

(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2019-531741  
(P2019-531741A)

(43) 公表日 令和1年11月7日(2019.11.7)

(51) Int.Cl.	F I	テーマコード (参考)
<b>C12Q 1/6886 (2018.01)</b>	C12Q 1/6886 ZNAZ	2G045
<b>GO1N 33/50 (2006.01)</b>	GO1N 33/50 P	4B029
<b>GO1N 33/53 (2006.01)</b>	GO1N 33/53 M	4B063
<b>C12M 1/00 (2006.01)</b>	C12M 1/00 A	
<b>C12Q 1/686 (2018.01)</b>	C12Q 1/686 Z	
審査請求 有 予備審査請求 未請求 (全 68 頁) 最終頁に続く		

(21) 出願番号 特願2019-520030 (P2019-520030)  
 (86) (22) 出願日 平成30年4月24日 (2018. 4. 24)  
 (85) 翻訳文提出日 平成31年4月12日 (2019. 4. 12)  
 (86) 国際出願番号 PCT/KR2018/004732  
 (87) 国際公開番号 WO2018/199589  
 (87) 国際公開日 平成30年11月1日 (2018. 11. 1)  
 (31) 優先権主張番号 10-2017-0052365  
 (32) 優先日 平成29年4月24日 (2017. 4. 24)  
 (33) 優先権主張国・地域又は機関 韓国 (KR)

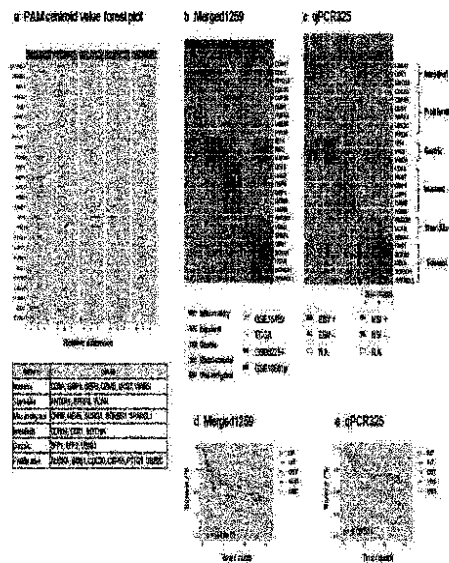
(71) 出願人 517227297  
 ノヴォミクス・カンパニー・リミテッド  
 大韓民国・ソウル・ヨンドンポーク・タン  
 サンーロ・171・クムガン・ペンテリウム・  
 アイティー・タワー・タンサンードン  
 ・4ーガ・306-1、306、305、  
 304  
 (74) 代理人 100108453  
 弁理士 村山 靖彦  
 (74) 代理人 100110364  
 弁理士 実広 信哉  
 (74) 代理人 100133400  
 弁理士 阿部 達彦

最終頁に続く

(54) 【発明の名称】 胃癌の生物学的特性に基づく群区分および予後予測システム

(57) 【要約】

本発明は、胃癌の生物学的特性に基づく群区分および予後予測システムに関し、進行性胃癌の予後標的遺伝子群のmRNA発現水準の定量的検査結果を利用して生存率の側面から予後を予測できるアルゴリズムを開発して、胃癌患者の治療方法を決定するのに補助的情報として活用することができる。



## 【特許請求の範囲】

## 【請求項 1】

TFF1、TFF2、VSIG1、CNN1、NEXN、SCRG1、SORBS1、SPARCL1、AURKA、BUB1、CDC20、CEP55、PTTG1、UBE2C、CD8A、GBP1、GBP5、GZMB、NKG7、WARS、ANTXR1、SFRP4、VCAN、CDH17、CDX1およびMYO1Aを含む標的遺伝子群のmRNAの発現水準を測定する製剤；および

ACTB、ATP5E、GPX1、UBBおよびHPRT1を含む参考遺伝子群のmRNAの発現水準を測定する製剤を含む胃癌2期および3期の進行性胃癌の予後予測用組成物。

10

## 【請求項 2】

標的遺伝子群または参考遺伝子群のmRNAの発現水準を測定する製剤は、前記mRNAに相補的な配列を有するオリゴヌクレオチドを含む、請求項1に記載の胃癌2期および3期の進行性胃癌の予後予測用組成物。

## 【請求項 3】

標的遺伝子群または参考遺伝子群のmRNAの発現水準を測定する製剤は、SEQ ID NOS：1～62に記載されたプライマーセット；またはSEQ ID NOS：63～93に記載されたプローブを含む、請求項1に記載の胃癌2期および3期の予後予測用組成物。

## 【請求項 4】

前記組成物は、全体生存率の側面で胃癌2期および3期の進行性胃癌の予後を予測するものである、請求項1に記載の胃癌2期および3期の進行性胃癌の予後予測用組成物。

20

## 【請求項 5】

請求項1に記載の胃癌2期および3期の進行性胃癌の予後予測用組成物を含む胃癌2期および3期の進行性胃癌の予後予測用キット。

## 【請求項 6】

キットは、qPCR (Quantitative real-time polymerase chain reaction) キットを含む、請求項5に記載の胃癌2期および3期の進行性胃癌の予後予測用キット。

## 【請求項 7】

統計的有意値を示すことができる程度のサンプル数を有する胃癌2期および3期の進行性胃癌患者から得た基準サンプルと生物学的サンプルで、

TFF1、TFF2およびVSIG1からなる胃 (gastric) シグネチャー；CNN1、NEXN、SCRG1、SORBS1およびSPARCL1からなる間葉 (mesenchymal) シグネチャー；AURKA、BUB1、CDC20、CEP55、PTTG1およびUBE2Cからなる増殖 (proliferative) シグネチャー；CD8A、GBP1、GBP5、GZMB、NKG7およびWARSからなる免疫 (immune) シグネチャー；ANTXR1、SFRP4およびVCANからなる幹細胞様 (stem-like) シグネチャー；およびCDH17、CDX1およびMYO1Aからなる腸 (intestinal) シグネチャーを含む標的遺伝子群と、ACTB、ATP5E、GPX1、UBBおよびHPRT1を含む参考遺伝子群のmRNA発現水準を測定する段階；

30

下記の式1により基準サンプルと生物学的サンプルの標的遺伝子群のCq値を計算して、コンピュータプログラムに入力する段階；および

前記コンピュータプログラムに入力した値に対してNMF (Non-negative Matrix Factorization) およびNMF基盤クラスタリングを行い、複数個のクラスターに分類し、各クラスターで標的遺伝子群のスコア

40

## 【数 1】

$$(d_{ik})$$

を下記の式 2 に適用して SV (Score Value) を計算し、SV によって腸 (Intestinal) 分子亜型、炎症 (Inflammatory) 分子亜型、間葉 (Mesenchymal) 分子亜型、混合型間質 (Mixed-stromal) 分子亜型および胃 (Gastric) 分子亜型に分類し、全体生存率 (Overall Survival) の側面で予後を分析して、生物学的サンプルが属する分子亜型の予後を予測する段階を含み、

前記胃癌の分子亜型は、胃 (gastric) シグネチャーの SV が最大値を有するクラスターを胃 (Gastric) 分子亜型に定め；胃 (Gastric) 分子亜型に決定されたクラスターを除いたクラスターのうち間葉 (mesenchymal) シグネチャーの SV が最大値を有し、増殖 (proliferative) シグネチャーの SV が最小値を有するクラスターを間葉 (Mesenchymal) 分子亜型に定め；胃 (Gastric) 分子亜型および間葉 (Mesenchymal) 分子亜型に決定されたクラスターを除いたクラスターのうち免疫 (immune) シグネチャーの SV が最大値を有し、腸 (intestinal) シグネチャーの SV が最小値を有するクラスターを炎症 (Inflammatory) 分子亜型に定め；胃 (Gastric) 分子亜型、間葉 (Mesenchymal) 分子亜型および炎症 (Inflammatory) 分子亜型に決定されたクラスターを除いたクラスターのうち幹細胞様 (stem-like) シグネチャーの SV が最大値を有する場合、混合型間質 (Mixed-stromal) 分子亜型に定め；最終的に残ったクラスターを腸 (Intestinal) 分子亜型に定めて分類し、

前記胃癌の分子亜型が炎症 (Inflammatory) 分子亜型である場合、全体生存率の側面で良い予後群；腸 (Intestinal) 分子亜型および胃 (Gastric) 分子亜型である場合、中間予後群；混合型間質 (Mixed-stromal) 分子亜型および間葉 (Mesenchymal) 分子亜型である場合、悪い予後群として予測する、胃癌 2 期および 3 期の予後予測のための情報を提供する方法；

## 【数 2】

[式 1]

$$\Delta C_q = (\text{標的遺伝子の } C_q \text{ 値}) - (\text{参考遺伝子群の } C_q \text{ 平均値})$$

ここで、参考遺伝子群の  $C_q$  平均値は、ACTB、ATP5E、GPX1、UBB および HPR1 を含む参考遺伝子群の  $C_q$  値の平均値を意味する。

## 【数 3】

[式 2]

$$SV (\text{Score Value}) = \frac{1}{t} \sum_{i \in SN_{\theta}} d_{ik}$$

ここで、SV は、NMF 基盤クラスタリングに得たクラスターで各シグネチャーの発現

平均値、 $t$  は、当該シグネチャーに属する遺伝子 ( $i$ ) の数、

【数 4】

$$SN_{\theta}$$

は、シグネチャー ( $\theta = 6$ )、 $k$  は、クラスターの数であって、 $2 \sim 7$  の整数であり、

【数 5】

10

$$d'_{ik}$$

は、各遺伝子の全体平均と各クラスター平均の間の距離をベースとするスコアであり、下記の式 3 により求める。

【数 6】

[式 3]

20

$$d'_{ik} = \text{sign}(d_{ik})(|d_{ik}| - \Delta)_+$$

ここで、臨界値 ( $\Delta$ ) は、 $0 \sim 1$  に定めて、子垂型別特異性がない遺伝子は 0 に収斂させ、

【数 7】

30

$$\text{sign}(d_{ik})$$

は、

【数 8】

$$d_{ik}$$

40

の符号を意味し、

【数 9】

$$d_{ik}$$

50

は、下記の式 4 により求める。

【数 1 0】

[式 4]

$$d_{ik} = \frac{\bar{x}_{ik} - \bar{x}_i}{m_k(s_i + s_0)}$$

$$s_i^2 = \frac{1}{n-k} \sum_{k=1}^5 \sum_{j \in C_k} (x_{ij} - \bar{x}_{ik})^2 \quad 10$$

ここで、各分子垂型 (k) 内の当該遺伝子 (i) 平均値

【数 1 1】

$$\bar{x}_{ik} = \sum_{j \in C_k} \frac{x_{ij}}{n_k} \quad 20$$

当該遺伝子 (i) の全体平均値

【数 1 2】

$$\bar{x}_i = \sum_{j=1}^n \frac{x_{ij}}{n} \quad 30$$

であり、

【数 1 3】

$$m_k$$

40

は、

【数 1 4】

$$\bar{x}_{ik} - \bar{x}_i$$

の標準誤差を補正するための自由度

【数 1 5】

$$m_k = \sqrt{\frac{1}{n_k} + \frac{1}{n}}$$

10

であり、

【数 1 6】

 $S_i$ 

は、遺伝子 ( i ) の全体サンプルに対する当該分子型 ( k ) 内の標準偏差であり、

【数 1 7】

20

 $S_0$ 

は、

【数 1 8】

 $S_i$ 

30

の中央値である。

【請求項 8】

統計的有意値を示すことができる程度のサンプルの数は、300～10,000である、請求項 7 に記載の胃癌 2 期および 3 期の予後予測のための情報を提供する方法。

【請求項 9】

生物学的サンプルは、新鮮腫瘍組織、新鮮凍結腫瘍組織、パラフィン包埋腫瘍組織、微細針吸引液、腹水、管洗浄液および胸膜液で構成される群から選択される、請求項 7 に記載の胃癌 2 期および 3 期の予後予測のための情報を提供する方法。

40

【請求項 10】

標的遺伝子群または参考遺伝子群の mRNA の発現水準測定は、qPCR ( Quantitative real-time polymerase chain reaction ) により行われる、請求項 7 に記載の胃癌 2 期および 3 期の予後予測のための情報を提供する方法。

【請求項 11】

統計的有意値を示すことができる程度のサンプル数を有する胃癌 2 期および 3 期の進行性胃癌患者から得た基準サンプルと生物学的サンプルで、

TFF1、TFF2 および VSI G1 からなる胃 ( gastric ) シグネチャー ; CNN1、NEXN、SCR G1、SORBS1 および SPARCL1 からなる間葉 ( me

50

senchymal)シグネチャー；AURKA、BUB1、CDC20、CEP55、PTTG1およびUBE2Cからなる増殖(proliferative)シグネチャー；CD8A、GBP1、GBP5、GZMB、NKG7およびWARSからなる免疫(immune)シグネチャー；ANTXR1、SFRP4およびVCANからなる幹細胞様(stem-like)シグネチャー；およびCDH17、CDX1およびMYO1Aからなる腸(intestinal)シグネチャーを含む標的遺伝子群と、ACTB、ATP5E、GPX1、UBBおよびHPRT1を含む参考遺伝子群のmRNA発現水準を測定する段階；

下記の式1により基準サンプルと生物学的サンプルの標的遺伝子群のCq値を計算して、コンピュータプログラムに入力する段階；および

前記コンピュータプログラムに入力した値に対してNMF(Non-negative Matrix Factorization)およびNMF基盤クラスタリングを行い、複数個のクラスターに分類し、各クラスターで標的遺伝子群のスコア

【数19】

10

(*dik*)

20

を下記の式2に適用してSV(Score Value)を計算して、SVによって腸(Intestinal)分子亜型、炎症(Inflammatory)分子亜型、間葉(Mesenchymal)分子亜型、混合型間質(Mixed-stromal)分子亜型および胃(Gastric)分子亜型で分類して、全体生存率(Overall Survival)の側面で予後を分析して生物学的サンプルが属する分子亜型の予後を予測する段階を含み、

前記胃癌の分子亜型は、胃(gastric)シグネチャーのSVが最大値を有するクラスターを胃(Gastric)分子亜型に定め；胃(Gastric)分子亜型に決定されたクラスターを除いたクラスターのうち間葉(mesenchymal)シグネチャーのSVが最大値を有し、増殖(proliferative)シグネチャーのSVが最小値を有するクラスターを間葉(Mesenchymal)分子亜型に定め；胃(Gastric)分子亜型および間葉(Mesenchymal)分子亜型に決定されたクラスターを除いたクラスターのうち免疫(immune)シグネチャーのSVが最大値を有し、腸(intestinal)シグネチャーのSVが最小値を有するクラスターを炎症(Inflammatory)分子亜型に定め；胃(Gastric)分子亜型、間葉(Mesenchymal)分子亜型および炎症(Inflammatory)分子亜型に決定されたクラスターを除いたクラスターのうち幹細胞様(stem-like)シグネチャーのSVが最大値を有する場合、混合型間質(Mixed-stromal)分子亜型に定め；最終的に残ったクラスターを腸(Intestinal)分子亜型に定めて分類し、

30

40

前記胃癌の分子亜型が炎症(Inflammatory)分子亜型である場合、全体生存率の側面で良い予後群；腸(Intestinal)分子亜型および胃(Gastric)分子亜型である場合、中間予後群；混合型間質(Mixed-stromal)分子亜型および間葉(Mesenchymal)分子亜型である場合、悪い予後群として予測する、胃癌2期および3期の予後予測方法；

【数 2 0】

[式 1]

$$\Delta C_q = (\text{標的遺伝子の } C_q \text{ 値}) - (\text{参考遺伝子群の } C_q \text{ 平均値})$$

ここで、参考遺伝子群の  $C_q$  平均値は、ACTB、ATP5E、GPX1、UBBおよびHPR1を含む参考遺伝子群の  $C_q$  値の平均値を意味する。

【数 2 1】

10

[式 2]

$$SV (\text{Score Value}) = \frac{1}{t} \sum_{i \in SN_\theta} d'_{ik}$$

20

ここで、 $SV$  は、NMF 基盤クラスタリングに得たクラスターで各シグネチャーの発現平均値、 $t$  は、当該シグネチャーに属する遺伝子 ( $i$ ) の数、

【数 2 2】

$$SN_\theta$$

は、シグネチャー ( = 6 )、 $k$  は、クラスターの数であって、2 ~ 7 の整数であり

30

【数 2 3】

$$d'_{ik}$$

は、各遺伝子の全体平均と各クラスター平均の間の距離をベースとするスコアであり、下記の式 3 により求める。

【数 2 4】

40

[式 3]

$$d'_{ik} = \text{sign}(d_{ik})(|d_{ik}| - \Delta)_+$$

ここで、臨界値 ( ) は、0 . 1 に定めて、分子亜型別の特異性がない遺伝子は、0 に収斂させ、

【数 2 5】

$$\text{sign}(d_{ik})$$

は、

【数 2 6】

$$d_{ik}$$

10

の符号を意味し、

【数 2 7】

$$d_{ik}$$

20

は、下記の式 4 により求める。

【数 2 8】

[式 4]

$$d_{ik} = \frac{\bar{x}_{ik} - \bar{x}_i}{m_k(s_i + s_0)}$$

$$s_i^2 = \frac{1}{n-k} \sum_{k=1}^5 \sum_{j \in C_k} (x_{ij} - \bar{x}_{ik})^2$$

