

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
7 August 2008 (07.08.2008)

PCT

(10) International Publication Number  
**WO 2008/095050 A1**

(51) International Patent Classification:  
A61K 49/00 (2006.01) G01N 33/574 (2006.01)  
C12Q 1/68 (2006.01)

(74) Agent: **HOSTETLER, Michael, J.**; Wilson Sonsini  
Goodrich & Rosati, 650 Page Mill Road, Palo Alto,  
California 94304-1050 (US).

(21) International Application Number:  
PCT/US2008/052540

(81) Designated States (unless otherwise indicated, for every  
kind of national protection available): AE, AG, AL, AM,  
AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA,  
CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE,  
EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID,  
IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC,  
LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN,  
MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH,  
PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV,  
SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN,  
ZA, ZM, ZW.

(22) International Filing Date: 30 January 2008 (30.01.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/887,318 30 January 2007 (30.01.2007) US  
60/911,855 13 April 2007 (13.04.2007) US

(84) Designated States (unless otherwise indicated, for every  
kind of regional protection available): ARIPO (BW, GH,  
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,  
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,  
FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL,  
NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG,  
CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (for all designated States except US): **PHAR-  
MACYCLICS, INC.** [US/US]; 995 East Arques Avenue,  
Sunnyvale, California 94085 (US).

(72) Inventors; and

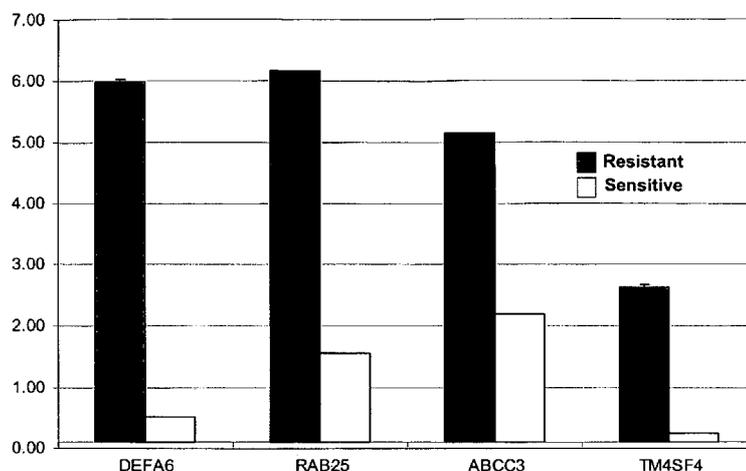
(75) Inventors/Applicants (for US only): **BUGGY, Joseph,  
J.** [US/US]; 858 Cuesta Drive, Mountain View, California  
94040 (US). **BALASUBRAMANIAN, Sriram** [US/US];  
1623 Howard Ave, San Carlos, California 94070 (US).

Published:  
— with international search report

[Continued on next page]

(54) Title: METHODS FOR DETERMINING CANCER RESISTANCE TO HISTONE DEACETYLASE INHIBITORS

**Fig. 6**



(57) Abstract: [00153] Described herein are methods and compositions for determining whether a particular cancer is resistant to or susceptible to a histone deacetylase inhibitor or to histone deacetylase inhibitors. The methods include analysis of the expression levels of at least four biomarker genes associated with response to a histone deacetylase inhibitor. Also described herein are methods and compositions for increasing the likelihood of a therapeutically effective treatment in a patient, comprising an analysis of the expression levels of at least four biomarker genes associated with response to a histone deacetylase inhibitor. Also described herein are isolated populations of nucleic acids derived from a cancer sensitive to or resistant to a histone deacetylase inhibitor. Further described are kits and indications that are used in conjunction with the aforementioned methods and compositions.

WO 2008/095050 A1



- 
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

## METHODS FOR DETERMINING CANCER RESISTANCE TO HISTONE DEACETYLASE INHIBITORS

### RELATED APPLICATIONS

5 [0001] This application claims the benefit of U.S. Provisional Patent Application No. 60/887,318, entitled "Methods for determining cancer resistance to histone deacetylase inhibitors," filed January 30, 2007, and U.S. Provisional Patent Application No. 60/911,855 entitled "Methods for determining cancer resistance to histone deacetylase inhibitors," filed April 13, 2007, the contents of both of which are incorporated by reference in their entirety.

### BACKGROUND OF THE INVENTION

10 [0002] The highly heterogeneous response of the same type of cancer (e.g., colon cancer) to a given anti-cancer compound in different patients is one of the most vexing and tragic problems of modern medicine. It is widely thought that human genetic and epigenetic diversity underlies much of the variation in response to chemotherapy. Thus, there is an ongoing effort to identify in the human  
15 population the molecular genetic correlates (i.e., molecular signatures) of cancer resistance and sensitivity to specific therapeutic agents. It is hoped that such efforts will ultimately enable physicians to predetermine the likelihood that a patient's cancer can be effectively treated with a particular anti-cancer compound.

### SUMMARY OF THE INVENTION

20 [0003] Described herein are methods and compositions for classifying a cancer in a patient as resistant or sensitive to a histone deacetylase inhibitor (HDACi) compound by (i) comparing the expression levels of at least four biomarker genes to a first set of biomarker gene expression level values, which was determined in cancer cells known to be resistant to the HDACi compound, or by comparing the expression levels to a second set of biomarker gene expression level values, which was  
25 determined in cancer cells known to be sensitive to the HDACi compound, and (ii) indicating that the cancer is sensitive to the HDACi compound if the biomarker gene expression levels are significantly lower than the first set of expression level values, or indicating that the cancer is resistant to the HDACi compound if the biomarker gene expression levels are greater than the second set of expression level values. The referred-to biomarker genes include PTPN3, ABCC3, SARG, PPAP2C,  
30 NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBP1, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMPRSS2, GDA, MST1R, ITGB4, ANXA3, CCL15, DPEP1, NOXO1, IFI27, CYP3A43, and PKP2.

[0004] Accordingly, in one aspect provided herein is a method for classifying a cancer in a patient, comprising comparing the expression levels of at least four biomarker genes in the cancer to expression level to a first or second set of expression level threshold values for the biomarker genes, and indicating that the cancer is sensitive to a HDAC inhibitor if the expression levels of the

5 biomarker genes are lower than the first set of expression level threshold values, or indicating that the cancer is resistant to a HDAC inhibitor if the expression levels are greater than the second set of expression level threshold values, wherein the at least four biomarker genes are selected from PTPN3, ABCC3, SARG, PPAP2C, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA,

10 FGFBP1, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMPRSS2, GDA, MST1R, ITGB4, ANXA3, CCL15, DPEP1, NOXO1, IFI27, CYP3A43, and PKP2. In some embodiments, the at least four marker genes are selected from DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP1. In some embodiments, the at least four biomarker genes include at least one of

15 DEFA6, RAB25, TM4SF4, or IL18. In some embodiments, the at least four biomarker genes include DEFA6, ITGB4, TM4SF3, SYK, PPAP2C, and RAB25. In some embodiments, the at least four biomarker genes include DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP1. In some embodiments, one or more of the above-mentioned expression levels is an mRNA expression level. In some

20 embodiments, one or more of the expression levels is a polypeptide expression level. In some embodiments, the patient's cancer is a colon cancer. In some embodiments, the method for classifying the cancer further comprises determining the level of expression of the at least four biomarker genes in the cancer prior to the step of comparing. In some embodiments, the referred-to HDAC inhibitor is PCI-24781. In some embodiments, the expression levels of the at least four

25 biomarker genes are compared to the first set and the second set of biomarker gene expression level threshold level values.

[0005] In another aspect provided herein is a method for classifying a cancer in a patient, comprising determining the expression levels of at least four biomarker genes in the cancer, comparing the expression levels of the at least four biomarker genes in the cancer to expression level

30 to a first or second set of expression level threshold values for the biomarker genes, and indicating that the cancer is sensitive to a HDAC inhibitor if the expression levels of the biomarker genes are lower than the first set of expression level threshold values, or indicating that the cancer is resistant to a HDAC inhibitor if the expression levels are greater than the second set of expression level threshold values, wherein the at least four biomarker genes are selected from PTPN3, ABCC3, SARG,

35 PPAP2C, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBP1, PTK6,

EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMPRSS2, GDA, MST1R, ITGB4, ANXA3, CCL15, DPEP1, NOXO1, IFI27, CYP3A43, and PKP2.

[0006] In some embodiments, at least one of the at least four marker genes are selected from DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP1. In some embodiments, the at least four biomarker genes include at least one of DEFA6, RAB25, TM4SF4, or IL18. In some embodiments, the at least four biomarker genes include DEFA6, ITGB4, TM4SF3, SYK, PPAP2C, and RAB25. In some embodiments, the at least four biomarker genes include DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP1.

In some embodiments, wherein one or more of the expression levels of the referred-to biomarker genes is an mRNA expression level. In some embodiments, one or more of the expression levels is a polypeptide expression level. In some embodiments, the patient's cancer is a colon cancer. In some embodiments, the HDAC inhibitor is PCI-24781. In some embodiments, the method further comprises prescribing or administering an HDAC inhibitor to the patient based on the comparison of the biomarker gene expression levels. In some embodiments, the expression levels of the at least four biomarker genes are compared to the first set and the second set of biomarker gene expression level threshold level values.

[0007] In a further aspect provided herein is an isolated population of nucleic acids comprising a plurality of nucleic acids derived from a cancer cell, wherein the cancer cell is a type of cancer cell that is sensitive to an HDAC inhibitor compound. In some embodiments, the isolated population contains RNAs. In some embodiments, the isolated population contains cDNAs. In some embodiments, the referred-to HDAC inhibitor is PCI-24781. In some embodiments, the referred-to cancer cell was isolated from a population of cells grown *in vitro*. In some embodiments, the cancer cell is a colon carcinoma cell. In some embodiments, the colon carcinoma cell is derived from colon carcinoma R1059261097, R4498160614, R5456781761, R7424107588, or R0948311023. In some embodiments, the nucleotide sequences of at least four of DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, or DPEP1 are represented in the isolated population of nucleic acids.

[0008] In a related aspect provided herein is an isolated population of nucleic acids comprising a plurality of nucleic acids derived from a cancer cell, wherein the cancer cell is a type of cancer cell that is resistant to an HDAC inhibitor compound. In some embodiments, the isolated population contains RNAs. In some embodiments, the isolated population contains cDNAs. In some embodiments, the referred-to HDAC inhibitor is PCI-24781. In some embodiments, the referred-to cancer cell was isolated from a population of cells grown *in vitro*. In some embodiments, the cancer cell is a colon carcinoma cell. In some embodiments, the colon carcinoma cell is derived from colon carcinoma R1059261097, R4498160614, R5456781761, R7424107588, or R0948311023. In some embodiments, the nucleotide sequences of at least four of DEFA6, ITGB4, TM4SF4, SYK, PPAP2C,

RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, or DPEP1 are represented in the isolated population of nucleic acids.

[0009] In some embodiments provided herein is a kit comprising the above referred-to isolated population of nucleic acids and an insert indicating the ratio of a biomarker gene nucleic acid level in the population to an internal expression control gene nucleic acid level in the population.

[0010] In some embodiments provided herein is a kit comprising the above referred-to isolated population of nucleic acids and an insert indicating the ratio of a biomarker gene nucleic acid level in the population to a nucleic acid level of the biomarker gene in a population of nucleic acids derived from a cancer cell, wherein the cancer cell is a type of cancer cell that is sensitive to the HDAC inhibitor compound.

[0011] In another aspect provided herein is a method for generating an expression level reference population of nucleic acids for expression profiling, comprising deriving an isolated population of nucleic acids from a cancer cell, wherein the cancer cell is a type of cancer cell that is sensitive to an HDAC inhibitor compound. In some embodiments, the isolated population contains RNAs. In some embodiments, the isolated population contains cDNAs. In some embodiments, the just-referred to HDAC inhibitor compound is PCI-24781. In some embodiments, the cancer cell is present in a biopsy sample. In some embodiments, the cancer cell is present in a population of cells grown *in vitro*. In some embodiments, the cancer cell is a colon carcinoma cell. In some embodiments, the carcinoma cell is derived from colon carcinoma R1059261097, R4498160614, R5456781761, R7424107588, or R0948311023. In some embodiments, the nucleotide sequences of at least four of DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, or DPEP1 are represented in the above referred-to isolated population of nucleic acids. In some embodiments, the method further comprises determining, prior to the isolating step, that the type of cancer cell is sensitive to an HDAC inhibitor compound. In some embodiments, the type of cancer cell determined to be sensitive to an HDAC inhibitor compound HDAC inhibitor compound *in vitro*. In some embodiments, the HDAC inhibitor compound is PCI-24781.

[0012] In a related aspect provided herein is a method for generating an expression level reference sample for expression profiling, comprising deriving an isolated population of nucleic acids from a cancer cell, wherein the cancer cell is a type of cancer cell that is resistant to an HDAC inhibitor compound. In some embodiments, the isolated population contains RNAs. In some embodiments, the isolated population contains cDNAs. In some embodiments, the just-referred to HDAC inhibitor compound is PCI-24781. In some embodiments, the cancer cell is present in a biopsy sample. In some embodiments, the cancer cell is present in a population of cells grown *in vitro*. In some embodiments, the cancer cell is a colon carcinoma cell. In some embodiments, the carcinoma cell is derived from colon carcinoma R1059261097, R4498160614, R5456781761, R7424107588, or R0948311023. In some embodiments, the nucleotide sequences of at least four of

DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, or DPEP1 are represented in the above referred-to isolated population of nucleic acids. In some embodiments, the method further comprises determining, prior to the isolating step, that the type of cancer cell is resistant to an HDAC inhibitor compound. In some  
5 embodiments, the type of cancer cell determined to be resistant to an HDAC inhibitor compound HDAC inhibitor compound *in vitro*. In some embodiments, the HDAC inhibitor compound is PCI-24781.

[0013] In another aspect provided herein is a human cancer cell line that is resistant to an HDAC inhibitor compound *in vitro*. In some embodiments, the human cell line expresses DEFA6,  
10 ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP1. In some embodiments, the HDAC inhibitor compound to which the referred-to human cancer cell line is resistant is PCI 24781. In some embodiments, the PCI 24781-resistant human cancer cell line is resistant to a PCI 24781 concentration of at least about 1  $\mu$ M. In some embodiments, the human cancer cell line is a colon carcinoma cell line. In some  
15 embodiments, the colon carcinoma cell line is R5247682266, R9866135153, R1078103114, or R4712781606.

[0014] In a further aspect provided herein is a method for increasing the likelihood of therapeutically effective treatment of a cancer with an HDAC inhibitor, comprising providing an indication that a cancer in a patient is sensitive to treatment with an HDAC inhibitor if expression  
20 levels of at least four biomarker genes in a sample from the patient's cancer are lower than expression level threshold values for the four biomarker genes, or providing an indication that the cancer is resistant to treatment with the HDAC inhibitor if the expression levels of the biomarker genes are higher than the expression level threshold values, wherein the at least four biomarker genes are selected from PTPN3, ABCC3, SARG, PPAP2C, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3,  
25 GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBP1, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, Tmprss2, GDA, MST1R, ITGB4, ANXA3, CCL15, DPEP1, NOXO1, IFI27, CYP3A43, and PKP2, whereby the likelihood of therapeutically effective treatment of the cancer with the HDAC inhibitor is increased. In some embodiments, the indication is provided in a digital  
30 medium. In some embodiments, the indication is provided in a hardcopy medium. In some embodiments, the indication is a biomedical publication reference. In some embodiments, the indication refers to expression levels of at least two of the biomarker genes. In some embodiments, the at least four biomarker genes include DEFA6, RAB25, TM4SF4, or IL18. In some embodiments, the at least four biomarker genes include DEFA6, ITGB4, TM4SF3, SYK, PPAP2C, and RAB25. In  
35 some embodiments, the at least four biomarker genes include DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and

DPEP1. In some embodiments, the cancer is colon cancer. In some embodiments, the HDAC inhibitor is PCI-24781.

[0015] In yet another aspect provided herein is a method for optimizing selection of an anti-cancer agent for treating a cancer in combination with an HDAC inhibitor compound, by: (i) comparing a first set of biomarker genes the expression of which is correlated to resistance or sensitivity of the cancer to the anti-cancer agent to a second set of biomarker genes the expression of which is correlated with resistance to the HDAC inhibitor compound; and (ii) selecting the anti-cancer agent for treatment of the cancer in combination with the HDAC inhibitor if the biomarker genes in the first set are different from the biomarker genes in the second set, where the biomarker genes in the second set are DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP1. In some embodiments, the method further comprises comparing the expression level of the second set of biomarker genes in a plurality of cancer cells treated with the HDAC inhibitor together with a second anti-cancer agent.

[0016] In a further aspect provided herein is an indication of the likelihood of a therapeutically effective treatment of a cancer with an HDAC inhibitor compound, comprising a means of communicating an interpretation of expression levels of at least four biomarker genes selected from DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP. In some embodiments, the indication further comprises the expression levels of the at least four biomarker genes. In some embodiments, the means of communicating is a paper document or an electronic document. In some embodiments, the interpretation includes a biomedical publication reference. In some embodiments, the interpretation includes a graph. In some embodiments, the interpretation includes information that indicates that a cancer in a patient is sensitive to treatment with an HDAC inhibitor if expression levels of the biomarker genes in a sample from the patient's cancer are lower than expression level threshold values for the four biomarker genes, or information that indicates that the cancer is resistant to treatment with the HDAC inhibitor if the expression levels of the biomarker genes are higher than the expression level threshold values.

[0017] In another aspect provided herein is a method for determining the likelihood of effectively treating a cancer in a patient with an HDAC inhibitor compound, comprising: (i) determining in the cancer the expression levels of at least four biomarker genes selected from DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP; and (ii) comparing the expression levels of that at least four biomarker genes in the cancer to expression levels of the at least four biomarker genes in an expression level reference sample derived from cancer cells previously determined to be resistant to the HDAC inhibitor compound, wherein the likelihood of effectively treating the cancer is higher if the expression level of the at least four biomarkers in the cancer from the patient is lower than the expression levels of the biomarker genes in the expression level reference sample. In some

embodiments, the method further comprises selecting an anti-cancer agent other than an HDAC inhibitor compound for treating the cancer.

**[0018]** In yet another aspect provided herein is a method for classifying a cancer in a patient, comprising comparing the expression levels of at least four biomarker genes in the cancer to a first or second set of expression level values for the biomarker genes, and for each comparison assigning a probability to the biomarker gene expression level that the cancer in the patient is resistant to a histone deacetylase inhibitor compound, where: (i) the first set of expression level values were measured in cancer cells determined to be resistant to the histone deacetylase inhibitor compound; (ii) the second set of expression level values were measured in cancer cells determined to be sensitive to the histone deacetylase inhibitor compound; (iii) the assigned probability is inversely proportional to a negative deviation of the biomarker gene expression level from the first set of expression level values and directly proportional to a positive deviation of the biomarker gene expression level from the second set of expression level values; and

(iv) the at least four biomarker genes are selected from PTPN3, ABCC3, SARG, PPAP2C, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGF1P, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMRSS2, GDA, MST1R, ITGB4, ANXA3, CCL15, DPEP1, NOXO1, IFI27, CYP3A43, and PKP2.

**[0019]** In another aspect provided herein is a method for classifying a population of cells, comprising comparing the expression levels of at least four biomarker genes in the population of cells to a first or second set of expression level threshold values for the biomarker genes, and indicating that the population of cells is sensitive to a HDAC inhibitor if the expression levels of the biomarker genes are lower than the first set of expression level threshold values, or indicating that the population of cells is resistant to a HDAC inhibitor if the expression levels are greater than the second set of expression level threshold values, wherein the at least four biomarker genes are selected from PTPN3, ABCC3, SARG, PPAP2C, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGF1P, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMRSS2, GDA, MST1R, ITGB4, ANXA3, CCL15, DPEP1, NOXO1, IFI27, CYP3A43, and PKP2.

**[0020]** In another aspect provided herein is a method for determining HDAC inhibition *in vivo*, comprising determining the expression level of an HDAC inhibitor-responsive biomarker gene in a biological sample obtained from a subject after the subject had been administered an HDAC inhibitor compound, wherein the HDAC inhibitor-responsive biomarker genes are any of the genes listed in Table 5.

[0021] In another aspect provided herein is a method for determining the most responsive tissues and the tumors derived therefrom to an HDAC inhibitor, comprising: (i) providing a first tissue of the tissue type (including blood) at a first time point and administration of HDAC inhibitor compound to the first tissue by any applicable route at a first time point, (ii) providing a second tissue of the tissue type (including blood) at a second time point and administration of HDAC inhibitor compound to the second tissue by any applicable route at a second time point, and (iii) determining expression profiles in the first and second tissues for any of the genes listed in Table 5.

[0022] In a further aspect provided herein is a method for classifying one or more cells, comprising determining the expression levels of no more than four to fifty biomarker genes in the one or more cells, wherein at least four of the biomarker genes are selected from PTPN3, ABCC3, SARG, PPAP2C, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBP1, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMPRSS2, GDA, MST1R, ITGB4, ANXA3, CCL15, DPEP1, NOXO1, IFI27, CYP3A43, and PKP2. In some embodiments, the method further comprises comparing the expression levels of the four to fifty biomarker genes to a first or second set of expression level threshold values for the biomarker genes, and indicating that the cancer is sensitive to a HDAC inhibitor if the expression levels of the biomarker genes are lower than the first set of expression level threshold values, or indicating that the cancer is resistant to a HDAC inhibitor if the expression levels are greater than the second set of expression level threshold values. In some embodiments, the one or more cells are cancer cells. In some embodiments, the at least four biomarker genes are selected from DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP. In some embodiments, the method further comprises determining the expression levels of no more than four to twenty biomarker genes. In some embodiments, the method comprises determining the expression levels of no more than four biomarker genes. In some embodiments, the four biomarker genes consist of DEFA6, RAB25, TM4SF4, and IL18.

[0023] In yet another aspect provided herein is a nucleic acid hybridization array comprising nucleic acid probes that hybridize under high stringency hybridization conditions to nucleic acids of no more than four to fifty biomarker genes, wherein at least four of the biomarker genes are selected from PTPN3, ABCC3, SARG, PPAP2C, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBP1, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMPRSS2, GDA, MST1R, ITGB4, ANXA3, CCL15, DPEP1, NOXO1, IFI27, CYP3A43, and PKP2. In some embodiments, the nucleic acid hybridization array comprises at least four biomarker genes selected from DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP. In some embodiments, the at least four biomarker genes consist of DEFA6, RAB25, TM4SF4, and IL18.

[0024] It is to be understood that the methods and compositions described herein are not limited to the particular methodology, protocols, cell lines, constructs, and reagents described herein and as such may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the methods and compositions described herein, which will be limited only by the appended claims.

[0025] As used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly indicates otherwise.

[0026] The term "biomarker gene" refers to a gene whose expression or activity yields at least one expression product the level of which is quantitatively correlated to a phenotypic state of interest (e.g., drug resistance, pathology).

[0027] The term "detectable label" refers to a label which is observable using analytical techniques including, but not limited to, fluorescence, chemiluminescence, electron-spin resonance, ultraviolet/visible absorbance spectroscopy, mass spectrometry, nuclear magnetic resonance, magnetic resonance, and electrochemical methods.

[0028] The terms "differentially expressed gene," "differential gene expression," and their synonyms, which are used interchangeably, refer to a gene whose expression is upregulated or downregulated in a first cell population relative to the expression of the same gene in a second population of cells. Such differences are evidenced by, e.g., a change in mRNA levels, surface expression, secretion or other partitioning of a polypeptide. Differential gene expression includes, in some embodiments, a comparison of expression between two or more genes or their gene products, or a comparison of the ratios of the expression between two or more genes or their gene products, or even a comparison of two differently processed products of the same gene, which differ between two populations of cells. Differential expression includes both quantitative, as well as qualitative, differences in the temporal or cellular expression pattern in a gene or its expression products among, for example, normal and diseased cells, or among cells which have undergone different disease events or disease stages, or cells that are significantly sensitive or resistant to certain therapeutic drugs.

[0029] The term "fluorophore" refers to a molecule which upon excitation emits photons and is thereby fluorescent.

[0030] The phrase "gene amplification" refers to a process by which multiple copies of a gene or gene fragment are formed in a particular cell or cell line. The duplicated region (a stretch of amplified DNA) is often referred to as "amplicon." Frequently, the amount of the messenger RNA (mRNA) produced, i.e., the level of gene expression, also increases in proportion to the number of copies made of the particular gene.

[0031] The term "gene expression profiling," unless otherwise specified, is used in the broadest sense, and includes methods of quantification of a gene's mRNA or nucleic acids derived therefrom, and/or protein levels or peptides derived therefrom and/or protein functions in a biological sample.

[0032] The term "high stringency hybridization" refers to hybridization conditions of incubating at 68 °C for an hour, followed by washing 3 times for 20 minutes each at room temperature in 2X SSC and 0.1% SDS and twice at 50 °C in 0.1X SSC and 0.1% SDS, or any art-recognized equivalent hybridization conditions.

5 [0033] The term "internal expression control gene" refers to a gene the expression level of which is known to or expected to be very similar in cells that differ in one or more phenotypes, or which have been subjected to differing experimental treatments. For example, the expression of the gene HDAC3 is shown to be to very similar in colon cancer cells that are resistant or sensitive to treatment with an HDACi compound.

10 [0034] The term "isolated" refers to separating and removing a component of interest from components not of interest. Isolated substances are optionally in either a dry or semi-dry state, or in solution, including but not limited to an aqueous solution. The isolated component is optionally in a homogeneous state or the isolated component is optionally a part of a pharmaceutical composition that comprises additional pharmaceutically acceptable carriers and/or excipients. Purity and  
15 homogeneity are determined, for example, using analytical chemistry techniques including, but not limited to, polyacrylamide gel electrophoresis or high performance liquid chromatography. In addition, when a component of interest is isolated and is the predominant species present in a preparation, the component is described herein as substantially purified. The term "purified," as used herein, refers to a component of interest which is at least 85% pure, at least 90% pure, at least 95%  
20 pure, at least 99% or greater pure. By way of example only, nucleic acids or proteins are "isolated" when such nucleic acids or proteins are free of at least some of the cellular components with which it is associated in the natural state, or that the nucleic acid or protein has been concentrated to a level greater than the concentration of its *in vivo* or *in vitro* production.

[0035] The term "label" refers to a substance which is incorporated into a compound and is readily detected, whereby its physical distribution is detected and/or monitored.

[0036] The term "microarray" refers to an ordered arrangement of hybridizable array elements, preferably polynucleotide probes, on a substrate.

[0037] The term "nucleic acid" or "nucleic acid probe," when used in singular or plural, generally refers to any polyribonucleotide or polydeoxribonucleotide, which includes unmodified  
30 RNA or DNA or modified RNA or DNA. Thus, for instance, nucleic acids as defined herein include, without limitation, single- and double-stranded DNA, DNA including single- and double-stranded regions, single- and double-stranded RNA, and RNA including single- and double-stranded regions, hybrid molecules comprising DNA and RNA that are optionally single-stranded or, more typically, double-stranded or include single- and double-stranded regions. In addition, the term "nucleic acid" as  
35 used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions are optionally from the same molecule or from different molecules. The regions optionally include all of one or more of the molecules, but more typically involve only a

region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide. The term "nucleic acid" specifically includes cDNAs. The term includes DNAs (including cDNAs) and RNAs that contain one or more modified bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "nucleic acids" as referred to herein. DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritiated bases, are included within the term "nucleic acid" as defined herein. In general, the term "nucleic acid" embraces all chemically, enzymatically and/or metabolically modified forms of unmodified polynucleotides, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including simple and complex cells.

**[0038]** The term "oligonucleotide" refers to a relatively short polynucleotide, including, without limitation, single-stranded deoxyribonucleotides, single- or double-stranded ribonucleotides, RNA:DNA hybrids and double-stranded DNAs. Oligonucleotides, such as single-stranded DNA probe oligonucleotides, are often synthesized by chemical methods, for example using automated oligonucleotide synthesizers that are commercially available. However, oligonucleotides are optionally made by a variety of other methods, including in vitro recombinant DNA-mediated techniques and by expression of DNAs in cells and organisms.

**[0039]** The terms "prediction," "predicting," "prognostic," or "prognosis" are used herein to refer to the likelihood that a patient will respond either favorably or unfavorably to a drug (e.g., an anti-cancer compound) or set of drugs, and also the extent of those responses. The predictive methods of described herein are valuable tools in predicting if a patient suffering from a cancer is likely to respond favorably to an HDAC inhibitor compound treatment regimen alone or in combination with another therapeutic agent (e.g., a second anti-cancer compound).

**[0040]** The term "subject" or "patient" refers to an animal which is the object of treatment, observation or experiment. By way of example only, a subject includes, but is not limited to, a mammal including, but not limited to, a human.

**[0041]** The term "substantially purified" refers to a component of interest that is substantially or essentially free of other components which normally accompany or interact with the component of interest prior to purification. By way of example only, a component of interest is "substantially purified" when the preparation of the component of interest contains less than about 30%, less than about 25%, less than about 20%, less than about 15%, less than about 10%, less than about 5%, less than about 4%, less than about 3%, less than about 2%, or less than about 1% (by dry weight) of contaminating components. Thus, a "substantially purified" component of interest optionally has a purity level of about 70%, about 75%, about 80%, about 85%, about 90%, about 95%, about 96%, about 97%, about 98%, about 99% or greater.

[0042] The term “therapeutically effective amount” refers to the amount of a composition administered to a patient already suffering from a disease, condition or disorder, sufficient to cure or at least partially arrest, or relieve to some extent one or more of the symptoms of the disease, disorder or condition being treated. The effectiveness of such compositions depend conditions including, but not limited to, the severity and course of the disease, disorder or condition, previous therapy, the patient's health status and response to the drugs, and the judgment of the treating physician. By way of example only, therapeutically effective amounts are determined by methods, including but not limited to a dose escalation clinical trial.

[0043] The terms “treat,” “treating” or “treatment,” include alleviating, abating or ameliorating a disease or condition symptoms, preventing additional symptoms, ameliorating or preventing the underlying metabolic causes of symptoms, inhibiting the disease or condition, e.g., arresting the development of the disease or condition, relieving the disease or condition, causing regression of the disease or condition, relieving a condition caused by the disease or condition, or stopping the symptoms of the disease or condition. The terms “treat,” “treating” or “treatment”, include, but are not limited to, prophylactic and/or therapeutic treatments.

[0044] The term “tumor” or “cancer” refers to all neoplastic cell growth and proliferation, whether malignant or benign, and all pre-cancerous and cancerous cells and tissues.

[0045] Unless otherwise indicated, conventional methods of cell culture, protein chemistry, biochemistry, recombinant DNA techniques including gene amplification and hybridization techniques, mass spectroscopy, and pharmacology, are employed.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0046] FIG. 1 is an illustrative schematic flow diagram of a method for identifying biomarker genes for HDACi compound resistance in cancer cells based on gene expression profiling, and the clinical application of expression profiling of the identified biomarker genes.

[0047] FIG. 2 is an illustrative graph showing *in vitro* inhibition of cell proliferation versus concentration of the HDACi compound PCI-24781 for a series of colon carcinoma cell lines.

[0048] FIG. 3 is an illustrative flow diagram illustrating the statistical approach used to analyze microarray data to identify differentially expressed genes in populations of cancer cells resistant to a HDACi compound versus cancer cells that are sensitive to the compound.

[0049] FIG. 4 is an illustrative scatter plot illustrating principal component analysis of gene expression microarray data in HDACi compound-treated and untreated cancer cells, and sensitive and resistant cancer cells.

[0050] FIG. 5 is an illustrative bar graph comparing the results of a microarray method versus TaqMan<sup>®</sup> quantitative RT-PCR method for determining the ratio of mRNA expression levels for a series of identified HDACi compound resistance biomarker genes in PCI-24781-resistant versus PCI-24781 colon carcinoma cells.

[0051] FIG. 6 is an illustrative bar graph comparing relative expression levels of four HDACi compound resistance biomarker genes in cancer cells that are resistant to the HDAC inhibitor compound (PCI-24781) versus expression of the biomarker genes in cancer cells that are sensitive to the compound.

5 [0052] FIG. 7 (A) is an illustrative bar graph showing the time course of tubulin acetylation in peripheral blood mononuclear cells from mice treated with the HDAC inhibitor compound PCI-24781; (B) is a time course of the expression profile of genes whose mRNA levels are correlated with changes in tubulin acetylation.

10 [0053] FIG. 8 is an illustrative set of two line graphs illustrating the expression profiles of two HDAC inhibitor-responsive biomarker genes as determined by microarray analysis, quantitative RT-PCR, and immunoblotting.

[0054] FIG. 9 is an illustrative bar graph showing average *in vivo* mRNA levels in various tissues of five of the HDAC inhibitor-responsive biomarker genes at 3 and 8 hours post-HDAC inhibitor treatment.

15 [0055] FIG. 10 is an illustrative series of dose response curves for the effect of the HDAC inhibitor PCI-24781 on tumors derived from the indicated tumors.

[0056] FIG. 11 (A) is a series of line graphs illustrating the amount of *in vitro* growth inhibition by the HDAC inhibitor PCI-24781 of primary colon tumor cells derived from newly diagnosed, naive colon cancer patients; (B) is a series of line graphs illustrating the amount of *in vitro* growth inhibition by the HDAC inhibitor PCI-24781 of colon cancer cells derived from patients having advanced, metastatic colon tumors; (C) is a bar graph illustrating the correlation between tumor cell resistance to an HDAC inhibitor *in vitro* and the mRNA expression level of the HDAC resistance biomarker gene *DEFA6*.

#### DETAILED DESCRIPTION OF THE INVENTION

25 [0057] The methods described herein include classifying a cancer in a patient as resistant or sensitive to a histone deacetylase inhibitor (HDACi) compound by comparing the expression levels of at least four biomarker genes expressed in the cancer to biomarker gene expression level threshold values, as described herein. Where the expression levels of at least four biomarker genes are greater than the expression level threshold values, the cancer is indicated as being resistant to the HDACi compound. Conversely, if the expression levels of the at least four biomarker genes are lower than the expression level threshold values, the cancer is indicated to be sensitive to the HDACi compound.

30 [0058] Also described herein is a population of nucleic acids derived from a cancer cell, where the cancer cell is a type of cancer cell that is resistant to an HDACi compound. Further described herein is a population of nucleic acids derived from a cancer cell, where the cancer cell is a type of cancer cell that is sensitive to an HDACi compound. Also described herein are methods for generating these populations of nucleic acids. Such populations of nucleic acids are optionally used as

expression level reference standards for setting biomarker gene expression threshold levels as described herein. Further described herein are cell lines determined to be resistant to an HDACi compound. Also described herein are cell lines determined to be sensitive to an HDACi compound.

[0059] Also described herein is a method for increasing the likelihood of therapeutically effective treatment of a cancer with an HDACi compound by providing an indication that a cancer is sensitive to treatment with an HDACi compound if the expression levels of at least four of the biomarker genes described herein are lower than the expression level threshold values for those biomarker genes, or providing an indication that a cancer is resistant to treatment with an HDACi compound if the expression levels of at least four of the biomarker genes described herein are higher than the expression level threshold values for those biomarker genes.

[0060] Further described herein are methods for optimizing selection of an anti-cancer agent for treating cancer in combination with an HDACi compound by comparing a first set of biomarker genes the expression of which is correlated to resistance or sensitivity of the cancer to the anti-cancer agent to a second set of biomarker genes the expression of which is correlated with resistance to the HDACi compound, and then selecting the anti-cancer agent for treatment of the cancer in combination with the HDAC inhibitor only if all of the biomarker genes in the first set are different from the biomarker genes in the second set.

#### Identification of HDACi compound resistance biomarker genes (HDACiR-BGs)

[0061] Described herein are methods for identifying genes whose expression levels in cancer cells are significantly and consistently correlated with resistance of the cells to an HDACi compound. Such genes are termed HDACi compound resistance biomarker genes (HDACiR-BGs). In an exemplary embodiment, HDACiR-BGs are identified as follows.

[0062] The *ex-vivo* response of primary tumor cells (e.g., colon cancer cells) from various patients to an HDAC inhibitor is determined by culturing the cells in the presence of varying concentrations of the HDACi compound.

[0063] After determining the HDACi compound sensitivity the cancer cells from each patient, mRNA expression profiles are determined for HDACi-resistant and sensitive tumors. Total RNA is isolated and fluorescent probes are prepared and hybridized to a whole genome cDNA microarray (e.g., Codelink Human Whole Genome oligonucleotide microarrays containing ~55,000 unique probes; GE Healthcare Bio-Sciences Corp., Piscataway, NJ) according to the manufacturer's instructions. Following hybridization, the microarrays are scanned (e.g., in a GenePix 4000B scanner; Molecular Devices Corporation, Sunnyvale CA). The images are then processed with Codelink software and the data are normalized to the median.

[0064] The median-normalized microarray data are imported into a microarray data analysis program for principal component analysis (PCA) and hierarchical clustering analysis (e.g., Genespring software from Agilent). Multiple analysis methods are employed to provide additional confidence in the mRNA expression analysis. For multiple hypothesis correction, the q-values approach for false discovery rates (FDR) are optionally used as described in Storey *et al.* (2003), *Proc. Nat. Acad. Sci. USA*, 100:9440-9445. As a second analytical approach the Bayesian ANOVA approach described in Ishwaran *et al.* (2003), *J. Amer. Stat. Assoc.*, 98:438-455 is optionally used.

[0065] In the Bayesian ANOVA method, the contributions of irrelevant genes to the ANOVA model are selectively shrunk to balance total false detections against total false non-detections. The output is a Zcut score which identifies genes whose contribution to the ANOVA model is larger than the standard z-score. See Ishwaran *et al.*, *ibid.*, and the website at bamarray.com.

[0066] The just-described method and variants thereof is optionally used to identify biomarker genes for other specific phenotypic states, e.g., resistance to anti-cancer agents other than HDACi compounds.

[0067] HDACiR-BGs identified by the just-described methods include those listed in Table 1. The sequence for the mRNA of each of the listed genes is included herein in an appendix .

<b>Gene Name</b>	<b>Gene Symbol</b>	<b>GenBank Accession #</b>	<b>SEQ ID NO</b>
<b>PTPN3</b>	<b>PTPN3</b>	<b>AK096975</b>	<b>1</b>
<b>ATP-binding cassette, sub-family C (CFTR/MRP), member 3</b>	<b>ABCC3</b>	<b>NM_020037</b>	<b>2</b>
<b>specifically androgen-regulated protein</b>	<b>SARG</b>	<b>NM_023938</b>	<b>3</b>
<b>phosphatidic acid phosphatase type 2C</b>	<b>PPAP2C</b>	<b>NM_177526</b>	<b>4</b>
<b>neural proliferation, differentiation and control, 1</b>	<b>NPDC1</b>	<b>NM_015392</b>	<b>5</b>
<b>C-terminal tensin-like</b>	<b>CTEN</b>	<b>NM_032865</b>	<b>6</b>
<b>RAB25, member RAS oncogene family</b>	<b>RAB25</b>	<b>NM_020387</b>	<b>7</b>
<b>Hephaestin</b>	<b>HEPH</b>	<b>NM_138737</b>	<b>8</b>
<b>thiopurine S-methyltransferase</b>	<b>TPMT</b>	<b>NM_000367</b>	<b>9</b>
<b>plakophilin 3</b>	<b>PKP3</b>	<b>NM_007183</b>	<b>10</b>
<b>UDP-N-acetyl-alpha-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase 5 (GalNAc-T5)</b>	<b>GALNT5</b>	<b>NM_014568</b>	<b>11</b>
<b>calmodulin-like 4</b>	<b>CALML4</b>	<b>NM_033429</b>	<b>12</b>

<b>UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 12 (GalNAc-T12)</b>	<b>GALNT12</b>	<b>AK024865</b>	<b>13</b>
<b>thiamin pyrophosphokinase 1</b>	<b>TPK1</b>	<b>NM_022445</b>	<b>14</b>
<b>defensin, alpha 6, Paneth cell-specific</b>	<b>DEFA6</b>	<b>NM_001926</b>	<b>15</b>
<b>epithelial protein lost in neoplasm beta</b>	<b>EPLIN</b>	<b>NM_016357</b>	<b>16</b>
<b>chloride intracellular channel 5</b>	<b>CLIC5</b>	<b>NM_016929</b>	<b>17</b>
<b>PERP, TP53 apoptosis effector</b>	<b>PERP</b>	<b>NM_022121</b>	<b>18</b>
<b>spleen tyrosine kinase</b>	<b>SYK</b>	<b>NM_003177</b>	<b>19</b>
<b>solute carrier family 12 (sodium/potassium/chloride transporters), member 2</b>	<b>SLC12A2</b>	<b>NM_001046</b>	<b>20</b>
<b>guanylate cyclase 2C (heat stable enterotoxin receptor)</b>	<b>GUCY2C</b>	<b>NM_004963</b>	<b>21</b>
<b>transmembrane 4 superfamily member 4</b>	<b>TM4SF4</b>	<b>NM_004617</b>	<b>22</b>
<b>transforming growth factor, alpha</b>	<b>TGFA</b>	<b>NM_003236</b>	<b>23</b>
<b>fibroblast growth factor binding protein 1</b>	<b>FGFBP1</b>	<b>NM_005130</b>	<b>24</b>
<b>PTK6 protein tyrosine kinase 6</b>	<b>PTK6</b>	<b>NM_005975</b>	<b>25</b>
<b>epithelial V-like antigen 1</b>	<b>EVA1</b>	<b>NM_005797</b>	<b>26</b>
<b>EPH receptor A2</b>	<b>EPHA2</b>	<b>NM_004431</b>	<b>27</b>
<b>integrin, alpha 6</b>	<b>ITGA6</b>	<b>NM_000210</b>	<b>28</b>
<b>Table 1 (continued)</b>			
<b>Gene Name</b>	<b>Gene Symbol</b>	<b>GenBank Accession #</b>	<b>SEQ ID NO</b>
<b>tumor necrosis factor receptor superfamily, member 21</b>	<b>TNFRSF21</b>	<b>NM_014452</b>	<b>29</b>
<b>transmembrane 4 superfamily member 3</b>	<b>TM4SF3</b>	<b>NM_004616</b>	<b>30</b>
<b>interleukin 18 (interferon-gamma-inducing factor)</b>	<b>IL18</b>	<b>NM_001562</b>	<b>31</b>
<b>bone morphogenetic protein 4</b>	<b>BMP4</b>	<b>NM_130850</b>	<b>32</b>
<b>sphingomyelin phosphodiesterase, acid-like 3B</b>	<b>SMPDL3B</b>	<b>NM_014474</b>	<b>33</b>
<b>transmembrane protease, serine 2</b>	<b>TMPRSS2</b>	<b>NM_005656</b>	<b>34</b>
<b>guanine deaminase</b>	<b>GDA</b>	<b>NM_004293</b>	<b>35</b>
<b>macrophage stimulating 1 receptor (c-met-related tyrosine</b>	<b>MST1R</b>	<b>NM_002447</b>	<b>36</b>

<b>kinase)</b>			
<b>integrin, beta 4</b>	ITGB4	NM_000213	<b>37</b>
<b>annexin A3</b>	ANXA3	NM_005139	<b>38</b>
<b>chemokine (C-C motif) ligand 15</b>	CCL15	NM_032965	<b>39</b>
<b>dipeptidase 1 (renal)</b>	DPEP1	NM_004413	<b>40</b>
<b>NADPH oxidase organizer 1</b>	NOXO1	NM_172167	<b>41</b>
<b>interferon, alpha-inducible protein 27</b>	IFI27	NM_005532	<b>42</b>
<b>cytochrome P450, family 3, subfamily A, polypeptide 43</b>	CYP3A43	NM_057095	<b>43</b>
<b>plakophilin 2</b>	PKP2	NM_004572	<b>44</b>

Classification of individual patient cancers as resistant or sensitive to an HDACi compound

[0068] In some embodiments, gene expression profiling is performed on a biological sample obtained from an individual patient suffering from a cancer (e.g., a colon cancer tumor) to classify the cancer in the patient as resistant or sensitive to an HDACi compound. The gene expression profiling includes profiling the expression of at least one of the HDACi compound resistance biomarker genes (HDACiR-BGs) listed in Table 1, which were identified as described herein.

[0069] In some embodiments the HDACiR-BG is selected from among DEFA6, TM4SF4, TGFA, FGFBP1, EPHA2, TNFRSF2, TM4SF3, IL18, Tmprss2, and CCL15.

[0070] In some embodiments, at least four of the HDACiR-BGs are expression profiled. In some embodiments, at least one of the four HDACiR-BGs are selected from among DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF3, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, or DPEP1. In some embodiments, all of the at least four HDACiR-BGs are selected from among DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF3, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, or DPEP1.

[0071] In some embodiments, the expression of at least sixteen of the HDACiR-BGs is profiled. In some embodiments, the at least sixteen HDACiR-BGs include one or more of DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF3, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, or DPEP1. In some embodiments, the at least 16 HDACiR-BGs include DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF3, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, or DPEP1.

