

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2019-129850
(P2019-129850A)

(43) 公開日 令和1年8月8日(2019.8.8)

(51) Int.Cl.	F I	テーマコード (参考)
C 1 2 Q 1/68 (2018.01)	C 1 2 Q 1/68	2 G O 4 5
C 1 2 Q 1/02 (2006.01)	C 1 2 Q 1/02	4 B O 6 3
G O 1 N 33/50 (2006.01)	G O 1 N 33/50	P
G O 1 N 33/68 (2006.01)	G O 1 N 33/68	

審査請求 有 請求項の数 16 O L (全 69 頁)

(21) 出願番号 特願2019-87335 (P2019-87335)
 (22) 出願日 令和1年5月7日 (2019.5.7)
 (62) 分割の表示 特願2017-167175 (P2017-167175) の分割
 原出願日 平成24年7月19日 (2012.7.19)
 (31) 優先権主張番号 61/509,137
 (32) 優先日 平成23年7月19日 (2011.7.19)
 (33) 優先権主張国・地域又は機関 米国 (US)
 (31) 優先権主張番号 11178148.0
 (32) 優先日 平成23年8月19日 (2011.8.19)
 (33) 優先権主張国・地域又は機関 欧州特許庁 (EP)

(71) 出願人 590000248
 コーニンクレッカ フィリップス エヌ
 ヴェ
 KONINKLIJKE PHILIPS
 N. V.
 オランダ国 5656 アーエー アイン
 ドーフェン ハイテック キャンパス 5
 High Tech Campus 5,
 NL-5656 AE Eindhove
 n
 (74) 代理人 100107766
 弁理士 伊東 忠重
 (74) 代理人 100070150
 弁理士 伊東 忠彦

最終頁に続く

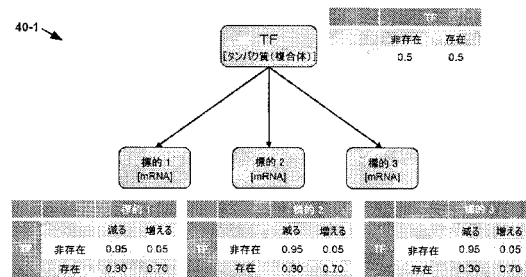
(54) 【発明の名称】 標的遺伝子発現の確率モデルを使用した細胞シグナル経路の活性の評価

(57) 【要約】

【課題】 ゲノム解析及びプロテオミクス解析をそれぞれ行うための適した方法及び手段を提供すること。

【解決手段】 本願は、主に、抽出された医療対象者の組織の試料において測定された1つ又は複数の細胞シグナル経路の1つ又は複数の標的遺伝子の1つ又は複数の発現レベルに少なくとも基づき、医療対象者の組織における1つ又は複数の細胞シグナル経路の活性を推量するための特定の方法及び、そのような方法を行うように構成されたデジタルプロセッサを含む装置、及び、そのような方法を行うためのデジタル処理装置によって実行可能な指示を記憶する非一過性記憶媒体に関する。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

医療対象者の組織における 1 つ又は複数の細胞シグナル経路の活性を、抽出された前記医療対象者の組織の試料において測定される前記 1 つ又は複数の細胞シグナル経路の 1 つ又は複数の標的遺伝子の 1 つ又は複数の発現レベルに少なくとも基づき推量するステップを含む方法であって、前記推量するステップは、

前記抽出された医療対象者の組織の試料において測定される 1 つ又は複数の細胞シグナル経路の 1 つ又は複数の標的遺伝子の 1 つ又は複数の発現レベルを少なくとも含む入力組に対する 1 つ又は複数の細胞シグナル経路を表す、確率モデル、好ましくはベイジアンネットワークの少なくとも一部を評価することによって、前記医療対象者の組織における前記 1 つ又は複数の細胞シグナル経路の活性を推量するステップ、

前記医療対象者の組織における、前記 1 つ又は複数の細胞シグナル経路の 1 つ又は複数の標的遺伝子の転写を制御する少なくとも 1 つの転写因子 (TF) 要素のレベルを推定するステップであり、前記少なくとも 1 つの TF 要素、及び、前記抽出された医療対象者の組織の試料において測定される 1 つ又は複数の細胞シグナル経路の 1 つ又は複数の標的遺伝子の 1 つ又は複数の発現レベルを関連づける条件付き確率に少なくとも部分的に基づくステップ、及び

前記組織の試料における前記転写因子の前記推定されたレベルに基づき、前記 1 つ又は複数の細胞シグナル経路の活性を推量するステップ、

を含み、

前記推量するステップは、前記 1 つ又は複数の細胞シグナル経路の前記確率モデルを使用してデジタル処理装置によって行われる、方法。

【請求項 2】

前記推量するステップは、

前記確率モデルの TF ノードによって表される前記医療対象者の組織における少なくとも 1 つの転写因子 (TF) 要素のレベルを推定するステップを含み、前記 TF 要素は、前記 1 つ又は複数の細胞シグナル経路の 1 つ又は複数の標的遺伝子の転写を制御し、当該推定するステップは、前記 TF ノード、及び、前記抽出された医療対象者の組織の試料において測定される 1 つ又は複数の細胞シグナル経路の 1 つ又は複数の標的遺伝子を表す前記確率モデルにおけるノードを関連づける、前記確率モデルの条件付き確率に少なくとも部分的に基づき、

前記推量するステップは、前記 1 つ又は複数のシグナル経路に関する情報を表すノードを含むベイジアンネットワーク、及び、ベイジアンネットワークの接続されたノード間の条件付き確率の関係を使用することによって好ましくは行われる、請求項 1 に記載の方法。

【請求項 3】

前記 1 つ又は複数の細胞シグナル経路は、Wnt 経路、ER 経路、AR 経路、及び/又は、ヘッジホッグ経路を含む、請求項 1 又は 2 に記載の方法。

【請求項 4】

前記推量するステップは、

KIAA1199、AXIN2、RNF43、TBX3、TDGF1、SOX9、ASCL2、IL8、SP5、ZNF3、KLF6、CCND1、DEFA6、及びFZD7を含む群から選択される前記抽出された医療対象者の組織の試料において測定されるWnt経路の1つ又は複数の、好ましくは少なくとも3つの標的遺伝子の発現レベルに少なくとも基づいて前記医療対象者の組織におけるWnt経路の活性を推量するステップを含む、請求項3に記載の方法。

【請求項 5】

前記推量するステップは、

CDH26、SGK3、PGR、GREB1、CA12、XBP1、CELSR2、WISP2、DSCAM、ERBB2、CTSD、TFF1、及びNRIP1を含む群から

選択される前記抽出された医療対象者の組織の試料において測定されるER経路の1つ又は複数の、好ましくは少なくとも3つの標的遺伝子の発現レベルに少なくとも基づいて前記医療対象者の組織におけるER経路の活性を推量するステップを(も)含む、請求項3又は4に記載の方法。

【請求項6】

前記推量するステップは、

GLI1、PTCH1、PTCH2、IGFBP6、SPP1、CCND2、FST、FOX L1、CFLAR、TSC22D1、RAB34、S100A9、S100A7、MYCN、FOX M1、GLI3、TCEA2、FYN、及びCTSL1を含む群から選択される前記抽出された医療対象者の組織の試料において測定されるヘッジホッグ経路の1つ又は複数の、好ましくは少なくとも3つの標的遺伝子の発現レベルに少なくとも基づいて前記医療対象者の組織におけるヘッジホッグ経路の活性を推量するステップを(も)含む、請求項3乃至5のいずれか一項に記載の方法。

10

【請求項7】

前記推量するステップは、

KLK2、PMEPA1、TMPRSS2、NKX3-1、ABCC4、KLK3、FKBP5、ELL2、UGT2B15、DHCR24、PPAP2A、NDRG1、LRIG1、CREB3L4、LCP1、GUCY1A3、AR、及びEAF2を含む群から選択される前記抽出された医療対象者の組織の試料において測定されるAR経路の1つ又は複数の、好ましくは少なくとも3つの標的遺伝子の発現レベルに少なくとも基づいて前記医療対象者の組織におけるAR経路の活性を推量するステップを(も)含む、請求項3乃至6のいずれか一項に記載の方法。

20

【請求項8】

前記推量するステップがさらに、NKD1、OAT、FAT1、LEF1、GLUL、REG1B、TCF7L2、COL18A1、BMP7、SLC1A2、ADRA2C、PPARG、DKK1、HNF1A、及びLECT2を含む群から選択される前記抽出された医療対象者の組織の試料において測定されるWnt経路の少なくとも1つの標的遺伝子の発現レベルに基づく、請求項4に記載の方法。

【請求項9】

前記推量するステップがさらに、AP1B1、ATP5J、COL18A1、COX7A2L、EBAG9、ESR1、HSPB1、IGFBP4、KRT19、MYC、NDUFV3、PISD、PRDM15、PTMA、RARA、SOD1、及びTRIM25を含む群から選択される前記抽出された医療対象者の組織の試料において測定されるER経路の少なくとも1つの標的遺伝子の発現レベルに基づく、請求項5に記載の方法。

30

【請求項10】

前記推量するステップがさらに、BCL2、FOXA2、FOXF1、H19、HIP、IL1R2、JAG2、JUP、MIF、MYLK、NKX2-2、NKX2-8、PITRM1、及びTOM1を含む群から選択される前記抽出された医療対象者の組織の試料において測定されるヘッジホッグ経路の少なくとも1つの標的遺伝子の発現レベルに基づく、請求項6に記載の方法。

40

【請求項11】

前記推量するステップがさらに、APP、NTS、PLAU、CDKN1A、DRG1、FGF8、IGF1、PRKACB、PTPN1、SGK1、及びTACC2を含む群から選択される前記抽出された医療対象者の組織の試料において測定されるAR経路の少なくとも1つの標的遺伝子の発現レベルに基づく、請求項7に記載の方法。

【請求項12】

抽出された前記医療対象者の組織の試料において測定されるWnt経路の標的遺伝子の組のうち2つ、3つ、若しくはそれ以上の標的遺伝子の発現レベルに少なくとも基づき医療対象者の組織におけるWnt経路の活性を推量するステップ、及び/又は、

抽出された前記医療対象者の組織の試料において測定されるER経路の標的遺伝子の組

50

のうち2つ、3つ、若しくはそれ以上の標的遺伝子の発現レベルに少なくとも基づき医療対象者の組織におけるER経路の活性を推量するステップ、及び/又は、

抽出された前記医療対象者の組織の試料において測定されるヘッジホッグ経路の標的遺伝子の組のうち2つ、3つ、若しくはそれ以上の標的遺伝子の発現レベルに少なくとも基づき医療対象者の組織におけるヘッジホッグ経路の活性を推量するステップ、及び/又は、

抽出された前記医療対象者の組織の試料において測定されるAR経路の標的遺伝子の組のうち2つ、3つ、若しくはそれ以上の標的遺伝子の発現レベルに少なくとも基づき医療対象者の組織におけるAR経路の活性を推量するステップ、
を含む、請求項1乃至11のいずれか一項に記載の方法。

10

【請求項13】

前記Wnt経路の標的遺伝子の組は、KIAA1199、AXIN2、RNF43、TBX3、TDGF1、SOX9、ASCL2、IL8、SP5、ZNRF3、KLF6、CCND1、DEFA6、及びFZD7を含む群から選択される少なくとも9つ、好ましくは全ての標的遺伝子を含み、及び/又は、

前記ER経路の標的遺伝子の組は、CDH26、SGK3、PGR、GREB1、CA12、XBP1、CELSR2、WISP2、DSCAM、ERBB2、CTSD、TF1、及びNRI1を含む群から選択される少なくとも9つ、好ましくは全ての標的遺伝子を含み、及び/又は、

前記ヘッジホッグ経路の標的遺伝子の組は、GLI1、PTCH1、PTCH2、IGFBP6、SPP1、CCND2、FST、FOX1、CFLAR、TSC22D1、RAB34、S100A9、S100A7、MYCN、FOX1、GLI3、TCEA2、FYN、及びCTSL1を含む群から選択される少なくとも9つ、好ましくは全ての標的遺伝子を含み、及び/又は、

20

前記AR経路の標的遺伝子の組は、K12、PMEPA1、TMPRSS2、NKX3-1、ABCC4、K13、FKBP5、ELL2、UGT2B15、DHCR24、PPAP2A、NDRG1、LRIG1、CREB3L4、LCP1、GUCY1A3、AR、及びEAF2を含む群から選択される少なくとも9つ、好ましくは全ての標的遺伝子を含む、請求項12に記載の方法。

【請求項14】

前記Wnt経路の標的遺伝子の組が、NKD1、OAT、FAT1、LEF1、GLUL、REG1B、TCF7L2、COL18A1、BMP7、SLC1A2、ADRA2C、PPARG、DKK1、HNF1A、及びLECT2を含む群から選択される少なくとも1つの標的遺伝子をさらに含み、及び/又は、

30

前記ER経路の標的遺伝子の組が、AP1B1、ATP5J、COL18A1、COX7A2L、EBAG9、ESR1、HSPB1、IGFBP4、KRT19、MYC、NDUFV3、PISD、PRDM15、PTMA、RARA、SOD1、及びTRIM25を含む群から選択される少なくとも1つの標的遺伝子をさらに含み、及び/又は、

前記ヘッジホッグ経路の標的遺伝子の組が、BCL2、FOXA2、FOXF1、H19、HHIP、IL1R2、JAG2、JUP、MIF、MYLK、NKX2-2、NKX2-8、PITRM1、及びTOM1を含む群から選択される少なくとも1つの標的遺伝子をさらに含み、及び/又は、

40

前記AR経路の標的遺伝子の組が、APP、NTS、PLAU、CDKN1A、DRG1、FGF8、IGF1、PRKACB、PTPN1、SGK1、及びTACC2を含む群から選択される少なくとも1つの標的遺伝子をさらに含む、請求項13に記載の方法。

【請求項15】

前記医療対象者の組織における推量された前記1つ又は複数の細胞シグナル経路の活性に基づき、前記1つ又は複数の細胞シグナル経路が前記医療対象者の組織において異常に機能しているかどうかを決定するステップをさらに含む、請求項1乃至14のいずれか一項に記載の方法。

50

【請求項 16】

請求項 1 乃至 15 のいずれか一項に記載の方法を行うように構成されたデジタルプロセッサを含む装置、又は、請求項 1 乃至 15 のいずれか一項に記載の方法を行うためにデジタル処理装置により実行可能な命令を記憶する非一過性の記憶媒体、又は、デジタル処理装置に請求項 1 乃至 15 のいずれか一項に記載の方法を行わせるプログラムコード手段を含むコンピュータプログラム。

【発明の詳細な説明】**【技術分野】****【0001】**

本明細書において記載される対象は、主に、生物情報学、ゲノム処理技術、プロテオミクス処理技術、及び関連技術に関する。

【背景技術】**【0002】**

種々の癌が、例えば細胞増殖及び転移等の癌の増殖及び進化において役割を果たすゲノム変異/変種の特定の組み合わせ、並びに/又は、特定の遺伝子に対する高い若しくは低い発現レベルと付随していると既知である腫瘍学等の医学分野における臨床適用に対して、ゲノム及びプロテオミクス解析は、実質的な実現化される且つ有望な見込みを有する。例えば、Wntシグナル経路は、細胞増殖の調節に影響を与え、且つ、高度に調節される。調節の損失による高度なWnt経路の活性は癌と関連づけられ、中でも、悪性の結腸腫瘍を有した癌と関連づけられてきた。いかなる特定の動作原理にも限定されないけれども、悪性の結腸細胞におけるWnt経路の調節解除は高度なWnt経路の活性を生じ、続いて、悪性の結腸細胞の細胞増殖、すなわち、結腸癌の転移を引き起こすと信じられている。一方、異常に低い経路の活性も、例えば骨粗鬆症の場合に関心のあるものであり得る。

【0003】

ゲノム及びプロテオミクスデータを取得するための技術は、臨床設定において容易に利用可能になってきている。例えば、マイクロアレイによる測定は、遺伝子発現レベル、タンパク質レベル、及びメチル化等を評価するために日常的に利用されている。自動遺伝子配列決定は、対費用効果の高いDNA及びmRNAにおける遺伝的変異の同定を可能にする。遺伝子配列決定の間のmRNAレベルの定量評価は、遺伝子発現レベルを評価するためのさらに別の臨床ツールとして見込みを持つ。

【0004】

これらの進歩にもかかわらず（又或いは、これらの進歩のため）、ゲノム及びプロテオミクス解析の臨床適用は、実質的な困難、データのオーバーロードに直面している。例えば、1つの臨床試料における同定可能な変異の数は、数十万以上に達し得る。これらの変異のほとんどは、癌増殖への特定の寄与はないいわゆるバスタンダー変異であり、少数の変異のみが癌増殖及び機能の発展に対して寄与し、これらは効果的な治療に対する標的をもたらす。1つのマイクロアレイが、数万の遺伝子に対する遺伝子発現レベルを生じ得る。これらの多量なデータを処理して、例えば正しい療法を選ぶことの適用における臨床的に有用な情報を同定することは、困難である。

【0005】

1つのアプローチは、アメリカ食料医薬品局（U.S. Food and Drug Administration（FDA））によって認められているテスト等、いくつかの基準のテスト又は標準化されたテストに解析を限定することである。そのようなアプローチにおいて、特定の指標又は指標の組み合わせ（例えば、変異及び/又は特定化された高い若しくは低い遺伝子発現レベル）が、示された疾患の条件（例えば、特定のタイプの癌）に対する「陽性」をテストするために検出される。基準のテストは、疾患の条件又は治療の有効性との強い相互関係を示してきた臨床研究によって支持される。このアプローチは、例えば、特定の疾患の診断又は特定のステージの特定の癌のタイプにおける薬物への応答を予測する等、基準のテストが開発されてきた臨床条件に対してのみ有用であり、また標準的条件に対してのみ適用可能であるように柔軟性に欠ける。

10

20

30

40

50

【0006】

別のアプローチは、機能的に関連するゲノム又はプロテオミクス指標の群の同定に基づいている。例えば、Wnt経路は、プロテオミクス反応のカスケードを含む。この鎖の主要構成成分は、(それだけに限定されないが)ディシブルタンパク質ファミリーのタンパク質の活性化を引き起こし、次に細胞核におけるβ-カテニン/TCF4ベースのタンパク質複合体等の転写作用剤のレベルに影響を与える細胞のFrizzled表面受容体へのWntシグナルタンパク質の結合を含む。これらの転写作用剤は、次に、標的mRNA分子の転写を制御し、標的mRNA分子は、次に、Wnt経路の標的タンパク質に翻訳される。臨床研究は、Wnt経路の調節タンパク質とWnt経路の活性との相互関係をいくつか示してきた。

10

【0007】

しかし、そのような臨床研究の結果を、特定の患者の診断及び臨床評価に適用することは、例えばWnt経路等のシグナル経路の複雑さのため困難である。シンプルな例として、Wnt経路において「上流」であるタンパク質の発現レベルの測定は、Wnt経路において「下流」であるタンパク質の異常行動を検出できない場合がある。Wnt経路は多数のフィードバック機構を含み、且つ、単純化された「上流」及び「下流」の概念は、Wnt経路の重要な部分に対して不適當でありえ、さらに一般的には、Wnt経路を含むタンパク質カスケードの一部における異常行動はタンパク質カスケードの他の部分に対して、及び、全体としてWnt経路の活性に対して多かれ少なかれ影響を有し得ると信じられている。他の一部の臨床研究において、シグナルカスケードの調節タンパク質に対するタンパク質発現レベルは、調節タンパク質をコードする遺伝子のmRNA発現レベルを測定することによって評価される。これは、調節タンパク質の発現レベルを正確に評価することができない間接測定であり、(リン酸化のような特定の翻訳後修飾の後の)活性タンパク質の量をこれまでほとんど反映しない。

20

【発明の概要】

【発明が解決しようとする課題】

【0008】

本発明の根底にある主な問題は、従って、ゲノム解析及びプロテオミクス解析をそれぞれ行うための適した方法及び手段を提供することであった。本発明と関連した、根底にある問題だけでなく、さらなる異論の特定の態様は、本明細書において提供される説明及び実施例を熟読した場合に、特に、付随の特許請求の範囲を熟読した場合に明らかになる。

30

【課題を解決するための手段】

【0009】

本発明は、本明細書において開示される新たな且つ改善された方法及び装置を提供する。

【0010】

本発明の主な態様によると、上記の問題は、標的遺伝子発現の確率モデルを使用して細胞シグナル経路の活性を評価するための特定の方法、すなわち、

抽出された医療対象者の組織の試料において測定される1つ又は複数の細胞シグナル経路の1つ又は複数の標的遺伝子の1つ又は複数の発現レベル(特に、mRNA及び/又はタンパク質レベル)に少なくとも基づき、医療対象者の組織における1つ又は複数の細胞シグナル経路の活性を推量するステップを含む方法によって解決され、前記推量するステップは、

40

抽出された医療対象者の組織の試料において測定される1つ又は複数の細胞シグナル経路の1つ又は複数の標的遺伝子の1つ又は複数の発現レベルを少なくとも含む入力組に対する1つ又は複数の細胞シグナル経路を表す、確率モデル、好ましくはベイジアンネットワークの少なくとも一部を評価することによって、医療対象者の組織における1つ又は複数の細胞シグナル経路の活性を推量するステップ、

医療対象者の組織における、少なくとも1つの転写因子(TF)要素のレベルを推定するステップであり、前記少なくとも1つのTF要素は、前記1つ又は複数の細胞シグナル

50

経路の1つ又は複数の標的遺伝子の転写を制御し、当該推定するステップは、前記少なくとも1つのTF要素、及び、抽出された医療対象者の組織の試料において測定される1つ又は複数の細胞シグナル経路の1つ又は複数の標的遺伝子の1つ又は複数の発現レベルを関連づける条件付き確率に少なくとも部分的に基づくステップ、並びに、

組織試料における転写因子の前記推定されたレベルに基づき、1つ又は複数の細胞シグナル経路の活性を推量するステップ、さらに、

医療対象者の組織において推量された1つ又は複数の細胞シグナル経路の活性に基づき、1つ又は複数の細胞シグナル経路が医療対象者の組織において異常に機能しているかどうかを決定するステップ、

を含み、前記推量するステップは、1つ又は複数の細胞シグナル経路の確率モデルを使用してデジタル処理装置によって行われる。

10

【0011】

「1つ又は複数の標的遺伝子」は、(本明細書において記載される場合)「直接の標的遺伝子」及び/又は「間接の標的遺伝子」であり得る。

【0012】

好ましくは、前記推量するステップは、確率モデルのTFノードによって表される医療対象者の組織における少なくとも1つの転写因子(TF)要素のレベルを推定するステップを含み、TF要素は、1つ又は複数の細胞シグナル経路の1つ又は複数の標的遺伝子の転写を制御し、前記推定するステップは、前記TFノード、及び、抽出された医療対象者の組織の試料において測定される1つ又は複数の細胞シグナル経路の1つ又は複数の標的遺伝子を表す前記確率モデルにおけるノードを関連づける、確率モデルの条件付き確率に少なくとも部分的に基づく。

20

【0013】

確率モデルは、ベイジアンネットワークモデルであり得る。従って、好ましい実施形態によると、前記推量するステップは、1つ又は複数のシグナル経路に関する情報を表すノードを含むベイジアンネットワーク、及び、ベイジアンネットワークの接続されたノード間の条件付き確率の関係を使用することによって行われる。

【0014】

1つ又は複数の細胞シグナル経路は、Wnt経路、ER(エストロゲン受容体)経路、AR(アンドロゲン受容体)経路、及び/又は、ヘッジホッグ経路であり得る。従って、好ましい実施形態によると、1つ又は複数の細胞シグナル経路は、Wnt経路、ER経路、AR経路、及び/又は、ヘッジホッグ経路を含む。

30

【0015】

特に、適した標的遺伝子は、以下のテキストの一節並びに下記の実施例(例えば表1から9を参照)において記載されている。

【0016】

従って、好ましい実施形態によると、1つ又は複数の標的遺伝子は、(Wnt経路に対して)表1又は表6において収載された標的遺伝子、(ER経路に対して)表2、表5、又は表7において収載された標的遺伝子、(ヘッジホッグ経路に対して)表3又は表8において収載された標的遺伝子、及び、(AR経路に対して)表4又は表9において収載された標的遺伝子を含むか若しくは該標的遺伝子から成る群から選択される。

40

【0017】

特に好ましいのは、前記推量するステップが、

KIAA1199、AXIN2、RNF43、TBX3、TDGF1、SOX9、ASC2、IL8、SP5、ZNF3、KLF6、CCND1、DEFA6、及びFZD7を含むか又は該標的遺伝子から成る群から選択される抽出された医療対象者の組織の試料において測定されるWnt経路の1つ又は複数の、好ましくは少なくとも3つの標的遺伝子の発現レベルに少なくとも基づいて医療対象者の組織におけるWnt経路の活性を推量するステップを含む方法である。

【0018】

50

さらに好ましいのは、前記推量するステップがさらに、NKD1、OAT、FAT1、LEF1、GLUL、REG1B、TCF7L2、COL18A1、BMP7、SLC1A2、ADRA2C、PPARG、DKK1、HNF1A、及びLECT2を含むか又は該標的遺伝子から成る群から選択される抽出された医療対象者の組織の試料において測定されるWnt経路の少なくとも1つの標的遺伝子の発現レベルに基づく方法である。