30

ここで、各分子亜型 (k) 内の当該遺伝子 (i) 平均値

【数 2 9】

$$\bar{x}_{ik} = \sum_{j \in C_k} \frac{x_{ij}}{n_k}$$

( )、

40

当該遺伝子 (i) の全体平均値

【数 3 0】

$$\bar{x}_i = \sum_{j=1}^n \frac{x_{ij}}{n}$$

であり、

【数 3 1】

$$m_k$$

は、

【数 3 2】

$$\bar{x}_{ik} - \bar{x}_i$$

の標準誤差を補正するための自由度

【数 3 3】

$$m_k = \sqrt{\frac{1}{n_k} + \frac{1}{n}}$$

であり、

【数 3 4】

$$s_i$$

は、遺伝子 ( i ) の全体サンプルに対する当該分子垂型 ( k ) 内の標準偏差であり、

【数 3 5】

$$s_0$$

は、

10

20

30

40

【数 3 6】

$S_i$

の中央値である。

【請求項 1 2】

統計的有意値を示すことができる程度のサンプルの数は、300～10,000である、請求項 1 1 に記載の胃癌 2 期および 3 期の予後予測方法。 10

【請求項 1 3】

生物学的サンプルは、新鮮腫瘍組織、新鮮凍結腫瘍組織、パラフィン包埋腫瘍組織、微細針吸引液、腹水、管洗浄液および胸膜液で構成される群から選択される、請求項 1 1 に記載の胃癌 2 期および 3 期の予後予測方法。

【請求項 1 4】

標的遺伝子群または参考遺伝子群の mRNA の発現水準測定は、qPCR (Quantitative real-time polymerase chain reaction) により行われる、請求項 1 1 に記載の胃癌 2 期および 3 期の予後予測方法。

【発明の詳細な説明】

【技術分野】 20

【0001】

本発明は、胃癌の生物学的特性に基づく群区分および予後予測システムに関する。

【背景技術】

【0002】

全世界的に、胃癌は、癌による死亡率のうち三番目に高い癌であり、特に韓国内では予後が比較的良いと知られている甲状腺癌を除いて最もありふれた癌である。韓国内では、国家次元の健康診断を通じた早期発見と手術の標準化および抗癌治療剤の発見などで胃癌患者の生存率が大きく向上したが、現在標準化された治療にもかかわらず、依然として 2 期、3 期進行性胃癌の場合、約半分程度の患者が再発を経験する。

【0003】 30

癌は、遺伝体疾患と認識されており、NGS (Next Generation Sequencing) 等の遺伝体検査技術の発展に伴い、癌を既存の解剖学的、病理学的表現型による分類でなく、分子的生物学的特性によって分類しようとする努力があった。TCGA (The Cancer Genome Atlas) プロジェクトにおいて胃癌がその多様な分子的特徴によって大きく 4 種類の形態に分けられることが、最近報告されたことがある。これは、解剖学的に同じ病期であるとしても、分子的生物学的特徴によって予後と抗癌剤に対する反応程度が異なり得ることを意味する。

【0004】

最近に発表された 295 人の胃癌患者の TCGA プロジェクト結果を見れば、胃癌は、(1) EBV 陽性胃癌 (Epstein-Barr virus positive, EBV positive) (2) マイクロサテライト不安定型 (Microsatellite instability-high, MSI-H)、(3) 染色体不安定型 (Chromosomal instability, CIN) (4) 遺伝体安定型胃癌 (Genomically stable, GS) の 4 種類に区分される。このような膨大な癌ゲノムシーケンシング (Cancer Genome Sequencing) を通じて胃癌も一つの単一の癌腫でなく、分子遺伝学的に区別される異質な小グループに分けられるという点に分かる。したがって、胃癌の個人に合ったオーダーメイド治療を実現するためには、分子遺伝学的特徴および病理学的特徴に基づくサブタイプを区別してそれぞれのターゲット遺伝子を発掘、適用が必要であることを示唆する。また、胃癌の研究の側面において、胃癌の亜型区分によって予後が区分され得る結果が報告されている。 40

## 【 0 0 0 5 】

胃癌手術による抗癌剤 k 治療後、患者の予後を予測できるとすれば、各予後に応じてこれに合った治療戦略を樹立できる根拠資料になる。現在の標準化された治療慣行上、2、3期の進行性胃癌では、手術後に補助抗癌療法をすべての患者に使用されている。これは、予後が悪い群に対しては、過小治療 (under treatment) でありうる。すなわち、予後が良くない患者群に対して現在の標準治療以外に他の追加的な治療方法に対する戦略を開発できる臨床学的な意味を有しているといえる。

## 【 0 0 0 6 】

2010年以後、現在2期、3期の進行性胃癌の場合、標準化されたD2胃切除手術以後、補助抗癌療法が胃癌患者の生存率を高めることを発見し、現在、これは、標準治療に該当する。伝統的に、胃癌は、その解剖学的病理学的表現型によって分類し、TNM病期分類法によって2期以上の場合、抗癌治療をしているが、抗癌治療による予後を予測できる方法がTNM病期以外にはない状況である。

10

## 【 発明の概要 】

## 【 発明が解決しようとする課題 】

## 【 0 0 0 7 】

本発明の目的は、進行性胃癌 (2期 - 3期 : AJCC 6版基準) 患者の手術後、予後を予測できる標的遺伝子群の mRNA 発現水準の定量値を基盤とする進行性胃癌の予後予測用組成物を提供することにある。

## 【 0 0 0 8 】

本発明の他の目的は、進行性胃癌患者の手術後、予後を予測できる標的遺伝子群の mRNA 発現水準の定量値を基盤として患者の生存率の側面で予後を予測するための情報を提供する方法を提供することにある。

20

## 【 0 0 0 9 】

本発明のさらに他の目的は、進行性胃癌患者の手術後、予後を予測できる標的遺伝子群の mRNA 発現水準の定量値を基盤として患者の生存率の側面で予後を予測する方法を提供することにある。

## 【 課題を解決するための手段 】

## 【 0 0 1 0 】

前記目的を達成するために、本発明は、TFF1、TFF2、VSI G1、CNN1、NEXN、SCR G1、SORBS1、SPARCL1、AURKA、BUB1、CDC20、CEP55、PTTG1、UBE2C、CD8A、GBP1、GBP5、GZMB、NKG7、WARS、ANTXR1、SFRP4、VCAN、CDH17、CDX1およびMYO1Aを含む標的遺伝子群の mRNA の発現水準を測定する製剤 ; およびACTB、ATP5E、GPX1、UBBおよびHPR T1を含む参考遺伝子群の mRNA の発現水準を測定する製剤を含む胃癌2期および3期の進行性胃癌の予後予測用組成物を提供する。

30

## 【 0 0 1 1 】

また、本発明は、前記胃癌2期および3期の予後予測用組成物を含む胃癌2期および3期の進行性胃癌の予後予測用キットを提供する。

40

## 【 0 0 1 2 】

また、本発明は、統計的有意値を示すことができる程度のサンプル数を有する胃癌2期および3期の進行性胃癌患者から得た基準サンプルと生物学的サンプルで、

TFF1、TFF2およびVSI G1からなる胃 (gastric) シグネチャー ; CNN1、NEXN、SCR G1、SORBS1およびSPARCL1からなる間葉 (mesenchymal) シグネチャー ; AURKA、BUB1、CDC20、CEP55、PTTG1およびUBE2Cからなる増殖 (proliferative) シグネチャー ; CD8A、GBP1、GBP5、GZMB、NKG7およびWARSからなる免疫 (immune) シグネチャー ; ANTXR1、SFRP4およびVCANからなる幹細胞様 (stem-like) シグネチャー ; およびCDH17、CDX1およびMYO1Aが

50

らなる腸 ( i n t e s t i n a l ) シグネチャーを含む標的遺伝子群と、ACTB、ATP5E、GPX1、UBBおよびHPRT1を含む参考遺伝子群のmRNA発現水準を測定する段階；

下記の式1により基準サンプルと生物学的サンプルの標的遺伝子群のCq値を計算して、コンピュータプログラムに入力する段階；および

前記コンピュータプログラムに入力した値に対してNMF ( N o n - n e g a t i v e M a t r i x F a c t o r i z a t i o n ) およびNMF基盤クラスタリングを行い、複数個のクラスターに分類し、各クラスターで標的遺伝子群のスコア

【数1】

10

$$\left( \frac{d_{ik}}{\sum_j d_{ij}} \right)$$

を下記の式2に適用してSV ( S c o r e V a l u e ) を計算し、SVによって腸 ( I n t e s t i n a l ) 分子亜型、炎症 ( I n f l a m m a t o r y ) 分子亜型、間葉 ( M e s e n c h y m a l ) 分子亜型、混合型間質 ( M i x e d - s t r o m a l ) 分子亜型および胃 ( G a s t r i c ) 分子亜型に分類し、全体生存率 ( O v e r a l l S u r v i v a l ) の側面で予後を分析して、生物学的サンプルが属する分子亜型の予後を予測する段階を含み、

20

前記胃癌の分子亜型は、胃 ( g a s t r i c ) シグネチャーのSVが最大値を有するクラスターを胃 ( G a s t r i c ) 分子亜型に定め；胃 ( G a s t r i c ) 分子亜型に決定されたクラスターを除いたクラスターのうち間葉 ( m e s e n c h y m a l ) シグネチャーのSVが最大値を有し、増殖 ( p r o l i f e r a t i v e ) シグネチャーのSVが最小値を有するクラスターを間葉 ( M e s e n c h y m a l ) 分子亜型に定め；胃 ( G a s t r i c ) 分子亜型および間葉 ( M e s e n c h y m a l ) 分子亜型に決定されたクラスターを除いたクラスターのうち免疫 ( i m m u n e ) シグネチャーのSVが最大値を有し、腸 ( i n t e s t i n a l ) シグネチャーのSVが最小値を有するクラスターを炎症 ( I n f l a m m a t o r y ) 分子亜型に定め；胃 ( G a s t r i c ) 分子亜型、間葉 ( M e s e n c h y m a l ) 分子亜型および炎症 ( I n f l a m m a t o r y ) 分子亜型に決定されたクラスターを除いたクラスターのうち幹細胞様 ( s t e m - l i k e ) シグネチャーのSVが最大値を有する場合、混合型間質 ( M i x e d - s t r o m a l ) 分子亜型に定め；最終的に残ったクラスターを腸 ( I n t e s t i n a l ) 分子亜型に定めて分類し、

30

前記胃癌の分子亜型が炎症 ( I n f l a m m a t o r y ) 分子亜型である場合、全体生存率の側面で良い予後群；腸 ( I n t e s t i n a l ) 分子亜型および胃 ( G a s t r i c ) 分子亜型である場合、中間予後群；混合型間質 ( M i x e d - s t r o m a l ) 分子亜型および間葉 ( M e s e n c h y m a l ) 分子亜型である場合、悪い予後群として予測する、胃癌2期および3期の予後予測のための情報を提供する方法、または、胃癌2期および3期の予後予測方法を提供する；

40

【0013】

【数2】

[式1]

$$\Delta Cq = (\text{標的遺伝子の} Cq \text{値}) - (\text{参考遺伝子群の} Cq \text{平均値})$$

【0014】

50

ここで、参考遺伝子群の C q 平均値は、ACTB、ATP5E、GPX1、UBBおよびHPR1を含む参考遺伝子群の C q 値の平均値を意味する。

【0015】

【数3】

[式2]

$$SV (\text{Score Value}) = \frac{1}{t} \sum_{i \in SN_{\theta}} d'_{ik}$$

10

【0016】

ここで、SVは、NMF基盤クラスタリングに得たクラスターで各シグネチャーの発現平均値、tは、当該シグネチャーに属する遺伝子(i)の数、

【数4】

20

$SN_{\theta}$

は、シグネチャー( = 6)、kは、クラスターの数であって、2~7の整数であり、

【数5】

$d'_{ik}$

30

は、各遺伝子の全体平均と各クラスター平均との間の距離をベースとするスコアであって、下記の式3により求める。

【0017】

【数6】

[式3]

$$d'_{ik} = \text{sign}(d_{ik})(|d_{ik}| - \Delta)_{+}$$

40

【0018】

ここで、臨界値( )は、0.1に定めて、分子亜型別の特異性がない遺伝子は、0に収斂させ、

【数 7】

$$\text{sign}(d_{ik})$$

は、

【数 8】

$$d_{ik}$$

10

の符号を意味し、

【数 9】

$$d_{ik}$$

20

は、下記の式 4 により求める。

【0019】

【数 10】

[式 4]

$$d_{ik} = \frac{\bar{x}_{ik} - \bar{x}_i}{m_k(s_i + s_0)}$$

$$s_i^2 = \frac{1}{n-k} \sum_{k=1}^5 \sum_{j \in C_k} (x_{ij} - \bar{x}_{ik})^2$$

30

【0020】

ここで、各分子垂型 (k) 内の当該遺伝子 (i) 平均値

【数 11】

40

$$\bar{x}_{ik} = \sum_{j \in C_k} \frac{x_{ij}}{n_k}$$

( )、

当該遺伝子 (i) の全体平均値

50

【数 1 2】

$$\bar{x}_i = \sum_{j=1}^n \frac{x_{ij}}{n}$$

であり、

【数 1 3】

$$m_k$$

は、

【数 1 4】

$$\bar{x}_{ik} - \bar{x}_i$$

の標準誤差を補正するための自由度

【数 1 5】

$$m_k = \sqrt{\frac{1}{n_k} + \frac{1}{n}}$$

であり、

【数 1 6】

$$s_i$$

は、遺伝子 ( i ) の全体サンプルに対する当該分子亜型 ( k ) 内の標準偏差であり、

10

20

30

40

【数 17】

 $S_0$ 

は、

【数 18】

 $S_i$ 

10

の中央値である。

【発明の効果】

【0021】

本発明は、進行性胃癌の予後を予測できる標的遺伝子群の mRNA 発現水準の定量値を利用して全体生存率の側面で予後を予測できるアルゴリズムを開発して、胃癌患者の治療方法を決定するのに補助的情報として活用することができる。

【図面の簡単な説明】

20

【0022】

【図1】図1は本発明の実験分析流れ図を示すものであり、段階Iは、一連の分析結果、胃癌(GC)分子亜型、分類器(classifier)およびGCシグネチャーを確認する過程(NMF、non-negative matrix factorization; SAM、significance analysis of microarrays; PAM、prediction analysis of microarrays; GSEA、gene-set enrichment analysis; WGCNA、weighted gene co-expression network analysis)であり、段階IIは、患者サンプルでGCシグネチャーの発現をプロファイリングするための一連のプロブ、miniClassifier-26の構築を通じて臨床効用を最大化するためのGCシグネチャーと分子亜型の同定過程である。

30

【図2】図2は、Classifier-PAM932でNMFコンセンサスクラスタリングを行った結果を示すものであり、(a)は、Molecular Signatures Database(MSigDB)および以前の研究で分析された遺伝子セット(消化、痙攣抑制ポリペプチド-発現化生(SPEM)、腸上皮化生(intestinal metaplasia(IM)、免疫系、基質、上皮間葉転換(EMT)、および細胞周期)で特性糾明されたトレーニングセットIのヒートマップ(GSE13861p)を示す。GST、胃亜型; INF、炎症; MSC、間葉; INT、臓器; MXD、混合基質。独立テストセットIとClassifier-PAM932を使用したNMFコンセンサスクラスタリングの検証結果である。(b)GSE62254(アジア癌研究グループ、ACRG)、(c)TCGAおよび(d)GSE15459(シンガポール)のヒートマップを示す。以前の亜型情報は、各ヒートマップの上に共に提供される。

40

【図3】図3は、NMFクラスタリングを示すものであり、 $SD = 0.8$ 、 $SD = 0.9$ 、 $SD = 1.0$ での(a)クラスター番号(k)2~7および次の分散カットオフがあるコンセンサスマップ、(b)コーフェン相関係数および(c)シルエット幅を使用したサンプル指定を示す。

【図4】図4は、WGCNA分析により発見された亜型を定義する胃癌シグネチャーを示すものであり、(a)は、GSE13861pの系統数であり、(b)は、GSE13861pで探知されたモジュールおよびGSE62254、TCGAおよびGSE15459の当該モジュールマップである。独立的なテストセットである各コホートのモジュール

50

が G S E 1 3 8 6 1 p のモジュールと共通遺伝子を共有すれば、モジュールの色を視覚的便宜のために一対一で調整され、(c)は、トレーニングセットで各亜型の上位 25% 上位点数 P A M g e n e s マッピング結果である。(d)は、GCで一致するモジュールの組合せで表現される6個のGCシグネチャーを示す。(e)は、5個のGC亜型と6個のGCシグネチャーの関係性を分析するために、スピアマンの相関関係 ( S p e a r m a n ' s c o r r e l a t i o n ) を使用して示すものである。赤色は、当該亜型と正の相関関係があるモジュールを示し、青色は、負の相関関係があるモジュールを示す。

【図5】図5は、段階 I I のトレーニングセットである M e r g e d 1 2 5 9 ( G S E 1 3 8 6 1 p、G S E 6 2 2 5 4 ( A C R G )、T C G A、および G S E 1 5 4 5 9 ( シンガポール ) ) を使用して5個の亜型による全体生存率 ( O S ) を示すものである ( l i k e l i h o o d r a t i o t e s t ; p = 3 . 4 2 e - 0 9 ) 。 M e r g e d 1 2 5 9 コホートのサンプルのうち臨床情報がない61個のサンプルを除いた1198個のサンプルに対する全体生存曲線である：各分子亜型別の5年生存率は、I N F の場合、76.1% ( 95% 信頼区間 67.7 - 85.7 )、I N T の場合、65.1% ( 95% 信頼区間 56.2 - 75.4 )、G S T の場合、64.6% ( 95% 信頼区間 55.0 - 75.9 )、M X D の場合、51.3% ( 95% 信頼区間 42.1 - 62.4 )、M S C の場合、46.3% ( 95% 信頼区間 38.0 - 56.5 ) である。