[0072] In various embodiments, the types of cancers and tumors that are optionally classified (from individual patients) for resistance or sensitivity to an HDACi compound include, but are not limited to, colorectal cancer, ovarian cancer, pancreatic cancer biliary tract cancer; bladder cancer; bone cancer; brain and CNS cancer; breast cancer; cervical cancer; choriocarcinoma; connective

tissue cancer; cancer of the digestive system; endometrial cancer; esophageal cancer; eye cancer; cancer of the head and neck; gastric cancer; intra-epithelial neoplasm; kidney cancer; larynx cancer; leukemia; liver cancer; lung cancer (e.g., small cell and non-small cell); lymphoma including Hodgkin's and non-Hodgkin's lymphoma; melanoma; myeloma; neuroblastoma; oral cavity cancer (e.g., lip, tongue, mouth, and pharynx); prostate cancer; retinoblastoma; rhabdomyosarcoma; rectal cancer; renal cancer; cancer of the respiratory system; sarcoma; skin cancer; stomach cancer; testicular cancer; thyroid cancer; uterine cancer; cancer of the urinary system, as well as other carcinomas and sarcomas.

[0073] Types of cancer cells that are optionally classified in various embodiments include, but are not limited to, squamous cell papilloma, squamous cell carcinoma, basal cell tumor, basal cell carcinoma, transitional cell papilloma, transitional cell carcinoma, glandular epithelium adenoma, melanocytes glomus tumor, melanocytic nevus, malignant melanoma, fibroma, fibrosarcoma, an adenocarcinoma, gastrinoma, malignant gastrinoma, an oncocytoma, cholangiocellular adenoma, cholangiocellular carcinoma, hepatocellular adenoma, hepatocellular carcinoma, renal tubular adenoma, renal cell carcinoma (Grawitz tumor), myxoma, myxosarcoma, lipoma, liposarcoma, leiomyoma, leiomyosarcoma, rhabdomyoma, rhabdomyosarcoma, benign teratoma, malignant teratoma, hemangioma, hemangiosarcoma, Kaposi sarcoma, lymphangioma, lymphangiosarcoma, an osteoma, an osteosarcoma, an osteogenic sarcoma, cartilage chondroma, chondrosarcoma, meninges meningioma, malignant meningioma, oligoastrocytoma, an ependymoma, an astrocytoma, pilocytic astrocytoma, glioblastomamultiforme, an oligodendroglioma, neuroblastoma, schwannoma, retinoblastoma, or neurofibroma. Other types of cancers and tumors include those described in reference sources, e.g., the "International Classification of Diseases for Oncology," 3rd Edition, International Association of Cancer Registries.

[0074] A biological sample is any biological sample that includes cellular material from which DNA, RNA or protein are optionally isolated, e.g., solid tissue samples, such as a biopsy specimen or tissue cultures or cells derived therefrom and the progeny thereof, blood and other liquid samples of biological origin, e.g., sputum (including saliva, buccal wash, or bronchial brush), stool, semen, urine, ascitic fluid, cerebral spinal fluid, bladder wash, or pleural fluid. The term "biological sample" also encompasses samples that have been manipulated in any way after their procurement, such as by treatment with reagents, solubilization, or enrichment for certain components. The term encompasses a clinical sample, and also includes cells in cell culture, cell supernatants, cell lysates, serum, plasma, biological fluids, and tissue samples, e.g., freshly collected tissue, frozen tissue, archived tissue, or biological fluids

[0075] In some embodiments, the biological sample is a tumor biopsy (e.g., a core biopsy, a needle biopsy, or an excisional biopsy) containing one or more cancer cells. In one embodiment the biological sample is a population of cancer cells obtained by laser capture dissection from a tumor tissue section as described in, e.g., U.S. Patent No. 6,040,139. Methods for optimizing tissue sample

preparation and processing for expression profiling include, e.g., Bova *et al.* (2005), *Methods Mol. Med.*, 103:15-66.

[0076] In some embodiments, one or more cells (e.g., from a cultured cancer cell line), are classified by determining the expression levels of no more than four to fifty biomarker genes described herein., e.g., 5, 6, 7, 8, 9, 10, 12, 16, 18, 20, 24, 30, 32, 35, 40, 44, 45, 47, or any other number of biomarker genes from four to fifty. In some embodiments, four to forty four of the biomarker genes are selected from Table 3, e.g., 5, 6, 7, 8, 9, 10, 12, 16, 18, 20, 24, 30, 32, 35, 40, or any other number of biomarker genes from four to forty four is selected from Table 3. In some embodiments, at least four of the biomarker genes are selected from PTPN3, ABCC3, SARG, PPAP2C, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGF1P, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMPRSS2, GDA, MST1R, ITGB4, ANXA3, CCL15, DPEP1, NOXO1, IFI27, CYP3A43, and PKP2. In some embodiments, the four to fifty biomarker comprises one or more genes selected from DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGF1P, ABCC3, TPMT, IL18, and DPEP. In some embodiments, classification of the cells comprises comparing the determined expression levels to a first or second set of expression level threshold values for the biomarker genes, and indicating that the one or more cells are sensitive to a HDAC inhibitor if the expression levels of the biomarker genes are lower than the first set of expression level threshold values, or indicating that the one or more cells are resistant to a HDAC inhibitor if the expression levels are greater than the second set of expression level threshold values. In some embodiments, the expression of no more than four to twenty biomarker genes is determined. In some embodiments, the expression levels of no more than four biomarker genes is determined. In some embodiments, the four biomarker genes the expression level of which is determined are: DEFA6, RAB25, TM4SF4, and IL18.

#### Methods for HDACiR-BG Expression Profiling

[0077] HDACiR-BG expression profiles are optionally generated by any convenient means for determining differential gene expression between two samples, e.g. quantitative hybridization of mRNA, labeled mRNA, amplified mRNA, cRNA, etc., quantitative PCR, ELISA for protein quantitation, and the like.

[0078] In some embodiments, HDACiR-BG mRNA levels (including cDNA copy or aRNA copies) are quantified. The expression profile is optionally generated from the initial nucleic acid sample using any convenient protocol. While a variety of different manners of generating expression profiles are known, such as those employed in the field of differential gene expression analysis, one representative and convenient type of protocol for generating expression profiles is array based gene expression profile generation protocols. Such applications are hybridization assays in which a nucleic acid that displays "probe" nucleic acids for each of the genes to be assayed/profiled in the profile to be generated is employed. In these assays, a sample of target nucleic acids is first prepared from the

initial nucleic acid sample being assayed, where preparation optionally includes labeling of the target nucleic acids with a label, e.g., a member of signal producing system. Following target nucleic acid sample preparation, the sample is contacted with the array under hybridization conditions, whereby complexes are formed between target nucleic acids that are complementary to probe sequences attached to the array surface. HDACiR-BG hybridization complexes are then detected and quantified.

5 [0079] Specific hybridization technologies which are optionally practiced to generate the HDACiR-BG expression profiles employed in the methods described herein includes the technology described in U.S. Pat. Nos. 5,143,854; 5,288,644; 5,324,633; 5,432,049; 5,470,710; 5,492,806; 5,503,980; 5,510,270; 5,525,464; 5,547,839; 5,580,732; 5,661,028; 5,800,992; as well as WO 10 95/21265; WO 96/31622; WO 97/10365; WO 97/27317; EP 373 203; and EP 785 280. In these methods, an array of "probe" nucleic acids that includes a probe for each of the phenotype determinative genes whose expression is being assayed is contacted with target nucleic acids as described above. Contact is carried out under hybridization conditions, e.g., stringent hybridization conditions as those conditions are practiced in the art, and unbound nucleic acid is then removed. The resultant pattern of hybridized nucleic acid provides quantitative information regarding expression for each of the HDACiR-BGs that have been probed.

15 [0080] Evaluation of differences in expression values is optionally performed using any convenient methodology, e.g., by comparing digital images of the expression profiles, by comparing databases of expression data, etc. Patents describing ways of comparing expression profiles include, but are not limited to, U.S. Pat. Nos. 6,308,170 and 6,228,575 and U.S. Patent Application Serial No. 20 10/858,867.

[0081] In some embodiments, the methods described herein are performed on nucleic acid hybridization arrays comprising nucleic acid probes that hybridize under high stringency hybridization conditions to nucleic acids of no more than four to fifty biomarker genes, e.g., 5, 6, 7, 8, 9, 10, 12, 16, 18, 20, 24, 30, 32, 35, 40, 44, 45, 47, or any other number of biomarker genes from four to fifty. In some embodiments, four to forty four of the biomarker genes are selected from Table 3, e.g., 5, 6, 7, 8, 9, 10, 12, 16, 18, 20, 24, 30, 32, 35, 40, or any other number of biomarker genes from four to forty four is selected from Table 3. In some embodiments, at least four of the biomarker genes for the array probes are selected from PTPN3, ABCC3, SARG, PPAP2C, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBP1, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, Tmprss2, GDA, MST1R, ITGB4, ANXA3, CCL15, DPEP1, NOXO1, IFI27, CYP3A43, and PKP2. In some embodiments, the at least four biomarker genes are selected from DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP. In some 35 embodiments, the at least four biomarker genes are DEFA6, RAB25, TM4SF4, and IL18.

[0082] Alternatively, non-array based methods for quantitating the levels of one or more nucleic acids in a sample are employed, including quantitative PCR, and the like.

[0083] In some embodiments, expression profiling of HDACiR-BGs expressed in a biological sample (e.g., a tumor biopsy) is done by a quantitative reverse transcription PCR assay (qRT-PCR). In this method, RNA from a biological sample is reverse transcribed to generate segments of cDNA which are then be amplified by gene-specific quantitative PCR. The rate of accumulation of specific PCR products is optionally correlated to the abundance of the corresponding RNA species in the original sample and thereby provide an indication of gene expression levels.

[0084] In one embodiment, the qPCR assay is a TaqMan™ assay. In brief, PCR typically utilizes the 5' exonuclease activity of Taq or Tth polymerase to hydrolyze a fluorescently-labelled hybridization probe bound to its target amplicon, but any enzyme with equivalent 5' exonuclease activity is optionally used. Two oligonucleotide primers are used to generate an amplicon typical of a PCR reaction. A third oligonucleotide, or probe, is designed to hybridize to a nucleotide sequence located between the two PCR primers. The probe is non-extendible by Taq DNA polymerase enzyme, and is 5' labeled with a reporter fluorescent dye and a 3' labeled with a quencher fluorescent dye. Any laser-induced emission from the reporter dye is quenched by the quenching dye when the two dyes are located close together as they are on the probe. During the amplification reaction, the Taq DNA polymerase enzyme cleaves the probe in a template-dependent manner. The resultant probe fragments disassociate in solution, and signal from the released reporter dye is free from the quenching effect of the second chromophore. One molecule of reporter dye is liberated for each new molecule synthesized, and detection of the unquenched reporter dye provides the basis for quantitative interpretation of the data.

[0085] qRT-PCR is optionally performed using commercially available equipment, such as, for example, the ABI PRISM 7900™ Sequence Detection System™ (Perkin-Elmer-Applied Biosystems, Foster City, CA), or LightCycler™. (Roche Molecular Biochemicals, Mannheim, Germany). In one embodiment, the 5' exonuclease procedure is run on a real-time quantitative PCR device such as the ABI PRISM 7900™ Sequence Detection System™ or one of the similar systems in this family of instruments. The system consists of a thermocycler, laser, charge-coupled device (CCD), camera and computer. The system amplifies samples in 96-well or 384 well formats on a thermocycler. During amplification, laser-induced fluorescent signal is collected in real-time through fiber optic cables for all reaction wells, and detected at the CCD. The system includes software for running the instrument and for analyzing the data.

[0086] Exonuclease assay data are initially expressed as a  $C_T$  value, i.e., the PCR cycle at which the fluorescent signal is first recorded as statistically significant.

[0087] In order to minimize errors and the effects of sample-to-sample variation and process variability mRNA level measurements are generally normalized to the expression level of an internal expression control gene. Methods for normalizing qPCR assays include, see, e.g., the website at normalisation.gene-quantification.info. The ideal internal expression control gene is one that is expressed at a relatively constant level among different patients or subjects, and is unaffected by the experimental treatment.

[0088] In some embodiments, the internal expression control gene is RNA polymerase II (GenBank Accession No. X74870).

[0089] In other embodiments, the internal expression control gene is HDAC3 (NM\_003883).

[0090] In further embodiments, the internal expression control gene is ZNF217 (NM\_006526).

[0091] In some embodiments, HDACiR-BG mRNA expression levels for each sample are normalized by the total amount of RNA in each sample. The amount of RNA in a sample is optionally determined, e.g., by UV-spectrophotometry or by using an RNA detection reagent, e.g., RiboGreen<sup>®</sup> from Invitrogen (Carlsbad, CA).

[0092] Where the HDACiR-BG expression profile to be determined is a protein expression profile, any convenient protein quantitation protocol is optionally employed, where the levels of one or more proteins in the assayed sample are determined. Representative methods include, but are not limited to; proteomic arrays, mass spectrometry, or standard immunoassays (e.g., RIA or ELISA). See, e.g., the methods set forth in R. Scopes, *Protein Purification*, Springer-Verlag, N.Y. (1982); Sandana (1997) *Bioseparation of Proteins*, Academic Press, Inc.; Bollag *et al.* (1996) *Protein Methods*. 2nd Edition Wiley-Liss, NY; Walker (1996) *The Protein Protocols Handbook* Humana Press, NJ, Harris and Angal (1990) *Protein Purification: Principles and Practice* 3rd Edition Springer Verlag, NY; Janson and Ryden (1998) *Protein Purification: Principles, High Resolution Methods and Applications*, Second Edition Wiley-VCH, NY; and Satinder Ahuja ed., *Handbook of Bioseparations*, Academic Press (2000); Harlow *et al.*, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 353-355 (1988).

[0093] Proteomic expression profiling methods detection methods include various multidimensional electrophoresis methods (e.g., 2-D gel electrophoresis), mass spectrometry based methods e.g., SELDI, MALDI, electrospray, etc.), or surface plasmon resonance methods. For example, in MALDI, a sample is usually mixed with an appropriate matrix, placed on the surface of a probe and examined by laser desorption/ionization. See, e.g., U.S. Pat. Nos. 5,045,694, 5,202,561, and 6,111,251. Similarly, for SELDI, a first aliquot is contacted with a solid support-bound (e.g., substrate-bound) adsorbent. A substrate is typically a probe (e.g., a biochip) that is optionally positioned in an interrogatable relationship with a gas phase ion spectrometer. SELDI has been applied to diagnostic proteomics. See, e.g. Issaq *et al.* (2003), *Anal. Chem.* 75: 149A-155A.

[0094] In one embodiment, any of the just-described protein detection methods are used to determine the expression level of one or more HDACiR-BG proteins that are known to be secreted proteins, e.g., DEFA6, TM4SF4, TM4SF3, TGFA, FGF1P1, EPHA2, TNFRSF2, IL18, CCL15, or TMPRSS2.

5

#### Expression Level Reference Samples

[0095] In some embodiments, expression profiles of HDACiR-BGs in a biological sample of interest (e.g., a colon cancer biopsy) are compared to HDACiR-BG expression profiles in an expression level reference sample. The expression level reference sample is a biological sample derived from one or more cancer patients determined to be suffering from a particular cancer or tumor for which sensitivity or resistance to treatment with an HDACi compound (e.g., PCI-24781) has been determined. In other words, the expression level reference sample serves as a standard with which to compare expression level values for each HDACiR-BG in a test sample. The deviation of HDACiR-BG expression levels from the expression level values in a reference sample indicates whether the cancer in the patient from the biological sample was derived is sensitive or resistant to treatment with an HDACi compound. In some embodiments, HDACiR-BG threshold expression level values are optionally set based on one or more statistical criteria for deviation from HDACiR-BG expression level values in an expression level reference sample, e.g., two or more SDs away from the value for a reference sample HDACiR-BG expression level.

10

15

20

[0096] In some embodiments, the expression level reference sample is a “negative” reference sample, i.e., a sample derived from a patient having a cancer or tumor determined to be sensitive to an HDACi compound. Thus, where expression levels of multiple HDACiR-BGs (e.g. at least 4, 5, 6, 8, 10, 12, or 16) are significantly greater than the threshold expression level values based on the negative reference sample, the patient’s cancer is indicated as resistant to the HDACi compound.

25

[0097] In some embodiments, the expression level reference sample is a “positive” reference sample, i.e., a sample derived from a patient having a cancer or tumor determined to be resistant to an HDACi compound. Thus, where expression levels of multiple HDACiR-BGs (e.g. at least 4, 5, 6, 8, 10, 12, or 16) are significantly lower than the threshold expression level values based on the negative reference sample, the patient’s cancer is indicated as sensitive to the HDACi compound.

30

[0098] In some embodiments, HDACiR-BG expression profiles are compared to those in both positive and negative reference samples.

[0099] In some embodiments, HDACiR-BGs expression level measurements are performed in parallel for the biological sample of interest and the (positive or negative) expression level reference. For example, where an array hybridization method is used, HDACiR-BG mRNA levels in the biological sample of interest and in an expression level reference sample are optionally measured simultaneously by separately labeling nucleic acid populations (e.g., mRNA, cDNA, aRNA

35

populations) from each with a detectably distinct fluorophore, and then hybridizing the fluorescently labeled nucleic acids to the same array.

[00100] In some embodiments an expression level reference sample is a population of nucleic acids (e.g., mRNAs, aRNAs, cDNAs, or aRNAs) derived from a cancer biopsy sample within which the sequences of at least four HDACiR-BGs are represented, and for which sensitivity to an HDACi compound has been determined. In some embodiments, the population of nucleic acids is derived from patient tumor cells cultivated in culture. In other embodiments, the population is derived directly from a biopsy without a cell culture step.

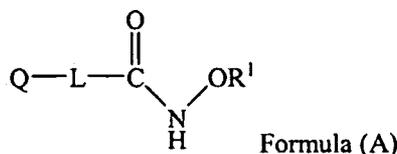
[00101] In some embodiments, the population of nucleic acids serving as an expression level reference sample is generated as follows. A cancer biopsy is obtained from a patient as described above, and afterwards viable tumors cells are then isolated and grown in culture as described in, e.g., Kern *et al.* (1990), *J. Natl. Cancer Inst.*, 82:582-588. In order to determine if cancer cells are sensitive to an HDACi compound, they are then grown in the presence of the HDACi compound at a range of concentrations, e.g., (0-10  $\mu$ M), and cell proliferation is measured by any number of methods, e.g., tritiated thymidine incorporation. Inhibition of tumor cell proliferation by the HDACi compound is measured relative to tumor cell proliferation in the absence of the compound (i.e., no inhibition). Assignment of the cancer as sensitive or resistant is optionally determined based on a number of cell proliferation criteria. For example, if the  $IC_{50}$  of the HDACi compound in the tested cancer cells is significantly lower (e.g., by 2 SDs) than that observed for cells known to be sensitive to the compound, the cancer is characterized as resistant. Thus, cells derived from the resistant cancer (e.g., directly or after passage in culture) are optionally used to generate a population of nucleic acids serving as an expression level (positive) reference sample used for setting HDACiR-BG expression level threshold values as described above. Conversely, tumor cells found to be sensitive to an HDACi compound are used generate a population of nucleic acids serving as an expression level (negative) reference sample.

[00102] Methods for obtaining RNA from biological samples (e.g., tissues or cells) including linear aRNA amplification from single cells include, e.g., Luzzi *et al.* (2005), *Methods Mol Biol.*, 293:187-207. Further, diverse kits for high quality RNA purification are available commercially, e.g., from Qiagen (Valencia, CA), Invitrogen (Carlsbad, CA), Clontech (Palo Alto, CA), and Stratagene (La Jolla, CA).

[00103] In some embodiments, the expression level reference sample is an RNA sample isolated from one or more HDACi compound-resistant colon cancer cells. In one embodiment, the cells were derived from colon carcinoma biopsy R5247682266, R9866135153, R1078103114, or R4712781606 described herein.

HDACi inhibitor Compounds

[00104] In another embodiment, HDACi inhibitor tumor compounds for which cancer resistance or sensitivity include, but are not limited to carboxylates, short-chain fatty acids, hydroxamic acids, electrophilic ketones, epoxides, cyclic peptides, and benzamides. In a further embodiment, HDACi inhibitor tumor compounds for which cancer resistance or sensitivity include, but are not limited to hydroxamic acids having the structure of Formula (A):



wherein

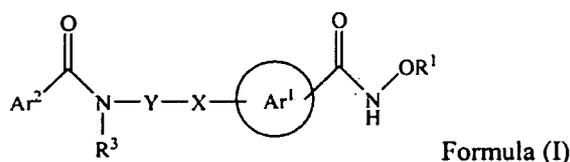
Q is an optionally substituted C<sub>5-12</sub> aryl or an optionally substituted C<sub>5-12</sub> heteroaryl;

L is a linker having at least 4 atoms;

R<sup>1</sup> is H or alkyl;

and a pharmaceutically acceptable salt, pharmaceutically acceptable N-oxide, pharmaceutically active metabolite, pharmaceutically acceptable prodrug, pharmaceutically acceptable solvate thereof.

[00105] HDACi inhibitor tumor compounds for which cancer resistance or sensitivity include, but are not limited to compounds having the structure of Formula (I):



wherein:

R<sup>1</sup> is hydrogen or alkyl;

X is -O-, -NR<sup>2</sup>-, or -S(O)<sub>n</sub> where n is 0-2 and R<sup>2</sup> is hydrogen or alkyl;

Y is alkylene optionally substituted with cycloalkyl, optionally substituted phenyl, alkylthio, alkylsulfinyl, alkylsulfonyl, optionally substituted phenylalkylthio, optionally substituted phenylalkylsulfonyl, hydroxy, or optionally substituted phenoxy;

Ar<sup>1</sup> is phenylene or heteroarylene wherein said Ar<sup>1</sup> is optionally substituted with one or two groups independently selected from alkyl, halo, hydroxy, alkoxy, haloalkoxy, or haloalkyl;

R<sup>3</sup> is hydrogen, alkyl, hydroxyalkyl, or optionally substituted phenyl; and

Ar<sup>2</sup> is aryl, aralkyl, aralkenyl, heteroaryl, heteroaralkyl, heteroaralkenyl, cycloalkyl, cycloalkylalkyl, heterocycloalkyl, or heterocycloalkylalkyl;

and individual stereoisomers, individual geometric isomers, or mixtures thereof; or a pharmaceutically acceptable salt thereof.

[00106] In another embodiment, HDACi inhibitor tumor compounds for which cancer resistance or sensitivity include, but are not limited to, PCI-24781.

[00107] In some embodiments, a patient is prescribed or administered an HDAC inhibitor to the patient based on a classification of the patient's cancer as being sensitive or resistant to an HDAC inhibitor according to the methods described herein.

[00108] In some embodiments, the methods described herein are used to optimize the selection of an anti-cancer agent for use in combination with an HDACi compound. In some embodiments, optimized selection of the second anti-cancer agent is performed by first comparing the set of known biomarker genes for resistance to the HDACi compound to sets of biomarker genes identified for other anti-cancer agents. The second anti-cancer agent is then selected for use in combination with the HDACi compound based on minimal overlap of the respective sets of resistance biomarker genes.

[00109] Examples of anti-cancer agents that are optionally used in combination with an HDACi compound include, but are not limited to, any of the following: gossypol, genasense, polyphenol E, Chlorofusin, all trans-retinoic acid (ATRA), bryostatins, tumor necrosis factor-related apoptosis-inducing ligand (TRAIL), 5-aza-2'-deoxycytidine, all trans retinoic acid, doxorubicin, vincristine, etoposide, gemcitabine, imatinib (Gleevec®), geldanamycin, 17-N-Allylamino-17-Demethoxygeldanamycin (17-AAG), flavopiridol, LY294002, bortezomib, trastuzumab, BAY 11-7082, PKC412, or PD184352, Taxol™, also referred to as "paclitaxel", is an anti-cancer drug which acts by enhancing and stabilizing microtubule formation, and analogs of Taxol™, such as Taxotere™. Compounds that have the basic taxane skeleton as a common structure feature, have also been shown to have the ability to arrest cells in the G2-M phases due to stabilized microtubules and are optionally useful for treating cancer in combination with the compounds described herein.

[00110] Further examples of anti-cancer agents for use in combination with an HDACi compound include inhibitors of mitogen-activated protein kinase signaling, e.g., U0126, PD98059, PD184352, PD0325901, ARRY-142886, SB239063, SP600125, BAY 43-9006, wortmannin, or LY294002.

[00111] Other anti-cancer agents that are optionally employed in combination with an HDACi compound include Adriamycin, Dactinomycin, Bleomycin, Vinblastine, Cisplatin, acivicin; aclarubicin; acodazole hydrochloride; acronine; adozelesin; aldesleukin; altretamine; ambomycin; ametantrone acetate; aminoglutethimide; amsacrine; anastrozole; anthramycin; asparaginase; asperlin; azacitidine; azetepa; azotomycin; batimastat; benzodepa; bicalutamide; bisantrene hydrochloride; bisnafide dimesylate; bizelesin; bleomycin sulfate; brequinar sodium; bropirimine; busulfan; cactinomycin; calusterone; caracemide; carbetimer; carboplatin; carmustine; carubicin hydrochloride; carzelesin; cedefingol; chlorambucil; cirolemycin; cladribine; crisnatol mesylate; cyclophosphamide; cytarabine; dacarbazine; daunorubicin hydrochloride; decitabine; dexormaplatin; dezaguanine; dezaguanine mesylate; diaziquone; doxorubicin; doxorubicin hydrochloride; droloxifene; droloxifene citrate; dromostanolone propionate; duazomycin; edatrexate; eflornithine hydrochloride; elsamitrucin; enloplatin; enpromate; epipropidine; epirubicin hydrochloride; erbulozole; esorubicin hydrochloride;

estramustine; estramustine phosphate sodium; etanidazole; etoposide; etoposide phosphate; etoprine;  
 fadrozole hydrochloride; fazarabine; fenretinide; floxuridine; fludarabine phosphate; fluorouracil;  
 flurocitabine; fosquidone; fostriecin sodium; gemcitabine; gemcitabine hydrochloride; hydroxyurea;  
 idarubicin hydrochloride; ifosfamide; iimofosine; interleukin II (including recombinant interleukin II,  
 5 or rIL2), interferon alfa-2a; interferon alfa-2b; interferon alfa-n1 ; interferon alfa-n3; interferon beta-1  
 a; interferon gamma-1 b; iproplatin; irinotecan hydrochloride; lanreotide acetate; letrozole; leuprolide  
 acetate; liarozole hydrochloride; lometrexol sodium; lomustine; losoxantrone hydrochloride;  
 masoprocol; maytansine; mechlorethamine hydrochloride; megestrol acetate; melengestrol acetate;  
 melphalan; menogaril; mercaptopurine; methotrexate; methotrexate sodium; metoprine; meturedepa;  
 10 mitindomide; mitocarcin; mitocromin; mitogillin; mitomalcin; mitomycin; mitosper; mitotane;  
 mitoxantrone hydrochloride; mycophenolic acid; nocodazole; nogalamycin; ormaplatin; oxisuran;  
 pegaspargase; peliomycin; pentamustine; peplomycin sulfate; perfosfamide; pipobroman; pipsulfan;  
 piroxantrone hydrochloride; plicamycin; plomestane; porfimer sodium; porfiromycin; prednimustine;  
 procarbazine hydrochloride; puromycin; puromycin hydrochloride; pyrazofurin; riboprine;  
 15 rogletimide; safingol; safingol hydrochloride; semustine; simtrazene; sparfosate sodium;  
 sparsomycin; spirogermanium hydrochloride; spiromustine; spiroplatin; streptonigrin; streptozocin;  
 sulofenur; talisomycin; tecogalan sodium; tegafur; teloxantrone hydrochloride; temoporfin;  
 teniposide; teroxirone; testolactone; thiamiprine; thioguanine; thiotepa; tiazofurin; tirapazamine;  
 toremifene citrate; trestolone acetate; tricirbine phosphate; trimetrexate; trimetrexate glucuronate;  
 20 triptorelin; tubulozole hydrochloride; uracil mustard; uredepa; vapreotide; verteporfin; vinblastine  
 sulfate; vincristine sulfate; vindesine; vindesine sulfate; vinepidine sulfate; vinglycinate sulfate;  
 vinleurosine sulfate; vinorelbine tartrate; vinrosidine sulfate; vinzolidine sulfate; vorozole; zeniplatin;  
 zinostatin; zorubicin hydrochloride.

**[00112]** Other anti-cancer agents that are optionally employed in combination with an HDACi  
 25 compound include: 20-epi-1 ,25 dihydroxyvitamin D3; 5-ethynyluracil; abiraterone; aclarubicin;  
 acylfulvene; adecyphenol; adozelesin; aldesleukin; ALL-TK antagonists; altretamine; ambamustine;  
 amidox; amifostine; aminolevulinic acid; amrubicin; amsacrine; anagrelide; anastrozole;  
 andrographolide; angiogenesis inhibitors; antagonist D; antagonist G; antarelix; anti-dorsalizing  
 morphogenetic protein-1 ; antiandrogen, prostatic carcinoma; antiestrogen; antineoplaston; antisense  
 30 oligonucleotides; aphidicolin glycinate; apoptosis gene modulators; apoptosis regulators; apurinic  
 acid; ara-CDP-DL-PTBA; arginine deaminase; asulacrine; atamestane; atrimustine; axinastatin 1 ;  
 axinastatin 2; axinastatin 3; azasetron; azatoxin; azatyrosine; baccatin III derivatives; balanol;  
 batimastat; BCR/ABL antagonists; benzochlorins; benzoylstauroporine; beta lactam derivatives;  
 beta-alethine; betaclamycin B; betulinic acid; bFGF inhibitor; bicalutamide; bisantrene;  
 35 bisaziridinylspermine; bisnafide; bistratene A; bizelesin; breflate; bropirimine; budotitane; buthionine  
 sulfoximine; calcipotriol; calphostin C; camptothecin derivatives; canarypox IL-2; capecitabine;  
 carboxamide-amino-triazole; carboxyamidotriazole; CaRest M3; CARN 700; cartilage derived

inhibitor; carzelesin; casein kinase inhibitors (ICOS); castanospermine; cecropin B; cetorelix;  
 chlorlins; chloroquinoxaline sulfonamide; cicaprost; cis-porphyrin; cladribine; clomifene analogues;  
 clotrimazole; collismycin A; collismycin B; combretastatin A4; combretastatin analogue; conagenin;  
 crambescidin 816; crisnatol; cryptophycin 8; cryptophycin A derivatives; curacin A;  
 5 cyclopentantraquinones; cycloplam; cypemycin; cytarabine ocfosfate; cytolytic factor; cytostatin;  
 dacliximab; decitabine; dehydrodidemnin B; deslorelin; dexamethasone; dexifosfamide; dexrazoxane;  
 dexverapamil; diaziquone; didemnin B; didox; diethylnorspermine; dihydro-5-azacytidine; 9-  
 dioxamycin; diphenyl spiromustine; docosanol; dolasetron; doxifluridine; droloxifene; dronabinol;  
 duocarmycin SA; ebselen; ecomustine; edelfosine; edrecolomab; eflornithine; elemene; emitefur;  
 10 epirubicin; epristeride; estramustine analogue; estrogen agonists; estrogen antagonists; etanidazole;  
 etoposide phosphate; exemestane; fadrozole; fazarabine; fenretinide; filgrastim; finasteride;  
 flavopiridol; flezelastine; fluasterone; fludarabine; fluorodaunorubicin hydrochloride; forfenimex;  
 formestane; fostriecin; fotemustine; gadolinium texaphyrin; gallium nitrate; galocitabine; ganirelix;  
 gelatinase inhibitors; gemcitabine; glutathione inhibitors; hepsulfam; heregulin; hexamethylene  
 15 bisacetamide; hypericin; ibandronic acid; idarubicin; idoxifene; idramantone; ilmofofosine; ilomastat;  
 imidazoacridones; imiquimod; immunostimulant peptides; insulin-like growth factor-1 receptor  
 inhibitor; interferon agonists; interferons; interleukins; iobenguane; iododoxorubicin; ipomeanol, 4-;  
 iroplact; irsogladine; isobengazole; isohomohalicondrin B; itasetron; jasplakinolide; kahalalide F;  
 lamellarin-N triacetate; lanreotide; leinamycin; lenograstim; lentinan sulfate; leptolstatin; letrozole;  
 20 leukemia inhibiting factor; leukocyte alpha interferon; leuprolide+estrogen+progesterone; leuprorelin;  
 levamisole; liarozole; linear polyamine analogue; lipophilic disaccharide peptide; lipophilic platinum  
 compounds; lissoclinamide 7; lobaplatin; lombricine; lometrexol; lonidamine; losoxantrone;  
 lovastatin; loxoribine; lurtotecan; lutetium texaphyrin; lysofylline; lytic peptides; maitansine;  
 mannostatin A; marimastat; masoprocòl; maspin; matrilysin inhibitors; matrix metalloproteinase  
 25 inhibitors; menogaril; merbarone; meterelin; methioninase; metoclopramide; MIF inhibitor;  
 mifepristone; miltefosine; mirimostim; mismatched double stranded RNA; mitoguazone; mitolactol;  
 mitomycin analogues; mitonafide; mitotoxin fibroblast growth factor-saporin; mitoxantrone;  
 mofarotene; molgramostim; monoclonal antibody, human chorionic gonadotrophin; monophosphoryl  
 lipid A+myobacterium cell wall sk; mopidamol; multiple drug resistance gene inhibitor; multiple  
 30 tumor suppressor 1 -based therapy; mustard anticancer agent; mycaperoxide B; mycobacterial cell  
 wall extract; myriaporone; N-acetyldinaline; N-substituted benzamides; nafarelin; nagrestip;  
 naloxone+pentazocine; napavin; naphterpin; nartograstim; nedaplatin; nemorubicin; neridronic acid;  
 neutral endopeptidase; nilutamide; nisamycin; nitric oxide modulators; nitroxide antioxidant;  
 nitrullyn; O6-benzylguanine; octreotide; okicenone; oligonucleotides; onapristone; ondansetron;  
 35 ondansetron; oracin; oral cytokine inducer; ormaplatin; osaterone; oxaliplatin; oxaunomycin;  
 palauamine; palmitoylrhizoxin; pamidronic acid; panaxytriol; panomifene; parabactin; pazelliptine;  
 pegaspargase; peldesine; pentosan polysulfate sodium; pentostatin; pentozole; perflubron;

perfosfamide; perillyl alcohol; phenazinomycin; phenylacetate; phosphatase inhibitors; picibanil; pilocarpine hydrochloride; pirarubicin; piritrexim; placetin A; placetin B; plasminogen activator inhibitor; platinum complex; platinum compounds; platinum-triamine complex; porfimer sodium; porfiromycin; prednisone; propyl bis-acridone; prostaglandin J2; proteasome inhibitors; protein A-based immune modulator; protein kinase C inhibitor; protein kinase C inhibitors, microalgal; protein tyrosine phosphatase inhibitors; purine nucleoside phosphorylase inhibitors; purpurins; pyrazoloacridine; pyridoxylated hemoglobin polyoxyethylene conjugate; raf antagonists; raltitrexed; ramosetron; ras farnesyl protein transferase inhibitors; ras inhibitors; ras-GAP inhibitor; retelliptine demethylated; rhenium Re 186 etidronate; rhizoxin; ribozymes; RII retinamide; rogletimide; rohitukine; romurtide; roquinimex; rubiginone B1 ; ruboxyl; safingol; saintopin; SarCNU; sarcophytol A; sargramostim; Sdi 1 mimetics; semustine; senescence derived inhibitor 1 ; sense oligonucleotides; signal transduction inhibitors; signal transduction modulators; single chain antigen-binding protein; sizofiran; sobuzoxane; sodium borocaptate; sodium phenylacetate; solverol; somatomedin binding protein; sonermin; sparfosic acid; spicamycin D; spiromustine; splenopentin; spongistatin 1 ; squalamine; stem cell inhibitor; stem-cell division inhibitors; stipiamide; stromelysin inhibitors; sulfinosine; superactive vasoactive intestinal peptide antagonist; suradista; suramin; swainsonine; synthetic glycosaminoglycans; tallimustine; tamoxifen methiodide; tauromustine; tazarotene; tecogalan sodium; tegafur; tellurapyrylium; telomerase inhibitors; temoporfin; temozolomide; teniposide; tetrachlorodecaoxide; tetrazomine; thaliblastine; thiocoraline; thrombopoietin; thrombopoietin mimetic; thymalfasin; thymopoietin receptor agonist; thymotrinan; thyroid stimulating hormone; tin ethyl etiopurpurin; tirapazamine; titanocene bichloride; topsentin; toremifene; totipotent stem cell factor; translation inhibitors; tretinoin; triacetyluridine; triciribine; trimetrexate; triptorelin; tropisetron; turosteride; tyrosine kinase inhibitors; tyrphostins; UBC inhibitors; ubenimex; urogenital sinus-derived growth inhibitory factor; urokinase receptor antagonists; vapreotide; variolin B; vector system, erythrocyte gene therapy; velaresol; veramine; verdins; verteporfin; vinorelbine; vinxaltine; vitaxin; vorozole; zanoterone; zeniplatin; zilascorb; and zinostatin stimalamer.

**[00113]** Yet other anticancer agents that are optionally employed in combination with an HDACi compound include alkylating agents, antimetabolites, natural products, or hormones, e.g., nitrogen mustards (e.g., mechloroethamine, cyclophosphamide, chlorambucil, etc.), alkyl sulfonates (e.g., busulfan), nitrosoureas (e.g., carmustine, lomustine, etc.), or triazines (decarbazine, etc.). Examples of antimetabolites include but are not limited to folic acid analog (e.g., methotrexate), or pyrimidine analogs (e.g., Cytarabine), purine analogs (e.g., mercaptopurine, thioguanine, pentostatin).

**[00114]** Examples of natural products useful in combination with an HDACi compound include but are not limited to vinca alkaloids (e.g., vinblastin, vincristine), epipodophyllotoxins (e.g., etoposide), antibiotics (e.g., daunorubicin, doxorubicin, bleomycin), enzymes (e.g., L-asparaginase), or biological response modifiers (e.g., interferon alpha).

[00115] Examples of alkylating agents that are optionally employed in combination an HDACi compound include, but are not limited to, nitrogen mustards (e.g., mechloroethamine, cyclophosphamide, chlorambucil, meiphalan, etc.), ethylenimine and methylmelamines (e.g., hexamethylmelamine, thiotepa), alkyl sulfonates (e.g., busulfan), nitrosoureas (e.g., carmustine, lomusitne, semustine, streptozocin, etc.), or triazenes (decarbazine, etc.). Examples of antimetabolites include, but are not limited to folic acid analog (e.g., methotrexate), or pyrimidine analogs (e.g., fluorouracil, floxouridine, Cytarabine), purine analogs (e.g., mercaptopurine, thioguanine, pentostatin.

[00116] Examples of hormones and antagonists useful in combination with an HDACi compound include, but are not limited to, adrenocorticosteroids (e.g., prednisone), progestins (e.g., hydroxyprogesterone caproate, megestrol acetate, medroxyprogesterone acetate), estrogens (e.g., diethylstilbestrol, ethinyl estradiol), antiestrogen (e.g., tamoxifen), androgens (e.g., testosterone propionate, fluoxymesterone), antiandrogen (e.g., flutamide), gonadotropin releasing hormone analog (e.g., leuprolide). Other agents that are optionally used in the methods and compositions described herein for the treatment or prevention of cancer include platinum coordination complexes (e.g., cisplatin, carboplatin), anthracenedione (e.g., mitoxantrone), substituted urea (e.g., hydroxyurea), methyl hydrazine derivative (e.g., procarbazine), adrenocortical suppressant (e.g., mitotane, aminoglutethimide).

[00117] Examples of anti-cancer agents which act by arresting cells in the G2-M phases due to stabilized microtubules and which are optionally used in combination with an HDACi compound include without limitation the following marketed drugs and drugs in development: Erbulozole (also known as R-55104), Dolastatin 10 (also known as DLS-10 and NSC-376128), Mivobulin isethionate (also known as CI-980), Vincristine, NSC-639829, Discodermolide (also known as NVP-XX-A-296), ABT-751 (Abbott, also known as E-7010), Altorhyrtins (such as Altorhyrtin A and Altorhyrtin C), Spongistatins (such as Spongistatin 1, Spongistatin 2, Spongistatin 3, Spongistatin 4, Spongistatin 5, Spongistatin 6, Spongistatin 7, Spongistatin 8, and Spongistatin 9), Cemadotin hydrochloride (also known as LU-103793 and NSC-D-669356), Epothilones (such as Epothilone A, Epothilone B, Epothilone C (also known as desoxyepothilone A or dEpoA), Epothilone D (also referred to as KOS-862, dEpoB, and desoxyepothilone B), Epothilone E, Epothilone F, Epothilone B N-oxide, Epothilone A N-oxide, 16-aza-epothilone B, 21-aminoepothilone B (also known as BMS-310705), 21-hydroxyepothilone D (also known as Desoxyepothilone F and dEpoF), 26-fluoroepothilone), Auristatin PE (also known as NSC-654663), Soblidotin (also known as TZT-1027), LS-4559-P (Pharmacia, also known as LS-4577), LS-4578 (Pharmacia, also known as LS-477-P), LS-4477 (Pharmacia), LS-4559 (Pharmacia), RPR-112378 (Aventis), Vincristine sulfate, DZ-3358 (Daiichi), FR-182877 (Fujisawa, also known as WS-9885B), GS-164 (Takeda), GS-198 (Takeda), KAR-2 (Hungarian Academy of Sciences), BSF-223651 (BASF, also known as ILX-651 and LU-223651), SAH-49960 (Lilly/Novartis), SDZ-268970 (Lilly/Novartis), AM-97 (Armad/Kyowa Hakko), AM-132 (Armad), AM-138 (Armad/Kyowa Hakko), IDN-5005 (Indena), Cryptophycin 52 (also known as LY-

355703), AC-7739 (Ajinomoto, also known as AVE-8063A and CS-39.HCl), AC-7700 (Ajinomoto, also known as AVE-8062, AVE-8062A, CS-39-L-Ser.HCl, and RPR-258062A), Vitilevuamide, Tubulysin A, Canadensol, Centaureidin (also known as NSC-106969), T-138067 (Tularik, also known as T-67, TL-138067 and TI-138067), COBRA-1 (Parker Hughes Institute, also known as DDE-261 and WHI-261), H10 (Kansas State University), H16 (Kansas State University), Oncocidin A1 (also known as BTO-956 and DIME), DDE-313 (Parker Hughes Institute), Fijianolide B, Laulimalide, SPA-2 (Parker Hughes Institute), SPA-1 (Parker Hughes Institute, also known as SPIKET-P), 3-IAABU (Cytoskeleton/Mt. Sinai School of Medicine, also known as MF-569), Narcosine (also known as NSC-5366), Nascapine, D-24851 (Asta Medica), A-105972 (Abbott), Hemiasterlin, 3-BAABU (Cytoskeleton/Mt. Sinai School of Medicine, also known as MF-191), TMPN (Arizona State University), Vanadocene acetylacetonate, T-138026 (Tularik), Monsatrol, Inanocine (also known as NSC-698666), 3-IAABE (Cytoskeleton/Mt. Sinai School of Medicine), A-204197 (Abbott), T-607 (Tularik, also known as T-900607), RPR- 115781 (Aventis), Eleutherobins (such as Desmethyleleutherobin, Desaeyleleutherobin, Isoeleutherobin A, and Z-Eleutherobin), Caribaeside, Caribaecin, Halichondrin B, D-64131 (Asta Medica), D-68144 (Asta Medica), Diazonamide A, A-293620 (Abbott), NPI-2350 (Nereus), Taccalonolide A, TUB-245 (Aventis), A-259754 (Abbott), Diozostatin, (-)-Phenylahistin (also known as NSCL-96F037), D-68838 (Asta Medica), D-68836 (Asta Medica), Myoseverin B, D-43411 (Zentaris, also known as D-81862), A-289099 (Abbott), A-318315 (Abbott), HTI-286 (also known as SPA-110, trifluoroacetate salt) (Wyeth), D-82317 (Zentaris), D-82318 (Zentaris), SC-12983 (NCI), Resverastatin phosphate sodium, BPR-OY-007 (National Health Research Institutes), and SSR-250411 (Sanofi).

#### Applications of HDACiR-BGs

[00118] The methods and compositions described herein are optionally used to increase the likelihood of a therapeutically effective treatment of a patient's cancer with an HDACi compound by providing an indication (e.g. by oral or written communication in any analog or digital medium) of which genes are HDACiR-BGs, as well as HDACiR-BG expression level reference values (e.g., expression level threshold values) above which HDACi compound resistance is likely (i.e., greater than the probability by chance) or below which HDACi compound sensitivity is likely.

[00119] In some embodiments, the indication includes a document with an interpretation of expression levels of at least four biomarker genes selected from Table 1 as to the likelihood that a patient's cancer is resistant or sensitive to treatment with an HDACi compound.

[00120] In some embodiments, the document includes an interpretation of the expression levels of at least one HDACiR-BG selected from DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP1.

[00121] In some embodiments an indication is provided in one or more databases containing information concerning one or more HDACiR-BGs, including one or more expression level threshold values that permit the interpretation of the effect of HDACiR-BG expression levels on the resistance or sensitivity of a cancer to an HDACi compound according to any of the methods described herein.

5 Such expression level threshold values include those set based on, e.g., deviation of HDACiR-BG expression levels in a test sample from the corresponding HDACiR-BG expression levels in an expression level (positive or negative) reference sample as described herein. Alternatively, or in addition, expression level threshold values are optionally set based on deviation of the expression ratios of HDACiR-BGs to one or more internal expression control genes (e.g., RNA polymerase II, HDAC3, or ZNF217). For example, as described herein, the mean expression ratio (based on TaqMan fluorescence intensity) of the HDACiR-BG DEFA6 to the internal expression control gene ZNF217 is 10 5.83 in HDACi-resistant colon cancer cells and 0.24 in HDACi-sensitive colon cancer cells.

