する。

【0148】

本明細書に開示されている全ての特徴は、任意の組み合わせ方法で、組み合わせ可能である。本明細書に開示されている特徴の各々は、同一の、同等の、又は類似の目的に寄与する代替的な特徴で置換され得る。このように、別途明示的に説明がなされない限り、開示された特徴(例えば、化合物構造)は、それと同等の又は類似の特徴を持つ属の例である。

10

20

30

40

50

【0149】

本明細書での用法のとおり、単数形の「a」、「and」及び「the」には、文脈上、明確に他の意味に解される場合を除き、複数形のものへの言及が含まれる。このように、例えば、単数の「第1の予測遺伝子、第2の予測遺伝子、第3の予測遺伝子、第4の予測遺伝子、第5の予測遺伝子」等又は単数の「陽性又は陰性予測遺伝子」に言及する場合には、複数の第1の予測遺伝子、第2の予測遺伝子、第3の予測遺伝子、第4の予測遺伝子、第5の予測遺伝子等、又は複数の陽性及び/又は陰性予測遺伝子への言及が含まれる。

【0150】

本明細書において引用される、全ての出願、公報、特許及びこれら以外の文献、ジェンバンクの引用及びATCCの引用は、参照により、全体として本明細書に組み込まれる。食い違いがある場合には、定義を含め本明細書が優先する。

10

【0151】

本明細書での用法のとおり、全ての数値又は数値範囲には、文脈上、明らかに別に解すべき場合を除き、かかる範囲内の整数及び当該数値の端数が含まれる。したがって、例示として、90 - 100%の範囲に言及した場合には、91%、92%、93%、94%、95%、95%、97%、等が含まれ、91.1%、91.2%、91.3%、91.4%、91.5%等、92.1%、92.2%、92.3%、92.4%、92.5%等が含まれる。

【0152】

ある数字より大きい又はある数字より小さいという場合には、その基準となる数字よりも大きい任意の数字、又は、その基準となる数字よりも小さい任意の数字が、それぞれ含まれる。したがって、例えば、30,000より小さいという場合には、29,999、29,998、29,997等の数字1に至るまでの全てが含まれ、及び20,000より小さいという場合には、19,999、19,998、19,997等の数字1に至るまでの全てが含まれる。

20

【0153】

範囲又は一連の範囲に言及する場合には、その範囲、その部分的な範囲、及びその一連の範囲の組み合わせに含まれる整数への言及が含まれる。例えば、5~10の範囲には、1、2、3、4、5、6、7、8、9、10が含まれ、5~10の範囲には、5~7、5~8、6~8、5~9、7~9、5~10の部分的な範囲が含まれる。一連の範囲への言及には、当該葉にの上端及び下端の組み合わせが含まれる。例えば、1~10、10~20、20~30、30~40、40~50、60~70、70~80、80~90、90~100という一連の範囲に言及した場合には、5~20、5~50、5~100、20~50、20~100、30~50、30~100、40~60、40~70、40~80等の範囲への言及が含まれる。

30

【0154】

本発明は、本明細書において、多数の実施形態を説明するために、概して断定的な言葉を用いて開示される。実施形態においては、発明主題、例えば、物質又は材料、方法の工程及び条件、プロトコル、又は手順が全体として又は部分的に除外される場合があるが、本発明は、このような実施形態も含む。このように、本発明は全般に、本発明が何を含まないかという観点では説明されていないものの、本発明から明示的に除外されていない態様は本明細書に開示されているものとする。

40

【0155】

本発明の多数の実施形態が説明されるが、当業者は、本発明の趣旨や範囲から逸脱することなく、本発明を様々な使用方法及び条件に適応させるために、本発明に様々な変更及び改変を行うことができる。したがって、以下の実施例は、本発明の説明のためのものであり、請求項に記載された発明の範囲を限定することを意図するものではない。

【実施例】

【0156】

実施例1。本実施例は、物質及び方法の説明を含む。

50

【 0 1 5 7 】

サンプルソース：全 1 2 2 の移植前、凍結（液体窒素）ドナーPBMC（末梢血単核細胞）サンプル及び対応するレシピエントの G V H D 歴が、NMDP（全米骨髄バンク）の凍結移植ドナー血液保存サンプル及び情報データベースから取得された。調査された122のHCTの全てがHLA10/10適合非血縁ドナー移植に対応するものであり、米国の47の異なる移植センターから得られたものである（表4）。調査されたHCTは、NMDPが選択したALL、AML、CML、又はMDSを有する患者の治療に用いられたものである。これらの122のサンプルが分析され、これらのサンプルから得られた、例示的な陽性及び陰性の G V H D 予測遺伝子が表 1（RNA 1538）に記載されている。

【 0 1 5 8 】

6つの異なる G V H D 転帰グループ（大文字の「G」）が、各センターについて、比較的均一に分散している。これにより、HCTサンプルソース母集団が高い多様性を有することが分かり、移植臨床センターの処理ソースサンプル及び臨床転帰属性の最も発生しやすいバイアスが除去される。表4は臨床センターを示す。

（表4）

【 0 1 5 9 】

患者の G V H D 関連疾患転帰defined groups：各移植の G V H D 転帰は、グループ1からグループ6と命名された、臨床的に関連する6つのグループに分けられる。これらの転帰グループは、急性グレード3又は4の G V H D（最も激しく及び生命を脅かすもの）及び急性グレード1又は2の G V H D（重症度が低く及び軽度と考えられることがあるもの）と広範囲な慢性 G V H D との複数の組み合わせを含む（表5）。表5はグループを示す。

（表5）

【 0 1 6 0 】

ラボ内 selection of blood-derived specific T-lymphocytes for RNA発現analysis：商業的に利用可能な磁気マイクロビーズ技術（Miltenyi Corp.）を用いて、民間の研究所（Southern Research Institute（SRI），Birmingham, AL）との契約に基づいて、CD4陽性T細胞がドナーのPBMC凍結血液サンプルから分離された。他の契約研究所において、商業的に利用可能なRNeasy kit（Qiagen）を用いて、精製されたCD4陽性細胞からRNAが抽出された。

【 0 1 6 1 】

このようにして得られた122のドナーのRNAサンプルの各々について、遺伝子発現（すなわち、細胞内RNA存在量）が、20,000以下の遺伝子（48,803のヒトゲノムプローブによって表わされ、データの平均化の確実性のために、各々を20回以下の独立の技術的測定のために複製した）について、商業的に利用可能なIllumina HT-12 BeadArrays v3.0 microarrays（Illumina Corp.）（Illumina mRNA Expression Analysis, Customer Solutions, Illumina, Inc., San Diego, CA, esp. section Illumina Whole-Genome Gene Expression BeadChips, pp. 5-7）を用いて分析された。このRNA抽出及び定量的な遺伝子発現測定は、イルミナ社のマイクロアレイを用い、契約の元で研究室内の測定値の数値化が行われた（Expression Analysis Inc. (EA), Durham, NC）。

【 0 1 6 2 】

実施例2。この実施例には、数学的安定性及び数値安定性並びにバックグラウンド除去のためのデータ変換についての説明を含む。

【 0 1 6 3 】

一般に、分散安定化（variance stabilization）は、多くの異なる種類の実データに対して、数学的に及び統計学的に適切に一般現象を軽減することを目的とする。測定の分散又は標準偏差は、本質的に測定レベルから独立するというよりも、測定レベルと共に本質的に増加する。こうした“レベル依存（level-dependent）”分散又は標準偏差を、理論上又は経験上調整され一定の許容できる特性を統計学的に有する適切な数学的変換によって、軽減又はさらに取り除くことが望ましい（Durbin, et al., Bioinformatics, 18, Suppl.

10

20

30

40

50

1,S105-S110 (2002))。

【 0 1 6 4 】

最も標準的なパラメータテスト（例えば、T検定）及び一部の非パラメータテスト（例えば、ランク型統計的検定（rank-based statistical tests））は、理論上及び実用上、1セットの分散又は標準偏差がこのような測定の平均値に依存しないことを前提として、分散安定化は第一に必要とされる。同様に、これは、分散、標準偏差、又は単一の測定の測定エラー及び単一の測定のレベルに対しても適用される。測定の分散又は標準偏差が測定のレベルに依存する限りにおいて、実際問題として、レベル非依存の分散又は標準偏差に基づく統計的検定は、信頼性に劣る。このように、本質的な（一定の適切なデータのサブセットの平均値間の差に関係する）分散又は標準偏差のレベル依存性が存在する場合には、統計的検定を行う前に、分散安定化数学的 / 数字上の変換をデータに適用することが推奨される（Sheskin, David J. Handbook of Parametric and Nonparametric Statistical Procedures, 3rd Edition, Chapman & Hall / CRC Press, Boca Raton, Florida, 2004, esp. pp.404-409.）

10

【 0 1 6 5 】

特有のVST（分散安定化変換）を用いた定量化した測定値の数学的又は数値的な安定化HCTがGVHDを引き起こすか否かを予測する陽性及び陰性の遺伝子を同定する定量化された遺伝子発現データについてカスタマイズされた前処理又は後処理を行う前に、製造者から提供された公知のIllumina Bead Studioソフトウェアを用いて、イルミナ社のマイクロアレイ遺伝子発現データにバックグラウンド除去を行った。この処理は、Expression Analysis Inc.によって実施された。全てのマイクロアレイ由来の遺伝子発現データは、サンプルごとに、Illumina社の測定値に対してカスタマイズされたコンピュータ処理分散Stabilizing Transformation（VST）を実行し、線形に最大で4.5倍にリスケールした後、分位正規化した。（Lin, et al., Nucleic Acids Research 36(2):e11, 1-9 (2008)) for the バックグラウンド mathematical statistics of VST as applied to Illumina Bead Array microarray data）。

20

【 0 1 6 6 】

Lin, et al., Nucleic Acids Research 36(2):e11, 1-9 (2008))における一部の数值的な細部は以下に説明するとおり更新され、その後技術的有効性が確認された。カスタマイズされたMatlabプログラム言語（The MathWorks, Inc., 3 Apple Hill Dr., Natick, MA 01760）において、VST改善計算及び実装が開発され実装された。

30

【 0 1 6 7 】

具体的には、Illumina Bead Array特有の数学的、統計的、数値的VST（Lin, et al., Nucleic Acids Research 36(2):e11, 1-9 (2008) with refinements）がサンプルごと（すなわち、各HCTドナーについて個別に）に、Illumina社のプラットフォーム由来の遺伝子発現の測定値に適用された。この遺伝子発現測定値に対して行われる「前処理的な」数学的、統計的、数値的データ処理工程は、転帰予測データ分析の前に行われた。

【 0 1 6 8 】

イルミナ社プラットフォーム由来の遺伝子発現データのために設計されたVSTは、一般的ではなく、広くしようされているものではないが、科学文献（Durbin, et al., Bioinformatics, 18, Suppl. 1, S105-S110 (2002); Dunning, et al., BMC (Biomed Central) Bioinformatics 9, #85, doi:10.1186/1471-2105-9-85, 1-15 (2008)）において確立されたものである。イルミナ社の影響がある公開文献（Illumina mRNA Expression Analysis, Customer Solutions, Illumina, Inc., San Diego, CA, esp. section Illumina Whole-Genome Gene Expression BeadChips, pp. 5-7; Dunning, et al., BMC (Biomed Central) Bioinformatics 9, #85, doi:10.1186/1471-2105-9-85, 1-15 (2008)）及びこの研究から得られたイルミナ社のデータの厳密な調査からは、遺伝子発現測定に内在する技術的な標準偏差又は技術的な「誤差」、すなわち、イルミナ社のプラットフォームにおけるいわゆるビーズ標準誤差（ビーズ標準誤差）（Illumina mRNA Expression Analysis, Customer S

40

50

olutions, Illumina, Inc., San Diego, CA, esp. section Illumina Whole-Genome Gene Expression BeadChips, pp.5-7, Dunning, et al., BMC (Biomed Central) Bioinformatics 9, #85, doi:10.1186/1471-2105-9-85, 1-15 (2008)) が発現測定の大きさに応じて大幅に大きくなるという深刻な傾向がある。したがって、本研究でLDA (線形判別分析) 及びT検定を使用するためもあって、本研究で得られたイルミナ社のデータにVSTを適用した。

【0169】

Lin, et al. (Lin, et al., Nucleic Acids Research 36(2):e11, 1-9 (2008)) によって公表されているように、イルミナ社に関し、VSTの最新のバージョンの改良が開発され、遺伝子発現データに適用された。

10

【0170】

数値データのテーブル配置

以下の説明においては、ラボ内イルミナ社プラットフォーム生成データは2次元テーブル又は行列に配置されている。サンプル、すなわち、ドナーHCT又は対照サンプルは列を参照し、特に分子プローブが定義されるオン-プラットフォーム (on-platform) (すなわち、イルミナ社のビーズアレイ) は行を参照する。所定の行は、それゆえ、同一のプローブに関係するデータ、及び全ての列に渡るデータ、すなわち、全てのHCTサンプル又は対照サンプルに渡るデータを表す。所定の列は、所定のサンプル、即ち、所定のドナーHCT又は対照サンプルを表し、各行は、一行ずつそれぞれ、全てのサンプルに対し同じ行方向順序で (row-wise ordering) 、特定のプローブに関係する測定値を表す。例えば、

20

【0171】

イルミナ社のアレイは、遺伝子発現測定シグナルを用いて3つの数値を提供する。イルミナ社のプラットフォームによって4つの別個の (しかし相関のある) 種類の測定値、すなわち遺伝子発現シグナル、ビーズ標準誤差、シグナル測定に関連するビーズの平均数、及びイルミナ社が定義及び計算するシグナル「検出p値」が提供される。このビーズ標準誤差、ビーズの平均数、及びシグナル検出p値は、遺伝子発現シグナル測定値の各々と関連付けられる (BeadStudio Gene Expression Module v3.2 User Guide, Part # 11279596 Rev. A, Illumina, Inc., San Diego, CA, esp., Detection P-value section, Normalization and Differential Analysis, Ch. 4.)。以下の説明において、遺伝子発現測定値はシグナルと称され、ビーズ標準誤差はbead_stderr (又はこれと同様の名称、小文字又は大文字のいずれか) と省略され、ビーズの平均数はavg_nbeads (又はこれと同様の名称、小文字又は大文字のいずれか) と省略され、検出p値は、検出p値 (又はこれと同様の名称、小文字又は大文字のいずれか) と省略される。

30

【0172】

バックグラウンド除去 - 機器及び後続の文脈 : 発現の研究において、シグナルとは、イルミナ社のプラットフォームによって提供される未加工のシグナル数値から、イルミナ社のプラットフォームによって所与のサンプルのプローブの各々について評価および算出された、いわゆる数値的に推定された機器の「局所バックグラウンド蛍光」を差し引いたもの (すなわち、当研究において、シグナルは未加工のシグナルからイルミナ社のプラットフォームにより提供される「バックグラウンド」除去を差し引いたもの) として開始する。また当研究において、イルミナ社のプラットフォームによって提供されるシグナルは、イルミナ社のプラットフォームによって「正規化」されていない。

40

【0173】

VSTのための数学的、統計的、および数値的なコンピュータによる処理の過程において、マイナスまたはゼロ付近のシグナル値は、それら自体がVST手続きの中で考慮され調整される遺伝子発現測定「バックグラウンド」と考えられている。このことは後の説明の中

50

で明示的かつ明確になる。

【0174】

イルミナ社によって提供される数量の表配置：上記の通り、シグナルについての数値的データ行列（行にプローブ、列にサンプル）が存在する。また、3つの非シグナルのイルミナ社によって提供される測定値、すなわち、シグナル測定のプロブ標準誤差、シグナル測定に関わるプロブの平均個数、及びイルミナ社によって定義及び提供される「検出p値」のそれぞれについて、そのような類似かつレジスタにおいて行方向および列方向の数値データ行列が存在する。このように根本的なレベルにおいて、所与のサンプルの所与のプロブにおける所与のシグナル値に対応するこれら3種の研究所機器レベルの数量が存在する。この3つ組の情報はイルミナ社のデータのためのVST (Lin, et al., *Nucleic Acids Research* 36(2):e11, 1-9 (2008)) 及びイルミナ社のデータのためのVSTのカスタマイズされた変更及び実行の中で利用されている。

10

【0175】

データは転帰予測データ分析が実行される前に変換される。VSTは各サンプルに対して個別に、すべてのプローブにわたって適用される。48個のサンプルのそれぞれに対する48803箇所のプローブにわたるサンプル単位の計算VSTパラメータc1、c2、及びc3の実際の結果の例が下記表6に示されている。それでも、以下の記述において、数学的及び統計的な方法及びコンピュータによる手続きの説明は、個別のサンプルの一つ一つを例として限定的に着目していない。このように、下記におけるデータ及び図表は、多くの場合すべてのプローブ（48803箇所のプローブ）及びすべてのHCTサンプル（例えば、当研究のある特定の段階における48個のサンプル）の観点で参照される。それは例えば、対応するbead_stderr、avg_nbeads、及び検出p値の $48803 \times 48 = 2,342,544$ 個のシグナルの数値を含む。

20

【0176】

シグナルヒストグラム：48803 プローブ \times 48サンプルのイルミナ社のシグナル値の典型的未平滑化ヒストグラムが図1に示されている (Plot 5,1)。イルミナ社の「バックグラウンド」が除去された可能な未加工の値の幅が約マイナス数百から約プラス40,000であるため、シグナル値のlog10をヒストグラム化すると経験分布はより明確に視覚化される。当然、対数であるため、プラスの未加工シグナル値のみがそのようなヒストグラムに表すことができる。図1 (Plot 5,1) は、すべてのプラスのシグナル値の基礎的な経験分布図である。なお、このlog10 (プラスシグナル) のヒストグラムは厳密には単一ガウスのではない (すなわち、メインピークの右側に認識可能な広い肩を有している)。

30

【0177】

プロブ標準誤差、プロブの標準個数、シグナル、及びヒストグラム：サンプルごとの何千もの遺伝子発現を計測するイルミナ社のBeadArray技術の利点は、プロブ標準誤差の提供である (Dunning, et al., *BMC (Biomed Central) Bioinformatics* 9, #85, doi:10.1186/1471-2105-9-85, 1-15 (2008); *BeadStudio Gene Expression Module v3.2 User Guide, Part # 11279596 Rev. A, Illumina, Inc., San Diego, CA, esp., Detection P-value section, Normalization and Differential Analysis, Ch. 4.*)。イルミナ社のプラットフォームの物理的性質のため、プロブ標準誤差は物理的に及び統計的に所与のシグナル測定値の (機器レベルの) 標準誤差と考えることができる。すなわち、所与のサンプルの所与のプローブのプロブ標準誤差は、対応するシグナル値付近の標準誤差の半分の幅を有する従来のエラーバーと考えることができる。(すなわち、測定値の標準偏差を測定反復回数の平方根で割った値であり、イルミナ社のプラットフォームはbead_stderr = 測定値標準偏差を測定に関わったavg_nbeadsの平方根で割った値である) (Durbin, et al., *Bioinformatics*, 18, Suppl. 1, S105-S110 (2002); Dunning, et al., *BMC (Biomed Central) Bioinformatics* 9, #85, doi:10.1186/1471-2105-9-85, 1-15 (2008))。イルミナ社のプラットフォームは、イルミナ社の「プロブ」個別の測定を別々に行い、最終的には報告されたシグナル値及びプロブ標準誤差を、報告されたシグナル及びプロブ標準誤差を得る

40

50

ことに関わったビーズの平均個数とともに構成する (BeadStudio Gene ExpressionModule v3.2 User Guide, Part # 11279596 Rev. A, Illumina, Inc., San Diego, CA, esp., Detection P-value section, Normalization and Differential Analysis, Ch. 4.; Lin, et al., Nucleic Acids Research 36 (2):e11, 1-9 (2008))。典型的にはこれらの研究において見られるとおり、平均して約20個の別個のビーズが所与の遺伝子プローブのシグナル測定に関わっている。これらの原則のため、所与のシグナルは実際に個別のビーズ幅平均シグナルである。よって、「シグナル」とは実際は本来的に「平均シグナル」である。以下の説明において、シグナル及び平均シグナルの用語は、同一の数量について交換が可能である。つまり、シグナル、すなわち、所与のサンプルの所与のプローブについてイルミナ社のプラットフォームにより報告された平均シグナルである。より具体的には、イルミナ社により提供されたシグナル(すなわち、「平均シグナル」)測定値の標準偏差は、測定に関わったビーズの平均個数の平均個数の平方根(すなわち、 $\sqrt{\text{avg_nbeads}}$ × 対応する報告されたビーズ標準誤差) (Durbin, et al., Bioinformatics, 18, Suppl. 1, S105-S110 (2002); Lin, et al., Nucleic Acids Research 36 (2):e11, 1-9 (2008))。かすかにガウスのな分布ではあるが、ビーズ標準誤差は顕著に大きく増加し、シグナルの増加を伴う。これが、振る舞いのよい、かつ実行可能な分散安定化変換がイルミナ社のデータに適用される根本的な理由である。この現象は、後続のGVHD転帰予測分析及び発見の前にVSTをデータに適用することを必要とし、それにより、ビーズ標準誤差又は標準偏差の顕著なレベル依存、またはシグナルレベルの分散を軽減する。

10

20

30

40

50

【0178】

上記の発見を支持するものとして、ビーズ標準誤差の(及び本質的に分散の)観察されたシグナルレベル依存性が、図2 (Plot 6,5)に明確にかつ劇的に示されている。図2は $\log_{10}(\text{bead_stderr})$ 対 $\log_{10}(\text{プラスシグナル})$ の散布図である。明確に述べれば、 bead_stderr は、そのままであっても、あるいは図2におけるように \log_{10} 単位であっても、定数対シグナルではない。VSTに関わるイルミナ社の検出p値について、および報告されたイルミナ社の検出p値がなぜ高度に限定された方法で採用されたかについての見解を述べる。VSTの実行には「近似シグナル検出高品質」閾値が採用されるが、それはイルミナ社のプラットフォーム高品質シグナル検出(後述)に基づくデータからVSTパラメータ自体を計算するのに用いられるシグナルを要求することによってなされる。これは公開されたVSTのバージョンに加えられた追加工程であり (Lin, et al., Nucleic Acids Research 36 (2):e11, 1-9 (2008))、この工程により、データ由来の計算されたVSTパラメータが、機器のバックグラウンドノイズ近傍ではない技術的に信頼できるシグナル測定値に基づいていることが保証される。ただし、これに引き続いてデータ由来のパラメトリックな、いわゆるVSTはすべてのデータに適用される。すなわち、データ依存の計算されたVSTパラメータ c_1 、 c_2 、及び c_3 自体(下記参照)を、技術的品質が劣る(イルミナ社のプラットフォームにより提供される検出p値に技術的に基づく)シグナルがVSTパラメータ計算処理自体から除外されているシグナル値の組に、非常に慎重に基づかせることにより、計算されたVSTの信頼性は原則として保証される。報告された検出p値を生成する際にイルミナ社のプラットフォームによって採用されるアルゴリズムは複雑である (Illumina mRNA Expression Analysis, Customer Solutions, Illumina, Inc., San Diego, CA, esp. section Illumina Whole-Genome Gene Expression BeadChips, pp. 5-7, especially BeadStudio Gene ExpressionModule v3.2 User Guide, Part # 11279596 Rev. A, Illumina, Inc., San Diego, CA, esp., Detection P-value section, Normalization and Differential Analysis, Ch. 4.参照)。実際には、検出p値の経験的な分散特性は、実際のイルミナ社のデータの集合から計算することができ、特定の目的において、特に、ある計算処理から技術的に信頼性の低いシグナルを選択し除外する際に、報告された検出p値の使用に関する実際の判断の指針となる。

【0179】

よって、VSTパラメータ自体の計算において採用されるデータの、0.5未満の検出p値限界を設定する際に、それらの計算はデータのうちの最も信頼性の高い60%に技術的に基づ

いている。それはすなわち、シグナルデータのうちの技術的に最も信頼性の高い過半部分であり、このことは、そのようなデータに関わる可能性のある実験的又は生物学的解釈とは無関係である。従って、計算によって導かれたデータ依存のVSTパラメータの値自体は、技術的に非常に信頼性の高いデータから導かれたという点で、信頼できるものである。約0.5より大きい検出p値は圧倒的にゼロ付近のシグナル、及び特にマイナスのシグナル、すなわちレベルが低レベルバックグラウンドノイズと対等なシグナルに対応している。(注：生物学的には、そのような遺伝子は合理的な機器検出限界にあるか、またはそれよりも低いかのいずれかに解釈できる。よって、遺伝子発現において事実上信頼性が「ない」。機器検出限界付近で計測された(よってその数値は小さいが「ノイズ」とは異なる定量的数値ほどには信頼性がない)遺伝子発現シグナルは、生物学的に解釈すれば、遺伝子発現においてまさに信頼性が「ない」。これは作為的な結果ではなく、物理的な現実の事実である。つまり、機器検出限界付近で計測された数量は、信頼性のある数値ではなく、物理的、化学的、または生物学的な解釈において、信頼性の点でまだ「存在していない」。) 10