【図6】図6は、胃癌細胞株を利用して調査された胃癌基質シグネチャー ( n = 26 ) を示すものであり、(a)は、癌細胞株データと併合されたトレーニングセットでGC亜型を示すヒートマップである。(b)は、基質モジュール e i g e n g e n e と共に配列されたヒートマップである。H s 7 4 6、S N U - 4 8 4 ( M S C )、M K N - 4 5 および N C I - N 8 7 ( I N T ) 細胞株は、矢先で表示される。

【図7】図7は、胃癌細胞株の M S C および I N T タイプは、(a)インビトロスクラッチ創傷治癒法、(b)浸潤分析法、(c)インビトロ腫瘍スフェロイド形成分析法 (スケールバー、100 μ m )、および (d) 生体内同所腫瘍形成 ( n = 3 ) に比較される。H s 7 4 6 T および S N U - 4 8 4 腫瘍の拡散成長および M K N - 4 5 および N C I - N 8 7 腫瘍の閉じ込めは、M R I イメージ (軸方向の断面) で白色点線で境界付けられる。黒い点線からなるボックスの写真は、定位モデル構成を描写する。T G F - 抑制剤 ( L Y 2 1 5 7 2 9 9 ( L Y ) ) で処理した H s 7 4 6 T 細胞の基質挙動の抑制は、試験管内スクラッチ創傷治癒法 ( e )、浸潤法 ( f ) およびインビトロ腫瘍スフェロイド形成分析 ( g ) を通じて観察する。複合薬物治療 ( オキサリプラチンおよびフルオロウラシル ) 中に T G F - 抑制剤の同時投与下でマウス異種移植モデル ( n = 8 ) で H s 7 4 6 t 腫瘍 ( h ) および N C I - N 8 7 腫瘍 ( i ) の腫瘍成長を測定する生体内薬物抵抗分析結果である ( P < 0 . 0 5 ) 。

【図8】図8は、GCで m i n i C l a s s i f i e r 2 6 で区分された5個の分子亜型を示すものであり、(a)は、P A M で5個の分子亜型に対する26個の代表遺伝子の相対的差異を示すものである。(b)は、GCのマイクロアレイデータに対する N M F 基盤クラスタリング ( 26個の遺伝子 ) のヒートマップ ( C o m B a t 方法を利用して併合された N = 1 2 5 9 ; G S E 1 3 8 6 1 p、G S E 1 5 4 5 9、T C G A および G S E 6 2 2 5 4 ) である。(c)は、GCの q P C R データに対する N M F 基盤クラスタリング ( 26個の遺伝子 ) のヒートマップである。(d)と(e)は、(b)と(c)でクラスタリングされた5個の分子亜型の全体生存率 ( O S ) を示すものである。(d)は、M e r g e d 1 2 5 9 コホートのサンプルのうち臨床情報がない61個のサンプルを除いた1198個のサンプルに対する全体生存曲線である：各分子亜型別の5年生存率は、I N F の場合、67.3% ( 95% 信頼区間 61.3 ~ 73.9% )、I N T の場合、58.8% ( 95% 信頼区間 52.9 ~ 65.4% )、G S T の場合、55.3% ( 95% 信頼区間 48.2 ~ 63.4% )、M X D の場合、45.0% ( 95% 信頼区間 36.5 ~ 55.4% )、M S C の場合、33.0% ( 95% 信頼区間 27.3 ~ 40.0% ) である。(e)は、q P C R 3 2 5 コホートの q P C R 測定結果から26個の遺伝子で区分した群の5年全体生存曲線である：各分子亜型別の5年生存率は、I N F の場合、78.5% ( 9

10

20

30

40

50

5%信頼区間69.4 - 88.8%)、INTの場合、70.7%(95%信頼区間61.1 - 81.8%)、GSTの場合、68.4%(95%信頼区間55.1 - 84.9%)、MXDの場合、54.6%(95%信頼区間41.7 - 71.4%)、MSCの場合、57.5%(95%信頼区間48.4 - 68.2%)である。

【図9】図9は、新鮮凍結組織のマイクロアレイデータとqPCRデータ、およびパラフィン包埋サンプルのqPCRで安定性を有する遺伝子26個(miniClassifier-26)を選定および分析する詳細なフローチャートを示すものである。

【図10】図10は、NMFクラスタリングを通じて得た5個のクラスターを分子亜型で命名する方法を示すものである。

【発明を実施するための形態】

【0023】

以下、本発明の構成を具体的に説明する。

【0024】

本発明は、TFF1、TFF2、VSIG1、CNN1、NEXN、SCRG1、SORBS1、SPARCL1、AURKA、BUB1、CDC20、CEP55、PTTG1、UBE2C、CD8A、GBP1、GBP5、GZMB、NKG7、WARS、ANTXR1、SFRP4、VCAN、CDH17、CDX1およびMYO1Aを含む標的遺伝子群のmRNAの発現水準を測定する製剤；および

ACTB、ATP5E、GPX1、UBBおよびHPRT1を含む参考遺伝子群のmRNAの発現水準を測定する製剤を含む胃癌2期および3期の進行性胃癌の予後予測用組成物に関するものである。

【0025】

本発明の胃癌2期および3期の予後予測用組成物は、標的遺伝子群のmRNA発現水準を測定して、進行性胃癌患者の予後を生存率の側面で予測するための用途で使用できることを特徴とする。

【0026】

本明細書で、用語「進行性胃癌」は、AJCC6版を基準として2期～3期に該当する胃癌を意味する。

【0027】

本明細書で、用語「標的遺伝子」または「マーカー遺伝子」は、明細書内で混用されて使用され、正常や病的な状態を区分できたり、治療後、5年生存率を予測したり、治療反応の予測を客観的に測定できる標識子を意味する。本発明では、進行性胃癌の予後を予測するのに使用できる遺伝子であり、予後に対して増加したり減少する差等的なmRNA発現水準を示す遺伝子である。本発明の一具体例によれば、異形質性を有する胃癌に対して新鮮凍結組織のマイクロアレイデータとqPCRデータ、およびパラフィン包埋サンプル検体のqPCRデータで統計的有意性を確保して、5個の分子亜型、すなわち、腸(Intestinal)分子亜型、炎症(Inflammatory)分子亜型、間葉(Mesenchymal)分子亜型、混合型間質(Mixed-stromal)分子亜型および胃(Gastric)分子亜型に区分できる26種の遺伝子、すなわち、TFF1、TFF2、VSIG1、CNN1、NEXN、SCRG1、SORBS1、SPARCL1、AURKA、BUB1、CDC20、CEP55、PTTG1、UBE2C、CD8A、GBP1、GBP5、GZMB、NKG7、WARS、ANTXR1、SFRP4、VCAN、CDH17、CDX1およびMYO1Aを選定した。

【0028】

本明細書で、用語「参考遺伝子(reference gene)」は、常に安定して発現する遺伝子を指す。すなわちどんな組織でも一定に発現する遺伝子であって、参考遺伝子の発現量とマーカー遺伝子の発現量を比較することにより、マーカー遺伝子の発現量を調査するとき使用する。すなわち、サンプルごとに定性(quality)的差異、保管機関による変移が存在するので、遺伝子の発現量を測定しても、その測定量が生物学的変移だと判断し難い。したがって、標準化(normalization)を通じてサ

10

20

30

40

50

ンプル間の遺伝子発現量 ( C q ) を決定する。通常、標準化方法には、 Q u a n t i l e による方法、 G l o b a l N o r m a l i z a t i o n 方法、参考遺伝子による方法などがあるが、本発明は、参考遺伝子による標準化を使用する。また、単一遺伝子を参考遺伝子として活用することは、正確度に劣ることがあり、多数の遺伝子を選定し、変異度を調査して組織の特性に適合した参考遺伝子を選定することができる。本発明では、胃癌と関連して文献に開示されていたり、既存の商用化製品で活用されている遺伝子を選定し、選定された遺伝子を対象として適格の可否を立証して参考遺伝子として使用する。本発明の一具体例によれば、文献に開示された 21 個の参考遺伝子を対象として食道癌、すい臓癌、胃癌、大腸癌などの癌組織と正常組織を比較して、 q P C R を通じて最も変異度が小さい遺伝子を参考遺伝子として選定した。次に、商用化製品で使用する参考遺伝子として、 A C T B 、 A T P 5 E 、 G P X 1 、 U B B および H P R T 1 を選定して q P C R を行い、最終的に、本発明の進行性胃癌の再発または抗癌剤反応の可能性を予測するのに使用する参考遺伝子として A C T B 、 A T P 5 E 、 G P X 1 、 U B B および H P R T 1 からなる遺伝子群を使用した。

10

#### 【 0 0 2 9 】

本明細書で、用語「 mRNA の発現水準測定」とは、進行性胃癌の再発を予測するために生物学的試料で予後マーカー遺伝子の mRNA 発現程度を確認する過程であり、 mRNA の量を測定することを意味する。例えば、 q P C R ( Q u a n t i t a t i v e r e a l - t i m e p o l y m e r a s e c h a i n r e a c t i o n ) が使用できるが、これに限定されるものではない。

20

#### 【 0 0 3 0 】

本発明による組成物において、予後マーカー遺伝子の mRNA の発現水準を測定する製剤は、予後マーカー遺伝子の mRNA に特異的に結合するプライマー、プローブまたはアンチセンスヌクレオチドを含む。本発明による予後マーカー遺伝子の情報は、 G e n B a n k 、 U n i P r o t など知られているので、当業者ならこれを基に遺伝子の mRNA に特異的に結合するプライマー、プローブまたはアンチセンスヌクレオチドを容易にデザインすることができる。

#### 【 0 0 3 1 】

本明細書で、用語「プライマー」は、標的遺伝子配列を認知する断片であって、正方向および逆方向のプライマー対を含むが、好ましくは、特異性および感受性を有する分析結果を提供するプライマー対である。プライマーの核酸配列が試料内存在する非 - 標的配列と不一致する配列なので、相補的なプライマー結合部位を含有する標的遺伝子配列のみを増幅し、非特異的増幅を誘発しないプライマーであるとき、高い特異性が付与され得る。本発明の一具体例によれば、 S E Q I D N O S : 1 ~ 6 2 に記載されたプライマーセットを使用することができる。各標的遺伝子群および参考遺伝子群のプライマーセットは、下記の表 1 および 2 に羅列される。

30

#### 【 0 0 3 2 】

本明細書で、用語「プローブ」とは、試料内の検出しようとする標的物質と特異的に結合できる物質を意味し、前記結合を通じて特異的に試料内の標的物質の存在を確認できる物質を意味する。プローブの種類は、当業界で通常使用される物質であって、制限はないが、好ましくは、 P N A ( p e p t i d e n u c l e i c a c i d ) 、 L N A ( l o c k e d n u c l e i c a c i d ) 、ペプチド、ポリペプチド、蛋白質、RNA または DNA であってもよい。より具体的に、前記プローブは、バイオ物質として生物に由来するか、これと類似なものまたは生体他で製造されたものを含むものであり、例えば、酵素、蛋白質、抗体、微生物、動植物細胞および器官、神経細胞、DNA、および RNA であってもよく、DNA は、cDNA、ゲノム DNA、オリゴヌクレオチドを含み、RNA は、ゲノム RNA、mRNA、オリゴヌクレオチドを含み、蛋白質の例としては、抗体、抗原、酵素、ペプチドなどを含むことができる。本発明の一具体例によれば、 S E Q I D N O S : 6 3 ~ 9 3 に記載されたプローブが使用できる。好ましくは、前記プローブは、蛍光標識されたものであってもよい。各標的遺伝子群および参考遺伝子群のプローブ

40

50

は、下記表 1 および表 2 に羅列される。

【 0 0 3 3 】

本明細書で、用語「アンチセンス」は、アンチセンスオリゴマーがワトソン・クリック型塩基対の形成によってRNA内の標的配列と混成化されて、標的配列内で典型的にmRNAとRNA：オリゴマーヘテロ二重体の形成を許容する、ヌクレオチド塩基の配列およびサブユニット間バックボーンを有するオリゴマーを意味する。オリゴマーは、標的配列に対する正確な配列相補性または近似相補性を有することができる。

【 0 0 3 4 】

本明細書で、用語「予後予測」は、特定の疾病または疾患に対する対象 ( s u b j e c t ) の感受性 ( s u s c e p t i b i l i t y ) を判定すること、特定の疾病または疾患にかかった対象の予後 ( p r o g n o s i s ; 例えば、前 - 転移性または転移性癌状態の同定、癌の段階決定または治療に対する癌の反応性決定 ) を判定すること、またはセラメトリックス ( t h e r a m e t r i c s ; 例えば、治療効能に対する情報を提供するためにオブジェクトの状態をモニタリングすること ) を含む。本発明の目的上、手術後、胃癌患者の予後を全体生存率 ( O v e r a l l S u r v i v a l ) の側面で予測することである。

【 0 0 3 5 】

本発明の一具体例によれば、前記標的遺伝子は、次のように選別する。まず、進行性胃癌組織を対象にコンセンサス基盤のNMFを進めて、高い関連性を示す5個の分子的特徴が区分される亜型を決定し、亜型特異的な遺伝子群を選別する。

【 0 0 3 6 】

前記NMF ( 非負値行列因子分解、 n o n n e g a t i v e m a t r i x f a c t o r i z a t i o n ) は、次元縮小方法 ( D i m e n s i o n R e d u c t i o n ) で発現パターンの類似性が高いサンプルを亜型に分類する方法である。

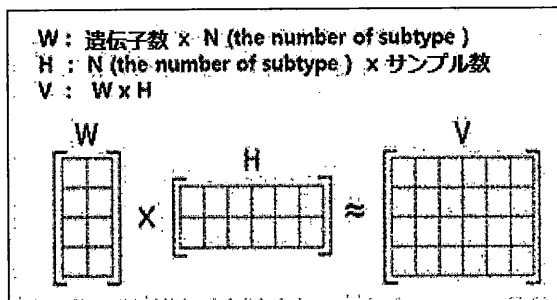
【 0 0 3 7 】

前記胃癌分子亜型を決定するためには、資源縮小方法で一つのマトリックスを二つの非負値行列で因数分解することにより、NMFアルゴリズムを通じて共通した部分 ( b a s i s ) を分離することである。実際の情報集団をV、分離しようとする行列をW、Hという。V = WHを満たすようになり、Wは、基底行列を、Hは、符号化行列を示す。Vは、Wの基底の和で構成、Vは、( n x m ) サイズ、Wは、( n x r )、Hは、( r x m ) サイズの行列である。高次元のデータ行列を低次元の係数行列と基底行列に分解し、各行列が希少特性を有しているので、部分基盤表現が可能である。このアルゴリズムにより、胃癌は、類似性が高いもの同士の群を形成して胃癌亜型分類に適用することができる。

【 0 0 3 8 】

【表 A】

[参考図 1]



【 0 0 3 9 】

RプログラムのNMFパッケージを利用して、コンセンサスマップ ( c o n s e n s u s m a p ) とコーフェン ( c o p h e n e t i c ) グラフを根拠に胃癌亜型の適合した

クラスターを選定する。それぞれのSD (Standard Deviation = 0.8、0.9、1.0) でk (クラスターの数) 値にコンセンサスクラスタリング (consensus clustering) されることを確認し、最終的に、胃癌亜型の個数を選択する (図3参照)。本発明の一具体例によれば、Kは、5に定めた。

**【0040】**

前記各亜型別の特徴を生物学的遺伝子セットから導き出して分子的異形質性を区分した結果、5個の亜型、すなわち、腸 (Intestinal) 分子亜型、炎症 (Inflammatory) 分子亜型、間葉 (Mesenchymal) 分子亜型、混合型間質 (Mixed-stromal) 分子亜型および胃 (Gastric) 分子亜型に区分される。

10

**【0041】**

コックス比例ハザードモデル (Cox proportional hazard model) を利用して前記5個の分子亜型格別の生存分析を行った結果、炎症 (Inflammatory) 分子亜型である場合、全体生存率の側面で良い予後、腸 (Intestinal) 分子亜型および胃 (Gastric) 分子亜型である場合、中間予後、混合型間質 (Mixed-stromal) 分子亜型および間葉 (Mesenchymal) 分子亜型である場合、悪い予後を示す。

**【0042】**

一方、胃癌の生物学的特性を分析するためには、WGCNA (Weighted correlation network analysis) を使用する。WGCNAは、発現様相が類似に現れる遺伝子間の相関関係を通じて発現関連性によって遺伝子をクラスタリングする分析技法であって、WGCNAを通じてモジュール (clusters of highly interconnected genes) を探し、モジュールの性格とモジュールの相関関係によって胃癌特性を反映したシグネチャーを選別する。本発明の一具体例によれば、胃癌で安定的に保存されている6個のシグネチャー、すなわち、TFF1、TFF2およびVSI G1からなる胃 (gastric) シグネチャー；CNN1、NEXN、SCR G1、SORBS1およびSPARCL1からなる間葉 (mesenchymal) シグネチャー；AURKA、BUB1、CDC20、CEP55、PTTG1およびUBE2Cからなる増殖 (proliferative) シグネチャー；CD8A、GBP1、GBP5、GZMB、NKG7およびWARSからなる免疫 (immune) シグネチャー；ANTXR1、SFRP4およびVCANからなる幹細胞様 (stem-like) シグネチャー；およびCDH17、CDX1およびMYO1Aからなる腸 (intestinal) シグネチャーを確認した。