【0019】

特に好ましいのは、前記推量するステップが、

CDH26、SGK3、PGR、GREB1、CA12、XBP1、CELSR2、WISP2、DSCAM、ERBB2、CTSD、TFF1、及びNRIP1を含むか又は該標的遺伝子から成る群から選択される抽出された医療対象者の組織の試料において測定されるER経路の1つ又は複数の、好ましくは少なくとも3つの標的遺伝子の発現レベルに少なくとも基づいて医療対象者の組織におけるER経路の活性を推量するステップを(も)含む方法である。

10

【0020】

さらに好ましいのは、前記推量するステップがさらに、AP1B1、ATP5J、COL18A1、COX7A2L、EBAG9、ESR1、HSPB1、IGFBP4、KRT19、MYC、NDUFV3、PISD、PREDM15、PTMA、RARA、SOD1、及びTRIM25を含むか又は該標的遺伝子から成る群から選択される抽出された医療対象者の組織の試料において測定されるER経路の少なくとも1つの標的遺伝子の発現レベルに基づく方法である。

20

【0021】

前記推量するステップが、

GLI1、PTCH1、PTCH2、IGFBP6、SPP1、CCND2、FST、FOX L1、CFLAR、TSC22D1、RAB34、S100A9、S100A7、MYCN、FOX M1、GLI3、TCEA2、FYN、及びCTSL1を含むか又は該標的遺伝子から成る群から選択される抽出された医療対象者の組織の試料において測定されるヘッジホッグ経路の1つ又は複数の、好ましくは少なくとも3つの標的遺伝子の発現レベルに少なくとも基づいて医療対象者の組織におけるヘッジホッグ経路の活性を推量するステップを(も)含む方法も好ましい。

30

【0022】

さらに好ましいのは、前記推量するステップがさらに、BCL2、FOXA2、FOX F1、H19、HHIP、IL1R2、JAG2、JUP、MIF、MYLK、NKX2.2、NKX2.8、PITRM1、及びTOM1を含むか又は該標的遺伝子から成る群から選択される抽出された医療対象者の組織の試料において測定されるヘッジホッグ経路の少なくとも1つの標的遺伝子の発現レベルに基づく方法である。

【0023】

前記推量するステップが、

KLK2、PMEPA1、TMPRSS2、NKX3__1、ABCC4、KLK3、FKBP5、ELL2、UGT2B15、DHCR24、PPAP2A、NDRG1、LRIG1、CREB3L4、LCP1、GUCY1A3、AR、及びEAF2を含むか又は該標的遺伝子から成る群から選択される抽出された医療対象者の組織の試料において測定されるAR経路の1つ又は複数の、好ましくは少なくとも3つの標的遺伝子の発現レベルに少なくとも基づいて医療対象者の組織におけるAR経路の活性を推量するステップを(も)含む方法も好ましい。

40

【0024】

さらに好ましいのは、前記推量するステップがさらに、APP、NTS、PLAU、CDKN1A、DRG1、FGF8、IGF1、PRKACB、PTPN1、SGK1、及びTACC2を含むか又は該標的遺伝子から成る群から選択される抽出された医療対象者の組織の試料において測定されるAR経路の少なくとも1つの標的遺伝子の発現レベルに基づく方法である。

50

【 0 0 2 5 】

本発明の別の態様は、（本明細書において記載される）方法に関し、当該方法はさらに、

1つ又は複数の細胞シグナル経路の異常な機能を矯正する薬物を医療対象者に対して処方することを勧めるステップを含み、

該勧めるステップは、前記1つ又は複数の細胞シグナル経路が、医療対象者の組織において異常に機能していると前記1つ又は複数の細胞シグナル経路の推量された活性に基づき決定される場合のみに行われる。

【 0 0 2 6 】

本発明は、（本明細書において記載される）方法にも関し、当該方法は、

抽出された医療対象者の組織の試料において測定されるW n t経路の標的遺伝子の組のうち2つ、3つ、若しくはそれ以上の標的遺伝子の発現レベルに少なくとも基づき医療対象者の組織におけるW n t経路の活性を推量するステップ、

及び/又は、

抽出された医療対象者の組織の試料において測定されるE R経路の標的遺伝子の組のうち2つ、3つ、若しくはそれ以上の標的遺伝子の発現レベルに少なくとも基づき医療対象者の組織におけるE R経路の活性を推量するステップ、

及び/又は、

抽出された医療対象者の組織の試料において測定されるヘッジホッグ経路の標的遺伝子の組のうち2つ、3つ、若しくはそれ以上の標的遺伝子の発現レベルに少なくとも基づき医療対象者の組織におけるヘッジホッグ経路の活性を推量するステップ、

及び/又は、

抽出された医療対象者の組織の試料において測定されるA R経路の標的遺伝子の組のうち2つ、3つ、若しくはそれ以上の標的遺伝子の発現レベルに少なくとも基づき医療対象者の組織におけるA R経路の活性を推量するステップ、を含む。

【 0 0 2 7 】

好ましくは、

前記W n t経路の標的遺伝子の組は、K I A A 1 1 9 9、A X I N 2、R N F 4 3、T B X 3、T D G F 1、S O X 9、A S C L 2、I L 8、S P 5、Z N R F 3、K L F 6、C C N D 1、D E F A 6、及びF Z D 7を含むか若しくは該標的遺伝子から成る群から選択される少なくとも9つ、好ましくは全ての標的遺伝子を含み、

並びに/又は、

前記E R経路の標的遺伝子の組は、C D H 2 6、S G K 3、P G R、G R E B 1、C A 1 2、X B P 1、C E L S R 2、W I S P 2、D S C A M、E R B B 2、C T S D、T F F 1、及びN R I P 1を含むか若しくは該標的遺伝子から成る群から選択される少なくとも9つ、好ましくは全ての標的遺伝子を含み、

並びに/又は、

前記ヘッジホッグ経路の標的遺伝子の組は、G L I 1、P T C H 1、P T C H 2、I G F B P 6、S P P 1、C C N D 2、F S T、F O X L 1、C F L A R、T S C 2 2 D 1、R A B 3 4、S 1 0 0 A 9、S 1 0 0 A 7、M Y C N、F O X M 1、G L I 3、T C E A 2、F Y N、及びC T S L 1を含むか若しくは該標的遺伝子から成る群から選択される少なくとも9つ、好ましくは全ての標的遺伝子を含み、

並びに/又は、

前記A R経路の標的遺伝子の組は、K L K 2、P M E P A 1、T M P R S S 2、N K X 3 _ 1、A B C C 4、K L K 3、F K B P 5、E L L 2、U G T 2 B 1 5、D H C R 2 4、P P A P 2 A、N D R G 1、L R I G 1、C R E B 3 L 4、L C P 1、G U C Y 1 A 3、A R、及びE A F 2を含むか若しくは該標的遺伝子から成る群から選択される少なくとも9つ、好ましくは全ての標的遺伝子を含む。

【 0 0 2 8 】

10

20

30

40

50

前記 Wnt 経路の標的遺伝子の組が、NKD1、OAT、FAT1、LEF1、GLUL、REG1B、TCF7L2、COL18A1、BMP7、SLC1A2、ADRA2C、PPARG、DKK1、HNF1A、及びLECT2を含むか若しくは該標的遺伝子から成る群から選択される少なくとも1つの標的遺伝子をさらに含み、並びに/又は、

前記 ER 経路の標的遺伝子の組が、AP1B1、ATP5J、COL18A1、COX7A2L、EBAG9、ESR1、HSPB1、IGFBP4、KRT19、MYC、NDUFV3、PISD、PREDM15、PTMA、RARA、SOD1、及びTRIM25を含むか若しくは該標的遺伝子から成る群から選択される少なくとも1つの標的遺伝子をさらに含み、並びに/又は、前記ヘッジホッグ経路の標的遺伝子の組が、BCL2、FOXA2、FOXF1、H19、HHIP、IL1R2、JAG2、JUP、MIF、MYLK、NKX2.2、NKX2.8、PITRM1、及びTOM1を含むか若しくは該標的遺伝子から成る群から選択される少なくとも1つの標的遺伝子をさらに含み、並びに/又は、

前記 AR 経路の標的遺伝子の組が、APP、NTS、PLAU、CDKN1A、DRG1、FGF8、IGF1、PRKACB、PTPN1、SGK1、及びTACC2を含むか若しくは該標的遺伝子から成る群から選択される少なくとも1つの標的遺伝子をさらに含む方法が特に好ましい。

【0029】

本発明に従って使用されることになる1つ又は複数の試料は、好ましくは生検手順又は他の試料抽出手順を介した、例えば、乳房病変から、又は、結腸癌を有すると知られるか若しくは結腸癌を有するのではないかと疑われる医療対象者の結腸から、又は、肝癌を有すると知られるか若しくは肝癌を有するのではないかと疑われる医療対象者の肝臓等から得られる試料であり得る。試料が抽出される組織は、例えば、結腸、乳房、肝臓、又は、結腸、乳房、肝臓、若しくは他の臓器の外に広がる他の臓器から生じる(疑われる)悪性の組織等、転移性の組織であってもよい。一部の例において、組織試料は、循環性腫瘍細胞、すなわち、血流に入った腫瘍細胞であってもよく、さらに、適した単離技術を使用して抽出される組織試料として抽出してもよい。

【0030】

本発明の別の開示される態様は、特定の疾患の診断、又は、特定のステージの特定の癌タイプにおける薬物に対する応答の予測に対する、本明細書において記載される非一過性の記憶媒体、又は、本明細書において記載されるコンピュータプログラムの使用に係する。

【0031】

別の開示される態様によると、デジタルプロセッサを含む装置が、本明細書において記載される本発明による方法を行うように構成される。

【0032】

別の開示される態様によると、非一過性の記憶媒体は、本明細書において記載される本発明による方法を行うためにデジタル処理装置により実行可能な指示を記憶する。非一過性の記憶媒体は、ハードドライブ又は他の磁気記憶媒体、光ディスク又は他の光記憶装置、ランダムアクセスメモリ(RAM)、リードオンリーメモリ(ROM)、フラッシュメモリ、又は他の電子記憶媒体、ネットワークサーバー等のコンピュータ読み取り可能な記憶媒体であってもよい。デジタル処理装置は、ハンドヘルドの装置(例えば携帯情報端末又はスマートフォン等)、ノートブック型コンピュータ、デスクトップ型コンピュータ、タブレット型のコンピュータ又は装置、遠隔ネットワークサーバー等であってもよい。

【0033】

別の開示される態様によると、コンピュータプログラムは、デジタル処理装置に、本明細書において記載される本発明による方法を行わせるためのプログラムコード手段を含む。デジタル処理装置は、ハンドヘルドの装置(例えば携帯情報端末又はスマートフォン等)、ノートブック型コンピュータ、デスクトップ型コンピュータ、タブレット型のコンピュータ又は装置、遠隔ネットワークサーバー等であってもよい。

10

20

30

40

50

【0034】

1つの利点は、例えば、Wnt経路、ER経路、AR経路、及び/又は、ヘッジホッグ経路のベイジアンネットワークモデルを使用した、1つ又は複数の細胞シグナル経路の確率の解析に基づき臨床勧告を提供する臨床判断支援(CDS)システムに属する。

【0035】

別の利点は、エラーに感受性の低い改善された細胞シグナル経路活性の評価に属する。

【0036】

別の利点は、細胞シグナル経路の調節の損失に対する標的治療を勧告するCDSシステムを提供することに属する。

【0037】

別の利点は、Wnt経路、ER経路、AR経路、又は、ヘッジホッグ経路等の特定の細胞シグナル経路に対する調節の損失を検出するように設計され、且つ、その特定の細胞シグナル経路によって出所が明らかになる種々のタイプの癌に対する勧告を提供するように容易に適応するCDSシステムを提供することに属する。

10

【0038】

本明細書において記載される本発明は、例えば、

- 予測される(推量される)活性に基づく診断、
- 予測される(推量される)活性に基づく予後、
- 予測される(推量される)活性に基づく薬物処方、
- 予測される(推量される)活性に基づく薬効の予測、
- 予測される(推量される)活性に基づく副作用の予測、
- 薬効のモニタリング、
- 医薬品開発、
- アッセイ開発、
- 経路の研究、
- 癌進行度分類、
- 予測される(推量される)活性に基づく臨床試験における対象の登録、
- 行われることになる後の試験の選択、及び/又は、
- コンパニオン診断試験の選択、

20

に関しても有利に使用することができる。

30

【0039】

さらなる利点が、付随の図面、以下の説明を読み且つ理解した後、特に、本明細書において以下に提供される詳細な実施例を読んだ後に当業者には明らかになる。

【図面の簡単な説明】

【0040】

【図1】細胞シグナル経路の一部を表したシンプルなベイジアンネットワークを示した図である。細胞シグナル経路は、転写因子(TF)複合体、及び、存在する転写因子複合体の結果として産生される標的遺伝子によって象徴化されている。二進数離散化(Binary discretization)の場合のTF要素と標的遺伝子との確率の関係は、図において描かれている条件付き確率の表によって表すことができる。

40

【図2】仮説の細胞シグナル経路を記載した例示的なベイジアンネットワークを示した図である。上流のタンパク質も下流の標的mRNAノードも図に描かれている。上流のタンパク質は、転写因子複合体における入力として役立つ一方、標的mRNAは転写因子複体の出力ノードである。

【図3】ベイジアンネットワークの例示的な例を示した図であり、多数の転写因子複体を有した1つの細胞シグナル経路、若しくは、1つのベイジアンネットワークに組み合わせられるその独自の転写因子複体を有した多数の細胞シグナル経路、又は、その組み合わせを表している。

【図4】ベイジアンネットワークの例を示した図であり、図1に類似の細胞シグナル経路のシンプルな説明を例示している。ここでは、さらなるノードが付け加えられて、標的タ

50

ンパク質への標的 mRNA の翻訳を表している。

【図 5】ベイジアンネットワークの例示を示した図であり、細胞シグナル経路の別のシンプルな説明を例示している。経路は、転写因子複合体及びその標的タンパク質レベルを使用して表されている。

【図 6】さらなるノードの層を有した図 1 の例示的なベイジアンネットワークを示した図であり、プローブの強度を対応する標的 mRNA レベルに関係づけるマイクロアレイチップ上のプローブセットを表している。

【図 7】図 1 のベイジアンネットワークの異なる実施形態の例示的な例を示した図であり、この特定の例において、含まれる標的 mRNA レベルのいずれにも対するさらなる情報のノードに対しての例としてメチル化及びコピー数多型を表したノードを含んでいる。

【図 8】結腸試料のデータセット (GSE20916) における本明細書において記載されるベイジアンネットワーク及び最近隣セントロイド法 (nearest centroid method) の予測される Wnt 経路の活性を示した図である。

【図 9】結腸試料のデータセット (GSE4183) における本明細書において記載されるベイジアンネットワーク及び最近隣セントロイド法の予測される Wnt 経路の活性を示した図である。

【図 10】結腸試料のデータセット (GSE15960) における本明細書において記載されるベイジアンネットワーク及び最近隣セントロイド法の予測される Wnt 経路の活性を示した図である。

【図 11】乳癌試料のデータセット (GSE12777) における本明細書において記載されるベイジアンネットワーク及び最近隣セントロイド法の予測される Wnt 経路の活性を示した図である。

【図 12】乳癌試料のデータセット (GSE21653) における本明細書において記載されるベイジアンネットワーク及び最近隣セントロイド法の予測される Wnt 経路の活性を示した図である。

【図 13】肝癌試料のデータセット (GSE9843) における本明細書において記載されるベイジアンネットワーク及び最近隣セントロイド法の予測される Wnt 経路の活性を示した図である。

【図 14 - I】結腸試料のデータセット (GSE20916) における本明細書において記載される広範な文献によるリスト (broad literature list) の標的遺伝子と比較される証拠により精選されたリスト (evidence curated list) の標的遺伝子を使用したベイジアンネットワークを使用し予測される Wnt 経路の活性を示した図である。

【図 14 - II】結腸試料のデータセット (GSE20916) における本明細書において記載される広範な文献によるリスト (broad literature list) の標的遺伝子と比較される証拠により精選されたリスト (evidence curated list) の標的遺伝子を使用したベイジアンネットワークを使用し予測される Wnt 経路の活性を示した図である。

【図 15 - I】結腸試料のデータセット (GSE4183) における本明細書において記載される広範な文献によるリストの標的遺伝子と比較される証拠により精選されたリストの標的遺伝子を使用したベイジアンネットワークを使用し予測される Wnt 経路の活性を示した図である。

【図 15 - II】結腸試料のデータセット (GSE4183) における本明細書において記載される広範な文献によるリストの標的遺伝子と比較される証拠により精選されたリストの標的遺伝子を使用したベイジアンネットワークを使用し予測される Wnt 経路の活性を示した図である。

【図 16 - I】結腸試料のデータセット (GSE15960) における本明細書において記載される広範な文献によるリストの標的遺伝子と比較される証拠により精選されたリストの標的遺伝子を使用したベイジアンネットワークを使用し予測される Wnt 経路の活性を示した図である。

10

20

30

40

50

【図16 - I I】結腸試料のデータセット (G S E 1 5 9 6 0) における本明細書において記載される広範な文献によるリストの標的遺伝子と比較される証拠により精選されたりリストの標的遺伝子を使用したベイジアンネットワークを使用し予測されるW n t 経路の活性を示した図である。

【図17 - I】乳癌試料のデータセット (G S E 1 2 7 7 7) における本明細書において記載される広範な文献によるリストの標的遺伝子と比較される証拠により精選されたりリストの標的遺伝子を使用したベイジアンネットワークを使用し予測されるW n t 経路の活性を示した図である。

【図17 - I I】乳癌試料のデータセット (G S E 1 2 7 7 7) における本明細書において記載される広範な文献によるリストの標的遺伝子と比較される証拠により精選されたりリストの標的遺伝子を使用したベイジアンネットワークを使用し予測されるW n t 経路の活性を示した図である。

10

【図18 - I】肝癌試料のデータセット (G S E 9 8 4 3) における本明細書において記載される広範な文献によるリストの標的遺伝子と比較される証拠により精選されたりリストの標的遺伝子を使用したベイジアンネットワークを使用し予測されるW n t 経路の活性を示した図である。

【図18 - I I】肝癌試料のデータセット (G S E 9 8 4 3) における本明細書において記載される広範な文献によるリストの標的遺伝子と比較される証拠により精選されたりリストの標的遺伝子を使用したベイジアンネットワークを使用し予測されるW n t 経路の活性を示した図である。

20

【図19 - I】髄芽腫試料のデータセット (G S E 1 0 3 2 7) における本明細書において記載される広範な文献によるリストの標的遺伝子と比較される証拠により精選されたりリストの標的遺伝子を使用したベイジアンネットワークを使用し予測されるW n t 経路の活性を示した図である。

【図19 - I I】髄芽腫試料のデータセット (G S E 1 0 3 2 7) における本明細書において記載される広範な文献によるリストの標的遺伝子と比較される証拠により精選されたりリストの標的遺伝子を使用したベイジアンネットワークを使用し予測されるW n t 経路の活性を示した図である。

【図20】本明細書において開示される (W n t 経路に対して例証的に示されている) 1 つ又は複数の細胞シグナル経路を評価するように構成される臨床判断支援 (C D S) システムを示した概略図である。

30

【図21】G S E 4 1 8 3 からの結腸試料における予測されるW n t 経路の活性を示した図である。

【図22】G S E 1 0 3 2 7 からの髄芽腫試料における予測されるW n t 経路の活性を示した図である。

【図23】G S E 9 8 4 3 からの肝癌試料における予測されるW n t 経路の活性を示した図である。

【図24】G S E 1 2 7 7 7 からの乳癌細胞株における予測されるW n t 経路の活性を示した図である。

【図25】G S E 1 2 7 7 7 からの乳癌細胞株における予測されるE R 経路の活性を示した図である。

40

【図26】G S E 1 2 2 7 6 からの乳癌試料における予測されるE R 経路の活性を示した図である。

【図27】G S E 3 6 1 3 3 からの癌細胞株における予測されるE R 経路の活性を示した図である。

【図28】G S E 3 4 2 1 1 からの癌細胞株における予測されるヘッジホッグ経路の活性を示した図である。

【図29】G S E 1 0 3 2 7 からの髄芽腫試料における予測されるヘッジホッグ経路の活性を示した図である。

【図30】G S E 1 2 2 7 6 からの乳癌試料における予測されるヘッジホッグ経路の活性

50

を示した図である。

【図31】GSE21618からのMCF7及びタモキシフェン耐性細胞株における予測されるER経路の活性を示した図である。

【図32】GSE11324からの時系列のエストロゲンにより刺激されたMCF7細胞株試料における予測されるER経路の活性を示した図である。

【図33】GSE12276のルミナル(luminal)A試料におけるWnt、ER及びヘッジホッグ経路の活性を示した図である。

【図34】GSE12276の基底試料(basal sample)におけるWnt、ER及びヘッジホッグ経路の活性を示した図である。

【図35】GSE20916からの結腸試料における予測されるWnt経路の活性を示した図である。

10

【図36】エストロゲンで刺激されたMCF7細胞株(E2)又は負の対照(EtOH)における予測されるER経路の活性を示した図である(GSE9253)。

【図37】経路の活性に従ってグループ化されたGSE12276データセットからの患者のカプラン・マイヤーの生存曲線を示した図である。

【図38】GSE7708からの異なる処理レジメで処理したLNCaP細胞株における予測されるAR経路の活性を示した図である。

【図39】GSE17951からの前立腺癌試料における予測されるAR経路の活性を示した図である。

【図40】GSE12276からの乳癌試料における予測されるAR経路の活性を示した図である。

20

【図41】種々の癌タイプを表す細胞株試料を含有するGSE36133データセットにおける予測されるAR経路の活性を示した図である。

【図42】種々の癌タイプを表す細胞株試料を含有するGSE34211データセットにおける予測されるAR経路の活性を示した図である。

【発明を実施するための形態】

【0041】

以下の例は、特に好ましい方法、及び、それに関して選択される態様を単に例示している。そこに提供される教示は、いくつかのテスト及び/又はキットを構築して、例えば、1つ又は複数の細胞シグナル経路の異常な活性を検出、予測、及び/又は診断するために使用してもよい。さらに、本明細書において記載される方法を使用すると、薬の処方有利に導くことができ、薬の予測並びに薬効(及び/又は副作用)のモニタリングをすることができ、薬物耐性を予測及びモニターして、例えば、(コンパニオン診断テストのような)行われることになる後の1つ又は複数のテストを選択することができる。以下の例は、本発明の範囲を限定するとして解釈されることはない。

30

【実施例1】

【0042】

ベイジアンネットワークの構築

本明細書で開示するように、確率モデル(例:図6に示す例示的なベイジアンモデル)を構築し、且つ多くの異なる標的遺伝子の発現と細胞シグナル経路の活性レベルとの間の条件付き確率の関係を取り入れることによって、かかるモデルは高い確度をもって、細胞シグナル経路の活性の決定に用いることができる。また、確率モデルは、条件付き確率を調整し且つ/又は新しいノードをモデルに追加して追加的な情報源を表すようにすることで、即座に更新されて、後の臨床研究で得られる追加的な知識を取り入れることができる。このようにして、確率モデルは、最新の医学知識を取り入れるよう適切に更新することができる。

40

【0043】

細胞シグナル経路を表すもっとも単純なベイジアンネットワークモデルの一つは、転写因子要素及び関連する標的遺伝子を含む2レベルモデルである(図1参照)。転写因子複合体要素は、転写因子複合体のレベルを表す。転写因子要素のタンパク質レベルは、転写

50

因子の標的遺伝子の多くの mRNA レベルに結びつけられる（当該例示的なベイジアンネットワークにおいては、標的遺伝子が3つのみ表示されているが、これらは転写因子が利用可能である場合に組織内に発現することが知られている）。多くの、ほとんどの又はすべての経路の標的遺伝子（Wnt、ER、ヘッジホッグ及びAR経路の場合、特に表1、表2、表3及び表4にそれぞれ言及されている標的遺伝子）は、TF要素によって同じように調節される。TF要素のレベルと標的遺伝子の mRNA レベルとの間の関係は、エッジによって、ベイジアンネットワークにおいてモデル化される。各標的遺伝子に対し、条件付き確率の分布が、遺伝子の mRNA レベルがいかに関与するTF要素のレベルに依存しているかを特定する。

【0044】

TF要素及び標的遺伝子のレベルは、多様に表されてもよい。1つの選択肢は、二進数離散化を使用してTF要素の「非存在」及び「存在」を、標的遺伝子の mRNA レベルの「減る」「増える」の状態を示すことである（図1参照）。TF要素及び標的遺伝子間の確率の関係は、条件付き確率の表によって表すことができる（同図参照）。レベルは、二進数離散化の代わりに、連続的なレベル値又は3つ以上の量子化レベル（例：標的遺伝子について、「減る」「標準」及び「増える」）を有する量子化値で表すこともできる。

【0045】

上述の単純なベイジアンネットワークの例示は、ベイジアンネットワークモデルの例示的な実施形態に過ぎない（図1）。一般的には、ベイジアンネットワークモデルは、エッジによって接続されたノードを有する非循環有向グラフを有する。各ノードは、間近の経路（又は、より一般的には細胞シグナル経路）に関する情報アイテムを表す。各経路要素ノードは、細胞シグナル経路のゲノム又はプロテオミクス要素を表す。例示的な例として、経路要素ノードは、タンパク質、タンパク質複合体、細胞シグナル経路の標的遺伝子から転写された標的 mRNA 分子、メチル化遺伝子、リン酸化タンパク質、リン酸化タンパク質複合体等のうち一つを表してもよいが、それらに限定されない。後述するとおり、例示のものに限定されないが、幾つかのほかの種類ノードがベイジアンネットワークに含まれて、特定の測定データ要素や遺伝子の変異の出現等の他のタイプの情報を表してもよい。

