

(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2000-308500**

(P2000-308500A)

(43)公開日 平成12年11月7日(2000.11.7)

(51) Int.Cl <sup>7</sup>	識別記号	庁内整理番号	F I	技術表示箇所
C 1 2 Q 1/68			C 1 2 Q 1/68	A
C 1 2 N 5/10			G 0 1 N 33/15	Z
15/09			33/50	Z
G 0 1 N 33/15			33/53	M
33/50			33/566	

審査請求 未請求 請求項の数 65〇L (全 77数) 最終頁に続く

(21)出願番号 特願2000-64155(P2000-64155)

(22)出願日 平成12年2月2日(2000.2.2)

(31)優先権主張番号 09/256,301

(32)優先日 平成11年2月24日(1999.2.24)

(33)優先権主張国 米国(US)

(71)出願人 399125757

アフィメトリックス インコーポレイテッド

アメリカ合衆国 カリフォルニア州 9505  
1 サンタ クララ セントラル エクスプレスウェイ 3380

(72)発明者 オリナー ジョナサン

アメリカ合衆国 カリフォルニア州 マウンテン ピュー ユニット 22 シエラビスタ アベニュー 173

(74)代理人 100102978

弁理士 清水 初志 (外1名)

最終頁に続く

(54)【発明の名称】癌抑制遺伝子WT1の下流遺伝子

(57)【要約】

【課題】WT1遺伝子に関する癌の診断、薬物スクリーニング、および変異の機能分析方法を提供することを課題とする。

【解決手段】本方法は、WT1、およびEWSと融合したWT1によって調節される、新規に同定された遺伝子セットの使用を包含する。これらの遺伝子セットの発現レベルのモニターは、遺伝子の遺伝的状態の指標として使用できる。また、下流遺伝子に対してどれが同様な影響を持つかの同定もできる。

## 【特許請求の範囲】

【請求項1】 下記の段階を含む、標的細胞においてWT1遺伝子の機能変異を検出する方法：

(a) 標的細胞、および(b) 野生型WT1遺伝子を持ち、その他の点では標的細胞に実質的に類似した参照細胞のサンプル中で、少なくとも3つのWT1の下流遺伝子の発現を検出する段階であって、下流遺伝子は野生型WT1遺伝子にアップレギュレーションまたはダウンレギュレーションされる検出段階；および標的細胞および参照細胞における下流遺伝子の発現を比較する段階であって、標的細胞と参照細胞とで発現に差があれば、標的細胞にWT1機能変異が存在することが示唆される段階。

【請求項2】 下流遺伝子が、野生型WT1遺伝子によって転写調節されており、該下流遺伝子の発現が、参照細胞および標的細胞中の該下流遺伝子の転写産物の量を測定することによって検出される、請求項1記載の方法。

【請求項3】 転写産物の量が、高密度核酸アレイを用いて測定される、請求項1記載の方法。

【請求項4】 下流遺伝子が、ナチュラルキラー細胞プロテイン4前駆体 (M59807)、葉酸結合蛋白質 (M2531 7)、チューブリンのHALPHA44遺伝子 (X06956)、熱ショック蛋白質HSP70B (X51758)、90K産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、アンフィレグリン (M3 0704)、プロコラーゲン 1 (T51558)、遊走性プラスミノーゲン活性化因子阻害剤1 (M14083)、放射線照射ケラチノサイト由来のシステインプロテアーゼ阻害剤 (X0 5978)、脳性ナトリウム利尿蛋白質(brain natriuretic protein)(BNP)(M31766)、白血病ウイルス受容体1 (GLVR 1)(L20859)、1型細胞骨格17ケラチン (R71870)、チューブリン 5鎖 (H45051)、プリンヌクレオシドホスホリラーゼ (T47964)、Gem GTPアーゼ (U10550)、アドレノメダリン (D14874)、GPIアンカー型ウロキナーゼプラスミノーゲン活性化因子表面受容体 (R38636)、組織型プラスミノーゲン活性化因子 (X13097)、グラビン (M9632 2)、jun-B (X51345)、延長因子1 -2 (X79490)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57 348)、インターフェロン 処理誘導可能mRNA (M2668 3)、膜貫通受容体蛋白質 (Z17227)、酸性FGF (X6577 9)、軟骨結合蛋白質-1 (X17405)、IL-11 (X58377)、およびミトコンドリアリン酸輸送蛋白質 (R49231) からなる群より選択されるWT1にアップレギュレーションされる遺伝子を、少なくとも1つ含む、請求項1記載の方法。

【請求項5】 下流遺伝子が、カベオリンおよび延長因子1 -2からなる群より選択される、WT1にダウンレギュレーションされる遺伝子を少なくとも1つ含む、請求項1記載の方法。

【請求項6】 下流遺伝子がp21またはEGFRをさらに含む、請求項4記載の方法。

【請求項7】 WT1にアップレギュレーションされる遺伝子の発現が、参照細胞と比較して、標的細胞で少なく

とも2分の1以下であるか、WT1にダウンレギュレーションされる遺伝子の発現が、参照細胞と比較して標的細胞で少なくとも2倍以上である場合に、標的細胞中のWT1遺伝子が機能損失変異であることを示す段階をさらに含む、請求項1記載の方法。

【請求項8】 下記の段階を含む、標的細胞中でEWS-WT1遺伝子融合体を検出する方法：

(a) 標的細胞、および(b) 野生型EWSおよびWT1遺伝子を持ち、その他の点では標的細胞に実質的に類似した参照細胞のサンプル中で、1つまたは複数のEWS-WT1融合体の下流遺伝子の発現を検出する段階であって、該下流遺伝子は野生型EWSおよびWT1遺伝子にアップレギュレーションまたはダウンレギュレーションされる検出段階；および標的細胞および参照細胞における該下流遺伝子の発現を比較する段階であって、標的細胞と参照細胞とで発現に差があれば、標的細胞中にEWS-WT1融合体が存在することが示唆される段階。

【請求項9】 下流遺伝子が、仮性狂犬病ウイルス核抗原 (R60906およびT79475)、インターロイキン-2受容体鎖 (M26062)、G1/S-特異的サイクリンD1 (H2052 9)、ニューロン特異的 -2工ノラーゼ (M22349)、転座t(3;5)(q25.1;p34)により得られる融合遺伝子NPM-MLF1 (L49054)、コラーゲン 3(IV)鎖 (H62466)、T-リンパ球特異的蛋白質チロシンキナーゼp56lck (lck) (U238 52)、インスリン様成長因子結合蛋白質-4 (U20982)、カテーテンB (L16510)、シスタチンC (T51534)、二重特異的マイトジエン活性化蛋白質キナーゼキナーゼ2 (R42291)、神経グリア細胞接着分子 (T53118)、PIM-1癌原遺伝子セリン/スレオニン-蛋白質キナーゼ (H1092 5)、カルボキシルエステラーゼ前駆体 (R54359)、蛋白質-チロシンホスファターゼPAC-1 (L11329)、およびインターフェロン誘導蛋白質6-16 (X02492) からなる群より選択される、EWS-WT1にアップレギュレーションされる遺伝子を少なくとも一つ含む、請求項8記載の方法。

【請求項10】 下記の段階を含む、WT1配列変化の細胞内機能アッセイ法：WT1配列変化を有する標的細胞の標的サンプル、および野生型WT1遺伝子を持ち、その他の点では標的細胞に実質的に類似した参照細胞の参照サンプルにおいて、少なくとも3つの下流遺伝子の発現を検出する段階であって、該下流遺伝子は野生型WT1遺伝子にアップレギュレーションまたはダウンレギュレーションされる検出段階；および該標的サンプル中の該発現と該参照サンプル中の該発現とを比較する段階であって、2つのサンプル間で発現に差があれば、該WT1配列変化がWT1の生物学的機能に影響を与えることが示唆される段階。

【請求項11】 下流遺伝子が、野生型WT1遺伝子によって転写調節されており、該下流遺伝子の発現が、参照細胞および標的細胞中の該下流遺伝子の転写産物の量を

測定することによって検出される、請求項10記載の方法。

【請求項12】 転写産物の量が、高密度核酸アレイを用いて測定される、請求項11記載の方法。

【請求項13】 下流遺伝子が、ナチュラルキラー細胞プロテイン4前駆体(M59807)、葉酸結合蛋白質(M25317)、チューブリンのHALPHA44遺伝子(X06956)、熱ショック蛋白質HSP70B(X51758)、90K産物(H17969)、熱ショック70kd蛋白質1(T66307)、アンフィレグリン(M30704)、プロコラーゲン1(T51558)、遊走性プラスミノーゲン活性化因子阻害剤1(M14083)、放射線照射ケラチノサイト由来のシスティンプロテアーゼ阻害剤(X05978)、脳性ナトリウム利尿蛋白質(BNP)(M31766)、白血病ウイルス受容体1(GLVR1)(L20859)、1型細胞骨格17ケラチン、チューブリン5鎖(H45051)、プリンヌクレオシドホスホリラーゼ(T47964)、Gem GTPアーゼ(U10550)、アドレノメダリン(D14874)、GPIアンカー型ウロキナーゼプラスミノーゲン活性化因子表面受容体(R38636)、組織型プラスミノーゲン活性化因子(X13097)、グラビン(M96322)、jun-B(X51345)、ホメオティック遺伝子調節因子(R16977)、クローン9112(X57348)、インターフェロン処理誘導可能mRNA(M26683)、膜貫通受容体蛋白質(Z17227)、酸性FGF(X65779)、軟骨結合蛋白質-1(X17405)、IL-11(X58377)、およびミトコンドリアリン酸輸送蛋白質(R49231)からなる群より選択されるWT1にアップレギュレーションされる遺伝子を少なくとも1つ含む、請求項10記載の方法。

【請求項14】 下流遺伝子が、カベオリンおよび延長因子1-2(X79490)からなる群より選択される、WT1にダウンレギュレーションされる遺伝子を少なくとも1つ含む、請求項10記載の方法。

【請求項15】 下流遺伝子がp21またはEGFRをさらに含む、請求項13記載の方法。

【請求項16】 WT1にアップレギュレーションされる遺伝子の発現が、参照細胞と比較して、標的細胞で少なくとも2分の1以下であるか、WT1にダウンレギュレーションされる遺伝子の発現が、参照細胞と比較して標的細胞で少なくとも2倍以上である場合に、WT1配列変化が野生型の機能損失変異であることを示す段階をさらに含む、請求項10記載の方法。

【請求項17】 下記の段階を含む、コンピュータを用いた標的WT1遺伝子中の変異の検出方法：野生型WT1遺伝子を含む野生型サンプル中の、該野生型WT1遺伝子によって転写調節を受ける少なくとも3つの下流遺伝子の野生型発現データをコンピュータに入力する段階；標的WT1遺伝子を含む標的サンプル中の複数の該下流遺伝子の標的の発現データをコンピュータに入力する段階；ならびに標的および野生型の発現データを比較して、差を測定する段階であって、差があれば、該標的WT1遺伝子中に変異が存在することが示唆される段階。

【請求項18】 下流遺伝子が、ナチュラルキラー細胞

プロテイン4前駆体(M59807)、葉酸結合蛋白質(M25317)、チューブリンのHALPHA44遺伝子(X06956)、熱

ショック蛋白質HSP70B(X51758)、90K産物(H17969)、熱ショック70kd蛋白質1(T66307)、アンフィレグリン(M30704)、プロコラーゲン1(T51558)、遊走性プラスミノーゲン活性化因子阻害剤1(M14083)、放射線照射ケラチノサイト由来のシスティンプロテアーゼ阻害剤(X05978)、脳性ナトリウム利尿蛋白質(BNP)(M31766)、白

血病ウイルス受容体1(GLVR1)(L20859)、1型細胞骨格17ケラチン、チューブリン5鎖(H45051)、プリンヌクレオシドホスホリラーゼ(T47964)、Gem GTPアーゼ(U10550)、アドレノメダリン(D14874)、GPIアンカー型ウロキナーゼプラスミノーゲン活性化因子表面受容体(R38636)、組織型プラスミノーゲン活性化因子(X13097)、グラビン(M96322)、jun-B(X51345)、ホメオティック遺伝子調節因子(R16977)、クローン9112(X57348)、インターフェロン処理誘導可能mRNA(M26683)、膜貫通受容体蛋白質(Z17227)、酸性FGF(X65779)、軟骨結合蛋白質-1(X17405)、IL-11(X58377)、およびミトコンドリアリン酸輸送蛋白質(R49231)からなる群より選択されるWT1にアップレギュレーションされる遺伝子を少なくとも1つ含む、請求項17記載の方法。

【請求項19】 下流遺伝子が、カベオリンおよび延長因子1-2(X79490)からなる群より選択される、WT1にダウンレギュレーションされる遺伝子を少なくとも1つ含む、請求項17記載の方法。

【請求項20】 下流遺伝子がp21またはEGFRをさらに含む、請求項18記載の方法。

【請求項21】 WT1にアップレギュレーションされる遺伝子の発現が、参照細胞と比較して、標的細胞で少なくとも2分の1以下であるか、WT1にダウンレギュレーションされる遺伝子の発現が、参照細胞と比較して標的細胞で少なくとも2倍以上である場合に、標的WT1遺伝子中に変異があること示す段階をさらに含む、請求項17記載の方法。

【請求項22】 下記の段階を含む、コンピュータを用いたEWSとWT1遺伝子とを融合する転座の検出方法：野生型EWSおよびWT1遺伝子を含む野生型サンプル中の、EWS-

WT1融合蛋白質によって転写調節を受ける複数の下流遺伝子の野生型発現データをコンピュータに入力する段階；EWS-WT1融合蛋白質の存在を試験する標的サンプル中の複数の該下流遺伝子の標的の発現データをコンピュータに入力する段階；ならびに標的および野生型の発現データを比較して、差を測定する段階であって、差があればEWSとWT1遺伝子を融合する転座があることが示唆される段階。

【請求項23】 下流遺伝子が、仮性狂犬病ウイルス核抗原(R60906およびT79475)、インターロイキン-2受容

体鎖(M26062)、G1/S-特異的サイクリンD1(H2052

9)、ニューロン特異的 -2エノラーゼ (M22349)、転座t(3;5)(q25.1;p34)により得られる融合遺伝子NPM-MLF1 (L49054)、コラーゲン 3(IV)鎖 (H62466)、T-リンパ球特異的蛋白質チロシンキナーゼp56lck (lck) (U23852)、インスリン様成長因子結合蛋白質-4 (U20982)、カテーテンB、シスタチンC (T51534)、二重特異的マイトジエン活性化蛋白質キナーゼキナーゼ2 (R42291)、神経グリア細胞接着分子 (T53118)、PIM-1癌原遺伝子セリン/スレオニン-蛋白質キナーゼ (H10925)、カルボキシルエステラーゼ前駆体 (R54359)、蛋白質-チロシンホスファターゼPAC-1 (L11329)、およびインターフェロン誘導蛋白質6-16 (X02492) からなる群より選択される、EWS-WT1にアップレギュレーションされる遺伝子を少なくとも一つ含む、請求項22記載の方法。

【請求項 24】 下記の段階を含む、試験細胞の新生物形成性の診断方法：mRNA、cDNAおよびcRNAからなる群より選択される試験細胞の転写指標を、核酸プローブのセットにハイブリダイズさせる段階であって、核酸プローブのセットが、ナチュラルキラー細胞プロテイン4前駆体 (M59807)、熱ショック蛋白質HSP70B (X51758)、90K 産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、プロコラーゲン 1 (T51558)、放射線照射ケラチノサイト由来のシステインプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質 (BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン (R71870)、プリンヌクレオシドホスホリラーゼ (T47964)、アドレノメダリン (D14874)、グラビン (M96322)、jun-B (X51345)、延長因子1 -2 (X79490)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57348)、インターフェロン 処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、酸性FGF (X65779)、軟骨結合蛋白質-1 (X17405)、IL-11 (X58377)、およびミトコンドリアリン酸輸送蛋白質 (R49231) からなる群より選択される、各々がWT1に活性化または抑制される遺伝子の一部である複数の核酸分子を含むハイブリダイズ段階；各々の該核酸プローブセットにハイブリダイズする転写指標の量を検出する段階；ならびに(1) WT1に活性化される遺伝子であるプローブへの試験細胞の転写指標のハイブリダイゼーションが、正常細胞の転写指標のハイブリダイゼーションよりも低レベルである、または(2) WT1に抑制される遺伝子であるプローブへの試験細胞の転写指標のハイブリダイゼーションが、正常細胞の転写指標のハイブリダイゼーションよりも高レベルである場合に、試験細胞が新生物形成性であると同定する段階。

【請求項 25】 試験細胞が腎臓細胞である、請求項24記載の方法。

【請求項 26】 プローブの少なくとも4つが、WT1に活性化またはWT1に抑制される遺伝子の一部を含んでいる、請求項24記載の方法。

(4)  
10 6

【請求項 27】 プローブの少なくとも10個が、WT1に活性化またはWT1に抑制される遺伝子の一部を含んでいる、請求項24記載の方法。

【請求項 28】 プローブの少なくとも15個が、WT1に活性化またはWT1に抑制される遺伝子の一部を含んでいる、請求項24記載の方法。

【請求項 29】 核酸プローブが固体支持体に結合している、請求項24記載の方法。

10 【請求項 30】 核酸プローブがアレイ上に配置されている、請求項24記載の方法。

【請求項 31】 アレイが少なくとも250の異なる遺伝子の一部である核酸プローブを含んでいる、請求項30記載の方法。

【請求項 32】 アレイが少なくとも6000の異なる遺伝子の一部である核酸プローブを含んでいる、請求項30記載の方法。

【請求項 33】 細胞のWT1遺伝子型の状態を決定するために、試験細胞中でWT1遺伝子の配列を決定する段階をさらに含む、請求項24記載の方法。

20 【請求項 34】 比較されるサンプル間でハイブリダイゼーションに少なくとも2倍の差がある場合に試験細胞が新生物形成性であると同定される、請求項24記載の方法。

【請求項 35】 下記の段階を含む、抗癌剤の同定方法：試験化合物をヒト細胞と接触させる段階；WT1にアップレギュレーションされる少なくとも1つの遺伝子、またはWT1にダウンレギュレーションされる少なくとも1つの遺伝子の、細胞における発現を測定する段階であって、WT1にアップレギュレーションされる遺伝子が、ナ

30 チュラルキラー細胞プロテイン4前駆体 (M59807)、熱ショック蛋白質HSP70B (X51758)、90K 産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、プロコラーゲン 1 (T51558)、放射線照射ケラチノサイト由来のシステインプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質 (BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン (R71870)、プリンヌクレオシドホスホリラーゼ (T47964)、アドレノメダリン (D14874)、グラビン (M96322)、jun-B (X51345)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57348)、インターフェロン 処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、酸性FGF (X65779)、軟骨結合蛋白質-1 (X17405)、IL-11 (X58377)、およびミトコンドリアリン酸輸送蛋白質 (R49231) からなる群より選択され、WT1にダウンレギュレーションされる遺伝子が延長因子1 -2 (X70940) である測定段階；および試験化合物が、ヒト細胞においてWT1にアップレギュレーションされる少なくとも1つの遺伝子の発現を上昇させるか、WT1にダウンレギュレーションされる少なくとも1つの遺伝子の発現を低下させる場合に、抗癌剤の可能性があると同定する段階。

40 50

【請求項36】 細胞が腎臓癌細胞である、請求項35記載の方法。

【請求項37】 細胞が腫瘍細胞である、請求項35記載の方法。

【請求項38】 細胞に野生型WT1遺伝子が含まれない、請求項35記載の方法。

【請求項39】 WT1にアップレギュレーションおよびダウンレギュレーションされる遺伝子の発現が、細胞中の該遺伝子の転写産物の量を測定することにより検出される、請求項35記載の方法。

【請求項40】 転写産物の量が高密度核酸アレイを用いて測定される、請求項39記載の方法。

【請求項41】 下記の段階を含む、試験細胞の新生物形成性の診断方法：mRNA、cDNAおよびcRNAからなる群より選択される試験細胞の転写指標を、核酸プローブのセットにハイブリダイズさせる段階であって、核酸プローブのセットは、各々がEWS-WT1融合蛋白質に活性化または抑制される遺伝子の一部である複数の核酸分子を含むハイブリダイズ段階；各々の該核酸プローブセットにハイブリダイズする転写指標の量を検出する段階；ならびに(1) EWS-WT1に活性化される遺伝子であるプローブへの試験細胞の転写指標のハイブリダイゼーションが、正常細胞の転写指標のハイブリダイゼーションよりも低レベルである、または(2) EWS-WT1に抑制される遺伝子であるプローブへの試験細胞の転写指標のハイブリダイゼーションが、正常細胞の転写指標のハイブリダイゼーションよりも高レベルである場合に、試験細胞が新生物形成性であると同定する段階。

【請求項42】 試験細胞がユーイング肉腫細胞である、請求項41記載の方法。

【請求項43】 プローブの少なくとも4つが、EWS-WT1に活性化またはEWS-WT1に抑制される遺伝子の一部を含んでいる、請求項41記載の方法。

【請求項44】 プローブの少なくとも10個が、EWS-WT1に活性化またはEWS-WT1に抑制される遺伝子の一部を含んでいる、請求項41記載の方法。

【請求項45】 プローブの少なくとも15個が、EWS-WT1に活性化またはEWS-WT1に抑制される遺伝子の一部を含んでいる、請求項41記載の方法。

【請求項46】 核酸プローブが固体支持体に結合している、請求項41記載の方法。

【請求項47】 核酸プローブがアレイ上に配置されている、請求項41記載の方法。

【請求項48】 アレイが少なくとも250の異なる遺伝子の一部である核酸プローブを含んでいる、請求項47記載の方法。

【請求項49】 アレイが少なくとも6000の異なる遺伝子の一部である核酸プローブを含んでいる、請求項47記載の方法。

【請求項50】 少なくとも1つの核酸プローブが、仮

性狂犬病ウイルス核抗原 (R60906およびT79475)、インターロイキン-2受容体 鎖 (M26062)、G1/S-特異的サイクリンD1 (H20529)、ニューロン特異的 -2エノラーゼ (M22349)、転座t(3;5)(q25.1;p34)により得られる融合遺伝子NPM-MLF1 (L49054)、コラーゲン 3(IV)鎖 (H62466)、T-リンパ球特異的蛋白質チロシンキナーゼp56lck(lck) (U23852)、インスリン様成長因子結合蛋白質-4 (U20982)、カテプシンB、シスタチンC (T51534)、二重特異的マイトジエン活性化蛋白質キナーゼキナーゼ2 (R42291)、神経グリア細胞接着分子 (T53118)、PIM-1癌原遺伝子セリン/スレオニン-蛋白質キナーゼ (H10925)、カルボキシリエステラーゼ前駆体 (R54359)、蛋白質-チロシンホスファターゼPAC-1 (L11329)、およびインターフェロン誘導蛋白質6-16 (X02492)からなる群より選択される、EWS-WT1にアップレギュレーションされる遺伝子の一部を含む、請求項41記載の方法。

【請求項51】 比較されるサンプル間でハイブリダイゼーションに少なくとも2倍の差がある場合に試験細胞が新生物形成性であると同定される、請求項41記載の方法。

【請求項52】 比較されるサンプル間でハイブリダイゼーションに少なくとも3倍の差がある場合に試験細胞が新生物形成性であると同定される、請求項41記載の方法。

【請求項53】 比較されるサンプル間でハイブリダイゼーションに少なくとも5倍の差がある場合に試験細胞が新生物形成性であると同定される、請求項41記載の方法。

30 【請求項54】 下記の段階を含む、抗癌剤の同定方法：試験化合物をヒト細胞と接触させる段階；EWS-WT1にアップレギュレーションされる少なくとも1つの遺伝子、またはEWS-WT1にダウンレギュレーションされる少なくとも1つの遺伝子のヒト細胞における発現を測定する段階；および試験化合物が、EWS-WT1にアップレギュレーションされる遺伝子の発現を低下させるか、EWS-WT1にダウンレギュレーションされる遺伝子の発現を上昇させる場合に、抗癌剤であると同定する段階。

【請求項55】 細胞が腎臓癌細胞である、請求項54記載の方法。

【請求項56】 細胞が腫瘍細胞である、請求項54記載の方法。

【請求項57】 WT1にアップレギュレーションおよびダウンレギュレーションされる遺伝子の発現が、細胞中の該遺伝子の転写産物の量を測定することにより検出される、請求項54記載の方法。

【請求項58】 転写産物の量が高密度核酸アレイを用いて測定される、請求項57記載の方法。

50 【請求項59】 WT1にアップレギュレーションされる遺伝子が、仮性狂犬病ウイルス核抗原 (R60906およびT7

9475)、インターロイキン-2受容体鎖(M26062)、G1/S-特異的サイクリンD1(H20529)、ニューロン特異的-2工ノラーゼ(M22349)、転座t(3;5)(q25.1;p34)により得られる融合遺伝子NPM-MLF1(L49054)、コラーゲン3(IV)鎖(H62466)、T-リンパ球特異的蛋白質チロシンキナーゼp56lck(lck)(U23852)、インスリン様成長因子結合蛋白質-4(U20982)、カテプシンB、シスタチンC(T51534)、二重特異的マイトジエン活性化蛋白質キナーゼキナーゼ2(R42291)、神経グリア細胞接着分子(T53118)、PIM-1癌原遺伝子セリン/スレオニン-蛋白質キナーゼ(H10925)、カルボキシリエステラーゼ前駆体(R54359)、蛋白質-チロシンホスファターゼPAC-1(L11329)、およびインターフェロン誘導蛋白質6-16(X02492)からなる群より選択される、請求項54記載の方法。

【請求項60】下記の段階を含む、標的細胞中でWT1遺伝子の機能変異を検出する方法：

(a) 標的細胞、および(b) 野生型WT1遺伝子を持ち、その他の点では標的細胞に実質的に類似した参照細胞のサンプル中で、少なくとも1つのWT1の下流遺伝子の発現を検出する段階であって、該下流遺伝子が、ナチュラルキラー細胞プロテイン4前駆体(M59807)、葉酸結合蛋白質(M25317)、チューブリンのHALPHA44遺伝子(X06956)、熱ショック蛋白質HSP70B(X51758)、90K産物(H17969)、熱ショック70kd蛋白質1(T66307)、アンフィレグリン(M30704)、プロコラーゲン1(T51558)、遊走性プラスミノーゲン活性化因子阻害剤1(M14083)、放射線照射ケラチノサイト由来のシステインプロテアーゼ阻害剤(X05978)、脳性ナトリウム利尿蛋白質(BNP)(M31766)、白血病ウイルス受容体1(GLVR1)(L20859)、1型細胞骨格17ケラチン、チューブリン5鎖(H45051)、プリンヌクレオシドホスホリラーゼ(T47964)、Gem GTPアーゼ(U10550)、アドレノメダリン(D14874)、GPIアンカー型ウロキナーゼプラスミノーゲン活性化因子表面受容体(R38636)、組織型プラスミノーゲン活性化因子(X13097)、グラビン(M96322)、jun-B(X51345)、延長因子1-2(X79490)、ホメオティック遺伝子調節因子(R16977)、クローン9112(X57348)、インターフェロン処理誘導可能mRNA(M26683)、膜貫通受容体蛋白質(Z17227)、酸性FGF(X65779)、軟骨結合蛋白質-1(X17405)、IL-11(X58377)、およびミトコンドリアリン酸輸送蛋白質(R49231)からなる群より選択される少なくとも1つの下流遺伝子の発現を検出する段階であって、該下流遺伝子は野生型WT1遺伝子にアップレギュレーションまたはダウンレギュレーションされる検出段階；ならびに該標的サンプル中の該発現と該参照サンプル中の該発現を比較する段階であって、2つのサンプル間で発現に差があれば、該WT1配列変化がWT1の生物学的機能に影響を与えることが示唆される段階。

【請求項61】下記の段階を含む、WT1配列変化の細胞内機能アッセイ法：WT1配列変化を持つ標的細胞の標的サンプル、および野生型WT1遺伝子を持ち、その他の点では標的細胞に実質的に類似した参照細胞の参照サン

プルにおいて、ナチュラルキラー細胞プロテイン4前駆体(M59807)、葉酸結合蛋白質(M25317)、チューブリンのHALPHA44遺伝子(X06956)、熱ショック蛋白質HSP70B(X51758)、90K産物(H17969)、熱ショック70kd蛋白質1(T66307)、アンフィレグリン(M30704)、プロコラーゲン1(T51558)、遊走性プラスミノーゲン活性化因子阻害剤1(M14083)、放射線照射ケラチノサイト由来のシステインプロテアーゼ阻害剤(X05978)、脳性ナトリウム利尿蛋白質(BNP)(M31766)、白血病ウイルス受容体1(GLVR1)(L20859)、1型細胞骨格17ケラチン、チューブリン5鎖(H45051)、プリンヌクレオシドホスホリラーゼ(T47964)、Gem GTPアーゼ(U10550)、アドレノメダリン(D14874)、GPIアンカー型ウロキナーゼプラスミノーゲン活性化因子表面受容体(R38636)、組織型プラスミノーゲン活性化因子(X13097)、グラビン(M96322)、jun-B(X51345)、延長因子1-2(X79490)、ホメオティック遺伝子調節因子(R16977)、クローン9112(X57348)、インターフェロン処理誘導可能mRNA(M26683)、膜貫通受容体蛋白質(Z17227)、酸性FGF(X65779)、軟骨結合蛋白質-1(X17405)、IL-11(X58377)、およびミトコンドリアリン酸輸送蛋白質(R49231)からなる群より選択される少なくとも1つの下流遺伝子の発現を検出する段階であって、該下流遺伝子は野生型WT1遺伝子にアップレギュレーションまたはダウンレギュレーションされる検出段階；ならびに該標的サンプル中の該発現と該参照サンプル中の該発現を比較する段階であって、2つのサンプル間で発現に差があれば、該WT1配列変化がWT1の生物学的機能に影響を与えることが示唆される段階。