20

30

**【0043】**

これらの6個のシグネチャーを代表する遺伝子と5個の胃癌分子亜型の間をスピアマン相関関係で比較分析したとき、相互関連性を示した。これを通じて、共通して保存される遺伝子を選別した後、これらをサンプル方法 (Fresh Frozen、FFPE) によって、そして、発現測定プラットフォーム (microarray、qPCR) によって安定的に発現結果を示す26個の遺伝子、TFF1、TFF2、VSI G1、CNN1、NEXN、SCR G1、SORBS1、SPARCL1、AURKA、BUB1、CDC20、CEP55、PTTG1、UBE2C、CD8A、GBP1、GBP5、GZMB、NKG7、WARS、ANTXR1、SFRP4、VCAN、CDH17、CDX1およびMYO1Aを選別した。

40

**【0044】**

前記選別された26個の遺伝子を標的遺伝子群に定め、各標的遺伝子のスコア

【数 19】

( *d<sub>ik</sub>* )

を利用して標的遺伝子群の各シグネチャーの発現平均値を計算し、これから分類された分子亜型を確認する。

10

【0045】

前記で胃癌の分子亜型が決定されれば、各分子亜型の生存曲線を通じて予後を予測することである。例えば、前記分子亜型は、全体生存率の側面で良い予後群、中間予後群および悪い予後群に分類される。具体的に、胃癌分子亜型が炎症 (Inflammatory) 分子亜型である場合、全体生存率の側面で良い予後群；腸 (Intestinal) 分子亜型および胃 (Gastric) 分子亜型である場合、中間予後群；混合型間質 (Mixed-stromal) 分子亜型および間葉 (Mesenchymal) 分子亜型である場合、悪い予後群として予測することである。

【0046】

本発明の胃癌 2 期および 3 期の予後予測用組成物は、薬剤学的に許容可能な担体をさらに含むことができる。

20

【0047】

前記薬剤学的に許容可能な担体は、医薬分野において通常使用される担体およびピヒクルを含み、具体的にイオン交換樹脂、アルミナ、アルミニウムステアレート、レシチン、血清蛋白質 (例、ヒト血清アルブミン)、緩衝物質 (例えば、各種燐酸塩、グリシン、ソルビン酸、カリウムソルベート、飽和植物性脂肪酸の部分的なグリセリド混合物)、水、塩または電解質 (例、プロタミンサルフェート、リン酸水素二ナトリウム、リン酸水素カリウム、塩化ナトリウムおよび亜鉛塩)、膠質性シリカ、マグネシウムトリシリケート、ポリビニルピロリドン、セルロース系基質、ポリエチレングリコール、ナトリウムカルボキシメチルセルロース、ポリアリレート、ワックス、ポリエチレングリコールまたは羊毛脂などを含むが、これに制限されない。

30

【0048】

また、本発明の組成物は、前記成分以外に、潤滑剤、湿潤剤、乳化剤、懸濁剤、または保存剤などをさらに含むことができる。

【0049】

また、本発明は、胃癌 2 期および 3 期の予後予測用組成物を含む胃癌 2 期および 3 期の予後予測用キットに関する。

【0050】

好ましくは、前記キットは、qPCR (Quantitative real-time polymerase chain reaction) キットなどであってもよい。

40

【0051】

前記胃癌 2 期および 3 期の予後予測用キットは、分析方法に適合した一種類またはそれ以上の他の構成成分組成物、溶液または装置をさらに含むことができる。好ましくは、前記診断用キットは、qPCR 反応を行うために必要な必須要素をさらに含むことができる。qPCR キットは、マーカー蛋白質を暗号化する遺伝子に対して特異的なプライマー対を含む。プライマーは、前記遺伝子の核酸配列に特異的な配列を有するヌクレオチドであって、約 7 bp ~ 50 bp の長さ、より好ましくは、約 10 bp ~ 30 bp の長さを有することができる。また、対照群遺伝子の核酸配列に特異的なプライマーを含むことができる。その他、qPCR キットは、テストチューブまたは他の適切な容器、反応バッファー

50

(pHおよびマグネシウム濃度は多様)、デオキシヌクレオチド(dNTPs)、Taq-ポリメラーゼおよび逆転写酵素のような酵素、DNase、RNase抑制剤DEPC-水(DEPC-water)、滅菌水などを含むことができる。

【0052】

また、本発明の胃癌2期および3期の予後予測用キットは、DNAチップを行うために必要な必須要素を含むことができる。DNAチップキットは、遺伝子またはその断片に該当するcDNAまたはオリゴヌクレオチド(oligonucleotide)が付着している基板、および蛍光標識プローブを製作するための試薬、製剤、酵素などを含むことができる。また、基板は、対照群遺伝子またはその断片に該当するcDNAまたはオリゴヌクレオチドを含むことができる。

10

【0053】

また、本発明は、統計的有意値を示すことができる程度のサンプル数を有する胃癌2期および3期の進行性胃癌患者から得た基準サンプルと生物学的サンプルで、

TFF1、TFF2およびVSI G1からなる胃(gastric)シグネチャー；CNN1、NEXN、SCRG1、SORBS1およびSPARCL1からなる間葉(mesenchymal)シグネチャー；AURKA、BUB1、CDC20、CEP55、PTTG1およびUBE2Cからなる増殖(proliferative)シグネチャー；CD8A、GBP1、GBP5、GZMB、NKG7およびWARSからなる免疫(immune)シグネチャー；ANTXR1、SFRP4およびVCANからなる幹細胞様(stem-like)シグネチャー；およびCDH17、CDX1およびMYO1Aからなる腸(intestinal)シグネチャーを含む標的遺伝子群と、ACTB、ATP5E、GPX1、UBBおよびHPRT1を含む参考遺伝子群のmRNA発現水準を測定する段階；

20

下記の式1により基準サンプルと生物学的サンプルの標的遺伝子群のCq値を計算して、コンピュータプログラムに入力する段階；および

前記コンピュータプログラムに入力した値に対してNMF(Non-negative Matrix Factorization)およびNMF基盤クラスタリングを行い、複数個のクラスターに分類し、各クラスターで標的遺伝子群のスコア

【数20】

30

(*d<sub>ik</sub>*)

を下記の式2に適用してSV(Score Value)を計算し、SVによって腸(Intestinal)分子亜型、炎症(Inflammatory)分子亜型、間葉(Mesenchymal)分子亜型、混合型間質(Mixed-stromal)分子亜型および胃(Gastric)分子亜型に分類し、全体生存率(Overall Survival)の側面で予後を分析して、生物学的サンプルが属する分子亜型の予後を予測する段階を含み、

40

前記胃癌の分子亜型は、胃(gastric)シグネチャーのSVが最大値を有するクラスターを胃(Gastric)分子亜型に定め；胃(Gastric)分子亜型に決定されたクラスターを除いたクラスターのうち間葉(mesenchymal)シグネチャーのSVが最大値を有し、増殖(proliferative)シグネチャーのSVが最小値を有するクラスターを間葉(Mesenchymal)分子亜型に定め；胃(Gastric)分子亜型および間葉(Mesenchymal)分子亜型に決定されたクラスターを除いたクラスターのうち免疫(immune)シグネチャーのSVが最大値を有し、腸(intestinal)シグネチャーのSVが最小値を有するクラスターを炎症(

50

Inflammatory) 分子亜型に定め；胃 (Gastric) 分子亜型、間葉 (Mesenchymal) 分子亜型および炎症 (Inflammatory) 分子亜型に決定されたクラスターを除いたクラスターのうち幹細胞様 (stem-like) シグネチャーの SV が最大値を有する場合、混合型間質 (Mixed-stromal) 分子亜型に定め；最終的に残ったクラスターを腸 (Intestinal) 分子亜型に定めて分類し、

前記胃癌の分子亜型が炎症 (Inflammatory) 分子亜型である場合、全体生存率の側面で良い予後群；腸 (Intestinal) 分子亜型および胃 (Gastric) 分子亜型である場合、中間予後群；混合型間質 (Mixed-stromal) 分子亜型および間葉 (Mesenchymal) 分子亜型である場合、悪い予後群として予測する、胃癌 2 期および 3 期の予後予測のための情報を提供する方法、または胃癌 2 期および 3 期の予後予測方法に関する：

10

【0054】

【数21】

[式1]

$$\Delta C_q = (\text{標的遺伝子の } C_q \text{ 値}) - (\text{参考遺伝子群の } C_q \text{ 平均値})$$

【0055】

ここで、参考遺伝子群の  $C_q$  平均値は、ACTB、ATP5E、GPX1、UBB および HPR1 を含む参考遺伝子群の  $C_q$  値の平均値を意味する。

20

【0056】

【数22】

[式2]

$$SV (\text{Score Value}) = \frac{1}{t} \sum_{i \in SN_\theta} d'_{ik}$$

30

【0057】

ここで、SV は、NMF 基盤クラスタリングに得たクラスターで各シグネチャーの発現平均値、 $t$  は、当該シグネチャーに属する遺伝子 ( $i$ ) の数、

【数23】

$SN_\theta$

40

は、シグネチャー ( $\theta = 6$ )、 $k$  は、クラスターの数であって、2 ~ 7 の整数であり

【数 2 4】

$$d'_{ik}$$

は、各遺伝子の全体平均と各クラスター平均の間の距離をベースとするスコアであり、下記の式 3 により求める。

【0058】

10

【数 2 5】

[式 3]

$$d'_{ik} = \text{sign}(d_{ik})(|d_{ik}| - \Delta)_+$$

【0059】

20

ここで、臨値 ( $\Delta$ ) は、0 . 1 に定めて、分子亜型別の特異性がない遺伝子は、0 に収斂させ、

【数 2 6】

$$\text{sign}(d_{ik})$$

は、

30

【数 2 7】

$$d_{ik}$$

の符号を意味し、

【数 2 8】

40

$$d_{ik}$$

は、下記の式 4 により求める。

【0060】

【数 2 9】

[式 4]

$$d_{ik} = \frac{\bar{x}_{ik} - \bar{x}_i}{m_k(s_i + s_0)}$$

$$s_i^2 = \frac{1}{n-k} \sum_{k=1}^5 \sum_{j \in C_k} (x_{ij} - \bar{x}_{ik})^2$$

10

【0 0 6 1】

ここで、各分子亜型 (k) 内の当該遺伝子 (i) 平均値

【数 3 0】

$$\bar{x}_{ik} = \sum_{j \in C_k} \frac{x_{ij}}{n_k}$$

20

当該遺伝子 (i) の全体平均値

【数 3 1】

$$\bar{x}_i = \sum_{j=1}^n \frac{x_{ij}}{n}$$

30

であり、

【数 3 2】

$$m_k$$

40

は、

【数 3 3】

$$\bar{x}_{ik} - \bar{x}_i$$

の標準誤差を補正するための自由度

50

【数 3 4】

$$m_k = \sqrt{\frac{1}{n_k} + \frac{1}{n}}$$

であり、

【数 3 5】

 $s_i$ 

は、遺伝子 ( i ) の全体サンプルに対する当該分子型 ( k ) 内の標準偏差であり、

【数 3 6】

 $s_0$ 

は、

【数 3 7】

 $s_i$ 

の中央値である。

【0062】

本発明の胃癌 2 期および 3 期の予後予測のための情報を提供する方法を段階別に具体的に説明すれば、次のとおりである。

【0063】

第 1 段階で、胃癌 2 期および 3 期腫瘍で得た一定の数の基準サンプルで標的遺伝子群の mRNA の発現水準を測定し、胃癌 2 期および 3 期腫瘍で得た生物学的サンプルの標的遺伝子群の mRNA の発現水準を測定して、基準サンプルと生物学的サンプルの標的遺伝子群の mRNA の発現水準に該当する値をコンピュータプログラムに入力する段階である。

【0064】

前記一定の数、すなわち、統計的有意値を示すことができる程度の基準サンプル数は、標的遺伝子群の mRNA 発現水準に対して NMF ( Non - negative Matrix Factorization ) を進める場合、有意値の p 値が 0 . 01 未満になり得るサンプル数を意味する。

【0065】

また、前記基準サンプル数は、NMF 基盤クラスタリングによって差別化された遺伝子発現様相を示しながら、複数個のクラスターに分類され得る程度の数でなければならない。

【0066】

10

20

30

40

50

このような条件を充足する基準サンプル数は、好ましくは、300～10,000であってもよい。

【0067】

前記標的遺伝子群のmRNA発現水準は、好ましくは、qPCRを通じて測定することができ、qPCRによるmRNA発現水準は、Cq値に定める。前記Cq値は、95で10分（初期変性）；95で10秒、40～45回（変性）、60で5秒（アニーリング）および72で25秒（延長）のPCR過程中に増幅が明確に増加し始めたサイクルの数値をいう。

【0068】

標的遺伝子群と参考遺伝子群のCq平均値を利用して下記の式1によりCq値を計算する。

10

【0069】

【数38】

[式1]

$$\Delta Cq = (\text{標的遺伝子の} Cq \text{ 値}) - (\text{参考遺伝子群の} Cq \text{ 平均値})$$

【0070】

ここで、参考遺伝子群のCq平均値は、ACTB、ATP5E、GPX1、UBBおよびHPRT1を含む参考遺伝子群のCq値の平均値を意味する。

20

【0071】

したがって、コンピュータプログラムに入力する標的遺伝子群のmRNA発現水準に該当する値は、前記Cq値を意味する。

【0072】

第2段階は、コンピュータプログラムに入力された基準サンプルと生物学的サンプルの標的遺伝子群のCq値に対してNMFおよびNMF基盤クラスタリングを行い、複数個のクラスターに分類し、各クラスターで標的遺伝子群のスコア

【数39】

30

( $d_{ik}$ )

を下記の式2に適用してSV (Score Value) を計算し、SVによって腸 (Intestinal) 分子亜型、炎症 (Inflammatory) 分子亜型、間葉 (Mesenchymal) 分子亜型、混合型間質 (Mixed-stromal) 分子亜型および胃 (Gastric) 分子亜型に分類し、全体生存率 (Overall Survival) の側面で予後を分析して生物学的サンプルが属する分子亜型の予後を予測する段階である。

40

【0073】

【数 4 0】

[式 2]

$$SV (\text{Score Value}) = \frac{1}{t} \sum_{i \in SN_{\theta}} d'_{ik}$$

10

【0074】

ここで、SVは、NMF基盤クラスタリングに得たクラスターで各シグネチャーの発現  
 平均値、tは、当該シグネチャーに属する遺伝子(i)の数、

【数 4 1】

$$SN_{\theta}$$

20

は、シグネチャー( = 6)、kは、クラスターの数であって、2~7の整数であり

【数 4 2】

$$d'_{ik}$$

は、各遺伝子の全体平均と各クラスター平均の間の距離をベースとするスコアである。

【0075】

30

本発明によれば、前記kは、好ましくは、5であってもよい。すなわち、NMF基盤ク  
 ラスタリングを通じて得た5個のクラスターを意味する。

【0076】

前記スコア

【数 4 3】

$$(d'_{ik})$$

40

は、各遺伝子の全体平均と各クラスター平均の間の距離をベースとするスコア、すなわ  
 ち、当該遺伝子(i)のt-統計値であり、同じ分子亜型(subtype)内のサンプ  
 ル遺伝子発現量に加重値を与えて標準化するもので、線形判別分析(LDA; Linea  
 r Discriminant Analysis)で頻繁に使用される統計方法であっ  
 て、下記の式3により求める：

【0077】

【数 4 4】

[式 3]

$$d'_{ik} = \text{sign}(d_{ik})(|d_{ik}| - \Delta)_+$$

10

【0078】

ここで、臨界値 ( ) は、0 . 1 に定めて、分子垂型別の特異性がない遺伝子は、0 に収斂させ、

【数 4 5】

$$\text{sign}(d_{ik})$$

20

は、

【数 4 6】

$$d_{ik}$$

の符号を意味し、

【数 4 7】

30

$$d_{ik}$$

は、下記の式 4 により求める：

【0079】

【数 4 8】

40

[式 4]

$$d_{ik} = \frac{\bar{x}_{ik} - \bar{x}_i}{m_k(s_i + s_0)}$$

$$s_i^2 = \frac{1}{n-k} \sum_{k=1}^5 \sum_{j \in C_k} (x_{ij} - \bar{x}_{ik})^2$$

50

【 0 0 8 0 】

ここで、各分子型 ( k ) 内の当該遺伝子 ( i ) 平均値

【 数 4 9 】

$$\bar{x}_{ik} = \sum_{j \in C_k} \frac{x_{ij}}{n_k}$$

10

当該遺伝子 ( i ) の全体平均値

【 数 5 0 】

$$\bar{x}_i = \sum_{j=1}^n \frac{x_{ij}}{n}$$

20

であり、

【 数 5 1 】

$$m_k$$

30

は、

【 数 5 2 】

$$\bar{x}_{ik} - \bar{x}_i$$

の標準誤差を補正するための自由度

【 数 5 3 】

40

$$m_k = \sqrt{\frac{1}{n_k} + \frac{1}{n}}$$

であり、

50

【数 5 4】

 $S_i$ 

は、遺伝子 (  $i$  ) の全体サンプルに対する当該分子亜型 (  $k$  ) 内の標準偏差であり、  
【数 5 5】

10

 $S_0$ 

は、  
【数 5 6】

 $S_i$ 

20

の中央値である。  
【0081】  
前記スコア  
【数 5 7】

 $(d_{ik})$ 

30

で正数は、遺伝子の高発現、負数は、遺伝子の低発現、0は、遺伝子の発現変化がないことを意味する。したがって、正数の値が大きいほど発現がさらに高くなり、負数の値が大きいほど発現はさらに低くなる。