[00122] In some embodiments, the databases include HDACiR-BG expression level profiles or thresholds associated with resistance to one or more HDACi compounds for one or more types of 15 cancer.

[00123] Other information that is optionally included in the databases or in other types of indication include, but are not limited to, HDACiR-BG sequence information, frequency distributions of HDACiR-BG expression levels in a particular cancer population, descriptive information concerning the clinical status of a biological sample analyzed for HDACiR-BG expression profiles, or 20 the clinical status of the patient from which the sample was derived. The database is optionally be designed to include different parts, for instance an HDACiR-BG list database, and an informative HDACiR-BG expression profile database, e.g., a database associating with each HDACiR-BG expression profile record the probability that the expressin profile is associated with resistance to an HDACi compound. Methods for the configuration and construction of databases are widely available, 25 for instance, see U.S. Pat. No. 5,953,727.

[00124] The databases described herein are optionally linked to an outside or external database. In some embodiments, the database optionally communicates with outside data sources, such as database of the developmental therapeutics program of the national cancer institute or the National Center for Biotechnology Information through the internet.

30 [00125] Any appropriate computer platform is used to perform the methods for interpreting one or more HDACiR-BG expression profiles by the methods described herein. In some embodiments, the computer platform receive direct input from a database, e.g., one of the databases described herein. For example, a large number of computer workstations are available from a variety of manufacturers, such has those available from Silicon Graphics. Client-server environments, 35 database servers and networks are also widely available and are appropriate platforms for the databases described herein.

[00126] The databases described herein are optionally used to present information identifying a set of HDACiR-BG expression profiles in an individual and such a presentation is optionally used to predict or diagnose the likelihood of a effective therapeutic treatment of the individual's cancer with a particular HDACi compound based on a statistical comparison of the individual's expression profile to HDACiR-BG expression level thresholds as described herein. Accordingly, one chooses to partition cancer patients into subgroups at any threshold value of the measured HDACiR-BG expression, where all patients with expression values above the threshold have higher risk, and all patients with expression values below the threshold have lower risk, of and HDACi compound-resistant cancer resistance or vice versa, depending on whether the expression level threshold is based on an expression level in a cancer determined to be resistant to an HDACi compound treatment (i.e., a positive reference sample) or sensitive to the HDACi compound treatment (i.e., a negative reference sample). Alternatively, HDACiR-BG expression profiles ranked on a probability continuum, where the more an HDACiR-BG expression level deviates negatively from (i.e., is less than) an expression level positive reference value, the higher the probability that the cancer is sensitive to treatment with an HDACi compound. Conversely, the more an HDACiR-BG expression level deviates positively from (i.e., is greater than) an expression level negative reference value, the higher the probability that the cancer is resistant to treatment with an HDACi compound.

#### EXAMPLES

[00127] The following specific examples are to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever. Without further elaboration, it is believed that one skilled in the art can, based on the description herein, utilize the present invention to its fullest extent.

#### Example 1: mRNA expression profiling of HDACi sensitive versus resistant colorectal tumor cells *ex vivo*

[00128] We and others previously developed several pharmacodynamic markers for HDACi compounds (such as tubulin or histone acetylation, p21 expression etc). However, there is currently no clinically predictive biomarker for response to these agents available. In this work, we developed a strategy to identify such biomarkers for the HDACi compound PCI-24781 in primary human colorectal tumors.

[00129] The method used soft agar chemosensitivity assays in which primary human tumors were exposed in culture to PCI-24781. Either a tritiated thymidine or alamar blue assay was then used to estimate the percentage of resistance to PCI-24781. For example in the tritiated thymidine assay, sensitive tumor cells affected by the drug divided less and therefore incorporated less thymidine, whereas resistant tumor cells continued to grow and divide and therefore incorporated more thymidine into their DNA. It has been shown historically that under the optimized conditions of

this assay, a patient whose tumor is classified as resistant to a given drug has < 1% probability of response to that drug in the clinic (in published correlations to clinical outcome, these assays predicted resistance with an accuracy of 99% in solid cancers and 92% in blood cancers). For example, a recent paper correlated *in vitro* sensitivity or resistance to fludarabine in the DiSC assay in B-cell CLL patients with clinical outcome (median survival 7.9 months in resistant vs 41.7 months in sensitive patients). Similar data has also been published for solid tumors : e.g., for sensitivity or resistance to Pt in ovarian tumors, and to CPX and DOX in breast tumors.

5  
10  
[00130] After determining *ex vivo* sensitivity or resistance to PCI-24781 for each tumor, RNA isolated from tumor cells was then profiled on microarrays and a marker set was identified by statistical analysis of the data. This marker set was validated by RT-PCR (TaqMan™) analysis. Such pharmacogenomic biomarkers that are used for patient stratification in the clinic provide a competitive advantage in the development of PCI-24781. A graphic summary of the method and its clinical applications is illustrated in Fig. 1.

15  
[00131] We examined the *ex-vivo* response of primary colorectal tumors from various patients to an HDAC inhibitor, PCI-24781, and subsequently determined whether there were robust differences in the mRNA expression profiles of sensitive versus resistant tumor cells prior to HDACi treatment.

20  
25  
[00132] Primary colorectal cancer (CRC) samples were obtained from patient biopsies (Table 2). Viable tumor cells were plated and cultured in soft agar as described in Kern *et al.* (1990), *J. Natl. Cancer Inst.*, 82:582-588, and were treated with a range of PCI-24781 concentrations (0.01-2μM). Tritiated thymidine was added to the culture after 3 days of exposure to the drug, and the amount of radioactivity incorporated into the cells after a further 2 days was quantified. The percentage of cell growth inhibition (%GI) was calculated by comparing the treated cells to the control cells, and from these growth profiles the tumors were classified as either sensitive or resistant based on deviation from the median profile. As shown in Fig. 2, primary tumors displayed a spectrum of growth inhibition phenotypes from 100% to 0% relative to control at the PCI-24781 concentrations tested (up to 2 μM).

Research ID	Cancer Name	Age	Sex	Site	Clinical Diagnosis	Histology	Specimen Type
R1078103114	Colon Carcinoma	54	F	R Ovary	Colon Carcinoma	ADENOCARCINOMA	Solid Tumor Biopsy
R1105698672	Colon Carcinoma	72	F	Portion of Terminal Ileum	Colon Carcinoma	NA	Solid Tumor Biopsy
R2163660366	Colon Carcinoma	58	F	Uterus	Rectal Cancer	NA	Solid Tumor Biopsy
R4712781806	Colon Carcinoma	59	M	Colon Resection	Colon Carcinoma	NA	Solid Tumor Biopsy
R5247682266	Colon Carcinoma	51	F	Upper Lobe Lung	Colon Carcinoma	ADENOCARCINOMA	Solid Tumor Biopsy
R5891015174	Colon Carcinoma	43	F	Colon	Cecal Carcinoma	NA	Solid Tumor Biopsy
R6173297194	Colon Carcinoma	65	M	Omentum	Colon Carcinoma	ADENOCARCINOMA	Solid Tumor Biopsy
R7103644976	Colon Carcinoma	52	F	R Tube & Ovary	Colon Cancer	NA	Solid Tumor Biopsy
R886135153	Colon Carcinoma	55	F	R Hepatic Lobe	Colon Carcinoma	ADENOCARCINOMA	Solid Tumor Biopsy
R2881036089	Colon Carcinoma	79	F	Colon	Colon Carcinoma	CARCINOMA, PD	Solid Tumor Biopsy
R5492724373	Colon Carcinoma	55	F	Cecum	Colon Carcinoma	COLON CARCINOMA	Solid Tumor Biopsy
R8624442989	Colon Carcinoma	47	F	Brain	Colon Carcinoma	NA	Solid Tumor Biopsy
R0948311023	Colon Carcinoma	33	F	L Lower Lung Lobe Nodule	Colon Carcinoma	ADENOCARCINOMA	Solid Tumor Biopsy
R1059261097	Colon Carcinoma	50	M	Liver	Colon Cancer	ADENOCARCINOMA	Solid Tumor Biopsy
R2191729233	Colon Carcinoma	62	F	Ovary	Colon Cancer	ADENOCARCINOMA	Solid Tumor Biopsy
R4498160514	Colon Carcinoma	40	F	L Ovary	Colon Carcinoma	ADENOCARCINOMA	Solid Tumor Biopsy
R4891777011	Colon Carcinoma	53	F	R Abdominal Sidewall	Colon Carcinoma	ADENOCARCINOMA	Solid Tumor Biopsy
R5456781761	Colon Carcinoma	65	F	Liver Lobes 5&6	Met. Colon CA to L	NA	Solid Tumor Biopsy
R5978110794	Colon Carcinoma	63	F	Sigmoid Rectum	Colon Carcinoma	NA	Solid Tumor Biopsy
R6289195776	Colon Carcinoma	56	M	Liver	Colon Carcinoma	ADENOCARCINOMA	Solid Tumor Biopsy
R6324805249	Colon Carcinoma	55	F	Ovary	Colon Carcinoma	ADENOCARCINOMA	Solid Tumor Biopsy
R7424107588	Colon Carcinoma	48	M	Lumbar/Spine Biopsy	Colon Carcinoma	NA	Solid Tumor Biopsy
R6701041232	Colon Carcinoma	65	M	Sigmoid Colon	Colon Carcinoma	ADENOCARCINOMA	Solid Tumor Biopsy
R9418488310	Colon Carcinoma	55	F	Cecum	Colon Carcinoma	ADENOCARCINOMA	Solid Tumor Biopsy

Table 2

[00133] After determining tumor sensitivity to PCI-24781, gene expression profiles were determined for resistant and sensitive tumors that were treated with PCI-24781 (2  $\mu$ M) or untreated. Total RNA was isolated using Qiagen procedures (Qiagen, Inc., Valencia, CA) and fluorescent probes were prepared and hybridized to Codelink Human Whole Genome oligonucleotide microarrays  
5 containing ~55,000 unique probes (GE Healthcare Bio-Sciences Corp., Piscataway, NJ) according to the manufacturer's instructions. The microarrays were scanned in a GenePix 4000B scanner (Molecular Devices Corporation, Sunnyvale CA). The images were processed with Codelink software and the exported data was analyzed as follows.

[00134] The median-normalized microarray data were imported into Genespring software  
10 (Agilent), and principal component analysis (PCA) and hierarchical clustering analysis were performed. We looked for consistent results from multiple analysis methods to provide additional confidence in our results. For multiple hypothesis correction, we used the q-values approach for false discovery rates (FDR) as described in Storey *et al.* (2003), *Proc. Nat. Acad. Sci. USA*, 100:9440-9445. As a second analytical approach we adopted the Bayesian ANOVA approach described in Ishwaran *et al.*,  
15 *(2003), J. Amer. Stat. Assoc.*, 98:438-455.

[00135] In the Bayesian ANOVA method, the contribution of irrelevant genes to the ANOVA model are selectively shrunk to balance total false detections against total false non-detections. The output is a Zcut score which identifies genes whose contribution to the ANOVA model is larger than the standard z-score. See Ishwaran *et al.*, *ibid.*, and the website at bamarray.com. For the  
20 identification of biomarkers predictive of PCI-24781 resistance, we used only the untreated control samples divided into pools based on the sensitivity or resistance classification in the assay described above. This analytical approach is summarized in Fig. 3.

[00136] As shown in Fig. 4, principal components analysis clearly distinguished untreated cell expression profiles from treated cell expression profiles. Controls (arrowhead) are more similar to  
25 each other and well separated from the treated samples. The major component PCA1 clearly resolves treated from control samples. Interestingly, the resistant cell expression profiles (circled in both the treated and untreated samples) clustered together before and after treatment, whereas the sensitive samples varied widely in their profiles after treatment with PCI-24781. This suggested that it is easier to identify patients with the most resistant tumors and exclude them from a clinical trial rather than to  
30 identifying patients with sensitive tumors.

[00137] Based on the microarray analysis, we identified a total of 44 genes (see table 3) whose level of expression was significantly higher (z-score greater than 3.5) in PCI-24781 resistant cells than in PCI-24781 sensitive cells (data not shown). Of note, the expression of the identified biomarker genes was not altered by treatment with PCI-24781.

<b>Gene Name</b>	<b>Gene Symbol</b>	<b>GenBank Accession #</b>	<b>z-score</b>	<b>Res./Sens. Fold Expression Difference</b>
PTPN3	PTPN3	AK096975	14.19	2.58
ATP-binding cassette, sub-family C (CFTR/MRP), member 3	ABCC3	NM_020037	13.24	2.37
specifically androgen-regulated protein	SARG	NM_023938	13.04	4.00
phosphatidic acid phosphatase type 2C	PPAP2C	NM_177526	12.95	4.75
neural proliferation, differentiation and control, 1	NPDC1	NM_015392	11.88	2.45
C-terminal tensin-like	CTEN	NM_032865	11.32	3.83
RAB25, member RAS oncogene family	RAB25	NM_020387	10.96	3.51
hephaestin	HEPH	NM_138737	10.49	3.38
thiopurine S-methyltransferase	TPMT	NM_000367	9.97	2.56
plakophilin 3	PKP3	NM_007183	9.31	3.13
UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 5 (GalNAc-T5)	GALNT5	NM_014568	9.31	2.54
calmodulin-like 4	CALML4	NM_033429	9.14	3.51
UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 12 (GalNAc-T12)	GALNT12	AK024865	8.86	2.51
thiamin pyrophosphokinase 1	TPK1	NM_022445	8.81	3.55
defensin, alpha 6, Paneth cell-specific	DEFA6	NM_001926	8.58	12.92
epithelial protein lost in neoplasm beta	EPLIN	NM_016357	8.49	2.33
chloride intracellular channel 5	CLIC5	NM_016929	7.20	3.60
PERP, TP53 apoptosis effector	PERP	NM_022121	6.94	2.60
spleen tyrosine kinase	SYK	NM_003177	6.90	3.59
solute carrier family 12 (sodium/potassium/chloride transporters), member 2	SLC12A2	NM_001046	6.75	4.85
guanylate cyclase 2C (heat stable enterotoxin receptor)	GUCY2C	NM_004963	6.72	3.53
transmembrane 4 superfamily member 4	TM4SF4	NM_004617	6.54	12.09
transforming growth factor, alpha	TGFA	NM_003236	6.44	3.11

fibroblast growth factor binding protein 1	FGFBP1	NM_005130	6.27	5.35
PTK6 protein tyrosine kinase 6	PTK6	NM_005975	6.24	3.10
epithelial V-like antigen 1	EVA1	NM_005797	5.96	4.55
EPH receptor A2	EPHA2	NM_004431	5.90	2.18
integrin, alpha 6	ITGA6	NM_000210	5.53	4.09
tumor necrosis factor receptor superfamily, member 21	TNFRSF21	NM_014452	5.47	2.16
transmembrane 4 superfamily member 3	TM4SF3	NM_004616	5.32	3.75
interleukin 18 (interferon-gamma-inducing factor)	IL18	NM_001562	5.24	5.22
bone morphogenetic protein 4	BMP4	NM_130850	4.82	3.91
sphingomyelin phosphodiesterase, acid-like 3B	SMPDL3B	NM_014474	4.62	5.49
transmembrane protease, serine 2	TMPRSS2	NM_005656	4.62	3.51
guanine deaminase	GDA	NM_004293	4.56	6.52
macrophage stimulating 1 receptor (c-met-related tyrosine kinase)	MST1R	NM_002447	4.49	4.52
integrin, beta 4	ITGB4	NM_000213	4.41	3.98
annexin A3	ANXA3	NM_005139	4.11	3.34
chemokine (C-C motif) ligand 15	CCL15	NM_032965	3.87	3.74
dipeptidase 1 (renal)	DPEP1	NM_004413	3.72	5.53
NADPH oxidase organizer 1	NOXO1	NM_172167	3.71	8.92
interferon, alpha-inducible protein 27	IFI27	NM_005532	3.69	3.65
cytochrome P450, family 3, subfamily A, polypeptide 43	CYP3A43	NM_057095	3.65	3.40
plakophilin 2	PKP2	NM_004572	3.54	3.45

**[00138]** Analysis of the biological pathways associated with these genes showed that homologous recombination, nucleotide excision repair, cell cycle, and apoptosis were among those that affect sensitivity to PCI-24781.

5

[00139] In order to validate the higher expression of each resistance biomarker gene identified by microarray analysis, we analyzed the expression of each biomarker gene by the TaqMan<sup>®</sup> quantitative RT-PCR method as described below.

[00140] TaqMan<sup>®</sup> Gene Expression Assays for selected genes were obtained from Applied Biosystems (Foster City, CA). One-step-RT-PCR was carried out in triplicate on 25 ng of total RNA from each sample on an ABI PRISM<sup>®</sup> 7900HT sequence detection system. The mRNA levels for each gene were normalized to the amount of RNA in the well as measured in parallel using Ribogreen (Invitrogen, Inc., Carlsbad, CA). We then calculated the ratios of expression levels of the biomarker genes in the resistant & sensitive samples (R/S) and compared them to the corresponding ratios obtained from the microarray analysis. The comparative analysis for 16 of the biomarker genes listed in Table 3 is shown in Table 4. As a further validation of our microarray analysis, we performed TaqMan assays for three genes whose expression, as measured by microarray hybridization, was not found to correlate with PCI-24781 resistance (see last three genes in Table 3).

**Table 4**  
**Microarray vs TaqMan Analysis of Genes Upregulated in**  
**PCI-24781-Resistant vs Sensitive Colorectal Tumor Cells**

GeneName	GeneCards	Microarrays				Taqman				
		Zcut	Resist. mean	Sens. mean	Ratio AR/S	Sens Ct	ResistAvg	SensAvg	Ratio TaqR/S Taq/Arr	
defensin, alpha 6, Paneth cell-specific	DEFA6	8.88	8.57	0.65	12.92	37.20	1.34	0.06	23.94	1.85
integrin, beta 4	ITGB4	4.41	0.67	0.17	3.98	28.99	86.18	16.59	5.20	1.31
transmembrane 4 superfamily member 3	TM4SF3	5.32	239.99	65.01	3.76	29.21	108.96	14.30	7.62	2.03
spleen tyrosine kinase	SYK	6.90	5.15	1.48	3.59	35.45	1.50	0.19	7.90	2.20
phosphatidic acid phosphatase type 2C	PPAP2C	12.95	5.35	1.14	4.76	36.45	1.26	0.09	13.31	2.80
RAB25, member RAS oncogene family	RAB25	10.96	55.31	15.92	3.51	32.56	16.97	1.40	12.10	3.46
hephaeslin	HEPH	10.49	8.11	2.46	3.38	32.90	4.34	1.11	3.93	1.16
NADPH oxidase organizer 1	NOXO1	3.71	0.98	0.11	8.92	35.41	4.60	0.19	23.76	2.66
transmembrane 4 superfamily member 4	TM4SF4	6.54	2.06	0.18	12.09	40.00	0.22	0.01	27.22	2.25
PTPN3	PTPN3	14.19	5.45	2.16	2.68	30.71	6.60	5.04	1.31	0.51
EPH receptor A2	EPHA2	5.90	29.27	13.49	2.18	31.91	25.80	2.20	11.73	5.37
fibroblast growth factor binding protein 1	FGFBP1	6.27	27.93	5.30	5.35	37.76	0.84	0.04	22.08	4.13
ATP-binding cassette, sub-family C, member 3	ABCC3	13.24	4.14	1.82	2.37	40.00	0.01	0.01	0.96	0.41
thiopurine S-methyltransferase	TPMT	9.97	26.21	10.11	2.56	40.00	0.01	0.01	0.96	0.38
interleukin 18 (interferon-gamma-inducing factor)	IL18	6.24	26.57	5.04	5.22	40.00	0.62	0.01	77.06	14.77
dipeptidase 1 (renal)	DPEP1	3.72	2.93	0.54	5.63	40.00	0.01	0.01	0.96	0.17
HDAC3	HDAC3		Not significant			25.66	141.70	167.11	0.85	
Zinc Finger Protein znf217	ZNF217		Not significant			35.07	0.23	0.25	0.93	
TSG101	TSG101		Not significant			40.00	0.01	0.01	0.96	

[00141] The comparison of microarray versus results is graphically summarized in Fig. 2. As shown in Table 4 and Fig. 2, genes found to be significantly upregulated by the microarray method were also found to be upregulated by the TaqMan method, though the latter generally yielded higher R/S ratios. Likewise, three genes whose expression did not differ significantly in the microarray analysis also showed no significant difference in the TaqMan assay.

[00142] Interestingly, several of the identified biomarker genes have previously been studied in relation to cancer, e.g., DEFA6, RAB25 small GTPase, MRP3 (ABCC3), and TM4SF4. Further, a number of the identified genes encode secreted proteins or transmembrane proteins that shed their extracellular domains. Genes encoding secretable proteins include, e.g., DEFA6 (NM\_001926), TM4SF4 (NM\_004617), TGFA (NM\_003236), FGFBP1 (NM\_005130), EPHA2 (NM\_004431), TNFRSF21 (NM\_014452), TMF4SF3 (NM\_004616), IL18 (NM\_001562), TMPRSS2 (NM\_005656), and CCL15 (NM\_032965).

[00143] Based on these data, we concluded that the expression pattern of subsets (e.g., four or more) of the identified biomarker genes provide "resistance signatures" that are optionally used to reliably identify colorectal tumors that are resistant or susceptible to the HDAC inhibitor PCI-24781.

[00144] In a validation experiment, we found that *ex vivo* cultured primary colon tumor cells from twelve newly diagnosed, naive patients were all sensitive to growth inhibition by the HDAC inhibitor PCI-24781 (Fig. 11A). In contrast, we found that in a number of cases, advanced metastatic colon tumor cells were resistant to growth inhibition by the HDAC inhibitor PCI-24781 (Fig. 11B), and the DEFA6 mRNA expression levels were higher in HDAC-resistant cells than in HDAC-sensitive cells (Fig. 11C).

#### Example 2: Identification and cross-validation of functional biomarkers for HDAC inhibitor compounds and selection of clinical indications

[00145] In order to determine relevant tumor types and to identify pharmacodynamic (PD) markers that are useful in the clinic, we first identified biomarkers of HDAC inhibition in mice and used these to identify HDACi- "sensitive" tissues. This was done by identifying, in HDACi-treated mice, genes in peripheral blood mononuclear cells (PBMC) whose mRNA levels showed the same timecourse as acetylated tubulin levels, an index of HDAC inhibition. These biomarker genes were then used to identify HDACi responsive mouse tissues. Primary human tumors corresponding to sensitive tissues were then tested *ex-vivo* with PCI-24781, and it was found that tumors from tissues that showed higher levels of activity were sensitive to inhibition by PCI-24781, thus validating that this technique does indeed predict sensitive tumor types.

[00146] In brief, female BALB/c mice were injected IV with 50 mg/kg PCI-24781 or vehicle. Blood and various tissues were collected at 0.25, 0.5, 1, 2, 3 & 8 hours after dosing. For acetylated histone and tubulin detection, organs/tissues were pooled for each vehicle and drug-treated organ group. RNA and protein were extracted from the samples with the PARIS Protein and RNA Isolation

System (Ambion). Levels of acetylated and total a-tubulin & histones were evaluated by immunoblotting.

[00147] RNA expression profiles were determined using a GE-Codelink Mouse Uniset1 10K oligonucleotide arrays in duplicate. Each treated sample was normalized to the corresponding vehicle control.

5

In order to validate the expression profile of HDADi-responsive genes identified by the gene expression array assays, Taqman gene expression assays were performed using Applied Biosystems Inc. assays. One-step RT-PCR was carried out in triplicate on 25 ng of total RNA from each sample on a ABI PRISM 7700 instrument. The mRNA levels for each gene were normalized to the amount of RNA in the well as measured in parallel using Ribogreen (Molecular Probes). The treated samples were then normalized to the vehicle control at that time point.

10

[00148] A set of 16 genes (Table 5) whose expression profile in PBMC (Fig. 7A) closely tracked increases in tubulin acetylation levels (Fig. 7B) following treatment with the HDAC inhibitor PCI-24781.

15

**Table 5: HDAC Inhibitor (HDACi)-Responsive Biomarker Genes**

Common	Description	Function
Sic9a3r1	solute carrier family 9 isoform 3 regulator 1	ION TRANSPORT
Ing1	inhibitor of growth family, member 1-like	CELL PROLIFERATION AND DIFFERENTIATION
Gadd45g	growth arrest and DNA-damage-inducible 45 gamma	CELL PROLIFERATION AND DIFFERENTIATION;APOPTOSIS
Plaur	urokinase plasminogen activator receptor	MULTIPLE
EST	RIKEN cDNA 2810405O22 gene	UNKNOWN
Ins16	insulin-like 6	BIOLOGICAL PROCESS UNKNOWN
Luc71	Luc7 homolog (S. cerevisiae)-like	RNA PROCESSING
Taf9	TAF9 RNA polymerase II	MRNA TRANSCRIPTION
Gadd45b	growth arrest and DNA-damage-inducible 45 beta	CELL PROLIFERATION AND DIFFERENTIATION
Syng12	synaptogyrin 2	UNKNOWN
Poir2e	polymerase (RNA) II (DNA directed) polypeptide E	MRNA TRANSCRIPTION
Kras2	Mouse c-Ki-ras oncogene	ONCOGENE
Hspa5	heat shock 70kD protein 5	STRESS RESPONSE
Fgf15	fibroblast growth factor 15	CELL PROLIFERATION AND DIFFERENTIATION
Tuba4	tubulin, alpha 4	CELL STRUCTURE
H2afz	H2A histone family, member Z	CHROMATIN PACKAGING

[00149] Subsequently, we validated the expression profile of two of HDACi-responsive genes, *Fgf15* and *Syng1*, by quantitative RT-PCR and immunoblotting. As shown in Fig. 8, the expression profiles obtained the three different methods closely matched one another, suggesting that the microarray analysis identified HDACi-responsive genes reliably.

5 [00150] We then determined the *in vivo* expression levels for five of the HDACi-responsive biomarker genes in various tissues following 3 hours or 8 hours following administration of PCI-24781 (50 mg/kg). A Taqman assay was performed to determine mRNA expression levels in brain, colon, kidney, liver, stomach, ovary, uterus, mammary, muscle, heart, lung, spleen, and pancreas. The mean and SD for mRNA expression levels of all 5 genes in each tissue at each time point are shown in Fig. 9. The tissue distribution pattern was very reproducible across the biomarker set. Ovary showed the highest level of induction, followed by uterus.

10 [00151] Subsequently, primary human tumor samples were obtained and viable tumor cells were plated in soft agar and treated with the HDAC inhibitor PCI-24781. Tritiated thymidine was added after 3 days, and 2 days later the radioactivity incorporated into the DNA was quantified. The tumors were then classified as either resistant (EDR: Extreme Drug Resistance), sensitive (LDR) or intermediate (IDR) based on deviation from the median profile (Oncotech, Inc. Tustin, CA). As predicted based on the HDACi responsive biomarker gene profiles hematopoietic tumors had the lowest proportion of resistant (EDR) tumors, and colon the most (38%). See Fig. 10 and Table 6. Among the solid tumors, ovarian had the lowest proportion of resistant tumors, consistent with the high HDACi- biomarker responsiveness of this tissue.

20

**Table 6: Tumor Resistance to HDAC Inhibitor PCI-24781**

Tumor Type	Resistant		Intermediate		Sensitive		%Resistance
	EDR	IDR	LDR	Total	LDR	Total	
AML	1	4	5	10	10	10	10
Multiple Myeloma	2	0	4	6	33	33	33
Ovarian	3	4	5	12	25	25	25
Glioblastoma	2	1	4	7	29	29	29
Colon	9	3	12	24	38	38	38

Note: EDR/LDR status as determined by Oncotech's algorithm from their assay data

**[00152]** Based on the above results, we concluded that expression profiles of the orthologous human biomarkers will reflect PCI-24781 activity in human blood, and serve as PD markers in the clinic. Further, the identified set of HDACi-responsiveness biomarker genes accurately predicts tumor sensitivity to treatment with HDAC inhibitors.

5

**APPENDIX**

**Nucleotide Sequences for HDACi Compound Resistance Biomarker Genes**

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
-----------	-------------	---------------------	-----------

5

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>PTPN3</b>	<b>PTPN3</b>	<b>AK096975</b>	<b>1</b>

10     1 tgaatagttt gctggtagca agacggatga agacctatat gggagattct ttatctctag  
 61 agctagcata ttacttgca tactttgttt cttttccaca tggatatttt actgctaaat  
 121 ggcagaggtg ggagggagat gtcacacagt accataacc catattgaaa acaagaaacc  
 181 accagaaagt ttgcagctaa ggggcagggg attcagttcc tacgcccact cagcactaac  
 241 tacttgcggg cctggttgct tagaagctct acctctctt cattatctgt aaaatagaaa  
 15     301 caatacttag gactttagtt ggaacatgag gattgaataa gatcacgcta ttcatgtgac  
 361 tttttatcgg ctagaacagc aacagacact gctgtgggtg agttacttag aaaagtttag  
 421 ttatcagtga ttagcccaaa aacacatcag tcaaaaatag aatccactgg atttttgtct  
 481 ctcttttttag agacagggtc tcaactgctc ccaggctgga gtacagtggc atgatcattg  
 541 tcaactgcag cctcaaatc ctgggctcaa gcaatcctcg cacctcagc tctgagtag  
 20     601 cgggactat aggcacatgc cacctcacct ggcttgtgtg tgtgtgtgtg tgtgtgtgtg  
 661 tgtgtgtgtg tgtgtgtgta gagacaggat cttgatgtgt cgctaggct ggtctcaaac  
 721 tcttgccctc aagtgatctt cccacctcag cctccaaaac tgttgggatt ataggcgtga  
 781 gccactgtgc ccagcctaac tgggttttta tgagaggaaa atagaaaatg ctcttctaga  
 841 agagagagaa caagagcaca aaataatctg gactcaciaa aattcagcaa gctccaagaa  
 25     901 agggggatgg agggaaacgct ggcaaaaatt taaatgccat taggatattt agcaagttat  
 961 tactgtttgg taaaaatgca tcatcaccct gtgtgcaaaa tgcttgcaaa gtagtctaaa  
 1021 tgtctttgga gatgggtgtt ttactgcttt tttccaaaa caaattgttt attatggttg  
 1081 cagaaatgca gccattacgg tcacataaat ttctaaaaag cctaccaaa gttgcaagca  
 1141 gtcttctgcc actgggcagg ccagcagttc agaccagcg aggttgccag gaacaaatcc  
 30     1201 aggaaatact ggggaagaaca agacaagaga attacctaaa agagcaaca attcaagtaa  
 1261 atcctgtage tattaccact taaaatcctg agctcaagat tctgtttca ccaccttata  
 1321 cacttaagca attatactta agcctttttt tagtcctaag tgaagaacta catcagaatc  
 1381 aggataagta ttttgccctg gaaatttggc tgcataatgaa tggagaagac atttacctcc  
 1441 tatgttctgg cactttctga aagatctaata taaacatgtt gatgtgcaa tttaatcaag  
 35     1501 atgagagatc cctgctggtg tcaccctcta gaacctgcac ttgggtttt gactttccag  
 1561 aagaaaaaaa tgcaacttg gttagggggc agtgggtgga tcacacagtt gtctttcgtt  
 1621 tctaccaca gtaattcata tttaaatatg ctttttagatt agtgtggata ctattgctgc  
 1681 tgtgttgcta cctgacctt ttctggggg ggtacctcag aatgagcat ttgagggcaa  
 1741 gcaaaaaagc cctcttcac ctccagagc aacaaagagg cagcagaaat ggggaaagat  
 40     1801 tgtgagaggc agggcttggg tctagacctg gacttaggca agatatgttg cctcaacc  
 1861 tgagttttct tatatgtaaa aagggaaagt tgggctggac tagatgaggt caagatttgc  
 1921 cattctggga ggctgatatt ccagagaatc aaaattaatc ctaaaccaaa gctttatggc  
 1981 tgctacagag acatgtcaca tttctgagac ttgtcaccaa gagttgtcc ctgagacttt  
 2041 ggcgctgttg aatgcaaaga caaggatggc caccttctg ttcttgctg ttgtcctcag

2101 ctgagagcag tctcggtaaa ggtggcaaag attctgtgac ctcagaccgg ggaccaaagt  
 2161 cttgggagtc tgatggccgg gctgggccac cattctcata gctctcattc tgtttggagc  
 2221 aaccaaagga tttgtgtgaa gttatttggg aaaggacctt aactgagcag taatcttttt  
 2281 tctgtatatt tggaatgttt ttcattctga cctgttctgt cagtgattct actgaaaaac  
 5 2341 aatttaataca atataaaaat gttcaagcta tgcaac

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
-----------	-------------	---------------------	-----------

ATP-binding cassette, sub-family C (CFTR/MRP), member 3	ABCC3	NM_020037	2
---	-------	-----------	---

10

1 ctcgggccc cgtctgccc gccgctgggt cggaccgcgc tcgccttctc tgcagccgcg  
 61 cctcggcccc atggacgccc tgtgcggttc cggggagctc ggctccaagt tctgggactc  
 121 caacctgtct gtgcacacag aaaaccgga cctcactccc tgcttcaga actccctgct  
 15 181 ggctgggtg cctgcatct acctgtgggt cgccctgccc tgctactgct tctacctgct  
 241 gcaccattgt cgtggctaca tcatcctctc ccacctgtcc aagctcaaga tggctctggg  
 301 tgtcctgctg tgggtgctct cctgggcgga ccttttttac tccttccatg gcctgggtcca  
 361 tggccgggccc cctgcccctg ttttctttgt caccctctg gtggtggggg tcaccatgct  
 421 gctggccacc ctgctgatac agtatgagcg gctgcagggc gtacagtctt cgggggtcct  
 20 481 cattatcttc tggttcctgt gtgtggctct cgccatcgct ccattccgct ccaagatcct  
 541 tttagccaag gcagagggtg agatctcaga ccccttcgct tcaccacct tctacatcca  
 601 ctttgcccctg gtactctcta cctcatctt ggctgcttc agggagaaac ctccattttt  
 661 ctcgcgcaaag aatgtcgacc ctaaccctta ccttgagacc agcgtggct tctctcccg  
 721 cctgtttttc tgggtggttca caaagatggc catctatggt taccggcatc cctggagga  
 25 781 gaaggacctg tggtccttaa aggaagagga cagatcccag atggtggtgc agcagctgct  
 841 ggaggcatgg aggaagcagg aaaagcagac ggcacgacac aaggcttcag cagcacctgg  
 901 gaaaaatgcc tccggcgagg acgaggtgct gctgggtgcc cggcccaggc cccggaagcc  
 961 ctccttctctg aaggccctgc tggccacctt cggctccagc ttcctcatca gtgctgctt  
 1021 caagcttata caggacctgc tctccttcat caatccacag ctgctcagca tcctgatcag  
 30 1081 gtttatctcc aaccctatgg cccctcctg gtggggcttc ctggtggctg ggctgatggt  
 1141 cctgtgctcc atgatgcagt cgtgatctt acaacactat taccactaca tctttgtgac  
 1201 tggggtgaag tttcgtactg ggatcatggg tgtcatctac aggaaggctc tggttatcac  
 1261 caactcagtc aaacgtgcgt ccaactgtggg ggaaattgtc aacctcatgt cagtggatgc  
 1321 ccagcgcttc atggacctg ccccttctc caatctgctg tggtcagcac ccctgcagat  
 35 1381 catcctggcg atctacttcc tctggcagaa cctaggtccc tctgtcctgg ctggagtcgc  
 1441 tttcatggtc ttgctgatte cactcaacgg agctgtggcc gtgaagatgc gcgcctcca  
 1501 ggtaaagcaa atgaaattga aggactcgcg catcaagctg atgagtgaga tcctgaacgg  
 1561 catcaaggctg ctgaagctgt acgcctggga gccagcttc ctgaagcagg tggagggcat  
 1621 caggcagggt gagctccagc tgctgcgcac ggcggcctac ctccacacca caaccacctt  
 40 1681 cacctggatg tgcagcccct tcctggtgac cctgatcacc ctctgggtgt acgtgtacgt  
 1741 ggacccaaac aatgtgctgg acgcccagaa ggcctttgtg tctgtgtcct tgtttaatat  
 1801 cttaagactt cccctcaaca tgetgcccc gttaatcagc aacctgactc aggccagtgt  
 1861 gtctctgaaa cggatccagc aattcctgag ccaagaggaa cttgaccccc agagtgtgga  
 1921 aagaaagacc atctccccg gctatgccat caccatacac agtggcactc tcacctgggc  
 45 1981 ccaggacctg cccccactc tgcacagcct agacatccag gtcccgaagc gggactgggt  
 2041 ggccgtgggtg gggcctgtgg gctgtgggaa gtccctcctg gtgtctgccc tctctgggaga  
 2101 gatggagaag ctagaaggca aagtgcacat gaaggctcc gtggcctatg tgccccagca  
 2161 ggcattggatc cagaactgca ctcttcagga aaactgtctt ttcggcaagc cctgaacc

2221 caagcgctac cagcagactc tggaggcctg tgcccttgcta gctgacctgg agatgctgcc  
 2281 tgggtggggat cagacagaga ttggagagaa gggcattaac ctgtctgggg gccagcggca  
 2341 gcggttcagt ctggctcgag ctgtttacag tgatgccgat attttcttgc tggatgacct  
 2401 actgtccgcg gtggactctc atgtggccaa gcacatcttt gaccacgtca tcgggcccaga  
 5 2461 aggcgtgctg gcaggcaaga cgcgagtgct ggtgacgcac ggcattagct tcctgcccc  
 2521 gacagacttc atcattgtgc tagctgatgg acaggtgtct gagatgggccc cgtaccaggc  
 2581 cctgctgcag cgcaacggct cctttgccaa ctttctctgc aactatgccc ccgatgagga  
 2641 ccaagggcac ctggaggaca gctggaccgc gttggaaggt gcagaggata aggaggcact  
 2701 gctgattgaa gacacactca gcaaccacac ggatctgaca gacaatgatc cagtcaacct  
 10 2761 tgtggtccag aagcagttta tgagacagct gactgccttg tcctcagatg gggagggaca  
 2821 gggtcggcct gtaccccgga ggcacctggg tccatcagag aaggtgcagg tgacagaggc  
 2881 gaaggcagat ggggactga cccaggagga gaaagcagcc attggcactg tggagctcag  
 2941 tgtgttctgg gattatgcca aggcctggg gctctgtacc acgctggcca tctgtctcct  
 3001 gtatgtgggt caaagtgcgg ctgccattgg agccaatgtg tggctcagtg cctggacaaa  
 15 3061 tgatgccatg gcagacagta gacagaacaa cacttccttg aggctgggccc tctatgtctg  
 3121 tttaggaatt ctgcaagggt tcttgggtgat gctggcagcc atggccatgg cagcgggtgg  
 3181 catccaggct gcccgtgtgt tgcaccaggg actgctgcac aacaagatac gctcggcaca  
 3241 gtccttcttt gacaccacac catcaggccg catcctgaac tgcttctcca aggacatcta  
 3301 tgccttcttt gacaccacac catcaggccg catcctgaac tgcttctcca aggacatcta  
 20 3361 tgctgtgat gaggttctgg cccctgtcat cctcatgctg ctcaattcct tcttcaacgc  
 3421 cctggctgtg ctctacacct tagtgcagcg cttctatgca gccacatcac ggcaactgaa  
 3481 gcggtcgaa tcagtcagcc gctcacctat ctactcccac ttttcggaga cagtgactgg  
 3541 tgccagtgtc atccgggccc acaaccgcag ccgggatttt gagatcatca gtgatactaa  
 3601 ggtggatgcc aaccagagaa gctgctaccc ctacatcacc tccaaccggg cagaagccgc  
 25 3661 ctccctcget cctgctcct ccaggaattc ccagcaggct ctctgggtgt cagggtcctt  
 3721 gtccctcctt tcccctaage agaaaactgg ccctgcctg cccctgcccc atttctcct  
 3781 catctgatcc cccataggcg gctgagcacc ggagtggagt tcgtggggaa ctgctgggtg  
 3841 ctctttgctg cactatttgc cgtcatcggg aggagcagcc tgaaccgggg gctggtgggc  
 3901 ctttctgtgt cctactcctt gcaggtgaca tttgctctga actggatgat acgaatgatg  
 30 3961 tcagatttgg aatctaacat cgtggctgtg gagagggcca aggagtactc caagacagag  
 4021 acagaggcgc cctgggtggt ggaaggcagc cgccctccc aaggttggcc cccacgtggg  
 4081 gaggtggagt tccggaatta ttctgtgctg taccggccgg gcttagacct ggtgctgaga  
 4141 gacctgagtc tgcatgtgca cgggtggcag aaggtgggga tcgtgggccc cactggggct  
 4201 ggcaagtctt ccatgacctt ttgcctgttc cgcactctgg aggcggcaaa ggggtgaaatc  
 35 4261 cgcattgatg gcctcaatgt ggcagacatc ggccctccatg acctgcgctc tcagctgacc  
 4321 atcatcccgc aggaccccat cctgttctcg gggacctgc gcatgaacct ggacccttc  
 4381 ggcagctact cagaggagga ctttgggtgg gctttggagc tgtcccacct gcacacgttt  
 4441 gtgagctccc agccggcagg cctggacttc cagtgtcag agggcgggga gaatctcagc  
 4501 gtgggcccaga ggcagctcgt gtgctggccc cgagccctgc tccgcaagag ccgcatcctg  
 40 4561 gttttagacg aggccacagc tgccatcgac ctgggactg acaacctcat ccaggctacc  
 4621 atccgcaccc agtttgatac ctgcactgtc ctgacctcg cacaccggct taacactatc  
 4681 atggactaca ccagggtcct ggtcctggac aaaggagtag tagctgaatt tgattctcca  
 4741 gccaacctca ttgcagctag aggcactctc tacgggatgg ccagagatgc tggacttgcc  
 4801 taaaatata tcttgagatt tccctctggc ctttctctgg tttcatcagg aaggaaatga  
 45 4861 caccaaata gtccgcagaa tggacttgat agcaaacact gggggcacc taagattttg  
 4921 cacctgtaaa gtgccttaca gggtaactgt gctgaatgct ttagatgagg aatgatccc  
 4981 caagtgggtg atgacacgcc taaggtcaca gctagtgtga gccagttaga ctagtcccc  
 5041 ggtctcccga ttcccactg agtgttattt gcacactgca ctgttttcaa ataacgattt  
 5101 tatgaaatga cctctgtcct cctctgatt tttcatattt tccaaagt tctgttctgt  
 50 5161 tttttaataa aaagcttttt cctcctggaa cagaagacag ctgctgggtc aggccacccc  
 5221 taggaactca gtctgtact ctgggtgct gcctgaatcc attaaaaatg ggagtactga  
 5281 tgaataaaaa ctacatggtc aacagtaaaa

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
specifically androgen-regulated protein	SARG	NM_023938	3

```

1  gtgggggcca ggcagcacag atgaagcatt tacctatcta ggtaagtcag gaggagctca
61  aaaggagaag aaaacagtag gaggcagggg aagcagcctc tgtctccatc tctgcccttt
121  gaaacaaaag ggtatttctt ttctctcttc agcccccaac ccagtggagg cccggcttgg
5  181  gacattgttc acttcccctc gtttcccctc tagaagcccc ctttgccatc cctgcacctt
241  gtttcgggtg atgcccgaga gggagctgtg gccagcgggg actggctcag aaccgctgac
301  ccgtgtcggc agctgtgaca gcatgatgag cagcacctcc acccgctctg gatctagtga
361  tagcagctac gacttctgtt ccactgaaga gaaggagtgt ctgctcttcc tggaggagac
421  cattggctca ctggacacgg aggctgacag ggactgttcc actgacgagt ctgagccagc
10  481  cacaactccc agaggtttcc gagcactgcc cataacccaa cccactcccc ggggaggtcc
541  agaggagacc atcactcagc aaggacgaac gccaaaggaca gtaactgagt ccagctcatc
601  ccaccctcct gagccccagg gcctaggcct caggctctggc tcctacagcc tccttaggaa
661  tatccacatt gccagaagcc agaacttcag gaaaagcacc acccaggcta gcagtcacaa
721  ccttgagaaa ccggggaggc ttgctgccaga gcctgagaaa gaacaggtca gccagagcag
15  781  ccaaccaggc caggcacctg ccagccccca ggaggctgcc cttgacttgg acgtggtgct
841  catccctccg ccagaagctt tccgggacac ccagccagag cagtgtaggg aagccagcct
901  gccogagggg ccaggacagc agggccacac accccagctc cacacaccat ccagctccca
961  ggaagagag cagactcctt cagaagccat gtcccaaaaa gccaaagaaa cagtctcaac
1021  caggtacaca caaccacagc ctctctctgc agggttgcct cagaatgcaa gagctgaaga
20  1081  tgctcccctc tcatcagggg aggacccaaa cagccgacta gtccccctca caacccttaa
1141  gccccggaag ctgccaccta atattgttct gaagagcagc cgaagcagtt tccacagtga
1201  cccccagcac tggctgtccc gccacactga ggctgcccct ggagattctg gcctgatctc
1261  ctgttcaactg caagagcaga gaaaagcagc taaagaagct ctagagaagc tggggctacc
1321  ccaggatcaa gatgagcctg gactccactt aagtaagccc accagctcca tcagacccaa
25  1381  ggagacacgg gcccagcatc tgtccccagc tccaggctctg gctcagcctg cagctccagc
1441  ccaggcctca gcagctattc ctgctgtctg gaaggctctg gctcaagctc cggctccagc
1501  tccaggtcca gctcagggac ctttgccaat gaagtctcca gctccaggca atgttcagc
1561  tagcaaatct atgccaattc ctatocctaa ggccccagg gcaaacagtg ccoctgactcc
1621  accgaagcca ggtcagggc tgactctcca ggagagcaac acccctggcc tgagacagat
30  1681  gaacttcaag tccaacactc tggagcgtctc aggcgtggga ctgagcagct acctttcaac
1741  tgagaaaagat gccagcccc aaccagcac ttctctggga aagggtcct tcttgacaa
1801  gatctcgccc agtgtcttac gtaattctcg gccccgccc gcctcctgg gcacggggaa
1861  agattttgca ggtatccagg taggcaagct ggctgacctg gagcaggagc agagctccaa
1921  gcgctgttcc taccaaggac agagcctgca caagcttctc cgccccctc gtgtcagtgt
35  1981  caagatctcc ccaaaggggtg tccccaatga acacagaagg gaggccctga agaagctggg
2041  actgttgaag gagtagactc tgcgaccagt acagaccctg tcctggctga acaagaagag
2101  acacatgctc cacttgggag cctttgccac cacgcaactc agggctcaag atgaatggga
2161  gggagagatt tgagtccaag catacattta tattcagtg tgtgccattg agttcccatg
2221  tggatcattc tgaaggtgat ctccacaaga ggggtgtgtg gtgtgtgttt ggtgtgtgtg
40  2281  tggagggggg gccgctggat acatcactga agctattgat ataacacaat gactcactgt
2341  tcagaatttt gctcttgta gatgttttct tacattgggt agagtccagc ctagtgagag
2401  ctgagtgaag gggctggcca tgctgagac aaaaagtcaa atgagacaat ggacgtgtca
2461  atgacttgaa aaaaagtcac atccagcaaa tgcagggtca catgaaatat gggcctcctg
2521  gaatccctac agtggatgga gactggctca taccttgcca gatccctctc tcagttccag
45  2581  ccttctggac aaggcctggg ctaagaggag ctgattcgtt atctcttca cactgacctt
2641  ctcagtatca ccagtcctca agacaggata cgtccctgta acccaatctc tgggtgatt
2701  gatagcagaa cagctcttgt ttgtctgaga aggcaggata agtgaccaca tatttatgcc
2761  actaccctca ccagggagag tccttctcca caggcttgat aaattcaatc accaactgtg
2821  ctgtcgtccc tgactctgct actcccgttc ttctgcttt cctgctccgt atctcagctt
50  2881  gcactgacct cagggctggg ctgacatcaa gatgggagcc cagcccacgg gctttataaa
2941  caccacaaga ccgtttcaga tcttctctgt gctgatgcag gtagttttaa atttttctca