【0180】

従って、報告されたイルミナ社の検出p値は、データ依存のVSTパラメータ自体を計算するための非常に慎重な方法においてのみ採用される。その方法とはすなわち、そのような計算を、シグナル値における機器のバックグラウンド「ノイズ」付近である可能性が低い、データの過半部分にのみ偏向させることである。

【0181】

特にVSTを計算するために必要なイルミナ社の測定データ： 20

【0182】

バックグラウンドシグナルについての考慮事項：Lin, et al., Nucleic Acids Research 36(2):e11, 1-9 (2008)によって公開されたVSTは、研究所の定量機器のための分析化学による、広く受け入れられかつ長期にわたって存在するエラーモデルまたは「ノイズ」モデルに基づいている(Durbin, et al., Bioinformatics, 18, Suppl. 1, S105-S110 (2002))。それは、機器によって計測されたシグナルは通常、いわゆる加法的な及び乗法的なノイズの組み合わせによって、非常に小さい、又は時に大きい程度のエラーが避けられないからである。 30

【0183】

したがって、信頼性をもって「バックグラウンドノイズ」が優位にあるとみなされる低レベルシグナルは「バックグラウンドシグナル」と呼ばれ、報告された検出p値が約0.5より大きいイルミナ社の報告されたシグナルとして、実際的にも運用上も定義される。よって、イルミナ社の報告された検出p値が0.5より大きいシグナルは、経験的に「バックグラウンドシグナル」と定義される。

【0184】

特にVSTのための「バックグラウンドシグナル」分散の考慮事項：上記されているように、非常に信頼性の高いシグナル標準偏差は、一般に $\text{stddev} = \sqrt{\text{avg_nbeads}} * \text{bead_stderr}$ である。よって、非常に信頼性の高いシグナル分散は、一般に $\text{stddev}^2 = \text{avg_nbeads} * \text{bead_stderr}^2$ である。「バックグラウンドシグナル」分散はこのように stddev の2乗として計算され、「バックグラウンドシグナル」 stddev は avg_nbeads 及び bead_stderr から計算される。 40

【0185】

VSTパラメータc3は「バックグラウンドシグナル」分散から計算される：Lin, et al., Nucleic Acids Research 36(2):e11, 1-9 (2008)の定義によって、データ由来のVSTパラメータc3は、「バックグラウンドシグナル」分散の算術平均として計算される。すなわち、検出p値が0.50より大きいシグナルである。

【0186】

「バックグラウンドシグナル」分散の観察される分布は、より大きい分散に向かって右 50

方向に歪んでいるため、平均値は「バックグラウンドシグナル」分散の中心傾向を過大評価したものとなる。すなわち、平均値は中央値よりも約10%大きい。実際的な観点では、「バックグラウンドシグナル」分散をわずかに過大評価する方が、それを過小評価するよりも、安全でかつ慎重である。このように、VSTパラメータc3について「バックグラウンドシグナル」分散の数値評価において、中央値よりも平均値を採用する方が実際にはよい。c3の計算は次のように行われる。

【0187】

公開されたVSTの手続き (Lin, et al., Nucleic Acids Research 36 (2):e11, 1-9 (2008)) に対するもう一つの変更をアルゴリズムによって行うには、c1及びc2の計算から、シグナルのうち最大のもの2%を除外することになる。実際には、観察されたシグナル値の98パーセントは未加工シグナル値が約1500未満であり、例えば、98パーセントシグナルは、48803個の遺伝子×48個のサンプル集合について、約1200である。これはすなわち、イルミナ社のプラットフォームからのシグナルのわずか約2%が約1500より大きい観察値を有することになる。図3 (Plot 5,1) は重要である：(1) 図3はイルミナ社の未加工のシグナルデータについて、数万のレベルにおいて多くのシグナルが発生しているものの、大部分は約1500未満のレベルにおいて発生することを経験的に示している。(2) 大部分のシグナル、すなわち、98%のシグナル(なぜならば、最大のもの2%はVSTのc1及びc2パラメータの算出から除外されているため)については、標準偏差又は分散のシグナルレベルに対する明確かつ顕著な依存が依然として存在する。そして(3) 各サンプルについて別個にVSTデータ依存のパラメータc1及びc2を計算する際に採用されるのは、まさに図3 (Plot 5,1) に表されているようなデータである。

【0188】

VSTデータ依存のパラメータc1及びc2の計算：計算の対象となるシグナルデータのうち最小の98%に対する制限を導入することを除けば、手続きはLin, et al., Nucleic Acids Research 36 (2):e11, 1-9 (2008) に従う。VSTを計算する基礎となる所与のデータの組について、c1及びc2の計算から除外されるのはシグナルのうち最大のもの2%である。実行される手続きはその後以下の通りである。上記のようにc3が「バックグラウンドシグナル」平均分散として計算された後、データの(x, y)ペアを考慮する。ここで、yはsqrt(シグナル分散-c3)であり、シグナル分散はc3を超過し、また、xは対応するシグナル値である。すなわち、yがプラスであり、その分散がc3を超過し、またxが98パーセントのシグナル値未満である(x, y)のみを考慮する。その後、(x, y)データ点の組に対する形式 $y=c1*x + c2$ の線形近似を計算する。すなわち、c1は線形近似の傾きと定義され、c2は線形近似のy切片と定義される。よって、データ依存のVSTパラメータc1及びc2の数値は、(x, y)データ点からそのように得られる。図3 (Plot 5,1) における計算された近似線においては、slope c1 は約0.2であり、y-軸切片 c2は約9.0である。

【0189】

公式としての分散安定化変換(VST)等式：VSTは、イルミナ社のプラットフォームによって提供されるシグナルデータに対して作用する、特定の3パラメータ非線形関数fである。関数fはいかなるシグナル値sも(ただしlogシグナル値は除く)新しいシグナル値tに変換する。すなわち、実際および実行においてVSTは単純に $t = f(s; c1, c2, c3)$ である。ここで、c1、c2、及びc3は、VSTが適用されているイルミナ社のデータの所与の組から導かれた数値定数である。この解説および図表は、実データの48803個の遺伝子×48個のサンプル集合の例を採用する。しかしながら、実際においては、VSTは各サンプルに個別に適用され、1回に1サンプルずつ、48803遺伝子プローブのすべてにわたって適用される。実際においては、48個のサンプルは48回の個別のVSTの適用を必要とする。各サンプルごとの適用はそれぞれ、3つの必要なデータ依存定数c1、c2、及びc3の計算を必要とする(表6参照、例えば、48個のサンプル及び48803個のプローブ全体にわたるサンプル計算のそれぞれについての、実際のc1、c2、及びc3の数値結果)。上記のデータ行列については、VSTは各列に個別に適用される。それが行われるのは、定数c1、c2、及びc3のサンプル個別の値を導出し、その後所与の列、すなわち所与のサンプルのすべてのシグナル値sを、V

ST関数 $t=f(s; c_1, c_2, c_3)$ を当該所与のデータ列に適用することによって変換する際である。

【0190】

未加工の遺伝子発現シグナル s を対応する分散安定化され変換されたシグナル $t = f(s; c_1, c_2, c_3)$ に変換するための基礎的なVST式は、パラメータ c_1 、 c_2 、及び c_3 のデータ由来の数値に関わっている (Lin, et al., *Nucleic Acids Research* 36 (2):e11, 1-9 (2008))。上記VST式は以下の通り表される。

【数1】

$$t = f(s; c_1, c_2, c_3) = \frac{1}{c_1} \operatorname{arcsinh} \left(\frac{c_1 s + c_2}{\sqrt{c_3}} \right) \text{ and where}$$

10

ここで、出力 t はいわゆる分散安定化され変換された (すなわち、「VST」された) シグナルであり、

入力については、

$\operatorname{arcsinh}(z)$ は実数の (複素数ではない)、マイナス、ゼロ、又はプラスの引数 z の、従来の数学的逆双曲線正弦関数であり、

s は未加工のプラットフォームにより決定されるシグナル (log変換されておらず、プラットフォームによって決定される機器バックグラウンドが除去されている、あるいはされていない) であり、マイナス、ゼロ、プラスのいずれでもよい。

20

また、 c_1 、 c_2 、 c_3 は数学的に明確に定義されているが、サンプル又はサンプル集合 S のVSTの数値定数の全体にわたってデータから導出されており、すべてプラスである。(注：数学的には c_3 はゼロでもよいが、その場合、引数 z の代数形式が変化する。しかしながら、実際には c_3 は常にプラスである。なぜなら、物理的に機器から、シグナルの分散は、 bead_stderr からのものであれ、他の何らかの物理的分散の定量化からのものであれ、決してゼロにはならないからである。)

【0191】

所与のサンプル又はサンプル集合 S の所与の遺伝子プローブについて： c_3 はシグナル s のデータ由来の分散の算術平均であり、シグナル s は「バックグラウンドシグナル s 」とみなされ、所与のサンプル又はサンプル集合のすべての遺伝子プローブのほとんど又はすべてにわたって計算される。 c_1 と c_2 はそれぞれ、所与の (すなわち、上記のデータ行列において、ある特定の列に沿う、または多数の列の集合に沿う) サンプルの遺伝子プローブのほとんど又はすべてにわたる (x, y) データ y 対 x の線形近似の傾きと y 切片である。また、 x はプラスのシグナル s であり、 s ($\text{var}(s)$ と表される) の分散は所与のサンプル又はサンプル集合についてすでに計算された c_3 を超過する。また、次の式が成り立つ。

30

【数2】

$$\sqrt{\text{var}(s) - c_3}$$

40

(注：よって y は常にプラスである)

また、 $\text{var}(s)$ はシグナル s の bead_stderr 由来の分散、または他の物理的に導出された分散である。典型的には、イルミナ社のプラットフォームにおける、 $\text{var}(s)$ で表されるシグナルのビーズレベルの技術的分散は bead_stderr 由来であり、すなわち、 $\text{var}(s) = \text{stddev}^2(s) = (\text{sqrt}(\text{avg_nbeads})) (\text{bead_stderr})^2$ (Durbin, et al., *Bioinformatics*, 18, Suppl. 1, S105-S110 (2002); *BeadStudio Gene Expression Module*

50

v3.2 User Guide, Part # 11279596 Rev. A, Illumina, Inc., San Diego, CA, esp., Detection P-value section, Normalization and Differential Analysis, Ch. 4.) であり、当然、 $(\text{avg_nbeads}) (\text{bead_stderr})^2$ とも等しい。2つの式のうち先に記述された式は数値的に $\text{var}(s)$ を計算するために実行することを意図している。というのも、先に記述された式は、後に記述された式よりも数値的精度に関して多少振る舞いがよいからである。しかしながら、いずれの式も実行は可能である。

$\text{Arcsinh}(z) = \ln(2z) + (1/(2*2*z^2)) - 1*3/(2*4*4*z^4) + 1*3*5/(2*4*6*6*z^6) - \dots$ 、ここで $|z| > 1$ であり、 $\ln(z)$ は自然対数の従来の記法、すなわち、基数 e による対数である。

【 0 1 9 2 】

遺伝子発現数値分析に対する真の利点及び利益は、概して、遺伝子発現シグナルデータの数値変換の適用が必要な場合に、逆正弦 (arcsinh) 関数を採用して対数関数と置換することである。また、 $\text{arcsinh}(z)$ は $\log(2z)$ とは1パーセントの4/10未満の差があり、大きさが5を超える引数 z について、増加とともにその差は小さくなる。対数に対する逆正弦 (arcsinh) の有利な数値的特性は数学的及び固有的なものであり、逆正弦 (arcsinh) の、例えばVSTの分散安定化特性自体によるものではない。その数学的利点が得られるのは、VSTが遺伝子発現シグナルデータに適用され、技術的シグナル測定分散の安定化から生じる後続の統計分析のために主に採用されるときである。

【 0 1 9 3 】

基礎的なVST式 $t=f(s; c_1, c_2, c_3)$ の基礎の数学的純化: Lin, et al., Nucleic Acids Research 36(2):e11, 1-9 (2008)において、受け入れられた加法的及び乗算的な機器エラーモデルに適用される、正当に一般化された分散安定化積分の左辺は、不定の積分変換(すなわち、論文において積分の特定の下界については明示も暗示もされていない)として正式に記述されている。正式にそのように記述されていることは、完全に正しいわけではない。むしろ、まったく同じ被積分関数が等式6において、定積分内で、すなわち、現れる積分法の明示的な下界を有する積分内で、現われるべきである。数学的な、かつコンピュータプログラム化されたVSTの実行法は、適切かつ正確な一定の積分バージョンを採用する。

【 0 1 9 4 】

遺伝子発現シグナルデータにおけるVSTのシグナル分布効果の包括的視覚化: 図4 (Plot 7,4) は48803 x 48個のサンプル集合が「VST」された後の、すべてのシグナル値のヒストグラムを示す。変換されたデータのヒストグラムに含まれる二次ガウスの分布を示唆する認識可能な隆起は存在しない(図1 (Plot5,1)の「VST以前の」データにおいては見られるが)。すなわち、経験的には、「VSTされた」シグナルデータは、VSTによる変換の前よりも、包括分布的によりよく振舞う。図4 (Plot 7,4) はまた、高度に発現した、または「過剰に発現した」、すなわち非常に大きな遺伝子発現シグナル s の相対的小部分を捉える、右側に曲がった尾を示している。右側の尾はまた、全体の経験分布に対して、分布的によい振る舞いをしている。横軸に沿ってヒストグラムをより見やすくするために、図4にプロットされたシグナルデータは、VSTのシグナル s のうち最大の2%を除去している。

【 0 1 9 5 】

VSTパラメータ c_1, c_2, c_3 のサンプルごとの数値の表: 表6には、VSTを48803個の遺伝子 x 48個のサンプル集合の全体に不統一にVSTを適用した場合の、VSTパラメータ c_1, c_2, c_3 の典型的な研究結果が示されており、これは上記の説明で用いたものと同様である。むしろ、表6はVSTパラメータの数値結果、及び他の関連する数量を示しており、(48個のサンプルにわたる)パラメータの記述的統計を含んでいる。それらは、48個のサンプルの研究集合から一つ一つ各サンプルにつき計算したものである。48個のサンプルの研究集合から(48803個の遺伝子プローブのすべてにわたって)所与のサンプルのシグナル値が、基礎的なVST関数によって変換され(しかし、基数 e による対数の単位ではなく \log_{10} 単位にするために $1/\ln(10)$ で乗算され)たときに、(c_1, c_2, c_3)の3つ組の値は、その所与のサンプルに対して基礎的なVST式の中で用いられる。結果として生じる「VSTされた」シ

10

20

30

40

50

グナル値 (log10単位になっていると考えられる) は、後続の統計分析及び転帰予測分析の手続きに用いられる。

(表6)

実施例3

【0196】

この実施例は、GVHDの分類区分、及び、HCTドナー遺伝子発現測定値に基づいて患者のGVHD転帰における差違を決定するために用いられる統計的なT検定に関する説明を含む。「クラス分割」は、それぞれの移植患者が一又は複数の特定のGVHD転帰グループ(例えば、クラス1)を見せたドナーの定量的な遺伝子発現と、それぞれの移植患者が一又は複数の他の特定のGVHD転帰グループ(例えば、クラス2。クラス2は、2つのクラスの比較の定義上、クラス1と異なるものである)を見せたドナーの定量的な遺伝子発現との、数值的、数学的、統計的、又はコンピュータによる直接の比較を意味する。

10

【0197】

クラス分割は、2つのクラスの間での比較を行うが、2以上のクラスの間では同時に比較を行わない。表7に示されているように、所定の明確に定義されたクラスは、定義上、対応する移植患者のGVHDに関連する転帰の種類について、1以上のいわゆる「グループ(Group)」を含む。このように、クラス分割には、常に、正確に2つの定義されたクラスが関連するが、所定の定義されたクラスは、1つ以上の定義されたGVHD転帰グループを含むことができる。表7は、T検定分割を示す。

20

(表7)

【0198】

HCTレシピエントのGVHD転帰を予測するドナーの移植前CD4陽性T細胞のRNA発現プロファイルを特定するために、従来の単一遺伝子発現分析、すなわち1変量の統計的T検定(すなわち、線形判別分析(LDA)の一次元の形態)(Sheskin, David J. Handbook of Parametric and Nonparametric Statistical Procedures, 3rd Edition, Chapman & Hall / CRC Press, Boca Raton, Florida, 2004, esp. pp.404-409)が、5つの2クラス分割(表7)について、が実行された。ここでは、グループ1(急性及び慢性GVHDのいずれも存在しない)のサンプルを、GVHD陽性グループの様々な組み合わせと比較した。

30

【0199】

2クラス分割の各々について2種類のT検定が実行された。1つ目は、標準的な不均一の両側T検定とした。2つ目は、測定された(measured)遺伝子発現シグナル「予測重み付け(正確性重み付け)T検定」(不均一の両側検定でもある)とした。これは、イルミナ社のプラットフォーム(EAとの契約に基づく)によって提供されるいわゆる「ビーズ標準誤差」変数において報告されているように、イルミナ社のビーズアレイ(BeadArray)測定誤差の内在的な推定値(inherent numerical estimates)を、測定された遺伝子発現の各々について考慮に入れている。2つのT検定について用いられる式は以下のとおりである。

【0200】

標準T検定2クラス両側不均一T検定は、クラスP(GVHD転帰陽性)及びクラスN(GVHD転帰陰性)のプロープシグナル平均値、それぞれの不等分散、及びそれぞれのクラスごとのサンプル総数、値t及び値DF(自由度)に関して確立されている標準的な統計の式に従って実行された。また、p値は、標準的なコンピュータソフトウェア(Excel又はMatlab)のT検定機能(すなわち、標準的なT分布表における参照と同等)をコンピュータで実行することにより決定された。上記のプロープシグナル平均値は以下のとおり表される。

40

【数 3】

$$\overline{P}, \overline{N}$$

上記の不等分散は以下のとおり表される。

【数 4】

$$S_P^2, S_N^2$$

10

上記のサンプル総数は以下のとおり表される。

【数 5】

$$n_P, n_N$$

【数 6】

$$(1) \quad t = \frac{\overline{P} - \overline{N}}{\sqrt{\frac{S_P^2}{n_P} + \frac{S_N^2}{n_N}}}$$

20

30

【数 7】

$$(2) \quad DF = \frac{\left(\frac{S_P^2}{n_P} + \frac{S_N^2}{n_N}\right)^2}{\frac{\left(\frac{S_P^2}{n_P}\right)^2}{n_P - 1} + \frac{\left(\frac{S_N^2}{n_N}\right)^2}{n_N - 1}}$$

40

【0201】

正確性重み付け T 検定 2 クラスプローブシグナル測定値の正確性重み付け (precision-weighted) T 検定が、値 t 及び値 DF を決定するために、重み付け平均値 (式 (13) 及び式 (14)) 及び複合不等分散 (式 (20) 及び式 (21)) を用い、標準的な両側不

50

均一 T 検定と同じ基礎的な統計の式を用いて実行された。上記の重み付け平均値は以下のとおり表される。

【数 8】

$$\overline{P_w}, \overline{N_w}$$

上記の複合不等分散は以下のとおり表される。

【数 9】

$$S_{CP}^2, S_{CN}^2$$

10

【数 10】

$$(3) \quad t = \frac{\overline{P_w} - \overline{N_w}}{\sqrt{\frac{S_{CP}^2}{n_P} + \frac{S_{CN}^2}{n_N}}}$$

20

30

【数 11】

$$(4) \quad DF = \frac{\left(\frac{S_{CP}^2}{n_P} + \frac{S_{CN}^2}{n_N}\right)^2}{\frac{\left(\frac{S_{CP}^2}{n_P}\right)^2}{n_P - 1} + \frac{\left(\frac{S_{CN}^2}{n_N}\right)^2}{n_N - 1}}$$

40

【0202】

正確性重み付け T 検定において用いられる重みは、イルミナ社のデータの VST 処理から、各サンプルについて提供されるビーズ標準誤差 (be) の逆数に基づいて、以下のようにし

50

で定められる。これにより、1つのクラスについて全ての重みと同じであれば、重み付けされた発現量及びその平均が各クラスにおいて重み付けされていない対応する値と変わらないことが確実になる。

【0203】

各クラスにおける i 番目のサンプルについてのプリウエイト (pre-weight) の定義は以下のとおりである。

【数12】

$$PW_{Pi}, PW_{Ni}$$

10

ここで、

【数13】

$$(5) PW_{Pi} = \frac{1}{be_{Pi}}$$

20

【数14】

$$(6) PW_{Ni} = \frac{1}{be_{Ni}}$$

【0204】

各クラスについての平均プリウエイトの定義は以下のとおりである。

【数15】

$$\overline{PW_P}, \overline{PW_N}$$

30

ここで、

【数16】

$$(7) \overline{PW_P} = \frac{\sum_i^{n_P} be_{Pi}}{n_P}$$

40

【数 17】

$$(8) \overline{PW_N} = \frac{\sum_i^{N} be_{Ni}}{N}$$

10

【0205】

各クラスの i 番目のサンプルについての重みの定義は以下のとおりである。

【数 18】

$$W_{Pi}, W_{Ni}$$

ここで、

【数 19】

20

$$(9) W_{Pi} = \frac{PW_{Pi}}{PW_P}$$

【数 20】

$$(10) W_{Ni} = \frac{PW_{Ni}}{PW_N}$$

30

【0206】

プローブシグナル値の重み付けされた個別の値とクラス平均の定義 この定義は、各サンプル値、サンプル重みを蒸散し、各クラスについて重み付けされた値を平均かするという単純な考えに基づいている。プローブシグナル値の重み付けされた個別の値は以下のよう表される。

40

【数 21】

$$P_{Pi}, N_{Ni}$$

プローブシグナル値のクラス平均は以下のように表される。

【数 2 2】

$$\overline{P_w} \overline{N_w}$$

ここで、

【数 2 3】

$$(11) P_{wPi} = W_{Pi} \cdot P_{Pi}$$

10

【数 2 4】

$$(12) N_{wNi} = W_{Ni} \cdot N_{Ni}$$

20

【数 2 5】

$$(13) \overline{P_w} = \frac{\sum_i^{N_P} P_{wPi}}{N_P}$$

30

【数 2 6】

$$(14) \overline{N_w} = \frac{\sum_i^{N_N} N_{wNi}}{N_N}$$

40

【0207】

重み付けされたクラス分散の決定 (S^2_{wP} , S^2_{wN})。コンセプトは、(1) 計測されたシグナル及び重み付けされた平均 (各サンプルの分散寄与を反映) 間で、差が形成され、2乗され、及び重み付けされ (dev^2_{wPi} , dev^2_{wNi})、そして(2) この重み付けされ、2乗された偏差 (分散寄与) はその後、重み付けされた分散の合計を生成するために平均化される (S^2_{wP} , S^2_{wN})。

【数 27】

$$(15) dev_{wNi}^2 = w_{Ni} \cdot [(N)_{Ni} - \overline{N_w}]^2$$

【数 28】

$$(16) s_{wP}^2 = \frac{\sum_i^{n_P} dev_{wPi}^2}{n_P - 1}$$

10

【数 29】

$$(17) s_{wN}^2 = \frac{\sum_i^{n_N} dev_{wNi}^2}{n_N - 1}$$

20

【0208】

クラス内の分散に対するピーズ分散寄与の決定 (S^2_{beP}, S^2_{beN})。これは、クラス間のピーズ標準誤差の総体的な差を反映するために重要である。重み付けされたクラス分散 (S^2_{wP}, S^2_{wN}) に対して、ここまで全ての重み付けは各クラス内に制限されるので、クラス間のピーズ標準誤差自体の差のいずれか又は大部分を反映せず、このことは重要にもクラス分離の信頼性の定量化に寄与することができる。

【数 30】

30

$$(18) s_{beP}^2 = \frac{\sum_i^{n_P} be_{Pi}^2}{n_P - 1}$$

【数 31】

$$(19) s_{beN}^2 = \frac{\sum_i^{n_N} be_{Ni}^2}{n_N - 1}$$

40

【0209】

重み付けされたシグナル 分散及びクラス分散に対するピーズ分散寄与の和としての複合分散の決定。

これらの2つのクラスに関する分散 (S^2_{be} 及び S^2_w) を追加することによって、測

50

定エラー (be) に関する個別のサンプル測定における信頼性が (S^2_{be} を介して)、各クラスの平均測定エラー (クラス内の重み付けされた分散 S^2_w) と同様に考慮される。それゆえ、より確かで、数字上低く (すなわち、より統計学的に意味がある) より信頼できる p 値をもたらす、一層の確からしさが、より小さな測定エラー (be) を有するサンプル測定及びクラス内に配置される。

