

(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2013-513387

(P2013-513387A)

(43) 公表日 平成25年4月22日(2013.4.22)

(51) Int.Cl.	F I	テーマコード (参考)
C 1 2 Q 1/68 (2006.01)	C 1 2 Q 1/68 A	2 G 0 4 5
G 0 1 N 33/68 (2006.01)	G 0 1 N 33/68	4 B 0 2 4
C 1 2 N 15/113 (2010.01)	C 1 2 N 15/00 G	4 B 0 6 3

審査請求 未請求 予備審査請求 未請求 (全 96 頁)

(21) 出願番号	特願2012-543298 (P2012-543298)	(71) 出願人	512151539
(86) (22) 出願日	平成22年12月9日 (2010.12.9)		アヴィール インコーポレイテッド
(85) 翻訳文提出日	平成24年8月7日 (2012.8.7)		アメリカ合衆国 94303 カリフォル
(86) 国際出願番号	PCT/US2010/059781		ニア州 パロアルト フェイバープレイス
(87) 国際公開番号	W02011/072177		2463
(87) 国際公開日	平成23年6月16日 (2011.6.16)	(74) 代理人	100095407
(31) 優先権主張番号	61/285,121		弁理士 木村 満
(32) 優先日	平成21年12月9日 (2009.12.9)	(74) 代理人	100109449
(33) 優先権主張国	米国 (US)		弁理士 毛受 隆典
		(74) 代理人	100132883
			弁理士 森川 泰司
		(74) 代理人	100123618
			弁理士 雨宮 康仁
		(74) 代理人	100148633
			弁理士 桜田 圭

最終頁に続く

(54) 【発明の名称】 循環器疾患の診断と分類のためのバイオマーカーアッセイ

(57) 【要約】

本開示の方法、アッセイおよびキットは、ヒトの心臓血管の健康を評価するため、バイオマーカー、特に *miRNA* および / またはタンパク質バイオマーカーを同定する。特定の実施形態において、方法、アッセイおよびキットは、ヒトの心臓血管の健康を評価するために、血液循環する *miRNA* および / またはタンパク質バイオマーカーを識別する。

【選択図】なし

【特許請求の範囲】**【請求項 1】**

ヒトの心臓血管の健康を評価するための方法であって、

a) ヒトから生物学的サンプルを得ること、

b) 前記生物学的サンプル中の、表 20 に記載の miRNA の中から選択される少なくとも 2 つの miRNA マーカーのレベルを決定すること、

c) 各 miRNA マーカーのレベルを有するデータセットを得ること、

d) 前記生物学的サンプルを分類するために前記データを使用する分析・分類プロセスに、前記データを入力することであって、ここで、分類は、アテローム性動脈硬化心血管疾患の分類、健康の分類、薬剤曝露の分類、薬剤曝露無しの分類から成る群から選択されること、及び、

e) ステップ (d) の分類に基づいてヒトのための治療計画を決定することを含み、

当該ヒトの心臓血管の健康が評価されること、

を特徴とする、方法。

【請求項 2】

前記少なくとも 2 つの miRNA マーカーは、miR - 378、miR - 497、miR - 21、miR - 15b、miR - 99a、miR - 29a、miR - 24、miR - 30b、miR - 29c、miR - 331.3p、miR - 19a、miR - 22、miR - 126、let - 7b、miR - 502.3、および miR - 652 から成る群から選択される、請求項 1 に記載の方法。

【請求項 3】

前記少なくとも 2 つの miRNA のマーカーは、miR - 378、miR - 497、miR - 21、miR - 15b、miR - 99a、および miR - 652 から成る群から選択される、請求項 2 に記載の方法。

【請求項 4】

前記アテローム性動脈硬化心血管疾患の分類は、冠動脈疾患、心筋梗塞 (MI)、不安定狭心症から成る群から選択される、請求項 1 に記載の方法。

【請求項 5】

さらに、アテローム性動脈硬化症の診断、アテローム性動脈硬化症のステージング、アテローム性動脈硬化症の予後、血管の炎症レベル、アテローム性動脈硬化症の進行の程度を、決定すること、治療反応をモニタリングすること、冠動脈カルシウムスコアを予測すること、アテローム性動脈硬化症症状の安定性のものと不安定なものとを区別すること、および、それらの組み合わせのために分類を使用することを特徴とする、請求項 1 に記載の方法。

【請求項 6】

前記データセットはさらに、1 つ以上の臨床的しるし (clinical indicia) のためのデータを含む、請求項 1 に記載の方法。

【請求項 7】

前記 1 個以上の臨床的しるしは、年齢、性別、LDL 濃度、HDL 濃度、トリグリセリド濃度、血圧、ボディマス指数、CRP 濃度、冠動脈カルシウムスコア、ウエスト周囲径、喫煙状況、心血管疾患の既往歴、心血管疾患の家族歴、心拍数、空腹時インスリン濃度、空腹時グルコース濃度、糖尿病の状態、高血圧薬の使用、およびそれらの組合せ、から成る群から選択される、請求項 6 に記載の方法。

【請求項 8】

前記選択された臨床的しるしは、年齢、性別、糖尿病、および M1 の家族歴である、請求項 7 に記載の方法。

【請求項 9】

前記生物学的サンプルは、血液、血清、血漿、唾液、尿、汗、母乳、及びそれらの組合せを含む、請求項 1 に記載の方法。

【請求項 10】

さらに、前記生物学的サンプル中の少なくとも一つのタンパク質バイオマーカーのレベルを決定することを含む、請求項 1 に記載の方法。

【請求項 1 1】

前記少なくとも一つのタンパク質バイオマーカーは、IL - 16、sFAS、Fasリガンド、MCP - 3、HGF、CTACK、エオタキシン (EOTAXIN)、アディポネクチン (adiponectin)、IL - 18、TIMP . 4、TIMP . 1、CRP、VEGFおよびEGFから成る群から選択される、請求項 10 に記載の方法。

【請求項 1 2】

前記少なくとも一つのタンパク質バイオマーカーは、IL - 16、エオタキシン、Fasリガンド、CTACK、MCP - 3、HGF、およびsFASから成る群から選択される、請求項 1 1 に記載の方法。

10

【請求項 1 3】

3 またはそれ以上の蛋白質バイオマーカーのレベルが決定される、請求項 1 1 に記載の方法。

【請求項 1 4】

前記分析・分類プロセスは、予測モデルの使用を含む、請求項 1 に記載の方法

【請求項 1 5】

前記分析・分類プロセスは、得られた前記データセットを参照データセットと比較することを含む、請求項 1 に記載の方法。

【請求項 1 6】

前記予測モデルは、分類のために少なくとも 0 . 68 の少なくとも一つの品質メトリック (metric) を含む、請求項 1 3 に記載の方法。

20

【請求項 1 7】

前記品質メトリックは、AUCと精度から選択される、請求項 1 5 に記載の方法。

【請求項 1 8】

前記分析・分類プロセスは、線形判別分析モデル、サポートベクターマシン分類 (support vector machine classification) アルゴリズム、帰納的な特徴除去 (recursive feature elimination) モデル、マイクロアレイモデルの予測分析、ロジスティック回帰モデル、CARTアルゴリズム、フレックス・ツリー・アルゴリズム、LARTアルゴリズム、ランダムフォレスト・アルゴリズム、MARTアルゴリズム、機械学習アルゴリズム、ペナルティ付き回帰法 (a penalized regression method)、およびそれらの組み合わせから成る群から選択される 1 以上のものの使用を含む、請求項 1 に記載の方法。

30

【請求項 1 9】

前記分析・分類プロセスは、少なくとも 0 . 68 の品質メトリックを提供するように選択されるターム (terms) を含む、請求項 1 8 に記載の方法。

【請求項 2 0】

前記分析・分類プロセスは、0 . 70 の品質メトリックを提供するように選択されるタームを含む、請求項 1 8 に記載の方法。

40

【請求項 2 1】

前記分析・分類プロセスは、分類のための少なくとも 0 . 70 の少なくとも一つの品質メトリックを含む、請求項 1 8 に記載の方法。

【請求項 2 2】

前記治療計画は、追加検査、薬理的介入、無治療、及びそれらの組合せからなる群から選択される 1 つまたは複数を含む、請求項 1 に記載の方法。

【請求項 2 3】

ヒトの心臓血管の健康を評価するための方法であって、

a) ヒトから生物学的サンプルを得ること、

b) 前記生物学的サンプル中の、IL - 16、sFas、Fasリガンド、MCP - 3

50

、HGF、CTACK、エオタキシン(EOTAXIN)、アディポネクチン(adiponectin)、IL-18、TIMP-4、TIMP-1、CRP、VEGF、およびEGFから成る群から選択される少なくとも3つのタンパク質マーカーのレベルを決定すること、

c) 各タンパク質マーカーのレベルを有するデータセットを得ること、

d) 生物学的サンプルを分類するために前記データを使用する分析・分類プロセスに、前記データを入力することであって、ここで、分類は、アテローム性動脈硬化症心血管疾患分類、健康分類、薬剤曝露分類、薬剤非曝露分類から成る群から選択されること、及び、

e) ステップ(d)の分類に基づいてヒトのための治療計画を決定することを含み、当該ヒトの心臓血管の健康が評価されること、
を特徴とする、方法。

10

【請求項24】

前記少なくとも3つのタンパク質マーカーは、IL-16、エオタキシン(EOTAXIN)、Fasリガンド、CTACK、MCP-3、HGF、およびsFasからなる群から選択される、請求項23に記載の方法。

【請求項25】

前記データセットは、さらに、年齢、性別、LDL濃度、HDL濃度、トリグリセリド濃度、血圧、ボディマス指数、CRP濃度、冠動脈カルシウムスコア、ウエスト周囲径、喫煙状況、心血管疾患の既往歴、心血管疾患の家族歴、心拍数、空腹時インスリン濃度、空腹時グルコース濃度、糖尿病の状態、高血圧薬の使用、およびそれらの組合せから成る群から選ばれる1以上の臨床的しるしのためのデータを含む、請求項23に記載の方法。

20

【請求項26】

治療計画の必要性または有効性を決定するためにヒトの心臓血管の健康を評価するための方法であって、

ヒトから生物学的サンプルを得ること、

前記生物学的サンプルでの、表20に記載のmiRNAの中から選択される少なくとも2つのmiRNAマーカーのレベルを決定すること、

前記生物学的サンプルでの、IL-16、sFas、Fasリガンド、MCP-3、HGF、CTACK、エオタキシン(EOTAXIN)、アディポネクチン(adiponectin)、IL-18、TIMP-4、TIMP-1、CRP、VEGF、およびEGFから成る群から選択される少なくとも3つのタンパク質マーカーのレベルを決定すること、

30

個々のmiRNAマーカーおよびタンパク質マーカーのレベルを構成するデータセットを得ること、

前記生物学的サンプルを分類するために前記データを使用する分析・分類プロセスに、前記データを入力することであって、ここで、該分類は、アテローム性動脈硬化症心血管疾患分類、健康分類、薬剤曝露分類、薬剤非曝露分類から成る群から選択されること、および、

前記分類プロセスの出力に応じて前記生物学的サンプルを分類し、前記分類に基づいてヒトのための治療計画を決定すること、

40

を特徴とする、方法。

【請求項27】

前記miRNAマーカーは、miR-378、miR-497、miR-21、miR-15b、miR-99a、miR-29a、miR-24、miR-30b、miR-29c、miR-331.3p、miR-19a、miR-22、miR-126、let-7b、miR-502.3、およびmiR-652から成る群から選択される、請求項26に記載の方法。

【請求項28】

前記タンパク質バイオマーカーは、IL-16、エオタキシン、Fasリガンド、CT

50

A C K、M C P - 3、H G F、および s F A S から成る群から選択される、請求項 26 に記載の方法。

【請求項 29】

治療計画の必要性または有効性を決定するためにヒトの心臓血管の健康を評価するためのキットであって、

生物学的サンプル中の表 20 に記載の m i R N A の中から選択される少なくとも 2 つの m i R N A マーカーのレベルを決定するためのアッセイと、

使用説明書であって、

m i R N A マーカーの個々のレベルで構成されるデータセットを取得すること、

前記生物学的サンプルを分類するために前記データを使用する、分析・分類プロセスに前記データを入力することであって、前記分類は、アテローム性動脈硬化心血管疾患の分類、健康の分類、薬剤曝露の分類、薬剤曝露無しの分類から成る群から選択されること、

前記分類プロセスの出力に応じて前記生物学的サンプルを分類すること、および、

前記分類に基づいてヒトのための治療計画を決定すること、のための使用説明書と、を含む、キット。

【請求項 30】

さらに、前記生物学的サンプルでの、I L - 16、s F a s、F a s リガンド、M C P - 3、H G F、C T A C K、エオタキシン (E O T A X I N)、アディポネクチン (a d i p o n e c t i n)、I L - 18、T I M P . 4、T I M P . 1、C R P、V E G F、および E G F から成る群から選択される少なくとも 3 つのタンパク質マーカーのレベルを決定するためのアッセイと、

使用説明書であって、

前記タンパク質マーカーの個々のレベルで構成されるデータセットを取得すること、

前記生物学的サンプルを分類するために前記データを使用する、分析・分類プロセスに前記 m i R N A と前記タンパク質マーカーの前記データを入力することであって、前記分類は、アテローム性動脈硬化症心血管疾患分類、健康分類、薬剤曝露分類、薬剤非曝露分類から成る群から選択されること、

前記分類プロセスの出力に応じて前記生物学的サンプルを分類すること、および、

前記分類に基づいてヒトのための治療計画を決定すること、のための使用説明書と、を含む、請求項 29 に記載のキット。

【請求項 31】

ヒトの心血管イベントのリスクを評価するための方法であって、

a) ヒトから生物学的サンプルを得ること、

b) 前記生物学的サンプル中の、表 20 に記載の m i R N A の中から選択される少なくとも 2 つの m i R N A マーカーのレベルを決定すること、

c) 各 m i R N A マーカーのレベルを構成するデータセットを得ること、

d) 前記データセットに基づいて心血管イベントのリスクを決定するために、リスク予測分析プロセスに前記データを入力すること、及び、

e) ステップ (d) の心血管イベントの予測リスクに基づいてヒトのための治療計画を決定することを含み、

当該ヒトの心臓血管イベントが評価されること、

を特徴とする、方法。

【請求項 32】

前記心血管イベントのリスクは、サンプルが得られた日から、約 1 年、約 2 年、約 3 年、約 4 年、および約 5 年から成る群から選択された期間に対して決定される、請求項 31 に記載の方法。

【請求項 33】

前記生物学的サンプルでの、3 つ以上のタンパク質バイオマーカーのレベルを決定することを、さらに含む、請求項 31 に記載の方法。

【請求項 34】

10

20

30

40

50

前記 3 つ以上のタンパク質バイオマーカーは、IL - 16、sFas、Fas リガンド、MCP - 3、HGF、CTACK、エオタキシン (EOTAXIN)、アディポネクチン (adiponectin)、IL - 18、TIMP . 4、TIMP . 1、CRP、VEGF、および EGF から成る群から選択される、請求項 33 に記載の方法。

【請求項 35】

前記 3 つ以上のタンパク質バイオマーカーは、IL - 16、エオタキシン、Fas リガンド、CTACK、MCP - 3、HGF、および sFAS から成る群から選択される、請求項 34 に記載の方法。

【請求項 36】

ヒトの心臓血管イベントのリスクを評価するための方法であって、

10

a) ヒトから生物学的サンプルを得ること、

b) 前記生物学的サンプルでの、IL - 16、sFas、Fas リガンド、MCP - 3、HGF、CTACK、エオタキシン (EOTAXIN)、アディポネクチン (adiponectin)、IL - 18、TIMP . 4、TIMP . 1、CRP、VEGF、および EGF から成る群から選択される 1 以上のタンパク質マーカーのレベルを決定すること、

c) 各タンパク質マーカーのレベルを構成するデータセットを得ること、

d) 前記データセットに基づいて心血管イベントのリスクを決定するために、リスク予測分析プロセスに前記データを入力こと、及び、

e) ステップ (d) の心血管イベントの予測リスクに基づいて、ヒトのための治療計画を決定することを含み、

20

当該ヒトの心血管イベントのリスクが評価されること、

を特徴とする、方法。

【請求項 37】

心血管イベントのリスクは、サンプルが得られた日から、約 1 年、約 2 年、約 3 年、約 4 年、および約 5 年から成る群から選択された期間に対して決定される、請求項 36 に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

30

(関連出願のためのクロスリファレンス)

この出願では、2009 年 12 月 9 日に提出された、米国仮特許出願第 61 / 285 , 121 号を基礎に優先権を主張しており、その全体が参照により本明細書に組み込まれている。

【背景技術】

【0002】

アテローム性動脈硬化症 (ASCVD) は、世界的に、罹患率と死亡率の主な原因である。心筋梗塞 (MI) の約 60% は、危険因子が 0 または 1 の人に発生する。すなわち、心臓イベントが発生した人々の大多数は、現在の方法で評価した場合、低めの中間または中間のリスクのカテゴリにある。

40

【発明の概要】

【発明が解決しようとする課題】

【0003】

遺伝的要因と環境的要因の組み合わせは、疾患の開始および進行に責任がある。アテローム性動脈硬化症はしばしば無症候性であり、現在の診断法によって検出されない。実際、多くの場合、アテローム性動脈硬化症の最初の症状は、心臓発作や心臓突然死である。

【0004】

循環器疾患やその進展を、正確に予測し、診断することができるアッセイや方法は、非常に望ましい。

【課題を解決するための手段】

50

【0005】

この開示は、人間の心臓血管系の健康を評価するための方法、アッセイおよびキットを提供する。一実施形態では、以下の工程を含む、人間の心臓血管系の健康を評価するための方法が提供されている。 a) 人間から生物学的サンプルを得ること； b) 前記生物学的サンプルで、表20に記載されているmiRNAの中から選択される少なくとも2つのmiRNAマーカーのレベルを決定すること； c) 各miRNAマーカーのレベルを有するデータセットを得ること； d) 前記生物学的サンプルを分類するためにデータを使用する分析・分類プロセスに、前記データを入力すること； ここで該分類は、アテローム性動脈硬化症、健康、薬剤曝露、薬剤非曝露からなる分類群から選択されるものであり、そして、 e) ステップ(d)の分類に基づいて、人間のための治療法を決定することであって、ここでヒトの心臓血管系の健康を評価すること。

10

【0006】

ヒトの心臓血管系の健康を評価するための方法であって、 a) ヒトから生物学的サンプルを得ること； b) 前記生物学的サンプル中の、IL-16、sFAS、Fasリガンド、MCP-3、HGF、CTACK、エオタキシン(EO TAXIN)、アディポネクチン、IL-18、TIMP-4、TIMP-1、CRP、VEGF、及びEGFから成る群から選択される少なくとも3つのタンパク質バイオマーカーのレベルを決定すること； c) 各タンパク質マーカーのレベルを有するデータセットを得ること； d) 前記生物学的サンプルを分類するために前記データを使用する分析・分類プロセスに、前記データを入力すること； ここで該分類は、アテローム性動脈硬化症分類、健康分類、薬剤曝露分類、薬剤非曝露分類からなる分類群から選択されること、および、 e) ステップ(d)の分類に基づいて、ヒトのための治療計画を決定することであって、ヒトの心臓血管系の健康を評価すること、を特徴とする、方法。

20

【0007】

治療計画の必要性や有効性を判断するためにヒトの心臓血管系の健康を評価するための方法であって、以下を含む。ヒトから生物学的サンプルを得ること、前記生物学的サンプル中の、表20に記載されているmiRNAの中から選ばれる少なくとも2つのmiRNAのマーカーのレベルを決定することと、前記生物学的サンプル中の、IL-16、sFAS、Fasリガンド、MCP-3、HGF、CTACK、エオタキシン(EO TAXIN)、アディポネクチン、IL-18、TIMP-4、TIMP-1、CRP、VEGF、及びEGFから成る群から選択される少なくとも3つのタンパク質バイオマーカーのレベルを決定すること、各miRNAマーカーとタンパク質バイオマーカーのレベルを有するデータセットを取得すること、前記生物学的サンプルを分類するために前記データを使用して、分析・分類プロセスに前記データを入力することであって、ここで該分類は、アテローム性動脈硬化症分類、健康分類、薬剤曝露分類、薬剤非曝露分類からなる分類群から選択されること、および、前記分類プロセスの出力に応じて前記生物学的サンプルを分類し、前記分類に基づいてヒトのための治療計画を決定すること。

30

【0008】

さらに別の実施形態では、治療計画の必要性や有効性を決定するためにヒトの心臓血管系の健康を評価するためのキットが提供されている。キットは以下を含む：生物学的サンプル中の表20に記載されているmiRNAの中から選ばれる少なくとも2つのmiRNAマーカーのレベルを決定するためのアッセイと、及び/又は、生物学的サンプル中の、IL-16、sFAS、Fasリガンド、MCP-3、HGF、CTACK、エオタキシン、アディポネクチン、IL-18、TIMP-4、TIMP-1、CRP、VEGF及びEGFからなる群から選択される少なくとも3つのタンパク質マーカーのレベルを決定するためのアッセイと、以下のための使用説明書とを含み、使用説明書は、(1)各miRNA及び/又はタンパク質マーカーのレベルを有するデータセットを取得すること、(2)生物学的サンプルを分類するためのデータを使用して、分析・分類プロセスにデータを入力することであって、前記分類は、アテローム性動脈硬化症心血管疾患分類、健康分類、薬剤曝露分類、薬剤非曝露分類から成る群から選択されること、(3)前記分類に基

40

50

づいてヒトのための治療計画を決定すること、のためのものである。

【0009】

さらに別の実施形態は、ヒトの心血管イベントのリスクを評価するための方法であって、a) ヒトから生物学的サンプルを得ること、b) 前記サンプル中の、IL-16、sFAS、Fasリガンド、MCP-3、HGF、CTACK、エオタキシン、アディポネクチン、IL-18、TIMP-4、TIMP-1、CRP、VEGFおよびEGFからなる群から選ばれる3以上のタンパク質バイオマーカーのレベル、及び/又は、表20の中の2つ以上のmiRNAのレベルを決定すること、c) 各タンパク質バイオマーカー及び/又はmiRNAバイオマーカーのレベルを有するデータセットを得ること、d) 前記データセットに基づいて心血管イベントのリスクを決定するために、リスク予測分析プロセスにデータを入力すること、e) ステップ(d)の心血管イベントの予測リスクに基づいてヒトのための治療計画を決定することであって、ここでは、ヒトの心血管イベントのリスクが評価される、方法、である。

10

【図面の簡単な説明】

【0010】

【図1】ロジスティック回帰分析アプローチに基づいた52検体セット(26症例(cases)と26のコントロール)の予想される分類性能を示すグラフである。予想されるAUCとそれに対応する95%信頼区間は、52の個々人またはプールされたサンプルのいずれかのセット分類の500のシミュレーションから得られた。破線のエラーバー上の白丸は、対数正規分布に従うと仮定されたバイオマーカー濃度やスコア値と共に、期待値と、プールされたサンプル(各プール内の5サンプル)を用いた信頼区間を表している。実線のエラーバー上の白丸は、期待値と同じ分布から個々の試料を用いて信頼区間を表している。実線の黒い点は、理論的な結果を表している。x軸は、症例とコントロールのバイオマーカーやスコア分布の平均の差を表している。

20

【0011】

【図2】ロジスティック回帰分析アプローチに基づいた52検体セット(26症例と26コントロール)の予想される分類性能を示すグラフである。予想されるAUCとそれに対応する95%信頼区間は、52の個々人またはプールされたサンプルのいずれかのセット分類の500のシミュレーションから得られた。破線のエラーバー上の白丸は、「正規分布」に従うと仮定されたバイオマーカー濃度やスコア値と共に、期待値と、プールされたサンプル(各プール内の5サンプル)を用いた信頼区間を表している。実線のエラーバー上の白丸は、期待値と同じ分布から個々の試料を用いて信頼区間を表している。実線の黒い点は、理論的な結果を表している。x軸は、症例とコントロールのバイオマーカーやスコア分布の平均の差を表している。

30

【0012】

【図3】44のmiRのセットから共変量を選択するモデルに基づいてプールされたサンプルの分類のためのAUC値の分布のグラフである。AUC値の計算は、ペナルティ付きロジスティック回帰モデル(L1ペナルティ付き)のデータへの適合を介して、100の事前検証済み(prevalidated)分類スコアベクトルを得ることに基づいている。X軸はAUCを表し、Y軸は頻度を表している。示されているように、平均AUCは0.68である。

40

【0013】

【図4】44のmiRのセットから共変量を選択するモデルに基づいた、個々のサンプル分類のためのAUC値分布のグラフである。AUC値の計算は、ペナルティ付きロジスティック回帰モデル(L1ペナルティ付き)のデータへの適合を介して、100の事前検証済み(prevalidated)分類スコアベクトルを得ることに基づいている。示されているように、平均AUCは0.78である。

【0014】

【図5】44のmiRのセットと47のタンパク質バイオマーカーから共変量を選択するモデルに基づいた、個々のサンプル分類のためのAUC値分布のグラフである。AUC値

50

の計算は、ペナルティ付きロジスティック回帰モデル (L 1 ペナルティ付き) のデータへの適合を介して、100の事前検証済み (p r e v a l i d a t e d) 分類スコアベクトルを得ることに基づいている。示されているように、平均AUCは0.75である。

【0015】

【図6】垂直線で示された、最も高い負の相関と、最も高い正の相関を含めた、miRとタンパク質との相関分布を示すグラフである。

【0016】

【図7】miR s 単独での相関の分布を示すグラフである。

【0017】

【図8】単独の、タンパク質バイオマーカーのデータに基づいて算出され、事前検証済みのスコア (500リポート) に基づいたAUCの分布を示すグラフである。

10

【0018】

【図9】コントロールの平均値と標準偏差に正規化されたタンパク質バイオマーカーの単変量ハザード比を示すグラフである。

【0019】

【図10】タンパク質バイオマーカーの調整ハザード比 (a d j u s t e d h a z a r d r a t i o 、 H R) を示すグラフである。調整は、従来の危険因子 (T R F s) である、年齢、性別、収縮期血圧 (B P) 、拡張期血圧、コレステロール、高比重リポ蛋白 (H D L) 、高血圧症、高血圧薬の使用、高脂血症、糖尿病、喫煙状態、に基づいて行った。

20

【0020】

【図11A】最も高い時間依存性AUCを有するマーカー群と、5年までのフォローアップ対応値を示すグラフである。s F A S 、 N T . p r o B N P 、 M I G 、 I L . 1 6 、 M I G 、 及び A N G 2 に対するAUCが示されている。

【図11B】最も高い時間依存性AUCを有するマーカー群と、5年までのフォローアップ対応値を示すグラフである。F a s リガンド、S C D 4 0 L 、アディポネクチン、M C P . 3 、レプチン及びR A N T E S に対するAUCが示されている。

【0021】

【図12】C o x 比例ハザード回帰モデルの用語の数の関数として、絶対値とドロップ・イン・逸脱 (d r o p - i n - d e v i a n c e) の標準誤差を示すグラフである。モデルに含まれるマーカーの最適な数は、「1 - 標準誤差ルール」を使用して、選択されている。

30

【0022】

【図13A】コントロールの、M a r s h f i e l d サンプルセットでの4 C o x P H モデルから得られた線形予測のカーネル密度推定を示すグラフである。

【図13B】症例の、M a r s h f i e l d サンプルセットでの4 C o x P H モデルから得られた線形予測のカーネル密度推定を示すグラフである。

【0023】

【図14A】コントロールの、M E S A サンプルセットでの4 C o x P H モデルから得られた線形予測のカーネル密度推定を示すグラフである。

40

【図14B】症例の、M E S A サンプルセットでの4 C o x P H モデルから得られた線形予測のカーネル密度推定を示すグラフである。

【発明を実施するための形態】

【0024】

本開示は、ヒトの心臓血管系の健康を評価するための方法、アッセイおよびキットを提供し、特に、ヒトでのアテローム性動脈硬化症 (A S C V D) を、予測、診断、および監視するために、提供する。開示される方法、アッセイおよびキットは、ヒトの心臓血管の健康を評価するための、循環するマイクロリボ核酸 (m i R N A) バイオマーカー及び/又は蛋白質バイオマーカーを、識別 (同定) する。これら方法、アッセイおよびキットの特定の実施例では、循環するm i R N A 及び/又はタンパク質バイオマーカーは、ヒトの

50

心臓血管系の健康を評価するために識別される。

【0025】

一実施形態において、本開示は、治療計画のための必要性、またはその有効性を決定するためにヒトの心臓血管系の健康を評価するための方法を提供し、それは、以下の、ヒトから生物試料を得ること、前記生物試料で、表20に記載されている群から選択された少なくとも2つのmiRNAマーカーのレベルを決定すること、各miRNAマーカーのレベルを含むデータセットを得ること、前記生物試料を分類するためにデータを使用する分析・分類プロセスに、前記データを入力すること、ここで該分類は、アテローム性動脈硬化症分類、健康分類、薬剤曝露分類、薬剤非曝露分類からなる分類群から選択されるものであって、分類プロセスの出力に応じて前記生物学的サンプルを分類すること、および、前記分類に基づいて、ヒトのための治療計画を決定すること、からなる。

10

【0026】

特定の実施形態では、治療計画の必要性、または有効性を判断するために、以下(の工程)、つまり、ヒトから生物学的サンプルを得ること(工程)、前記生物学的サンプル中の、IL-16、sFAS、Fasリガンド、MCP-3、HGF、CTACK、エオタキシン(EOTAXIN)、アディポネクチン、IL-18、TIMP-4、TIMP-1、CRP、VEGF、及びEGFから成る群から選択される少なくとも3つのタンパク質マーカーのレベルを決定すること、各タンパク質バイオマーカーのレベルを含むデータセットを得ること、前記生物学的サンプルを分類するために前記データを使用する分析・分類プロセスに、前記データを入力すること；ここで該分類は、アテローム性動脈硬化症分類、健康分類、薬剤曝露分類、薬剤非曝露分類からなる分類群から選択されること、および、前記分類プロセスの出力に応じて前記生物学的サンプルを分類し、前記分類に基づいて人間のための治療計画を決定すること、を含む、ヒトの心臓血管系の健康を評価するための方法が、開示されている。

20

【0027】

別の実施形態においては、ヒトの心臓血管系の健康を評価するための方法が、提供されている。特定の実施形態においては、前記評価は、治療計画の必要性や有効性を決定するために使用することができる。当該方法は、ヒトから生物学的サンプルを得ること、前記生物学的サンプルで、表20に記載されているmiRNAの中から選択される少なくとも2つのmiRNAマーカーのレベルを決定すること、前記生物学的サンプル中の、IL-16、sFAS、Fasリガンド、MCP-3、HGF、CTACK、エオタキシン(EOTAXIN)、アディポネクチン、IL-18、TIMP-4、TIMP-1、CRP、VEGF、及びEGFから成る群から選択される少なくとも3つのタンパク質バイオマーカーのレベルを決定すること、miRNAマーカーとタンパク質バイオマーカーの個々のレベルで構成されるデータセットを取得すること、前記生物学的サンプルを分類するために前記データを使用して分析・分類プロセスに前記データを入力することであって、ここで該分類は、アテローム性動脈硬化症分類、健康分類、薬剤曝露分類、薬剤非曝露分類からなる分類群から選択されること、および、前記分類プロセスの出力に応じて前記生物学的サンプルを分類し、前記分類に基づいてヒトのための治療計画を決定すること、である。

30

40

【0028】

さらに別の実施形態では、ヒトの心血管系イベントのリスクを評価するための方法を提供する。この方法は、以下を含む。つまり、ヒトから生物学的サンプルを得ること、当該サンプル中の、(1)IL-16、sFAS、Fasリガンド、MCP-3、HGF、CTACK、エオタキシン(EOTAXIN)、アディポネクチン、IL-18、TIMP-4、TIMP-1、CRP、VEGF、及びEGFから成る群から選択される3つまたはそれ以上のタンパク質バイオマーカーのレベル、及び又は、(2)表20のmiRNAの中から選ばれる2つまたはそれ以上のmiRNAのマーカーのレベル、を決定すること。当該方法では、各タンパク質バイオマーカー及び又は各miRNAバイオマーカーのレベルを含むデータセットが得られる。前記データセットに基づいて心血管イベントの

50

スクを予測するリスク予測分析プロセスへデータが入力され、この心血管イベントの予測リスクに基づいて、ヒトのための治療計画が決定される。心血管系イベントのリスクは、サンプルが得られ、及び又は、分析された日から、約1年、約2年、約3年、約4年、約5年、又はそれ以上の期間、予測することができる。当該予測された心血管系イベントは、後述するように、アテローム性動脈硬化疾患の進展、MI（心筋梗塞、myocardial infarction）等を予測できる。

【0029】

用語「マーカー」と「バイオマーカー」は、本開示を通して同義語として使われている。

【0030】

本開示の方法では、検出され、そのレベルが決定されたmiRNAマーカーの数は、1つ、又は、2、3、4、5、6、7、8、9、10やそれ以上であり得る。特定の実施形態では、検出されたmiRNAのマーカーの数は、3、または5以上である。検出され、そのレベルが決定されているタンパク質バイオマーカー数は、1か、またはそれより多く、例えば2、3、4、5、6、7、8、9、10以上である。特定の実施形態では、1、2、3、または5以上のmiRNAマーカーが検出されて、そのレベルが決定され、そして、1、2、3、または5以上のタンパク質バイオマーカーが検出され、そのレベルが決定されている。

【0031】

本開示の方法は、アテローム性動脈硬化症の診断とモニタリングのために有用である。アテローム性動脈硬化症は、また、アテローム性動脈硬化症、動脈硬化症、アテローム性血管疾患、動脈閉塞性疾患、または心臓血管疾患として知られており、血管壁へのプラーク蓄積や、血管の炎症によって特徴づけられる。血管の炎症は、アクティブなアテローム性動脈硬化症、不安定プラーク、または脆弱性プラークの特質である。プラークは、蓄積された、細胞内および細胞外脂質、平滑筋細胞、結合組織、炎症性細胞、およびグリコサミノグリカンから構成されている。特定のプラークはまた、カルシウムを含む。不安定な、アクティブな、または、脆弱なプラークは、炎症性細胞で濃縮される。

【0032】

例によって、本発明は、サンプルに関連付けられたデータセットであって、少なくともmiRNAのマーカー単独か、またはアテローム性動脈硬化症の予測因子として同定されているタンパク質バイオマーカーとの組み合わせによる定量的データを含むデータセットを取得すること、前記データセットを、アテローム性動脈硬化症の診断とモニタリングに有用な結果を生成するために前記データセットを使用する分析プロセスに前記データセットを入力することによって、アテローム性動脈硬化症の診断とモニタリングに有用な結果を生成するための方法を含む。この定量的データは、DNA、RNA、タンパク質の発現レベル、およびそれらの組み合わせを含めることができる。

【0033】

本開示の、方法、アッセイおよびキットはまた、心筋梗塞（MI）、急性冠症候群、脳卒中、心不全、および狭心症を含む心血管疾患の合併症の診断およびモニタリングのために有用である。一般的な合併症の例としては、通常、心筋のセグメントへの冠（冠状動脈）血流量の急激な減少に起因する虚血性心筋壊死に関連するMIである。急性MIの患者の大多数で、しばしばプラーク破裂に伴う急性血栓が、損傷部位を供給する動脈を閉塞する。プラークの破裂は、炎症性細胞が濃縮されたアテローム性動脈硬化プラークによって以前に部分的に遮られた動脈で一般的に発生する。一般的なアテローム性動脈硬化の合併症の他の例は、胸部の痛みや心臓への不十分な血流に起因する不快感の症状を伴う状態の、狭心症である。

【0034】

本開示は、与えられた個々人のために、採血から特定の期間内に、心血管イベント（例えば、MI）のリスク予測と同様に、アテローム性動脈硬化症の診断と分類に使用することができる炎症のバイオマーカーのプロファイルを識別する。本開示でアッセイされたm

10

20

30

40

50

miRNAおよび蛋白質バイオマーカーは、別のアテローム性動脈硬化症の分類、例えば、診断、ステージング、予後、モニタリング、治療反応、擬似冠動脈カルシウムスコアの予測を区別することが可能である学習アルゴリズムを用いて同定されるものである。臨床的しるし（例えば、従来の危険因子）のような、アテローム性動脈硬化症の分類をするために有用なその他のデータはまた、アテローム性動脈硬化症分類のための有用な結果を生成するために使用されるデータセットの一部であり得る。

【0035】

様々なmiRNAマーカーおよび蛋白質バイオマーカーの定量的なデータを含むデータセットは、単独または併用で、本明細書に開示され、他のデータセットコンポーネント（例えば、DNA、RNA、臨床的しるしの程度）の定量的データは、分析プロセスに入力され、結果を生成するために使用することができる。分析プロセスは、定義されたパラメータ、言い換えれば、予測モデルでの、任意の学習アルゴリズムのタイプであり得る。予測モデルは、学習アルゴリズムを参照または対照データの適切な型に適用することによって、様々なアテローム性動脈硬化症の分類やリスク予測のために開発することができる。分析プロセス/予測モデルの結果は、適切な個人が、適切な処置を取るために使用することができる。例えば、分類が、「健康」や「アテローム性動脈硬化症」である場合、その結果は、個々の治療の適切な臨床経路を決定するために使用することができる。

【0036】

マイクロRNA（本明細書では、また、miRNA、pRNA、miRと称す）は、約17~27ヌクレオチドの長さの一本鎖RNA分子の形態で、遺伝子発現を調節する。miRNAは、遺伝子によってコードされ、そのDNAから転写される。しかし、miRNAは、タンパク質に翻訳されず（すなわち、miRNAはノンコーディングRNAである）、代わりに、各一次転写産物（pri-miRNA）は、pre-miRNAと呼ばれる短いステムループ構造になり、最終的には機能的なmiRNAへとプロセスされる。

【0037】

ヒトの（心血管系の）炎症に関連し、心血管系の健康を評価するために有用なmiRNAマーカーは、以下に限定されはしないが、1つ以上の、miR-26a、miR-16、miR-222、miR-10b、miR-93、miR-192、miR-15a、miR-125-a.5p、miR-130a、miR-92a、miR-378、miR-20a、miR-20b、miR-107、miR-186、hsa.let.7f、miR-19a、miR-150、miR-106b、miR-30c、及び、let7bを含む。特定の実施形態では、miRNAマーカーは、1つ以上のmiR-26a、miR-16、miR-222、miR-10b、miR-93、miR-192、miR-15a、miR-125-a.5p、miR-130a、miR-92a、miR-378、及び、let7bを含む。特に、表20に記載されているmiRNAは、ヒトの心血管系の健康を評価するのに有用である。

【0038】

ヒトの（心血管系の）炎症に関連し、心血管系の健康を評価するために有用なタンパク質バイオマーカーは、以下に限定されないが、1つ以上のRANTES、TIMP1、MCP-1、MCP-2、MCP-3、MCP-4、エオタキシン（eotaxin）、IL-10、M-CSF、IL-3、TNFα、Ang-2、IL-5、IL-7、IGF-1、sVCAM、sICAM-1、E-selectin、P-selectin、インターロイキン-6、インターロイキン-18、クレアチンキナーゼ、LDL、oxLDL、LDL粒子サイズ、リポプロテイン(a)トロポニンI、トロポニンT、LpPLA2、CRP、HDL、トリグリセリド、インスリン、BNP、フラクタルキン、オステオポンチン、オステオプロテゲリン、オンコスタチン-M、ミエロペルオキシダーゼ、ADMA、PAI-1（プラスミノゲンアクチベーター阻害剤）、SAA（循環アミロイドA）、t-PA（組織型プラスミノゲン活性化因子）、sCD40リガンド、フィブリノーゲン、ホモシステイン、D-ダイマー、白血球数、心臓型脂肪酸結合タンパク質、MMP1、プラスミノゲン、葉酸、ビタミンB6、レブチン、可溶性トロンボモジュリン、P

10

20

30

40

50

APP A、MMP 9、MMP 2、VEGF、PIGF、HGF、vWF、及び、シスタチン C、を含む。特定の実施形態では、タンパク質バイオマーカーは、1つ以上の、IL - 16、sFas、Fasリガンド、MCP - 3、HGF、CTACK、エオタキシン (EOTAXIN)、アディポネクチン、IL - 18、TIMP . 4、TIMP . 1、CRP、VEGF、及び、EGFを含む。特定のバイオマーカーに加えて、本開示は、さらに、例示配列と、約90%、約95%、または約97%同一であるバイオマーカー変異体を含む。本明細書で使用される変異体は、多型、スプライシング (splice) 変異体、遺伝子変異などが含まれている。