```

3001 gttccagtga tagaaaaccc acacaataca tctcttgcca gtcttaatag aatatcagag  
 3061 gtaagagggg cctcagagaa gctctgacgc agtgctgctg gggaagggaa gtgactaacc  
 3121 ccgggtcagc ctgccattta gggaaagagc tgaggttctt acccttggtg catgctgcca  
 3181 cctctcctta gccagtgtct ttgtacatcc acacagcacc ctaaggagcc atagtcacca  
 5 3241 tcaaagactc aaccctaagg cccttcaaga tctcaaagtg cttctgaag catcagagat  
 3301 taaatattgt tcaaactaat agttattgct gtggctttta atttatctt tgggaagatag  
 3361 ctatatggta actcatcatt aaccagaaca cctctcccct caaattccgt gaccaagtgtg  
 3421 tgcagcttga gcaaatgccg aaagagggta ttatgggtgg gtgggtgtggg cttgcaaata  
 3481 caagcttgga ggtgagacat ggccagacat gactcctgct tccccttagg aagtaaatct  
 10 3541 tacttatggg tgtgaactgc ttggagtcca ggatgccag atgtgagggg cagatgaagg  
 3601 gaatgttgct ggaaaggtgc cttttaaggc tgctgagaat ttctggactg tgcctgatg  
 3661 gacgcagcac catcaaagcc cagaatttct gaaaacgggtg acaaggtaa cataaggaca  
 3721 acaaaactc cacctgtca tggatgtga ggtgtgggtg tggcggttc tgtgtacgtt  
 3781 tgctcataca cgcacatcca aaagcctgtg cctcattcct ggccatgggt gaggacttg  
 15 3841 tctgtcacgg ctgatgagga ctcccacaac cgtgaggac tgattccaca tgctatttc agctgagtgtg  
 3901 agaaagagag aaagctgcct tctgaagcc aagaatgatc acctattcc ttattcatta  
 3961 atttctgtgt ctatttcaac ccataacctg aagaatgatc acctattcc ttattcatta  
 4021 atttctgtga ttaatagga aacttgggaa tagctataaa gtaaaacttg ggtggaacct  
 4081 ggggcccctg catcacacaa gtgtgattag gatggtcaag gtcacagga gtacagccta  
 20 4141 ttatattccc acatcctgag aaaggtcatt tctcccacac acgacaaagt cacagacatc  
 4201 ctgcaacctgc cactaggcat cctcatccta ctgacatgcc catttctoca gttttcttaa  
 4261 tctgagactc cttcccttg ttttttaaag ataccgtgct tctccacatc ctcatccttc  
 4321 aaggagcata ttttgcctt aggatggtct ttgggattca agaatagaat aataaatcca  
 4381 aacttggtca ttcccatttt gaagagatgc aagagggccc agtgaggaca tccgctccc  
 25 4441 tgaaagtggg gctagacaga gctgaggtca ttgtatctgt gtatccacat aggatattctc  
 4501 ttaattcagc ttgaattgat ggggagggag gtaagagtag ggtcagagtt actcatcct  
 4561 tttcaaagaa ttgtgggtgg aagtttgtaa aggccattca tttgattttc aaaatcaaag  
 4621 cgacagctct acttccactt ggccttagat ctctgctata ccctgccata gccttgatgc  
 4681 cactgggcac aagccacctg ccaaatacag gagtggcctc tcccagcctg gcatgatagg  
 30 4741 ggggtctgtg ccctcagatg tgttgacagc tgctctctctg aattgccaca cctgtgctac  
 4801 acttgggaatt ctgtgctctg actctgcagg gtaggaccac gtgccatctc acacagaggt  
 4861 caaccgatga gcccactcac tegtacatgc cttcttccac agtgggaagc atgatctggc  
 4921 aggggccgccc ctgtaggctg gggatgggct gctgtgtgaa tgttgacgtt cgtttcatgg  
 4981 agaaagggga ggtgaaagat tgaagagcag gttcctgtca atgttctgag ttcgagctgg  
 35 5041 aggtgtagat tgaatagtct acatggtctg tgagtgtgtg agatgaacct ttccatcctt  
 5101 tgacacctgg ttgtatgtgt aggctaagaa ggaaggacc tctgtcagt gtgcaaagct  
 5161 gtaactctcat ggactagagg agagggggcc aaggggatgg acaggagaag tcatgcagaa  
 5221 tctaagcagg aatgcagata gaacacatct aggctctttt cccaggaga gtgatgatgg  
 5281 agcatataga tctggctcaa attcagcctc catcacttac cagtcaggaa ccctggcgat  
 40 5341 atcactttaa ctttctgaac ctgagagtct tcacctataa gacggggaaa ataataccac  
 5401 cttttcaaga ttgttgagat aaataagtga tataaacat gtaaagctta gttctggcca  
 5461 cagtgtagct actcaataaa tgataaact

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
phosphatidic acid phosphatase type 2C	PPAP2C	NM_177526	4

45 1 ctctctccg cgcggggcgg gctccgcgcc acgtgactcc gcgggcgggc cgggacgcga  
 61 cgggacgcgc tgggaccggc gtcgggggtc gcgggacca tgcagcggag cctccctgcc  
 121 cttcgctatc ctgacgctgg tgaacgcccc gtacaagcga ggattttact gcggggatga  
 181 ctccatccgg taccctacc gtccagatac catcaccac gggctcatgg ctgggggtcac  
 241 catcaaggcc accgtcatcc ttgtctcggc cggggaagcc tacctgggtg acacagaccg  
 50 301 gctctattct cgtcggact tcaacaacta cgtggctgct gtatacaagg tgctggggac

361 cttcctgttt ggggctgccg tgagccagtc totgacagac ctggccaagt acatgattgg  
 421 gcgtctgagg cccaacttcc tagccgtctg cgaccccgcac tggagccggg tcaactgctc  
 481 ggtctatgtg cagctggaga aggtgtgcag gggaaaccct gctgatgtca ccgaggccag  
 541 gttgtctttc tactcgggac actcttcctt tgggatgtac tgcattggtg tcttggcgct  
 5 601 gtatgtgcag gcacgactct gttggaagtg ggcacggctg ctgacacca cagtccagtt  
 661 cttcctggtg gcctttgccc tctacgtggg ctacaccgcg gtgtctgatt acaaacacca  
 721 ctggagcgat gtccttgttg gcctcctgca gggggcactg gtggctgcc tcaactgtctg  
 781 ctacatctca gacttcttca aagcccagcc cccacagcac tgtctgaagg aggaggagct  
 841 ggaacggaag cccagcctgt cactgacgtt gaccctgggc gaggctgacc acaaccacta  
 10 901 tggatacccg cactcctcct cctgaggccg gaccccgcg aggcagggag ctgctgtgag  
 961 tccagctgag gccacccag gtggtccctc cagccctggt taggactga gggctctgga  
 1021 cgggctccag gaacctggg ctgatgggag cagtgagcgg gctccgctgc ccctgcct  
 1081 gcaactggacc aggagtctgg agatgcctgg gtagccctca gcatttggag gggaaacctgt  
 1141 tcccgctcgt ccccaaatat ccccttctt ttatgggggt aaggaagga ccgagagatc  
 15 1201 agatagttgc tgttttgtaa aatgtaatgt atatgtggtt ttagtaaaa tagggcacct  
 1261 gtttcacaaa

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
neural proliferation, differentiation and control, 1	NPDC1	NM_015392	5

20 1 ggcgcctcgc ccggcgccctc catcccggat ccttgcctgca gcgtcagcgc gcgcgccctg  
 61 gcctttcttc ttctctctcc tctctcttgg catccgcctc ttcttctctc tgcgtctctc  
 121 cccgctgcct ccgctgctcc cgacgcggag cccggagccc gcgcccagcc cctggcctcg  
 181 cgggtccatg ctgccccggc ggcggcgctg aaggatggcg acgcccgtgc ctccgccctc  
 25 241 cccgcggcac ctgcggtgctg tgcggctgct gctctccggc ctctgctctg gcgcccctc  
 301 gcgtggagcc gccgcccggc acccgatgt agccgcctgt cccgggagcc tggactgtgc  
 361 cctgaagagg cgggcaaggt gtcctcctgg tgcacatgcc tgtgggccc gcttcagcc  
 421 cttccaggag gaccagcaag ggctctgtgt gccaggatg cgcggcctc caggcggggg  
 481 ccggccccag cccagactgg aagatgagat tgacttctg gccaggagc ttgcccggaa  
 30 541 ggagtctgga cactcaactc cgcccctacc caaggaccga cagcggctcc cggagcctgc  
 601 caccctgggc ttctcggcac gggggcaggg gctggagctg ggctccctc ccaactccag  
 661 aacccccacg cccacgcccc acaactccat gggctccct gtgtcatccg acccgggtga  
 721 catgtcgcct ctggagcccc ggggagggca aggcgacggc ctgcctctg tgcgtatcct  
 781 ggcgttctgt gtggccgggt cagccgcct ctccgtagcc tccctctgt ggtgcaggct  
 35 841 gcatcgtgag atccgctga ctcagaaggc cgactacgcc actgcgaagg cccctggctc  
 901 acctgcagct ccccgatct cgctggggga ccaacggctg gcacagagcg cggagatgta  
 961 ccactaccag caccaacggc aacagatgct gtgcctggag cggcataaag agccacccaa  
 1021 ggagctggac acggctcct cggatgagga gaatgaggac ggagacttca cgggtgtacga  
 1081 gtgccccggc ctggccccga cgggggaaat ggaggtgcgc aacctctgt tcgaccacgc  
 40 1141 cgcactgtcc gcgcccctgc cggccccag ctaccgcct gcaactgcat gacctggagg  
 1201 cagacagacg cccactgct ccccgacctc gagcccccg gggaggggca gggcctggag  
 1261 cttcccacta aaaacatggt ttgatgctgt gtgcttttgg ctgggctcgc ggctccagcc  
 1321 cctgggacct cttgccaggg agacccccga acctttgtgc caggacacct cctggctccc  
 1381 tgcacctctc ctgttcgggt tagaccccc aactggagg ggcattggaga accgtagagc  
 45 1441 gcaggaacgg gtgggtaatt ctagagacaa aagccaatta aagtccattt cagaaaaaaa

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
C-terminal tensin-like	CTEN	NM_032865	6

```

1 gggcaacagt ctgccacct gtggacacca gatcctggga gtcctgggtt agcaagtgag
61 atctctggga tgtcagtgag gctggttgaa gaccagaggt aaactgcaga ggtcaccacc
5 121 cccaccatgt cccaggtgat gtccagccca ctgctggcag gaggccatgc tgtcagcttg
181 ggcgcttggt atgagcccag gaggaccctg caccagcac ccagcccag cctgccacc
241 cagtgttctt actacaccac ggaaggctgg ggagcccagg ccctgatggc ccccgtgccc
301 tgcattggggc ccctggccg actccagcaa gccccacagg tggaggccaa agccacctgc
361 ttctgcccgt ccctgggtga gaaggccttg gggaccccag aggacctga ctctacatt
10 421 gacttctcac tggagagcct caatcagatg atcctggaac tggaccccac ctccagctg
481 cttccccccag ggactggggg ctcccaggt gagctggccc agagcaccat gtcaatgaga
541 aagaaggagg aatctgaagc cttggacata aagtacatcg agtgacctc cgccagatca
601 aggtgccacg attggcccc gactgctcc agcccctctg tcacccgcc cttcggctcc
661 cctcgcagtg gtggcctcct ctttccaga gacgtcccc gagagacacg aagcagcagt
15 721 gagagcctca tcttctctgg gaaccagggc agggggcacc agcgcctct gccccctca
781 gagggtctct cccctcgacc cccaaattcc ccagcatct caatcccttg catggggagc
841 aaggcctcga gccccatgg tttgggctcc cgctgggtg cttctccaag actggagaag
901 cggctgggag gcttgcccc acaggggggc agcaggatct ctgtgctgtc agccagccca
961 gtgtctgatg tcagctatat gtttgggaag agccagtccc tctgcactc cagcaactcc
20 1021 agccatcagt catcttccag atccttggaa agtccagcca actcttctcc cagcctccac
1081 agccttggct cagtgtccct gtgtacaaga cccagtgact tccaggtccc cagaaacccc
1141 accctaacca tgggccaacc cagaacaccc cactctccac cactggccaa agaacatgcc
1201 agcatctgcc ccccatccat caccaactcc atggtggaca taccattgt gctgatcaac
1261 ggctgccacg aaccagggtc ttctccacc cagcggacc caggacacca gaactccgtt
25 1321 caacctggag ctgcttctcc cagcaacccc tgtccagcca ccaggagcaa cagccagacc
1381 ctgtcagatg ccccttttac cacatgccc gagggteccg ccaggacat gcagcccacc
1441 atgaagtctg tgatggacac atctaatac tggtttaagc caaacatcac ccgagagcaa
1501 gcaatcgagc tgctgaggaa ggaggagcca ggggctttg tcataagggg cagctcttca
1561 taccgaggct cctcggcct ggcctgaag gtgcaggagg ttcccgcgtc tgctcagaat
30 1621 cgaccagggtg aggacagcaa tgacctcacc cgaccttcc tcatcgagtc gtctgcaaaa
1681 ggagtgcatt tcaaaggagc agatgaggag ccctactttg ggagcctctc tgcttctgtg
1741 tgccagcatt ccatcatggc cctggccctg ccctgcaaac tcaccatccc acagagagaa
1801 ctgggagggtg cagatggggc ctcgactct acagacagcc cagcctctct ccagaagaaa
1861 tctgcgggct gccacaccct gtacctgagc tcagtgageg tggagacct gactggagcc
35 1921 ctggccgtgc agaaagccat ctccaccacc tttgagaggg acatcctccc cagcccacc
1981 gtggtccact tcgaagtcac agagcagggc atcactctga ctgatgtcca gaggaaggtg
2041 ttttccggc gccattacc actcaccacc ctccgcttct gtggtatgga ccctgagcaa
2101 cggagtggc agaagtactg caaacctcc tggatcttg gtttggggc caagagccag
2161 acagagcctc aggagaacgt atgccacctc tttgaggagt atgacatggt ccagccagcc
40 2221 tcgcaggtca tcggcctggt gactgctctg ctgcaggagc cagaaaggat gtaggggaga
2281 gactgcctgt gcacctaac aacacctcca ggggctcgt aaggagcccc cctccacccc
2341 ctgaatgggt gtggcttgtg gccatattga cagaccaatc tatgggacta gggggattgg
2401 catcaagttg acacccttga acctgctatg gccttcagca gtcaccatca tccagacccc
2461 ccgggctca gtttctcaa tcatagaaga agaccaatag acaagatcag ctgttcttag
45 2521 atgctgggtg gcatttgaac atgctcctcc atgattctga agcatgcaca cctctgaaga
2581 cccctgcatg aaaataacct ccaaggacct tctgaccca tcgacctggg ccctgcccac
2641 acaacagtct gagcaagaga cctgcagccc ctgttctgtg gcagacagca ggtgcctggc
2701 ggtgaccac ggggctcctg gcttgagct ggtgatggtc aagaactgac taaaaacag
2761 gaatggatag actctatttc ctccatatac tgttctctg tctctttcc cactttctgg
50 2821 gtggcttttt ggggtccacc agccaggatg ctgcaggcca agctgggtgt ggtatttagg
2881 gcagctcagc agggggaact tgtccccatg gtcagaggag acccagctgt cctgcacccc

```

2941 cttgcagatg agtatcacc cactctttct ttccacttgg tttttatatt tattttttt  
 3001 gagacagagt ctcaactgtca cccaggctga actgcagtgg tgtgatctag gctcaactgca  
 3061 acctccacct cccaggttca agcaattatc ctgcctcagg ctcccagagta gctgggatta  
 3121 caggcatgtg caactcacc agctaatttt gtatttttag tagagacagg gtttcacat  
 5 3181 gttggccagg ctggtcttga actcctgacc gcaggtaatc cacctgcttc ggctcccaa  
 3241 agtgctggga ttacaggcgc aagccacca gccagcttc tttccattcc ttgataggcg  
 3301 agtattccaa agctggtatc gtagctgccc taatggtgca tattaggcgg cgggggcaga  
 3361 gataagggcc atctctctgt gattctgctt cagctcctgt cttgctgagc cctcccccaa  
 3421 cccacgetcc aacacacaca cacacacaca cacacacaca cacacacaca cacacacaca  
 10 3481 cacgcccctc tactgctatg tggcttcaac cagcctcaca gccacacggg ggaagcagag  
 3541 agtcaagaat gcaaagaggc cgcttcccta agaggcttgg aggagctggg ctctatccca  
 3601 caccaccccc caccacccc ccaccagcc tccagaagct ggaaccattt ctcccgcagg  
 3661 cctgagttcc taaggaaacc accctaccgg ggtggaaggg agggctcagg aagaaacca  
 3721 ctcttgctct acgaggagca agtgctgccc ccctcccagc agccagcctt gccaaagtgt  
 15 3781 cattatcttt ggccaaggct gggcctgacg gttatgattt cagcctgagg cctgcaggag  
 3841 aggctgagat cagcccacc agccagtggc cgagcactgc cccgccgcca aagtctgcag  
 3901 aatgtgagat gaggttctca aggtcacagg ccccagctcc agcctggggg ctggcagagg  
 3961 cccccatata ctctgctaca gctcctatca tgaaaaataa aatgt

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>RAB25, member RAS oncogene family</b>	RAB25	NM_020387	7

20

1 ctctgcttcc ttacagcacc cccacctgcc agagctgatc ctccctagcc cctgacctaac  
 61 cttgagttgg cccccaatcc ctctggtctgc agaagtcccc ttacccccaa tgagaggagg  
 121 ggcaggacca gatcttttga gagctgaggg ttgagggcat tgagccaaca  
 cacagatttg  
 25 181 tcgctctgt cccgaagac acctgcacc tccatgcgga gccaagatgg  
 ggaatggaac  
 241 tgaggaagat tataactttg tcttcaagg ggtgctgatc ggcgaatcag  
 gtgtgggaa  
 301 gaccaatcta ctctccgat tcacgcgcaa tgagttcagc cacgacagcc  
 30 gcaccacat  
 361 cgggggtgag ttctccacc gcaactgtgat gttgggcacc gctgctgtca  
 aggtcagat  
 421 ctgggacaca gctggcctgg agcggtaacc agccatcacc tcggcgact  
 atcgtggtgc  
 35 481 agtgggggcc ctctggtgt ttgacctaac caagcaccag acctatgctg  
 tgggtggagc  
 541 atggctgaag gagctctatg acctgctga agccacgatc gtcgtcatgc  
 tcgtgggtaa  
 601 caaaagtgac ctacagccagg cccgggaagt gccactgag gaggccccgaa  
 40 tgttcgctga  
 661 aaacaatgga ctgctcttcc tggagacctc agcctggac tctaccaatg  
 ttgagctagc  
 721 ctttgagact gtctgaaag aaatctttgc gaagggttcc aagcagagac  
 agaacagcat  
 45 781 ccgaccaat gccatcactc tgggcagtgc ccaggctgga caggagcctg  
 gccctgggga  
 841 gaagagggcc tgttgcata gcctctgacc ttggccagca ccacctgccc  
 ccaactggctt  
 901 tttggtgccc cttgtcccca cttcagcccc aggaccttc cttgacctt  
 50 ggttccagat

961 atcagactgt tccctgttca cagcaccctc agggctctta ggtcttcatg  
 ccctatcaca  
 1021 aatacctctt ttatctgtcc acccctcaca gactaggacc ctcaaataaa  
 gctgttttat  
 5 1081 atcaaaaaaa

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
hephaestin	HEPH	NM_138737	8

1 gcccagcctg cctggagaaa agtgtctgct cctagccaag atctctcat  
 10 cacaaaagta  
 61 atgtgggcca tggagtcagg ccacctcctc tgggctctgc tgttcatgca  
 gtccttgtgg  
 121 cctcaactga ctgatggagc cactcgagtc tactacctgg gcatccggga  
 tgtgcagtgg  
 15 181 aactatgctc ccaagggag aaatgtcctc acgaaccagc ctctggacag  
 tgacatagtg  
 241 gcttccagct tcttaaagtc tgacaagaac cggatagggg gaacctaca  
 gaagaccatc  
 301 tataaagaat acaaggatga ctcatacaca gatgaagtgg cccagcctgc  
 20 ctggttgggc  
 361 ttcttggggc cagtgttgca ggctgaagtg ggggatgtca ttcttattca  
 cctgaagaat  
 421 tttgccactc gtcctatac catccacct catggtgtct tctacgagaa  
 ggactctgaa  
 25 481 ggttccctat acccagatgg ctctctggg ccaactgaaag ctgatgactc  
 tgttcccccg  
 541 gggggcagcc atatctaca ctggaccatt ccagaaggcc atgcaccac  
 cgatgctgac  
 601 ccagcgtgcc tcacctggat ctaccattct catgtagatg ctccacgaga  
 30 cattgcaact  
 661 ggctaattg ggctctcat cacctgtaa agaggagccc tggatgggaa  
 ctccccctc  
 721 caacgccagg atgtagacca tgatttctc ctctcttca gtgtggtaga  
 tgagaacctc  
 35 781 agctggcatc tcaatgagaa cattgccact tactgctcag atctgcttc  
 agtggacaaa  
 841 gaagatgaga catttcagga gagcaatagg atgcatgcaa tcaatggctt  
 tgtttttggg  
 901 aatttacctg agctgaacat gtgtgcacag aaacgtgtgg cctggcactt  
 40 gtttggcatg  
 961 ggcaatgaaa ttgatgtcca cacagcattt ttccatggac agatgctgac  
 taccgtgga  
 1021 caccacactg atgtggctaa catctttcca gccaccttg tgactgctga  
 gatggtgccc  
 45 1081 tgggaacctg gtacctggtt aattagctgc caagtgaaca gtcactttcg  
 agatggcatg  
 1141 caggcactct acaaggtaaa gtcttctctc atggccctc ctgtggacct  
 gctcacaggc  
 1201 aaagtctgac agtacttcat tgaggcccat gagattcaat gggactatgg  
 50 cccgatgggg

1261 catgatggga gtactgggaa gaatttgaga gagccaggca gtatctcaga  
 taagtttttc  
 1321 cagaagagct ccagccgaat tgggggcact tactggaaag tgcgatatga  
 agcctttcaa  
 5 1381 gatgagacat tccaagagaa gatgcatttg gaggaagata ggcattcttg  
 aatcctgggg  
 1441 ccagtgatcc gggctgaggt gggtgacacc attcaggtgg tcttctacaa  
 ccgtgcctcc  
 1501 cagccattca gcatgcagcc ccatggggtc ttttatgaga aagactatga  
 10 aggcactgtg  
 1561 tacaatgatg gctcatctta ccctggcttg gttgccaagc cctttgagaa  
 agtaacatac  
 1621 cgctggacag tccccctca tgccgggtccc actgctcagg atcctgcttg  
 tctcacttgg  
 15 1681 atgtacttct ctgctgcaga tcccataaga gacacaaatt ctggcctggt  
 gggcccgtg  
 1741 ctggtgtgca gggctggtgc cttgggtgca gatggcaagc agaaaggggt  
 ggataaagaa  
 1801 ttctttcttc tcttactgt gttggatgag aacaagagct ggtacagcaa  
 20 tgccaatcaa  
 1861 gcagctgcta tgttggattt ccgactgctt tcagaggata ttgagggctt  
 ccaagactcc  
 1921 aatcggatgc atgccattaa tgggtttctg ttctctaacc tgcccaggct  
 ggacatgtgc  
 25 1981 aagggtgaca cagtggcctg gcacctgctc ggccctgggca cagagactga  
 tgtgcatgga  
 2041 gtcgatgttc agggcaacac tgtgcagctt cagggcatga ggaaggggtgc  
 agctatgctc  
 2101 tttcctcata cctttgtcat ggccatcatg cagcctgaca accttgggac  
 30 atttgagatt  
 2161 tattgccagg caggcagcca tcgagaagca gggatgaggg caatctataa  
 tgtctcccag  
 2221 tgcctggcc accaagccac ccctcgcaa cgctaccaag ctgcaagaat  
 ctactatatac  
 35 2281 atggcagaag aagtagagtg ggactattgc cctgaccgga gctgggaacg  
 ggaatggcac  
 2341 aaccagtctg agaaggacag ttatggttac attttctga gcaacaagga  
 tgggctcctg  
 2401 ggttccagat acaagaaagc tgtattcagg gaatacactg atggtacatt  
 40 caggatccct  
 2461 cggccaagga ctggaccaga agaacacttg ggaatcttgg gtccacttat  
 caaaggtgaa  
 2521 gttggtgata tcctgactgt ggtattcaag aataatgcca gccgcccta  
 ctctgtgcat  
 45 2581 gctcatggag tgctagaatc tactactgtc tggccactgg ctgctgagcc  
 tggtaggtg  
 2641 gtcacttatac agtggaaacat cccagagagg tctggccctg ggcccaatga  
 ctctgcttgt  
 2701 gtttcttggga tctattatct tgcagtggat cccatcaagg acatgtatag  
 50 tggcctggtg  
 2761 gggcccttgg ctatctgcca aaagggcatc ctggagcccc atggaggacg  
 gagtgacatg  
 2821 gatcgggaat ttgcattggt gttcttgatt tttgatgaaa ataagtcttg  
 gtatttggag

2881 gaaaatgtgg caacccatgg gtcccaggat ccaggcagta ttaacctaca  
 ggatgaaact  
 2941 ttcttgagaga gcaataaaat gcatgcaatc aatgggaaac tctatgcaa  
 ccttaggggt  
 5 3001 cttaccatgt accaaggaga acgagtggcc tggtagatgc tggccatggg  
 ccaagatgtg  
 3061 gatctacaca ccatccactt tcatgcagag agcttcctct atcggaatgg  
 cgagaactac.  
 3121 cgggcagatg tggtagatct gttcccaggg acttttgagg ttgtggagat  
 10 ggtggccagc  
 3181 aaccctggga catggctgat gcactgccat gtgactgacc atgtccatgc  
 tggcatggag  
 3241 accctcttca ctgttttttc tcgaacagaa cacttaagcc ctctcaccgt  
 catcaccaaa  
 15 3301 gagactgaaa aagcagtgcc ccccagagac attgaagaag gcaatgtgaa  
 gatgctgggc  
 3361 atgcagatcc ccataaagaa tgtagatg ctggcctctg ttttgggtgc  
 cattagtgtc  
 3421 accctctgctc tcgttgttct ggctcttggg ggagtggttt ggtaccaaca  
 20 tcgacagaga  
 3481 aagctacgac gcaataggag gtccatcctg gatgacagct tcaagcttct  
 gtctttcaaa  
 3541 cagtaacatc tggagcctgg agatctctc aggaagcaca tctgtagtgc  
 actcccagca  
 25 3601 ggccatggac tagtcactaa ccccacactc aaaggggcat gggtagtggg  
 gaagcagaag  
 3661 gagcaatcaa gcttatctgg atatttcttt ctttatttat tttacatgga  
 aataatatga  
 3721 tttcactttt tctttagttt ctttgcctca cgtgggcacc tggcactaag  
 30 ggagtacctt  
 3781 attatcctac atcgcaaatt tcaacagcta cattatattt cttctgaca  
 cttggaaggt  
 3841 attgaaattt ctagaaatgt atccttctca caaagtagag accaagagaa  
 aaactcattg  
 35 3901 attgggtttc tacttctttc aaggactcag gaaatttcac tttgaactga  
 ggccaagtga  
 3961 gctgttaaga taaccacac ttaaactaaa ggctaagaat ataggcttga  
 tgggaaattg  
 4021 aaggtaggct gagtattggg aatccaaatt gaattttgat tctccttggc  
 40 agtgaactac.  
 4081 tttgaagaag tggtagatgc gttgttgcct ccatgagcat gtacaacctc  
 tggagctaga  
 4141 agctcctcag gaaagccagt tctccaagtt cttaacctgt ggcactgaaa  
 ggaatgttga  
 45 4201 gttacctctt catgttttag acagcaaacc ctatccatta aagtacttgt tagaacactg

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
thiopurine S-methyltransferase	TPMT	NM_000367	9

```

1  gcggggcggag gcggggcgcg gagaagtggc ggaggtggaa gcggaggcgt
accgcccct
5   61  ggggacgtca ttggtggcgg aggcaatggc cggcaaccag ctgtaagcga
ggcacggaag
121 acatatgctt gtgagacaaa ggtgtctctg aaactatgga tggtaacaaga
acttcacttg
181 acattgaaga gtactcggat actgaggtac agaaaaacca agtactaact
10  ctggaagaat
241 ggcaagacaa gtgggtgaac ggcaagactg cttttcatca ggaacaagga
catcagctat
301 taagaagca ttagatact ttccttaaag gcaagagtgg actgagggta
tttttctctc
15  361 tttgoggaaa agcggttgag atgaaatggt ttgcagaccg gggacacagt
gtagtgggtg
421 tggaaatcag tgaacttggg atacaagaat tttttacaga gcagaatctt
tcttactcag
481 aagaaccaat caccgaaatt cctggaacca aagtatttaa gagttcttcg
20  gggaacatth
541 cattgtactg ttgcagtatt tttgatcttc ccaggacaaa tattggcaaa
tttgacatga
601 tttgggatag aggagcatta gttgccatca atccaggtga tcgcaaatgc
tatgcagata
25  661 caatgttttc cctcctggga aagaagtttc agtatctcct gtgtgttctt
tcttatgatc
721 caactaaaca tccaggtcca ccattttatg ttccacatgc tgaattgaa
aggttgtttg
781 gtaaaatatg caatatacgt tgtcttgaga aggttgatgc ttttgaagaa
30  cgacataaaa
841 gttggggaat tgactgtctt tttgaaaagt tatatctact tacagaaaag
taaatgagac
901 atagataaaa taaatcaca ctgacatggt tttgaggaat tgaaaattat
gctaaagcct
35  961 gaaaatgtaa tggatgaatt tttaaaattg tttataaatc atatgataga
tctttactaa
1021 aatggcctt ttagtaaagc catttacttt ttctaaaaaa gttttagaag
aaaagatgt
1081 aactaaactt ttaaagtagc tcctttggag aggagattat gatgtgaaag
40  attatgccta
1141 tgtgtcttgc agattgcaag atatthtacc aatcagcatg tgttacctgt
acaattaaaa
1201 aatattttca aatgcaatg catattaaat ataatacaca cagaaaaact
ggcatttatt
45  1261 ttgttttatt tttttgagat ggagtttctg tcttggtgcc caacctggag
tgcaatgggtg
1321 caatctcagc tcaactgcaac ctctgcctcc caggttcagg tgattctcct
gcctcagcct
1381 cctgagtagc tgggattaca ggtgtgcgcc accacgccca gctaattttt
50  tgtatthtta

```

1441 gtagagacag ggtttcacca tgttggtcag gctgatctcg agctcctgac  
 ctcaggtgat  
 1501 ctaccacact cggcctccca aagtgctggg attacaggcg tgagccactg  
 cacctggcct  
 5 1561 gacattcttt atgaaattta gaattggtga agaactataa catttcagta  
 gggttcaagg  
 1621 tgggtcccaaa agttatataa aagattagtt ttactataa acccttgtct  
 ttactcaga  
 1681 tcttagcatc ccttttcaca tggtttctcc atgtatataa cagaatcaag  
 10 aaacaaat  
 1741 taattaaaca atctgtaaca gaatcaagaa acaatacat ttaattaa  
 caatctatat  
 1801 ggaacaaaca ttcccaaatt ctaagaataa attttcttt aagtttctc  
 tgagtttggc  
 15 1861 aattggtggt ttttataatt taatctggtt aaatcatcag gtcttataaa  
 atataatgta  
 1921 cttagagctg gattcatggc tgtttattat gaaaggtag atttctcagt  
 tcttctttaa  
 1981 ccacatthtgg ttatatcaga cagtcctcta taactctgta ctaccaaca  
 20 actaaatggg  
 2041 ttagattggt tagctcatgt taataggatg gttgtgtatt ataaaaaacg  
 agttacgtgt  
 2101 gtgtgtgcac gcatgcacgc acatgtgctg gcttaaagggt tgtaaatgca  
 aggtttgggg  
 25 2161 tcccctttaa cactggtgaa agctacggta ctctcccag agatatgtct  
 tgtcagctc  
 2221 tctagttccc cttggcctgc atgtacaaac ttctacccta gaagctctct  
 gccatcgatg  
 2281 tattctaata gatttgaag gctattaatt tgaagcaact ccttgctcac  
 30 agtgattctt  
 2341 gcttctctga gacctgctcc cagtcgatac tgtgggcttc agaagccatg  
 actccccaac  
 2401 tctgcctgta tcaccggttg aatggacaac taacccgagc tggaccaaca  
 caattctctc  
 35 2461 cagagacttt tgattttact tttatgtaga gacagggtct cactttggtg  
 cccacgctga  
 2521 tgttgaactt gacgtgaggc ctcaagcagt cctcctgtct tggccacca  
 aagtgctagg  
 2581 attacaggta tgagccattg cgctggccct cttcataggc ttttggactt  
 40 gggaaatagaa  
 2641 aagcaacccc gtctctacta aaaatacaaa aaaattagcc aggcgtggg  
 gcacgtgcct  
 2701 gtaatcccag ctacttggga ggctgaggca ggagaatcac ttgaacctag  
 gaggcggagg  
 45 2761 ttgcagtgag ctgagatcat gccactgcac gcaagcctgg gcaacagagc  
 aagactctgt  
 2821 ctcaaaagaa agaaaaagaa aagaaaaaaa agaaaggcaa gttgactgct  
 gaaaggggaa  
 2881 tctgtgtacg cctgggagct gtggggcagc cacattccag cacatggatc  
 50 tgagaaacag  
 2941 aacgctgatc tgcagaaaga gatgagaacc aaagagaggc cacctgcgtc  
 ctgggtccat  
 3001 tttcatcctc cctgaagccc agctgcccag ggtggggaga aacaccctgt  
 gtccatggga

3061 tagagtcctt tccgcttgca gttgtgcca aagaatctta aatacaaatg  
 agatatecctt  
 3121 aggtagtga tcatttatgt aatatgtgtc ttcactgggg aatactgact  
 tcctaaaatc  
 5 3181 tcaagatgga agatatacca catgtaaatt attttagagc aattaaattg  
 ttttcaggat  
 3241 tttccaaaa

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>plakophilin 3</b>	<b>PKP3</b>	<b>NM_007183</b>	<b>10</b>

10 1 ggcctcgagg gacaggacgt gaagatagtt gggtttgag gggccgcca  
 ggcccaggcc  
 61 cggtaggacct gccgccatgc aggacggtaa cttcctgctg tcggccctgc  
 agcctgaggc  
 121 cggcgtgtgc tcctgggcgc tgccctctga cctgcagctg gaccgccggg  
 15 gcgccgaggg  
 181 gccggaggcc gagcggctgc gggcagcccg cgtccaggag caggtcgcg  
 cccgctctt  
 241 gcagctggga cagcagccgc ggcacaacgg ggccgctgag cccgagcctg  
 aggccgagac  
 20 301 tgccagaggc acatccaggg ggcagtacca caccctgcag gctggcttca  
 gctctcgtc  
 361 tcagggcctg agtggggaca agacctggg cttccggccc atcgccaagc  
 cggcctacag  
 421 cccagcctcc tggctctccc gctccgcccgt ggatctgagc tgcagtcgga  
 25 ggctgagttc  
 481 agcccacaac gggggcagcg cctttggggc cgctgggtac gggggtgcc  
 agcccacccc  
 541 tcccatgccc accaggcccg tgctcttcca tgagcgcggt ggggttggga  
 gccgggcca  
 30 601 ctatgacaca ctctccctgc gctcgtgctg gctggggccc gggggcctgg  
 acgaccgcta  
 661 cagcctggtg tctgagcagc tggagcccgc ggccacctcc acctacaggg  
 cctttgcgta  
 721 cgagcgccag gccagctcca gctccagccg ggcagggggg ctggactggc  
 35 ccgaggccac  
 781 tgaggtttcc ccgagccgga ccatccgtgc ccctgccgtg cggaccctgc  
 agcattcca  
 841 gagcagccac cggagccgcg gggtaggcgg ggcagtgccg ggggcccgtcc  
 tggagccagt  
 40 901 ggctcgagcg ccatctgtgc gcagcctcag cctcagcctg gctgactcgg  
 gccacctgcc  
 961 ggacgtgcat gggttcaaca gctacggtag ccaccgaacc ctgcagagac  
 tcagcagcgg  
 1021 ttttgatgac attgacctgc cctcagcagt caagtacctc atggcttcag  
 45 accccaacct  
 1081 gcagggtgctg ggagcggcct acatccagca caagtgttac agcgatgcag  
 ccgccaagaa  
 1141 gcaggcccgc agccttcagg ccgtgcctag gctggtgaag ctcttcaacc  
 acgccaacca

1201 ggaagtgcag cgccatgccca caggtgccat gcgcaacctc atctacgaca  
 acgctgacaa  
 1261 caagctggcc ctggtggagg agaacgggat cttcgagctg ctgcggacac  
 tgcgggagca  
 5 1321 ggatgatgag cttcgcaaaa atgtcacagg gatcctgtgg aacctttcat  
 ccagcgacca  
 1381 cctgaaggac cgcctggcca gagacacgct ggagcagctc acagacctgg  
 tgttgagccc  
 1441 cctgtcgggg gctgggggtc cccccctcat ccagcagaac gcctcggagg  
 10 cggagatctt  
 1501 ctacaacgcc accggettcc tcaggaacct cagctcagcc tctcaggcca  
 ctcgccagaa  
 1561 gatgcgggag tgccacgggc tgggtggacgc cctggtcacc tctatcaacc  
 acgcctgga  
 15 1621 cgcgggcaaa tgcgaggaca agagcgtgga gaacgcggtg tgcgtcctgc  
 ggaacctgtc  
 1681 ctaccgcctc tacgacgaga tgccgccgtc cgcgctgcag cggctggagg  
 gtcgcggccg  
 1741 cagggacctg gcgggggccc cgcggggaga ggtcgtgggc tgcttcacgc  
 20 cgcagagccg  
 1801 gcggctgcgc gagctgcccc tcgccgccga tgcgctcacc ttcgaggagg  
 tgtccaagga  
 1861 cccaagggc ctcgagtggc tgtggagccc ccagatcgtg gggctgtaca  
 accggtgct  
 25 1921 gcagcgtgc gagctcaacc ggcacacgac ggaggcggcc gccggggcgc  
 tgcagaacat  
 1981 cacggcaggc gaccgcaggt gggcgggggt gctgagccgc ctggccctgg  
 agcaggagcg  
 2041 tattctgaac ccctgctag accgtgtcag gaccgccgac caccaccagc  
 30 tgcgctcact  
 2101 gactggcctc atccgaaacc tgtctcggaa cgctaggaac aaggacgaga  
 tgtccacgaa  
 2161 ggtggtgagc cacctgatcg agaagctgcc gggcagcgtg ggtgagaagt  
 cgccccagc  
 35 2221 cgagggtgctg gtcaacatca tagctgtgct caacaacctg gtggtggcca  
 gccccatcg  
 2281 tgcccagagc ctgctgtatt ttgacggact ccgaaagctc atcttcatca  
 agaagaagcg  
 2341 ggacagcccc gacagtgaga agtcctcccg ggcagcatcc agcctcctgg  
 40 ccaacctgtg  
 2401 gcagtacaac aagctccacc gtgacttccg ggcgaagggc tatcggaagg  
 aggacttct  
 2461 gggcccatag gtgaagcctt ctggaggaga aggtgacgtg gcccagcgtc  
 caagggacag  
 45 2521 actcagctcc aggctgcttg gcagcccagc ctggaggaga aggctaatga  
 cggaggggccc  
 2581 cctcgctggg gccctgtgt gcatctttga gggcctcggg ccaccaggag  
 gggcagggtc  
 2641 ttatagctgg ggacttggt tccgcagggc aggggggtgg gcagggctca  
 50 aggctgctct  
 2701 ggtgtatggg gtggtgaccc agtcacattg gcagaggtgg gggttggctg  
 tggcctggca  
 2761 gtatcttggg atagccagca ctgggaataa agatggccat gaacagtcaa

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 5 (GalNAc-T5)	GALNT5	NM_014568	11

```

1 agtgtttatac agaacttagc cagggccagc caagcaggca cagatgctct
gctatgaaat
5 61 gccacgcagg cagagactga caagcggtag gaactgagct ttccccttgg
actgctgctt
121 cctgctgtgt tcaggggagg gggtcacttt ctggcaactc tgctgctgct
gctgctgctg
181 ctgctacttc agcttcctct ccaactcaagg taagcaggct aaggaggaggc
aggctgctag
10 241 ggaaagcttt gtacatgaa caggatccga aagtttttcc gaggaagtgg
gagagtcttg
301 gcatttatct ttgtagcttc tgtcatctgg ctccctcttg acatggcagc
tctccgctc
361 tcattcagtg agatcaacac tcgggtcadc aaggaagaca ttgtgaggag
15 ggagcggata
421 ggattcagag ttcagccaga ccaaggaaaa attttttaca gcagcataaa
agagatgaaa
481 cctcccctaa ggggacatgg gaaaggggca tggggcaaag agaatgtag
aaaaactgag
20 541 gagagtgtgc tcaaggttga ggtggacttg gaccaaacc agagggaaag
aaaaatgcag
601 aatgccctgg gaaggggcaa ggttgtgccc ttgtggcadc ctgcacatct
gcagaccctc
661 cctgtgactc ctaacaagca gaagacagac gggagaggca ccaaacctga
25 agcctcctct
721 caccagggga caccaaagca aacgacagct cagggggctc caaagacctc
attcatagca
781 gcaaaaggaa ctcaggtagt caaaatatca gtacacatgg gacgtgtcag
tttaaaacag
30 841 gagccccgga agagtcatag tcccagcagt gacacatcaa aactagcagc
tgaaagggac
901 ttgaatgtga ccatcagtct tagtactgat agaccaaagc agcgatcaca
ggcagtagca
961 aacgagaggg cacaccctgc cagcacagca gtgccgaagt ctggggaagc
35 catggcctta
1021 acaaaaacta agactcagag caaagaagtc aatgcaaata aacacaaagc
caatacagag
1081 cttccttttc ctaagttcac tgtcaattca aatcgcttaa ggaagcaatc
tattaatgag
40 1141 acacctttgg gaagtttgtc aaaggatgat ggagctagag gggctcatgg
gaagaaactc
1201 aattttctctg aaagccatct tgtgattata accaaagagg aagagcaaaa
ggcagacccc
1261 aaagaggtct ctaattctaa aacaaaaca atatttcta aagtattggg
45 taaaagccaa
1321 agtaaacaca tttccaggaa tagaagtgag atgtcttctt cttcacttgc
tccacataga
1381 gtgccactgt cccaaactaa ccatgcttta actggagggc tagagccagc
aaaaatcaac