【数 3 2】

$$(20) s_{cp}^2 = s_{wp}^2 + s_{bep}^2$$

10

【数 3 3】

$$(21) s_{cn}^2 = s_{wn}^2 + s_{ben}^2$$

20

実施例 4

【0 2 1 0】

この実施例は、遺伝子発現分析及びドナー HCT の GVHD 転帰 - 予測の説明を含む。

【0 2 1 1】

「クラス分割」の概要及び詳細、異なる考えられる患者 GVHD 転帰、及び、真及び偽陽性及び陰性の符号 (tallies)、コンピュータによる / 統計学による転帰 - 予測の分類グループ: 各クラス分割及びプローブに対し、予測された GVHD 転帰分類精度を得るために、LDA が実行される (conventional lineardiscriminant analysis, the associated p-value being equivalent to a T-test whensingle-variate; Richard O. Duda, Peter E. Hart, & David G. Stork, PatternClassification, 2nd Ed., John Wiley & Sons, NY, 2001)。LDA は、RNA 発現量が閾値 (この特定の研究においては、LDA のための閾値は、厳密に、関係するそれぞれの 2 つのクラスからの陽性及び陰性の GVHD サンプル RNA 発現量の平均値間の中間であった) より大きい又は小さいかどうか依存して各サンプル (すなわち、ドナー) を GVHD 陽性 (すなわちレシピエント内に GVHD を誘発する)、又は GVHD 陰性 (すなわちレシピエント内に GVHD を誘発しない) として分類するために用いられる。サンプルがコンピュータにより / 統計学により、正確に又は不正確に分類されているかどうか依存して、それらは 4 つの異なるカテゴリ又はグループ ("group") (小文字 "g" の "group" であり、GVHD に関する大文字の "g" の "Group" と混同されない) の 1 つに分類される。1. TN (真 陰性)、計算により 陰性として分類された実際上の GVHD 陰性サンプル

30

40

2. FN (偽 陰性)、計算により 陰性として分類された実際上の GVHD 陽性サンプル

3. FP (偽 陽性)、計算により 陽性として分類された実際上の GVHD 陰性サンプル

4. TP (真 陽性)、計算により 陽性として分類された実際上の GVHD 陽性サンプル

【0 2 1 2】

同じ専門用語 (nomenclature)、TN、FN、FP、TP は、そのように分類されたドナーの数 (numbers) (すなわち、総数) と同様に、上述の分類状態を定義するために用い

50

られ、その用法は文脈から明確である。用語“サンプル”は、“ドナー”に対して又はデータを記述するドナーの分析された定量的な遺伝子発現に対して、互換的に用いられる。サンプル総数の合計は、各グループ（小文字“g”の“group”）に対してその後合計される。1. $TN_{tot} = TN$ サンプル合計 2. $FN_{tot} = FN$ サンプル合計 3. $FP_{tot} = FP$ サンプル合計 4. $TP_{tot} = TP$ サンプル合計

【0213】

GVHD 陰性合計、 N_{tot} 、及びGVHD 陽性、 P_{tot} 、研究に貢献するサンプルは発生（occurrence）の和であり、次のように定義される。1. $N_{tot} = TN_{tot} + FP_{tot}$ 2. $P_{tot} = TP_{tot} + FN_{tot}$

【0214】

統計、分類統計、及びデータマイニング分野における確立された慣習は、“バイナリ”転帰カテゴリが考慮され、即ち、“真”又は“偽”、さらに“偽陽性”イベントは“陰性”イベントとしてカウントされる。なぜなら、 FP_{tot} が TN_{tot} に加えられるという上に見られる N_{tot} の定義により、それは“陽性”ではないからである。“偽陰性”イベント及びすぐ上の2ラインにおける P_{tot} の定義に対しても、同様である。実施例5

【0215】

この実施例は、例示的な遺伝子発現投票モデルRNA20の説明を含む。

【0216】

LDAモデルのRNA20投票スキーム：例示的なRNA20モデル（あらゆるGVHD vs. 非GVHDディビジョン）にとって、各コンポーネントRNAマーカーLDAモデルは、yes/no（1/0）予測、すなわち、各サンプルに対するGVHD 陰性転帰にとっての投票を提供する（各20RNA種の122HCTドナーを通した一連のGVHD 陰性投票は、表8における分離する行に表示される）。20RNA種を通した全てのGVHD 陰性投票は、個別のマーカー投票プロファイル（表8）の下に表示される“GVHD 陰性スコア”を得るために、各サンプルに対してカウントされ、RNA種の総数（すなわち、20）で除される。

【0217】

サンプルは最終的に、GVHD 陰性スコアが（ユーザに選択された）0.77という閾値より大きい（above）場合には、この20マーカーモデルによってGVHD 陰性として分類され、すなわち、少なくとも20RNA種に基づく投票の合計の77%という大部分が、“GVHD 陰性”（黒又はダークグレー地に白文字の値）として分類されるサンプルとして要求される。同様に、GVHD 陰性スコアが0.77という閾値より小さい（below）場合（白又はライトグレー地に黒文字の値）には、サンプルはGVHD 陽性として分類される。GVHD 陰性スコア、グループ1としてレポートされる、真・陰性、偽・陽性及び N_{tot} の合計数（集計行）、及び、各グループ2から6としてレポートされる、偽・陰性、真・陽性及び P_{tot} の合計数（集計行）が表8に示される。表8は、RNA20投票スキームである。

（表8）

【0218】

RNA種に寄与する20の詳細及びそれらの個別のLDA分類パフォーマンスのリストが表9に示される。表9は、RNA20リスト

（表9）

【0219】

真及び偽、陽性及び陰性に関する異なる分類の表現に関わるドナー数の固有差に起因するバランス効果：GVHD 陽性及び陰性サンプルグループからの寄与を等しく数字上バランスさせるように、全4転帰分類グループの関係する貢献が決定され、GVHD 陽性及び陰性グループのサイズの合計、すなわち、関係する各サンプルの数における固有の不等に対してバランスさせる。1. $TN_{bal} = 0.5(TN_{tot} / N_{tot})$ 2. FN_{bal}

10

20

30

40

50

$$a1 = 0.5 (FN_{tot} / P_{tot})^3 \cdot FP_{bal} = 0.5 (FP_{tot} / N_{tot})^4 \cdot TP_{bal} = 0.5 (TP_{tot} / P_{tot})$$

【0220】

バランスされたGVHD陽性及び陰性サンプルの寄与は、ここで、それぞれ0.5に等しく、和は1となる：1. $N_{bal} = TN_{bal} + FN_{bal} = 0.5$ 2. $P_{bal} = TP_{bal} + FP_{bal} = 0.5$ 3. $P_{bal} + N_{bal} = 1$

【0221】

4つのバランスされた転帰分類グループを用いることにより、5つの異なるバランスされた転帰予測測定が決定される（以下、用語TN、FN、FP、及びTPの全ての使用は、それぞれ、 TN_{bal} 、 FN_{bal} 、 FP_{bal} 及び TP_{bal} のことを指す）：1. バランスされたNPV（陰性予測値）= $TN / (TN + FN)$ ・本当は陰性であり陰性として分類されたサンプルの分数 2. バランスされたTNR（真 - 陰性率）又は特殊性 = $TN / (TN + FP)$ ・正しく分類された総陰性サンプル数の分数 3. バランスされたPPV（陽性予測値）= $TP / (TP + FP)$ ・本当は陽性であり陽性として分類されたサンプルの分数 4. バランスされたTPR（真 - 陽性率）又は特殊性 = $TP / (TP + FN)$ ・正しく分類された総陽性サンプル数の分数 5. バランスされた精度 = $(TP + TN) / (TP + FP + TN + FN)$ ・正しく分類された総サンプル数の分数

【0222】

NPV、TNR、PPV、及びTPRの基本的な定義は、統計学、分類統計学、データマイニングにおける標準的な従来からの定義である。バランスされたバージョンは、これゆえ、標準的なバージョンに依存する。しかしながら、それらは、TN、FN、FP、及びTPの類似のバランスされたバージョンを用いる。以下、特に明記しない限り、用語NPV、TNR、PPV、TPR及び精度（Accuracy）の全ての使用は、バランスされたNPV、バランスされたTNR、バランスされたPPV、バランスされたTPR及びバランスされた精度（Accuracy）のことを指す。

実施例6

【0223】

この実施例は、例示的なRNA20モデルに基づく遺伝子発現の分析結果及びGVHD転帰の予測についての説明を含む。

【0224】

本研究における122のHCTはすべて、HLA10/10適合の非血縁ドナー移植に対応しており、このことは、米国における毎年の移植の過半数を反映している。上述のとおり、移植GVHD転帰は、6つの異なるグループに分類された（表5）。

【0225】

グループは、GVHDの重症度が増加する順序で番号が付けられており、急性GVHDも慢性GVHDも示さないグループ1から始まり、グレード3又はグレード4の重度の急性GVHD及び広範囲の慢性GVHDを示すグループ6で終わる。グループ5は、グレード3又はグレード4のGVHDを示すが、慢性GVHDは示さない。グループ4及びグループ4は、グレード1又はグレード2の急性GVHDを示し、それぞれ、慢性GVHD示す場合と示さない場合に対応する。グループ2は、慢性GVHDのみを示し、急性GVHDは示さない。グレード3又はグレード4の急性GVHDは、最も激しく及び生命を脅かすGVHDの形態である。グレード1又はグレード2の急性GVHDは、重症度がより低く、軽度と考えられることもある。急性GVHDのグレードの分類は、広範囲の又は広範囲ではない慢性GVHDについても同様であるが、GVHDの重症度を分類する医師にとって臨床又はがん実務において確立された複数の症状の診断的な分類である。グループの定義は本明細書に独自のものであるが、グループは、NMDPの専門家によって確立された医学的に重要なGVHD重症度のグループである。

【0226】

異なるGVHDグループに関連付けられたHCTドナー間での巨視的な発現レベルの傾向ドナーCD4陽性T細胞RNA発現プロファイルとHCTレシピエントGVHD転帰との間

10

20

30

40

50

の関係を特徴付けるとともに、データセット中の潜在的なバイアスを、生物学的に根拠のある関係と区別するために、G V H Dグループ平均RNA発現レベルの全体的な挙動傾向が、マイクロアレイ遺伝子発現プロープについて並べられているランクとして分析された(表10)。G V H DグループRNA発現の順位が、各遺伝子プロープについて、6つのグループの各々についての登り順で(すなわち、発現結果のレベルが高いとランクが高くなる)、平均遺伝子発現レベルのランクとして、決定された。データセットは、その後、G V H D陰性サンプル(グループ1)の平均RNA発現がG V H D陽性サンプル(グループ2~グループ6)よりも高いか低いかに応じて、「N>P Subset」及び「P>N Subset」で示される2つのサブセットに分けられる。これらの2つのサブセットについて、2つの異なる定義のプロープ内、すなわち、(1)全48,803のプロープ、すなわち「N>P Total Subset」及び「P>N Total Subset」、及び(2)不均一分散T検定及び上述した正確性重み付け(正確性重み付け)T検定の両方についてT検定p値 $s \leq 0.05$ を有する1024のプロープの選択サブセット(Select Subsets)、内で、G V H DグループRNA発現ランクの中央値、すなわち「ランク(Rank)」が決定される。これらの処理が、任意のG V H Dと非G V H Dの分類区分(表6)、すなわち「N>P Select Subset」及び「P>N Select Subset」について実行される。表10は、グループ順位を示す。

(表10)

【0227】

「N>P Total Subset」及び「Select Subsets」については、グループ1についてのランク(Rank)(RNA発現ランクの中央値)は、グループ2からグループ6のランク(Rank)の各々と比較すると、「N>P Subset」よりも例外なく一貫して大きい。同様に、「P>N Total Subset」及び「Select Subsets」については、グループ1のランク(Rank)は、グループ2からグループ6のランク(Rank)の各々と比較すると、「P>N Subset」よりも例外なく一貫して小さい(表10)。「N>P Total Subset」及び「P>N Total Subset」の選択基準は、グループ1の平均を、サンプル of グループ2~6の組み合わせられたサンプルの平均と比較する場合に限定され、グループ2~6の個別の平均とは比較されない。したがって、(1)「N>P Subset」及び「P>N Subset」について適用される選択基準は、グループ2~6のランク(Rank)がグループ1のランク(Rank)よりも一様に大きいことや小さいことを必ずしも保証するものではなく、(2)実際にそうであるという事実は、グループ2~6のメンバーの任意のものに他のグループ2~グループ6のランク(Rank)と比較されるグループ1のランク(Rank)の他の側にそのランクを配置するという強いバイアスが存在しないことを示す。

【0228】

また、「N>P Total Subset」及び「P>N Total Subset」について、グループ2~6のランク(Rank)は全て同じである。これは、調査された48,803のマイクロアレイプロープ(表10)のすべてにわたって、グループ2~6のランク(Rank)の一様性が高いレベルにあることを示している。これとは対照的に、プロープに関連する選択された1,024のG V H D転帰については、グループ2~6内でのランク(Rank)の順位は、「N>P Select Subset」について、ランク(Rank)4~2まで明確な下降トレンドを示している。また、「P>N Select Subset」については、G V H Dグループの数及び関連付けられたG V H Dの重症度の上昇と平行して、ランク(Rank)2~5までの上昇トレンドを示している。グループ2~6の「Rank」一様性の分散(48,803プロープの総セットについて測定される)は、恣意的なバイアスを示していない。むしろ、きちんとした平行トレンドを示している。遺伝子発現の大きさがG V H Dの重症度と相関する場合、「N>P Select Subset」については-0.95、Rank順位とG V H Dグループメンバーとの間において「P>N Select Subset」については+0.92という極めて高いピアソン相関(R)により計測可能に裏付けられている。

【0229】

要するに、重要な洞察は、「N>P Select Subset」及び「P>N Select Subset」(表10)の1,024のプロープについての、G V H Dグループ疾患の重症度の順位とランク(Rank)

10

20

30

40

50

(RNA発現ランクの中央値)の順位との強い相関は、G V H D (any G V H D) 対非 G V H D (no G V H D) のクラス分割についての T 検定の結果に従ってこれらのプローブを選択すること自体では、5つの G V H D 陽性グループ 2 ~ 6 内において、順位や区別される特徴を選択できなかったということである。これは、グループ 2 ~ 6 のサンプルが単純に T 検定分析用に集められており、そのために特定の G V H D 陽性グループソースに関する全ての情報が失われていたためである。したがって、G V H D 陽性グループ内での観測されたランク (Rank) の順位は、RNAプロファイルの選択セットによって示された内在的で天然の生物的特徴であり、統計分析においてプローブを選択した方法とは独立である。換言すれば、グループ 2 ~ 6 のランク (Rank) と G V H D グループメンバーとの強いピアソン相関は、統計的作為としてデータの分析や処理によってもたらされたものではなく、ラン

10

【0230】

G V H D 転帰予測は移植センターについてのバイアスを示さない。以下に、具体的な例として、2つの特定の個別的なRNA発現予測について、G V H D 転帰予測挙動及びサンプル移植センターソースの分散が調査された。具体的には、発現レベルが G V H D の強度とともに増加する傾向のある「CTCF」(CCCTC結合因子)、発現レベルが G V H D 強度とともに減少する傾向のある「BLVRA」(ピリベルジン還元酵素 A)、及び、例示的な20のRNA発現セット「投票」モデル(方法及び表8)である「RNA20」(表9に記載されているRNA種)について、調査を行った。

20

【0231】

RNA発現の測定値を、全122サンプルについて、6つの G V H D 転帰クラスの各々について登り順でプロットし、サンプル移植センターソース(TCS)(図5~7、CTCF TCS、BLVRA TCS、RNA20 TCS)に応じてラベルをつけた。少なくとも4サンプルを提供している移植センターは、別の色でラベルをつけた(n=85)。2~3のサンプルを提供するセンターは are labeled by 三角形でラベルをつけた(n=13)。また、サンプルを1つだけ提供しているセンターは、四角でラベルをつけた(n=24)。各データ点には、移植センターの番号でラベルをつけ(センターの実際の名前は伏せた)、ダッシュのあとに当該センターによって提供されたサンプルの数を記載した。これらの3つの例の全て(CTCF、BLVRA及びRNA20モデル)において、複数のセンターからのサンプル及び単一のセンターからのサ

30

ンプルは均一に分散されており、特定の発現量領域内でのクラスター、又は、toward 特定の G V H D 転帰グループに対する関連するバイアスは示されていない。遺伝子発現解析又は G V H D 転帰属性を歪めたりバイアスをかける可能性のある様々な移植センターによって変動が導入される可能性があるにもかかわらず、3つの具体例の各々は、T検定性能及びLDA精度の観点から(表11)、強い G V H D 転帰予測能力を示している。表11は、性能(performance)を示す。

(表11)

【0232】

異なるRNAモデルCTCF、BLVRA及びRNA20に関する転帰予測観察他の系列のプロットにおいて、サンプルには、6つの G V H D 転帰グループに応じてラベルが付されている。各グループについて平均RNA発現量を挿入して、G V H D 転帰グループ及び付随する G V H D 臨床強度の番号に対する遺伝子発現の増加傾向又は減少傾向を特に示している。図8においては(CTCF GROUPS)、G V H D グループの番号に対して安定して一本調子に増加する G V H D グループの平均値の系列が見られた。図9(BLVRA GROUPS)は、G V H D グループの平均値が、G V H D グループの番号に対して、安定して下降傾向にあることを示している。グループ1においては、非常に低い検出限界値は存在せず、グループ6においては、G V H D の最も深刻な態様を表す非常に高い値が存在しない。

40

【0233】

図10においては(RNA20 GROUPS)、20の良好な個別のLDAモデルから G V H D 陰性の票(negative vote)の相対スコアをプロットしており、G V H D の重症度が増加するに

50

つれてGVHDグループの平均スコアが安定的に減少する傾向が見られる。非GVHD (no GVHD) グループ (グループ1) の平均と5つのGVHD陽性グループ (グループ1~グループ6) の平均との間にはかなり大きなギャップがある。(命名の注意として、1つのクラスは1つのグループ、典型的にはグループ1を含むことができ、他のクラスは複数のグループの含むことができる。例えば、「anyGVHD」を表すクラスはグループ2~6を含む)。

【0234】

「anyGVHD」対「noGVHD」(表6)について、サンプル特有のGVHD転帰予測が、2つの個別のRNA発現マーカーに対応するLDAモデルについて詳細にプロットされている。また、図11~図13(CTCF LDA、BLVRA LDA及びRNA20 LDA-A)には、「20 RNA」マーカー投票モデルについて、サンプル特有のGVHD転帰予測がプロットされている。本質的に、サンプルは、それぞれの発現レベルが当該クラスの平均観測発現としての区分線の同じ側にあるか否か(上又は下)に応じて、GVHD陰性又は陽性クラスに分類されている。すなわち、図11(CTCF LDA)においては、サンプルは、区分線の下側においてGVHD陰性として分類されており、図12及び図13(BLVRA LDA及びRNA 20 LDA)においては、サンプルは、区分線の上にあるときに、GVHD陰性と分類されている。次に、分類が正確か否かについて、当該サンプルの既知の正しいクラスと比較し、次に、各サンプルを、TN(真陰性)、FN(偽陰性)、FP(偽陽性)、又はTP(真陽性)のいずれかに採点した。図11及び12(CTCF LDA及びBLVRA LDA)に示されているように、圧倒的多数のGVHD陽性サンプルが正しく分類されていたが、陰性サンプルの多くは不正確に分類されていた(偽陽性)。この非対称性の一部は、陽性サンプルの数(n=96)と陰性サンプルの数(n=26)との非対称性の直接的な反映である。すなわち、少ない相対的な偽陰性の分類率(陽性サンプルを示す偽陰性)であるが、この率は、 $96/26=3.7$ 倍することにより、陰性として誤分類された陽性サンプルの推定観測数と同じになる。注目すべきは、図13(RNA20 LDA-A)において、陽性及び陰性の非対称的な表現にもかかわらず、わずかに4つのGVHD陽性サンプルしか陰性と誤分類されていないことである(NPV = 0.94)。20個の良好なRNA発現LDAモデルの単純な投票モデルが、様々なソースの偶然によるシグナル変動及びノイズによって起こる誤分類の大部分を克服できることを示している。

【0235】

この例示的なRNA20投票スキームは、20個のRNAマーカーRNAモデルの各々について、「サンプルがGVHD陰性として分類されるには、GVHD陰性票の少なくとも77%の多数が必要」というGVHD陰性予測閾値と組み合わせて(参照方法及び表7)、GVHD陰性予測の単純な集合、すなわち投票を反映している点に留意されたい。

【0236】

本来、投票スキームは、ソースのうち20個の情報を組み込むことにより、各構成要素マーカーのソースの限界及び誤差を克服するためのものであり、投票閾値の設定を通じてGVHD陰性転帰予測の厳密さ(stringency)の定義に柔軟性を提供するためのものである。例えば、この例において用いられた0.77という閾値は、真陰性の相対的に高い数及び高いTNRを維持する一方で、偽陰性を最小化し陰性予測値(Negative Predictive Value)を最大化するために人為的に選択された(図18、RNA20 LDA PERFORMANCE - A)。これに対応して、表7において、6つのGVHDグループの各々について、例示的なRNA20モデルによって分類された偽陰性の数が非常に少ないことが観察される。グループ5及び6の両方とも、24の異なる米国臨床センターから得られた合計39個のサンプルであるが、1つの偽陰性予測も観測されなかった点に注目されたい。臨床応用については、低い偽陰性率が非常に望まれる。すなわち、GVHDを引き起こさないために、移植前にHCTに適したドナーと宣言したにもかかわらず、実際にはそのドナーのHCTによって移植後に患者が不幸にもGVHDを避けられないということが起こる率を極めて低くすることが望ましい。

【0237】

「anyGVHD」対「noGVHD」転帰の予測をうまく行うことに加えて、20 RNA発現LDA投票モ

デルは、慢性及び急性のGVHDの異なる種類及び強度を「no GVHD」転帰からうまく区別することもできる（表6及び11参照）。慢性GVHD（単独で又は任意の形態の急性GVHDと組み合わせて）を「no GVHD」転帰から区別することに関し（「c GVHD」対「no GVHD」）、わずか2例の偽陰性分類が報告された（図14、RNA20 LDA - B）（NPV = 0.95）。急性GVHD（単独で又は任意の形態の慢性GVHDと組み合わせて）を「no GVHD」転帰から区別することに関し（「a GVHD」対「no GVHD」）、僅かに3例の偽陰性分類が報告された（図15、RNA20 LDA - C）（NPV = 0.94）。急性GVHD（任意の形態）と組み合わせられた慢性GVHDを「no GVHD」転帰と区別することに関し（「a&c GVHD」対「no GVHD」）、わずかに1例の偽陰性分類が報告された（図16、RNA20 LDA - D）（NPV = 0.96）。注目すべきことに、グレード3又はグレード4急性GVHD（単独で又は任意の形態の慢性GVHDと組み合わせて）を「no GVHD」転帰から区別することに関し（「a34 GVHD」対「no GVHD」）、1例の偽陰性分類も報告されなかった（図17、RNA20 LDA - E）。

10

実施例7

【0238】

この実施例は、GVHD転帰予測性能のサマリーを含む。

【0239】

GVHD転帰予測結果のサマリー 単一の及び20個のRNAマーカーモデルの全てについての数値的な転帰分類の結果は、上述され、また、表11にまとめられている。LDA転帰予測測定（LDA outcome predictivemeasure）は、均衡（balanced）LDAモデルに基づくものであり、GVHD陰性及び陽性転帰サンプル（本発明の方法において説明した）のサンプルに応じた（サンプル-wise）貢献値と同数を表すように調整される。クラスをはっきりと識別する $6.6E-03$ 及び $4.3E-03$ のT検定p値（表記法： xEy は、 x に10の y 乗を乗じるという意味であり、科学分野において標準的な表記法である）が「any GVHD」対「no GVHD」分割の単一値モデルについて報告されているが、20RNAマーカーモデルについての当該p値は、数オーダーの大きさで小さく、 $1.1E-08$ である。単一RNAマーカーモデルについて、69%という全体として高い正確性が報告されているが、20マーカーモデルについての正確性はもっと高く81~83%の範囲である。

20

【0240】

単一マーカーモデルにおける67~69%からRNA20投票モデルにおいて90%以上となる陰性予測値（Negative Predictive Value、NPV）の上昇が、臨床的な重要性が最も大きなものの1つである。NPVは、サンプルthat were classified as GVHD陰性と分類されたサンプルのうちどれだけが実際に陰性となるかを表す。例えば、本研究のデータを用いれば、「34 GVHD」対「no GVHD」転帰予測についての100%のNPVは、このモデルを用いてGVHD陰性に分類されたドナーを使用する場合には、その移植のうちグレード3又はグレード4の急性GVHDになる例は1つもないということである。このことは、GVHDの発生が完全に、100%除去されることと対応する。現在は50%程度のGVHDが観察されている。このことは、HCTレシピエントとのHLA適合に基づくHCTドナーの選択と比べた重要な改善である。

30

【0241】

転帰予測におけるトレードオフは、閾値の意図的な再設定を行った場合に発生しえる（すなわち、区分線の再配置）。GVHD転帰予測についてのLDAモデルにおける正確性を調べる際には、区分線がどこに位置するかに応じて、1つの達成指標におけるメリットが他の達成指標におけるデメリットとなることがある。例示的なRNA20モデルについて、陰性予測値（Negative Predictive Value）（NPV）が増加すると、真陰性率（True Negative Rate、TNR、特異性）が減少する（図18、RNA20 LDA PERFORMANCE - A）。このような「検出閾値依存性（detector threshold-dependent）」のトレードオフ現象は、一般に知られており、「ROC」曲線（Receiver Operating Characteristic曲線）と省略されることがある。GVHD陰性に分類されたサンプルの多くが真に陰性（NPV）であるため、より少ない数の総陰性サンプルが正しく分類されている（TNR）。換言すれば、移植前に誤