【0046】

経路の（活性状態又は非活性状態の）調節タンパク質を表す追加的な「上流」レベルが、かかるタンパク質のレベルの知識が臨床判断支援の勧告を決定する証明になる場合に、典型的に加えられる。例えば、ベイジアンネットワークにおいて転写因子の元のタンパク質又は転写因子の上流にある必須タンパク質を含むこと（図2参照）は、経路全体ではなくかかるタンパク質を特異的に標的とする薬物が利用可能な場合、有用であり得る。転写因子（TF）要素は、シグナル経路の大部分においてタンパク質複合体（すなわち、標的遺伝子からの転写を調節する機能を実行する特定の構造において、合わせて結合されているタンパク質の組み合わせ）であると考えられている。他の経路については、TF要素は単一のタンパク質であってよい。加えて、シグナル経路は、1つより多い転写因子を介して活性であってもよく、その結果、多数の転写因子を標的遺伝子に送り込む、より複雑なベイジアンネットワークとなる（多数の転写因子要素が標的遺伝子転写に影響を及ぼしている仮定図について、図3を参照）。かかる多数の転写因子のベイジアンネットワークは、1つのベイジアンネットワークに結合された複数の経路の組み合わせの結果であってもよい。

【0047】

標的遺伝子のさらに下流の追加的な情報ノードもまた、ベイジアンネットワークに含まれてもよい。この例示的な例が、タンパク質への標的遺伝子の mRNA の翻訳（図4）又は標的遺伝子の mRNA レベルの代替ノードとしての標的遺伝子のタンパク質レベルノードである（図5）。標的遺伝子の標的 mRNA 分子は、リボソーム分子との相互作用によって翻訳されて mRNA 分子と標的遺伝子とに対応するタンパク質を形成する。これは、タンパク質レベルでの標的遺伝子の発現である。たとえば、これらに限定はされないが、

10

20

30

40

50

質量分析、免疫組織化学、ゲル電気泳動技術によるタンパク質レベルの測定は、これらの標的タンパク質レベルの証拠として作用し得る。

【0048】

標的遺伝子の発現レベルは、マイクロアレイの対応するプローブセットの測定された強度に基づいて計算されてもよく、たとえば加算平均又は他の技術の他の手段（たとえば、RNA配列）によって計算されてもよい。幾つかの実施例において、この計算は、図6に示し本明細書に記載のとおり、使用される各プローブセットについてのノードでベイジアンネットワークを延ばし、且つ対応する標的遺伝子ノードから各「測定」ノードに及ぶエッジを含むことによって、ベイジアンネットワークに組み込まれる。

【0049】

確率モデルは、任意で、変異、コピー数多型、遺伝子発現、メチル化、転座情報などの、ゲノム配列を変更する追加のゲノム情報を取り入れてもよく、当該ゲノム配列は、経路のシグナルカスケードに関連して経路の活性を推量し且つ（活性又は不活性のいずれかの）機能異常を引き起こすWnt経路における欠陥を特定し、このことはメチル化及びコピー数データの例示的なケースについての図7の例示的な参照に記載の通りである。しかしながら、標的遺伝子に関する他の種類の情報は、同じように情報ノードに翻訳されることが理解される。かかるゲノム情報は、RNA配列およびSNP分析を介して利用可能であるが、これらに限定されない。

【0050】

また、本明細書に後述されるWnt、ER、AR及びヘッジホッグ経路に関する実施例が、例示的な例として記載されているが、本明細書に開示されている細胞シグナル経路分析のアプローチは、上述の経路に加えて他の細胞内シグナル経路にも容易に適用される。他の細胞内シグナル経路とは、細胞膜内に受容体を備える細胞内シグナル経路（例：NOTCH、HER2/P13K、TGFbeta、EGF、VEGFおよびTNF-NFkappaB細胞シグナル経路）及び細胞内に受容体を備える細胞内シグナル経路（例：プロゲステロン、レチノイン酸及びビタミンD細胞シグナル経路）等である。

【実施例2】

【0051】

機械学習方法の比較

二種類の機械学習技術の実行を、Wnt経路を例として互いに比較する。最近隣セントロイド法を用いたWnt活性の予測を、本発明によれば最適の方法、たとえばベイジアンネットワークを用いたものと比較する。

【0052】

上記で検証したとおり、ベイジアンネットワークアプローチを、確率論的アプローチにおいて存在する利点に基づいて選択した。確率論的アプローチは、利用可能な情報を「ソフト」、たとえば証明となる特性を呈する研究対象のパーセンテージ、及び条件付き確率の関係を用いた「ハード」形式のどちらにおいても採用することができる。加えて、確率モデルは、根本的な細胞シグナル経路の（包括的というよりはむしろ）部分的な知識に基づいて情報が取り入れられることを可能にし、この場合もやはり条件付き確率の表を用いる。

【0053】

ここで、他の機械学習方法、たとえば既知の方法である最近隣セントロイド分類（nearest centroid classification）と比較して、本発明者等は、既知の生物学的特性及びベイジアンネットワークを用いたソフトエビデンスの利用可能性を含んだやり方で値を加えることが実証される。最近隣セントロイド分類は、機械学習方法であり、訓練試料の各クラスについて平均プロファイル（=セントロイド）が計算され、次に、分類される試料について、ラベルが最近隣のセントロイドに基づいて予測される（最近隣のセントロイドのラベルが、予測結果である）。2つのセントロイドが、ベイジアンネットワークに用いられるプローブセットの同一リスト上で計算され、「Wntオン」及び「Wntオフ」セントロイドについては、GSE8671の同一のfRMA

10

20

30

40

50

処理データの腺腫試料及び正常な結腸試料にそれぞれ基づく。試料と2つのセントロイドとの間の2つのユークリッド距離の \log_2 比を続けて用いて、種々のデータセットから試料を分類して試料の分類を推量した。このことは、0の \log_2 比が試料の2つのセントロイドに対して等しい距離と対応し、値 > 0 が活性Wntシグナル伝達として分類された試料に相当するが、値 < 0 は、不活性Wntシグナル伝達を有すると識別された試料に相当することを意味する。

【0054】

ベイジアンネットワークを、図6及び本明細書に記載された手順と同様に構築した。Wntベイジアンネットワークの記載と同様、プローブセット間のエッジの条件付き確率の表及びそれらのそれぞれの遺伝子を、遺伝子発現情報データベース(Gene Expression Omnibus)(<http://www.ncbi.nlm.nih.gov/geo/>、最終アクセス日2011年7月13日)からのデータセットGSE8671からの32の正常な結腸試料及び32の腺腫試料のfRMA処理データを使用して、訓練した。訓練したベイジアンネットワークを次に、種々のデータセットに対してテストして、Wnt経路が「オン」である、すなわち活性であるという確率 $P(Wnt \text{ オン})$ を推量したが、これはWnt経路転写複合体が「存在」するという推量された確率と等しいと解釈する。

【0055】

訓練されたベイジアンネットワーク及び最近隣セントロイドモデルを、次に、種々のfRMA処理されたマイクロアレイデータセットに対してテストしてWnt経路が「オン」である確率を推量したが、それは $P(Wnt \text{ オン})$ 及び距離の \log_2 比によって測定した。ベイジアンネットワーク及び最近隣セントロイドモデルの結果のまとめを、図8及び図13に示す。読者は、2つの方法の出力メトリックは1対1の関係ではないが、方法における出力メトリックの相対的な大きさとサインは、比較可能であることに留意されたい。

【0056】

結腸(癌)試料(GSE20916、GSE4183)の圧倒的多数が、活性Wnt経路と不活性Wnt経路との間に等しく分類されるが、最近隣セントロイド法(偽陰性)において高いフラクシオンの誤って分類されたネガティブな試料を有したGSE15960は、除く。高いフラクシオンのこの偽陰性の認知は、他の癌のタイプにおいても維持される。これは特に乳癌試料(GSE12777、GSE21653)及び肝癌(GSE9843)にあてはまり、いくつかの例外を除いて、すべての試料が不活性のWnt経路を有すると予測されるが、基底乳癌及びCTNNB1肝癌試料の場合は正しくないことが知られている。一部の例では、例えばGSE15960において明白であるが、最近隣セントロイド分類の閾値を増減することで、分類を修正することができる。この考えの背景には、Wnt活性の閾値が、異なる組織タイプにおいて変更され得ることがある。しかしながら、このことは、他の組織タイプに適応可能な最近隣セントロイド法の追加的な訓練を伴う。ベイジアンネットワークモデルの強みの一つは、組織タイプに関して非特異的であることが確立されているため、組織用の訓練が要求されないことである。

【実施例3】

【0057】

標的遺伝子の選択

転写因子(TF)は、特定のDNA配列に結合することで標的遺伝子からの転写を調節することができるタンパク質の複合体(すなわち、特定の構造において、合わせて結合されているタンパク質の組み合わせ)または1つのタンパク質であり、それによりDNAからmRNAへの遺伝子情報の転写を制御する。転写複合体のこの作用により直接生成されたmRNAを、本明細書では「直接の標的遺伝子」と呼ぶ。経路の活性化は、より二次的な遺伝子転写をもたらしてもよく、これを「間接の標的遺伝子」と呼ぶ。以下、経路の活性とmRNAレベルとの間の直接リンクとして、直接の標的遺伝子を含むか又は該直接の標的遺伝子からなる(例示的な確率モデルとしての)ベイジアンネットワークモデルが望

10

20

30

40

50

ましいが、直接の標的遺伝子及び間接の標的遺伝子の区別が、必ずしも明白であるとは限らない。ここで、利用可能な文献データに基づくスコアリング機能を用いて、直接の標的遺伝子を選択する方法を提示する。とはいえ、限られた情報及び生物学的な変種及び不確実性のせいで、間接の標的遺伝子を誤って選択してしまうことは排除できない。

【0058】

特定経路の mRNA 標的遺伝子を、科学文献から選択したが、その選択は、証拠が蓄積された科学的実験のタイプに応じて、特定の標的遺伝子についての科学的証拠を格付けした格付け法を用いることによって行った。幾つかの実験的証拠は、例えばヘッジホッグ経路が活性であると知られている胚のマイクロアレイ上で mRNA が増加している、といった、遺伝子が標的遺伝子であることを示唆するに過ぎないものであるが、他の証拠は非常に強力であり得る。例えば、識別された経路の転写因子結合部位、及び、細胞における特定の経路の刺激の後のクロマチン免疫沈降 (ChIP) アッセイにおける当該部位の取得と、細胞株における経路の特定の刺激後の mRNA の増加との組み合わせである。

10

【0059】

特定経路の標的遺伝子を見つけるための実験を、科学文献において数種類識別することができる。

1. 経路 - 転写因子をゲノム上のその結合部位へ直接結合させる ChIP 実験が実証される。例：クロマチン免疫沈降 (ChIP) 技術を用いることによって、活性 Wnt 経路有り無し両方の結腸細胞株の DNA における、後に推定される機能的 TCF4 転写因子結合部位を、純粋にヌクレオチド配列に基づいて認識された結合部位の部分集合として、識別した。推定される機能性を、転写因子が DNA 結合部位に結合することが判明した ChIP 由来の証拠として、識別した。

20

2. 結合配列を含有する DNA 断片に転写因子を *in vitro* で結合することを示す、ゲルシフトアッセイ (EMSA)。EMSA に基づく証拠は *in vivo* 状況に翻訳することができないので、ChIP に基づく証拠と比べて強くない。

3. 経路の刺激及びマイクロアレイ上の mRNA プロファイルの測定又は RNA 配列決定の使用、経路 - 誘導性細胞株の使用、並びに誘導後に数回測定される mRNA プロファイルの測定であって、タンパク質への翻訳を妨げるシクロヘキシミドがある状態で測定され、それにより誘導された mRNA が直接の標的遺伝子と仮定される。

30

4. 3 と同様であるが、定量的 PCR を用いて mRNA の量を測定する。

5. 生物情報学アプローチを用いた、ゲノムにおける転写因子結合部位の識別。Wnt 経路の例：既知の TCF4 - カテニン転写因子 DNA 結合配列を用いて、ソフトウェアプログラムをヒトゲノム配列に対して実行し、遺伝子プロモータ領域及び他のゲノム領域の両方において、潜在的な結合部位を識別した。

6. 3 と同様であるが、シクロヘキシミドは存在しない。

7. 4 と同様であるが、シクロヘキシミドは存在しない。

8. 経路が活性であることが知られている特定の組織又は細胞試料の mRNA 発現プロファイリング。しかしながら、適切なネガティブコントロール条件は存在しない。

【0060】

最も単純な形式では、標的 mRNA を識別したこれらの各実験的アプローチにつき、潜在的な標的 mRNA 毎に 1 ポイント付与することができる。

40

【0061】

或いは、1 の技術に 1 ポイント、2 の技術に 2 つめのポイントを加える等といった意味で、ポイントを追加的に付与することができる。この相対的な格付けストラテジーを使用することで、最も信頼性のある標的遺伝子のリストを作成することができる。

【0062】

或いは、別の格付け法を用いて直接の標的遺伝子である可能性が高い標的遺伝子を識別してもよい。その方法は、*in vivo* の直接の標的遺伝子について最大限の証拠を提供する技術により高い点数を付与することによってなされ、上記リストにおいては実験的アプローチ 1) 7 から 2) に対する 8 ポイントを意味し、実験的アプローチ 8 に対して 1

50

ポイントに下がる。かかるリストを、「一般的標的遺伝子リスト」と称してもよい。

【0063】

生物学的な変種及び不確実性にもかかわらず、本発明者等は、直接の標的遺伝子は組織に依存しないやり方で誘導される可能性が高いと仮定した。これらの標的遺伝子のリストを、「証拠により精選された標的遺伝子リスト」と称してもよい。これらの精選された標的リストを用いて、異なった組織起源からの試料に適用可能な計算モデルを構築してきた。

【0064】

「一般的標的遺伝子リスト」は、恐らくより組織特異的な遺伝子を含み、且つ乳癌試料等の特異的組織からの試料で適用するためのモデルの感度及び特異性を最適化させるよう潜在的に用いることができる。

10

【0065】

以下に、証拠により精選された標的遺伝子リストの選択が、具体的にどのようにER経路に対して構築されたかという例示を示す。

【0066】

「モデル」のための入力として用いられるER標的遺伝子を選択する目的で、以下の3つの基準を用いた。

1. 遺伝子プロモータ/エンハンサー領域が、エストロゲン応答要素(ERE)モチーフを含有する。

a. EREモチーフは、例えば、特定のEREモチーフがレポーター遺伝子にリンクされている一過性導入アッセイを用いて、エストロゲンに応答することを証明されなければならない。

20

b. EREモチーフの存在は、例えば遺伝子プロモータ/エンハンサー領域の強化モチーフ分析によって確認されなければならない。

2. 例えばChIP/CHIP実験又はクロマチン免疫沈降アッセイによって実証されることによって、ERは(差次的に)、問題となっている遺伝子のプロモータ/エンハンサー領域に*in vivo*で結合する。

a. ER経路が活性であるときに、ERは遺伝子のプロモータ/エンハンサー領域に結合すると証明される。

b. ER経路が活性でない場合、(好ましくは)遺伝子の遺伝子プロモータ/エンハンサー領域に結合しない(又は弱く結合する)。

30

3. ER経路が活性であるとき、遺伝子は差次的に転写され、例えば、以下a. b. によって実証される。

a. リアルタイムPCR又はマイクロアレイ実験を介した、問題となっている遺伝子のmRNAのfold enrichment、または

b. RNA Pol IIが、免疫沈澱アッセイを介して遺伝子のプロモータ領域に結合することの実証。

【0067】

ER標的遺伝子であって、上述した3つの全ての基準を満たしたことを証明した、十分かつ適切に文書化された実験的証拠が集められたER標的遺伝子を定めることによって、選択を行った。ER差次的結合の証拠を収集するための好適な実験は、例えば、エストロゲンに暴露した場合とされない場合のエストロゲン(例: MCF-7細胞株)に反応する癌細胞株におけるChIP/CHIP実験結果を比較することである。mRNA転写の証拠の収集についても、同様である。

40

【0068】

以下は、包括的なアプローチ、及び上述したアプローチを用いて発見された証拠に基づいて多くの標的遺伝子を選択するために用いられてきた、標的遺伝子選択手順のより具体的な例について検討している。例示的な経路、すなわちWnt、ER、ヘッジホッグ及びAR経路についてベイジアンネットワークモデルにおいて用いられた標的遺伝子のリストを、表1、表2、表3及び表4にそれぞれ示す。

50

【 0 0 6 9 】

本明細書に記載の(表2参照)ER経路のベイジアンネットワークモデルに用いられたER経路の標的遺伝子は、文献証拠スコアに基づく標的遺伝子の選択を含有し、最も高い証拠スコアを有する標的遺伝子(本発明による好適な標的遺伝子)のみを、このショートリストに加えた。低い証拠スコアを有する遺伝子を含むER標的遺伝子の完全なリストを、表5に示す。

【 0 0 7 0 】

表1、表2、表3及び表4に示したWnt、ER、ヘッジホッグ及びAR経路の標的遺伝子のさらなる下位選択又は格付けを、プローブセットノードを対応する標的遺伝子ノードにリンクさせる訓練された条件付き確率の表を用いて計算したオッズ比と文献証拠スコアの組み合わせに基づいて行った。オッズ比は、経路の活性を推量するにあたっての標的遺伝子の重要性の評価である。一般的に、経路の全体的な活性について、より低いオッズ比を有する標的遺伝子と比べて、より高いオッズ比を有する標的遺伝子の発現レベルのようが、より参考になる可能性が高いことが予想される。しかしながら、細胞シグナル経路の複雑さのせいで、標的遺伝子と経路の活性との間により複雑な相関関係が存在し得ることが理解される。例えば、より高いオッズ比を有する標的遺伝子を単体で考慮するより、低いオッズ比を有する標的遺伝子の種々の組み合わせの発現レベルを考慮することが、より証明力があるかもしれない。本明細書で報告したWnt、ER、ヘッジホッグ及びARモデリングにおいて、表6、表7、表8及び表9に示された標的遺伝子が、Wnt、ER、ヘッジホッグ及びAR経路活性の予測に対して、より低い格付けの標的遺伝子と比べてより高い証明性があることが判った(したがって、本発明によると、表6及び9に示された標的遺伝子が、特に好ましい)。とはいえ、マイクロアレイなどの取得技術が多くの遺伝子の組についての発現レベルを取得できるという相対的な容易さを前提とすると、表6、表7、表8及び表9の幾つかもしくは全ての標的遺伝子を利用し、且つ任意で追加で、表1、表2、表3及び表4に示される1つ、2つ、幾つか又は全ての追加の標的遺伝子の格付けを、図6に記載のベイジアンモデルに用いることが考えられる。

10

20

【 0 0 7 1 】

【表 1】

表1 ベイジアンネットワークに用いられるWnt経路の標的遺伝子、及び標的遺伝子のmRNA発現レベルを測定するために用いられる関連するプローブセットの、証拠により精選されたリスト（# =添付の配列表における配列番号）

標的遺伝子	プローブセット	#	標的遺伝子	プローブセット	#
ADRA2C	206128_at	4	HNF1A	210515_at	102
ASCL2	207607_at	10		216930_at	
	229215_at		IL8	202859_x_at	110
AXIN2	222695_s_at	13		211506_s_at	
	222696_at		KIAA1199	1554685_a_at	119
	224176_s_at			212942_s_at	
	224498_x_at		KLF6	1555832_s_at	121
BMP7	209590_at	17		208960_s_at	
	209591_s_at			208961_s_at	
	211259_s_at			211610_at	
	211260_at			224606_at	
CCND1	208711_s_at	27	LECT2	207409_at	129
	208712_at		LEF1	210948_s_at	130
	214019_at			221557_s_at	
CD44	1557905_s_at	30		221558_s_at	
	1565868_at		LGR5	210393_at	131
	204489_s_at			213880_at	
	204490_s_at		MYC	202431_s_at	142
	209835_x_at			244089_at	
	210916_s_at		NKD1	1553115_at	150
	212014_x_at			229481_at	
	212063_at			232203_at	
	216056_at		OAT	201599_at	157
	217523_at		PPARG	208510_s_at	173
	229221_at		REG1B	205886_at	184
	234411_x_at		RNF43	218704_at	189
	234418_x_at		SLC1A2	1558009_at	200
COL18A1	209081_s_at	40		1558010_s_at	
	209082_s_at			208389_s_at	
DEFA6	207814_at	52		225491_at	
DKK1	204602_at	54	SOX9	202935_s_at	209
EPHB2	209588_at	67		202936_s_at	
	209589_s_at		SP5	235845_at	210

10

20

30

40

	210651_s_at		TBX3	219682_s_at	215
	211165_x_at			222917_s_at	
EPHB3	1438_at	68		225544_at	
	204600_at			229576_s_at	
FAT1	201579_at	72	TCF7L2	212759_s_at	219
FZD7	203705_s_at	90		212761_at	
	203706_s_at			212762_s_at	
GLUL	200648_s_at	95		216035_x_at	
	215001_s_at			216037_x_at	
	217202_s_at			216511_s_at	
	217203_at			236094_at	
	242281_at		TDGF1	206286_s_at	220
			ZNRF3	226360_at	248

10

【 0 0 7 2 】

20

【表 2】

表2 ベイジアンネットワークに用いられるER経路の標的遺伝子、及び標的遺伝子のmRNA発現レベルを測定するために用いられる関連するプローブセットの、証拠により精選されたリスト（# = 添付の配列表における配列番号）

標的遺伝子	プローブセット	#	標的遺伝子	プローブセット	#
AP1B1	205423_at	5	RARA	1565358_at	183
ATP5J	202325_s_at	12		203749_s_at	
COL18A1	209081_s_at	40		203750_s_at	
	209082_s_at			211605_s_at	
COX7A2L	201256_at	41		216300_x_at	
CTSD	200766_at	46	SOD1	200642_at	205
DSCAM	211484_s_at	59	TFF1	205009_at	221
	237268_at		TRIM25	206911_at	230
	240218_at			224806_at	
EBAG9	204274_at	61	XBP1	200670_at	244
	204278_s_at			242021_at	
ESR1	205225_at	70	GREB1	205862_at	97
	211233_x_at			210562_at	
	211234_x_at			210855_at	
	211235_s_at		IGFBP4	201508_at	106
	211627_x_at		MYC	202431_s_at	142
	215551_at			244089_at	
	215552_s_at		SGK3	227627_at	196

10

20

30

	217163_at			220038_at	
	217190_x_at		WISP2	205792_at	241
	207672_at		ERBB2	210930_s_at	69
HSPB1	201841_s_at	103		216836_s_at	
KRT19	201650_at	124		234354_x_at	
	228491_at		CA12	203963_at	22
NDUFV3	226209_at	148		204508_s_at	
	226616_s_at			204509_at	
NRIP1	202599_s_at	154		210735_s_at	
	202600_s_at			214164_x_at	
PGR	208305_at	162		215867_x_at	
	228554_at			241230_at	
PISD	202392_s_at	164	CDH26	232306_at	32
PRDM15	230553_at	174		233391_at	
	230777_s_at			233662_at	
	231931_at			233663_s_at	
	234524_at		CELSR2	204029_at	36
	236061_at			36499_at	
PTMA	200772_x_at	179			
	200773_x_at				
	208549_x_at				
	211921_x_at				

10

20

30

【 0 0 7 3 】

【表 3】

表 3 ベイジアンネットワークに用いられるヘッジホッグ経路の標的遺伝子、及び標的遺伝子の mRNA 発現レベルを測定するために用いられる関連するプローブセットの、証拠により精選されたリスト（# = 添付の配列表における配列番号）

標的遺伝子	プローブセット	#	標的遺伝子	プローブセット	#
GLI1	206646_at	93	CTSL1	202087_s_at	47
PTCH1	1555520_at	177	TCEA2	203919_at	216
	208522_s_at			238173_at	
	209815_at			241428_x_at	
	209816_at		MYLK	1563466_at	145
	238754_at			1568770_at	
PTCH2	221292_at	178		1569956_at	
HHIP	1556037_s_at	101		202555_s_at	
	223775_at			224823_at	
	230135_at		FYN	1559101_at	88
	237466_s_at			210105_s_at	
SPP1	1568574_x_at	212		212486_s_at	
	209875_s_at			216033_s_at	
TSC22D1	215111_s_at	232	PITRM1	205273_s_at	165
	235315_at			239378_at	
	243133_at		CFLAR	208485_x_at	37
	239123_at			209508_x_at	
CCND2	200951_s_at	28		209939_x_at	
	200952_s_at			210563_x_at	
	200953_s_at			210564_x_at	
	231259_s_at			211316_x_at	
H19	224646_x_at	253		211317_s_at	
	224997_x_at			211862_x_at	
IGFBP6	203851_at	107		214486_x_at	
TOM1	202807_s_at	229		214618_at	
JUP	201015_s_at	117		217654_at	
FOXA2	210103_s_at	82		235427_at	
	214312_at			237367_x_at	
	40284_at			239629_at	
MYCN	209756_s_at	144		224261_at	
	209757_s_at		IL1R2	205403_at	108
	211377_x_at			211372_s_at	

10

20

30

40

	234376_at		S100A7	205916_at	254
	242026_at		S100A9	203535_at	255
NKX2_2	206915_at	249	CCND1	208711_s_at	27
NKX2_8	207451_at	250		208712_at	
RAB34	1555630_a_at	182		214019_at	
	224710_at		JAG2	209784_s_at	115
MIF	217871_s_at	134		32137_at	
GLI3	1569342_at	94	FOXM1	202580_x_at	85
	205201_at		FOXF1	205935_at	83
	227376_at		FOXL1	216572_at	84
FST	204948_s_at	87		243409_at	
	207345_at				
	226847_at				
BCL2	203684_s_at	14			
	203685_at				
	207004_at				
	207005_s_at				