```

1441 ataactgcc aagccccctc tacagaatac aaccagagtc atataaaagc  
 ctttttacct  
 1501 gaagacagtg gaacgcacca ggtgtaaga attgatgtga cactttctcc  
 aagggacccc  
 5 1561 aaagctccag ggcagtttgg gcgtcctgta gttgtcccc atggaaagga  
 gaaggaggca  
 1621 gaaagaagat ggaaagaagg aaacttcaat gtctacctta gcgatttgat  
 cccagtgat  
 1681 agagccattg aagacaccag acctgctgga tgtgcagagc agctagtcca  
 10 caataacctc  
 1741 ccaaccacca gtgtcatcat gtgctttgtg gatgaagtgt ggtccactct  
 cctgagatct  
 1801 gttcacagtg tcatcaatcg ctctcctcca cacctcatca aggagattct  
 gctggtagat  
 15 1861 gacttcagca ccaaagacta tctaaaagat aatttgata aatacatgct  
 ccagtttcca  
 1921 aaagttcggga ttcttcgct caaagagaga catggcttaa taagggccag  
 gctggcagga  
 1981 gcacagaatg caacaggtga tgtgttgaca tttttagatt ctcatgtgga  
 20 atgtaacggt  
 2041 ggttggttg aacctctctt ggaagagatt tattaagta gaaagaaagt  
 ggctgtcca  
 2101 gtaatcgaag tcatcaatga taaggatatg agttacatga cagtggataa  
 ctttcaaaga  
 25 2161 ggcattcttg tgtggccat gaactttggt tggagaacaa tctctccaga  
 tgtcattgca  
 2221 aaaaacagaa ttaaagaaac tgatacaata aggtgcctg tcatggctgg  
 tggattgttt  
 2281 tctattgaca aaagttactt ttttgaactt ggaacatacg acctggcct  
 30 tgatgtttgg  
 2341 ggtggggaaa atatggagct ctcaattcaag gtgtggatgt gtgggtgga  
 aattgagatc  
 2401 attccctgct cccgagtggg ccatatattc agaaatgaca atccatattc  
 cttccccaaa  
 35 2461 gaccggatga agacagtgga gcggaacttg gtgcgggttg ccgaggtctg  
 gctggatgag  
 2521 tataaggagc tgttctatgg ccacggagac cacctcatcg accaagggt  
 agatgttggc  
 2581 aacctcacc agcaaagga gctgcgaaag aaactgaagt gcaaaagttt  
 40 caaatggtac  
 2641 ttggagaatg tctttctga ctttaaggct cccattgtga gagctagtgg  
 tgtgcttatt  
 2701 aatgtggctt tgggtaaag catttccatt gaaaacta cagtcattct  
 ggaagactgc  
 45 2761 gatgggagca aagagcttca acaatttaac tacacctggt taagacttat  
 taaatgtgga  
 2821 gaatggtgta tagccccat cctgataaa ggagccgtaa ggctgcaccc  
 ttgtgataac  
 2881 agaaacaaag ggctaaaatg gctgcataaa tcaacatcag tctttcatcc  
 50 agaactggtg  
 2941 aatcacattg tttttgaaa caatcagcaa ttattatgct tggaaggaaa  
 tttttctcaa  
 3001 aagatcctga aagtagctgc ctgtgacca gtgaagccat atcaaaagt  
 gaaatttgaa

3061 aaatattatg aagcctgaag tgtaactgat gtttttatat agtaaacca  
 ttaaatactg  
 3121 tgaaaataac a

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
calmodulin-like 4	CALML4	NM_033429	12

5

1 ggggctgagg gtggagagag gaaggaagg aagaaaagg gagccttctt  
 ggccagggtg  
 61 accggcacta agaggcctca ctccaagccc ccgaggagcc tgtggtgggg  
 ctggagacct  
 10 121 ggctcaggcc cctccaccac ccttaaagtc ctcagaaggt gggaaactgaa  
 ctggcacagg  
 181 atgggaaccg gctgtgcgct ggccaattga ttttgccagc tgcctgtaa  
 ttcagctggt  
 241 gaggaaactg aggcacagac tgaggtagaa tgattcgcca gtcactcagc  
 aagtcagcag  
 15 301 acggggagga ctgaatccca gcctgagagc accgaagctt gtatccctgc  
 aataccgagc  
 361 cccaagcctg cgagccccgg tgcccatttc tgagttaggc cgtcttgaa  
 gggttccctt  
 20 421 cctcctacaa gatggtgtgt gaggagcctt caatacgacc cggggtgtaa  
 agtgtccaac  
 481 tctagtaggg gcctgatggc atccccgccg agtcccagga gagagagaga  
 agacccttc  
 541 ctggagtcca gggctcccgg gaagaaacac tggcatttgt ccctttgctt  
 cggcttctgg  
 25 601 aggcagagac tctgagccca gggagagcct tctgcagccc catttcctca  
 aaaatccaac  
 661 ctgcccaggt ggcggtcat gagctgtgct caggaagctg gaatctgacc  
 ctggtggcgt  
 30 721 cgggcccagt ctccatggca gccgagcatt tattaccgg gcctccacc  
 agcttggcag  
 781 actttagact tgaggctgga ggaaagggaa ctgaacgcgg ttctgggagc  
 agcaagccca  
 841 cgggtagcag ccgaggcccc agaatggcca agtttcttc ccaagaccaa  
 attaatgagt  
 35 901 acaaggaatg cttctcctg tatgacaagc agcagagggg gaagataaaa  
 gccaccgacc  
 961 tcatggtggc catgaggtgc ctgggggcca gcccgacgc aggggaggtg  
 cagcggcacc  
 40 1021 tgcagacca cgggatagac ggaaatggag agctggattt ctccactttt  
 ctgaccatta  
 1081 tgcacatgca aataaaacaa gaagaccaa agaaagaat tcttctagcc  
 atgttgatg  
 1141 tggacaagga gaagaaaggt tacgtcatgg cgtccgacct gcggtcacaa  
 ctcacgagtc  
 45 1201 tgggggagaa gtcacccac aaggaagtgg atgatctctt cagggagca  
 gatatcgaac  
 1261 ccaatggcaa agtgaagtat gatgaattta tccacaagat cacccttct  
 ggacgggact

1321 attgaaggag gagaatggga gagcctcccc tgggcctgaa aacttgggagc  
 aattaatttt  
 1381 ttttaaaaag tgttcttttc acttgggaga gatggcaaac acagtggcaa  
 gacaacatta  
 5 1441 cccaactata gaagagaggc taactagcaa caataataga tgatttcagc  
 catgggatga  
 1501 gtagatcttt aataaaagat ttgtattgat tttattaact accgtgagtc  
 cggccctttc  
 1561 aagcatggaa ggagcctgcg gtttggagtc tggcctgggt tccagtctg  
 10 gctctgctgc  
 1621 ttcccactgt gactttgggc aaatcatttc actcctcaaa gccccccac  
 acaagctgga  
 1681 ttcccacttc ttacctcatg gagcctggtg aggaaggatt gagctgatga  
 ctttaaggga  
 15 1741 atctaccaag agacttattc tgtatttggg ggctagaacc atcttcata  
 tttccaagat  
 1801 tttccaagat gaagccagtg ctagctgaga agcagcaatg aacagaaagc  
 tgtaacactt  
 1861 atgacaacaa ttcttgcaag gccagaggcc catttacaac ttctcatttc  
 20 catctcaaca  
 1921 gatatagtga catagctcag gctattcatt cataaacaca gagtgtagag  
 tgaaaacact  
 1981 agagtgaaaa cacatgctac aatgaggcag catcagctga gagcaggaag  
 agcgatctac  
 25 2041 tttacacccc acaccaaagg aaaccagatg tgagctgcta aattgactgg  
 ccttgcagag  
 2101 ctcaagaagg gggcttccaa tgctgtgaga attccgagct gttccctggg  
 ctctgttaac  
 2161 aggcagagag gttccgggat ggtctgctca agtggccac actggtcatt  
 30 gccttaagcc  
 2221 acctccccag gacttacgga gagaaataag gggatgtaac cagcaatggc  
 cagggtaaca  
 2281 cagccctgga aaacagtagt aggagcacta ggctttctgg gagtccatcc  
 agctggagtg  
 35 2341 gctttgagtg agttacacag ctagaagggtg ccagggttgg gctgccagag  
 attcagaggt  
 2401 gccatacact tgtcaaatct ggatcattcg tagtgccagc acagtcctaa  
 aagggtgga  
 2461 gtaccacacc aacacaggta ggggtgcagg gcttcaagta caaagatttg  
 40 catccatgta  
 2521 tgtatcaaaa gtgggttctc tgggctgtgg ctttgtctag tagtaccaca  
 gtggctaaag  
 2581 tagaagaaaa ccaaatcaaa tgggatgtgt cttttgggag gatgtacaag  
 acacaaatct  
 45 2641 ttcactaggc accgggcaca gggaaaactg caggaacaa gagttgtagt  
 gttagtcaa  
 2701 ctgtctcaac gatgctgtgt ggcttcagac ccaacaagg ccctgaggaa  
 ggagactctc  
 2761 atttcccaa gcataactgc aaggagagga ggaattccta ggagccaaag  
 50 agttttgtgg  
 2821 ggtgagggta aataaatggc ccaaatgcca actaggtgaa gttgtgacca  
 tctggctggg  
 2881 aagcccaggt ccacacagtg taggagcaga tgttttgtgg ggtctgaggt  
 ttacgagatt

2941 tggctgcctt aagaatacaa aaacagaaat gcagaatttc tggggctgct  
 cctaggacca  
 3001 gaacaagtga agggctctgg tgcttaaact tcattacctt catggtaaat  
 ccaccagagg  
 5 3061 gccggttaga tgctggcccc gccgagagaa ctgctgtcac tttcaggcaa  
 agctcaaagg  
 3121 tectaggccc acagttcttt tgagctccag tcatggacat taggaagtaa  
 atcctgcaca  
 3181 gccaacctgg aataccaaag attagatggg agatagatac caatgattta  
 10 gatggcacag  
 3241 gaagagcaag ttctggatat aataaatgag ggtactttcc gtcaaagctt  
 ttctatgtct  
 3301 atatttatca ctgaatagtc ccagtatggg tttaaagcaa gttttatgaa  
 tctcatttgc  
 15 3361 ctaacaggaa tctgaaatat aacttgocaa aaacacacag ttgggtgtgga  
 atggtcatta  
 3421 gaacctgggg ctctcttca cggactccct gctcattaag ggattcagtg  
 gtccagagtc  
 3481 taagatccta ttaagtgttt gattcaaacc tctacccgag gaagggtgt  
 20 taccttactc  
 3541 ctggctctgg tttcaagctc attcctgaaa ttccagctgg tttctctage  
 acctagtgtt  
 3601 gtttacaaga aggccacggg gctcttagca ttcaaactgc agatactaaa  
 cagatgctgt  
 25 3661 gatttattaa agagttagcc atatttcaac aagaaagga aatgatggct  
 atattcatta  
 3721 cttacctcaa agcatgctgc aagaaaatta gttagttact tgtcatgctt  
 tgaatctct  
 3781 ggatgaaagg tgctttggaa gcacaaacca ttatcacttg tctcataggg  
 30 attgtcccct  
 3841 tgaacatcca gcagtgttat tttacagaag acaaattaac tgaaggcttt  
 tcttttatta  
 3901 catctaaaga gctctacata aacaggtaac attcaatagg taaacaattt  
 tttccaatg  
 35 3961 catgtaataa atattttcac ttgggtacttt tatacaaaact gacattgtct  
 actatacatt  
 4021 tttaaaagcc attttactgg tttggcatgc ggtatggaaa ttctaagaga  
 gaaagtttta  
 4081 aggcaatgaa tcacagattt aagttcatgg aatttatggg aactttatct  
 40 gtttatgtac  
 4141 attttcccct ttgttaaaca attaacagca gcacactctg ggaccaccag  
 ctattttccc  
 4201 tctctttctg aatctaage tttgtattta attaaaaaac agaattcaac  
 atctattgat  
 45 4261 aaaacaaaat tcttactaaa ataatttcaa atgtgcttta aaaagtctctg  
 aagatcttga  
 4321 aagttttatg tgtttaaaat tgaattgtc taaaaaatg ctctttccac  
 attaatttag  
 4381 ttaggatata ttttactcc atttcagaca cttgactcaa aggaaaatct  
 50 gccaaagaat  
 4441 ccgatttttc agagcttacg tgaatctttc ctcagtaaag atacagaatt  
 gtgatcatgt  
 4501 ctaaataatt agtaaagcaa ttttaatgct caaaatagtc aaccaagtat  
 ggcatggttc

4561 tggttcagat tttttttttt taagatgtat ccaataacac tcacgaagta  
 attaaaagcc  
 4621 actttaaccc tgctaaaaaa

5

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 12 (GalNAc-T12)	GALNT12	AK024865	13

1 cattttataa tgaagcctgg tcaactctcc ttcggacagt ttacagtgtc  
 cttgagacat  
 61 ccccggatat cctgctagaa gaagtgatcc ttgtagatga ctacagtgat  
 10 agagagcacc  
 121 tgaaggageg cttggccaat gagctttcgg gactgcccac ggtgcgctg  
 atccgcgcca  
 181 acaagagaga gggcctggtg cgagcccggc tgctgggggc gtctgcggcg  
 aggggcgatg  
 15 241 ttctgacctt cctggactgt cactgtgagt gccacgaag gtggctggag  
 ccgctgctgc  
 301 agaggatcca tgaagaggag tcggcagtgg tgtgcccggg gattgatgtg  
 atcgactgga  
 361 acaccttoga atacctgggg aactccgggg agccccagat cggcggtttc  
 20 gactggaggc  
 421 tgggtttcac gtggcacaca gttcctgaga gggagaggat acggatgcaa  
 tccccgctg  
 481 atgtcatcag gtctccaaca atggctggtg ggctgtttgc tgtgagtaag  
 aaatattttg  
 25 541 aatatctggg gtcttatgat acaggaatgg aagtttgggg aggagaaaac  
 ctcgaatttt  
 601 cctttaggat ctggcagtgt ggtggggttc tggaaacaca cccatgttcc  
 catgttgcc  
 661 atgttttccc caagcaagct ccctactccc gcaacaaggc tctggccaac  
 30 agtgttcgtg  
 721 cagctgaagt atggatggat gaatttaaag agcttacta ccatcgcaac  
 ccccgtgccc  
 781 gcttggaaacc ttttggggat gtgacagaga ggaagcagct ccgggacaag  
 ctccagtgta  
 35 841 aagacttcaa gtggttcttg gagactgtgt atccagaact gcatgtgcct  
 gaggacaggc  
 901 ctggcttctt cgggatgctc cagaacaaag gactaacaga ctactgcttt  
 gactataacc  
 961 ctcccgatga aaaccagatt gtgggacacc aggtcattct gtacctctgt  
 40 catgggatgg  
 1021 gccagaatca gtttttcgag tacacgtccc agaaagaaat acgctataac  
 accaccagc  
 1081 ctgagggctg cattgctgtg gaagcaggaa tggataccct tatcatgcat  
 ctctgcaag  
 45 1141 aaactgcccc agagaatcag aagttcatct tgcaggagga tggatcttta  
 tttcacgaac

1201 agtccaagaa atgtgtccag gctgcgagga aggagtcgag tgacagtttc  
 gttccactct  
 1261 tacgagactg caccaactcg gatcatcaga aatggttcctt caaagagcgc  
 atgttatgaa  
 5 1321 gcctcgtgta tcaaggagcc catcgaagga gactgtggag ccaggactct  
 gcccaacaaa  
 1381 gacttagcta agcagtgacc agaaccacc aaaaactagg ctgcattgct  
 ttgaagaggc  
 1441 aatcattttg ccatttgtga aagttgtggt ggatttagta aaaatgtgaa  
 10 taagctttgt  
 1501 acttattttg agaacttttt aatggttcca aaatacccta ttttcaaagg  
 gtaatcgtaa  
 1561 gatgtaacc cttggtattt agaaaattaa aacctataa tatttttcta  
 tcaagatgta  
 15 1621 tattttacag tcgtgccttt tactctcatt agcaaaaaag ataaagattt  
 tattttggta  
 1681 ttacaagaa ttcccaggta cgaagatata tgcattgggtg gaaatcaggt  
 tcaagcaacg  
 1741 tactttgcat taactgataa tacctcagct gcgggggtaa agttttccca  
 20 gtatagagag  
 1801 actgtcacta ggaacattgt attgatttat tcaggtcatt gagatcttct  
 agatgtattt  
 1861 taaaaagaat gctttttggt tatgtgttgc taccacagtt aacctccat  
 aatgttcatg  
 25 1921 tcagccaaag aggactaacc aaagctgaaa tctcagagaa caatttgctt  
 tactaagctg  
 1981 agtcaacttg agagcgaact tctaacaatg ccgcaactgta gtgtggctgg  
 ttctaccact  
 2041 atgactttaa aacatgttta tatcattttt aatttttatg atacggtagt  
 30 gtcagggaga  
 2101 aatgtaatgt tctatatgaa attccttttt caagtttggt cattaataac  
 agttattaat  
 2161 taaatcagc gttagagttt gtgctgctgc aactgcttg aaattttctc  
 tgagtaattc  
 35 2221 tgatttgtga atgatcccag accaaccctg agattttgtc aacctgatta  
 agtcaatatg  
 2281 aatgattaaa aagatgtgag

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>thiamin pyrophosphokinase 1</b>	TPK1	NM_022445	14

40 1 aaggctcctc agccgagcgc cgagcggctg atcgccgtag ctcccgcagc  
 ctgcatctc  
 61 cagtctgtgg ctctaccag ccattgtagg ccaataatcc gttatggagc  
 atgcctttac  
 121 ccggttgag ccctgcttt ccaactgggaa tttgaagtac tgcttgtaa  
 45 ttcttaataca  
 181 gcctttggac aactattttc gtcacttttg gaacaaagct cttttaagag  
 cctgtgccga  
 241 tggagggtgc aaccgcttat atgatatac cgaaggagag agagaaagct  
 ttttgctga

301 attcatcaat ggagactttg attctattag gcctgaagtc agagaatact  
 atgctactaa  
 361 gggatgtgag ctcatctcaa ctctgatca agaccacact gactttacta  
 agtgccttaa  
 5 421 aatgctccaa aagaagatag aagaaaaaga cttaaagggt gatgtgatcg  
 tgacactggg  
 481 aggccttgct gggcgttttg accagattat ggcactctgtg aataccttgt  
 tccaagcgac  
 541 tcacatcact ccttttccaa ttataataat ccaagaggaa tcgctgatct  
 10 acctgctcca  
 601 accaggaaag cacaggttgc atgtagacac tggaaatggag ggtgattggt  
 gtggccttat  
 661 tctgtttgga cagccttgta tgcaggttac aaccacaggc ctcaagtgga  
 acctcacaaa  
 15 721 tgatgtgctt gcttttgga cattggtcag tacttccaat acctacgacg  
 ggtctgggtg  
 781 tgtgactgtg gaaactgacc acccactcct ctggaccatg gccatcaaaa  
 gctaacctgt  
 841 tgactggcat ccataagtgt gcctctgctt tatctcattt ctcaacagtt  
 20 cattgctcaa  
 901 caagaacgat tcacctgggt ttgcaagaat ctaaacctct ctaggggaag  
 cccactgggt  
 961 ttaaagatgt tagtgttttag ataatacagg taacattata aatgacagat  
 ctcaatttta  
 25 1021 tagtagtggg aaagatacat gctaagaaag caaataagct ctattatatt  
 cggttggaac  
 1081 ctaatgggaa tcattccact atacaattca gtactgatta ttcttcttac  
 attattaatc  
 1141 attccattta tcttagaaaa ttgttttta tttgaatcag agaaaactgt  
 30 tgaggttct  
 1201 cttggagtct agaacatect taaatgtcta acaacaaggg ctacctctga  
 gtacctttta  
 1261 gtattagttt tctgtatatg atatatatta tcttatactg aaaaaaatt  
 cctttcagat  
 35 1321 tggggtgtta gaagtgcacc aggtcactct gacctatta ctgtctttgg  
 tattgtctta  
 1381 aataaatcaa gaatcattga cctaattggt aaatttaaaa ataggtagtt  
 agcaataggt  
 1441 ggaaagagaa atgatgtgaa agataaatga tgattcgtgg agccctactc  
 40 acacattaac  
 1501 ccccaaattc aaaagtaaga atgcaaaaagt ctagaggggg taacagtctg  
 catcatcatc  
 1561 acaacctaaa tggagaaagc tgtgcagagg aaacttaagc ataaaaattg  
 aattcgtttc  
 45 1621 tgacatacct tagactgaaa aactgttgg tcatccagaa gtgtattcat  
 attaccagaa  
 1681 aatgagtttg tctatgggga tacatgaact tcatatacta aggagcctaa  
 ctccaaagcc  
 1741 tgcgtttctca tcccagtctg atattcacct aagtttcegg acccttttcc  
 50 ttagctgtaa  
 1801 aatggaagcg gttggactga tgggtgtctga ggttctttcc cacactgaaa  
 ttctaaatat  
 1861 tgacacttag cagtcatagg gctgataata cacacagtta ctgacttagc  
 ctaaacaacc

1921 tgggtgcatcg aaatgtattc acctttcttt tgtaaagaga ccatcttcta  
 tcttctttcc  
 1981 acctttctct gttttatgaa accaactggt gacatacaaa ccatgattga  
 aggagaacct  
 5 2041 gtccaacatg ttttatgtac acaaatccct atgttgctat aagaaaagtg  
 aaagtaactg  
 2101 ttttcttctt ggtgctatga cagtgtgaga ctcaggttgt ctgtagagaa  
 tgaaaggagc  
 2161 agtggcccgc gtgattgtgg catttaagga gcagtggccc atgtgactgt  
 10 ggcattttcg  
 2221 gcacttttca ttactttctg cttgaccgga agttgaggct tagctatgtt  
 tccatcttca  
 2281 gtttctgaag actagttata tattccttac tagaaatata ttcataatat  
 ataaaagaaa  
 15 2341 tatatctgtg attttaaaat tttgctacca aagaatgcat gttctgtgtg  
 ccctgaaaat  
 2401 gttaccagtg ttaataaatg gatacttattc aaaaaagaaa

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>defensin, alpha 6, Paneth cell-specific</b>	DEFA6	NM_001926	15

20

1 acacatctgc tctgctctc tctcctccag cgaccctagc catgagaacc  
 ctcaccatcc  
 61 tcaactgctgt tctcctcgtg gcctccagg ccaaggctga gccactccaa  
 gctgaggatg  
 25 121 atccactgca ggcaaaagct tatgaggctg atgccagga gcagcgtggg  
 gcaaatgacc  
 181 aggactttgc cgtctccttt gcagaggatg caagctcaag tcttagagct  
 ttgggctcaa  
 241 caagggcttt cacttgccat tgcagaaggt cctgttattc aacagaatat  
 30 tcctatggga  
 301 cctgcaactgt catgggtatt aaccacagat tctgctgcct ctgagggatg  
 agaacagaga  
 361 gaaatatatt cataatttac tttatgacct agaaggaaac tgcgtgtgtg  
 cccatacatt  
 35 421 gccatcaact ttgtttctc atctcaaata aagtcctttc agcaaaaaaa aaaaa

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
epithelial protein lost in neoplasm beta	EPLIN	NM_016357	16

```

1 ggcctaggta gagcgccggg acctgtgaca gggctggtag cagcgcagag
gaaagggcggc
5 61 ttttagccag gtatttcagt gtctgtagac aagatggaat catctccatt
taatagacgg
121 caatggacct cactatcatt gagggtaaca gccaaagaac tttctcttgt
caacaagaac
181 aagtcacatcg ctattgtgga aatattctcc aagtaccaga aagcagctga
10 agaaacaaac
241 atggagaaga agagaagtaa caccgaaaat ctctcccagc actttagaaa
ggggaccctg
301 actgtgttaa agaagaagtg ggagaacca gggctgggag cagagtctca
cacagactct
15 361 ctacggaaca gcagcactga gattaggcac agagcagacc atcctcctgc
tgaagtgaca
421 agccacgctg cttctggagc caaagctgac caagaagaac aaatccacc
cagatctaga
481 ctcaggtcac ctctgaagc cctcgttcag ggtcgatc cccacatcaa
20 ggacggtgag
541 gatcttaaag accactcaac agaaagtaaa aaaatggaaa attgtctagg
agaatccagg
601 catgaagtag aaaaatcaga aatcagtgaa aacacagatg cttcgggcaa
aatagagaaa
25 661 tataatgttc cgctgaacag gcttaagatg atgtttgaga aaggtgaacc
aactcaaact
721 aagattctcc gggcccaaag ccgaagtgca agtggaagga agatctctga
aaacagctat
781 tctctagatg acctggaaat aggcccaggc cagttgtcat cttctacatt
30 tgactcggag
841 aaaaatgaga gtagacgaaa tctggaactt ccacgcctct cagaaacctc
tataaggat
901 cgaatggcca agtaccagc agctgtgtcc aaacaaagca gctcaaccaa
ctatacaaat
35 961 gagctgaaag ccagtgggtg cgaaatcaaa attcataaaa tggagcaaaa
ggagaatgtg
1021 ccccaggtc ctgaggtctg catcacccat caggaagggg aaaagatttc
tgcaaatgag
1081 aatagcctgg cagtccgttc caccctgcc gaagatgact cccgtgactc
40 ccaggttaag
1141 agtgaggttc aacagcctgt ccatcccaag ccactaagtc cagattccag
agcctccagt
1201 ctttctgaaa gttctcctcc caaagcaatg aagaagtttc aggcacctgc
aagagagacc
45 1261 tgcgtggaat gtcagaagac agtctatcca atggagcgtc tcttgccaa
ccagcaggtg
1321 tttcacatca gctgcttccg ttgctcctat tgcaacaaca aactcagtct
aggaacatat
1381 gcattcttac atggaagaat ctattgtaag cctcacttca atcaactctt
50 taaatctaag

```

1441 ggcaactatg atgaaggctt tgggcacaga ccacacaagg atctatgggc  
 aagcaaaaat  
 1501 gaaaacgaag agattttggg gagaccagcc cagcttgcaa atgcaagggg  
 gaccctcac  
 5 1561 agcccagggg tagaagatgc ccctattgct aagggtgggtg tcttggtgctgc  
 aagtatggaa  
 1621 gccaaaggcct cctctcagca ggagaaggaa gacaagccag ctgaaaccaa  
 gaagctgagg  
 1681 atcgcttggc cccccccac tgaacttggg agttcaggaa gtgccttggg  
 10 ggaagggatc  
 1741 aaaatgtcaa agcccaaagc gcctcctgaa gacgaaatca gcaagcccga  
 agttcctgag  
 1801 gatgtcgatc tagatctgaa gaagctaaga cgatcttctt cactgaagga  
 aagaagccgc  
 15 1861 ccattcactg tagcagcttc atttcaaagc acctctgtca agagcccaaa  
 aactgtgtcc  
 1921 ccacctatca ggaaaggctg gagcatgtca gagcagagtg aagagtctgt  
 ggggtggaaga  
 1981 gttgcagaaa ggaaacaagt ggaaaatgcc aaggcttcta agaagaatgg  
 20 gaatgtggga  
 2041 aaaacaacct ggcaaaaaca agaatctaaa ggagagacag ggaagagaag  
 taaggaaggt  
 2101 catagtttgg agatggagaa tgagaatctt gtagaaaatg gtgcgactc  
 cgatgaagat  
 25 2161 gataacagct tctcaaaca acaatctcca caagaacca agtctctgaa  
 ttggtcgagt  
 2221 tttgtagaca acacctttgc tgaagaattc actactcaga atcagaaatc  
 ccaggatgtg  
 2281 gaactctggg agggagaagt ggtcaaagag ctctctgtgg aagaacagat  
 30 aaagagaaat  
 2341 cggattatg atgaggatga ggatgaagag tgacaaattg caatgatgct  
 gggccttaaa  
 2401 ttcattgtag tgtagcgag ccaactgcct ttgtcaaat gtgatgcaca  
 taagcaggta  
 35 2461 tcccagcatg aatgtaatt tacttggag taactttgga aaagaattcc  
 ttcttaaaat  
 2521 caaaaacaaa acaaaaaaac acaaaaaaca cattctaat actagagata  
 actttactta  
 2581 aattcttcat ttagcagtg atgatatgag taagtgtgtg aaggcttgta  
 40 actggggaaa  
 2641 tattccacct gataatagcc cagattctac tgtattccca aaaggcaata  
 ttaaggtaga  
 2701 tagatgatta gtagtatatt gttacacact attttggat tagagaacat  
 acagaaggaa  
 45 2761 tttaggggct taaacattac gactgaatgc acttttagtat aaagggcaca  
 gtttgtatat  
 2821 ttttaaatga ataccaattt aattttttag tatttacctg ttaagagatt  
 atttagtctt  
 2881 taaatttttt aggttaattt tcttgctgtg atatatatga ggaatttact  
 50 actttatgtc  
 2941 ctgctctcta aactacatcc tgaactcgac gtcttgaggt ataatacaac  
 agagcacttt  
 3001 ttgaggcaat tgaaaaacca acctacactc ttcgggtgctt agagagatct  
 gctgtctccc

3061 aaataagctt ttgtatctgc cagtgaattt actgtactcc aaatgattgc  
 tttcttttct  
 3121 ggtgatatct gtgcttctca taattactga aagctgcaat attttagtaa  
 taccttcggg  
 5 3181 atcaactgtcc cccatcttcc gtgtagagc aaagtgaaga gtttaaagga  
 ggaagaagaa  
 3241 agaactgtct tacaccactt gagctcagac ctctaaacc tgtatttccc  
 ttatgatgtc  
 3301 ccctttttga gacactaatt tttaaatact tactagctct gaaatatatt  
 10 gatttttattc  
 3361 acagtattct cagggtgaaa ttaaaccaac tataggcctt tttcttggga  
 tgattttcta  
 3421 gtcttaaggt ttggggacat tataaacttg agtacatttg ttgtacacag  
 ttgatattcc  
 15 3481 aaattgtatg gatgggaggg agaggtgtct taagctgtag gcttttcttt  
 gtactgcatt  
 3541 tatagagatt tagctttaat attttttaga gatgtaaac attctgcttt  
 cttagtctta  
 3601 cctagtctga aacattttta ttcaataaag attttaatta aaatttgaaa  
 20

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
chloride-intracellular channel 5	CLIC5	NM_016929	17

1 gacagtcgcg gatcctgtga cacctccggg cagcccggca cttgttgctc  
 ccacgacctg  
 25 61 ttgtcattcc ctaaacccgg ctttccccgt ggccccccgc ctctcccgg  
 cttcgtctct  
 121 tttcatgtga gcatctggga cactgatctc tcagaccccg ctgctcgggc  
 tggagaatag  
 181 atggttttgt gaaaaattaa acaccgccct gaagaggagc cccgctgggc  
 30 agcggcagga  
 241 ggcagagtg ctggcccagg tgctgcagag gtggcgctc cccggcccgg  
 gacggtagcc  
 301 ccgggcgcca acggcatgac agactcggcg acagctaacg gggacgacag  
 ggaccccag  
 35 361 atcgagctct ttgtgaaggc tggaatcgat ggagaaagca tcggcaactg  
 tcctttctct  
 421 cagcgcctct tcatgatcct ctggctgaaa ggagtcgtgt tcaatgtcac  
 cactgtggat  
 481 ctgaaaagaa agccagctga cctgcacaac ctagcccccg gcacgcacc  
 40 gcccttctg  
 541 accttcaacg gggacgtgaa gacagacgtc aataagatcg aggagtctct  
 ggaggagacc  
 601 ttgaccttg aaaagtacc caaactggct gcaaacacc gggaatccaa  
 cacagcgggc  
 45 661 atcgacatct tttccaagtt ttctgcctac atcaaaaata ccaagcagca  
 gaacaatgct  
 721 gctcttgaag gaggcctaac caaggctcta aagaaattgg atgactacct  
 gaacaccct

781 ctaccagagg agattgacgc caacacttgt ggggaagaca aggggtcccg  
 gcgcaagttc  
 841 ctggatgggg atgagctgac cctggctgac tgcaatctgt tgcccaagct  
 ccatgtggtc  
 5 901 aagattgtgg ccaagaaata cgcgaactat gatatcccgg ctgagatgac  
 aggcctgtgg  
 961 cggtaacctca agaacgccta tgcccgtgat gagttcacca acacctgtgc  
 agctgacagt  
 1021 gagatcgagt tggcctacgc tgatgtcgcc aaacgcctca gccgatcctg  
 10 agcacagcca  
 1081 ttttgcccca tccccgtgc agaaggactc aacctctcc ctaagactcc  
 agcttcatag  
 1141 actcctctgt atcactgctt tgaggcgcac tttttataat caagcctcat  
 cttgctggta  
 15 1201 tcatgggaac tccagcctgc tatctttcat gaaggtcagc accatcctg  
 gcctcctcac  
 1261 ataggaatct agcagaaatg atagacacag tccacctttc ggccggccag  
 cctgatctgg  
 1321 gctcagcatg tttggggca gtcagtgttg gagagcccac atatgggatt  
 20 gccactagct  
 1381 tcttctgcca atatcaaaat accttctcag atgctttaga aacatgcaac  
 accaactcct  
 1441 tttctaccct cctctcctgc catacctaca aggccaagga caaacgccat  
 cttcatcctt  
 25 1501 cttagaaaga gatctattac cccattaggg gagacagaga gagtgaatgg  
 aggagtaccg  
 1561 agctggctat ggacttgggt gtctggcaaa cacagcttca gtctcactac  
 ttctgacact  
 1621 ctggttattg ggcactaagg gccagactgg aaagtcactt gagacacatt  
 30 ctcaagttgt  
 1681 tgcagtgcc ggaatgctgc gctgctgctg ctgctcacct ggcccatgct  
 gtccctggct  
 1741 tccatgccgt ccaggccctg ccagaaaagg aaattggcat gcaattctaa  
 actgcagtga  
 35 1801 ctgggatggg aggggagggg agcagtgttg atgccaaaat acccacgggg  
 tctaccagcc  
 1861 atggggtttg cttgcttagg agtagttgtt tcagaggtga ttacaggcct  
 gggtttgact  
 1921 gtgcttacca atgagtgggt tttgagctat gagaaagtgg atgggagtgg  
 40 gaggaggaga  
 1981 gatgggtgaa gacaaaagag ttctttatga gcctcgatgt tccctggtaa  
 acttttaaaa  
 2041 aggccctctc tcatgatcta agtcttggac tgggtggcatc atgtaactgc  
 taaccttaca  
 45 2101 gtaaaaacc aagaatgggt caaaaatgct ttcccagttt ctccaagctg  
 cttctggaat  
 2161 gcaggtctgt cggctgggtg ctctccagca gctgctcctg cctgattcaa  
 ctgtagcctg  
 2221 taatgggtaa aagccacatt taggaggtgg tctgatcata gaacacctta  
 50 ggaagaaagt  
 2281 ccatgagact ttctgactag gaaacctatg ggtttgaact tgaagaaaaa  
 tntagacca  
 2341 tctgggttaa ttttctaca atctgactca actgccaggt gaaaaaaaaa  
 aggaaaaatt

2401 ttttaagctaa tatttcactc ttttgtcatt ctccttaagt ttcattctct  
 aaaaagctta  
 2461 cccagcctga gcttggggac ctgtgcagag gaaactaaga aaaatgcact  
 catcaactcc  
 5 2521 ttctcccagt gaacgcccgg tgagaaaatc catttgccac aggcccttac  
 cttcaacaat  
 2581 cccccttcta tagtgttcgc tggtaaaggg tgaggctccc aagtgtgga  
 aagcccctgg  
 2641 acttggtctca tttctcagca agggcaggat agcacgggtc ctttccatag  
 10 aatatcaac  
 2701 aaattctaac ccaagcaatc cctggaccta cctgcctcca gggatctctg  
 aagaaaaaaa  
 2761 gtaaccctt gatcaaatca gaggagagga agcaggaggc ctcttagagc  
 ccattgagga  
 15 2821 agaggaactt tctcagtagg acactttata agcctgagaa agctttgaaa  
 aggcggaatg  
 2881 agttgattca tttccacctc aaaaggaacc tttccaggtc ccctggaaa  
 ttgtgcctcg  
 2941 gagatgttta acaaggagaa ctggtgagga aagagtcctt ttttactgta  
 20 gggaaaagcc  
 3001 ccaaactggc ctctggggg atgagggctg aaatgatccc gaaggccttt  
 taattagtgt  
 3061 gaaatcctgc tgtactcaga aatccttccc cgaatttaca gcacaggcag  
 gatgacctaa  
 25 3121 gaggcagttt acttcctga gaccacagc tgggctgttc tggaaacaca  
 tctgtgaatc  
 3181 atagccaatt gccacagaga aacagaacc aagcctcgg tgaggccact  
 ccacccaga  
 3241 gaagtctgca gaattccaag gactcggatt ggatgttcag aattcagcaa  
 30 ctggaaagtc  
 3301 cttaaaaaca aacaggccaa accaaatcaa tattgctgtt tctagatgtc  
 ccttctgtgg  
 3361 ttgagctagt tttacagaga taaatatatt aagacaagga ggtgggggtg  
 ttatatgatc  
 35 3421 aatgatagcc atttgaaaga gagggaggag tacagaagga aggcacttct  
 ggtacttaa  
 3481 ttcagaaatt tctttatatt tcagcactgg attatcatat aatgcaagtg  
 actatggact  
 3541 aagagttagt tatggtgtct tatgactaga tttattatgg tatattaag  
 40 taacaataat  
 3601 attaataatta ccttcctttt tttttttgtt tcaaaagaga tctttctcca  
 gatgcttcag  
 3661 cctgtctggc cttcttatca tatgtgcagc acatcatgtc tcagcaacag  
 tgtggtgagg  
 45 3721 tccttaggtg tcccagaac aactcagga gcacgggagg gtctgcagtt  
 gggacccac  
 3781 aactatacag ctatagggtg ggaggcttc ttttcattgg tcctgaatga  
 atacaaatcg  
 3841 ctcagaaagc attttgggtg cacagaaagg ggatgtattt gtgttgagat  
 50 cttattttat  
 3901 tttgtattta tttatcttct ttgacttgca cagcactatt ggggggtggg  
 gaagcagggt  
 3961 agtgggagac gaaggcagaa gcaagagtca aactcagaat gactgagttg  
 aattcactgt

4021 ctagtcagca atgcctgctt ctgagtttgg cccagagaga aggtattgag  
 taagatttta  
 4081 ataactgtaa aaagtaagct ggataagtaa aatcatgatg gatccaaagc  
 acagtttctt  
 5 4141 catctcctga taaagaaagt caaatgcttg ataaattcag agtcacagat  
 gtgagcatag  
 4201 ctatattctt ttaaacgaga ggtagagtga cctagcacta agcaaagtgag  
 ctgaaatgct  
 4261 ggaaacagag tccatcagct tatttggcca cacgatccca aactagtttt  
 10 atcttgggaa  
 4321 atggccctgt cctcagcatt cccttcttgt gctggtgggg ccagtgaagt  
 cttgatctta  
 4381 tcagaaaaag gccacaccaa gtgcgagttt tcccaggctg actttccagg  
 cccttatcaa  
 15 4441 atgaaacaac agaagctctt cacagttctg tgcccatgg ccaactccaca  
 gacagacaat  
 4501 accaagcatc ttagaactgt cataagatag gtcatgcctg aaatagatct  
 tgaccatag  
 4561 agagtcccag aaatcagcaa ggccctggaca aatagaacta agagagagggc  
 20 agaggcagga  
 4621 agctgcgggt ctatcttgta aagagtttag catcactgtg agagtgtgtg  
 tctaaaatta  
 4681 aattaaacta gaagcagcag gtgagtatct ggtaagtact tctgtgactc  
 gcctcaattc  
 25 4741 ccaactggcca ggggccatct caactgcacg gtgaatcaag atgctggtgt  
 catcctcctt  
 4801 ggaaaaagga aatgttaact catggttaaa actaagtaca atgattccca  
 agggatcact  
 4861 ttcttatttt tttaaagac attaaggaga atcttaagaa agcatcagag  
 30 aaagacatgt  
 4921 gcatgtgaag caccctgatt ctgatgtag gaaaacttaa gcgaacagga  
 cctgctgcac  
 4981 acagcccat tgtcttctat ccatttctct ttatcattca aatcaagcaa  
 catgtgccct  
 35 5041 cctcatcaac acacattctt cccctttgct agtatgcac tcccagctta  
 gtgtcaggat  
 5101 actttcgatt cataattatg tatgatccaa agtgtgcata atttcattta  
 acgttaaaga  
 5161 aatagatcca attcctttct tgcaacccaaa aataaataaa atacgttgcc  
 40 tcaatataag  
 5221 gtttgggcta ttctgtgttt ctatagaagc aatctgtttt tggtaaaatg  
 tacttttaag  
 5281 gatccagtca tctgaagtat tttatgtaga gttagagatt tcacaatatt  
 gactatacat  
 45 5341 atatttaaaa tataaattat ccagctgatg tttgaatttg tcttactttc  
 ctggccacct  
 5401 cgttgtccta tttataagc tggggagtta actagcttaa caaaagatgc  
 ttagcttttg  
 5461 taaaagaaca agtgtttcat tttacaaaga cactccaaat gatagttact  
 50 tgattttctc  
 5521 gagaccttta actatggtga tgaataacag gacttgcttt caagccttaa  
 taaatgtaaa  
 5581 atgcctttta atgaagatac agctgagtgt tttcctcatg aatctgaacc  
 aattaccaat

5641 ttgtgttcca gtcttgattg gtattgactg attcaaataa agttggttta  
 ttttcaaata  
 5701 tta

5

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
PERP, TP53 apoptosis effector	PERP	NM_022121	18

1 gcttttgtgg cggcgcccgc gctcgcaggc cactctctgc tgtcgccegt  
 cccgcgcgct  
 61 cctccgaccc gctccgctcc gctccgctcg gccccgcgcc gcccgtaac  
 10 atgatccgct  
 121 ggggctggc ctgcgagcgc tgcgctgga tcctgccct gctcctaact  
 agcgccatcg  
 181 ccttcgacat catcgcgctg gccggccgcg gctggttgca gtctagcgac  
 cacggccaga  
 15 241 cgtcctcgtc gtgggtgaaa tgcctccaag agggcggcgg cagcgggtcc  
 tacgaggagg  
 301 gctgtcagag cctcatggag tacgcgtggg gtagagcagc ggctgccatg  
 ctcttctgtg  
 361 gcttcatcat cctgggtgatc tgtttcatcc tctccttctt cgccctctgt  
 20 ggaccccaga  
 421 tgcttgtctt cctgagagtg attggaggtc tccttgcctt ggctgctgtg  
 ttccagatca  
 481 tctccctggc aatttaccct gtgaagtaca cccagacctt cacccttcat  
 gccaacctg  
 25 541 ctgtcactta catctataac tgggcctacg gctttgggtg ggcagccaag  
 attatcctga  
 601 ttggctgtgc cttcttcttc tgcctcctcc ccaactacga agatgacctt  
 ctgggcaatg  
 661 ccaagcccag gtacttctac acatctgcct aacttgggaa tgaatgtggg  
 30 agaaaatcgc  
 721 tgctgctgag atggactcca gaagaagaaa ctgtttctcc aggcgacttt  
 gaaccattt  
 781 tttggcagtg ttcattat taaactagtc aaaaatgcta aaataatttg  
 ggagaaaata  
 35 841 ttttttaagt agtgttatag tttcatgttt atcttttatt atgttttgtg  
 aagttgtgtc  
 901 ttttactaa ttacctatac tatgccaata tttcctatct atccataaca  
 tttatactac  
 961 atttgtaaga gaatatgcac gtgaaactta acactttata aggtaaaaat  
 40 gaggtttcca  
 1021 agatttaata atctgatcaa gttcttgta tttccaaata gaatggactc  
 ggtctgttaa  
 1081 gggctaagga gaagaggaag ataaggtaa aagttgttaa tgaccaaaca  
 ttctaaaaga  
 45 1141 aatgcaaaaa aaaagtttat tttcaagcct tcgaactatt taaggaaagc  
 aaaatcattt  
 1201 cctaaatgca tatcatttgt gagaatttct cattaatata ctgaatcatt  
 catttttagct  
 1261 aaggcttcat gttgactcga tatgtcatct aggaaagtac tatttcatgg  
 50 tccaaacctg

1321 ttgccatagt tggtaaggct ttcctttaag tgtgaaatat ttagatgaaa  
 ttttctcttt  
 1381 taaagttctt tataggggta ggggtgtggga aaatgctata ttaataaatc  
 tgtagtgttt  
 5 1441 tgtgtttata tgttcagaac cagagtagac tggattgaaa gatggactgg  
 gtctaattta  
 1501 tcatgactga tagatctggt aagttgtgta gtaaagcatt aggaggggtca  
 ttcttgtcac  
 1561 aaaagtgcc aaaaacagc ctcaggagaa taaatgactt gcttttctaa  
 10 atctcagggt  
 1621 tatctgggct ctatcatata gacaggcttc tgatagtttg caactgtaag  
 cagaaacctt  
 1681 catatagtta aaatcctggt ctttcttggg aacacagattt taaatgtctg  
 atataaaaca  
 15 1741 tgccacagga gaattcgggg atttgagttt ctctgaatag catatatatg  
 atgcatcggg  
 1801 taggtcatta tgatttttta ccatttcgac ttacataatg aaaaccaatt  
 cattttaaat  
 1861 atcagattat tattttgtaa gttgtggaaa aagctaattg tagttttcat  
 20 tatgaagttt  
 1921 tccaataaaa ccaggatctt t