40

50

って陰性と分類されたGVHD陽性移植の数を最小化するコストは、臨床的には、GVHD陰性移植に向かう一部のドナーが、実際には移植をしないということである。しかしながら、例示的なRNA20モデルにおいては、区分線は、相対的なGVHD陰性投票スコア0.77に位置し、94%のNPVが65%のTNRと共に得られている。換言すれば、GVHDを避けられることが94%確かであれば、GVHD陽性に誤分類された35%のドナーの口スを受け入れるであろうということである。しかしながら、HCTドナー候補を失うことを除けば実質的な害はない。5つのLDA正確性指標の詳細な挙動は、陽性評価値(Positive Predict Value、PPV)及び真陽性率(True Positive Rate、TPR、Sensitivity)を含み、図19に示されている(RNA20 LDA PERFORMANCE - B)。

実施例 8

【0242】

この実施例は、GVHD転帰の決定及び予測における比較用の正規化又は標準化された値に対する遺伝子発現の比の使用についての説明を含む。

【0243】

分析化学、分析物理学、及び定量測定においては、レシオメトリック分析は、ただ1つの検体又は測定量を用いる類似の解析法よりも正確である。レシオメトリック分析の内在的に正確性が優れている点は、有利性がある場合には、測定値が同一の器具を用いて同一のサンプルについて測定された適切な基準シグナルに対する対象シグナルの比として報告されるときには、幾つかの信頼できる参照基準に対して実質的に自己校正を行う点である。

【0244】

一般に、レシオメトリック分析は、(1)対象測定量Aについての測定基準値で対象測定量Aを割った比。(2)2つの区別され異なる(関連するものであっても独立のものであってもよい)測定量A及びBの比、又は、それ以上の複雑な比、例えば、Aの基準値に対するAの比をBの基準値に対するBの比で割ったもの。(3)(1)及び(2)で説明されたのと同種の比。ここで、分子及び分母は、測定シグナルと測定されたバックグラウンドシグナルとの差である。(4)上記(1)~(3)で説明された種類の比の組み合わせ。

【0245】

2つの異なる結果の予測又は決定の信頼性は、2つの分離された単一検体解析を用いるよりも、2つの検体又は予測因子を用いるレシオメトリック分析に基づく場合に優れている。レシオメトリックを用いた場合の正確性が増す現象は、比を考慮することによって2つの異なる結果の差違(すなわち、評価又は予測結果)が内在的に増加するときに発生する。例えば、転帰1の原型が遺伝子Aの発現が高くかつ遺伝子Bの発現が低いことで特徴付けられるが、転帰2の原型は遺伝子Aの発現が低くかつ遺伝子Bの発現が高いことで特徴付けられると仮定する。この状況においては、典型的には、AのBに対する発現の比は、1つの転帰を他のものから区別するとき又は他のもので

はなく1つの転帰を予測するときには、1つの遺伝子を単独で又は両方の遺伝子を別々に用いた場合よりも、遺伝子発現比を用いた場合に、より正確で、より感度が高くなる。

【0246】

繰り返しになるが、遺伝子発現のレシオメトリック分析(すなわち、遺伝子発現比)は、以下のように定式化される。(1)特定の対象遺伝子(すなわち、予測遺伝子)の発現のハウスキーピング遺伝子、又は、1セットのハウスキーピング遺伝子について評価された集計値に対する比、又は、(2)2つの異なる特定の対象遺伝子発現の比、すなわち、2つの異なる予測遺伝子発現の比。上述のとおり、(1)のバージョンは、予測因子の測定値の内部的な自己校正という特徴を有しており、(2)のバージョンは、自己校正の特徴及び2つの異なる対照結果(上述した)に対して差違が増加するという特徴がある。

10

20

30

40

50

【0247】

比を用いることが別個の測定と比べて明確に有利な点の1つは、測定値が、所定の測定値を得た器具について定数である不確定の倍数因子又はゲインによって拡大される場合である。比は、個別の測定値よりも特に閾値からの差違に関してより正確となる。これは、定数（しかし、未知）の倍数因子が、2つの測定値の比における分子及び分母の両方で等しく相殺されるためである。

【0248】

また、比には、ある状況において得られる、測定値ノイズ除去（技術的ノイズ、統計的ノイズ、又は生物学的ノイズ又はランダム変動）の有利な特徴もある。これは、別途用いられる単一遺伝子（単一予測因子）測定では得られない利点である。具体的には、2つの異なる遺伝子の発現が平均して逆相関である場合（すなわち、一方が増加する傾向のときに、他方が減少する傾向、またはその逆の場合）には、対数形式（又は対数状の形式、例えば、逆双曲正弦変換）に再計算されたときに、2つの別個の遺伝子発現測定値の擬似ランダム要素は、当該2つの遺伝子発現の測定値が本来的に比で構成されている場合には、相殺される傾向がある。これは、一般的な数学もしくは計算又は統計事象であり（すなわち、遺伝子発現のみに限定されるものではない）、「対照変量法（the method of antithetic variables）」として知られている（C Eisenhart & M. Zelin, Elements of Probability, Ch. 12, in EU Condon & H Odishaw, eds., Handbook of Physics, McGraw-Hill Book Co., NY, 1958, pp. 1-143, P. Kevin MacKeown, Stochastic Simulation in Physics, Springer-Verlag, Singapore, 1997, esp. p.21 & p. 212）。2つの変数が逆相関のときには、各変数のノイズ又は変化の合計よりも、2つの変数の合計の方がノイズ又は変化が少ない。

10

20

【0249】

比の対数関数又は逆双曲正弦等の準対数関数が用いられる場合には、対数比は、分子の対数と分母の対数との差となる。分子及び分母の変数がより多く又は少なく逆相関している場合には、「対照変量（antithetic variables）」は部分的にノイズを除去し、又は変化の減少が起こる。すなわち、「対照変量」によるノイズの部分的な除去の現象は、レシオメトリック測定を用いるときに得られる追加的で副次的な利点である。

実施例9

30

【0250】

この実施形態は、単一又は複数の予測遺伝子の測定された発現レベルに基づいて、連続的な尺度、スコア、又は指標でのGVHDリスクの予測に関する説明を含む。

【0251】

予測遺伝子の遺伝子発現、又は予測遺伝子及び基準遺伝子（例えば、ハウスキーピング遺伝子）を含む発現比、又はHCTドナーにおける2つの異なる予測遺伝子の発現比は、原則として、連続した数値を表す。各測定値、測定値の間隔、又は測定値の範囲（その連続値からの）は、実験的に、計算的に、又は統計的に裏付けされた方法で、移植前に、ドナーのHCTが、HCTレシピエント（患者）において、移植後に、GVHDを引き起こすか否かの具体的に予測されるリスクと関連付けられる。

40

【0252】

GVHDの番号又は値が基準発現測定値よりも低いか高い可能性あるリスクの1つの例、GVHDの発生リスクが所定の可能性よりも低い又は高い、又はその逆。このことは、GVHDのリスクが低いこと又は高いこととして説明される。GVHD転帰予測は、本質的に、その後、2つの方法で実行される。（1）GVHDのリスクが低いこと又は高いことの報告、又は特定の可能性又は可能性の間隔又は範囲の報告、又は（2）GVHDの可能性又はGVHDのリスクのスコアの連続した値としての報告。このようなリスクスコアは、測定された遺伝子発現に適用されるGVHD転帰予測の数学的、統計的、又は数値的モデルの元で、連続値から得られると考えられ、単純な数学の式の使用により実行される。または、このリスクスコアは、同一の数値的入力（遺伝子発現）及び出力（GVHDの

50

リスク) マッピングを有する事前計算された数値的な参照テーブルから得られる。

【0253】

また、複数の異なる遺伝子又は複数の発現比が関連する測定された発現量は、さらに複数の方法、例えば、単純加算又は事前定義された定数による再スケーリングをしての加算、又はこれ以外の単純で明確な数学的処理において組み合わせられ、GVHDリスク、又はGVHDの可能性、又はGVHDのリスクスコア等、又は具体的に定義されたGVHDのリスク又はGVHDの可能性のためのしきい基準値の連続値と関連付けられる連続値出力変数に到達する。

10

【0254】

また、GVHD転帰予測遺伝子の発現測定値又は当該予測遺伝子及び基準(例えば、ハウスキーピング)遺伝子が関連する比、又は2つの異なる遺伝子が関連する比は、各々が個別に高GVHDリスク票(例えば、1又は1よりわずかに小さい数値)、又は低GVHDリスク票(例えば、0又は0よりわずかに大きい数値)を投じることができ、このような測定値又は測定比に基づく投票は、全体としての投票スコア又は指標を形成するために集計される。投票スコア又は指標は、0から1の範囲で広がる連続値(continuum)又は準連続値(quasi-continuum)を形成すると考えられる。投票スコア又は指標は、GVHDのリスク又はGVHDの可能性の連続値(continuum)と関連付けられる、又はGVHDが起きやすいか置きにくいかということについての所定の閾値よりも上又は下に収まり、又はGVHDのリスク度合いを定性的に知らせる所定の間隔(0から1の範囲で広がるスコアを区切る)の中に収まる。

20

実施例10

【0255】

この実施例は、上述した122ドナーPBMCサンプルと組み合わせられる、120ドナーPBMC(末梢血単核細胞)サンプルの追加的な分析についての説明を含む。合計では242ドナーのPBMCサンプルとなる。GVHD予測遺伝子を同定するための対応するレシピエントGVHD歴も用いられる。

30

【0256】

従来のコンピュータによる相互確認は、単一遺伝子及び投票スキームの転帰予測性能を評価するために適用されていた。転帰予測能力に関して遺伝子を他の遺伝子との間でランキングするために、限定線形判別分析の一形態を採用し、それぞれのドナーに関連付けられた患者について異なる程度のGVHD対非GVHD(「no GVHD」)を識別する点について、遺伝子の性能を評価した。上述した改善されたコンピュータ統計分析を、合計で242のドナーのサンプル遺伝子発現プロファイルについて、GVHD転帰情報(上述したように分析された最初の122のドナーサンプルに追加された120の新しいドナーサンプル)を用いて実行した結果、追加的な121の遺伝子が、GVHD転帰予測として同定された。また、23の遺伝子がハウスキーピング(「HSK」)遺伝子(当初の122サンプルの分析に基づいて以前は予測遺伝子として挙げられていた真核生物翻訳開始因子4H(EIF4H)の転写変異体の1遺伝子を含む)として同定された。合計で、143の新しい遺伝子(表2A、RNA143)が診断試験候補遺伝子(diagnostic test candidate gene)として同定された(121の転帰予測遺伝子、22のハウスキーピング遺伝子、全てがRNA192リストにおいても完全に列挙されている。表2B)。これは、当初の122サンプルの分析結果には含まれていなかった。また、その結果、以前のRNA1546リストにも含まれていなかった。完全な242ドナーのサンプル/GVHD歴(GVHD history)のデータセットについてのより深い分析の結果、192遺伝子(表2B、169の予測遺伝子及び23のハウスキーピング遺伝子)は、さらに高い信頼性のRT-PCR遺伝子発現解析のための例示的な「RNA192」遺伝子リストとして選択された。

40

50

実施例 1 1

【 0 2 5 7 】

この実施例は、a description of リアルタイム逆転写 (RT) 定量ポリメラーゼ連鎖反応 (PCR) 測定 of N及びP 予測遺伝子発現候補及びN及びP 予測遺伝子発現データ。

【 0 2 5 8 】

人間の臨床的診断に応用するために、遺伝子発現の測定のためのRT-PCR (逆転写ポリメラーゼ連鎖反応)、例えば、TaqManリアルタイムRT-PCRプラットフォーム (ABI, Applied Biosystems Inc.) において実行されるものが、一般的に正確性が低く、感度も低いと考えられているマイクロアレイ遺伝子発現分析 (例えば、48,000程度の異なる遺伝子特異プローブを使用できるイルミナ社のHT12 v3.0マイクロアレイプラットフォームを用いる、本明細書において説明した、GVHDドナーサンプルからのmRNAレベルの調査において用いられるもの)と比較して、忠実性が高い定量的な遺伝子発現レベル解析のための代表的な手法と考えられている。

10

【 0 2 5 9 】

リアルタイムRT-PCR (逆転写ポリメラーゼ連鎖反応) 遺伝子発現データは、192の特定の遺伝子 (表 2 B に「RNA 192」として記載されている) 及び180の異なるドナーの凍結血液サンプルについて取得された。従来のTaqManプラットフォームが用いられた。1反応あたり100ngのcDNAを使用した (cDNAは上述のようにして得られたRNAサンプルから誘導された)。各遺伝子は、off-the-shelf TaqMan assay ID (ABI, Applied Biosystems Inc.) に応じた固有のプライマーによって決定された。技術的に有効で商業的に利用可能なTaqManアッセイは、遺伝子固有のプライマー及び加水分解プローブのヌクレオチド配列を決定する。しかしながら、正確なヌクレオチド配列は、ABIに独占的に所有されている。ABI TaqManアッセイは、PCR増幅効率が2となるように最適化されている。

20

【 0 2 6 0 】

以下は、従来の工業標準となっていて、長年の使用で、本質的にコモディティ化されている。RT-PCRは、最初のRT (逆転写) 工程を含む。この逆転写工程は、RNAをcDNAに変換し、続いて、cDNAのPCR (ポリメラーゼ連鎖反応) 増幅が行われる。リアルタイムRT-PCRは、遺伝子配列に特有のプライマーに加えて、遺伝子配列に特有の内部 (internal) 加水分解プローブの使用を含む。この内部加水分解プローブは、蛍光プローブで化学的に修飾されたヌクレオチドを含み、加水分解プローブ鎖の近くに配置されるときに意図的に他の蛍光を抑制する。内部加水分解プローブが増幅された遺伝子特異的なcDNAの単鎖に結合するときに、PCRにおいて用いられる熱安定DNAポリメラーゼの5' -3' エクソヌクレアーゼ活性は、加水分解プローブを構成成分ヌクレオチドに分解する。これらの蛍光標識化された構成成分ヌクレオチドが解放されると、もはや互いに蛍光を抑制することはなく、遺伝子特異的なmRNAコピー数に比例する定量化可能な蛍光シグナルがRT-PCR反応に現れる。このシグナルは、遺伝子特異的な方法で、増幅されたcDNAの量に比例して増加する。

30

【 0 2 6 1 】

TaqMan測定出力は、各 サンプル及び各遺伝子について、アッセイの所定の蛍光シグナル (STV、シグナルスレッシュホールド値) が実現されるRT-PCRサイクル数として定義される C_t 値を報告する。全ての器具レベルでの C_t 値測定出力データは、以後、RWCT、すなわち「生 C_t 」と呼ばれる。RWCT値は、cDNAの開始量に反比例する (cDNAの開始量が少ないと、RT-PCRサイクルが長くなる、すなわち、STVに到達するまでの C_t が大きくなるためである)。

40

【 0 2 6 2 】

PCR増幅前の開始物質の量、すなわち、S (シグナル)、を決定する工程が、文献及び実務において確立された対数形式で表される以下の式 (対数は常に \log_{10} 、すなわち底が10の対数として定義される。) で実行される。 $\log S = \log STV - C_t * \log E$ ここで、S

50

は開始シグナル (initial signal) 又はサンプル中の遺伝子量 (アッセイにより imputed される)、STVは、Signal Threshold Value of 遺伝子増幅 for determining C_t (所定量)、及び E は、PCR増幅の効率 (Efficiency) (本質的に 2 と同程度となるように PBD によって確認される)。

【 0 2 6 3 】

「未決定 (UNDETERMINED)」という語は、TaqMan出力において 40 より大きい RWCT 値に割り当てられている。40 サイクルを超える場合のシグナルは、STV に到達するために非常に多くの増幅が必要であったということなので、信頼性がないと考えられるためである。標準的なリアルタイム RT-PCR では、増幅処理は 40 サイクルで終了する。

10

【 0 2 6 4 】

本明細書において用いられる分析について、20 より小さい RWCT 値は、非現実的に高い遺伝子特異的な開始量であることを示しており、信頼できるアッセイ出力とは考えられないので、「OUTLIERS」と記されている。

【 0 2 6 5 】

本明細書で実現しているように、RT-PCR データの 4 工程の前処理で RRCF 値が得られる。GVHD 転帰予測は、この値に基づく。RRCF は、「RT-PCR, Relative, logsignal, E=2, Clear, with Floor values replaced」として定義される。1. OUTLIER 値の置換:

RCTC (RT-PCR, C_t , Clean)

20

全ての OUTLIER 値は、データセット中の 180 サンプル全てについての遺伝子の RWCT 中央値で置換される。「Clean」は、中央値で置換された OUTLIER 値を意味する。2. PCR 効率 E = 2 について、RT-PCR シグナルの対数を生成

RL2C (RT-PCR, LogSignal, E=2, Clean)

RL2C 値は、以下の式に従って、RCTC 値に基づいて生成される。

$$RL2C = \log STV - RCTC * \log (2)$$

注記: $\log STV$ は 1.4 である。P2 ファイリングにおける GVHD 予測遺伝子リストに繋がるイルミナ社のマイクロアレイデータと比較できる数値範囲において RT-PCR 出力値を得ようとした結果、 $\log STV$ は、PBD によって、任意のユニットにおいて、1.4 に等しいと定義される (RT-PCR 実務においては特定のユニットやスケージングの推奨は存在しない)。

30

3. 「UNDETERMINED」フロア値 (Floor Value) の置換

RL2F (RT-PCR, Log Signal, E=2, clean, Floor Value replaced)

UNDETERMINED のフロア値は、以下の値で置き換えられる (RCTC が 40 についての RL2C 値を表す)。

RL2F (UNDETERMINED 値) = [14 - 40 * $\log (2)$] = 1.958800173440754 . ハウスキーピング (HSK) 遺伝子の相対平均シグナルを除去することによるバックグラウンドシグナルの訂正を通じた RT-PCR シグナルの相対化 (相対定量化)

RRCF (RT-PCR, Relative, log Signal, E=2, Clean, Floor Value replaced) バックグラウンドシグナルの訂正を通じた RT-PCR シグナルの相対定量化は、各サンプルの 6 つのハウスキーピング (RHSKAG6) 遺伝子の相対平均シグナルの除算により、以下の式に従って

40

$$RRCF = RL2F - RHSKAG6$$

注記: RHSKAG6 は、各サンプルについて、6 つの PBD 選択 HSK 遺伝子 (HSK6) の Relative HSK Average Signal として定義され、全 180 の RHSKAG6 値について 0 に中心がある。

$$RHSKAG6 = HSKAG6 - AVGHKAG6$$

AVGHKAG6 は、全 180 サンプルに特異的な HSKAG6 の決定の平均値を表す定数として定義される。HSKAG6 は、各サンプルについて、6 つのハウスキーピング遺伝子の RL2F 値の平均として定義される。

【 0 2 6 6 】

RRCF 決定に用いられる HSKAG6 値の決定のための例示的な 6 つのハウスキーピング遺伝子 (H

50

SK6) の定義表 1 2 は、HSKAG6及びRHSKAG6決定(上述した)に用いられるHSK6リストを説明する。RHSKAG6 = HSKAG6 - [average HSKAG6 over all 180 サンプル]

(表 1 2)

表 2 Bにおける総数192の遺伝子からQC基準を満たす175の遺伝子の選択データサンプルの55%以上の遺伝子で検出されなければならない基礎的なQC(品質制御)基準(すなわち、UNDETERMINEDがなく、RWCTが40以上ではない)の定義。表 1 3は、表 2 Bに記載された総数192の遺伝子のうち175の遺伝子(「SG175」)を記載している。

(表 1 3)

実施例 1 2

10

【0267】

この実施例は、GVHD予測に関するRGP(レシオメトリック遺伝子対)候補の評価の説明を含む。

【0268】

導入レシオメトリック遺伝子対(RGP)は、追加的な転帰予測の信頼性を、(1)バックグラウンド変位を分割することによる自己校正、及び(2)発現レベルにおける遺伝子間の優れた相互作用経路効果の捕捉、を通じて提供する。RGPは、選択単一遺伝子の発現レベルを他の選択単一遺伝子の発現レベルで割ることによって決定される。

【0269】

RGPの決定PBD技術の実行において、RRCFデータは、mRNA濃度測定レベル(上述)の対数形式で表現される、すなわち、 $RRCF_X$ は $\log(\text{gene } X)$ 及び $RL2F_X$ は $\log(\text{gene } X)$ なので、遺伝子Xの発現/遺伝子Yの発現の比は、対数形式では $\log(\text{遺伝子}X/\text{遺伝子}Y)$ であり、差の形式、すなわち $\log(\text{gene } X) - \log(\text{gene } Y)$ でも表される。これは、 $RRCF_X - RRCF_Y$ と同等である。したがって、以降は、RRCFデータのRGP値は、以下のように定義される。 $RGP = RRCF_X - RRCF_Y$

20

【0270】

HSKバックグラウンド除去はRGPの計算で実行されるため、RGP値は、HSK遺伝子のバックグラウンド除去前に、RL2Fデータから直接計算することもできる点に注意されたい。 $RRCF = RL2F - RHSKAG6$ (上述)であれば、 $RGP = (RL2F_X - RHSKAG6) - (RL2F_Y - RHSKAG6)$ となる。したがって、代替的に、 $RGP = RL2F_X - RL2F_Y$ と表される。

30

【0271】

全180サンプルについてのRGP値は、QC(上記)を満たした全175の選択されたSGs(単一遺伝子)のRRCFデータに基づいて、15,225の固有のRGPの完全なセットについて決定された。175のSGs(単一遺伝子)について、RGPの総数、すなわち固有ペアごとの(unique pair-wise)SG組み合わせは、 $(175^2 - 175) / 2 = 15,225$ と決定される。

【0272】

RGPの転帰予測性能の決定GVHD転帰予測性能を、以下のようにクラス分離(class separation)を決定することにより、15,225個のRGPの各々について評価した。(1)両側不均一T検定を使用したp値(2)LDA(線形判別分析)を使用した正確性(ACC) T検定p値及びLDA正確性の計算は、上記のとおり実行された。

40

【0273】

180のサンプルデータセットは、以下に示す4つのサンプルクラスを含む。(1)Gneg: 移植レシピエントにおいて、いかなる態様の急性又は慢性のGVHDも観察されなかった59のサンプル。(2)Gpos: 移植レシピエントにおいて任意の急性GVHD又は慢性GVHDが観察された121のサンプル(3)Gag2: 移植レシピエントにおいてグレードII、III又はIVの急性GVHDが観察された110のサンプル。慢性GVHDを伴う場合と伴わない場合がある。(4)Gag3: 移植レシピエントにおいてグレードIII又はIVの急性重傷GVHDが観察された77のサンプル。慢性GVHDを伴う場合と伴わない場合がある。

【0274】

50

グレードII、III又はIVの急性GVHDの有病率P（移植全体における発生）を仮定すると（Gag2）、一般に受け入れられている範囲は35%～55%であり、中点は45%である。110のGag2ケースは、合計で244（110/0.45）の移植における観察結果と推測される。概ね合計244の移植のうちの77のグレードIII又はIVの急性GVHD（Gag3）の割合は、0.315、すなわち31.5%に相当し、これは、グレードIII又はIVの急性GVHDの有病率範囲として一般に受け入れられている範囲である。定義に基づけば、合計77のGag3ケースは、110のGag2ケースの一部である。

【0275】

要するに、Gag2のケース内におけるGag3ケースの割合は、Gag2の35%～55%の範囲の有病率及びGag3の15%～35%の範囲の有病率と概ね調和している。したがって、Gag2及びGag3転帰についてのGVHDの減少予想は、ドナーを分析によってGVHDを起こさないと予測されたドナーのみに限定するGVHD転帰予測を用いる場合、一般に許容されている範囲のGag2有病率及びGag3有病率を示すGag2サンプルとGag3サンプルとのバランスの取れた比率を用いてトレーニングされた予測モデルに基づくものとなる。

10

【0276】

GVHD転帰予測性能（T検定及びLDA）を、以下のクラス分割（class division）について決定した。（1）Gneg対pos

（2）Gneg対Gag2

（3）Gneg対Gag3

【0277】

全ての3つの異なるクラス分割の各々に関する全てのLDAの計算について、Gneg対Gpos分割に基づくLDA区分線を用いた。この区分線は、59のGnegサンプルの平均RGP値と121のGposサンプルの平均RGP値との中点として決定された。

20

【0278】

正確性は、陰性に分類されたサンプルと及び陽性に分類されたサンプルとの「均衡（balanced）」比例表現（proportional representation）を用い、GVHD陽性ケースの均衡（balanced）有病率を $P_b = 0.5$ （50%）として決定された。4つの混同行列分類値（混同行列 classification value、CMCV）、TN、FP、TP、FNの全ては、合計1の割合として表される点に留意されたい。すなわち、4つの値の合計が必ず1にならなければならない。均衡CMCV（Balanced CMCV、下付文字「b」で表される）は、開始（initial）CMCV（下付文字「0」で示される）から、開始有病率 P_0 に基づいて、以下の式に従って決定される。（1）