【0082】

前記胃癌の分子亜型は、胃 (  $g a s t r i c$  ) シグネチャーの  $S V$  が最大値を有するクラスターを胃 (  $G a s t r i c$  ) 分子亜型に定め；胃 (  $G a s t r i c$  ) 分子亜型に決定されたクラスターを除いたクラスターのうち間葉 (  $m e s e n c h y m a l$  ) シグネチャーの  $S V$  が最大値を有し、増殖 (  $p r o l i f e r a t i v e$  ) シグネチャーの  $S V$  が最小値を有するクラスターを間葉 (  $M e s e n c h y m a l$  ) 分子亜型に定め；胃 (  $G a s t r i c$  ) 分子亜型および間葉 (  $M e s e n c h y m a l$  ) 分子亜型に決定されたクラスターを除いたクラスターのうち免疫 (  $i m m u n e$  ) シグネチャーの  $S V$  が最大値を有し、腸 (  $i n t e s t i n a l$  ) シグネチャーの  $S V$  が最小値を有するクラスターを炎症 (  $I n f l a m m a t o r y$  ) 分子亜型に定め；胃 (  $G a s t r i c$  ) 分子亜型、間葉 (  $M e s e n c h y m a l$  ) 分子亜型および炎症 (  $I n f l a m m a t o r y$  ) 分子亜型に決定されたクラスターを除いたクラスターのうち幹細胞様 (  $s t e m - l i k e$  ) シグネチャーの  $S V$  が最大値を有する場合、混合型間質 (  $M i x e d - s t r o m a l$  ) 分子亜型に定め；最終的に残ったクラスターを腸 (  $I n t e s t i n a l$  ) 分子亜型に定めて分類

40

50

する。

【0083】

前記で分類された分子亜型に対して各分子亜型の生存曲線を通じて全体生存率の側面で予後を予測する。

【0084】

前記胃癌の分子亜型は、生存曲線で差別化された予後様相を示し、炎症 (Inflammatory) 分子亜型である場合、全体生存率の側面で良い予後群；腸 (Intestinal) 分子亜型および胃 (Gastric) 分子亜型である場合、中間予後群；混合型間質 (Mixed-stromal) 分子亜型および間葉 (Mesenchymal) 分子亜型である場合、悪い予後群として予測することができる。

10

【0085】

したがって、生物学的サンプルが属する分子亜型を定め、前記分子亜型の生存曲線を通じて予後を予測することができる。

【0086】

前記生物学的サンプルは、新鮮腫瘍組織、新鮮凍結腫瘍組織、パラフィン包埋腫瘍組織、微細針吸引液、腹水、管洗浄液または胸膜液などであってもよく、好ましくは、パラフィン包埋腫瘍組織であってもよい。

【0087】

また、前記標的遺伝子群の mRNA の発現水準の測定は、qPCR によって行われ得る。

20

【0088】

本発明の利点および特徴、そしてそれらを達成する方法は、詳細に後述されている実施例を参照すれば明確になるだろう。しかし、本発明は、以下で開示される実施例に限定されるものではなく、互いに異なる多様な形態で具現されるものであり、単に本実施例は、本発明の開示が完全になりようにし、本発明の属する技術分野における通常の知識を有する者に発明の範疇を完全に知らせるために提供されるものであり、本発明は、請求項の範疇によって定義されるだけである。

【0089】

<実施例>

(患者およびサンプル)

2000年から2010年まで延世大学セブランス病院で1次治療時に胃癌切除手術を受けた胃癌患者の新鮮-凍結腫瘍標本および臨床データを確保した。すべてのサンプルは、患者から書面同意を得た後に収集され、研究は、YUSHの機関検討委員会の承認を受けた。サンプルに注釈を付けたが、患者識別が可能な情報からそれらを分離した。後ろ向きに臨床データを得た。OS (overall survival) は、手術から死亡までの時間として定義し、再発ない生存は、手術後に最初の再発までの時間として定義した。最後の接触時に再発なしに患者が生きているとき、データが検閲された。

30

【0090】

48,803個の遺伝子特徴を有する Human HT-12 v3.0 Expression Bead Chip array (Illumina) を使用して外科的に除去された凍結 GC 腫瘍組織の 497 個のサンプルからの遺伝子発現プロファイルを測定した (GSE13861p)。簡単に話せば、RecoverAll (登録商標) total nucleic acid isolation kit (Ambion) または miRvana RNA Isolation Labeling Kit (Ambion) を使用して新鮮-凍結組織でトータル RNA を抽出した。RNA 濃度と純度は、NanoDrop 2000 (Thermo Fischer Scientific) を使用して 260nm と 280nm (A260:A280 = 1.8) で測定した。RNA の完全性 (integrity) は、RNA Nano 6000 チップ (Agilent) (RIN > 7) を使用して評価した。TotalPrep (登録商標) RNA Amplification Kit (Illumina) を使用して製造業者のプロトコルにより 5

40

50

00 ngのトータルRNAを標識した後、BeadChipアレイプラットフォームを使用して遺伝子発現水準を測定した。

【0091】

(GCランニングのためのトレーニングセットおよびテストセット)

GC分子垂型を探すためのトレーニングセットIは、GSE13861p (n=497、Illumina HumanHT-12 v3.0 Expression BeadChipアレイ)で構成した。これを確認するためのテストセットIは、GSE15459 (n=200、Affymetrix Human Genome U133 plus 2.0 Array)、TCGA (n=262、Illumina HiSeq2000)およびGSE62254 (n=300、Affymetrix Human Genome U133 plus 2.0 Array)のデータセットで構成した。

10

【0092】

最終miniClassifier-26 (26個の遺伝子)を選定するために、トレーニングセットIIは、GSE13861p (n=497)、GSE15459 (n=200)、TCGA (n=262)、GSE62254 (n=300)をComBat方法で併合されたMerged1259 (n=1259)で構成し、qPCR基盤測定可能な26個の遺伝子は、qPCR325 (n=325)をテストセットIIで構成して確認した。qPCR325を得るために、MasterPure (登録商標) Complete DNAとRNA Purification Kit (Epicenter)を使用してトータルRNAを抽出した。cDNAは、M-MLV Reverse Transc

20

【0093】

(データ加工)

データ前加工：マイクロアレイデータセットは、主にR言語環境で処理された。正規化は、Illumina BeadChipアレイプラットフォームのデータセットに対する「マイクロアレイデータ用線形モデル (limma)」パッケージのBetween-Array Normalization (quantile)により行われた。GSE15459およびGSE62254は、堅固な多重アレイ平均正規化を含むR「affy」パッケージを使用して標準化された。qPCRデータセットは、内部標準によって標準化された。遺伝子フィルタリングのためには、プラットフォーム由来プローブ有効性と遺伝子発現の差異を考慮しなければならない。トレーニングセットの場合、データセットを一括調整した後、「遺伝子発現マイクロアレイデータ (ComBat)の配置を結合するとき、一括処理効果防止」方法と結合した。

30

【0094】

NMF基盤分類：ComBat併合データは、「Nonnegative Matrix Factorization (NMF) アルゴリズムおよびフレームワーク」パッケージを使用して分類された。クラスター数kは、2から7に設定された。Brunet方法は、反復近似のためのアップデートアルゴリズムで使用された。NMF派生クラスターを特性化する前に、「Silhouette」Rパッケージを使用して各クラスターで異常値サンプルを除いた。それぞれのNMFクラスターを代表する遺伝子を定義するために、SAMおよびPAMは、それぞれBioconductorパッケージ「siggenes」および「pamr」を使用して行われた。

40

【0095】

WGCNA分析：WGCNAは、R「wgcna」パッケージを使用して497GCサンプルで行われた。ネットワーク構成のために加重ネットワーク隣接性は、6の自乗で表

50

現された類似性によって定義される。任意カットオフ選択を避けるために、WGCNAで提供した「ソフト臨界設定手続」に従った。モジュール-探知方法でR「dynamicTreeCut」パッケージの動的混成化が適用された。各モジュールが生存および臨床病理学的変数と関連があるかを評価するために、連結性、モジュール重要性およびモジュールeigen geneのような多様なモジュール特性が活用された。

【0096】

Gene set enrichment analysis (GSEA): GSEAはR「GSEABase」パッケージを使用して行われた。事前定義された遺伝子セットは、Molecular Signatures Database (MSigDB; [www.broadinstitute.org/msigdb](http://www.broadinstitute.org/msigdb)) から持ってきた。エンリッチメント分析は、Gene Ontology Consortium (<http://geneontology.org>) の遺伝子オントロジーを使用して行われた。

10

【0097】

生存分析: 生存分析は、R「survival」および「meta」パッケージでHRのCox比例危険モデルおよびメタ分析を使用して行われた。

【0098】

遺伝子発現プロファイルの階層的クラスタリングおよびイラストレーション: マイクロアレイデータセットの遺伝子クラスタリングは、Gene Cluster 3.0を使用して行われた。階層的クラスタリング結果は、R「gplots」パッケージを使用して視覚的に視覚化された。

20

【0099】

(統計的分析)

統計的試験として超幾何分布検定/Fisher's exact test、Pearson's correlation、Spearman's correlationおよびWilcoxon rank-sum testを使用した。

【0100】

(プライマリ細胞 - 特異的遺伝子発現プロファイリング)

プライマリ細胞培養: 一次組織を2%抗生剤 (Welgene LS203-01) を含有したDulbecco's phosphate-buffered saline (Welgene LB00-02) で洗浄し、滅菌ブレードで細切れにした。0.2 μm注射器で濾過した後、細切れにされた組織をMEM (Gibco A10490) および150 U/mL Collagenase I (Thermo Fisher Scientific) と37 °Cで24時間5% CO<sub>2</sub>の加湿の雰囲気中で培養した。培養された組織を200 × gで5分間遠心分離した後、新鮮な培地に移した。収穫された細胞を5% CO<sub>2</sub>大気下で37 °Cで2~3日間培養した。

30

【0101】

Illumina HiSeq 2500シーケンシングシステムを利用したトータルRNAシーケンシング: RNA純度は、NanoDrop 8000分光光度計で1 μLのトータルRNA抽出物を分析して決定した。トータルRNA完全性は、RNA完全性番号 (RIN) 値とRNA断片 > 200 nt断片分配値 (DV200) の百分率でAgilent Technologies 2100 Bioanalyzerを使用して確認した。トータルRNAシーケンシングライブラリーは、メーカーの指針 (Illumina TruSeq RNA Access Library kit) によって準備された。その後、100 ngのトータルRNAを高温で2価カチオンを使用して小さい断片で分裂させた。第1および第2ストランド合成の間に、ランダムプライミングを使用して切断されたRNA断片からcDNAを生成し、シーケンシングアダプタを生成された二本鎖cDNA断片にライゲーションさせた。転写体のコーディング領域は、最終ライブラリーを生成するために配列特異的プローブを使用してこのライブラリーからキャプチャーされた。増幅されたライブラリーの品質は、毛細管電気泳動 (Bioanalyzer, Agilent) により確認された。SYBR Green PCR Master Mix (

40

50

Applied Biosystems) を使用して qPCR 後、同量のモルのタグがついた索引されたライブラリーをプールに結合させた。クラスターの生成は、cBot 自動化クラスター生成システム (Illumina) のフローセルで実現された。引き続き、前記フローセルを HiSeq 2500 シーケンシングシステム (Illumina) 上にローディングさせ、配列分析は、 $2 \times 100$  bp 読み取り長さを使用して行った。

#### 【0102】

(インビトロおよびインビボ実験検証)

細胞株：ヒト胃癌細胞株 SNU - 1、SNU - 5、SNU - 16、SNU - 216、SNU - 484、SNU - 520、SNU - 601、SNU - 620、SNU - 638、SNU - 668、SNU - 719、MKN、MKN - 45、MKN - 74、KATO III、NCI - N87 および Hs 746 T は、韓国細胞株銀行 (ソウル、韓国) で購入した；また、YCC - 1、YCC - 2、YCC - 3、YCC - 6、YCC - 7、YCC - 9、YCC - 10、YCC - 11 および YCC - 16 は、延世癌研究所 (ソウル、韓国) で購入した。SNU - 1、SNU - 5、SNU - 16、SNU - 216、SNU - 484、SNU - 520、SNU - 601、SNU - 620、SNU - 638、SNU - 668、SNU - 719、MKN - 28、MKN - 45、MKN - 74、KATO III および NCI - N87 は、RPMI 1640 (Welgene、大邱、韓国) で成長させた；Hs 746 T は、Dulbecco's modified Eagle's medium (DMEM；Welgene、大邱、韓国) で成長させた；また、YCC - 1、YCC - 2、YCC - 3、YCC - 6、YCC - 7、YCC - 9、YCC - 10、YCC - 11、および YCC - 16 は、Minimum essential media Eagle (MEM；Welgene、大邱、韓国) で成長させた。すべての細胞を  $37^{\circ}\text{C}$  で 10% FBS (Gibco) と 1% 抗生剤 - 抗菌溶液 (10,000 単位ペニシリン、10 mg ストレプトマイシン、 $25 \mu\text{g}$  アムホテリシン B / mL を含む、Sigma - Aldrich) が補充された完全培地で 5%  $\text{CO}_2$  を含有する加湿の雰囲気中で培養した。すべての細胞は、e - Myco (登録商標) と Mycoplasma PCR Detection Kit (iNtRON Biotechnology、城南、韓国) によりマイコプラズマに対して陰性であると確認された。

10

20

#### 【0103】

浸湿分析：分析のために培養培地 (M199) に  $2 \times 10^4$  HUV EC 細胞をピプロネクティンがコートされたトランスウェルに入れ、底を 0.2% ゼラチンでコートした後、単層の形成まで 48 時間培養した。次に、FBS が無い Cell Tracker (登録商標) (Molecular Probes、C2925) が含まれた  $1 \times 10^5 / 50 \mu\text{L}$  の Hs 746 T および NCI - N87 細胞を別にトランスウェルに添加した。10% FBS を含有する培養培地を下部チャンパーに添加した。48 時間培養した後、メンブレインの上部細胞を綿棒で除去した。下部メンブレインの細胞を室温で 2 ~ 3 時間  $200 \mu\text{L}$  の溶解バッファーで溶解させた。Ex / Em 492 / 517 で蛍光を測定した。TGF - 抑制剤が細胞の侵入能力に及ぼす影響を調べるために、LY 2157299 (AdooQ、カリフォルニア、米国) を  $50 \mu\text{M}$  投与した。

30

#### 【0104】

移動分析：Hs 746 T および NCI - N87 細胞を 10% FBS および 1% 抗生剤が含有された培養培地で単層に成長させた。密集度 (confluency) が 70% に到達すれば、細胞単層を  $100 \mu\text{L}$  のピペットの先でかき出した。72 時間後に創傷の幅を測定し、かき出した直後に測定した創傷の幅で標準化した。TGF - 抑制剤が細胞移動に及ぼす影響を評価するために、LY 2157299 ( $50 \mu\text{M}$ ) を投与した。

40

#### 【0105】

腫瘍スフェロイド形成分析：96 - ウェルプレートで、10 個の細胞を bFGF、EGF、B27、10% FBS および 1% 抗生剤が添加された  $50 \mu\text{L}$  の DMEM / F12 (Gibco) で培養した。培養 30 日後、各ウェルで腫瘍スフェロイドを計数した。また、LY 2157299 ( $50 \mu\text{M}$ ) を投与して、TGF - 抑制剤が腫瘍スフェロイド形

50

成に及ぼす影響を調査した。

【0106】

同所マウスモデルでインビボ腫瘍形成：すべての動物実験は、国際実験動物管理評価認証協会（A A A L A C）の承認を受けて行われた。同所移植マウスモデルを確立するために、約1時間略5mmで上部の正中線に沿って皮膚と腹膜を切開して、BALB/cヌードマウス（男性）の外装型胃壁に $1 \times 10^7$  GC細胞（Hs746TおよびNCI-N87）を移植した。胃を腹膜に戻して、腹壁を1層の創傷縫合で閉鎖した。モデルで腫瘍成長を観察するために、Bruker動物性コイル（RF SUC 400 1H M-BR-LIN ROAD、Bruker Medical Systems）が装着された9.4T動物磁気共鳴映像（MRI）装置を使用して癌の有無およびサイズを追跡した（磁気共鳴映像測定条件：Echo = 1、TR = 2300ms、TE = 22.0ms、FA = 180deg、TA = Oh 4m 54s 400ms、NEX = 2、およびFOV = 4.00cm）。