【請求項62】下記の段階を含む、コンピュータを用いた標的WT1遺伝子中の変異の検出方法：野生型WT1遺伝子を含む野生型サンプル中の、少なくとも1つのWT1下流遺伝子の野生型発現データをコンピュータに入力する段階であって、下流遺伝子が、ナチュラルキラー細胞プロテイン4前駆体(M59807)、葉酸結合蛋白質(M25317)、チューブリンのHALPHA44遺伝子(X06956)、熱ショック蛋白質HSP70B(X51758)、90K産物(H17969)、熱ショック70kd蛋白質1(T66307)、アンフィレグリン(M30704)、プロコラーゲン1(T51558)、遊走性プラスミノーゲン活性化因子阻害剤1(M14083)、放射線照射ケラチノサイト由来のシステインプロテアーゼ阻害剤(X05978)、脳性ナトリウム利尿蛋白質(BNP)(M31766)、白血病ウイルス受容体1(GLVR1)(L20859)、1型細胞骨格17ケラチン、チューブリン5鎖(H45051)、プリンヌクレオシドホスホリラーゼ(T47964)、Gem GTPアーゼ(U10550)、アドレノメダリン(D14874)、GPIアンカー型ウロキナーゼプラスミノーゲン活性化因子表面受容体(R38636)、組織型プラスミノーゲン活性化因子(X13097)、グラビン(M96322)、jun-B(X51345)、延長因子1-2(X79490)、ホメオティック遺伝子調節因子(R16977)、クローン9112(X57348)、インターフェロン処理誘導可能mRNA(M26683)、膜貫通受容体蛋白質(Z17227)、酸性FGF(X65779)、軟骨結合蛋白質-1(X17405)、IL-11(X58377)、ミトコンドリアリン酸輸送蛋白質(R49231)、およびカベオリンからなる群より選択される検出段階；ならびに標的細胞および該参照細胞における少なくとも1つの下流遺伝子の発現を比較する段階であって、標的細胞と参照細胞とで該発現に差があれば、標的細胞にWT1機能変異が存在することが示唆される段階。

【請求項63】下記の段階を含む、WT1配列変化の細胞内機能アッセイ法：WT1配列変化を持つ標的細胞の標的サンプル、および野生型WT1遺伝子を持ち、その他の点では標的細胞に実質的に類似した参照細胞の参照サン

ーン9112 (X57348)、インターフェロン 処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、カベオリン、酸性FGF (X65779)、軟骨結合蛋白質-1 (X17405)、IL-11 (X58377)、およびミトコンドリアリン酸輸送蛋白質 (R49231) からなる群より選択される入力段階；標的WT1遺伝子を含む標的サンプル中の複数の該下流遺伝子の標的発現データをコンピュータに入力する段階；ならびに標的および野生型の発現データを比較して、差を測定する段階であって、差があれば、該標的WT1遺伝子中に変異が存在することが示唆される段階。

【請求項63】 下記の段階を含む、試験細胞の新生物形成性の診断方法：ナチュラルキラー細胞プロテイン4前駆体 (M59807)、熱ショック蛋白質HSP70B (X51758)、90K産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、プロコラーゲン 1 (T51558)、放射線照射ケラチノサイト由来のシステインプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質 (BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン (R71870)、プリンヌクレオシドホスホリラーゼ (T47964)、アドレノメダリン (D14874)、グラビン (M96322)、jun-B (X51345)、延長因子1 -2 (X79490)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57348)、インターフェロン 処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、酸性FGF (X65779)、軟骨結合蛋白質-1 (X17405)、IL-11 (X58377)、およびミトコンドリアリン酸輸送蛋白質 (R49231) からなる群より選択される、WT1に活性化または抑制される少なくとも1つの遺伝子の試験細胞および対照細胞における発現レベルを測定する段階；試験細胞で測定された発現レベルを対照細胞で測定された発現レベルと比較する段階；ならびに(1) WT1に活性化される遺伝子である少なくとも1つの遺伝子の発現レベルが、対照細胞よりも試験細胞のほうが低い、または(2) WT1に抑制される遺伝子の少なくとも1つの発現レベルが、対照細胞よりも試験細胞のほうが高い場合に、試験細胞が新生物形成性であると同定する段階。

【請求項64】 下記の段階を含む、試験細胞の新生物形成性の診断方法：EWS-WT1融合蛋白質に活性化または抑制される少なくとも1つの遺伝子の試験細胞および対照細胞における発現レベルを測定する段階；試験細胞の測定された発現レベルを対照細胞の測定された発現レベルと比較する段階；ならびに(1) EWS-WT1に活性化される遺伝子の試験細胞における発現が、正常細胞における発現よりも低い、または(2) EWS-WT1に抑制される遺伝子の試験細胞における発現が、正常細胞における発現よりも高い場合に、試験細胞が新生物形成性であると同定する段階。

【請求項65】 下記の段階を含む、試験細胞の新生物形成性の診断方法：WT1に活性化または抑制される少なくとも3つの遺伝子の試験細胞および対照細胞における

発現レベルを測定する段階；試験細胞にて測定された発現レベルを対照細胞にて測定される発現レベルと比較する段階；ならびに(1) WT1に活性化される遺伝子の試験細胞における発現が、正常細胞における発現よりも低い、または(2) WT1に抑制される遺伝子の試験細胞における発現が、正常細胞における発現よりも高い場合に、試験細胞が新生物形成性であると同定する段階。

【発明の詳細な説明】

【0001】

10 【発明の属する技術分野】本発明は、WT1遺伝子に関する癌の診断、薬物スクリーニング、および変異の機能分析方法に関する。

【0002】

【従来の技術】多くの生物学的機能は、転写の制御および/または翻訳の制御（例、開始の制御、RNA前駆体の提供、RNAプロセッシングなど）によって種々の遺伝子の発現を変化させることにより達成される。例えば、細胞周期の調節、細胞の分化、および細胞死のような基本的な生物学的過程は、しばしば遺伝子群の発現レベルの変化により特徴づけられる。

20 【0003】遺伝子発現は、病因にも関連している。例えば、機能的な癌抑制遺伝子の十分な発現不足および/または癌遺伝子/癌原遺伝子の過剰発現は、腫瘍形成を引き起こす可能性がある（Marshall, Cell, 64:313-326 (1991); Weinberg, Science, 254:1138-1146 (1991)、参照として本明細書に組み入れられる）。したがって、特定の遺伝子（例、癌遺伝子または癌抑制遺伝子）の発現レベルの変化は、種々の疾患の存在および進行を知る指標となる。当技術分野では、診断および新しい治療法の探索に利用するために、特定の癌抑制因子および癌遺伝子によってどの遺伝子が影響されるかを知る必要がある。

【0004】

【発明が解決しようとする課題】本発明の目的は、WT1遺伝子の機能変異を同定する方法を提供することである。

【0005】EWS-WT1遺伝子融合を同定する方法を提供することも、本発明のもう一つの目的である。

【0006】WT1変異の生物学的効果を試験する細胞内アッセイ法を提供することも本発明のさらに別の目的である。

【0007】コンピュータを用いてWT1変異またはEWS-WT1融合を検出する方法を提供することも、本発明のさらに別の目的である。

【0008】新生物形成性を診断する方法を提供することも、本発明のさらなる目的である。

【0009】新生物形成性の治療に有用な薬物を同定する方法を提供することも、本発明のもう一つの目的である。

50 【0010】

【課題を解決するための手段】本発明のこれらおよびその他の目的は、以下の態様の1つまたは複数によって、提供される。1つの態様では、標的細胞中でWT1遺伝子の機能変異を検出する方法が提供される。(a)標的細胞、および(b)野生型WT1遺伝子を持つ参照細胞のサンプル中で、WT1の少なくとも3つの下流の遺伝子の発現を検出す。参照細胞は、その他の点で標的細胞と実質的に類似した細胞である。この下流遺伝子は、野生型WT1遺伝子によって、アップレギュレーションまたはダウンレギュレーションされる。標的細胞と参照細胞中の下流遺伝子の発現を比較し、標的細胞と参照細胞とで発現の差があれば、標的細胞中にWT1機能変異があることが示唆される。

【0011】さらに別の態様では、標的細胞中のEWS-WT1遺伝子融合を検出する方法が提供される。EWS-WT1融合の1つまたは複数の下流遺伝子の発現を、(a)標的細胞、および(b)野生型のEWSおよびWT1遺伝子を持つ参照細胞のサンプル中で検出する。参照細胞は、その他の点では、標的細胞と実質的に類似している。下流の遺伝子は、野生型EWSおよびWT1遺伝子によって、アップレギュレーションまたはダウンレギュレーションされる。標的細胞と参照細胞中での下流遺伝子の発現を比較する。標的細胞と参照細胞とで発現の差があれば、標的細胞中でEWS-WT1融合が存在することが示唆される。

【0012】本発明の別の局面は、WT1配列変化の細胞内機能アッセイ法である。WT1配列変化を持つ標的細胞由来の標的サンプル、および野生型WT1遺伝子を持つ参照細胞由来の参照サンプル中で、少なくとも3つの下流遺伝子の発現を検出する。参照細胞は、その他の点では、標的細胞と実質的に類似している。下流の遺伝子は、野生型WT1遺伝子によって、アップレギュレーションまたはダウンレギュレーションされる。標的サンプル中の発現と参照サンプル中の発現を比較する。2つのサンプルで発現の差があれば、WT1配列の変化が、WT1の生物学的機能に影響を与えたことが示唆される。

【0013】本発明の別の局面に従い、コンピュータを使って標的WT1遺伝子中の変異を検出する方法が提供される。野生型WT1遺伝子を含む野生型サンプル中の少なくとも3つの下流遺伝子の野生型発現データを、コンピュータに入力する。下流遺伝子は、野生型WT1遺伝子によって転写調節されている。標的WT1遺伝子を含む標的サンプル中の複数の下流遺伝子の標的発現データも、コンピュータに入力する。標的および野生型の発現データを比較して、差異を決定する。差があれば、標的WT1遺伝子に変異があることが示唆される。

【0014】また別の態様では、コンピュータを用いたEWSとWT1遺伝子とを融合する転座の検出方法が提供される。野生型EWSおよびWT1遺伝子を含む野生型サンプル中の複数の下流遺伝子の野生型発現データをコンピュータに入力する。下流遺伝子は、EWS-WT1融合蛋白質によっ

て転写調節されている。EWS-WT1融合蛋白質の存在を試験する標的サンプル中の複数の下流遺伝子の標的発現データも、コンピュータに入力する。標的と野生型の発現データを比較して差異を決定する。差異があれば、EWSとWT1遺伝子とを融合する転座があることが示唆される。

【0015】本発明の別の態様では、試験細胞の新生物形成性を診断する方法が提供される。試験細胞の転写指標を、核酸プローブのセットにハイブリダイズさせる。

10 転写指標は、mRNA、cDNA、およびcRNAからなる群より選択される。核酸プローブのセットは、それぞれが、WT1によって活性化または抑制され、ナチュラルキラー細胞プロテイン4前駆体 (M59807)、熱ショック蛋白質HSP70B (X51758)、90 K 産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、プロコラーゲン 1 (T51558)、放射線照射ケラチノサイト由来のシステインプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質 (brain natriuretic protein: BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン (R71870)、プリンヌクレオシドホスホリラーゼ (T47964)、アドレノメダリン (D14874)、グラビン (M96322)、jun-B (X51345)、延長因子1 -2 (X79490)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57348)、インターフェロン 処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、ミトコンドリアリン酸輸送タンパク質 (R49231) からなる群より選択される遺伝子の一部である、複数の核酸分子を含む。各核酸プローブのセットにハイブリダイズする転写指標の量を検出する。試験細胞は、(1) WT1に活性化される遺伝子であるプローブに30 対して、試験細胞の転写指標がハイブリダイズする量が、正常細胞由来の転写指標がハイブリダイズする量よりも少ない、または、(2) WT1に抑制される遺伝子であるプローブに対して、試験細胞の転写指標がハイブリダイズする量が、正常細胞由来の転写指標がハイブリダイズする量よりも多い場合に、試験細胞は新生物形成性であると同定される。

【0016】本発明の別の局面は、抗癌剤の同定方法によって提供される。試験化合物を、ヒト細胞に接触させる。細胞の、WT1によってアップレギュレーションまたはダウンレギュレーションされる少なくとも1つの遺伝子の発現に対する、試験化合物の効果を決定する。調節される遺伝子は、ナチュラルキラー細胞プロテイン4前駆体 (M59807)、熱ショック蛋白質HSP70B (X51758)、90 K 産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、プロコラーゲン 1 (T51558)、放射線照射ケラチノサイト由来のシステインプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質 (BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン、プリンヌクレオシドホスホリラーゼ (T47964)、アドレノメダリン (D14874)、グラビン (M96322)、jun-B (X5150)

345)、延長因子1 -2 (X79490)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57348)、インターフェロン 処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、ミトコンドリアリン酸輸送タンパク質 (R49231)からなる群から選択される。試験化合物は、ヒト細胞中で、WT1にアップレギュレーションされる少なくとも1つの遺伝子の発現を上昇させるか、WT1にダウンレギュレーションされる遺伝子の発現を低下させれば、抗癌剤の可能性があると同定される。

【0017】本発明のさらに別の局面は、試験細胞の新生物形成性の診断方法である。試験細胞の転写指標を、核酸プローブのセットにハイブリダイズさせる。転写指標は、mRNA、cDNA、およびcRNAからなる群より選択される。核酸プローブのセットは、それぞれが、EWS-WT1融合蛋白質によって活性化または抑制される遺伝子の一部である、複数の核酸分子を含む。各核酸プローブのセットにハイブリダイズする転写指標の量を検出する。(1) EWS-WT1に活性化される遺伝子であるプローブに対して、試験細胞の転写指標がハイブリダイズする量が、正常細胞由来の転写指標がハイブリダイズする量よりも少ない、または、(2) EWS-WT1に抑制される遺伝子であるプローブに対して、試験細胞の転写指標がハイブリダイズする量が、正常細胞由来の転写指標がハイブリダイズする量よりも多い場合に、試験細胞は新生物形成性であると同定される。

【0018】本発明のさらに別の態様に従い、抗癌剤を同定する方法が提供される。試験化合物を、ヒト細胞に接触させる。EWS-WT1にアップレギュレーションされる遺伝子またはEWS-WT1にダウンレギュレーションされる遺伝子の発現を決定する。試験化合物は、ヒト細胞中で、EWS-WT1にアップレギュレーションされる遺伝子の発現を低下させるか、EWS-WT1にダウンレギュレーションされる遺伝子の発現を上昇させれば、抗癌剤の可能性があると同定される。

【0019】標的細胞中でWT1遺伝子の機能変異を検出する方法も提供される。(a) 標的細胞、および(b) 野生型WT1遺伝子を持つ参照細胞のサンプル中で、少なくとも1つのWT1の下流遺伝子の発現を検出する。参照細胞は、他の点では標的細胞と実質的に同様である。下流遺伝子は、ナチュラルキラー細胞プロテイン4前駆体 (M59807)、葉酸結合蛋白質 (M25317)、チューブリンのHALPHA44遺伝子 (X06956)、熱ショック蛋白質HSP70B (X51758)、90K産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、アンフィレグリン (M30704)、プロコラーゲン 1 (T51558)、遊走性プラスミノーゲン活性化因子阻害剤1 (M14083)、放射線照射ケラチノサイト由来のシステインプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質 (BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン、チューブリン 5鎖 (H45051)、プリンヌクレオシドホスホリラーゼ (T47964)、Gem GTPアーゼ (U10550)、アドレノメダリン (D14874)、GPIアンカー型ウロキナーゼプラスミノーゲン活性化因子表面受容体 (R38636)、組織型プラスミノーゲン活性化因子 (X13097)、グラビン (M96322)、jun-B (X51345)、延長因子1 -2 (X79490)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57348)、インターフェロン 処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、ミトコンドリアリン酸輸送タンパク質 (R49231)からなる群より選択される少なくとも1つの下流遺伝子の発現を検出する。参照細胞は、他の点では標的細胞と実質的に同様である。下流遺伝子は、野生型WT1遺伝子によってアップレギュレーションまたはダウンレギュレーションされる。標的細胞における発現と参照細胞における発現とを比較する。2つのサンプルで発現に差異があれば、WT1配列の変化が、WT1の生物学的機能に影響を与えることが示唆される。

【0020】WT1配列変化の細胞内機能アッセイ法も本発明によって提供される。WT1配列変化を持つ標的細胞由来の標的サンプルと、野生型WT1遺伝子を持つ参照細胞由来の参照サンプル中で、ナチュラルキラー細胞プロテイン4前駆体 (M59807)、葉酸結合蛋白質 (M25317)、チューブリンのHALPHA44遺伝子 (X06956)、熱ショック蛋白質HSP70B (X51758)、90K産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、アンフィレグリン (M30704)、プロコラーゲン 1 (T51558)、遊走性プラスミノーゲン活性化因子阻害剤1 (M14083)、放射線照射ケラチノサイト由来のシステインプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質 (BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン、チューブリン 5鎖 (H45051)、プリンヌクレオシドホスホリラーゼ (T47964)、Gem GTPアーゼ (U10550)、アドレノメダリン (D14874)、GPIアンカー型ウロキナーゼプラスミノーゲン活性化因子表面受容体 (R38636)、組織型プラスミノーゲン活性化因子 (X13097)、グラビン (M96322)、jun-B (X51345)、延長因子1 -2 (X79490)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57348)、インターフェロン 処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、カベオリンおよびミトコンドリアリン酸輸送タンパク質 (R49231)からなる群より選択される少なくとも1つの下流遺伝子の発現を検出する。参照細胞は、他の点では標的細胞と実質的に同様である。下流遺伝子は、野生型WT1遺伝子によってアップレギュレーションまたはダウンレギュレーションされる。標的細胞における発現と参照細胞における発現とを比較する。2つのサンプルで発現に差異があれば、WT1配列の変化が、WT1の生物学的機能に影響を与えることが示唆される。

【0021】本発明の別の態様では、コンピュータを使った標的WT1遺伝子中の変異の検出方法が提供される。野生型WT1遺伝子を含む野生型サンプル中の少なくとも1つのWT1の下流遺伝子の野生型の発現データをコンピュータに入力する。下流遺伝子は、ナチュラルキラー細胞プロテイン4前駆体 (M59807)、葉酸結合蛋白質 (M25317)

7)、チューブリンのHALPHA44遺伝子 (X06956)、熱ショック蛋白質HSP70B (X51758)、90K産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、アンフィレグリン (M30704)、プロコラーゲン 1 (T51558)、遊走性プラスミノーゲン活性化因子阻害剤1 (M14083)、放射線照射ケラチノサイト由来のシステインプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質 (BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン、チューブリン 5鎖 (H45051)、プリンヌクレオシドホスホリラーゼ (T47964)、GMP GTPアーゼ (U10550)、アドレノメダリン (D14874)、GPIアンカー型ウロキナーゼプラスミノーゲン活性化因子表面受容体 (R38636)、組織型プラスミノーゲン活性化因子 (X13097)、グラビン (M96322)、jun-B (X51345)、延長因子1 -2 (X79490)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57348)、インターフェロン 処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、カベオリン、およびミトコンドリアリン酸輸送タンパク質 (R49231)からなる群より選択される。標的WT1遺伝子を含む標的サンプル中の複数の下流遺伝子の標的発現データも、コンピュータに入力する。標的と野生型の発現データを比較して、差異を決定する。差異があれば、標的WT1遺伝子に変異があることが示唆される。

【0022】本発明のさらに別の局面に従い、試験細胞の新生物形成性の診断方法が提供される。試験細胞と、対照細胞中で、WT1によって活性化または抑制される少なくとも1つの遺伝子の発現レベルを決定する。遺伝子は、ナチュラルキラー細胞プロテイン4前駆体 (M59807)、熱ショック蛋白質HSP70B (X51758)、90K産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、プロコラーゲン 1 (T51558)、放射線照射ケラチノサイト由来のシステインプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質 (BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン (R71870)、プリンヌクレオシドホスホリラーゼ (T47964)、アドレノメダリン (D14874)、グラビン (M96322)、jun-B (X51345)、延長因子1 -2 (X79490)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57348)、インターフェロン 処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、ミトコンドリアリン酸輸送タンパク質 (R49231)からなる群より選択される。標的細胞で決定された発現レベルを、対照細胞で決定された発現レベルと比較する。(1) WT1に活性化される少なくとも1つの遺伝子の発現レベルが、対照細胞よりも試験細胞の方が低い、または、(2) WT1に抑制される少なくとも1つの遺伝子の発現レベルが、対照細胞よりも試験細胞の方が高い場合に、試験細胞は新生物形成性であると同定される。

【0023】本発明の別の局面では、試験細胞の新生物形成性の診断方法が提供される。試験細胞と、対照細胞中で、EWS-WT1融合蛋白質によって活性化または抑制さ

れる少なくとも1つの遺伝子の発現レベルを決定する。試験細胞で決定された発現レベルを、対照細胞で決定された発現レベルと比較する。(1) EWS-WT1に活性化される遺伝子の発現が、正常細胞よりも試験細胞の方が少ない、または、(2) EWS-WT1に抑制される遺伝子の発現が、正常細胞よりも試験細胞の方が高い場合に、試験細胞は新生物形成性であると同定される。

【0024】試験細胞の新生物形成性を診断するための別の方法も提供される。試験細胞と対照細胞で、WT1に活性化または抑制される少なくとも3つの遺伝子の発現レベルを決定する。試験細胞で決定された発現レベルを、対照細胞で決定された発現レベルと比較する。(1) WT1に活性化される遺伝子の発現が、正常細胞よりも試験細胞の方が少ない、または、(2) WT1に抑制される遺伝子の発現が、正常細胞よりも試験細胞の方が高い場合に、試験細胞は新生物形成性であると同定される。

【0025】本発明に係る検出方法においては、(1) 下記の段階を含む、標的細胞においてWT1遺伝子の機能変異を検出する方法を特徴とする：

20 (a) 標的細胞、および(b) 野生型WT1遺伝子を持ち、その他の点では標的細胞に実質的に類似した参照細胞のサンプル中で、少なくとも3つのWT1の下流遺伝子の発現を検出する段階であって、下流遺伝子は野生型WT1遺伝子にアップレギュレーションまたはダウンレギュレーションされる検出段階；および標的細胞および参照細胞における下流遺伝子の発現を比較する段階であって、標的細胞と参照細胞とで発現に差があれば、標的細胞にWT1機能変異が存在することが示唆される段階。

【0026】また、本発明に係る方法においては、(2) 下流遺伝子が、野生型WT1遺伝子によって転写調節されており、該下流遺伝子の発現が、参照細胞および標的細胞中の該下流遺伝子の転写産物の量を測定することによって検出される、上記(1)記載の方法であることを特徴とする。

【0027】また、本発明に係る方法においては、(3) 転写産物の量が、高密度核酸アレイを用いて測定される上記(1)記載の方法であることを特徴とする。

【0028】また、本発明に係る方法においては、(4) 下流遺伝子が、ナチュラルキラー細胞プロテイン4前駆体 (M59807)、葉酸結合蛋白質 (M25317)、チューブリンのHALPHA44遺伝子 (X06956)、熱ショック蛋白質HSP70B (X51758)、90K産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、アンフィレグリン (M30704)、プロコラーゲン 1 (T51558)、遊走性プラスミノーゲン活性化因子阻害剤1 (M14083)、放射線照射ケラチノサイト由来のシステインプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質(brain natriuretic protein) (BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン (R71870)、チューブリン 5鎖 (H45051)、プリンヌクレオシドホスホリラーゼ

(T47964)、Gem GTPアーゼ (U10550)、アドレノメダリン (D14874)、GPIアンカー型ウロキナーゼプラスミノーゲン活性化因子表面受容体 (R38636)、組織型プラスミノーゲン活性化因子 (X13097)、グラビン (M96322)、jun-B (X51345)、延長因子1 -2 (X79490)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57348)、インターフェロン 处理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、酸性FGF (X65779)、軟骨結合蛋白質-1 (X17405)、IL-11 (X58377)、およびミトコンドリアリン酸輸送蛋白質 (R49231) からなる群より選択されるWT1にアップレギュレーションされる遺伝子を、少なくとも1つ含む、上記(1)記載の方法であることを特徴とする。

【0029】また、本発明に係る方法においては、(5)下流遺伝子が、カベオリンおよび延長因子1 -2 からなる群より選択される、WT1にダウンレギュレーションされる遺伝子を少なくとも1つ含む、上記(1)記載の方法であることを特徴とする。

【0030】また、本発明に係る方法においては、(6)下流遺伝子がp21またはEGFRをさらに含む、上記(4)記載の方法であることを特徴とする。

【0031】また、本発明に係る方法においては、(7)WT1にアップレギュレーションされる遺伝子の発現が、参照細胞と比較して、標的細胞で少なくとも2分の1以下であるか、WT1にダウンレギュレーションされる遺伝子の発現が、参照細胞と比較して標的細胞で少なくとも2倍以上である場合に、標的細胞中のWT1遺伝子が機能損失変異であることを示す段階をさらに含む上記(1)記載の方法であることを特徴とする。

【0032】また、本発明に係る方法においては、(8)下記の段階を含む、標的細胞中でEWS-WT1融合体を検出する方法であることを特徴とする：

(a) 標的細胞、および(b) 野生型EWSおよびWT1遺伝子を持ち、その他の点では標的細胞に実質的に類似した参照細胞のサンプル中で、1つまたは複数のEWS-WT1融合体の下流遺伝子の発現を検出する段階であって、該下流遺伝子は野生型EWSおよびWT1遺伝子にアップレギュレーションまたはダウンレギュレーションされる検出段階；および標的細胞および参照細胞における該下流遺伝子の発現を比較する段階であって、標的細胞と参照細胞とで発現に差があれば、標的細胞中にEWS-WT1融合体が存在することが示唆される段階。

【0033】また、本発明に係る方法においては、(9)下流遺伝子が、仮性狂犬病ウイルス核抗原 (R60906およびT79475)、インターロイキン-2受容体鎖 (M26062)、G1/S-特異的サイクリンD1 (H20529)、ニューロン特異的 -2エノラーゼ (M22349)、転座t(3;5)(q2 5.1;p34)により得られる融合遺伝子NPM-MLF1 (L49054)、コラーゲン 3(IV)鎖 (H62466)、T-リンパ球特異的蛋白質チロシンキナーゼp56lck (lck) (U23852)、イ

ンスリン様成長因子結合蛋白質-4 (U20982)、カテプシンB (L16510)、シスタチンC (T51534)、二重特異的マイトジエン活性化蛋白質キナーゼ2 (R42291)、神経グリア細胞接着分子 (T53118)、PIM-1癌原遺伝子セリン/スレオニン-蛋白質キナーゼ (H10925)、カルボキシルエステラーゼ前駆体 (R54359)、蛋白質-チロシンホスファターゼPAC-1 (L11329)、およびインターフェロン誘導蛋白質6-16 (X02492) からなる群より選択される、EWS-WT1にアップレギュレーションされる遺伝子を少なくとも一つ含む、上記(8)記載の方法であることを特徴とする。

【0034】また、本発明に係る方法においては、(10)下記の段階を含む、WT1配列変化の細胞内機能アップセイ法であることを特徴とする：WT1配列変化を有する標的細胞の標的サンプル、および野生型WT1遺伝子を持ち、その他の点では標的細胞に実質的に類似した参照細胞の参照サンプルにおいて、少なくとも3つの下流遺伝子の発現を検出する段階であって、該下流遺伝子は野生型WT1遺伝子にアップレギュレーションまたはダウンレギュレーションされる検出段階；および該標的サンプル中の該発現と該参照サンプル中の該発現とを比較する段階であって、2つのサンプル間で発現に差があれば、該WT1配列変化がWT1の生物学的機能に影響を与えることが示唆される段階。

【0035】また、本発明に係る方法においては、(11)下流遺伝子が、野生型WT1遺伝子によって転写調節されており、該下流遺伝子の発現が、参照細胞および標的細胞中の該下流遺伝子の転写産物の量を測定することによって検出される、上記(10)記載の方法であることを特徴とする。

【0036】また、本発明に係る方法においては、(12)転写産物の量が、高密度核酸アレイを用いて測定される、上記(11)記載の方法であることを特徴とする。

【0037】また、本発明に係る方法においては、(13)下流遺伝子が、ナチュラルキラー細胞プロテイン4前駆体 (M59807)、葉酸結合蛋白質 (M25317)、チューブリンのHALPHA44遺伝子 (X06956)、熱ショック蛋白質HSP70B (X51758)、90K産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、アンフィレグリン (M30704)、プロコラーゲン 1 (T51558)、遊走性プラスミノーゲン活性化因子阻害剤1 (M14083)、放射線照射ケラチノサイト由来のシスティンプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質 (BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン、チューブリン 5鎖 (H45051)、プリンヌクレオシドホスホリラーゼ (T47964)、Gem GTPアーゼ (U10550)、アドレノメダリン (D14874)、GPIアンカー型ウロキナーゼプラスミノーゲン活性化因子表面受容体 (R38636)、組織型プラスミノーゲン活性化因子 (X13097)、グラビン (M96322)、jun-B (X51345)、ホメオティック遺伝子調節因子

(R16977)、クローン9112 (X57348)、インターフェロン処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、酸性FGF (X65779)、軟骨結合蛋白質-1 (X17405)、IL-11 (X58377)、およびミトコンドリアリン酸輸送蛋白質 (R49231) からなる群より選択されるWT1にアップレギュレーションされる遺伝子を少なくとも1つ含む、上記(10)記載の方法であることを特徴とする。

【0038】また、本発明に係る方法においては、(14)下流遺伝子が、カベオリンおよび延長因子1-2 (X79490) からなる群より選択される、WT1にダウントレギュレーションされる遺伝子を少なくとも1つ含む、上記(10)記載の方法であることを特徴とする。