10

20

【 0 0 7 4 】

【表 4】

表 4 ベイジアンネットワークに用いられるAR経路の標的遺伝子、及び標的遺伝子の mRNA発現レベルを測定するために用いられる関連するプローブセットの、証拠により精選されたリスト（# =添付の配列表における配列番号）

標的遺伝子	プローブセット	#	標的遺伝子	プローブセット	#
ABCC4	1554918_a_at	2	LCP1	208885_at	128
	1555039_a_at		LRIG1	211596_s_at	132
	203196_at			238339_x_at	
APP	200602_at	7	NDRG1	200632_s_at	147
	211277_x_at		NKX3_1	209706_at	251
	214953_s_at			211497_x_at	
AR	211110_s_at	8		211498_s_at	
	211621_at		NTS	206291_at	155
	226192_at		PLAU	205479_s_at	167
	226197_at			211668_s_at	
CDKN1A	1555186_at	34	PMEPA1	217875_s_at	169
	202284_s_at			222449_at	
CREB3L4	226455_at	42		222450_at	
DHCR24	200862_at	53	PPAP2A	209147_s_at	171
DRG1	202810_at	58		210946_at	

10

20

EAF2	1568672_at	60	PRKACB	202741_at	175
	1568673_s_at			202742_s_at	
	219551_at			235780_at	
ELL2	214446_at	65	KLK3	204582_s_at	123
	226099_at			204583_x_at	
	226982_at		PTPN1	202716_at	180
FGF8	208449_s_at	75		217686_at	
FKBP5	204560_at	77	SGK1	201739_at	195
	224840_at		TACC2	1570025_at	214
	224856_at			1570546_a_at	
GUCY1A3	221942_s_at	99		202289_s_at	
	227235_at			211382_s_at	
	229530_at		TMPRSS2	1570433_at	225
	239580_at			205102_at	
IGF1	209540_at	105		211689_s_at	
	209541_at			226553_at	
	209542_x_at		UGT2B15	207392_x_at	236
	211577_s_at			216687_x_at	
KLK2	1555545_at	122			
	209854_s_at				
	209855_s_at				
	210339_s_at				

10

20

30

【 0 0 7 5 】

【表 5】

表 5 顕著な文献証拠を有すると認められる E R 標的遺伝子の遺伝子記号
 (= E R 標的遺伝子ロングリスト) (# = 添付の配列表における配列番号)

遺伝子記号	#	遺伝子記号	#	遺伝子記号	#	遺伝子記号	#
AP1B1	5	SOD1	205	MYC	142	ENSA	66
COX7A2L	41	TFF1	221	ABCA3	1	KIAA0182	118
CTSD	46	TRIM25	230	ZNF600	247	BRF1	19
DSCAM	59	XBP1	245	PDZK1	160	CASP8AP2	25
EBAG9	61	GREB1	97	LCN2	127	CCNH	29
ESR1	70	IGFBP4	106	TGFA	222	CSDE1	43
HSPB1	103	SGK3	196	CHEK1	38	SRSF1	213
KRT19	124	WISP2	241	BRCA1	18	CYP1B1	48
NDUFV3	148	ERBB2	69	PKIB	166	FOXA1	81
NRIP1	154	CA12	22	RET	188	TUBA1A	235
PGR	162	CELSR2	36	CALCR	23	GAPDH	91
PISD	164	CDH26	32	CARD10	24	SFI1	194
PRDM15	174	ATP5J	12	LRIG1	132	ESR2	258
PTMA	179	COL18A1	40	MYB	140	MYBL2	141
RARA	183	CCND1	27	RERG	187		

10

20

【 0 0 7 6 】

30

【表 6】

表 6 文献証拠スコア及びオッズ比に基づく Wn t 標的遺伝子のショートリスト
 (# = 添付の配列表における配列番号)

標的遺伝子	#
KIAA1199	119
AXIN2	13
CD44	30
RNF43	189
MYC	142
TBX3	215
TDGF1	220
SOX9	209
ASCL2	10
IL8	110
SP5	210
ZNRF3	248
EPHB2	67
LGR5	131
EPHB3	68
KLF6	121
CCND1	27
DEFA6	52
FZD7	90

10

20

30

【 0 0 7 7 】

【表 7】

表 7 文献証拠スコア及びオッズ比に基づく ER 標的遺伝子のショートリスト
 (# = 添付の配列表における配列番号)

標的遺伝子	#
CDH26	32
SGK3	196
PGR	162
GREB1	97
CA12	22
XBP1	244
CELSR2	36
WISP2	241
DSCAM	59
ERBB2	69
CTSD	46
TFF1	221
NRIP1	154

10

20

【 0 0 7 8 】

【表 8】

表 8 文献証拠スコア及びオッズ比に基づくヘッジホッグ標的遺伝子のショートリスト（# = 添付の配列表における配列番号）

標的遺伝子	#
GLI1	93
PTCH1	177
PTCH2	178
IGFBP6	107
SPP1	212
CCND2	28
FST	87
FOXL1	84
CFLAR	37
TSC22D1	232
RAB34	182
S100A9	255
S100A7	254
MYCN	144
FOXM1	85
GLI3	94
TCEA2	216
FYN	88
CTSL1	47

10

20

30

【 0 0 7 9 】

【表 9】

表 9 文献証拠スコア及びオッズ比に基づく AR 標的遺伝子のショートリスト
(# = 添付の配列表における配列番号)

標的遺伝子	#
KLK2	122
PMEPA1	169
TMPRSS2	225
NKX3_1	251
ABCC4	2
KLK3	123
FKBP5	77
ELL2	65
UGT2B15	236
DHCR24	53
PPAP2A	171
NDRG1	147
LRIG1	132
CREB3L4	42
LCP1	128
GUCY1A3	99
AR	8
EAF2	60

10

20

30

【実施例 4】

【0080】

証拠により精選されたリストと広範な文献によるリストとの比較

本明細書において記載される手順に従い文献証拠に基づき構築された Wnt 標的遺伝子のリスト(表 1)が、上述の手順に従わない別の標的遺伝子のリストと比較される。もう一つの別のリストは、分子生物学及び Wnt 経路の領域における専門家には既知の名高い研究室によって 3 つの公然の情報源において発表された Wnt 標的遺伝子になることになる、いろいろな実験的アプローチからの種々のデータによって示された遺伝子の寄せ集めである。もう一つの別のリストは、Hatzis 等(Hatzis P、2008)からの表 S3、de Sousa e Melo(de Sousa E Melo F、2011)からのテキスト及び表 S1A、並びに、Wnt シグナルの分野におけるパイオニアである Roel Nusse によって集められ且つ主張された標的遺伝子のリスト(Nusse、2012)に記載の遺伝子の組み合わせである。これら 3 つの情報源の組み合わせは、124 個の遺伝子のリストを生じた(= 広範な文献によるリスト、表 10 を参照)。ここでは、このもう一つの別のリストから得られるアルゴリズムによる臨床試料における Wnt 活性を予測することの実行が、既存の遺伝子のリスト(= 証拠により精選されたリスト、表 1)に基づき構築されたモデルと比較して同様に又はより優れて実行するこ

40

50

とかどうかという問題が論議されている。

【 0 0 8 1 】

【表 10】

表 10 Wnt 標的遺伝子のもう1つの別のリスト (=広範な文献リスト)
(# =添付の配列表における配列番号)

標的遺伝子	参照	#	標的遺伝子	参照	#
ADH6	de Sousa e Melo et al.	3	L1CAM	Nusse	125
ADRA2C	Hatzis et al.	4	LBH	Nusse	126
APCDD1	de Sousa e Melo et al.	6	LEF1	Hatzis et al., de Sousa e Melo et al., Nusse	130
ASB4	de Sousa e Melo et al.	9	LGR5	de Sousa e Melo et al., Nusse	131
ASCL2	Hatzis et al., de Sousa e Melo et al.	10	LOC283859	de Sousa e Melo et al.	260
ATOH1	Nusse	11	MET	Nusse	133
AXIN2	Hatzis et al., de Sousa e Melo et al., Nusse	13	MMP2	Nusse	135
BIRC5	Nusse	15	MMP26	Nusse	136
BMP4	Nusse	16	MMP7	Nusse	137
BMP7	Hatzis et al.	17	MMP9	Nusse	138
BTRC	Nusse	20	MRPS6	Hatzis et al.	139
BZRAP1	de Sousa e Melo et al.	21	MYC	Hatzis et al., Nusse	142
SBSPON	de Sousa e Melo et al.	259	MYCBP	Nusse	143
CCL24	de Sousa e Melo et al.	26	MYCN	Nusse	144
CCND1	Nusse	27	NANOG	Nusse	146
CD44	Nusse	30	NKD1	de Sousa e Melo et al.	150
CDH1	Nusse	31	NOS2	Nusse	151
CDK6	Hatzis et al.	33	NOTUM	de Sousa e Melo et al.	152
CDKN2A	Nusse	35	NRCAM	Nusse	153
CLDN1	Nusse	39	NUAK2	Hatzis et al.	156
COL18A1	Hatzis et al.	40	PDGFB	Hatzis et al.	159
CTLA4	Nusse	44	PFDN4	Hatzis et al.	161
CYP4X1	de Sousa e Melo et al.	49	PLAUR	Nusse	168
CYR61	Nusse	50	POU5F1	Nusse	170
DEFA5	de Sousa e Melo et al.	51	PPARD	Nusse	172
DEFA6	de Sousa e Melo et al.	52	PROX1	de Sousa e Melo et al.	176
DKK1	de Sousa e Melo et al., Nusse	54	PTPN1	Hatzis et al.	180

10

20

30

40

DKK4	de Sousa e Melo et al.	55	PTTG1	Nusse	181
DLL1	Nusse	56	REG3A	de Sousa e Melo et al.	185
DPEP1	de Sousa e Melo et al.	57	REG4	de Sousa e Melo et al.	186
EDN1	Nusse	62	RPS27	Hatzis et al.	190
EGFR	Nusse	64	RUNX2	Nusse	191
EPHB2	Hatzis et al., de Sousa e Melo et al., Nusse	67	SALL4	Nusse	192
EPHB3	Hatzis et al., Nusse	68	SLC1A1	de Sousa e Melo et al.	199
ETS2	Hatzis et al.	71	SLC7A5	Hatzis et al.	201
FAT1	Hatzis et al.	72	SNAI1	Nusse	202
FGF18	Nusse	73	SNAI2	Nusse	203
FGF20	Nusse	74	SNAI3	Nusse	204
FGF9	Nusse	76	SIK1	Hatzis et al.	261
FLAD1	Hatzis et al.	78	SOX17	Nusse	206
AK122582	Hatzis et al.	262	SOX2	de Sousa e Melo et al.	207
FN1	Nusse	79	SOX4	Hatzis et al.	208
FOSL1	Nusse	80	SOX9	Nusse	209
FOXP1	Nusse	86	SP5	Hatzis et al., de Sousa e Melo et al.	210
FST	Nusse	87	SP8	Hatzis et al.	211
FZD2	de Sousa e Melo et al.	89	TCF3	Nusse	217
FZD7	Nusse	90	TDGF1	Hatzis et al.	220
GAST	Nusse	92	TIAM1	Nusse	224
GMDS	Hatzis et al.	96	TNFRSF19	Nusse	227
GREM2	Nusse	98	TNFSF11	Nusse	228
HES6	Hatzis et al.	100	TRIM29	de Sousa e Melo et al.	231
HNF1A	Nusse	102	TSPAN5	de Sousa e Melo et al.	233
ID2	Nusse	104	TTC9	de Sousa e Melo et al.	234
IL22	de Sousa e Melo et al.	109	VCAN	Nusse	237
IL8	Nusse	110	VEGFA	Nusse	238
IRX3	de Sousa e Melo et al.	111	VEGFB	Nusse	239
IRX5	de Sousa e Melo et al.	112	VEGFC	Nusse	240
ISL1	Nusse	113	WNT10A	Hatzis et al.	242
JAG1	Nusse	114	WNT3A	Nusse	243
JUN	Nusse	116	ZBTB7C	de Sousa e Melo et al.	246
KIAA1199	de Sousa e Melo et al.	119	PATZ1	Hatzis et al.	263
KLF4	Hatzis et al.	120	ZNRF3	Hatzis et al.	248

hip Human Genome U133 Plus 2.0アレイのプローブセットを見つけることから成り立った。このプロセスは、Rのバイオコンダクタープラグイン (Bioconductor plugin)、及び、UCSCゲノムブラウザに基づくプローブセットの関連性に対する手動の精選を使用して行い、その結果、例えば反対の鎖上又は外側の遺伝子エクソン領域のプローブセットを除去した。124個の遺伝子のうち2つの遺伝子に対して、このマイクロアレイチップ上で利用可能なプローブセットがなく、従って、ベイジアンネットワークに挿入することはできなかった。これらの遺伝子は、LOC283859及びWNT3Aである。全部で287個のプローブセットが、残りの122個の遺伝子に相当すると分かった (表11)。

【0082】

【表 1 1】

表 1 1 広範な文献遺伝子リストにおける W n t 標的遺伝子と付随するプローブセット
(# = 添付の配列表における配列番号)

遺伝子 記号	プローブ セット	#	遺伝子 記号	プローブ セット	#	遺伝子 記号	プローブ セット	#
ADH6	207544_s_at	3	FAT1	201579_at	72	PFDN4	205360_at	161
	214261_s_at		FGF18	206987_x_at	73		205361_s_at	
ADRA2C	206128_at	4		211029_x_at			205362_s_at	
APCDD1	225016_at	6		211485_s_at		PLAUR	210845_s_at	168
ASB4	208481_at	9		231382_at			211924_s_at	
	217228_s_at		FGF20	220394_at	74		214866_at	
	217229_at		FGF9	206404_at	76	POU5F1	208286_x_at	170
	235619_at			239178_at		PPARD	208044_s_at	172
	237720_at		FLAD1	205661_s_at	78		210636_at	
	237721_s_at			212541_at			37152_at	
ASCL2	207607_at	10	AK122582	235085_at	262		242218_at	
	229215_at		FN1	1558199_at	79	PROX1	207401_at	176
ATOH1	221336_at	11		210495_x_at			228656_at	
AXIN2	222695_s_at	13		211719_x_at		PTPN1	202716_at	180
	222696_at			212464_s_at			217686_at	
	224176_s_at			214701_s_at			217689_at	
	224498_x_at			214702_at		PTTG1	203554_x_at	181
BIRC5	202094_at	15		216442_x_at		REG3A	205815_at	185
	202095_s_at		FOSL1	204420_at	80		234280_at	
	210334_x_at		FOXN1	207683_at	86	REG4	1554436_a_at	186
BMP4	211518_s_at	16	FST	204948_s_at	87		223447_at	
BMP7	209590_at	17		207345_at		RPS27	200741_s_at	190
	209591_s_at			226847_at		RUNX2	216994_s_at	191
	211259_s_at		FZD2	210220_at	89		221282_x_at	
	211260_at			238129_s_at			232231_at	
BTRC	1563620_at	20	FZD7	203705_s_at	90		236858_s_at	
	204901_at			203706_s_at			236859_at	
	216091_s_at		GAST	208138_at	92	SALL4	229661_at	192
	222374_at		GMDS	204875_s_at	96	SLC1A1	206396_at	199
	224471_s_at			214106_s_at			213664_at	
BZRAP1	205839_s_at	21	GREM2	220794_at	98	SLC7A5	201195_s_at	201
SBSPON	214725_at	259		235504_at		SNAI1	219480_at	202
	235209_at			240509_s_at		SNAI2	213139_at	203

10

20

30

40

	235210_s_at		HES6	226446_at	100	SNAI3	1560228_at	204
CCL24	221463_at	26		228169_s_at		SIK1	208078_s_at	261
CCND1	208711_s_at	27	HNFI1A	210515_at	102		232470_at	
	208712_at			216930_at		SOX17	219993_at	206
	214019_at		ID2	201565_s_at	104		230943_at	
CD44	1557905_s_at	30		201566_x_at		SOX2	213721_at	207
	204489_s_at			213931_at			213722_at	
	204490_s_at		IL22	221165_s_at	109		228038_at	
	209835_x_at			222974_at		SOX4	201416_at	208
	210916_s_at		IL8	202859_x_at	110		201417_at	
	212014_x_at			211506_s_at			201418_s_at	
	212063_at		IRX3	229638_at	111		213668_s_at	
	217523_at		IRX5	210239_at	112	SOX9	202935_s_at	209
	229221_at		ISL1	206104_at	113		202936_s_at	
CDH1	201130_s_at	31	JAG1	209097_s_at	114	SP5	235845_at	210
	201131_s_at			209098_s_at		SP8	237449_at	211
	208834_x_at			209099_x_at			239743_at	
CDK6	207143_at	33		216268_s_at		TCF3	209151_x_at	217
	214160_at		JUN	201464_x_at	116		209152_s_at	
	224847_at			201465_s_at			209153_s_at	
	224848_at			201466_s_at			210776_x_at	
	224851_at		KIAA1199	1554685_a_at	119		213730_x_at	
	231198_at			212942_s_at			213811_x_at	
	235287_at		KLF4	220266_s_at	120		215260_s_at	
	243000_at			221841_s_at			216645_at	
CDKN2A	207039_at	35	L1CAM	204584_at	125	TDGF1	206286_s_at	220
	209644_x_at			204585_s_at		TIAM1	206409_at	224
	211156_at		LBH	221011_s_at	126		213135_at	
CLDN1	218182_s_at	39	LEF1	210948_s_at	130	TNFRSF19	223827_at	227
	222549_at			221557_s_at			224090_s_at	
COL18A1	209081_s_at	40		221558_s_at		TNFSF11	210643_at	228
	209082_s_at		LGR5	210393_at	131		211153_s_at	
CTLA4	221331_x_at	44		213880_at		TRIM29	202504_at	231
	231794_at		MET	203510_at	133		211001_at	
	234362_s_at			211599_x_at			211002_s_at	
	236341_at			213807_x_at		TSPAN5	209890_at	233
CYP4X1	227702_at	49		213816_s_at			213968_at	
CYR61	201289_at	50	MMP2	1566678_at	135		225387_at	

10

20

30

40

	210764_s_at			201069_at			225388_at	
DEFA5	207529_at	51	MMP26	220541_at	136	TTC9	213172_at	234
DEFA6	207814_at	52	MMP7	204259_at	137		213174_at	
DKK1	204602_at	54	MMP9	203936_s_at	138	VCAN	204619_s_at	237
DKK4	206619_at	55	MRPS6	224919_at	139		204620_s_at	
DLL1	224215_s_at	56	MYC	202431_s_at	142		211571_s_at	
	227938_s_at		MYCBP	203359_s_at	143		215646_s_at	
DPEP1	205983_at	57		203360_s_at			221731_x_at	
EDN1	218995_s_at	62		203361_s_at		VEGFA	210512_s_at	238
	222802_at		MYCN	209756_s_at	144		210513_s_at	
EGFR	1565483_at	64		209757_s_at			211527_x_at	
	1565484_x_at			211377_x_at			212171_x_at	
	201983_s_at			234376_at		VEGFB	203683_s_at	239
	201984_s_at		NANOG	220184_at	146	VEGFC	209946_at	240
	210984_x_at		NKD1	1553115_at	150	WNT10A	223709_s_at	242
	211550_at			229481_at			229154_at	
	211551_at			232203_at		ZBTB7C	217675_at	246
	211607_x_at		NOS2	210037_s_at	151	ZBTB7C	227782_at	246
EPHB2	209588_at	67	NOTUM	228649_at	152	PATZ1	209431_s_at	263
	209589_s_at		NRCAM	204105_s_at	153		211391_s_at	
	210651_s_at			216959_x_at			210581_x_at	
	211165_x_at		NUAK2	220987_s_at	156		209494_s_at	
EPHB3	1438_at	68	PDGFB	204200_s_at	159	ZNRF3	226360_at	248
	204600_at			216061_x_at				
ETS2	201328_at	71		217112_at				
	201329_s_at							

10

20

30

後に、図 6 及び本明細書において説明される手順と同様にベイジアンネットワークを構築した。証拠により精選されたリストに基づく W n t ベイジアンネットワークの記載と同様に、証拠により精選されたリストも広範な文献によるリストも、プローブセットとそのそれぞれの遺伝子との間のエッジの条件付き確率の表を、(2011年7月13日に最後にアクセスした <http://www.ncbi.nlm.nih.gov/geo/> にて入手可能な) 遺伝子発現オムニバスからのデータセット GSE8671 由来の 32 の正常な結腸試料及び 32 の腺腫試料の fRMA 処理データを使用して訓練した。

40

【0083】

訓練したベイジアンネットワークを、次に、種々のデータセットに対してテストし、W n t 経路が「オン」、すなわち活性である確率 $P(W n t \text{ オン})$ を推量し、該確率は、W n t 経路転写複合体が「存在」という推量した確率に等しいと解される。訓練した広範な文献によるモデル及び証拠により精選されたモデルの要約された結果が、図 14 乃至 19 において示されている。

【0084】

50

明らかに、広範な文献によるモデルは、一般的に、オン又はオフであるWntシグナルに対してより極端な確率を予測するということが推定することができる。加えて、もう1つの別のモデルは、結腸癌のデータセット(GSE20916、GSE4183、GSE15960)に対して類似の結果を予測するが、予想を超える試料が、乳癌(GSE12777)、肝癌(GSE9843)、及び髄芽腫試料(GSE10327)のデータセットにおいて予測された活性のWntシグナルを有する。

【0085】

結論として、広範な文献による標的遺伝子のリストは、一方で結腸癌におけるWnt活性をほぼ等しく優れて予測するが、もう一方で他の癌タイプにおいてはより劣って予測する(偽陽性が多すぎる)。これは、特異的に結腸細胞に偏り過ぎている、従って組織特異的過ぎるもう1つの別の標的遺伝子のリストの結果であり得る。de Sousa E Melo等もHatzis等も、主な関心は、非結腸特異的Wnt標的遺伝子が含まれ得るけれども、結腸直腸癌であった。加えて、これらのリストにおそらく含まれる非Wnt特異的な標的遺伝子は、他の癌タイプにおける悪化したWnt活性の予測の原因であり得る。もう1つの別のリストは、より間接的に制御される標的遺伝子を含有する可能性が高く、おそらくこのリストをより組織特異的にする。もともとのリストは、全ての組織においてWnt感受性であり、従って、組織特異性を減らす遺伝子を表す可能性が最もある直接的な標的遺伝子を含有することに対して調整される。

10

【実施例5】

【0086】

ベイジアンネットワークの訓練及び使用

ベイジアンネットワーク使用して、テスト試料における経路の活性を推量することができる前に、ネットワーク要素間の確率の関係を描くパラメータを決定しなければならない。さらに、入力する測定値が分散した状態である場合に、離散化の仕方を描く閾値を設定しなければならない。

20

【0087】

典型的に、ベイジアンネットワークは、好ましくは全ネットワークノードの全ての状態が既知である代表的な訓練試料のセットを使用して訓練される。しかし、モデルをすることになる経路のどの活性状態であるのか既知である多くの異なる種類の癌から訓練試料を得ることは非実用的である。結果として、入手可能な訓練セットは、限定された数の試料から成り、典型的には1つのタイプの癌のみからの試料から成る。ベイジアンネットワークが他のタイプの試料によく一般化するのを可能にするために、従って、パラメータが決定される方法に特別注意を払わなくてはならず、好ましくは本明細書において記載されるアプローチにおいて以下のように行われる。

30

【0088】

TFノードに対しては、「非存在」及び「存在」という状態にある(無条件の)確率は、多くの試料のセットにおける予想される出現によって与えられる。或いは、正の結果又は負の結果にたいしていかなる偏りも有さないように、図1において行われているように0.5に設定することができる。

【0089】

標的遺伝子ノードに対しては、条件付き確率が、図1のように設定される。TF要素が「非存在」である場合、標的遺伝子は「減る」可能性が最も高く、従って、0.95の確率がこれに対して選ばれ、標的遺伝子が「増える」確率は0.05である。後者の(ゼロではない)確率は、他の因子によって標的遺伝子が制御されるか、又は、(例えば測定ノイズのために)偶然に「増える」が観察される(まれな)可能性に相当することになる。TF要素が「存在」する場合、従って、0.70というまずまずの確率を有して標的遺伝子は「増え」、0.30という確率を有して標的遺伝子は「減る」。後者の値は、例えば遺伝子のプロモータ領域がメチル化されているため等、TF要素は存在するけれども標的遺伝子が強く発現されない理由がいくつかあり得るため、このように選ばれる。標的遺伝子がTF要素によって上方制御されないが下方制御される場合、確率は同様に選ばれ

40

50

るが、TF要素の存在による下方制御を反映している。

【0090】

プローブセットの強度が入力測定値を編成する図6において与えられているベイジアンネットワークモデルに対して、最終的に、プローブセットの強度をそれぞれの標的遺伝子のmRNAレベルに関係づける離散化及び条件付き確率の表に対するパラメータを決定しなければならない。これらのうちどちらも、本発明における訓練データに基づいている。プローブセットの強度のレベルの、「低い」及び「高い」という状態への離散化に対して、経路が活性化されている訓練試料（「オン」試料）のセットにおける強度値を、経路が活性化されていない訓練試料（「オフ」試料）のセットにおける強度値から最良に分ける適した閾値が決定される。最後に、それぞれの標的遺伝子の「減る」又は「増える」という状態に応じて「低い」又は「高い」強度を有するというプローブセットの確率を記載する条件付き確率の表が、それぞれの閾値の下及び上のプローブセットの強度値を用いて「オン」及び「オフ」試料の数を数えることによって行われる。これは、頻度論的アプローチとして文献において既知である。ゼロの値を有した条件付き確率の表への移行を防ぐ、ベイジアンネットワークの極端なふるまいを防ぐために、ダミーの数値（dummy count）が各群に加えられる。