30

$$TN_b = (1 - P_b) / (1 - P_0) * TN_0$$

$$(2) \quad FP_b = (1 - P_b) / (1 - P_0) * FP_0$$

$$(3) \quad TP_b = P_b / P_0 * TP_0$$

$$(4) \quad FN_b = P_b / P_0 * FN_0$$

正確性（ACC）は、均衡有病率 P_b に関して調整され、以下のように定義される（P1及びP2も参照）。

$$ACC = (TN_b + TP_b) / (TN_b + FN_b + TP_b + FP_b)$$

【0279】

RGPの転帰予測性能の評価クラス識別分析は、6つの性能変数、すなわち3つの分割の各々についてのp値及び正確性の出力を、全15,225 RGP（表RGP348参照）について提供する。この6つの転帰予測性能変数は、以下に基づいて、1～15,225までそれぞれ順位付けされる。（1）3つの分割の各々について、最も優れた（最小の）p値から最も劣った（最大の）p値に応じて

40

（2）3つの分割の各々について、最も優れた（最大の）転帰予測LDA正確性から最も劣った（最小の）転帰予測LDA正確性に応じて

【0280】

GVHD転帰予測プロファイルの高性能化のための候補RGPの最初の絞り込みとして、6つの予測性能変数の全てが、（1～15,225の中で）最大範囲で2000以上であり最小範囲でのランクで200以上であることを要求することにより、348のRGP（RGP348、表14

50

参照)を選択した。換言すれば、全ての6つの転帰予測性能変数がトップ2000(13.1%)であり、6つの転帰予測性能変数のうち少なくとも1つがトップ200であること(1.3%)が要求される。RGP348リスト(表14)において、175のSGのうちの128がare represented、ranging from participation of各SG in 1 to 53 differentRGPs(表15参照、SG128)。表14は、高い転帰予測性能ランクを有する348RGPs(RGP348)のリストを示す。

(表14)

表15は、348RGP内で高い転帰予測性能ランクを有する175のSGのリストを示す。

(348RGPのリスト内で、175のSGのうちの128が示されている。)

10

(表15)

【0281】

PRGPの決定(RGP対)GVHD転帰予測プロファイルの高性能化のための候補RGPの別の絞り込みとして、RGPは(セミアディティブ)複数RGP投票モデルにおいてうまく機能し、追加RGP対(PRGP、以下で定義する)モデルでもうまく機能すると予測された。PRGPでうまく働くRGP事前に選択することによって、複数RGP投票モデルは、強化された傾向を有する候補RGPから選択できるようになり、相乗的に複数RGPシナリオにおけるGVHD転帰予測の改善の方向に作用する。

【0282】

競合的な相互作用(上述)は、数学的に競合的で、レシオメトリック関係(x/y 、又は対数表記で $\log(x/y)$ 、又はこれと同等の $\log x - \log y$)にあると表されていたので、相乗的な相互作用は、数学的に相乗的で、乗法の関係($x*y$ 、又は対数表現では $\log(x*y)$ 、又はこれと同等の $\log x + \log y$)と表される。本明細書において説明される技術的な実現においては、RRCF及びRGP値は、mRNA濃度測定レベルの対数形式で表現されているので、このような相乗的な相互作用は、RGP値に対して追加的な形式で表される。

20

【0283】

RGP/RRCFデータについてのPRGP値は以下のとおり定義される。 $PRGP = RGP X + RGP Y$ 。

【0284】

SG測定するとき、RRCFレベルで(HSKシグナル減算によって調整補正されたRL2F)、PRGPは定義される。 $RGP X = RRCF A - RRCF B$

30

$RGP Y = RRCF C - RRCF D$

$PRGP = (RRCF A - RRCF B) + (RRCF C - RRCF D)$

【0285】

代替的に、SG測定ときに、RL2Fレベルでは、PRGPは以下のように定義される。 $RGP X = RL2F A - RL2F B$

$RGP Y = RL2F C - RL2F D$

$PRGP = (RL2F A - RL2F B) + (RL2F C - RL2F D)$

【0286】

PRGPに決定に関して、175のRGPを、RGP348リスト(表14)から、以下の基準で選択した。(1) 両方の遺伝子メンバーが最小範囲の性能ランクが100を示したRGPのみを考慮する。すなわち最小範囲のランクが100以下でフィルターする。

40

(2) 中央値ランクに応じてフィルターされたRGPに優先順位をつけ、最も優れた175の中央値ランクのRGPを選択する。

【0287】

PRGP決定のために選択された175のRGPの全てが、最小範囲のランクが100以下で、中央値ランクが464以下で、最大範囲のランクが1380以下となる。

【0288】

得られた選択175RGPに基づいて、PRGPの完全なセットについて、全ての180サンプルについてのPRGP値を決定した(RGPについて上述したのと同様)。RGPについて説明したのと同

50

様に、G V H D 転帰予測性能及びランクを15,225のPRGPについて評価した。

【0289】

PRGPの転帰予測性能の評価348のPRGPのセット（PRGP348、表16には特定のPRGPを記載している）を、以下のように、15,225のPRGPの完全なセットから選択した。(1) 各PRGPが全ての6つの予測性能変数ランクについて、(1~15,225の中から)最大範囲のランクで5000以下で最小範囲のランクで500以下を有することを要求する。その結果、890のPRGPが得られた。

(2) 890のPRGPのセット内で最も優れた(best ranking)348の中央値ランクのPRGPを優先する。

【0290】

性能値及びランキングを、RGPに関して上述したのと同様に、PRGPについて決定した。

表16は、348 PRGPのリストを示す（PRGP348）。

(表16)

実施例13

【0291】

この実施例は、G V H D 予想についての代替的なRGP Vmod（投票モデル）の実装、最も優れたRGP Vmodの定義、及び他の優れた代替的なRGP Vmodの説明を含む。

【0292】

以上で選択されたRGPのレシオメトリックG V H D 転帰 予想と自己校正特徴の組み合わせを利用することに加えて、個別のRGP投票によってもたらされる誤差を平均することによって、複数RGP 投票モデル(Vmods)を通じて、G V H D 転帰予測におけるさらなる正確性と信頼性の実現が期待される。このようなG V H D 転帰予測Vmodにおいて、RGPの優先的なサブセットが、個別の「N」（Nは、レシピエントにG V H Dを引き起こさない）転帰予想投票を提供するために用いられ、これらの投票は、G V H D N 転帰スコア、又はGNOSとして集計され平均化される。その後、GNOSが所定の「GNOS閾値」レベル以上のときに、ドナーサンプルが最終的にN、又は「移植で用いられるとレシピエントにG V H D 陰性転帰を引き起こす可能性が高い」とされる。

【0293】

代替的なRGP Vmodの選択RGP（又は間接的なRGP選択についてはPRGP）をGNOS決定のためのVmodに集約するための複数の理にかなった方法が適用され、6のHSK（表12）を含む総計48のSGを用いて、最初のSG校正のために、G V H D 転帰予測を得る。異なるVmodに寄与するRGP及びSGのリストは、以下で詳細に説明される（表17及び18参照、それぞれVmodRGP100及びVmodSG64）。

【0294】

全Vmodsについて、Vmod GNOSの計算は、常に、RGP値に対して直接実行され、PRGPやSG値には直接実行されない点に留意されたい（たとえ、PRGPがRGPの選択に寄与するとしても、SG値がRGP値の決定に利用されるとしても同じである）。

【0295】

3つの基本的な方法の概要が以下に示される。RGP348リスト（表14）から3つの代替的なVmodsへのRGPの優先順位付け。(1) Vmod: SG43RGP46-GPperformance: RGP348リスト（表14）中のRGPが性能ランクの中央値のみに応じて優先順位付けされる。寄与SGが複数のGPのメンバーであることの制約はない。最高の中央値性能ランクの46のRGPは、43の固有のSGを有し、HSK6SGs（表12）の1つを含んでいる。これらの43のSGを残りの5つのHSK6SGと組み合わせることにより、G V H D 転帰予測試験として実施するために、合計で48のSGが得られる。

(2) Vmod: SG42RGP21-GPminimalist:

RGP348リスト（表14）中のRGPが、中央値性能ランクに応じて優先順位付けされる。この最も上位ランクのRGPがVmodのために選択されたのち、Vmodのために既に選択されているSGも含む全てのRGPが候補リストから除外され、次の最も上位ランクのRGPが選択される

10

20

30

40

50

。上位21のランクのRGPについては、いかなるSGも1つ以上のRGP内に現れることが許されず、合計で42の固有のSGが得られる。これらの42のSGを残りの6つのHSK6SGと組み合わせることにより、GVHD転帰予測試験として実施するために、合計48のSGが得られる。

(3) Vmod:SG43RGP37-GPconnectivity:

RGP348リスト(表14)に寄与する121のSGが2つの工程で優先順位付けされる。すなわち、

1. 特定のSGがメンバーであるRGPの最も高い数、すなわち、the highest SG connectivities in the RGPネットワーク内で最も高いSG連結性(SGconnectivity)

2. 最もよい中央値転帰予測性能ランク(6つの標準性能ランクについて、上記参照)。

この方法に従って、43の組み合わせられた上位のSGを選択した。RGPには、43のSGのリストから唯一のSGメンバーを含み、中央値性能ランクに応じて優先順位付けをした。43のSGの各々が少なくとも1度現れるという限定された選択に基づく最小限の数の上位ランクRGPは、次に、中央値性能ランク順位に優先順位付けされたRGPを集計することによって選択された。しかしながら、集計処理においては、新しいRGPは、Vmodとして既に選択されたRGP内にSGメンバーが存在しないかただ1つ見つかる場合に、含めることがゆるされた。(注記:この制限がなければ、100以上のGPが43の最も結合されている(connected)SGをカバーするために必要となる。これは、最も結合されているSGの多くは、低ランクのRGPに参加しているためである。)合計で37のRGPがこのVmodに寄与する。43の固有のSGを含み、HSK6SGsのうちの1つを含んでいる。これらの43のSGを残りの5つのHSK6SGと結合することにより、GVHD転帰予測試験として実施するために、合計で48のSGが得られる。

【0296】

PRGPでうまくいくRGPを事前選択することにより、Vmodは、強化された傾向を有する候補RGPから選択できるようになり、相乗的に複数RGPシナリオにおけるGVHD転帰予測の改善の方向に作用する。2つの基本的な方法がPRGP(黙示的(implicit)寄与RGP及びSG)について以下にまとめられる。PRGP348リスト(表16)から2つの代替的なVmodへの優先順位付け。

(1) Vmod:SG43RGP51-PRGPminranksort:

PRGP348リスト(表16)中のPRGPを、最大範囲の(maximal)(最悪の)性能ランクによって優先順位付けし、続いて中央値ランクによって優先順位付けした。次に最小範囲の(minimal)(最高の)ランクで優先順位付けし、最後の優先順位付け基準が最もよいもの、すなわち、最小範囲の(minimal)性能ランク、となるようにした。SG又はRGPが複数のPRGPのメンバーとなることに制約は課さなかった。最高の中央値性能ランクの45のPRGPは、51の固有のRGP及び43の固有のSGを含み、これにはHSK6SGの1つを含む。これらの43のSGを残りの5つのHSK6SGを組み合わせることにより、GVHD転帰予測試験として実施するために、合計で48のSGが得られる。

(2) Vmod:SG43RGP55-PRGPmedranksort:

PRGP348リスト(表16)中のPRGPを最初に最大範囲の(maximal)(最悪の)性能ランクによって優先順位付けし、次に、最小範囲の(minimal)(最高の)ランクで優先順位付けし、続いて、中央値ランクによって優先順位付けした。これにより、最後の優先順位付け基準が、中央値性能ランクとなるようにした。最高の中央値性能ランクの60のPRGPは、55の固有のRGP及び43の固有のSGを含み、これにはHSK6SGの1つを含む。これらの43のSGを残りの5つのHSK6SGを組み合わせることにより、GVHD転帰予測試験として実施するために、合計で48のSGが得られる。

【0297】

上述したVmodの5つのセットは、合計で100の固有RGP及び64の固有SGを有しており、これは、それぞれ、表17及び表18に記載されている、VmodRGP100及びVmodSG64である。表17は、代替的なVmodGVHD転帰予測の実施に参加したVmodメンバーシップの100のRGP(「VmodRGP100」)(列中の「x」は、対応する行のRGPが当該Vmodのメンバーであることを示す。他方、「-」は、RGPがVmodの構成要素でないことを示す。)を示す。

(表17)

表 1 8 は、代替的なVmod G V H D 転帰予測の実施に参加したVmodメンバーシップの64 SG (「VmodSG64」)を示す(列中の「x」は、対応する行のRGPが当該Vmodのメンバーであることを示す。他方、「-」は、RGPがVmodの構成要素でないことを示す。)
(表 1 8)

【 0 2 9 8 】

一般に、各Vmod、最終的な G V H D N 転帰スコア (GNOS) は、RGP値を元に、各サンプルについて、以下のように計算される(一般投票モデル (Vmod) の実施については上述した)。各RGPについて、その値が所定の閾値 (LDA区分線、G V H D 陰性平均RGP値と陽性平均RGP値との間の midpoint) として定義される。) よりも上又は下にある場合に、当該サンプルは、N (陰性)、すなわち、移植において G V H D を引き起こさない、と分類される。N 票は「1」とカウントされ、それ以外は「0」とカウントされる。特定の投票モデルにおける RGPの全てについての平均は、GNOSに到達する。

10

【 0 2 9 9 】

RGP Vmodsの最長マッチ (greedy) 最適化実験によって実証可能な、データに由来する証拠の複数のソースが、48のSGの選択とは異なる G V H D 転帰予測の実施のための 転帰予測の高い性能を持つRGP候補として、RGPの変数を組み込んだVmodを通じて、。100のRGP (表 1 7、VmodRGP100) の選択に貢献した。これらの証拠のソースは、個別のRGP performance、PRGP performance中の構成要素RGP、及び代替的なRGP選択基準とともにデザインされたVmodsにおける統合されたRGP性能を含む。

20

【 0 3 0 0 】

このように、G V H D 転帰予測Vmodの設計において適用されるRGPの最適化 (上述したRGP優先順位付けの代わり) 手順によって、優れたVmod性能が得られる可能性がある。したがって、以下で述べる、単純な「最長マッチ (greedy)」サーチ/集計手順を、以下で述べるように、Vmod最適化に適用した。(1) Vmod:SG43RGP36-RGPgreedysearch.

最長マッチ探索は、表 1 7 (VmodRGP100リスト) から最も優れたRGPを選択することで開始する。これは、Vmod(1)を定義する。100マイナス1の代替的Vmodが、Vmod(1)のRGPを残りの100マイナス1のRGPと組み合わせることによって評価され、転帰予測性能がこれらの100マイナス1のVmodの各々について決定される。100マイナス1のVmodのリストから、最も優れたVmodがVmod(2)として選択される。次に、100マイナス2の代替的なVmodが、Vmod(2)中のRGPを残りの100マイナス2のRGPと組み合わせることにより評価され、転帰予測性能がこれらの100マイナス2のVmodの各々について決定される。この100マイナス2のVmodから、最も優れたVmodがVmod(3)として選択される。一般化すれば、Vmod(i)中のRGPを残りの100マイナスiのRGPと組み合わせることによって、100マイナスiの代替的なVmodが評価され、転帰予測性能がこの100マイナスiのVmodの各々について決定される。この100マイナスiのVmodのリストから、最も優れたVmodがVmod(i+1)として選択される。インデックスiはその後1ずつ、i=101まで増加され、このとき探索は終了し、全てのRGPが集計される。

30

Vmod Gneg対 Gag3の T検定 p値は、各サンプルのGNOS から決定される。全てのVmodは、GNOSの閾値0.55で評価される。すなわち、各Vmodにおいて、構成要素RGP票の少なくとも55%が票1、すなわち、所定のサンプルについてN票 (noGVHD) を投じなければならない。均衡CMCVは、Gnegサンプル及びGag3サンプルの各々について、Vmod候補によって決定される1(N)値及び0(N以外)値に基づいて決定される。

40

各繰り返しにおいて、iのVmod候補から最も優れた (best ranking) Vmodが、以下の基準に基づいて選択される。

1. Gneg対 Gag3の分割について、TNR (真陰性率、特異性) が、考慮されるVmodについて0.4以上でなければならない。
2. Gneg対 Gag3分割について、NPV (クラス番号ごとに均衡化された均衡陰性予測値、すなわち $P_b=0.5$) の最もよい平均ランク、ランク、及びp値ランクに応じて、最も優れた (best ranking) Vmodが選択される。

最長マッチ探索の36回の繰り返しの後、合計で36のRGPが選択された。これには、43の固

50

有SGを含み、この中にはHSK6 SG (表 1 9 参照、VmodGreedySearch) の 1 つが含まれている。これらの43のSGを残し5つのHSK6 SGと組み合わせることにより、GVHD転帰予測試験として実施するために、合計で48のSGが得られる。

(2) Vmod:SG21RGP28-RGPmaxgreedysearch.

上述した 最長マッチ探索における最も優れているVmodは、Gneg対 Gag3については、NPV (0.96)、ACC (0.90) 及び p 値 (2.04×10^{-23}) と組み合わせられて、RGP選択の 2 1 回の繰り返しにおいて発生する、すなわち、28の異なるRGP (SG21RGP28、表 1 9 参照、VmodGreedySearch) に参加する21の固有SGを代表することに留意されたい。このVmodは、サポート性能として、GVHD転帰予測試験として実施され得るが、SG21RGP28-RGPmaxgreedysearchが、半分より少ないSGを用いてSG43RGP36-RGPgreedysearchと比較されることで、通常

10

のGVHD転帰予測実務におけるSG43RGP36-RGPgreedysearchと同程度に一貫し又は信頼できるように動作するのかわかりしない。

表 1 9 は、Vmods SG43RGP36及びSG21RGP28についての最長一致 (Greedy) サーチ性能を示す (「VmodGreedySearch」)。

(表 1 9)

【 0 3 0 1 】

RGP Vmodの転帰予測性能の決定及び評価 7 つのVmodの全てについての全体的なGVHD転帰予測性能が表 2 0 (VmodSpecs) にまとめられている。全ての 3 つの分割について (Gneg対Gpos、Gneg対Gag2、Gneg対Gag3)、転帰予測性能についての以下の標準的な明細 (specification) (SSPCs) が報告されている。(1) 各サンプルについて、Vmod出力において報告されたGNOS値に基づく不均一両側 T 検定 p 値。

20

(2) 5 つの異なるGNOS分離区分線 (それ以上でドナーサンプルが最終的に N に分類される、すなわちレシピエントにおいてGVHDを引き起こさないGNOS閾値値) である0.50、0.55、0.65、0.75及び0.85について、5 つの追加的な転帰予測SSPC (例えば、上記のCMCVの定義を参照)。

1. NPV、negativeprediction value、均衡型、すなわち、実 N 転帰 (TN + FP) 及び実 P 転帰 (TP +FN) の均等な割合に調節される、すなわち、均衡有病率 $P_b=0.5$ について $NPV = TN_b / (TN_b + FN_b)$

2. TNR、truenegative rate、特異性 (実 N 転帰及び実 P 転帰の割合に影響されない)

30

$TNR = TN / (TN +FP)$

3. PPV、positiveprediction value、均衡型、すなわち、実 N 転帰 (TN +FP) 及び実 P 転帰 (TP +FN) の均等な割合に調節される、すなわち、均衡有病率 $P_b=0.5$ について $PPV = TP_b / (TP_b + FP_b)$

4. TPR、truepositive rate、感度 (実 N 転帰及び実 P 転帰の割合によって影響されない)

$TPR = TP / (TP +FN)$

5. ACC、accuracy、均衡型、すなわち、すなわち、実 N 転帰 (TN +FP) 及び実 P 転帰 (TP +FN) の均等な割合に調節される、すなわち、均衡有病率 $P_b=0.5$ について

$ACC = (TN_b + TP_b) / (TN_b +FN_b + TP_b + FP_b)$

40

【 0 3 0 2 】

以下の考察は表 2 0 (VmodSpecs) に関する。(1) Gneg対Gag3分割について、0.75又は0.85のGNOS区分線でのNPVは、しばしば0.90以上となるが、このときしばしばTNR値が0.25以下となる。

(2) Gneg対Gag3分割について、VmodSG43RGP36-RGPgreedysearchは、GNOS閾値0.55において、42又は43のSGを用いる任意の分割でのVmodの最も高い正確性 (0.87) に到達し、NPV値0.92をTNR (特異性) 値0.80及びTPR (感度) 値0.94と組み合わせる。

(3) 全ての 3 つの分割について、及び、42又は43のSGを有する全てのVmodについて、VmodSG43RGP36-RGPgreedysearchは、抜群に最も低い (最適) T 検定 p 値を示し、Gneg対Gpos分割についての 1.1×10^{-18} からGneg対Gag3分割についての 3.3×10^{-19} まで広がっている。

50

(4) SG21RGP28-RGPmaxgreedysearchは、全てのカテゴリで抜群に最も優れたSSPCを示す。このVmodはサポート性能と共にG V H D転帰予測試験として実行され得るが、SG21RGP28-RGPmaxgreedysearchは、SG43RGP36-RGPgreedysearchと比較して半分より少ないSGを用いており、通常のG V H D転帰予測実務におけるSG43RGP36-RGPgreedysearchと同程度には一貫して又は安定して機能できない可能性がある。

表20は、選択された7つのVmod(「VmodSpecs」)の有病率、 $P=0.5$ における全体的な転帰予測性能を示す

(表20)

【0303】

観察され及び均衡化された(実N転帰及び実P転帰の均等な割合に調節される)、絶対又は相対の(調整された値について)「混同行列」は、5つの異なるGNOS区分線(0.50、0.55、0.65、0.75及び0.85)について、正確に及び不正確に分類されたサンプル(TN、FP、TP、FN; bTN、bFP、bTP、bFN)をカウントする。これにより、5つの転帰予測正確性/割合(NPV、TNR、PPV、TPR、ACC)を計算したので、表21(VmodCounts)で報告する。表21は、選択された7つのVmodsについて、正確な及び不正確なサンプル分類のカウント及び均衡化された相対カウント(“VmodCounts”)を示す。

(表21)

【0304】

G V H D転帰予測を用いる場合の予測される臨床的メリットG V H D転帰の減少による予測される臨床実務及び患者死亡率及び死亡数のメリットは、予想されるG V H D・Nドナーキャプチャー(G V H D・Nドナーは、移植においてG V H Dを引き起こさない実際のドナーと決定される。すなわち真陰性と偽陽性の合計)に不随されるが、グレードII、III又はIVの急性G V H D(35%~55%)及びグレードIII又はIVの急性G V H D(15%~35%)の現実的な有病率推定値について、異なるGNOS区分線について、表22及び23(VmodMedGainGag3及びVmodMedGainGag2)にまとめられている。G V H Dの減少予測は、レシピエントにG V H Dを引き起こさないと予想されるドナーのみが移植において使用されるという前提に基づいている。換言すれば、差し迫った「N」ドナー選択実務を仮定すれば、グレードII、III、又はIVの急性G V H Dが依然として起こりえるのは、残っている偽陰性予測によるものである。

【0305】

表22及び23(VmodMedGainGag3及びVmodMedGainGag2)において、様々な代替の有病率をカバーする予測が、選択された6つのVmodについて、直接に及び完全に、表21(VmodCounts)に記載された値、すなわち、サンプルの分類カウント及び均衡CMCV(混同行列分類値)から誘導されている。変動させたGNOS区分線も考慮に入れられている。また、これらの表は、また、それぞれの代替的な有病率 P_a 及びGNOS区分線において、NPV、TNR(特異性)及びTPR(感度)のSSPCを報告している。

【0306】

代替的な(下付記号「a」で示されている)有病率 P_a について選択SSCVを決定するため、まず4つのCMCVが以下に示すように調節される必要がある。(1) $TN_a = (1 - P_a) / (1 - P_0) * TN_0$

$$(2) FP_a = (1 - P_a) / (1 - P_0) * FP_0$$

$$(3) TP_a = P_a / P_0 * TP_0$$

$$(4) FN_a = P_a / P_0 * FN_0$$

【0307】

表21「VmodCounts」に記載された均衡CMCVを変換するために、 $P_0 = P_b = 0.50$ である。

【0308】

P_a について調節されたCMCVについて、以下の5つのSSPCが以下のように決定される。(1) NPV、negative prediction value、代替的な有病率 P_a に関して調整される。

10

20

30

40

50

$$NPV = TN_a / (TN_a + FN_a)$$

(2) TNR、true negative rate、特異性（有病率に影響されない）

$$TNR = TN / (TN + FP)$$

(3) PPV、positive prediction value、代替的な有病率 P_a に関して調整される。

$$PPV = TP_a / (TP_a + FP_a)$$

(4) TPR、true positive rate、すなわち、感度（有病率に影響されない）

$$TPR = TP / (TP + FN)$$

(5) ACC、accuracy、代替的な有病率 P_a に関して調整される。

$$ACC = (TN_a + TP_a) / (TN_a + FN_a + TP_a + FP_a)$$

【0309】

これらの表で報告されているGVHD減少値は、それぞれの陰性予想値（NPV）及び代替的な有病率（ P_a ）から、以下の式にしたがって計算されている。GVHD減少 = $1 - (1 - NPV) / P_a$ 。