【0039】また本発明に係る方法においては、(15)下流遺伝子がp21またはEGFRをさらに含む、上記(13)記載の方法であることを特徴とする。

【0040】また、本発明に係る方法においては、(16)WT1にアップレギュレーションされる遺伝子の発現が、参照細胞と比較して、標的細胞で少なくとも2分の1以下であるか、WT1にダウントレギュレーションされる遺伝子の発現が、参照細胞と比較して標的細胞で少なくとも2倍以上である場合に、WT1配列変化が野生型の機能損失変異であることを示す段階をさらに含む、上記(10)記載の方法であることを特徴とする。

【0041】また、本発明に係る方法においては、(17)下記の段階を含む、コンピュータを用いた標的WT1遺伝子中の変異の検出方法であることを特徴とする：野生型WT1遺伝子を含む野生型サンプル中の、該野生型WT1遺伝子によって転写調節を受ける少なくとも3つの下流遺伝子の野生型発現データをコンピュータに入力する段階；標的WT1遺伝子を含む標的サンプル中の複数の該下流遺伝子の標的発現データをコンピュータに入力する段階；ならびに標的および野生型の発現データを比較して、差を測定する段階であって、差があれば、該標的WT1遺伝子中に変異が存在することが示唆される段階。

【0042】また、本発明に係る方法においては、(18)下流遺伝子が、ナチュラルキラー細胞プロテイン4前駆体 (M59807)、葉酸結合蛋白質 (M25317)、チューブリンのHALPHA44遺伝子 (X06956)、熱ショック蛋白質HSP70B (X51758)、90K産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、アンフィレグリン (M30704)、プロコラーゲン1 (T51558)、遊走性プラスミノーゲン活性化因子阻害剤1 (M14083)、放射線照射ケラチノサイト由来のシスティンプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質 (BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン、チューブリン5鎖 (H45051)、プリンヌクレオシドホスホリラーゼ (T47964)、Gem GTPアーゼ (U10550)、アドレノメタリン (D14874)、GPIアンカー型ウロキナーゼプラスミノーゲン活性化因子表面受容体 (R38636)、組織型プラスミノーゲン活性化因子 (X13097)、グラビン (M963250)

2)、jun-B (X51345)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57348)、インターフェロン処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、酸性FGF (X65779)、軟骨結合蛋白質-1 (X17405)、IL-11 (X58377)、およびミトコンドリアリン酸輸送蛋白質H (R49231) からなる群より選択されるWT1にアップレギュレーションされる遺伝子を少なくとも1つ含む、上記(17)記載の方法であることを特徴とする。

【0043】また、本発明に係る方法においては、(19)下流遺伝子が、カベオリンおよび延長因子1-2 (X79490) からなる群より選択される、WT1にダウントレギュレーションされる遺伝子を少なくとも1つ含む、上記(17)記載の方法であることを特徴とする。

【0044】また、本発明に係る方法においては、(20)下流遺伝子がp21またはEGFRをさらに含む、上記(18)記載の方法であることを特徴とする。

【0045】また、本発明に係る方法においては、(21)WT1にアップレギュレーションされる遺伝子の発現が、参照細胞と比較して、標的細胞で少なくとも2分の1以下であるか、WT1にダウントレギュレーションされる遺伝子の発現が、参照細胞と比較して標的細胞で少なくとも2倍以上である場合に、標的WT1遺伝子中に変異があることを示す段階をさらに含む、上記(17)記載の方法であることを特徴とする。

【0046】また、本発明に係る方法においては、(22)下記の段階を含む、コンピュータを用いたEWSとWT1遺伝子とを融合する転座の検出方法を特徴とする：野生型EWSおよびWT1遺伝子を含む野生型サンプル中の、EWS-WT1融合蛋白質によって転写調節を受ける複数の下流遺伝子の野生型発現データをコンピュータに入力する段階；EWS-WT1融合蛋白質の存在を試験する標的サンプル中の複数の該下流遺伝子の標的発現データをコンピュータに入力する段階；ならびに標的および野生型の発現データを比較して、差を測定する段階であって、差があればEWSとWT1遺伝子を融合する転座があることが示唆される段階。

【0047】また、本発明に係る方法においては、(23)下流遺伝子が、仮性狂犬病ウイルス核抗原 (R60906) およびT79475)、インターロイキン-2受容体鎖 (M26040)、G1/S-特異的サイクリンD1 (H20529)、ニューヨン特異的-2エノラーゼ (M22349)、転座t(3;5)(q25.1;p34)により得られる融合遺伝子NPM-MLF1 (L49054)、コラーゲン3(IV)鎖 (H62466)、T-リンパ球特異的蛋白質チロシンキナーゼp56lck (lck) (U23852)、インスリン様成長因子結合蛋白質-4 (U20982)、カテプシンB、シスタチンC (T51534)、二重特異的マイトジエン活性化蛋白質キナーゼ2 (R42291)、神経グリア細胞接着分子 (T53118)、PIM-1癌原遺伝子セリン/スレオニン-蛋白質キナーゼ (H10925)、カルボキシリエステラーゼ前駆体 (R54359)、蛋白質-チロシンホスファ

ターゲットPAC-1 (L11329)、およびインターフェロン誘導蛋白質6-16 (X02492) からなる群より選択される、EWS-WT1にアップレギュレーションされる遺伝子を少なくとも一つ含む、上記(22)記載の方法であることを特徴とする。

【0048】また、本発明に係る方法においては、(24)下記の段階を含む、試験細胞の新生物形成性の診断方法であることを特徴とする：mRNA、cDNAおよびcRNAからなる群より選択される試験細胞の転写指標を、核酸プローブのセットにハイブリダイズさせる段階であって、核酸プローブのセットが、ナチュラルキラー細胞プロテイン4前駆体 (M59807)、熱ショック蛋白質HSP70B (X5158)、90K産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、プロコラーゲン 1 (T51558)、放射線照射ケラチノサイト由来のシスティンプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質 (BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン (R71870)、プリンヌクレオシドホスホリラーゼ (T47964)、アドレノメダリン (D14874)、グラビン (M96322)、jun-B (X51345)、延長因子1 -2 (X79490)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57348)、インターフェロン 処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、酸性FGF (X65779)、軟骨結合蛋白質-1 (X17405)、IL-11 (X58377)、およびミトコンドリアリン酸輸送蛋白質 (R49231) からなる群より選択される、各々がWT1に活性化または抑制される遺伝子の一部である複数の核酸分子を含むハイブリダイズ段階；各々の該核酸プローブセットにハイブリダイズする転写指標の量を検出する段階；ならびに(1) WT1に活性化される遺伝子であるプローブへの試験細胞の転写指標のハイブリダイゼーションが、正常細胞の転写指標のハイブリダイゼーションよりも低レベルである、または(2) WT1に抑制される遺伝子であるプローブへの試験細胞の転写指標のハイブリダイゼーションが、正常細胞の転写指標のハイブリダイゼーションよりも高レベルである場合に、試験細胞が新生物形成性であると同定する段階。

【0049】また、本発明に係る方法においては、(25)試験細胞が腎臓細胞である上記(24)記載の方法であることを特徴とする。

【0050】また、本発明に係る方法においては、(26)プローブの少なくとも4つが、WT1に活性化またはWT1に抑制される遺伝子の一部を含んでいる、上記(24)記載の方法であることを特徴とする。

【0051】また、本発明に係る方法においては、(27)プローブの少なくとも10個が、WT1に活性化またはWT1に抑制される遺伝子の一部を含んでいる、上記(24)記載の方法であることを特徴とする。

【0052】また、本発明に係る方法においては、(28)プローブの少なくとも15個が、WT1に活性化またはWT1に抑制される遺伝子の一部を含んでいる、上記(24)記載の方法であることを特徴とする。

T1に抑制される遺伝子の一部を含んでいる、上記(24)記載の方法であることを特徴とする。

【0053】また、本発明に係る方法においては、(29)核酸プローブが固体支持体に結合している、上記(24)記載の方法であることを特徴とする。

【0054】また、本発明に係る方法においては、(30)核酸プローブがアレイ上に配置されている、上記(24)記載の方法であることを特徴とする。

【0055】また、本発明に係る方法においては、(31)アレイが少なくとも250の異なる遺伝子の一部である核酸プローブを含んでいる、上記(30)記載の方法であることを特徴とする。

【0056】また、本発明に係る方法においては、(32)アレイが少なくとも6000の異なる遺伝子の一部である核酸プローブを含んでいる、上記(30)記載の方法であることを特徴とする。

【0057】また、本発明に係る方法においては、(33)細胞のWT1遺伝子型の状態を決定するために、試験細胞中でWT1遺伝子の配列を決定する段階をさらに含む、上記(24)記載の方法であることを特徴とする。

【0058】また、本発明に係る方法においては、(34)比較されるサンプル間でハイブリダイゼーションに少なくとも2倍の差がある場合に試験細胞が新生物形成性であると同定される、上記(24)記載の方法であることを特徴とする。

【0059】また、本発明に係る方法においては、(35)下記の段階を含む、抗癌剤の同定方法であることを特徴とする：試験化合物をヒト細胞と接触させる段階；WT1にアップレギュレーションされる少なくとも1つの遺伝子、またはWT1にダウンレギュレーションされる少なくとも1つの遺伝子の、細胞における発現を測定する段階であって、WT1にアップレギュレーションされる遺伝子が、ナチュラルキラー細胞プロテイン4前駆体 (M59807)、熱ショック蛋白質HSP70B (X51558)、90K産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、プロコラーゲン 1 (T51558)、放射線照射ケラチノサイト由来のシスティンプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質 (BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン (R71870)、プリンヌクレオシドホスホリラーゼ (T47964)、アドレノメダリン (D14874)、グラビン (M96322)、jun-B (X51345)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57348)、インターフェロン 処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、酸性FGF (X65779)、軟骨結合蛋白質-1 (X17405)、IL-11 (X58377)、およびミトコンドリアリン酸輸送蛋白質 (R49231) からなる群より選択され、WT1にダウンレギュレーションされる遺伝子が延長因子1 -2 (X70940) である測定段階；および試験化合物が、ヒト細胞においてWT1にアップレギュレーションされる少なくとも1つの遺伝子の

40

発現を上昇させるか、WT1にダウンレギュレーションされる少なくとも1つの遺伝子の発現を低下させる場合に、抗癌剤の可能性があると同定する段階。

【0060】また、本発明に係る方法においては、(36)細胞が腎臓癌細胞である上記(35)記載の方法であることを特徴とする。

【0061】また、本発明に係る方法においては、(37)細胞が腫瘍細胞である上記(35)記載の方法であることを特徴とする。

【0062】また、本発明に係る方法においては、(38)細胞に野生型WT1遺伝子が含まれない上記(35)記載の方法であることを特徴とする。

【0063】また、本発明に係る方法においては、(39)WT1にアップレギュレーションおよびダウンレギュレーションされる遺伝子の発現が、細胞中の該遺伝子の転写産物の量を測定することにより検出される上記(35)記載の方法であることを特徴とする。

【0064】また、本発明に係る方法においては、(40)転写産物の量が高密度核酸アレイを用いて測定される上記(39)記載の方法であることを特徴とする。

【0065】また、本発明に係る方法においては、(41)下記の段階を含む、試験細胞の新生物形成性の診断方法であることを特徴とする：mRNA、cDNAおよびcRNAからなる群より選択される試験細胞の転写指標を、核酸プローブのセットにハイブリダイズさせる段階であって、核酸プローブのセットは、各々がEWS-WT1融合蛋白質に活性化または抑制される遺伝子の一部である複数の核酸分子を含むハイブリダイズ段階；各々の該核酸プローブセットにハイブリダイズする転写指標の量を検出する段階；ならびに(1)EWS-WT1に活性化される遺伝子であるプローブへの試験細胞の転写指標のハイブリダイゼーションが、正常細胞の転写指標のハイブリダイゼーションよりも低レベルである、または(2)EWS-WT1に抑制される遺伝子であるプローブへの試験細胞の転写指標のハイブリダイゼーションが、正常細胞の転写指標のハイブリダイゼーションよりも高レベルである場合に、試験細胞が新生物形成性であると同定する段階。

【0066】また、本発明に係る方法においては、(42)試験細胞がユーイング肉腫細胞である上記(41)記載の方法であることを特徴とする。

【0067】また、本発明に係る方法においては、(43)プローブの少なくとも4つが、EWS-WT1に活性化またはEWS-WT1に抑制される遺伝子の一部を含んでいる上記(41)記載の方法であることを特徴とする。

【0068】また、本発明に係る方法においては、(44)プローブの少なくとも10個が、EWS-WT1に活性化またはEWS-WT1に抑制される遺伝子の一部を含んでいる上記(41)記載の方法であることを特徴とする。

【0069】また、本発明に係る方法においては、(45)プローブの少なくとも15個が、EWS-WT1に活性化ま

たはEWS-WT1に抑制される遺伝子の一部を含んでいる上記(41)記載の方法であることを特徴とする。

【0070】また、本発明に係る方法においては、(46)核酸プローブが固体支持体に結合している上記(41)記載の方法であることを特徴とする。

【0071】また、本発明に係る方法においては、(47)核酸プローブがアレイ上に配置されている上記(41)記載の方法であることを特徴とする。

【0072】また、本発明に係る方法においては、(48)アレイが少なくとも250の異なる遺伝子の一部である核酸プローブを含んでいる上記(47)記載の方法であることを特徴とする。

【0073】また、本発明に係る方法においては、(49)アレイが少なくとも6000の異なる遺伝子の一部である核酸プローブを含んでいる上記(47)記載の方法であることを特徴とする。

【0074】また、本発明に係る方法においては、(50)少なくとも1つの核酸プローブが、仮性狂犬病ウイルス核抗原(R60906およびT79475)、インターロイキン

20 -2受容体鎖(M26062)、G1/S-特異的サイクリンD1(H20529)、ニューロン特異的-2エノラーゼ(M22349)、転座t(3;5)(q25.1;p34)により得られる融合遺伝子NPM-MLF1(L49054)、コラーゲン3(IV)鎖(H62466)、T-リンパ球特異的蛋白質チロシンキナーゼp56lck(lck)(U23852)、インスリン様成長因子結合蛋白質-4(U20982)、カテプシンB、シスタチンC(T51534)、二重特異的マイトジエン活性化蛋白質キナーゼ2(R42291)、神経グリア細胞接着分子(T53118)、PIM-1癌原遺伝子セリン/スレオニン-蛋白質キナーゼ(H10925)、カルボキシルエステラーゼ前駆体(R54359)、蛋白質-チロシンホスファターゼPAC-1(L11329)、およびインターフェロン誘導蛋白質6-16(X02492)からなる群より選択される、EWS-WT1にアップレギュレーションされる遺伝子の一部を含む上記(41)記載の方法であることを特徴とする。

【0075】また、本発明に係る方法においては、(51)比較されるサンプル間でハイブリダイゼーションに少なくとも2倍の差がある場合に試験細胞が新生物形成性であると同定される上記(41)記載の方法であることを特徴とする。

【0076】また、本発明に係る方法においては、(52)比較されるサンプル間でハイブリダイゼーションに少なくとも3倍の差がある場合に試験細胞が新生物形成性であると同定される上記(41)記載の方法であることを特徴とする。

【0077】また、本発明に係る方法においては、(53)比較されるサンプル間でハイブリダイゼーションに少なくとも5倍の差がある場合に試験細胞が新生物形成性であると同定される上記(41)記載の方法であることを特徴とする。

【0078】また、本発明に係る方法においては、(54)下記の段階を含む、抗癌剤の同定方法であることを特徴とする：試験化合物をヒト細胞と接触させる段階；EWS-WT1にアップレギュレーションされる少なくとも1つの遺伝子、またはEWS-WT1にダウンレギュレーションされる少なくとも1つの遺伝子のヒト細胞における発現を測定する段階；および試験化合物が、EWS-WT1にアップレギュレーションされる遺伝子の発現を低下させるか、EWS-WT1にダウンレギュレーションされる遺伝子の発現を上昇させる場合に、抗癌剤であると同定する段階。

【0079】また、本発明に係る方法においては、(55)細胞が腎臓癌細胞である上記(54)記載の方法であることを特徴とする。

【0080】また、本発明に係る方法においては、(56)細胞が腫瘍細胞である上記(54)記載の方法であることを特徴とする。

【0081】また、本発明に係る方法においては、(57)WT1にアップレギュレーションおよびダウンレギュレーションされる遺伝子の発現が、細胞中の該遺伝子の転写産物の量を測定することにより検出される上記(54)記載の方法であることを特徴とする。

【0082】また、本発明に係る方法においては、(58)転写産物の量が高密度核酸アレイを用いて測定される上記(57)記載の方法であることを特徴とする。

【0083】また、本発明に係る方法においては、(59)WT1にアップレギュレーションされる遺伝子が、仮性狂犬病ウイルス核抗原 (R60906およびT79475)、インターロイキン-2受容体 鎮 (M26062)、G1/S-特異的サイクリンD1 (H20529)、ニューロン特異的 -2エノラーゼ (M22349)、転座t(3;5)(q25.1;p34)により得られる融合遺伝子NPM-MLF1 (L49054)、コラーゲン 3(IV)鎖 (H62466)、T-リンパ球特異的蛋白質チロシンキナーゼ p56lck (lck) (U23852)、インスリン様成長因子結合蛋白質-4 (U20982)、カテーテンB、シスタチンC (T51534)、二重特異的マイトジエン活性化蛋白質キナーゼキナーゼ2 (R42291)、神経グリア細胞接着分子 (T53118)、PIM-1癌原遺伝子セリン/スレオニン-蛋白質キナーゼ (H10925)、カルボキシルエステラーゼ前駆体 (R54359)、蛋白質-チロシンホスファターゼPAC-1 (L11329)、およびインターフェロン誘導蛋白質6-16 (X02492)からなる群より選択される上記(54)記載の方法であることを特徴とする。

【0084】また、本発明に係る方法においては、(60)下記の段階を含む、標的細胞中でWT1遺伝子の機能変異を検出する方法であることを特徴とする：

(a) 標的細胞、および(b) 野生型WT1遺伝子を持ち、その他の点では標的細胞に実質的に類似した参照細胞のサンプル中で、少なくとも1つのWT1の下流遺伝子の発現を検出する段階であって、該下流遺伝子が、ナチュラルキラー細胞プロテイン4前駆体 (M59807)、葉酸結合蛋白質

(M25317)、チューブリンのHALPHA44遺伝子 (X06956)、熱ショック蛋白質HSP70B (X51758)、90K産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、アンフィレグリン (M30704)、プロコラーゲン 1 (T51558)、遊走性プラスミノーゲン活性化因子阻害剤1 (M14083)、放射線照射ケラチノサイト由来のシスティンプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質 (BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン、チューブリン 5鎖 (H45051)、プリンヌクレオシドホスホリラーゼ (T47964)、Gem GTPアーゼ (U10550)、アドレノメダリン (D14874)、GPIアンカー型ウロキナーゼプラスミノーゲン活性化因子表面受容体 (R38636)、組織型プラスミノーゲン活性化因子 (X13097)、グラビン (M96322)、jun-B (X51345)、延長因子1 -2 (X79490)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57348)、インターフェロン 処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、酸性FGF (X65779)、軟骨結合蛋白質-1 (X17405)、IL-11 (X58377)、ミトコンドリアリン酸輸送蛋白質 (R49231)、およびカベオリンからなる群より選択される検出段階；ならびに標的細胞および該参照細胞における少なくとも1つの下流遺伝子の発現を比較する段階であって、標的細胞と参照細胞とで該発現に差があれば、標的細胞にWT1機能変異が存在することが示唆される段階。

【0085】また、本発明に係る方法においては、(61)下記の段階を含む、WT1配列変化の細胞内機能アッセイ法であることを特徴とする：WT1配列変化を持つ標的細胞の標的サンプル、および野生型WT1遺伝子を持ち、その他の点では標的細胞に実質的に類似した参照細胞の参照サンプルにおいて、ナチュラルキラー細胞プロテイン4前駆体 (M59807)、葉酸結合蛋白質 (M25317)、チューブリンのHALPHA44遺伝子 (X06956)、熱ショック蛋白質HSP70B (X51758)、90K産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、アンフィレグリン (M30704)、プロコラーゲン 1 (T51558)、遊走性プラスミノーゲン活性化因子阻害剤1 (M14083)、放射線照射ケラチノサイト由来のシスティンプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質 (BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン、チューブリン 5鎖 (H45051)、プリンヌクレオシドホスホリラーゼ (T47964)、Gem GTPアーゼ (U10550)、アドレノメダリン (D14874)、GPIアンカー型ウロキナーゼプラスミノーゲン活性化因子表面受容体 (R38636)、組織型プラスミノーゲン活性化因子 (X13097)、グラビン (M96322)、jun-B (X51345)、延長因子1 -2 (X79490)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57348)、インターフェロン 処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、カベオリン、酸性FGF (X65779)、軟骨結合蛋白質-1 (X17405)、IL-11 (X58377)、およびミトコンドリアリン酸輸

送蛋白質 (R49231) からなる群より選択される少なくとも1つの下流遺伝子の発現を検出する段階であって、該下流遺伝子は野生型WT1遺伝子にアップレギュレーションまたはダウンレギュレーションされる検出段階；ならびに該標的サンプル中の該発現と該参照サンプル中の該発現を比較する段階であって、2つのサンプル間で発現に差があれば、該WT1配列変化がWT1の生物学的機能に影響を与えることが示唆される段階。

【0086】また、本発明に係る方法においては、(62)下記の段階を含む、コンピュータを用いた標的WT1遺伝子中の変異の検出方法であることを特徴とする：野生型WT1遺伝子を含む野生型サンプル中の、少なくとも1つのWT1下流遺伝子の野生型発現データをコンピュータに入力する段階であって、下流遺伝子が、ナチュラルキラー細胞プロテイン4前駆体 (M59807)、葉酸結合蛋白質 (M25317)、チューブリンのHALPHA44遺伝子 (X06956)、熱ショック蛋白質HSP70B (X51758)、90K産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、アンフィレグリン (M30704)、プロコラーゲン 1 (T51558)、遊走性プラスミノーゲン活性化因子阻害剤1 (M14083)、放射線照射ケラチノサイト由来のシステインプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質 (BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン、チューブリン 5鎖 (H45051)、プリンスクレオシドホスホリラーゼ (T47964)、GTPアーゼ (U10550)、アドレノメダリン (D14874)、GPIアンカー型ウロキナーゼプラスミノーゲン活性化因子表面受容体 (R38636)、組織型プラスミノーゲン活性化因子 (X13097)、グラビン (M96322)、jun-B (X51345)、延長因子1

-2 (X79490)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57348)、インターフェロン 処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、カベオリン、酸性FGF (X65779)、軟骨結合蛋白質-1 (X17405)、IL-11 (X58377)、およびミトコンドリアリン酸輸送蛋白質 (R49231) からなる群より選択される入力段階；標的WT1遺伝子を含む標的サンプル中の複数の該下流遺伝子の標的発現データをコンピュータに入力する段階；ならびに標的および野生型の発現データを比較して、差を測定する段階であって、差があれば、該標的WT1遺伝子中に変異が存在することが示唆される段階。

【0087】また、本発明に係る方法においては、(63)下記の段階を含む、試験細胞の新生物形成性の診断方法であることを特徴とする：ナチュラルキラー細胞プロテイン4前駆体 (M59807)、熱ショック蛋白質HSP70B (X51758)、90K産物 (H17969)、熱ショック70kd蛋白質1 (T66307)、プロコラーゲン 1 (T51558)、放射線照射ケラチノサイト由来のシステインプロテアーゼ阻害剤 (X05978)、脳性ナトリウム利尿蛋白質 (BNP) (M31766)、白血病ウイルス受容体1 (GLVR1) (L20859)、1型細胞骨格17ケラチン (R71870)、プリンスクレオシドホスホリラー

ゼ (T47964)、アドレノメダリン (D14874)、グラビン (M96322)、jun-B (X51345)、延長因子1 -2 (X79490)、ホメオティック遺伝子調節因子 (R16977)、クローン9112 (X57348)、インターフェロン 処理誘導可能mRNA (M26683)、膜貫通受容体蛋白質 (Z17227)、酸性FGF (X65779)、軟骨結合蛋白質-1 (X17405)、IL-11 (X58377)、およびミトコンドリアリン酸輸送蛋白質 (R49231) からなる群より選択される、WT1に活性化または抑制される少なくとも1つの遺伝子の試験細胞および対照細胞における発現レベルを測定する段階；試験細胞で測定された発現レベルを対照細胞で測定された発現レベルと比較する段階；ならびに(1) WT1に活性化される遺伝子である少なくとも1つの遺伝子の発現レベルが、対照細胞よりも試験細胞のほうが低い、または(2) WT1に抑制される遺伝子の少なくとも1つの発現レベルが、対照細胞よりも試験細胞のほうが高い場合に、試験細胞が新生物形成性であると同定する段階。

【0088】また、本発明に係る方法においては、(64)下記の段階を含む、試験細胞の新生物形成性の診断方法であることを特徴とする：EWS-WT1融合蛋白質に活性化または抑制される少なくとも1つの遺伝子の試験細胞および対照細胞における発現レベルを測定する段階；試験細胞の測定された発現レベルを対照細胞の測定された発現レベルと比較する段階；ならびに(1) EWS-WT1に活性化される遺伝子の試験細胞における発現が、正常細胞における発現よりも低い、または(2) EWS-WT1に抑制される遺伝子の試験細胞における発現が、正常細胞における発現よりも高い場合に、試験細胞が新生物形成性であると同定する段階。

【0089】また、本発明に係る方法においては、(65)下記の段階を含む、試験細胞の新生物形成性の診断方法であることを特徴とする：WT1に活性化または抑制される少なくとも3つの遺伝子の試験細胞および対照細胞における発現レベルを測定する段階；試験細胞にて測定された発現レベルを対照細胞にて測定される発現レベルと比較する段階；ならびに(1) WT1に活性化される遺伝子の試験細胞における発現が、正常細胞における発現よりも低い、または(2) WT1に抑制される遺伝子の試験細胞における発現が、正常細胞における発現よりも高い場合に、試験細胞が新生物形成性であると同定する段階。

【0090】

【発明の実施の形態】WT1は遺伝子セットの転写を調節しており、そのセットのサイズは以前に示唆されたよりもはるかに大きいというのが、本発明の所見である。WT1によってアップレギュレーションおよびダウンレギュレーションされる遺伝子のセットが同定されたが、後者のセットの方が、はるかに小さい。さらに、転座によって生じるEWS-WT1遺伝子融合によってアップレギュレーションされる遺伝子が同定されている。この転座は、ユ

ーイング肉腫のような線維形成性の円形細胞腫で起こる。これらの遺伝子セットを同定すると、様々な診断方法および分析方法に利用できる。また、WT1と同一または類似した効果を持つ薬物、およびEWS-WT1遺伝子融合と反対の効果を持つ薬物のための化合物をスクリーニング\*

WT1標的の候補	
M59807	ナチュラルキラー細胞プロテイン4前駆体(ヒト)；エレメントMSR1反復エレメントを含む
M25317	ヒト葉酸結合蛋白質(FBP)mRNA、3'末端
X06956	ヒト $\alpha$ チューブリンのHALPHA44遺伝子、エクソン1~3
X51758	ヒト熱ショック蛋白質HSP70B'のmRNA
H17969	ヒト90K産物のmRNA
T66307	熱ショック70kDa蛋白質1(ヒト)
M30704	ヒトアンフィレグリン(AR)mRNA、完全cds、クローン $\lambda$ -AR1および $\lambda$ -AR2
T51558	プロコラーゲン $\alpha$ 1(I)鎖前駆体(ヒト)
M14083	ヒト $\beta$ 遊走性プラスミノーゲンアクチベーター阻害剤I mRNA、3'末端
X05978	ヒト照射ケラチノサイトのシスティンプロテアーゼ阻害剤のmRNA
Z18951	ヒトカベオリンのmRNA
M31776	ヒト脳ナトリウム排泄増加性蛋白質(BNP)遺伝子、完全cds
L20859	ヒト白血病ウイルス受容体1(GLVR1)mRNA、完全cds
R71870	ケラチン、I型細胞骨格17(ヒト)
H45051	チューブリン $\alpha$ -5鎖(ニワトリ)
T47964	ブリンヌクレオンドホスホリラーゼ(ヒト)
U10550	ヒトGem GTPアーゼ(gem)mRNA、完全cds
D14874	ヒトアドレノメダリンmRNA、完全cds
R38636	ウロキナーゼプラスミノーゲンアクチベーター表面受容体、GPI-アンカー型(ヒト)
X13097	ヒト組織型プラスミノーゲンアクチベーターmRNA
M96322	ヒトグラビン(重症筋無力症血清が認識する細胞質抗原)mRNA、3'末端
X51345	ヒトJUN-B蛋白質のjun-B mRNA
X70940	ヒト延長因子1 $\alpha$ -2のmRNA
R16077	ホメオティック遺伝子調節因子(ショウジョウバエ)
X57348	ヒトmRNA(クローン9112)
M26683	インターフェロン $\gamma$ 処理で誘導されるヒトmRNA
Z17227	ヒト膜貫通受容体蛋白質のmRNA
R49231	ミトコンドリアン酸輸送蛋白質前駆体(ヒト)