【0039】

タンパク質バイオマーカーは、種々の方法で検出することができる。たとえば、in vivo イメージングは、心臓組織におけるアテローム性動脈硬化症関連タンパク質の存在を検出するために利用することができる。このような方法は、例えば、その様なタンパク質に特異的な標識抗体またはリガンドを利用することができる。これらの実施形態において、ポリペプチドに対して特異的である検出可能に標識された部分、例えば、抗体、リガンド等は、(例えば、注射によって) 個々人に投与され、そして、これらに標識された細胞は、限定はされないが、磁気共鳴イメージング、コンピュータ断層撮影スキャンなどを含む、標準的なイメージング技術を使用して、検出される。しかし、磁気共鳴イメージング、コンピュータ断層撮影スキャン、などが挙げられるが、これらに限定されない。検出には、イメージング試薬の1つか、またはカクテルを利用することができる。

【0040】

追加のマーカーは、以下に限定されないが、年齢、性別、LDL濃度、HDL濃度、トリグリセリド濃度、血圧、ボディマス指数 (体格指数)、CRP濃度、冠動脈カルシウムスコア、ウエスト周囲径、タバコ喫煙状況、心血管疾患の既往歴、心血管疾患の家族歴、心拍数、空腹時インスリン濃度、空腹時グルコース濃度、糖尿病の状態、および高血圧の薬の使用を含む、1つ以上の臨床的しるしから選択することができる。アテローム性動脈硬化症の分類を行うために有用な追加の臨床的しるしは、線形判別分析、サポートベクターマシン分類 (support vector machine classification)、帰納的な特徴除去 (recursive feature elimination)、マイクロアレイ予測分析、ロジスティック回帰、CART、フレックスツリー (FlexTree)、LART、ランダムフォレスト、MART及び/又は、生存分析回帰など、当技術分野で知られているこれらの学習アルゴリズムを用いて、同定することができる。これらは、当業者に知られており、さらに本明細書に記載されている。

【0041】

本明細書に開示された分析分類は、予測モデルの使用を含むことができる。予測モデルは、さらに、分類のために、少なくとも約0.68またはそれ以上の品質要求基準 (品質メトリック、a quality metric) を備えている。特定の実施形態では、分類のために、少なくとも約0.70またはそれ以上の品質要求基準を備えている。特定の実施形態では、品質要求基準 (品質メトリック) は、曲線 (AUC)、ハザード比 (HR)、相対リスク (RR)、再分類、陽性適中率 (PPV)、陰性適中率 (NPV)、精度、感度と特異性、ネット再分類指数 (Net reclassification Index)、臨床ネット分類インデックス (Clinical Net reclassification Index) 下の領域から、選択される。これらおよびその他のメトリックは、本明細書中に記載されるように使用することができる。さらに、様々な用語が、品質要求基準 (品質メトリック) を提供するように選択することができる。

【0042】

定量的データは、データセットの各コンポーネントに対して得られ、以前に定義されたパラメータでの分析プロセス (予測モデル) へ入力され、それから、結果を生成するために使用される。

【0043】

データは、サンプルに関連したデータを受け取る個人に帰着する任意の技術を介して、

取得することができる。たとえば、個々人は、当業者に公知の方法によって、自分自身でデータセットを生成することにより、データセットを取得することができる。また、データセットは、他の個人または団体からのデータセット、または、1つ以上のデータ値を受け取ることにより、取得することができる。例えば、ある研究室専門家は、特定のデータ値を生成することができ、一方、医療専門家のような他の個人は、そのデータセットの全部又は一部を、分析プロセスに入力して、結果を出すことができる。

【0044】

本開示を通して、参照が「サンプル」になされるが、定量的データは、調達方法、調達時間、組織の由来など、特性の任意の数が増加する複数のサンプルから得ることができることを、当業者は理解すべきである。

10

【0045】

アテローム性動脈硬化症の分類に有用な結果を生成する方法で、本明細書で提示されたタンパク質マーカーの血液、血清、その他での発現パターンが、得られる。興味あるタンパク質マーカーに関連する定量的データは、そのようなマーカーに関連するDNAまたはRNAレベルの測定を含むアテローム性動脈硬化症の分類に有用な結果の生成を許可する如何なるデータでもあり得るが、一般的には、タンパク発現パターンである。タンパク質レベルは、個別に、または発現プロファイルの一部のような、ハイスループット法を介して定量的測定を生成するような、当業者に公知の任意の方法を介して測定することができる。たとえば、血液由来の患者サンプル、例えば血液、血漿、血清などが、興味のあるタンパク質マーカーの存在と量を決定するために、特定の結合剤や、特定の結合剤のパネルに適用され得る。

20

【0046】

血液サンプル、または、例えば、血漿、血清などの血液由来のサンプルは、(複数の)miRNAマーカーだけで、または関心ある(複数の)タンパク質マーカーとの組み合わせで、それらの発現レベルの存在についてアッセイされる。一般的には、血液サンプルが採血され、血漿または血清のような血液由来物が、テストされる。加えて、サンプルは、唾液、尿、精液、母乳や汗の様な他の体液由来のものでも可能である。さらに、サンプルは、動脈、静脈、毛細血管のような組織由来でも可能である。さらに、miRNAマーカー、タンパク質バイオマーカーの両方を測定する場合、それらは同じサンプルからでも、または異なるサンプルからでも派生させることができる。例えば、miRNAバイオマーカーは、血液由来のサンプルでアッセイし、タンパク質バイオマーカーは、組織サンプルでアッセイすることもできる。

30

【0047】

興味あるmiRNAマーカーとタンパク質マーカーに関連付けられる定量的データは、通常、発現プロファイルの形式をとる。発現プロファイルは、多数の評価されたマーカーに対応した、多くのmiRNAまたはタンパク質産物の、相対的または絶対的な発現値のセットを構成している。様々な実施形態において、少なくとも約2、3、4、5、6、7またはそれ以上のマーカーの発現パターンを含む発現プロファイルが作られる。発現プロファイルの個々の発現が異なる各構成メンバーの発現パターンは、例えば、診断、予後、治療のモニタリングのための予測値に関して、特定の特異性と感度を提供することができる。

40

【0048】

発現データを取得するための多数の方法が知られており、これらテクニックのうちの1つかまたは複数は、単独でまたは組み合わせで、本開示での発現パターンとそのプロファイルを決定するのに適している。

【0049】

例えば、DNAとRNA(mRNA、pri-miRNA、pre-miRNA、miRNA、前駆体ヘアピンRNA、マイクロRNPなど)の発現パターンは、ノーザン解析、PCR、RT-PCR、TaqMan分析、FRET検出、1つまたは複数の分子標識(ビーコン)のモニタリング、オリゴヌクレオチドアレイのハイブリダイゼーション、

50

cDNAアレイのハイブリダイゼーション、ポリヌクレオチドアレイへのハイブリダイゼーション、液体マイクロアレイへのハイブリダイゼーション、マイクロ電気アレイへのハイブリダイゼーション、cDNAシーケンシング、クローン・ハイブリダイゼーション、cDNA断片フィンガープリント、遺伝子発現の連続解析(SAGE法)、サブトラクティブハイブリダイゼーション、ディファレンシャルディスプレイ及び/又はディファレンシャルスクリーニングによって、評価することができる。これらや他のテクニックは、当業者に周知の技術である。

【0050】

本開示は、好ましくは単離された形態の核酸分子を含む。本明細書中で使用されているように、核酸分子が、実質的に他のポリペプチドをコードする核酸分子のコンタミから分離されている場合に、核酸分子は、「単離」されるべきである。用語「核酸」は、コーディングまたは非コーディングのRNAまたはDNAとして定義される。相補的で、即ちハイブリダイズし、適切なストリンジェンシーな条件下で、分子に安定的に結合した状態を保つ核酸は、本開示の範囲内に含まれる。このような配列は、本明細書に開示されたRNAとのヌクレオチド配列の同一性が、少なくとも50%、60%、70%または75%、好ましくは少なくとも約80~90%、より好ましくは少なくとも約92~94%、さらにより好ましくは少なくとも約95%、98%、99%かそれ以上で、また、挿入、欠失、ゆらぎ塩基(wobble bases)、置換なども含む。さらに、本明細書に開示のタンパク質バイオマーカーの配列との同一性を、少なくとも約50%、60%、70%または75%、好ましくは少なくとも約80~90%、より好ましくは少なくとも約92~94%、そして、最も好ましくは少なくとも約95%、98%、99%以上共有する配列が考えられている。

【0051】

ゲノムDNA、cDNA、RNA(mRNA、pri-miRNA、pre-miRNA、miRNA、ヘアピン前駆体RNA、RNPなど)の分子が、天然物由来かまたは合成で、代替バックボーンに基づく核酸や別の塩基を含む核酸と同様に、具体的な開示の範囲内で意図されている。

【0052】

ヌクレオチドまたはアミノ酸配列レベルでの相同性または同一性は、配列類似性検索用に調整されている、blastp、blastn、blastx、tblastnおよびtblastxプログラムで採用されたアルゴリズムを用いたBLAST(Basic Local Alignment Search Tool)解析によって決定される。BLASTプログラムによって使用されるアプローチは、まず最初に、検索配列とデータベース配列との間で、ギャップの有無で類似セグメントを考慮し、その後、同定された一致の全ての統計学的意義を評価し、最後に、選択された有意なしきい値を満たす一致データのみサマライズすることである。ヒストグラム、説明、アライメント、期待(すなわち、データベース配列に対する一致をレポートするための統計学的に有意なしきい値)、カットオフ、マトリックスおよびフィルタ(低複雑さ)のための検索パラメータは、デフォルト設定である。blastp、blastx、tblastnおよびtblastxで使用するデフォルトのスコアリングマトリックスは、85ヌクレオチドまたはアミノ酸長さ以上の検索配列に推奨される、BLOSUM62マトリックスである。

【0053】

blastnについては、スコアリングマトリックスは、N(即ち、ミスマッチ残基ペアに対するペナルティスコア)に対するM(即ち、マッチ残基ペアに対する報酬スコア)の比によって設定されるが、ここで、MとNのデフォルト値は、それぞれ、5と-4である。4つのblastnパラメータは、以下のように調整された。つまり、Q=10(ギャップ生成ペナルティ)、R=10(ギャップ伸長ペナルティ)、wink=1(クエリに沿ってすべてのwinkth位置で単語のヒットを生成する)、及びgapw=16(ギャップアライメントが生成される範囲内でウィンドウ幅を設定する)。同等のBlastpパラメータセッティングは、Q=9、R=2、wink=1およびgapw=32で

ある。G C G パッケージのバージョン 1 0 . 0 で利用可能な、配列間のベストフィットな比較は、D N A パラメータの G A P = 5 0 (ギャップ生成ペナルティ)と、L E N = 3 (ギャップ伸長ペナルティ)を使用し、および、タンパク質比較での同等の設定は、G A P = 8 と L E N = 2 である。

【 0 0 5 4 】

「ストリンジェントな条件」は以下のものであって、(1) 低イオン強度および高温を採用する、例えば、5 0 で 0 . 0 1 5 M N a C l / 0 . 0 0 1 5 M クエン酸ナトリウム / 0 . 1 % S D S での洗浄、(2) ハイブリダイゼーション中にホルムアミドなどの変性剤を採用、例えば、5 0 % (体積 / 体積) のホルムアミドに、0 . 1 % ウシ血清 a l b u m i n / 0 . 1 % F i c o l l / 0 . 1 % p o l y v i n y l p y r r o l i d o n e / p H 6 . 5 の 5 0 m M リン酸ナトリウム緩衝液、7 5 0 m M N a C l 、7 5 m M クエン酸ナトリウム、4 2 で。別の例では、5 0 % ホルムアミド中、5 x S S C (0 . 7 5 M の N a C l 、0 . 0 7 5 M クエン酸ナトリウム)、5 0 m M リン酸ナトリウム (p H 6 . 8) 、0 . 1 % ピロリン酸ナトリウム、5 x D e n h a r d t 液、超音波処理サケ精子 D N A (5 0 p g / m l) 、0 . 1 % S D S 、1 0 % 硫酸デキストランで 4 2 でのハイブリダイゼーション後、4 2 、0 . 2 x S S C 及び 0 . 1 % S D S で洗浄。当業者であれば、明確で検出可能なハイブリダイゼーションシグナルを得るために、適切なストリンジェンシー条件を、容易に決定し、変えることができる。

10

【 0 0 5 5 】

本開示はさらに、開示の核酸分子のフラグメント (断片) を提供する。本明細書中で使用されるように、核酸分子のフラグメントとは、コーディングまたは非コーディング配列の小さな部分を指している。フラグメントのサイズは、使用目的によって決定される。例えば、フラグメントがタンパク質の活性部分をエンコードするように選択されている場合、フラグメントは、タンパク質の機能的領域をエンコードするのに十分な大きさである必要がある。例えば、予測抗原領域に対応するペプチドをコードするフラグメントを、調製することができる。フラグメントが、核酸プローブまたは P C R プライマーとして使用される場合、フラグメントの長さは、プライミング / プロービング時の偽陽性の数が比較的小さくなるように、選択される。

20

【 0 0 5 6 】

タンパク質の発現パターンは、定量的な尺度を提供し、以下の 1 つかまたはそれ以上の方法のように、サンプルから抽出された多数のマーカーの評価に適している、当業者に公知の任意の方法によって評価することができる。つまり、タンパク質アレイ (例えば、抗体アレイ) または蛍光活性化セルソーティング (F A C S) に結合する、E L I S A サンドイッチアッセイ、フローサイトメトリー、質量分析検出、熱量測定アッセイ、など。

30

【 0 0 5 7 】

一実施形態では、あるアプローチは、E L I S A 、抗体標識蛍光ビーズアレイ、抗体アレイ、または F A C S スクリーンで、1 つまたは複数のタンパク質産物のエピトープを認識する、標識されたアフィニティー試薬 (例えば、抗体、小分子など) の使用を含む。抗体の産生・評価方法は、当該分野で周知である。

【 0 0 5 8 】

沢山の適切なハイスループットフォーマットが、開示されたバイオマーカーの発現パターンとプロファイルを評価するために存在している。一般的に、用語「ハイスループット」は、一日に、少なくとも約 1 0 0 アッセイ、または少なくとも約 5 0 0 アッセイ、または少なくとも約 1 0 0 0 アッセイ、または少なくとも約 5 0 0 0 アッセイ、または少なくとも約 1 0 , 0 0 0 アッセイ、またはそれ以上を実行するフォーマットを指している。アッセイを列挙するとき、サンプル数やアッセイされるマーカー数のどちらかを考慮することができる。

40

【 0 0 5 9 】

ハイスループット発現解析を行うための多数の技術的プラットフォームが知られている。一般的に、このような方法は、対象サンプル、タンパク質マーカー、またはその両方の

50

うちのいずれかの論理的アレイまたは物理的なアレイを含んでいる。一般的なアレイフォーマットは、液体相、固体相の両方のアレイを含む。例えば、核酸のハイブリダイゼーション、リガンドへの抗体または他の受容体の結合などのための、液相アレイを採用するアッセイは、マルチウェルまたはマイクロタイタープレートで行うことができる。96、384または1536ウェルのマイクロタイタープレートは、広く入手可能であり、例えば、3456や9600のより多くのウェルでさえ、使用することができる。一般的には、マイクロタイタープレートの選択は、例えば、サンプルの調製および分析のために使用される、ロボットによるハンドリングやローディングシステムのような、方法や装置によって決定される。典型的なシステムは、例えば、ルミネックス社 (Austin, TX) の xMAP (登録商標) テクノロジー、メソスケールディスカバリー社 (Gaithersburg, MD) の MULTI-ARRAY (登録商標) や MULTI-SPOT (登録商標) テクノロジーを伴う SECTOR (登録商標) イメージャ、Beckman-Coulter 社 (Fullerton, Calif.) の ORCA (商標) システム、及び、Zymark 株式会社 (Hopkinton, MA) の ZYMATE (商標) システム、miRCURY LNA (商標) microRNA アレイ (Exiqon 社、Woburn, MA)、である。

10

20

30

40

50

【0060】

また、様々な固相アレイが、好ましくは、本開示の方法、アッセイおよびキットの属性内で、発現パターンを決定するために用いることができる。典型的なフォーマットは、膜またはフィルタアレイ (例えば、ニトロセルロース、ナイロンの)、ピンアレイ、およびビーズアレイ (例えば、液体の「スラリー」内で) を含む。一般的に、候補ライブラリーのメンバーに対応する発現産物と特異的に相互作用する (例えば、ハイブリダイズするか結合する) 核酸またはタンパク質試薬に対応するプローブは、固体支持体に、例えば直接または間接的に架橋するなどして、固定化されている。本質的には、特定の発現アッセイを行うために必要な試薬および条件に耐えることができる任意の固体支持体を利用することができる。例えば、機能化ガラス、シリコン、二酸化ケイ素、変性シリコン、(ポリ)テトラフルオロエチレン、(ポリ)フッ化ビニリデン (vinylidene difluoride)、ポリスチレン、ポリカーボネートのような様々なポリマーのいずれか、またはそれらの組み合わせは全て、固相アレイの基質として役立つことができる。

【0061】

一実施形態において、アレイは、例えば、上記指定された材料の一つで構成される「チップ」である。候補となるライブラリーの個々のコンポーネントの発現産物と特異的に相互作用する、例えば、cDNA、合成オリゴヌクレオチドなどのような、例えばRNAまたはDNAの、ポリヌクレオチドプローブ、または、抗体や抗原結合フラグメントまたはそれらの誘導体のような、結合タンパク質が、論理的に順序付けられた方法、すなわち、アレイで、チップに固定されている。さらに、(サンプル標識のデザインに応じて) マーカーの塩基配列のセンスまたはアンチセンス配列のいずれかに特異的親和性を持つ分子は、マーカーに対して特異的親和性を失うことなく、アレイ表面に固定することができ、アレイ作成のため取得し製造することができる。それらは、例えば、マーカーの特定の核酸配列を認識するタンパク質、リボザイム、ペプチド核酸 (PNA)、または特定の親和性を有する他の化学物質や分子である。

【0062】

マイクロアレイの発現は、様々なレーザーやCCDベースのスキャナでマイクロアレイをスキャンし、そして、例えば、IMAGENE (商標) (バイオディスカバリー社)、Feature Extractionソフトウェア (アジレント社)、SCANLYZE (商標) (スタンフォード大学、Stanford, CA)、GENEPIX (商標) (アクソン・インスツルメンツ社) の様な、多数のソフトウェアパッケージで、特徴を抽出することによって、検出することができる。

【0063】

ハイスループットのタンパク質システムは、Ciphergen Biosystem

s社(Fremont, Calif.)のPROTEIN CHIP(商標)アレイや、S&S Biosciences社(Keene, N.H., U.S.)のFASTQUANT(商標)ヒトケモカインタンパク質マイクロアレイの様に市販のシステムを含む。

【0064】

臨床的しるし(clinical indicia)、代謝的測定(metabolic measures)、遺伝子解析など、他のデータセット構成要素に関する定量的データは、当業者に公知の方法によって決定することができる。

【0065】

したがって、miRNA、タンパク質マーカー、及び他のデータセット構成要素(すなわち、臨床的しるしなど)について得られた定量的データは、予測モデルに入力された学習アルゴリズムを使用して以前に決定されたパラメータを用いた分析プロセスに供される。分析プロセスのパラメータは、本明細書に開示されたもの、または本明細書に記載のガイドラインを使用して導き出されたものかもしれない。線形判別分析、回帰的特徴の排除、マイクロアレイの予測分析、ロジスティック回帰、CART、FlexTree、LART、ランダムフォレスト、MART、または他の機械学習アルゴリズムの様な学習アルゴリズムは、様々なアテローム性動脈硬化症分類に適している分析プロセスのためのパラメータを決定するために、適切な参照またはトレーニングデータに適用される。

【0066】

結果(分類、生存/タイム・トゥ・イベント(出来事までの期間)など)を生成するために使用される分析的プロセスは、サンプルを分類するための有用な結果、例えば、取得したデータセットと参照データセットとの比較、線形アルゴリズム、二次式アルゴリズム、決定ツリーアルゴリズム(a decision tree algorithm)、または投票アルゴリズムを用いて得られたデータセットの比較など、を提供することができるプロセスの任意のタイプかもしれない。

【0067】

アテローム性動脈硬化症の分類を行うための有用な結果を得るための様々な分析的プロセスが、本明細書に記載されているが、当業者であれば、容易に、分析プロセスに適切な任意のタイプは、この開示の範囲内であることを理解するであろう。

【0068】

分析プロセスへの入力前に、各データセット内のデータは、通常、2回、3回、または複数回反復して各マーカーの値を測定することにより、収集される。データは、操作することができ、例えば、生データは、標準曲線や、各患者の平均と標準偏差を計算するために使用される反復測定の平均を用いて、変換できる。それから、これらの値は、例えば、対数変換、Box-Cox変換などのモデルで使用される前に、変換することができる。このデータは、定義されたパラメータを使用して、分析プロセスへ入力することができる。

【0069】

分析プロセスでは、サンプルが、ある与えられたクラスに属する確率を決定するためのしきい値を、設定することができる。確率は、好ましくは少なくとも50%、または少なくとも60%または少なくとも70%または少なくとも80%、少なくとも90%、またはそれ以上である。

【0070】

他の実施形態では、分析プロセスは、得られたデータと参照データセット間の比較で、統計学的に有意な差をもたらすかどうかを決定する。もしそうであれば、データセットが取得されたサンプルは、参照データセットのクラスには属しないと分類される。逆に、このような比較で、参照データセットとは統計学的に有意な差がない場合、データセットが得られたサンプルは、参照データセットのクラスに属するものとして分類される。

【0071】

一般的には、分析プロセスは、以下で説明するものの様に、統計学的分析法によって生成されたモデルの形式である。そのような分析プロセスの例としては、線形アルゴリズム

10

20

30

40

50

、二次式アルゴリズム、多項式アルゴリズム、決定ツリーアルゴリズム (a d e c i s i o n t r e e a l g o r i t h m)、投票アルゴリズム (a v o t i n g a l g o r i t h m) を含めることができる。線形アルゴリズムは次式を有することができる：

【数 1】

$$R = C_0 + \sum_{i=1}^N C_i x_i$$

(ここで、Rは、得られた有用な結果、 C_0 はゼロになるかもしれない定数、 C_i および x_i は、それぞれ、定数と、適用可能なバイオマーカーや臨床的しるしの値、及び、Nはマーカーの合計数、である。)

10

【0072】

二次式アルゴリズムは、次式を有することができる：

【数 2】

$$R = C_0 + \sum_{i=1}^N C_i x_i^2$$

(ここで、Rは、得られた有用な結果、 C_0 はゼロになるかもしれない定数、 C_i および x_i は、それぞれ、定数と、適用可能なバイオマーカーや臨床的しるしの値、及び、Nはマーカーの合計数、である。)

20

【0073】

多項式アルゴリズムは、線形または二次式アルゴリズムの、より一般化された式である次式を有することができる：

【数 3】

$$R = C_0 + \sum_{i=0}^N C_i x_i^y$$

(ここで、Rは、得られた有用な結果、 C_0 はゼロになるかもしれない定数、 C_i および x_i は、それぞれ、定数と、適用可能なバイオマーカーや臨床的しるしの値、 y は、 x_i が上昇するパワーであり、Nはマーカーの合計数、である。)

30

【0074】

適切な参照 (基準) またはトレーニングデータセットは、任意の適切な学習アルゴリズムを使い、分類に用いる分析プロセスのパラメータを決定するため、すなわち、予測モデルを開発するために、使用することができる。使用する参照またはトレーニングデータセットは、決定すべき望ましいアテローム性動脈硬化症の分類に依存する。データセットは、2、3、4又はそれ以上のクラスからのデータを含めることができる。例えば、アテローム性動脈硬化症を診断するために用いる分析プロセスのパラメータを決定するために監督下にある学習アルゴリズムを使用するには、コントロールと患者の各サンプルを含むデータセットが、トレーニングセットとして使用される。あるいは、管理された学習アルゴリズムを、アテローム性動脈硬化症のステージ分類のための予測モデルを開発するために使用する場合は、トレーニングセットは、心血管疾患のさまざまなステージのそれぞれのデータを含めることができる。

40

【0075】

以下は、開示された方法、アッセイおよびキットの実施を支援するために、当業者に利用可能な統計学的解析方法のタイプの例である。統計分析は、2つのタスクのいずれかまたは両方に適用される可能性がある。まず、これらおよびその他の統計的方法は、好適なデータセットを形成する、マーカーや他のしるし (i n d i c a) の好ましいサブセットを識別するために使用されることがある。さらに、これらと他の統計的方法は、結果を生

50

み出すデータセットで使用する分析プロセスを生成するために使用されることがある。本明細書中に提示された、あるいは先行技術で使用可能な統計的手法のいくつかは、これらのタスクの両方を実行し、本明細書に開示される方法の実施のための分析プロセスとしての使用に適するモデルを提供するだろう。

【0076】

その対応する特徴的な値（例えば、濃度、発現レベル）で、例えば健康とアテローム性動脈硬化症との間を識別することができるバイオマーカーが、本明細書で同定されている。これらのマーカーと、それに対応する特徴（例えば、濃度、発現量）の同定は、患者のクラスを区別する1つの分析プロセス、または複数の分析プロセスを開発するために使用することができる。以下の（実施）例では、データ解析アルゴリズムが、そのような沢山の分析プロセスを構築するためにどのように使用することができるかを示している。実施例に記載のデータ解析アルゴリズムの各々は、健康とアテローム性動脈硬化症患者を含むトレーニング集団を交えて本明細書で同定された、マーカーのサブセットの特徴（例えば、発現値）を、使用する。本明細書に開示の、被験者を区別する1つ又は複数の分析プロセスを構築するための特定のデータ解析アルゴリズムは、以下の節で説明する。分析プロセスは、これらの例示的なデータ解析アルゴリズム、または当該技術分野で公知の他の技術を使用して構築された後、分析プロセスは、2つまたはそれ以上の表現型のいずれかのクラス（例えば、健康やアテローム性動脈硬化症患者）に被験者を分類するために、及び又は生存/タイム・トゥ・イベント（出来事までの期間）を予測するために使用することができる。これは、被験者から得られた1つ以上のマーカープロファイルに1つまたは複数の分析プロセスを適用することによって達成される。このような分析プロセスは、したがって、診断指標として非常に大きな価値を有する。

【0077】

本開示の方法、アッセイおよびキットは、一態様では、トレーニング集団から得られるマーカープロファイルに対する、被験者から1つまたは複数のマーカープロファイルの評価のために、提供される。いくつかの実施形態では、被験者と同様に、トレーニング集団の被験者から得られた各マーカーのプロファイルは、多くの異なるマーカーのそれぞれの特徴を構成する。いくつかの実施形態では、この比較は、(i) トレーニング集団からマーカープロファイルを用いた分析プロセスを開発し、(ii) 被験者からのマーカープロファイルに分析プロセスを適用することによって、達成される。このように、本明細書に開示のメソッドのいくつかの実施形態に適用される分析プロセスは、テスト被験者がアテローム性動脈硬化症を持っているかどうかを判断するために使用される。代替の実施形態では、本明細書に開示の方法は、被験者がMIを経験するか否かを決定し、および/または（例えば、MIおよび/または生存の）イベントまでの時間（タイム・トゥ・イベント）を予測することができる。

【0078】

本明細書に開示の方法のいくつかの実施形態では、分析プロセスのアプリケーションの結果が、対象（被験者）が、MIを経験する（に見舞われる）可能性が高いことを示しているときは、対象（被験者）は、「MI」対象（被験者）として診断/分類される。また、もし、例えば、分析プロセスの結果が、対象がアテローム性動脈硬化症に発展する可能性が高いことを示していれば、対象は、「アテローム性動脈硬化症」対象として診断される。分析プロセスの適用の結果が、対象がアテローム性動脈硬化症に発展しないことを示している場合は、対象は健常と診断される。したがって、いくつかの実施形態において、上述した二元的な決定状況の結果は、4つの可能な結果を有する。すなわち、(i) 真のアテローム性動脈硬化症：分析プロセスは、対象（被験者）がアテローム性動脈硬化症に発展することを予想し、実際に対象は、一定期間内にアテローム性動脈硬化症に進展する場合（真陽性、TP）。(ii) 偽（仮性）アテローム性動脈硬化症：分析プロセスは、対象が、アテローム性動脈硬化症に発展することを予想したが、実際は、対象は、一定期間内にアテローム性動脈硬化症に進展しない場合（偽陽性、FP）。(iii) 真の健康：分析プロセスは、対象が、アテローム性動脈硬化症に発展しないことを予想し、実際に

、一定期間内にアテローム性動脈硬化症に進展しない場合（真陰性、TN）。（iv）偽（仮性）健康：分析プロセスは、対象が、アテローム性動脈硬化症に進展しないと予想したが、実際には、一定期間内にアテローム性動脈硬化症に進展する場合（偽陰性、FN）である。

【0079】

これらTP（真陽性）、FP（偽陽性）、TN（真陰性）、FN（偽陰性）のための他の定義がなされ得ることが理解されるであろう。そのようなすべての選択的な定義は、本開示の方法、アッセイ、キットの範囲内であるが、理解を容易にするため、TP、FP、TN、FNの定義は、特に明記しない限り、上記（i）から（iv）によるものを、本明細書では使用する。

10

【0080】

当業者によって理解されるように、多くの定量的な判断基準は、あるテストマーカのプロファイルとリファレンスマーカースのプロファイル間での比較のパフォーマンス（例えば、被験者からのマーカースプロファイルへの分析プロセスの適用）を伝えるために使用することができる。これらは、陽性予測値（PPV）、陰性予測値（NPV）、特異性、感度、精度、および確実性を含んでいる。さらに、受信者操作曲線（ROC曲線）の様な他の構築物は、分析プロセスのパフォーマンスを評価するために使用することができる。本明細書中で使用されるのは： $PPV = TP / (TP + FP)$ 、 $NPV = TN / (TN + FN)$ 、 $特異性 = TN / (TN + FP)$ 、 $感度 = TP / (TP + FN)$ 、及び、 $精度 = 確実性 = (TP + TN) / N$ 。

20

【0081】

ここで、Nは、比較したサンプルの数（例えば、アテローム性動脈硬化か健康かの判定が求められているテストのサンプル数）である。例えば、この分類が求められている10の被験者がある場合を考えてみよう。各マーカースプロファイルが、10の被験者のそれぞれのために構築される。次に、各マーカースプロファイルは、トレーニング集団から得られたマーカースプロファイルに基づいて開発された、分析プロセスを適用することによって、評価される。この例では、上記の式のNは10に等しい。一般的に、Nは、サンプルのある集団の異なるメンバーから収集したサンプルの数である。この集団は、実際に、二つの異なるタイプからなる。1つのタイプでは、集団は、そのサンプルと表現型（例えば、マーカースの特徴値や、対象（被験者）がアテローム性動脈硬化症に進展するか否かの指標）のデータが、分析プロセスを構築または改良するために使用された対象で構成されている。この様な集団は、本明細書では、トレーニング集団と称する。他のタイプでは、集団は、分析プロセスを構築するために使用されていない対象で構成されている。このような集団は、検証集団と称する。特に断りのない限り、Nによって表される集団は、もっぱら、トレーニング集団か、または、検証集団であり、2つの集団タイプの混合物とは対照的である。検証集団とは対照的に、トレーニング集団に基づいている場合、精度などのスコアは、より高値（単一（unity）に近い）となることが理解されるであろう。それにもかかわらず、そうでなければ特に断りのない限り、確実性（精度）を含む1つの分析プロセス（または被験者からのバイオマーカースのプロファイルの評価の他の形態）のパフォーマンスを評価するのに使用された全ての基準は、トレーニング集団または検証集団のいずれかへの基準に関連する分析プロセスを適用することによって測定される基準に、対応している。

30

40

【0082】

いくつかの実施形態では、Nは、1より大きく、5より大きく、10より大きく、20より大きく、10から100の間、或いは、100より大きく、1000より小さい対象（被験者）である。ある実施形態では、分析プロセス（または他の比較形態）は、トレーニング集団または検証集団に対して、少なくとも約99%の確実性、あるいはそれ以上を有することができる。他の実施形態では、確実性は、トレーニング集団または検証集団に対して、少なくとも約97%、少なくとも約95%、少なくとも約90%、少なくとも約85%、少なくとも約80%、少なくとも約75%、少なくとも約70%、少なくとも約

50

65%、少なくとも約60%である。确实性の有用度は、特定の方法に依存して変化する。本明細書中で使用されるように、「确实性」は、「精度」を意味している。一実施形態では、感度および/または特異性は、トレーニング集団または検証集団に対して、少なくとも約97%、少なくとも約95%、少なくとも約90%、少なくとも約85%、少なくとも約80%、少なくとも約75%、少なくとも約70%である。いくつかの実施形態では、このような分析プロセスは、定められた精度で、アテローム性動脈硬化症の進展を予測するために使用されている。いくつかの実施形態では、このような分析プロセスは、定められた精度で、アテローム性動脈硬化症を診断するために使用されている。いくつかの実施形態では、このような分析プロセスは、定められた精度で、アテローム性動脈硬化症のステージを決定するために使用されている。

10

【0083】

十分な确实性を有する被験者を分類するための分析プロセスで使用される可能性のある特徴の数は、2またはそれ以上である。いくつかの実施形態では、それが3以上、4以上、10以上、または10から200の間である。しかしながら、求められる确实性の程度に応じて、分析プロセスで使用される特徴の数は、多かれ少なかれではあるが、すべてのケースで、少なくとも2であることができる。一実施形態では、被験者を分類するための分析プロセスによって使用され得る特徴の数は、高い确实性で、テスト対象（被験者）の分類を可能にするために最適化されている。

【0084】

特定の実施形態では、分析プロセスが生存率を予測するために利用される。生存分析は、タイムトゥイベント・データのモデリングを含む。比例ハザードモデルは、統計学における生存モデルの1つのクラスである。生存モデルは、いくつかのイベントがその量に関連付けられているかもしれない1つ以上の共変量を発生する前に経過する時間と、関連している。比例ハザードモデルでは、共変量の単位増加のユニークな効果は、ハザード率に関して掛け算の関係である。生存モデルは、以下の二つの部分から成ると見なすことができる：つまり、しばしば $A_0(t)$ と記述され、ハザード（リスク）が共変量のベースラインレベルで経時的にどのように変化するかを表す、基本的なハザード関数、及び、ハザードが、説明的な共変量に応じてどのように変化するかを表す、エフェクト・パラメータ、である。典型的な医療の例は、変動の減少および/または交絡に対する制御のために、年齢や、性別の様な患者の特性と同様に、治療の指定（割当て）や、他の疾患の存在の様な共変量を含む。

20

30

【0085】

比例ハザードの仮説は、共変数がハザードを増加させる仮説である。ベースラインハザードが異なるかもしれないが、固定係数の最も単純なケースでは、例えば、薬剤による治療は、言ってみれば、任意の時刻 t における被験者のハザードを半減させることができる。ただし、共変数は、二元的予測に限定されないことに、注意されたい：連続的な共変数 x の場合には、ハザードは、対数的に反応し、 x での各ユニットの増加は、結果としてハザードの比例スケーリングとなる。通常は完全に一般的なCoxモデルの下では、ベースラインハザードは、「統合外（integrated out）」とされ、またはヒューリスティックに考慮から除外され、残りの部分的尤度（likelihood）が最大化される。任意の比例ハザードモデルにより推定された共変数の影響が、こうしてハザード比として報告される。Coxモデルでは、比例ハザード仮説が成り立つ場合、それはハザード関数を考慮せずにエフェクト・パラメータを推定可能であることを前提としている。

40

【0086】

分析プロセスを開発するための適切なデータ解析アルゴリズムとしては、例えば、分類、回帰ツリー（Classification and Regression Tree）（CART）およびその変異体の様なツリーベースアルゴリズム、一般化された加算モデル、ペナルティー付き回帰手法の様に、線形、ロジスティック、より柔軟な識別技術を含む判別分析を含有するが、これらに限定されない。

【0087】

50

一実施形態では、トレーニング集団から得られたマーカープロファイルへの被験者のマーカープロファイルの比較が行われ、分析プロセスが適用される。分析プロセスは、コンピュータのパターン認識アルゴリズムの様なデータ解析アルゴリズムを使用して構築される。分析プロセスを構築するための他の適切なデータ解析アルゴリズムは、ロジスティック回帰または特徴値の分布の違いを検出するノンパラメトリックなアルゴリズム（例えば、ウィルコクソンの符号順位検定（Wilcoxon Signed Rank Test）（未調整と調整済み））を含むが、これらに限定されない。分析プロセスは、1、2、3、4、5、10、20及びそれ以上のマーカーから測定観測量に対応する2、3、4、5、10、20及びそれ以上の特徴に基づることができる。一実施形態では、分析プロセスは、数百以上の特徴に基づいている。分析プロセスは、また、分類ツリーアルゴリズムを使用して構築することもできる。たとえば、トレーニング集団からの各マーカーのプロファイルは、少なくとも3つの特徴を含むことができ、そこでは、特徴が、分類ツリーアルゴリズムの予測因子である。分析プロセスは、少なくとも約70%、少なくとも約75%、少なくとも約80%、少なくとも約85%、少なくとも約90%、少なくとも約95%、少なくとも約97%、少なくとも約98%、少なくとも約99%、または約100%の精度で、集団（またはクラス）内でのメンバーシップを予測する。

10

20

30

40

50

【0088】

適切なデータ解析アルゴリズムは、当該技術分野では知られている。一実施形態では、開示のデータ解析アルゴリズムは、分類と回帰ツリー（Classification and Regression Tree）（CART）、多重加法回帰ツリー（Multiple Additive Regression Tree）（MART）、マイクロアレイの予測分析（PAM）、またはランダムフォレスト分析で構成されている。このようなアルゴリズムは、対象（被験者）を、正常として、または、特定の疾患状態に特徴的なバイオマーカーレベルを有するとして、判別するために、血液サンプルなどの生体材料から複雑なスペクトル进行分类する。他の実施形態では、本開示のデータ解析アルゴリズムは、分散分析（ANOVA）とノンパラメトリック同等物、線形判別分析、ロジスティック回帰分析、最近傍分類分析（nearest neighbor classifier analysis）、ニューラルネットワーク、主成分分析、二次判別分析、回帰分類とサポートベクトルマシン（support vector machine）で構成されている。そのようなアルゴリズムは、分析プロセスを構築され、および/または分析プロセスのアプリケーションの速度と効率を向上させ、研究者のバイアスを避けるために使用できるかもしれないが、当業者は、コンピュータベースのアルゴリズムが、本開示の方法を実行する必要はないことを認識するだろう。

【0089】

バイオマーカーのプロファイルを生成するために使用することができるにもかかわらず、分析プロセスは、マーカーのプロファイルを評価するために使用することができる。例えば、適切な分析プロセスは、ガスクロマトグラフィーによって生じるマーカープロファイルや、静的な飛行時間型二次イオン質量分析（static time-of-flight secondary ion mass spectrometry）（TOF-MS）により得られるスペクトラ、を評価するために使用することができ、MALDI-TOF-MSスペクトラの解析により、高い確実性（79-89%の正しい分類率）で菌株を識別し、複雑な生体サンプル中のバイオマーカープロファイルを分類するために、MALDI-TOF-MSと液体クロマトグラフィー・エレクトロスプレーイオン化質量分析（LC/ESI-MS）を使用できる。

【0090】

本明細書に開示されるマーカーの発現レベルを使用して、分析プロセスを開発するための一つのアプローチは、最近の重心分類（the nearest centroid classifier）である。このような技術は、各クラス（例えば、健康とかアテローム性動脈硬化症）や、クラス内のマーカーの平均発現レベルによって与えられた重心のために、計算し、その重心に最も近いクラスに新しいサンプルを割り当てる。クラスタが

既知のクラスで置き換えられます場合を除き、このアプローチは、k - m e a n s クラスタリングと似ている。このアルゴリズムは、マーカーの多数が使用されているとき、ノイズに敏感となり得る。それらが偶然によるものである可能性が高いと判断される場合は、テクニックの1つのエンハンスメントは、シュリンケージ (s h r i n k a g e) を使用することであり、そこでは、各マーカーに対して、クラスの重心の違いが、ゼロに設定されている。このアプローチは、マイクロアレイ、またはPAMの予測分析に実装されている。シュリンケージ (s h r i n k a g e) は、差異がノイズとみなされるしきい値下によって制御される。ノイズレベル以上の有意差を示さないマーカーは削除される。しきい値は、相互検証 (クロスバリデーション) によって選択することができる。しきい値が減少するにつれて、より多くのマーカーが含まれ、ノイズマーカーの結果として、底に到達して、登山を再開する (過剰適合としてしられる現象) まで、推定される分類エラーが減少する。

10

【 0 0 9 1 】

多重加法回帰ツリー (M u l t i p l e A d d i t i v e R e g r e s s i o n T r e e) (M A R T) は、本明細書に開示される方法で使用することができる分析プロセスを構築するための他の方法を表している。 M A R T のための一般的なアルゴリズムは以下のとおりである。

1 . 初期化

【 数 4 】

$$F_0(x) = \operatorname{argmin}_y \sum_{i=1}^N L(y_i, y)$$

20

2 . m = 1 ~ M に対し :