【0310】

NPVが1のとき、すなわち陰性の分類が100%正しいとき、GVHDの減少は1、すなわち100%となる。NPVが $[1 - P_a]$ と1の間にある場合、GVHDの減少範囲は0~1、すなわち0%~100%となる。NPVが0と $[1 - P_a]$ の間にある場合には、GVHDの減少範囲は、 $[1 - 1 / P_a]$ （GVHDの減少の下限、 $P_a < 1$ のときに負となる）から0である。NPVが $[1 - P_a]$ よりも小さいとき、対追いする陰性GVHDの減少は、実際にはGVHDの増加を意味する点に留意されたい。したがって、GVHD転帰予測試験がGVHDの減少に効果があるというためには、NPVが $[1 - P_a]$ よりも大きいことが常に要求される。

【0311】

上記の表で報告されているGVHD・Nドナーキャプチャ値は、TNR値と同じであるが、百分率で報告されている。この値は、GVHD転帰予測試験で捕捉される「実際に」利用できる陰性ドナーの率を強調している。

【0312】

以下の考察は、最も低いGag3 有病率15%における、表22（VmodMedGainGag3）に関するものである。(1) GNOS区分線0.75又は0.85での予測されるGVHDの減少は、しばしば5%以上に達するが、しばしばGVHD・Nドナーキャプチャーが25%以下になる。(2) 42又は43のSGを有するVmod、VmodSG43RGP36-RGPgreedysearchは、GNOS閾値0.55で、最も結合して91%のGVHDの減少、80%のGVHD・Nドナーキャプチャとなる。(3) VmodSG42RGP21-RGPminimalistについて、GNOS閾値0.85では、100%のGVHD減少が予想されるが、GVHD・Nドナーキャプチャは25%である。また、将来的により多くのサンプルが、今までの調査よりも複雑な母集団を利用してより挙動を分散させて試験されることにより、100%という不自然に高い値は90%台のどこかに落ち着くであろう。(4) SG21RGP28-RGPmaxgreedysearchは、他の全てのVmodと比べて、最も高いGVHDの減少（95%）及びGVHD・Nドナーキャプチャー（83%）。このVmodはサポート性能と共にGVHD転帰予測試験として実行され得るが、SG21RGP28-RGPmaxgreedysearchは、SG43RGP36-RGPgreedysearchと比較して半分より少ないSGを用いており、通常のGVHD転帰予測実務におけるSG43RGP36-RGPgreedysearchと同程度には一貫して又は安定して機能できない可能性がある。

【0313】

以下の考察は、最も低いGag2 有病率35%を有する表23（VmodMedGainGag2）に関するものである。(1) 42又は43のSGを有するVmodについて、VmodSG43RGP36-RGPgreedysearchは、GNOS閾値0.55において、最も高い結合の73%GVHDの減少及び80%GVHD・Nドナーキャプチャに到達する。(2) VmodSG42RGP21-RGPminimalistは、GNOS閾値0.85で、84%のGVHDの減少を示すが、GVHD・Nドナーキャプチャは25%を示す。(3) SG21RGP28-RGPmaxgreedysearchは、他のVmodと比較して、最も高いGVHDの減少（78%）及びGVHD・Nドナーキャプチャー（83%）を示す。

表 2 2 は、有病率が15%~35%の場合における、GVHD 転帰の減少及びグレードIII又はIVの急性GVHDのGVHD・Nドナーキャプチャー（「VmodMedGainGag3」）において予測されるメリットを示す。

（表 2 2）

表 2 3 は、有病率が35%~55%の場合におけるGVHD 転帰の減少及びグレードII、III又はIVの急性GVHD（「VmodMedGainGag2」）において予測されるメリットを示す。

（表 2 3）

10

—

【 0 3 1 4 】

VmodSG43RGP36-RGPgreedysearchについての転帰予測性能決定の詳細表 2 4 及び 2 5（SG43RGP36exampleGneg及びSG43RGP36exampleGag3）は、Vmod SG43RGP36-RGPgreedysearch、Gneg及びGag3サンプル及び36のRGPの全てについて、GVHD・N 転帰に関する個別のRGP投票、各サンプルについてのGNOS値、及び0.55以上のGNOSについての最終のGVHD・N 転帰予測を示す。

【 0 3 1 5 】

表 2 4（SG43RGP36exampleGneg）に示されるように、59のGnegサンプルのうち計47個は、正しく、すなわち真陰性（TN）に分類されている。このように、特異性又は真陰性率 $TN / (TN + FP)$ は、0.80（0.7966）である。表 2 5（SG43RGP36exampleGag3）に示されているように、77のGag3サンプルのうち計72個は、正しく、すなわちTPに分類されている。このように、感度又は真陽性率、 $TP / (TP + FN)$ 、は0.94（0.9350）である。

20

【 0 3 1 6 】

Gag3（一般に許容されている推定値15%~35%の midpoint）について有病率が25%とし、(1) 特異性（TNR）が上にある場合、TNの全体の割合は59.8%（0.7966 * 75%）であり、FPの割合は15.2%（75% - 59.8%）であり、

(2) 感度（TPR）が上にある場合、TPの全体の割合は23.4%（0.9350 * 25%）であり、FNの割合は、1.6%（25% - 23.4%）であり、

30

(3) したがって、陰性予測値、 $TN / (TN + FN)$ は、0.97（0.974、すなわち、 $0.598 / (0.598 + 0.016)$ ）である。

【 0 3 1 7 】

GVHD 転帰予測試験を用いた結果、GVHD・N分類ドナーが移植に使用された場合にのみ、移植の97%では、グレードIII又はIVの急性GVHDが起こらない。予測分析を用いない場合は75%である。したがって、GVHD 転帰予測試験を用いなければ、移植の25%がグレードIII又はIVの急性GVHDを引き起こしたと考えられるが、GVHD・Nドナーを本試験を用いて選択すれば3%となる。換言すれば、GVHD 転帰予測試験を用いた後でもグレードIII又はIVの急性GVHD 転帰の12%は起こりえるが、but usage of the test for GVHD・Nドナー選定に本テストをしようすると、GVHDを89%も減らすことが期待できる（表 2 2 参照、VmodMedGainGag3、概要及び詳細）。

40

【 0 3 1 8 】

表 2 4 及び 2 5（SG43RGP36exampleGneg及びSG43RGP36exampleGag3）の両方において、移植のサンプルは、以下を用いた移植からのものである。(1) 骨髄（BM）及び末梢血幹細胞（PBSC）ソースが表されており、

(2) HLA 9 / 10 及び HLA 10 / 10 適合ドナーレシピエント対が表されている。

【 0 3 1 9 】

BM、PBSC、HLA9/10及びHLA 10/10サンプルは、目視によれば、GVHD・N（陰性）に分類されたか否かによらず、また、正しく分類されたか否かによらず、全てのサンプルにわたって本質的に均一に分散している。換言すれば、大多数のケースにおいて、independe

50

ntly of whether the 幹細胞がBMであるかPBSCであるかにかかわらず、また、移植がHLA 9/10適合又はHLA10/10適合のドナー・レシピエント対であるか否かによらず、GVHD転帰予測試験はGVHD・N（陰性）ドナーを正しく予測している。表24は、0.55以上のGNOSについて、VmodSG43RGP36-RGPgreedysearchの59のGnegサンプルについて、個別のRGP投票、得られたGNOS値、及び最終のGVHD・N転帰予測を示している（「SG43RGP36exampleGneg」）。

（表24）

表25は、0.55以上のGNOSについて、VmodSG43RGP36-RGPgreedysearchの77のGag3サンプルについて、個別のRGP投票、得られたGNOS値、及び最終のGVHD・N転帰予測を示している（「SG43RGP36exampleGag3」）

10

（表25）

実施例14

【0320】

この実施例は、GVHD転帰予測について、SG（単一遺伝子）Vmod性能よりも改善されたRGP Vmod性能の説明を含む。

【0321】

レシオメトリック遺伝子対（RGP）は、追加的な転帰予測の信頼性を、（1）バックグラウンド変位を分割することによる自己校正、及び（2）発現レベルにおける遺伝子間の優れた相互作用経路効果の捕捉、を通じて提供する。したがって、GVHD転帰予測遺伝子セットの性能を評価する際には、RGP投票モデルを実行することにより、単純なSG投票モデルを実施する場合よりも優れた性能を示すことが着たいされる。

20

【0322】

SG43RGP36-RGPgreedysearchについてSGとRGPとの比較例えば、表26を参照すると、Gneg対Gag3転帰予測性能についてのSG43RGP36compSGRGP、上述した最も優れた48の遺伝子のGVHD転帰予測の実行、SG43RGP36-RGPgreedysearchは、（1）単純な43のSG Vmod構成においては、43の予測遺伝子をGNOS値に対する個別の投票者及び貢献者として使用し、（2）優れた36のRGP Vmod構成では、36の異なる対ごとに、RGPの組み合わせをGNOS値に対する個別の投票者及び貢献者として使用している（配列の詳細及びRGPVmod性能の詳細については上述した）。

30

【0323】

36のRGP Vmodの利用（表26、SG43RGP36compSGRGP）は、43のSG Vmodの利用を全ての性能カテゴリにおいて上回っている点に留意されたい。T検定から説明すると、36のRGPのp値は10のオーダーの大きさで43のSGのp値よりも低い（良い）。GVHDの減少及びGVHD・Nドナーキャプチャは、5つの異なるGNOS閾値レベルにおいて、10%~20%、36のRGPモデルの方が43のSGの利用よりも優れている。表26は、Gneg対Gag3分割について、有病率P=0.25において、SG43RGP36-RGPgreedysearchのSG43投票モデルとRGP36投票モデルとの比較を示す（「SG43RGP36compSGRGP」）。

（表26）

40

実施例15

【0324】

この実施例は、Vmodsが厳格な最新のブートストラップ相互検証を受ける場合に、GVHD転帰予測について、信頼性のある統計RGP Vmod性能についての説明を含む。 [0104]

ブートストラップ相互検証は、レシオメトリック遺伝子対投票モデル（RGP Vmods）の転帰予測性能を評価するための計算集約的アプローチとして用いられた。ブートストラップ相互検証は、従来の相互検証法よりも技術的により高度で、モデル性能推定において信頼性が高い（Bradley Efron & Robert J. Tibshirani, An Introduction to the ブートストラップ, Chapman & Hall / CRC, Boca Raton, Florida, 1998, esp. pp, 247-255

50

; A. C. Davison & D. V. Hinkley, *ブートストラップMethods and TheirApplications*, Cambridge University Press, Cambridge, UK, 1997, esp. pp. 292-298) ブートストラップ相互検証は、従来の相互検証よりも本質的に効果があり、以下のものを含む。(i) データのブートストラップサンプリングが、1000又はそれ以上のオーダーの独立したブートストラップサンプル(トレーニングモデルにおいて陰性及び陽性サンプルの従来からの数を備える各独立したブートストラップサンプル、例えば、RRCFデータに対する59 G n e g及び121 G p o s)のnB数に対して何度も、トレーニングモデルに取り出される場合には、トレーニングに用いられるサンプルの結果としての経験上の分布は、データの単一のセットよりも、データによって表される自然状態の潜在的な分布により近づく(この現象は、ブートストラップサンプリングに対して固有である); (ii) また、ブートストラップサンプリング(これは復元(replacement)を伴うサンプリングである)に対して固有である、データの約37%が、いずれの所定のデータのブートストラップサンプリングによっても選択されず(確率論によって、所定のブートストラップサンプリングによって選択されないデータの分数は、 $(1 - 1/nB)^{nB}$ であるため。およそ=0.367、nB>100の場合; Efron & Tibshirani, pp. 281-282; Davison & Hinkley, p. 114)、これにより、本質的に、トレーニングモデルのために用いられるデータの所定のブートストラップサンプルにおいて用いられない相互検証フェーズにおいて用いられるサンプルのテストセットとして、対応する補足的なデータのセットを提供し; そして(iii) nB>1000を伴うブートストラップサンプルから経験的に決定される統計的な信頼区間は、信頼できるものであり容易に得られる。分析された各事例において、nB=10000の独立したデータのブートストラップサンプリングを伴うブートストラップ相互検証を適用した。このように、対応するnB=10000のテストセットの全体(ensemble)が生成される。

【0325】

RRCF又はRL2F RT-PCRデータ(SG43RGP46-performance及びSG43RGP36-RGPgreedysearch)に基づく、2つの異なるRGP Vmodについてのブートストラップ相互検証の結果を以下に示す。いずれも、GNOS投票閾値として0.5を用い、合理的で発生しやすい状況であるグレードIII又はIVの急性GVDH(Gag3)の30%の疾患有病率についてのものである。実験的に誘導された任意の達成指標の周囲の90パーセントの信頼区間は括弧の中に示した。

【0326】

VmodSG43RGP46-performance: 0.79平均感度(0.67,0.90)、0.75平均特異性(0.58,0.90)、0.76平均正確性(0.65,0.86)、0.59平均陽性予測値(0.45,0.75)、及び0.89平均陰性予測値(0.84,0.94)。

【0327】

VmodSG43RGP36-RGPgreedysearch: 0.84平均感度(0.73,0.94)、0.84平均特異性(0.71,0.95)、0.84平均正確性(0.75,0.92)、0.71平均陽性予測値(0.56,0.88)、及び0.93平均陰性予測値(0.88,0.97)。

【0328】

このように、これらの2つのVmodは、計算機処理によるブートストラップ相互検証で、実用レベルまでの非常に優れた有効性が確認された。特にNegative Predictive Value(陰性予測値)は、ドナーがGVHDを引き起こさないと予測された場合には、GVHD転帰の正しい予測である率は、臨床的には非常に重要である。

実施例16

【0329】

この実施例は、(a)異なるアッセイプラットフォームから得られた遺伝子発現データ測定、及び、(b)混乱する可能性のある測定及びサンプルの挙動の変化を反映するためにノイズで変更された入力データに関するGVHD転帰予測例についての一貫して信頼性のあるRGP Vmod性能についての説明を含む。

【0330】

この実施例は、(a)異なるアッセイプラットフォームから得られた遺伝子発現データ測

定、及び、(b) 混乱する可能性のある測定及びサンプルの挙動の変化を反映するためにノイズで変更された入力データに関する G V H D 転帰予測例についての一貫して信頼性のある RGP Vmod 性能についての説明を含む。

【0331】

レシオメトリック G V H D 転帰予測と RGP の自己校正特徴とを組み合わせることに加え、G V H D 転帰予測におけるさらなる正確性及び信頼性が、複数 RGP 投票モデル (Vmods) の使用によって、個別の RGP 投票によって加えられる誤差を平均化することにより実現されることが期待される。多重遺伝子複数 RGP 投票モデルの安定化、誤差補償及び誤差希釈化を組み合わせた特徴は、たとえ以下の場合であっても、全般に信頼性の高い転帰予測を提供できると思われる。(1) RT-PCR プラットフォームではなくてマイクロアレイ

10

プラットフォームを用いて遺伝子発現が測定される場合。すなわち

1. RGPVmod に関して、トレーニング用及びオリジナルの性能測定が実行するデータを提供する方法と比較して、全く別の方法を用いる場合

2. 正確性が低く感度も低いと一般に考えられている別の方法を用いる場合 (マイクロアレイ遺伝子発現アッセイはノイズが多くサーベイ目的により適していると考えられている一方、RT-PCR は、特に人間の臨床診断に関する場合には、定量的な遺伝子発現アッセイの標準的な方法と考えられている)、及び

(2) 遺伝子発現データが、異なるプラットフォーム (例えば、マイクロアレイ及び RT-PCR) から得られたものであり、サンプルの取り扱い、ラボ内処理、器具ノイズ、生物学的多様性等によって、様々な変動源によって歪められている場合 (既存の測定データにコンピュータで生成されたランダムノイズを高レベルで加えることによりシミュレート可能である)。

20

【0332】

G V H D 転帰 予測性能は、全 180 サンプル (上記の RRCF 又は RL2F データ) についての TaqMan リアルタイム RT-PCR 測定、及び、RT-PCR アッセイと全く同じ遺伝子 180 サンプル (上記の VQLS) 中の 163 サンプルについてのイルミナ社の HT12 v3.0 マイクロアレイ測定に基づいて、3 つの異なる RGP Vmods について決定された。データが RGP についてはことなるスケールであるため、RGP 区分線は、RT-PCR とマイクロアレイ データセットとで別個に定めた。GNOS 値及び GNOS 閾値は、両方のデータセットについて同じ方法で定めた。

【0333】

ノイズ存在下での RT-PCR 及びマイクロアレイ測定データについての Vmod G V H D 転帰予測の信頼性同じ遺伝子の RT-PCR 測定データ及びマイクロアレイ測定データについて、ノイズ (コンピュータでランダムに生成された) を測定値に加えた。対応する RGP 値を算出する前に、SG 標準偏差を有する $\pm 0.1x \sim \pm 10x$ の範囲の均一なランダムノイズを、各サンプルについて各 SG 値に加えた。SG 標準偏差は、全 180 の RT-PCR 及び 163 のマイクロアレイ測定に対して、各 SG について特別に決定された。問題となっているノイズには、サンプルのクラスの違いによる生物学的変動だけでなく具体的ではない生物学的及び測定アッセイ変動も含むため、SG 標準偏差の推定は誤差が大きくなるように設計されている。シミュレートされたランダムノイズは、乱数発生器でサンプル化され、SG 測定値に 1000 回加えられた。繰り返しごとに、各 Vmod について、RGP 及び対応する GNOS 値が報告された。各ノイズサンプリングごとの GNOS 値の各セットについて、5 つの SSPC (転帰予測についての標準的な明細 (specification)) が決定された。7 つの異なるレベルのノイズの各々について、SSPC の平均値及び標準偏差が、3 つの異なる分割について、異なる GNOS 閾値及び G V H D 有病率ごとに決定された。

30

40

【0334】

G V H D 転帰 予測性能の例は、RT-PCR 又はマイクロアレイ 測定データのいずれかに加えられた $0.1x \sim 10x$ 標準偏差のノイズによって歪められた SG 入力に基づいて、3 つの G V H D 転帰予測試験について、Gneg 対 Gag3 分割について、GNOS 閾値 0.55、Gag3 有病率 25% (一般に許容されている 15% ~ 35% の範囲の midpoint である) として、以下のとおり報告される。

【0335】

50

表 2 7 (SG43RGP36noisecompRRCFVQLS) は、RT-PCR及びマイクロアレイデータについて、これまで最も優れている48のgeneGVHD転帰予測使用、SG43RGP36-RGPgreedysearchについて、ノイズによる改変に対してGVHD転帰予測がどの程度ロバスト性を有するかの比較を示す。標準偏差の0.1倍のノイズでは、RT-PCRは、マイクロアレイ由来のVmodの結果と比較して、低い(良い)log10 p 値(4桁)及び高い(良い)GVHDの減少(10%程度)、及び75%程度がそれ以上のGVHDの減少及びGVHD・Nドナーキャプチャーが得られた。しかしながら、標準偏差以上のノイズでは、RT-PCR由来のVmodとマイクロアレイ由来のVmodとの間でSSPCはほとんど全て区別できなかった(両者について、SSPCの平均及び標準偏差)。標準偏差の2倍程度のノイズでも、GVHDの減少及びGVHD・Nドナーキャプチャーについては、RT-PCR及びマイクロアレイについて45%程度又はそれ以上であった。

10

【0336】

要するに、これらの結果から、代替的な測定プラットフォーム(例えば、マイクロアレイ、RT-PCR)及び入力データの高レベルの改変に対して、RGP Vmodに内在するロバスト性が確認された。少なくとも50%程度のGVHDの減少及びGVHD・Nドナーキャプチャーが、大きな入力データ改変に繋がるような極端な状況であっても、SG43RGP36-RGPgreedysearchのGVHD転帰予測使用によって実現される。

【0337】

表 2 8 (3VmodnoisecompRRCF) は、3つの代替的なVmodGVHD転帰予測使用、SG43RGP46-RGPperformance、SG43RGP36-RGPgreedysearch、及びSG21RGP28-RGPmaxgreedysearchについて、ノイズによる改変に対して、GVHD転帰予測がどの程度のロバスト性を有しているかの比較をRT-PCRデータについて示す。標準偏差の0.1倍のノイズでは、SG21RGP28-RGPmaxgreedysearchが、最も優れた性能を示し、次がSG43RGP36-RGPgreedysearchであった。しかしながら、標準偏差の1倍のノイズでは、SG21RGP28-RGPmaxgreedysearchのSSPCは、SG43RGP46-RGPperformanceとほとんど区別できなかったが、最も優れた性能を示したVmodは、SG43RGP36-RGPgreedysearchであった。要するに、RT-PCRデータをGVHD転帰予測への入力として用いた場合には、SG21RGP28-RGPmaxgreedysearchは低レベルノイズでは優れた性能を発揮するが、SG43RGP36-RGPgreedysearchの方がロバスト性が強く、改変するようなノイズの存在下でより優れた性能を発揮する。

20

【0338】

表 2 9 (3VmodnoisecompVQLS) は、3つの代替的なVmodGVHD転帰予測使用、SG43RGP46-RGPperformance、SG43RGP36-RGPgreedysearch、及びSG21RGP28-RGPmaxgreedysearchについて、ノイズによる改変に対して、GVHD転帰予測がどの程度のロバスト性を有しているかの比較をマイクロアレイデータについて示す。全てのノイズレベルにおいて、SG21RGP28-RGPmaxgreedysearchのSSPCは、SG43RGP46-RGPperformanceとほとんど区別できなかった。一方、最も優れた能力を発揮したVmodは、SG43RGP36-RGPgreedysearchによる表現であった。要するに、マイクロアレイデータをGVHD転帰予測への入力として用いた場合には、SG43RGP36-RGPgreedysearchは、SG21RGP28-RGPmaxgreedysearch及びSG43RGP46-RGPperformanceと比較した場合、最もロバスト性に優れており、改変するようなノイズの存在下で最も優れた性能を発揮する。

30

40

【0339】

表 A (Graph 1、3VmodnoisecompTtest) は、3つの代替的なVmodGVHD転帰予測使用、SG43RGP46-RGPperformance、SG43RGP36-RGPgreedysearch、及びSG21RGP28-RGPmaxgreedysearchについて、ノイズによる改変に対して、GVHD転帰予測 p 値がどの程度のロバスト性を有しているかの比較をRT-PCRデータ及びマイクロアレイデータについて示す。最も低い p 値は、SG43RGP36-RGPgreedysearch及びSG21RGP28-RGPmaxgreedysearchについて、RT-PCRデータを用いた場合にのみ観測されたがマイクロアレイデータを用いた場合には観測されなかった。また、全モデルについて、p 値は、RT-PCRデータに用いるかマイクロアレイデータに用いるかによらず、標準偏差の0.5倍まで、本質的に、ノイズ中の摂動に対して高いロバスト性を有しており、1倍のノイズでも極めて低い $10^{-6} \sim 10^{-8}$ 程度の範

50

困にある。しかしながら、p値は、標準偏差の2倍のノイズで顕著に劣化したが、GVHD転帰予測試験については使用可能性がある範囲である。標準偏差の5倍以上のノイズ（これは、非常に大きなノイズである）では、転帰予測p値は、本質的に完全に改変される。また、標準偏差の1倍程度のノイズで、VmodSG43RGP36-RGPgreedysearchは、他のVmodと比べて、最も低いp値、すなわち 10^{-8} より小さい値を示した。これは、RT-PCRデータを使うか又はマイクロアレイデータを使うかによらない。また、VmodSG43RGP36-RGPgreedysearchは、SG21RGP28-RGPmaxgreedysearchと比較して、ノイズによる改変に対して優れたロバスト性を示した。

【0340】

表B（Graph 2、3VmodnoisecompGVHDred）は、3つの代替的なVmodGVHD転帰予測使用、SG43RGP46-RGPperformance、SG43RGP36-RGPgreedysearch、及びSG21RGP28-RGPmaxgreedysearchについて、ノイズによる改変に対して、予想GVHD減少がどの程度のロバスト性を有しているかの比較をRT-PCRデータ及びマイクロアレイデータについて示す。最も高い予想GVHD減少は、80%~90%の範囲であり、SG43RGP36-RGPgreedysearch及びSG21RGP28-RGPmaxgreedysearchについて、RT-PCRデータを用いた場合にのみ観測されたがマイクロアレイデータを用いた場合には観測されなかった。予想GVHD減少は、ノイズの0.1倍~0.2倍では、改変に対するロバスト性が高い。0.5倍~1倍では、全てのVmodにおいて、50%~75%のレンジでGVHDの減少が示された。標準偏差の2倍のノイズでも、GVHDの減少は、依然として35%~50%の範囲で予想された。5倍では（これは非常に大きなノイズである）、予想GVHD減少は、ほとんど完全に改変された。