【表1 b】 WT1(-KTS)によって誘導されるさらに別\*の標的遺伝子

寄託番号	説明	チップによる誘導倍率
X65779	酸性FGF	22.1
X17405	軟骨結合蛋白質1	46.1
X58377	IL-11	18.8

これらの遺伝子のうち、カベオリンおよび延長因子1-2のみが、WT1にダウンレギュレーションされる。他の遺伝子は、WT1にアップレギュレーションされる。

【0092】表2は、WT1に調節される遺伝子の中で、以

\*グするのにも使用できる。

【0091】表1aおよび1bは、WT1に調節される遺伝子を示す。

【表1 a】

前に癌との関連が示唆されている遺伝子のサブセットと、その関連を示す。

40 【表2】 WT1によって誘導または抑制される遺伝子

遺伝子	遺伝子クラス	誘導倍率	記述の癌との関係
$\alpha$ チューブリン	細胞骨格	12	移動、浸潤
組織型プラスミノーゲンアクチベーター	プロテアーゼ系	4	移動、浸潤、転移
ウロキナーゼプラスミノーゲンアクチベーター阻害剤	プロテアーゼ系	4	移動、浸潤、転移
ウロキナーゼプラスミノーゲンアクチベーター受容体	プロテアーゼ系	3	移動、浸潤、転移
シスタチン	プロテアーゼ系	6	接着、運動性、レベルは癌の予後と相關
カベオリン	膜	(-4)	足場非依存性増殖
アンフィレグリン	増殖因子	8	新生物、浸潤、転移
Mac-2結合蛋白質	分泌	8	細胞接着、多くの腫瘍により分泌
葉酸受容体	受容体	10	卵巣癌の90%で過剰発現、腫瘍の増殖を促進
GEM	グアニンヌクレオチド結合	4	癌遺伝子を介した形質転換により誘導

【0093】表3は、EWS-WT1融合遺伝子に調節される遺伝子を示す。

50

【表3】

EWS	
R60906	推定核抗原 (仮性狂犬病ウイルス)
M26062	インターロイキン-2受容体 $\beta$ 鎮前駆体 (ヒト)
H20529	G1/S-特異的サイクリンD1 (ヒト)
T79475	推定核抗原 (仮性狂犬病ウイルス)
M22349	ヒトニューロン特異的 $\gamma$ -2エノラーゼ、完全cds
L49054	ヒトt(3;5)(q25.1;p34)融合遺伝子NPM-MLF1 mRNA、完全cds
H62466	コラーゲン $\alpha$ 3(VI)鎖 (ニワトリ)
U23852	ヒトTリンパ球特異的蛋白質チロシンキナーゼp56lck (lck)異常型mRNA、完全cds
U20982	ヒトインスリン様成長因子結合蛋白質-4(IGFBP4) 遺伝子、プロモーターおよび完全cds
L16510	ヒトカテプシンmRNA、完全cds
T61534	シスタチンC前駆体 (ヒト)
R42291	二重特異的マイトジエン活性化蛋白質キナーゼキナーゼ2 (ヒト)
T53118	神経グリア細胞接着分子前駆体 (ニワトリ)
H10925	PIM-1癌原遺伝子セリン/スレオニン-蛋白質キナーゼ(ハツカネズミ)
R54359	カルボキシルエステラーゼ前駆体(ハツカネズミ)
L11329	蛋白質-チロシンホスファターゼPAC-1(ヒト)
X02492	インターフェロンに誘導される蛋白質6-16前駆体 (ヒト) ; L1反復エレメントを含む

これらの遺伝子は、すべて融合遺伝子にアップレギュレーションされる。間違いない、これらの遺伝子によって他の遺伝子も調節されており、それらは、実施例に記述される技術と同様の技術を用いて同定できる。または、遺伝子発現の連続分析(SAGE)のような他の技術を使って、まだ同定されていないか公開データベースにない、調節される遺伝子を同定することもできる。

【0094】WT1またはEWS-WT1の下流で調節される遺伝子の発現は、当技術分野で公知の任意のテクニックを用いてモニターできる。これには、SAGE、アレイ上のプローブへのハイブリダイゼーション、ウェスタンプロット、ノーザンプロット、ドットプロット、スロットプロット、ELISA、ラジオイムノアッセイの使用が含まれるが、これらに限定されるわけではない。mRNAまたは蛋白質のいずれでも、遺伝子産物の定量に使用される任意の技術が使用できる。本発明の方法は、主に、対照サンプルと比較した、試験サンプル中の変化を分析する。精密な定量を実施する必要はない。多くの場合、遺伝子産物の絶対量を決定しなくとも、比較で十分である。

【0095】本発明で使用される標的細胞は、試験される細胞である。典型的には、その変異の状態は不明である。標的細胞で変異が同定されているが、機能面での効果は不明の場合もある。参照細胞は、WT1遺伝子またはEWS-WT1遺伝子に起因する差異が容易に検出できるように、好ましくは標的細胞とできるだけ類似したものである。通常は、参照細胞は野生型のWT1対立遺伝子を含む。しばしば、参照細胞は、同一個体、および試験する腫瘍サンプルに隣接する組織から採取した細胞である。

【0096】標的細胞と参照細胞との差異は、小さい場合も大きい場合もある。いくつかの小さな差異には、非常に再現性が高く、それ故有用なものもある。他の目的のためには、検出が容易なため、大きな差異の方が望しい可能性がある。アッセイ法によっては、変化の閾値レベルを設定することが有用なことがある。目的によつ

ては、そのような閾値レベルは0.5、2、3、5、7、または10倍であり得る。これは、検出に使用する技術、および試験する遺伝子の数に依存する。当業者は、特定の方

20 法を用いたときに好ましい閾値レベルを容易に設定できる。

【0097】分析する下流遺伝子の数は、例えば、誘導される変化の程度およびその再現性などによって様々である。分析する特定の遺伝子の発現には、他の遺伝子も影響する可能性があるため、WT1の変化をより正確に反映するために、より多くの遺伝子を分析する場合もある。通常、WT1またはEWS-WT1の下流の、少なくとも1、2、3、4、5、7、10、または15の遺伝子が分析される。複数の遺伝子の分析は、同時に、連続的に、または追加方式で実施できる。

【0098】本発明では、データの保存、数学的比較、結果の保存などのためにコンピュータを使用できる。データは、手作業で入力するか、光学スキャナー、シンチレーションカウンターなどのような他の機械から送ることができる。したがって、高処理量スクリーニングは効率良く実施し、分析することができる。

#### 【0099】1. 定義

特異的にハイブリダイズする：「特異的にハイブリダイズする」という用語は、あるヌクレオチド配列がDNAまたはRNAの複雑な混合物（例、全細胞内）中に存在するときに、ある分子がストリンジメントな条件下で、実質的に一つまたは複数の特定のヌクレオチド配列に、または該配列にのみ結合、二本鎖形成、またはハイブリダイズすることを指す。

【0100】mRNAまたは転写産物：「mRNA」という用語は、遺伝子の転写産物を指す。転写産物は、例えばそのまま翻訳可能な成熟したメッセンジャーRNA、転写産物のプロセッシングの種々の段階の産物などを含むRNAである。転写産物のプロセッシングには、スプライシング、編集、および分解が含まれ得る。

【0101】核酸：「核酸」または「核酸分子」という用語は、一本鎖または二本鎖のデオキシリボヌクレオチドまたはリボヌクレオチドポリマーを指し、特に制限がないかぎり、天然に存在するヌクレオチドと同様に機能できる天然ヌクレオチドの類似体も含む。オリゴヌクレオチドは、2~n塩基の一本鎖核酸であり、nは500~1000以上である可能性がある。核酸は当技術分野で公知の任意の技術を用いて、クローニングもしくは合成できる。天然に存在しないヌクレオチド類似体およびペプチド核酸も含み得る。

【0102】調節分子をコードする核酸：調節分子は、DNA、RNA、または蛋白質であり得る。したがって、例えば蛋白質または他の核酸分子に結合するDNA部位は、核酸によってコードされる調節分子のクラスに含まれる。

【0103】プローブ：本明細書で用いられる「プローブ」とは、通常は水素結合の形成による、通常は相補的な塩基の対合による、1つまたは複数種類の化学結合によって、相補的な配列を持つ標的核酸に結合する能力のある核酸と定義される。本明細書では、プローブには天然（すなわちA、G、U、C、もしくはT）、または修飾塩基（7-デアザグアノシン、イノシン等）が含まれ得る。また、プローブ中の塩基は、ハイブリダイゼーションに干渉しないかぎり、ホスホジエステル結合以外の結合によって結びついている可能性がある。したがって、プローブは、構成する塩基がホスホジエステル結合ではなくペプチド結合によって結びついているペプチド核酸であり得る。

【0104】標的核酸：「標的核酸」という用語は、プローブが特異的にはハイブリダイズするように設計されている核酸（しばしば生体サンプル由来）を指す。標的核酸の有無を検出するか、標的核酸の量を定量する。標的核酸は、標的に向けられる対応するプローブの核酸配列に相補的な配列を持つ。標的核酸という用語は、プローブが向けられる、より大きな核酸中の特定のサブ配列、または発現レベルを検出することが望ましい全体の配列（例、遺伝子またはmRNA）を指す場合もある。使用法の違いは、文脈から明らかであると考えられる。

【0105】ストリンジエントな条件：「ストリンジエントな条件」という用語は、プローブが標的のサブ配列にはハイブリダイズするが、他の配列には実質的にハイブリダイズせず、その差異が同定できる条件を指す。ストリンジエントな条件は、配列に依存し、状況が異なれば、条件も異なる。長い配列の方が、高い温度で特異的にハイブリダイズする。一般にストリンジエントな条件には、定められたイオン強度とpHにおける特異的な配列の融解温度（T<sub>m</sub>）よりも、約5~低い温度が選択される。

【0106】サブ配列：「サブ配列」は、より長い核酸配列の一部を含む核酸配列を指す。

【0107】融解温度（T<sub>m</sub>）：融解温度は、定められたイオン強度、pH、および核酸濃度で、標的配列に相補的なプローブの50%が標的配列に平衡状態でハイブリダイズする温度である。標的配列は通常過剰に存在するので、T<sub>m</sub>では、プローブの50%が平衡状態にある。典型的には、ストリンジエントな条件とは、pH 7.0~8.3で塩濃度が少なくとも約0.01~1.0 M Naイオン濃度（または他の塩）で、短いプローブ（例、10~50ヌクレオチド）の場合に温度は少なくとも約30~である。ストリンジエントな条件は、フォルムアミドのような不安定化剤を添加しても達成し得る。

【0108】定量：遺伝子の転写レベルの定量をするという文脈で使用される場合の「量を計る」または「定量」という用語は、絶対量または相対量の定量を指すことができる。絶対量の定量は、既知の濃度の1つまたは複数の標的核酸（Bio Bのような対照核酸、または既知の量の標的核酸自身）を含め、既知の標的核酸と未知のもののハイブリダイゼーション強度を比較することによって（例、標準曲線の作製による）実行できる。または、相対量の定量は、2つもしくはそれ以上の遺伝子、または2つもしくはそれ以上の処理の間のハイブリダイゼーションシグナルの比較をして、ハイブリダイゼーション強度と、したがって転写レベルの変化を定量することで実行できる。

【0109】上流または下流遺伝子：第1の遺伝子の発現が第2の遺伝子によって調節されている場合、第2の遺伝子は第1の遺伝子の「上流遺伝子」と呼ばれ、第1の遺伝子は第2の遺伝子の「下流」遺伝子と呼ばれる。第2の遺伝子による第1の遺伝子の調節は、トランス活性化による可能性がある。例えば、第1の遺伝子が第2の遺伝子の発現をコントロールする転写因子をコードする場合である。調節は、シス作用機構による場合もある。例えば、第1の遺伝子が第2の遺伝子に近接しており、第2の遺伝子の発現に位置効果を発揮する場合である。この場合、第1の遺伝子は、第2の遺伝子に影響するために、発現される必要はない。

【0110】遺伝子の活性は、その産物、すなわち遺伝子によってコードされる蛋白質または他の分子の活性に反映される。このような産物分子は、生物学的機能を担っている。しかし、遺伝子産物の活性を直接に測定することは、一部の遺伝子についてはしばしば困難である。その代わりに、遺伝子活性の尺度として、最終産物またはそのペプチドプロセッシング中間体の免疫活性または量が決定される。多くの場合、転写産物、RNAプロセッシング中間体、または成熟mRNAのような中間体の量または活性が、遺伝子活性の尺度として検出される。

【0111】多くの場合、遺伝子の最終産物の形態および機能は不明である。そのような場合、遺伝子活性は、転写産物、RNAプロセッシング中間体、成熟mRNA、もしくはその蛋白質産物の量または活性、または蛋白質産物

の機能的活性によって、便宜的に測定される。

【0112】遺伝子の活性を測定する任意の方法は、本発明の少なくともいくつかの態様に有用である。例えば、一般的なノーザンプロッティングおよびハイブリダイゼーション、スクレアーゼ保護、RT-PCR、ならびにディファレンシャルディスプレイなどが、遺伝子活性の検出に使用されてきた。これらの方法は、本発明のいくつかの態様に有用である。しかし、本発明は、多数の遺伝子の発現を検出する方法と併用すると、最も有用である。

【0113】高密度アレイは、転写、RNAプロセッシングおよび分解レベルで発現コントロールをモニターするために、特に役立つ。遺伝子発現のモニタリングにおける高密度アレイの製作と適用は、例えば国際公開公報第97/10365号、国際公開公報第92/10588号、1996年12月23日出願の米国特許出願第08/772,376号、1995年9月15日出願の米国特許出願第08/529,115号、1993年12月15日出願の米国特許出願第08/168,904号、1990年12月6日出願の米国特許出願第07/624,114号、1990年6月7日出願の米国特許出願第07/362,901号などにすでに開示されており、これらはすべて参考として本明細書に組み入れられる。高密度アレイを用いるいくつかの態様では、参考として本明細書に組み入れられる米国特許第5,445,934号に開示される超大規模固定化ポリマー合成(VLSIPS)のような方法を用いて、高密度オリゴスクレオチドアレイが合成される。各オリゴスクレオチドは、基質上の既知の位置を占めている。核酸標的サンプルをオリゴスクレオチドの高密度アレイにハイブリダイズさせ、アレイ上の各プローブにハイブリダイズした標的核酸の量を定量する。好ましい1つの定量方法は、共焦点顕微鏡および蛍光標識を使用することである。GeneChip(登録商標)システム(Affymetrix、カリフォルニア州サンタクララ)は、ハイブリダイゼーションの定量に特に適している；しかし、任意の類似システムまたは他の実際上同等の検出方法が使用できることは、当業者には明らかであると思われる。

【0114】高密度アレイは大量の不均質な核酸群の存在下で、ある遺伝子の発現量の小さな変動を定量するために適している。そのような高密度アレイは、基質上で新たに合成するか、基質の特定の位置に核酸配列をスポットするまたは運ぶことによって、製作できる。核酸は興味のある配列のクローニングされたセグメントを含むバクテリアのプラスミドのような、生体材料から精製および/または単離される。適当な核酸は、鋳型の増幅によっても生産できる。限定することのない実例としては、ポリメラーゼ連鎖反応および/またはインビトロ転写は、適当な核酸増幅法である。

【0115】合成されたオリゴスクレオチドアレイは、本発明に特に好ましい。他の方法と比較して、オリゴスクレオチドアレイには、生産効率、アレイ内およびアレ

イ間の変動の低下、情報内容の増加、および高い信号対雑音比のような、数多くの利点がある。

【0116】遺伝子機能の同定および遺伝ネットワークマッピングに好ましい高密度アレイには、好ましくは1cm<sup>2</sup>以下の表面積に、約100以上、好ましくは約1000以上、さらに好ましくは約16,000以上、および最も好ましくは65,000もしくは250,000以上、または約1,000,000以上の異なるオリゴスクレオチドプローブが含まれている。オリゴスクレオチドプローブの長さは、約5から約5

10または約500スクレオチド、さらに好ましくは約10から約40スクレオチド、および最も好ましくは約15から40スクレオチドである。

【0117】一般に、遺伝子発現のモニタリング法には、(a) 1つまたは複数の標的遺伝子のRNA転写産物を含む標的核酸またはRNA転写産物に由来する核酸のプールを提供する段階；(b) 核酸サンプルをプローブの高密度のアレイにハイブリダイズする段階、および(c) ハイブリダイズした核酸を検出し、相対的および/または絶対的な発現(転写、RNAプロセッシングまたは分解)レベルを計算する段階、が含まれる。

【0118】(A) 核酸サンプルの提供

当業者は、興味のある転写産物を反映する標的核酸配列を含む核酸サンプルがあることが望ましいことを認識すると思われる。したがって、適切な核酸サンプルには、興味のある転写産物が含まれ得る。しかし、適切な核酸サンプルには、興味のある転写産物に由来する核酸が含まれることもある。本明細書では、転写産物に由来する核酸とは、mRNA転写産物またはそのサブ配列が、その合成のために最終的に鋳型として使われた核酸を指す。したがって、転写産物から逆転写されたcDNA、そのcDNAから転写されたRNA、そのcDNAから増幅されたDNA、その増幅されたDNAから転写されたRNA等は、すべて転写産物に由来するものであり、そのような産物が検出されることにより、サンプル中の元の転写産物の存在および/またはその量が示される。したがって、適当なサンプルには、遺伝子の転写産物、転写産物から逆転写されたcDNA、そのcDNAから転写されたcRNA、その遺伝子から増幅されたDNA、増幅されたDNAから転写されたRNA等が含まれるが、これに限定されない。

【0119】本明細書では、転写産物にはプレmRNA新生転写産物、転写産物プロセッシング中間体、成熟mRNA、および分解産物が含まれるが、これに限定されない。本発明の実施のためにすべての種類の転写産物をモニターする必要はない。例えば、成熟mRNAレベルのみを測定して本発明を実施することができる。

【0120】1つの態様では、そのようなサンプルは細胞もしくは組織または他の生体サンプルのホモジネートである。好ましくは、そのようなサンプルは生体サンプルの総mRNA調製物である。いくつかの態様でさらに好ましくは、そのような核酸サンプルは、生体サンプルから

単離された総mRNAである。当業者は、大部分の方法で調製される総mRNAには、成熟mRNAのみならずRNAプロセッシング中間体および新生プレmRNA転写産物も含まれることを認識すると思われる。例えば、ポリ(dT)カラムによって精製された総mRNAは、ポリ(A)テールを持つRNA分子が含まれている。このポリA<sup>+</sup> RNA分子は、mRNA、RNAプロセッシング中間体、新生転写産物、または分解中間体である可能性がある。

【0121】生体サンプルは、任意の生物体由来の任意の生体組織、体液または細胞でよい。しばしば、サンプルは患者から得られた「臨床サンプル」である。臨床サンプルは、遺伝子ネットワークまたは遺伝子発現の種々の状態についての豊かな情報源を提供する。本発明のいくつかの態様は、変異の検出、および変異の表現型の同定に利用される。そのような態様は、臨床診断および臨床研究に、広範囲に適用できる。典型的な臨床サンプルには、痰、血液、血液細胞（例、白血球）、組織、または細針生検サンプル、尿、腹水、胸水、またはそれらの細胞が含まれるが、これに限定されない。生体サンプルには、組織学的目的のために採取した凍結切片またはホルマリン固定切片のような組織切片も含まれ得る。

【0122】生体サンプルの別の典型的な供給源は、遺伝子間の関係を探るために遺伝子発現状態を操作できる培養細胞である。本発明の1つの局面では遺伝子ネットワークの多岐にわたる状態を反映する生体サンプルを作製するための方法が提供される。

【0123】当業者は、ホモジネートをハイブリダイゼーションに使用する前に、ホモジネート中に存在するRNアーゼを阻害または破壊することが望ましいことを認識すると思われる。ヌクレアーゼを阻害または破壊する方法は、当業者に周知である。いくつかの好ましい態様では、ヌクレアーゼを阻害するために、カオトロピック剤の存在下で細胞または組織をホモジナイズする。別の態様では、熱処理後、プロテイナーゼ処理を行って、RNアーゼを阻害または破壊する。

【0124】総mRNAの単離方法も、当業者に周知である。例えば、核酸の単離および精製方法は、P. Tijssen編、Elsevier、ニューヨーク（1993）「生化学と分子生物学の実験室テクニック：核酸プローブとのハイブリダイゼーション、第1部、理論と核酸の調製（Laboratory Techniques in Biochemistry and Molecular Biology: Hybridization With Nucleic Acid Probes, Part I. Theory and Nucleic Acid Preparation）」の第3章に詳述されている。

【0125】好ましい態様では、与えられたサンプルから、例えばAGPC（acid guanidinium-phenol-chloroform）抽出法を用いて総RNAを単離し、オリゴ（dT）カラムクロマトグラフィーまたは磁性ビーズ上の（dT）を用いてポリA<sup>+</sup> mRNAを単離する（例えば、Sambrookら、「分子クローニング：実験室マニュアル（Molecular Cloning 50

g: A Laboratory Manual）第2版」1～3巻、Cold Spring Harbor Laboratory（1989）、またはAusubelら編「分子生物学の最新プロトコール（Current Protocols in Molecular Biology）」Greene PublishingおよびWiley-Interscience、New York（1987）参照）。

【0126】しばしば、ハイブリダイゼーションの前に核酸サンプルを増幅することが望ましい。当業者は、どのような増幅方法を使うにしても、定量的な結果が望ましい場合には、定量的な増幅をするために増幅される核酸の相対的な頻度を維持または制御する方法を、注意深く使う必要があることを認識すると思われる。

【0127】「定量的な」増幅方法は、当業者に周知である。例えば、定量的PCRには、同一のプライマーを使用して、既知の量の対照配列を同時に増幅することが含まれる。これにより、PCR反応の較正に使用できる内部標準が提供される。そのため、高密度アレイには、増幅された核酸の定量のための内部標準に特異的なプローブが含まれる可能性がある。

【0128】好ましい内部標準の1つは、合成AW106 cRNAである。AW106 cRNAを、当業者に公知の標準的なテクニックにしたがってサンプルから単離されたRNAに添加する。その後、RNAを逆転写酵素を用いて逆転写して、コピーDNAを作製する。その後、標識プライマーを用いて、このcDNA配列を増幅する（例、PCRによる）。増幅産物を、通常は電気泳動により分離し、放射活性（増幅産物の量に比例）の量を決定する。その後、既知のAW106 RNA標準の生成するシグナルと比較して、サンプル中のmRNAの量を計算する。定量的PCRの詳細なプロトコールは、「PCRプロトコール、方法と適用の手引き（PCR Protocols, A Guide to Methods and Applications）」Innisら、Academic Press, Inc. ニューヨーク（1990）に記載されている。

【0129】他の適当な増幅方法には、ポリメラーゼ連鎖反応（PCR）（Innisら、「PCRプロトコール、方法と適用の手引き（PCR Protocols, A Guide to Methods and Applications）」Academic Press, Inc. サンディエゴ（1990））、リガーゼ連鎖反応（LCR）（WuおよびWallace, Genomics, 4: 560 (1989)、Landergrenら、Science, 241: 1077 (1988)、およびBarringerら、Gene 89: 117 (1990)参照）、転写増幅（Kwohら、Proc. Natl. Acad. Sci. USA 86: 1173 (1989)参照）、および自己持続性配列複製（Guatelliら、Proc. Natl. Acad. Sci. USA 87: 1864 (1990)）が含まれるが、これに限定されない。

【0130】細胞溶解物または組織ホモジネートには、しばしばポリメラーゼ活性の阻害剤がいくつも含まれている。したがって、通常RT-PCRには、後に増幅錆型として使用するための総RNAまたはmRNAを単離する予備的段階が含まれる。1チューブmRNA補足法を用いて、同じチューブでそのままRT-PCRに使用できるポリ(A)<sup>+</sup> RNAサンプルを調製することができる（Boehringer Mannhei

m)。補足したmRNAは、逆転写反応液と、その後にPCR反応液を添加することによって、直接RT-PCRにかけることができる。

【0131】1つの特に好ましい態様では、逆転写酵素、オリゴ(dT)およびT7ファージプロモーターをコードする配列を含むプライマーを用いてサンプルmRNAを逆転写して、一本鎖DNA鑄型を提供する。第2のDNA鎖は、DNAポリメラーゼを用いて重合させる。二本鎖cDNAの合成後、T7 RNAポリメラーゼを添加して、cDNA鑄型からRNAを転写する。その後の各一本鎖cDNA鑄型からの転写サイクルの繰り返しによって、増幅されたRNAが得られる。インビトロ重合の方法は、当業者に周知であり(例、Sambrook前記参照)、この方法は、本方法にしたがったインビトロ増幅が、種々のRNA転写産物の相対的な頻度を保存することを示したVan Gelderら、Proc.Natl. Acad. Sci. USA 87: 1663-1667 (1990)に詳述されている。さらに、Eberwineら、Proc. Natl. Acad. Sci. USA 89: 3010-3014は、インビトロ転写による2サイクルの増幅を用いて、開始材料を10<sup>6</sup>以上に増幅し、生体サンプルが限られているときにも発現のモニタリングを可能にするプロトコールを提供している。

【0132】(B) 高密度アレイへの核酸のハイブリダイゼーション

### 1. プローブのデザイン

当業者は、本発明の実施には、多数のアレイデザインが適していることを認めると思われる。高密度アレイには、通常、興味のある配列に特異的にハイブリダイズするいくつかのプローブが含まれている。さらに、1つの好ましい態様では、アレイには1つまたは複数の対照プローブが含まれている。

【0133】高密度アレイチップには「試験プローブ」が含まれている。試験プローブは、約5~約45または5~約500ヌクレオチド、さらに好ましくは約10~約40ヌクレオチド、および最も好ましくは約15~約40ヌクレオチドの長さのオリゴヌクレオチドである。特に好ましい別の態様では、プローブの長さは20または25ヌクレオチドである。別の好ましい態様では、試験プローブは二本鎖または一本鎖DNA配列である。DNA配列は、天然の供給源から単離またはクローニングされるか、天然の核酸を鑄型として天然の供給源から増幅される。これらのプローブは、それが発現を検出するように設計された遺伝子の特定のサブ配列に相補的な配列を持つ。したがって、試験プローブは、検出する標的核酸に特異的にハイブリダイズする能力を持つ。

【0134】興味のある標的核酸に結合する試験プローブのほかに、高密度アレイにはいくつかの対照プローブが含まれ得る。対照プローブは、本明細書で1) 正規化対照；2) 発現レベル対照；および3) ミスマッチ対照と呼ばれる3種類に分類される。

【0135】正規化対照は、核酸サンプルに添加される

標識参照オリゴヌクレオチドまたは他の核酸配列に相補的なオリゴヌクレオチドプローブまたは他の核酸プローブである。ハイブリダイゼーション後に正規化対照から得られるシグナルは、完全なハイブリダイゼーションのシグナルをアレイごとに変動させる可能性のある、ハイブリダイゼーション条件、標識の強度、「読み取り」効率、および他の要素の変動の対照を提供する。好ましい態様では、アレイの他のすべてのプローブから読まれるシグナル(例、蛍光強度)を、対照プローブのシグナル(例、蛍光強度)で割り、測定値を正規化する。

【0136】実質上、任意のプローブが正規化対照になり得る。しかし、塩基組成とプローブの長さによってハイブリダイゼーションの効率が変化することが認識されている。好ましい正規化プローブは、アレイ中に存在する他のプローブの平均の長さを反映するように選択されるが、ある範囲の長さをカバーするように選択することもできる。正規化対照は、アレイ中の他のプローブの(平均)塩基組成を反映するように選択できるが、好ましい態様では、1つまたは少数の正規化プローブのみが使用され、これは効率良くハイブリダイズし(すなわち二次構造を取らない)、標的に特異的なプローブのいずれにも合致しないように選択される。

【0137】発現レベル対照は、生体サンプル中で構成的に発現される遺伝子に特異的にハイブリダイズするプローブである。実質上、任意の構成的に発現される遺伝子が、発現レベル対照の適当な標的を提供する。典型的には、発現レベル対照プローブは、アクチン遺伝子、トランスフェリン受容体遺伝子、GAPDH遺伝子などを含むがこれらに限定される訳ではない構成的に発現する「ハウスキーピング遺伝子」のサブ配列に相補的な配列を持つ。

【0138】標的遺伝子のプローブ、発現レベル対照、または正規化対照に対して、ミスマッチ対照も提供されることがある。ミスマッチ対照は、1つまたは複数のミスマッチ塩基が存在するという以外は、対応する試験または対照プローブに同一のオリゴヌクレオチドプローブまたは他の核酸プローブである。ミスマッチする塩基は、プローブが特異的にハイブリダイズするはずの標的配列中の対応する塩基に相補的ではないように、選択された塩基である。適当なハイブリダイゼーション条件下(例、ストリングエントな条件)では、試験プローブまたは対照プローブが標的配列にハイブリダイズするが、ミスマッチプローブはハイブリダイズしない(またはかなり低いレベルでハイブリダイズする)と期待されるよう、1つまたは複数のミスマッチが選択される。好ましいミスマッチプローブには、1つの中心ミスマッチが含まれている。したがって、例えば、プローブが20量体の場合、対応するミスマッチプローブは6から14の位置の任意の位置(中心ミスマッチ)にある単一塩基のミスマッチ(例、AをG、C、またはTが置換)を除いては、同

一の配列を持つ。

【0139】したがって、ミスマッチプローブは、サンプル中の、プローブが対応する標的以外の核酸に対する、非特異的結合または交差ハイブリダイゼーションの対照を提供する。したがって、ミスマッチプローブは、ハイブリダイゼーションが特異的であるか否かを示す。例えば、標的が存在すれば、完全にマッチするプローブは、ミスマッチプローブよりも常に明るくなくてはいけない。また、すべての中心ミスマッチが存在する場合、ミスマッチプローブを用いて変異を検出することができる。完全なマッチとミスマッチプローブとの強度の差(1(PM)-1(MM))は、ハイブリダイズした材料の濃度の良い尺度となる。

【0140】高密度アレイには、サンプルの調製/増幅対照プローブも含まれていることがある。これらは、解析する特定の生体サンプルの核酸には通常存在しないため選択された対照遺伝子のサブ配列に、相補的なプローブである。適当なサンプル調製/増幅対照プローブには、例えば、問題のサンプルが真核細胞の生体サンプルである場合の細菌遺伝子に対するプローブ(例、Bio B)が含まれる。