(a) 1 = 1 , 2 , . . . , N を計算。

【 数 5 】

$$r_{im} = - \left| \frac{\partial L(y_i, f(x_i))}{\partial f(x_i)} \right|_{f=f_{im-1}}$$

30

(b) 末端領域の R J M を与えるターゲットリムへ回帰ツリー (r e g r e s s i o n t r e e) をフィットさせる。ここで、 j = 1 , 2 , . . . , J m である。

(c) j = 1 , 2 , . . . , J m に対して計算する。

【 数 6 】

$$\gamma_{jm} = \operatorname{argmin}_{\gamma \in R_{jm}} \sum_{i \in R_{jm}} L(y_i, f_{m-1}(x_i) + \gamma)$$

$$(d) \text{ Update } f_m(x) = f_{m-1}(x) + \sum_{j=1}^{J_m} \gamma_{jm} I(x \in R_{jm})$$

40

3 . アウトプット。 f (x) = f _ M (x)

【 0 0 9 2 】

特定のアルゴリズムが、さまざまな損失基準 L (y , f (x)) の挿入によって、得られる。アルゴリズムの最初のラインは、単に単一のターミナルノードツリー (a s i n g l e t e r m i n a l n o d e t r e e) である最適な定数モデルに初期化する。ライン 2 (a) に計算された負の勾配の成分は、一般的な疑似残差、 r と呼ばれている。一般的に使用される損失関数の勾配は、当技術分野で知られている。MARTの手順に関連付けられているチューニングパラメータは、反復数 M と構成するツリー J . サブ m (m = 1 , 2 , . . . , M) のそれぞれのサイズである。

【 0 0 9 3 】

50

いくつかの実施形態では、対象（被験者）を分類するために使用される分析プロセスは、回帰分析を用いて構築されている。このような実施形態では、分析プロセスは、回帰分類（`regression classifier`）、好ましくはロジスティック回帰分類として特徴づけることができる。このような回帰分類は、分類（`classifier`）を構築するために使用されるマーカー（例えば、そのような各マーカーの発現レベル）の各係数を含む。このような実施形態では、回帰分類の係数は、例えば、最尤法を用いて計算される。このような計算では、バイオマーカーの特徴（例えば、`RT-PCR`、マイクロアレイデータ）が使用されている。特定の実施形態では、2つだけの形質のサブグループからの分子マーカーデータが使用される（例えば、健康な対象やアテローム性動脈硬化症患者で）と、従属変数は、マーカーのデータが利用可能である被験者の特定の形質の存在の有無である。

10

【0094】

別の実施形態では、トレーニング集団は多くの形質のサブグループ（例えば、3つ以上の形質のサブグループ、4つ以上の特定形質のサブグループ、等）から構成されている。これら多くの形質のサブグループは、トレーニング集団での、健康体から、軽度や中程度のアテローム性動脈硬化症等、表現型の進行での個別のステージに対応することができる。この実施形態では、マルチカテゴリーの応答を処理するロジスティック回帰モデルの一般化は、トレーニング集団で見つかった、さまざまな形質の任意のサブグループ間を区別する意思決定を開発するために使用することができる。たとえば、選択した分子マーカーの測定データは、トレーニング集団で表される多くの形質のサブグループのいずれかの間を区別することができる分類（`classifier`）を開発するために、マルチカテゴリーロジットモデルのいずれにも適用することができる。

20

【0095】

いくつかの実施形態では、分析プロセスは、回帰モデル、好ましくは、ロジスティック回帰モデルに基づいている。このような回帰モデルでは、本明細書に開示されるマーカーの選択されたセット内のマーカーのそれぞれの係数が含まれている。このような実施形態では、回帰モデルの係数は、例えば、最尤法を用いて計算される。特定の実施形態では、2つのグループ（例えば、健康と病気）の分子マーカーデータが使用され、従属変数は、マーカーの特性データに対応する患者の状態である。

【0096】

本開示の方法、アッセイおよびキットのいくつかの実施形態は、マルチカテゴリーの（多くの部分に分かれた）応答を処理するロジスティック回帰モデルの一般化を提供している。このような実施形態は、1つまたは3つまたはそれ以上の分類に生体を区別するために使用することができる。このような回帰モデルは、同時に全てのカテゴリペアを参照し、他方のカテゴリではなく、ある一方のカテゴリの応答の可能性（`odds`）を記述する、マルチカテゴリーのロジットモデルを使用している。モデルカテゴリのある特定の（ $J - 1$ ）ペアのロジット（`logits`）を指定したら、残りは不要になる。

30

【0097】

線形判別分析（`LDA`）は、特定の客体の特性に基づいて、2つのカテゴリのいずれかに対象を分類しようとする。言い換えれば、実験で測定された客体が寄与しているか否かの`LDA`テストは、対象の分類（`categorization`）を予測する。`LDA`は、通常、連続的な独立変数と二分カテゴリ従属変数を必要とする。本開示の方法で使用するために、トレーニング集団のサブセット間のマーカーの選択された一連の式の値が、必要な連続的な独立変数として機能する。トレーニング集団のメンバーのそれぞれのグループの分類は、二分カテゴリ従属変数としての役割を果たす。

40

【0098】

`LDA`は、グループ情報を使用して、群間分散と群内分散との比を最大にする、変数の線形結合（`the linear combination`）を求める。暗黙のうちに、`LDA`で使用される線形重みは、トレーニング・セット全体でのあるマーカーの発現は、両群（例えば、アテローム性動脈硬化症を持つグループ、アテローム性動脈硬化症を持た

50

ないグループ)にどの様に分離するか、また、この発現は、他のマーカーの発現とどの様に相関するかに、依存する。いくつかの実施形態では、LDAは、本開示に記載された遺伝子の組合せ中、K個の遺伝子によって、トレーニングサンプルのN数のメンバーのデータマトリックスに適用される。その後、トレーニング集団の各メンバーの線形判別式(discriminant)が、プロットされる。理想的には、第1のサブグループ(アテローム性動脈硬化症を持たない被験者など)を表すトレーニング集団のそれらのメンバーは、線形判別値(例えば、負の値)の第1の範囲にクラスタし、第2のサブグループ(アテローム性動脈硬化症を持つ被験者など)を表すトレーニング集団のそれらのメンバーは、線形判別値(例えば、正の値)の第2の範囲にクラスタするだろう。判別値のクラスタ間の分離が大きい場合には、LDAは、より成功したとみなされる。

10

【0099】

二次判別分析(QDA)は、LDAと同じ入力パラメータを受け取り、同じ結果を返す。QDAは、結果を生成するために、線形方程式ではなく、二次方程式を使用している。LDAとQDAは、(要求される対象数に関連した違いがあるが)ほぼ互換性があり、かつ、使用するかは、好み、および/または分析をサポートするためのソフトウェアの可用性の、問題である。ロジスティック回帰は、LDAとQDAと同じ入力パラメータを受け取り、同じ結果を返す。

【0100】

本明細書で同定されたマーカーの発現レベルを用いて構築することができる分析プロセスの一つのタイプは、決定ツリー(decision tree)である。ここで、「データ解析アルゴリズム」は、最終「決定ツリー」が、分析プロセスであるのに対し、分析プロセスを構築することができる任意のテクニックである。分析プロセスは、トレーニング集団と、特定のデータ解析アルゴリズムとを使用して、構築される。ツリーベースの方法は、特徴空間を四角形のセットに分かち、その後、あるモデル(定数など)をそれぞれ1つに適合させる。

20

【0101】

トレーニング集団のデータは、トレーニングのセット集団全体のマーカーのための特徴(例えば、発現値、または他のいくつかの観察可能なもの)を含む。分析プロセスを構築するために使用することができる、ある特定のアルゴリズムは、分類・回帰ツリー(CART)である。他の特定の決定ツリーアルゴリズム(decision tree algorithms)は、ID3、C4.5、MART、及びランダム森林を含むが、これらに限定されない。すべてのそのようなアルゴリズムは、当技術分野で公知である。

30

【0102】

本開示の方法、アッセイおよびキットのいくつかの実施形態では、決定ツリー(decision trees)は、マーカーの選択されたセットのための発現データを用いて患者を分類するために使用される。決定ツリーアルゴリズムは、監督下にある学習アルゴリズムのクラスに属する。決定ツリーの目的は、実世界の実施例データから分析プロセス(ツリー)を誘導することである。このツリーは、決定ツリーを導き出すために使用されていない、目に見えないサンプルを分類するために、使用することができる。

40

【0103】

決定ツリー(decision tree)はトレーニングデータから導出される。ある例は、様々な属性に対する値と、その例がどのクラスに属するかを、含んでいる。一実施形態では、トレーニングデータは、トレーニング集団全体に亘って本明細書に記載のマーカーの組合せに対する、発現データである。

【0104】

次のアルゴリズムは、決定ツリーの導出について説明している。

ツリー(事例、クラス、属性)。

ルートノード(root node)を作成する。

すべての事例で、同じクラス値を使用している場合は、ルートにこのラベルを与える

50

でなければ、属性が空の場合は、最も一般的な値に基づいてルートにラベルを付ける。

でなければ、開始する。

各属性の情報利得を計算する。

最高情報利得を持つ属性 A を選択し、このルートの属性を作る。

この属性の各々の値 v のために

A = v に対応して、ルートの下に新しい枝 (ブランチ) を追加する。

事例 (v) は、A = v の事例とする。

事例 (v) が空の場合、新しい枝に、複数の事例の中で最も一般的な値で標識されたリーフノードを作る。

10

でなければ、この新しい枝は、ツリー (事例 (v) 、クラス、属性 { A }) によって作成されたツリーとする。

終了。

【 0 1 0 5 】

情報利得 (information gain) の計算の詳細については、以下に示されている。事例の有り得べきクラス v_i が、確率 $P(v_i)$ を持つ場合、実際の答えの情報コンテンツ I は次式で与えられる：

【 数 7 】

$$I(P(V_1), \dots, P(V_n)) = \sum_{i=1}^n -P(v_i) \log_2 P(v_i)$$

20

I 値は、使用される特定のデータセットの分類の結果を記述できるようにするためには、情報がどの程度必要であることを示している。仮に、データセットが、p ポジティブ (例えば、アテローム性動脈硬化症を持っている) や、n ネガティブ (例えば、健康) の事例 (例えば、個人) を含むと仮定すると、正解に含まれる情報は以下の通りである：

【 数 8 】

$$I\left(\frac{p}{p+n}, \frac{n}{p+n}\right) = -\frac{p}{p+n} \log_2 \frac{p}{p+n} - \frac{n}{p+n} \log_2 \frac{n}{p+n}$$

30

ここで、 \log_2 は、ベース 2 を使用する対数である。単一の属性をテストすることによって、正しい分類を行うために必要な情報量を減らすことができる。ある特定の属性 A (例えば、マーカー) の残りの部分は、必要とされる情報をどの位減らすことができるかを示している。

【 数 9 】

$$\text{Remainder}(A) = \sum_{i=1}^v \frac{p_i + n_i}{p + n} I\left(\frac{p_i}{p_i + n_i}, \frac{n_i}{p_i + n_i}\right)$$

40

ここで「v」は、特定のデータセットにおける、属性 A に固有な属性値の数である。

「i」は、ある特定の属性値である。「 p_i 」は、分類が正 (ポジティブ) (例えば、アテローム性動脈硬化症) である、属性 A の事例数である。「 n_i 」は、分類が負 (ネガティブ) (例えば、健康) である、属性 A の事例の数である。

【 0 1 0 6 】

特定の属性 A の情報利得 (information gain) は、属性 A の特定クラスと属性 A の残りのものとの情報量の差として計算される。

【数 10】

$$\text{Gain}(A) = I\left(\frac{p}{p+n}, \frac{n}{p+n}\right) - \text{Remainder}(A)$$

情報利得は、異なる複数の属性（どの様にうまく複数の事例を分割するか）、そして最高の情報を有する属性は、分類のためにどの程度重要かを評価するために使用される。

【0107】

分類・回帰ツリー（CART）、多変量の決定ツリー、ID3及びC4.5を含み、これらに限定されないが、一般的に、沢山の様々な決定ツリーアルゴリズムが、存在する。

10

【0108】

一実施形態では、決定ツリーを用いた場合には、トレーニング集団全体のある選択されたマーカーセットのための発現データは、平均ゼロ、単位分散（unit variance）を持つことが標準化されている。トレーニング集団のメンバーは、ランダムに、トレーニングセットとテストセットに分割される。例えば、一実施形態では、トレーニング集団のメンバーの3分の2は、トレーニング・セットに配置され、トレーニング集団のメンバーの3分の1は、テストセットに配置される。本明細書中に記載のマーカーの選択・組合せのための発現値は、分析プロセスを構築するために使用されている。その後、テスト・セットのメンバーを正しく分類するための、分類プロセスの能力が、決定される。いくつかの実施形態では、この計算は、ある与えられたマーカーの組合せで複数回実行される。各々の計算の繰り返しにおいて、トレーニング集団のメンバーは、ランダムにトレーニングセットとテストセットに割り当てられる。その後、分子マーカーの組合せの品質が、分析プロセス計算の個々のそのような反復の平均値として取り扱われる。

20

【0109】

分割が、本明細書に開示マーカーのセットの中で対応するマーカーの発現レベル、または2つのそのようなマーカーの発現レベルに基づいている、一変量の決定ツリーに加えて、多変量の決定ツリーを、分析プロセスとして実装することができる。このような多変量の決定ツリーでは、意思決定の一部またはすべては、実際には、マーカーの複数の発現レベルの線形結合で構成されている。このような線形結合は、分類上の勾配降下（gradient descent）などの公知の技法や、二乗和誤差基準（sum-square error criterion）を用いて訓練することができる。

30

【0110】

このような分析プロセスを説明するために、式： $0.04x_1 + 0.16x_2 < 500$ を考えてみよう。ここでは、 x_1 と x_2 は、本明細書に開示のマーカーの中から、二つの異なるマーカーの2つの異なる特徴を、参照している。分析プロセスを調査する（poll）ために、特徴 x_1 と x_2 の値は、分類されていない被験体から得られた測定値から取得される。これらの値は、式に挿入される。500未満の値が計算されている場合は、意思決定ツリー（決定ツリー）の最初の分岐（ブランチ）が取られる。それ以外の場合は、意思決定ツリーの2番目のブランチが取得される。

【0111】

40

本開示で使用可能なもう一つのアプローチは、多変量適応型回帰スプライン（MARS）である。MARSは、回帰の適応手順で、本明細書に開示される方法によって処理される高次元の問題によく適している。MARSは、段階的な線形回帰の一般化、または、回帰の設定でCARTのパフォーマンスを向上させるためのCART法の修正、とみなすことができる。

【0112】

いくつかの実施形態では、マーカーの選択されたセットのための発現値は、トレーニングセットをクラスタするために使用される。たとえば、10個のマーカーが使用されている場合を考える。トレーニング集団の各メンバーは10個のマーカーのそれぞれの発現値を持つ。トレーニング集団のあるメンバーmからのそのような値は、ベクトルを定義する

50

。

$$x_{1m} \ x_{2m} \ x_{3m} \ x_{4m} \ x_{5m} \ x_{6m} \ x_{7m} \ x_{8m} \ x_{9m} \ x_{10m}$$

ここで、 x_{im} は、被検体 m の i 番目のマーカーの発現レベルである。トレーニングセットに m の生物が存在する場合、 i マーカーの選択は、 m のベクトルを定義する。本明細書に開示される方法は、ベクトルで使用するすべての単一マーカーの発現値はすべての単一ベクトル m で表されることを必要としないことに注意されたい。言い換えれば、 i 番目のマーカーのいずれかが見つからない被検体（対象）からのデータでも、まだクラスタリングに使用することができる。このような場合には、ミスしている発現値が、「ゼロ」または他の基準化された値が割り当てられている。いくつかの実施形態では、クラスタリングの前に、発現値の式の値がゼロと、単位分散の平均値を持つように基準化されている。

10

【0113】

トレーニング集団全体で同様の発現パターンを示すトレーニング集団のそれらのメンバーは、一緒にクラスタする傾向がある。マーカーの特定の組み合わせは、ベクトルがトレーニング集団に見られる形質群にクラスタ化するとき、本明細書に開示される方法のこの局面では良い分類指標であると考えられている。例えば、トレーニング集団が、健康な対象やアテローム性動脈硬化症患者を含む場合、クラスタリングの分類指標は、その集団を、健康な対象とアテローム性動脈硬化症患者のいずれかをそれぞれ一意的に表す、2つのグループにクラスタ化するだろう。

20

【0114】

クラスタリングの問題は、データセット内のナチュラルなグループを見つけることの一つとして記載されている。ナチュラルなグループを識別するために、2つの問題点が扱われる。まず、2つのサンプル間の類似度（または非類似度）を測定する方法が決定される。このメトリック（類似度測定）は、1つのクラスタ内のサンプル同士は、他のクラスタ内のサンプルに対してよりも互いに似ていることを確認するために使用される。第二に、類似性の尺度を使用して各クラスタにデータを分割するためのメカニズムが決定される。

【0115】

クラスタリング調査を開始する一つの方法は、距離関数を定義し、データセット内のサンプルのすべてのペア間の距離の行列を計算することである。距離が類似性の良い指標である場合は、同じクラスタ内のサンプル間の距離は、異なるクラスタ内のサンプル間の距離よりも大幅に少ない。しかし、クラスタリングは、距離メトリックを使用する必要はない。たとえば、ノンメトリックな類似度関数 $s(x, x')$ は、2つのベクトル x と x' を比較するために使用することができる。従来、 $s(x, x')$ は、 x と x' が幾分「類似」している場合にその値が大きい、対称関数（symmetric function）である。

30

【0116】

データセット内のポイント間の「類似性」または「類似度」を測定するための方法が選択された後、クラスタリングは、データの任意のパーティションのクラスタリング品質を測定する、評価関数を必要とする。評価関数を極端化する（extremize）データセットのパーティションは、データをクラスタ化するのに使用される。しかしながら、本明細書に開示の方法で使用する特定の例示的なクラスタリング技術は、以下に限定はされないが、階層的クラスタリング（最近傍アルゴリズム（nearest-neighbor algorithm）を用いた凝集クラスタリング（agglomerative clustering）、最遠傍アルゴリズム（farthest-neighbor algorithm）、平均連鎖アルゴリズム、重心アルゴリズム、または二乗和（平方和）アルゴリズム（sum-of-squares algorithm）、K平均法（k-means clustering）、ファジィK平均法アルゴリズム（fuzzy k-means clustering algorithm）、及び、ジャーヴィス・パトリック・クラスタリング（Jarvis-Patrick clusteri

40

50

ng)を含有する。

【0117】

主成分分析(PCA)は、バイオマーカーデータを分析することを提案している。より一般的には、PCAは、患者の1クラスとその他(例えば、アテローム性動脈硬化症を持っている人とそうでない人)とを判別する分析プロセスを構築するために、本明細書に開示されるマーカーの特徴値のデータを分析するために使用することができる。主成分分析は、データの特徴を要約し、データを、その変数(主成分)の新しいセットに変換することによって、データセットの次元数を削減するための、古典的な手法である。

【0118】

PCAのいくつかの非限定的な例は次のとおりである。主要なコンポーネント(PC)に相関が無く、k番目のPCは、PCの中でk番目の最大分散を有するように順序付けされる。k番目のPCは、最初のk-1のPCに直交するデータポイントの予測の変動を最大限に高める方向として解釈することができる。最初のいくつかのPCは、データセット内の変動の大部分をキャプチャする。対照的に、最後のいくつかのPCは、データ内の残りの「ノイズ」のみをキャプチャするために度々仮定される。

【0119】

PCAはまた、本明細書に開示されているように分析プロセスを作成するために使用することができる。そのようなアプローチでは、マーカーの選択セットのためのベクトルは、クラスタリングのために説明したのと同じ方法で構築することができる。実際には、ベクトルのセットは、各ベクトルが、トレーニング集団の特定メンバーからの選択マーカーの発現値を表す、ある行列(matrix)で表すことができる。いくつかの実施形態では、このマトリックスは、モノマーの定性バイナリ記述のフリー・ウィルソン法(Free-Wilson method)で表され、第1主成分(PC)が、可能な分散情報の最大量をキャプチャし、第2主成分(PC)が、すべての分散情報の二番目の最大量をキャプチャして、マトリックス内のすべての分散情報が占められるまで、PCAを使用して最大限に圧縮された空間に分布される。

【0120】

次に、ベクトル(各ベクトルは、トレーニング集団のメンバーを表す)の各々は、プロットされる。多くの異なる種類のプロットが可能である。いくつかの実施形態では、一次元のプロットが行われる。プロットのこの形式では、予想(the expectation)は、最初のグループ(例えば、健康な対象)のメンバーが、第1主成分の値の1つの範囲にクラスタし、第2のグループ(例えば、アテローム性動脈硬化症患者)のメンバーが、第1主成分の値の2番目の範囲にクラスタするだろう(当業者は、マーカー値の分布は、これが有効であるためにいかなる変数も伸長がないことを示す必要があることを、理解されよう。)。

【0121】

一例では、トレーニング集団は、2つのグループ：健康な対象とアテローム性動脈硬化症の患者：で構成されている。第1主成分は、トレーニング集団のデータセット全体にわたって、選択マーカーのマーカー発現値を使用して、計算される。その後、トレーニングセットの各メンバーは、第1主成分の値の関数としてプロットされる。この例では、第1主成分が正であるトレーニング集団のそれらのメンバーは、健康な対象者であり、第1主成分が負であるトレーニング集団のそれらのメンバーは、アテローム性動脈硬化症患者である。

【0122】

いくつかの実施形態では、トレーニング集団のメンバーは、複数の主成分に対してプロットされている。たとえば、いくつかの実施形態では、トレーニング集団のメンバーは、第1の次元が第1の主要成分(principal component)であり、第2の次元は、第2の主成分である、2次元プロット上にプロットされる。そのような二次元プロットでは、予想(the expectation)は、トレーニング集団で表される各々のサブグループのメンバーが、個別のグループにクラスタすることである。たとえ

ば、2次元プロットのメンバーの第1のクラスは、軽度のアテローム性動脈硬化症を有する被験者を表し、2次元プロットのメンバーの第2のクラスは、中等度のアテローム性動脈硬化症を有する被験者を表す、などである。

【0123】

いくつかの実施形態では、トレーニング集団のメンバーは、2つ以上の主成分に対してプロットされており、トレーニング集団のメンバーが、それぞれがトレーニング集団で見つかったサブグループをユニークに表す複数のグループにクラスタリングされているかどうか、判定される。いくつかの実施形態では、主成分分析は、当業者に公知の、R m v a パッケージ（統計解析言語）を使用して実行される。

【0124】

最近傍分類指標（Nearest neighbor classifiers）は、メモリベースのものであり、適合モデルは必要ない。クエリポイント（query point） x_0 を考えると、 k トレーニングポイント $x_{(r)}$ 、 x_0 までの距離が最も近い $r \dots k$ が識別され、その後、ポイント x_0 は、 k 最近傍を使用して分類される。結合はランダムに分けることができる。いくつかの実施形態では、特徴空間におけるユークリッド距離は、以下のように距離を決定するために使用される。

$$d_{(r)} = ||x_{(r)} - x_{(0)}||$$

【0125】

一般的に、最近傍アルゴリズムを使用する場合は、線形判別式を計算するのに使われる発現データは、平均がゼロで分散が1であるよう標準化されている。本開示の方法については、トレーニング集団のメンバーは、ランダムに、トレーニングセットとテストセットに分割される。例えば、一実施形態では、トレーニング集団のメンバーの3分の2は、トレーニング・セットに配置され、トレーニング集団のメンバーの3分の1は、テストセットに配置される。本明細書に開示されるマーカーの選択されたセットのプロファイルは、テスト・セットのメンバーがプロットされている特徴空間を表している。次に、テスト・セットのメンバーを正しく特徴づけるためのトレーニングセットの能力が計算される。いくつかの実施形態では、最近傍計算（nearest neighbor computation）は、マーカーのある特定の組合せで複数回実行される。計算の各繰り返しで、トレーニング集団のメンバーは、ランダムにトレーニングセットとテストセットに割り当てられる。その後、マーカーの組み合わせの品質が、最近傍計算（nearest neighbor computation）の各繰り返しの平均値として扱われる。

【0126】

最近傍ルールは、不平等クラスの事前確率、差動の誤分類コスト、特徴選択の問題に対処するために改良することができる。これらの改良の多くは、近傍の加重投票のいくつかのフォームを含んでいる。

【0127】

生物進化の過程に触発され、分類指標の設計の進化的な方法は、分析プロセスのための確率的探索を採用している。概観では、このような方法は、本明細書に開示されたデータセットを生成するバイオマーカーのようなものの測定から、いくつかの集団分析プロセスを作成する。各分析プロセスは、他から多少異なる。次に、分析のプロセスは、トレーニングデータセット全体のデータに記録される。生物進化とのアナロジーを維持するために、得られた（スカラー）のスコアは、時々フィットネス（適応）と呼ばれている。分析プロセスは、そのスコアに基づいてランク付けされ、最良の分析プロセスは、（分析プロセスの総集団のある部分）保持される。再び、生物学的な用語に合わせて、これは適者生存と呼ばれている。分析プロセスは、確率的に次の世代の子供や子孫で変更される。いくつかの子孫の分析プロセスは、前の世代の親よりも高いスコアを持っているが、いくつかは低いスコアを持っている。全体的なプロセスは、その後の世代のために繰り返される。つまり、分析プロセスがスコアリングされて、ベストのものが保持され、別の世代へ与えるためにランダムに変更される。部分的には、ランキングが故に、各世代が、平均して以前

10

20

30

40

50

のものよりわずかに高いスコアを有する。プロセスは、ある世代内の単一の最良の分析プロセスが、所望の基準値を超えるスコアを持つ場合、停止される。

【0128】

バギング、ブースティング、ランダム部分空間法、及び付加的ツリーは、脆弱な分析プロセスを改善するために使用することができる技術の組合せとして知られている、データ解析のアルゴリズムである。これらの技術は、前に述べた意思決定ツリーのような決定ツリーのために設計され、通常、適用される。さらに、このような技術はまた、線形判別分析のような他のタイプのデータ解析アルゴリズムを使用して開発された分析プロセスにも有用である。

【0129】

バギングにおいて、トレーニングデータセットを試し、ランダムな独立したブートストラップを生成、複製し、これらの各々の分析プロセスを構築し、最終的な分析プロセスの単純多数決によってそれらを集約する。ブースティングでは、分析プロセスは、以前の分析プロセスの結果に依存するトレーニング・セットの重み付けのバージョン上に構築される。最初に、すべてのオブジェクトが等しい重みを持っており、最初の分析プロセスは、このデータセット上に構築される。次に、重みは、分析プロセスのパフォーマンスに応じて変更される。誤って分類されたオブジェクトは、より大きな重みを取得し、次の分析プロセスは、再重み付けされた (reweight) トレーニングセットにブースト (boost) される。このように、一連のトレーニングセットと分類が、単純な多数決によって、または最終決定での加重多数決によって結合され、取得される。

【0130】

ブースティングを説明するために、研究下の集団によって示された2つの表現型のグループ、表現型1（例えば、予後不良の患者）、および表現型2（例えば、予後良好の患者）が存在する場合を考える。分子マーカーXのベクトルを考えると、分類G(X)は、上記2つの値のセット：{表現型1、表現型2}でのタイプ値のいずれか1つを取って予測を生成する。トレーニングサンプルのエラー率は以下で、

【数11】

$$err = 1/N \sum_{i=1}^N I(y_i \neq G(x_i))$$

ここで、Nはトレーニングセットでの被験者の数（表現型1または2のいずれかの表現型を持つ被験者の合計）である。例えば、35人の健康者と46人の（アテローム性動脈）硬化症の患者が存在する場合、Nは81である。

【0131】

弱い分析プロセスは、そのエラー率が、ランダム推測よりわずかに優れているに過ぎないものである。ブースティングのアルゴリズムでは、弱い分析プロセスを、繰り返し、データの修正バージョンに適用し、これにより、一連の弱い分類指標 $G_m(x)$ 、 $m = 1, 2, \dots, M$ のシーケンスを生成する。このシーケンスの分類のすべての予測は、最終的な予測を生成するために加重多数決を介して結合される：

【数12】

$$G(x) = \text{sign} \left(\sum_{m=1}^M \alpha_m G_m(x) \right)$$

1. 観察加重 $w_i = 1/N$ 、 $i = 1, 2, \dots, N$ を初期化

2. $m = 1$ から M 、に対して：

(a) 分析プロセス $G_m(x)$ を、重み w_i を使用するトレーニングセットへ適合

(b) 計算

【数 1 3】

$$err = \frac{\sum_{i=1}^N w_i I(y_i \neq G_m(x_i))}{\sum_{i=1}^N w_i}$$

(c) 計算 $a_m = \log((1 - err_m / err_m))$ (d) w_i をセット $w_i \exp[-m I(y_i \neq G_m(X_i))]$, $i = 1, 2, \dots, N$ 3. アウトプット

10

【0 1 3 2】

ここで a_1, a_2, \dots, a_m は、ブースティング・アルゴリズムによって計算され、その目的は、それぞれの $G_m(x)$ の寄与を比較検討することである。その効果は、シーケンス内のより正確な分類に高い影響を与えることである。

【0 1 3 3】

各ブースティング段階でのデータ変更は、重み w_1, w_2, \dots, w_n を、各トレーニング観察結果 (x_i, y_i) , $i = 1, 2, \dots, N$ のそれぞれに適用することから構成されている。最初に、すべてのウェイトは、 $w_i = 1/N$ にセットされ、その結果、最初のステップは、単に通常の方法でデータでの分析プロセスを養成する。それぞれの連続した反復 $m = 2, 3, \dots, M$ のために、観察結果加重は個別に変更され、分析プロセスは、加重された観察結果 (observation) に再適用される。幹 m で、前のステップで誘導される分析プロセス $G_{m-1}(x)$ で誤って分類されたとの観察結果は、それらの加重は増加し、一方、正しく分類された観察結果は、その加重は減少する。反復が進むにつれてこのように、正しく分類するのが困難な観察結果は、増え続ける影響を享受する。それぞれの連続した分析プロセスは、それによってそのシーケンスの前の時に見逃されているトレーニング観測結果に集中するように強制される。

20

【0 1 3 4】

典型的なブースティングアルゴリズムは以下のとおり：

1. 観察加重 $w_i = 1/N$ 、 $i = 1, 2, \dots, N$ を初期化2. $m = 1 \sim M$ 、に対して：(a) 分析プロセス $G_m(x)$ を、重み w_i を使用するトレーニング・セットへ適合

(b) 計算

30

【数 1 4】

$$err = \frac{\sum_{i=1}^N w_i I(y_i \neq G_m(x_i))}{\sum_{i=1}^N w_i}$$

(c) 計算 $a_m = \log((1 - err_m / err_m))$ (d) w_i をセット： $w_i \exp[-m I(y_i \neq G_m(X_i))]$, $i = 1, 2, \dots, N$

40

3. アウトプット

【数 1 5】

$$G(x) = \text{sign} \left[\sum_{m=1}^M a_m G_m(x) \right]$$

【0 1 3 5】

アルゴリズム m で、現在の分類指標 $G_m(x)$ は、ライン 2 a で加重観察に誘導される。結果の重み付き誤差率は、ライン 2 b で計算される。ライン 2 c は、最終的な分類 G_M

50

(ライン3)の生成で、 $G_m(x)$ に与えられる重み a_m を計算する。それぞれの観察の個々の重みは、ライン2dでの次の反復のために更新される。 $G_m(x)$ によって誤って分類された観察結果は、シーケンスでの次の分類 $G_{m+1}(x)$ を誘導するための相対的な影響力を増加させる因子 $e^{x p(m)}$ によって、その重みがスケーリングされている。いくつかの実施形態では、ブースティングまたは適応ブースティング方法が使用されている。

【0136】

いくつかの実施形態では、特徴の事前選択は、ノンパラメトリック・スコアリング方法の様な技術を使用して実行される。特徴の事前選択は、分類指標に使用されるために選択されるマーカーが、分類間でベストなものを識別する、二次元的削減の一形式です。その後、ロジット・ブースト(Logit Boost)手法は、ブースティング手順よりむしろ、使用されている。いくつかの実施形態では、ブースティングおよび他の分類方法が、開示された方法で使用されている。

10

【0137】

ランダム部分空間法では、分類指標は、データ特徴空間のランダム部分空間で構築されている。これらの分類指標は、通常、最終決定ルールで単純過半数の議決権によって結合される(すなわち、分析プロセスで)。

【0138】

示されているように、本明細書に記載の統計的手法は、データセットに含めるべき好ましいマーカーグループを同定するために使用でき、そのデータセットを使用して結果を生成するために使用できる分析プロセスを生成するために、使用できる、アルゴリズムやモデルの種類の例を単に示しているに過ぎない。また、上記や別の場所で説明した技法の組み合わせは、同じタスクのため、あるいは異なるタスクのためのいずれにも使用できる。決定ツリーとブースティングの組合せの使用の様な、いくつかの組合せを、記載している。しかし、他の多くの組み合わせが可能である。例によって、射影追跡(Projection Pursuit)と重み付き多数決(Weighted Voting)の様な、技術上の他の統計的手法は、データセットに含めると、マーカーの優先グループを識別するために使用することができ、また、そのデータセットを使用して結果を生成するために使用できる分析プロセスを生成するために使用することができる。

20

【0139】

分析プロセスで評価されるべきデータセットコンポーネントの最適な数を決定することができる。ある予測モデルを開発するために、上記の学習アルゴリズムを使用した場合、当業者は、分析プロセスを定義するために、マーカーのサブセットを、すなわち、少なくとも3つ、少なくとも4つ、少なくとも5つ、少なくとも6つ、マーカーの完全なセットとなるまで、選択することができる。通常、マーカーのサブセットは、精度の高い予測モデルを維持しながら、例えば、試薬の有用性、定量等の利便性等、定量的なサンプル分析のニーズに合わせて用意されて選択される。

30

【0140】

分類モデル構築のための、多くの有益なマーカーの選択は、パフォーマンス・メトリックの定義と、このメトリックに基づいた有用な予測能力を持つモデルを生成するためのユーザー定義しきい値(user-defined threshold)とを、必要とする。たとえば、パフォーマンス・メトリックは、AUC、予測性の感度および/または特異性と同様に、予測モデルの全体的な精度である。

40

【0141】

あるモデルの予測能力は、例えば、特定の値や値の範囲のAUCまたは正確さなど、品質メトリックを提供する能力に基づいて評価することができる。いくつかの実施形態では、所望の品質のしきい値は、少なくとも約0.7、少なくとも約0.75、少なくとも約0.8、少なくとも約0.85、少なくとも約0.9、少なくとも約0.95、またはそれ以上の精度で、サンプルを分類する予測モデルである。代替措置として、所望の品質のしきい値は、少なくとも約0.7、少なくとも約0.75、少なくとも約0.8、少なく

50

とも約 0.85、少なくとも約 0.9、またはそれ以上の精度で、ある AUC でサンプルを分類する予測モデルである。

【0142】

当技術分野で知られているように、予測モデルの相対的な感度と特異度は、互いに反比例の関係を持つ、選択性メトリックまたは感受性メトリックのいずれかを、優先するように「調整」することができる。上記のようにモデルでの制限は、実行されるテストの特定の要件に応じて、選択した感度や特異性のレベルを提供するために調整することができる。感度と特異性のいずれかまたは両方は少なくとも約 0.7、少なくとも約 0.75、少なくとも約 0.8、少なくとも約 0.85、少なくとも約 0.9、またはそれ以上とすることができる。

10

【0143】

様々な方法がトレーニングモデルで使用されている。マーカーのサブセットの選択は、マーカーのサブセットの順方向選択、または逆方向選択を経由するかもしれない。選択されるマーカーの数は、すべてのマーカーを使用しなくても、モデルのパフォーマンスを最適化する数である。用語の最適な数を定義する一つの方法は、ある与えられたアルゴリズムのために使用される用語の任意の組み合わせと数を使用して、このメトリックのために得られた最大値から 1 標準誤差以下の範囲にある、所望の予測能力（例えば、 $AUC > 0.75$ 、または感度 / 特異性の同等の測定）を持つモデルを生成するいくつかの用語を選択することである。

【0144】

上述したように、データセットのコンポーネントの定量的データは、分析プロセスに入力され、結果を生成するために使用されている。この結果は、例えば、分類、連続的な変数、またはベクトルといった、アテローム性動脈硬化症の分類を行うために有用な、任意のタイプの情報であり得る。例えば、連続変数またはベクトルの値は、あるサンプルが特定の分類に関連付けられる可能性（*the likelihood*）を決定するために使用されることがある。

20

【0145】

アテローム性動脈硬化症の分類は、あらゆるタイプの情報またはアテローム性動脈硬化症状態に関連付けられた任意の種類の情報の生成を参照する。例えば、診断、ステージング、アテローム性動脈硬化進行の評価範囲、予後、モニタリング、処置に対する治療的反応、既知のアテローム性動脈硬化症の治療と同様のメカニズムを介して作用する化合物を同定するためのスクリーニング、擬似冠動脈カルシウムスコアの予測、安定（すなわち、心筋梗塞）か不安定（すなわち、狭心症）か、アテローム性動脈硬化症の合併症の識別などである。

30

【0146】

好ましい実施形態では、結果は、アテローム性動脈硬化症発生の診断や、検出のために使用され、その様なアテローム性動脈硬化症は、特に、心筋梗塞、心不全などの傾向の指標となる。この実施形態では、「健康」と「アテローム性動脈硬化症」のサンプルを含む基準またはトレーニングセットが、予測モデルを開発するために使用されている。データセットは、好ましくは、アテローム性動脈硬化症を示すマーカーのタンパク質発現レベルを含み、結果を生成するために、予測モデルに入力される。その結果、「健康」または「アテローム性動脈硬化症」のいずれかとしてサンプルを分類することができる。他の実施形態では、結果は、サンプルを分類するための有用な情報を提供する連続的な変数であり、例えば、高値は「アテローム性動脈硬化症」のサンプルである確率が高いことを示し、低値は「健康」のサンプルであることの確率が低いことを示す。

40

【0147】

他の実施形態では、結果は、アテローム性動脈硬化症のステージ分類のために使用される。この実施形態では、さまざまなステージで疾患を持つ個人からのサンプルを含む基準またはトレーニングデータセットが、予測モデルを開発するために使用される。モデルは、既知のステージの病気のサンプルや、1つまたは複数のデータセットに対する個々のデ

50

ータセットの単純な比較であり得るし、より複雑な多変量分類モデルでもあり得る。特定の実施形態では、モデルにデータセットを入力することは、データセットが、指定された心血管疾患のステージであるものとして生成されるから、サンプルの分類結果を生成する。同様のメソッドは、参照（基準）セットまたはトレーニングセットが、病気に進展し、後に病気に進展しなかった個人から得られたデータを含む場合を除いて、アテローム性動脈硬化症の予後を提供するために使用することができる。

【0148】

他の実施形態では、結果は、アテローム性動脈硬化症の治療への応答を決定するために使用される。この実施形態では、参照（基準）データセットまたはトレーニングデータセットと予測モデルは、アテローム性動脈硬化症（疾患を有する患者と、疾患を有さない個人からのサンプル）を診断するために使用されるものと同じである。しかし、診断不明とされた個人からのサンプルで構成されるデータセットを入力する代わりに、データセットは、特定の治療が施されている既知の疾患を持つ個人で構成され、サンプルが、「正常・健康の分類」対「アテローム性動脈硬化症の分類」の傾向があるか、その範囲内に置かれているかが、決定される。

10

【0149】

本明細書中で使用される様に、治療は、制限なしに、3, 6, 又は12ヶ月後のフォローアップ検診を含めることができ、例えば、それは、遮断薬、カルシウムチャンネル遮断薬、アスピリン、コレステロール低下剤、等の薬理学的介入、および/または心血管疾患/状況の存在または程度を決定するためのさらなるテストなどである。特定の例では、即時の治療は必要ない。

20

【0150】

別の実施形態では、結果は、薬剤スクリーニング、すなわち、既知のアテローム性動脈硬化症の薬物治療と同様のメカニズムを介して作用する化合物を識別するのに、使用される。この実施形態では、既知のアテローム性動脈硬化症の薬物治療を受けた個人や、特定の治療を受けていない個人を含む基準またはトレーニングセットは、予測モデルの開発に使用することができる。未知のメカニズムを持つ化合物で処理された個人からのデータセットは、モデルに入力される。結果が、サンプルは、既知のアテローム性動脈硬化症の薬物治療で投与された対象から来たものとして分類することができることを示している場合は、新しい化合物は、同じメカニズムを介して作用する可能性がある。