【0341】

興味深いことに、予想GVHD減少は、0.5倍以上のノイズを加えたときに、全てのVmodについて、RT-PCRデータは、RGPを選択しVmodをデザインするために使用されるにもかかわらず、マイクロアレイを用いたときの方がRT-PCRデータを用いたときよりもロバスト性が高い。また、1倍のノイズでは、全てのVmodについて、SG43RGP36-RGPgreedysearchが最も高い、すなわち65%程度の予想GVHD減少を、RT-PCR及びマイクロアレイデータに対して示した。

【0342】

臨床利用の結論及び優先順位概して、SG43RGP36-RGPgreedysearchは、SG21RGP28-RGPmaxgreedysearch及びSG43RGP46-RGPperformanceと比較して、標準偏差の1倍までの中程度のノイズの存在下では、RT-PCRとマイクロアレイデータの両方について最もよいSSPCを有しており、最もロバスト性が高い。しかしながら、RT-PCRデータを使うときの低ノイズにおいては、SG21RGP28-RGPmaxgreedysearchが最も良いSSPCを示す。また、3つのGVHD転帰予測Vmodの間の相違は、マイクロアレイデータを用いるときよりもRT-PCRデータを用いるときにより顕著である。これは、RT-PCRデータの忠実性及び正確性がマイクロアレイデータよりも高いためと考えられる。このように、マイクロアレイデータ及びRT-PCRデータを用いたGVHD転帰予測使用は、いずれも可能であるが、RT-PCRの方が高い忠実性を提供し、全体的に優れたGVHD転帰予測性能が得られる。

【0343】

GVHD転帰予測の実際の臨床応用について、SG43RGP36-RGPgreedysearchは、GVHD・N転帰予測性能及びロバスト性の両方に優れているので、使用の利点がある。SG21RGP28-RGPmaxgreedysearchVmodの全てのSG及びRGP（低ノイズで優れている可能性がある）もSG43RGP36-RGPgreedysearchに含まれているので、SG21RGP28-RGPmaxgreedysearchの結果は、SG43RGP36-RGPgreedysearchと同じ測定値に基づいて決定できる可能性がある。したがって、SG43RGP36-RGPgreedysearchを用いた、GVHD・N転帰予測は、最も信頼できると考えられ、平行して行うSG21RGP28-RGPmaxgreedysearchの調査評価によって、GVHD転帰予測のためにVmodSG43RGP36-RGPgreedysearchで用いたのと同じ測定データの関連するサブセットの処理から、このVmodの利点及び応用が決められるであろう。

（表27）

10

20

30

40

50

(表 28)

(表 29)

(表 A)

(表 B)

実施例 17

【0344】

この実施例は、絶対的なRT-PCR遺伝子発現データを相対的なものと比較することにより、GVHD転帰予測及びGVHD転帰予測分析の生物学的な基礎に関連する追加的な証拠の説明を含む。

【0345】

絶対的な遺伝子発現は、RL2F (outlier値及びnondetectable値の置換を含む。「RT-PCRデータの4工程の前処理でRRCF値が得られる。」という上記の説明を参照。GVHD転帰予測決定に基づくもの。)に関して上述したように、RT-PCR測定使用の出力から直接評価される。したがって、絶対的な遺伝子発現解析は、多くの変動源(出発物質の変動、サンプルの操作及び処理、細胞の代謝状態、器具のキャリブレーション)の影響を受ける。このような変動源は、相対定量化処理によって補償可能である。such as carried out forRRCF及びRGP(「RT-PCRデータの4工程の前処理でRRCF値が得られる。」という上記の説明を参照。GVHD転帰予測決定に基づくもの。また、「RGPの決定」という項も参照)。

【0346】

したがって、絶対的な遺伝子発現は、一般に、人間の診断用とは用いられない。しかしながら、所定の統計的/数値的セーフガード及び追加的なQC確認項目があれば、絶対的な遺伝子発現は、信頼できる人間の診断用途にも用いることができる。

【0347】

注記: GVHD転帰予測試験において用いられるRGPへの適用については、相対的なRRCFと絶対的なRL2FデータのいずれをGVHD転帰予測試験への入力として用いるかは重要ではない(上記の「RGPの決定」)。

【0348】

RL2Fデータに基づくGVHD転帰予測RL2Fデータ(絶対的なRT-PCR量)に基づいてGVHD転帰予測を評価する場合には、0.05以下のp値を有する175の選択遺伝子が2倍も多く観察される(表30参照、RL2FRRCFSGcomp)。Gneg対Gag2についてのT検定p値のゲノム平均は0.0458であり、対応するRRCF値と比較して非常に小さい。(注記:ゲノム平均は、伝統的に推奨されている統計的なp値を平均化する方法である。例えば、 $p_1=0.00001$ 及び $p_2=0.1$ のゲノム平均は $p=0.001$ である。算術平均はミスリーディングであり $p=0.05$ となる。)

【0349】

注目すべき事に、RL2F遺伝子(175のセットから、表13参照、SG175)の95%は、P配向(P-directional)である。これは、Gpos、Gag2又はGag3サンプルの平均遺伝子発現レベルがGnegサンプルにおけるものよりも高いことを意味する。比較として、P配向遺伝子は、RRCFデータセットの49%しかない。

【0350】

この観察は、GVHD陽性転帰に関連付けられるドナーのCD4陽性T細胞の潜在的な生物学的特徴が存在することを示唆している。すなわち、遺伝子発現レベルは、一般に、GVHD陰性転帰に関連付けられるドナーと比較して、GVHDを引き起こすドナーのCD4陽性T細胞における遺伝子の圧倒的多数に関して非常に高くなる。このことは、よりアロ反応性のCD4陽性T細胞内で代謝及び転写活性が高まっているためである可能性がある。しかしながら、科学文献においては、このような代謝活性の相違に関するより詳細な

10

20

30

40

50

研究が見られない。

【 0 3 5 1 】

RL2FデータからのSGが、GVHD転帰予測に関して、個別のSGレベルでいかによく働くかによって、上記で調べたような種類のSG Vmod内でもよく働く可能性がある。しかしながら、表 3 1 (RL2FRRCFSGVmodcomp) で見たように、RL2Fデータは、VmodSG43RGP36-RGPgreedysearchの43個のSGに関しては、RRCFデータと比較して、非常に働きが悪い。

【 0 3 5 2 】

表 3 1 がVmod SG43RGP36-RGPgreedysearchの43個のSG使用に関して示唆するように、SGレベルでは明らかに、RL2Fデータは、GVHD転帰予測用途に上記で選択されたVmodにおいては、RRCFデータを代替することはできない。しかしながら、SGの優先順位がRL2Fデータ (レシオメトリック又はその他の自己校正方法を含まないもの) について特別に選択されるならば、現在検討されているRL2F SG Vmodバージョンよりも高いGVHD転帰予測性能を有するRL2F SG Vmodをデザインすることも考えられる。校正困難な絶対的RL2F RT-PCRデータのラボ内測定信頼性に内在的な欠点を前提として、RL2Fデータに基づいてGVHD転帰予測試験をデザインすることは、原理的には可能であるが、信頼性に関するリスクが高く、それゆえにこれ以上探求されないかもしれないが、開発の優先度を置くべきかもしれない。

(表 3 0)

(表 3 1)

10

20

実施例 1 8

【 0 3 5 3 】

この実施例は、GVHDの重症度の増加を反映するGVHD転帰予測グループにおけるGNOS値のランク順位に関する説明を含む。

【 0 3 5 4 】

本明細書で分析されたGVHDグループは、GVHDの様々な強度を反映しており、Gnegすなわち no GVHDからGag3すなわち重傷でしばしば致命的となるグレードIII又はIV急性GVHD、及びその間にある様々な疾患の強度の変化を反映している。具体的には、GVHD転帰のクラスは6つの異なるグループ (上記のグループとは別のもの) をカバーしている。医学的に受容されているGVHD重症度の順序は以下のとおり。(1) Gneg (急性GVHDも慢性GVHDも発生しない)、

(2) cGonly (急性GVHDは発生せず慢性GVHDのみ発生)、

(3) ag2 (グレードII急性GVHDが発生するが、グレードIII又はIVの急性GVHDは発生しない。慢性GVHDは発生してもしなくともよい。)

(4) Gpos (慢性GVHD及びグレードII、III又はIV急性GVHDを含む任意のGVHDが発生)

(5) Gag2 (グレードII、III又はIVの急性GVHDが発生する。慢性GVHDは発生してもしなくともよい。)

(6) Gag3 (グレードIII又はIVの急性GVHDが発生する。慢性GVHDは発生してもしなくともよい。)

【 0 3 5 5 】

これらの6つのグループの各々におけるサンプルについて、変動する数値的ノイズの存在下におけるデータから、GNOS値を3つの異なるVmodについて平均化した (上記の「ノイズ存在下でのRT-PCR及びマイクロアレイ測定データについてのVmod GVHD転帰予測の信頼性」参照)。これらのGNOS平均から、ランクを下り順で決定した。すなわち、最も高いGNOS平均は、1.0にランクされ、最も低いGNOS平均は 6.0にランクされた。

【 0 3 5 6 】

表 3 2 (GNOSrankorder) において、最も性能のよいVmods SG43RGP36-RGPgreedysearch及びSG21RGP28-RGPmaxgreedysearch、またこれに近いVmodSG43RGP46-RGPperformanceについ

30

40

50

て、我々は、一貫して、GNOS平均の同じランク順位を観察し、上記の医学的に認められた疾患重症度についてもそのようにした。これは、これらのVmodの入力として用いられるRT-PCR及びマイクロアレイデータに一貫して適用された。標準偏差の2倍のノイズの存在下でも、Gneg及びGag3グループは、一貫して、極端なランクを反映し、他のグループは、概して、上記の順位における中間のどこかに落ち着いた。

【0357】

結論として、GNOS値は、6つのGVHDグループ平均のランクに反映されるので、GVHDの重症度に対応して直接に変化する内在的で、統合された、真の生物学的シグナルを反映する可能性が非常に高い。この対応は、医学的に認められている上記のGVHD重症度の順位にも示されている。統合的で潜在的な生物学的シグナルの反映は、使用されるVmodの種類によらず、RT-PCRデータとマイクロアレイデータのいずれが使用されるかによらず、また測定データに僅かなレベルから極端なレベルまでの数値的にランダムなノイズが加えられても、ロバスト性が高い。このように、ドナーのCD4陽性細胞遺伝子発現プロファイルに基づく、レシピエントのGVHDの転帰予測は、根本的には、これらの細胞における複雑な生物学的な遺伝子活性及び抑制パターンによるものであり、レシピエントのGVHDの重症度に対応して変化し、本明細書で説明した、レシオメトリック遺伝子対の例示した投票モデルにおいて情報として補足されるものである。

(表32)

実施例19

【0358】

この実施例は、高い転帰予測性能及び有効可能性を有する遺伝子の複数の選択肢及び投票モデル選択の説明を含む。

【0359】

任意のGVHD転帰予測の単一の分類者又は投票者は、RT-PCR又はマイクロアレイに基づく遺伝子発現データが、単一の遺伝子、遺伝子対、又は統合された投票モデル(例えば、RL2F、RRCF、VQLS、SG、RGP、GNOS等)のどのレベルで処理されたかによらず、分類される各サンプルについて、連続的な分類者レベル(CL)となる。GnegサンプルについてのCL平均がGposサンプルについてのものよりも高い場合には、この分類者は、N配向(N-directional)、又は N_d (GVHD陰性のN)と考えられる。GposサンプルについてのCL平均がGnegサンプルについてのものよりも高い場合には、この分類者はP配向(P-directional)、又は P_d (GVHD陽性のP)と考えられる。Gnegサンプル及びGposサンプルのそれぞれのCL平均の間の中点は、各CLについての区分線と定義される。 N_d 分類者について、そのCLが区分線よりも高いか等しい場合には、値1で表されるGVHDのN票が投票される。投票値は逆に0に設定されてもよい。 P_d 投票者について、そのCLが区分線よりも低い場合には、値1で表されるGVHDのN票が投票される。

【0360】

分類者の任意の組のGVHD・N転帰の票は、上述した分類者の極めて多数の組み合わせから得られ、投票モデル(Vmod)に統合される。この投票モデルは、本明細書において説明したように、単純にGVHDのN票の平均を生成する。この平均は、GNOS(GVHD陰性転帰スコア)と呼ばれる。しかしながら、投票者及び分類者は、信頼できるGVHD転帰予測を得るための他の好適な方法を用いて統合されてもよい(以下の「多変量転帰予測モデルの代替手段」を参照)。GNOSは、1に設定された「GVHD・N」票について定義されたので、「非GVHD・N」票は0に設定され、GNOSの分類者は、常にN配向である。また、GNOSに応じてサンプルの分類の最終のGVHD転帰を決定するときには、所望のGVHD転帰予測性能目標に応じて、しばしば区分線ではなく「N投票閾値」(例えば、最も性能がよいSG43RGP36-RGPgreedysearchについては55%)が選択的に適用される。

【0361】

GVHD転帰予測分類者の定義一般に、表13(RNA175)、表2B(RNA192list)(表13のRNA175遺伝子は、表2BのRNA192配列表にも記載されている)、又はRNA1546配列

10

20

30

40

50

表又はRNA1538配列表（表13におけるRNA175遺伝子の全てがRNA1546又はRNA1538配列表に記載されているとは限らない）中の遺伝子を用いて、複数の、ほぼ無限定の（上記の及び一般のレシオメトリック遺伝子対、又は一般の遺伝子対の様々な組み合わせサブセットに基づいて、又はSGを直接分類者として使用して、例えばRL2F、RRCF、VQLS、SG、RGP等のデータに反映されて）、GVHD転帰予測のためのVmodが生成され、有効化される。そして、 N_d 分類者及び P_d 分類者の様々な組み合わせが許容される。すなわち、(1) N_d 分類者と P_d 分類者の混合、 N_d 分類者及び P_d 分類者の相対的な表現を変更できる。

(2) P_d 分類者のみ

(3) N_d 分類者のみ

【0362】

RGP Vmodすなわちレシオメトリック遺伝子対投票モデルに関しては、相対的なSG測定（RT-PCRデータ又はマイクロアレイデータを用いる）に基づいて、転帰予測シグナル（ X/Y 又は同様の $\log [X/Y]$ 又は $\log X - \log Y$ ）遺伝子対レベルでラボ内での解析に用いることができるように、（RGPの内在的な自己校正の特徴に加えて）、大多数のケースではRGPメンバー遺伝子の方向性（directionality）は逆向きになり、すなわち、遺伝子Xが N_d であれば遺伝子Yは P_d となり、遺伝子Xが P_d であれば遺伝子Yは N_d となる。これは、遺伝子Xの「活性化経路（activator pathway）」活性が遺伝子Yの「抑制経路（inhibitor pathway）」活性よりも相対的に高い場合（及びこの逆）にのみ、生物学的応答が起こることの生物学的な理由の解釈に沿うものである。

【0363】

また、レイヤー化された競合的な経路の場合もある。例えば、経路X及び経路Yは、絶対的な意味で生物学的応答のために高められ得るが、生物学的応答が完全に起こるには、経路Xは経路Yよりも相対的に高くなければならない。このように、RGPを用いる転帰予測について、遺伝子X-Y、 P_d - P_d 又は N_d - N_d 対が起こりえる。すなわち、 P_d 及び N_d 配向は、相対定量化RTPCR又はマイクロアレイデータに関してSGレベルで決定される。しかしながら、概して、RGP投票モデルに寄与するRGPは、 P_d 状態又は N_d 状態を有するSGの数に関して、SGレベルで均等に釣り合っていない（相対的に定量化された遺伝子発現データについては）。

【0364】

しかしながら、RL2F絶対SG測定に基づくRGP Vmodに関しては、RL2F遺伝子の大部分がP配向（上述）にバイアスされているので、大部分の場合、RGP X-Y対、 P_d - P_d 対は、転帰予測分類者として、Vmodにおいて用いられる。また、有効なRGP値は、入力としてRRCF又はRL2Fデータが用いられるかに依存しない。

【0365】

SGを構成要素分類者として使用するVmodに関しては、特に遺伝子発現データの相対定量化、例えばRRCFを用いる場合には、基本原理に従って、SGのP配向又はN配向はVmodが有効であるために要求されない。表13（RNA175）のSGは、RRCF及びVQLSデータから決定され、配向に関して、相対的に均等に釣り合っている（表30も参照）。

【0366】

しかしながら、上記のとおり、GVHD転帰予測のために相対ではなく絶対RT-PCR定量化RL2Fデータを用いる場合には、GVHD転帰予測に関し、SGの能力において有病率について支配的な天然のP配向遺伝子へのバイアスが存在する（データで示された生物学的傾向）。このように、絶対定量化のレベルで、RL2FベースのGVHD転帰予測について、任意の良好な投票モデルの候補は、RL2Fレベルで、P配向SGの大多数に基づくものとなると思われる。絶対RT-PCR定量化に基づくかかる統合化された P_d 配向SG Vmodは、GVHD転帰予測において非常に効果的である、おそらくはGVHD転帰予測試験として開発される。しかしながら、RT-PCRの診察用途における現在の実務を前提とすれば、絶対定量化は今のところ人間の診断については信頼できる解析とは考えられておらず、絶対RT-PCR定量化に基づくSG Vmodは、開発の現時点での優先事項ではない。しかしながらかかるモデルは、将来において開発の優先事項となり得る。

10

20

30

40

50

実施例 2 0

【 0 3 6 7 】

この実施例は、多変量転帰予測モデルについて開示された例示的なVmodの代替の説明を含む。

【 0 3 6 8 】

個別のRGP票の選択セットをGNOS値に集約化又は平均化することは、複数の遺伝子の個別のmRNA測定レベルにおける情報を効率的に、実際に、確実に、また透明に使用し、GVHD・N転帰スコアを提供する最も単純な方法である。しかしながら、多くの現存の代替的な方法（一般に分類者と呼ばれる）は、複数のRGP Vmodに加えて、多変量予測モデルを生成するものである。このような代替的な分類者（Richard O. Duda, Peter E. Hart, & David G. Stork, *Pattern Classification*, Second Edition, John Wiley & Sons, Inc, NY, 2001）には、個別の変数の重み付け平均、変数のペアごとの（pari-wise）組み合わせの重み付け平均、又は多変量組み合わせの重み付け平均、又はLDA（線形判別分析）に組み込まれる非線形のもの、QDA（二次線形判別分析）、ディビジョンツリー、SVM（support vector machines）、k近傍法、ニューラルネットワーク等、又は一般化された多変量線形及び非線形モデルの様々な実装、変化する自由度を伴うもの、賢明な検索（judicious search）、及び、最適化アルゴリズム（例えば、いわゆる遺伝的アルゴリズムのような、古典的な最適化アルゴリズム又は微分情報を用いない（derivative-free）アルゴリズム）と結合したもの、に基づくものが含まれる。このような代替的な方法は、ここにリストされるSGs、RGP s及びPRGP sのリストからGNOS値を引き出すために用いられ得る。しかしながら、このようなモデルの比較上の複雑さ（comparative complexity）及び自由度に依存して、より観察的な結合ドナー遺伝子発現測定及び関係するレシピエントGVHD臨床転帰データサンプルが、このような代替的な、より複雑な分類者の実装に対して適切な統計学上のサポートを提供するために、必要とされ得る。

10

20

【 図 1 】

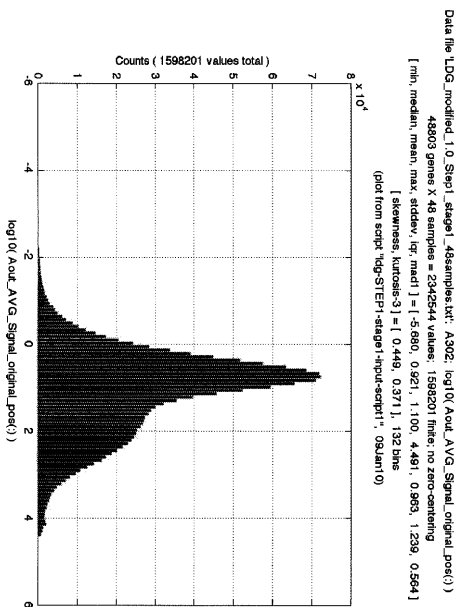


FIGURE 1: (Plot 5.1): Histogram of log10 AVG_Signal_original
Extracted from 48903 Genes by 48 Samples.

【 図 2 】

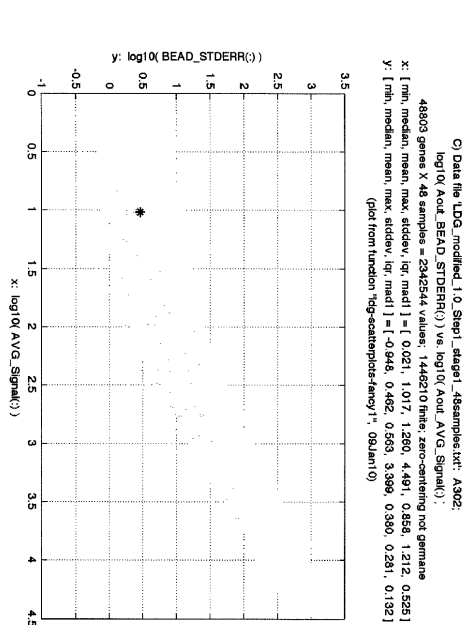


FIGURE 2: (Plot 6.5): Scatterplot of log10 BEAD_STDErr vs. log10 AVG_Signal
Extracted from 48903 Genes by 48 Samples.

FIGURE 4: (Plot 7.4): Histogram of 0.4343*c1*VST-modified SIGNAL (for Signal restricted to < 0.95 quantile of all Signal values) Extracted from 48803 Genes by 48 Samples.

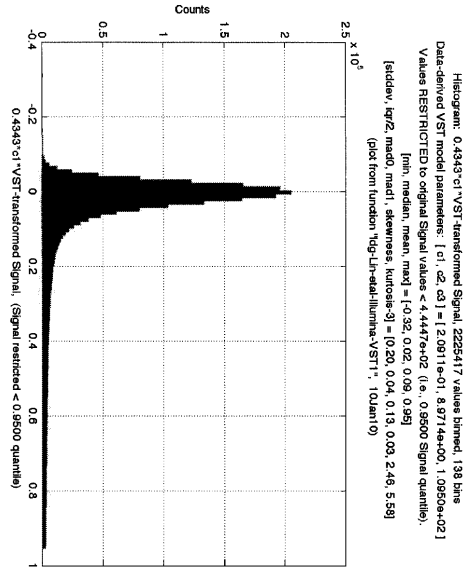


FIGURE 7: RNA20 TCS

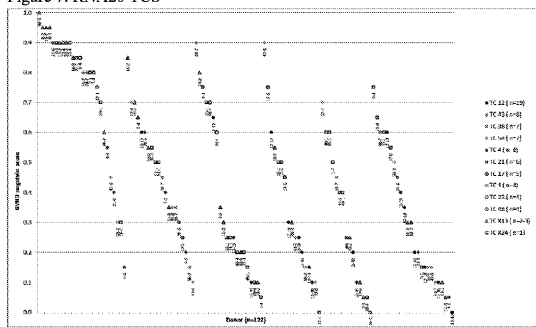


FIGURE 8: CTCF GROUPS

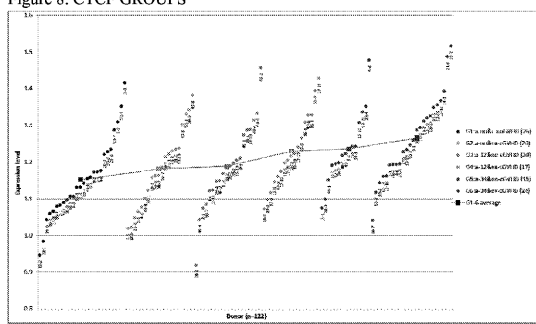


FIGURE 3: (Plot 5.1): SQRT(Positive Signal VARIANCE - c3) vs. Positive SIGNAL, (and for which Signal Variance > c3=109.5042) Extracted from 48803 Genes by 48 Samples.