【0141】処理の前に、サンプルの調製/増幅対照プローブの対応する、既知の量の核酸をRNAサンプルに添加する。サンプルの調製/増幅対照プローブのハイブリダイゼーションを定量すれば、処理段階(例、PCR、逆転写、インビトロ転写等)で生じた核酸量の変化が測定できる。

【0142】好ましい態様において高密度アレイ中のオリゴヌクレオチドプローブは、使用する特定のハイブリダイゼーション条件下では、最低限の非特異的結合または交差ハイブリダイゼーションしかせずに、対応する核酸標的に特異的に結合するように選択される。本発明の高密度アレイには、1,000,000以上もの異なるプローブが含まれるので、特定の核酸配列に結合する、特有の長さのすべてのプローブを提供することが可能である。

【0143】また、好ましい態様では、発現モニタリングアレイを用いて、数百塩基対の長さの遺伝子の存在および発現(転写)レベルが同定される。たいていの用途では、数線から10万遺伝子の有無、または発現レベルを同定することが有用である。好ましい態様では、1つのアレイあたりのオリゴヌクレオチドの数が限定されているので、発現を検出する各遺伝子に特異的な限定されたプローブセットのみを含めることが望ましい。

【0144】米国特許出願第08/772,376号に記述されるように、わずか15、20、または25ヌクレオチドのプローブでも、遺伝子のサブ配列にハイブリダイズするのに十分であり、大部分の遺伝子では、広範囲の標的核酸濃度で良好な働きをするプローブセットが存在する。好ましい態様では、高密度アレイを合成する前に、各遺伝子の好ましいまたは「最適な」プローブサブセットを選択す

ることが望ましい。

【0145】2. ハイブリダイゼーション

核酸のハイブリダイゼーションは、単に、プローブとそれに相補的な標的とが、相補的な塩基対合によって安定なハイブリッド二本鎖分子を形成できる条件下で、プローブと標的の核酸を接触させることを含む。その後、ハイブリッド二本鎖分子を形成しない核酸を洗い流して、ハイブリダイズした核酸を残し、この核酸は通常、結合した検出可能標識を検出することによって検出される。核酸は、これを含む緩衝液の塩濃度を低下させるか温度を上昇させると変性するということは、一般に認識されている。低ストリンジエンシー条件(例、低い温度および/または高塩濃度)では、アニーリングする配列が完全に相補的でなくても、ハイブリッド二本鎖分子(例、DNA:DNA、RNA:RNA、またはRNA:DNA)が形成される。したがって、ストリンジエンシーが低いとハイブリダイゼーションの特異性は低下する。逆に、高ストリンジエンシー条件では(例、高温、または低塩濃度)、ハイブリダイゼーションが起きるためには、ミスマッチが少ない必要がある。

【0146】当業者は、任意の程度のストリンジエンシーを提供するために、ハイブリダイゼーション条件が選択できることを認識すると思われる。好ましい態様では、ハイブリダイゼーションを確実に起こすために、この場合は37で6 X SSPE-T(0.005%トリトンX-100)中で低ストリンジエンシーにてハイブリダイゼーションを行い、より高いストリンジエンシーでその後の洗浄を行い(例、1 X SSPE-T、37)、ミスマッチしたハイブリッド二本鎖分子を除去する。その後の洗浄は、望ましいレベルのハイブリダイゼーション特異性が得られるまで、徐々にさらにストリンジエンシーを高くして行うことができる(例、37~50にて0.25 X SSPE-Tまで下げる)。別の好ましい態様では、1 X MES緩衝液が使用される。この緩衝液は、0.1 M 2-[N-モルホリニオ]エタンスルホン酸、1.0 M NaCl、および0.01%トリトン-X 100(登録商標)を含んでいる。ストリンジエンシーは、ホルムアミドのような薬剤を添加することによって上昇させられる。ハイブリダイゼーション特異性は、試験プローブへのハイブリダイゼーションと、存在し得る種々の対照(例、発現レベル対照、正規化対照、ミスマッチ対照など)へのハイブリダイゼーションとを比較することによって、評価できる。

【0147】一般に、ハイブリダイゼーション特異性(ストリンジエンシー)とシグナル強度とは、同時には満たし得ない。したがって、好ましい態様では、洗浄は一貫した結果が得られ、バックグラウンドよりも約10%以上強いシグナルが得られる最もストリンジエンシーの高い条件で行われる。したがって、好ましい態様では、ハイブリダイズしたアレイは、次第にストリンジエンシーを高くした溶液で順次洗浄しながら、各洗浄の間に測

定をすることができる。このようにして得られたデータセットを分析すると、ハイブリダイゼーションパターンが検出可能な程は変化しないながら、興味のある特定のオリゴヌクレオチドプローブが適当なシグナルを持つような上記の洗浄のストリングエンサーがわかる。

【0148】RNAまたはDNA間に形成された二本鎖分子の安定性は、溶液中では通常、RNA : RNA > RNA : DNA > DNA : DNAの順である。長いプローブは標的と安定した二本鎖分子を形成するが、短いプローブよりはミスマッチの区別が困難になる（ミスマッチの区別とは、完全にマッチするプローブと1塩基のミスマッチのあるプローブとの間に測定されるハイブリダイゼーションシグナル比である）。短いプローブ（例、8量体）はミスマッチをはっきりと区別することができるが、二本鎖分子の全体の安定性は低くなる。

【0149】標的とプローブとの間に形成される二本鎖分子の熱安静性（T<sub>m</sub>）を、例えば既知のオリゴヌクレオチド類似体などを用いて変化させると、二本鎖分子の安定性とミスマッチの区別の最適化が可能になる。T<sub>m</sub>の変化の有用な1つの局面は、アデニン-チミン（A-T）二本鎖分子は1塩基対あたり2本の水素結合を持つが、グアニン-シトシン（G-C）二本鎖分子は1塩基対あたり3つの水素結合を持つという理由にもよって、A-T二本鎖分子がG-C二本鎖分子よりも低いT<sub>m</sub>を持つということに由来する。塩基が不均一に分布している異種オリゴヌクレオチドアレイでは、各オリゴヌクレオチドプローブのハイブリダイゼーションを同時に最適化することは、通常不可能である。したがって、いくつかの態様では、G-C二本鎖分子を選択的に不安定化する、および/またはA-T二本鎖分子の安定性を上昇させることができ、望ましい。これは、例えば、アレイ中のプローブでG-C二本鎖分子を形成するグアニン残基をヒポキサンチンで置換するか、プローブ中でA-T二本鎖分子を形成するアデニン残基を2,6ジアミノプリンで置換するか、またはNaClの代わりに塩化テトラメチルアンモニウム（TMACl）を使用することによって、実行できる。

【0150】オリゴヌクレオチド類似体プローブを用いて得られる二本鎖分子の安定性の変化は、例えば、標的オリゴヌクレオチドとハイブリダイズしたオリゴヌクレオチド類似体アレイの蛍光シグナル強度を時間を追って追跡すると、確認できる。このデータを用いると、例えば、室温（将来の、単純化した診断用途のため）における特定のハイブリダイゼーション条件の最適化が可能になる。

【0151】二本鎖分子の安定性の変化を確認する別の方法は、ハイブリダイゼーションで生成されるシグナル強度を時間を追って追跡することである。DNA標的およびDNAチップを用いた以前の実験では、シグナル強度は時間が経つと増加し、安定度の高い二本鎖分子の方が、安定度の低い二本鎖分子と比べて、シグナル強度の増加

が速く起こることが示された。シグナルは、すべての結合部位が占領されたために一定時間後にはプラトーに達する、または「飽和する」。これらのデータを用いて、ハイブリダイゼーションの最適化、および特定の温度での最適条件を決定することができる。

【0152】ハイブリダイゼーション条件の最適化の方法は当技術分野で周知である（「生化学と分子生物学の実験室テクニック（Laboratory Techniques in Biochemistry and Molecular Biology）」24巻「核酸プローブとのハイブリダイゼーション（Hybridization With Nucleic Acid Probes）」P. Tijssen編、Elsevier、ニューヨーク（1993）参照）。

#### 【0153】(C) シグナルの検出

好ましい態様では、ハイブリダイズした核酸の検出は、サンプル核酸に結合した1つまたは複数の標識を検出することによって行う。標識は、当技術分野で周知のいくつかの方法のいずれを使っても組み込むことができる。しかし、好ましい態様では、標識はサンプル核酸の調製時の増幅段階に、同時に組み込まれる。したがって、例えば、標識されたプライマーまたは標識されたヌクレオチドを用いたポリメラーゼ連鎖反応（PCR）は、標識された増幅産物を提供する。好ましい態様では、上述のような、標識ヌクレオチド（例、蛍光標識UTPおよび/またはCTP）を用いた転写増幅で、標識が転写核酸に取り込まれる。

【0154】または、標識は元の核酸サンプル（例、mRNA、ポリA mRNA、cDNAなど）、または増幅終了後に増幅産物に、直接添加することもできる。核酸に標識を結合する方法は、当技術分野で周知で、例えば、ニックトランスレーション、または核酸をリン酸化した後に核酸リンカーニーを結合（連結）して、サンプル核酸を標識（例、蛍光体）に結合する末端標識（例、標識RNA使用）などが含まれる。

【0155】本発明に使用が適した検出可能標識には、分光的、光化学的、生化学的、免疫化学的、電気的、光学的、または化学的手法で検出可能な任意の組成物が含まれる。本発明で有用な標識には、標識ストレプトアビジン結合体による染色のためのビオチン、磁性ビーズ（例、Dynabeads（登録商標））、蛍光色素（例、フルオレセイン、テキサスレッド、ローダミン、緑色蛍光蛋白質（GFP）など）、放射性標識（例、<sup>3</sup>H、<sup>125</sup>I、<sup>35</sup>S、<sup>14</sup>C、または<sup>32</sup>P）、酵素（例、セイヨウワサビペルオキシダーゼ、アルカリホスファターゼ、およびELISAで通常使用される他の酵素）、およびコロイド金もしくは着色ガラスのような比色標識、またはプラスチック（例、ポリスチレン、ポリプロピレン、ラテックス等）のビーズが含まれる。そのような標識の使用を開示している特許には、米国特許第3,817,837号；第3,850,752号；第3,939,350号；第3,996,345号；第4,277,437号；第4,275,149号；および第4,366,241号が含まれる。

【0156】そのような標識の検出方法は、当業者には周知である。したがって、例えば、放射性標識は写真フィルムまたはシンチレーションカウンターを用いて検出でき、蛍光マーカーは、光検出器を用いて発光を検出することによって検出できる。酵素標識の検出は、通常、酵素に基質を提供し、基質に対する酵素の作用によって生成する反応産物を検出して行い、比色標識は着色した標識を基に可視化することにより検出される。1つの特定の好ましい方法は、散乱光を測定することで検出できるコロイド金標識を用いるものである。

【0157】標識は、ハイブリダイゼーションの前または後に標的（サンプル）核酸に添加できる。いわゆる「直接標識」は、ハイブリダイゼーションの前に標的（サンプル）核酸に直接に結合または組み込まれた検出可能標識である。反対に、いわゆる「間接標識」は、ハイブリダイゼーション後にハイブリッド二本鎖分子に結合される。しばしば、間接標識はハイブリダイゼーションの前に標的核酸に結合された結合性成分に結合される。したがって、例えば、ハイブリダイゼーション前に、標的核酸をビオチン化することができる。ハイブリダイゼーション後に、アビジンの結合した蛍光体はビオチンを持つハイブリッド二本鎖分子に結合するため、簡単に検出できる標識が得られる。核酸の標識およびハイブリダイズした標識核酸の検出の詳細な総説は、「生化学と分子生物学の実験室テクニック(Laboratory Techniques in Biochemistry and Molecular Biology)」24巻「核酸プローブとのハイブリダイゼーション(Hybridization With Nucleic Acid Probes)」P. Tijssen編、Elsevier、ニューヨーク(1993)を参照のこと。

【0158】蛍光標識は好ましく、インピトロ転写反応中に簡単に添加することができる。好ましい態様では、上述のようにインピトロ転写反応で生成するRNAに、フルオレセイン標識UTPおよびCTPを組み込ませる。

【0159】高密度アレイのプローブにハイブリダイズした標識された標的（サンプル）核酸の検出方法は、当技術分野で公知である。したがって、例えば、比色標識を使用する場合には、標識を単に可視化すれば十分である。放射性標識プローブが使用される場合には、放射線の検出（例、写真フィルムまたは固体検出器による）で十分である。

【0160】しかし、好ましい態様では、標的核酸は蛍光標識され、プローブアレイ上の標識の局在は、蛍光顕微鏡で検出される。その特定の蛍光標識の励起波長の光源でハイブリダイズしたを励起し、得られた発光波長の蛍光を検出する。特に好ましい態様では、励起光源は蛍光標識の励起に適当なレーザーである。

【0161】共焦点顕微鏡を、コンピュータ制御ステージと共に自動化し、高密度アレイ全体を自動的にスキャンすることができる。同様に、自動データ収集システムに接続したフォトトランスデューサー（例、光電子増倍

管、固体アレイ、CCDカメラなど）を顕微鏡に取り付け、アレイ上の各オリゴヌクレオチドプローブへのハイブリダイゼーションによって生成した蛍光シグナルを自動的に記録するようになる。そのような自動化システムは、米国特許第5,143,854号、PCT出願20 92/10092号、および1994年2月10日出願の同時係属中の米国特許出願第08/195,889号に詳しく説明されている。自動化共焦点顕微鏡と合わせてレーザー照射を使用すると、約10  $\mu\text{m}$ 以上、さらに好ましくは約50  $\mu\text{m}$ 以上、および最も好ましくは約25  $\mu\text{m}$ 以上の解像度で検出が可能になる。

【0162】当業者は、ハイブリダイゼーション結果の評価方法は、使用した特定のプローブ核酸および提供される対照の性質によって異なることを認識すると思われる。最も単純な態様では、各プローブの蛍光強度が単純に定量される。これは、高密度アレイ上の各位置（異なるプローブを表す）におけるプローブシグナルの強度を測定するだけで実施できる（例、標識が蛍光標識の場合、アレイ上の各位置における固定した励起照射によって生成する蛍光の量（強度）を検出する）。「試験」サンプルの核酸にハイブリダイズしたアレイの絶対強度を、「対照」サンプルが生成する強度と比較すると、各プローブにハイブリダイズした核酸の相対的発現が測定できる。

【0163】しかし、当業者は、ハイブリダイゼーションシグナルの強度は、ハイブリダイゼーション効率、サンプル核酸上の標識の量、およびサンプル中の特定の核酸の量によって、変わることを認識すると思われる。通常、非常に低レベル（例、<1pM）で存在する核酸は、非常に弱いシグナルを発する。ある低濃度では、シグナルはバックグラウンドとほぼ区別が不可能になる。ハイブリダイゼーションデータを評価する際には、それ以下のシグナルがバックグラウンドと基本的に区別不能であるとする閾値強度を選択することができる。

【0164】低レベルの核酸発現を検出することが望ましい場合には、低い閾値を選択する。逆に、高レベルの発現のみを評価する場合には、高い閾値レベルを選択する。好ましい態様では、適切な閾値は平均バックグラウンドシグナルを約10%上回るレベルである。

【0165】また、適当な対照を提供すると、ハイブリダイゼーション条件、細胞の健康状態、非特異的結合などの変動要因を抑えた、さらに詳しい分析が可能になる。したがって、例えば、好ましい態様では、ハイブリダイゼーションアレイには正規化対照を提供する。正規化対照は、サンプルに既知の濃度で加えた対照配列に相補的なプローブである。全体のハイブリダイゼーション条件が良くない場合には、正規化対照はハイブリダイゼーションの低下を反映して、より弱いシグナルを示す。逆に、ハイブリダイゼーション条件が良い場合には、正規化対照はハイブリダイゼーションの向上を反映して、より高いシグナルを提供する。アレイの他のプローブの

シグナルを正規化対照に対して正規化すると、ハイブリダイゼーション条件の変動の対照が提供される。通常、正規化はアレイ上の他のプローブの測定されたシグナルを、正規化対照の生成する平均シグナルで除することを行われる。正規化には、サンプル調製および増幅による変動の修正も含まれ得る。そのような正規化は、サンプル調製/増幅対照プローブ(例、Bio Bプローブ)の平均シグナルで、測定シグナルを除することによって実行できる。得られた値に定数を乗じて、結果を概算することができる。

【0166】上述のように、高密度アレイには、ミスマッチ対照も含まれ得る。好ましい態様では、アレイの各プローブ(正規化対照を除く)の中心ミスマッチを持つミスマッチ対照がある。ストリンジエントな条件で洗浄した後、完全な対合はプローブにハイブリダイズするが、ミスマッチはハイブリダイズしないと予測されるので、ミスマッチ対照のシグナルは、非特異的結合のみ、またはミスマッチにハイブリダイズする核酸がサンプル中に存在することを示すはずである。問題のプローブおよび対応するミスマッチ対照のいずれもが高いシグナルを示す場合、またはミスマッチが対応する試験プローブよりも高いシグナルを示す場合には、ハイブリダイゼーションに問題があるので、これらのプローブのシグナルは無視される。標的特異的プローブと、それに対応するミスマッチ対照の間のハイブリダイゼーションシグナルの強度の差は、標的特異的プローブの区別の尺度である。したがって、好ましい態様では、ミスマッチプローブのシグナルを、対応する試験プローブのシグナルから差し引くと、試験プローブの特異的な結合に起因するシグナルの量が得られる。

【0167】その後、遺伝子に特異的に結合するプローブの各々のシグナル強度を測定し、正規化対照に対して正規化することによって、特定の配列の濃度が決定できる。プローブのシグナルがミスマッチよりも大きい場合には、ミスマッチを差し引く。ミスマッチの強度が対応する試験プローブと等しいかそれを上回る場合には、シグナルを無視する。その後、特定の遺伝子の発現レベルは、陽性シグナルの数(絶対値または閾値を超える部分のいずれか)、陽性シグナルの強度(絶対値または選択された閾値を超える部分のいずれか)、またはその両方の組みあわせ(例、加重平均)で計算できる。

【0168】好ましい態様では、コンピュータシステムを用いて、各ペアの完全なマッチとミスマッチプローブのハイブリダイゼーション強度を比較する。その遺伝子が発現しているならば、ペアの完全マッチプローブのハイブリダイゼーション強度(または親和性)は、対応するミスマッチプローブよりも、高いことが認識できるはずである。一般に、一対のプローブのハイブリダイゼーション強度が実質的に同一ならば、その遺伝子が発現されていないことを示す可能性がある。しかし、遺伝子が

発現しているかどうかの決定は、一対のプローブのみに基づくのではなく、多くの対のプローブの分析に基づいて行う。

【0169】システムは、完全マッチプローブとミスマッチプローブのハイブリダイゼーション強度を比較した後、遺伝子の発現を示す。例えば、システムはユーザに対して、遺伝子が存在する(発現されている)、下限に近い、または存在しない(発現されていない)のいずれかを示す。データ分析の具体的な手順は、すべての目的のため以前に本明細書に組み入れられた米国特許出願第08/772,376号に開示されている。

【0170】高密度核酸アレイの他に、他の方法も遺伝子発現の大規模モニタリングに有用である。Liang, P. およびPardee, A.B. (「ポリメラーゼ連鎖反応による真核生物のメッセンジャーRNAのディファレンシャルディスプレイ(Differential Dispolay of eukaryotic messenger RNA by means of the polymerase chain reaction.)」 *Science* 257:967-971, 1992、これはすべての目的のために参照として本明細書に組み入れられる)によって記述されたディファレンシャルディスプレイにより、2つのサンプル間の遺伝子発現を区別するために有用な方法が提供される。Velculescuら(「遺伝子発現の連続分析(Serial Analysis of Gene Expression.)」 *Science*, 270:484-487, 1995、これはすべての目的のために参照として本明細書に組み入れられる)によって記述された遺伝子発現の連続分析により、遺伝子発現の定量および定性分析のための別の方法が提供される。Fergusonら(「遺伝子発現分析のための光ファイバーDNAバイオセンサーマイクロアレイ(A Fiber-optic DNA biosensor microarray for the analysis of gene expression.)」 *Nature-Biotechnology* 14:1681-1684, 1996)によって記述された光ファイバーオリゴヌクレオチドセンサーも、遺伝子発現のモニタリングに使用できる。

【0171】本明細書中の実施例および態様は例示のためのみに記載されており、当業者にはそれらに照らし合わせた種々の修正または変更が示唆され、これらは本出願の精神および範囲内であり、添付の請求の範囲内であることが理解される。本明細書に引用されるすべての出版物、特許、および特許出願は、すべての目的のために参考として本明細書に組み入れられる。

【0172】

【実施例】WT1およびEWS-WT1の機能的性質を探り、内因性標的遺伝子に対するその効果を調べるために、本発明者らは厳密に調節された誘導可能プロモーターにより全長蛋白質が発現される細胞を開発した。これらの細胞内でWT1およびEWS-WT1を強制発現すると、アポトーシスが誘導された。オリゴヌクレオチドアレイに基づく発現の概略分析によって、WT1およびEWS-WT1の誘導後に、少数の遺伝子の発現が変化したことが示された。

【0173】WT1およびEWS-WT1の機能的性質を調べるた

めに、本発明者らは骨肉腫細胞U2OS中でWT1およびEWS-WT1の誘導可能なテトラサイクリン調節発現系を確立した (GossenおよびBujard, 1992)。本発明者らは、WT1およびEWS-WT1の誘導発現後に発現レベルが変化する可能性のある内因性遺伝子を探した。WT1およびEWS-WT1誘導の12時間後に細胞からメッセンジャーRNAを単離し、ビオチン化し、6,800の既知の転写産物および発現された配列タグを持つ高密度オリゴヌクレオチドアレイ (Lockhartら、1996; Wodickaら、1997) にハイブリダイズさせた。遺伝子は、WT1およびEWS-WT1誘導後に少なくとも2回の実験の繰り返しで再現的に誘導または抑制され、少なくとも3倍の発現量の変化が見られれば、その遺伝子はWT1およびEWS-WT1標的の候補と見なされた。

【0174】このいわゆる「テト・オフ (tet-off)」誘導系では、組織培養培地からテトラサイクリンを除去して、組換え遺伝子を誘導する。この実験の目的は、これらの癌抑制因子の機能的仲介因子の候補 (CFM) を同定することであった。WT1またはEWS-WT1発現の誘導後、再現的に発現が3倍またはそれ以上に変化する内因性遺伝子が、(モニターした7000の遺伝子の内) それぞれ28個および17個同定された。

【0175】WT1およびEWS-WT1の誘導発現を厳密に調節した細胞株を作製して、本発明者らは癌抑制因子としての機能に貢献すると考えられる下流経路を同定した。

【0176】発現プロファイルの作製のためのハイブリダイゼーションに基づくアッセイ法  
何10万もの化学合成したオリゴヌクレオチドを含むアレイのセットに全mRNA集団をハイブリダイズさせることにより、mRNAレベルが決定される。オリゴヌクレオチドは、光指向性 (light-directed) 固相コンビナトリアルケミストリーを用いて、ガラス支持体上にてインサイチューで合成される。アレイは配列情報のみに基づいて設計および合成されるため、ゲノム配列と遺伝子発現の差の尺度とを直接結び付けるものとなる。アレイは1.28 x 1.28 cmで測定され、50 x 50ミクロンの合成物を65,000個以上含む。各合成物は特定の25量体オリゴヌクレオ\*

\*チドを10<sup>7</sup>以上のコピー数含む。

【0177】各々のmRNAサンプルにつき、6,800の全長ヒト遺伝子の発現レベルがモニターされる。これらの遺伝子をカバーするオリゴヌクレオチド全体のセットを、合計約1平方インチの面積を含む4つのアレイに分割する。各遺伝子につき、自動化した選択基準に基づいて、約20の相補的25量体が選択される。この基準には、ゲノムの他の部分と比較した配列の独自性、およびアレイ上のハイブリダイゼーションに悪影響を与えると決定された特徴的配列 (例、自己相補性または単一ヌクレオチドの集団) が存在しないことに関する試験が含まれる。各遺伝子にオリゴヌクレオチドのセットを使用すると、検出およびデータ分析に冗長性が得られ、時折発生する交差ハイブリダイゼーションによる混乱の可能性を低減するため、定量的情報を得るために、すべてのオリゴヌクレオチドが同一レベルでハイブリダイズする必要がなくなる。検出の感度および特異性をさらに向上させるために、各相補的オリゴヌクレオチド (完全なマッチ、またはPM) には、物理的に隣接した位置に、密接に関連したミスマッチ (MM) パートナーを合成して配置する。ミスマッチパートナーは、25量体の中央の位置の単一の塩基以外は、同一のものである。各対のMMオリゴヌクレオチドは、均一なハイブリダイゼーションパターン (PMシグナルが対応するMMシグナルよりも強いというパターン) を認識するための内部対照の役割を果たす。定量画像分析は、PMとMMパートナーとの平均の差に基づくもので、各遺伝子のオリゴヌクレオチド対のセットを通じて、非特異的なバックグラウンドの影響は相殺され、特異的なハイブリダイゼーションシグナルが建設的に加算される傾向がある。このようなハイブリダイゼーションシグナルは、1:300,000から1:300まで、3桁にわたって定量的である。

【0178】

【発明の効果】本発明により、WT1遺伝子に関する癌の診断、薬物スクリーニング、および変異の機能分析方法が提供された。

## フロントページの続き

(51) Int.CI.<sup>7</sup> 識別記号  
G 0 1 N 33/53  
33/566  
33/574  
//(C 1 2 N 5/10  
C 1 2 R 1:91)

F I	テ-マコ-ト <sup>®</sup> (参考)
G 0 1 N 33/574	D
C 1 2 N 5/00	B
15/00	A

(72)発明者 トルオン ピビ  
アメリカ合衆国 カリフォルニア州 サン  
ホセ キンドラ ヒル ドライブ 7082

(72)発明者 ヘイバー ダニエル  
アメリカ合衆国 マサチューセッツ州 チ  
エストナット ヒル モナドンク ロード  
34

(72)発明者 リー シーン  
アメリカ合衆国 マサチューセッツ州 マ  
ルデン ダニエルズ ストリート 30 ア  
パートメント 101

## 【外国語明細書】

## 1 Title of Invention

***DOWNSTREAM GENES OF TUMOR  
SUPPRESSOR WT1***

## 2 Claims

1. A method for detecting a WT1 gene functional mutation in target cells comprising the steps of:  
detecting expression of at least three down-stream genes of WT1 in a sample of (a) target cells, and (b) reference cells having a wild-type WT1 gene, the reference cells are otherwise substantially similar to the target cells, the down-stream genes are up- or down-regulated by the wild-type WT1 gene; and  
comparing the expression of the down-stream genes in the target cells and the reference cells, wherein a difference in the expression between the target cells and reference cells suggests a WT1 functional mutation in the target cells.
2. The method of claim 1, wherein said down-stream genes are transcriptionally regulated by said wild-type WT1 gene and the expression of said down-stream genes is detected by measuring amounts of transcripts of said down-stream genes in said reference and target cells.
3. The method of claim 1, wherein said amounts of transcripts are measured with high-density nucleic acid array.

4. The method of claim 1, wherein said down-stream genes comprise at least one WT1 up-regulated gene selected from the group consisting of: natural killer cells protein 4 precursor (M59807), folate binding protein (M25317), HALPHA44 gene for alpha-tubulin (X06956), heat shock protein HSP70B (X51758), 90K product (H17969), heat shock 70kd protein 1 (T66307), amphiregulin (M30704), procollagen alpha 1 (T51558), beta-migrating plasminogen activator inhibitor 1 (M14083), cysteine protease inhibitor from radiated keratinocytes (X05978), brain natriuretic protein (BNP) (M31766), leukemia virus receptor 1 (GLVR1) (L20859), type 1 cytoskeletal 17 keratin (R71870), tubulin alpha-5 chain (H45051), purine nucleoside phosphorylase (T47964), GTPase (U10550), adrenomedullin (D14874), GPI-anchored urokinase plasminogen activator surface receptor (R38636), tissue type plasminogen activator (X13097), gravin (M96322), jun-B (X51345), elongation factor 1 alpha-2 (X79490), homeotic gene regulator (R16977), clone 9112 (X57348), interferon gamma treatment inducible mRNA (M26683), transmembrane receptor protein (Z17227), acidic FGF (X65779), cartilage linking protein-1 (X17405), IL-11 (X58377) and mitochondrial phosphate carrier protein (R49231).
5. The method of claim 1 wherein said down-stream genes comprise at least one WT1 down-regulated gene selected from the group consisting of: caveolin and elongation factor 1 alpha-2.
6. The method of claim 4 wherein said down-stream genes further comprise p21 or EGFR.
7. The method of claim 1 further comprising the step of: indicating a loss of function mutation in the WT1 gene in the target cells if the expression of said WT1 up-regulated genes is at least two times less in said target cells than in said reference cells or if the expression of said WT1 down-regulated genes is at least two times more in said target cells than in said reference cells.