30

【0151】

好ましい実施形態では、結果は冠動脈カルシウムスコア（CCS）に相関する定量的な尺度である、「擬似冠動脈カルシウムスコア」を決定するために使用される。CCSは、全体的な動脈硬化性プラークの負荷を測定する、臨床的心血管疾患のスクリーニング手法である。イメージング技術の様々な異なるタイプを、アテローム性動脈硬化プラークのカルシウム面積と密度を定量化するために使用することができる。電子ビームCTとマルチスライスCTが使用されている場合、CCSは、X線減衰係数とカルシウム沈着の面積の関数である。通常、0のスコアが、動脈硬化性プラークの負荷が無いことを示し、>0から10は、プラークの負荷の証拠は最小限のもので、11から100は、少なくとも軽度で、101から400は、少なくとも中等度で、400を超えると、プラーク負荷の大規模な証拠を示すと考えられる。従来危険因子と組み合わせて使用するCCSは、心血管疾患の合併症の予測能力を向上させる。さらに、CCSはまた、心血管疾患の合併症に依存しない予測因子として作用することができる。

40

【0152】

冠動脈カルシウムスコアが高値の人と低値の人とを含む基準セットまたはトレーニングセットは、個々の擬似冠動脈カルシウムスコアを予測するモデルを開発するために使用することができる。この予測された擬似冠動脈カルシウムスコアは、アテローム性動脈硬化症の診断およびモニタリングに有用である。いくつかの実施形態では、擬似冠動脈カルシウムスコアは、そのような心血管疾患の診断およびモニターするためのイメージング技術から派生した実際の冠動脈カルシウムスコアのように、他の既知の心血管疾患の診断とモ

50

ニタリング手法と組み合わせて使用される。

【0153】

当業者は、これらのメソッドを使用して生成された結果は、心血管疾患の診断およびモニタリングするための、当業者に公知の様々な他の多数の方法と組み合わせて使用することができることを認識するであろう。

【0154】

当業者であれば、これらのメソッドを使用して生成された結果は、心血管疾患の診断およびモニタリングのために、当業者に公知の様々な他の任意の数の方法と組み合わせて使用することもできることを、認識するだろう。

【0155】

一実施形態では、必要性や治療法の有効性を判断するために人間の心臓血管の健康を評価するためのキットが提供され、これは、生物学的サンプル中の表20のmiRNAから選択される少なくとも2種のmiRNAマーカーのレベルを決定するためのアッセイ、各miRNAマーカーのレベルから成るデータセットを取得するための手順、を含み、そして、生物学的サンプルを分類するためのデータを使用する分析・分類プロセスに、データを入力すること（ここで、分類は、アテローム性動脈硬化症分類、健康分類、薬剤曝露分類、非薬剤曝露分類がないからなる群から選択される）、分類プロセスのアウトプットに従って生物学的サンプルを分類すること、そして、分類に基づいて治療法を決定すること、を含む。

10

【0156】

特定の実施形態では、キットはさらに、生物学的サンプル中の、IL-16、sFAS、Fasリガンド、MCP-3、HGF、CTACK、エオタキシン、アディポネクチン、IL-18、TIMP-4、TIMP-1、CRP、VEGF、及びEGFからなる群から選択される少なくとも3つのタンパク質バイオマーカーのレベルを決定するためのアッセイと、タンパク質マーカーの個々のレベルから構成されるデータセットを取得するための手順とを含み、そして、生物学的サンプルを分類するためのデータを使用する、分析・分類プロセスにmiRNAとタンパク質マーカーのデータを入力すること（ここで、分類は、アテローム性動脈硬化症分類、健康分類、薬剤曝露分類、非薬剤曝露分類がないからなる群から選択される）、分類プロセスのアウトプットに従って生物学的サンプルを分類すること、そして、分類に基づいて治療法を決定すること、を含む。

20

30

【0157】

そのような試薬の一つのタイプは、興味の対象のマーカーセットに結合する抗体のアレイまたはキットである。様々な異なるアレイフォーマットが、様々な広範囲に異なるプロープ構造、基板組成物および接着技術と共に、当該分野で知られている。代表的なアレイや興味の対象であるキット組成物は、少なくとも2、少なくとも3、少なくとも4、少なくとも5かそれ以上のmiRNAマーカーの単独、またはタンパク質マーカーとの組合せによる定量用試薬を含有するかまたはそれから構成されている。この点で、試薬は、表1にリストされたmiRNA、そして好ましくは表20でリストされたmiRNAから選択された、少なくとも1、少なくとも2、少なくとも3、少なくとも4、少なくとも5のmiRNAマーカーの定量のために可能である。

40

表1（表1-1～表1-15）

【表 1 - 1】

TABLE 1

Coverage Human microRNA	Target sequence	SEQ ID No:	Target sequence accession
hsa-miR-155*	CUCCUACAUUUAGCAUUAACA	1	MIMAT0004658
hsa-miR-486-5p	UCCUGUACUGAGCUGCCCCGAG	2	MIMAT0002177
hsa-miR-596	AAGCCUGCCCCGGCUCUCGGG	3	MIMAT0003264
hsa-miR-532-3p	CCUCCCACACCCAAGGCUUGCA	4	MIMAT0004780
hsa-miR-1238	CUUCCUCGUCUGUCUGCCCC	5	MIMAT0005593
hsa-miR-34b	CAAUCACUAAUCCACUGCCAU	6	MIMAT0004676
hsa-miR-151-5p	UCGAGGAGCUCACAGUCUAGU	7	MIMAT0004697
hsa-miR-361-3p	UCCCCAGGUCUGAUUUCUGAUUU	8	MIMAT0004682
hsa-miR-211	UUCCCUUUGUCAUCCUUCGCCU	9	MIMAT0000268
hsa-miR-217	UACUGCAUCAGGAACUGAUUGGA	10	MIMAT0000274
hsa-miR-370	GCCUGCUGGGGUGGAACCUUGU	11	MIMAT0000722
hsa-miR-483-3p	UCACUCCUCUCCUCCCGUCUU	12	MIMAT0002173
hsa-miR-520e	AAAGUGCUUCCUUUUUGAGGG	13	MIMAT0002825
hsa-miR-409-5p	AGGUUACCCGAGCAACUUUGCAU	14	MIMAT0001638
hsa-miR-186	CAAAGAAUUCUCCUUUUGGGCU	15	MIMAT0000456
hsa-miR-519c-3p	AAAGUGCAUCUUUUUAGAGGAU	16	MIMAT0002832
hsa-miR-330-3p	GCAAAGCACACGGCCUGCAGAGA	17	MIMAT0000751
hsa-miR-187	UCGUGUCUUGUGUUGCAGCCGG	18	MIMAT0000262
hsa-miR-623	AUCCCUUGCAGGGGUGUUGGGU	19	MIMAT0003292
hsa-miR-106b*	CCGCACUGUGGGUACUUGCUGC	20	MIMAT0004672
hsa-miR-583	CAAAGAGGAAGGUCCAUUAC	21	MIMAT0003248
hsa-miR-135a*	UAUAGGGAUUGGAGCCGUGGCG	22	MIMAT0004595
hsa-miR-30d*	CUUUCAGUCAGAUUUUGCUGC	23	MIMAT0004551
hsa-miR-671-3p	UCCGGUUCUCAGGGCUCACC	24	MIMAT0004819
hsa-miR-1270	CUGGAGAUUGGAAGAGCUGUGU	25	MIMAT0005924
hsa-miR-129-3p	AAGCCCUUACCCCAAAAAGCAU	26	MIMAT0004605
hsa-miR-647	GUGGCUGCACUCACUCCUUC	27	MIMAT0003317
hsa-miR-934	UGUCUACUACUGGAGACACUGG	28	MIMAT0004977
hsa-miR-519e*	UUCUCCAAAAGGGAGCACUUUC	29	MIMAT0002828
hsa-miR-524-3p	GAAGCGCUUCCCUUGGAGU	30	MIMAT0002850
hsa-miR-25*	AGGCGGAGACUUGGGCAAUUG	31	MIMAT0004498
hsa-miR-221*	ACCUGGCAUACAAUGUAGAUUU	32	MIMAT0004568
hsa-miR-302d*	ACUUUAACAUGGAGGCACUUGC	33	MIMAT0004685
hsa-miR-455-3p	GCAGUCCAUGGGCAUUAACAC	34	MIMAT0004784
hsa-miR-433	AUCAUGAUGGGCUCUCCGUGU	35	MIMAT0001627
hsa-miR-139-5p	UCUACAGUGCACGUGUCUCCAG	36	MIMAT0000250
hsa-miR-425*	AUCGGGAUUGUCGUGUCCGCC	37	MIMAT0001343
hsa-miR-30a	UGUAAACAUCCUCGACUGGAAG	38	MIMAT0000087
hsa-miR-520d-3p	AAAGUGCUUCUCUUUGGUGGGU	39	MIMAT0002856
hsa-miR-611	GCGAGGACCCUCGCGGUCUGAC	40	MIMAT0003279
hsa-miR-410	AAUAUAACACAGAUGGCCUGU	41	MIMAT0002171
hsa-miR-502-3p	AAUGCACCUGGGCAAGGAUUA	42	MIMAT0004775
hsa-miR-1200	CUCCUGAGCCAUUCUGAGCCUC	43	MIMAT0005863
hsa-miR-1224-3p	CCCCACCUCUCUCUCCUCAG	44	MIMAT0005459
hsa-miR-511	GUGUCUUUUGCUCUGCAGUCA	45	MIMAT0002808
hsa-miR-148b	UCAGUGCAUCACAGAACUUUGU	46	MIMAT0000759
hsa-miR-127-3p	UCGGAUCCGUCUGAGCUUGGCU	47	MIMAT0000446
hsa-miR-485-3p	GUCAUACACGGCUCUCCUCUCU	48	MIMAT0002176
hsa-miR-1181	CCGUCGCCGCCACCCGAGCCG	49	MIMAT0005826
hsa-miR-518e	AAAGCGCUUCCCUUCAGAGUG	50	MIMAT0002861
hsa-miR-20a*	ACUGCAUUAUGAGCACUUAAG	51	MIMAT0004493

10

20

30

40

【表 1 - 2】

Coverage Human microRNA	Target sequence	SEQ ID No:	Target sequence accession
hsa-miR-492	AGGACCUGCGGGACAAGAUUCUU	52	MIMAT0002812
hsa-miR-654-3p	UAUGUCUGCUGACCAUCACCUU	53	MIMAT0004814
hsa-miR-520g	ACAAAGUGCUUCCCUUAGAGUGU	54	MIMAT0002858
hsa-miR-1264	CAAGUCUUAUUUGAGCACCUGUU	55	MIMAT0005791
hsa-miR-324-5p	CGCAUCCCCUAGGGCAUUGGUGU	56	MIMAT0000761
hsa-miR-129*	AAGCCCUUACCCCAAAAGUAU	57	MIMAT0004548
hsa-miR-1256	AGGCAUUGACUUCACUAGCU	58	MIMAT0005907
hsa-miR-937	AUCCGCGCUCUGACUCUCUGCC	59	MIMAT0004980
hsa-miR-369-5p	AGAUCGACCGUGUUAUUAUUCGC	60	MIMAT0001621
hsa-miR-519d	CAAAGUGCCUCCCUUAGAGUG	61	MIMAT0002853
hsa-miR-103	AGCAGCAUUGUACAGGGCUAUGA	62	MIMAT0000101
hsa-miR-99b*	CAAGCUCGUGUCUGUGGGUCCG	63	MIMAT0004678
hsa-miR-193b*	CGGGGUUUUGAGGGCGAGAUGA	64	MIMAT0004767
hsa-miR-15a	UAGCAGCACAUAAUGGUUUUGUG	65	MIMAT0000068
hsa-miR-551b	GCGACCCAUACUUGGUUUUCAG	66	MIMAT0003233
	GCUGGGCAGGGCUUCUGAGCUCC		
hsa-miR-612	UU	67	MIMAT0003280
hsa-miR-1237	UCCUUCUGCUCCGUCCCCCAG	68	MIMAT0005592
hsa-miR-595	GAAGUGGCCGUGGUGUGUCU	69	MIMAT0003263
hsa-miR-765	UGGAGGAGAAGGAAGGUGAUG	70	MIMAT0003945
hsa-miR-582-3p	UACUGGUUGAACAACUGAACC	71	MIMAT0004797
hsa-let-7b	UGAGGUAGUAGGUUGUGUGUU	72	MIMAT0000063
hsa-miR-520a-3p	AAAGUGCUUCCCUUUGGACUGU	73	MIMAT0002834
hsa-miR-604	AGGCUGCGGAUUCAGGAC	74	MIMAT0003272
hsa-miR-600	ACUUCAGACAAGAGCCUUGCUC	75	MIMAT0003268
hsa-miR-508-5p	UACUCCAGAGGGCGUCACUCAUG	76	MIMAT0004778
hsa-miR-27a	UUCACAGUGGCUAAGUCCGC	77	MIMAT0000084
hsa-miR-31*	UGCUAUGCCAACAUAUUGCCAU	78	MIMAT0004504
hsa-miR-194	UGUACAGCAACUCCAUUGGA	79	MIMAT0000460
hsa-miR-490-5p	CCAUGGAUCUCCAGGUGGGU	80	MIMAT0004764
hsa-miR-1265	CAGGAUGUGGUCAAGUGUUGUU	81	MIMAT0005918
hsa-miR-593	UGUCUCUGCUGGGGUUUUCU	82	MIMAT0004802
hsa-miR-18b	UAAGGUGCAUCUAGUGCAGUUAG	83	MIMAT0001412
hsa-miR-323-5p	AGGUGGUCCGUGGCGCGUUCGC	84	MIMAT0004696
hsa-miR-33a*	CAAUGUUUCCACAGUGCAUCAC	85	MIMAT0004506
hsa-miR-185*	AGGGGCGUGGCUUCCUCUGGUC	86	MIMAT0004611
hsa-miR-720	UCUCGCGUGGGGCCUCCA	87	MIMAT0005954
hsa-miR-18b*	UGCCCUAAAUGCCCCUUCUGGC	88	MIMAT0004751
hsa-miR-122	UGGAGUGUGACAAUGGUGUUUG	89	MIMAT0000421
hsa-miR-1178	UUGCUCACUGUUCUCCCUAG	90	MIMAT0005823
hsa-miR-892a	CACUGUGUCCUUCUGCGUAG	91	MIMAT0004907
hsa-miR-149*	AGGGAGGGACGGGGCUGUGC	92	MIMAT0004609
hsa-miR-940	AAGGCAGGGCCCCCGCUCCCC	93	MIMAT0004983
hsa-let-7f-2*	CUAUACAGUCUACUGUCUUUCC	94	MIMAT0004487
hsa-miR-154*	AAUCAUACACGGUUGACCUAUU	95	MIMAT0000453
	ACUGGGGGCUUUCGGGCUCUGCG	96	
hsa-miR-637	U		MIMAT0003307
hsa-miR-182*	UGGUUCUAGACUUGCCAACUA	97	MIMAT0000260
hsa-miR-192,	CUGACCUAUGAAUUGACAGCC	98	MIMAT0000222
hsa-miR-519a*, hsa-			
miR-518e*, hsa-miR-	CUCUAGAGGGAAGCGCUUUCUG	99	MIMAT0005452
519b-5p, hsa-miR-519c-			
5p, hsa-miR-522* & hsa-			

10

20

30

40

【表 1 - 3】

Coverage Human microRNA	Target sequence	SEQ ID No:	Target sequence accession
miR-523*			
hsa-miR-202	AGAGGUUAUAGGGCAUGGGAA	100	MIMAT0002811
hsa-miR-499-5p	UUAAGACUUGCAGUGAUGUUU	101	MIMAT0002870
hsa-miR-548i	AAAAGUAAUUGCGGAUUUUGCC	102	MIMAT0005935
hsa-miR-769-3p	CUGGGAUUCUCCGGGUCUUGGUU	103	MIMAT0003887
hsa-miR-337-3p	CUCCUAUAUGAUGCCUUUCUUC	104	MIMAT0000754
hsa-miR-522	AAAUGGUUCCCUUUAGAGUGU	105	MIMAT0002868
hsa-miR-486-3p	CGGGGCAGCUCAGUACAGGAU	106	MIMAT0004762
hsa-miR-17	CAAAGUGCUUACAGUGCAGGUAG	107	MIMAT0000070
hsa-miR-891b	UGCAACUUACCUGAGUCAUUGA	108	MIMAT0004913
hsa-miR-181a*	ACCAUCGACCGUUGAUUGUACC	109	MIMAT0000270
hsa-miR-525-3p	GAAGGCGCUUCCCUUUAGAGCG	110	MIMAT0002839
hsa-miR-603	CACACACUGCAAUUACUUUUGC	111	MIMAT0003271
hsa-miR-889	UUAAUAUCGGACAACCAUUGU	112	MIMAT0004921
hsa-miR-338-5p	AACAAUAUCCUGGUGCUGAGUG	113	MIMAT0004701
	AGCAGAAGCAGGGAGGUUCUCCA	114	
hsa-miR-298			MIMAT0004901
hsa-miR-616	AGUCAUUGGAGGGUUUGAGCAG	115	MIMAT0004805
hsa-miR-26b*	CCUGUUCUCCAUUACUUGGCUC	116	MIMAT0004500
	AAAGGAUUCUGCUGUCGGUCCAC	117	
hsa-miR-541*	U		MIMAT0004919
hsa-miR-28-3p	CACUAGAUUGUGAGCUCUGGA	118	MIMAT0004502
	GACCUGGACAUGUUUGUGCCCAGU	119	
hsa-miR-619			MIMAT0003288
hsa-miR-148a	UCAGUGCACUACAGAACUUUGU	120	MIMAT0000243
hsa-miR-1249	ACGCCCUCUCCCCCUUCUUA	121	MIMAT0005901
hsa-miR-1204	UCGUGGCCUGGUCUCCAUUAU	122	MIMAT0005868
hsa-let-7d	AGAGGUAGUAGGUUGCAUAGUU	123	MIMAT0000065
hsa-miR-429	UAAUACUGUCUGGUAAAACCGU	124	MIMAT0001536
hsa-miR-453	AGGUUGUCCUGGUGAGUUCGCA	125	MIMAT0001630
hsa-miR-195*	CCAAUAUUGGUGUGCUGCUCC	126	MIMAT0004615
hsa-miR-132	UACAGUCUACAGCCAUGGUCG	127	MIMAT0000426
hsa-miR-135b	UAUGGCUUUUUAUCCUAUGUGA	128	MIMAT0000758
hsa-miR-32	UAUUGCACAUAACUAAGUUGCA	129	MIMAT0000090
hsa-miR-29c*	UGACCGAUUUCUCCUGGUGUUC	130	MIMAT0004673
hsa-miR-100	AACCCGUAUAUCCGAACUUGUG	131	MIMAT0000098
hsa-miR-512-5p	CACUCAGCCUUGAGGGCACUUUC	132	MIMAT0002822
hsa-miR-524-5p	CUACAAAGGGAAGCACUUUCUC	133	MIMAT0002849
hsa-miR-885-3p	AGGCAGCGGGGUGUAGUGGAUA	134	MIMAT0004948
hsa-miR-372	AAAGUGCUGCGACAUAUUGAGCGU	135	MIMAT0000724
hsa-miR-518a-5p, hsa-miR-527,	CUGCAAAGGGAAGCCCUUUC	136	MIMAT0005457
hsa-miR-1185	AGAGGAUACCCUUUGUAUGUU	137	MIMAT0005798
hsa-miR-518f	GAAAGCGCUUCUCUUUAGAGG	138	MIMAT0002842
hsa-miR-627	GUGAGUCUCUAAGAAAAGAGGA	139	MIMAT0003296
hsa-miR-181a-2*	ACCACUGACCGUUGACUGUACC	140	MIMAT0004558
hsa-miR-1205	UCUGCAGGGUUUGCUUUGAG	141	MIMAT0005869
hsa-miR-200b*	CAUCUUACUGGGCAGCAUUGGA	142	MIMAT0004571
hsa-miR-645	UCUAGGCUGGUACUGCUGA	143	MIMAT0003315
hsa-miR-649	AAACCUGUGUUGUUAAGAGUC	144	MIMAT0003319
hsa-miR-1206	UGUUCAUGUAGAUGUUUAAGC	145	MIMAT0005870
hsa-miR-1255b	CGGAUGAGCAAAGAAAGUGGUU	146	MIMAT0005945
hsa-miR-329	AACACACCUGGUUAACCUCUUU	147	MIMAT0001629

10

20

30

40

【表 1 - 4】

Coverage Human microRNA	Target sequence	SEQ ID No:	Target sequence accession
hsa-miR-498	UUUCAAGCCAGGGGCGUUUUUC	148	MIMAT0002824
hsa-miR-335	UCAAGAGCAAUACGAAAAUGU	149	MIMAT0000765
hsa-miR-199b-5p	CCCAGUGUUUAGACUAUCUGUUC	150	MIMAT0000263
hsa-miR-339-5p	UCCCUGUCCUCCAGGAGCUCACG	151	MIMAT0000764
hsa-miR-320a	AAAAGCUGGGUUGAGAGGGCGA	152	MIMAT0000510
hsa-miR-181d	AACAUUCAUUGUUGUCGGUGGGU	153	MIMAT0002821
hsa-miR-331-3p	GCCCCUGGGCCUAUCCUAGAA	154	MIMAT0000760
hsa-miR-302a	UAAGUGCUUCCAUUUUUGGUGA	155	MIMAT0000684
hsa-miR-548k	AAAAGUACUUGCGGAUUUUGCU	156	MIMAT0005882
hsa-miR-924	AGAGUCUUGUGAUGUCUUGC	157	MIMAT0004974
hsa-miR-339-3p	UGAGCGCCUCGACGACAGAGCCG	158	MIMAT0004702
hsa-miR-127-5p	CUGAAGCUCAGAGGGCUCUGAU	159	MIMAT0004604
hsa-miR-133b	UUUGGUCCCCUUAACCAGCUA	160	MIMAT0000770
hsa-miR-220a	CCACACCGUAUCUGACACUUU	161	MIMAT0000277
hsa-miR-422a	ACUGGACUUAGGGUCAGAAGGC	162	MIMAT0001339
hsa-miR-567	AGUAUGUUCUUGCGACAGAAC	163	MIMAT0003231
hsa-miR-493*	UUGUACAUGGUAGGCUUUCUU	164	MIMAT0002813
hsa-miR-216a	UAAUCUCAGCUGGCAACUGUGA	165	MIMAT0000273
hsa-miR-589	UGAGAACCACGUCUGCUCUGAG	166	MIMAT0004799
hsa-miR-382	GAAGUUGUUCGUGGUGGAUUCG	167	MIMAT0000737
hsa-miR-212	UAACAGUCUCCAGUCACGGCC	168	MIMAT0000269
hsa-miR-26b	UUCAAGUAAUUCAGGAUAGGU	169	MIMAT0000083
hsa-miR-363*	CGGGUGGAUCACGAUGCAAUUU	170	MIMAT0003385
hsa-miR-1263	AUGGUACCCUGGCAUACUGAGU	171	MIMAT0005915
hsa-miR-873	GCAGGAACUUGUGAGUCUCCU	172	MIMAT0004953
	CACUGUAGGUGAUGGUGAGAGUG	173	
hsa-miR-1183	GGCA		MIMAT0005828
hsa-miR-517c	AUCGUGCAUCCUUUUAGAGUGU	174	MIMAT0002866
hsa-miR-501-3p	AAUGCACCCGGGCAAGGAUUCU	175	MIMAT0004774
hsa-miR-378	ACUGGACUUGGAGUCAGAAGG	176	MIMAT0000732
hsa-miR-662	UCCACGCUUGUGGCCAGCAG	177	MIMAT0003325
hsa-miR-552	AACAGGUGACUGGUUAGACAA	178	MIMAT0003215
hsa-miR-134	UGUGACUGGUUGACCAGAGGGG	179	MIMAT0000447
hsa-miR-591	AGACCAUGGGUUCUAUUGU	180	MIMAT0003259
hsa-miR-26a-1*	CCUAUUCUUGGUUACUUGCACG	181	MIMAT0004499
hsa-miR-936	ACAGUAGAGGGAGGAUUCGCAG	182	MIMAT0004979
hsa-miR-195	UAGCAGCACAGAAUUAUUGGC	183	MIMAT0000461
hsa-miR-24-2*	UGCCUACUGAGCUGAAACACAG	184	MIMAT0004497
hsa-miR-148a*	AAAGUUCUGAGACACUCCGACU	185	MIMAT0004549
hsa-miR-450b-5p	UUUUGCAAUAUGUCCUGAAUA	186	MIMAT0004909
hsa-miR-143	UGAGAUGAAGCACUGUAGCUC	187	MIMAT0000435
hsa-miR-145*	GGAUUCCUGGAAUACUGUUCU	188	MIMAT0004601
hsa-miR-105*	ACGGAUGUUUGAGCAUGUGCUA	189	MIMAT0004516
hsa-miR-302c*	UUUAACAUGGGGGUACCUGCUG	190	MIMAT0000716
hsa-miR-576-3p	AAGAUGUGGAAAAUUGGAAUC	191	MIMAT0004796
hsa-miR-191*	GCUGCGCUUGGAUUUCGUCCCC	192	MIMAT0001618
hsa-miR-770-5p	UCCAGUACCACGUGUCAGGGCCA	193	MIMAT0003948
hsa-miR-542-5p	UCGGGGAUCAUCAUGUCACGAGA	194	MIMAT0003340
hsa-miR-659	CUUGGUUCAGGGAGGGUCCCCA	195	MIMAT0003337
hsa-miR-1227	CGUGCCACCCUUUCCCCAG	196	MIMAT0005580
hsa-miR-452*	CUCAUCUGCAAAGAAGUAAGUG	197	MIMAT0001636
hsa-miR-491-3p	CUUAUGCAAGAUCCCCUUCUAC	198	MIMAT0004765
hsa-miR-380*	UGGUUGACCAUAGAACAUGCGC	199	MIMAT0000734

10

20

30

40

【表 1 - 5】

Coverage Human microRNA	Target sequence	SEQ ID No:	Target sequence accession
hsa-miR-194*	CCAGUGGGGCGUCUGUUAUCUG	200	MIMAT0004671
hsa-miR-586	UAUGCAUUGUAUUUUUAGGUCC	201	MIMAT0003252
hsa-miR-668	UGUCACUCGGCUCGGCCACUAC	202	MIMAT0003881
hsa-miR-18a	UAAGGUGCAUCUAGUGCAGAUAG	203	MIMAT0000072
hsa-miR-29b-2*	CUGGUUUCACAUGGUGGCUUAG	204	MIMAT0004515
hsa-let-7b*	CUAUACAACCUACUGCCUCCCC	205	MIMAT0004482
hsa-miR-629*	GUUCUCCCAACGUAAGCCCAGC	206	MIMAT0003298
hsa-miR-1243	AACUGGAUCAAUUUAAGGAGUG	207	MIMAT0005894
hsa-miR-933	UGUGCGCAGGGAGACCUCCCC	208	MIMAT0004976
hsa-miR-181c*	AACCAUCGACCGUUGAGUGGAC	209	MIMAT0004559
hsa-miR-505	CGUCAACACUUGCUGGUUUCCU	210	MIMAT0002876
hsa-miR-562	AAAGUAGCUGUACCAUUUGC	211	MIMAT0003226
	CUGAAGUGAUGUGUAACUGAUCAG	212	
hsa-miR-573			MIMAT0003238
hsa-let-7a*	CUAUACAACCUACUGUCUUUC	213	MIMAT0004481
hsa-miR-376b	AUCAUAGAGGAAAUCUAGUU	214	MIMAT0002172
hsa-miR-27b*	AGAGCUUAGCUGAUUGGUGAAC	215	MIMAT0004588
hsa-miR-891a	UGCAACGAACCUAGGCCACUGA	216	MIMAT0004902
hsa-miR-532-5p	CAUGCCUUGAGUGUAGGACCGU	217	MIMAT0002888
hsa-miR-590-5p	GAGCUUAUUCAUAAAAGUGCAG	218	MIMAT0003258
hsa-miR-302b	UAAGUGCUUCCAUGUUUUAGUAG	219	MIMAT0000715
hsa-miR-589*	UCAGAACAAGUCCGGUUCUCCAGA	220	MIMAT0003256
hsa-miR-558	UGAGCUGCUGUACCAAAAU	221	MIMAT0003222
hsa-miR-193b	AACUGGCCCUCUAAAGUCCCGCU	222	MIMAT0002819
hsa-miR-126	UCGUACCGUGAGUAAUUAUGCG	223	MIMAT0000445
hsa-miR-634	AACCAGCACCCCAACUUUGGAC	224	MIMAT0003304
hsa-miR-1245	AAGUGAUCUAAAGGCCUACAU	225	MIMAT0005897
hsa-miR-21	UAGCUUAUCAGACUGAUGUUGA	226	MIMAT0000076
hsa-miR-875-3p	CCUGGAAACACUGAGGUUGUG	227	MIMAT0004923
hsa-miR-556-3p	AUAUUAACAUUAGCUCAUCUUU	228	MIMAT0004793
hsa-miR-650	AGGAGGCAGCGCUCUCAGGAC	229	MIMAT0003320
	AGGGAUCGCGGGCGGGUGGCGGC	230	
hsa-miR-638	CU		MIMAT0003308
hsa-miR-518a-3p	GAAAGCGCUUCCCUUUGCUGGA	231	MIMAT0002863
hsa-miR-31	AGGCAAGAUGCUGGCAUAGCU	232	MIMAT0000089
hsa-miR-1258	AGUUAGGAUUAGGUCGUGGAA	233	MIMAT0005909
hsa-miR-767-5p	UGCACCAUGGUUGUCUGAGCAUG	234	MIMAT0003882
hsa-miR-188-5p	CAUCCCUUGCAUGGUGGAGGG	235	MIMAT0000457
hsa-miR-556-5p	GAUGAGCUCAUUGUAAUUAUGAG	236	MIMAT0003220
hsa-miR-361-5p	UUUUCAGAAUCUCCAGGGGUAC	237	MIMAT0000703
	GAUGAUGAUGGCAGCAAAUUCUGA	238	
hsa-miR-1272	AA		MIMAT0005925
hsa-miR-15b	UAGCAGCACAUAUGGUUUACA	239	MIMAT0000417
	AAGUAGUUGGUUUUGAUGAGAUGG		
hsa-miR-1244	UU	240	MIMAT0005896
hsa-miR-767-3p	UCUGCUCAUACCCCAUGGUUUUCU	241	MIMAT0003883
hsa-let-7i*	CUGCGCAAGCUACUGCCUUGCU	242	MIMAT0004585
hsa-miR-920	GGGAGCUGUGGAAGCAGUA	243	MIMAT0004970
hsa-miR-587	UUUCCAUAAGGUGAUGAGUCAC	244	MIMAT0003253
hsa-miR-340*	UCCGUCUCAGUUACUUUAUAGC	245	MIMAT0000750
hsa-miR-875-5p	UAUACCUCAGUUUUUAUCAGGUG	246	MIMAT0004922
hsa-miR-27b	UUCACAGUGGCUAAGUUCUGC	247	MIMAT0000419
hsa-miR-1248	ACCUUCUUGUAUAAGCACUGUGCU	248	MIMAT0005900

10

20

30

40

【表 1 - 6】

Coverage Human microRNA	Target sequence	SEQ ID No:	Target sequence accession
	AAA		
hsa-miR-582-5p	UUACAGUUGUUAACAGUUACU	249	MIMAT0003247
hsa-miR-22*	AGUUCUUCAGUGGCAAGCUUUA	250	MIMAT0004495
hsa-miR-223	UGUCAGUUGUCAAUACCCCA	251	MIMAT0000280
hsa-miR-548c-5p	AAAAGUAAUUGCGGUUUUUGCC	252	MIMAT0004806
hsa-miR-92a	UAUUGCACUUGUCCCGGCCUGU	253	MIMAT0000092
hsa-miR-526b	CUCUUGAGGGAAGCACUUUCUGU	254	MIMAT0002835
hsa-miR-24	UGGCUCAGUUCAGCAGGAACAG	255	MIMAT0000080
	GCUGGUUUAUAUGGUGGUUAGA	256	
hsa-miR-29b-1*			MIMAT0004514
hsa-miR-526b*	GAAAGUGCUUCCUUUUAGAGGC	257	MIMAT0002836
hsa-miR-877*	UCCUCUUCUCCUCCUCCAG	258	MIMAT0004950
hsa-miR-182	UUUGGCAAUGGUAGAACUCACACU	259	MIMAT0000259
hsa-miR-133a	UUUGGUCCCCUUAACAGCUG	260	MIMAT0000427
hsa-miR-124*	CGUGUUCACAGCGGACCUUGAU	261	MIMAT0004591
hsa-miR-1236	CCUCUCCCCUUGUCUCCAG	262	MIMAT0005591
hsa-miR-578	CUUCUUGUGCUUUGGAUUGU	263	MIMAT0003243
hsa-miR-769-5p	UGAGACCUCUGGGUUCUGAGCU	264	MIMAT0003886
hsa-miR-599	GUUGUGUCAGUUUAUCAAAC	265	MIMAT0003267
hsa-miR-192*	CUGCCAAUCCAUAAGGUCACAG	266	MIMAT0004543
hsa-miR-614	GAACGCCUGUUCUUGCCAGGUGG	267	MIMAT0003282
hsa-miR-643	ACUUGUAUGCUAGCUCAGGUAG	268	MIMAT0003313
hsa-miR-541	UGGUGGGCACAGAAUCUGGACU	269	MIMAT0004920
hsa-miR-92a-2*	GGUGGGGAUUUGUUGCAUUAAC	270	MIMAT0004508
hsa-miR-323-3p	CACAUUACACGGUCGACCUCU	271	MIMAT0000755
hsa-miR-454*	ACCCUAUCAAUUUGUCUCUGC	272	MIMAT0003884
hsa-miR-518c*	UCUCUGGAGGGAAGCACUUUCUG	273	MIMAT0002847
	CUAGUGAGGGACAGAACCAGGAUU	274	
hsa-miR-921	C		MIMAT0004971
hsa-miR-566	GGGCGCCUGUGAUCCCAAC	275	MIMAT0003230
hsa-miR-520f	AAGUGCUUCCUUUAGAGGGUU	276	MIMAT0002830
hsa-miR-663	AGGCGGGGCGCCGCGGGACCGC	277	MIMAT0003326
hsa-miR-203	GUGAAUGUUUAGGACCACUAG	278	MIMAT0000264
	AGGGUGGUGUUGGGACAGCUCC	279	
hsa-miR-608	GU		MIMAT0003276
hsa-miR-513c	UUCUCAAGGAGGUGUCGUUUAU	280	MIMAT0005789
hsa-miR-95	UUCAACGGGUUUUUUAGGCA	281	MIMAT0000094
hsa-miR-216b	AAUUCUCUGCAGGCAAUUGUGA	282	MIMAT0004959
hsa-let-7d*	CUAUACGACCUGCUGCCUUUCU	283	MIMAT0004484
hsa-miR-142-3p	UGUAGUGUUUCCUACUUUAUGGA	284	MIMAT0000434
hsa-miR-20a	UAAAGUGCUUAUAGUGCAGGUAG	285	MIMAT0000075
hsa-miR-505*	GGGAGCCAGGAAGUAUUGAUGU	286	MIMAT0004776
hsa-miR-152	UCAGUGCAUGACAGAACUUGG	287	MIMAT0000438
hsa-miR-125b-2*	UCACAAGUCAGGCUCUUGGGAC	288	MIMAT0004603
hsa-miR-379	UGGUAGACUAUGGAACGUAGG	289	MIMAT0000733
hsa-miR-20b	CAAAGUGCUCAUAGUGCAGGUAG	290	MIMAT0001413
hsa-miR-636	UGUGCUUGCUCGUCCCGCCGCA	291	MIMAT0003306
hsa-miR-371-3p	AAGUGCCGCCAUCUUUUGAGUGU	292	MIMAT0000723
hsa-miR-302e	UAAGUGCUUCCAUGCUU	293	MIMAT0005931
hsa-miR-452	AACUGUUUGCAGAGGAAACUGA	294	MIMAT0001635
hsa-miR-21*	CAACACCAGUCGAUGGGCUGU	295	MIMAT0004494
hsa-miR-324-3p	ACUGCCCCAGGUGCUGCUGG	296	MIMAT0000762
hsa-miR-140-3p	UACCACAGGGUAGAACCACGG	297	MIMAT0004597

10

20

30

40

【表 1 - 7】

Coverage Human microRNA	Target sequence	SEQ ID No:	Target sequence accession
hsa-miR-516b*, hsa-miR-516a-3p,	UGCUUCCUUUCAGAGGGU	298	MIMAT0002860
hsa-miR-191	CAACGGAAUCCCAAAAGCAGCUG	299	MIMAT0000440
hsa-miR-621	GGCUAGCAACAGCGCUUACCU	300	MIMAT0003290
hsa-miR-155	UUAUUGCUAAUCGUGAUAGGGGU	301	MIMAT0000646
hsa-miR-16-2*	CCAAUAAUACUGUGCUGCUUUA	302	MIMAT0004518
hsa-miR-19b-1*	AGUUUUGCAGGUUUGCAUCCAGC	303	MIMAT0004491
hsa-miR-302d	UAAGUGCUUCCAUGUUUGAGUGU	304	MIMAT0000718
hsa-miR-631	AGACCUGGCCAGACCUCAGC	305	MIMAT0003300
hsa-miR-550*	UGUCUUACUCCUCAGGCACAU	306	MIMAT0003257
hsa-miR-222*	CUCAGUAGCCAGUGUAGAUCU	307	MIMAT0004569
hsa-let-7g*	CUGUACAGGCCACUGCCUUGC	308	MIMAT0004584
hsa-miR-602	GACACGGGCGACAGCUGCGGCC	309	MIMAT0003270
hsa-miR-130b	CAGUGCAAUGAUGAAAGGGCAU	310	MIMAT0000691
hsa-miR-34a*	CAAUCAGCAAGUAUACUGCCCU	311	MIMAT0004557
hsa-miR-124	UAAGGCACGCGGUGAAUGCC	312	MIMAT0000422
hsa-miR-598	UAAAGUGCUGACAGUGCAUCA	313	MIMAT0003266
hsa-miR-149	UCUGGCUCGUGUCUUCACUCCC	314	MIMAT0000450
hsa-miR-28-5p	AAGGAGCUCACAGUCUAUUGAG	315	MIMAT0000085
hsa-let-7f-1*	CUAUACAUCUAUUGCCUCCC	316	MIMAT0004486
hsa-miR-19b-2*	AGUUUUGCAGGUUUGCAUUA	317	MIMAT0004492
hsa-miR-135a	UAUGGCUUUUUUAUCCUAUGUGA	318	MIMAT0000428
hsa-let-7a	UGAGGUAGUAGGUUGUAUAGUU	319	MIMAT0000062
hsa-miR-106b	UAAAGUGCUGACAGUGCAGAU	320	MIMAT0000680
hsa-miR-2110	UUGGGGAAACGGCCGCUGAGUG	321	MIMAT0010133
hsa-miR-130a*	UUCACAUUGUGCUACUGUCUGC	322	MIMAT0004593
hsa-miR-1184	CCUGCAGCGACUUGAUGGCUUCC	323	MIMAT0005829
hsa-miR-551a	GCGACCCACUCUUGGUUUCCA	324	MIMAT0003214
hsa-miR-519b-3p	AAAGUGCAUCCUUUUAGAGGUU	325	MIMAT0002837
hsa-miR-210	CUGUGCGUGUGACAGCGGCUGA	326	MIMAT0000267
hsa-miR-503	UAGCAGCGGGAAACAGUUCUGCAG	327	MIMAT0002874
hsa-miR-549	UGACAACUAUGGAUGAGCUCU	328	MIMAT0003333
hsa-miR-517*	CCUCUAGAUGGAAGCACUGUCU	329	MIMAT0002851
hsa-miR-425	AAUGACACGAUCACUCCGUUGA	330	MIMAT0003393
hsa-miR-153	UUGCAUAGUCACAAAAGUGAUC	331	MIMAT0000439
hsa-miR-125a-5p	UCCUGAGACCCUUUAACCUGUGA	332	MIMAT0000443
hsa-miR-520a-5p	CUCCAGAGGGAAGUACUUUCU	333	MIMAT0002833
hsa-miR-198	GGUCCAGAGGGGAGAUAGGUUC	334	MIMAT0000228
hsa-miR-571	UGAGUUGGCCAUCUGAGUGAG	335	MIMAT0003236
hsa-miR-30b	UGUAAACAUCUACACUCAGCU	336	MIMAT0000420
hsa-miR-1	UGGAAUGUAAAGAAGUAUGUAU	337	MIMAT0000416
hsa-miR-379*	UAUGUAACAUGGUCCACUAACU	338	MIMAT0004690
hsa-miR-557	GUUUGCACGGGUGGGCCUUGUCU	339	MIMAT0003221
hsa-miR-378*	CUCCUGACUCCAGGUCCUGUGU	340	MIMAT0000731
hsa-miR-490-3p	CAACCUGGAGGACUCCAUGCUG	341	MIMAT0002806
hsa-miR-510	UACUCAGGAGAGUGGCAUACAC	342	MIMAT0002882
hsa-miR-1201	AGCCUGAUUAAACACAUGCUCUGA	343	MIMAT0005864
hsa-miR-1271	CUUGGCACCUAGCAAGCACUCA	344	MIMAT0005796
hsa-miR-200a*	CAUCUUACCGGACAGUGCUGGA	345	MIMAT0001620
hsa-miR-758	UUUGUGACCUGGUCCACUAACC	346	MIMAT0003879
hsa-miR-497	CAGCAGCACACUGUGGUUUUGU	347	MIMAT0002820
hsa-miR-525-5p	CUCCAGAGGGAGUCACUUUCU	348	MIMAT0002838
hsa-miR-220c	ACACAGGGCUGUUGUGAAGACU	349	MIMAT0004915