10

【0107】

移植マウスモデルで薬物反応：異種移植マウスモデルを確立するために、BALB/cヌードマウス（雄性）の近位大腿部位に $1 \times 10^7$  GC細胞（Hs746TおよびNCI-N87）を移植した。腫瘍体積が $400 \text{ mm}^3$ に増加するとき、腫瘍保有マウスを3個の他の治療群（PBS対照群、Oxal + 5FU / PBS処理群およびOxal + 5FU / LY2157299（TBF 阻害剤）処理群、各群当たり $n = 8$ ）で無作為抽出した。オキサリプラチン（単一用量当たり $60 \mu\text{g}$ ）とフルオロウラシル（単一用量当たり1mg）を混合して、マウスに一週間に3回腹腔注射した。LY2157299（1.5mg / マウス）を週2回腫瘍内注射モデルに投与した。移植された腫瘍のサイズは、週当たり3回検査し、腫瘍のサイズは、 $(4/3) \times \times (\text{短軸}/2)^2 \times (\text{長軸}/2) \text{ mm}^3$ で計算した。

20

【0108】

qPCR分析のためのminiClassifier-26の異質性検査：空間的異質性がFFPE標本で組織サンプルを準備する本発明の手に影響を及ぼす可能性を調査するために、RNA抽出のために単一FFPE標本で3個の組織サンプルを得た。cDNAを製造した後、qPCRを3回施行して、単一腫瘍標本で3個の各標本に対して平均を求めた。各遺伝子に対する分散係数を評価して、三種類の標本の異質性を決定した。

30

【0109】

（組織マイクロアレイ構築）

各ホルマリン固定、パラフィン内蔵基本腫瘍で代表的な3mm直径腫瘍組織コア2個を組織マイクロアレイ（TMA）ブロックで組み立てた。各TMAブロックは、標識および内部統制として14個の腫瘍および一つの正常胃粘膜組織コアを含有する。引き続き、免疫組織化学（IHC）分析のために、各TMAブロックから $4 \mu\text{m}$ 厚さのセクションを準備した。

【0110】

（免疫組織化学分析）

IHCは、MutLホモログ1（MLH1、使用準備、Roche、Basel、Switzerland）、MutS蛋白質ホモログ2（MSH2、使用準備、複製G219-1129）用抗体があるVentanaXTシステム（Ventana Corporation）、Roche）、IHCは、前述したように行った。MLH1およびMSH2の場合、腫瘍細胞で核染色がないものは発現消失であり、正常発現は、腫瘍細胞での核発現の存在として定義された。すべてのIHC結果は、臨床病理学的特徴に対する知識なしに評価された。

40

【0111】

（エプスタイン・バーウイルス - 暗号化されたRNAインシチュ（insitu）ハイブリダイゼーション（EBER IISH））

EBER IISHは、Ventana Bench Markシステム（IISH iV

50

iewキット、Ventana Corporation、AZ、米国)で行った。パラフィン包埋された組織切片をEZ Prepバッファー(Ventana Corporation)で脱パラフィン化させ、プロテアーゼIで4分間分解させた。引き続いて、EBERに対するプローブを85℃で10分間変性させた後、37℃で1時間混成化させた。混成化した後、組織を57℃で2×SSCバッファーで洗浄した。引き続いて、抗フルオレシン単クローン抗体とのインキュベーションを20分間行った後、アルカリブルー検出キット(Ventana Corporation)をメーカーのプロトコルによって使用した。スライドを10分間ニュークリア・ファスト・レッドで対照染色した。

【0112】

【表 1 - 1】

表 1 標的遺伝子群および参考遺伝子群、プライマーセットおよびプローブのリスト

No.	遺伝子名	プライマー番号	登録番号	産物サイズ	start	stop	プライマー/プローブ配列 (SEQ ID NO.) (5' -3' )	
1	TFF1	N643-2	NM_003225	75	196	270	正方向 プライマー ー	aaataagggc tgctgtttcg (SEQ ID NO: 1)
							逆方向 プライマー ー	gggacgtcga tggtattagg (SEQ ID NO: 2)
							プローブ	acgacaccgt tcgtggggtc (SEQ ID NO: 63)
2	TFF2	N644-2	NM_005423	67	357	423	正方向 プライマー ー	ccctccaaa gcaagagtc (SEQ ID NO: 3)
							逆方向 プライマー ー	gggtagccac agtttcttcg (SEQ ID NO: 4)
							プローブ	tcagtgcgtc atggaggctc ca (SEQ ID NO: 64)
3	VSIG1	N649-1	NM_001170 553	64	812	875	正方向 プライマー ー	catcgtgcca gtgaaagaaa (SEQ ID NO: 5)
							逆方向 プライマー ー	tgtcagattt ccaatgacca a (SEQ ID NO: 6)
							プローブ	tcaaccaac caccgggatt (SEQ ID NO: 65)

10

20

30

40

【表 1 - 2】

4	CNN1	N185-5	NM_001299 .4	68	526	593	正方向 プライマ ー	agtccaccct cctggettt (SEQ ID NO: 7)
							逆方向 プライマ ー	cttcactecc acgttcacct t (SEQ ID NO: 8)
							プローブ	cctttcgtct tcgcatgct gg (SEQ ID NO: 66)
5	NEXN	N797-3	NM_144573 .3	66	1152	1217	正方向 プライマ ー	gcgccaaatg gtaaatgaag (SEQ ID NO: 9)
							逆方向 プライマ ー	gggcgtacc ctttaaaaat (SEQ ID NO: 10)
							プローブ	tgaggaaaac caagacacag caaa (SEQ ID NO: 67)
6	SCRG1	N710-2	NM_007281 .2	72	338	409	正方向 プライマ ー	cccagtgagt gtgacattt (SEQ ID NO: 11)
							逆方向 プライマ ー	gcttttgcc cttttcttc (SEQ ID NO: 12)
							プローブ	tggtcttggc agaggatgct tc (SEQ ID NO: 68)

10

20

30

40

【表 1 - 3】

7	SORBS 1	N835-3	NM_015385 .3	75	263	337	正方向 プライマ ー	gctgtgatga atggcttgg (SEQ ID NO: 13)
							逆方向 プライマ ー	cccagtcag attttttag g (SEQ ID NO: 14)
							プローブ	ttgtcttgcc cattgctgcc (SEQ ID NO: 69)
8	SPARC L1	N454-3	NM_004684 .2	69	401	469	正方向 プライマ ー	cattccaaac caactgetga (SEQ ID NO: 15)
							逆方向 プライマ ー	agcttcagcc cataaactgg (SEQ ID NO: 16)
							プローブ	cggtagcacc tgacaacact gc (SEQ ID NO: 70)
9	AURKA	N471-2	NM_198434 .1	69	1074	1142	正方向 プライマ ー	gcagattttg ggtggtcagt (SEQ ID NO: 17)
							逆方向 プライマ ー	gtagtccagg gtgccacaga (SEQ ID NO: 18)
							プローブ	ctccatcttc caggaggacc a (SEQ ID NO: 71)

10

20

30

40

【表 1 - 4】

10	BUB1	N245-1	NM_004336 .2	75	1776	1850	正方向 プライマ ー	ccttcaaac caaaggagga (SEQ ID NO: 19)	10
							逆方向 プライマ ー	gcagcgaata ccccataca (SEQ ID NO: 20)	
							プローブ	ccaaaaactc ttcagcatga ggca (SEQ ID NO: 72)	
11	CDC20	N504-3	NM_001255	71	607	677	正方向 プライマ ー	cttcctgcc agaccgtat (SEQ ID NO: 21)	20
							逆方向 プライマ ー	ccaatccaca aggttcaggt (SEQ ID NO: 22)	
							プローブ	cctggatgcg cctgaaatcc (SEQ ID NO: 73)	
12	CEP55	N731-2	NM_018131 .4	74	438	511	正方向 プライマ ー	caagtgggaa aggaaagctg (SEQ ID NO: 23)	30
							逆方向 プライマ ー	ctcagcctca aggactcgaa (SEQ ID NO: 24)	
							プローブ	ttttctcaa aagtctgtgt ctctc (SEQ ID NO: 74)	

【表 1 - 5】

13	PTTG1	N813-2	NM_004219 .3	67	166	232	正方向 プライマ ー	ctgaagctgg ggtctgga (SEQ ID NO: 25)	10
							逆方向 プライマ ー	aacgtggtgt tgaacctgga (SEQ ID NO: 26)	
							プローブ	ccttcaatca aagccttaga tggga (SEQ ID NO: 75)	
14	UBE2C	N716-3	NM_181802 .1	73	1021	1093	正方向 プライマ ー	ccctgctatc accccaac (SEQ ID NO: 27)	20
							逆方向 プライマ ー	gggcagacca cttttccttc (SEQ ID NO: 28)	
							プローブ	caccagggt aacatatgcc tgg (SEQ ID NO: 76)	
15	CD8A	N205-5	NM_171827 .3	70	1745	1814	正方向 プライマ ー	cagagctacc cgcagagttc (SEQ ID NO: 29)	30
							逆方向 プライマ ー	aagagttga gatggcatgg (SEQ ID NO: 30)	
							プローブ	tgcctccagc t ctctcagcat ga(SEQ ID NO: 77)	

【 0 1 1 3 】



【表 2 - 2】

19	NKG7	N705-2	NM_005601.3	66	641	706	正方向 プライマー	gtccccgtcctgg ctatg (SEQ ID NO: 37)
							逆方向 プライマー	aacgctcaaaact catcttgc (SEQ ID NO: 38)
							プローブ	cgctcttgccctc tgctcaca (SEQ ID NO: 81)
20	WARS	N717-3	NM_173701.1	73	408	480	正方向 プライマー	ttgtggaccatg gacagta (SEQ ID NO: 39)
							逆方向 プライマー	ccaaaccgaacaa tgagctt (SEQ ID NO: 40)
							プローブ	tgccttttgcaact gcttgctg (SEQ ID NO: 82)
21	ANTXR1	N722-2	NM_053034.2	67	558	624	正方向 プライマー	cagttggctcaca aatcacc (SEQ ID NO: 41)
							逆方向 プライマー	ttcctcgggtgga gaaaac (SEQ ID NO: 42)
							プローブ	aaaggacattctc aactgtgggc (SEQ ID NO: 83)

10

20

30

40

【表 2 - 3】

22	SFRP4	N187-3	NM_003014.2	64	1298	1361	正方向 プライマー	ggagacttccgac ttccttaca (SEQ ID NO: 43)
							逆方向 プライマー	tggecttacatag gctgtcc (SEQ ID NO: 44)
							プローブ	aggcaatgcc cagcctcatc (SEQ ID NO: 84)
23	VCAN	N595-3	NM_001126336	74	1830	1903	正方向 プライマー	tttgagcatgact tccgttg (SEQ ID NO: 45)
							逆方向 プライマー	ctgtctggctggt tgggtct (SEQ ID NO: 46)
							プローブ	tggcagcactg caatacga (SEQ ID NO: 85)
24	CDH17	N729-3	NM_004063.3	68	1882	1949	正方向 プライマー	gcaatgtgactgc caaggat(SEQ ID NO: 47)
							逆方向 プライマー	acctttgtgtct cccctca (SEQ ID NO: 48)
							プローブ	ccagaaggtctgg acataagc (SEQ ID NO: 86)

10

20

30

40

【表 2 - 4】

25	CDX1	N730-2	NM_001804.2	67	1319	1385	正方向 プライマー	agggaggaacgtg gtcaact (SEQ ID NO: 49)
							逆方向 プライマー	tatgatgggggca ggtagaa (SEQ ID NO: 50)
							プローブ	tgccctcttctgc agcctca (SEQ ID NO: 87)
26	MY01A	N793-3	NM_005379.3	71	1034	1374	正方向 プライマー	ccgcctctttgac tggatag (SEQ ID NO: 51)
							逆方向 プライマー	ccttcttcttttc cccgatg (SEQ ID NO: 52)
							プローブ	cccaccttgatgc tctcattgatc (SEQ ID NO: 88)
27	ACTB	N037	NM_001101	72	278	349	正方向 プライマー	tcaccctgaagta cccatc (SEQ ID NO: 53)
							逆方向 プライマー	tgtggtgccagat tttctcc (SEQ ID NO: 54)
							プローブ	cgcatcgtcacc aactggg (SEQ ID NO: 89)

10

20

30

40

【表 2 - 5】

28	ATP5E	N041	NM_006886	74	117	189	正方向 プライマー	atggtgcctact ggagaca (SEQ ID NO: 55)	10
							逆方向 プライマー	Ctctcactgcttt tgcacaga (SEQ ID NO: 56)	
							プローブ	tggactcagctac atccgatactccc a (SEQ ID NO: 90)	
29	GPX1	N201-5	NM_000581.2	71	308	378	正方向 プライマー	cccgtgcaaccag tttgg (SEQ ID NO: 57)	20
							逆方向 プライマー	ggacgtacttgag ggaattcaga (SEQ ID NO: 58)	
							プローブ	ctcttcgcttctg ggttct cctgatg (SEQ ID NO: 91)	
30	UBB	N203-5	NM_018955.2	78	61	138	正方向 プライマー	tgggtgagcttgt ttgtgtc (SEQ ID NO: 59)	30
							逆方向 プライマー	tttgacctgtag cggatacc (SEQ ID NO: 60)	
							プローブ	caccaaccacgtc cacccac (SEQ ID NO: 92)	

【表 2 - 6】

31	HPRT1	N049-2	NM_000194.1	67	531	597	正方向 プライマー	tggtcaggcagta taatccaa (SEQ ID NO: 61)
							逆方向 プライマー	cttcgtgggtcc ttttcac (SEQ ID NO: 62)
							プローブ	tgcaagcttgc gacctgacc (SEQ ID NO: 93)

10

## 【0114】

## &lt; 実施例 1 &gt; 進行性胃癌の 5 個の分類

本発明の実験フローチャートは、図 1 に示した。本発明者らは、コンセンサス基盤の NMF をベースに 5 個の GC 分子亜型を確認した。本発明者らは、GC (GSE13861 pトレーニングセット I; HumanHT-12 v3.0 Array (Illumina)) 患者で胃の切除手術サンプル (n = 497) の遺伝子発現プロファイルを調査した。5 個の分子亜型に対する GC の分類は、高い一致を示し、ヒートマップは、独特の遺伝子発現様相を示した (図 2 a および図 3)。本発明者らは、マイクロアレイ (SAM、false discovery rate (FDR) = 0) の有意性分析に引き続いて、マイクロアレイの予測分析 (PAM; overall error rate = 0.10) を使用して 932 個の亜型特異遺伝子を確認した (Classifier - PAM932)。

20

Classifier - PAM932 を使用して GC 患者の独立的な遺伝子発現データセット (Test set I) で分子亜型化を安定的に確認した。

30

## 【0115】

## &lt; 実施例 2 &gt; GC 亜型を記述する 6 個の分子シグネチャー確認

先験的に定義された遺伝子セットの表記法を拡張するために監督されない更なる遺伝子格別のクラスタリングを行った。GSE13861p で WGCNA は、32 個の遺伝子モジュール (高度で相互連結された遺伝子のクラスター) を検出した (図 4 a、b)。全体的に、(i) いくつかのモジュールがコホートで顕著に保存されたという事実が観察された (超幾何分布検定;  $P < 0.01$ ) (図 4 b)、(ii) 保存されたモジュールは、PAM genes の上位 25% PAM 分析の相対的な差異による (図 4 c)、(iii) このような関連性は、GC 生物学と有意な関連があった (図 4 d)。保存されたモジュールをベースに 6 個の GC シグネチャーは、5 個の GC 亜型と顕著に関連している。本発明者らは、GC シグネチャーの遺伝子を再抽出してスピアマンの相関関係を使用して GC シグネチャーの特定の組合せを 5 個の NMF 派生亜型に変換することができることを示した (図 4 e)。5 個の亜型に対する本発明者らの注釈は、ネットワーク分析に基づく亜型の特性により生物学的にさらに関連がある。

40

## 【0116】

## &lt; 実施例 3 &gt; 5 個の GC 亜型の臨床的特性糾明

GC 亜型と臨床病理学的情報 (年齢、性別、腫瘍位置、AJCC stage (6th)、WHO 分類、Lauren type) 間の関係を調査した。5 個の亜型の生存分析は、亜型と全体生存率の間の有意な相関関係を確認した ( $P = 3.42e - 09$ 、図 5)

50

。各亜型の5年生存率を決定した：各分子亜型別の5年生存率は、INFの場合、76.1% (95%信頼区間67.7~85.7)、INTの場合、65.1% (95%信頼区間56.2~75.4)、GSTの場合、64.6% (95%信頼区間55.0~75.9)、MXDの場合、51.3% (95%信頼区間42.1~62.4)、MSCの場合、46.3% (95%信頼区間38.0~56.5)である。INF亜型は、トレーニングセットIのMXDおよびMSC亜型より有意に低い死亡危険と関連があった。

#### 【0117】

最後に、前記分類の類似点と差異点をアジア癌研究グループ(ACRG)(GSE62254)、Cancer Genome Atlas(TCGA)およびシンガポール研究グループ(GSE15459)により報告されたGC亜型と比較した：すなわち、i)最悪の臨床結果を示すMSC亜型は、ACRG EMT亜型、TCGA GS亜型およびシンガポール研究グループ浸潤性亜型の4個の分類体系で共通した亜型であり、ii)最上の臨床結果を有するINF亜型多くのエプスタイン-バーウイルス(EBV)陽性患者およびTCGAおよびACRGにより確認された部分的なマイクロサテライトの高い不安定性(MSI)グループが含まれた。そうではない場合、GSTおよびINT亜型は、シンガポール研究グループで部分的に説明した。MXD亜型は、TCGAによる構造的染色体不安定性と関連している。non-MSCおよびnon-INF GC集団で亜型マッチングは、病理学的に曖昧な傾向があったが、おそらくTCGAとACRGが体細胞コピー数およびTP53活性の程度を分類器(Classifier)として使用したからであると認められる。