8. A method for detecting a EWS-WT1 gene fusion in target cells comprising the steps of:  
detecting expression of one or more down-stream genes of an EWS-WT1 fusion in a sample of (a) target cells, and (b) reference cells having a wild-type EWS and WT1 gene, said reference cells being otherwise substantially similar to said target cells, said down-stream genes being up- or down-regulated by said wild-type EWS and WT1 gene; and  
comparing the expression of said down-stream genes in said target cells and said reference cells, wherein a difference in said expression between the target cells and reference cells suggests a EWS-WT1 fusion in the target cells.
9. The method of claim 8, wherein said down-stream genes comprise at least one EWS-WT1 up-regulated gene selected from the group consisting of: Pseudorabies virus nuclear antigen (R60906 and T79475), beta chain of interleukin-2 receptor (M26062), G1/S-specific cyclin D1 (H20529), neuron-specific gamma-2 enolase (M22349), fusion gene NPM-MLF1 resulting from translocation t(3;5)(q25.1;p34) (L49054), collagen alpha 3(IV) chain (H62466), T-lymphocyte specific protein tyrosine kinase p56lck(lck) (U23852), insulin-like growth factor binding protein-4 (U20982), cathepsin B (L16510), cystatin C (T51534), dual specificity mitogen-activated protein kinase kinase 2 (R42291), neuronal-glial cell adhesion molecule (T53118), PIM-1 proto-oncogene serine/threonine-protein kinase (H10925), carboxylesterase precursor (R54359), protein-tyrosine phosphatase PAC-1 (L11329), and interferon-induced protein 6-16 (X02492).
10. An in-cell functional assay for a WT1 sequence alteration comprising the steps of:  
detecting expression of at least three down-stream genes in a target sample from target cells having a WT1 sequence alteration and in a reference sample from reference cells having a wild-type WT1 gene, said reference cells being otherwise substantially similar to said target cells, said down-stream genes being up- or down-regulated by said wild-type WT1 gene; and  
comparing said expression in said target sample to said expression in said reference sample, wherein a difference in the expression between said two samples suggests that said WT1 sequence alteration affects the biological function of WT1.

11. The method of claim 10, wherein said down-stream genes are transcriptionally regulated by said wild-type WT1 gene and the expression of said down-stream genes is detected by measuring amounts of transcripts of said down-stream genes in said reference and target cells.
12. The method of claim 11, wherein said amounts of transcripts are measured with a high-density nucleic acid array.
13. The method of claim 10, wherein said down-stream genes comprise at least one WT1 up-regulated gene selected from the group consisting of: natural killer cells protein 4 precursor (M59807), folate binding protein (M25317), HALPHA44 gene for alpha-tubulin (X06956), heat shock protein HSP70B (X51758), 90K product (H17969), heat shock 70kd protein 1 (T66307), amphiregulin (M30704), procollagen alpha 1 (T51558), beta-migrating plasminogen activator inhibitor 1 (M14083), cysteine protease inhibitor from radiated keratinocytes (X05978), brain natriuretic protein (BNP) (M31766), leukemia virus receptor 1 (GLVR1) (L20859), type 1 cytoskeletal 17 keratin, tubulin alpha-5 chain (H45051), purine nucleoside phosphorylase (T47964), GTPase (U10550), adrenomedullin (D14874), GPI-anchored urokinase plasminogen activator surface receptor (R38636), tissue type plasminogen activator (X13097), gravin (M96322), jun-B (X51345), homeotic gene regulator (R16977), clone 9112 (X57348), interferon gamma treatment inducible mRNA (M26683), transmembrane receptor protein (Z17227), acidic FGF (X65779), cartilage linking protein-1 (X17405), IL-11 (X58377) and mitochondrial phosphate carrier protein (R49231).
14. The method of claim 10 wherein said down-stream genes comprise at least one WT1 down-regulated gene selected from the group consisting of caveolin and elongation factor 1 alpha-2 (X79490).
15. The method of claim 13 wherein said down-stream genes further comprise p21 or EGFR.

16. The method of claim 10 further comprising the steps of:

indicating that said WT1 sequence alteration is a loss of wild-type function mutation if the expression of said WT1 up regulated genes is at least two times less in said target cells than in said reference cells or if the expression of said WT1 down regulated genes is at least two times more in said target cells than in said reference cells.

17. A method for detecting a mutation in a target WT1 gene using a computer comprising:

inputting into a computer wild-type expression data of at least three down-stream genes in a wild-type sample containing a wild-type WT1 gene, said down-stream genes being transcriptionally regulated by said wild-type WT1 gene;

inputting into a computer target expression data of said plurality of down-stream genes in a target sample containing said target WT1 gene;

comparing the target and wild-type expression data to determine differences, wherein differences suggest a mutation in said target WT1 gene.

18. The method of claim 17, wherein said down-stream genes comprise

at least one WT1 up-regulated gene selected from the group consisting of: natural killer cells protein 4 precursor (M59807), folate binding protein (M25317), HALPHA44 gene for alpha-tubulin (X06956), heat shock protein HSP70B (X51758), 90K product (H17969), heat shock 70kd protein 1 (T66307), amphiregulin (M30704), procollagen alpha 1 (T51558), beta-migrating plasminogen activator inhibitor 1 (M14083), cysteine protease inhibitor from radiated keratinocytes (X05978), brain natriuretic protein (BNP) (M31766), leukemia virus receptor 1 (GLVR1) (L20859), type 1 cytoskeletal 17 keratin, tubulin alpha-5 chain (H45051), purine nucleoside phosphorylase (T47964), GTPase (U10550), adrenomedullin (D14874), GPI-anchored urokinase plasminogen activator surface receptor (R38636), tissue type

plasminogen activator (X13097), gravin (M96322), jun-B (X51345), homeotic gene regulator (R16977), clone 9112 (X57348), interferon gamma treatment inducible mRNA (M26683), transmembrane receptor protein (Z17227), acidic FGF (X65779), cartilage linking protein-1 (X17405), IL-11 (X58377) and mitochondrial phosphate carrier protein H (R49231).

19. The method of claim 17 wherein said down-stream genes comprise at least one WT1 down-regulated gene selected from the group consisting of caveolin and elongation factor 1 alpha-2 (X79490).
20. The method of claim 18 wherein said down-stream genes further comprise p21 or EGFR.
21. The method of claim 17 further comprising the step of: indicating a mutation in said target WT1 gene if the expression of said WT1 up-regulated genes is at least two times less in said target cells than in said reference cells or if the expression of said WT1 down-regulated genes is at least two times more in said target cells than in said reference cells.
22. A method for detecting a translocation fusing EWS and WT1 genes using a computer comprising:
  - inputting into a computer wild-type expression data of a plurality of down-stream genes in a wild-type sample containing wild-type EWS and WT1 genes, said down-stream genes being transcriptionally regulated by a EWS-WT1 fusion protein;
  - inputting into a computer target expression data of said plurality of down-stream genes in a target sample being tested for the presence of a EWS-WT1 fusion protein;
  - comparing the target and wild-type expression data to determine differences, wherein differences suggest a translocation fusing the EWS and WT1 genes.

23. The method of claim 22, wherein said down-stream genes comprise at least one EWS-WT1 up-regulated gene selected from the group consisting of: Pseudorabies virus nuclear antigen (R60906 and T79475), beta chain of interleukin-2 receptor (M26062), G1/S-specific cyclin D1 (H20529), neuron-specific gamma-2 enolase (M22349), fusion gene NPM-MLF1 resulting from translocation t(3;5)(q25.1;p34) (L49054), collagen alpha 3(IV) chain (H62466), T-lymphocyte specific protein tyrosine kinase p56lck(lck) (U23852), insulin-like growth factor binding protein-4 (U20982), cathepsin B, cystatin C (T51534), dual specificity mitogen-activated protein kinase kinase 2 (R42291), neuronal-glial cell adhesion molecule (T53118), PIM-1 proto-oncogene serine/threonine-protein kinase (H10925), carboxylesterase precursor (R54359), protein-tyrosine phosphatase PAC-1 (L11329), and interferon-induced protein 6-16 (X02492).
24. A method of diagnosing neoplasia of a test cell comprising: hybridizing a transcription indicator of a test cell to a set of nucleic acid probes, wherein the transcription indicator is selected from the group consisting of mRNA, cDNA and cRNA, wherein the set of nucleic acid probes comprises a plurality of nucleic acid molecules each of which is a portion of a gene which is activated by or repressed by WT1 selected from the group consisting of: natural killer cells protein 4 precursor (M59807), heat shock protein HSP70B (X51758), 90K product (H17969), heat shock 70kd protein 1 (T66307), procollagen alpha 1 (T51558), cysteine protease inhibitor from radiated keratinocytes (X05978), brain natriuretic protein (BNP) (M31766), leukemia virus receptor 1 (GLVR1) (L20859), type I cytoskeletal 17 keratin (R71870), purine nucleoside phosphorylase (T47964), adrenomedullin (D14874), gravin (M96322), jun-B (X51345), elongation factor 1 alpha-2 (X79490), homeotic gene regulator (R16977), clone 9112 (X57348), interferon gamma treatment inducible mRNA (M26683), transmembrane receptor protein (Z17227), acidic FGF (X65779), cartilage linking protein-1(X17405), IL-11 (X58377) and mitochondrial phosphate carrier protein (R49231); detecting amounts of transcription indicator which hybridize to each of said set of nucleic acid probes;

identifying a test cell as neoplastic if (1) hybridization of the transcription indicator of the test cell to a probe which is a WT1-activated gene is lower than hybridization using a transcription indicator from a normal cell, or (2) hybridization of the transcription indicator of the test cell to a probe which is a WT1-repressed gene is higher than hybridization using a transcription indicator from a normal cell.

25. The method of claim 24 wherein the test cell is a kidney cell.
26. The method of claim 24 wherein at least 4 of said probes comprise portions of genes which are WT1-activated or WT1-repressed.
27. The method of claim 24 wherein at least 10 of said probes comprise portions of genes which are WT1-activated or WT1-repressed.
28. The method of claim 24 wherein at least 15 of said probes comprise portions of genes which are WT1-activated or WT1-repressed.
29. The method of claim 24 wherein the nucleic acid probes are attached to a solid support.
30. The method of claim 24 wherein the nucleic acid probes are arranged in an array.
31. The method of claim 30 wherein the array comprises nucleic acid probes which are portions of at least 250 different genes.
32. The method of claim 30 wherein the array comprises nucleic acid probes which are portions of at least 6000 different genes.
33. The method of claim 24 further comprising the step of:  
determining the sequence of WT1 genes in the test cell to determine the WT1 genotypic status of the cell.
34. The method of claim 24 wherein a test cell is identified as neoplastic if hybridization is at least 2-fold different between compared samples.

35. A method of identifying anti-cancer drugs, comprising the step of:  
contacting a test compound with a human cell;  
determining the expression of the cell of at least one WT1-up  
regulated gene or at least one WT1-down regulated gene, wherein the WT1  
up-regulated genes are selected from the group consisting of: natural killer  
cells protein 4 precursor (M59807), heat shock protein HSP70B (X51758),  
90K product (H17969), heat shock 70kd protein 1 (T66307), procollagen  
alpha 1 (T51558), cysteine protease inhibitor from radiated keratinocytes  
(X05978), brain natriuretic protein (BNP) (M31766), leukemia virus  
receptor 1 (GLVR1) (L20859), type 1 cytoskeletal 17 keratin, purine  
nucleoside phosphorylase (T47964), adrenomedullin (D14874), gravin  
(M96322), jun-B (X51345), homeotic gene regulator (R16977), clone 9112  
(X57348), interferon gamma treatment inducible mRNA (M26683),  
transmembrane receptor protein (Z17227), acidic FGF (X65779), cartilage  
linking protein-1 (X17405), IL-11 (X58377) and mitochondrial phosphate  
carrier protein (R49231), wherein the WT1-down regulated gene is  
elongation factor 1 alpha-2 (X70940); and  
identifying a test compound as a potential anti-cancer drug if it increases  
expression of at least one WT1 up-regulated gene or decreases expression of  
a WT1 down-regulated gene in the human cell.
36. The method of claim 35 wherein the cell is a kidney cancer cell.
37. The method of claim 35 wherein the cell is a tumor cell.
38. The method of claim 35 wherein the cell contains no wild-type WT1  
genes.
39. The method of claim 35, the expression of said WT1 up-regulated  
and down-regulated genes is detected by measuring amounts of  
transcripts of said genes in said cell.
40. The method of claim 39, wherein said amounts of transcripts are  
measured using a high-density nucleic acid array.
41. A method of diagnosing neoplasia of a test cell comprising:

- 10 -

hybridizing a transcription indicator of a test cell to a set of nucleic acid probes, wherein the transcription indicator is selected from the group consisting of mRNA, cDNA and cRNA, wherein the set of nucleic acid probes comprises a plurality of nucleic acid molecules each of which is a portion of a gene which is activated by or repressed by a EWS-WT1 fusion protein;

detecting amounts of transcription indicator which hybridize to each of said set of nucleic acid probes;

identifying a test cell as neoplastic if (1) hybridization of the transcription indicator of the test cell to a probe which is a EWS-WT1-activated gene is lower than hybridization using a transcription indicator from a normal cell, or (2) hybridization of the transcription indicator of the test cell to a probe which a EWS-WT1-repressed gene is higher than hybridization using a transcription indicator from a normal cell.

42. The method of claim 41 wherein the test cell is a Ewing sarcoma cell.
43. The method of claim 41 wherein at least 4 of said probes comprise portions of genes which are EWS-WT1-activated or EWS-WT1-repressed.
44. The method of claim 41 wherein at least 10 of said probes comprise portions of genes which are EWS-WT1-activated or EWS-WT1-repressed.
45. The method of claim 41 wherein at least 15 of said probes comprise portions of genes which are EWS-WT1-activated or EWS-WT1-repressed.
46. The method of claim 41 wherein the nucleic acid probes are attached to a solid support.
47. The method of claim 41 wherein the nucleic acid probes are arranged in an array.

48. The method of claim 47 wherein the array comprises nucleic acid probes which are portions of at least 250 different genes.
49. The method of claim 47 wherein the array comprises nucleic acid probes which are portions of at least 6000 different genes.
50. The method of claim 41 wherein at least one of the nucleic acid probes comprises a portion of a EWS-WT1 up-regulated gene selected from the group consisting of: Pseudorabies virus nuclear antigen (R60906 and T79475), beta chain of interleukin-2 receptor (M26062), G1/S-specific cyclin D1 (H20529), neuron-specific gamma-2 enolase (M22349), fusion gene NPM-MLF1 resulting from translocation t(3;5)(q25.1;p34) (L49054), collagen alpha 3(IV) chain (H62466), T-lymphocyte specific protein tyrosine kinase p56lck(lck) (U23852), insulin-like growth factor binding protein-4 (U20982), cathepsin B, cystatin C (T51534), dual specificity mitogen-activated protein kinase kinase 2 (R42291), neuronal-glial cell adhesion molecule (T53118), PIM-1 proto-oncogene serine/threonine-protein kinase (H10925), carboxylesterase precursor (R54359), protein-tyrosine phosphatase PAC-1 (L11329), and interferon-induced protein 6-16 (X02492).
51. The method of claim 41 wherein a test cell is identified as neoplastic if hybridization is at least 2-fold different between compared samples.
52. The method of claim 41 wherein a test cell is identified as neoplastic if hybridization is at least 3-fold different between compared samples.
53. The method of claim 41 wherein a test cell is identified as neoplastic if hybridization is at least 5-fold different between compared samples.

- 12 -

54. A method of identifying anti-cancer drugs, comprising the steps of:  
contacting a test compound with a human cell;  
determining expression of at least one EWS-WT1 up-regulated genes  
or at least one EWS-WT1 down-regulated gene in the human cell;  
identifying a test compounds an anti-cancer drug if it reduced  
expression of an EWS-WT1 up-regulated gene or increases expression of an  
EWS-WT1 down-regulated gene.
55. The method of claim 54 wherein the cell is a kidney cancer cell.
56. The method of claim 54 wherein the cell is a tumor cell.
57. The method of claim 54, the expression of said WT1 up-regulated  
and down-regulated genes is detected by measuring amounts of  
transcripts of said genes in said cell.
58. The method of claim 57, wherein said amounts of transcripts are  
measured using a high-density nucleic acid array.
59. The method of claim 54, wherein said WT1 up-regulated gene is  
selected from the group consisting of Pseudorabies virus nuclear  
antigen (R60906 and T79475), beta chain of interleukin-2 receptor  
(M26062), G1/S-specific cyclin D1 (H20529), neuron-specific  
gamma-2 enolase (M22349), fusion gene NPM-MLF1 resulting from  
translocation t(3;5)(q25.1;p34) (L49054), collagen alpha 3(IV) chain  
(H62466), T-lymphocyte specific protein tyrosine kinase p56lck(lck)  
(U23852), insulin-like growth factor binding protein-4 (U20982),  
cathepsin B, cystatin C (T51534), dual specificity mitogen-activated  
protein kinase kinase 2 (R42291), neuronal-glial cell adhesion  
molecule (T53118), PIM-1 proto-oncogene serine/threonine-protein  
kinase (H10925), carboxylesterase precursor (R54359), protein-  
tyrosine phosphatase PAC-1 (L11329), and interferon-induced  
protein 6-16 (X02492).
60. A method for detecting a WT1 gene functional mutation in target  
cells comprising the steps of:

detecting expression of at least one down-stream gene of WT1 in a sample of (a) target cells, and (b) reference cells having a wild-type WT1 gene, said reference cells being otherwise substantially similar to said target cells, said down-stream gene being selected from the group consisting of: natural killer cells protein 4 precursor (M59807), folate binding protein (M25317), HALPHA44 gene for alpha-tubulin (X06956), heat shock protein HSP70B (X51758), 90K product (H17969), heat shock 70kd protein 1 (T66307), amphiregulin (M30704), procollagen alpha 1 (T51558), beta-migrating plasminogen activator inhibitor 1 (M14083), cysteine protease inhibitor from radiated keratinocytes (X05978), brain natriuretic protein (BNP) (M31766), leukemia virus receptor 1 (GLVR1) (L20859), type 1 cytoskeletal 17 keratin, tubulin alpha-5 chain (H45051), purine nucleoside phosphorylase (T47964), GTPase (U10550), adrenomedullin (D14874), GPI-anchored urokinase plasminogen activator surface receptor (R38636), tissue type plasminogen activator (X13097), gravin (M96322), jun-B (X51345), elongation factor 1 alpha-2 (X79490), homeotic gene regulator (R16977), clone 9112 (X57348), interferon gamma treatment inducible mRNA (M26683), transmembrane receptor protein (Z17227), acidic FGF (X65779), cartilage linking protein-1 (X17405), IL-11 (X58377) mitochondrial phosphate carrier protein (R49231), and caveolin; and

comparing the expression of said at least one down-stream gene in said target cells and said reference cells, wherein a difference in said expression between the target cells and reference cells suggests a WT1 functional mutation in the target cells.

61. An in-cell functional assay for a WT1 sequence alteration comprising the steps of:

detecting expression of at least one down-stream gene selected from the group consisting of: natural killer cells protein 4 precursor (M59807), folate binding protein (M25317), HALPHA44 gene for alpha-tubulin (X06956), heat shock protein HSP70B (X51758), 90K product (H17969), heat shock 70kd protein 1 (T66307), amphiregulin (M30704), procollagen alpha 1 (T51558), beta-migrating plasminogen activator inhibitor 1 (M14083), cysteine protease inhibitor from radiated keratinocytes (X05978), brain natriuretic protein (BNP) (M31766), leukemia virus receptor 1 (GLVR1) (L20859), type 1 cytoskeletal 17 keratin, tubulin alpha-5 chain (H45051),

purine nucleoside phosphorylase (T47964), Gem GTPase (U10550), adrenomedullin (D14874), GPI-anchored urokinase plasminogen activator surface receptor (R38636), tissue type plasminogen activator (X13097), gravin (M96322), jun-B (X51345), elongation factor 1 alpha-2 (X79490), homeotic gene regulator (R16977), clone 9112 (X57348), interferon gamma treatment inducible mRNA (M26683), transmembrane receptor protein (Z17227), caveolin , acidic FGF (X65779), cartilage linking protein-1 (X17405), IL-11 (X58377) and mitochondrial phosphate carrier protein (R49231) in a target sample from target cells having a WT1 sequence alteration and in a reference sample from reference cells having a wild-type WT1 gene, said reference cells being otherwise substantially similar to said target cells, said down-stream genes being up- or down-regulated by said wild-type WT1 gene; and

comparing said expression in said target sample to said expression in said reference sample, wherein a difference in the expression between said two samples suggests that said WT1 sequence alteration affects the biological function of WT1.

62. A method for detecting a mutation in a target WT1 gene using a computer comprising:

inputting into a computer wild-type expression data of at least one WT1-down-stream gene in a wild-type sample containing a wild-type WT1 gene, wherein said down-stream gene is selected from the group consisting of: natural killer cells protein 4 precursor (M59807), folate binding protein (M25317), HALPHA44 gene for alpha-tubulin (X06956), heat shock protein HSP70B (X51758), 90K product (H17969), heat shock 70kd protein 1 (T66307), amphiregulin (M30704), procollagen alpha 1 (T51558), beta-migrating plasminogen activator inhibitor 1 (M14083), cysteine protease inhibitor from radiated keratinocytes (X05978), brain natriuretic protein (BNP) (M31766), leukemia virus receptor 1 (GLVR1) (L20859), type 1 cytoskeletal 17 keratin, tubulin alpha-5 chain (H45051), purine nucleoside phosphorylase (T47964), Gem GTPase (U10550), adrenomedullin (D14874), GPI-anchored urokinase plasminogen activator surface receptor (R38636), tissue type plasminogen activator (X13097), gravin (M96322), jun-B (X51345), elongation factor 1 alpha-2 (X79490), homeotic gene regulator (R16977), clone 9112 (X57348), interferon gamma treatment

inducible mRNA (M26683), transmembrane receptor protein (Z17227), caveolin, acidic FGF (X65779), cartilage linking protein-1(X17405), IL-11 (X58377) and mitochondrial phosphate carrier protein (R49231); and

inputting into a computer target expression data of said plurality of down-stream genes in a target sample containing said target WT1 gene;

comparing the target and wild-type expression data to determine differences, wherein differences suggest a mutation in said target WT1 gene.

63. A method of diagnosing neoplasia of a test cell comprising:

determining expression levels in a test cell and in a control cell of at least one gene which is activated by or repressed by WT1 selected from the group consisting of: natural killer cells protein 4 precursor (M59807), heat shock protein HSP70B (X51758), 90K product (H17969), heat shock 70kd protein 1 (T66307), procollagen alpha 1 (T51558), cysteine protease inhibitor from radiated keratinocytes (X05978), brain natriuretic protein (BNP) (M31766), leukemia virus receptor 1 (GLVR1) (L20859), type 1 cytoskeletal 17 keratin (R71870), purine nucleoside phosphorylase (T47964), adrenomedullin (D14874), gravin (M96322), jun-B (X51345), elongation factor 1 alpha-2 (X79490), homeotic gene regulator (R16977), clone 9112 (X57348), interferon gamma treatment inducible mRNA (M26683), transmembrane receptor protein (Z17227), acidic FGF (X65779), cartilage linking protein-1(X17405), IL-11 (X58377) and mitochondrial phosphate carrier protein (R49231);

comparing the determined expression levels in the test cell to the determined expression levels in the control cell;

identifying the test cell as neoplastic if (1) the level of expression of at least one of said genes which is a WT1-activated gene is lower in the test cell than in the control cell, or (2) the level of expression of at least one of said genes which is WT1-repressed is higher in the test cell than in the control cell.

64. A method of diagnosing neoplasia of a test cell comprising:

determining expression levels in a test cell and in a control cell of at least one gene which is activated by or repressed by a EWS-WT1 fusion protein;

comparing the determined expression levels in the test cell to the determined expression levels in the control cell;

identifying a test cell as neoplastic if (1) expression in the test cell of a EWS-WT1-activated gene is lower than expression in a normal cell, or (2) expression in the test cell of a EWS-WT1-repressed gene is higher than expression in the normal cell.

65. A method of diagnosing neoplasia of a test cell comprising:

determining expression levels in a test cell and in a control cell of at least three genes which are activated by or repressed by WT1;

comparing the determined expression levels in the test cell to the determined expression levels in the control cell;

identifying a test cell as neoplastic if (1) expression in the test cell of a WT1-activated gene is lower than expression in a normal cell, or (2) expression in the test cell of a WT1-repressed gene is higher than expression in the normal cell.

### 3 Detailed Description of Invention

#### BACKGROUND OF THE INVENTION

Many biological functions are accomplished by altering the expression of various genes through transcriptional (e.g. through control of initiation, provision of RNA precursors, RNA processing, etc.) and/or translational control. For example, fundamental biological processes such as cell cycle regulation, cell differentiation and cell death, are often characterized by the variations in the expression levels of groups of genes.

Gene expression is also associated with pathogenesis. For example, the lack of sufficient expression of functional tumor suppressor genes and/or the over expression of oncogene/protooncogenes could lead to tumorigenesis (Marshall, *Cell*, 64: 313-326 (1991); Weinberg, *Science*, 254: 1138-1146 (1991), incorporated herein by reference for all purposes). Thus, changes in the expression levels of particular genes (e.g. oncogenes or tumor suppressors) serve as signposts for the presence and progression of various diseases. There is a need in the art for discovering which genes are affected by particular tumor suppressors and oncogenes, so that they can be used diagnostically as well as in the search for new therapeutics.

#### SUMMARY OF THE INVENTION

It is an object of the invention to provide methods for identifying a functional mutation in a WT1 gene.

It is another object of the invention to provide methods for identifying a EWS-WT1 gene fusion.

It is still another object of the invention to provide an in-cell assay to test the biological effect of a WT1 mutation.

It is yet another object of the invention to provide methods to detect a WT1 mutation or a EWS-WT1 fusion using a computer.

It is a further object of the invention to provide methods to diagnose neoplasia.

It is another object of the invention to provide methods for identifying drugs useful for treating neoplasias.

These and other objects of the invention are provided by one or more of the following embodiments. In one embodiment a method is provided for detecting a WT1 gene functional mutation in target cells. Expression of at least three down-stream genes of WT1 is detected in a sample of (a) target cells, and (b) reference cells having a wild-type WT1 gene. The reference cells are otherwise substantially similar to the target cells. The down-stream genes are up- or down-regulated by the wild-type WT1 gene. The expression of the down-stream genes in the target cells and the reference cells is compared; a difference in the expression between the target cells and reference cells suggests a WT1 functional mutation in the target cells.

According to yet another embodiment a method is provided for detecting a EWS-WT1 gene fusion in target cells. Expression of one or more down-stream genes of an EWS-WT1 fusion is detected in a sample of (a) target cells, and (b) reference cells having a wild-type EWS and WT1 gene. The reference cells are otherwise substantially similar to the target cells. The down-stream genes are up- or down-regulated by the wild-type EWS and WT1 gene. The expression of the down-stream genes in the target cells and the reference cells is compared. A difference in the expression between the target cells and reference cells suggests a EWS-WT1 fusion in the target cells.

Another aspect of the invention is an in-cell functional assay for a WT1 sequence alteration. The expression is detected of at least three down-stream genes in a target sample from target cells having a WT1 sequence alteration and in a reference sample from reference cells having a wild-type WT1 gene. The reference cells are otherwise substantially similar to the target cells. The down-stream genes are up- or down-regulated by the wild-type WT1 gene. The expression in the target sample is compared to the expression in the reference sample. A difference in the expression between the two samples suggests that the WT1 sequence alteration affects the biological function of WT1.

According to another aspect of the invention a method is provided for detecting a mutation in a target WT1 gene using a computer. Wild-type expression data of at least three down-stream genes in a wild-type sample containing a wild-type WT1 gene is input into a computer. The down-stream

genes are transcriptionally regulated by the wild-type WT1 gene. Target expression data of the plurality of down-stream genes in a target sample containing the target WT1 gene is also input into a computer. The target and wild-type expression data are compared to determine differences. Differences suggest a mutation in the target WT1 gene.

In still another embodiment a method is provided for detecting a translocation fusing EWS and WT1 genes using a computer. Wild-type expression data of a plurality of down-stream genes in a wild-type sample containing wild-type EWS and WT1 genes is input into a computer. The down-stream genes are transcriptionally regulated by a EWS-WT1 fusion protein. Target expression data of the plurality of down-stream genes in a target sample which is being tested for the presence of a EWS-WT1 fusion protein is also input into a computer. The target and wild-type expression data are compared to determine differences. Such differences suggest a translocation fusing the EWS and WT1 genes.

Another embodiment of the invention provides a method of diagnosing neoplasia of a test cell. A transcription indicator of a test cell is hybridized to a set of nucleic acid probes. The transcription indicator is selected from the group consisting of mRNA, cDNA and cRNA. The set of nucleic acid probes comprises a plurality of nucleic acid molecules each of which is a portion of a gene which is activated by or repressed by WT1 selected from the group consisting of: natural killer cells protein 4 precursor (M59807), heat shock protein HSP70B (X51758), 90K product (H17969), heat shock 70kd protein 1 (T66307), procollagen alpha 1 (T51558), cysteine protease inhibitor from radiated keratinocytes (X05978), brain natriuretic protein (BNP) (M31766), leukemia virus receptor 1 (GLVRI) (L20859), type 1 cytoskeletal 17 keratin (R71870), purine nucleoside phosphorylase (T47964), adrenomedullin (D14874), gravin (M96322), jun-B (X51345), elongation factor 1 alpha-2 (X79490), homeotic gene regulator (R16977), clone 9112 (X57348), interferon gamma treatment inducible mRNA (M26683), transmembrane receptor protein (Z17227) and mitochondrial phosphate carrier protein (R49231). Amounts of transcription indicator which hybridize to each of the set of nucleic acid probes are detected. A test cell is identified as neoplastic if (1) hybridization of the transcription indicator of the test cell to a probe which is a WT1-activated gene is lower than hybridization using a transcription indicator from a normal cell, or (2) hybridization of the transcription indicator of the test cell to a probe which a

WT1-repressed gene is higher than hybridization using a transcription indicator from a normal cell.