10

【0091】

図6のベイジアンネットワークを考慮し、且つ、プローブセットに関するマイクロアレイ測定が利用可能であると仮定して、ベイジアンネットワークが訓練された後に、以下のようにテスト試料に対して適用することができる。第一のステップは、上記のようにテスト試料における各プローブセットの強度をそれぞれの閾値と比較することによって、入力測定値を離散化することである。この比較は、各プローブセットを「低い」若しくは「高い」強度に設定したハードな方法（「ハードエビデンス（hard evidence）」と呼ばれる）で行ってもよく、又は、測定におけるある不確かな状態（ノイズ）を仮定し、各プローブセットに対して、「低い」若しくは「高い」という確率を設定したソフトな方法（「ソフトエビデンス（soft evidence）」と呼ばれる）で行うことができる。例えば、閾値よりすぐ下の強度を有したプローブセットのソフトエビデンスは、ノイズの適した推定及び閾値との差に基づき0.8の「低い」という確率、及び、0.2の「高い」という確率であってもよい。

20

【0092】

次に、このハードエビデンス又はソフトエビデンスが、例えばジャンクションツリーアルゴリズムに基づき、ベイジアンネットワークに対する適した推量エンジンに供給される（（Neapolitan, 2004）を参照）。そのようなエンジンは、従って、「非存在」又は「存在」である、TF要素の更新された確率を推量することができ、提供される証拠が与えられる。「存在」であるTF要素の推量された確率は、次に、それぞれの経路が活性であると推定された確率として解釈される。

30

【0093】

好ましくは、Wnt、ER、ヘッジホッグ及びAR経路のベイジアンネットワークモデルの訓練は、（上記の<http://www.ncbi.nlm.nih.gov/geo/>にて入手可能な）Gene Expression Omnibusで入手可能な公衆データを使用して行われる。

40

【0094】

Wntベイジアンネットワークを、不活性のWnt経路を有するとして考慮される32の正常な結腸試料、及び、活性のWnt経路を有するとして知られる32の確認済み腺腫試料（GSE8671データセット）を使用して例証的に訓練した。

【0095】

ER経路のベイジアンネットワークモデルを、同様にGene Expression Omnibusにて入手可能なGSE8597データセット由来の、不活性のER経路を有するとして知られる4つのエストロゲン欠乏性MCF7試料、及び、活性のER経路を有するとしてみなされた4つのエストロゲンで刺激されたMCF7試料を使用して例証

50

的に訓練した。

【0096】

ヘッジホッグ経路のベイジアンネットワークモデルを、GSE7553データセットにおいて入手可能な、活性のヘッジホッグ経路を有するとして確認された15の基底細胞癌試料、及び、不活性のヘッジホッグ経路を有する試料を表す4つの正常な皮膚細胞試料を使用して例証的に訓練した。

【0097】

AR経路のベイジアンネットワークモデルを、正のAR活性を有する3つの試料、ジヒドロテストステロン(DHT)で刺激されたLNCaP細胞株、強力なAR経路活性化因子、及び、不活性のAR経路のケースを表す3つの無刺激のLNCaP細胞株を使用して例証的に訓練した。

10

【0098】

図35及び図36を参考にすると、訓練したWnt及びER経路のベイジアンネットワークモデルを使用して、本明細書に記載の訓練手順において使用されていない類似の試料(Wnt及びERベイジアンネットワークに対して、それぞれ結腸試料及びMCF7乳癌細胞株)における経路の活性を予測した(ヘッジホッグベイジアンネットワークに対する適切なデータセットは見つからなかった)。大多数の試料の予測した経路の活性は、検証されることになるモデルに対して臨床的に予想される経路の活性と一致しているべきである。

【0099】

図35は、GSE20916データセットにおける、バーの色によって示された分類によってグループ化された結腸試料の予測したWnt活性を示しており、試料に対して、縦軸上のロジットP(Wntオン)として描かれており、横軸上のバーによって例示されている。全ての正常な結腸試料は、健康な組織の試料であることに基づき、不活性の経路を有する(スコア<0)と正しく予測されている。4つの試料以外の、活性の経路を有すると疑わしい試料は全て、活性のWnt経路を有すると予測されている。

20

【0100】

図36において、訓練したERベイジアンネットワークモデルの検証結果が、MCF7乳癌細胞株試料を使用して測定された2つのマイクロアレイに対して示されており、1つは、エストロゲンで刺激され(E2)、もう1つは、負の対照を有しており(EtOH)、どちらもGSE9253データセットから生じている。疑わしいER活性との同意において、エストロゲンで刺激した試料は、活性のER経路を有すると予測される一方、負の対照は、不活性のER経路を予測する。

30

【0101】

訓練した(例えばWnt、ER、AR及びヘッジホッグの経路の)ベイジアンネットワークを使用して、それぞれの経路の活性を予測するためのさらなる詳細及び例が、以下の実施例6において説明される。

【0102】

上述の訓練プロセスは、臨床適用の他のベイジアンネットワークに利用することができる。ここでは、細胞シグナル経路、さらに具体的にはWnt、ER、AR及びヘッジホッグの経路を表す本明細書において開示される方法を使用して構築したベイジアンネットワークモデルに対して機能することが示され且つ証明される。

40

【実施例6】

【0103】

(異常な)経路の活性の診断

以下において、どのようにして例えばベイジアンネットワークモデルを使用して、細胞シグナル経路の活性を診断するかが例証的に例示されている。

【0104】

本明細書において記載される図6と類似し、且つ、本明細書において記載されるように訓練した、転写因子の存在に対するノード、標的遺伝子のmRNAを表すノードの層、及

50

び、標的遺伝子（表 1、表 2、表 3、表 4）に対応するプローブセットの強度を表すノードの層を使用して構築した Wnt、ER、ヘッジホッグ及び AR の経路のベイジアンネットワークを使用し、種々の以前に訓練には使用されていないデータセットにおいて、活性である「オン」又は不活性の「オフ」として経路の活性を予測し、どのくらい適切に推量成分が機能しているかを推量した。予測した経路の活性のスコアは、臨床知識と関係づけられる。テストランの選択に対する結果の要約は、以下の図 2 1 において示されている。

【 0 1 0 5 】

以下の図 2 1 を参考にして、本明細書において記載されるベイジアンネットワークモデルを使用した医学的な組織試料に対する経路の活性の推量結果が示される。

【 0 1 0 6 】

図 2 1 は、結腸試料のデータセット GSE 4 1 8 3 に対する Wnt 活性テストの結果を示している。ベイジアンネットワークモデルは、腺腫細胞に対して高い値の P (Wnt オン)、及び、正常な試料に対して低い値を生じ、それは、腺腫及び健康な組織の (病態) 生理学に一致している。速い細胞増殖を有し、従って高い Wnt 活性を有する腺腫性組織とは相対的に、健康な組織は、ゆっくりとした細胞増殖を有し、従って、低い Wnt 活性を有している。IBD 試料に対して、ベイジアンネットワークモデルは、1 つの試料以外の全ての試料に対して、低い Wnt 経路の活性 (P (Wnt オン) ~ 0) を示した。ここでも、これは、速い細胞増殖を生じない IBD 試料と一貫している。結腸直腸癌細胞試料に対して、結果はまちまちであり、これらの試料のうち約半分において高い Wnt 経路の活性が検出された。しかし、これは、良性の腺腫性組織が悪性の癌にかかった組織になった場合の腫瘍ドライバーの役割を仮定する他の経路の結果、又は、例えば多すぎる非腫瘍組織を含有する試料、若しくは、mRNA が部分的に分解されている等、試料解析問題の結果であり得る。

【 0 1 0 7 】

本明細書において報告される実験において使用されるベイジアンネットワークモデルは、結腸試料のデータセット GSE 8 6 7 1 を使用して訓練した。しかし、Wnt 経路は、(おそらく不活性にもかかわらず) 他の細胞タイプにおいて存在している。従って、おそらく、ベイジアンネットワークは、他のタイプの癌と相関的な異常に高い Wnt 経路の活性を推量するために適用可能であり得るということを考慮した。これに対する理論的根拠は、結腸試料を使用してベイジアンネットワークモデルを訓練したけれども、他の細胞タイプにおける (おそらく不活性であるにもかかわらず) Wnt 経路が存在するということの働きの第一原理に基づいているということである。図 2 2 乃至 2 4 は、そのような「交差組織タイプ」の推量を調査するいくつかの結果を示している。

【 0 1 0 8 】

図 2 2 は、髄芽腫試料 (データセット GSE 1 0 3 2 7) における Wnt 経路の活性を推量するために適用されている、結腸試料を使用して訓練したベイジアンネットワークモデルを使用したテストに対する結果を示している。このデータセットに含まれる試料を、いくつかのサブセットにおいてさらに特徴づけた。前記サブセットの 1 つは、Wnt 経路が活性の試料である。Wnt 経路のベイジアンネットワークは、活性の Wnt 経路を有した Wnt - 活性試料の群を予測する一方、他の試料を不活性の Wnt 経路を有するとして正確に予測する。

【 0 1 0 9 】

肝癌試料を含有するデータセット (GSE 9 8 4 3) において Wnt ベイジアンネットワークモデルを使用したテスト結果が、図 2 3 において示されている。ここで、試料は、GSE 9 8 4 3 データセットによって割り当てられる以下の事前の注釈：「CTNNB1」、「炎症」、「多染色体性 chr 7」、「増殖」及び「注釈なし」によってグループ化されている。「炎症」グループの試料は、炎症状態が速い細胞増殖を引き起こさないために予想されるように、異常に高い Wnt 経路の活性は有していないと一様に推量されている。「多染色体性 chr 7」とラベルされた試料も、異常に高い Wnt 経路の活性は有していないと一様に推量されている。染色体数 7 の多染色体性は、3 つ以上の第 7 染色

10

20

30

40

50

体があるということの意味している。この多染色体性状態がWnt経路に影響を与えると予想する理由がないため、これらの試料が異常に高いWnt経路の活性を有さないということは予想外ではない。

【0110】

「増殖」とラベルされた試料のうち約5分の1が、 $P(Wntオン) > 0.5$ を有する。増殖によって、速い細胞分裂増殖の状態が示唆される。そのような状態は、異常に高いWnt経路の活性と付随し得るが、細胞増殖の可能性のある数多くの他の原因とも付随し得る。従って、異常に高いWnt経路の活性を有するこれらの試料のうち約5分の1が、不合理な結果ではない。

【0111】

「CTNNB1」グループの試料のうち約半分が、異常に高いWnt経路の活性を有しているとベイジアンネットワークによって推量されている。CTNNB1遺伝子は、Wnt経路の調節タンパク質であるベータカテニンタンパク質をコード化し、さらに、この遺伝子において変異を活性化することによって、異常なWntの活性化が引き起こされる。このように、「CTNNB1」グループと高いWnt経路の活性との相互関係は、整合した予想である。

【0112】

図24は、乳癌試料のセットに対する、本明細書において記載されるWntベイジアンネットワークモデルのテスト結果を描いている。この場合、乳癌細胞株の3つのグループがテストされている。1つのグループは、Wnt経路が異常に高いレベルで機能していると事前に知られており(Wntオングループ)、1つのグループは、Wnt経路が異常に高いレベルで機能していないと事前に知られており(Wntオフグループ)、さらに、もう1つのグループは、Wnt経路の活性が事前に知らされていない(不明のグループ)。加えて、活性のWnt経路を有し得るという文献において矛盾する報告があるけれども(しかし、これは少数の報告であり、より多くの研究論文が不活性のWnt経路を報告している)、低レベルのWnt活性化を有していると疑われている試料も1つある(Wntの疑いあり)。図24において見られるように、ベイジアンネットワークによって提供される、事前の知識との推量の相互関係は、Wntオン及びオフのグループに対して強いものである。また、図の一番右の試料(Wntの疑いあり)は、Wnt経路はオフであると述べる文献における大部分の報告に一致する推量を示している。Wnt経路の活性における事前の知識がない図24において示されている不明のグループのケースにおいては、ベイジアンネットワークは、 $P(Wntオン) > 0.5$ の1つの例以外は、Wnt経路に対して低い活性を推量している。文献によると、この細胞株は共受容体LRP6の高い発現を有していることが示されており、Wnt経路はオンであると説明することができる。

【0113】

図25は、同じ乳癌細胞株のデータセットに対する結果を示しているが、ここでは、本明細書において記載されるMCF7乳癌細胞株を使用して訓練したERベイジアンネットワークを使用して、ER活性に対してテストした。活性のWnt経路を有すると事前に知らされた試料を、不活性のER経路を有すると予測した。これは、Wnt経路はすでに速い細胞分裂増殖を駆り立てているため、驚くべきことではない。もう一方で、ER陽性試料を、Wntオフ試料及び不明の試料の中から見つけた。図24を考慮すると、これは驚くべきことではない。

【0114】

癌試料のセット(GSE12276)に対する、乳癌細胞株で訓練したERベイジアンネットワークの予測のテスト結果が、図26において示されている。乳癌試料を、よく知られた分類である、ルミナルA(LumA)、ルミナルB(LumB)、ヒト上皮増殖因子受容体2陽性(HER2)及び基底乳癌サブタイプに細分した。ルミナルA及びルミナルBサブタイプにおける組織試料は、ERを発現すると知られている。大部分の試料が高いER経路の活性を有すると予測されるのもこれらのサブタイプにおけることである。もう一方で、基底のサブタイプのものであるとして分類される試料は、ERの発現

10

20

30

40

50

を有さない又はERの低い発現を有するとして知られ、これは、基底グループ試料において予測される活性でないER経路との相互関係を正確に示す。HER2グループにおいて、3つの試料のみが、 $P(ERオン) > 0.5$ を有する一方で、大多数の試料が、不活性のER経路を有すると予測される。これは、これらの試料は増幅されるHER2発現を有するという事実に対して分類が行われているという事実との相関関係をよく示しており、制御されない細胞複製が、おそらく、ER経路以外の細胞シグナル経路を介したHER2シグナルを介して駆り立てられている（例えば、図24におけるWnt活性乳癌細胞株、又は、図30におけるヘッジホッグ活性乳癌試料を参照）。

【0115】

本明細書に記載される構築且つ訓練したERベイジアンネットワークモデルを使用して、種々の癌の細胞株大パネルにおけるER経路の活性が予測され、結果が図27において示されている。予想されたように、ER活性が予測された試料のみを乳癌細胞株において見つけた。他のタイプの癌細胞株全てを不活性なER経路を有すると予測し、これは予想通りである。

10

【0116】

本明細書において記載されるヘッジホッグ経路に対して構築且つ訓練したベイジアンネットワークモデルを使用して、GSE34211データセットにおける種々の癌タイプの細胞株に対するヘッジホッグ経路の活性が予測される。ヘッジホッグの活性予測は、図28において示されている。正の予測されたヘッジホッグ活性の最も高い部分は、中枢神経系(CNS)、皮膚、子宮内膜及び子宮癌タイプにおいて見付き、これは、これらの細胞タイプにおけるヘッジホッグ依存性細胞増殖に関する文献の知識とよく一致している。

20

【0117】

図29は、本明細書において記載されるWntベイジアンネットワークモデルを使用してすでに解析した髄芽腫試料(GSE10327)の予測されたヘッジホッグ活性を示している。髄芽腫試料は、サブクラスにおいて特徴づけられており、それらのうち1つは、活性のヘッジホッグシグナル経路を有している（識別子：SHH）。SHHサブタイプにおける試料のうち全てが、活性のヘッジホッグシグナルを有すると予測されている。さらに、Wntサブタイプにおける髄芽腫試料も、活性のヘッジホッグ経路を有すると予測した。これは、どちらの経路もこれらの腫瘍において活性であることが多いことを示す臨床的証拠と一致している。それにもかかわらず、Wntベイジアンネットワークは、WntサブタイプにおいてのみWnt活性を正確に予測することが明らかにできた。このように、Wnt及びヘッジホッグベイジアンネットワークの組み合わせは、これら2つのサブタイプの分類を正確にすることができる。

30

【0118】

ERベイジアンネットワークモデルを使用してER活性を予測するために以前に使用したGSE12276乳癌試料における、ヘッジホッグベイジアンネットワークモデルを使用して予測したヘッジホッグ活性が、図30において示されている。ヘッジホッグ経路は、全てのサブタイプの試料のうち一部分において活性であると予測される。これは奇妙に思えるが、図26において示されたER経路の予測に一致しており、ヘッジホッグの活性は、活性のER経路を有さない試料においてのみ予測されると見ることができる。これは、（乳房）組織における制御されない細胞増殖は種々のシグナル経路によって駆り立てられ得るという仮説とよく一致している。

40

【0119】

要約すれば、図21～30に示された種々の癌性組織試料及び細胞に対するテスト結果は、組織/経路特異的試料に対して訓練したWnt、ER及びヘッジホッグモデルのベイジアンネットワークが、他のタイプの組織の試料における解析に対して利用可能であることを強く示唆している。これは、細胞シグナル経路解析が「交差組織タイプ」に適用されるのを可能にし得る。このように、（本明細書において記載される）CDSシステム10は、ベイジアンネットワークモデル40を訓練するために使用された試料の組織タイプ以外の組織タイプの範囲における経路の活性を評価するために容易に適用される（例えば、

50

本明細書において開示される（Wnt 経路に対して例証的に示される）1つ又は複数の細胞シグナル経路を評価するために構成された臨床判断支援（CDS）システムを概略的に示す図20を参照）。推量成分40、44、46、48が、解析下の組織が異常に高いWnt、ER又はヘッジホッグ経路の活性を示すと表すが、組織特異的な薬物が利用可能でない場合、一般的なWnt、ER又はヘッジホッグ経路の抑制薬物又は機能障害特異的な薬物が、CDSシステム10により提供される勧告28又は勧告26にそれぞれ基づき、医師によって考慮されてもよい。

【0120】

図21～30の結果は、Wnt、ER及びヘッジホッグ経路に対するベイジアンネットワークモデルにおける交差組織タイプの適用性を表すけれども、臨床的適用に対して、ベイジアンネットワークモデルは、解析下の特定の組織タイプ（例えば乳房組織若しくは肝臓組織等）に対するその適用性を最大にするために選択的に更新又は適応されてもよいということが予想される。そのような更新又は適応は、例えば、解析下の組織タイプの臨床研究に基づき条件付き確率を調整することを必要とし得るか、又は、調査下の1つ又は複数の経路の組織特異的標的遺伝子を有する本明細書において記載される証拠により精選した標的遺伝子のリストの質を高め得る。加えて、解析下の組織に対してベイジアンネットワークモデルをより良く調整するために、ノードを追加又は除去してもよい。或いは、異なるベイジアンネットワークモデルを、組織タイプに応じて異なる訓練セットを使用して最初から訓練してもよい。さらに、図21～30の結果は、本明細書に記載されるプロセスの、Wnt、ER及びヘッジホッグ以外の経路の証拠により精選した標的遺伝子のリストを使用してベイジアンネットワークモデルを開発及び訓練し、経路の活性を予測及び診断する能力を例示している。

10

20

【0121】

本明細書において記載される構築及び訓練したARベイジアンネットワークモデルのテスト結果を例証的に使用して、種々の治療レジメで治療されたLNCaP前立腺癌細胞株（GSE7708）におけるARの活性を予測した（図38を参照）。予想したように、DHTで刺激されていないLNCaP細胞は、予測された不活性のAR経路を生じる一方で、LNCaP刺激細胞を、活性のAR経路を有するとして、さらに、抗アンドロゲン薬物であるピカルタミドで治療されたLNCaP細胞を、抑制されたAR経路を有するとして正確に予測した。

30

【0122】

本明細書において記載される訓練したAR経路のベイジアンネットワークを使用して、GSE17951データセットからの前立腺癌試料においてAR経路が活性である確率も予測した（結果は図39において示されている）。大部分の前立腺生検及び腫瘍を、対照試料と比較してより高いARの活性の確率を有するとしては意外なことに予測しなかった。

【0123】

ARベイジアンネットワークモデルを、交差組織テスト、すなわち、GSE12276データセットに含まれる乳癌試料にも適用した。このテストに対する結果は、図40において示されている。全てのサブグループにおいて見つかった、試料のうちごく一部分が、活性の経路を有すると予測される一方で、大多数の試料が、不活性のAR経路を有した。意外なことに、最も高い割合の活性のAR経路を有した試料は、HER2サブグループにおいて見つけられ、これは、HER2とAR経路との間に掛け合い応答があり、さらに、AR経路はHER2シグナル伝達によっても誘発され得ると文献から既知のように予想外のことでない。

40

【0124】

図41及び図42において描かれているように、上述のARARベイジアンネットワークモデルを使用して、種々の癌タイプの細胞株試料の2つのセット（GSE36133及びGSE34211）におけるAR経路の活性も予測した。予想したように、大部分の細胞株が、不活性のAR経路を有すると分かった。これに対する例外は、AR経路の活性を

50

発現する癌細胞株試料をいくつか有する前立腺癌試料である。表 1 2 において、前立腺癌試料における A R 経路の活性の予測全てが既知の A R 活性と一致していることが示されている。

【 0 1 2 5 】

【表 1 2】

表 1 2 G S E 3 6 1 3 3 及び G S E 3 4 2 1 1 データセットの前立腺癌細胞株における既知及び予測した A R の活性

データセット	試料識別子	前立腺細胞株	活性であると既知?	P(AR オン)
36133	GSM886837	22Rv1	YES	0.698127
	GSM886988	DU 145	NO	0.001279
	GSM887271	LNcaP clone FGC	YES	1
	GSM887302	MDA PCa 2b	YES	1
	GSM887440	NCI-H660	NO	1.25E-05
	GSM887506	PC-3	NO	0.009829
	GSM887731	VCaP	YES	1
34211	GSM843494	DU145T	NO	0.005278
	GSM844559	HPET11	NO	0.005602
	GSM844560	HPET13 replicate 1	NO	0.003382
	GSM844561	HPET13 replicate 2	NO	0.000501
	GSM844562	HPET5	NO	0.007673
	GSM844579	LNCAP	YES	1
	GSM844674	PC3 PFIZER	NO	0.004066
	GSM844675	PC3 Good_NCI50_WYETH	NO	0.006163

10

20

30

【実施例 7】

【 0 1 2 6 】

経路の活性に基づく予後

W n t やヘッジホッグといった初期発生経路は、癌幹細胞と称される、より幹細胞様の表現型に帰属する癌細胞により引き起こされた転移において役割を果たすと考えられている。確かに、W n t 経路等の初期発生経路が癌転移における役割を果たすことについて、十分な証拠が入手可能であり、転移性癌細胞が別の臓器又は組織におけるシーディング位置において分裂し始めることを可能にする。転移は、悪い予後と関連するので、癌細胞における W n t 及びヘッジホッグ経路等の初期発生経路の活性が、悪い予後に対する前兆となることが予想される。このことは、本明細書に記載のベイジアンネットワークモデルを用いて活性 E R 経路を有するが活性 W h t 又はヘッジホッグ経路を有しないと識別された G S E 1 2 2 7 6 データセットからの乳癌患者が、図 3 7 の Kaplan-Meier プロットに例示された活性のヘッジホッグ又は W n t 経路のいずれか又はその両方を有すると識別された患者よりも良好な予後を有したといった事実によって、支持される。

40

【実施例 8】

【 0 1 2 7 】

療法計画、薬効の予測、副作用の予測及び薬効のモニタリング

療法計画、薬効の予測、薬効の予測のモニタリング及び関連する活動について、いかに

50

確率モデル、特にベイジアンネットワークモデルを用いるかを以下に例示する。

【0128】

図6と同様、また本明細書に記載したように訓練したように、転写因子の存在についてノードを用いて構築されたER経路のベイジアンネットワークモデル、標的遺伝子 mRNAレベル(表2)を表すノードの層、及び標的遺伝子(表2)に対応するプローブセットの強度を表すノードの層を用いて、経路の活性を予測した。経路の活性は、薬効の予測又は薬効の予測のモニタリングに関連づけられるよう、後に実証する。結果のまとめを、図31及び32に示す。

【0129】

タモキシフェンは、ER+ (エストロゲン受容体陽性) 乳癌の治療について現在用いられている薬物である。ERシグナリングによって誘発されたと考えられた、制御されていない細胞増殖を妨げるエストロゲン受容体の部分アンタゴニストとして作用する。残念なことに、癌組織スライドのルーチン的な組織病理学分析による癌細胞のERタンパク質の存在の実証にもかかわらず、全ての乳癌がタモキシフェンを伴う治療に応答するわけではない。このいわゆるタモキシフェン耐性を調査するため、多くの研究がなされてきた。公的に入手可能なGSE21618データセットは、かかる研究の一つの結果であり、異なる処理レジメのもとでのタモキシフェン耐性及び野生型MCF7細胞株のマイクロアレイデータを含む。本明細書に記載された、構築され且つ訓練されたERベイジアンネットワークモデルを用いて、タモキシフェン耐性及びMCF7細胞株を異なる処理レジメで分析し、その結果を図31に示す。

10

20

【0130】

TamR. Ctrlにより示される制御タモキシフェン耐性は、タモキシフェン追加(1、2、3、6、12、24、及び48h)後、各時点について不活性ER経路を有すると予測される。TamR. Tamによって指摘された、タモキシフェンを伴うタモキシフェン治療に感度の悪いタモキシフェン耐性細胞株の治療が、効果のないことであるのは、驚きではない。タモキシフェン耐性細胞株(TamR. Ctrl)の分析によれば、制御されていない細胞増殖の推進力は、活性ERシグナリングに因るものではない。したがって、ERアンタゴニストで処理することは、細胞増殖を妨げない。このことが例示するのは、タモキシフェンを伴う治療は、負の予測されるER経路の活性の場合には、推奨されないことである。