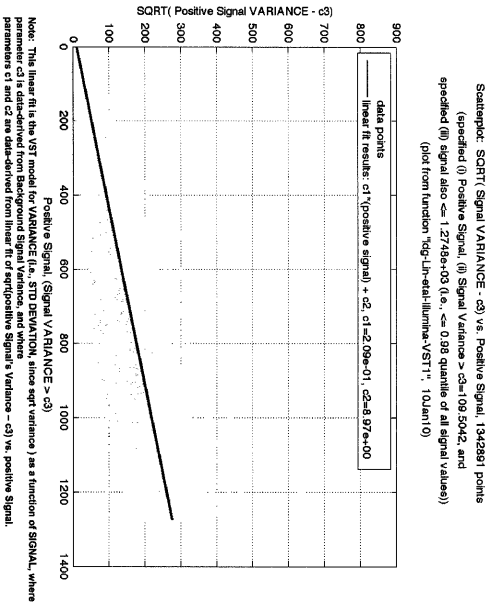


FIGURE 5: CTCF TCS

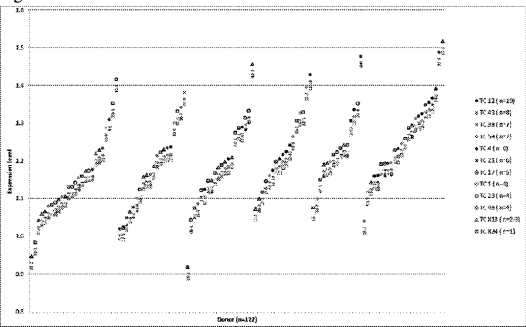
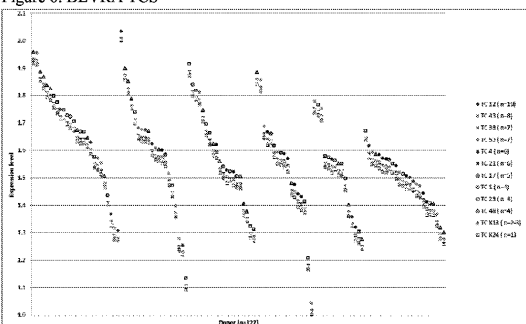
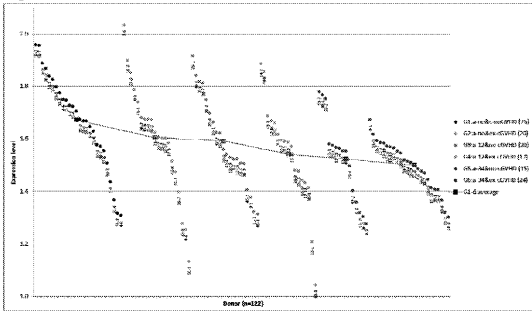


FIGURE 6: BLVRA TCS



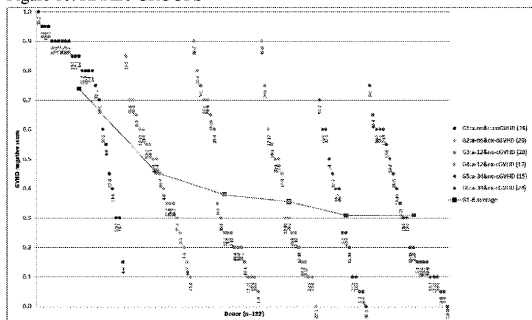
【 図 9 】

Figure 9: BLVRA GROUPS



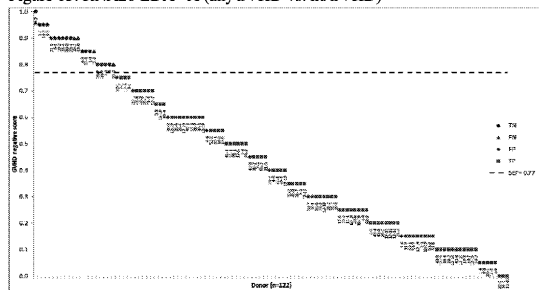
【 図 10 】

Figure 10: RNA20 GROUPS



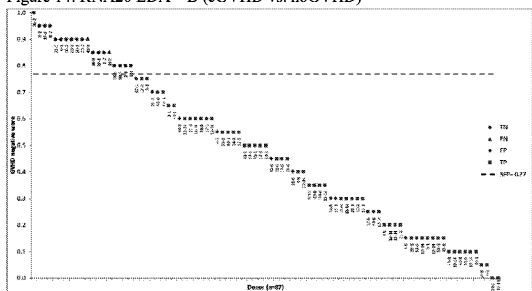
【 図 13 】

Figure 13: RNA20 LDA - A (anyGVHD vs. noGVHD)



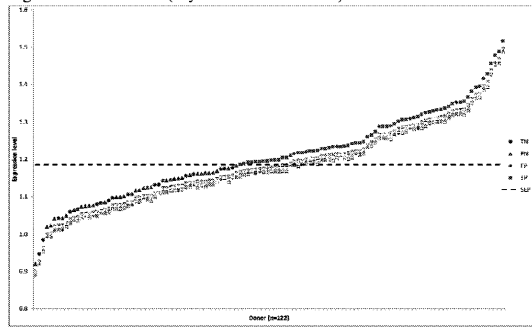
【 図 14 】

Figure 14: RNA20 LDA - B (cGVHD vs. noGVHD)



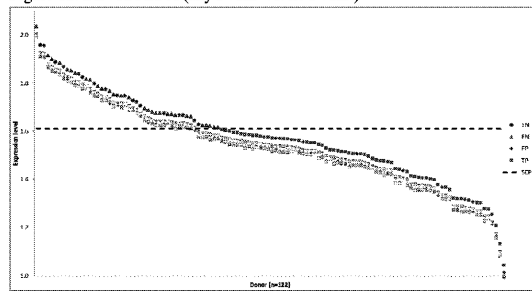
【 図 11 】

Figure 11: CTCF LDA (anyGVHD vs. noGVHD)



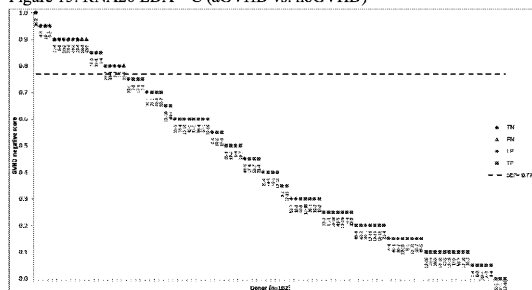
【 図 12 】

Figure 12: BLVRA LDA (anyGVHD vs. noGVHD)



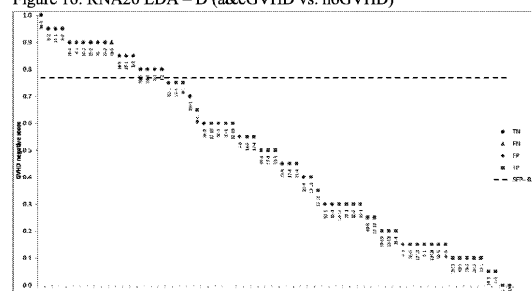
【 図 15 】

Figure 15: RNA20 LDA - C (aGVHD vs. noGVHD)



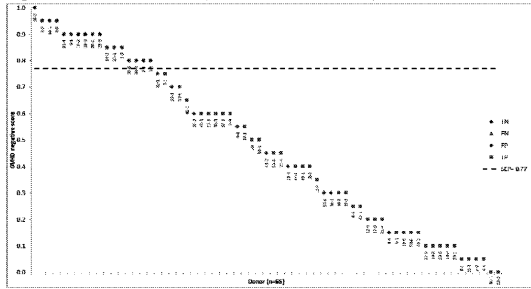
【 図 16 】

Figure 16: RNA20 LDA - D (a&cGVHD vs. noGVHD)



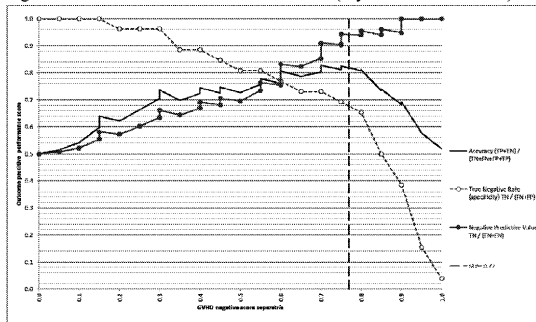
【 図 1 7 】

Figure 17: RNA20 LDA – E (a34GVHD vs. noGVHD)



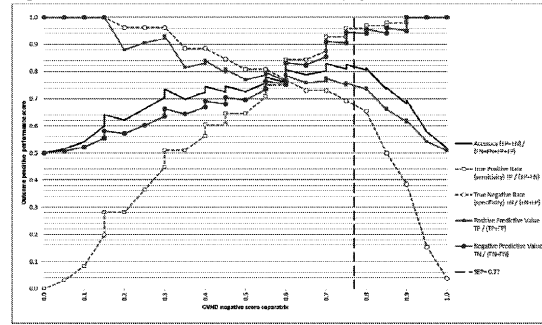
【 図 1 8 】

Figure 18: RNA20 LDA PERFORMANCE – A (anyGVHD vs. noGVHD)



【 図 1 9 】

Figure 19: RNA20 LDA PERFORMANCE – B (anyGVHD vs. noGVHD)



【 配 列 表 】

2014501098000001.app

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US 11/58669
A. CLASSIFICATION OF SUBJECT MATTER IPC(8) - C12Q 1/68; G01N 33/53; C12P 19/34 (2012.01) USPC - 435/6.1, 6.11, 6.12, 7.1, 91.2 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) USPC: 435/6.1, 6.11, 6.12, 7.1, 91.2 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched USPC: 435/6.1, 6.11, 6.12, 7.1, 91.2 (keyword limited; terms below) Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) PubWEST (PGPB,USPT,USOC,EPAB,JPAB); Google; PubMed Search terms: GVHD, graft-versus-host disease, T cells, CD8+, CD4+, risk, predict, marker, biomarker, donor, recipient, IL12RB1, VAMP2, ICOS, CD28, CD86, B7-2, BDNF, FOXP3, PREX1, NOTCH2, AEBP		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2007/0264272 A1 (PERREAULT et al.) 15 November 2007 (15.11.2007) para [0006], [0010]-[0011], [0014], [0016]-[0017], [0021], [0024]-[0025], [0050]-[0051], [0085]-[0086], [0089], [0091]-[0092], [0094], [0096], [0098], [0100], [0102], [0104]-[0106], [0112], [0127], [0129], [0134], [0136], [0140], [0146], [0148], [0150], [0152]	1-16, 23-28, 36-52, 54-59, 61-65, 67-114
Y	US 2002/0006403 A1 (YU et al.) 17 January 2002 (17.01.2002) para [0030], [0047], [0248]-[0251]; SEQ ID NO:1	1-16, 23-28, 36-52, 54-59, 61-65, 67-114
Y	OGAWA et al. "Opposing effects of anti-activation-inducible lymphocyte-immunomodulatory molecule/inducible costimulator antibody on the development of acute versus chronic graft-versus-host disease" J. Immunol; 15 November 2001 (15.11.2001); Vol. 167, No. 10; pages 5741-5748; abstract	3, 98-104, 113
Y	SCHMID et al. "Comparison of normalization methods for Illumina BeadChip HumanHT-12 v3. BMC Genomics" 2 June 2010 (02.06.2010); Vol. 11, No. 349; pages 1-17; pg 2, para 1	6-7, 36-38, 46-50, 88-97, 109
A	MUTIS et al. "The association of CD25 expression on donor CD8+ and CD4+ T cells with graft-versus-host disease after donor lymphocyte infusions" Haematologica; October 2005 (10.2005); Vol. 90, No. 10; pages 1389-1395.	1-16, 23-28, 36-52, 54-59, 61-65, 67-114
A	PABST et al. "The graft content of donor T cells expressing gamma delta TCR+ and CD4+foxp3+ predicts the risk of acute graft versus host disease after transplantation of allogeneic peripheral blood stem cells from unrelated donors" Clin. Cancer Res.; 15 May 2007 (15.05.2007); Vol. 13, No. 10; pages 2916-2922.	1-16, 23-28, 36-52, 54-59, 61-65, 67-114
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/>		
* 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 24 April 2012 (24.04.2012)		Date of mailing of the international search report 23 MAY 2012
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201		Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4900 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 11/58669

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	REZVANI et al. "High donor FOXP3-positive regulatory T-cell (Treg) content is associated with a low risk of GVHD following HLA-matched allogeneic SCT" 15 August 2006 (15.08.2006); Vol. 108, No. 4; pages 1291-1297.	1-16, 23-28, 36-52, 54-59, 61-65, 67-114
A	MAGENAU et al. "Frequency of CD4(+)CD25(hi)FOXP3(+) regulatory T cells has diagnostic and prognostic value as a biomarker for acute graft-versus-host-disease" Biol. Blood Marrow Transplant.; July 2010 (07.2010); Vol. 16, No. 7; pages 907-914.	1-16, 23-28, 36-52, 54-59, 61-65, 67-114
A	PACZESNY et al. "A biomarker panel for acute graft-versus-host disease" Blood; 8 January 2009 (08.01.2009); Vol. 113, No. 2; pages 273-278.	1-16, 23-28, 36-52, 54-59, 61-65, 67-114
A	SCHULTZ et al. "Toward biomarkers for chronic graft-versus-host disease: National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: III. Biomarker Working Group Report" Biol. Blood Marrow Transplant.; February 2006 (02.2006); Vol. 12, No. 2; pages 126-137.	1-16, 23-28, 36-52, 54-59, 61-65, 67-114
A	US 2009/0011456 A1 (SMITH et al.) 8 January 2009 (08.01.2009)	1-16, 23-28, 36-52, 54-59, 61-65, 67-114
A	US 2006/0105637 A1 (MISCHAK) 18 May 2006 (18.05.2006)	1-16, 23-28, 36-52, 54-59, 61-65, 67-114

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 11/58669

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.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. Claims Nos.: 17-22, 29-35, 53, 60, 66, 115
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:
Group I: Claims 1, 3, 5-16, 23-28, 36-52, 54-59, 61-65, 67-71, 73-87 and 114, drawn to a method for predicting or determining the risk of a hematopoietic cell transplant (HCT) from a candidate donor to induce or not to induce graft vs. host disease (GVHD) in a HCT recipient; AND a method for classifying a hematopoietic cell transplant (HCT) from a candidate donor for risk of inducing graft vs. host disease (GVHD) in a HCT recipient.

Group II: Claims 2, 5-16, 23-28, 36-52, 54-59, 61-65, 67-87 and 114, drawn to a method for predicting or determining the risk of the HCT from the candidate donor to induce or not to induce graft vs. host disease (GVHD) in a HCT recipient comprising: contacting CD4+ T cells or CD8+ T cells, or nucleic acid or protein expressed by CD4+ T cells or CD8+ T cells, from a candidate donor with an analyte that detects expression of one or more positive or negative GVHD predictor genes listed in Tables 1 (RNA 1538), 2A (RNA 143), 2B (RNA 143), 3, 13 (SD175), 15 (SG128) or 18 (SG64), or a polymorphism thereof.

(Continued on extra sheet)

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 additional fees, this Authority did not invite payment of additional fees.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

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

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 11/58669

Continuation of Box III - observations where unity of invention is lacking:

Group III: Claims 4-16, 23-28, 36-38, 46-47, 49-52, 54-59, 61-65, 67-71, 73-81, 83-87, 98-103 and 114, drawn to a method of producing a database or organizational construct comprising a plurality of actual or candidate HCT donors each assigned a score based upon the probability or degree of risk of the actual or candidate donor HCT to induce or not to induce graft vs. host disease (GVHD) in a HCT recipient; AND a database or organization construct, comprising gene expression profile of two or more positive or negative predictor genes from a plurality of actual or candidate HCT donors, wherein the two or more positive or negative predictor genes are listed in Tables 1 (RNA 1538), 2A (RNA 143), 2B (RNA 143), 3, 13 (SD175), 15 (SG128) or 18 (SG64), or a polymorphism thereof, and wherein the database or organizational construct associates the gene expression profile with each of the actual or candidate HCT donors.

Group IV: Claims 88-97 and 104-113 and , drawn to a kit, comprising two or more primer pairs, wherein each primer pair is oppositely oriented to each other, wherein the first primer pairs hybridizes to RNA or cDNA produced from one of the positive or negative predictor genes and the second hybridizes to a housekeeping gene listed in Table 1 (RNA 1538), 2, 2A, 2B (RNA 192), 3 and/or 12; AND a kit, comprising one or more nucleic acid probes, wherein said one or more nucleic acid probes hybridizes to RNA or cDNA of one or more of the positive or negative predictor genes listed in Table 1 (RNA 1538), 2, 2A (RNA 143), 2B (RNA 192), 3, 13 (SG175), 15(SG128) or 18 (SG64), or a polymorphism thereof; AND an array of primers, comprising two or more primer pairs, wherein each primer pair is oppositely oriented to each other, wherein each of the primer pairs hybridizes to RNA or cDNA produced from one of the positive or negative predictor genes listed in Tables 1 (RNA 1538), 2A (RNA 143), 2B (RNA 143), 3, 13 (SD175), 15 (SG128) or 18 (SG64), or a polymorphism thereof, and wherein each primer pair is affixed to or contained in a support or substrate; AND an array of probes, wherein each probe hybridizes to RNA or cDNA of one or more of the positive or negative predictor genes listed in Table 1 (RNA 1538), 2, 2A (RNA 143), 2B (RNA 192), 3, 13 (SG175), 15(SG128) or 18 (SG64), or a polymorphism thereof, and wherein each probe is affixed to or contained in a support or substrate.

The inventions listed as Groups I-IV do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Groups I-III do not include the inventive concept of primer pairs or nucleic probes hybridizing to one or more of the positive or negative predictor gene, as required by Group IV.

Groups I-II and IV do not include the inventive concept of a method of producing a database or organizational construct comprising a plurality of actual or candidate HCT donors each assigned a score based upon the probability or degree of risk of the actual or candidate donor HCT to induce or not to induce graft vs. host disease (GVHD) in a HCT recipient; OR a database or organization construct, comprising gene expression profile of two or more positive or negative predictor genes from a plurality of actual or candidate HCT donors, as required by Group III.

Groups I and III-IV do not include the inventive concept of a method for predicting or determining the risk of the HCT from the candidate donor to induce or not to induce graft vs. host disease (GVHD) in a HCT recipient comprising: contacting CD4+ T cells or CD8+ T cells, or nucleic acid or protein expressed by CD4+ T cells or CD8+ T cells, from a candidate donor with an analyte that detects expression of one or more positive or negative GVHD predictor genes listed in Tables 1 (RNA 1538), 2A (RNA 143), 2B (RNA 143), 3, 13 (SD175), 15 (SG128) or 18 (SG64), or a polymorphism thereof, as required by Group II.

Group IV do not include the inventive concept of a method for predicting or determining the risk of a hematopoietic cell transplant (HCT) from a candidate donor to induce or not to induce graft vs. host disease (GVHD) in a HCT recipient; OR a method for classifying a hematopoietic cell transplant (HCT) from a candidate donor for risk of inducing graft vs. host disease (GVHD) in a HCT recipient, as required by Group I.

Groups I-IV share the technical feature of the one or more positive or negative GVHD predictor genes listed in Tables 1 (RNA 1538), 2A (RNA 143), 2B (RNA 143), 3, 13 (SD175), 15 (SG128) or 18 (SG64).

Groups I-III share the technical feature of Claim 1.

However, these shared technical features do not represent a contribution over the prior art, specifically, Claim 1 lacks an inventive step under PCT Article 33(3) as being obvious over US 2007/0264272 A1 to Perreault et al. (Perreault) in view of US 2002/0006403 A1 to Yu et al. (hereinafter 'Yu').

(Continued on extra sheet)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 11/58669

Continuation of Box III - observations where unity of invention is lacking (first extra sheet):

Regarding Claim 1, Parreault teaches a method for predicting or determining the risk of a hematopoietic cell transplant (HCT) from a candidate donor to induce or not induce graft vs. host disease (GVHD) in an HCT recipient (para [0010]) comprising:

(a) measuring expression of one or more positive or negative GVHD predictor genes (para [0011]) in CD4+ T cells (para [0024]) or CD8+ T cells (para [0025]) from a candidate donor (para [0011]);

(b) obtaining an expression value from the positive or negative GVHD predictor genes based on the expression measured in (a) (para [0010]);

(c) comparing the expression value for the positive or negative GVHD predictor gene to a predefined expression value for the positive or negative GVHD predictor gene (para [0011]);

wherein an expression value for the positive GVHD predictor gene greater or less than the predefined reference expression value for the positive GVHD predictor gene indicates that the HCT from the candidate donor is at higher or lower risk, respectively, of inducing GVHD in an HCT recipient (para [0106]), or

wherein an expression value for the negative GVHD predictor gene greater or less than the predefined reference expression value for the negative GVHD predictor gene indicates that the HCT from the candidate donor is at lower or higher risk, respectively, of inducing GVHD in an HCT recipient (para [0105]); and

(d) predicting or determining the risk of the HCT from the candidate donor to induce or not induce GVHD in an HCT recipient, based on an evaluation of expression values of positive or negative GVHD predictor genes, that the HCT from the candidate donor is at higher or lower risk of inducing GVHD in an HCT recipient (para [0014]).

Parreault does not expressly disclose that said one or more positive or negative GVHD predictor gene in CD4+ T cells or CD8+ T cells from the candidate donor is selected from Tables 1 (RNA 1538), 2A (RNA 143), 2B (RNA 143), 3, 13 (SD175), 15 (SG128) or-18 (SG64) such as CD28 (listed in Table 1; pg 58 of the Current Application). However, Yu discloses that a donor CD4+ T cell or CD8+ T cell that expresses less CD28 than wild-type T cells induces less severe GVHD in a recipient (para [0030]-'CD28-deficient donor T cells induced less-severe GVHD than wild-type donor T cells', [0248]-[0251]), based on this disclosure, one of ordinary skill in the art would have found it obvious that said CD28 could be used as a positive GVHD predictor gene in the method disclosed by Parreault so as to expand the repertoire of relevant predictor GVHD genes in order to improve prognostic accuracy by employing multiple predictors and to provide one of ordinary skill in the art more options in choosing proper predictors.

As said shared technical features would have been obvious to one of ordinary skill in the art at the time of the invention, they cannot be considered special technical features that would otherwise unify the groups.

Groups I-IV therefore lack unity under PCT Rule 13 because they do not share a same or corresponding special technical feature.

フロントページの続き

(51)Int.Cl. F I テーマコード(参考)
 G 0 1 N 33/53 M
 G 0 1 N 37/00 1 0 2

(81)指定国 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, T
 J, 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, R
 O, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA,
 BB, BG, BH, 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, H
 U, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI
 , NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
 UZ, VC, VN

(特許庁注: 以下のものは登録商標)

1 . F L A S H

(74)代理人 100125195

弁理士 尾畑 雄一

(72)発明者 ソモギ, ロランド

カナダ国 オンタリオ州 ケー0エイチ 2ティー0 シデナム, アールアール1ピーオー クリ
 フサイド レーン 1090

(72)発明者 グレラー, ラリー, デイビッド

アメリカ合衆国 メリーランド州 21403-3168 アナポリス 42号, アメリカーナ
 ドライブ 670

Fターム(参考) 4B024 AA11 CA04 CA09 HA12

4B029 AA07 BB20 CC01 CC03 FA15

4B063 QA19 QQ43 QQ53 QR32 QR56 QR62 QS25 QS34 QX02

(54)【発明の名称】造血細胞移植のドナーをスクリーニング、予想、特定する方法、キット及びアレイ、並びに、造
 血細胞移植(HCT)が移植片対宿主病(GVHD)を引き起こすリスクを予測する方法、キッ
 ト及びアレイ

专利名称(译)	用于筛选, 预测和预测造血细胞移植供体的方法, 试剂盒和阵列, 以及用于预测造血细胞移植 (HCT) 引起移植物抗宿主病 (GVHD) 的风险的方法, 试剂盒和阵列		
公开(公告)号	JP2014501098A	公开(公告)日	2014-01-20
申请号	JP2013536914	申请日	2011-12-22
[标]申请(专利权)人(译)	PB迪生物诊断有限责任公司 PBD生物诊断		
申请(专利权)人(译)	PB迪生物诊断, LLC		
[标]发明人	ソモギロランド グレラーラリーデイビッド		
发明人	ソモギ,ロランド グレラー,ラリー,デイビッド		
IPC分类号	C12Q1/68 C12N15/09 C12M1/00 G01N33/53 G01N37/00		
FI分类号	C12Q1/68.ZNA.A C12N15/00.A C12N15/00.F C12M1/00.A G01N33/53.D G01N33/53.M G01N37/00.102		
F-TERM分类号	4B024/AA11 4B024/CA04 4B024/CA09 4B024/HA12 4B029/AA07 4B029/BB20 4B029/CC01 4B029/CC03 4B029/FA15 4B063/QA19 4B063/QQ43 4B063/QQ53 4B063/QR32 4B063/QR56 4B063/QR62 4B063/QS25 4B063/QS34 4B063/QX02		
优先权	61/498965 2011-06-20 US		
外部链接	Espacenet		

摘要(译)

本发明涉及预测或测定来自供体的造血细胞移植 (HCT) 在HCT受体中诱导移植物抗宿主病 (GVHD) 的风险;根据在HCT接受者中诱导GVHD的风险从候选供者中分类HCT;以及根据在HCT接受者中诱导GVHD的风险对一个或多个候选供者的HCT进行分类或评分的组织构造 (例如数据库) 和产生组织构造 (例如, 数据库) 的方法。本发明还涉及用于预测或确定来自候选供体以诱导HCT接受者中的GVHD的HCT风险, 以及根据在HCT接受者中诱导GVHD的风险对这样的供体进行分类或评分的试剂盒和阵列。

$$t = f(s, c_1, c_2, c_3) = \frac{1}{c_1} \operatorname{arcsinh} \left(\frac{c_1 s + c_2}{\sqrt{c_3}} \right) \text{ and where}$$