A further aspect of the invention is provided by a method of identifying anti-cancer drugs. A test compound is contacted with a human cell. The effect of the test compound on the expression by the cell of at least one WT1 up- or down-regulated gene is determined. The regulated gene is selected from the group consisting of: natural killer cells protein 4 precursor (M59807), heat shock protein HSP70B (X51758), 90K product (H17969), heat shock 70kd protein 1 (T66307), procollagen alpha 1 (T51558), cysteine protease inhibitor from radiated keratinocytes (X05978), brain natriuretic protein (BNP) (M31766), leukemia virus receptor 1 (GLVR1) (L20859), type 1 cytoskeletal 17 keratin, purine nucleoside phosphorylase (T47964), adrenomedullin (D14874), gravin (M96322), jun-B (X51345), elongation factor 1 alpha-2 (X79490), homeotic gene regulator (R16977), clone 9112 (X57348), interferon gamma treatment inducible mRNA (M26683), transmembrane receptor protein (Z17227) and mitochondrial phosphate carrier protein (R49231). A test compound is identified as a potential anti-cancer drug if it increases expression of at least one WT1 up-regulated gene or decreases expression of a WT1 down-regulated gene in the human cell.

Another aspect of the invention is a method of diagnosing neoplasia of a test cell. A transcription indicator of a test cell is hybridized to a set of nucleic acid probes. The transcription indicator is selected from the group consisting of mRNA, cDNA and cRNA. The set of nucleic acid probes comprises a plurality of nucleic acid molecules each of which is a portion of a gene which is activated by or repressed by a EWS-WT1 fusion protein. Amounts of transcription indicator which hybridize to each of the set of nucleic acid probes are detected. A test cell is identified as neoplastic if (1) hybridization of the transcription indicator of the test cell to a probe which is a EWS-WT1-activated gene is lower than hybridization using a transcription indicator from a normal cell, or (2) hybridization of the transcription indicator of the test cell to a probe which a EWS-WT1-repressed gene is higher than hybridization using a transcription indicator from a normal cell.

According to still another embodiment of the invention a method of identifying anti-cancer drugs is provided. A test compound is contacted with a human cell. Expression of an EWS-WT1 up-regulated gene or a EWS-WT1 down-regulated gene is determined. A test compound is identified as a potential

anti-cancer drug if it decreases expression of a EWS-WT1 up-regulated genes or increases expression of a EWS-WT1 down-regulated gene in the human cell.

A method is also provided for detecting a WT1 gene functional mutation in target cells. Expression of at least one down-stream gene of WT1 is detected in a sample of (a) target cells, and (b) reference cells having a wild-type WT1 gene. The reference cells are otherwise substantially similar to the target cells. The down-stream genes are selected from the group consisting of: natural killer cells protein 4 precursor (M59807), folate binding protein (M25317), HALPHA44 gene for alpha-tubulin (X06956), heat shock protein HSP70B (X51758), 90K product (H17969), heat shock 70kd protein 1 (T66307), amphiregulin (M30704), procollagen alpha 1 (T51558), beta-migrating plasminogen activator inhibitor 1 (M14083), cysteine protease inhibitor from radiated keratinocytes (X05978), brain natriuretic protein (BNP) (M31766), leukemia virus receptor 1 (GLVR1) (L20859), type 1 cytoskeletal 17 keratin, tubulin alpha-5 chain (H45051), purine nucleoside phosphorylase (T47964), GTPase (U10550), adrenomedullin (D14874), GPI-anchored urokinase plasminogen activator surface receptor (R38636), tissue type plasminogen activator (X13097), gravin (M96322), jun-B (X51345), elongation factor 1 alpha-2 (X79490), homeotic gene regulator (R16977), clone 9112 (X57348), interferon gamma treatment inducible mRNA (M26683), transmembrane receptor protein (Z17227) mitochondrial phosphate carrier protein (R49231), and caveolin. The expression of the at least one down-stream gene in the target cells and the reference cells is compared. A difference in the expression between the target cells and reference cells suggests a WT1 functional mutation in the target cells.

An in-cell functional assay for a WT1 sequence alteration is also provided by the present invention. Expression is detected in a target sample from target cells having a WT1 sequence alteration and in a reference sample from reference cells having a wild-type WT1 gene of at least one down-stream gene selected from the group consisting of: natural killer cells protein 4 precursor (M59807), folate binding protein (M25317), HALPHA44 gene for alpha-tubulin (X06956), heat shock protein HSP70B (X51758), 90K product (H17969), heat shock 70kd protein 1 (T66307), amphiregulin (M30704), procollagen alpha 1 (T51558), beta-migrating plasminogen activator inhibitor 1 (M14083), cysteine protease inhibitor from radiated keratinocytes (X05978), brain natriuretic protein (BNP) (M31766), leukemia virus receptor 1 (GLVR1)

(L20859), type 1 cytoskeletal 17 keratin, tubulin alpha-5 chain (H45051), purine nucleoside phosphorylase (T47964), Gem GTPase (U10550), adrenomedullin (D14874), GPI-anchored urokinase plasminogen activator surface receptor (R38636), tissue type plasminogen activator (X13097), gravin (M96322), jun-B (X51345), elongation factor 1 alpha-2 (X79490), homeotic gene regulator (R16977), clone 9112 (X57348), interferon gamma treatment inducible mRNA (M26683), transmembrane receptor protein (Z17227), caveolin and mitochondrial phosphate carrier protein (R49231). The reference cells are otherwise substantially similar to the target cells. The down-stream genes are up- or down-regulated by the wild-type WT1 gene. Expression in the target sample is compared to the expression in the reference sample. A difference in the expression between the two samples suggests that the WT1 sequence alteration affects the biological function of WT1.

Another embodiment of the invention provides a method for detecting a mutation in a target WT1 gene using a computer. Wild-type expression data of at least one WT1-down-stream gene in a wild-type sample containing a wild-type WT1 gene is input into a computer. The down-stream gene is selected from the group consisting of: natural killer cells protein 4 precursor (M59807), folate binding protein (M25317), HALPHA44 gene for alpha-tubulin (X06956), heat shock protein HSP70B (X51758), 90K product (H17969), heat shock 70kd protein 1 (T66307), amphiregulin (M30704), procollagen alpha 1 (T51558), beta-migrating plasminogen activator inhibitor 1 (M14083), cysteine protease inhibitor from radiated keratinocytes (X05978), brain natriuretic protein (BNP) (M31766), leukemia virus receptor 1 (GLVR1) (L20859), type 1 cytoskeletal 17 keratin, tubulin alpha-5 chain (H45051), purine nucleoside phosphorylase (T47964), Gem GTPase (U10550), adrenomedullin (D14874), GPI-anchored urokinase plasminogen activator surface receptor (R38636), tissue type plasminogen activator (X13097), gravin (M96322), jun-B (X51345), elongation factor 1 alpha-2 (X79490), homeotic gene regulator (R16977), clone 9112 (X57348), interferon gamma treatment inducible mRNA (M26683), transmembrane receptor protein (Z17227), caveolin, and mitochondrial phosphate carrier protein (R49231). Target expression data of the plurality of down-stream genes in a target sample containing the target WT1 gene is also input into a computer. The target and wild-type expression data are compared to determine differences. Such differences suggest a mutation in the target WT1 gene.

According to yet another aspect of the invention a method of diagnosing neoplasia of a test cell is provided. Expression levels are determined in a test cell and in a control cell of at least one gene which is activated by or repressed by WT1. The genes are selected from the group consisting of: natural killer cells protein 4 precursor (M59807), heat shock protein HSP70B (X51758), 90K product (H17969), heat shock 70kd protein 1 (T66307), procollagen alpha 1 (T51558), cysteine protease inhibitor from radiated keratinocytes (X05978), brain natriuretic protein (BNP) (M31766), leukemia virus receptor 1 (GLVR1) (L20859), type 1 cytoskeletal 17 keratin (R71870), purine nucleoside phosphorylase (T47964), adrenomedullin (D14874), gravin (M96322), jun-B (X51345), elongation factor 1 alpha-2 (X79490), homeotic gene regulator (R16977), clone 9112 (X57348), interferon gamma treatment inducible mRNA (M26683), transmembrane receptor protein (Z17227) and mitochondrial phosphate carrier protein (R49231). The determined expression levels in the test cell is compared to the determined expression levels in the control cell. The test cell is identified as neoplastic if (1) the level of expression of at least one of the genes which is a WT1-activated gene is lower in the test cell than in the control cell, or (2) the level of expression of at least one of the genes which is WT1-repressed is higher in the test cell than in the control cell.

Another aspect of the invention provides a method of diagnosing neoplasia of a test cell. Expression levels in a test cell and in a control cell are determined of at least one gene which is activated by or repressed by a EWS-WT1 fusion protein. The determined expression levels in the test cell are compared to the determined expression levels in the control cell. A test cell is identified as neoplastic if (1) expression in the test cell of a EWS-WT1-activated gene is lower than expression in a normal cell, or (2) expression in the test cell of a EWS-WT1-repressed gene is higher than expression in the normal cell.

Also provided is an additional method of diagnosing neoplasia of a test cell. Expression levels are determined in a test cell and in a control cell of at least three genes which are activated by or repressed by WT1. The determined expression levels in the test cell are compared to the determined expression levels in the control cell. A test cell is identified as neoplastic if (1) expression in the test cell of a WT1-activated gene is lower than expression in a normal cell, or (2) expression in the test cell of a WT1-repressed gene is higher than expression in the normal cell.

## DETAILED DESCRIPTION OF THE INVENTION

It is a discovery of the present inventors that WT1 regulates the transcription of a set of genes, the size of the set being far greater than previously suggested. Sets of genes have been identified which are up-regulated and down-regulated by WT1, the latter set being far smaller. In addition, genes have been identified which are up-regulated by the EWS-WT1 gene fusion caused by a translocation. This translocation occurs in desmoplastic round cell tumors, such as Ewing's Sarcomas. Identification of these sets of genes permits their use in a variety of diagnostic and analytical methods. It also permits their use to screen compounds for drugs which have the same or similar effects as WT1 and for drugs which have the opposite effect from the EWS-WT1 gene fusion.

Tables 1a and 1b show the genes which are regulated by WT1. Of those genes only caveolin and elongation factor 1 alpha-2 are down-regulated by WT1. The others are up-regulated by WT1. Table 2 shows a subset of the WT1-regulated genes which were known to have a previously suggested link to cancer and what that link is. Table 3 shows the genes which are regulated by the EWS-WT1 fusion gene. All of these genes are up-regulated by the fusion gene. Other genes are undoubtedly regulated by these genes, and they can be identified using similar techniques to those described in the example. Alternatively, other techniques such as Serial Analysis of Gene Expression (SAGE) can be used to identify regulated genes which may not yet have been identified or made part of the public databases.

Expression of genes which are regulatorily downstream of WT1 or EWS-WT1 can be monitored by any techniques known in the art. These include but are not limited to the use of SAGE, hybridization to probes on an array, Western blot, Northern blot, dot blot, slot blot, ELISA, radioimmunoassay. Any technique which can be used to quantitate a gene product, whether mRNA or protein, can be used. The methods of the present invention predominantly analyze changes in a test sample as compared to a control sample. Precise quantitation need not be accomplished. In many cases comparisons will suffice without ever determining an absolute amount of a gene product.

Target cells, as used in the present invention are those cells which are being tested. Typically they are of an unknown mutational status. In some instances, a mutation is identified in the target cells, but its functional effect is unknown. Reference cells are preferably as similar as possible to the target cells, to permit the

earliest detection of differences which are due to the WT1 gene or the EWS-WT1 gene. Reference cells typically comprise wild-type WT1 alleles. Often the reference cells are taken from the same individual and from the tissue adjacent to a tumor sample which is being tested.

Differences between target cells and reference cells may be small or large. Some small differences may be very reproducible and therefore useful. For other purposes, large differences may be desirable for ease of detection. For some assays it may be useful to set threshold levels of change. For some purposes such threshold levels may be 0.5, 2, 3, 5, 7, or 10 fold. This may depend on the technique being used for detection as well as on the number of genes which are being tested. One of skill in the art can readily set threshold levels which are preferred in a particular methodological context.

The number of downstream genes analyzed can vary depending, for example, on the magnitude of induced changes and their reproducibility. In some instances, larger numbers of genes may be analyzed to more accurately reflect a WT1 change, since other genes may also affect expression of a particular gene being analyzed. Typically at least 1, 2, 3, 4, 5, 7, 10, or 15 downstream genes of WT1 or EWS-WT1 are analyzed. Analysis of a number of genes greater than 1 can be accomplished simultaneously, sequentially, or cumulatively.

Computers can be used in the present invention to store data, to perform mathematical comparisons, to store results. Data can be manually input or may be fed from other machines, such as optical scanners, scintillation counters, and the like. High through-put screens can thus be efficiently run and analyzed.

## I. Definitions

Hybridizing specifically to: The phrase "hybridizing specifically to" refers to the binding, duplexing, or hybridizing of a molecule substantially to or only to a particular nucleotide sequence or sequences under stringent conditions when that sequence is present in a complex mixture (e.g., total cellular) DNA or RNA.

mRNA or transcript: The term "mRNA" refers to transcripts of a gene. Transcripts are RNA including, for example, mature messenger RNA ready for translation, products of various stages of transcript processing. Transcript processing may include splicing, editing and degradation.

Nucleic Acid: The terms "nucleic acid" or "nucleic acid molecule" refer to a deoxyribonucleotide or ribonucleotide polymer in either single- or double-stranded form, and unless otherwise limited, would encompass analogs of natural nucleotide that can function in a similar manner as naturally occurring nucleotide. An oligonucleotide is a single-stranded nucleic acid of 2 to n bases, where n may be greater than 500 to 1000. Nucleic acids may be cloned or synthesized using any technique known in the art. They may also include non-naturally occurring nucleotide analogs, such as those which are modified to improve hybridization and peptide nucleic acids.

Nucleic acid encoding a regulatory molecule: The regulatory molecule may be DNA, RNA or protein. Thus for example DNA sites which bind protein or other nucleic acid molecules are included within the class of regulatory molecules encoded by a nucleic acid.

Probe: As used herein a "probe" is defined as a nucleic acid, capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a probe may include natural (i.e. A, G, U, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, the bases in probes may be joined by a linkage other than a phosphodiester bond, so long as it does not interfere with hybridization. Thus, probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages.

Target nucleic acid: The term "target nucleic acid" refers to a nucleic acid (often derived from a biological sample), to which the probe is designed to specifically hybridize. It is either the presence or absence of the target nucleic acid that is to be detected, or the amount of the target nucleic acid that is to be quantified. The target nucleic acid has a sequence that is complementary to the nucleic acid sequence of the corresponding probe directed to the target. The term target nucleic

acid may refer to the specific subsequence of a larger nucleic acid to which the probe is directed or to the overall sequence (e.g., gene or mRNA) whose expression level it is desired to detect. The difference in usage will be apparent from context.

Stringent conditions: The term "stringent conditions" refers to conditions under which a probe will hybridize to its target subsequence, but with only insubstantial hybridization to other sequences or to other sequences such that the difference may be identified. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. Generally, stringent conditions are selected to be about 5 °C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH.

Subsequence: "Subsequence" refers to a sequence of nucleic acids that comprise a part of a longer sequence of nucleic acids.

Thermal melting point (Tm): The Tm is the temperature, under defined ionic strength, pH, and nucleic acid concentration, at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. As the target sequences are generally present in excess, at Tm, 50% of the probes are occupied at equilibrium). Typically, stringent conditions will be those in which the salt concentration is at least about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30 °C for short probes (e.g., 10 to 50 nucleotide). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide.

Quantifying: The term "quantifying" or "quantitating" when used in the context of quantifying transcription levels of a gene can refer to absolute or to relative quantification. Absolute quantification may be accomplished by inclusion of known concentration(s) of one or more target nucleic acids (e.g. control nucleic acids such as Bio B or with known amounts the target nucleic acids themselves) and referencing the hybridization intensity of unknowns with the known target nucleic acids (e.g. through generation of a standard curve). Alternatively, relative quantification can be accomplished by comparison of hybridization signals between two or more genes, or between two or more treatments to quantify the changes in hybridization intensity and, by implication, transcription level.

Up-stream or down-stream gene. If the expression of a first gene is regulated by a second gene, the second gene is called an "up-stream gene" for the first gene and the first gene is the "down-stream" gene of the second gene. The regulation of the first gene by second gene could be through trans-activation. For example, the first gene encodes a transcriptional factor that controls the expression of the second gene. The regulation can also be exerted by a cis-acting mechanism. For example, the first gene is in the proximity of the second gene and exerts a positional effect on the expression of the second gene. In this case, the first gene does not have to be expressed in order to have an influence on the second gene.

Activity of a gene is reflected by the activity of its product(s): the proteins or other molecules encoded by the gene. Those product molecules perform biological functions. Directly measuring the activity of a gene product is, however, often difficult for certain genes. Instead, the immunological activities or the amount of the final product(s) or its peptide processing intermediates are determined as a measurement of the gene activity. More frequently, the amount or activity of intermediates, such as transcripts, RNA processing intermediates, or mature mRNAs are detected as a measurement of gene activity.

In many cases, the form and function of the final product(s) of a gene is unknown. In those cases, the activity of a gene is measured conveniently by the amount or activity of transcript(s), RNA processing intermediate(s), mature mRNA(s) or its protein product(s) or functional activity of its protein product(s).

Any methods that measure the activity of a gene are useful for at least some embodiments of this invention. For example, traditional Northern blotting and hybridization, nuclease protection, RT-PCR and differential display have been used for detecting gene activity. Those methods are useful for some embodiments of the invention. However, this invention is most useful in conjunction with methods for detecting the expression of a large number of genes.

High-density arrays are particularly useful for monitoring the expression control at the transcriptional, RNA processing and degradation level. The fabrication and application of high-density arrays in gene expression monitoring have been disclosed previously in, for example, WO 97/10365, WO 92/10588, U.S. Application Ser. No. 08/772,376 filed December 23, 1996; serial number 08/529,115 filed on September 15, 1995; serial number 08/168,904 filed December

15, 1993; serial number 07/624,114 filed on December 6, 1990, serial number 07/362,901 filed June 7, 1990, all incorporated herein for all purposes by reference. In some embodiment using high-density arrays, high-density oligonucleotide arrays are synthesized using methods such as the Very Large Scale Immobilized Polymer Synthesis (VLSIPS) disclosed in U.S. Pat. No. 5,445,934 incorporated herein for all purposes by reference. Each oligonucleotide occupies a known location on a substrate. A nucleic acid target sample is hybridized with a high-density array of oligonucleotides and then the amount of target nucleic acids hybridized to each probe in the array is quantified. One preferred quantifying method is to use confocal microscope and fluorescent labels. The GeneChip® system (Affymetrix, Santa Clara, CA) is particularly suitable for quantifying the hybridization; however, it will be apparent to those of skill in the art that any similar systems or other effectively equivalent detection methods can also be used.

High-density arrays are suitable for quantifying a small variations in expression levels of a gene in the presence of a large population of heterogeneous nucleic acids. Such high-density arrays can be fabricated either by de novo synthesis on a substrate or by spotting or transporting nucleic acid sequences onto specific locations of substrate. Nucleic acids are purified and/or isolated from biological materials, such as a bacterial plasmid containing a cloned segment of sequence of interest. Suitable nucleic acids are also produced by amplification of templates. As a nonlimiting illustration, polymerase chain reaction, and/or in vitro transcription, are suitable nucleic acid amplification methods.

Synthesized oligonucleotide arrays are particularly preferred for this invention. Oligonucleotide arrays have numerous advantages, as opposed to other methods, such as efficiency of production, reduced intra- and inter array variability, increased information content and high signal-to-noise ratio.

Preferred high-density arrays for gene function identification and genetic network mapping comprise greater than about 100, preferably greater than about 1000, more preferably greater than about 16,000 and most preferably greater than 65,000 or 250,000 or even greater than about 1,000,000 different oligonucleotide probes, preferably in less than 1 cm<sup>2</sup> of surface area. The oligonucleotide probes range from about 5 to about 50 or about 500 nucleotides, more preferably from about 10 to about 40 nucleotide and most preferably from about 15 to about 40 nucleotides in length.

Generally methods of monitoring gene expression involve (a) providing a pool of target nucleic acids comprising RNA transcript(s) of one or more target gene(s), or nucleic acids derived from the RNA transcript(s); (b) hybridizing the nucleic acid sample to a high-density array of probes and (c) detecting the hybridized nucleic acids and calculating a relative and/or absolute expression (transcription, RNA processing or degradation) level.

(A) Providing a Nucleic Acid Sample

One of skill in the art will appreciate that it is desirable to have nucleic samples containing target nucleic acid sequences that reflect the transcripts of interest. Therefore, suitable nucleic acid samples may contain transcripts of interest. Suitable nucleic acid samples, however, may contain nucleic acids derived from the transcripts of interest. As used herein, a nucleic acid derived from a transcript refers to a nucleic acid for whose synthesis the mRNA transcript or a subsequence thereof has ultimately served as a template. Thus, a cDNA reverse transcribed from a transcript, an RNA transcribed from that cDNA, a DNA amplified from the cDNA, an RNA transcribed from the amplified DNA, *etc.*, are all derived from the transcript and detection of such derived products is indicative of the presence and/or abundance of the original transcript in a sample. Thus, suitable samples include, but are not limited to, transcripts of the gene or genes, cDNA reverse transcribed from the transcript, cRNA transcribed from the cDNA, DNA amplified from the genes, RNA transcribed from amplified DNA, and the like.

Transcripts, as used herein, may include, but not limited to pre-mRNA nascent transcript(s), transcript processing intermediates, mature mRNA(s) and degradation products. It is not necessary to monitor all types of transcripts to practice this invention. For example, one may choose to practice the invention to measure the mature mRNA levels only.

In one embodiment, such sample is a homogenate of cells or tissues or other biological samples. Preferably, such sample is a total RNA preparation of a biological sample. More preferably in some embodiments, such a nucleic acid sample is the total mRNA isolated from a biological sample. Those of skill in the art will appreciate that the total mRNA prepared with most methods includes not only the mature mRNA, but also the RNA processing intermediates and nascent pre-mRNA transcripts. For example, total mRNA purified with a poly (dT) column contains RNA molecules with poly (A) tails. Those polyA<sup>+</sup> RNA

molecules could be mature mRNA, RNA processing intermediates, nascent transcripts or degradation intermediates.

Biological samples may be of any biological tissue or fluid or cells from any organism. Frequently the sample will be a "clinical sample" which is a sample derived from a patient. Clinical samples provide a rich source of information regarding the various states of genetic network or gene expression. Some embodiments of the invention are employed to detect mutations and to identify the phenotype of mutations. Such embodiments have extensive applications in clinical diagnostics and clinical studies. Typical clinical samples include, but are not limited to, sputum, blood, blood cells (e.g., white cells), tissue or fine needle biopsy samples, urine, peritoneal fluid, and pleural fluid, or cells therefrom. Biological samples may also include sections of tissues, such as frozen sections or formalin fixed sections taken for histological purposes.

Another typical source of biological samples are cell cultures where gene expression states can be manipulated to explore the relationship among genes. In one aspect of the invention, methods are provided to generate biological samples reflecting a wide variety of states of the genetic network.

One of skill in the art would appreciate that it is desirable to inhibit or destroy RNase present in homogenates before homogenates can be used for hybridization. Methods of inhibiting or destroying nucleases are well known in the art. In some preferred embodiments, cells or tissues are homogenized in the presence of chaotropic agents to inhibit nuclease. In some other embodiments, RNase is inhibited or destroyed by heat treatment followed by proteinase treatment.

Methods of isolating total mRNA are also well known to those of skill in the art. For example, methods of isolation and purification of nucleic acids are described in detail in Chapter 3 of *Laboratory Techniques in Biochemistry and Molecular Biology: Hybridization With Nucleic Acid Probes, Part I. Theory and Nucleic Acid Preparation*, P. Tijssen, ed. Elsevier, N.Y. (1993) and Chapter 3 of *Laboratory Techniques in Biochemistry and Molecular Biology: Hybridization With Nucleic Acid Probes, Part I. Theory and Nucleic Acid Preparation*, P. Tijssen, ed. Elsevier, N.Y. (1993)).

In a preferred embodiment, the total RNA is isolated from a given sample using, for example, an acid guanidinium-phenol-chloroform extraction method and polyA<sup>+</sup> mRNA is isolated by oligo(dT) column chromatography or by using (dT) on magnetic beads (see, e.g., Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual* (2nd ed.), Vols. 1-3, Cold Spring Harbor Laboratory, (1989), or *Current*

*Protocols in Molecular Biology*, F. Ausubel *et al.*, ed. Greene Publishing and Wiley-Interscience, New York (1987)).

Frequently, it is desirable to amplify the nucleic acid sample prior to hybridization. One of skill in the art will appreciate that whatever amplification method is used, if a quantitative result is desired, care must be taken to use a method that maintains or controls for the relative frequencies of the amplified nucleic acids to achieve quantitative amplification.

Methods of "quantitative" amplification are well known to those of skill in the art. For example, quantitative PCR involves simultaneously co-amplifying a known quantity of a control sequence using the same primers. This provides an internal standard that may be used to calibrate the PCR reaction. The high-density array may then include probes specific to the internal standard for quantification of the amplified nucleic acid.

One preferred internal standard is a synthetic AW106 cRNA. The AW106 cRNA is combined with RNA isolated from the sample according to standard techniques known to those of skilled in the art. The RNA is then reverse transcribed using a reverse transcriptase to provide copy DNA. The cDNA sequences are then amplified (e.g., by PCR) using labeled primers. The amplification products are separated, typically by electrophoresis, and the amount of radioactivity (proportional to the amount of amplified product) is determined. The amount of mRNA in the sample is then calculated by comparison with the signal produced by the known AW106 RNA standard. Detailed protocols for quantitative PCR are provided in *PCR Protocols, A Guide to Methods and Applications*, Innis *et al.*, Academic Press, Inc. N.Y., (1990).

Other suitable amplification methods include, but are not limited to polymerase chain reaction (PCR) (Innis, *et al.*, *PCR Protocols. A guide to Methods and Application*. Academic Press, Inc. San Diego, (1990)), ligase chain reaction (LCR) (see Wu and Wallace, *Genomics*, 4: 560 (1989), Landegren, *et al.*, *Science*, 241: 1077 (1988) and Barringer, *et al.*, *Gene*, 89: 117 (1990), transcription amplification (Kwoh, *et al.*, *Proc. Natl. Acad. Sci. USA*, 86: 1173 (1989)), and self-sustained sequence replication (Guatelli, *et al.*, *Proc. Nat. Acad. Sci. USA*, 87: 1874 (1990)).

Cell lysates or tissue homogenates often contain a number of inhibitors of polymerase activity. Therefore, RT-PCR typically incorporates preliminary steps to isolate total RNA or mRNA for subsequent use as an amplification template. A one-tube mRNA capture method may be used to prepare poly(A)<sup>+</sup>RNA samples

suitable for immediate RT-PCR in the same tube (Boehringer Mannheim). The captured mRNA can be directly subjected to RT-PCR by adding a reverse transcription mix and, subsequently, a PCR mix.

In a particularly preferred embodiment, the sample mRNA is reverse transcribed with a reverse transcriptase and a primer consisting of oligo(dT) and a sequence encoding the phage T7 promoter to provide single stranded DNA template. The second DNA strand is polymerized using a DNA polymerase. After synthesis of double-stranded cDNA, T7 RNA polymerase is added and RNA is transcribed from the cDNA template. Successive rounds of transcription from each single cDNA template results in amplified RNA. Methods of *in vitro* polymerization are well known to those of skill in the art (see, e.g., Sambrook, *supra*) and this particular method is described in detail by Van Gelder, *et al.*, *Proc. Natl. Acad. Sci. USA*, 87: 1663-1667 (1990) who demonstrate that *in vitro* amplification according to this method preserves the relative frequencies of the various RNA transcripts. Moreover, Eberwine *et al.* *Proc. Natl. Acad. Sci. USA*, 89: 3010-3014 provide a protocol that uses two rounds of amplification via *in vitro* transcription to achieve greater than  $10^6$  fold amplification of the original starting material, thereby permitting expression monitoring even where biological samples are limited.

(B) Hybridizing nucleic acids to high-density arrays

1. Probe design

One of skill in the art will appreciate that an enormous number of array designs are suitable for the practice of this invention. The high-density array will typically include a number of probes that specifically hybridize to the sequences of interest. In addition, in a preferred embodiment, the array will include one or more control probes.

The high-density array chip includes "test probes." Test probes could be oligonucleotides that range from about 5 to about 45 or 5 to about 500 nucleotides, more preferably from about 10 to about 40 nucleotides and most preferably from about 15 to about 40 nucleotides in length. In other particularly preferred embodiments the probes are 20 or 25 nucleotides in length. In another preferred embodiment, test probes are double or single strand DNA sequences. DNA sequences are isolated or cloned from nature sources or amplified from nature sources using nature nucleic acid as templates. These probes have sequences complementary to particular subsequences of the genes whose expression they are

designed to detect. Thus, the test probes are capable of specifically hybridizing to the target nucleic acid they are to detect.

In addition to test probes that bind the target nucleic acid(s) of interest, the high-density array can contain a number of control probes. The control probes fall into three categories referred to herein as 1) normalization controls; 2) expression level controls; and 3) mismatch controls.

Normalization controls are oligonucleotide or other nucleic acid probes that are complementary to labeled reference oligonucleotides or other nucleic acid sequences that are added to the nucleic acid sample. The signals obtained from the normalization controls after hybridization provide a control for variations in hybridization conditions, label intensity, "reading" efficiency and other factors that may cause the signal of a perfect hybridization to vary between arrays. In a preferred embodiment, signals (e.g., fluorescence intensity) read from all other probes in the array are divided by the signal (e.g., fluorescence intensity) from the control probes thereby normalizing the measurements.

Virtually any probe may serve as a normalization control. However, it is recognized that hybridization efficiency varies with base composition and probe length. Preferred normalization probes are selected to reflect the average length of the other probes present in the array, however, they can be selected to cover a range of lengths. The normalization control(s) can also be selected to reflect the (average) base composition of the other probes in the array, however in a preferred embodiment, only one or a few normalization probes are used and they are selected such that they hybridize well (i.e. no secondary structure) and do not match any target-specific probes.