30

【0131】

一方、タモキシフェン感受性であることが知られている、17β-エストラジオール(wt1.E2)で処理される野生型MCF7細胞株は、ゆっくりとホルモン治療に反応し、このことは、増加するER陽性活性予測において見ることができる。エストロゲン生成を妨げることが知られているアロマターゼ阻害剤でかかる細胞株を処理することは、ER経路を妨げ、これは、やがて減少するER経路予測によって例示される。このことは、GSE11324データセットにおいて時間を増やすためにエストロゲンで処理するMCF7試料からのマイクロアレイデータに基づくER経路予測によって、支持される。結果を、図32に示す。

40

【0132】

上記は、療法計画、薬効の予測及び薬効の予測のモニタリングに用いる確率モデルの能力、特に、ベイジアンネットワークモデルを、例示する。しかしながら、同じ手法を、副作用を予測しモニタリングすることに適用することが理解される。

【実施例9】

【0133】

医薬品開発

療法応答モニタリングと同様、経路モデルを医薬品開発に用いて、種々の推定される化合物の有効性を評価することができる。例えば、癌細胞株における或る経路の起こりうる効果について多くの化合物をスクリーニングするとき、それぞれの経路モデルを用いて、化合物の適用後に経路の活性が上昇するか下降するかを決定することができる。しばしば

50

、このチェックは、経路の活性の1つ又は幾つかの推定マーカのみを用いてなされるが、このことは、治療効果の有効でないモニタリングの可能性を上げる。さらに、動物又は患者対象についての追跡研究において、経路モデルを同様に用いて候補薬物の有効性を評価し、且つ経路の活性に最大限に影響を与えるための最適な量を決定することができる。

【0134】

新規の薬剤化合物の有効でないモニタリングの例を、図38に示すようにGSE7708試料における予測されるAR経路活性によって例示する。本研究において、ポリアミド1及びポリアミド2によって示す、AR経路活性を妨げる可能性のある2つの薬剤化合物を、開発した。これらの2つのポリアミドが、AR経路の妨げを可能とすることを実証し、この実証は、ポリアミドがアンドロゲン応答配列 (Androgen Response Element (ARE)) に結合し、本明細書に記載の標的遺伝子選択にも含まれているAR活性についての周知のマーカであるKLK3 (=PSA) の発現を妨げているという発見、並びにDHTによって誘発された35%の転写に基づいている。対称的に、AR経路のベイジアンネットワークモデルが、これらの試料が依然として活性AR経路を有することを予測した。発見によれば、ARベイジアンネットワークモデルを用いた上方制御された標的遺伝子の推量された確率の調査は、KLK3が他の標的遺伝子とは対照的に下方制御されたことを示した。一方、全ての他の標的遺伝子(ポリアミド1の場合AR、GUCY1A3及びTMPRSS2を除く)は、明らかに差次的にポリアミド1及びポリアミド2処理された試料において発現した。換言すれば、AR活性についての1つのマーカ、KLK3を下方制御したが、識別された大部分の標的遺伝子は、依然として上方制御され、AR経路が依然として無損傷であり従って活性であることを示した。文献証拠に基づいて多数の標的遺伝子を考慮にいと、本発明者等は、ポリアミドのAR活性の妨げが限定されていて、且つKLK3発現のみがこれらのポリアミドを使用して明らかに下方制御されたことを示すことができた。さらに、このことは、医薬品開発における還元主義的なアプローチと比べ、ベイジアンネットワークモデルを用いた体系化されたアプローチの値を例示している。

10

20

【実施例10】

【0135】

アッセイの開発

マイクロアレイ又はRNA配列決定から生じるmRNA入力データに対して上述のベイジアンネットワークを適用する代わりに、例えばqPCRを使用する統合的プラットフォームに対して同じ測定を行って標的遺伝子のmRNAレベルを決定するために献身的なアッセイを開発することが、臨床的適用において有益であり得る。従って、開示される標的遺伝子のRNA/DNA配列を使用して、そのようなプラットフォームに対してどのプライマー及びプローブを選択するかを決定することができる。

30

【0136】

そのような献身的なアッセイの検証を、マイクロアレイに基づくベイジアンネットワークを参照モデルとして使用し、さらに、開発したアッセイが検証試料のセットに対して類似の結果を与えるかどうかを検証することによって行うことができる。献身的なアッセイの次に、これは、mRNA配列決定データを入力測定値として使用する類似のベイジアンネットワークモデルを組み立てる及び較正するために行うこともできる。

40

【実施例11】

【0137】

経路の研究及び癌病態生理学の研究

以下において、どのようにしてベイジアンネットワークモデルを、(臨床的な)経路の研究、すなわち、どの経路が特定の疾患に関与するかを見つけることに関心がある研究において利用することができるが例示され、これは、さらなる詳しい研究に対して進めて、例えばシグナルタンパク質における変異を(モデルを用いて測定された)経路の活性化における変化に結びつけることができる。これは、特定の癌の始まり、増殖、並びに、進化及び転移(病態生理学)を調査することに関連している。

50

【0138】

本明細書において記載される図6に類似の、転写因子の存在に対するノード、標的遺伝子のmRNAレベル(表1、表2、表3及び表4)を表すノードの層、並びに、標的遺伝子(表1、表2、表3及び表4)に対応するプロブセットの強度を表すノードの層を使用して構築し、且つ、本明細書において記載されるように訓練したWnt、ER、ヘッジホッグ及びAr経路のベイジアンネットワークモデルを使用して、乳癌試料から成るデータセット(GSE12276)の経路の活性を予測した。

【0139】

研究者は、1つ又は複数の細胞シグナル経路、及び、制御されない細胞増殖を駆り立てる1つ又は複数の特異的調節解除を調べることに関心があると想定する。研究者は、上述の確率モデル、特にベイジアンネットワークモデルを使用してマイクロアレイデータを解析し、どの経路がおそらく制御されない細胞増殖の原因であるかを見つけることができる。図33及び図34において示され、Wnt、ER及びヘッジホッグの活性のケースに対するそのような解析の例示を見ることができる(GSE12276データセットの基底及びルミナルA試料)。後に、研究者はより詳細に調査して、経路の調節解除の正確な原因を見つけることができる。

10

【0140】

図34を参考にして、基底試料は、3つの負の受容体の状態(ER、PR及びHER2)を有するとして既知であり、従って、全試料が不活性のER経路を有すると予測されることを見ることは驚くべきことではない。もう一方で、試料の一部は、図34において示されるように、Wnt若しくはヘッジホッグの活性、又は、両方の活性を有すると予測される。これらの予測した経路の活性は研究者を説得して、例えばWnt及び/又はヘッジホッグ経路における既知の変異又は他の既知の調節解除に対して等、これらの試料をより詳細に調査させる。

20

【0141】

別の例が図33において与えられており、GSE12276データセットのルミナルA試料におけるWnt、ER及びヘッジホッグの活性が例示されている。ルミナルA試料はERを発現するとして知られているが、これは、癌特性が活性のERシグナル伝達によるものであるとは必ずしも意味していない。予測した経路の活性から、試料の半分未満が活性のERシグナル伝達を有すると推量することができる。しかし、活性のERシグナル伝達を有さない試料の一部が、活性のWnt及び/又はヘッジホッグ経路を有するとわかる。これは、それぞれWnt及び/又はヘッジホッグシグナル経路における欠陥に対して、研究者により徹底的にこれらの試料を調査させるもとであり得る。試料の一部は、含まれた3つの経路のいずれも活性であるとは予測せず、おそらく、他の経路が制御されない細胞増殖を引き起こしている。これもまた、研究者に他の経路における欠陥を調査するためのさらなる情報を与えている。

30

【0142】

要約すれば、本明細書において記載される例示は、(上記の)訓練したベイジアンネットワークモデルの、より管理された方法で、制御されない細胞増殖の原因を見つけることのプロセスを支援する能力を表している。ベイジアンネットワークを利用して、経路の活性に対して試料を選別することによって、予測した経路の活性は、細胞増殖に対して可能性のある経路を正確に指摘することができ、これは、さらなる詳しい研究に対して進めて、例えばシグナルタンパク質における変異又は他の既知の調節解除を(モデルを用いて測定された)活性化における変化に結びつけることができる。

40

【0143】

本明細書において記載されるように、細胞シグナル経路のベイジアンネットワークを開発及び訓練するためのプロセスを使用して、本発明と提携して利用することもできる他の経路に対するベイジアンネットワークモデルを構築することができる。

【実施例12】

【0144】

50

予測した活性に基づく臨床試験における被験者の登録

候補の薬物が、例えば、腫瘍の増殖を駆り立てる特定の経路の活性を遮断するために開発され、さらに、この薬物が、臨床試験に入る場合、従って、そのような試験に登録する被験者の適切な選択が、薬物の潜在的な有効性を提供するためには必要である。そのような場合、その腫瘍において活性化されたそれぞれの経路を有さない患者は、そもそも経路が活性化されていない場合に薬物が効果的ではないということは明らかのため、試験から排除されるべきである。従って、経路の活性を予測することができる経路モデルを選択ツールとして使用して、活性化されたそれぞれの経路を有すると予測される患者のみを選択することができる。

【実施例 13】

【0145】

行われることになる後の1つ又は複数のテストの選択

異なる経路モデルを使用して腫瘍が解析され、さらに、モデルが特定の経路の調節解除を予測する場合、従って、これは、行われることになる後のテストの選択を導くことができる。例えば、近接ライゲーションアッセイ (PLA) を実施して、それぞれの転写複合体の存在を確かめてもよい (Soderberg O, 2006)。そのような PLA は、TF 複合体における2つの鍵となるタンパク質、例えば、Wnt 経路の TF 複合体におけるベータカテニンとTCF4 が実際に共に結合した場合に、肯定的な結果を与えるように設計することができる。

【0146】

別の例は、調節解除されていると予測した経路が、シグナルカスケードに関してより詳細に解析されることである。例えば、この経路における鍵となるタンパク質を解析して、そのそれぞれの遺伝子をコードするDNA領域において変異があるかどうかを決定してもよく、又は、これらのタンパク質の存在量に対してテストし、正常値よりも高いか若しくは低いかどうかを見てもよい。そのようなテストは、どのような根本的原因が、経路の調節解除の原因となっているのかを表し、さらに、どの利用可能な薬物を使用して経路の活性を減らすことができたかを洞察することができる。

【0147】

これらのテストは、ベイジアンモデルを使用して同定される経路の活性を確かめるために選択される。しかし、コンパニオン診断テストの選択も可能である。モデルを使用した経路の同定後、標的療法の選択に対して、同定した経路に適用可能であるコンパニオン診断テストのみが行われることが必要とされる (選択)。

【実施例 14】

【0148】

コンパニオン診断テストの選択

先の実施例と同様に、腫瘍が解析され、さらに、経路モデルが特定の経路の調節解除を予測し、さらに任意選択で、多くのさらなるテストが調節解除の原因を調査するために行われた場合、従って、腫瘍学者は、多くの候補の薬物を選択して患者を治療してもよい。しかし、そのような薬物を用いた治療は、例えば臨床ガイドラインを満たすか、若しくは、治療費の払い戻しを保証するために、第一に実行されることになるコンパニオン診断テストを必要とし得るか、又は、規制 (FDA) のため、薬物の投与に先立ちコンパニオン診断テストを行うことが必要とされる。そのようなコンパニオン診断テストの例は、薬物ハーセプチン (トラスツズマブ) を用いた乳癌患者の治療に対する Her2 テストである。従って、経路モデルの成果を使用して、候補の薬物、及び、行われることになるそれぞれのコンパニオン診断テストを選択することができる。

【実施例 15】

【0149】

CDS 適用

(本明細書において開示される (Wnt 経路に対して例証的に示された) 1つ又は複数の細胞シグナル経路を評価するために構成された臨床判断支援 (CDS) システムを概略

10

20

30

40

50

的に示す) 図 20 を参考にして、臨床判断支援 (CDS) システム 10 が、適切に構成されたコンピュータ 12 として実施される。コンピュータ 12 は、ハードドライブ若しくは他の磁気記憶媒体、光ディスク若しくは他の光学記憶媒体、ランダムアクセスメモリ (RAM)、リードオンリーメモリ (ROM)、フラッシュメモリ、若しくは他の電子記憶媒体、又は、ネットワークサーバー等、非一過性の記憶媒体 (図示せず) 上に記憶された適したソフトウェア、ファームウェア、又は他の支持を実行することによって、CDS システム 10 として作動するように構成されてもよい。例示的な CDS システム 10 が例示的なコンピュータ 12 によって具体化されるけれども、より一般的には、CDS システムは、本明細書において定められる臨床判断支援方法を行うように構成されたデジタル処理装置又はデジタルプロセッサを含む器具によって具体化されてもよい。例えば、デジタル処理装置は、(例えば、携帯情報端末、又は、CDS アプリケーションを動かすスマートフォン等) ハンドヘルドの装置、ノートブックコンピュータ、デスクトップコンピュータ、タブレットコンピュータ若しくは装置、又は、遠隔ネットワークサーバー等であってもよい。コンピュータ 12 又は他のデジタル処理装置は、典型的に、ディスプレイ装置 14 を含むか、又は、該装置と動作可能に接続され、ディスプレイ装置 14 を介して、臨床判断支援の勧告を含む情報が、医療職員に表示される。コンピュータ 12 又は他のデジタル処理装置は、典型的に、例示的なキーボード 16、又は、マウス、トラックボール、トラックパッド、(おそらくディスプレイ装置 14 と統合される) タッチスクリーン、若しくは、他のポインターに基づくユーザ入力装置等の 1 つ若しくは複数のユーザ入力装置も含むか、又は、該装置と動作可能に接続され、ユーザ入力装置を介して、医療職員は、CDS システム 10、又は、CDS システム 10 による使用のためのデータ等を制御するためにオペレーションコマンド等の情報を入力することができる。

【0150】

CDS システム 10 は、(例えば、腫瘍学者、医師、若しくは他の医療職員によって治療されている病院患者又は外来患者、或いは、結腸癌、乳癌、若しくは肝癌等の特定タイプの癌を有するとして知られるか若しくは疑われる癌検診又はいくつか他の医療診断を受ける人等) 医療対象者に関連する入力情報として受ける。CDS システム 10 は、種々のデータ解析アルゴリズムをこの入力情報に適用して、ディスプレイ装置 14 を介して(又は、音声合成装置、若しくは、体感可能出力を提供する他の装置を介して) 医療職員に与えられる臨床判断支援の勧告を生じる。いくつかの実施形態において、これらのアルゴリズムは、臨床ガイドラインを患者に適用することを含んでもよい。臨床ガイドラインは、典型的には医療専門家のパネルの勧告に基づき構築され、さらに任意選択で、臨床ガイドラインを介して導くことを促進するために臨床的な「フローチャート」の形状でフォーマットされた標準的又は「基準」の治療勧告の記憶されたセットである。種々の実施形態において、本明細書において記載される機械学習法等の CDS 10 のデータ処理アルゴリズムは、さらに、又或いは、臨床判断支援を導き出すために入力情報に対して行われる種々の診断若しくは臨床試験のアルゴリズムを含んでもよい。

【0151】

本明細書において開示される例示的な CDS システム (例えば CDS システム 10) において、CDS データ解析アルゴリズムは、1 つ又は複数の医療研究所 18 によって取得された入力するゲノム及び/又はプロテオミクスの情報に対して行われる 1 つ又は複数の診断若しくは臨床試験アルゴリズムを含む。これらの研究所は、「オン・サイト」でさまざまに位置していてもよく、すなわち、病院、又は、医療対象者が医療検査及び/若しくは治療を受ける他の場所にて位置していてもよいが、或いは、例えば、医療対象者から抽出された医療対象者の組織の試料 (例えば、生検手順若しくは他の試料抽出手順を介して、乳房の病変部から、又は、結腸癌を有するとして知られる若しくは疑われる医療対象者の結腸から、又は、肝癌を有するとして知られる若しくは疑われる医療対象者の肝臓から得られた試料等) を (郵便若しくは別の送付サービスを介して) 受ける特定化及び集中化した研究所等、「オフ・サイト」でさまざまに位置していてもよい。試料が抽出される組織はまた、例えば結腸、乳房、肝臓、又は、結腸、乳房、肝臓若しくは他の臓器の外側で

広がる他の臓器から生じる（疑われる）悪性の組織等、転移性組織であってもよい。いくつかのケースにおいて、組織試料は、循環性腫瘍細胞、すなわち、血流に入った腫瘍細胞であってもよく、さらに、適した単離技術を使用して抽出される組織試料として抽出されてもよい。抽出された試料は、研究所によって処理され、ゲノム又はプロテオミクスの情報を生じる。例えば、抽出された試料は、（遺伝子チップ、DNAチップ、若しくはバイオチップ等としても当分野においてさまざまに呼ばれる）マイクロアレイを使用して、又は、定量的ポリメラーゼ連鎖反応（qPCR）処理によって処理して、例えば、遺伝子から転写されるメッセンジャーリボ核酸（mRNA）のレベル、又は、遺伝子から転写されたmRNAから翻訳されるタンパク質のレベルの形状で、関心のある遺伝子の発現レベル等の証拠となるゲノム又はプロテオミクスの情報を測定してもよい。別の例として、抽出された試料は、遺伝子配列決定の研究所によって処理して、デオキシリボ核酸（DNA）に対する配列を生じるか、又は、RNA配列若しくはコピー数多型等を生じてもよい。他の熟考される測定アプローチは、病理学スライド上で行われる、免疫組織化学（IHC）、細胞学、蛍光インサイツハイブリダイゼーション（FISH）、又は、近接ライゲーションアッセイ等を含む。マイクロアレイ処理、質量分析、遺伝子配列決定、又は、他の研究技術によって生じることができ他の情報は、メチル化情報を含む。そのようなゲノム及び/又はプロテオミクスの測定の種々の組み合わせもまた行ってもよい。

10

【0152】

一部の実施形態において、医療研究所18は、多量なゲノム及び/又はプロテオミクスのデータを生じるように、抽出された医療対象者の組織の試料に対して多くの標準化されたデータ取得を行う。例えば、標準化されたデータ取得技術は、1つ又は複数の染色体若しくは染色体の一部に対する、又は、組織のゲノム全体に対する（任意選択で整列させた）DNA配列を生じてもよい。標準的なマイクロアレイを適用することによって、多数の遺伝子に対する発現レベル及び種々のメチル化データ等、何千又は何万ものデータアイテムを生じることができる。このゲノム及び/又はプロテオミクスのデータ若しくはその選択された部分の過多は、臨床判断支援の勧告を考案するために臨床的に有用な情報を引き出すように、CDSシステム10に入力されて処理される。

20

【0153】

開示されるCDSシステム及び関連する方法は、種々の細胞シグナル経路の活性を評価するためのゲノム及び/又はプロテオミクスのデータの処理に関する。しかし、開示されるCDSシステム（例えばCDSシステム10等）は、任意選択で、バイタルサイン監視データ、患者の病歴データ、（例えば性別若しくは年齢等の）患者の人口統計学的データ、又は、患者の医用画像データ等、種々の患者データに基づく記憶された臨床ガイドラインに従って臨床判断支援の勧告を生じる等、多様なさらなる能力をさらにも含むことを理解されたい。或いは、一部の実施形態において、CDSシステム10の能力は、ゲノム及び/又はプロテオミクスのデータ解析を行って、本明細書において開示される細胞シグナル経路を評価することのみに限定されてもよい。

30

【0154】

例証的な図20を続けて参考にして、CDSシステム10は、抽出された試料において測定される細胞シグナル経路の標的遺伝子の発現レベルに、それに限定されないが、少なくとも部分的に基づき、医療対象者の組織における細胞シグナル経路の活性を推量し、さらに、この推量した活性に基づき、細胞シグナル経路が医療対象者の組織において異常に機能しているかどうかを決定する。本明細書において開示される例は、例示的な細胞シグナル経路としてWnt、ER、AR及びヘッジホッグ経路に関する。これらの経路は、経路の調節の損失が癌の増殖の原因であり得るため、腫瘍学の種々の領域において関心のあるものである。約10から15の関連するシグナル経路があり、さらに、各癌は、原則的に、調節解除されている1つの支配的な経路によって駆り立てられる。いかなる特定の機能の原理に限定されることなく、これらの経路は細胞増殖を調節し、結果として、癌細胞におけるこれらの経路の調節の損失は、「常にオン」である経路をもたらし、従って癌細胞の増殖を速め、次に、癌の増殖、侵入、又は転移（広がり）として現れ得る。

40

50

【0155】

細胞シグナル経路を形成するタンパク質カスケードの一部である中間タンパク質等の細胞シグナル経路の調節タンパク質をコード化する遺伝子の mRNA 発現レベルの測定は、調節タンパク質発現レベルの間接測定であり、さらに、(細胞シグナル経路の全活性とははるかに少なく)実際の調節タンパク質発現レベルとの強い相互関係を示してもよく、又は、示さなくてもよい。細胞シグナル経路は、標的遺伝子の転写を直接調節し、従って、標的遺伝子から転写される mRNA の発現レベルは、この調節活性の直接的な結果である。従って、CDS システム 10 は、細胞シグナル経路の標的遺伝子の発現レベル(代替の測定として mRNA 又はタンパク質レベル)に少なくとも基づき(例えば Wnt、ER、AR 及びヘッジホッグ経路等の)細胞シグナル経路の活性を推量する。これは、CDS システム 10 が、測定された標的遺伝子の発現レベルによって提供される直接的な情報に基づき経路の活性を推量するということを保証する。

10

【0156】

しかし、本明細書において開示されるように、全経路の活性を評価するのに効果的であるけれども、測定された経路の標的遺伝子の発現レベル 20 は、(実際にそうである場合に)なぜ経路は異常に機能しているのかに関しては特に情報価値のあるものではない。前記別の方法で、測定された経路の標的遺伝子の発現レベル 20 は、経路が異常に機能していることを表すことができるが、(例えば、十分な調節を欠く等)経路のどの部分が機能不全であって、全経路に異常に機能させるかは表さない。

20

【0157】

従って、CDS システム 10 が特定の経路の異常な活性を検出する場合、CDS システム 10 は、従って、任意選択で、整列した遺伝子配列 22、及び/又は、経路の1つ又は複数の調節遺伝子に対する測定された1つ又は複数の発現レベル 24 等、抽出された試料に対して医療研究所 18 によって提供される他の情報を使用するか、或いは、経路のどの部分が機能不全であるかを評価するために次に行われることになる診断テストを選択する。効率を最大にするために、一部の実施形態において、この任意選択の、なぜ経路が機能不全であるかという評価が、測定された経路の標的遺伝子の発現レベル 20 の解析が、経路が異常に機能していると表す場合にのみ行われる。他の実施形態において、この評価は、本明細書において記載される細胞シグナル経路の確率の解析に統合される。

30

【0158】

経路のどの部分が機能不全であるかを CDS システム 10 が評価し、さらに、それが成功している実施形態において、さらなる情報が、特定の機能不全を標的にする薬物を処方するのに CDS システム 10 が勧めるのを可能にする(図 20 において示された勧告 26)。特定の経路機能不全が、(任意選択のさらなる評価が行われていないため、又は、評価が、機能不全している経路のいかなる特定の部分の同定も失敗するために)同定されない場合、従って、CDS システム 10 は、(異常な経路の活性が過度に高い活性であることを仮定して)この特定の経路に対する一般的な抑制薬物の処方を勧めるデフォルトの勧告 28 を提供することができる。

【実施例 16】

【0159】

40

経路の活性を測定するキット及び解析ツール

ベイジアンモデルを使用したマイクロアレイ/RNA 配列決定ベースの調査に基づき特定の経路の活性を最も表すとわかった標的遺伝子のセットは、組織又は細胞の試料に対して行われることになる多重定量的 PCR アッセイに移すことができる。そのような FDA により承認されたテストを、経路の活性に対して開発するために、臨床試験において臨床的に検証して規制の許可を得る必要がある標準化されたテストキットの開発が必要とされる。

【0160】

一般に、Wnt、ER、AR 及び/又はヘッジホッグ経路に付随する例が例示的な例として提供されるけれども、本明細書において開示される細胞シグナル経路の解析に対する

50

アプローチが、細胞膜における受容体を用いた細胞内シグナル経路（上記参照）、及び、細胞内部の受容体を用いた細胞内シグナル経路（上記参照）に対して等、これらの経路以外の他の細胞シグナル経路に容易に適用されるということを理解されたい。加えて、本願は、いくつか好ましい実施形態を記載している。修正及び変更が、上述の詳細な説明を読み且つ理解した後で他のものに対して生じてもよい。本願は全てのそのような修正及び変更を、付随の特許請求の範囲又はその同等物にある限りにおいて、含むとして解釈されることが意図される。

【0161】

参考文献

de Sousa E Melo F, C. S. (2011). Methylation of cancer-stem-cell-associated Wnt target genes predicts poor prognosis in colorectal cancer patients. *Cell Stem Cell*, 476-485 10

Hatzis P, v. d. (2008). Genome-wide pattern of TCF7L2/TCF4 chromatin occupancy in colorectal cancer cells. *Mol Cell Biol*, 2732-2744

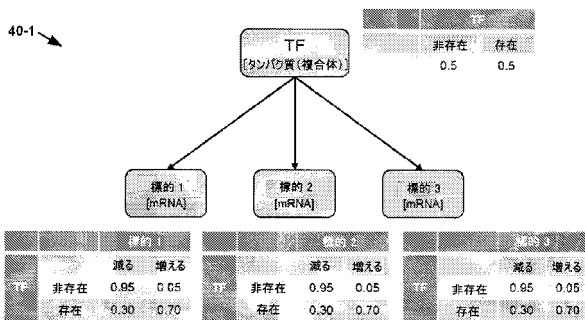
Neapolitan, R. (2004). Learning Bayesian networks. Pearson Prentice Hall

Nusse, R. (2012, May 1). Wnt target genes. Retrieved from The Wnt homepage: http://www.stanford.edu/group/nusselab/cgi-bin/wnt/target_genes

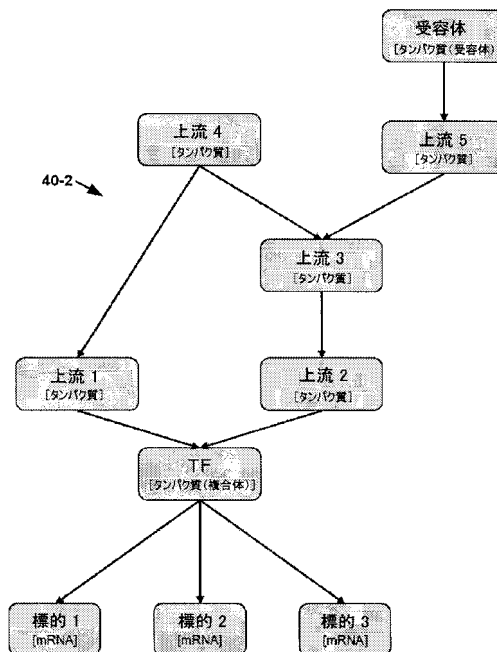
Soderberg O, G. M. (2006). Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat Methods.*, 995-1000 20

van de Wetering M, S. E.-P.-F. (2002). The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell*, 241-250.