10

20

30

40

【表 1 - 8】

Coverage Human microRNA	Target sequence	SEQ ID No:	Target sequence accession
hsa-miR-24-1*	UGCCUACUGAGCUGAUUACAGU	350	MIMAT0000079
hsa-miR-409-3p	GAAUGUUGCUCGGUGAACCCCU	351	MIMAT0001639
hsa-let-7f	UGAGGUAGUAGAUUGUAGUU	352	MIMAT0000067
hsa-miR-675*	CUGUAGCCCCUACCGCUCA	353	MIMAT00006790
hsa-miR-25	CAUUGCACUUGUCUGGUCUGA	354	MIMAT0000081
hsa-miR-375	UUUGUUCGUUCGGCUCGCGUGA	355	MIMAT0000728
hsa-miR-455-5p	UAUGUGCCUUUGGACUACAUCG	356	MIMAT0003150
hsa-miR-328	CUGGCCCUUCUGCCCUUCCGU	357	MIMAT0000752
hsa-miR-574-3p	CACGCUCAUGCACACACCCACA	358	MIMAT0003239
hsa-miR-671-5p	AGGAAGCCCUUGAGGGGCGUGAG	359	MIMAT0003880
hsa-miR-99b	CACCCGUAGAACCAGCUUGCG	360	MIMAT0000689
hsa-miR-147b	GUGUGCGGAAUUGCUUCUGCUA	361	MIMAT0004928
hsa-miR-450b-3p	UUGGGAUCAUUUUGCAUCCAUA	362	MIMAT0004910
hsa-miR-629	UGGGUUUACGUUGGGAGAACU	363	MIMAT0004810
hsa-miR-663b	GGUGGCCCGGCCGUGCCUGAGG	364	MIMAT0005867
hsa-miR-330-5p	UCUCUGGGCCUGUGUCUUAGGC	365	MIMAT0004693
hsa-miR-34c-3p	AAUCACUAACCACACGGCCAGG	366	MIMAT0004677
hsa-miR-146b-3p	UGCCUGUGGACUCAGUUCUGG	367	MIMAT0004766
hsa-miR-592	UUGUGUCAAUAUGCGAUGAU	368	MIMAT0003260
hsa-miR-30d	UGUAAACAUCGCCGACUGGAAG	369	MIMAT0000245
hsa-miR-555	AGGGUAAGCUGAACCCUCUGAU	370	MIMAT0003219
hsa-miR-23a	AUCACAUUGCCAGGGAUUUCC	371	MIMAT0000078
hsa-miR-101*	CAGUUAUCACAGUGCUGAUGCU	372	MIMAT0004513
hsa-miR-197	UUCACCACCUUCUCCACCCAGC	373	MIMAT0000227
hsa-miR-487a	AAUCAUACAGGGACAUCCAGUU	374	MIMAT0002178
hsa-miR-512-3p	AAGUGCUGUCAUAGCUGAGGUC	375	MIMAT0002823
hsa-miR-520h	ACAAAGUGCUUCCCUUUAGAGU	376	MIMAT0002867
hsa-miR-92b	UAUUGCACUCGUCCCGGCCUCC	377	MIMAT0003218
hsa-miR-138	AGCUGGUGUUGUGAAUCAGGCCG	378	MIMAT0000430
hsa-miR-196a	UAGGUAGUUUAUGUUGUUGGG	379	MIMAT0000226
hsa-miR-652	AAUGGCGCCACUAGGGUUGUG	380	MIMAT0003322
hsa-let-7a-2*	CUGUACAGCCUCCUAGCUUCC	381	MIMAT0010195
hsa-miR-105	UCAAAUGCUCAGACUCCUGUGGU	382	MIMAT0000102
hsa-miR-301b	CAGUGCAAUGAUUUUGUCAAGC	383	MIMAT0004958
hsa-miR-337-5p	GAACGGCUUCAUACAGGAGUU	384	MIMAT0004695
hsa-miR-630	AGUUAUCUGUACCAGGGAAGGU	385	MIMAT0003299
hsa-miR-296-3p	GAGGGUUGGGUGGAGGCUCUCC	386	MIMAT0004679
hsa-let-7i	UGAGGUAGUAGUUUGUGCUGUU	387	MIMAT0000415
hsa-miR-489	GUGACAUCAUAUACGGCAGC	388	MIMAT0002805
hsa-miR-504	AGACCCUGGUCUGCACUCUAUC	389	MIMAT0002875
hsa-miR-15b*	CGAAUCAUUUUUGCUGCUCUA	390	MIMAT0004586
hsa-miR-147	GUGUGUGGAAUUGCUUCUGC	391	MIMAT0000251
hsa-miR-376a*	GUAGAUUCUCCUUCUAUGAGUA	392	MIMAT0003386
hsa-miR-125b-1*	ACGGGUUAGGCUCUUGGGAGCU	393	MIMAT0004592
hsa-miR-146a*	CCUCUGAAAUUCAGUUCUUCAG	394	MIMAT0004608
hsa-miR-187*	GGCUACAACACAGGACCCGGGC	395	MIMAT0004561
hsa-miR-302c	UAAGUGCUUCCAUGUUUCAGUGG	396	MIMAT0000717
hsa-miR-520b	AAAGUGCUUCCUUUUAGAGGG	397	MIMAT0002843
hsa-miR-518b	CAAAGCGCUCUUUUUAGAGGU	398	MIMAT0002844
hsa-miR-886-5p	CGGGUCGGAGUUAGCUAAGCGG	399	MIMAT0004905
hsa-miR-34c-5p	AGGCAGUGUAGUUAGCUGAUUGC	400	MIMAT0000686
hsa-miR-16	UAGCAGCACGUAAAUUUGGCG	401	MIMAT0000069
hsa-miR-30e*	CUUUCAGUCGGAUGUUUACAGC	402	MIMAT0000693

10

20

30

40

【表 1 - 9】

Coverage Human microRNA	Target sequence	SEQ ID No.	Target sequence accession
hsa-miR-641	AAAGACAUAGGAUAGAGUACCCUC	403	MIMAT0003311
hsa-miR-188-3p	CUCCCAUAGCAGGGUUUGCA	404	MIMAT0004613
hsa-miR-1203	CCCGAGCCAGGAUGCAGCUC	405	MIMAT0005866
hsa-miR-92b*	AGGGACGGGACGCGGUGCAGUG	406	MIMAT0004792
hsa-miR-548a-5p	AAAAGUAAUUGCGAGUUUUACC	407	MIMAT0004803
hsa-miR-96	UUUGGCACUAGCACAUUUUUGCU	408	MIMAT0000095
hsa-miR-23b	AUCACAUUGCCAGGGAUUACC	409	MIMAT0000418
hsa-miR-219-1-3p	AGAGUUGAGUCUGGACGUCCCG	410	MIMAT0004567
hsa-miR-1266	CCUCAGGGCUGUAGAACAGGGCU	411	MIMAT0005920
hsa-miR-548j	AAAAGUAAUUGCGGUCUUUGGU	412	MIMAT0005875
hsa-miR-495	AAACAAACAUUGGUCACUUCUU	413	MIMAT0002817
hsa-miR-331-5p	CUAGGUUAGGUCCAGGGAUCC	414	MIMAT0004700
hsa-miR-34b*	UAGGCAGUGUCAUUAGCUGAUUG	415	MIMAT0000685
hsa-miR-500	UAAUCCUUGCUACCUGGGUGAGA	416	MIMAT0004773
hsa-miR-601	UGGUCUAGGAUUGUUGGAGGAG	417	MIMAT0003269
hsa-miR-135b*	AUGUAGGGCUAAAAGCCAUGGG	418	MIMAT0004698
hsa-let-7e	UGAGGUAGGAGGUUGUAUAGUU	419	MIMAT0000066
hsa-miR-876-3p	UGGUGGUUUACAAAGUAAUUA	420	MIMAT0004925
hsa-miR-29a*	ACUGAUUUUUUUGGUGUUCAG	421	MIMAT0004503
hsa-miR-515-5p	UUCUCCAAAAGAAAGCACUUUCUG	422	MIMAT0002826
hsa-miR-96*	AAUCAUGUGCAGUGCCAAUAUG	423	MIMAT0004510
hsa-miR-411*	UAUGUAAACACGGUCCACUAACC	424	MIMAT0004813
hsa-miR-15a*	CAGGCCAUAAUUGUGCUGCCUCA	425	MIMAT0004488
hsa-miR-296-5p	AGGGCCCCCCCUCAAUCCUGU	426	MIMAT0000690
hsa-miR-122*	AACGCCAUUAUCACACUAAUA	427	MIMAT0004590
hsa-miR-499-3p	AAACUACAGCAGUGUGUGCU	428	MIMAT0004772
hsa-miR-654-5p	UGGUGGCGCCGACGAACAUUGUC	429	MIMAT0003330
hsa-miR-942	UCUUCUCUGUUUUUGGCCAUGUG	430	MIMAT0004985
hsa-miR-496	UGAGUAAUACAUGGCCAAUUCUC	431	MIMAT0002818
hsa-miR-376c	AACAUAGAGGAAAUUCCACGU	432	MIMAT0000720
hsa-miR-106a*	CUGCAAUGUAAGCACUUCUAC	433	MIMAT0004517
hsa-let-7c	UGAGGUAGUAGGUUGUAUGGUU	434	MIMAT0000064
hsa-miR-615-5p	GGGGUCCCCGGUGUCUGGAUC	435	MIMAT0004804
hsa-miR-125a-3p	ACAGGUGAGGUUCUUGGGAGCC	436	MIMAT0004602
hsa-miR-543	AAACAUUCGCGGUGCACUUCUU	437	MIMAT0004954
hsa-miR-484	UCAGGCUCAGUCCCCUCCGAU	438	MIMAT0002174
hsa-miR-502-5p	AUCCUUGCUAUCUGGGUGCUA	439	MIMAT0002873
hsa-miR-19b	UGUGCAAUCCAUGCAAACUGA	440	MIMAT0000074
hsa-miR-523	GAACGCGCUUCCCUAUAGAGGGU	441	MIMAT0002840
hsa-miR-615-3p	UCCGAGCCUGGGUCUCCUUCU	442	MIMAT0003283
hsa-miR-564	AGGCACGGUGUCAGCAGGC	443	MIMAT0003228
hsa-miR-1269	CUGGACUGAGCCGUGCUACUGG	444	MIMAT0005923
hsa-miR-130b*	ACUCUUUCCUGUUGCACUAC	445	MIMAT0004680
hsa-miR-30a*	CUUUCAGUCGGAUGUUUGCAGC	446	MIMAT0000088
hsa-miR-509-3p	UGAUUGGUACGUCUGGGUAG	447	MIMAT0002881
hsa-miR-412	ACUUCACCUGGUCCACUAGCCGU	448	MIMAT0002170
hsa-miR-526a, hsa-miR-518d-5p & hsa-miR-520c-5p	CUCUAGAGGGAAGCACUUUCUG	449	MIMAT0002845
hsa-miR-33b*	CAGUGCCUCGGCAGUGCAGCCC	450	MIMAT0004811
hsa-miR-877	GUAGAGGAGAUGGCGCAGGG	451	MIMAT0004949
hsa-miR-325	CCUAGUAGGUGUCCAGUAAGUGU	452	MIMAT0000771
hsa-miR-125b	UCCUGAGACCCUAACUUGUGA	453	MIMAT0000423

10

20

30

40

【表 1 - 10】

Coverage Human microRNA	Target sequence	SEQ ID No:	Target sequence accession
hsa-miR-1182	GAGGGUCUUGGGAGGGAUGUGAC	454	MIMAT0005827
hsa-miR-107	AGCAGCAUUGUACAGGGCUAUCA	455	MIMAT0000104
hsa-miR-488	UUGAAAGGCUAUUUUCUUGGUC	456	MIMAT0004763
hsa-miR-93*	ACUGCUGAGCUAGCACUCCCCG	457	MIMAT0004509
hsa-miR-516a-5p	UUCUCGAGGAAAGAAGCACUUUC	458	MIMAT0004770
hsa-miR-887	GUGAACGGGCGCCAUCCCCGAGG	459	MIMAT0004951
hsa-miR-885-5p	UCCAUUACACUACCCUGCCUCU	460	MIMAT0004947
hsa-miR-888*	GACUGACACCUCUUUGGGUGAA	461	MIMAT0004917
hsa-miR-185	UGGAGAGAAAGGCAGUCCUGA	462	MIMAT0000455
hsa-miR-138-2*	GCUAUUUACGACACACAGGGUU	463	MIMAT0004596
hsa-miR-922	GCAGCAGAGAAUAGGACUACGUC	464	MIMAT0004972
hsa-miR-200c*	CGUCUUACCCAGCAGUGUUUGG	465	MIMAT0004657
hsa-miR-508-3p	UGAUUGUAGCCUUUUGGAGUAGA	466	MIMAT0002880
hsa-miR-449a	UGGCAGUGUAUUGUAGCUGGU	467	MIMAT0001541
hsa-miR-200c	UAAUACUGCCGGGUAUUGAUGGA	468	MIMAT0000617
hsa-miR-145	GUCCAGUUUUCCCAGGAAUCCCU	469	MIMAT0000437
hsa-miR-218	UUGUGCUUGAUCUAACCAUGU	470	MIMAT0000275
hsa-miR-548b-3p	CAAGAACCUCAGUUGCUUUUGU	471	MIMAT0003254
hsa-miR-34a	UGGCAGUGUCUAGCUGGUUGU	472	MIMAT0000255
hsa-miR-205	UCCUUCAUUCCACCGGAGUCUG	473	MIMAT0000266
hsa-miR-423-3p	AGCUCGGUCUGAGGCCCCUCAGU	474	MIMAT0001340
hsa-miR-487b	AAUCGUACAGGGUCAUCCACUU	475	MIMAT0003180
hsa-miR-708	AAGGAGCUUACAUCUAGCUGGG	476	MIMAT0004926
hsa-miR-519e	AAGUGCCUCCUUUUAGAGUGUU	477	MIMAT0002829
hsa-miR-610	UGAGCUAAAUGUGUGCUGGGA	478	MIMAT0003278
hsa-miR-371-5p	ACUCAAACUGUGGGGGCACU	479	MIMAT0004687
hsa-miR-199a-5p	CCCAGUGUUCAGACUACCGUUC	480	MIMAT0000231
hsa-miR-488*	CCCAGAUAAUGGCACUCUCAA	481	MIMAT0002804
hsa-miR-1260	AUCCACCUCUGCCACCA	482	MIMAT0005911
hsa-miR-520c-3p	AAAGUGCUUCCUUUUAGAGGGU	483	MIMAT0002846
hsa-miR-616*	ACUCAAAACCCUUCAGUGACUU	484	MIMAT0003284
hsa-miR-766	ACUCCAGCCCCACAGCCUCAGC	485	MIMAT0003888
hsa-miR-141*	CAUCUUCAGUACAGUGUUGGA	486	MIMAT0004598
hsa-miR-622	ACAGUCUGCUGAGGUUGGAGC	487	MIMAT0003291
hsa-miR-17*	ACUGCAGUGAAGGCACUUGUAG	488	MIMAT0000071
hsa-miR-509-3-5p	UACUGCAGACGUGGCAAUCAUG	489	MIMAT0004975
hsa-miR-141	UAACACUGUCUGGUAAAGAUGG	490	MIMAT0000432
hsa-miR-580	UUGAGAAUGAUGAAUCAUUAGG	491	MIMAT0003245
hsa-miR-517a	AUCGUGCAUCCUUUAGAGUGU	492	MIMAT0002852
hsa-miR-204	UUCCCUUUGUCAUCCUAUGCCU	493	MIMAT0000265
hsa-miR-376a	AUCAUAGAGGAAAUCCACGU	494	MIMAT0000729
hsa-miR-335*	UUUUUCAUUUUGCUCCUGACC	495	MIMAT0004703
hsa-miR-214	ACAGCAGGCACAGACAGGCAGU	496	MIMAT0000271
hsa-miR-342-3p	UCUCACACAGAAUCCGACCCGU	497	MIMAT0000753
hsa-miR-326	CCUCUGGGCCCUUCCUCCAG	498	MIMAT0000756
hsa-miR-9	UCUUUGGUUAUCUAGCUGUAUGA	499	MIMAT0000441
hsa-miR-10b*	ACAGAUUCGAUUCUAGGGGAU	500	MIMAT0004556
hsa-miR-23b*	UGGGUUCUGGCAUGCUGAUUU	501	MIMAT0004587
hsa-miR-342-5p	AGGGGUGCUAUCUGUGAUUGA	502	MIMAT0004694
hsa-miR-449b	AGGCAGUGUAUUGUAGCUGGC	503	MIMAT0003327
hsa-miR-154	UAGGUUAUCCGUGUUGCCUUCG	504	MIMAT0000452
hsa-miR-450a	UUUUGCGAUGUGUCCUAAUUAU	505	MIMAT0001545
hsa-miR-99a*	CAAGCUCGCUUCUAUGGGUCUG	506	MIMAT0004511

10

20

30

40

【表 1 - 1 1】

Coverage Human microRNA	Target sequence	SEQ ID No.	Target sequence accession
hsa-miR-99a	AACCCGUAGAUCCGAUCUUGUG GGCGGAGGGAAGUAGGUCCGUUG GU	507 508	MIMAT0000097
hsa-miR-658	ACUGCCCUAAGUGCUCCUUCUGG	509	MIMAT0003336
hsa-miR-18a*	AAAAGCUGGGUUGAGAGGGCAA	510	MIMAT0005792
hsa-miR-320b	AGAGAAGAAGAUCAAGCCUGCA	511	MIMAT0005904
hsa-miR-1253	UUAGGGCCCCUGGCUCUCCUCC	512	MIMAT0005794
hsa-miR-1296	UGGAUUUCUUUGUGAAUCACCA	513	MIMAT0004924
hsa-miR-876-5p	CUGUUGCCACUAACCUCUACCU	514	MIMAT0004946
hsa-miR-744*	CGUGUAUUUGACAAGCUGAGUU	515	MIMAT0004570
hsa-miR-223*	AACAUUCAUUGCUGUGGGU	516	MIMAT0000257
hsa-miR-181b	UAGUAGACCGUAUAGCGUACG	517	MIMAT0003329
hsa-miR-411	AGCUACAUUGUCUGUGGGUUUC	518	MIMAT0000278
hsa-miR-221	AUGAUCCAGGAACCUGCCUCU	519	MIMAT0003310
hsa-miR-640	CUUUUUGCGGUCUGGGCUUGC	520	MIMAT0000242
hsa-miR-100*	CAAGCUUGUAUCUAUAGGUAUG	521	MIMAT0004512
hsa-miR-199a-3p & hsa- miR-199b-3p	ACAGUAGUCUGCACAUUGGUUA	522	MIMAT0000232
hsa-miR-1208	UCACUGUUCAGACAGGCGGA	523	MIMAT0005873
hsa-miR-346	UGUCUGCCCGCAUGCCUGCCUCU	524	MIMAT0000773
hsa-miR-506	UAAGGCACCCUUCUGAGUAGA	525	MIMAT0002878
hsa-miR-140-5p	CAGUGGUUUUACCCUAUGGUAG	526	MIMAT0000431
hsa-miR-424*	CAAACGUGAGGCGCUGCUAU	527	MIMAT0004749
hsa-miR-632	GUGUCUGCUUCCUGUGGA	528	MIMAT0003302
hsa-miR-1267	CCUGUUGAAGUGUAUCCCA	529	MIMAT0005921
hsa-miR-299-5p	UGGUUUACCGUCCCAUACAU	530	MIMAT0002890
hsa-miR-943	CUGACUGUUGCCGUCCUCCAG	531	MIMAT0004986
hsa-miR-646	AAGCAGCUGCCUCUGAGGC	532	MIMAT0003316
hsa-miR-517b	UCGUGCAUCCCUUAGAGUGUU	533	MIMAT0002857
hsa-miR-760	CGGCUCUGGGUCUGUGGGA AGGCACCAGCCAGGCAUUGCUCAG C	534 535	MIMAT0004957 MIMAT0003261
hsa-miR-593*	AGCUACAUCUGGCUACUGGGU	536	MIMAT0000279
hsa-miR-222	ACCGUGGCUUUCGAUUGUUACU	537	MIMAT0004594
hsa-miR-132*	UGAGAACUGAAUCCAUAGGCU	538	MIMAT0002809
hsa-miR-146b-5p	CAAAGCGCUUCUCUUUAGAGUGU	539	MIMAT0002848
hsa-miR-518c	UAGGUAGUUUCCUGUUGUUGGG	540	MIMAT0001080
hsa-miR-196b	GCUAGUCCUGACUCAGCCAGU	541	MIMAT0003217
hsa-miR-554	UGAAGGUCUACUGUGUGCCAGG	542	MIMAT0003161
hsa-miR-493	AUCUGGAGGUAAGAAGCACUUU	543	MIMAT0002859
hsa-miR-516b	GGGUUCCUGGGGAUGGGAUUU	544	MIMAT0004496
hsa-miR-23a*	AGGUUGGGAUCGGUUGCAUGCU	545	MIMAT0004507
hsa-miR-92a-1*	CUUAGCAGGUUGUAUUUAUUAU	546	MIMAT0004956
hsa-miR-374b*	GCUACUUCACAACACAGGGGCC	547	MIMAT0004607
hsa-miR-138-1*	AAAAGUGCUUACAGUGCAGGUAG	548	MIMAT0000103
hsa-miR-106a	AGACUCCCAUUUGAAGGUGGC	549	MIMAT0003286
hsa-miR-617	UGAGGUAGUAGUUUGUACAGUU	550	MIMAT0000414
hsa-let-7g	AACAUUCAACGCUGUCGGUGAGU	551	MIMAT0000256
hsa-miR-181a	CAGGUCGUCUUGCAGGGCUUCU	552	MIMAT0004757
hsa-miR-431*	UUAUGGUUUUGCCUGGGACUGAG	553	MIMAT0003249
hsa-miR-584	ACUGUAGUAUGGGCACUCCAG	554	MIMAT0004752
hsa-miR-20b*	GGUGCAGUGCUGCAUCUCUGGU	555	MIMAT0004599
hsa-miR-143*	CGCGGGUGCUUACUGACCCUU	556	MIMAT0004906
hsa-miR-886-3p			

10

20

30

40

【表 1 - 1 2】

Coverage Human microRNA	Target sequence	SEQ ID No:	Target sequence accession
hsa-let-7c*	UAGAGUUACACCCUGGGAGUUA	557	MIMAT0004483
hsa-miR-941	CACCCGGCUGUGUGCACAUGUGC	558	MIMAT0004984
hsa-miR-214*	UGCCUGUCUACACUUGCUGUGC	559	MIMAT0004564
hsa-miR-151-3p	CUAGACUGAAGCUCCUUGAGG	560	MIMAT0000757
hsa-miR-1468	CUCGGUUUGCCUGUUUCGUG	561	MIMAT0006789
hsa-miR-639	AUCGUGCGGUUGCGAGCGCUGU	562	MIMAT0003309
hsa-miR-494	UGAAACAUACACGGGAAACCUC	563	MIMAT0002816
hsa-miR-183*	GUGAAUUACCGAAGGGCCAUAA	564	MIMAT0004560
hsa-miR-7-2*	CAACAAUCCAGUCUACCUAA	565	MIMAT0004554
hsa-miR-454	UAGUGCAAUUAUUGCUUAUAGGU	566	MIMAT0003885
hsa-miR-548o	CCAAACUGCAGUUACUUUUGC	567	MIMAT0005919
hsa-miR-126*	CAUUUUUACUUUUGGUACGCG	568	MIMAT0000444
hsa-miR-938	UGCCCUUAAAGGUGAACCCAGU	569	MIMAT0004981
hsa-miR-380	UAUGUAAUUGGUCCACAUCUU	570	MIMAT0000735
hsa-miR-1908	CGGCGGGGACGGCGAUUGGUC	571	MIMAT0007881
hsa-miR-345	GCUGACUCCUAGUCCAGGGCUC	572	MIMAT0000772
hsa-miR-548h	AAAAGUAAUCGCGGUUUUUGUC	573	MIMAT0005928
hsa-miR-193a-3p	AACUGGCCUACAAAGUCCAGU	574	MIMAT0000459
hsa-miR-7	UGGAAGACUAGUGAUUUUGUUGU	575	MIMAT0000252
hsa-miR-423-5p	UGAGGGGCAGAGAGCGAGACUUU	576	MIMAT0004748
hsa-miR-1259	AUAUAUGAUGACUUAGCUUUU	577	MIMAT0005910
hsa-miR-1911	UGAGUACCGCCAUGUCUGUUGG	578	MIMAT0007885
hsa-miR-605	UAAAUCCCAUGGUGCCUUCUCCU	579	MIMAT0003273
hsa-miR-513a-3p	UAAAUUUCACCUUUCUGAGAAGG	580	MIMAT0004777
hsa-miR-215	AUGACCUAUGAAUUGACAGAC	581	MIMAT0000272
hsa-miR-1911*	CACCAGGCAUUGUGGUCUCC	582	MIMAT0007886
hsa-miR-10a	UACCCUGUAGAUCGAAUUGUG	583	MIMAT0000253
hsa-miR-184	UGGACGGAGAACUGAUAAAGGU	584	MIMAT0000454
hsa-miR-576-5p	AUUCUAAUUUCUCCAGUCUUU	585	MIMAT0003241
hsa-miR-421	AUCAACAGACAUAUUUUGGGCGC	586	MIMAT0003339
hsa-miR-373	GAAGUGCUUCGAUUUUGGGGUGU	587	MIMAT0000726
hsa-miR-2053	GUGUUAUUAAACCUCUAUUUAC	588	MIMAT0009978
hsa-miR-22	AAGCUGCCAGUUGAAGAACUGU	589	MIMAT0000077
hsa-miR-30c	UGUAAACAUCUACACUCUCAGC	590	MIMAT0000244
hsa-miR-374b	AUAUAUAACAACCGUCUAAGUG	591	MIMAT0004955
hsa-miR-103-2*	AGCUUCUUUACAGUGCGCCUUG	592	MIMAT0009196
hsa-miR-10b	UACCCUGUAGAACCAGAAUUGUG	593	MIMAT0000254
hsa-miR-519a	AAAGUGCAUCCUUUUAGAGUGU	594	MIMAT0002869
hsa-miR-553	AAAACGGUGAGAUUUUGUUUU	595	MIMAT0003216
hsa-miR-609	AGGGUGUUUCUCUCAUCUCU	596	MIMAT0003277
hsa-miR-628-5p	AUGCUGACAUAUUUACUAGAGG	597	MIMAT0004809
hsa-miR-1538	CGGCCCGGGCUGCUGCUGUCCU	598	MIMAT0007400
hsa-miR-206	UGGAUGUAAGGAAGUGUGUGG	599	MIMAT0000462
hsa-miR-19a	UGUGCAAUUCUAUGCAAAACUGA	600	MIMAT0000073
	AAUCCUUGGAACCUAGGUGUGAGU	601	
hsa-miR-362-5p			MIMAT0000705
hsa-miR-196b*	UCGACAGCACGACACUGCCUUC	602	MIMAT0009201
hsa-miR-9*	AUAAAGCUAGAUAAACCGAAAGU	603	MIMAT0000442
hsa-miR-220b	CCACCACCGUGUCUGACACUU	604	MIMAT0004908
hsa-miR-365	UAAUGCCCCUAAAAUCCUUUU	605	MIMAT0000710
hsa-miR-1471	GCCCGCGUGUGGAGCCAGGUGU	606	MIMAT0007349
hsa-miR-1179	AAGCAUUCUUUCAUUGGUUGG	607	MIMAT0005824
hsa-miR-624*	UAGUACCAGUACCUUGUGUUA	608	MIMAT0003293

10

20

30

40

【表 1 - 13】

Coverage Human microRNA	Target sequence	SEQ ID No:	Target sequence accession
hsa-miR-128	UCACAGUGAACCGGUCUCUUU	609	MIMAT0000424
hsa-miR-579	UUCAUUUGGUUAAACCGCGAUU	610	MIMAT0003244
hsa-miR-518d-3p	CAAAGCGCUUCCCUUUGGAGC	611	MIMAT0002864
hsa-miR-224*	AAAAUGGUGCCCUAGUGACUACA	612	MIMAT0009198
hsa-miR-551b*	GAAAUCAAGCGUGGGUGAGACC	613	MIMAT0004794
hsa-miR-449b*	CAGCCACAACUACCCUGCCACU	614	MIMAT0009203
hsa-miR-33a	GUGCAUUGUAGUUGCAUUGCA	615	MIMAT0000091
hsa-miR-10a*	CAAAUUCGUAUCUAGGGGAAUA	616	MIMAT0004555
hsa-miR-890	UACUUGGAAAGGCAUCAGUUG	617	MIMAT0004912
hsa-miR-802	CAGUAAACAAAGAUUCAUCCUUGU	618	MIMAT0004185
hsa-miR-208b	AUAAGACGAACAAAAGGUUUGU	619	MIMAT0004960
hsa-miR-620	AUGGAGAUAGAUUAGAAAU	620	MIMAT0003289
hsa-miR-550	AGUGCCUGAGGGAGUAAGAGCCC	621	MIMAT0004800
hsa-miR-628-3p	UCUAGUAAGAGUGGCAGUCGA	622	MIMAT0003297
hsa-miR-98	UGAGGUAGUAAGUUGUAUUGUU	623	MIMAT0000096
hsa-miR-224	CAAGUCACUAGUGGUUCCGUU	624	MIMAT0000281
hsa-miR-30c-2*	CUGGGAGAAGGCUGUUUACUCU	625	MIMAT0004550
hsa-miR-448	UUGCAUAGUAGGAUGUCCCAU	626	MIMAT0001532
hsa-miR-1914*	GGAGGGGUCCCCGCACUGGGAGG	627	MIMAT0007890
hsa-miR-514	AUUGACACUUCUGUGAGUAGA	628	MIMAT0002883
hsa-miR-544	AUUCUGCAUUUUUAGCAAGUUC	629	MIMAT0003164
hsa-miR-625*	GACUAUAGAACUUUCCCCCUCA	630	MIMAT0004808
hsa-miR-501-5p	AAUCCUUUGUCCUGGGUGAGA	631	MIMAT0002872
hsa-miR-607	GUUCAAUCCAGAUUAUAAAC	632	MIMAT0003275
hsa-miR-200b	UAAUACUGCCUGGUAAUGAUGA	633	MIMAT0000318
hsa-miR-515-3p	GAGUGCCUUCUUUUGGAGCGUU	634	MIMAT0002827
hsa-miR-183	UAUGGCACUGGUAGAAUUCACU	635	MIMAT0000261
hsa-miR-297	AUGUAUGUGUGCAUGUGCAUG	636	MIMAT0004450
hsa-miR-365*	AGGGACUUUCAGGGGCGAGCUGU	637	MIMAT0009199
hsa-miR-137	UUUUUGCUUAAAGAAUACGCGUAG	638	MIMAT0000429
hsa-miR-588	UUGGCCACAAUGGGUUAGAAC	639	MIMAT0003255
	UGCCUGGGUCUCUGGCCUGCGCG	640	
hsa-miR-661	U		MIMAT0003324
hsa-miR-130a	CAGUGCAAUGUUAAAAGGGCAU	641	MIMAT0000425
hsa-miR-340	UUUAAAAGCAAUGAGACUGAUU	642	MIMAT0004692
hsa-miR-150	UCUCCCAACCCUUGUACCAGUG	643	MIMAT0000451
hsa-miR-1974	UGGUUGUAGUCCGUGCGAGAAUA	644	MIMAT0009449
hsa-miR-744	UGCGGGGCUAGGGCUAACAGCA	645	MIMAT0004945
hsa-miR-1979	CUCCACUGCUUCACUUGACUA	646	MIMAT0009454
hsa-miR-193a-5p	UGGGUCUUUGCGGGGCGAGAUGA	647	MIMAT0004614
hsa-miR-577	UAGAUAAAUAUUGGUACCUG	648	MIMAT0003242
hsa-miR-190b	UGAUUGUUUGAUUUUGGGUU	649	MIMAT0004929
hsa-miR-30b*	CUGGGAGGUGGAUGUUUACUUC	650	MIMAT0004589
hsa-miR-653	GUGUUGAAACAAUCUCUACUG	651	MIMAT0003328
hsa-miR-144*	GGAUUAUCAUAUACUGUAAG	652	MIMAT0004600
hsa-miR-518f*	CUCUAGAGGGAAGCACUUUCUC	653	MIMAT0002841
hsa-miR-1914	CCCUGUGCCCGGCCACUUCUG	654	MIMAT0007889
hsa-miR-1913	UCUGCCCCUCCGCGUCGCGCA	655	MIMAT0007888
hsa-miR-219-2-3p	AGAAUUGUGGCGUGGACAUCUGU	656	MIMAT0004675
hsa-miR-539	GGAGAAAUUAUCCUUGGUGUGU	657	MIMAT0003163
hsa-miR-26a-2*	CCUAUUCUUGAUUACUUGUUUC	658	MIMAT0004681
hsa-miR-888	UACUCAAAAAGCUGUCAGUCA	659	MIMAT0004916
hsa-miR-545	UCAGCAAACAUAUUUUGUGUGC	660	MIMAT0003165

10

20

30

40

【表 1 - 1 4】

Coverage Human microRNA	Target sequence	SEQ ID No:	Target sequence accession
hsa-miR-29b	UAGCACCAUUUGAAAUCAGUGUU	661	MIMAT0000100
hsa-miR-208a	AUAAGACGAGCAAAAAGCUUGU	662	MIMAT0000241
hsa-miR-708*	CAACUAGACUGUGAGCUUCUAG	663	MIMAT00004927
hsa-miR-1539	UCCUGCGCGUCCCAGAUGCCC	664	MIMAT0007401
hsa-miR-181c	AACAUUCAACCUGUCGGUGAGU	665	MIMAT0000258
hsa-miR-520d-5p	CUACAAAGGGAAGCCCUUUC	666	MIMAT0002855
	AGCCUGGAAGCUGGAGCCUGCAGU	667	
hsa-miR-1254			MIMAT0005905
hsa-miR-2113	AUUUGUGCUUGGCUCUGUCAC	668	MIMAT0009206
hsa-miR-301a	CAGUGCAAUAGUAUUGUCAAGC	669	MIMAT0000688
hsa-miR-146a	UGAGAACUGAAUCCAUGGGUU	670	MIMAT0000449
hsa-miR-548d-5p	AAAAGUAAUUGUGUUUUUGCC	671	MIMAT0004812
hsa-miR-381	UAUACAAGGGCAAGCUCUCUGU	672	MIMAT0000736
hsa-miR-218-1*	AUGGUUCCGUCAAGCACCAUGG	673	MIMAT0004565
hsa-miR-1912	UACCCAGAGCAUGCAGUGUGAA	674	MIMAT0007887
hsa-miR-1207-5p	UGGCAGGGAGGCUGGGAGGGG	675	MIMAT0005871
hsa-miR-570	CGAAAACAGCAAUUACCUUUGC	676	MIMAT0003235
hsa-miR-491-5p	AGUGGGGAACCCUCCAUGAGG	677	MIMAT0002807
hsa-miR-572	GUCCGCUCCGGGUGGCCCA	678	MIMAT0003237
hsa-miR-548c-3p	CAAAAUCUCAUUUACUUUUGC	679	MIMAT0003285
hsa-miR-29a	UAGCACCAUCUGAAAUCGGUUA	680	MIMAT0000086
hsa-miR-302a*	ACUUAACGUGGAUGUACUUGCU	681	MIMAT0000683
hsa-miR-1909	CGCAGGGGCCGGGUGUCACCG	682	MIMAT0007883
hsa-miR-1252	AGAAGGAAUUGAAUUAUUA	683	MIMAT0005944
hsa-miR-299-3p	UAUGUGGGAUGGUAACCGCUU	684	MIMAT0000687
hsa-miR-373*	ACUCAAAUUGGGGCGCUUCC	685	MIMAT0000725
hsa-miR-362-3p	AACACACCUAUUCAAGGAUUA	686	MIMAT0004683
hsa-miR-521	AACGCACUCCCCUUUAGAGUGU	687	MIMAT0002854
hsa-miR-200a	UAACACUGUCUGGUAACGAUGU	688	MIMAT0000682
hsa-miR-1972	UCAGGCCAGGCACAGUGGCUCA	689	MIMAT0009447
hsa-miR-665	ACCAGGAGGCUGAGGCCCU	690	MIMAT0004952
hsa-miR-548m	CAAAGGUAAUUGUGGUUUUG	691	MIMAT0005917
hsa-miR-626	AGCUGUCUGAAAUGUCUU	692	MIMAT0003295
hsa-miR-384	AUUCUAGAAUUGUUAUA	693	MIMAT0001075
hsa-miR-30e	UGUAAACAUCUUGACUGGAAG	694	MIMAT0000692
hsa-miR-93	CAAAGUGCUGUUCGUGCAGGUAG	695	MIMAT0000093
hsa-miR-383	AGAUCAGAAGGUGAUUGUGGCU	696	MIMAT0000738
hsa-miR-1537	AAAACCGUCUAGUUACAGUUGU	697	MIMAT0007399
hsa-miR-548l	AAAAGUAAUUGCGGGUUUUGUC	698	MIMAT0005889
hsa-miR-338-3p	UCCAGCAUCAGUGAUUUUGUUG	699	MIMAT0000763
hsa-miR-642	GUCCUCUCCAAUGUGUCUUG	700	MIMAT0003312
hsa-miR-30c-1*	CUGGGAGAGGUUGUUUACUCC	701	MIMAT0004674
hsa-miR-142-5p	CAUAAAGUAGAAAGCACUACU	702	MIMAT0000433
hsa-miR-7-1*	CAACAAUACAGUCUGCCAU	703	MIMAT0004553
hsa-miR-26a	UUCAAGUAAUCCAGGAUAGGCU	704	MIMAT0000082
hsa-miR-664	UAUUCAUUUUACCCCAGCCUACA	705	MIMAT0005949
hsa-miR-363	AAUUGCACGGUAUCCAUCUGUA	706	MIMAT0000707
hsa-miR-660	UACCAUUGCAUAUCGGAGUUG	707	MIMAT0003338
hsa-miR-561	CAAAGUUUAAGAUCUUGAAGU	708	MIMAT0003225
hsa-miR-29c	UAGCACCAUUUGAAUCCGGUUA	709	MIMAT0000681
hsa-miR-202*	UCCUAUGCAUAUACUUCUUG	710	MIMAT0002810
hsa-miR-432*	CUGGAUGGCUCUCCAUGUCU	711	MIMAT0002815
hsa-miR-675*	CUGUAUGCCCUCACCGCUCA	712	MIMAT0006790

10

20

30

40

【表 1 - 15】

Coverage Human microRNA	Target sequence	SEQ ID No:	Target sequence accession
hsa-miR-377	AUCACACAAAGGCAACUUUUGU	713	MIMAT0000730
hsa-miR-451	AAACCGUUACCAUUCUGAGUU	714	MIMAT0001631
hsa-miR-148b*	AAGUUCUGUUUAUACACUCAGGC	715	MIMAT0004699
hsa-miR-424	CAGCAGCAAUUCAGUUUUUGAA	716	MIMAT0001341
hsa-miR-431	UGUCUUGCAGGCCGUCAUGCA	717	MIMAT0001625
hsa-miR-1247	ACCCGUCCCGUUCGUCCCGGA	718	MIMAT0005899
hsa-miR-651	UUUAGGAUAAGCUUGACUUUUG	719	MIMAT0003321
hsa-miR-103-as	UCAUAGCCCUGUACAAUGCUGCU	720	MIMAT0007402

10

代わりに、または、加えて、試薬は、表 2 から選択された、少なくとも 1、少なくとも 2、少なくとも 3、少なくとも 4、少なくとも 5、少なくとも 6、少なくとも 7、少なくとも 8、少なくとも 9、少なくとも 10 のタンパク質バイオマーカーの定量のためにできる。

表 2 (表 2 - 1 ~ 表 2 - 5)