10

20

#### 【0118】

<実施例4> INT亜型と比較してMSC亜型の前臨床治療反応

モジュール基質分析で、基質特性は、Lauren分類で再発だけでなく、拡散タイプ(Diffuse type)と有意な関連があった。これは、GC細胞株を使用してMSC亜型の間葉および幹細胞類似の行動を検証するように刺激した。さらに、最近の証拠がEMTと関連した薬物耐性の獲得が多様な種類の癌で予後が良くないことを示すので、MSC亜型の前臨床治療反応を評価した。GC細胞株(n=26)は、細胞株の遺伝子発現データを患者GC腫瘍サンプルのデータと併合した後、5個の亜型に分類した(距離加重値識別方法)(図6a)。基質モジュールeigengeneで順位を付けることにより、Hs746TおよびSNU484 GC細胞株をMSC-亜型細胞株でモデル細胞株として選択した。INT亜型に割り当てられたNCI-N87およびMKN-45細胞を基質シグネチャーがない対照群として使用した(図6b)。試験管内浸潤および創傷治癒分析で、Hs746TおよびSNU484細胞は、NCI-N87およびMKN-45細胞より浸湿的な性能および運動性を示した(図7a、b)。3Dスフェロイド形成の分析結果、Hs746TおよびSNU484細胞は、幹細胞類似特性を示した(図7c)。生体内同所腫瘍モデルのT2加重軸磁気共鳴映像は、NCI-N87およびMKN-45細胞が制限された腫瘍を形成するのに対し、Hs746TおよびSNU484腫瘍が胃壁の壁に沿って拡散されることを示した(図7e、白色点線)。また、NCI-N87細胞と比較してHs746Tの基質特性に対するTGF-抑制剤(LY2157299)の影響を観察した。TGF-抑制剤を使用した治療は、試験管内でHs746T細胞の創傷治癒、侵犯および3Dスフェロイド形成能力を遅延させた(図7e~g)。EMT関連薬物耐性を確認するために、Hs746T細胞を使用して生体内異種移植マウスモデル(群別n=8)にTGF-抑制剤と抗癌剤の組合せ(オキサリプラチン+5-FU)を共同投与した。オキサリプラチン+5-FU治療がHs746Tモデルで腫瘍成長に対して単に若干効果的であったが、TGF-抑制剤/オキサリプラチン+5-FUの併用投与は、Hs746Tで薬物耐性と腫瘍の量を有意に減少させた(図7h)。その代わりに、抗癌剤の組合せだけでTGF-抑制剤の助けなく、非基質性NCI-N87腫瘍で腫瘍成長を減少させた(図7i)。

30

40

#### 【0119】

<実施例5> 臨床適用のためのqPCRプローブセットとしてGC miniClass

50

## s i f i e r - 2 6

Class i f i e r - P A M 9 3 2 を q P C R プローブセットで m i n i C l a s s i f i e r - 2 6 で精製して、安定的且つ臨床的に活用可能な分類システムを構築した ( 図 8 ) ( 図 9 は、m i n i C l a s s i f i e r - 2 6 で選定および分析する流れ図を示す)。分類器 ( C l a s s i f i e r ) 選別のために、G C 安定性の代表性の程度を考慮した。本発明者らは、胃 ( g a s t r i c ) シグネチャー、間葉 ( m e s e n c h y m a l ) シグネチャー、増殖 ( p r o l i f e r a t i v e ) シグネチャー、免疫 ( i m m u n e ) シグネチャーおよび腸 ( i n t e s t i n a l ) シグネチャーの 6 個の G C シグネチャーにより m i n i C l a s s i f i e r サブセットを分類して、亜型別およびコホート保存型モジュールで候補遺伝子を選別した。候補者は、i ) 亜型判別点数 ( P A M 分析 ) と i i ) モジュール内連結性により追加にフィルタリングされた ( W G C N A 分析 )。プローブ安定性は、プラットフォーム ( マイクロアレイおよび q P C R ) およびサンプリング方法 ( 新鮮 - 凍結および F F P E 標本 ) の独立性に基づいて評価された。最後に、癌生物学の先験的生物学的知識により遺伝子を減少させて、m i n i C l a s s i f i e r - 2 6 q P C R プローブセットを得た。また、選択された m i n i C l a s s i f i e r - 2 6 プローブセットが F F P E 標本の可能な空間異質性 ( 分散係数 5 % ) の影響を受けないことを確認した。

10

## 【 0 1 2 0 】

トレーニングセット I I ( n = 1 2 5 9、新鮮 - 凍結サンプルで様々なプラットフォームで得た M e r g e d 1 2 5 9 ) を使用して m i n i C l a s s i f i e r - 2 6 プローブセットを作った ( 図 8 )。m i n i C l a s s i f i e r - 2 6 プローブセットに分類された G C 亜型は、生存分析で高い予後と関連があった ( L R テスト、 $P = 2.48e-09$ ) ( 図 8 d )。C l a s s i f i e r - P A M 9 3 2 により分類された亜型と同様に、I N F 亜型は、最も良い予後 ( 5 年生存率 6 7 . 3 %、9 5 % C I 6 1 . 3 - 7 3 . 9 % )、そして M X D ( 5 年生存率 4 5 . 0 %、9 5 % C I 3 6 . 5 - 5 5 . 4 % ) と M S C ( 5 年生存率 3 3 . 0 %、9 5 % C I 2 7 . 3 ~ 4 0 . 0 % ) は、最悪の予後を示した。G C 亜型 ( 図 8 d ) の生存曲線の傾向の一貫性は、q P C R 3 2 5 テストセットでも確認した ( 図 8 e、 $P = 0.000534$ )。最も予後が良い I N F 亜型は、5 年生存率が 7 8 . 5 % ( 9 5 % C I 6 9 . 4 ~ 8 8 . 8 % ) であり、最も予後が良くない M X D、M S C 亜型 5 年生存率は、それぞれ 5 4 . 6 % ( 9 5 % C I 4 1 . 7 ~ 7 1 . 4 % )、5 7 . 5 % ( 9 5 % C I 4 8 . 4 ~ 6 8 . 2 % ) であった。

20

30

## 【 0 1 2 1 】

前記で確認された胃癌の 5 個の分子亜型に対して個人および公知のデータセット ( G S E 1 3 8 6 1 p、G S E 1 5 4 5 9、T C G A および S G E 6 2 2 5 4 ) で同定した。

## 【 0 1 2 2 】

図 8 のように、炎症 ( I n f l a m m a t o r y ) 分子亜型は、免疫 ( i m m u n e ) シグネチャーと関連し、腸 ( I n t e s t i n a l ) 分子亜型は、i n t e s t i n a l e p i t h e l i a l d i f f e r e n t i a t i o n 遺伝子の高発現と関連し、胃 ( G a s t r i c ) 分子亜型は、g a s t r i c m u c o s a - s p e c i f i c 遺伝子の高発現と関連する。混合型間質 ( M i x e d - s t r o m a l ) 分子亜型は、異質性 t r a n s i t - a m p l i f y i n g 特徴を示し、間葉 ( M e s e n c h y m a l ) 分子亜型は、E M T および間葉特性と関連する。

40

## 【 0 1 2 3 】

表 3 は、図 8 a の標的遺伝子発現パターンをスコア

【数 5 8】

$$(d'_{ik})$$

で示したものである。表 3 のスコア

【数 5 9】

10

$$(d'_{ik})$$

で正数は、遺伝子の高発現、負数は、遺伝子の低発現、0 は、遺伝子の発現変化がないことを意味する。したがって、正数の値が大きいほど発現がさらに高くなり、負数の値が大きいほど発現はさらに低くなる。

20

【0 1 2 4】

【表 3】

表 3 分子亜型分類基準表,  $d'_{ik}$

シグニチャー	遺伝子名	分子亜型				
		GST	MSC	INF	MXD	INT
Gastric	TFF1	5.79042	-0.77487	-1.60336	-3.74536	1.405637
	TFF2	7.261728	0.125744	-1.90636	-4.92322	0.282166
	VSIG1	5.43225	0.676749	0	-3.77475	-1.73528
Mesenchymal	CNN1	-0.38563	6.066372	-2.73467	-0.2289	-3.78749
	NEXN	-0.31125	5.523202	-1.84385	0.061904	-4.42691
	SCRG1	0.035119	6.184015	-2.58117	0.734977	-5.35065
	SORBS1	0	4.114709	-2.22306	0	-2.45196
	SPARCL1	0	5.482394	-2.1124	-0.47046	-3.84298
Proliferative	AURKA	0.06389	-1.39487	0.516858	0.548502	0.460776
	BUB1	0.007791	-2.98824	1.132822	0.703528	1.649687
	CDC20	0.595127	-3.58249	1.676981	0.679837	1.436201
	CEP55	0.471205	-2.79214	1.807234	0.623274	0.512529
	PTTG1	0	-2.97661	1.877958	0.617296	1.081789
	UBE2C	0	-3.15876	1.383331	0.71818	1.729608
Immune	WARS	0.374598	-0.90165	2.998455	-0.29396	-1.76155
	CD8A	0.476849	1.302821	2.354663	-0.13816	-4.13092
	NKG7	0.489444	0.701763	2.919867	-0.34839	-3.82988
	GBP1	-0.27312	0.323953	3.151314	0.114002	-3.44722
	GBP5	-0.09979	-0.20596	4.17765	-0.33953	-3.22494
	GZMB	0.700318	-0.7865	3.864089	0	-3.25334
Intestinal	CDH17	0.480373	-0.47566	-5.93442	2.960223	3.769466
	CDX1	-0.32734	-0.7605	-5.00726	3.781794	3.283526
	MYO1A	1.215205	-1.24902	-2.97939	1.264486	2.561301
Stem-like	ANTXR1	-0.2461	2.607131	-1.6872	0.768521	-1.73629
	SFRP4	-1.30339	3.262537	-1.03758	1.096524	-2.65513
	VCAN	-0.18046	1.739665	-0.52306	0.569731	-1.74826

10

20

30

40

【 0 1 2 5 】  
前記スコア  
【 数 6 0 】

$d'_{ik}$   
( )

は、下記の式 2 に適用して SV (Score Value) を計算し、SV によって腸 (Intestinal) 分子亜型、炎症 (Inflammatory) 分子亜型、間葉 (Mesenchymal) 分子亜型、混合型間質 (Mixed-stromal) 分子亜型および胃 (Gastric) 分子亜型に分類することができる：

50

【 0 1 2 6 】

【 数 6 1 】

[式2]

$$SV (\text{Score Value}) = \frac{1}{t} \sum_{i \in SN_{\theta}} d'_{ik}$$

10

【 0 1 2 7 】

ここで、SVは、NMF基盤クラスタリングに得たクラスターで各シグネチャーの発現  
 平均値、tは、当該シグネチャーに属する遺伝子(i)の数、

【 数 6 2 】

$$SN_{\theta}$$

20

は、シグネチャー( = 6)、kは、クラスターの数であって、2~7の整数であり

【 数 6 3 】

$$d'_{ik}$$

30

は、各遺伝子の全体平均と各クラスター平均の間の距離をベースとするスコアである。

【 0 1 2 8 】

図10は、5個のクラスターを分子亜型で命名する方法を図示するものであり、胃(gastric)シグネチャーのSVが最大値を有するクラスターを胃(Gastric)分子亜型に決めて、残りの4個のクラスターのうち間葉(mesenchymal)シグネチャーのSVが最大値を有し、且つ増殖(proliferative)シグネチャーのSVが最小値を有するクラスターを間葉(Mesenchymal)分子亜型に定め、残りの3個のクラスターのうち免疫(immune)シグネチャーのSVが最大値を有し、且つ腸(intestinal)シグネチャーのSVが最小値を有するクラスターを炎症(Inflammatory)分子亜型に定め、残りの2個のクラスターのうち幹細胞様(stem-like)シグネチャーのSVが最大値を有する場合、混合型間質(Mixed-stromal)分子亜型に定め、最終的に残ったクラスターを腸(Intestinal)分子亜型に決めて命名する。

40

【 0 1 2 9 】

表3のスコア

【数 6 4】

$$(d'_{ik})$$

は、当該遺伝子別 t - 統計値で臨界値に補正し、下記の式 3 および 4 により

【数 6 5】

10

$$d'_{ik}$$

値を計算して、分子垂型基準を設定することができる：

【0 1 3 0】

【数 6 6】

[式 3]

20

$$d'_{ik} = \text{sign}(d_{ik})(|d_{ik}| - \Delta)_+$$

【0 1 3 1】

ここで、各分子垂型 (k) 内の当該遺伝子 (i) 平均値

【数 6 7】

30

$$\bar{x}_{ik} = \sum_{j \in C_k} \frac{x_{ij}}{n_k}$$

大部分がノイズであり、当該遺伝子 (i) の全体平均値

【数 6 8】

40

$$\bar{x}_i = \sum_{j=1}^n \frac{x_{ij}}{n}$$

に近くなり得るので、臨界値 ( ) に補正して

50

【数 6 9】

$$d_{ik}$$

値が 0 に収斂する遺伝子を除外させる。これにより、各分子亜型 ( k ) 内遺伝子 ( i ) の信頼できる平均値を計算することができる。前記式 3 の

【数 7 0】

10

$$d_{ik}$$

を計算するために下記の式 4 に従う：

【 0 1 3 2】

【数 7 1】

[式 4]

20

$$d_{ik} = \frac{\bar{x}_{ik} - \bar{x}_i}{m_k(s_i + s_0)}$$

$$s_i^2 = \frac{1}{n-k} \sum_{k=1}^5 \sum_{j \in C_k} (x_{ij} - \bar{x}_{ik})^2$$

【 0 1 3 3】

ここで、

【数 7 2】

30

$$d_{ik}$$

は、当該遺伝子 ( i ) の t - 統計値であり、同じ分子亜型 ( subtype ) 内のサンプル遺伝子発現量に加重値を与えて標準化するもので、線形判別分析 ( LDA ; Linear Discriminant Analysis ) で頻繁に使用される統計方法である。

40

【数 7 3】

$$s_i$$

は、遺伝子 ( i ) の全体サンプルに対する当該分子亜型 ( k ) 内の標準偏差である。 j は、当該分子亜型 ( k ) 内のサンプルであり、

50

【数 7 4】

$$m_k \left( m_k = \sqrt{\frac{1}{n_k} + \frac{1}{n}} \right)$$

10

は、

【数 7 5】

$$\bar{x}_{ik} - \bar{x}_i$$

の標準誤差を補正するための自由度であり、

20

【数 7 6】

$$s_0$$

は、

【数 7 7】

$$s_i$$

30

の中央値である。

【0 1 3 4】

表 3 は、分子亜型分類基準表であり、赤色は、最大値、青色は、最小値を意味し、3 2 5 個のサンプルに対して 2 6 個の標的遺伝子のスコア

【数 7 8】

40

$$(d'_{ik})$$

から式 3 および 4 により計算された

【数 7 9】

$$d'_{ik}$$

を基盤として作成された分子亜型分類基準表である。

【0 1 3 5】

テストサンプルの分子亜型を把握しようとするとき、テストサンプルの標的遺伝子群の発現値に対して前記式 3 と 4 で計算された値を式 5 に適用して、表 3 のスコア

【数 8 0】

$$(d'_{ik})$$

を分子亜型分類基準表としてテストサンプル

【数 8 1】

20

$$(x^*)$$

の分子亜型 (k) を判別することができる：

【0 1 3 6】

【数 8 2】

[式 5]

30

$$\delta_k(x^*) = \sum_{i=1}^p \frac{(x_i^* - \bar{x}'_{ik})^2}{(s_i - s_0)} - 2 \log \pi_k, \min_k \delta_k(x^*)$$

$$\sum_{k=1}^K \pi_k = 1, \quad \pi_k = \frac{1}{K} ; \pi_k \text{ は分子亜型 (k) の事前確立}$$

40

【0 1 3 7】

ここで、テストサンプル

【数 8 3】

$$(x_i^*)$$

50

は、26個の標的遺伝子Cq値を325個のテストセットIIの26個の標的遺伝子Cq値とメディアンセンタリング (median centering) した値であり、  
【数84】

$$\delta_k(x^*)$$

は、テストサンプル  
【数85】

10

$$(x^*)$$

の分子亜型「k」を判別するスコアであって、テストサンプルの判別スコアが最も小さい値の分子亜型 (k) を選択して分類する。表4は、  
【数86】

20

$$\delta_k(x^*)$$

によって325個のサンプルに対する分子亜型決定に一致率およびエラー率である。  
【0138】  
【表4】

30

表4 分子亜型別NMFと判別スコア一致率およびエラー率, n=325

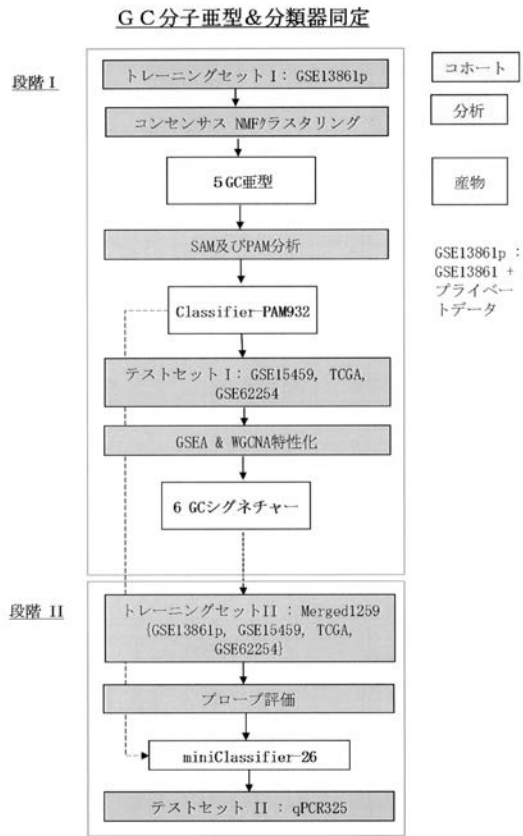
判別スコア \ NMF	Intestinal	Inflammatory	Mesenchymal	Mixed-stromal	Gastric	エラー率 (%)
Intestinal	38	0	0	0	0	0%
Inflammatory	0	46	0	0	0	0%
Mesenchymal	0	3	84	0	2	5.60%
Mixed-stromal	31	20	13	44	1	59.60%
Gastric	6	2	0	0	35	19.60%
総エラー率						16.96%