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
spleen tyrosine kinase	SYK	NM_003177	19

25 1 aggaagagcc gcgggcccgg cggctgaggc cccccggcg gcggtggag  
 agcgaggagg  
 61 agcgggtggc cccgcgctgc gcccgccctc gcctcacctg gcgaggtgg  
 acacctgcg  
 30 121 aggtgtgtgc cctccggccc ctgaagcatg gccagcagcg gcatggctga  
 cagcgccaac  
 181 cacctgcctt tctttttcgg caacatcacc cgggaggagg cagaagatta  
 cctggtccag  
 241 gggggcatga gtgatgggct ttatttgctg cgccagagcc gcaactacct  
 35 ggggtggcttc  
 301 gccctgtccg tggcccacgg gaggaaggca caccactaca ccatcgagcg  
 ggagctgaat  
 361 ggcacctacg ccatcgccgg tggcaggacc catgccagcc ccgcccacct  
 ctgccactac  
 40 421 cactcccagg agtctgatgg cctggtctgc ctctcaaga agcccttcaa  
 ccggcccaa  
 481 ggggtgcagc ccaagactgg gccctttgag gatttgaagg aaaacctcat  
 caggaatat  
 541 gtgaagcaga catggaacct gcagggtcag gctctggagc aggccatcat  
 45 cagtcagaag  
 601 cctcagctgg agaagctgat cgctaccaca gcccatgaaa aaatgccttg  
 gttccatgga  
 661 aaaatctctc gggaagaatc tgagcaaatt gtctgatag gatcaaagac  
 aaatggaaag  
 50 721 ttcctgatcc gagccagaga caacaacggc tctacgccc tgtgctgct  
 gcacgaaggg

781 aagggtgctgc actatcgcat cgacaaagac aagacagggg agctctccat  
 ccccgagggg  
 841 aagaagtctg acacgctctg gcagctagtc gagcattatt cttataaagc  
 agatggtttg  
 5 901 ttaagagttc ttactgtccc atgtcaaaaa atcggcacac agggaaatgt  
 taattttgga  
 961 ggccgtccac aacttcagg ttcccatcct gcgacttggc cagcgggtgg  
 aataatctca  
 1021 agaatcaaat catactcctt cccaaagcct ggccacagaa agtcctcccc  
 10 tgcccaaggg  
 1081 aaccggcaag agagtactgt gtcattcaat ccgtatgagc cagaacttgc  
 accctgggct  
 1141 gcagacaaaag gccccagag agaagcccta cccatggaca cagaggtgta  
 cgagagcccc  
 15 1201 tacgaggacc ctgaggagat caggcccaag gaggtttacc tggaccgaaa  
 gctgctgacg  
 1261 ctggaagaca aagaactggg ctctggtaat tttggaactg tgaaaaaggg  
 ctactacca  
 1321 atgaaaaaag ttgtgaaaac cgtggctgtg aaaatactga aaaacgaggg  
 20 caatgacccc  
 1381 gctcttaaag atgagttatt agcagaagca aatgtcatgc agcagctgga  
 caaccgctac  
 1441 atcgtgcgca tgatcgggat atgcgagggc gagtcctgga tgctagttat  
 ggagatggca  
 25 1501 gaacttggtc ccctcaataa gtatttgcag cagaacagac atgtcaagga  
 taagaacatc  
 1561 atagaactgg ttcacaggt ttccatgggc atgaagtact tggaggagag  
 caattttgtg  
 1621 cacagagatc tggtgcaag aatgtgttg ctagttacc aacattatgc  
 30 caagatcagt  
 1681 gatttcggac tctccaaagc actgcgtgct gatgaaaact actacaaggg  
 ccagacccat  
 1741 ggaaagtggc ctgtcaagtg gtaogctccg gaatgcatca actactacaa  
 gttctccagc  
 35 1801 aaaagcgatg tctggagctt tggagtgttg atgtgggaag cattctocta  
 tgggcagaag  
 1861 ccatatcgag ggatgaaagg aagtgaagtc accgctatgt tagagaaagg  
 agagcggatg  
 1921 gggtgccctg caggggtgtcc aagagagatg tacgatctca tgaatctgtg  
 40 ctggacatac  
 1981 gatgtggaaa acaggcccgg attcgcagca gtggaactgc ggctgcgcaa  
 ttactactat  
 2041 gacgtgggta actaaccgct ccgcacctg tcggtggctg cttttgatca  
 caggagcaat  
 45 2101 cacaggaaaa tgtatccaga ggaattgatt gtcagccacc tccctctgcc  
 agtcgggaga  
 2161 gccaggcttg gatggaacat gcccacaact tgcacccaa agcctgtccc  
 aggactcacc  
 2221 ctccacaaaag caaaggcagt cccgggagaa aagacggatg gcaggatcca  
 50 aggggctagc  
 2281 tggatttgtt tgttttcttg tctgtgtgat tttcatacag gttattttta  
 cgatctgttt  
 2341 ccaaatccct ttcattgtct tccacttctc tgggtcccgg ggtgcatttg  
 ttactcatcg

2401 ggcccaggga cattgcagag tggcctagag cactctcacc ccaagcggcc  
 ttttccaaat  
 2461 gcccaaggat gccttagcat gtgactcctg aagggaggc aaaggcagag  
 gaatttggt  
 5 2521 gcttctacgg ccatgagact gatccctggc cactgaaaag ctttctgac  
 aataaaaatg  
 2581 ttttgaggct ttaaaaagaa aatcaagttt gaccagtgc gtttctaage  
 atgtagccag  
 2641 ttaaggaaaag aaagaaaaaa

10

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>solute carrier family 12 (sodium/potassium/chloride transporters), member 2</b>	SLC12A2	NM_001046	20

1 ggtggcctct gtggccgtcc aggctagcgg cggcccgcag gcggcgggga  
 gaaagactct  
 15 61 ctcacctggt cttgcggctg tggccaccgc cggccagggg tgtggagggc  
 gtgctgccgg  
 121 agacgtccgc cgggctctgc agttccgccg ggggtcgggc agctatggag  
 ccgcggccca  
 181 cggcgcctc ctccggcgcc ccgggactgg ccggggtcgg ggagacgccg  
 20 tcagccgctg  
 241 cgctggccgc agccagggg gaactgcccg gcacggctgt gccctcggtg  
 ccggaggatg  
 301 ctgcgccgc gagccgggac ggcggcgggg tccgcgatga gggccccgcg  
 gcggccgggg  
 25 361 acgggctggg cagacccttg gggcccacc cagaccagag ccgtttccag  
 gtggacctg  
 421 tttccgagaa cgccgggcgg gccgctgctg cggcggcggc ggcggcggcg  
 gcagcggcgg  
 481 cggctggtgc tggggcgggg gccaaagcaga ccccccgga cggggaagcc  
 30 agcggcgaga  
 541 gcgagccggc taaaggcagc gaggaagcca agggccgctt ccgcgtgaac  
 ttcgtggacc  
 601 cagctgcctc ctcgctggct gaagacagcc tgtcagatgc tgccggggtc  
 ggagtcgacg  
 35 661 ggcccaacgt gagcttccag aacggcgggg acacggtgct gagcgagggc  
 agcagcctgc  
 721 actccggcgg cggcggcggc agtgggcacc accagcacta ctattatgat  
 acccacacca  
 781 acacctaeta cctgcgcacc ttcggccaca acaccatgga cgctgtgccc  
 40 aggatcgatc  
 841 actaccggca cacagccgcg cagctgggcg agaagctgct ccggcctagc  
 ctggcggagc  
 901 tccacgacga gctggaaaag gaaccttttg aggatggctt tgcaaatggg  
 gaagaaagta  
 45 961 ctccaaccag agatgctgtg gtcacgtata ctgcagaaag taaaggagtc  
 gtgaagtttg  
 1021 gctggatcaa ggggtgatta gtacgttgta tgttaaacad ttgggggtgtg  
 atgcttttca

1081 ttagattgtc atggattgtg ggtcaagctg gaataggtct atcagtcott  
 gtaataatga  
 1141 tggccactgt tgtgacaact atcacaggat tgtctacttc agcaatagca  
 actaatggat  
 5 1201 ttgtaagagg aggaggagca tattatttaa tatctagaag tctagggcca  
 gaatttgggtg  
 1261 gtgcaattgg tctaactcttc gcctttgcca acgctggtgc agttgctatg  
 tatgtggttg  
 1321 gatttgcaga aaccgtgggtg gagttgctta aggaacattc catacttatg  
 10 atagatgaaa  
 1381 tcaatgatat ccgaattatt ggagccatta cagtcgtgat tcttttaggt  
 atctcagtag  
 1441 ctggaatgga gtgggaagca aaagctcaga ttgttctttt ggtgatccta  
 cttcttgcta  
 15 1501 ttggtgattt cgtcatagga acatttatcc cactggagag caagaagcca  
 aaagggtttt  
 1561 ttggttataa atctgaaata tttaatgaga actttgggcc cgattttcga  
 gaggaagaga  
 1621 ctttcttttc tgtatttgcc atcttttttc ctgctgcaac tggattctg  
 20 gctggagcaa  
 1681 atatctcagg tgatcttgca gatcctcagt cagccatacc caaaggaaca  
 ctcttagcca  
 1741 ttttaattac tacattgggtt tacgtaggaa ttgcagtatc tgtaggttct  
 tgtgttggtc  
 25 1801 gagatgccac tggaaacggt aatgacacta tcgtaacaga gctaacaaac  
 tgtacttctg  
 1861 cagctgcaa attaaacttt gatttttcat cttgtgaaag cagtccttgt  
 tcctatggcc  
 1921 taatgaacaa cttccaggta atgagtatgg tgtcaggatt tacaccacta  
 30 atttctgcag  
 1981 gtatattttc agccactctt tcttcagcat tagcatccct agtgagtgct  
 cccaaaatat  
 2041 ttcaggctct atgtaaggac aacatctacc cagctttcca gatgtttgct  
 aaaggttatg  
 35 2101 ggaaaaataa tgaacctctt cgtggctaca tcttaacatt ctttaattgca  
 cttggattca  
 2161 tcttaattgc tgaactgaat gttattgcac caattatctc aaacttcttc  
 cttgcatcat  
 2221 atgcattgat caatttttca gtattccatg catcacttgc aaaatctcca  
 40 ggatggcgtc  
 2281 ctgcattcaa atactacaac atgtggatat cacttcttgg agcaattctt  
 tgttgcatag  
 2341 taatgttcgt cattaactgg tgggctgcat tgctaacata tgtgatagtc  
 cttgggctgt  
 45 2401 atatttatgt tacctacaaa aaaccagatg tgaattgggg atcctctaca  
 caagccctga  
 2461 cttacctgaa tgcactgcag cattcaattc gtctttctgg agtgggaagac  
 cacgtgaaaa  
 2521 actttaggcc acagtgctctt gttatgacag gtgctccaaa ctcacgtcca  
 50 gctttacttc  
 2581 atcttgttca tgatttcaca aaaaatggtg gtttgatgat ctgtggccat  
 gtacatatgg  
 2641 gtcctcgaag acaagccatg aaagagatgt ccatcgatca agccaaatat  
 cagcgatggc

2701 ttattaagaa caaatgaag gcattttatg ctccagtaca tgcagatgac  
 ttgagagaag  
 2761 gtgcacagta tttgatgcag gctgctggtc ttggctcgat gaagccaaac  
 acacttgcc  
 5 2821 ttggatttaa gaaagattgg ttgcaagcag atatgagggga tgtggatatg  
 tatataaact  
 2881 tatttcatga tgcttttgac atacaatatg gagtagtggt tattcgccta  
 aaagaaggtc  
 2941 tggatatatc tcatcttcaa ggacaagaag aattattgac atcacaagag  
 10 aatctcctg  
 3001 gcaccaagga tgtggtagta agtgtggaat atagtaaaaa gtccgattta  
 gatacttcca  
 3061 aaccactcag tgaaaaacca attacacaca aagttgagga agaggatggc  
 aagactgcaa  
 15 3121 ctcaaccact gttgaaaaaa gaatccaaag gccctattgt gcctttaaat  
 gtagctgacc  
 3181 aaaagcttct tgaagctagt acacagtttc agaaaaaca aggaaagaat  
 actattgatg  
 3241 tctggtggct ttttgatgat ggaggttga ccttattgat accttacctt  
 20 ctgacgacca  
 3301 agaaaaaatg gaaagactgt aagatcagag tattcattgg tggaaagata  
 aacagaatag  
 3361 accatgaccg gagagcgatg gctactttgc ttagcaagtt ccggatagac  
 ttttctgata  
 25 3421 tcatggttct aggagatate aataccaaac caaagaaaga aatattata  
 gcttttgagg  
 3481 aatcattga gccatacaga cttcatgaag atgataaaga gcaagatatt  
 gcagataaaa  
 3541 tgaagaaga tgaacatgg cgaataacag ataatgagct tgaactttat  
 30 aagaccaaga  
 3601 cataccggca gatcaggtta aatgagttat taaaggaaca ttcaagcaca  
 gctaataatta  
 3661 ttgtcatgag tctcccagtt gcacgaaaag gtgctgtgac tagtgctctc  
 tacatggcat  
 35 3721 ggttagaagc tctatctaag gacctaccac caatcctcct agttcgtggg  
 aatcatcaga  
 3781 gtgtccttac cttctattca taaatgttct atacagtgga cagccctcca  
 gaatggact  
 3841 tcagtgccta gtgtagtaac tgaaatcttc aatgacacat taacatcaca  
 40 atggcgaatg  
 3901 gtgacttttc tttcacgatt tcattaattt gaaagcacac aggaaagttg  
 ctccattgat  
 3961 aacgtgatg gagacttcgg ttttagtcaa ttccatatct caatcttaat  
 ggtgattctt  
 45 4021 ctctgttgaa ctgaagtttg tgagagtagt tttcctttgc tacttgaata  
 gcaataaaag  
 4081 cgtgttaact ttttgattga tgaaagaagt acaaaaagcc tttagccttg  
 aggtgccttc  
 4141 tgaattaac caaatttcat ccatatatcc tcttttataa acttatagaa  
 50 tgtcaaactt  
 4201 tgccttcaac tgtttttatt tctagtctct tccactttaa aacaaaatga  
 aactgcttg  
 4261 tcttcttcca ttgaccattt agtgttgagt actgtatgtg ttttgtaat  
 tctataaagg

4321 tatctgtag atattaaagg tgagaattag ggcaggtaa tcaaaaatgg  
 ggaaggggaa  
 4381 atggtaacca aaaagtaacc coatggtaag gtttatatga gtatatgtga  
 atatagagct  
 5 4441 aggaaaaaaaa gcccccccaa ataccttttt aaccctctg attggctatt  
 attactatat  
 4501 ttattattat ttattgaaac cttaggggaag attgaagatt catcccatac  
 ttctatatac  
 4561 catgcttaaa aatcacgtca ttctttaaac aaaaatactc aagatcattt  
 10 atatttattt  
 4621 ggagagaaaa ctgtcctaata ttagaatttc cctcaaactc gagggacttt  
 taagaaatgc  
 4681 taacagattt ttctggagga aatttagaca aaacaatgac atttagtaga  
 atatttcagt  
 15 4741 atttaagtgg aatttcagta tactgtacta tcctttataa gtcattaaaa  
 taatgtttca  
 4801 tcaaatgggt aaatggacca ctggtttctt agagaaatgt ttttaggctt  
 aattcattca  
 4861 attgtcaagt acacttagtc ttaatacact caggtttgaa cagattattc  
 20 tgaatattaa  
 4921 aatttaactc attcttaata ttttaaaact tttgttaaga aaaactgcca  
 gtttgtgctt  
 4981 ttgaaatgac tgttttgaca tcatagtcta gtaaaatctt gacagtgcac  
 atgtactgct  
 25 5041 actaaaagct ttatatgaaa ttattaatgt gaagtcttctc atttataatt  
 caaggaagga  
 5101 tttctgaaa acatttcaag ggatttatgt ctacatattt gtgtgtgtgt  
 gtgtatatac  
 5161 atgtaatatg catacacaga tgcatatgtg tatatataat gaaatttatg  
 30 ttgctgggtat  
 5221 tttgcatttt aaagtgatca agattcatta ggcaaacttt ggtttaagta  
 aacatatgct  
 5281 caaaatcaga ttaacagata caggtttcat agagaacaaa ggtgatcatt  
 tgaagggcat  
 35 5341 gctgtaattt cacacaattt tccagttcaa aatggagaa tacttgcct  
 aaaatactgt  
 5401 taagtgggtt aattgatata agtttctgtg gtggaaaatt tatgcaggtt  
 ttcacgaatc  
 5461 cttttttttt tttttttttt tttttgagac ggagtcttgc tctgttgcca  
 40 cgctggaatg  
 5521 cagtaacgtg atcttggtc actgcgacct ccacctccc agttcaagcg  
 attctcctgc  
 5581 ctacagctcc ctagtagctg ggactacggg tgcacgccac catgccagc  
 taatttttgt  
 45 5641 attttgagta gagacaggtt ttcaccgtgt tggctaggat ggtgtctatc  
 tcttgacctt  
 5701 gtgatccacc cgcctcagcc tcccagagtg ctgggattac aggtgcgagc  
 cactgcgctt  
 5761 ggctggtttt catgaatctt gatagacatc tataacgtta ttattttcag  
 50 tgggtgtgag  
 5821 catttttgct tcatgagtat gacctaggta tagagatctg ataacttgaa  
 ttcagaatat  
 5881 taagaaaatg aagtaactga ttttctaaaa aaaaaaaaa aaaaaatttc  
 tacattataa

5941 ctcacagcat tgttcattg caggttttgc aatgtttggg ggtaaagaca  
 gtagaaatat  
 6001 tattcagtaa acaataatgt gtgaactttt aagatggata atagggcatg  
 gactgagtgc  
 5 6061 tgctatcttg aaatgtgcac aggtacactt accttttttt ttttttttt  
 taagtttttc  
 6121 ccattcagga aaacaacatt gtgatctgta ctacaggaac caaatgtcat  
 gcgtcataca  
 6181 tgtgggtata aagtacataa aatatatcta actattcata atgtggggtg  
 10 ggtaatactg  
 6241 tctgtgaaat aatgtaagaa gcttttctact taaaaaaaaat gcattacttt  
 cacttaacac  
 6301 tagacaccag gtcgaaaatt ttcaagggtta tagtacttat ttcaacaatt  
 cttagagatg  
 15 6361 ctagctagtg ttgaagctaa aatagcttt atttatgctg aattgtgatt  
 tttttatgcc  
 6421 aaattttttt tagttctaata cattgatgat agcttggaat taaataatta  
 tgccatggca  
 6481 tttgacagtt cattattcct ataagaatta aattgagttt agagagaatg  
 20 gtgggtgtga  
 6541 gctgattatt aacagttact gaaatcaaat atttatttgt tacattattc  
 catttgtatt  
 6601 ttaggtttcc ttttacatcc tttttatatg cattctgaca ttacatattt  
 ttttaagacta  
 25 6661 tggaaataat ttaaagattt aagctctggg ggatgattat ctgctaagta  
 agtctgaaaa  
 6721 tgtaaatattt tgataatact gtaatatacc tgtcacacaa atgcttttct  
 aatgttttta  
 6781 ccttgagtat tgcagttgct gctttgtaca gaggttactg caataaagga  
 30 agtggattca  
 6841 ttaaaccctat ttaatgtcca

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>guanylate cyclase 2C (heat stable enterotoxin receptor)</b>	GUCY2C	NM_004963	21

1 cgcaaagcaa gtgggcacaa ggagtatggt tctaactgta ttgggggtcat  
 35 gaagacgttg  
 61 ctgttgact tggctttgtg gtcactgctc ttccagcccg ggtggctgtc  
 ctttagttcc  
 121 caggtgagtc agaactgccca caatggcagc tatgaaatca gcgtcctgat  
 gatgggcaac  
 40 181 tcagcctttg cagagcccct gaaaaacttg gaagatgcgg tgaatgaggg  
 gctggaaata  
 241 gtgagaggac gtctgcaaaa tgctggccta aatgtgactg tgaacgctac  
 tttcatgtat  
 301 tcggatggtc tgattcataa ctcaggcgac tgccggagta gcacctgtga  
 45 aggcctcgac  
 361 ctactcagga aaatttcaaa tgcacaacgg atgggctgtg tcctcatagg  
 gccctcatgt  
 421 acatactcca ccttccagat gtaccttgac acagaattga gotaccccat  
 gatctcagct

481 ggaagttttg gattgtcatg tgactataaa gaaaccttaa ccaggctgat  
 gtctccagct  
 541 agaaagttga tgtacttctt ggtaaacttt tggaaaacca acgatctgcc  
 cttcaaaact  
 5 601 tattcctgga gcacttcgta tgtttacaag aatggtacag aaactgagga  
 ctgtttctgg  
 661 taccttaaatg ctctggaggc tagcgtttcc tatttctccc acgaactcgg  
 ctttaaggtg  
 721 gtgtaagac aagataagga gtttcaggat atcttaatgg accacaacag  
 10 gaaaagcaat  
 781 gtgattatta tgtgtggtgg tccagagttc ctctacaagc tgaaggggtga  
 ccgagcagtg  
 841 gctgaagaca ttgtcattat totagtggat cttttcaatg accagtactt  
 ggaggacaat  
 15 901 gtcacagccc ctgactatat gaaaaatgtc cttgttctga cgctgtctcc  
 tgggaattcc  
 961 cttctaaata gctctttctc caggaatcta tcaccaacia aacgagactt  
 tgctcttgcc  
 1021 tatttgaatg gaatcctgct ctttggacat atgctgaaga tatttcttga  
 20 aatggagaa  
 1081 aatattacca ccccaaat tgctcatgct ttcaggaatc tcacttttga  
 agggatgac  
 1141 ggtccagtga ccttgatga ctggggggat gttgacagta ccatggtgct  
 tctgtatacc  
 25 1201 tctgtggaca ccaagaaata caaggttctt ttgacctatg ataccacgt  
 aaataagacc  
 1261 tatcctgtgg atatgagccc cacattcact tggagaact ctaaacttcc  
 taatgatatt  
 1321 acaggccggg gccctcagat cctgatgatt gcagtcttca ccctcactgg  
 30 agctgtggtg  
 1381 ctgctcctgc tegtcgctct cctgatgctc agaaaatata gaaaagatta  
 tgaacttctg  
 1441 cagaaaaaat ggtcccacat tccctcctgaa aatatcttcc ctctggagac  
 caatgagacc  
 35 1501 aatcatgtta gcctcaagat cgatgatgac aaaagacgag atacaatcca  
 gagactacga  
 1561 cagtgcaaat acgacaaaaa gcgagtgatt ctcaaagatc tcaagcacia  
 tgatggtaat  
 1621 ttcactgaaa aacagaagat agaattgaac aagttgcttc agattgacta  
 40 ttacaacctg  
 1681 accaagttct acggcacagt gaaacttgat accatgatct tcgggggtgat  
 agaatactgt  
 1741 gagagaggat ccctccggga agttttaaata gacacaattt cctaccctga  
 tggcacattc  
 45 1801 atggattggg agtttaagat ctctgtcttg tatgacattg ctaagggaaat  
 gtcatatctg  
 1861 cactccagta agacagaagt ccatggtcgt ctgaaatcta ccaactgcgt  
 agtggacagt  
 1921 agaatggtgg tgaagatcac tgattttggc tgcaattcca ttttacctcc  
 50 aaaaaaggac  
 1981 ctgtggacag ctccagagca cctccgcca gccaacatct ctcagaaagg  
 agatgtgtac  
 2041 agctatggga tcatcgcaca ggagatcatt ctgcggaaag aaaccttcta  
 cactttgagc

2101 tgtcgggacc ggaatgagaa gattttcaga gtggaaaatt ccaatggaat  
 gaaacccttc  
 2161 cgcccagatt tattcttggg aacagcagag gaaaaagagc tagaagtgta  
 cctacttgta  
 5 2221 aaaaactggt gggaggaaga tccagaaaag agaccagatt tcaaaaaaat  
 tgagactaca  
 2281 cttgccaaaga tatttggact ttttcatgac caaaaaaatg aaagctatat  
 ggataccttg  
 2341 atccgacgtc tacagctata ttctcgaaac ctggaacatc tggtagagga  
 10 aaggacacag  
 2401 ctgtacaagg cagagagggg cagggctgac agacttaact ttatgttgct  
 tccaaggcta  
 2461 gtggtaaagt ctctgaagga gaaaggcttt gtggagccgg aactatatga  
 ggaagttaca  
 15 2521 atctacttca gtgacattgt aggtttcact actatctgca aatacagcac  
 ccccatggaa  
 2581 gtggtggaca tgcttaatga catctataag agttttgacc acattgttga  
 tcatcatgat  
 2641 gtctacaagg tggaaacat cggatgatgag tacatggtgg ctagtggttt  
 20 gcctaagaga  
 2701 aatggcaatc ggcattgcaat agacattgcc aagatggcct tggaaatcct  
 cagcttcatg  
 2761 gggacctttg agctggagca tcttctcggc ctcccaatat ggattcgcac  
 tggagttcac  
 25 2821 tctggtcctt gtgctgctgg agttgtggga atcaagatgc ctcgttattg  
 tctatttggg  
 2881 gatacgggtca acacagcctc taggatggaa tccactggcc tccctttgag  
 aattcacgtg  
 2941 agtggctcca ccatagccat cctgaagaga actgagtgcc agttccttta  
 30 tgaagtgaga  
 3001 ggagaaacat acttaaaggg aagaggaaat gagactacct actggctgac  
 tgggatgaag  
 3061 gaccagaaat tcaacctgcc aacctcctc actgtggaga atcaacagcg  
 tttgcaagca  
 35 3121 gaattttcag acatgattgc caactcttta cagaaaagac aggcagcagg  
 gataagaagc  
 3181 caaaaacca gacgggtagc cagctataaa aaaggcactc tggaaactt  
 gcagctgaat  
 3241 accacagaca aggagagcac ctatttttaa acctaaatga ggtataagga  
 40 ctcacacaaa  
 3301 ttaaaaataca gctgcactga ggcagcgacc tcaagtgtcc tgaaagotta  
 cattttcctg  
 3361 agacctcaat gaagcagaaa tgtacttagg cttggctgcc ctgtctggaa  
 catggacttt  
 45 3421 cttgcatgaa tcagatgtgt gttctcagtg aaataactac cttccactct  
 ggaaccttat  
 3481 tccagcagtt gttccagga gcttctacct ggaaaagaaa agaaatgaat  
 agactatcta  
 3541 gaacttgaga agattttatt cttatttcat ttattttttg tttgtttatt  
 50 tttatcgttt  
 3601 ttgtttactg gctttccttc tgtattcata agatttttta aattgtcata  
 attatatttt  
 3661 aaatacccat cttcattaaa gtatatttaa ctcataattt ttgcagaaaa  
 tatgctatat  
 55 3721 attaggcaag aataaaagct aaagg

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>transmembrane 4 superfamily member 4</b>	TM4SF4	NM_004617	22

```

1  cttcagggtca gggagaatgt ataaatgtcc attgccatcg aggttctgct
   atttttgaga
5  61 agctgaagca actccaagga cacagttcac agaaatttgg ttctcagccc
   caaaatactg
   121 attgaattgg agacaattac aaggactctc tggccaaaaa cccttgaaga
   ggccccgtga
   181 aggaggcagt gaggagcttt tgattgctga cctgtgtcgt accaccccag
10  aatgtgcact
   241 gggggctgtg ccagatgcct gggggggacc ctcatcccc ttgctttttt
   tggcttcctg
   301 gctaacatcc tgttatTTTT tcttggagga aaagtgatag atgacaacga
   ccacctttcc
15  361 caagagatct ggtttttcgg aggaatatta ggaagcgggtg tcttgatgat
   cttccctgcg
   421 ctggtgttct tgggcctgaa gaacaatgac tgctgtgggt gctgcgcaa
   cgagggctgt
   481 ggggaagcgt ttgcgatggt cacctccacg atatttgctg tggttggatt
20  cttgggagct
   541 ggatactcgt ttatcatctc agccatttca atcaacaagg gtcttaaag
   cctcatggcc
   601 aatagtacat ggggctacct cttccacgac ggggattatc tcaatgatga
   ggcttatgg
25  661 aacaagtgcc gagagcctct caatgtgggt ccttggaaac tgaccctctt
   ctccatcctg
   721 ctggtcgtag gaggaatcca gatggttctc tgcgccatcc aggtgggtcaa
   tggcctcctg
   781 gggaccctct gtggggactg ccagtgttgt ggctgctgtg ggggagatgg
30  acccgtttaa
   841 acctccgaga tgagctgctc agactctaca gcatgacgac tacaatttct
   tttcataaaa
   901 cttcttctct tcttggaaatt attaattcct atctgcttcc tagctgataa
   agcttagaaa
35  961 aggcagttat tccttcttcc caaccagctt tgctcgagtt agaattttgt
   tattttcaaa
   1021 taaaaaatag tttggccact taacaaattt gattataaaa tctttcaaat
   tagttccttt
   1081 ttagaattta ccaacagggt caaagcatac ttttcatgat tttttatta
40  caaatgtaaa
   1141 atgtataaag tcacatgtac tgccatacta cttctttgta tataaagatg
   tttatatctt
   1201 tggaggtttt acataaatca aaggaagaaa gcacatttaa aatgagaaac
   taagaccaat
45  1261 ttctgttttt aagaggaaaa agaatgattg atgtatccta agtattgtta
   tttgttgtct
   1321 ttttttgctg ccttgcttga gttgcttgtg actgatcttt tgaggctgtc
   atcatggcta
   1381 gggttctttt atgtatgtta aattaaacc tgaattcaga ggtaacgt

```

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>transforming growth factor, alpha</b>	TGFA	NM_003236	23

```

1 ctggagagcc tgctgcccgc ccgcccgtaa aatggcctccc tcggctggac
agctcgcctt
5 61 gttcgtctctg ggtattgtgt tggctgcgtg ccaggccttg gagaacagca
cgtccccgct
121 gagtgcagac ccgcccgtgg ctgcagcagt ggtgtcccat tttaatgact
gcccagattc
181 ccacactcag ttctgcttcc atggaacctg caggtttttg gtgcaggagg
10 acaagccagc
241 atgtgtctgc cattctgggt acgttgggtgc acgctgtgag catgcccacc
tcttggccgt
301 ggtggctgcc agccagaaga agcaggccat caccgccttg gtggtggtct
ccatcgtggc
15 361 cctggctgtc cttatcatca catgtgtgct gatacactgc tgccaggtcc
gaaaacactg
421 tgagtgtgtc cgggccctca tctgccggca cgagaagccc agcgcctcc
tgaaggaag
481 aaccgcttgc tgccactcag aaacagtggc ctgaagagcc cagaggagga
20 gtttggccag
541 gtggactgtg gcagatcaat aaagaaaggc ttcttcagga cagcactgcc
agagatgcct
601 ggggtgtgcca cagaccttcc tacttggcct gtaatcacct gtgcagcctt
ttgtggcct
25 661 tcaaaactct gtcaagaact ccgtctgctt ggggttattc agtgtgacct
agagaagaaa
721 tcagcggacc acgatttcaa gacttgtaa aaaagaactg caaagagacg
gactcctgtt
781 cacctagggt aggtgtgtgc agcagttggt gtctgagtc acatgtgtgc
30 agttgtcttc
841 tgccagccat ggattccagg ctatatattt ctttttaatg ggccacctcc
ccacaacaga
901 attctgccc aacacaggaga tttctatagt tattgttttc tgcatttgc
ctactgggga
35 961 agaaagtgaa ggaggggaaa ctgtttaata tcacatgaag accctagctt
taagagaagc
1021 tgtatcctct aaccacgaga ctctcaacca gcccaacatc ttccatggac
acatgacatt
1081 gaagaccatc ccaagctatc gccacccttg gagatgatgt cttatttatt
40 agatggataa
1141 tggttttatt tttaatctct taagtcaatg taaaagtat aaaaccctt
cagacttcta
1201 cattaatgat gtatgtgttg ctgactgaaa agctatactg attagaaatg
tctggcctct
45 1261 tcaagacagc taaggcttgg gaaaagtctt ccagggtgcg gagatggaac
cagaggctgg
1321 gttactggta ggaataaagg taggggttca gaaatgggtgc cattgaagcc
acaaagccgg
1381 taaatgcctc aatacgttct gggagaaaac ttagcaaatc catcagcagg
50 gatctgtccc

```

1441 ctctggtggg gagagaggaa gagtgtgtgt gtctacacag gataaaccca  
 atacatattg  
 1501 tactgctcag tgattaaatg ggttcacttc ctcgtgagcc ctcggtaagt  
 atgttttagaa  
 5 1561 atagaacatt agccacgagc cataggcatt tcaggccaaa tccatgaaag  
 ggggaccagt  
 1621 catttatttt ccattttggt gcttgggttg tttgttgctt ttttttaaa  
 aggagaagtt  
 1681 taactttgct atttattttc gagcactagg aaaactattc cagtaatttt  
 10 tttttcctca  
 1741 tttccattca ggatgccggc tttattaaca aaaactctaa caagtcacct  
 ccactatgtg  
 1801 ggtcttcctt tcccctcaag agaaggagca attgttcccc tgacatctgg  
 gtccatctga  
 15 1861 cccatggggc ctgcctgtga gaaacagtgg gtcccttcaa atacatagtg  
 gatagctcat  
 1921 ccctaggaat tttcattaaa atttggaaac agagtaatga agaaataata  
 tataaactcc  
 1981 ttatgtgagg aaatgctact aatatctgaa aagtgaaga tttctatgta  
 20 ttaactctta  
 2041 agtgcaccta gcttattaca tcgtgaaagg tacatttaaa atatgttaaa  
 ttggcttgaa  
 2101 attttcagag aattttgtct tcccctaatt cttcttcctt ggtctggaag  
 aacaatttct  
 25 2161 atgaattttc tctttatttt tttttataa ttcagacaat tctatgacct  
 gtgtcttcat  
 2221 ttttggcact cttatttaac aatgccacac ctgaagcact tggatctggt  
 cagagctgac  
 2281 cccctagcaa cgtagttgac acagctccag gtttttaaat tactaaaata  
 30 agttcaagtt  
 2341 tacatccctt gggccagata tgtgggttga ggcttgactg tagcatcctg  
 cttagagacc  
 2401 aatcaatgga cactggtttt tagacctcta tcaatcagta gttagcatcc  
 aagagacttt  
 35 2461 gcagagggcgt aggaatgagg ctggacagat ggcggaacga gaggttccct  
 gcgaagactt  
 2521 gagatttagt gtctgtgaat gttctagttc ctaggctcag caagtcacac  
 ctgccagtgc  
 2581 cctcatcctt atgcctgtaa cacacatgca gtgagaggcc tcacatatac  
 40 gcctccctag  
 2641 aagtgccttc caagtcagtc ctttggaaac cagcaggtct gaaaaagagg  
 ctgcatcaat  
 2701 gcaagcctgg ttggaccatt gtccatgcct caggatagaa cagcctggct  
 ttttgggga  
 45 2761 tttttcttct agaaatcaaa tgactgataa gcattggctc cctctgccat  
 ttaatggcaa  
 2821 tggtagtctt tggttagctg caaaaatact ccatttcaag ttaaaaatgc  
 atcttcta  
 2881 ccactctctgc aagctccctg tgtttccttg cccttagaa aatgaattgt  
 50 tcaactacaat  
 2941 tagagaatca tttaacatcc tgacctggta agctgccaca cacctggcag  
 tggggagcat  
 3001 cgctgtttcc aatggctcag gagacaatga aaagcccca tttaaaaaaa  
 taacaaacat

3061 tttttaaag gcctccaata ctcttatgga gcctggattt ttcccactgc  
tctacaggct  
3121 gtgacttttt ttaagcatcc tgacaggaaa tgttttcttc tacatggaaa  
gatagacagc  
5 3181 agccaaccct gatctggaag acagggcccc ggctggacac acgtggaacc  
aagccagggga  
3241 tgggctggcc attgtgtccc cgcaggagag atgggcagaa tggccctaga  
gttcttttcc  
3301 ctgagaaagg agaaaaagat gggattgcca ctcaccacc cacactggta  
10 agggaggaga  
3361 atttgtgctt ctggagcttc tcaagggatt gtgttttgca ggtacagaaa  
actgcctggt  
3421 atcttcaagc caggttttcg agggcacatg ggtcaccagt tgctttttca  
gtcaatttgg  
15 3481 ccgggatgga ctaatgaggc tctaacactg ctcaggagac cctgccttc  
tagttggttc  
3541 tgggctttga tctcttcaa cctgcccagt cacagaagga ggaatgactc  
aatgcccga  
3601 aaccaagaac acattgcaga agtaagacaa acatgtatat ttttaaatgt  
20 tctaacataa  
3661 gacctgttct ctctagccat tgatttacca ggctttctga aagatctagt  
ggttcacaca  
3721 gagagagaga gactactgaa aaagcaactc ctcttcttag tcttaataat  
ttactaaaat  
25 3781 ggtcaacttt tcattatctt tattataata aacctgatgc ttttttttag  
aactccttac  
3841 tctgatgtct gtatatgttg cactgaaaag gttaatattt aatgttttaa  
tttattttgt  
3901 gtggtaagt aattttgatt tctgtaatgt gttaatgtga ttagcagtta  
30 ttttccttaa  
3961 tatctgaatt aacttaag agtagtgagc aatataagac gcaattgtgt  
ttttcagtaa  
4021 tgtgcattgt tattgagttg tactgtacct tatttggag gatgaaggaa  
tgaacctttt  
35 4081 tttcctaaaa

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>fibroblast growth factor binding protein 1</b>	FGFBP1	NM_005130	24

1 gaatagtcta ccccccttgc actctadctg acacagctgc agcctgcaat  
tcaactcgac  
40 61 tgcctgggat tgcactggat ccgtgtgctc agaacaaggt gaacgcccag  
ctgcagccat  
121 gaagatctgt agcctcacc tgctctcctt cctcctactg gctgctcagg  
tgctcctggt  
181 ggaggggaaa aaaaaagtga agaatggact tcacagcaaa gtggtctcag  
45 aacaaaagga  
241 cactctgggc aacaccaga ttaagcagaa aagcaggccc gggaaacaaag  
gcaagtttgt  
301 caccaaagac caagccaact gcagatgggc tgctactgag caggaggagg  
gcatctctct

361 caaggttgag tgcactcaat tggaccatga attttctctgt gtctttgctg  
 gcaatccaac  
 421 ctcatgccta aagctcaagg atgagagagt ctattggaaa caagttgccc  
 ggaatctgcg  
 5 481 ctcacagaaa gacatctgta gatattccaa gacagctgtg aaaaccagag  
 tgtgcagaaa  
 541 ggattttcca gaatccagtc ttaagctagt cagctccact ctatttggga  
 acacaaagcc  
 601 caggaaggag aaaacagaga tgtccccccag ggagcacatc aaaggcaaag  
 10 agaccacccc  
 661 ctctagccta gcagtgacce agaccatggc caccaaagct cccgagtgctg  
 tggaggaccc  
 721 agatatggca aaccagagga agactgccct ggagttctgt ggagagactt  
 ggagctctct  
 15 781 ctgcacattc ttcctcagca tagtgcagga cacgtcatgc taatgaggtc  
 aaaagagaac  
 841 gggttccctt aagagatgtc atgtcgtaag tccctctgta tactttaaag  
 ctctctacag  
 901 tcccccaaaa atatgaactt ttgtgcttag tgagtgcaac gaaatattta  
 20 aacaagtttt  
 961 gtattttttg cttttgtggt ttggaatttg cttatttttt cttggatgctg  
 atgttcagag  
 1021 gctgttttct gcagcatgta tttccatggc ccacacagct atgtgtttga  
 gcagcgaaga  
 25 1081 gtctttgagc tgaatgagcc agagtgataa tttcagtgca acgaactttc  
 tgctgaatta  
 1141 atggtaataa aactctgggt gtttttcaga aatacattca

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>PTK6 protein tyrosine kinase 6</b>	PTK6	NM_005975	<b>25</b>

30 1 gctggggccac agcctgggtcc tgccgctgctg cccgcccgcc atgggtgtccc  
 gggaccaggc  
 61 tcacctgggc cccaagtatg tgggcctctg ggacttcaag tcccggacgg  
 acgaggagct  
 121 gagcttccgc gcgggggacg tcttccacgt ggccaggaag gaggagcagt  
 35 ggtggtgggc  
 181 cacgctgctg gacgaggcgg gtggggccgt ggcccagggc tatgtgcccc  
 acaactacct  
 241 ggccgagagg gagacgggtg agtcggaacc gtggttcttt ggctgcatct  
 cccgctcgga  
 40 301 agctgtgctg eggctgcagg ccgagggcaa cgccacgggc gccttctga  
 tcagggtcag  
 361 cgagaagccg agtgccgact acgtcctgtc ggtgcgggac acgcaggctg  
 tgcggcacta  
 421 caagatctgg cggcgtgccc ggggcccggct gcacctgaac gaggcgggtg  
 45 ccttctcag  
 481 cctgcccagag cttgtgaact accacagggc ccagagcctg tcccacggcc  
 tgcggctggc  
 541 cgcgccctgc cgggaagcac agcctgagcc cctgcccctat tgggatgact  
 gggagaggcc

601 gagggaggag ttcacgctct gcaggaagct ggggtccggc tactttgggg  
 aggtcttcga  
 661 ggggctctgg aaagaccggg tccaggtggc cattaaggtg atttctcgag  
 acaacctcct  
 5 721 gcaccagcag atgctgcagt cggagatcca ggccatgaag aagctgcggc  
 acaaacacat  
 781 cctggcgtg tacgccgtgg tgtccgtggg ggaccccggtg tacatcatca  
 cggagctcat  
 841 ggccaagggc agcctgctgg agctgctccg cgactctgat gagaaagtcc  
 10 tgcccgtttc  
 901 ggagctgctg gacatgcct ggcaggtggc tgagggcatg tgttacctgg  
 agtcgcagaa  
 961 ttacatccac cgggacctgg ccgccaggaa catcctcgtc ggggaaaaca  
 cctctgcaa  
 15 1021 agttggggac ttcgggttag ccaggcttat caaggaggac gtctacctct  
 cccatgacca  
 1081 caatatcccc tacaagtgga cggcccctga agcgtctcc cgaggccatt  
 actccaccaa  
 1141 atccgacgtc tggtcctttg ggattctcct gcatgagatg ttcagcaggg  
 20 gtcaggtgcc  
 1201 ctaccaggc atgtccaacc atgaggcctt cctgaggggtg gacgccggct  
 accgcatgcc  
 1261 ctgccctctg gagtgcccgcc ccagcgtgca caagctgatg ctgacatgct  
 ggtgcagggg  
 25 1321 ccccgagcag agaccctgct tcaaggcctt gcgggagagg ctctccagct  
 tcaccagcta  
 1381 cgagaacctg acctgagctg ctgtggagcg ggcatggccg ggccctgctg  
 aggagggggc  
 1441 tgggcagagg gcctggacct gggatcaagg cccacgcgct tccctggggg  
 30 ttactgaggt  
 1501 gatgggtgca ggaaagggtc acaaatgtgg agtgtctgcg tccaatacac  
 gcgtgtgctc  
 1561 ctctccttac tccatcgtgt gtgccttggg tctcagctgc tgacacgcag  
 cctgctctgg  
 35 1621 agcctgcaga tgagatccgg gagactgaca cgaagccagc agaggtcaga  
 ggggactctg  
 1681 accacagccc gctctctggc tgtctgtctg cagtgcccg ctgaggggtgg  
 gaggcaaaca  
 1741 cgccttgctc ctgctcttcc cagttcagct tggggggaga aagtcattcg  
 40 cgtggctcgg  
 1801 gacgctcatg taaatttggg tttgggtgctc aagggttctt tctcccagg  
 ggcaggtggt  
 1861 tctttcctgt ttgtcttggt tcttgagagc ttggccttat gaccagtgag  
 aactctctcc  
 45 1921 ctgggtctctg ccagcccaag catcactgcc cgaggcgcca gctcagtttc  
 accgtccag  
 1981 tccacaaggg gcttttccca ccttcacctt tgtcgtggtg tcagtgtggtg  
 aaagcgcccc  
 2041 tcaactctgc gctgacaagg gcccttctct actgtctgtg ggggtgggtcc  
 50 gggctggggg  
 2101 ggctgcctcc tttgcacctg attttgaagg tgtctctttc atccatggtt  
 aagtcataaa  
 2161 aagcttattg gttttgggtt tgactcacct gaaagttttt ttgggttaaa  
 agaagaatag

2221 gcggggcacg gtggctcatg cctgtaatcc cagcactttg ggaggctgag  
 gcaggtggat  
 2281 cacgaggtca ggagatcgac accatcctgg ctaacacggt gaaaccccgt  
 ctctactaaa  
 5 2341 aaatacaaaa aattagctgg gtgtggtggt ggggggtgggc gcctgtagtc  
 ccagctacgt  
 2401 gggaggctga ggcagcagac tgggtgtgaac cggggaggtg gagcttgca  
 tgagccgaga  
 2461 tcgcgccact gcactccage ctgggcgaca gagcgagact ccatctcaaa

10

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
epithelial V-like antigen 1	EVA1	NM_005797	26