【表 2 - 1】

Protein	Gene
1 a2-Macroglobulin	A2M
2 a-Actinin-1	ACTN1
3 ABC Transporter	ABCG1
4 Adiponectin	PPARG, NR1C3
5 Adrenomedullin	ADM
6 CD166 Antigen	ALCAM
7 ANG-2, angiopoietin-2	TEK, TIE2
8 Annexin-2	ANXA2, ANX2
9 natriuretic peptide precursor A	ANP
10 apolipoprotein A1	APOA1
11 apolipoprotein A2	APOA2
12 apolipoprotein B	APOB
13 apolipoprotein C1	APOC1
14 apolipoprotein C3	APOC3
15 apolipoprotein E	APOE
16 apolipoprotein H (beta-2-glycoprotein I)	APOH
17 Clusterin, ApoJ	CLU
18 Antithrombin III	SERPINC1, AT3
19 B cell attracting chemokine 1	CXCL13, BCA-1
20 Nerve Growth Factor, beta polypeptide	NGFB
21 Complement protein C1Q	C1QA
22 Caspase 4	CASP1
23 CCL1	CCL1
24 CCL14	CCL14
25 CCL15	CCL15
26 CCL18	CCL18
27 CCL21	CCL21
28 CCL28	CCL28
29 CCL9	CCL9

20

30

【表 2 - 2】

Protein	Gene
30 CD40 Ligand	CD40LG
31 CD44	CD44
32 CD52	CD52
33 CD53	CD53
34 cytokine receptor-like factor 1	CRLF1
35 CRP	CRP
colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage)	CSF2RA
36 CTACK	CCL27
38 CXCL11	CXCL11
39 CXCL14	CXCL14
40 CXCL16	CXCL16
41 Cystatin C	CST3
42 D-dimer, fibrin degradation product	FGG, FGA, FGB
43 Epidermal growth factor	EGF
44 Endothelin-1	EDN1
45 En-RAGE, S100 calcium binding protein A12	S100A12
46 Eotaxin	CCL11
47 E-Selectin, endothelial adhesion molecule 1	SELE
48 fatty acid binding protein 3	FABP3
49 Factor II, thrombin	F2
50 Factor V	F5
51 Factor VII	F7
52 Factor VIII	F8
53 Fas, TNF receptor superfamily, member 6	FAS
54 Fas-Ligand, TNF superfamily, member 6	FASLG
55 Fc fragment of IgE	FCER1G
56 Fetuin A, alpha-2-HS-glycoprotein	AHSG
57 FGF-basic, fibroblast growth factor 2 (basic)	FGF2
58 Fibrinogen	FGG, FGA, FGB
59 fibronectin 1	FN1
60 Fractalkine	CX3CL1
61 frizzled-related protein	FRZB
62 Galectin-3	LGALS3
63 colony stimulating factor 3 (granulocyte)	CSF3
64 growth differentiation factor 15	GDF-15
65 Granulin	GRN
66 GROa	CXCL1
67 Haptoglobin	HP
68 fatty acid binding protein 3	FABP3
69 hepatocyte growth factor	HGF
70 Hsp-27, heat shock 27kDa protein 1	HSPB1
71 integrin-binding sialoprotein	IBSP
72 ICAM-1, intercellular adhesion molecule 1 (CD54)	ICAM1
73 interferon, alpha 2	IFNA2
74 interferon, gamma	IFNG
75 interferon gamma receptor 1	IFNGR1
76 IGF-1, insulin-like growth factor 1 (somatomedin C)	IGF1

10

20

30

【表 2 - 3】

Protein	Gene	
77 insulin-like growth factor binding protein 1	IGFBP1	
78 insulin-like growth factor binding protein 3	IGFBP3	
79 insulin-like growth factor binding protein 4	IGFBP4	
80 insulin-like growth factor binding protein 6	IGFBP6	
81 interleukin 10	IL10	
82 Interleukin 12b, IL-12(p40)	IL12B	
83 interleukin 16	IL16	
84 interleukin 18	IL18	
85 interleukin 1 alpha	IL1A	
86 Interleukin 1 beta	IL1B	
87 Interleukin 1 receptor-like 4	IL1RL1	
88 Interleukin 2 receptor alpha	IL2RA	10
89 interleukin 3	IL3	
90 interleukin 5	IL5	
91 interleukin 6	IL6	
92 interleukin 7	IL7	
93 interleukin 8	IL8	
94 IP-10	CXCL10	
95 I-TAC	CXCL11	
96 lymphocyte cytosolic protein 1	LCP1	
97 low density lipoprotein receptor	LDLR	
98 Leptin	LEP	
99 lectin, galactoside-binding, soluble, 3 binding protein	LGALS3BP	
100 leukemia inhibitory factor	LIF	
101 oxidised low density lipoprotein (lectin-like) receptor 1	OLR1	
102 lipoprotein, Lp(a)	LPA	20
103 LpPLA2, lipoprotein-associated phospholipase A2	PLA2G7	
104 L-Selectin, lymphocyte adhesion molecule 1	SELL	
105 Lysozyme	LYZ	
106 MCP-1	CCL2	
107 MCP-2	CCL8	
108 MCP-3	CCL7	
109 MCP-4	CCL13	
110 MCP-5	CCL12	
111 M-CSF, colony stimulating factor 1 (macrophage)	CSF1	
112 MDC, CCL22	CCL22	
113 matrix Gla protein	MGP	
114 macrophage migration inhibitory factor	MIF	
115 MIG	CXCL9	
116 MIP-1a, Macrophage inflammatory protein 1-alpha	CCL3	30
117 MIP-1 alpha P	CCL3L1	
118 MIP-1b	CXCL4	
119 MIP-2a, GROb	CXCL2	
120 MIP-2b, GROg	CXCL3	
121 MIP-3B, Macrophage inflammatory protein 3 beta	CCL19	
122 MMP-10, matrix metalloproteinase 10	MMP10	
123 MMP-2, matrix metalloproteinase 2	MMP2	
124 MMP-9, matrix metalloproteinase 9	MMP9	

【表 2 - 4】

Protein	Gene
125 MPO, myeloperoxidase	MPO
126 myelin protein zero-like 1	MPZL1
127 major histocompatibility complex, class I-related	MR1
128 NT-pro-BNP	NPPB
129 oncostatin M	OSM
130 Osteopontin	SPP1
Osteoprotegerin, Tumor necrosis factor receptor superfamily member 11B	TNFRSF11B
132 Ox-LDL receptor	OLR1
133 PAI-1, plasminogen activator inhibitor type 1	SERPINE1
134 PAI-1 (total)	SERPINE1
135 pregnancy-associated plasma protein A	PAPPA
136 proprotein convertase subtilisin/kexin type 9	PCSK9
137 platelet-derived growth factor beta	PDGFB
138 platelet derived growth factor C	PDGFC
139 platelet/endothelial cell adhesion molecule, CD31 antigen	PECAM1
140 phospholipase A2, group VII	PLA2G7
141 P-Selectin	SELP
142 prostaglandin D2 synthase	PTGDS
143 renal tumor antigen	RAGE
144 RANTES	CCL5
145 Renin, Angiotensinogenase	REN
146 Resistin	RETN
147 Rho GDP dissociation inhibitor (GDI) beta	ARHGDIB
148 regulator of G-protein signalling 1	RGS1
149 regulator of G-protein signalling 10	RGS10
150 S100 calcium binding protein A8	S100A8
151 S100 calcium binding protein A9	S100A9
152 serum amyloid A1	SAA
153 SAP, SH2 domain protein 1A	SH2D1A
154 SCF, KIT ligand	KITLG
155 SCGFb	CLEC11A
156 SDF-1	CXCL12
157 SDF-1a	CXCL12
158 group IID secretory phospholipase A2 (sPLA2)	PLA2G2D
159 frizzled-related protein	FRZB
160 solute carrier family 11	SLC11A1
161 suppressor of cytokine signaling 3	SOCS3
162 Thrombomodulin	THBD
163 Thrombospondin R, CD36 molecule (thrombospondin receptor)	CD36
164 Thrombospondin-1	THBS1
165 TIMP-1, metalloproteinase inhibitor 1	TIMP1
166 TIMP-2, metalloproteinase inhibitor 2	TIMP2
167 TIMP-3, metalloproteinase inhibitor 3	TIMP3
168 TIMP-4, metalloproteinase inhibitor 3	TIMP4
169 tenascin C	TNC
170 TNFa, tumor necrosis factor (TNF superfamily, member 2)	TNFA
171 tumor necrosis factor, alpha-induced protein 2	TNFAIP2

10

20

30

【表 2 - 5】

Protein	Gene
172 tumor necrosis factor, alpha-induced protein 6	TNFAIP6
173 TNFb, lymphotoxin alpha (TNF superfamily, member 1)	LTA
174 tumor necrosis factor receptor superfamily, member 1A, TNF-R1	TNFRSF1A
175 tumor necrosis factor receptor superfamily, member 1B, TNF-R2	TNFRSF1B
tumor necrosis factor (ligand) superfamily, member 11, TRANCE, RANKL	TNFSF11
176 TRAIL, tumor necrosis factor (ligand) superfamily, member 10	TNFSF10
177 plasminogen activator, urokinase	PLAU
178 Vasopressin-neurophysin 2-copeptin	AVP
180 vascular cell adhesion molecule 1	VCAM1
181 vascular endothelial growth factor	VEGF
182 von Willebrand factor	VWF
183 WARS, tryptophanyl-tRNA synthetase	WARS
184 WNT1 inducible signaling pathway protein 1	WISP1
185 wingless-type MMTV integration site family, member 4	WNT4

40

50

【 0 1 5 8 】

特定の実施形態では、タンパク質バイオマーカーは、IL - 16、sFAS、Fasリガンド、MCP - 3、HGF、CTACK、エオタキシン、アディポネクチン、IL - 18、TIMP . 4、TIMP . 1、CRP、VEGF、およびEGF、から選択されている。

【 0 1 5 9 】

キットはさらに、1つ以上の表現型の統計分析用のソフトウェアパッケージを含むことができ、分類の確率を計算するためのリファレンス・データベースが含まれる場合がある。キットは、血液サンプル、セカンドステージ抗体、ELISA試薬、チューブ、スピンカラム等の回収や処理のためのデバイスなど、様々な方法で用いられる試薬を含めることができる。

10

【 0 1 6 0 】

上記構成要素に加えて、対象キットはさらに、本発明の方法を実施するための使用説明書を含む。これら使用説明書は、様々な形態で対象キット中に、1つ以上存在し得る。これら使用説明が存在する1つの態様は、例えば、キットのパッケージや添付文書内の、情報が印刷された紙片など、適切な媒体や基質に関する印刷情報の様なものである。さらに別の手段は、例えば、フロッピーディスク、CDなど、情報が記録されたコンピュータ可読媒体であろう。また、存在し得る別の手段は、隔離されたサイト情報にアクセスするために、インターネットを介して使用することができるウェブサイトアドレスである。任意の便利な手段がキット内に存在し得る。

20

【 0 1 6 1 】

さらなる態様において、本明細書に開示される方法アッセイおよびキットは、プールされたサンプル中のバイオマーカーを検出するために使用することができる。この方法は、特に、複数のサンプルでわずかな量しか利用できない場合（例えば、保管されていた臨床サンプルセット）、及び/又は患者または対象集団に関連する有用なデータセットを生成する時に、有用である。この点で、（たとえば、約10 μ l、約15 μ l、約20 μ l、約30 μ l、約40 μ l、約50 μ l、またはそれ以上の）等量のサンプルが、約2、5、10、15、20、30、50、100またはそれ以上の）複数の個人から得ることができる。個人は、様々なしるし（indicia）によって一致させることができる。このしるしには、年齢、性別、疾患の既往歴、イベントまでの時間（time to event）などを含めることができる。各個人から得られたサンプルの等量をプールし、1つまたは複数のバイオマーカーの存在について分析することができる。結果は、本明細書に記載の予測・分類モデルを使用して、リファレンス・セットを生成し、予測を実行し、与えられた条件等に関連付けられるバイオマーカーを決定することができる。当業者であれば、本方法の多くの用途を容易に認識し、本明細書に開示されるmiRNA、タンパク質、疾患状態に限定されるものではないことを認識する。実際に、本方法は、様々な疾患や状況に関連付けられている、DNA、RNA（mRNA、miRNA、ヘアピン前駆体RNA、RNP）、タンパク質等を検出するために使用することができる。

30

【 0 1 6 2 】

定義

40

本明細書中で使用される用語は、特記のない限り、以下に記載するように定義されている。

【 0 1 6 3 】

本明細書で使用される用語「モニタリング」は、個人または個々人の健康や疾患の状況に関する有用な情報を提供するために、データセットから生成された結果を使用することを指す。「モニタリング」は、例えば、予後の判定、リスク層別化、薬物療法の選択、進行中の薬物療法の評価、治療の有効性の決定、結果の予測、治療への応答の決定、疾患または合併症の診断を含むことができ、疾患の進行の後に、または時間をかけて患者の健康状態に関連する情報を提供し、作用の既知の分子メカニズムを伴う実験的治療の恩恵を受ける可能性が最も高い患者を選択し、そのメカニズムが、投薬にラベルがない様なある疾

50

患の小さなサブセットで重要であるかもしれない既知の分子メカニズムを伴う承認薬の恩恵を受ける可能性が最も高い患者を選択し、例えば、非侵襲的な血液検査から、より侵襲的な生検などのオプションなど、より侵襲的な／高価なテストを決定を支援する患者集団をスクリーニングし、または別の指標を治療するために使用される薬の副作用を評価するテストを含めることができる。特に、用語「モニタリング」は、アテローム性動脈硬化症のステージング、アテローム性動脈硬化症の予後、血管炎症のレベルを参照することができ、アテローム性動脈硬化症の進行の程度を評価し、治療応答をモニタリングし、冠動脈カルシウムスコアを予測し、またはアテローム性動脈硬化症の不安定な症状から安定した症状を区別することができる。

【0164】

本明細書中で使用される用語「定量的なデータ」は、数値を特定することができる任意のデータセットコンポーネント（例えば、miRNAマーカー、タンパク質マーカー、臨床的しるし（臨床的標識）、代謝的計測、または遺伝的アッセイ）に関連付けられたデータを指す。定量的データは、マーカーのDNA、RNAまたはタンパク質レベルの尺度であることができ、モル濃度、重量濃度などのような測定の単位で表現することができる。例えば、マーカーがタンパク質である場合、そのマーカーの定量的データは、当業者に公知の方法を使用して測定され、mMまたはmg/dLの濃度単位で表現される、タンパク質の発現レベルである。

【0165】

本明細書で使用される用語「哺乳動物」は、ヒトおよび非ヒトの両方を含み、ヒト、非ヒト霊長類、イヌ、ネコ、マウス、ウシ、ウマ、およびブタも含むが、これに限定されるものではない。

【0166】

本明細書で使用される用語「擬似冠動脈カルシウムスコア」は、画像診断法による測定を通じてというよりむしろ、本明細書に開示されるような方法を使用して生成された冠動脈カルシウムスコアを指す。当業者であれば、擬似冠動脈カルシウムスコアは、画像診断法による測定により生成された冠動脈カルシウムスコアと交換可能に使用され得ることを認識するであろう。

【0167】

二つ以上の核酸またはポリペプチド配列の構成における用語であるパーセント「同一性」(percent identity)は、以下で説明する配列比較アルゴリズム（例えば、BLASTP及びBLASTNまたは当業者に利用可能な他のアルゴリズム）のいずれかを使用し、または目視検査によって測定されて、比較され、最大一致のために整列された場合に、二つ以上のシーケンスまたはサブシーケンスが、特定の比率で同一のヌクレオチドまたはアミノ酸残基を有することを、指す。アプリケーションに応じて、パーセント「同一性」は、比較される配列領域、例えば、機能ドメインに亘って存在することができ、あるいはまた、比較される2つの配列の全長にわたって存在することができる。

【0168】

特定の実施形態では、治療法の「効果」が決定される。治療法は、改善、改良、リスクの低減、または病状または疾患の進行の減速に基づいて有効であると認識される。このような決定は、当業者によって容易になされる。

【実施例】

【0169】

(実施例1)

プールされたサンプル中のmiRNA解析

本研究で利用されたプーリングアプローチ(pooling approach)は、a)血清中のmiRNAを識別するためにExiqon社のLocked Nucleic Acid (LNA (商標))技術の能力を調査すること、及び、b)テストのために貴重な保管臨床サンプルから最小量を利用することという、2つの目標を達成した。

【0170】

10

20

30

40

50

血清中のmiRNAを識別する、LNA（商標）技術の能力を評価するために、52のプールが、前向き研究（prospective study）から保管されていた血清サンプルを使用して作成された（Marshfield Clinical Personalized Medicine Research Project（PMRP）、Personalized Medicine, 2(1): 49-79 (2005)）。そのうち26プールは、患者ケースを、他の26プールは、コントロールを構成した。各プールは、年齢（40～80歳の間の5歳毎の8つの範囲の中から選択）、性別や事象までの時間（time to event）（即ち、0-6ヶ月内でMI（心筋梗塞）、6-12ヶ月内でMIなど）をマッチさせた、5人の各個人からの等量（50 μ L）の血清サンプルを含有した。以後のマッチングは、近似的であった。患者ケースは、採血から5年以内にMIや、入院の不安定狭心症を伴った被験者であった。コントロールは、採血から5年以内に上記のいずれの疾患イベントとならなかった被験者であった。これらサンプルは、分類問題として評価され、テストパフォーマンスは、曲線下面積（AUC）を用いて判定した。

10

【0171】

AUCの観点からテストパフォーマンスは、（個々のマーカーに対する）測定値の分布、または実験デザインの時点で不明であったスコアの分布に、依存する。実際の実験デザイン（26症例、26コントロール）と同様のサンプルサイズのセットに対するテストの期待される性能を見積もるために、多くのシミュレーションが、変数やプールしているサンプル数に対する、異なる仮定分布を用いて行なわれた。使用された仮定分布は、a) 正20
常、b) 無仮説（chisq）、および、c) 対数正規（log-normal）であった。あるプールサンプルの分布と数に対する適切な数の「コントロール」は、ランダムに選択されたものであり、対応する数の症例は、集団間の差異を表すために、既知の平均シフト分布から選択された。したがって、サイズMのプールのために、26 * Mコントロールと26 * M症例が選択され、プールサンプルの各々は、Mサンプルの値を平均することによって作成される。このプロセスは、500回繰り返され、予想されるAUCの分布は、プールされたサンプルの与えられた数と母集団の距離に対して推定された。

20

【0172】

図1は、個々のサンプル（白丸と実線の誤差範囲）と、プールされたサンプル（プール毎に5の固体サンプル）（白丸と破線の誤差範囲）を使用した際の、バイオマーカーの濃度やスコアの仮定対数正規分布の結果を示している。実線の黒のドットは、個々の測定のための理論的な答えを示している。予測AUCが、一貫して、個々のサンプルに対する真のAUCと予測AUCを過小評価していると観察されるが、（個々のサンプルの）不確かさの範囲は、プールされたサンプルのそれより小さい。図2は、測定値の想定正規分布の結果を表す。このケースでは、プールされたサンプルの結果は、理論的結果や個々のサンプル結果によく一致している。再び、プールされたサンプルの不確かさは、ヒトサンプルの対応する不確かさよりも小さい。仮定の無仮説分布は、対数正規分布から得られたものと一致する、多くのシミュレーション結果を提供した。これらのシミュレーションは、ヒトサンプルの分布が正規分布に従う場合は、プールされたサンプルの結果が、予測AUCの非常に良好な推定値を提供することを示しており、そうでなければ計算されたAUCは40
過小評価される。

30

40

【0173】

52のプールされたサンプルの38のmiRNAが、EXIQON社のUniRT（商標）LNA技術を用いて分析した。全RNAは、多少変更された、QIAGEN RNEASY（登録商標）ミニキットプロトコル（QIAGEN社、Valencia, CA）を使用して、提供血清サンプル（上記）から抽出した。

【0174】

全RNAはQIAGEN RNEASY（登録商標）ミニキットを用いて血清から抽出した。血清は、氷上で融解し、4-5分間1,000 \times gで遠心した。サンプルあたり200 μ Lの分量の血清を、新しいマイクロ遠心チューブへ移し、それに0.94 μ g / μ

50

LのMS2バクテリオファージを含むQiagol混合液750 μ Lを添加した。チューブを混合し、200 μ Lのクロロホルムを加えて、5分間インキュベートした。チューブを混合して2分間インキュベートし、415分間12,000 \times gで遠心した。上澄部分を新しいマイクロ遠心チューブに回収し、1.5倍量の100%エタノールを追加した。チューブを完全に混合し、750 μ Lのサンプルを、コレクションチューブ内のQIAGEN RNEASY（登録商標）ミニスピンカラムへ移し、室温で30秒、15,000 \times gで遠心分離した。

表3（表3-1～表3-2）

【表3-1】

	miRNA
1	hsa-let-7a
2	hsa-let-7b
3	hsa-let-7d
4	hsa-mir-1
5	hsa-mir-106b
6	hsa-mir-10b
7	hsa-mir-125b
8	hsa-mir-126
9	hsa-mir-146b-5p
10	hsa-mir-148a
11	hsa-mir-155
12	hsa-mir-15a
13	hsa-mir-16
14	hsa-mir-17
15	hsa-mir-182
16	hsa-mir-18a
17	hsa-mir-192
18	hsa-mir-200c
19	hsa-mir-205
20	hsa-mir-20a

10

20

【表3-2】

	miRNA
21	hsa-mir-20b
22	hsa-mir-21
23	hsa-mir-212
24	hsa-mir-218
25	hsa-mir-221
26	hsa-mir-222
27	hsa-mir-23a
28	hsa-mir-23b
29	hsa-mir-24
30	hsa-mir-26a
31	hsa-mir-27a
32	hsa-mir-32
33	hsa-mir-342-5p
34	hsa-mir-429
35	hsa-mir-451
36	hsa-mir-9
37	hsa-mir-103
38	hsa-mir-93

30

40

【0175】

各RNAサンプルは、3つの独立した逆転写（RT）反応でcDNAに逆転写（RT）され、singlicateリアルタイムPCRまたは定量PCR（qPCR）反応として実行される。

【0176】

50

各384ウェルプレートは、2つのmiRNAのアッセイのためのすべてのサンプルの反応液を含んでいた。RTステップでの無テンプレートコントロール(RNAを水で置換)、及び、RTステップでの無酵素コントロール(テンプレートとしてプールされたRNAを使用)を、ネガティブコントロールとして、実験に含めた。すべてのアッセイで、無テンプレートコントロール及び無酵素コントロールの結果が陰性であったという点で、この品質管理手順をパスした。

【0177】

リアルタイムPCR解析の追加のステップは、各反応の融解曲線を生成することにより、アッセイの特異性を評価するために行われた。融解曲線解析時に単一のピークの出現は、単一の特定の産物が定量PCRの過程で増幅されたことを示している。複数の融解曲線ピークの出現は、関連する複数の定量PCR増幅産物の指標を提供し、これは特異性の欠如の証拠である。複数のピークを示したアッセイは、いずれもデータセットから除外されている。Cp(交点、すなわち、測定されたシグナルが予め設定されたしきい値を超えて交差する点で、標的配列の測定可能な濃度を示す点)の決定のため(2次微分法による)、及び、溶解曲線分析のために、増幅曲線は、LIGHTCYCLER(登録商標)ソフトウェア(Roche, Indianapolis, IN)を用いて分析された。

10

【0178】

PCR効率は、また、LIN REG(登録商標)ソフトウェア(オープンソースソフトウェア)でPCR増幅曲線を分析することにより評価した。5つのハウスキーピングmiRNA(miR-16、miR-93、miR-103、miR-192、およびmiR-451)の成績が、付属の血清サンプルから抽出したRNAの品質を評価するために使用された。

20

【0179】

38のmiRNAターゲット中の24は、サンプルで検出された。50のサンプル(26症例と24コントロール)が、これらのサンプルの分類分析の成績を評価し、状態を予測するmiRNAを選択するために使用された。次の方法論、a)ロジスティック回帰アプローチ、およびb)(L1ペナルティ・ラッソ(L1 penalty-lasso))を使用したペナルティロジスティック回帰アプローチは、モデルを構築するために用いた。モデルでベストの分類を提供した用語の選択は、a)非ペナルティロジスティック回帰アプローチ(unpenalized logistic regression approach)のために、ベイズの情報基準(Bayesian Information criterion)を使用して前方選択を行うことにより、及び、b)のペナルティアプローチ(penalized approach)のために、最適なペナルティのクロスバリデーションに基づく選択により、完成した。後者では、ペナルティパラメータは、使用可能なパラメータの係数をゼロに追いやるため、結果として得られるモデルは、減少した数の予測miRNAだけを含む。パフォーマンスの客観的な尺度を評価するために、AUCは事前検証されたスコアを用いて算出した。事前検証は、クロスバリデーションのアプローチと非常によく似ており、与えられた結果を伴う「スコア」の関連付けは、与えられた検体のために、トレーニングセット内の特定の検体を使用せずにフィットしたモデルから予測されている値に基づいてされている。この分析では事前検証されたスコアは、以下の二つのアプローチに基づいて計算された。a)K分割交差検定(k-fold cross-validation)と、b)リーブ・ワン・アウト・交差検定(leave-one-out cross validation)である。事前検証の繰り返しは、N回(Nは、通常、100から1000まで)なされた。分析の完全なシーケンスは以下の通りである。

30

40

1)モデル選択のためのBICでの、ロジスティック回帰を使用して、またはトレーニングセットでの入れ子式の交差検定(nested cross-validation)を介して、ペナルティ関数を推定するペナルティロジスティック回帰を使用して、データのサブセットにモデルをフィットさせる。

2)K分割交差検定ために、モデルを、サンプルのK-1グループにフィットさせる

50

。

3) リーブ・ワン・アウト・交差検定 (leave-one-out cross validation) のために、M - 1 サンプル (ここでは、M = 50) にフィットさせる。

4) フィットモデルを使用して、抜き出しサンプル (left-out samples) のスコアを予測する (交差検定のためのグループ K と、リーブ・ワン・アウト・交差検定 (leave-one-out cross validation) のための、1 つの抜き出しサンプル (single left-out sample))。

5) すべてのスコアがすべてのサンプルに対して予測された後、分類問題のために AUC を計算する。

10

6) AUC の変動を評価するために、ステップを N 回 (1 から 3) 繰り返す。

【0180】

図 3 は、事前検定スコア計算 100 リピートで、ペナルティロジスティック回帰モデル (L1 ペナルティ・ラッソ (L1 penalty-lasso)) を使用して得られた AUC 値の分布を示す。表 4 は、モデル選択と、ペナルティロジスティック回帰 (L1 ペナルティ・ラッソ (L1 penalty-lasso)) とペナルティスコア計算のための 10 分割交差検定とを使用したフィッティングの過程で選択されたトップの miRNA を示す。あるマーカーが、この実行で選択され得る最大回数は、1000 である (1000 リピートのスコア事前検定と、各リピート間での 10 分割交差検定)。

表 4 (表 4 - 1 ~ 表 4 - 2)

20

【表 4 - 1】

miR	Counts
miR.16	999

【表 4 - 2】

miR.26a	998
miR.130a	981
miR.150	917
miR.222	856
miR.106b	836
miR.93	801
miR.10b	771
miR.30c	722
miR.192	717
let.7b	579
miR.20a	436
miR.107	313
miR.20b	239
hsa.let.7f	225
miR.186	208
miR.92a	157

30

【0181】

表 5 は、L1 ペナルティロジスティック回帰アプローチと組合せてリーブ・ワン・アウト・交差検定 (leave-one-out (LOOV) cross-validation) (LOOCV) を使用して選択した、マーカーの数を示す。この 2 つの方法は、ほぼ同じ順序で選択された、非常に重複したマーカーのセットを、提供している。カウントの違いは、セット内のサンプル数からくるものである。対応する AUC は 0.66 である。

40

。

表 5 (表 5 - 1 ~ 表 5 - 2)

【表 5 - 1】

miR	Counts
miR.26a	51
miR.16	51
miR.130a	51
miR.150	51
miR.106b	50
miR.93	50
miR.222	48
miR.192	47
miR.30c	47
miR.10b	40
let.7b	32
miR.20a	26
miR.20b	16
miR.107	16
hsa.let.7f	15
miR.186	14
miR.92a	12

10

【表 5 - 2】

miR.19a	3
---------	---

20

【 0 1 8 2 】

(実施例 2)

個々のサンプル中の m i R N A の評価

フォローアップの実験は、実施例 1 に記載の E X I Q O N L N A (登録商標)を使用して、個々の血清サンプル中の m i R N A の検出と成績 (2 6 症例、2 6 コントロール) の評価に集中した。合計 9 0 の m i R N A (表 6 を参照) が、スクリーニングされ、それには、プールされたサンプルでスクリーニングされた m i R N A を含んでいた。この 9 0 m i R N A のターゲットのうち 4 4 は、個々の血清サンプルで検出された。プールされたサンプルで検出された 2 4 の m i R N A は、個々のサンプルでも検出され、さらに追加的に、2 0 の m i R N A が、個々のサンプルで検出された。5 の m i R N A は、データ

30

表 6 (表 6 - 1 ~ 表 6 - 4)

【表 6 - 1】

	miRNA	Samples 1-52	Samples 53-104
1	hsa-let-7a	Yes*	Yes**
2	hsa-let-7b	Yes*	Yes**
3	hsa-let-7d	Yes*	Yes**
4	hsa-mir-1	No*	No**
5	hsa-mir-106b	Yes*	Yes**
6	hsa-mir-10b	Yes*	Yes**
7	hsa-mir-125b	No*	No**
8	hsa-mir-126	Yes*	Yes**
9	hsa-mir-146b-5p	No*	No**

40

50

【表 6 - 2】

	miRNA	Samples 1-52	Samples 53-104
46	hsa-mir-130a	Yes**	Yes**
47	hsa-mir-130b	No**	No**
48	hsa-mir-132	No**	No**
49	hsa-mir-135a	No**	No**
50	hsa-mir-136	No**	No**
51	hsa-mir-146a	Yes**	Yes**
52	hsa-mir-146b-3p	No**	No**
53	hsa-mir-150	Yes**	Yes**
54	hsa-mir-181a	No**	No**

10

【表 6 - 3】

	miRNA	Samples 1-52	Samples 53-104
10	hsa-mir-148a	Yes*	Yes**
11	hsa-mir-155	No*	No**
12	hsa-mir-15a	Yes*	Yes**
13	hsa-mir-16	Yes*	Yes**
14	hsa-mir-17	Yes*	Yes**
15	hsa-mir-182	No*	No**
16	hsa-mir-18a	No*	No**
17	hsa-mir-192	Yes*	Yes**
18	hsa-mir-200c	No*	No**
19	hsa-mir-205	No*	No**
20	hsa-mir-20a	Yes*	Yes**
21	hsa-mir-20b	Yes*	Yes**
22	hsa-mir-21	Yes*	Yes**
23	hsa-mir-212	No*	No**
24	hsa-mir-218	No*	No**
25	hsa-mir-221	Yes*	Yes**
26	hsa-mir-222	Yes*	Yes**
27	hsa-mir-23a	Yes*	Yes**
28	hsa-mir-23b	Yes*	Yes**
55	hsa-mir-186	Yes**	Yes**
56	hsa-mir-195	No**	No**
57	hsa-mir-196a	No**	No**
58	hsa-mir-199a-3p	Yes**	Yes**
59	hsa-mir-199a-5p	Yes**	Yes**
60	hsa-mir-19a	Yes**	Yes**
61	hsa-mir-19b	Yes**	Yes**
62	hsa-mir-208a	No**	No**
63	hsa-mir-208b	No**	No**
64	hsa-mir-210	No**	No**
65	hsa-mir-211	No**	No**
66	hsa-mir-214	No**	No**
67	hsa-mir-215	No**	No**
68	hsa-mir-22	Yes**	Yes**
69	hsa-mir-27b	No**	No**
70	hsa-mir-28-5p	No**	No**
71	hsa-mir-296-3p	No**	No**
72	hsa-mir-296-5p	No**	No**
73	hsa-mir-299-3p	No**	No**

20

30

40

【表 6 - 4】

	miRNA	Samples 1-52	Samples 53-104		miRNA	Samples 1-52	Samples 53-104
29	hsa-mir-24	Yes*	Yes**	74	hsa-mir-299-5p	No**	No**
30	hsa-mir-26a	Yes*	Yes**	75	hsa-mir-302a	No**	No**
31	hsa-mir-27a	Yes*	Yes**	76	hsa-mir-302b	No**	No**
32	hsa-mir-32	No*	No**	77	hsa-mir-302c	No**	No**
33	hsa-mir-342-5p	No*	No**	78	hsa-mir-30a	Yes**	Yes**
34	hsa-mir-429	No*	No**	79	hsa-mir-30c	Yes**	Yes**
35	hsa-mir-451	Yes*	Yes**	80	hsa-mir-30e	Yes**	Yes**
36	hsa-mir-9	No*	No**	81	hsa-mir-325	No**	No**
37	hsa-mir-103	Yes*	Yes**	82	hsa-mir-330-3p	No**	No**
38	hsa-mir-93	Yes*	Yes**	83	hsa-mir-330-5p	No**	No**
39	hsa-let-7c	Yes**	Yes**	84	hsa-mir-331-3p	Yes**	Yes**
40	hsa-let-7f	Yes**	Yes**	85	hsa-mir-331-5p	No**	No**
41	hsa-mir-107	Yes**	Yes**	86	hsa-mir-340	No**	No**
42	hsa-mir-125a-3p	No**	No**	87	hsa-mir-342-3p	Yes**	Yes**
43	hsa-mir-125a-5p	Yes**	Yes**	88	hsa-mir-34b	No**	No**
44	hsa-mir-129-3p	No**	No**	89	hsa-mir-378	Yes**	Yes**
45	hsa-mir-129-5p	No**	No**	90	hsa-mir-92a	Yes**	Yes**

* Assessed as part of Example 1, ** Assessed as part of Example 2

* 実施例 1 の一部として評価

* * 実施例 2 の一部として評価

【 0 1 8 3 】

実施例 1 での記載と同様の方法論 (methodology) は、このデータセットの分析のために使用された。リーブ・ワン・アウト・交差検定 (leave-one-out cross validation) (LOOCV) と共に、ペナルティロジスティック回帰を使用することで、0.778 に等しい AUC が提示された。個々の miRNA が事前検定スコアの計算に使用されるモデルで選択された回数は、表 7 に示されている (50 個のサンプルがあったので、合計 50 モデル)。平均モデルサイズは、~ 8 ターム (上位 8 の miRNA は「*」で示されている) であった。期待値は、プールされたデータに対して得られた対応値よりも高い。

【表 7】

miR	Counts
miR.378*	50
miR.92a*	50
miR.26a*	50
miR.130a*	48
miR.222*	41
miR.15a*	38
miR.125a.5p*	33
let.7b*	28
miR.331.3p	25
miR.221	18
miR.30e	9
miR.199a.3p	1
miR.22	1
miR.199a.5p	1
miR.20a	1
let.7a	1

10

【 0 1 8 4 】

表 8 は、4 分割交差検定での L 1 ペナルティロジスティック回帰アプローチが 50 の個人サンプルに適用された時に選択された miRNA を提供する。再び、マーカーと順番のかなりの重複は、この 2 つの方法間で観察されている。図 4 は、この分析から得られた AUC 値の分布を示す。

【表 8】

miR	Counts
miR.378	400
miR.92a	396
miR.26a	366
miR.130a	233
miR.125a.5p	172
miR.222	152
miR.15a	146

20

【 0 1 8 5 】

(実施例 3)

タンパク質バイオマーカーの分析

(実施例 1、2 で使用されたマーシュフィールド・コホート (Marshfield cohort) からの) タンパク質のデータのみを含むモデルが、開発された。総計 47 のユニークなタンパク質バイオマーカー (表 9) が分析された。血清サンプルは、回収され、- 80 で凍結保存され、その後、使用直前に解凍された。各サンプルは、2 つの異なる検出技術：ルミネックス社 (Luminex) (Austin, TX) の xMAP (登録商標) 技術と、メソスケールディスカバリ社 (Meso Scale Discovery) (MSD, Gaithersburg, MD) の MULTI-SPOT (登録商標) 技術と、を使用して、分析された。

30

【表 9】

Protein Biomarker				
Adiponectin	IFN-a2	LIF	MPO	sE-Selectin
ANG-2	IL-12p40	MCP-1	NTproBNP	sFas
b-NGF	IL-16	MCP-2	PAI-1	sICAM-1
CRP	IL-18	MCP-3	RANTES	sP-Selectin
CTACK	IL-1a	MCP-4	Resistin	TIMP-1
EGF	IL-2Ra	M-CSF	SCD40L	TIMP-4
Eotaxin	IL-3	MIF	SCF	TNF-b
FASLigand	IP-10	MIG	SCGF-b	TRAIL
GROa	I-TAC	MIP-1a	SDF-1a	VEGF
HGF	Leptin			

40

【 0 1 8 6 】

ルミネックスの xMAP (登録商標) 技術は、色分けされた微粒子にプレコートされた分析物特異的抗体を利用している。微粒子、スタンダード、サンプルをウェルに分注し、

50

固定化抗体は、対象分析物と結合する。適切な時間インキュベーションした後、粒子は、すべての非結合物質を除去するため、複数回、洗浄バッファー内で再懸濁される。対象分析物に特異的なビオチン化抗体カクテルを各ウェルに添加する。すべての非結合ビオチン化抗体を除去するための、第2のインキュベーション時間と洗浄を終えた後、ビオチン化検出抗体に結合する、ストレプトアビジン - フィコエリスリン結合体 (streptavidin-phycoerythrin conjugate) (ストレプトアビジン - PE) を、各ウェルに添加する。最後の洗浄で、非結合のストレプトアビジン - PE を除去し、微粒子を緩衝液に再懸濁して、ルミネックス・アナライザを使用して読み込む。アナライザは、マルチレーザー検出システムを介して微粒子を方向付けるためにフローセルを使用している。1つのレーザーは、微粒子特異的であり、対象物が検出されていることを決定する。他のレーザーは、結合した分析物の量に直接比例して、フィコエリトリン由来の信号の大きさを決定する。曲線は、スタンダードによって生成された信号を使用して構築され、サンプルのタンパク質バイオマーカの濃度は、それぞれの曲線から読み込まれる。47のルミネックス (Luminex) タンパク質バイオマーカアッセイの感度 (検出限界、Limit of Detection, LOD) と精度 (アッセイ内およびアッセイ間 CV %) は、表 10 に示されている。

表 10 (表 10 - 1 ~ 表 10 - 2)

【表 10 - 1】

Protein Biomarker	LOD (pg/mL)	Avg Intra Assay %CV	Avg Inter Assay %CV
Adiponectin	682	9%	11%
ANG-2	18	4%	7%
b-NGF	1	7%	13%
CRP	525	7%	9%
CTACK	25	10%	10%
EGF	9	5%	14%
Eotaxin	1	15%	16%
FASLigand	1	9%	12%
GROa	31	3%	6%
HGF	28	4%	11%
IFN- α 2	13	2%	9%
IL-12p40	144	5%	9%
IL-16	15	4%	8%
IL-18	3	5%	6%
IL-1a	1	5%	19%
IL-2Ra	13	4%	10%
IL-3	31	4%	4%
IP-10	0	5%	11%
I-TAC	2	10%	17%
Leptin	28	6%	8%
LIF	66	28%	31%
MCP-1	6	3%	8%
MCP-2	1	7%	10%
MCP-3	19	6%	12%
MCP-4	2	4%	11%
M-CSF	8	4%	7%
MIF	24	5%	12%
MIG	6	7%	7%
MIP-1a	54	7%	13%
MPO	156	7%	12%
NTproBNP	96	7%	55%
PAI-1	9	5%	6%

【表 10 - 2】

Protein Biomarker	LOD (pg/mL)	Avg Intra Assay %CV	Avg Inter Assay %CV
RANTES	4	7%	6%
Resistin	9	5%	8%
SCD40L	115	4%	11%
SCF	9	4%	7%
SCGF-b	1017	4%	9%
SDF-1a	23	8%	10%
sE-Selectin	7	3%	7%
sFas	6	5%	6%
sICAM-1	70	6%	7%
sP-Selectin	218	4%	9%
TIMP-1	17	5%	6%
TIMP-4	27	5%	41%
TNF-b	8	5%	13%
TRAIL	24	3%	8%
VEGF	5	7%	9%

10

【0187】

45のユニークなタンパク質バイオマーカーの中の10は、MSDプラットフォーム（表11）のに10プレックスアッセイ（10 - p l e x a s s a y）で分析した。