40

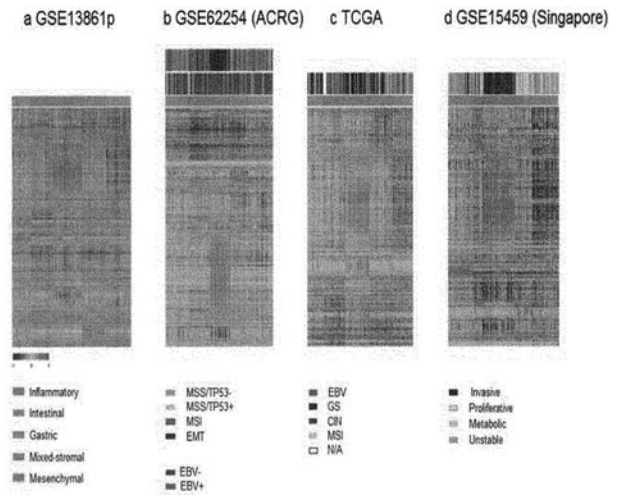
【産業上の利用可能性】  
【0139】

本発明は、胃癌予後予測分野に適用することができる。

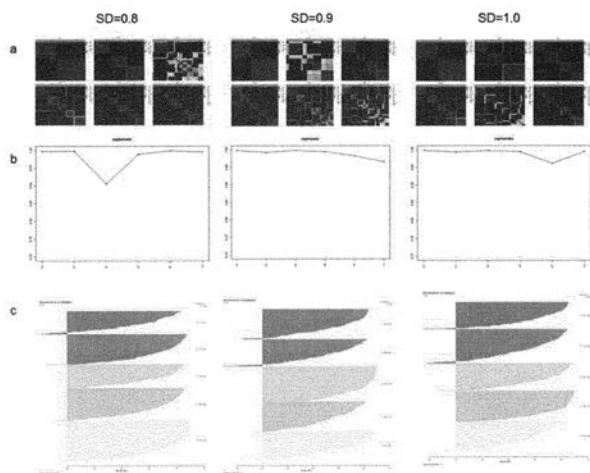
【 図 1 】



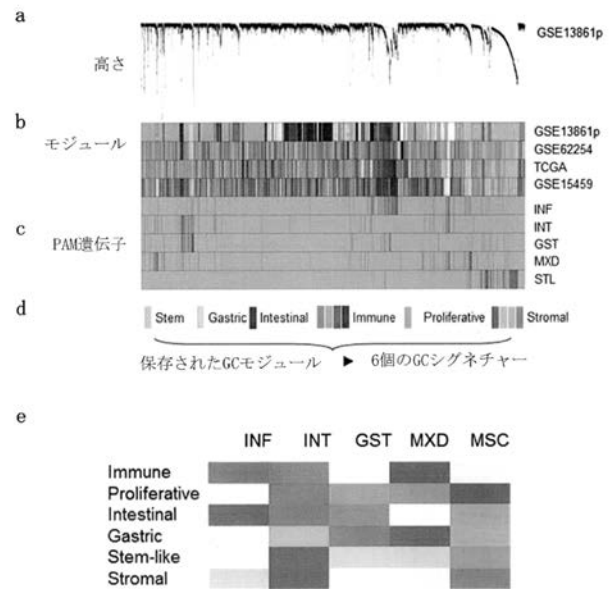
【 図 2 】



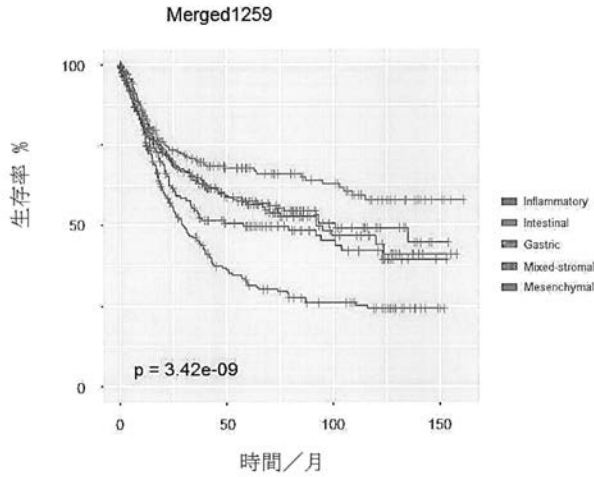
【 図 3 】



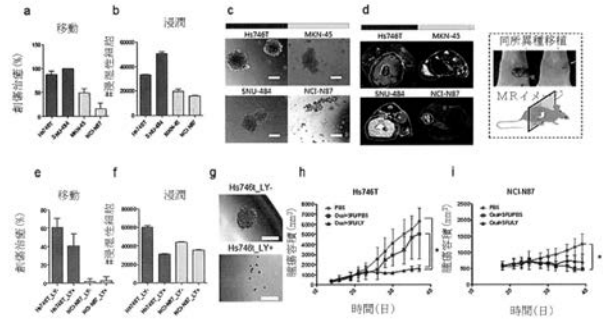
【 図 4 】



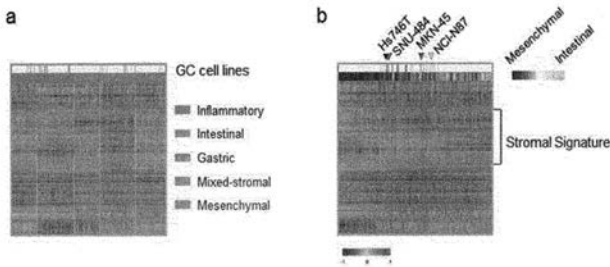
【 図 5 】



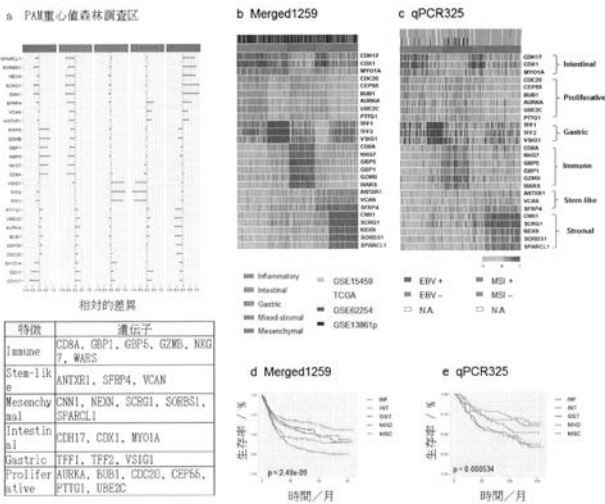
【 図 7 】



【 図 6 】



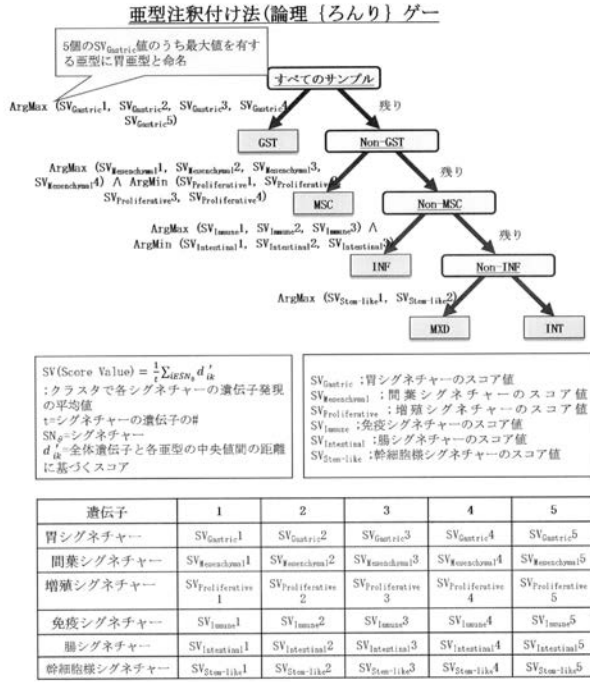
【 図 8 】



【 図 9 】




【 図 1 0 】



【 配 列 表 】

2019531741000001.app

## 【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. <b>PCT/KR2018/004732</b>
<b>A. CLASSIFICATION OF SUBJECT MATTER</b> <i>C12Q 1/6886(2018.01)i, G06F 19/12(2011.01)i, G06F 19/24(2011.01)i</i> According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) C12Q 1/6886; C12Q 1/68; G01N 33/574; G06F 19/12; G06F 19/24  Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Korean Utility models and applications for Utility models: IPC as above Japanese Utility models and applications for Utility models: IPC as above  Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) eKOMPASS (KIPO internal) & Keywords: advanced gastric cancer, mRNA, bio marker, qPCR		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2011-0059452 A1 (GOLDENRING, James R. et al.) 10 March 2011 See claims 1-26.	1-14
A	KR 10-2014-0121523 A (INDUSTRY-ACADEMIC COOPERATION FOUNDATION, YONSEI UNIVERSITY) 16 October 2014 See the entire document.	1-14
A	KR 10-2012-0063959 A (SAMSUNG LIFE PUBLIC WELFARE FOUNDATION) 21 June 2012 See the entire document.	1-14
A	TONG, Weihua et al., "Serum Biomarker Panels for Diagnosis of Gastric Cancer", OncoTargets and Therapy, 2016, vol. 9, pages 2455-2463 See the entire document.	1-14
A	WANG, Xiaoxiao et al., "Identification of Specific Biomarkers for Gastric Adenocarcinoma by ITRAQ Proteomic Approach", Scientific Reports, 2016, vol. 6, thesis no. 38871 (inner pages 1-15) See the entire document.	1-14
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search <p style="text-align: center;"><b>26 JULY 2018 (26.07.2018)</b></p>		Date of mailing of the international search report <p style="text-align: center;"><b>26 JULY 2018 (26.07.2018)</b></p>
Name and mailing address of the ISA/KR  Korean Intellectual Property Office Government Complex-Daejeon, 189 Seonsa-ro, Daejeon 302-701, Republic of Korea Facsimile No. +82-42-481-8578		Authorized officer   Telephone No.

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.

**PCT/KR2018/004732**

Patent document cited in search report	Publication date	Patent family member	Publication date
US 2011-0059452 A1	10/03/2011	NONE	
KR 10-2014-0121523 A	16/10/2014	KR 10-1504818 B1	24/03/2015
KR 10-2012-0065959 A	21/06/2012	CN 103459597 A	18/12/2013
		CN 103459597 B	30/03/2016
		EP 2653546 A2	23/10/2013
		KR 10-1437718 B1	11/09/2014
		US 2013-0337449 A1	19/12/2013
		US 9315869 B2	19/04/2016
		WO 2012-081898 A2	21/06/2012
		WO 2012-081898 A3	11/10/2012

국제조사보고서

국제출원번호  
**PCT/KR2018/004732**

<b>A. 발명이 속하는 기술분류(국제특허분류(IPC))</b> C12Q 1/6886(2018.01)i, G06F 19/12(2011.01)i, G06F 19/24(2011.01)i		
<b>B. 조사된 분야</b> 조사된 최소문헌(국제특허분류를 기재) C12Q 1/6886; C12Q 1/68; G01N 33/574; G06F 19/12; G06F 19/24 조사된 기술분야에 속하는 최소문헌 이외의 문헌 한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우)) eKOMPASS(특허청 내부 검색시스템) & 키워드: 진행성 위암, mRNA, 바이오 마커, qPCR		
<b>C. 관련 문헌</b>		
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
A	US 2011-0059452 A1 (GOLDENRING, JAMES R. 등) 2011.03.10 청구항 1-26 참조.	1-14
A	KR 10-2014-0121523 A (연세대학교 산학협력단) 2014.10.16 전체 문헌 참조.	1-14
A	KR 10-2012-0065959 A (사회복지법인 삼성생명공익재단) 2012.06.21 전체 문헌 참조.	1-14
A	TONG, WEIHUA 등, `Serum biomarker panels for diagnosis of gastric cancer` OncoTargets and Therapy, 2016, 9권, 페이지 2455-2463 전체 문헌 참조.	1-14
A	WANG, XIAOXIAO 등, `Identification of specific biomarkers for gastric adenocarcinoma by ITRAQ proteomic approach` Scientific Reports, 2016, 6권, 논문번호 38871 (내부 페이지 1-15) 전체 문헌 참조.	1-14
<input type="checkbox"/> 추가 문헌이 C(계속)에 기재되어 있습니다. <input checked="" type="checkbox"/> 대응특허에 관한 별지를 참조하십시오.		
* 인용된 문헌의 특별 카테고리: "A" 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌 "E" 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌 "L" 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌 "O" 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌 "P" 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌 "T" 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌 "X" 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다. "Y" 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다. "&" 동일한 대응특허문헌에 속하는 문헌		
국제조사의 실제 완료일 2018년 07월 26일 (26.07.2018)	국제조사보고서 발송일 2018년 07월 26일 (26.07.2018)	
ISA/KR의 명칭 및 우편주소 대한민국 특허청 (35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사) 팩스 번호 +82-42-481-8578	심사관 감유림 전화번호 +82-42-481-3516	

국제조사보고서  
대응특허에 관한 정보

국제출원번호  
**PCT/KR2018/004732**

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
US 2011-0059452 A1	2011/03/10	없음	
KR 10-2014-0121523 A	2014/10/16	KR 10-1504818 B1	2015/03/24
KR 10-2012-0065959 A	2012/06/21	CN 103459597 A	2013/12/18
		CN 103459597 B	2016/03/30
		EP 2653546 A2	2013/10/23
		KR 10-1437718 B1	2014/09/11
		US 2013-0337449 A1	2013/12/19
		US 9315869 B2	2016/04/19
		WO 2012-081898 A2	2012/06/21
		WO 2012-081898 A3	2012/10/11

## フロントページの続き

(51) Int.Cl.	F I	テーマコード (参考)
<b>C 1 2 Q 1/6851 (2018.01)</b>	C 1 2 Q 1/6851	Z
C 1 2 N 15/09 (2006.01)	C 1 2 N 15/09	Z

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

(72) 発明者 ヨン・ミン・ホ

大韓民国・ソウル・0 6 5 4 4 ・ソチョ - グ・シンバンポ - ロ・2 7 0 ・1 4 0 - ドン・9 0 2 -  
ホ・(バンポ - ドン・バンポ・シ・アパート)

F ターム(参考) 2G045 AA26 CB02 DA14 JA01  
4B029 AA07 BB11 BB20 CC01 FA15  
4B063 QA01 QA19 QQ03 QQ08 QQ42 QQ53 QQ58 QR08 QR32 QR36  
QR55 QR62 QS25 QS34 QX02

专利名称(译)	基于胃癌生物学特性的群体划分和预后预测系统		
公开(公告)号	<a href="#">JP2019531741A</a>	公开(公告)日	2019-11-07
申请号	JP2019520030	申请日	2018-04-24
发明人	ヨン・ミン・ホ		
IPC分类号	C12Q1/6886 G01N33/50 G01N33/53 C12M1/00 C12Q1/686 C12Q1/6851 C12N15/09		
CPC分类号	C12Q1/6886 C12Q2600/118 C12Q2600/158 G16B20/00 G16B25/10 G16B40/20 G16B40/30 C12Q2545/114 C12Q2561/113 G16B5/00 G16B40/00 G16B5/20		
FI分类号	C12Q1/6886.ZNA.Z G01N33/50.P G01N33/53.M C12M1/00.A C12Q1/686.Z C12Q1/6851.Z C12N15/09.Z		
F-TERM分类号	2G045/AA26 2G045/CB02 2G045/DA14 2G045/JA01 4B029/AA07 4B029/BB11 4B029/BB20 4B029/CC01 4B029/FA15 4B063/QA01 4B063/QA19 4B063/QQ03 4B063/QQ08 4B063/QQ42 4B063/QQ53 4B063/QQ58 4B063/QR08 4B063/QR32 4B063/QR36 4B063/QR55 4B063/QR62 4B063/QS25 4B063/QS34 4B063/QX02		
代理人(译)	村山彦 安倍晋三龙彦		
优先权	1020170052365 2017-04-24 KR		
外部链接	<a href="#">Espacenet</a>		

摘要(译)

本发明涉及基于胃癌的生物学特性的组划分和预后预测系统，并且可以通过利用晚期胃癌的预后目标基因组的mRNA表达水平的定量测试结果来根据存活率来预测预后。可以开发算法并将其作为辅助信息来决定如何治疗胃癌患者。