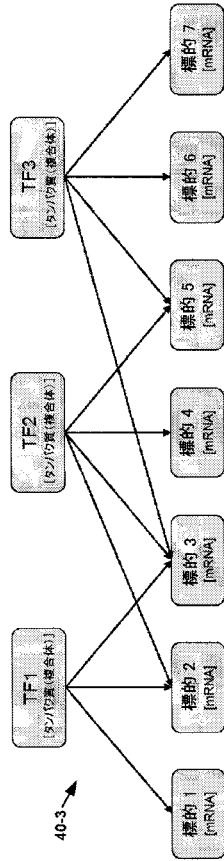
【図1】



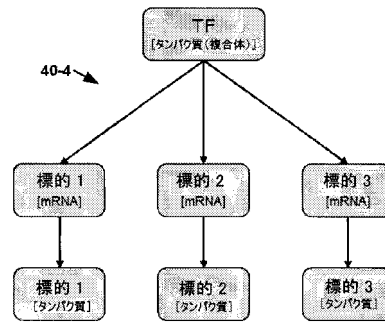
【図2】



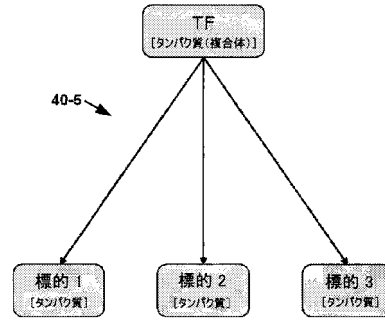
【 図 3 】



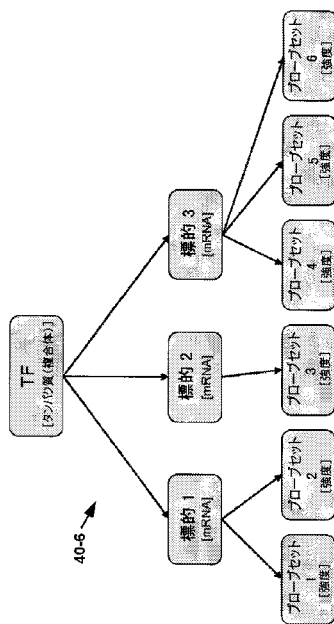
【 図 4 】



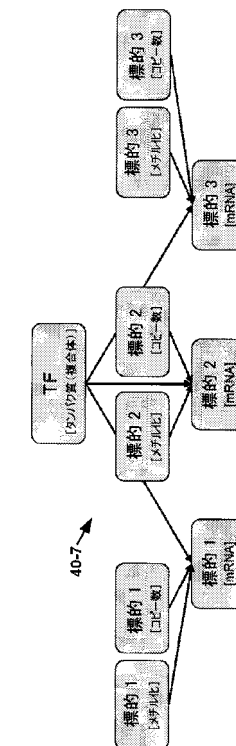
【 図 5 】



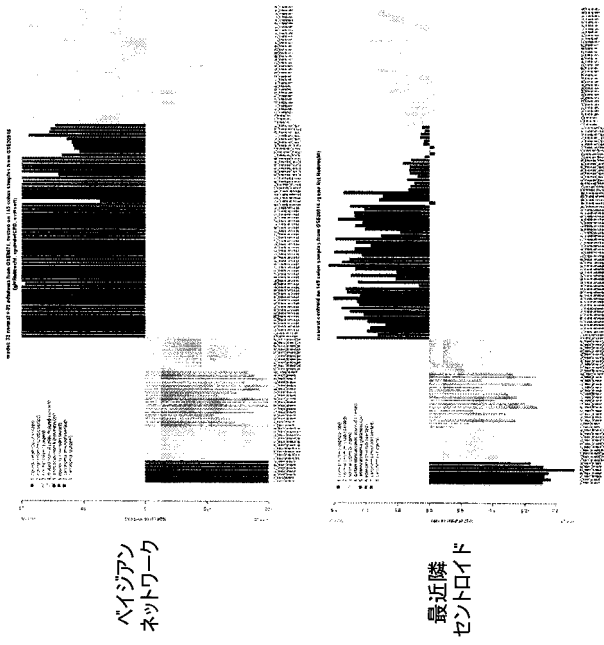
【 図 6 】



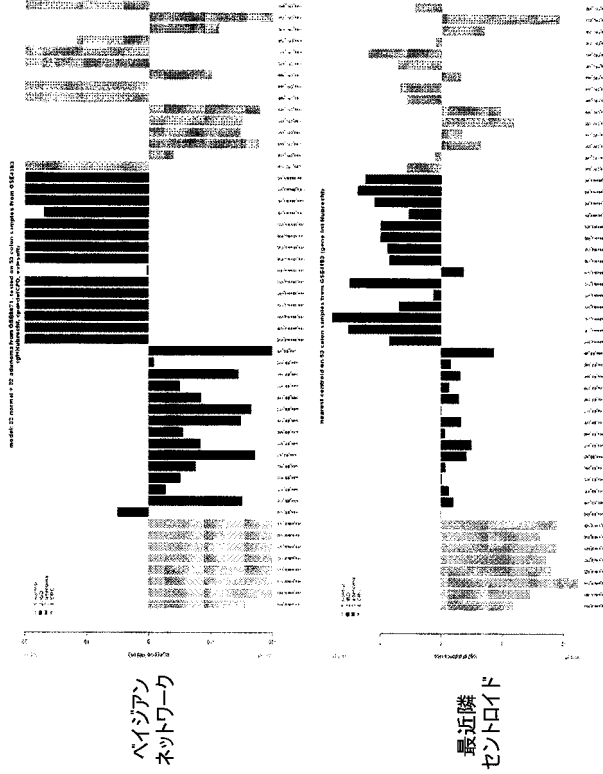
【 図 7 】



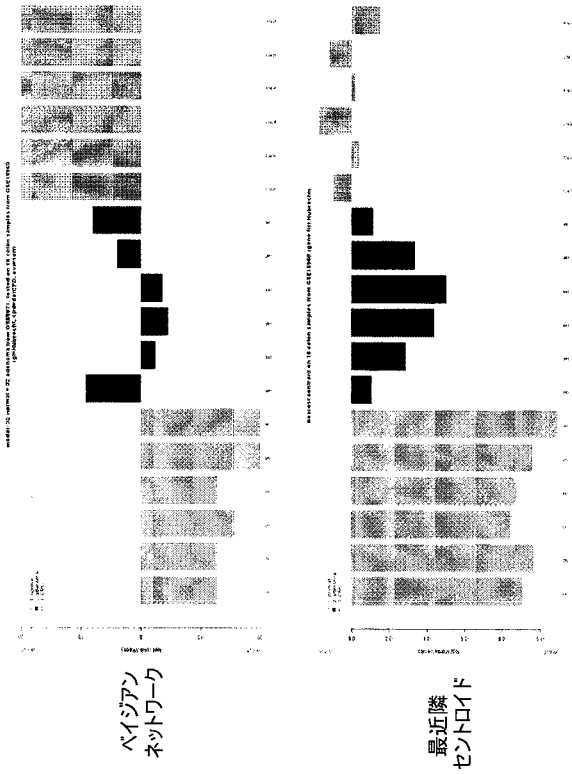
【図 8】



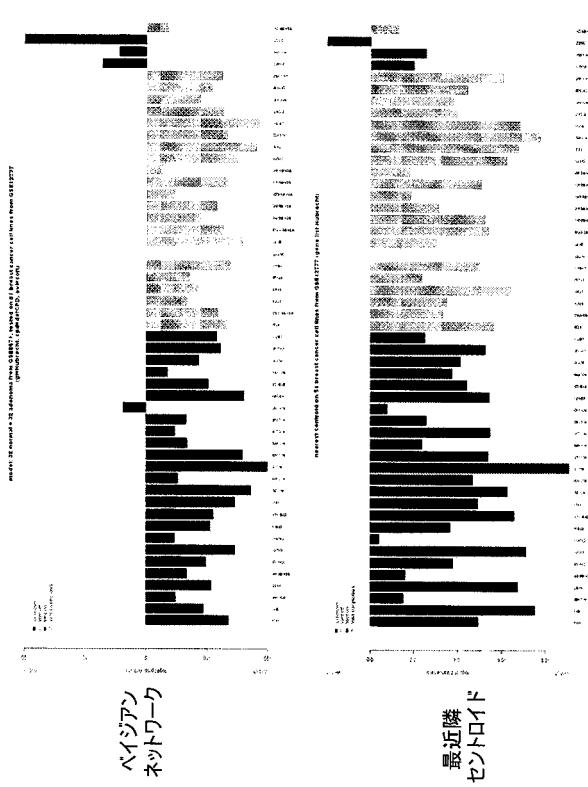
【図 9】



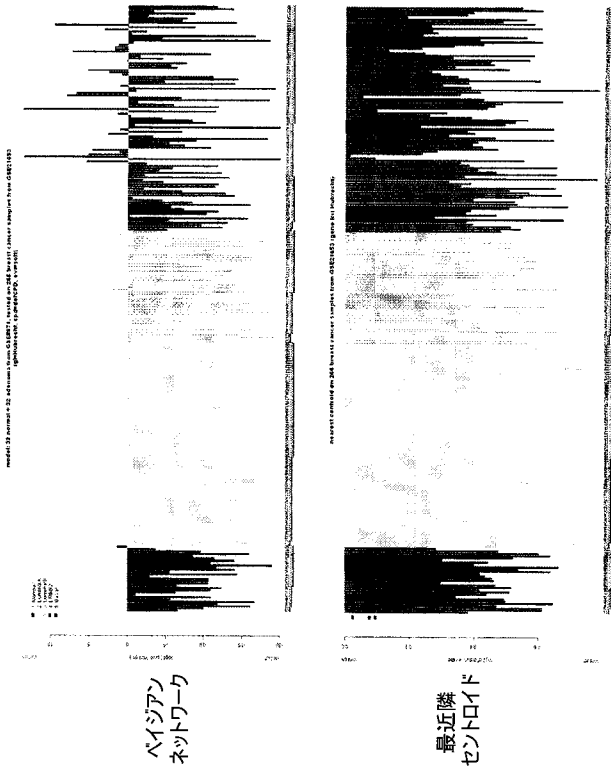
【図 10】



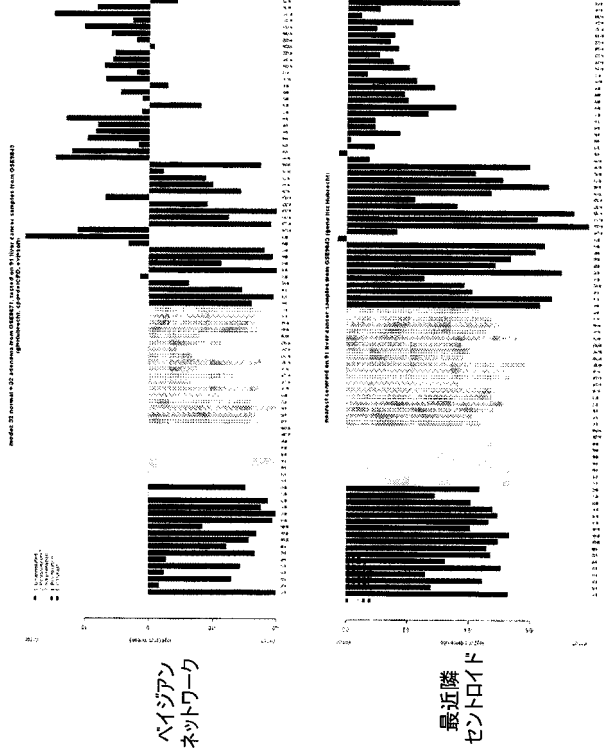
【図 11】



【 図 1 2 】

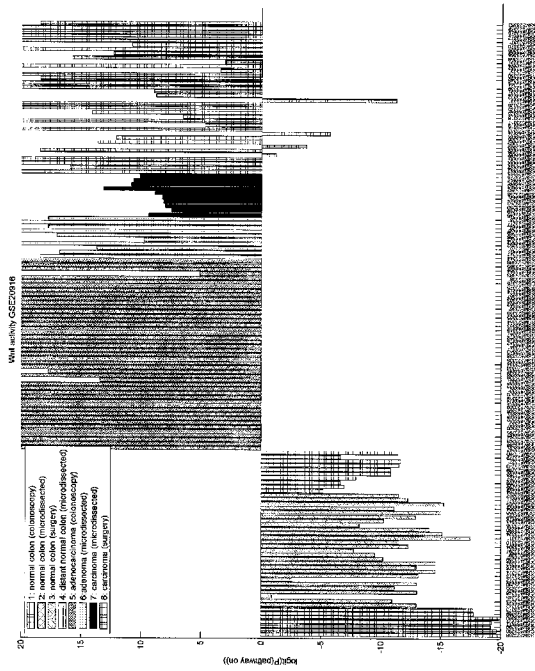


【 図 1 3 】



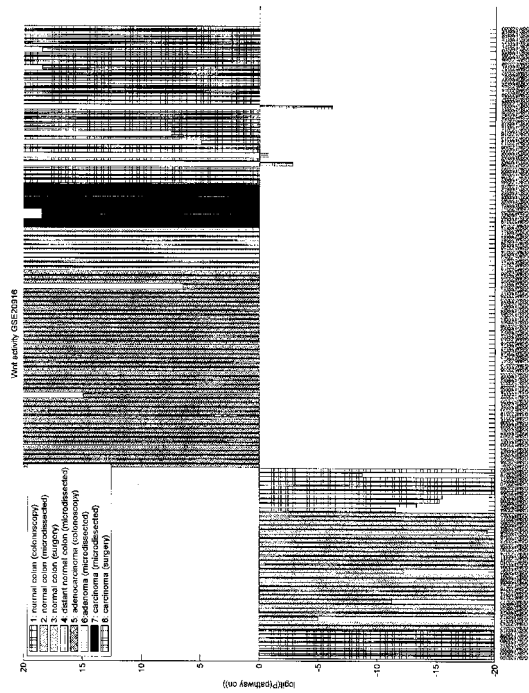
【 図 1 4 - I 】

証拠により精選したリスト



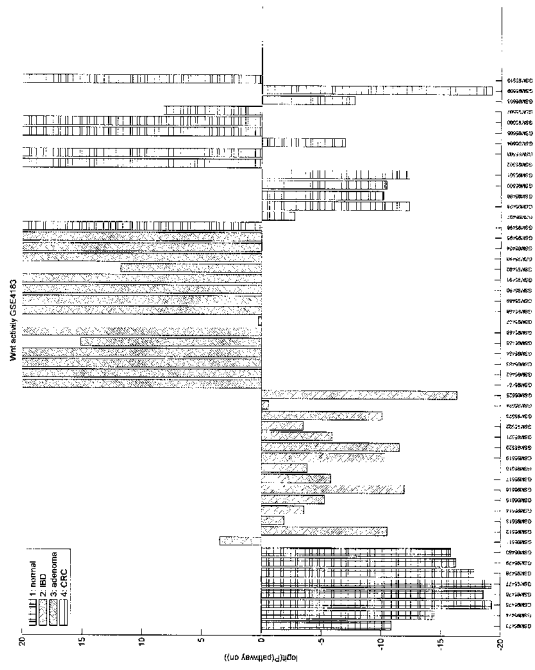
【 図 1 4 - II 】

広範な文献によるリスト



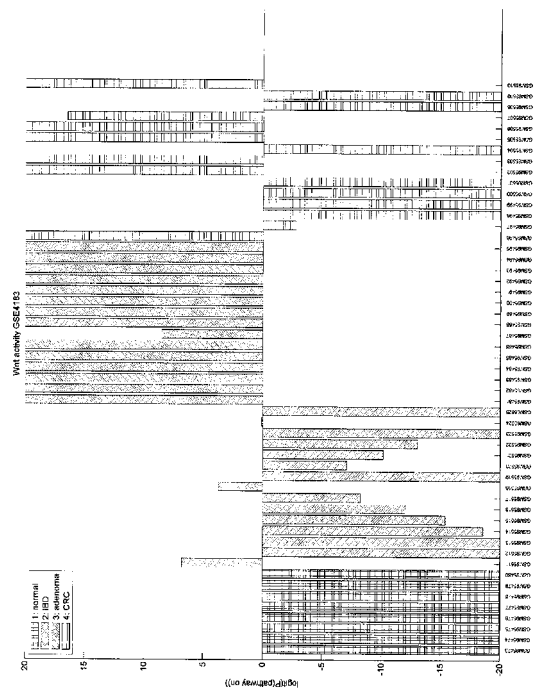
【図15 - I】

証拠により精選したリスト



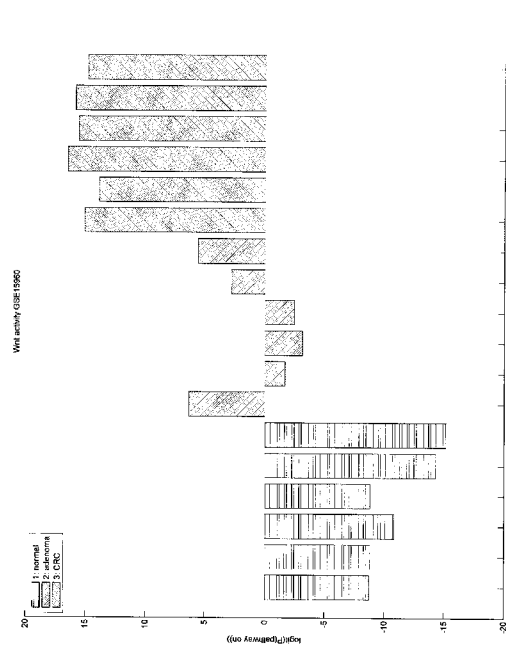
【図15 - II】

広範な文献によるリスト



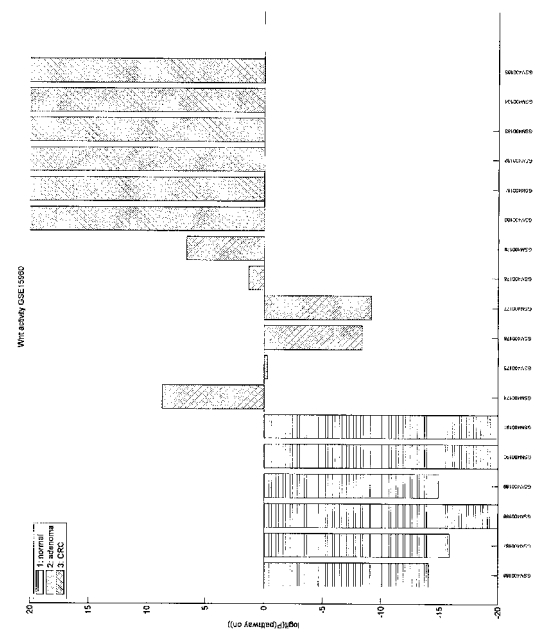
【図16 - I】

証拠により精選したリスト



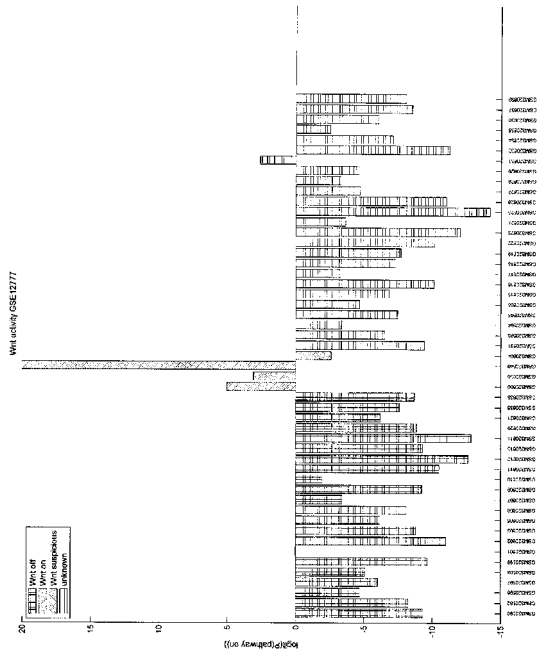
【図16 - II】

広範な文献によるリスト



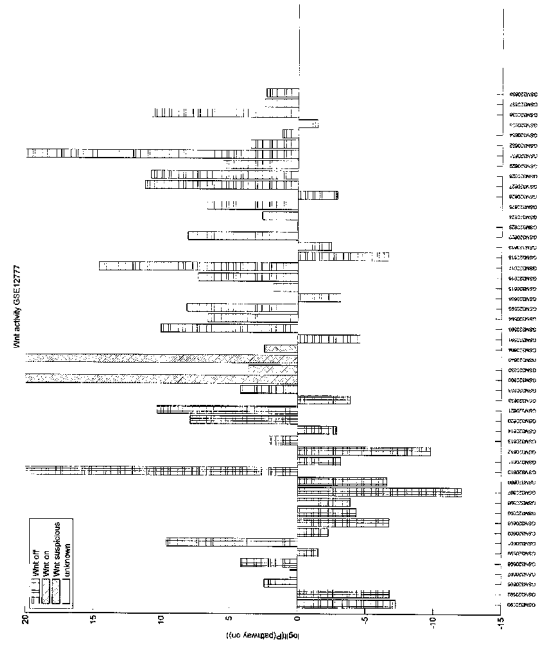
【 図 17 - I 】

証拠により精選したリスト



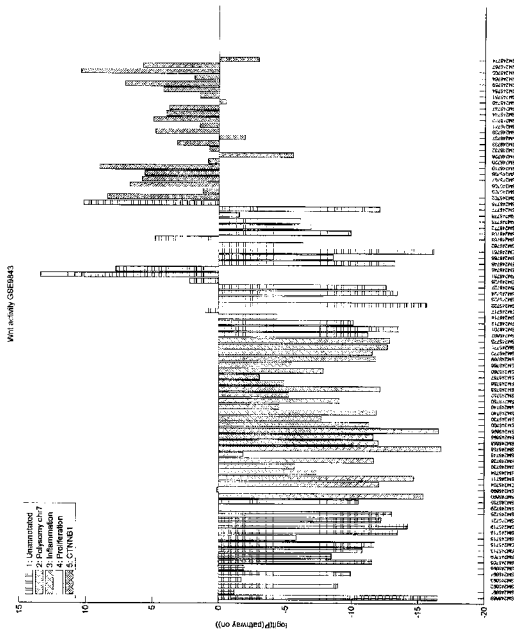
【 図 17 - II 】

広範な文献によるリスト



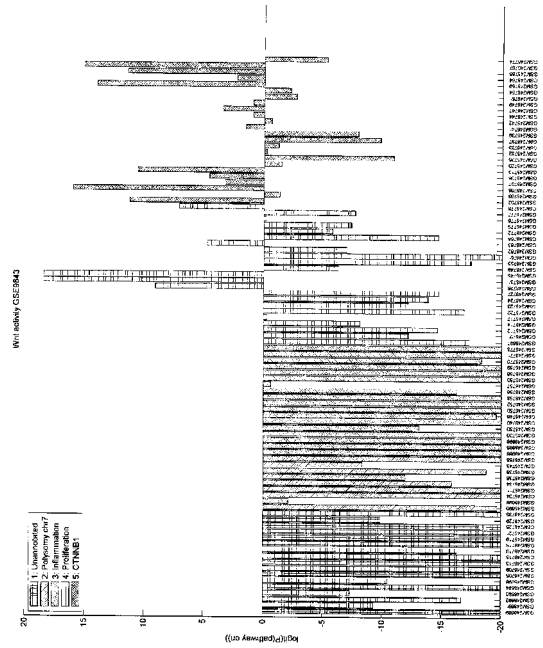
【 図 18 - I 】

証拠により精選したリスト



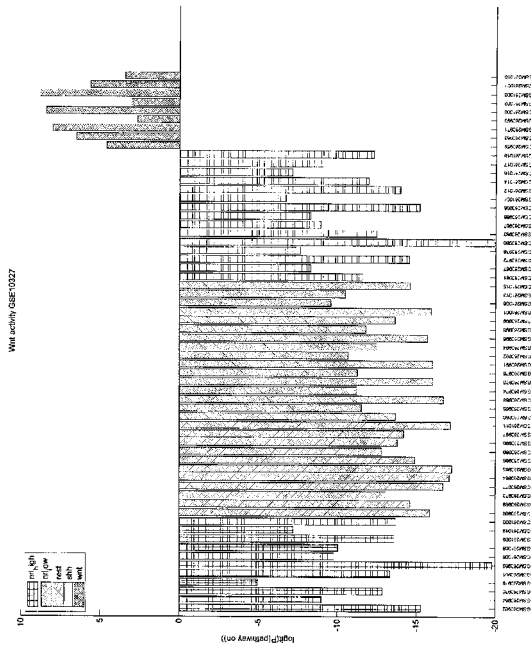
【 図 18 - II 】

広範な文献によるリスト



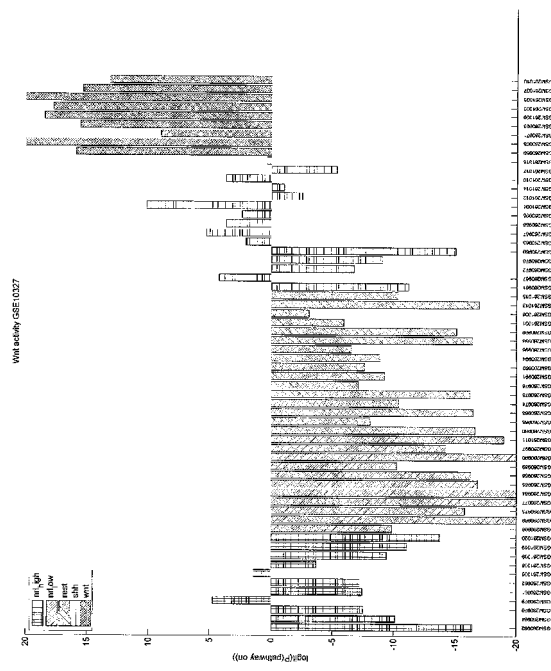
【図19-I】

証拠により精選したリスト

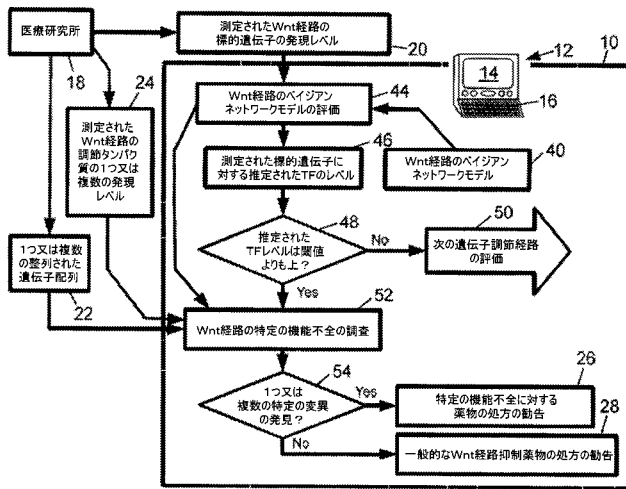


【図19-II】

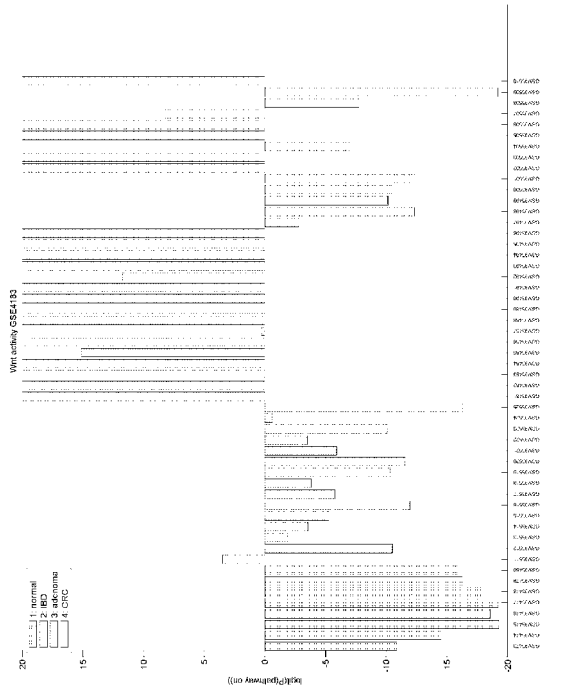
広範な文献によるリスト



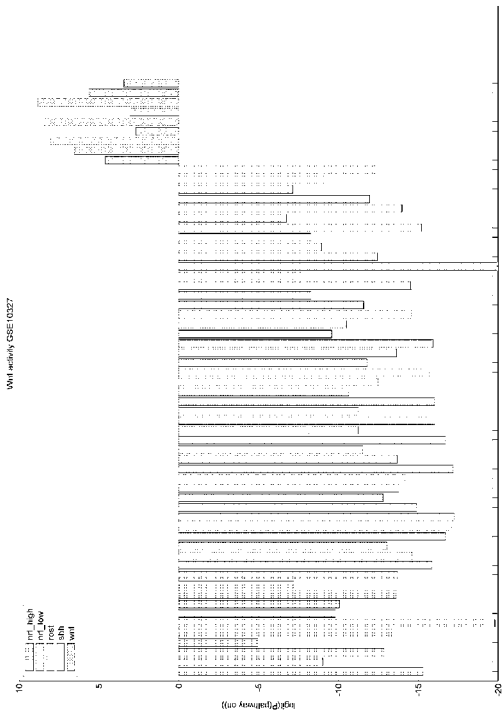
【図20】



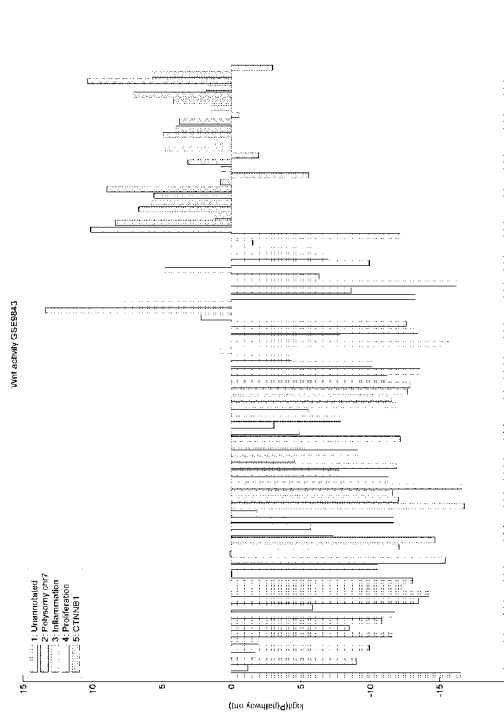
【図21】