【表 11】

20

Protein Biomarker
CTACK
HGF
IL-16
IL-18
MCP-3
M-CSF
MIF
MIG
NTproBNP
TRAIL

30

【0188】

MSD（メソスケールディスカバリー、M e s o S c a l e D i s c o v e r y）技術は、各プレートの底部を炭素表面で構築された、特殊な96ウェルマイクロタイタープレートを採用している。各タンパク質バイオマーカーに特異的な抗体は、マイクロタイタープレートの各ウェルの底部にアレイ（s p a t i a l a r r a y s）でスポットされている。スタンダードとサンプルを、プレコートプレートされたウェルに分注し、固定化抗体は、対象の分析物と結合する。適切な時間インキュベーションした後、プレートを複数回洗浄して、すべての非結合物質を除去する。S U L F O - T A G（登録商標）で標識された、分析物に特異的な二次抗体カクテルを各ウェルに添加する。2回目のインキュベーションの後、プレートを再度複数回洗浄して、任意の非結合物質を除去し、専用のリード・バッファ（R e a d B u f f e r）を各ウェルに添加する。次いで、プレートをS E C T O R（登録商標）イメージャーに設置すると、電流がマイクロタイタープレートの底部のカーボン電極に印加される。各スポットで、特異的二次抗体に結合したS U L F O - T A G（登録商標）標識は、この電気刺激時に発光し、高感度CCDカメラで検出される。曲線は、スタンダードで生成されたシグナルを使用して構築され、サンプルの蛋白質バイオマーカーの濃度は、それぞれの曲線で読み込まれる。10のMSD（メソスケールディスカバリー）タンパク質バイオマーカーアッセイの感度（検出限界、L O D）と精度（内およびアッセイ間C V %）は、表12に示されている。

40

【表 1 2】

Protein Biomarker	% Detected >LOD (pg/mL)	Avg Intra Assay %CV (FI)	Avg Inter Assay %CV (Conc)
CTACK	99%	9%	23%
HGF	99%	7%	15%
IL-16	99%	9%	11%
IL-18	99%	6%	8%
MCP-3	69%	6%	11%
M-CSF	99%	13%	34%
MIF	99%	5%	9%
MIG	99%	8%	14%
NTproBNP	99%	6%	27%
TRAIL	99%	9%	179%

10

【 0 1 8 9】

前述の事前検証済みスコア (prevalidated score) の計算に L O O V または K - 分割交差検定でロジスティック回帰アプローチを用いて、モデルが構築され、パフォーマンスが評価された。図 8 は、事前検証済みスコアを予測するために K - 分割交差検定法のみを用いたタンパク質ベースのモデルから得られた、A U C 値の分布を提供している。表 1 3 は、任意の交差検定モデルにおける、あるタンパク質マーカーの選択度数を提供する。高いカウントは、あるマーカーが一貫してコントロールから症例を分類する能力を持つことを示す。A U C は、事前検証済みスコアの計算に L O O V アプローチを使用して、0 . 6 9 8 であると計算され、および、表 1 4 は、L O O V 方法論を使用して構築した任意のモデル内のマーカーの、選択度数を提供している。後者の A U C は、k - 分割交差検定アプローチから計算された不確実性の限界内にある。どちらの手法も同じトップマーカーを選択している。

20

【表 1 3】

Marker	Counts
sP-Selectin	717
MPO	692
Eotaxin	536
IL-16	361
Resistin	249
VEGF	205
CRP	204
HGF	113

30

【表 1 4】

Marker	Counts
sP-Selectin	41
MPO	41
Eotaxin	38
IL-16	38

40

【 0 1 9 0】

(実施例 4)

m i R N A とタンパク質バイオマーカーの複合解析

(実施例 1、2 からの) タンパク質データと m i R N A のデータの両方を含むモデルが、開発された。(実施例 3 からの) 4 7 のバイオマーカー全体のタンパク質データは、ルミネックス (L u m i n e x C o r p , A u s t i n , T X) システムと、メソスケールディスカバリー (M e s o S c a l e D i s c o v e r y , M S D) システムの、2 つの異なる検出技術を用いて得られた。タンパク質と m i R N A のデータを組み合

50

わせたため、候補説明的変数 (candidate explanatory variables) の数がサンプルの数を超えている。このような状況では、非ペナルティメソッドの使用は適切ではないため、前述の事前検証済みスコア (prevalidated score) の計算に LOOV または K - 分割交差検定 (k-fold cross-validation) でペナルティ付きロジスティック回帰を用いて、モデルが構築され、パフォーマンスが評価された。図 5 は、miRNA とタンパク質の両方に基づいたモデルのための AUC 分布を提供する。AUC は、miRNA のみに対して得られたものと統計的に同等であるが、2 つの miRNA は一貫してモデルで選択されていた (表 15 を参照)。図 6 は、miRNA とタンパク質の相関の分布を示し、一方、図 7 は、miRNA の分布を示す。図 6 の 2 つの垂直線は、タンパク質と miRNA 間の最大と最小の相関関係を表している。いかなる特定の理論にも拘束されることを望まなければ、これらの相関関係は現在調査されていない調節作用に対応することができる。これら 2 つの図を比較すると、タンパク質は、このデータ・セット内の正の相関の高い数値を生成することを示している。

【表 15】

miR	Counts
miR.378	50
miR.26a	50
MPO	50
SP.SELECTIN	50
VEGF	50
EOTAXIN	48
M.HGF	44
miR.92a	32
RESISTIN	29
miR.125a.5p	25
M.IL.16	18
I.TAC	17

【0191】

(実施例 5)

miRNA バイオマーカーを用いた生存

本研究では、miRNA のレベルは、時間の経過とともに発生するあるイベント (ここでは M1) のリスクを説明している。112 の候補 miRNA マーカーの、単変量、多変量分類および生存分析を行った。分類結果は、実施例 2、3 に記載の手法に基づいて得られた。生存解析は Cox 比例ハザード回帰手法を用いて行った。後の分析のための応答変数は、時間を含み、それは、あるイベントが発生した時間、または、研究の終了までの時間、および、時間があるイベントまたは研究の終了 (打ち切り) に関連しているかどうかを示す指標、である。実施例 2 に記載の 52 サンプルについて、事象の時間またはフォローアップの終了時間が知られていた。研究の終了前にあるイベントを起こした 26 例については、ある事象のインジケータ変数は 1 に設定され、研究の期間内にある事象が起きなかった 26 例については、インジケータ変数は 0 に設定された。分析に含まれる説明変数は以下の通りで、: a) タンパク質レベル単独、b) miRNA レベル単独、および、c) miRNA レベル、タンパク質レベルのいずれか又は両方。モデルフィッティング (Model fitting) は、Cox 比例ハザードモデルのペナルティ付きバージョンとペナルティ無し (unpenalized) バージョンの両方を使用して行われた。モデルのペナルティ付きバージョンが適用される時は必ず、L1 - ペナルティ (ラッソ (Lasso)) を用いた。各モデルの変数選択 (variable selection) は、実施例 1 に記載したものと同一アプローチを使用して行った。すなわち、

a) モデルの、ペナルティ無し (unpenalized) バージョンのための前進選択 (forward selection) でのベイズ情報量規準 (the Bayesian information criterion) を使用して、及び、b) ペナルティ付きアプローチのための最適なペナルティの選択に基づいた交差検定 (cross-validation) を使用して。客観的な方法でこれらのモデルの性能を評価するために、実施例 1 に記載のものと同様の方法で得られた事前検証済みスコア (preva

l i d a t e d s c o r e) の計算が採用された。

【 0 1 9 2 】

最初の分析（分類）では、生存時間は無視され、イベントまでの時間（タイム・ツー・イベント（time - to - event））にかかわらず、すべてのケースが同じように処理された。表 1 6 は、単変量の分類分析（univariate classification analysis）の結果を示している。この表内のマーカーは、予測 AUC によって順位づけられている。表 1 8（表 1 7 のミス！）は、多変量分類モデルにおける miRNA の選択度数（the selection frequency）を示している。多重ロジスティック回帰モデルは、LOOV アプローチを通じて得られたトレーニングセットの事前検証プロセス中に構築され、無視されたサンプル（the left - out - sample）のスコアを提供した。モデルのサイズは、ベイズ情報量規準（the Bayesian Information Criterion）の使用によって決定した。平均的な分類性能は、事前検証済み分類スコアのベクトルに基づいており、0.7 に等しかった。

表 1 6（表 1 6 - 1 ~ 表 1 6 - 2）

【表 1 6 - 1】

	Estimate	Std. Error	z value	Pr(> z)	AUC
hsa.miR.378	-1.40	0.42	-3.33	0.00	0.84
hsa.miR.1974	0.68	0.30	2.29	0.02	0.76
hsa.miR.26a	0.74	0.28	2.61	0.01	0.76
hsa.miR.30b	0.95	0.35	2.75	0.01	0.74
hsa.miR.29c	-0.71	0.30	-2.34	0.02	0.74
hsa.miR.34a	-0.62	0.29	-2.11	0.03	0.73
hsa.miR.30c	0.71	0.31	2.28	0.02	0.72
hsa.miR.221	0.86	0.33	2.63	0.01	0.72
hsa.miR.192	-0.87	0.33	-2.60	0.01	0.72
hsa.miR.122	-0.76	0.30	-2.51	0.01	0.71
hsa.miR.19a	-0.54	0.29	-1.86	0.06	0.71
hsa.let.7a	0.67	0.31	2.15	0.03	0.71
hsa.miR.21	-0.77	0.33	-2.34	0.02	0.7
hsa.miR.497	-0.78	0.32	-2.45	0.01	0.7
hsa.miR.19b	-0.52	0.29	-1.79	0.07	0.7
hsa.miR.148a	-0.69	0.30	-2.29	0.02	0.7
hsa.miR.15b	-0.53	0.27	-1.94	0.05	0.69
hsa.miR.331.3p	0.65	0.30	2.19	0.03	0.69
hsa.miR.24	0.68	0.30	2.30	0.02	0.69
hsa.miR.142.5p	0.68	0.35	1.95	0.05	0.69

【表 1 6 - 2】

	Estimate	Std. Error	z value	Pr(> z)	AUC
hsa.miR.99a	-0.76	0.31	-2.42	0.02	0.69
hsa.miR.25	-0.47	0.29	-1.62	0.11	0.69
hsa.miR.29a	-0.86	0.36	-2.41	0.02	0.69
hsa.miR.22	-0.54	0.30	-1.77	0.08	0.68
hsa.miR.652	0.67	0.34	1.94	0.05	0.68
hsa.miR.92a	-0.40	0.28	-1.41	0.16	0.68
hsa.miR.140.3p	-0.48	0.29	-1.63	0.10	0.68

【表 1 7】

miRNA biomarker	Counts
hsa.miR.378	47
hsa.miR.497	47
hsa.miR.24	45
hsa.miR.126	45
hsa.miR.21	42
hsa.miR.15b	38
hsa.miR.652	33
hsa.miR.29a	26
hsa.miR.99a	17
hsa.miR.30b	10
hsa.miR.29c	6
hsa.miR.331.3p	4
hsa.miR.19a	4

【 0 1 9 3 】

表 1 8 は、単変量生存率分析の結果を示している。再び、このテーブル内のマーカーは

、予測 AUC によって順位づけられている。トップの選択マーカーは、分類分析から得られたものとほぼ同一であり、全体的なパフォーマンスは、時間依存性 AUC で評価されるように、分類アプローチから得られたそれと同等であった。表 19 は、Cox 比例ハザード回帰アプローチを用いた多変量生存率解析での miRNA マーカーの選択度数を示している。モデルにのみ基づいた miRNA のために期待されるパフォーマンスは、事前検証 (AUC = 0.78) を用いて推定した。トレーニングセットは、リーブ・ワン・アウト (leave-one-out) アプローチを介して構築され、各区画内でのモデルのサイズは、ベイズ情報量基準 (the Bayesian information criterion) に基づいて決定された。平均的なモデルのサイズは 8 であった。

【表 18】

	coef	exp(coef)	se(coef)	z	Pr(> z)	AUC
hsa.miR.378	-0.5	0.61	0.13	-3.68	0	0.62
hsa.miR.1974	0.24	1.27	0.15	1.62	0.11	0.74
hsa.miR.29c	-0.45	0.64	0.19	-2.4	0.02	0.74
hsa.miR.26a	0.36	1.44	0.17	2.09	0.04	0.74
hsa.miR.30b	0.42	1.52	0.19	2.2	0.03	0.72
hsa.miR.30c	0.33	1.39	0.19	1.76	0.08	0.72
hsa.miR.34a	-0.3	0.74	0.16	-1.85	0.06	0.71
hsa.miR.192	-0.4	0.67	0.19	-2.13	0.03	0.7
hsa.miR.122	-0.4	0.67	0.18	-2.23	0.03	0.7
hsa.miR.221	0.27	1.31	0.12	2.24	0.03	0.7
hsa.miR.331.3p	0.41	1.51	0.18	2.33	0.02	0.7
hsa.miR.497	-0.44	0.65	0.18	-2.44	0.01	0.7
hsa.miR.652	0.41	1.51	0.19	2.12	0.03	0.7
hsa.miR.21	-0.48	0.62	0.21	-2.3	0.02	0.7
hsa.let.7a	0.32	1.38	0.2	1.64	0.1	0.69
hsa.miR.148a	-0.29	0.75	0.15	-1.91	0.06	0.69
hsa.miR.29a	-0.58	0.56	0.21	-2.75	0.01	0.69
hsa.miR.19a	-0.26	0.77	0.18	-1.47	0.14	0.68
hsa.miR.19b	-0.19	0.83	0.17	-1.09	0.28	0.68
hsa.miR.15b	-0.34	0.71	0.17	-2.01	0.04	0.68

【表 19】

miRNA biomarker	Counts
hsa.miR.21	47
hsa.miR.378	47
hsa.miR.652	47
hsa.miR.497	47
hsa.miR.15b	47
hsa.miR.99a	41
hsa.miR.22	24
hsa.miR.126	13
hsa.miR.29a	7
hsa.let.7b	5
hsa.miR.502.3p	5

【0194】

(実施例 6)

拡張された miRNA スクリーニング

miRNA バイオマーカーの、症例 (case) 対コントロールを区別するための能力を調べるために、現在 miRBASE 1.3 に更新されている、Exiqon の水銀 LNA (登録商標) ユニバーサル RT マイクロ RNA PCR アレイ技術プラットフォームを使用して、以前に、実施例 2 の 52 の血清サンプルから得た RNA 抽出物を、表 1 に示す 720 の miRNA 標的配列の存在に関してスクリーニングした。

【0195】

多くの分析が、各 miRNA バイオマーカーの全体的な意義を提供するために結合された。単変量分類と生存分析は、重要な順に各ターゲットをランク付けするために使用された各個人の miRNA ターゲットのために、AUC 値を提供した。多変量解析もまた、47 の多変量モデルを生成するために実施された。miRNA ターゲットは、それらが選択されたモデルの数によってランク付けされた。t-検定分析 (1-tailed) もまた、症例 (case) 集団とコントロール集団における各 miRNA ターゲットのために測定された Cp 値を比較して行われた。最後に、四分位数分析 (a quartile a

10

20

30

40

50

analysis) が、データセットに対して行われた。各 miRNA ターゲットについては、すべてのサンプル (結合された症例集団とコントロール (対照) 集団) が、Cp 値に従って (低い値から高い値へ) ランク付けされた。ランク付けされた集団は、その後、各々が総集団の 25% を含む 4 つの四分位数 (four quartiles) に分割された。各四分位数での症例被験者と対照被験者の数は、その後記録された。もし、総数 26 症例のうち、65% より多いまたは 35% 未満が、「低い」四分位数にランクされた場合には、その miRNA ターゲットは、重要とみなされた。

【0196】

拡張セットの 720 の miRNA バイオマーカーの分析に基づいて、最終的な全体的なランクスコアが割り当てられるが、これは、miRNA ターゲットのセット全体がランク付けされたことにより、全体の重要度スコアの生成を説明している。表 20 は、トップ 50 にスコアされている miRNA を示している。

【表 20】

Biomarker	SCORE	Rank
miR-378	437	1
miR-497	411	2
miR-21	392	3
miR-15b	359	4
miR-99a	357	5
miR-652	356	6
miR-30b	345	7
miR-26a	335	8
miR-29a	329	9
miR-1974	327	10
miR-30c	325	11
miR-122	322	12
miR-29c	321	13
miR-192	321	14
miR-34a	319	15
miR-24	318	16
miR-221	317	17
miR-126	314	18
miR-331-3p	307	19
let-7a	299	20
miR-148a	296	21
let-7g	288	22
miR-19a	287	23
miR-142-5p	284	24
miR-22	283	25
miR-19b	272	26
miR-151-5p	262	27
miR-215	261	28
miR-25	258	29
let-7f	255	30
miR-10b	252	31
miR-423-3p	251	32
miR-502-3p	246	33
miR-140-3p	238	34
miR-92a	235	35
miR-660	233	36
miR-142-3p	229	37
miR-130a	218	38
miR-185	217	39
let-7c	215	40
miR-18a	210	41
miR-365	203	42
miR-26b	194	43
miR-125b	178	44
miR-297	171	45
miR-146a	151	46
miR-99b	104	47
miR-424	76	48
miR-93	60	49
let-7b	14	50

【0197】

(実施例 7)

タンパク質バイオマーカーベースの心血管リスクスコア

PMRP (Personalized Medicine, 2(1): 49-79 (2005)) からの 1123 人のサンプルによって、心血管リスクスコアの開発がなされた。セットは、症例コホート・デザイン (a case-cohort design) に基づいて選択された。ベースラインの採血の時に 40-80 歳であった場合、および、

事象 M I を有した、または 5 年フォローアップの間に不安定狭心症 (U A) のために入院した場合は、P M R P コホートからの被験者は、「症例」(c a s e s) とみなした。合計で、385 人の症例 (初期 M I の 164 例、および U A の 221 例) と 838 人のコントロールであった。利用可能なデータは、個々で測定された 59 (47 のユニークな) のタンパク質バイオマーカー、および 107 の臨床的特徴であって、人口統計 (年齢、性別、人種、糖尿病の状態、M I の家族歴、喫煙など) と、実験室での測定 (総コレステロール、H D L、L D L など) や薬の使用 (スタチン、降圧薬、血糖降下薬など) を、含んでいた。

【0198】

単変量解析

各バイオマーカーと患者の転帰との関連付けは、C o x 比例ハザード回帰と、H e a g e r t y ら (S u r v i v a l M o d e l P r e d i c t i v e A c c u r a c y a n d R O C C u r v e s B i o m e t r i c s, 61:92-105 (2005)) の K a p l a n - M e i e r 法を使用した曲線 (A U C) での時間依存的領域と、を使用して評価された。共通のスケール上で異なる濃度範囲のすべてのタンパク質バイオマーカーに亘ってハザード比 (h a z a r d r a t i o、H R) を提示するために、すべての被験者の値は、データをログ変換後コントロールの標準偏差で割ったコントロールの濃度の平均値を差し引くことにより、正規化された。ハザード比は、このように 1 つの標準偏差単位ごとに、表現された。図 9 は、リスクの多変量モデルを開発するための候補として使用された 35 のバイオマーカーの、未調整のハザード比と標準誤差を示す。22 のバイオマーカーは、統計的に有意な H R を有する。

【0199】

年齢、性別、収縮期血圧、拡張期血圧、コレステロール、H D L、高血圧症、高血圧症治療薬の使用、高脂血症、糖尿病、喫煙という、従来型危険因子 (t r a d i t i o n a l r i s k f a c t o r s、T R F s) に対して各バイオマーカーを調整しながら、同じ分析を繰り返した (図 10)。調整後、11 のバイオマーカーだけが、統計的有意性を維持していたが、選択された上記 T R F s が、心血管疾患に関連付けられていることは既知であったので、驚くべきことではない。図 11 A、図 11 B は、最も高い時間依存性 A U C と 5 年間のフォローアップでのその対応値を有するマーカーを示している。すべてのマーカーの A U C は、時間とともに減少していた N T - プロ B N P (N T - p r o B N P) アッセイの 2 つのバージョンを除き、時間とともに一定のままであった

【0200】

多変量解析: M I (心筋梗塞) および / または U A (不安定狭心症) の予後スコアの開発

予後スコアの開発は、タンパク質バイオマーカーと同様、T R F s (従来型危険因子) を含めることを基礎としていた。年齢、性別、糖尿病、および家族歴の心血管イベントとの関連性が既知であったならば、これら 4 つのパラメータは、モデルに含まれた。これら 4 つのパラメータを含めることは、沢山のフォワードマーカー選択アルゴリズム (f o r w a r d m a r k e r s e l e c t i o n a l g o r i t h m s) を実行することによって確認された。すべてのアルゴリズムは、最終的な多変量アルゴリズムで 4 つの変数を選択した。最適なモデルサイズの決定は、以下の基準の使用に基づいてなされた。(a) 赤池情報量基準 (A k a i k e i n f o r m a t i o n c r i t e r i o n)、(b) ベイズ情報量基準 (B a y e s i a n i n f o r m a t i o n c r i t e r i o n)、(c) ドロップ・イン・逸脱基準 (D r o p - i n - d e v i a n c e c r i t e r i o n)、である。最初の 2 つは、インサンプルエラー推定法 (i n - s a m p l e e r r o r e s t i m a t o r s) で知られており、3 番目は、適合度 (g o o d n e s s - o f - f i t) を推定する交差検証ループ (c r o s s - v a l i d a t i o n l o o p) を利用している。3 つのすべてのケースでは、モデルサイズは、最良のデータに適合し、過学習を回避するモデルに対して選ばれた。モデル選択のための特徴的なドロップイン・逸脱曲線 (d r o p - i n - d e v i a n c e c u r v e) (量の絶対値のプロット) を図 12 に示す。モデルのサイズは、曲線の最大値が同定され、その後

、最大値の下に1標準誤差ポイントからラインを引く、1標準誤差ルール(1 standard error rule)の使用に基づいて選択された。タンパク質バイオマーカーの最適な数は、対応する平均絶対逸脱値が前述のラインを超える最小数として選ばれた。その数は7つのタンパク質バイオマーカーと一致し、すなわち、最適ナリスクスコアは、それ故、4つのTRFsと7つのタンパク質バイオマーカーから構成されていた(図12)。すべての3つの方法は、モデル内のバイオマーカーの最適な数として5~7のマーカーを選択した。バイオマーカーのより小さなセットは、常により大きなセットのサブセットであった。表21は、年齢、性別、糖尿病、およびM1の家族歴が、モデル内に入れ込まれた後に、選択されたバイオマーカーの度数とランキングを示している。これらのカウントとランキングは、交差検定プロセスの間に構築された別のモデルから得られた。1つのモデルは、すべてのトレーニング集団(training fold)のために構築され、そのサイズは、前述のモデル選択法の1つによって選択される。交差検定プロセスは、順番に平均的に各被験者のメンバーシップの割り当てによって導入された変動にわたって繰り返された。

【表21】

Biomarker	Counts (out of 20)	Average Rank	Min Rank	Max Rank
EOTAXIN	20	3.7	2	7
IL16	20	1.05	1	2
MCP.3	20	4.4	2	7
CTACK	17	2.9	2	5
ADIPONECTIN	16	5.4	2	9
HGF	12	5.1	1	9
FASLIGAND	10	6.0	2	8
SFAS	10	6.6	5	8
IL18	9	7.7	4	12
TIMP.4	7	7.0	3	11
TIMP.1	5	8.4	5	12
CRP	4	6.3	4	9
HGF	4	7.5	3	11
VEGF	3	7.7	7	8
EGF	1	6.0	6	6

【0201】

表21は、5倍の事前検証(交差検証のフォーム)プロセスの4回繰り返し以上の各マーカーの、頻度選択(the frequency selection)、平均値、最小と最大のランクを示す。4つのTRFsは、各モデルに含めた。

【0202】

ドロップ・イン・逸脱アプローチ(the drop-in-deviance approach)で予測される最適なモデルサイズを使用して、Cox比例ハザードモデルは、異なる集団の検証を使用することができるモデルを得るために使用可能なすべてのデータにフィットした。この最終的なタンパク質ベースのモデルには、選択した順番で、次のタンパク質バイオマーカー、IL-16、エオタキシン、fasリガンド、CTACK、MCP-3、HGF、およびsFas、が含まれていた。

【0203】

(実施例8)

タンパク質モデルと他の標準的な予測モデルとの比較

心血管イベント(すなわち、M1またはUA)の予測リスクの開示モデルのトランスポータビリティは、45-84歳の米国の集団から選択された第2の多民族コホートで評価した(アテローム性動脈硬化症コホートの多民族的研究)[Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR, Greenland P, Jacob DR, Jr., Kronmal R, Liu K, Nelson JC, O'Leary D, Saad MF, Shea S, Szklo M, Tracy RP. アテローム性動脈硬化症の多民族的研究: 目的とデザイン Am J Epidemiol. 2002; 156(9): 871-88

1]。

【 0 2 0 4 】

開発に使用されたサンプルと、似て非なるサンプルでモデルの期待される性能を確立するために、事前検証メソッドが、第2の集団へモデルを適用する前に、再び使用された。2つのパフォーマンス・メトリック、純再分類指数 (the Net Reclassification Index、NRI) と臨床再分類指数 (the Clinical Net Reclassification Index、CNRI)、が使用された。純再分類指数の定義は次式で与えられる。

【 数 1 6 】

$$NRI = \frac{\text{Cases Up} - \text{Cases Down}}{\text{No. of cases in risk category}} - \frac{\text{Controls Up} - \text{Controls Down}}{\text{No. of controls in risk category}}$$

10

【 0 2 0 5 】

式は、パーセントの観点から、症例 (cases) とコントロールに対する改善を別々に評価し、単一の数値に結果を組み合わせる。症例に対する正パーセンタイル値とコントロールに対する負の正パーセンタイル値は、開示されたモデルで導入された性能の向上を表している。リスクカテゴリーは、既存の開示モデルによって予測されるリスクスコアのための、適切なしきい値を確立することによって定義されている。CNRIは同じ方法で定義されるが、改良された、グループ内の真のリスクの同定方法から得ることのできる集団のサブセットに適用される。心血管疾患については、例えばフラミンガム (Framingham) スコアによって定義された中程度リスク集団で、NRIメトリックのアプリケーションは、この基準を満たす。計算値は、中程度リスクカテゴリーのCNRIパフォーマンスを表している。

20

【 0 2 0 6 】

伝統的に、10年リスクに対するラミンガム・スコアで計算された中程度リスクカテゴリーは、10%から20%の間のリスク・スコアを持つ個人として定義されている。ここで示した結果は、中程度リスクカテゴリーを定義するために、以下のカットオフ値：< 3.5%、> 7.5%、を基礎としている。a) 開示モデルは5年という期間に焦点を当て、およびb) フラミンガム・スコアが開示された場合、現在の集団でのイベント率は、観察値よりも低いので、これらのより低いカットオフ値の使用が正当化される。

30

【 0 2 0 7 】

再分類比較は、与えられた対象に対して、各モデルからの絶対リスクの計算を必要とした。Cox比例ハザード (Cox PH) モデルを使用した各個人の絶対リスクの計算は、その特性とベースラインハザード推定とに基づいて、この個々人のための相対リスクの計算を必要とした。Cox PHモデルは相対的リスクを予測するために設計されているが、ハザード関数の特定を必要としない。Cox PHモデルから絶対リスク推定値を生成するために、我々は個々人の、または、「平均的」個人の絶対リスクを必要とし、この個人の、または平均人の、相対的なリスク推定を使用して、個々人の絶対リスクを計算した。平均人は、それぞれの予測因子に対する集団平均値を有する、仮想上の個人である。集団の真のベースライン・ハザードとそれに対応する「平均的な」人が知られていなければ、(心血管イベントのリスクを計算するための正しいモデルが未知であるため)、推定が提供される必要があった。R言語 [R: 統計コンピューティングのための言語と環境、R開発コアチーム、統計コンピューティングのためのR財団、ウィーン、オーストリア、2010] サーフフィット (survfit) 関数が、平均的な個人のベースラインハザードを計算するために使用された。サーフフィット (survfit) 関数は、計算に重みを使用している。つまり、集団の各メンバーは、平均と比較して推定されたりリスクスコアに応じて重みを受容し、加重ハザード推定値は、ベースラインハザードのために使用される。ベースラインハザードの推定は、使用されたモデルに依存し、予測された相対的リスクにも依存する。開示モデル対FRS及びTRFベースのモデルの、再分類性能の公正な比較を行うために、適切なベースラインハザード推定が、不当にいずれかのモデルを支

40

50

持していないことが必要とされた。比較対象の2つのモデルの平均スコアであるリスク・スコアを使用してベースラインハザードを計算するための好ましいアプローチは、以下に説明する。さらに、サーブフィット (survfit) 関数は、カプラン・マイヤー (Kaplan-Meier) とアーレンの、2つの異なる推定法を実行した。両方の推定法は、テストの結果、観測された差異はごくわずかなものであった。集団に我々の結論を拡張するために、ベースライン生存関数は、研究のケース・コホートの重みを使用して、共変量の集団平均で評価された。

【0208】

絶対的なリスク・スコアに関して、二つのモデルを比較するためのベースラインハザードの推定の選択は難しい問題であり、文献では取り上げていない。集団の真のベースラインハザードが未知であるため、各モデルによって異なる推定を使用すると、比較の結果に重大な影響が生じ得る。ベースラインハザード推定の影響を調べるために、2つの異なる方法を用いて、すべての計算が行われた。1) 線形予測スコアを使用した個々のベースライン生存者の推定値に基づいて、各モデルの絶対リスクスコアが、各モデルによって計算され、2) 2つのスコアから平均線形予測を計算することによって得られる共通のベースライン生存者の推定値に基づいた、絶対リスクスコアは、集団平均で中央を占める。

【0209】

表22、表23、および表24は、以下の3つの代替モデルに対して、バイオマーカーを含む事前検証モデルのパフォーマンスが期待された、NRIとCNRIを提示している。1) フラミンガム・リスクスコア (Framingham risk score、FRS)、2) 共変量として4つのTRFs (「4-TRF」：年齢、性別、糖尿病、M1の家族歴) を使用してマーシュフィールド・データ (Marshfield data) に適合されたモデル、3) 共変量として9のTRFs (「9-TRF」：年齢、性別、糖尿病、M1の家族歴、喫煙、総コレステロール、HDL、高血圧治療薬服用、収縮期血圧) を使用してマーシュフィールド・データに適合された代替モデル。

【0210】

全体的に、タンパク質バイオマーカーを含有するモデルは、心血管イベントに対する5年リスクの3.5 - 7.5%および3.5 - 10%の範囲で、FRSまたはTRFベースモデルの、より優れた再分類を提供した。表22は、事前検証 (マーシュフィールドのデータセット) に基づいて校正されたFRSスコアに対する、本開示のモデルのスコアの期待される再分類パフォーマンスを示している。表23および24は、それぞれ、事前検証 (マーシュフィールドのデータセット) に基づいた4-TRFと9-TRFモデルのスコアに対する、期待される再分類スコアを示している。

【0211】

NRIとCNRIの両面で、総合的な再分類は、ベースライン生存関数を計算するための2つの方法のいずれの使用においても、同程度であった。しかし、2つの方法の総NRIまたは総CNRIを構成する、症例 (cases) およびコントロールの再分類のバランスに、違いがあった。共通のベースライン生存関数法は、よりバランスのとれた再分類を提供してくれた。この結果は、モデルの相対的リスク予測のために得られた結果と一致した。図13A-Bは、FRS、本開示モデル (事前検証のアプローチの複数のリピートから得られる)、4-TRF、および9-TRFモデルのリニアスコア (linear scores) のカーネル密度推定の観点から、この比較を示している。本開示モデルのスコアは、他のモデルよりもより高く症例 (case) の相対リスクを提供した。コントロールの分布はまた、本開示モデルのスコアのためにより広く、他のスコアに比してコントロールのリスクの上下のバランスを示した。これらの結果は、本開示モデルスコアが、他のスコアに関連する症例を正しく分類したという強力な特徴を提供した。

【0212】

(平均スコアを使用した) 共通のベースライン生存関数法は、また、予測精度を向上させるための投票方式 (voting scheme) (すなわち、加重平均) を使用する多くの統計的アプローチと一致した。

10

20

30

40

50

【表 2 2】

	Range	Baseline Hazard calculation	NRI(sd)	NRI_case	NRI_ctrl	CNRI(sd)	CNRI_case	CNRI_ctrl
FRS	3.5-7.5 %	Individual	10.34% [1.85%]	6.1% [2.11%]	-4.24% [0.66%]	44.52% [4.5%]	2.95% [4.8%]	-41.56% [1.83%]
		Average	15.18% [2.26%]	23.23% [1.45%]	8.05% [1.42%]	48.51% [5.42%]	27.33% [3.31%]	-21.19% [4.05%]
	3.5-10.0 %	Individual	9.39% [2.1%]	5.41% [1.46%]	-3.98% [0.8%]	42.19% [4.92%]	1.74% [3.41%]	-40.45% [2.76%]
		Average	15.94% [1.2%]	24.23% [1.69%]	8.28% [0.88%]	44.07% [2.05%]	21.31% [3.06%]	-22.76% [2.59%]

事前検証（マーシュフィールドのデータセット）に基づいて校正されたフラミンガム・スコア（Framingham score）に対するアヴィール・スコア（Aviir score）の予測される再分類パフォーマンス

10

【表 2 3】

	Range	Baseline Hazard calculation	NRI(sd)	NRI_case	NRI_ctrl	CNRI(sd)	CNRI_case	CNRI_ctrl
4-TRF	3.5-7.5 %	Individual	6.92% [1.39%]	5.3% [1.71%]	-1.62% [0.69%]	33.42% [3.58%]	11.38% [3.99%]	-22.04% [3.12%]
		Average	13.24% [2.2%]	24.39% [1.86%]	11.15% [0.72%]	31.52% [4.72%]	34.64% [3.71%]	3.12% [3.04%]
	3.5-10.0 %	Individual	9.56% [2.4%]	7.32% [2.04%]	-2.24% [0.76%]	29.83% [3.84%]	6.61% [2.79%]	-23.22% [2.31%]
		Average	15.23% [1.86%]	25.91% [1.76%]	10.68% [0.48%]	31.66% [3.76%]	29.07% [3.27%]	-2.78% [1.7%]

事前検証（マーシュフィールドのデータセット）に基づいた4 - T R Fモデルスコアに対するアヴィール・スコア（Aviir score）の予測される再分類パフォーマンス

20

【表 2 4】

	Range	Baseline Hazard calculation	NRI(sd)	NRI_case	NRI_ctrl	CNRI(sd)	CNRI_case	CNRI_ctrl
9-TRF	3.5-7.5 %	Individual	-0.1% [1.52%]	-1.23% [1.69%]	-1.12% [0.81%]	29.86% [4.23%]	4.94% [3.53%]	-24.93% [2.73%]
		Average	3.95% [1.81%]	9.78% [1.77%]	5.83% [0.66%]	26.77% [3.78%]	19.95% [3.68%]	-8.82% [1.86%]
	3.5-10.0 %	Individual	1.9% [1.7%]	0.73% [1.71%]	-1.17% [0.73%]	28.25% [3.8%]	1.85% [2.67%]	-26.3% [2.46%]
		Average	7.19% [1.84%]	12.65% [1.54%]	5.46% [0.76%]	28.35% [3.83%]	16.32% [2.94%]	-12.03% [2.05%]

事前検証（マーシュフィールドのデータセット）に基づいた9 - T R Fモデルスコアに対するアヴィール・スコア（Aviir score）の予測される再分類パフォーマンス

30

【0 2 1 3】

（実施例 9）

30

第 2 集団への開示モデルのトランスポータビリティ（transportability）

複数の集団間の予後モデルのトランスポータビリティの問題は、予測モデルの有用性のための究極のテストを提供する。モデルの統計的および臨床的妥当性は、モデルのトランスポータビリティの、等しく重要な側面である。1) 内部検証、2) 時間的検証、3) 外部検証の、3 段階の検証アプローチが、新しいテストのために提案されている。モデリング手法を検証するために、事前検証のアプローチ（交差検定フォーム）を使用した第 1 ステップの完了は、上述した。第 2 ステップは、同じ集団や臨床センターからの異なる患者セットでのアルゴリズムテストをする必要がある。マーシュフィールド・スタディと現在の時間との間の、最後のイベントが起きた時間が、単に短い期間（2 年程度）であるなら、後発イベントの数は、同じ集団内の検証のためには小さすぎであった。したがって、外部検証ステップは、本開示のタンパク質モデルのトランスポータビリティのデモンストレーションとして設定された、MESA サンプルに、本開示のタンパク質モデルをテストすることによって実施された。

40

【0 2 1 4】

MESA コホートにおける本開示モデルのパフォーマンスを評価するために、824 のサンプル（222 人の症例および 602 人のコントロール）が、実施例 7 に記載の蛋白質バイオマーカー（IL - 16、エオタキシン（eotaxin）、fas リガンド、CTACK、MCP - 3、HGF、および sFas）のパネルを使用してアッセイされた。

50

【0 2 1 5】

マーシュフィールド訓練済みモデル (Marshfield-trained model) は、MESAの結果からの知見やインプットなしに、マーシュフィールド集団で実行されたマーカー選択とモデルフィッティングによって、MESAサンプルの各対象のスコアを予測するために使用された。

【0216】

すべてのモデルの絶対リスクスコアの計算は、上記のアプローチに基づくものであった。危険因子とバイオマーカーに対する幾らかの欠損値のため、コホートの重みは各比較のそれぞれの状態 (status) と性別の組み合わせに変更された。女性と男性の症例またはコントロールの再分類は、同じ重みを運ばないので、再分類の計算でも、同じ修正済みの重みを占めた。これは、欠損値はランダムな欠落によると仮定して、全集団に対して結果を適切に拡張する試みで行われた。

10

【0217】

表25、表26は、レイノルズスコア (Reynolds score) に対する比較と同様に、以前に提示したNRIとCNRIの観点から、本開示のモデルと、3つの他のモデルとの比較を示している [Ridker PM, Buring JE, Rifai N, et al. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score JAMA 2007; 297: 611-619]。比較はマーシュフィールド・セットからの予測パフォーマンスと一致していた。本開示モデルは、ここに提示の他のいかなるトランスポートモデルの上にも、より良い臨床ネット再分類 (clinical net reclassification) を提供した。ベースライン生存関数を推定するためのスコアの平均を用いる方法は、個々の推定値を用いる方法に比べて、症例とコントロール間の再分類によりよいバランスを提供した。これは、MESAサンプル (図14Aおよび14B) 上でこれらのモデルの相対的リスク予測に、再度一致した。これらの結果は、明らかに、MESAセット内の、低中間 (low intermediate) / 中間のリスク集団の本開示モデルの臨床的有用性とトランスポートビリティをサポートしている。非糖尿病集団でのモデルの予測能力は、NRIとCNRIの観点から、表27に示されている。以降では、リスクの中間範囲は、参照モデルに基づいて、3.5から7.5%の間隔に設定されている。ベースラインで糖尿病と診断されたすべての被験者は、比較から除外されている。結果は、再び、非糖尿病患者のための中間リスクカテゴリーのモデルの臨床的有用性を示している。

20

30

【表25】

	Baseline Hazard Calculation	NRI	NRI pval	NRI Case	NRI Ctrl	CNRI	CNRI pval	CNRI Case	CNRI Ctrl
FRS	individual	1.906%	0.3425	-3.568%	-5.474%	31.931%	0.0000	2.076%	-29.855%
	average	2.706%	0.2895	7.130%	4.424%	30.254%	0.0000	12.311%	-17.943%
4-TRFs	individual	6.071%	0.0650	-0.611%	-6.682%	23.566%	0.0000	2.198%	-21.368%
	average	12.266%	0.0025	19.505%	7.238%	23.932%	0.0000	20.426%	-3.505%
9-TRFs	individual	-0.289%	0.5269	-3.324%	-3.035%	20.211%	0.0002	2.407%	-17.804%
	average	2.257%	0.3033	4.479%	2.222%	18.404%	0.0012	8.400%	-10.004%
Reynolds	individual	-5.045%	0.8436	-6.102%	-1.057%	26.697%	0.0001	9.231%	-17.466%
	average	-8.490%	0.9606	-15.562%	-7.072%	25.202%	0.0003	3.380%	-21.822%

40

FRS、4-TRF、9-TRFモデルおよびレイノルズスコアモデル (Reynolds score models) に対してアヴィール・スコア (Aviir score) を比較した、MESAデータセットに対するNRIとCNRIの結果。CNRIは、参照モデルの3.5-10%リスクのベースラインの範囲に基づいている。バイオマーカーデータが不足している被験者は、比較から除外されている。

【表 2 6】

	Baseline Hazard Calculation	NRI	NRI pval	NRI Case	NRI Ctrl	CNRI	CNRI pval	CNRI Case	CNRI Ctrl
FRS-individ	individual	0.247%	0.4805	-9.878%	-10.125%	46.363%	0.0000	12.836%	-33.527%
FRS-average	average	0.657%	0.4477	4.875%	4.218%	39.596%	0.0000	24.328%	-15.268%
TRF4-individ	individual	2.703%	0.2660	-7.622%	-10.325%	30.501%	0.0000	4.666%	-25.834%
TRF4-average	average	2.902%	0.2520	10.940%	8.038%	anal	0.0269	19.772%	4.296%
TRFext-individ	individual	-3.249%	0.7582	-9.115%	-5.866%	32.157%	0.0001	11.602%	-20.556%
TRFext-average	average	-1.072%	0.5895	2.162%	3.234%	27.144%	0.0017	23.674%	-3.470%
Reynold-individ	individual	-3.951%	0.7919	-3.172%	0.779%	33.933%	0.0008	19.294%	-14.639%
Reynold-average	average	-6.377%	0.9229	-11.151%	-4.774%	22.063%	0.0257	2.718%	-19.345%