Expression level controls are probes that hybridize specifically with constitutively expressed genes in the biological sample. Virtually any constitutively expressed gene provides a suitable target for expression level controls. Typically expression level control probes have sequences complementary to subsequences of constitutively expressed "housekeeping genes" including, but not limited to the  $\beta$ -actin gene, the transferrin receptor gene, the GAPDH gene, and the like.

Mismatch controls may also be provided for the probes to the target genes, for expression level controls or for normalization controls. Mismatch controls are oligonucleotide probes or other nucleic acid probes identical to their corresponding test or control probes except for the presence of one or more mismatched bases. A mismatched base is a base selected so that it is not complementary to the

corresponding base in the target sequence to which the probe would otherwise specifically hybridize. One or more mismatches are selected such that under appropriate hybridization conditions (e.g. stringent conditions) the test or control probe would be expected to hybridize with its target sequence, but the mismatch probe would not hybridize (or would hybridize to a significantly lesser extent). Preferred mismatch probes contain a central mismatch. Thus, for example, where a probe is a 20 mer, a corresponding mismatch probe will have the identical sequence except for a single base mismatch (e.g., substituting a G, a C or a T for an A) at any of positions 6 through 14 (the central mismatch).

Mismatch probes thus provide a control for non-specific binding or cross-hybridization to a nucleic acid in the sample other than the target to which the probe is directed. Mismatch probes thus indicate whether a hybridization is specific or not. For example, if the target is present the perfect match probes should be consistently brighter than the mismatch probes. In addition, if all central mismatches are present, the mismatch probes can be used to detect a mutation. The difference in intensity between the perfect match and the mismatch probe ( $I(PM) - I(MM)$ ) provides a good measure of the concentration of the hybridized material.

The high-density array may also include sample preparation/amplification control probes. These are probes that are complementary to subsequences of control genes selected because they do not normally occur in the nucleic acids of the particular biological sample are assayed. Suitable sample preparation/amplification control probes include, for example, probes to bacterial genes (e.g., Bio B) where the sample in question is a biological from a eukaryote.

The RNA sample is then spiked with a known amount of the nucleic acid to which the sample preparation/amplification control probe is directed before processing. Quantification of the hybridization of the sample preparation/amplification control probe then provides a measure of alteration in the abundance of the nucleic acids caused by processing steps (e.g. PCR, reverse transcription, *in vitro* transcription, etc.).

In a preferred embodiment, oligonucleotide probes in the high-density array are selected to bind specifically to the nucleic acid target to which they are directed with minimal non-specific binding or cross-hybridization under the particular hybridization conditions utilized. Because the high-density arrays of this invention can contain in excess of 1,000,000 different probes, it is possible to provide every probe of a characteristic length that binds to a particular nucleic acid sequence.

In addition, in a preferred embodiment, expression monitoring arrays are used to identify the presence and expression (transcription) level of genes which are several hundred base pairs long. For most applications it would be useful to identify the presence, absence, or expression level of several thousand to one hundred thousand genes. Because the number of oligonucleotides per array is limited in a preferred embodiment, it is desired to include only a limited set of probes specific to each gene whose expression is to be detected.

As disclosed in U.S. Application Ser. No. 08/772,376, probes as short as 15, 20, or 25 nucleotide are sufficient to hybridize to a subsequence of a gene and that, for most genes, there is a set of probes that performs well across a wide range of target nucleic acid concentrations. In a preferred embodiment, it is desirable to choose a preferred or "optimum" subset of probes for each gene before synthesizing the high-density array.

## 2. Hybridization

Nucleic acid hybridization simply involves contacting a probe and target nucleic acid under conditions where the probe and its complementary target can form stable hybrid duplexes through complementary base pairing. The nucleic acids that do not form hybrid duplexes are then washed away leaving the hybridized nucleic acids to be detected, typically through detection of an attached detectable label. It is generally recognized that nucleic acids are denatured by increasing the temperature or decreasing the salt concentration of the buffer containing the nucleic acids. Under low stringency conditions (e.g., low temperature and/or high salt) hybrid duplexes (e.g., DNA:DNA, RNA:RNA, or RNA:DNA) will form even where the annealed sequences are not perfectly complementary. Thus specificity of hybridization is reduced at lower stringency. Conversely, at higher stringency (e.g., higher temperature or lower salt) successful hybridization requires fewer mismatches.

One of skill in the art will appreciate that hybridization conditions may be selected to provide any degree of stringency. In a preferred embodiment, hybridization is performed at low stringency in this case in 6X SSPE-T at 37 C (0.005% Triton X-100) to ensure hybridization and then subsequent washes are performed at higher stringency (e.g., 1 X SSPE-T at 37 C) to eliminate mismatched hybrid duplexes. Successive washes may be performed at increasingly higher stringency (e.g., down to as low as 0.25 X SSPE-T at 37 C to 50 C) until a desired

level of hybridization specificity is obtained. In another preferred embodiment 1 X MES buffer is used. The buffer comprises 0.1 M 2-[N-Morpholinio]ethanesulfonic acid, 1.0 M NaCl, and 0.01 % Triton-X 100<sup>TM</sup>. Stringency can also be increased by addition of agents such as formamide. Hybridization specificity may be evaluated by comparison of hybridization to the test probes with hybridization to the various controls that can be present (e.g., expression level control, normalization control, mismatch controls, etc.).

In general, there is a tradeoff between hybridization specificity (stringency) and signal intensity. Thus, in a preferred embodiment, the wash is performed at the highest stringency that produces consistent results and that provides a signal intensity greater than approximately 10% of the background intensity. Thus, in a preferred embodiment, the hybridized array may be washed at successively higher stringency solutions and read between each wash. Analysis of the data sets thus produced will reveal a wash stringency above which the hybridization pattern is not appreciably altered and which provides adequate signal for the particular oligonucleotide probes of interest.

The stability of duplexes formed between RNAs or DNAs are generally in the order of RNA:RNA > RNA:DNA > DNA:DNA, in solution. Long probes have better duplex stability with a target, but poorer mismatch discrimination than shorter probes (mismatch discrimination refers to the measured hybridization signal ratio between a perfect match probe and a single base mismatch probe). Shorter probes (e.g., 8-mers) discriminate mismatches very well, but the overall duplex stability is low.

Altering the thermal stability ( $T_m$ ) of the duplex formed between the target and the probe using, e.g., known oligonucleotide analogues allows for optimization of duplex stability and mismatch discrimination. One useful aspect of altering the  $T_m$  arises from the fact that adenine-thymine (A-T) duplexes have a lower  $T_m$  than guanine-cytosine (G-C) duplexes, due in part to the fact that the A-T duplexes have 2 hydrogen bonds per base-pair, while the G-C duplexes have 3 hydrogen bonds per base pair. In heterogeneous oligonucleotide arrays in which there is a non-uniform distribution of bases, it is not generally possible to optimize hybridization for each oligonucleotide probe simultaneously. Thus, in some embodiments, it is desirable to selectively destabilize G-C duplexes and/or to increase the stability of A-T duplexes. This can be accomplished, e.g., by substituting guanine residues in the probes of an array which form G-C duplexes with hypoxanthine, or by substituting adenine residues in probes which form A-T duplexes with 2,6

diaminopurine or by using the salt tetramethyl ammonium chloride (TMACl) in place of NaCl.

Altered duplex stability conferred by using oligonucleotide analogue probes can be ascertained by following, e.g., fluorescence signal intensity of oligonucleotide analogue arrays hybridized with a target oligonucleotide over time. The data allow optimization of specific hybridization conditions at, e.g., room temperature (for simplified diagnostic applications in the future).

Another way of verifying altered duplex stability is by following the signal intensity generated upon hybridization with time. Previous experiments using DNA targets and DNA chips have shown that signal intensity increases with time, and that the more stable duplexes generate higher signal intensities faster than less stable duplexes. The signals reach a plateau or "saturate" after a certain amount of time due to all of the binding sites becoming occupied. These data allow for optimization of hybridization, and determination of the best conditions at a specified temperature.

Methods of optimizing hybridization conditions are well known to those of skill in the art (see, e.g., *Laboratory Techniques in Biochemistry and Molecular Biology, Vol. 24: Hybridization With Nucleic Acid Probes*, P. Tijssen, ed. Elsevier, N.Y., (1993)).

### (C) Signal Detection

In a preferred embodiment, the hybridized nucleic acids are detected by detecting one or more labels attached to the sample nucleic acids. The labels may be incorporated by any of a number of means well known to those of skill in the art. However, in a preferred embodiment, the label is simultaneously incorporated during the amplification step in the preparation of the sample nucleic acids. Thus, for example, polymerase chain reaction (PCR) with labeled primers or labeled nucleotides will provide a labeled amplification product. In a preferred embodiment, transcription amplification, as described above, using a labeled nucleotide (e.g. fluorescein-labeled UTP and/or CTP) incorporates a label into the transcribed nucleic acids.

Alternatively, a label may be added directly to the original nucleic acid sample (e.g., mRNA, polyA mRNA, cDNA, etc.) or to the amplification product after the amplification is completed. Means of attaching labels to nucleic acids are well known to those of skill in the art and include, for example nick translation or

end-labeling (e.g. with a labeled RNA) by kinasing of the nucleic acid and subsequent attachment (ligation) of a nucleic acid linker joining the sample nucleic acid to a label (e.g., a fluorophore).

Detectable labels suitable for use in the present invention include any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels in the present invention include biotin for staining with labeled streptavidin conjugate, magnetic beads (e.g., Dynabeads<sup>TM</sup>), fluorescent dyes (e.g., fluorescein, texas red, rhodamine, green fluorescent protein, and the like), radiolabels (e.g., <sup>3</sup>H, <sup>125</sup>I, <sup>35</sup>S, <sup>14</sup>C, or <sup>32</sup>P), enzymes (e.g., horse radish peroxidase, alkaline phosphatase and others commonly used in an ELISA), and colorimetric labels such as colloidal gold or colored glass or plastic (e.g., polystyrene, polypropylene, latex, etc.) beads.

Patents teaching the use of such labels include U.S. Patent Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241.

Means of detecting such labels are well known to those of skill in the art. Thus, for example, radiolabels may be detected using photographic film or scintillation counters, fluorescent markers may be detected using a photodetector to detect emitted light. Enzymatic labels are typically detected by providing the enzyme with a substrate and detecting the reaction product produced by the action of the enzyme on the substrate, and colorimetric labels are detected by simply visualizing the colored label. One particular preferred methods uses colloidal gold label that can be detected by measuring scattered light.

The label may be added to the target (sample) nucleic acid(s) prior to, or after the hybridization. So called "direct labels" are detectable labels that are directly attached to or incorporated into the target (sample) nucleic acid prior to hybridization. In contrast, so called "indirect labels" are joined to the hybrid duplex after hybridization. Often, the indirect label is attached to a binding moiety that has been attached to the target nucleic acid prior to the hybridization. Thus, for example, the target nucleic acid may be biotinylated before the hybridization. After hybridization, an avidin-conjugated fluorophore will bind the biotin bearing hybrid duplexes providing a label that is easily detected. For a detailed review of methods of labeling nucleic acids and detecting labeled hybridized nucleic acids see *Laboratory Techniques in Biochemistry and Molecular Biology, Vol. 24: Hybridization With Nucleic Acid Probes*, P. Tijssen, ed. Elsevier, N.Y., (1993)).

Fluorescent labels are preferred and easily added during an *in vitro* transcription reaction. In a preferred embodiment, fluorescein labeled UTP and

CTP are incorporated into the RNA produced in an *in vitro* transcription reaction as described above.

Means of detecting labeled target (sample) nucleic acids hybridized to the probes of the high-density array are known to those of skill in the art. Thus, for example, where a colorimetric label is used, simple visualization of the label is sufficient. Where a radioactive labeled probe is used, detection of the radiation (e.g. with photographic film or a solid state detector) is sufficient.

In a preferred embodiment, however, the target nucleic acids are labeled with a fluorescent label and the localization of the label on the probe array is accomplished with fluorescent microscopy. The hybridized array is excited with a light source at the excitation wavelength of the particular fluorescent label and the resulting fluorescence at the emission wavelength is detected. In a particularly preferred embodiment, the excitation light source is a laser appropriate for the excitation of the fluorescent label.

The confocal microscope may be automated with a computer-controlled stage to automatically scan the entire high-density array. Similarly, the microscope may be equipped with a phototransducer (e.g., a photomultiplier, a solid state array, a CCD camera, etc.) attached to an automated data acquisition system to automatically record the fluorescence signal produced by hybridization to each oligonucleotide probe on the array. Such automated systems are described at length in U.S. Patent No: 5,143,854, PCT Application 2092/10092, and copending U.S. Application Ser. No. 08/195,889 filed on February 10, 1994. Use of laser illumination in conjunction with automated confocal microscopy for signal detection permits detection at a resolution of better than about 100  $\mu\text{m}$ , more preferably better than about 50  $\mu\text{m}$ , and most preferably better than about 25  $\mu\text{m}$ .

One of skill in the art will appreciate that methods for evaluating the hybridization results vary with the nature of the specific probe nucleic acids used as well as the controls provided. In the simplest embodiment, simple quantification of the fluorescence intensity for each probe is determined. This is accomplished simply by measuring probe signal strength at each location (representing a different probe) on the high-density array (e.g., where the label is a fluorescent label, detection of the amount of fluorescence (intensity) produced by a fixed excitation illumination at each location on the array). Comparison of the absolute intensities of an array hybridized to nucleic acids from a "test" sample with intensities produced by a "control" sample provides a measure of the relative expression of the nucleic acids that hybridize to each of the probes.

One of skill in the art, however, will appreciate that hybridization signals will vary in strength with efficiency of hybridization, the amount of label on the sample nucleic acid and the amount of the particular nucleic acid in the sample. Typically nucleic acids present at very low levels (e.g., < 1pM) will show a very weak signal. At some low level of concentration, the signal becomes virtually indistinguishable from background. In evaluating the hybridization data, a threshold intensity value may be selected below which a signal is not counted as are essentially indistinguishable from background.

Where it is desirable to detect nucleic acids expressed at lower levels, a lower threshold is chosen. Conversely, where only high expression levels are to be evaluated a higher threshold level is selected. In a preferred embodiment, a suitable threshold is about 10% above that of the average background signal.

In addition, the provision of appropriate controls permits a more detailed analysis that controls for variations in hybridization conditions, cell health, non-specific binding and the like. Thus, for example, in a preferred embodiment, the hybridization array is provided with normalization controls. These normalization controls are probes complementary to control sequences added in a known concentration to the sample. Where the overall hybridization conditions are poor, the normalization controls will show a smaller signal reflecting reduced hybridization. Conversely, where hybridization conditions are good, the normalization controls will provide a higher signal reflecting the improved hybridization. Normalization of the signal derived from other probes in the array to the normalization controls thus provides a control for variations in hybridization conditions. Typically, normalization is accomplished by dividing the measured signal from the other probes in the array by the average signal produced by the normalization controls. Normalization may also include correction for variations due to sample preparation and amplification. Such normalization may be accomplished by dividing the measured signal by the average signal from the sample preparation/amplification control probes (e.g., the Bio B probes). The resulting values may be multiplied by a constant value to scale the results.

As indicated above, the high-density array can include mismatch controls. In a preferred embodiment, there is a mismatch control having a central mismatch for every probe (except the normalization controls) in the array. It is expected that after washing in stringent conditions, where a perfect match would be expected to hybridize to the probe, but not to the mismatch, the signal from the mismatch controls should only reflect non-specific binding or the presence in the sample of

a nucleic acid that hybridizes with the mismatch. Where both the probe in question and its corresponding mismatch control both show high signals, or the mismatch shows a higher signal than its corresponding test probe, there is a problem with the hybridization and the signal from those probes is ignored. The difference in hybridization signal intensity between the target specific probe and its corresponding mismatch control is a measure of the discrimination of the target-specific probe. Thus, in a preferred embodiment, the signal of the mismatch probe is subtracted from the signal from its corresponding test probe to provide a measure of the signal due to specific binding of the test probe.

The concentration of a particular sequence can then be determined by measuring the signal intensity of each of the probes that bind specifically to that gene and normalizing to the normalization controls. Where the signal from the probes is greater than the mismatch, the mismatch is subtracted. Where the mismatch intensity is equal to or greater than its corresponding test probe, the signal is ignored. The expression level of a particular gene can then be scored by the number of positive signals (either absolute or above a threshold value), the intensity of the positive signals (either absolute or above a selected threshold value), or a combination of both metrics (e.g., a weighted average).

In some preferred embodiments, a computer system is used to compare the hybridization intensities of the perfect match and mismatch probes of each pair. If the gene is expressed, the hybridization intensity (or affinity) of a perfect match probe of a pair should be recognizably higher than the corresponding mismatch probe. Generally, if the hybridizations intensities of a pair of probes are substantially the same, it may indicate the gene is not expressed. However, the determination is not based on a single pair of probes, the determination of whether a gene is expressed is based on an analysis of many pairs of probes.

After the system compares the hybridization intensity of the perfect match and mismatch probes, the system indicates expression of the gene. As an example, the system may indicate to a user that the gene is either present (expressed), marginal or absent (unexpressed). Specific procedures for data analysis is disclosed in U.S. Application 08/772,376, previously incorporated for all purposes.

In addition to high-density nucleic acid arrays, other methods are also useful for massive gene expression monitoring. Differential display, described by Liang, P. and Pardee, A.B. (Differential Display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 257:967-971, 1992,

incorporated herein by reference for all purposes) provides a useful mean for distinguishing gene expression between two samples. Serial analysis of gene expression, described by Velculescu et al. (Serial Analysis of Gene Expression. *Science*, 270:484-487, 1995, incorporated herein by reference for all purposes) provides another method for quantitative and qualitative analysis of gene expression. Optical fiber oligonucleotide sensors, described by Ferguson et al. (A Fiber-optic DNA biosensor microarray for the analysis of gene expression. *Nature-Biotechnology* 14:1681-1684, 1996), can also be used for gene expression monitoring.

It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference for all purposes.

### Examples

To address the functional properties of WT1 and EWS-WT1 and examine their effects on endogenous target genes, we developed cells in which a tightly regulated inducible promoter allowed expression of the full length protein. Forced expression of WT1 and EWS-WT1 in these cells triggered apoptosis. Oligonucleotide array-based expression profiling revealed a small number of genes exhibiting altered expression following WT1 and EWS-WT1 induction.

To study the functional properties of WT1 and EWS-WT1, we established inducible, tetracycline-regulated expression (Gossen and Bujard, 1992) of WT1 and EWS-WT1 in U2OS osteosarcoma cells. We searched for endogenous genes whose expression levels might be altered following inducible expression of WT1 and EWS-WT1. Messenger RNA was isolated from cells at 12 hrs following WT1 and EWS-WT1 induction, biotinylated, and hybridized to high-density oligonucleotide arrays representing 6,800 known transcripts and expressed sequence tags (Lockhart et al., 1996; Wodicka et al., 1997). A gene was considered a candidate WT1 and EWS-WT1 target if it was reproducibly induced or repressed in at least two different experimental replicates following WT1 and EWS-WT1 induction, and it exhibited an expression changes of at least 3-fold.

In this so-called "tet-off" induction system, the recombinant gene is induced by withdrawal of tetracycline from the tissue culture medium. The purpose of these studies was to identify candidate functional mediators (CFMs) of these tumor suppressors. Upon induction of WT1 or EWS-WT1 expression, we identified 28 and 17 endogenous genes, respectively (of 7000 genes monitored), that displayed reproducible expression changes of 3-fold or greater.

By generating cell lines with tightly regulated inducible expression of WT1 and EWS-WT1, we have identified downstream pathways that are likely to contribute to its function as a tumor suppressor.

#### Hybridization-Based Assay for Generating Expression Profiles

Messenger RNA levels are determined by hybridization of complete mRNA populations to sets of arrays containing hundreds of thousands of chemically synthesized oligonucleotides. The oligonucleotides are synthesized *in situ* on glass supports using light-directed, solid-phase combinatorial chemistry. Because the arrays are designed and synthesized based on sequence information alone, they provide a direct link between genomic sequence and measurements of differential gene expression. The arrays measure 1.28 x 1.28 cm and contain more than 65,000 50 x 50 micron synthesis features. Each synthesis feature consists of more than  $10^7$  copies of a particular 25-mer oligonucleotide.

For each mRNA sample, the expression levels of 6,800 full-length human genes will be monitored. The full set of oligonucleotides covering these genes is divided over 4 different arrays, comprising a total area of approximately one square inch. For each gene, approximately 20 complementary 25-mers are chosen based on automated selection criteria. The criteria include tests for sequence uniqueness relative to the rest of the genome and the absence of sequence features (e.g., self-complementarity or clusters of single nucleotides) that have been determined to adversely affect hybridization behavior on arrays. The use of sets of oligonucleotides for each gene provides redundancy in the detection and analysis of the data, mitigates the potentially confounding effects of occasional cross-hybridization, and makes it so all oligonucleotides do not have to hybridize identically in order to obtain quantitative information. To further increase the sensitivity and specificity of detection, each complementary oligonucleotide (perfect match, or PM) is synthesized with a closely related mismatch (MM) partner in a physically adjacent position. The mismatch

partner is identical except for a single base difference at the central position of the 25-mer. The MM oligonucleotide of each pair serves as an internal control that allows consistent hybridization patterns (patterns of PM signals that are larger than the corresponding MM signals) to be recognized.

Quantitative image analysis is based on the average of the differences between the PM and MM partners, so that nonspecific and background contributions tend to cancel, while specific hybridization signals tend to add constructively across the set of oligonucleotide pairs for each gene. These hybridization signals are quantitative over three orders of magnitude, from 1:300,000 to 1:300.

Table 1b. Additional target genes induced by WT1(-KTS).

Accession number	Description	Fold induction by Chip
X65779	Acidic FGF	22.1
X17405	Cartilage linking protein-1	46.1
X58377	IL-11	18.8

Candidate WT1 Targets	
59807	NATURAL KILLER CELLS PROTEIN 4 PRECURSOR (HUMAN); contains element MSR1 repetitive element ;
25317	Human folate binding protein (FBP) mRNA, 3' end.
70956	Human HALPHA44 gene for alpha-tubulin, exons 1-3.
51758	Human mRNA for heat shock protein HSP70B.
17969	H.sapiens mRNA for 90K product.
36307	HEAT SHOCK 70 KD PROTEIN 1 (HUMAN);
30704	Human amphiregulin (AR) mRNA, complete cds, clones lambda-AR1 and lambda-AR2.
51558	PROCOLLAGEN ALPHA 1(I) CHAIN PRECURSOR (HUMAN).
14083	Human beta-migrating plasminogen activator inhibitor 1 mRNA, 3' end.
05878	Human radiated keratinocyte mRNA for cysteine protease inhibitor.
18951	H.sapiens mRNA for caveolin.
31776	Human brain natriuretic protein (BNP) gene, complete cds.
20859	Human leukemia virus receptor 1 (GLVR1) mRNA, complete cds.
71870	KERATIN, TYPE I CYTOSKELETAL 17 (HUMAN);
45051	TUBULIN ALPHA-5 CHAIN (Gallus gallus)
47864	PURINE NUCLEOSIDE PHOSPHORYLASE (HUMAN).
10550	Human Gm GTPase (gem) mRNA, complete cds.
14874	Human mRNA for adrenomedullin, complete cds.
38636	UROKINASE PLASMINOGEN ACTIVATOR SURFACE RECEPTOR, GPI-ANCHORED (HUMAN).
13097	Human mRNA for tissue type plasminogen activator.
496322	Human gravin (a cytoplasmic antigen recognized by myasthenia gravis sera) mRNA, 3' end.
51345	Human Jun-B mRNA for JUN-B protein.
70940	H.sapiens mRNA for elongation factor 1 alpha-2.
216077	HOMEOTIC GENE REGULATOR (Drosophila melanogaster)
57348	H.sapiens mRNA (clone 9112).
126683	Human interferon gamma treatment inducible mRNA.
317227	H.sapiens mRNA for transmembrane receptor protein.
R49231	MITOCHONDRIAL PHOSPHATE CARRIER PROTEIN PRECURSOR (HUMAN);

TABLE 3

	EWS
R60806	PROBABLE NUCLEAR ANTIGEN (Pseudorabies virus)
M28062	INTERLEUKIN-2 RECEPTOR BETA CHAIN PRECURSOR (HUMAN);
H20529	G1/S-SPECIFIC CYCLIN D1 (Homo sapiens)
T79475	PROBABLE NUCLEAR ANTIGEN (Pseudorabies virus)
M22348	Human neuron-specific gamma-2 enolase, complete cds.
L49054	Homo sapiens t(3;5)(q25.1;p34) fusion gene NPM-MLF1 mRNA, complete cds.
H62466	COLLAGEN ALPHA 3(VI) CHAIN (Gallus gallus)
U23852	Human T-lymphocyte specific protein tyrosine kinase p56lck (lck) aberrant mRNA, complete cds.
U20982	Human insulin-like growth factor binding protein-4 (IGFBP4) gene, promoter and complete cds.
L16510	Homo sapiens cathepsin B mRNA, complete cds.
T51534	CYSTATIN C PRECURSOR (HUMAN).
R42291	DUAL SPECIFICITY MITOGEN-ACTIVATED PROTEIN KINASE KINASE 2 (Homo sapiens)
T53118	NEURONAL-GLIAL CELL ADHESION MOLECULE PRECURSOR (Gallus gallus)
H10925	PIM-1 PROTO-ONCOGENE SERINE/THREONINE-PROTEIN KINASE (Mus musculus)
R54359	CARBOXYLESTERASE PRECURSOR (Mus musculus)
L11329	PROTEIN-TYROSINE PHOSPHATASE PAC-1 (HUMAN);
X02492	INTERFERON-INDUCED PROTEIN 6-16 PRECURSOR (HUMAN); contains L1 repetitive element ;

# Genes Induced or Repressed by WTI

TABLE 2

Gene	Gene Class	Fold Induction	Previous Cancer Link
Alpha-Tumornathin	Cytoskeleton	12	Migration, Invasion
Tissue Plasminogen Activator	Protease System	4	Migration, Invasion, metastasis
Urokinase Plasminogen Activator Receptor	Protease System	4	Migration, Invasion, metastasis
Urokinase Plasminogen Activator Receptor	Protease System	3	Migration, Invasion, metastasis
Cystatin	Protease System	6	Adherence, motility, levels correlated with cancer prognosis
Carcin	Mitofusin	74	Autocrine/independence growth
Amphiregulin	Growth Factor	8	Neoplasia, Invasion, metastasis
MUC-2 Binding Protein	Secreted	8	Cell adhesion, secreted by many tumors
Folate Receptor	Receptor	10	Overexpressed in 90% of ovarian cancers, enhances tumor growth
GEM	Cytosine Nucleotide Binding	4	Induced by oncogene-mediated transformation

**1 Abstract**

Methods are provided for diagnosing cancers, drug-screening, and functionally analyzing mutations involving the WT1 gene. The methods involve use of the newly identified set of genes which are regulated by WT1 as well as by the set of genes which are regulated by WT1 fusions to EWS. Monitoring expression levels of these sets of genes can be used as an indicator of the genetic status of the gene. It can also identify which have similar effects on down-stream genes.

**2 Representative Drawing**      None

专利名称(译)	肿瘤抑制基因WT1的下游基因		
公开(公告)号	<a href="#">JP20000308500A</a>	公开(公告)日	2000-11-07
申请号	JP2000064155	申请日	2000-02-02
[标]申请(专利权)人(译)	阿菲梅特里克斯公司		
申请(专利权)人(译)	Affymetrix公司公司		
[标]发明人	オリナージョナサン トルオンビビ ヘイバーダニエル リーシーン		
发明人	オリナージョナサン トルオンビビ ヘイバーダニエル リーシーン		
IPC分类号	G01N33/50 C07H21/02 C12N5/10 C12N15/00 C12N15/09 C12Q1/68 C12Q1/6827 C12R1/91 C40B30/06 C40B40/02 C40B50/06 G01N33/15 G01N33/53 G01N33/566 G01N33/574 G01N37/00 G06F19/00		
CPC分类号	C12Q1/6827 C12Q2600/158 Y02A90/26		
FI分类号	C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.M G01N33/566 G01N33/574.D C12N5/00.B C12N15/00.A C12N15/00.F C12N15/09.200 C12N5/00.102 C12N5/10 C12Q1/6827.C C12Q1/6827.Z C40B30/06 C40B40/02 C40B50/06 G01N37/00.102		
F-TERM分类号	2G045/AA25 2G045/AA26 2G045/AA35 2G045/CB01 2G045/CB02 2G045/DA12 2G045/DA13 2G045/DA14 2G045/DA78 2G045/FB02 2G045/JA01 4B024/AA01 4B024/AA12 4B024/AA19 4B024/AA20 4B024/CA04 4B024/CA09 4B024/CA12 4B024/DA02 4B024/HA14 4B063/QA06 4B063/QA07 4B063/QA08 4B063/QA17 4B063/QA19 4B063/QQ08 4B063/QQ43 4B063/QQ52 4B063/QQ53 4B063/QQ61 4B063/QR55 4B063/QR77 4B063/QR80 4B063/QR84 4B063/QS34 4B065/AA93X 4B065/AB01 4B065/AC20 4B065/BD50 4B065/CA44 4B065/CA46		
优先权	09/256301 1999-02-24 US		
外部链接	<a href="#">Espacenet</a>		

## 摘要(译)

要解决的问题：提供一种诊断癌症的方法，药物筛选以及有关WT1基因突变的功能分析方法。该方法涉及使用新鉴定的一组基因，这些基因受与EWS融合的WT1和WT1调控。监测这些基因集的表达水平可以用作基因遗传状态的指标。它还可以确定哪些对下游基因具有相似的作用。

KTS)によって誘導されるさらに別 \* \*の標的遺伝子

寄託番号	説明	チップによる 誘導倍率
X65779	酸性FGF	22.1
X17405	軟骨結合 蛋白質1	46.1
X58377	IL-11	18.8