1 acaggcacag gtgaggaact caactcaaac tcctctctct gggaaaacgc  
 ggtgcttgc  
 61 cctcccggag tggccttggc aggggtgttg agccctcggc ctgccccgtc  
 15 cggctctctgg  
 121 ggccaaggct gggtttcct catgtatggc aagagctcta ctcgtgcggt  
 gcttcttctc  
 181 cttggcatac agctcacagc tctttggcct atagcagctg tggaaattta  
 tacctcccgg  
 20 241 gtgctggagg ctgtaaatgg gacagatgct cggttaaaat gcactttctc  
 cagctttgcc  
 301 cctgtgggtg atgctctaac agtgacctgg aattttcgtc ctctagacgg  
 gggacctgag  
 361 cagtttgtat tctactacca catagatccc ttccaacca tgagtgggcg  
 25 gtttaaggac  
 421 cgggtgtctt gggatgggaa tcctgagcgg tacgatgcct ccatccttct  
 ctggaaactg  
 481 cagttcgacg acaatgggac atacacctgc caggtgaaga acccacctga  
 tgttgatggg  
 30 541 gtgatagggg agatccggct cagcgtcgtg cacactgtac gcttctctga  
 gatccacttc  
 601 ctggctctgg ccattggctc tgcoctgtgca ctgatgatca taatagtaat  
 tgtagtggtc  
 661 ctcttccagc attaccggaa aaagcgatgg gccgaaagag ctcataaagt  
 35 ggtggagata  
 721 aatcaaaaag aagaggaaag gctcaaccaa gagaaaaagg tctctgttta  
 tttagaagac  
 781 acagactaac aattttagat ggaagctgag atgatttcca agaacaagaa  
 ccctagtatt  
 40 841 tcttgaagtt aatggaaact tttctttggc ttttccagtt gtgaccggtt  
 ttccaaccag  
 901 ttctgcagca tattagattc tagacaagca acaccctct ggagccagca  
 cagtgtcct  
 961 ccatatcacc agtcatacac agcctcatta ttaaggtctt atttaatttc  
 45 agagtgtaaa  
 1021 ttttttcaag tgctcattag gttttataaa caagaagcta catttttggc  
 ctttaagacac  
 1081 tacttacagt gttatgactt gtatacacat atattggtat caaaagggat  
 aaaagccaat

1141 ttgtctgtta catttccttt cacgtatttc ttttagcagc acttctgcta  
 ctaaagttaa  
 1201 tgtgtttact ctctttcctt cccacattct caattaaaag gtgagctaag  
 cctcctcggg  
 5 1261 gtttctgatt aacagtaaat cctaaattca aactgttaaa tgacattttt  
 atttttatgt  
 1321 ctctccttaa ctatgagaca catcttgttt tactgaattt ctttcaatat  
 tccaggtgat  
 1381 agatTTTTgt tgTTTTgtta attaatccaa gatttacaat agcacaacgc  
 10 taaatcacac  
 1441 agtaactaca aaaggttaca tagatatgaa aagattggca gaggccattg  
 caggatgaat  
 1501 cacttgtcac ttttctctg tgctgggaaa aataatcaac aatgtgggtc  
 tttcatgagc  
 15 1561 agtgacggat agtttagctt actatgtttc cccccaatt caatgatcta  
 taacaacaga  
 1621 gcaaagtcta tgctcatttg cagactggaa tcattaagta atttaataaa  
 aaaattgtga  
 1681 aacagcatat tacaagtttg aaaattcagg gctggtgaaa aaaatcaact  
 20 ctaaattgatg  
 1741 ataattttgt acagttttat ataaaactct gagaactaga agaaattatt  
 aactttttt  
 1801 cttttttaat tctaattcac ttgtttattt tgggggagga agactttggg  
 atggagcaaa  
 25 1861 gaaataccaa aactacttta aatggaataa aaccaacttt attctttttt  
 tccccatac  
 1921 tggtagataa agcaaacttt ataagtgggc tattgaaaga aaagttacaa  
 gcttaagata  
 1981 cagaagcatt tgttcaaagg atagaaagca tctaaaagt taggctcaag  
 30 atcaatcttt  
 2041 acagattgat attttcagtt tttaatcgac tggactgcag atgttttttc  
 ttttaacaaa  
 2101 ctggaatttt caaacagatt atctgtattt aatgtatag acctgatat  
 tttccaata  
 35 2161 ctatttttta aaaaattgta tgatttacat atgaacctca gttctgaaat  
 tcattacata  
 2221 tctgtctcat tctgcctttt atactgtcta aaaaagcaaa gttttaaagt  
 gcaattttta  
 2281 aactgtaaat tacatctgaa ggctatatat ctttaataca cattttatat  
 40 ttttcttca  
 2341 caattctaac ctttgaaaat attataactg gatatttctt caaacagatg  
 tcctggatga  
 2401 tgggccataa gaataatgaa gaagtagtta aaaatgtatg gacagttttt  
 ccggcaaaat  
 45 2461 ttgtagctta tgtcttggct aaatagtcaa ggggtaatat gggcctgttg  
 tttagtgtct  
 2521 ccttcctaaa gagcactttt gtattgtaat ttatttttta ttatgcttta  
 aacactatgt  
 2581 aaataaacct ttagtaataa agaattatca gttataaaaa

50

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
EPH receptor A2	EPHA2	NM_004431	27

```

1  attaaggact cggggcagga ggggcagaag ttgcgcgcag gccggcgggc
   gggagcggac
5  61  accgaggccg gcgtgcaggc gtgcgggtgt gcgggagccg ggctcggggg
   gatcggaccg
   121 agagcgagaa gcgcggcatg gagctccagg cagcccgcgc ctgcttcgcc
   ctgctgtggg
   181 gctgtgcgct ggccgcggcc gcggcggcgc agggcaagga agtgggtactg
10  ctggactttg
   241 ctgcagctgg aggggagctc ggctggctca cacaccgta tggcaaaggg
   tgggacctga
   301 tgcagaacat catgaatgac atgccgatct acatgtactc cgtgtgcaac
   gtgatgtctg
15  361 gcgaccagga caactggctc cgcaccaact ggggtgtaccg aggagaggct
   gagcgtatct
   421 tcattgagct caagtttact gtacgtgact gcaacagctt ccctgggtggc
   gccagctcct
   481 gcaaggagac tttcaacctc tactatgccg agtcggacct ggactacggc
20  accaacttcc
   541 agaagcgcct gttaccaag attgacacca ttgcgcccga tgagatcacc
   gtcagcagcg
   601 acttcgaggc acgccacgtg aagctgaacg tggaggagcg ctccgtgggg
   ccgctcacc
25  661 gcaaaggctt ctacctggcc ttccaggata tcggtgctg tgtggcgctg
   ctctccgtcc
   721 gtgttacta caagaagtgc cccgagctgc tgcagggcct ggcccacttc
   cctgagacca
   781 tcgcccgtc tgatgcacct tccctggcca ctgtggccgg cacctgtgtg
30  gaccatgccg
   841 tggtgccacc ggggggtgaa gagccccgta tgcactgtgc agtggatggc
   gaggggctgg
   901 tgcccattgg gcagtgcctg tgccaggcag gctacgagaa ggtggaggat
   gcctgccagg
35  961 cctgctcgcc tggatTTTTT aagtttgagg catctgagag ccctgcttg
   gagtgcctg
   1021 agcacacgct gccatcccct gagggtgcca cctcctgca gtgtgaggaa
   ggcttcttcc
   1081 gggcacctca ggaccacgcg tcgatgcctt gcacacgacc cccctccgcc
40  ccacactacc
   1141 tcacagcct gggcatgggt gccaagggtg agctgcgctg gacgccccct
   caggacagcg
   1201 ggggccgca ggacattgtc tacagcgtca cctgcgaaca gtgctggccc
   gagtctgggg
45  1261 aatgcgggcc gtgtgaggcc agtgtgcgct actcggagcc tcctcacgga
   ctgaccgca
   1321 ccagtgtgac agtgagcgac ctggagcccc acatgaacta caccttcacc
   gtggaggccc
   1381 gcaatggcgt ctcaggcctg gtaaccagcc gcagcttccg tactgccagt
50  gtcagcatca

```

1441 accagacaga gcccccaag gtgaggctgg agggccgcag caccacctcg  
 cttagcgtct  
 1501 cctggagcat cccccgccg cagcagagcc gagtgtggaa gtacgaggtc  
 acttaccgca  
 5 1561 agaagggaga ctccaacagc tacaatgtgc gccgcaccga gggtttctcc  
 gtgaccctgg  
 1621 acgacctggc cccagacacc acctacctgg tccagggtgca ggcactgacg  
 caggagggcc  
 1681 agggggcccg cagcaagggtg cacgaattcc agacgctgtc cccggagggga  
 10 tctggcaact  
 1741 tggcgggtgat tggcggcgtg gctgtcggtg tggctctgct tctgggtgctg  
 gcaggagtgtg  
 1801 gcttctttat ccaccgcagg aggaagaacc agcgtgcccc ccagtccccg  
 gaggacgttt  
 15 1861 acttctccaa gtcagaacaa ctgaagcccc tgaagacata cgtggacccc  
 cacacatatg  
 1921 aggaccccaa ccaggctgtg ttgaagtca ctaccgagat ccatccatcc  
 tgtgtcactc  
 1981 ggcagaaggt gatcggagca ggagagttag gggagggtga caagggcatg  
 20 ctgaagacat  
 2041 cctcggggaa gaaggaggtg ccggtggcca tcaagacgct gaaagccggc  
 tacacagaga  
 2101 agcagcgagt ggacttcctc ggcgaggccg geatcatggg ccagttcagc  
 caccacaaca  
 25 2161 tcatccgcct agagggcgctc atctccaaat acaagcccat gatgatcatc  
 actgagtaca  
 2221 tggagaatgg ggccctggac aagttccttc gggagaagga tggcagagttc  
 agcgtgctgc  
 2281 agctgggtggg catgctgcgg ggcacgcag ctggcatgaa gtacctggcc  
 30 aacatgaact  
 2341 atgtgcaccg tgacctggct gcccgaaca tcctcgtcaa cagcaacctg  
 gtctgcaagg  
 2401 tgtctgactt tggcctgtcc cgcgtgctgg aggacgacc ccaggccacc  
 tacaccacca  
 35 2461 gtggcggcaa gatccccatc cgctggaccg ccccgaggc catttctac  
 cggaaagtca  
 2521 cctctgccag cgacgtgtgg agctttggca ttgtcatgtg ggagggtgatg  
 acctatggcg  
 2581 agcggcccta ctgggagttg tccaaccacg aggtgatgaa agccatcaat  
 40 gatggcttcc  
 2641 ggctccccac acctatggac tgcccctccg ccatctacca gctcatgatg  
 cagtgtggc  
 2701 agcaggagcg tgcccgcgc cccaagttag ctgacatcgt cagcatcctg  
 gacaagctca  
 45 2761 ttcgtgcccc tgactccctc aagaccctgg ctgactttga cccccgcgtg  
 tctatccggc  
 2821 tccccagcac gagcggctcg gaggggtgct ccttccgcac ggtgtccgag  
 tggctggagt  
 2881 ccatcaagat gcagcagtat acggagcact tcatggcggc cggctacact  
 50 gccatcgaga  
 2941 aggtggtgca gatgaccaac gacgacatca agaggattgg ggtgaggctg  
 cccggccacc  
 3001 agaagcgcac cgctacagc ctgctgggac tcaaggacca ggtgaacact  
 gtgggatcc

3061 ccatctgagc ctcgacaggg cctggagccc catcggccaa gaatacttga  
 agaaacagag  
 3121 tggcctccct gctgtgccat gctggggccac tggggacttt atttatttct  
 agttctttcc  
 5 3181 tccccctgca acttccgctg aggggtctcg gatgacacc tggcctgaac  
 tgaggagatg  
 3241 accagggatg ctgggctggg ccctctttcc ctgcgagacg cacacagctg  
 agcacttagc  
 3301 aggcaaccgc acgtcccagc atccctggag caggagcccc gccacagcct  
 10 tccgacagac  
 3361 atataggata ttcccaagcc gaccttcct cgccttctc ccacatgagg  
 ccatctcagg  
 3421 agatggaggg cttggcccag cgccaagtaa acagggtacc tcaagcccca  
 tttctcaca  
 15 3481 ctaagagggc agactgtgaa cttgactggg tgagacccaa agcggtcctt  
 gtccctctag  
 3541 tgccttcttt agaccctcg gcccctcct catccctgac tggccaaacc  
 cttgctttcc  
 3601 tgggcctttg caagatgctt ggttgtgttg aggtttttaa atatatatt  
 20 tgtactttgt  
 3661 ggagagaaatg tgtgtgtgtg gcagggggcc ccgccagggc tggggacaga  
 ggggtgcaaa  
 3721 cattcgtgag ctggggactc agggaccggt gctgcaggag tgtcctgccc  
 atgccccagt  
 25 3781 cggccccatc tctcctcctt ttggataagt ttctattctg tcagtgttaa  
 agattttgtt  
 3841 ttgttgga ttttttcga atcttaattt attatatttt ttatatttat  
 tgttagaaaa  
 3901 tgacttattt ctgctctgga ataaagttgc agatgattca aaccgaaaaa

30

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>integrin, alpha 6</b>	ITGA6	NM_000210	28

1 aacgggctca ttcagcggtc gcgagctgcc cgcgaggggg agcggccgga  
 cggagagcgc  
 61 gacccgtccc ggggggtggg ccgggcgcag cggcgagagg aggcgaaggt  
 35 ggctgcggta  
 121 gcagcagcgc ggcagcctcg gacccagccc ggagcgcagg gcggccgctg  
 caggtccccg  
 181 ctccccctccc cgtgcgtccg cccatggccg ccgccgggca gctgtgcttg  
 ctctacctgt  
 40 241 cggcggggct cctgtcccgg ctcggcgcag ccttcaactt ggacactcgg  
 gaggacaacg  
 301 tgatccgga atatggagac cccgggagcc tcttcggctt ctgctggcc  
 atgactggc  
 361 aactgcagcc cgaggacaag cggctgttgc tcgtgggggc cccgcgggca  
 45 gaagcgcttc  
 421 cactgcagag agccaacaga acgggagggc tgtacagctg cgacatcacc  
 gcccgggggc  
 481 catgcacgcg gatcgagttt gataacgatg ctgacccac gtcagaaagc  
 aaggaagatc

541 agtggatggg ggtcaccgtc cagagccaag gtccaggggg caaggctcgtg  
 acatgtgctc  
 601 accgatatga aaaaaggcag catgttaata cgaagcagga atccccgagac  
 atctttgggc  
 5 661 ggtgttatgt cctgagtcag aatctcagga ttgaagacga tatggatggg  
 ggagattgga  
 721 gcttttgtga tgggcgattg agaggccatg agaaatttgg ctcttgccag  
 caaggtgtag  
 781 cagctacttt tactaaagac tttcattaca ttgtatttgg agccccgggt  
 10 acttataact  
 841 ggaaagggat tgttcgtgta gagcaaaaga ataacacttt ttttgacatg  
 aacatctttg  
 901 aagatgggcc ttatgaagt ggtggagaga ctgagcatga tgaaagtctc  
 gttcctgttc  
 15 961 ctgctaacag ttacttaggt ttttctttgg actcagggaa aggtattggt  
 tctaaagatg  
 1021 agatcacttt tgtatctggt gctcccagag ccaatcacag tggagccgtg  
 gttttgctga  
 1081 agagagacat gaagtctgca catctcctcc ctgagcacat attcgatgga  
 20 gaaggtctgg  
 1141 cctcttcatt tggctatgat gtggcggtgg tggacctcaa caaggatggg  
 tggcaagata  
 1201 tagttattgg agccccacag tattttgata gagatggaga agttggaggt  
 gcagtgtatg  
 25 1261 tctacatgaa ccagcaaggc agatggaata atgtgaagcc aattcgtctt  
 aatggaacca  
 1321 aagattctat gtttggcatt gcagtaaaaa atattggaga tattaatcaa  
 gatggctacc  
 1381 cagatattgc agttggagct ccgatgatg acttgggaaa ggtttttatc  
 30 tatcatggat  
 1441 ctgcaaatgg aataaatacc aaaccaacac aggttctcaa gggatatatca  
 ccttattttg  
 1501 gatattcaat tgctggaaac atggaccttg atcgaaattc ctaccctgat  
 gttgctgttg  
 35 1561 gttccctctc agattcagta actattttca gatcccggcc tgtgattaat  
 attcagaaaa  
 1621 ccacacagt aactcctaac agaattgacc tccgccagaa aacagcgtgt  
 ggggcgccta  
 1681 gtgggatatg cctccagggt aaatcctggt ttgaatatac tgctaacccc  
 40 gctgggttata  
 1741 atccttcaat atcaattgtg ggcacacttg aagctgaaaa agaaagaaga  
 aatctgggc  
 1801 tatcctcaag agttcagttt cgaaaccaag gttctgagcc caaatatact  
 caagaactaa  
 45 1861 ctctgaagag gcagaaacag aaagtgtgca tggaggaaac cctgtggcta  
 caggataata  
 1921 tcagagataa actgcgtccc attcccataa ctgcctcagt ggagatccaa  
 gagccaagct  
 1981 ctcgtaggag agtgaattca cttccagaag ttcttccaat tctgaattca  
 50 gatgaacca  
 2041 agacagctca tattgatggt cacttcttaa aagagggatg tggagacgac  
 aatgtatgta  
 2101 acagcaacct taaactagaa tataaat ttt gcacccgaga aggaaatcaa  
 gacaaat ttt

2161 cttatttacc aattcaaaaa ggtgtaccag aactagttct aaaagatcag  
 aaggatattg  
 2221 ctttagaaat aacagtgaca aacagccctt ccaaccaag gaatcccaca  
 aaagatggcg  
 5 2281 atgacgccca tgaggctaaa ctgattgcaa cgtttccaga cactttaacc  
 tattctgcat  
 2341 atagagaact gagggctttc cctgagaaac agttgagttg tgttgccaac  
 cagaatggct  
 2401 cgcaagctga ctgtgagctc ggaaatcctt ttaaaagaaa ttcaaagtgc  
 10 actttttatt  
 2461 tggttttaag tacaactgaa gtcacctttg acaccccaga tctggatatt  
 aatctgaagt  
 2521 tagaaacaac aagcaatcaa gataatttgg ctccaattac agctaaagca  
 aaagtggtta  
 15 2581 ttgaactgct tttatcggtc tcgggagttg ctaaaccctc ccaggtgtat  
 tttggaggta  
 2641 cagttgttgg cgagcaagct atgaaatctg aagatgaagt gggagttta  
 atagagtatg  
 2701 aattcagggt aataaactta ggtaaacctc ttacaaacct eggcacagca  
 20 accttgaaca  
 2761 ttcagtggcc aaaagaaatt agcaatggga aatggttgct ttatttggtg  
 aaagtagaat  
 2821 ccaaaggatt ggaaaaggta acttgtgagc cacaaaagga gataaactcc  
 ctgaacctaa  
 25 2881 cggagtctca caactcaaga aagaaacggg aaattactga aaaacagata  
 gatgataaca  
 2941 gaaaattttc tttatttgc t gaaagaaaat accagactct taactgtagc  
 gtgaacgtga  
 3001 actgtgtgaa catcagatgc ccgctgcggg ggctggacag caaggcgtct  
 30 cttattttgc  
 3061 gctcgagggt atggaacagc acatttctag aggaatattc caaactgaac  
 tacttggaca  
 3121 ttctcatgcg agccttcatt gatgtgactg ctgctgccga aaatatcagg  
 ctgccaaatg  
 35 3181 caggcactca ggttcgagtg actgtgtttc cctcaaagac tgtagctcag  
 tattcgggag  
 3241 taccttggtg gatcatccta gtggctattc tcgctgggat cttgatgctt  
 gctttattag  
 3301 tgtttatact atggaagtgt ggtttcttca agagaaataa gaaagatcat  
 40 tatgatgcca  
 3361 catatcacia ggctgagatc catgctcagc catctgataa agagaggctt  
 acttctgatg  
 3421 catagtattg atctacttct gtaattgtgt ggattcttta aacgctctag  
 gtacgatgac  
 45 3481 agtgttccc gataccatgc tgtaaggatc cggaaagaag agcgagagat  
 caaagatgaa  
 3541 aagtatattg ataaccttga aaaaaaacag tggatcacia agtggaacga  
 aatgaaagc  
 3601 tactcatagc gggggcctaa aaaaaaaaaag cttcacagta cccaaactgc  
 50 tttttccaac  
 3661 tcagaaattc aatttggatt taaaagcctg ctcaatcctt gaggactgat  
 ttcagagtga  
 3721 ctacacacag tacgaaccta cagttttaac tgtggatatt gttacgtagc  
 ctaaggctcc

3781 tgttttgcac agccaaattt aaaactgttg gaatggattt ttctttaact  
 gccgtaattt  
 3841 aactttctgg gttgccttta tttttggcgt ggctgactta catcatgtgt  
 tggggaagg  
 5 3901 cctgcccagt tgcactcagg tgacatcctc cagatagtgt agctgaggag  
 gcacctacac  
 3961 tcacctgcac taacagagtg gccgtcctaa cctcgggcct gctgcgcaga  
 cgtccatcac  
 4021 gttagctgtc ccacatcaca agactatgcc attggggtag ttgtgtttca  
 10 acggaaagtg  
 4081 ctgtcttaaa ctaaagtgtc aatagaaggat gatgttgcca tctaccgtc  
 ttttctgtt  
 4141 tcttagctgt gtgaatacct gctcacgtca aatgcataca agtttcattc  
 tccctttcac  
 15 4201 taaaacacac aggtgcaaca gacttgaatg ctagttatac ttatttgtat  
 atggtattta  
 4261 ttttttcttt tctttacaaa ccattttggt attgactaac aggccaaaga  
 gtctccagtt  
 4321 tacccttcag gttggtttaa tcaatcagaa ttagagcatg ggaggtcatc  
 20 actttgacct  
 4381 aaattattta ctgcaaaaag aaaatcttta taaatgtacc agagagagtt  
 gttttaataa  
 4441 cttatctata aactataacc tctccttcat gacagcctcc accccacaac  
 ccaaaagggt  
 25 4501 taagaaatag aattataact gtaaagatgt ttatttcagg cattggatat  
 ttttacttt  
 4561 agaagcctgc ataatgtttc tggatttcat actgtaacat tcaggaattc  
 ttggagaaaa  
 4621 tgggtttatt cactgaactc tagtgccggt tactcactgc tgcaaatact  
 30 gtatattcag  
 4681 gacttgaaag aaatggtgaa tgcctatggt ggatccaaac tgatccagta  
 taagactact  
 4741 gaatctgcta ccaaaacagt taatcagtga gtcgatgttc tattttttgt  
 tttgtttcct  
 35 4801 cccctatctg tattcccaaa aattactttg gggctaattt aacaagaact  
 ttaaattgtg  
 4861 ttttaattgt aaaaatggca gggggtggaa ttattactct atacattcaa  
 cagagactga  
 4921 atagatatga aagctgattt tttttaatta ccatgcttca caatgttaag  
 40 ttatatgggg  
 4981 agcaacagca aacaggtgct aatttgtttt ggatatagta taagcagtgt  
 ctgtgttttg  
 5041 aaagaataga acacagtttg tagtgccact gttgttttg gggggctttt  
 ttcttttcgg  
 45 5101 aaatcttaaa ccttaagata ctaaggacgt tgttttggtt gtactttgga  
 attcttagtc  
 5161 acaaaatata ttttgtttac aaaaatttct gtaaaacagg ttataacagt  
 gtttaaagtc  
 5221 tcagtttctt gcttggggaa cttgtgtccc taatgtgttt agattgctag  
 50 attgctaagg  
 5281 agctgatact ttgacagtgt ttttagacct gtgttactaa aaaaaagatg  
 aatgtcctga  
 5341 aaagggtggt gggagggtgg ttcaacaaag aaacaaagat gttatggtgt  
 ttagatttat

5401 gggttgtaaa aatgtcatct caagtcaagt cactggctctg tttgcatttg  
 atacattttt  
 5461 gtactaacta gcattgtaaa attatttcat gattagaaat tacctgtgga  
 tatttgata  
 5 5521 aaagtgtgaa ataaattttt tataaaagtg ttcattgttt cgtaacacag  
 cattgtatat  
 5581 gtgaagcaaa ctctaaaatt ataaatgaca acctgaatta tctatttcat  
 caaaccaaag  
 5641 ttcagtgttt ttatttttgg tgtctcatgt aatctcagat cagccaaaga  
 10 tactagtgcc  
 5701 aaagcaatgg gattcggggg ttttttctgt tttcgtctta tgtaggtgat  
 cctcaagtct  
 5761 ttcattttcc ttctttatga ttaaaagaaa cctacaggta ttaacaacc

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
tumor necrosis factor receptor superfamily, member 21	TNFRSF21	NM_014452	29

15  
 1 gccaccacgt gtgtccctgc gcccggtggc caccgactca gtcacctgcc  
 gaccagtctg  
 61 ggcagcggag gagggtgggt ggcagtggt ggaagcttcg ctatgggaag  
 ttgttccttt  
 20 121 gctctctcgc gccagtcct cctccctggg tctcctcagc cgctgtcgga  
 ggagagcacc  
 181 cggagacgcg ggctgcagtc gcggcggctt ctccccgctt gggcggccgc  
 gccgctgggc  
 241 aggtgctgag cgcccctaga gcctcccttg ccgctcctt cctctgcccg  
 25 gccgcagcag  
 301 tgcacatggg gtgttgagg tagatgggt cccggcccgg gaggcggcgg  
 tggatgcggc  
 361 gctgggcaga agcagccgcc gattccagct gccccgcgc cccggggcgc  
 cctgcagagt  
 421 ccccggttca gccatgggga cctctccgag cagcagcacc gccctgcct  
 30 cctgcagccg  
 481 catgcccgc cgagccacag ccaogatgat cgcgggctcc cttctctgc  
 ttgattcct  
 541 tagcaccacc acagctcagc cagaacagaa ggctcgaat ctcatggca  
 35 cataccgcca  
 601 tgttgaccgt gccaccggcc aggtgctaac ctgtgacaag tgtccagcag  
 gaacctatgt  
 661 ctctgagcat tgtaccaaca caagcctgcg cgtctgcagc agttgcctg  
 tggggacctt  
 721 taccaggcat gagaatggca tagagaaatg ccatgactgt agtcagccat  
 40 gcccatggcc  
 781 aatgattgag aaattacctt gtgctgcctt gactgaccga gaatgcactt  
 gccacctgg  
 841 catgttccag tctaacgcta cctgtgcccc ccatacgggtg tgtcctgtgg  
 45 gttgggggtg  
 901 gcggaagaaa gggacagaga ctgaggatgt gcggtgtaag cagtgtgctc  
 ggggtacctt  
 961 ctcatgctg ccttctagtg tgatgaaatg caaagcatac acagactgtc  
 tgagtcagaa

1021 cctggtggtg atcaagccgg ggaccaagga gacagacaac gtctgtggca  
 cactcccgtc  
 1081 cttctccagc tccacctcac cttcccctgg cacagccatc tttccacgcc  
 ctgagcacat  
 5 1141 ggaaacccat gaagtcctt cctccactta tgttcccaa ggcatgaact  
 caacagaatc  
 1201 caactcttct gcctctgtta gaccaaaggt actgagtagc atccaggaag  
 ggacagtccc  
 1261 tgacaacaca agctcagcaa gggggaagga agacgtgaac aagaccctcc  
 10 caaaccttca  
 1321 ggtagtcaac caccagcaag gccccaccca cagacacatc ctgaagctgc  
 tgccgtccat  
 1381 ggaggccact gggggcgaga agtccagcac gccatcaag ggccccaga  
 ggggacatcc  
 15 1441 tagacagaac ctacacaagc attttgacat caatgagcat ttgccctgga  
 tgattgtgct  
 1501 tttctgctg ctggtgcttg tgggtgattgt ggtgtgcagt atccggaaaa  
 gctcgaggac  
 1561 tctgaaaaag gggccccggc aggatcccag tgccattgtg gaaaaggcag  
 20 ggctgaagaa  
 1621 atccatgact ccaaccaga accgggagaa atggatctac tactgcaatg  
 gccatggtat  
 1681 cgatatcctg aagctttag cagcccaagt ggaagccag tggaaagata  
 tctatcagtt  
 25 1741 tctttgcaat gccagtgaga gggagggtgc tgctttctcc aatgggtaca  
 cagccgacca  
 1801 cgagcgggcc tacgcagctc tgcagcactg gaccatccgg ggccccgagg  
 ccagcctcgc  
 1861 ccagctaatt agcgcctgc gccagcaccg gagaaacgat gttgtggaga  
 30 agattcgtgg  
 1921 gctgatggaa gacaccacc agctggaaac tgacaaacta gctctcccga  
 tgagccccag  
 1981 cccgcttagc ccgagcccca tccccagccc caacgcgaaa cttgagaatt  
 ccgctctcct  
 35 2041 gacggtggag cttccccac aggacaagaa caagggttc ttcgtggatg  
 agtcggagcc  
 2101 cttctccgc tgtgactcta catccagcgg ctctcccg ctgagcagga  
 acggttcctt  
 2161 tattacaaa gaaaagaagg acacagtgtt gcggcaggta cgctggacc  
 40 cctgtgactt  
 2221 gcagcctatc tttgatgaca tgctccactt tctaaatcct gaggagctgc  
 ggggtgattga  
 2281 agagattccc caggctgagg acaaactaga cggctatc gaaattattg  
 gagtcaagag  
 45 2341 ccaggaagcc agccagacc tcctggactc tgtttatagc catcttctg  
 acctgctgta  
 2401 gaacatagg atactgcatt ctggaaatta ctcaatttag tggcagggtg  
 gttttttaat  
 2461 tttctctgt ttctgatttt tgttgtttgg ggtgtgtgtg tgtgtttgtg  
 50 tgtgtgtgtg  
 2521 tgtgtgtgtg tgtgtgtgtg ttaacagag aatatggcca gtgcttgagt  
 tctttctct  
 2581 tctctctctc tctttttttt ttaaataact cttctgggaa gttggtttat  
 aagcctttgc

2641 caggtgtaac tgttgtaaa taccaccac taaagtttt taagttccat  
 attttctcca  
 2701 ttttgcttc ttatgtatt tcaagattat tctgtgcact ttaaatttac  
 ttaacttacc  
 5 2761 ataaatgcag tgtgactttt cccacacact ggattgtgag gctcttaact  
 tcttaaaagt  
 2821 ataatggcat cttgtgaatc ctataagcag tctttatgtc tcttaacatt  
 cacacctact  
 2881 ttttaaaaac aatattatt actattttta ttattgtttg tcctttataa  
 10 attttcttaa  
 2941 agattaagaa aatttaagac cccattgagt tactgtaatg caattcaact  
 ttgagttatc  
 3001 ttttaaatat gtcttgata gttcatattc atggctgaaa cttgaccaca  
 ctattgctga  
 15 3061 ttgtatgggt ttcacctgga caccgtgtag aatgcttgat tacttgtact  
 cttcttatgc  
 3121 taatatgtc tgggctggag aatgaaatc ctcaagccat caggatttgc  
 tatttaagtg  
 3181 gcttgacaac tgggccacca aagaactga acttcacctt ttaggatttg  
 20 agctgttctg  
 3241 gaacacattg ctgcactttg gaaagtcaaa atcaagtgcc agtggcgccc  
 tttccataga  
 3301 gaatttgccc agctttgctt taaaagatgt cttgtttttt atatacacat  
 aatcaatagg  
 25 3361 tccaatctgc tctcaaggcc ttggctctgg tgggatctct tcaccaatta  
 ctttaattaa  
 3421 aatggctgc aactgtaaga acccttgtct gatatatttg caactatgct  
 cccatttaca  
 3481 aatgtacctt ctaatgtca gttgccaggt tccaatgcaa aggtggcgtg  
 30 gactcccttt  
 3541 gtgtgggtgg ggtttgtggg tagtggtgaa ggaccgatat cagaaaaatg  
 cttcaagtg  
 3601 tactaattta ttaataaaca ttaggtgttt gttaaaaaaa

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>transmembrane 4 superfamily member 3</b>	TM4SF3	NM_004616	30

35 1 agtgccccag gagctatgac aagcaaagga acatacttgc ctggagatag  
 cctttgcat  
 61 atttaaatgt ccgtggatac agaaatctct gcaggcaagt tgctccagag  
 catattgcag  
 121 gacaagcctg taacgaatag ttaaattcac ggcactgga ttcctaatac  
 40 tttccgaaa  
 181 tggcaggtgt gagtgctgt ataaaatatt ctatgtttac cttcaacttc  
 ttgttctggc  
 241 tatgtggtat cttgatccta gcattagcaa tatgggtacg agtaagcaat  
 gactctcaag  
 45 301 caatttttgg ttctgaagat gtaggctcta gtcctacgt tgctgtggac  
 atattgattg  
 361 ctgtaggtgc catcatcatg attctgggct tctgggatg ctgcggtgct  
 ataaaagaaa  
 421 gtcgctgcat gcttctgttg tttttcatag gcttgcttct gatcctgctc  
 50 ctgcaggtgg

481 cgacaggtat cctaggagct gttttcaaat ctaagtctga tcgcattgtg  
aatgaaactc  
541 tctatgaaaa cacaaagctt ttgagcgcca caggggaaag tgaaaaacaa  
ttccaggaag  
5 601 ccataattgt gtttcaagaa gagtttaaat gctgcggtt ggtcaatgga  
gctgctgatt  
661 ggggaaataa ttttcaacac tatcctgaat tatgtgctg tctagataag  
cagagaccat  
721 gccaaagcta taatggaaaa caagtttaca aagagacctg tatttctttc  
10 ataaaagact  
781 tcttggcaaa aaatttgatt atagttattg gaatatcatt tggactggca  
gttattgaga  
841 tactgggtt ggtgttttct atggtcctgt attgccagat cggaacaaa  
tgaatctgtg  
15 901 gatgcatcaa cctatcgtca gtcaaaccct ttaaaatgt tgctttggct  
ttgtaaat  
961 aaatatgtaa gtgctatata agtcaggagc agctgtcttt ttaaaatgtc  
tcggctagct  
1021 agaccacaga tatcttctag acatattgaa cacatttaag atttgagggg  
20 tataagggaa  
1081 aatgatatga atgtgtattt ttactcaaaa taaaagtaac tgtttacggt

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>interleukin 18 (interferon-gamma-inducing factor)</b>	IL18	NM_001562	31

25 1 attctctccc cagcttgctg agccctttgc tcccctggcg actgcctgga  
cagtcagcaa  
61 ggaattgtct eccagtgcatt tttgccctcc tggctgcca ctctggctgc  
taaagcggct  
121 gccacctgct gcagtctaca cagcttcggg aagaggaaag gaacctcaga  
30 ccttccagat  
181 cgcttctct cgcaacaaac tatttgtcgc aggaataaag atggctgctg  
aaccagtaga  
241 agacaattgc atcaactttg tggcaatgaa atttattgac aatacgcttt  
actttatagc  
35 301 tgaagatgat gaaaacctgg aatcagatta ctttggcaag cttgaatcta  
aattatcagt  
361 cataagaaat ttgaatgacc aagttctctt cattgacca ggaaatcggc  
ctctatttga  
421 agatatgact gattctgact gtagagataa tgcaccccg accatattta  
40 ttataagtat  
481 gtataaagat agccagccta gaggtatggc tgtaactatc tctgtgaagt  
gtgagaaaat  
541 ttcaactctc tcctgtgaga acaaaattat tcctttaag gaaatgaatc  
ctcctgataa  
45 601 catcaaggat acaaaaagtg acatcatatt ctttcagaga agtgtcccag  
gacatgataa  
661 taagatgcaa tttgaatctt catcatacga aggatacttt ctagcttgtg  
aaaaagagag

721 agaccttttt aaactcattt tgaaaaaaga ggatgaattg ggggatagat  
 ctataatgtt  
 781 cactgttcaa aacgaagact agctattaa atttcatgcc gggcgcagtg  
 gctcacgcct  
 5 841 gtaatcccag ccctttggga ggctgaggcg ggcagatcac cagaggtcag  
 gtgttcaaga  
 901 ccagcctgac caacatggtg aaacctcatc tctactaaaa atacaaaaaa  
 ttagctgagt  
 961 gtagtgacgc atgccctcaa tcccagctac tcaagaggct gaggcaggag  
 10 aatcacttgc  
 1021 actccggagg tagaggttgt ggtgagccga gattgcacca ttgcgctcta  
 gcctgggcaa  
 1081 caacagcaaa actccatctc aaaaaataaa ataaataaat aaacaaataa  
 aaaattcata  
 15 1141 atgtg

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>bone morphogenetic protein 4</b>	<b>BMP4</b>	<b>NM_130850</b>	<b>32</b>

1 gagggagggg ccgccgggga agaggaggag gaaggaaaga aagaaagcga  
 20 gggagggaaa  
 61 gaggaggaag gaagatgcga gaaggcagag gaggagggag ggaggggaagg  
 agcgcggagc  
 121 ccggcccgga agctaggagc cattccgtag tgccatcccg agcaacgcac  
 tgctgcagct  
 25 181 tccctgagcc tttccagcaa gtttgttcaa gattggctgt caagaatcat  
 ggactgttat  
 241 tatatgcctt gttttctgtc aagacaccat gattcctggt aaccgaatgc  
 tgatggtcgt  
 301 tttattatgc caagtcctgc taggaggcgc gagccatgct agtttgatac  
 30 ctgagacggg  
 361 gaagaaaaaa gtcgccgaga ttcagggccca cgcgggagga cgcgcctcag  
 ggcagagcca  
 421 tgagctcctg cgggacttcg aggcgacact tctgcagatg tttgggctgc  
 gcgcgccccc  
 35 481 gcagcctagc aagagtgccg tcattccgga ctacatgcgg gatctttacc  
 ggcttcagtc  
 541 tggggaggag gaggaagagc agatccacag cactggtcct gagtatcctg  
 agcgcgccgc  
 601 cagccggggcc aacaccgtga ggagcttcca ccacgaagaa catctggaga  
 40 acatcccagg  
 661 gaccagtga aactctgctt ttcgtttctt cttaacctc agcagcatcc  
 ctgagaacga  
 721 ggcgatctcc tctgcagagc ttcggctctt ccgggagcag gtggaccagg  
 gccctgattg  
 45 781 ggaaaggggc ttcaccgta taaacattta tgaggttatg aagccccag  
 cagaagtgt  
 841 gcctgggcac ctcatcacac gactactgga cacgagactg gtccaccaca  
 atgtgacacg

901 gtgggaaact tttgatgtga gccctgctgg ccttcgctgg acccgggaga  
 agcagccaaa  
 961 ctatgggcta gccattgagg tgactcacct ccatcagact cggacccacc  
 agggccagca  
 5 1021 tgtcaggatt agccgatcgt tacctcaagg gagtgggaat tgggcccagc  
 tccggcccct  
 1081 cctggtcacc tttggccatg atggccgggg ccatgccttg acccgacgcc  
 ggagggccaa  
 1141 gcgtagccct aagcatcact cacagcgggc caggaagaag aataagaact  
 10 gccggcgcca  
 1201 ctcgctctat gtggacttca gcgatgtggg ctggaatgac tggattgtgg  
 ccccaccagg  
 1261 ctaccaggcc ttctactgcc atggggactg cccctttcca ctggetgacc  
 acctcaactc  
 15 1321 aaccaacat gccattgtgc agaccctggt caattctgtc aattccagta  
 tccccaaagc  
 1381 ctgttggtg cccactgaac tgagtgccat ctccatgctg tacctggatg  
 agtatgataa  
 1441 ggtggtactg aaaaattatc aggagatggt agtagagga tgtgggtgcc  
 20 gctgagatca  
 1501 ggcagtcctt gaggatagac agatatacac accacacaca cacaccacat  
 acaccacaca  
 1561 cacacgttcc catccactca cccacacact acacagactg cttccttata  
 gctggacttt  
 25 1621 tatttaaaaa aaaaaaaaaa aaaatggaaa aaatccctaa acattcacct  
 tgaccttatt  
 1681 tatgacttta cgtgcaaatg ttttgaccat attgatcata tattttgaca  
 aaatatattt  
 1741 ataactacgt attaaaagaa aaaaataaaa tgagtcatta ttttaaaggt  
 30

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>sphingomyelin phosphodiesterase, acid-like 3B</b>	<b>SMPDL3B</b>	<b>NM_014474</b>	<b>33</b>

1 ccagatcata cctgctggg caaaggagga agagccagag gatccagacg  
 35 ccttggagga  
 61 cttggaacac ctgtaacagg acaaggagtt ctgctcagge acgtggccac  
 agaaaactac  
 121 ttaggaagcc tgtggtgaga acaacaacag tgectgagaa tcccacggct  
 ctggggaagt  
 40 181 gagccccgag gatgaggctg ctgcctggc tgattttcct ggctaactgg  
 ggaggtgcca  
 241 gggctgaacc agggaagtte tggcacatcg ctgacctgca ccttgacct  
 gactacaagg  
 301 tatccaaaga ccccttccag gtgtgcccac cagctggatc ccagccagtg  
 45 cccgacgcag  
 361 gccctgggg tgactacctc tgtgattctc cctgggcccct catcaactcc  
 tccatctatg

421 ccatgaagga gattgagcca gagccagact tcattctctg gactggtgat  
gacacgcctc  
481 atgtgcccga tgagaaactg ggagaggcag ctgtactgga aattgtggaa  
cgctgacca  
5 541 agctcatcag agaggtcttt ccagatacta aagtctatgc tgctttggga  
aatcatgatt  
601 ttcaccccaa aaaccagttc ccagctggaa gtaacaacat ctacaatcag  
atagcagaac  
661 tatggaaacc ctggcttagt aatgagtcca tcgctctctt caaaaaaggt  
10 gccttctact  
721 gtgagaagct gccgggtccc agcggggctg ggcgaattgt ggtcctcaac  
accaatctgt  
781 actataccag caatgcgctg acagcagaca tggcggacce tggccagcag  
ttccagtggc  
15 841 tggaaagatg gctgaccgat gcatccaaag ctggggacat ggtgtacatt  
gtcggccacg  
901 tgcccccggg gttctttgag aagacgcaa acaaggcatg gttccgggag  
ggcttcaatg  
961 aaaaatacct gaaggtggtc cggaagcatc atcgcgtcat agcagggcag  
20 ttcttcgggc  
1021 accaccacac cgacagcttt cggatgctct atgatgatgc aggtgtcccc  
ataagcgcca  
1081 tgttcatcac acctggagtc accccatgga aaaccacatt acctggagtg  
gtcaatgggg  
25 1141 ccaacaatcc agccatccgg gtgttcgaat atgaccgagc cacactgagc  
ctgaaggaca  
1201 tggtgaccta ctcatgaac ctgagccagg cgaatgctca ggggacgccc  
cgctgggagc  
1261 tcgagtacca gctgaccgag gcctatgggg tgccggacgc cagcgcaccac  
30 tccatgcaca  
1321 cagtgctgga ccgcatcgtt ggcgaccaga gcacactgca gcgctactac  
gtctataact  
1381 cagtcagcta ctctgctggg gtctgcgacg aggcctgcag catgcagcac  
gtgtgtgcca  
35 1441 tgcgccaggt ggacattgac gcttacacca cctgtctgta tgctctggc  
accacgcccg  
1501 tgccccagct cccgctgctg ctgatggccc tgctgggcct gtgcacgctc  
gtgctgtgac  
1561 ctgccaggct caccttcttc ctggtaacgg gtaacggggg cagcgcaccag  
40 gatcaccag  
1621 agctgggcct tccaccattt cctccgcgcc tgaggagtga actgaaatag  
gacaaccgaa  
1681 tcaggaagcg aagccccagg agctgcagcc atccgtgatc gcgccactgc  
actccagcct  
45 1741 gggcgacaaa gccagactct ctccaaaaac aaaccagaaa cagaaaagaa  
atgacgaccc  
1801 aagaccccc tacaagcata cttcttttgc gtattatggt ttactcacia  
aacaagctc  
1861 atcatgctt tgaaaaaaaa

50

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
transmembrane protease, serine 2	TMPRSS2	NM_005656	34

```

1  cgcgagctaa gcaggaggcg gaggcggagg cggagggcga ggggcgggga
gcgccgcctg
5  61  gagcgcggca ggtcatattg aacattccag atacctatca ttactcgatg
ctgttgataa
121 cagcaagatg gctttgaact cagggtcacc accagctatt ggaccttact
atgaaaacca
181 tggataccia cggaaaacc cctatcccg cagcccact gtggteccca
ctgtctacga
10 241 ggtgcatccg gctcagtact acccgteccc cgtgccccag tacgccccga
gggtcctgac
301 gcaggcttcc aaccccgctg tctgcacgca gcccaaacc ccatccggga
cagtgtgcac
15 361 ctcaaagact aagaaagcac tgtgcatcac cttgaccctg gggaccttcc
tcgtgggagc
421 tgcgctggcc gctggcctac tctggaagtt catgggcagc aagtgtcca
actctgggat
481 agagtgcgac tcctcaggta cctgcatcaa ccctctaac tgggtgtgatg
gcgtgtcaca
20 541 ctgccccggc ggggaggacg agaatcgggtg tgttcgcctc tacggacca
acttcatcct
601 tcagatgtac tcatctcaga ggaagtcctg gcacctgtg tgccaagacg
actggaacga
25 661 gaactacggg cgggcggcct gcagggacat gggctataag aataatttt
actctagcca
721 aggaatagtg gatgacagcg gatccaccag ctttatgaaa ctgaacacaa
gtgccggcaa
781 tgtcgatata tataaaaaac tgtaccacag tgatgcctgt tcttcaaaag
cagtggtttc
30 841 tttacgctgt atagcctgcg gggtcactt gaactcaagc cgccagagca
ggatcgtggg
901 cggtgagagc gcgctcccgg gggcctggcc ctggcaggtc agcctgcacg
tccagaacgt
35 961 ccacgtgtgc ggaggtcca tcatcacc ctagtgatc gtgacagccg
cccactgct
1021 ggaaaaacct cttacaatc catggcattg gacggcattt gcggggattt
tgagacaatc
1081 tttcatgttc tatggagccg gataccaagt agaaaaagtg atttctcatc
caaattatga
40 1141 ctccaagacc aagaacaatg acattgctgt gatgaagctg cagaagcctc
tgactttcaa
1201 cgacctagtg aaaccagtgt gtctgccaa cccaggcatg atgctgcagc
cagaacagct
45 1261 ctgctggatt tccgggtggg gggccaccga ggagaaaggg aagacctcag
aagtgtgaa
1321 cgctgccaag gtgcttctca ttgagacaca gagatgcaac agcagatag
tctatgacaa