FRS、4 - TRF、9 - TRFモデルおよびレイノルズスコアモデル (Reynolds score models) に対してアヴィール・スコア (Aviir score) を比較した、MESAデータセットに対するNRIとCNRIの結果。CNRIは、参照モデルの3.5 - 7.5%リスクのベースラインの範囲に基づいている。バイオマーカーデータが不足している被験者は、比較から除外されている。

【表 2 7】

	Range	Baseline Hazard Calculation	NRI	NRI p-val	NRI_case	NRI_ctrl	CNRI	CNRI p-val	CNRI_case	CNRI_ctrl
FRS	3.5-7.5%	Individual	0.42%	0.472	-1.23%	-1.65%	38.42%	0.000	13.94%	-24.47%
		Average	4.64%	0.211	9.84%	5.21%	42.31%	0.000	23.28%	-19.02%
4-TRFs	3.5-7.5%	Individual	2.31%	0.324	-1.20%	-3.51%	23.48%	0.006	5.06%	-18.42%
		Average	9.44%	0.034	20.11%	10.67%	29.63%	0.001	34.91%	5.28%
9-TRFs	3.5-7.5%	Individual	3.69%	0.256	3.24%	-0.45%	30.17%	0.001	17.81%	-12.36%
		Average	6.78%	0.111	12.03%	5.25%	28.88%	0.003	26.59%	-2.29%

FRS、4 - TRF、および9 - TRFモデルに対してアヴィール・スコアを比較した、MESAデータセットに対するNRIとCNRIの結果。CNRIは、参照モデルの3.5 - 7.5%リスクのベースラインの範囲に基づいている。バイオマーカーデータが不足している被験者は、比較から除外されている。

【0218】

(実施例10)

ハイブリッド・バイオマーカー予後/診断モデル

タンパク質バイオマーカー/TRFに加えて、miRNAは、血液などの人間の体液で測定でき、ある被験者の将来の心血管イベントを予測するために使用することができる。

【0219】

表28で提示されているmiRNAセットから選択された共変量を有するハイブリッド予後モデルと、症例・コホート研究デザインを使用して単一スコアとして開示されたタンパク質・バイオマーカーモデル(実施例7-9を参照)とを、構築することによって、ハイブリッドのmiRNA/タンパク質バイオマーカーセットの予後パワーが決定される。コホート(cohort)は、対象とする時間フレーム内にM1に進展した全ての症例(N=200)と200のコントロールを含んでいる。より小さなコホートを効率的に利用するために、TRFsとタンパク質予測因子は、単一の計算されたスコア(単一変数)の観点から処理される、miRNAバイオマーカーの単変量関連づけ(univariate association)が、タンパク質バイオマーカーまたはTRFsについて観察されたそれよりも強力でない限り。後者のケースでは、多変量モデルが、使用可能なすべてのマーカー(TRFs、タンパク質バイオマーカー、miRNAs)から変数を選択する、ペナルティ付き回帰法の使用に基づいて構築される。前者の場合には、スコアの計算は、前述のように、より大きなコホートで事前に推定された係数を使用して行われる。交差検定(Cross-validation)とペナルティ付き回帰テクニックは、3タイプのモデルのモデルサイズとmiRNAマーカーを選択するために使用されている。a) miRNAのみ(miRNA-only)のモデル、b) TRF+miRNAベースモデル、および、c) TRF+タンパク質+miRNAバイオマーカーベース

モデル。適合されたモデルの期待されるパフォーマンスは、前述した T R F + タンパク質ベースモデル（実施例 8 - 9 を参照）と同様に、ハイブリッドモデルの時間依存性 A U C、N R I、および C N R I 特性、対、F R S に基づいて、評価される。

【表 2 8】

miRNAs	
miR-378	miR-19b
miR-497	miR-151-5p
miR-21	miR-215
miR-15b	miR-25
miR-99a	let-7f
miR-652	miR-10b
miR-30b	miR-423-3p
miR-26a	miR-502-3p
miR-29a	miR-140.3p
miR-1974	miR-92a
miR-30c	miR-660
miR-122	miR-142-3p
miR-29c	miR-130a
miR-192	miR-185
miR-34a	let-7c
miR-24	miR-18a
miR-221	miR-365
miR-126	miR-26b
miR-331-3p	miR-125b
let-7a	miR-297
miR-148a	miR-146a
let-7g	miR-99b
miR-19a	miR-424
miR-142-5p	miR-93
miR-22	let-7b

10

20

30

40

50

【0 2 2 0】

特に断りのない限り、明細書および特許請求の範囲で使用されている、分子量、反応条件の様に、成分、特性などの量を表現するすべての数字は、全ての例で、用語「約」で変更されるものとして理解されるべきである。したがって、これに反する指示がない限り、明細書及び添付の特許請求の範囲に記載の数値パラメータは、本開示によって得られるように求められた所望の特性に依存して変化しうる近似である。少なくとも、特許請求の範囲に均等論の適用を制限する試みとしてではなく、各数値パラメータは少なくとも、報告された有効桁数の光の中で、普通の丸め技術を適用することによって解釈されるべきである。開示の広い範囲を記載した数値範囲および数値パラメータは近似値であるにもかかわらず、特定の実施例に記載の数値は可能な限り正確に報告されている。しかし、任意の数値は、本質的には、それぞれの試験測定に見られる標準偏差に必然的に起因する特定のエラーが含まれている。

【0 2 2 1】

用語「a」、「an」、「the」、および本発明を説明する文脈（特に以下の特許請求の範囲の文脈）で使用される同様の指示は、特に文脈と明らかに矛盾したりまたはここに示さない限り、単数および複数の両方をカバーするために解釈されるべきである。本明細書での数値範囲の列挙は、単に、各個別の値がその範囲内にあることを個別に参照する簡単な方法として機能することを意図しているに過ぎない。そうでなければここに示されない限り、各個々の値は、それが本明細書に個々に記載されたかのように、本明細書に組み込まれている。そうでなければ本明細書に示された、あるいは明らかに文脈と矛盾しない限り、本明細書に記載のすべてのメソッドは、任意の適切な順序で実行することができる。本明細書で提供される任意のおよび全ての例、または典型的な言語（例えば、「など」）の使用は、本発明をより明らかにするためだけに意図され、それ以外の場合は、請求項に係る発明の範囲に制限をもたらすことはない。本明細書の言語は、本発明の実施に不可欠な任意のクレームされていない要素（non-claimed element）を示すものとして解釈されるべきではない。

【0 2 2 2】

本発明の別の要素または実施形態のグループ分けは、それに制限されて解釈されるべき

ではない。各グループのメンバーは、参照され、個々にクレームされ、グループの他のメンバーと組み合わせられ、または、本明細書にある他の要素と組み合わせられてクレームされることができる。利便性、および/または特許性の理由で、グループの1つまたは複数のメンバーが含まれるか、グループから削除されることは予想される。そのような包含または削除が発生した場合、明細書は、添付の特許請求の範囲で使用されているすべてのマーカッシュ群の記述要件を満たすよう修正されたグループが含まれているとみなされる。

【0223】

本発明の特定の実施形態は、本発明を実施するよう発明者に知られているベストモードを含んで、本明細書に記載されている。もちろん、これらの記載した実施形態のバリエーションは、上記の説明を読めば当業者に明らかになるであろう。本発明者は、当業者がそのようなバリエーションを適切に採用することを期待し、そして、さもなければ、本明細書に具体的に記載したこと以上に実施され则认为している。したがって、本発明は、適用される法律によって許可されているように、ここに添付した特許請求の範囲に記載されている主題のすべての改変および均等物を含んでいる。さらに、そうでなければここに示された、あるいは明らかに文脈と矛盾しない限り、すべての可能なバリエーションで、上記の要素の任意の組み合わせは、本発明に包含される。

10

【0224】

本明細書に開示される特定の実施形態は、言語からなる構成を使用する特許請求の範囲、または、本質的に言語からなる特許請求の範囲において、さらに制限を受けている。当初のクレーム又は補正の際に追加されるクレームで使用する場合、「～から成る (consisting of)」との遷移用語は、特許請求の範囲に指定されていない任意の要素、ステップ、または成分を除外する。「本質的に、～から成る (consisting essentially of)」との遷移用語は、指定された材料、ステップ、および基本的かつ新規な特徴に実質的に影響を与えないものに、特許請求の範囲が制限される。その様にクレームされた本発明の実施の形態は、本明細書に、本質的または明示的に記述され、使用可能である。

20

【0225】

さらに、本明細書を通して、多数の特許や刊行物が参照されている。上記の引用文献および刊行物の各々は、個別に、その全体が参照により本明細書に援用されている。

【0226】

30

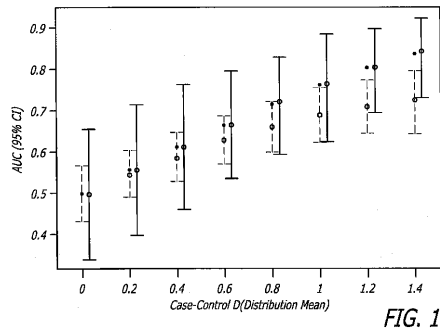
最後に、本明細書に開示の本発明の実施形態は、本発明の原理を例示するものであることは理解されるべきである。採用可能な他の修正は、本発明の範囲内である。この様に、例示の方法で、しかしそれに限定されることなく、本発明の他の構成は、本明細書の教示に従って利用することができる。したがって、本発明は、正確に示され記載されたものに、限定されない。

【0227】

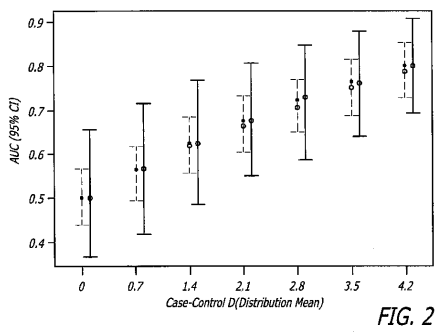
本明細書に開示される特定の実施形態は、言語からなる構成を使用する特許請求の範囲、または、本質的に言語からなる特許請求の範囲において、さらに制限を受けている。当初のクレーム又は補正の際に追加されるクレームで使用する場合、「～から成る (consisting of)」との遷移用語は、特許請求の範囲に指定されていない任意の要素、ステップ、または成分を除外する。「本質的に、～から成る (consisting essentially of)」との遷移用語は、指定された材料、ステップ、および基本的かつ新規な特徴に実質的に影響を与えないものに、特許請求の範囲が制限される。その様にクレームされた本発明の実施の形態は、本明細書に、本質的または明示的に記述され、使用可能である。

40

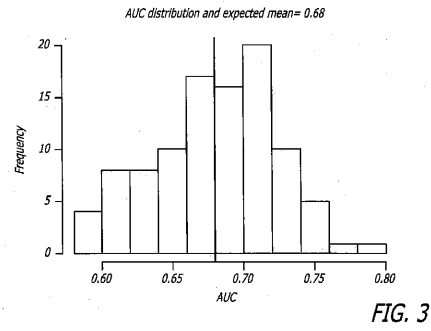
【 図 1 】



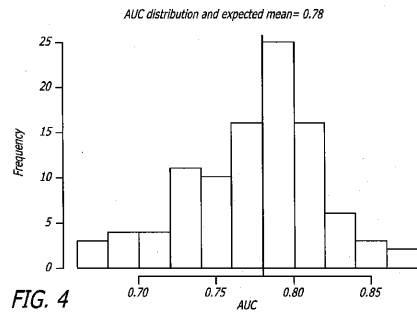
【 図 2 】



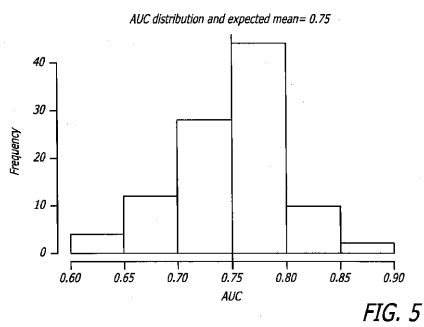
【 図 3 】



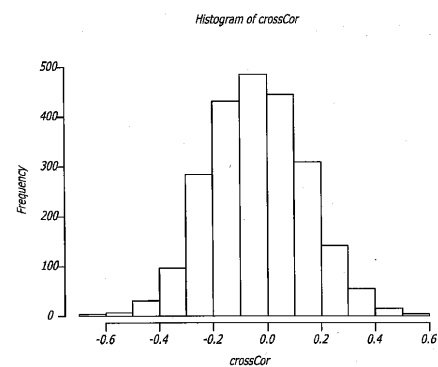
【 図 4 】



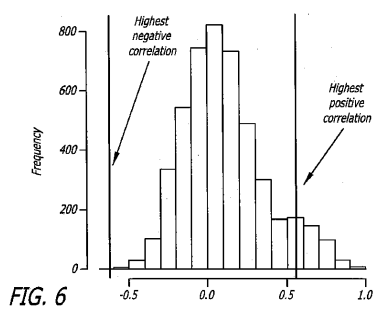
【 図 5 】



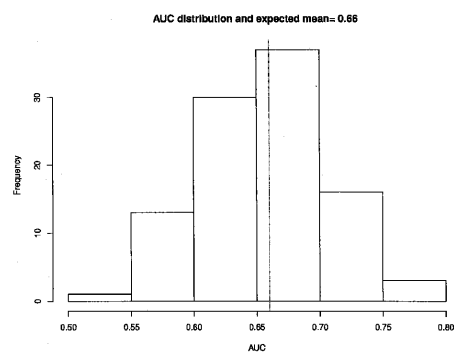
【 図 7 】



【 図 6 】

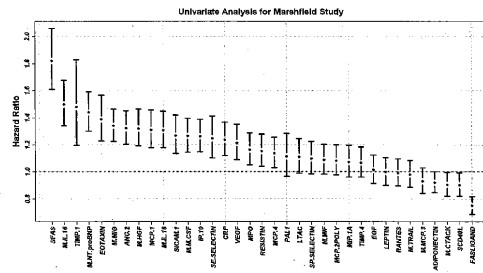


【 図 8 】



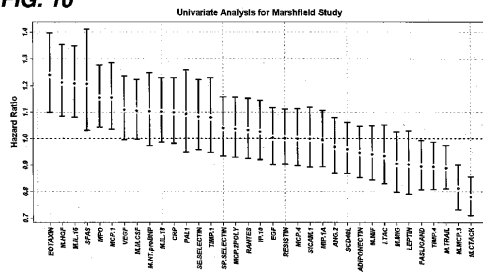
【 図 9 】

FIG. 9



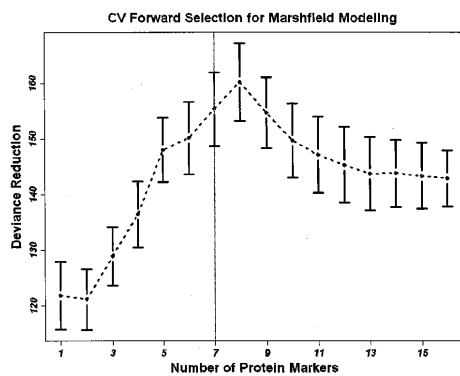
【 図 10 】

FIG. 10



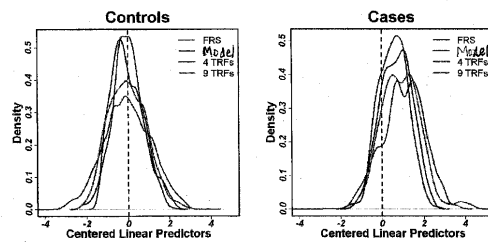
【 図 12 】

FIG. 12



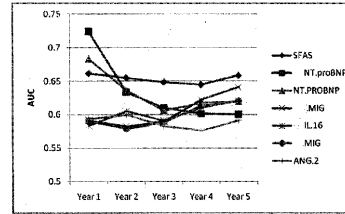
【 図 13 A - B 】

FIGS. 13 A & B



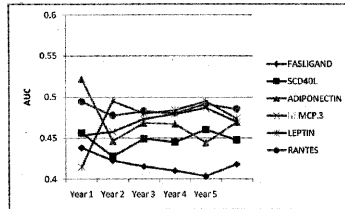
【 図 11 A 】

FIG. 11 A



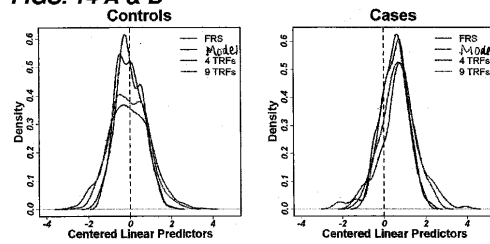
【 図 11 B 】

FIG. 11 B



【 図 14 A - B 】

FIGS. 14 A & B



【配列表】

2013513387000001.app

【手続補正書】

【提出日】平成24年9月28日(2012.9.28)

【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

治療計画の必要性または有効性および／または心血管イベントのリスクを決定するためにヒトの心血管の健康を評価する方法であって、

ヒトから生物学的サンプルを取得すること；

少なくとも1つのmiRNAがmiR10bまたはmiR126である、少なくとも2つのmiRNAマーカーのレベルを決定すること；

各miRNAマーカーのレベルを含むデータセットを取得すること；

前記生物学的サンプルを分類するためにデータを使用する分析・分類プロセスに、前記データを入力することであって、該分類は、アテローム性動脈硬化症分類、健康分類、薬剤曝露分類、または、薬剤非曝露分類であること；および、

前記分類に基づいて、治療計画の必要性または有効性および／または心血管イベントのリスクを決定するためにヒトの心血管の健康を評価すること、を含む方法。

【請求項2】

前記第2のマーカーは、miR10b、miR126、miR30a、miR92a、miR222、miR378、let-7b、miR-497、miR-21、miR-15b、miR-99a、miR-29a、miR-24、miR-30b、miR-29c、miR-331.3p、miR-19a、miR-22、miR-502.3またはmiR-652である、請求項1に記載の方法。

【請求項3】

生物学的サンプル中の少なくとも1つのタンパク質マーカーのレベルを決定すること；

少なくとも1つのタンパク質マーカーのレベルを含むデータセットを取得すること；

前記生物学的サンプルを分類するためにデータを使用する分析・分類プロセスに、前記データを入力することであって、該分類は、アテローム性動脈硬化症分類、健康分類、薬剤曝露分類、または、薬剤非曝露分類であること；および、

前記分類に基づいて、治療計画の必要性または有効性を決定するためにヒトの心血管の健康および／または心疾患イベントのリスクを評価すること、をさらに含む、請求項1に記載の方法。

【請求項4】

少なくとも1つのタンパク質マーカーが、エオタキシン、HGF、NTproBNP、VEGF、CRP、FASリガンド、sFAS、MPO、PAI-1、MCP-2、EGF、MCP-1、IL-16、ANG-2、MCP-3、IL-18、TIMP-4、Resistan、CTACK、MIF、TRAIL、アディポネクチン、IP-10またはレプチンである、請求項3に記載の方法。

【請求項5】

少なくとも1つのタンパク質マーカーが、IL-16、エオタキシン、FASリガンド、CTACK、MCP-3、HGF、またはsFASである、請求項4に記載の方法。

【請求項6】

少なくとも1つのタンパク質マーカーが、IL-18、CRP、VEGF、EGF、MPO、MCP-2、PAI-1、またはNTproBNPである、請求項4に記載の方法。

【請求項 7】

3 つ以上のタンパク質マーカーのレベルが決定される、請求項 3 に記載の方法。

【請求項 8】

3 つ以上のタンパク質マーカーが、エオタキシン、HGF、NTproBNP、VEGF、CRP、FASリガンド、sFAS、MPO、PAI-1、MCP-2、EGF、MCP-1、IL-16、ANG-2、MCP-3、IL-18、TIMP-4、Resistan、CTACK、MIF、TRAIL、アディポネクチン、IP-10 およびレプチンから選択される、請求項 7 に記載の方法。

【請求項 9】

少なくとも 3 つのタンパク質マーカーが、IL-16、エオタキシン、FASリガンド、CTACK、MCP-3、HGF、または sFAS から選択される、請求項 8 に記載の方法。

【請求項 10】

データセットが、1 つ以上の臨床的指標のデータをさらに含む、請求項 1 に記載の方法

。

【請求項 11】

前記臨床的指標が、年齢、性別、コレステロール値、LDL 濃度、HDL 濃度、トリグリセリド濃度、血圧、体格指数 (BMI)、CRP 濃度、冠動脈カルシウムスコア、腹囲、喫煙状況、心血管疾患の既往歴、心血管疾患の家族歴、心拍数、空腹時インスリン濃度、空腹時グルコース濃度、糖尿病の状態、高血圧薬の使用、またはそれらの組み合わせである、請求項 10 に記載の方法。

【請求項 12】

前記臨床的指標が、年齢、糖尿病の状態、喫煙状況、収縮期血圧、性別、心筋梗塞の家族歴、コレステロール値、HDL 濃度、及び、高血圧薬の使用である、請求項 11 に記載の方法。

【請求項 13】

前記生物学的サンプルは、血液、血清、血漿、唾液、尿、汗、母乳またはそれらの組合せを含む、請求項 1 に記載の方法。

【請求項 14】

前記分析・分類プロセスは、予測モデルの使用を含む、請求項 1 に記載の方法。

【請求項 15】

前記分析・分類プロセスは、取得したデータセットを参照データセットと比較することを含む、請求項 1 に記載の方法。

【請求項 16】

前記予測モデルは、分類のために少なくとも 0.68 の少なくとも 1 つの品質メトリックを備える、請求項 14 に記載の方法。

【請求項 17】

前記品質メトリックは、AUC 及び精度から選択される、請求項 16 に記載の方法。

【請求項 18】

前記分析・分類プロセスが、線形判別分析モデル、サポートベクトルマシン分類 (support vector machine classification) アルゴリズム、帰納的特徴除去 (recursive feature elimination) モデル、マイクロアレイモデルの予測分析、ロジスティック回帰モデル、CART アルゴリズム、フレックス・ツリー・アルゴリズム、LART アルゴリズム、ランダムフォレストアルゴリズム、MART アルゴリズム、機械学習アルゴリズムは、ペナルティ回帰法 (penalized regression method) またはその組み合わせである、請求項 1 に記載の方法。

【請求項 19】

前記分析・分類プロセスが、少なくとも 0.68 の品質メトリックを提供するように選択されるターム (term) を備える、請求項 18 に記載の方法。

【請求項 20】

前記分析・分類プロセスが、少なくとも0.70の品質メトリックを提供するように選択されるタームを備える、請求項18に記載の方法。

【請求項21】

前記分析・分類プロセスが、分類のために、少なくとも0.70の少なくとも1つの品質メトリックを備える、請求項18に記載の方法。

【請求項22】

前記アテローム性動脈硬化症分類は、冠動脈疾患、心筋梗塞または不安定狭心症である、請求項1に記載の方法。

【請求項23】

アテローム性動脈硬化症の診断、アテローム性動脈硬化症のステージング、アテローム性動脈硬化症の予後、血管炎症レベル、アテローム性動脈硬化症の進行の程度を決定するために分類を使用すること、治療反応をモニタリングすること、冠動脈カルシウムスコアを予測すること、アテローム性動脈硬化疾患の不安定症状から安定症状を区別すること、またはそれらの任意の組み合わせ、をさらに含む、請求項1に記載の方法。

【請求項24】

前記療計画は、1以上の更なる試験、3ヶ月、6ヶ月および/または12ヶ月のフォローアップ検診、遮断薬、カルシウムチャネル遮断薬、アスピリンおよび/またはコレステロール低下剤による薬理的介入、非治療またはそれらの組み合わせを含む、請求項1に記載の方法。

【請求項25】

心血管イベントの前記リスクが、サンプルを取得した日付から、約1年、約2年、約3年、約4年または約5年で決定される、請求項1に記載の方法。

【請求項26】

治療計画の必要性または有効性および/または心血管イベントのリスクを決定するためにヒトの心血管の健康を評価する方法であって、

ヒトから生物学的サンプルを取得すること；

ノーザン分析、ポリメラーゼ連鎖反応（PCR）、逆転写酵素（RT）-PCR法、蛍光共鳴エネルギー移動（FRET）検出、分子ビーコン検出、オリゴヌクレオチドアレイ、cDNAアレイ、ポリヌクレオチドアレイ、液体マイクロアレイおよび/またはマイクロエレクトリックアレイへのハイブリダイゼーション、cDNAシーケンシング、クローンハイブリダイゼーション、cDNA断片フィンガープリンティング、遺伝子発現連続分析（SAGE）、サブトラクティブハイブリダイゼーション、ディファレンシャルディスプレイおよび/またはディファレンシャルスクリーニングのうちの1以上を使用して、少なくとも1つのmiRNAがmiR10bまたはmiR126である、少なくとも2つのmiRNAマーカーのレベルを決定すること；

各miRNAマーカーのレベルを含むデータセットを取得すること；

ELISAサンドイッチアッセイ、フローサイトメトリー、質量分析検出、比色アッセイ、タンパク質アレイへの結合、蛍光活性化細胞選別（FACS）の一つ以上を使用して、前記生物学的サンプル中の少なくとも1つのタンパク質マーカーのレベルを決定することであって、少なくとも1つのタンパク質マーカーが、エオタキシン、HGF、NTproBNP、VEGF、CRP、FASリガンド、sFAS、MPO、PAI-1、MCP-2、EGF、MCP-1、IL-16、ANG-2、MCP-3、IL-18、TIMP-4、Resistin、CTACK、MIF、TRAIL、アディポネクチン、IP-10またはレプチンであること；

タンパク質マーカーのレベルを含むデータセットを得ること；

年齢、糖尿病の状態、タバコ喫煙状況、収縮期血圧、性別、心筋梗塞の家族歴、コレステロール値、HDL濃度、高血圧薬の使用のうち少なくとも2つを含む臨床的指標情報を決定すること；

前記臨床的指標情報を含むデータセットを取得すること；

前記各データセットを、分類のために少なくとも0.68の少なくとも1つの品質メ

リックを含む予測モデルおよび／または前記生物学的サンプルもしくは臨床的指標情報を分類するための参照データセットを用いて前記生物学的サンプルを分類するために、前記各データセットを使用する分析・分類プロセスに、入力することであって、前記分類が、アテローム性動脈硬化症分類、健康分類、薬物曝露分類または薬物非曝露分類であること

；
および、

治療計画の必要性または有効性および／または心血管イベントのリスクを決定するために、前記分類に基づいてヒトの心血管の健康を評価すること、を含む方法。

【請求項 27】

前記第2のmiRNAマーカーは、miR10b、miR126、miR30a、miR92a、miR222、miR378、let-7b、miR-497、miR-21、miR-15b、miR-99a、miR-29a、miR-24、miR-30b、miR-29c、miR-331.3p、miR-19a、miR-22、miR-502.3またはmiR-652である、請求項26に記載の方法。

【請求項 28】

少なくとも1つのタンパク質マーカーが、IL-16、エオタキシン、FASリガンド、CTACK、MCP-3、HGF、またはsFASである、請求項26に記載の方法。

【請求項 29】

少なくとも1つのタンパク質マーカーが、IL-18、CRP、VEGF、EGF、MPO、MCP-2、PAI-1、またはNTproBNPである、請求項26に記載の方法。

【請求項 30】

治療計画の必要性または有効性および／または心血管イベントのリスクを決定するためにヒトの心血管の健康を評価する方法であって、

ヒトから生物学的サンプルを取得すること；

ノーザン分析、ポリメラーゼ連鎖反応（PCR）、逆転写酵素（RT）-PCR法、蛍光共鳴エネルギー移動（FRET）検出、分子ビーコン検出、オリゴヌクレオチドアレイ、cDNAアレイ、ポリヌクレオチドアレイ、液体マイクロアレイおよび／またはマイクロエレクトリックアレイへのハイブリダイゼーション、cDNAシーケンシング、クローンハイブリダイゼーション、cDNA断片フィンガープリンティング、遺伝子発現連続分析（SAGE）、サブトラクティブハイブリダイゼーション、ディファレンシャルディスプレイおよび／またはディファレンシャルスクリーニングのうちの1以上を使用して、miR10bおよびmiR126のレベルを決定すること；

各miRNAマーカーのレベルを含むデータセットを取得すること；

ELISAサンドイッチアッセイ、フローサイトメトリー、質量分析検出、比色アッセイ、タンパク質アレイへの結合、蛍光活性化細胞選別（FACS）の一つ以上を使用して、前記生物学的サンプル中の少なくとも1つのタンパク質マーカーのレベルを決定することであって、少なくとも1つのタンパク質マーカーが、エオタキシン、HGF、NTproBNP、VEGF、CRP、FASリガンド、sFAS、MPO、PAI-1、MCP-2、EGF、MCP-1、IL-16、ANG-2、MCP-3、IL-18、TIMP-4、Resistan、CTACK、MIF、TRAIL、アディポネクチン、IP-10またはレプチンであること；

タンパク質マーカーのレベルを含むデータセットを得ること；

年齢、糖尿病の状態、タバコ喫煙状況、収縮期血圧、性別、心筋梗塞の家族歴、コレステロール値、HDL濃度、従前の高血圧薬使用のうち少なくとも1つを含む臨床的指標情報を決定すること；

前記臨床指標の情報を含むデータセットを取得すること；

前記各データセットを、分類のために少なくとも0.68の少なくとも1つの品質メトリックを含む予測モデルおよび／または参照データセットを用いて、前記生物学的サンプルと臨床的指標情報を分類するために前記各データセットを使用する分析・分類プロセス

に、入力することであって、前記分類が、アテローム性動脈硬化症分類、健康分類、薬物曝露分類または薬物非曝露分類であること；

および、

治療計画の必要性または有効性および／またはヒトの心血管イベントのリスクを決定するために、前記分類に基づいてヒトの心血管の健康を評価すること、を含む方法。

【請求項 3 1】

治療計画の必要性または有効性および／または心血管イベントのリスクを決定するためにヒトの心血管の健康を評価するためのキットであって、

ヒトから得た生物学的サンプル中の、少なくとも 1 つの *miRNA* は *miR10b* または *miR126* である少なくとも 2 つの *miRNA* マーカーのレベルを決定するためのアッセイと、

miRNA マーカーの個々のレベルを含むデータセットを取得するための、および、生物学的サンプルを分類するためにデータを使用する、分析・分類プロセスにデータを入力するための、使用説明書と、を備え、

前記分類は、アテローム性動脈硬化症分類、健康分類、薬剤曝露分類、薬剤非曝露分類から選択され、治療計画の必要性または有効性および／またはヒトの心血管イベントのリスクを決定するためにヒトの心血管の健康を評価するために使用されることを特徴とするキット。

【請求項 3 2】

前記生物学的サンプル中の、エオタキシン、*HGF*、*NTproBNP*、*VEGF*、*CRP*、*FAS* リガンド、*sFAS*、*MPO*、*PAI-1*、*MCP-2*、*EGF*、*MCP-1*、*IL-16*、*ANG-2*、*MCP-3*、*IL-18*、*TIMP-4*、*Resistan*、*CTACK*、*MIF*、*TRAIL*、アディポネクチン、*IP-10* または *レプチン* から選択される少なくとも 1 つのタンパク質マーカーのレベルを決定するためのアッセイと

；
少なくとも 1 つのタンパク質マーカーの個々のレベルを含むデータセットを得て、*miRNA* とタンパク質マーカーのデータを、前記生物学的サンプルを分類するためにデータを使用する分析・分類プロセスへ入力するための、前記使用説明書と、を更に備え、

前記分類は、アテローム性動脈硬化症分類、健康分類、薬剤曝露分類、薬剤非曝露分類から選択され、治療計画の必要性または有効性および／または心血管イベントのリスクを決定するためにヒトの心血管の健康を評価するために使用されることを特徴とする、請求項 3 1 に記載のキット。

【請求項 3 3】

追加的 *miRNA* マーカーは、*miR10b*、*miR126*、*miR30a*、*miR92a*、*miR222*、*miR378*、*let-7b*、*miR-497*、*miR-21*、*miR-15b*、*miR-99a*、*miR-29a*、*miR24*、*miR-30b*、*miR-29c*、*miR-331.3p*、*miR-19a*、*miR-22*、*miR-502.3* または *miR-652* から、選択される、請求項 3 2 に記載のキット。

【請求項 3 4】

エオタキシン、*HGF*、*NTproBNP*、*VEGF*、*CRP*、*FAS* リガンド、*sFAS*、*MPO*、*PAI-1*、*MCP-2*、*EGF*、*MCP-1*、*IL-16*、*ANG-2*、*MCP-3*、*IL-18*、*TIMP-4*、*Resistan*、*CTACK*、*MIF*、*TRAIL*、アディポネクチン、*IP-10* または *レプチン* から選択される少なくとも 3 つのタンパク質マーカーのレベルを決定するためのアッセイを含む、請求項 3 2 に記載のキット。

【請求項 3 5】

予測モデルを含む分析・分類プロセスで使用するための、請求項 3 2 に記載のキット。

【請求項 3 6】

取得したデータセットと参照データセットとの比較を含む分析・分類プロセスで使用するための、請求項 3 2 に記載のキット。

【請求項 37】

前記予測モデルは、分類のために少なくとも 0.68 の少なくとも 1 つの品質メトリックを含む、請求項 35 に記載のキット。

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2010/059781

A. CLASSIFICATION OF SUBJECT MATTER

INV. C12Q1/68 G06F19/18 G01N33/68
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12Q G06F G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS, Sequence Search, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	HAMER MARK ET AL: "Utility of C-reactive protein for cardiovascular risk stratification across three age groups in subjects without existing cardiovascular diseases." THE AMERICAN JOURNAL OF CARDIOLOGY 15 AUG 2009 LNKD- PUBMED:19660608, vol. 104, no. 4, 15 August 2009 (2009-08-15), pages 538-542, XP002624250, ISSN: 1879-1913 the whole document ----- -/--	1-22, 26-35

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document 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

26 May 2011

Date of mailing of the international search report

07/06/2011

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Gabriels, Jan

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2010/059781

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	SALOMAA VEIKKO ET AL: "Prospective Assessment of 30 Novel Biomarkers in the Prediction of Relative and Absolute Risk of Cardiovascular Events: The MORGAM Biomarker Study", CIRCULATION, vol. 119, no. 10, March 2009 (2009-03), page E342, XP002624251, & JOINT NUTRITION, PHYSICAL ACTIVITY AND METABOLISM CONFERENCE/49TH CARDIOVASCULAR DISEASE EPIDEMIOLOG; PALM HARBOR, FL, USA; MARCH 10 14, 2009 ISSN: 0009-7322 the whole document	1-22, 26-35
Y	THUM THOMAS ET AL: "MicroRNA-21 contributes to myocardial disease by stimulating MAP kinase signalling in fibroblasts", NATURE (LONDON), vol. 456, no. 7224, December 2008 (2008-12), page 980, XP002624252, ISSN: 0028-0836 the whole document	1-22, 26-35
Y	MISHRA PARAS KUMAR ET AL: "MicroRNAs as a therapeutic target for cardiovascular diseases", JOURNAL OF CELLULAR AND MOLECULAR MEDICINE, vol. 13, no. 4, April 2009 (2009-04), pages 778-789, XP002624253, ISSN: 1582-1838 the whole document	1-22, 26-35
Y	WO 2008/043521 A2 (UNIV WUERZBURG J MAXIMILIANS [DE]; THUM THOMAS [DE]; BAUERSACHS JOHANN) 17 April 2008 (2008-04-17) miR-378, miR-497, miR-21, miR-15b, miR-99a, miR-29a, miR-24, miR-30b, miR-29c, miR-331, miR-19a, miR-22, miR-126, let-7b, miR-502	1-22, 26-35
Y	WO 2009/012468 A2 (UNIV COLORADO [US]; PORT JONATHAN DAVID [US]; SUCHAROV CARMEN [US]; BR) 22 January 2009 (2009-01-22) miR-378,-22,let-7b	1-22, 26-35
Y	US 2007/042380 A1 (BENTWICH ITZHAK [IL] ET AL BENTWICH ITZHAK [IL] ET AL) 22 February 2007 (2007-02-22) Table 14 row 143table 14; sequences 2,5,6	1-22, 26-35
	----- -/--	

5

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

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2010/059781

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2009/114681 A2 (DHARMA CON INC [US]; SETO ANITA G [US]; BASKERVILLE SCOTT [US]; LEINWAN) 17 September 2009 (2009-09-17) hsa-miR-378, hsa-miR-99a, hsa-miR-497, hsa-miR-24, miR-29c claim 3 -----	1-22, 26-35
X	WO 2007/002677 A2 (UNIV LELAND STANFORD JUNIOR [US]; TABIBIAZAR RAYMOND [US]; TSAO PHILIP) 4 January 2007 (2007-01-04) claim 12 -----	23-25, 36,37
X	WO 2008/080126 A2 (AVIIR INC [US]; TABIBIAZAR RAYMOND [US]; HYTOPOULOS EVANGELOS [US]) 3 July 2008 (2008-07-03) claims 1,12 -----	23-25, 36,37
A	US 7 306 562 B1 (BAYKAL DEMIR [US]) 11 December 2007 (2007-12-11) the whole document -----	23-25

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2010/059781

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.:
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:

see additional 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 an additional fee, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

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

Remark on Protest

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

International Application No. PCT/ US2010/ 059781

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-22, 26-35

A method for assessing the cardiovascular health or risk of a cardiovascular event of a human by determining levels of at least 2 miRNA markers selected from miRNAs listed in Table 20 and inputting the data of the levels of each miRNA marker into an analytical classification process that uses the data to classify the biological sample, wherein the classification is selected from the group consisting of an atherosclerotic cardiovascular disease classification, a healthy classification, a medication exposure classification, a no medication exposure classification; and determining a treatment regimen for the human based on the classification or risk of a cardiovascular event.

2. claims: 23-25, 36, 37

A method for assessing the cardiovascular health or risk of a cardiovascular event of a human by determining levels of at least 3 protein markers selected from the group consisting of IL-16, sFas, Fas ligand, MCP-3, HGF, CTACK, EOTAXIN, adiponectin, IL-18, TIMP.4, TIMP.1, CRP, VEGF, and EGF and inputting the data of the levels of each protein marker into an analytical classification process that uses the data to classify the biological sample, wherein the classification is selected from the group consisting of an atherosclerotic cardiovascular disease classification, a healthy classification, a medication exposure classification, a no medication exposure classification; and determining a treatment regimen for the human based on the classification or risk of a cardiovascular event.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2010/059781

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2008043521 A2	17-04-2008	AU 2007306594 A1 EP 2094869 A2 EP 2208798 A1 JP 2010505427 T US 2010010073 A1	17-04-2008 02-09-2009 21-07-2010 25-02-2010 14-01-2010
WO 2009012468 A2	22-01-2009	AU 2008275877 A1 CN 101802227 A EP 2179060 A2 JP 2010533503 T KR 20100049079 A US 2010267804 A1	22-01-2009 11-08-2010 28-04-2010 28-10-2010 11-05-2010 21-10-2010
US 2007042380 A1	22-02-2007	NONE	
WO 2009114681 A2	17-09-2009	NONE	
WO 2007002677 A2	04-01-2007	AU 2006261779 A1 CA 2613584 A1 EP 1913388 A2 JP 2009501318 T US 2007099239 A1	04-01-2007 04-01-2007 23-04-2008 15-01-2009 03-05-2007
WO 2008080126 A2	03-07-2008	CA 2710286 A1 US 2008300797 A1	03-07-2008 04-12-2008
US 7306562 B1	11-12-2007	NONE	

フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, 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, HU, 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, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW

(特許庁注：以下のものは登録商標)

１．フロッピー

(74)代理人 100147924

弁理士 美恵 英樹

(72)発明者 ハリントン、ダグ

アメリカ合衆国 9 2 6 7 2 カリフォルニア州 サンクレメンテ カミノイビザ 9 0 6

(72)発明者 ハイトボウロス、エヴァンゲロス

アメリカ合衆国 9 4 4 0 3 カリフォルニア州 サンマテオ ヒルサイドコート 3 7 2 4

(72)発明者 フェルプス、ブルース

アメリカ合衆国 9 4 5 1 7 カリフォルニア州 クレイトン ラエンシナルコート 1 1 8

F ターム(参考) 2G045 AA25 CA25 CA26 CB03 CB07 CB12 DA14 DA36 DA54 DA64

DA70 DA77 JA01

4B024 AA11 CA01 CA11 HA14

4B063 QA19 QQ03 QQ52 QQ79 QR35 QR48 QR55 QR72 QS33 QS34

QX01