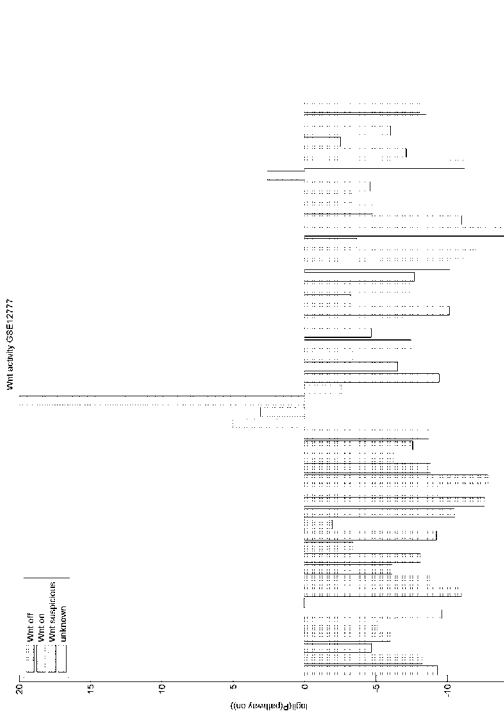
【 図 2 2 】



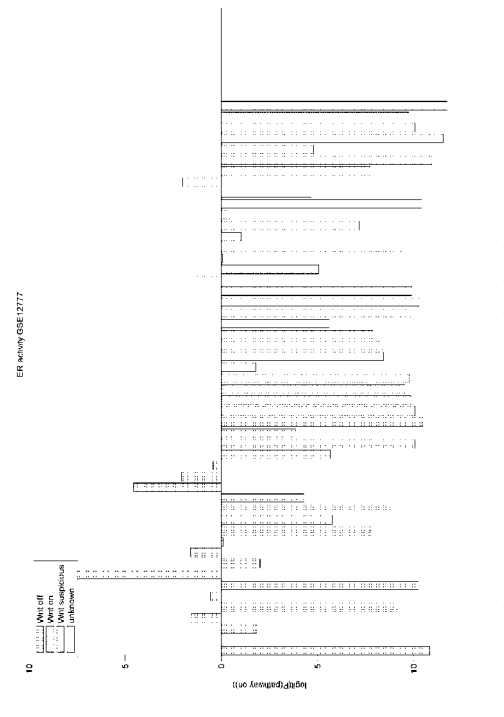
【 図 2 3 】



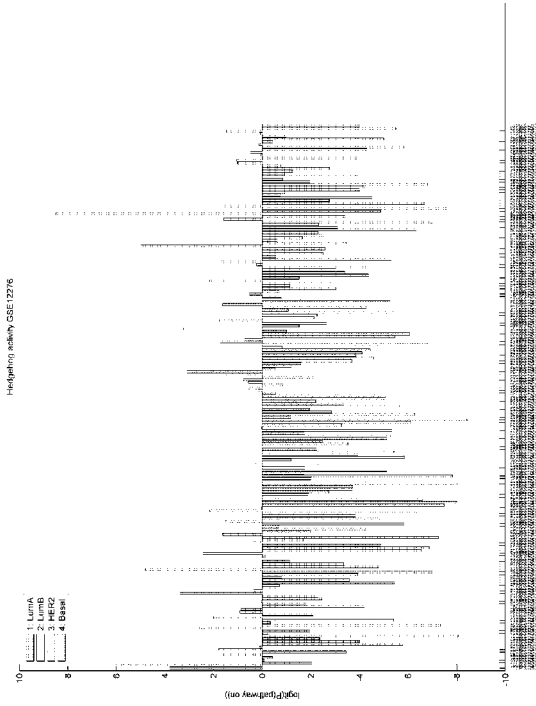
【 図 2 4 】



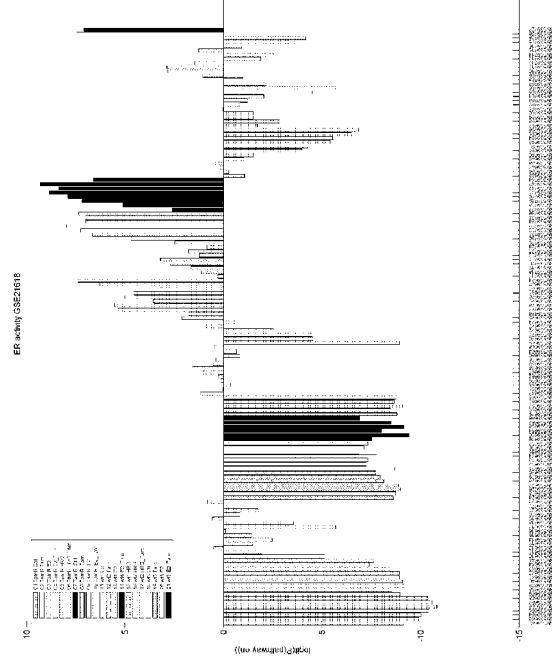
【 図 2 5 】



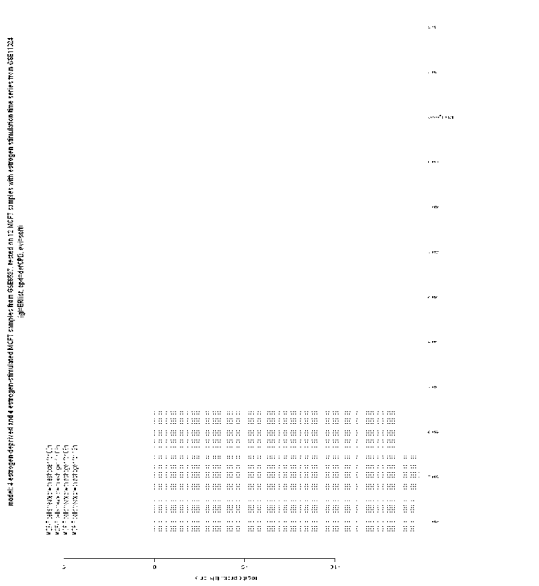
【 30 】



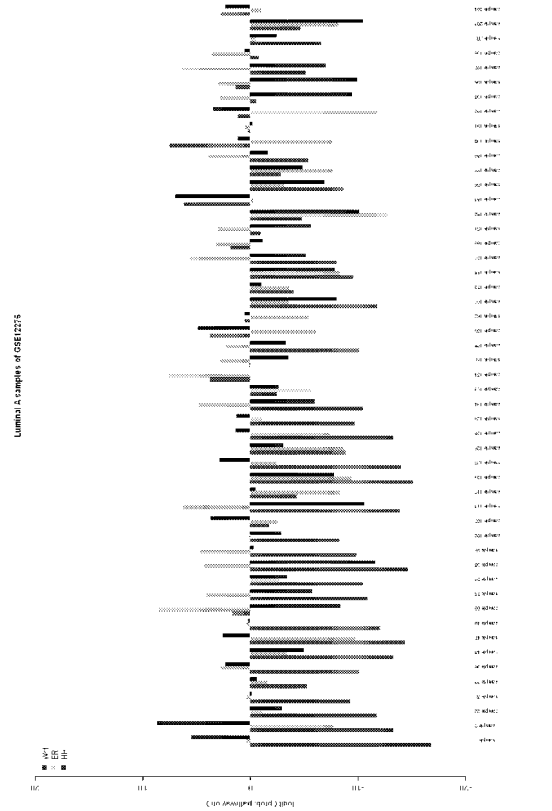
【 31 】



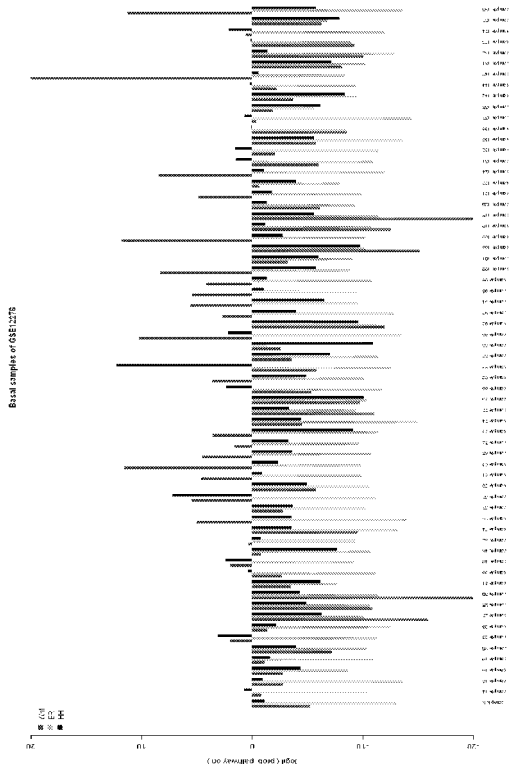
【 32 】



【 33 】



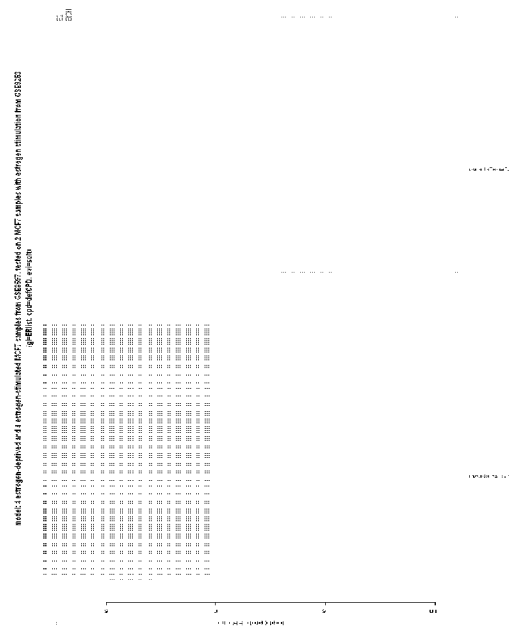
【 図 3 4 】



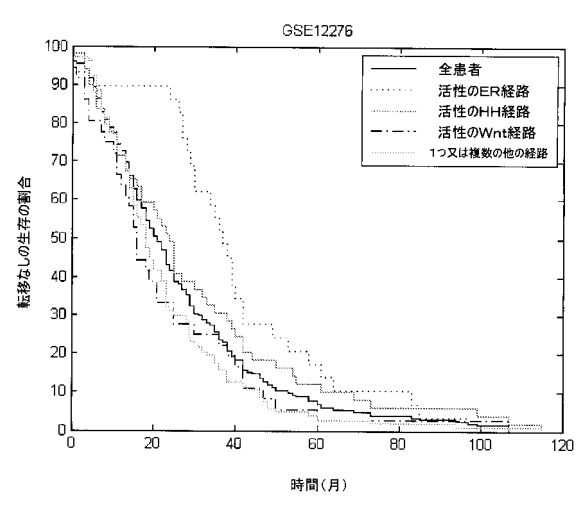
【 図 3 5 】



【 図 3 6 】

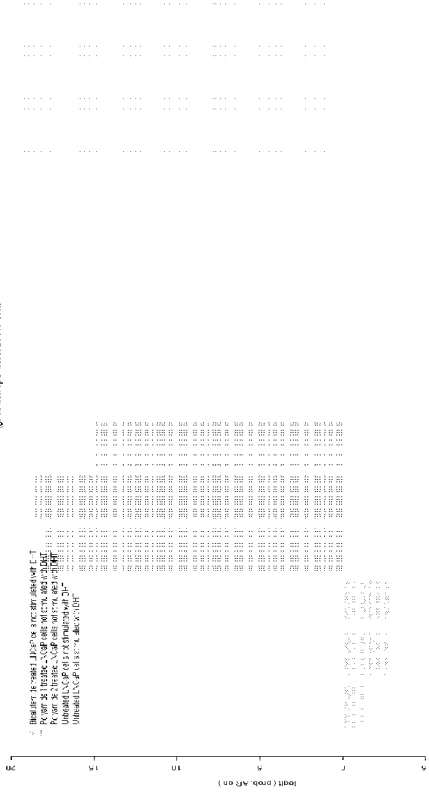


【 図 3 7 】



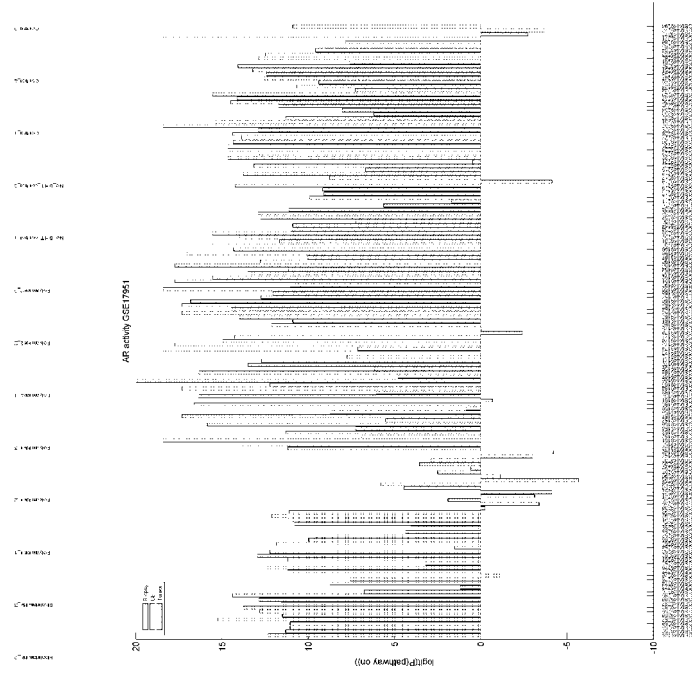
【 38 】

model 3 simulated and 3 non-simulated LUCIF samples from GSE7938. Mixed on is treated LUCIF samples from GSE7938 (log2(RPKM + 0.1))



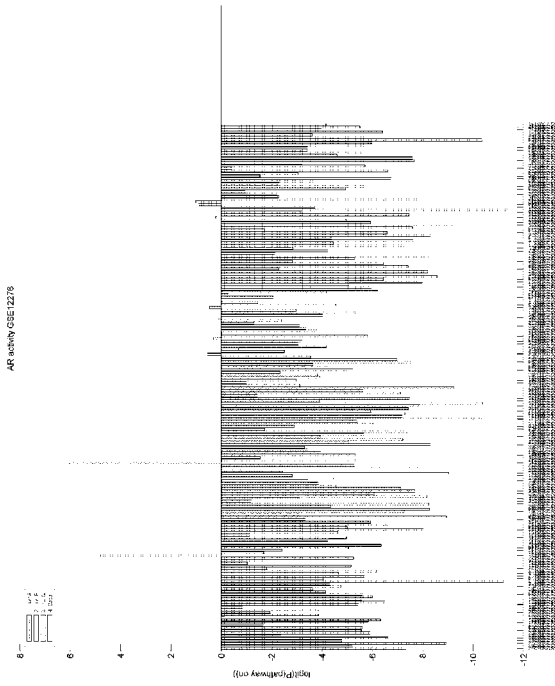
【 39 】

AP activity GSE17951



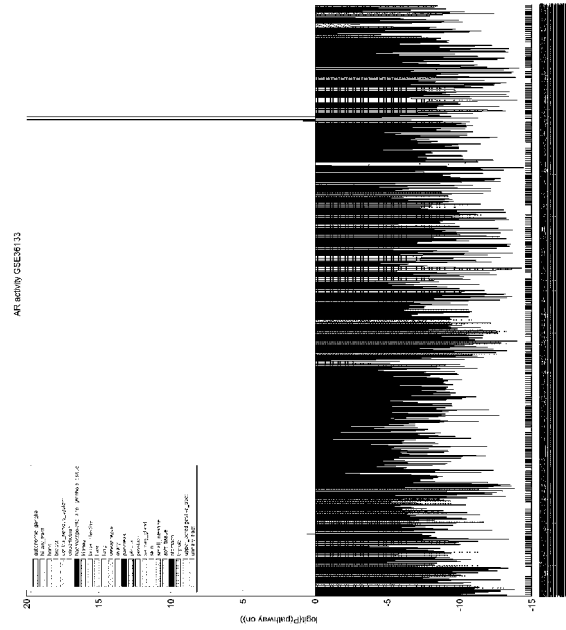
【 40 】

AR activity GSE12276

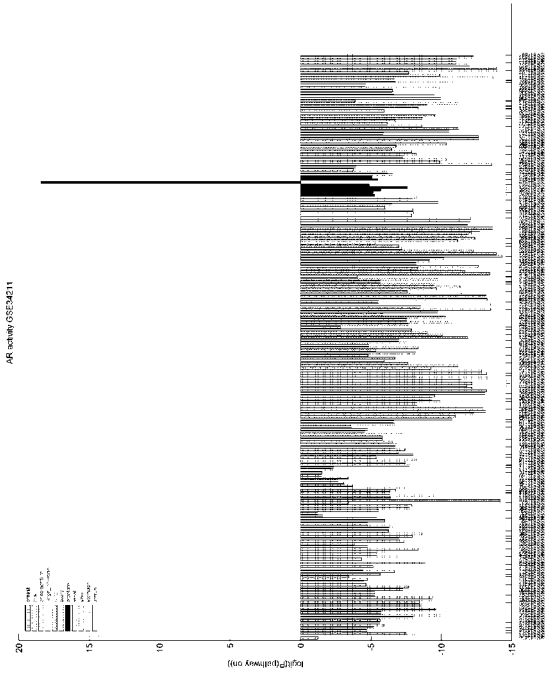


【 41 】

AR activity GSE6133



【 図 4 2 】



【 配列表 】

2019129850000001.app

フロントページの続き

(74)代理人 100091214

弁理士 大貫 進介

(72)発明者 フェルハーフ, ウィルヘルムス フランシスキュス ヨハネス

オランダ国, 5 6 5 6 アーエー アインドーフエン, ハイ・テク・キャンパス・ビルディング
4 4

(72)発明者 ファン デ ストルペ, アニャ

オランダ国, 5 6 5 6 アーエー アインドーフエン, ハイ・テク・キャンパス・ビルディング
4 4

(72)発明者 ファン オーイエン, ヘンドリック ヤン

オランダ国, 5 6 5 6 アーエー アインドーフエン, ハイ・テク・キャンパス・ビルディング
4 4

(72)発明者 ドゥラ, カリアナ チャクラヴァーティ

オランダ国, 5 6 5 6 アーエー アインドーフエン, ハイ・テク・キャンパス・ビルディング
4 4

(72)発明者 アルヴェス デ インダ, マルシア

オランダ国, 5 6 5 6 アーエー アインドーフエン, ハイ・テク・キャンパス・ビルディング
4 4

(72)発明者 ホフマン, ラルフ

オランダ国, 5 6 5 6 アーエー アインドーフエン, ハイ・テク・キャンパス・ビルディング
4 4

F ターム(参考) 2G045 AA24 AA25 AA26 CB01 DA14 DA36 JA01

4B063 QA01 QA05 QQ02 QQ08 QQ53 QR55 QR62 QR82