```

1381 cctgatcaca ccagccatga tctgtgccgg cttcctgcag gggaacgtcg  
 attccttgcca  
 1441 ggggtgacagt ggagggcctc tggtcacttc gaagaacaat atctgggtggc  
 tgatagggga  
 5 1501 tacaagctgg ggttctggct gtgccaaagc ttacagacca ggagtgtacg  
 ggaatgtgat  
 1561 ggtattcacg gactggattt atcgacaaat gagggcagac ggctaacca  
 catggtcttc  
 1621 gtccttgacg tcgttttaca agaaaacaat ggggctgggt ttgcttcccc  
 10 gtgcatgatt  
 1681 tactccttaga gatgattcag aggtcacttc atttttatta aacagtgaac  
 ttgtctggct  
 1741 ttggcactct ctgccattct gtgcaggctg cagtggctcc cctgcccagc  
 ctgctctccc  
 15 1801 taacccttg tccgcaaggg gtgatggccg gctggtttg ggcactggcg  
 gtcaagtgtg  
 1861 gaggagaggg gtggaggctg cccattgag atcttctgc tgagtccttt  
 ccaggggcca  
 1921 attttgatg agcatggagc tgtcacctct cagctgctgg atgacttgag  
 20 atgaaaaagg  
 1981 agagacatgg aaagggagac agccagggtg cacctgcagc ggctgcctc  
 tggggccact  
 2041 tggtagtgtc cccagcctac ctctccaca ggggattttg ctgatgggtt  
 cttagagcct  
 25 2101 tagcagcct ggatgggtggc cagaaataaa gggaccagcc cttcatgggt  
 ggtgacgtgg  
 2161 tagtcacttg taaggggaac agaaacattt ttgttcttat ggggtgagaa  
 tatagacagt  
 2221 gcccttggtg cgaggggaagc aattgaaaag gaacttgccc tgagcactcc  
 30 tgggtgcagg  
 2281 ctccacctgc acattgggtg gggctcctgg gagggagact cagccttct  
 cctcatcctc  
 2341 cctgaccctg ctctagcac cctggagagt gcacatgccc cttggtcctg  
 gcagggcgcc  
 35 2401 aagtctggca ccatggtggc ctcttcaggc ctgctagtca ctggaaattg  
 aggtccatgg  
 2461 gggaaatcaa ggatgctcag ttaaggtac actgtttcca tgttatgttt  
 ctacacattg  
 2521 ctacctcagt gtcctggaa acttagcttt tgatgtctcc aagtagtcca  
 40 ccttcattta  
 2581 actctttgaa actgtatcac ctttgccaag taagagtggg ggcctatctc  
 agctgctttg  
 2641 acaaaatgac tggctcctga cttaacgttc tataaatgaa tgtgctgaag  
 caaagtgccc  
 45 2701 atgggtggcg cgaagaagag aaagatgtgt tttgttttg actctctgtg  
 gtccttcca  
 2761 atgctgtggg tttccaacca ggggaagggt cccttttgca ttgccaagt  
 ccataacat  
 2821 gagcactact ctaccatggg tctgcctcct ggccaagcag gctggtttgc  
 50 aagaatgaaa  
 2881 tgaatgattc tacagctagg acttaacctt gaaatggaaa gtcttgcaat  
 ccatttgca  
 2941 ggatccgtct gtgcacatgc ctctgtagag agcagcattc ccagggacct  
 tggaacagt

3001 tggcactgta aggtgcttgc tccccaaagac acatcctaaa aggtgttgta  
 atggtgaaaa  
 3061 cgtcttccctt ctttattgcc ccttcttatt tatgtgaaca actgtttgtc  
 tttttttgta  
 5 3121 tcttttttaa actgtaaagt tcaattgtga aatgaatat catgcaaata  
 aattatgcga  
 3181 tttttttttc aaagcaaaaa

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>guanine deaminase</b>	GDA	NM_004293	35

10 1 gtagggagcc agcccctggg cgcggcctgc agggtagcgg caaccgcccg  
 ggtaagcggg  
 61 ggcaggacaa ggccggagcc tgtgtccgcc cggcagccgc cgcagctgc  
 agagagtccc  
 121 gctgcgtctc cgccgcgtgc gccctcctcg accagcagac ccgcgctgcg  
 15 ctccgccgct  
 181 gacatgtgtg ccgctcagat gccgcccctg gcgcacatct tccgagggac  
 gttcgtccac  
 241 tccacctgga cctgccccat ggaggtgctg cgggatcacc tcctcggcgt  
 gagcgacagc  
 20 301 ggcaaaatag tgtttttaga agaagcatct caacaggaaa aactggccaa  
 agaatgggtgc  
 361 ttcaagccgt gtgaaataag agaactgagc caccatgagt tcttcatgcc  
 tgggctggtt  
 421 gatacacaca tccatgcctc tcagtattcc tttgctggaa gtagcataga  
 25 cctgccactc  
 481 ttggagtggc tgaccaagta cacatttctt gcagaacaca gattccagaa  
 catcgacttt  
 541 gcagaagaag tatataccag agttgtcagg agaacactaa agaatggaac  
 aaccacagct  
 30 601 tgttactttg caacaattca cactgactca tctctgctcc ttgccgacat  
 tacagataaa  
 661 tttggacagc gggcatttgt gggcaaagt tgcattgatt tgaatgacac  
 ttttccagaa  
 721 tacaaggaga cactgagga atcgatcaag gaaactgaga gatttgtgtc  
 35 agaaatgctc  
 781 caaaagaact attctagagt gaagcccata gtgacaccac gtttttcct  
 ctctgctct  
 841 gagactttga tgggtgaact gggcaacatt gctaaaacc gtgatttgca  
 cattcagagc  
 40 901 catataagtg aaaatcgtga tgaagttgaa gctgtgaaaa acttataccc  
 cagttataaa  
 961 aactacacat ctgtgtatga taaaacaat cttttgacaa ataagacagt  
 gatggcacac  
 1021 ggctgctacc tctctgcaga agaactgaac gtattccatg aacgaggagc  
 45 atccatcgca  
 1081 cactgtccca attctaattt atcgctcagc agtggatttc taaatgtgct  
 agaagtctg  
 1141 aaacatgaag tcaagatagg gctgggtaca gacgtggctg gtggctattc  
 atattccatg

1201 cttgatgcaa tcagaagagc agtgatgggt tccaatatcc ttttaattaa  
 taaggtaa  
 1261 gagaaaagcc tcaccctcaa agaagtcttc agactagcta ctcttgagg  
 aagccaagcc  
 5 1321 ctggggctgg atgggtgagat tggaaacttt gaagtgggca aggaatttga  
 tgccatcctg  
 1381 atcaacccca aagcatccga ctctcccatt gacctgtttt atggggactt  
 ttttggtgat  
 1441 atttctgagg ctggtatcca gaagttcctc tatctaggag atgatcga  
 10 tattgaagag  
 1501 gtttatgtgg gcggaagca ggtggttccg tttccagct cagtgaaga  
 ccctcgggcg  
 1561 tctacaaagt tctcctggga ttagcgtggt tctgcatctc ccttgtgccc  
 aggtggagtt  
 1621 agaaagtcaa aaaatagtac cttgttcttg ggatgactat ccctttctgt  
 15 gtctagttac  
 1681 agtattcact tgacaaatag ttcgaaggaa gttgactaa ttctcaactc  
 tggttgagag  
 1741 ggttcataaa tttcatgaaa atatctcctt ttggagctgc tcagacttac  
 20 ttttaagctca  
 1801 aacagaaggg aatgctatta ctggtggtgt tcctacggta agacttaagc  
 aaagcctttt  
 1861 tcatatttga aaatgtggaa agaaaagatg ttctaaaag gttagatatt  
 ttgagctaat  
 1921 aattgcaaaa attagaagac tgaaaatgga cccatgagag tatattttta  
 25 tgagggagca  
 1981 aaagttagac tgagaacaaa cgtagaaaa tcaactcaga ttgtgtttga  
 aaattatata  
 2041 ctgagcatac taatttaaaa agagaacttg ttgaaatta aaacgtgttt  
 30 ctaggttgac  
 2101 cttgtgtttt agaaatttgc acttaatgga atttgcattt cagagatgtg  
 ttagtgttgt  
 2161 gctttgcctt ctttggcgat gaatgtcaga aattgaatgc cacatgcttt  
 cataatatag  
 2221 ttttgtgctt caaagtgttt gacagaagtt gggattaaa gatttaaagt  
 35 ctcttaggaa  
 2281 tattattcat gtaactccat ggcataaata gttgtatttt tgtgtacttt  
 aaaatcaact  
 2341 tataactgtg agatgttatt gcttccattt tattagaaga gaaacaaatt  
 40 ccatgcttta  
 2401 tggaaattat gtagactgga gtcttcgtga actggggcaa atgctggcat  
 ccaggagccg  
 2461 ccaatactaa caggacaggt tccattgcca tggcctattc cacccaaaca  
 atatgttgta  
 2521 gtttctggaa attccatact cagatatcag tctgctagaa ctttaaaatg  
 45 aaggacaaat  
 2581 cctgttaaag aaatattggt aaaaatcttt aaaccctgtg tattgaaagc  
 actctatttt  
 2641 ctaattttat ccagttttct gtttaactcc ttataatggt taggatatta  
 50 aaattttagg  
 2701 ataatagaaga gtacataatg tcctacttaa ttttatggt aataggactt  
 aattcttact  
 2761 agacatctag gaacattaca aagcaaagac ttttttatg cttccataac  
 ctagaattaa

2821 aaccaaatta tgaccttatg ataaatcttt aagtattggt gtgaatgtta  
 tttaaattct  
 2881 atatttttct tatttaatta caaatactat aaatgagcaa ggaaaaggaa  
 tagactttct  
 5 2941 taatatatta taacactcat tcctagagct taggggtgac tctttaatat  
 taccttatag  
 3001 tagaaacttt atgtaatata gctaactccg tatttacaga acaaaaaaac  
 acagttcccc  
 3061 ctctgtagt ataaatttta ttttcacata cttagctaat ttagcagtaa  
 10 ttggcccagt  
 3121 tttttcccta atagaaatac ttttagattt gattatgtat acatgacacc  
 taaagagggga  
 3181 acaaaagtta gttttatfff ttttaataaac aacagagttt gttttgtgag  
 ataagtatct  
 15 3241 tagtaaacc aatttccagt cttagtctgt attccaata tttctaattc  
 ctgagccacg  
 3301 tcaaagatgc cttgccaat ttctccccat ttctctacgg ggctagcaaa  
 aatcttcagc  
 3361 tttatcactc aaccctgcc aaaggaactt gattacatgg tgtctaacca  
 20 aatgagcagg  
 3421 cttaggaatt tagatgagat gtgtaagatt cacttacagg cagtagctgc  
 ttctagcatt  
 3481 tgcaagatcc tacactttta ccttctttaa ggggtgtacat tttgatgttg  
 aacatcagtt  
 25 3541 ttcattgtaga cttaggactc atgtgcagta aatataaata agtgtagcat  
 cagaagcagt  
 3601 aggaatggcc gtatacaacc atcctgttaa acatttaaat ttagctctga  
 tagtgtgta  
 3661 agacctgaat atctttccta gtaaaaatag gatgtgttga aatatttata  
 30 tgtactttga  
 3721 tctctccaca tcacttataa cttatgtgtt ttatttctcc aagtgcgggtg  
 ttcctgaatg  
 3781 ttatgtatgc ttttttttct gtaccacagg cattatctat acctggggcc  
 agattttctg  
 35 3841 cactttgaaa tgttgcttt gcctaatagta ggttgacttt ctgaattgtg  
 gagaggcact  
 3901 tttccaagcc aatcttattt gtcacttttt gtttaatat cttgctctct  
 gacaggaaag  
 3961 aaacaattca cttaccagcc tcctcaccac atcctccacc atttccttaa  
 40 tgttccatgg  
 4021 tattttcaac ggaatacact ttgaaaggta aaaacaattc aaaagtatcg  
 attatcataa  
 4081 attcacaataa tatttttgca accagaacac aaaagcaggc tagtcagcta  
 aggtaaattt  
 45 4141 cattttcaaa cgagagggaa acatgggaag taaaagatta ggatgtgaaa  
 ggttgtccta  
 4201 aacagaccaa ggagactggt ccctaattta ttctcttggc tggttctctc  
 attgaattat  
 4261 cagaccccaa gaggagatat tggaacaggc tccttcatg ccaagggctc  
 50 ttctaagtta  
 4321 atactgtgag cattgagccc ccattaaaac tcttttttac ttcagaaaga  
 attttacagg  
 4381 ttaaagggaa agaaatggtg ggaaactctc cccgtaatgc ttagccaact  
 ttaaagtga

4441 cccttcaata tccccattgg caactgcagc tgagatctta gagaggaaat  
 ataaccggtg  
 4501 tgagatctag caatgcattt tgaatcttca ctccctacca ggctcttctc  
 atttttaatc  
 5 4561 tcttcacctc agaactagac atatggagag cttaaaggc aagctggaag  
 gcacattgta  
 4621 tcaattctac cttgtgctat acgtaggaga gatccaaaat ttggatgctt  
 ctggagactc  
 4681 ttagacatct tttcattggt gtccattttt aaagttgatg attgctggaa  
 10 acattcacac  
 4741 gcttaaaaagc aatgggtgta gttattaatg ggtaaaactaa gaagtgttat  
 aggcaatgac  
 4801 ttgaaatggt ttttaaattg tatggattgt taagaattgt tgaaaaaaaa  
 tttttttttt  
 15 4861 ttggacagct tcaaggagat gttagcaatt tcagatatac tagccagttt  
 aggtatgact  
 4921 ttggaagtgc agaaacagaa ggatactggt agaaaatcct aacattggtc  
 tccgtgcatg  
 4981 tgttcacacc tggctcact gcctttcctt cccacagacc tgagtgtgaa  
 20 agactgagag  
 5041 ttgaggagtt actttgtgga tcttgtccaa atttagtgaa atgtggaagt  
 caaccagacc  
 5101 aatgatggaa ttaaattgaa attccaagag ggctttcaca gtccacaggg  
 ttcaaatgac  
 25 5161 ttgggtaaca gaagttattc ttagcttacc tgttatgtga cagtgattta  
 cctgtccatt  
 5221 tccaaccaa aagcctgtca gaaagcattc ttagagaaa accactttac  
 atttgttgtt  
 5281 aaactcctga tcgctactct taagaatata catgtatgta ttcataggaa  
 30 cattttttct  
 5341 caatatttgt atgattcgct tactgttatt gtgctgagtg agctcctgtg  
 tgcttcagac  
 5401 aaaaataaat gagactttgt gtttacgtta

35

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>macrophage stimulating 1 receptor (c-met-related tyrosine kinase)</b>	MST1R	NM_002447	36

1 ggatcctcta gggteccagc tcgctcgat ggagctctc ccgccgctgc  
 ctcagtcctt  
 61 cctgttctg ctgctgttgc ctgccaagcc cgcggcgggc gaggactggc  
 40 agtgcgccg  
 121 caccctac gcggcctctc gcgactttga cgtgaagtac gtggtgccca  
 gcttctccgc  
 181 cggaggcctg gtacaggcca tggtgacctc cgagggcgac agaaatgaga  
 gtgctgtgtt

241 tgtagccata cgcaatcgcc tgcattgtgt tgggcctgac ctgaagtctg  
 tccagagcct  
 301 ggccacgggc cctgctggag accctggctg ccagacgtgt gcagcctgtg  
 gccagggacc  
 5 361 ccacggccct cccgggtgaca cagacacaaa ggtgctgggt ctggatcccg  
 cgctgcctgc  
 421 gctggctcagt tgtggctcca gcctgcaggg ccgctgcttc ctgcatgacc  
 tagagcccca  
 481 agggacagcc gtgcatctgg cagcgcacgc ctgcctcttc tcagcccacc  
 10 ataaccggcc  
 541 cgatgactgc cccgactgtg tggccagccc attgggcacc cgtgtaactg  
 tggttgagca  
 601 aggccagggc tcctatttct acgtggcacc ctcaactggac gcagccgtgg  
 ctggcagctt  
 15 661 cagcccacgc tcagtgtcta tcaggcgtct caaggetgac gcctcgggat  
 tcgacccggg  
 721 ctttgtggcg ttgtcagtgc tgcccagca tcttgtctcc tacagtattg  
 aatacgtgca  
 781 cagcttccac acgggagcct tcgtatactt cctgactgta cagccggcca  
 20 gcgtgacaga  
 841 tgatcctagt gccctgcaca cagcctggc acggcttagc gccactgagc  
 cagagttggg  
 901 tgactatcgg gagctggtcc tcgactgcag atttgcctca aaacgcaggc  
 gccggggggc  
 25 961 cccagaaggc ggacagccct accctgtgct gcaggtggcc cactccgctc  
 cagtgggtgc  
 1021 ccaacttgc actgagctga gcacgcccga gggccaggaa gtactatttg  
 gggctcttct  
 1081 gactggcaag gatgggtggtc ctggcgtggg cccaactct gtcgtctgtg  
 30 ccttccccat  
 1141 tgacctgctg gacacactaa ttgatgaggg tgtggagcgc tgttgtgaat  
 ccccagtcca  
 1201 tccaggcctc eggcgaggcc tcgacttctt ccagtcgccc agtttttggc  
 ccaaccggcc  
 35 1261 tggcctggaa gccctcagcc ccaacaccag ctgccgccac ttcctctgc  
 tggtcagttag  
 1321 cagcttctca cgtgtggacc tattcaatgg gctgttggga ccagtacagg  
 tcaactgcatt  
 1381 gtatgtgaca gccttgaca acgtcacagt ggacacatg ggcacaatgg  
 40 atgggcgtat  
 1441 cctgcaggtg gagctggtca ggtcactaaa ctacttgctg tatgtgtcca  
 acttctcact  
 1501 gggtgacagt gggcagcccg tgcagcggga tgtcagtcgt cttggggacc  
 acctactctt  
 45 1561 tgcctctggg gaccaggttt tccaggtacc tatccgaggc cctggctgcc  
 gccacttct  
 1621 gacctgtggg cgttgcctaa gggcatggca tttcatgggc tgtggctggg  
 gtgggaacat  
 1681 gtgcggccag cagaaggagt gtccctggctc ctggcaacag gaccactgcc  
 50 cacctaagct  
 1741 tactgagttc cacccccaca gtggacctct aaggggcagt acaaggctga  
 cctgtgtgg  
 1801 ctccaacttc taccttccac cttctggctc ggtgcctgag ggaaccatc  
 aggtcactgt

1861 gggccaaagt ccctgccggc cactgcccaa ggacagctca aaactcagac  
 cagtgccccg  
 1921 gaaagacttt gtagaggagt ttgagtgtga actggagccc ttgggcaccc  
 aggcagtggg  
 5 1981 gcctaccaac gtcagcctca ccgtgactaa catgccaccg ggcaagcact  
 tccgggtaga  
 2041 cggcacctcc gtgctgagag gcttctcttt catggagcca gtgctgatag  
 cagtgcaacc  
 2101 cctctttggc ccacgggcag gaggcacctg tctcactctt gaaggccaga  
 10 gtctgtctgt  
 2161 aggaccagc cgggctgtgc tggatcaatgg gactgagtgt ctgctagcac  
 gggtcagtga  
 2221 ggggcagctt ttatgtgcca caccctctgg ggccacggtg gccagtgtcc  
 cccttagcct  
 15 2281 gcaggtgggg ggtgccaggg tacctgggtc ctggaccttc cagtacagag  
 aagaccctgt  
 2341 cgtgctaagc atcagcccca actgtggcta catcaactcc cacatcacca  
 tctgtggcca  
 2401 gcatctaact tcagcatggc acttagtgct gtcattccat gacgggctta  
 20 gggcagtggg  
 2461 aagcaggtgt gagaggcagc ttccagagca gcagctgtgc cgccttctg  
 aatatgtggg  
 2521 ccgagacccc cagggatggg tggcagggaa tctgagtgcc cgaggggatg  
 gagctgctgg  
 25 2581 ctttactctg cctggctttc gcttctctacc cccaccccat ccaccagtg  
 ccaacctagt  
 2641 tccactgaag cctgaggagc atgccattaa gtttgagtat attgggctgg  
 gcgctgtggc  
 2701 tgactgtgtg ggtatcaacg tgaccgtggg tgggtgagagc tgccagcacg  
 30 agttccgggg  
 2761 ggacatgggt gtctgcccc cgtccccatc cctgcagctt ggccaggatg  
 gtgccccatt  
 2821 gcaggtctgc gtagatgggt aatgtcatat cctgggtaga gtgggtgcggc  
 caggccaga  
 35 2881 tggggtccca cagagcacgc tccttgggat cctgctgcct ttgctgctgc  
 ttgtggctgc  
 2941 actggcgact gcactggctc tcagctactg gtggcggagg aagcagctag  
 ttcttctctc  
 3001 caacctgaat gacctggcat ccctggacca gactgctgga gccacacccc  
 40 tgctattct  
 3061 gtactcgggc tctgactaca gaagtggcct tgactccct gccattgatg  
 gtctggattc  
 3121 caccacttgt gtccatggag catccttctc cgatagtga gatgaatcct  
 gtgtgccact  
 45 3181 gctgcggaaa gagtccatcc agctaaggga cctggactct gcgctcttgg  
 ctgaggtcaa  
 3241 ggatgtgctg attccccatg agcgggtggg caccacagt gaccgagtca  
 ttggcaaagg  
 3301 ccactttgga gttgtctacc acggagaata catagaccag gcccagaatc  
 50 gaatccaatg  
 3361 tgccatcaag tcactaagtc gcatcacaga gatgcagcag gtggaggcct  
 tcctgcgaga  
 3421 ggggctgctc atgcgtggcc tgaaccaccc gaatgtgctg gctctcattg  
 gtatcatgtt

3481 gccacctgag ggctgcccc atgtgctgct gccctatatg tgccacggtg  
acctgctcca  
3541 gttcatccgc tcacctcage ggaacccac cgtgaaggac ctcatcagct  
ttggcctgca  
5 3601 ggtagcccgc ggcattggagt acctggcaga gcagaagttt gtgcacaggg  
acctggctgc  
3661 gcggaactgc atgctggacg agtcattcac agtcaagggtg gctgactttg  
gtttggccccg  
3721 cgacatcctg gacagggagt actatagtgt tcaacagcat cgccacgctc  
10 gcctacctgt  
3781 gaagtggatg gcgctggaga gcctgcagac ctatagattt accaccaagt  
ctgatgtgtg  
3841 gtcatttggg gtgctgctgt gggaaactgct gacacgggggt gccccacat  
accgccacat  
15 3901 tgaccctttt gaccttacc acttctctggc ccagggtcgg cgctgcccc  
agcctgagta  
3961 ttgccctgat tctctgtacc aagtgatgca gcaatgctgg gaggcagacc  
cagcagtgcg  
4021 acccaccttc agagtactag tgggggaggt ggagcagata gtgtctgcac  
20 tgcttgggga  
4081 ccattatgtg cagctgccag caacctacat gaacttgggc cccagcacct  
cgcatgagat  
4141 gaatgtgctt ccagaacagc cgcagttctc acctatgcca gggaaatgtac  
gccggccccg  
25 4201 gccactctca gacctctctc ggcccacttg acttagttct tgggctggac  
ctgcttagct  
4261 gccttgagct aacccaagg ctgcctctgg gccatgccag gccagagcag  
tgccctcca  
4321 ccttgttctt gccctttaa tttcagaggc aataggtaaa tgggcccatt  
30 aggtccctca  
4381 ctccacagag tgagccagtg agggcagtcc tgcaacatgt atttatggag  
tgctgctgt  
4441 ggaccctgtc ttctgggcac agtggactca gcagtgacca caccaacact  
gaccctttaa  
35 4501 ccaataaagg aacaaatgac tattaagca caaaaaaaaa a

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
integrin, beta 4	ITGB4	NM_000213	37

```

1  gcgctgcccg cctcgtcccc acccccccaa cccccgcgcc cgccctcgga
cagtcctctgc
5  61  tcgcccgcgc gctgcagccc catctcctag cggcagccca ggcgcggagg
gagcagagtc
121 gccccgaggt aggtccagga cgggcgcaca gcagcagccg aggctggccg
ggagagggag
181 gaagaggatg gcagggccac gccccagccc atgggccagg ctgctcctgg
10  cagccttgat
241 cagcgtcagc ctctctggga ccttgccaaa ccgctgcaag aaggccccag
tgaagagctg
301 cacggagtgt gtcctgtgtg ataaggactg cgctactgc acagacgaga
tgttcagggg
15 361 ccggcgctgc aacaccagc cggagctgct ggccgcgggc tgccagcggg
agagcatcgt
421 ggtcatggag agcagcttcc aatcacaga ggagaccag attgacacca
cctgcggcgc
20 481 cagccagatg tcccccaag gctgcgggt ccgtctgagg cccggtgagg
agcggcattt
541 tgagctggag gtgtttgagc cactggagag ccccgaggac ctgtacatcc
tcatggactt
601 ctccaactcc atgtccgatg atctggacaa cctcaagaag atggggcaga
acctggctcg
25 661 ggtcctgagc cagctacca ggcactacac tattggattt ggcaagtttg
tggacaaagt
721 cagcgtcccc cagacggaca tgaggcctga gaagctgaag gagccctggc
ccaacagtga
781 ccccccttc tccttaaga acgtcatcag cctgacagaa gatgtggatg
30  agttceggaa
841 taaactgcag ggagagcggg totcaggcaa cctggatgct cctgagggcg
gcttcgatgc
901 catcctgcag acagctgtgt gcaagagggg cattggctgg cccccggaca
gcacccacct
35 961 gctggctctt tccaccgagt cagccttcca ctatgaggct gatggcgcca
acgtgctggc
1021 tggcatcatg agccgcaacg atgaacggtg ccacctggac accacgggca
cctacacca
1081 gtacaggaca caggactacc cgtegggtgcc caccctggtg cgctgctcg
40  ccaagcacia
1141 catcatcccc atctttgctg tcaccaacta ctctatagc tactacgaga
agcttcacac
1201 ctatctccct gtctctcac tgggggtgct gcaggaggac tcgtccaaca
tcgtggagct
45 1261 gctggaggag gccttcaatc ggatccgctc caacctggac atccgggccc
tagacagccc
1321 ccgaggcctt cggacagagg tcacctcaa gatgttcag aagacgagga
ctgggtcctt
1381 tcacatccgg cggggggaag tgggtatata ccagggtcag ctgcgggccc
50  ttgagcacgt

```

1441 ggatgggacg cacgtgtgcc agctgccgga ggaccagaag ggcaacatcc  
 atctgaaacc  
 1501 ttccttctcc gacggcctca agatggacgc gggcatcatc tgtgatgtgt  
 gcacctgcca  
 5 1561 gctgcaaaaa gaggtgCGGT cagctcgctg cagcttcaac ggagacttcg  
 tgtgcggaca  
 1621 gtgtgtgtgc agcgagggt ggagtggcca gacctgcaac tgctccaccg  
 gctctctgag  
 1681 tgacattcag cctgcctgc gggagggcga ggacaagccg tgctccggcc  
 10 gtggggagtg  
 1741 ccagtgcggg cactgtgtgt gctacggcga aggccgctac gagggtcagt  
 tctgcgagta  
 1801 tgacaacttc cagtgtcccc gcaacttcgg gttcctctgc aatgaccgag  
 gacgctgctc  
 15 1861 catgggccag tgtgtgtgtg agcctgggtg gacaggccca agctgtgact  
 gtccccctcag  
 1921 caatgccacc tgcacgaca gcaatggggg catctgtaat ggacgtggcc  
 actgtgagtg  
 1981 tggccgctgc cactgccacc agcagtcgct ctacacggac accatctgcg  
 20 agatcaacta  
 2041 ctcgccgatc caccggggcc tctgcgagga cctacgctcc tgcgtgcagt  
 gccaggcgtg  
 2101 gggcaccggc gagaagaagg ggcgcacgtg tgaggaatgc aacttcaagg  
 tcaagatggt  
 25 2161 ggacgagctt aagagagccg aggaggtggt ggtgcgctgc tcttccggg  
 acgaggatga  
 2221 cgactgcacc tacagctaca ccatggaagg tgacggcgcc cctgggcccc  
 acagcactgt  
 2281 cctggtgcac aagaagaagg actgcccctc gggctccttc tgggtggtca  
 30 tccccctgct  
 2341 cctcctctc ctgcccctcc tggcccctgct actgctgcta tgctggaagt  
 actgtgcctg  
 2401 ctgcaaggcc tgccctggca ttctcccgtg ctgcaaccga ggtcacatgg  
 tgggtttta  
 35 2461 ggaagaccac tacatgctgc gggagaacct gatggcctct gaccacttg  
 acacgccccat  
 2521 gctgcgagc gggaaacctca agggccgtga cgtggctccg tggaaggtca  
 ccaacaacat  
 2581 gcagcggcct ggctttgcca ctcatgccgc cagcatcaac cccacagagc  
 40 tgggtgcccta  
 2641 cgggctgtcc ttgcgcctgg cccgcctttg caccgagaac ctgctgaagc  
 ctgacactcg  
 2701 ggagtgcgcc cagctgcgcc aggaggtgga ggagaacctg aacgaggtct  
 acaggcagat  
 45 2761 ctccggtgta cacaagctcc agcagaccaa gttccggcag cagcccaatg  
 ccgggaaaaa  
 2821 gcaagaccac accattgtgg acacagtgct gatggcgccc cgctcggcca  
 agccggccct  
 2881 gctgaagctt acagagaagc aggtggaaca gagggccttc cacgacctca  
 50 aggtggcccc  
 2941 cggctactac accctcactg cagaccagga cgccccggggc atggtggagt  
 tccaggaggg  
 3001 cgtggagctg gtggacgtac ggggtgccct ctttatccgg cctgaggatg  
 acgacgagaa

3061 gcagctgctg gtggaggcca tgcagctgcc cgcaggcact gccaccctcg  
 gccgccgcct  
 3121 ggtaaaccatc accatcatca aggagcaagc cagagacgtg gtgtcctttg  
 agcagcctga  
 5 3181 gttctcggtc agccgcgggg accagggtggc ccgcatccct gtcacccggc  
 gtgtcctgga  
 3241 cggcgggaag tcccaggtct cctaccgcac acaggatggc accgcgcagg  
 gcaaccggga  
 10 3301 ctacatcccc gtggagggtg agctgctggt ccagcctggg gaggcctgga  
 aagagctgca  
 3361 ggtgaagctc ctggagctgc aagaagttga ctccctcctg cggggccgcc  
 aggtcccgcg  
 3421 tttccacgtc cagctcagca accctaagtt tggggcccac ctggggccagc  
 cccactccac  
 15 3481 caccatcatc atcagggacc cagatgaact ggaccggagc ttcacgagtc  
 agatggtgtc  
 3541 atcacagcca cccctcacg gcgacctggg cgcgccgag aaccccaatg  
 ctaaggccgc  
 20 3601 tgggtccagg aagatccatt tcaactggct gccccttct ggcaagccaa  
 tggggtagac  
 3661 ggtaaagtac tggattcagg gtgactccga atccgaagcc cacctgctcg  
 acagcaaggt  
 3721 gccctcagtg gagctcacca acctgtacce gtattgcgac tatgagatga  
 aggtgtgcgc  
 25 3781 ctacggggct cagggcgagg gaccctacag ctccctggtg tctgcccga  
 cccaccagga  
 3841 agtgcccagc gagccagggc gtctggcctt caatgtcgtc tctccacgg  
 tgaccagct  
 3901 gagctgggct gagccggctg agaccaacgg tgagatcaca gcctacgagg  
 tctgctatgg  
 30 3961 cctggtcaac gatgacaacc gacctattgg gccatgaag aaagtgctgg  
 ttgacaaccc  
 4021 taagaaccgg atgctgctta ttgagaacct tcgggagtcc cagccctacc  
 gctacacggt  
 35 4081 gaaggcgcgc aacggggccg gctgggggcc tgagcgggag gccatcatca  
 acctggccac  
 4141 ccagcccaag aggcccatgt ccaccccat catccctgac atccctatcg  
 tggacgcca  
 4201 gagcggggag gactacgaca gcttccttat gtacagcgat gacgttctac  
 40 gctctccatc  
 4261 gggcagccag aggcccagcg tctccgatga cactggctgc ggctggaagt  
 tcgagcccct  
 4321 gctgggggag gagctggacc tgcggcgcgt cacgtggcgg ctgccccgg  
 agctcatccc  
 45 4381 gcgctgtcg gccagcagcg ggcgctcctc cgacgccgag gcgccccacg  
 ggccccgga  
 4441 cgacggcggc gcgggcggga agggcggcag cctgccccgc agtgcgacac  
 ccgggcccc  
 4501 cggagagcac ctggtgaatg gccggatgga ctttgccttc ccgggcagca  
 50 ccaactccct  
 4561 gcacaggatg accacgacca gtgctgctgc ctatggcacc cacctgagcc  
 cacacgtgcc  
 4621 ccaccgcgtg ctaagcacat cctccaccct cacacgggac tacaactcac  
 tgaccgcctc

4681 agaacactca cactcgacca cactgcccag ggactactcc accctcacct  
ccgtctcctc  
4741 ccacgactct cgectgactg ctgggtgtgcc cgacacgccc acccgctgg  
tgttctctgc  
5 4801 cctggggccc acatctctca gagtgagctg gcaggagccg cggtgcgagc  
ggccgctgca  
4861 gggctacagt gtggagtacc agctgctgaa cggcgggtgag ctgcatcggc  
tcaacatccc  
4921 caaccctgcc cagacctcgg tgggtgggga agacctcctg cccaaccact  
10 cctacgtggt  
4981 ccgctgctgg gccagagcc aggaaggctg gggccgagag cgtgaggggtg  
tcatcaccat  
5041 tgaatcccag gtgcaccgc agagcccact gtgtcccctg ccaggctccg  
ccttcacttt  
15 5101 gagcactccc agtgccccag gcccgctggt gttcactgcc ctgagcccag  
actcgtgca  
5161 gctgagctgg gagcggccac ggaggcccaa tggggatata gtcggctacc  
tggtagcctg  
20 5221 tgagatggcc caaggaggag ggccagccac cgcattccgg gtggatggag  
acagccccga  
5281 gagccggctg accgtgccgg gcctcagcga gaacgtgcc tacaagttca  
aggtgcaggc  
5341 caggaccact gagggcttcg ggccagagcg cgagggcatc atcacatag  
agtcccagga  
25 5401 tggaggacce ttcccgcagc tgggcagccg tgccgggctc ttccagcacc  
cgctgcaaag  
5461 cgagtacagc agcatcacca ccacccacac cagcgcacc gagcccttc  
tagtggatgg  
5521 gctgaccctg ggggcccagc acctggaggc aggcggtcc ctcaccggc  
30 atgtgacce  
5581 ggagtttgtg agccggacac tgaccaccag cggaacctt agcaccaca  
tggaccaaca  
5641 gttcttccaa acttgaccgc accctgcccc acccccgcca cgtcccacta  
ggcgtcctcc  
35 5701 cgactcctct cccggagcct ectcagctac tccatccttg caccctggg  
ggcccagccc  
5761 acccgcctgc acagagcagg ggctaggtgt ctctggggag gcatgaaggg  
ggcaaggctc  
5821 gtcctctgtg ggcccaaacc tatttgtaac caaagagctg ggagcagcac  
40 aaggaccag  
5881 cctttgttct gcaactaata aatggttttg ctactgctaa

45

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
annexin A3	ANXA3	NM_005139	38

1 ggggtggggaa gcttagagac cggtgagggga gcagagctgg ggcgcctgtg  
 tacaggata  
 5 61 gagcccggcg gcagcagggc gcggcttccc tttcccgggg cctggggccg  
 caatcaggtg  
 121 gagtgcgagag gccggaggag gggcaggagg aaggggtgcg gtcgcatcc  
 ggaccggag  
 181 ccagcgcgga gcacctgcgc ccgcggctga caccttcgct cgcagtttgt  
 10 tcgcagtta  
 241 ctgcacacacc agtttccccc accgcgcttt ggattagtgt gatctcagct  
 caaggcaaag  
 301 gtgggatatc atggcatcta tctgggttg acaccgagga acagtaagag  
 attatccaga  
 15 361 ctttagccca tcagtggatg ctgaagctat tcagaaagca atcagaggaa  
 ttggaactga  
 421 tgagaaaatg ctcatcagca ttctgactga gaggtcaaat gcacagcggc  
 agctgattgt  
 481 taaggaatat caagcagcat atggaaagga gctgaaagat gacttgaagg  
 20 gtgatctctc  
 541 tggccacttt gagcatctca tgggtggccct agtgactcca ccagcagtct  
 ttgatgcaaa  
 601 gcagctaaag aatccatga agggcgcggg aacaaacgaa gatgccttga  
 ttgaaatctt  
 25 661 aactaccagg acaagcaggc aatgaagga tatctctcaa gcctattata  
 cagtatacaa  
 721 gaagagtctt ggagatgaca ttagttccga aacatctggt gacttccgga  
 aagctctggt  
 781 gactttggca gatggcagaa gagatgaaag tctgaaagtg gatgagcatc  
 30 tggccaaaca  
 841 agatgccag attctctata aagctggtga gaacagatgg ggcacggatg  
 aagacaaatt  
 901 cactgagatc ctgtgtttaa ggagctttcc tcaattaa ctaacatttg  
 atgaatacag  
 35 961 aatatcagc caaaaggaca ttgtggacag cataaaagga gaattatctg  
 ggcattttga  
 1021 agacttactg ttggccatag ttaattgtgt gaggaacacg ccggcctttt  
 tagccgaaag  
 1081 actgcatcga gccttgaagg gtattggaac tgatgagttt actctgaacc  
 40 gaataatggt  
 1141 gtccagatca gaaattgacc ttttggacat tcgaacagag ttcaagaagc  
 attatggcta  
 1201 ttccttatat tcagcaatta aatcggatac ttctggagac tatgaaatca  
 cactcttaaa  
 45 1261 aatctgtggt ggagatgact gaaccaagaa gataatctcc aaaggccac  
 gatgggcttt  
 1321 cccaacagct ccaccttact tcttctcata ctatttaaga gaacaagcaa  
 atataaacag  
 1381 caacttgtgt tctaacagg aattttcatt gttctataac aacaacaaca  
 50 aaagcgatta

1441 ttatTTTtaga gcatctcatt tataatgtag cagctcataa atgaaattga  
 aaatggatt  
 1501 aaagatctgc aactactatc caacttatat ttctgcttcc aaagttaaga  
 atctttatag  
 5 1561 ttctactcca ttaaataataa agcaagataa taaaaattgt tgcttttggt aaaagtaaaa

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>chemokine (C-C motif) ligand 15</b>	CCL15	NM_032965	39

1 tgcagactga tatggattca ccaactgctaa cacctcctgg ttggaactac  
 10 aggaatagaa  
 61 ctggaaaggg aaaaaaggca gcattcacca catcccaatc ctgaatccaa  
 gagtctaaga  
 121 tagtccccca ctctatctc aggcttagag gattagatta atctcctgga  
 ggaagactc  
 15 181 ttccttgaaa catttttttt tatctgcctg tagctattgg gataattcgg  
 gaaatccaca  
 241 gggacagttc aagtcactct tgctctctac tttctgttgc actctcagcc  
 ttgttctctt  
 301 tttagaaaact gcatggtaac tattatatag ctaaagaaga gcattctgac  
 20 ctctgcctg  
 361 ggacttctg gatcctctc ttcttataaa tacaaggga gagctggtat  
 cccggggagc  
 421 caggaagcag tgagcccagg agtctctggc cagcctgccc tgcccaccag  
 gaggatgaag  
 25 481 gtctccgtgg ctgccctctc ctgcctcatg cttgttgctg tcttggatc  
 ccaggcccag  
 541 ttcataaatg atgcagagac agagttaatg atgtcaaagc ttcactgga  
 aatccagta  
 601 gttctgaaca gctttcactt tgctgctgac tgctgcacct cctacatctc  
 30 acaaagcatc  
 661 cgtgtttcac tcatgaaaag ttatTTTgaa acgagcagcg agtgctccaa  
 gccaggtgct  
 721 atattcctca ccaagaaggg gcggcaagtc tgtgccaac ccagtggctc  
 gggagttcag  
 35 781 gattgcatga aaaagctgaa gccctactca atataataat aaagagacia  
 aagaggccag  
 841 ccaccacct ccaacacctc ctgtgagttt cttggctctga aatacttaaa  
 aatatatat  
 901 attgttggtg ctggtaatga aagtaatgca tctaataaag agtattcaat ttttt  
 40

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
dipeptidase 1 (renal)	DPEP1	NM_004413	40

```

1  cgggggggta ctgtgcgagc cctcaaggag gtggctgttc tgtagctgga
gagctccgtg
5  61  ggtggcagga ctgaacttga acaccagaaa caacccccaa gccttgtgac
ctgggaggca
121 ggaggcgggt ctgtctccct gggacttggg tggctgagcc gaggtactcg
ggaccctgtc
181 cgcgcgatgg cagagtggct cctcacagcc tgaagtcac ctttctgcac
10  gggccagcca
241 ggccagcaca gaggcaccag ggcagcagtg cacacaggtc cccggggacc
ccaccatgtg
301 gagcggatgg tggctgtggc cccttgtggc cgtctgcact gcagacttct
ttcgggacga
15  361 ggcagagagg atcatgaggg actcccctgt cattgatggg cacaatgacc
tcccctggca
421 gctgctggat atgttcaaca accggctgca ggacgagagg gccaacctga
ccacctggc
481 cggcacacac accaacatcc ccaagctgag ggccggcttt gtgggaggcc
20  agttctggtc
541 cgtgtacacg ccctgcgaca cccagaacaa agacgccctg cggaggacgc
tggagcagat
601 ggacgtggtc caccgcatgt gccggatgta cccggagacc ttctgtatg
tcaccagcag
25  661 tgcaggcatt cggcaggcct tccgggaagg gaaggtggcc agcctgatcg
gcgtggaggg
721 cggccactcc attgacagca gtttgggcgt cctgcgggca ctctatcagc
tgggcatgcg
781 gtacctgacc ctcaccaca gctgcaacac gccttgggct gacaactggc
30  tggaggacac
841 gggagacagc gagccccaga gccaaaggctt gtcaccctt gggcagcgtg
tggagaagga
901 gctgaaccgt ctgggggtcc tcategactt ggctcacgtg tctgtggcca
ccatgaaggc
35  961 caccctgcag ctgtccagag ccccggtcat cttcagccac tcctcggcct
acagcgtgtg
1021 cgcaagccgg cgcaacgtgc ctgacgacgt cctgaggctg gtgaaacaga
cagacagcct
1081 ggtgatggtg aacttctaca acaattacat ttctgcacc aacaaggcca
40  acctgtccca
1141 agtggccgac catctggate acatcaagga ggtggcagga gccagagccg
tgggttttgg
1201 tggggacttt gatggtgttc caagggtccc tgaggggctg gaggacgtct
ccaagtatcc
45  1261 agacctgatc gctgagctgc tcaggaggaa ctggacggag gcggaggtca
agggcgact
1321 ggctgacaac ctgctgaggg tcttcgaggc tgtggaacag gccagcaacc
tcacacaggc
1381 tcccaggagg gagcccatcc cgctggacca gctgggtggc tcctgcagga
50  cccattacgg

```

1441 ctactcctct ggggcttcca gcctccatcg ccaactggggg ctctgctgg  
 cctccctcgc  
 1501 tcccctggtc ctctgtctgt ctctcctgtg aaacctggga gaccagagtc  
 ccctttaggg  
 5 1561 ttcccggagc tccgggaaga cccgcccac ccaggactcc agatgccagg  
 agccctgctg  
 1621 cccacatgca aggaccagca tctcctgaga ggacgcctgg gcttacctgg  
 ggggcaggat  
 1681 gcctgggggac agttcaggac acacacacag taggcccgc ataaaagcaa caccctt

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>NADPH oxidase organizer 1</b>	NOXO1	NM_172167	41

1 agccatggca ggccccgat acccagtttc agtgcaaggg gcagccctgg  
 tgcagatcaa  
 15 61 gaggctccaa acgtttgcct tctctgtgcg ctggtcagac ggcagcgaca  
 ccttcgtgcg  
 121 caggagttgg gacgaattca ggcagctcaa gaagaccctc aaggagacct  
 tcccggtgga  
 181. ggcgggcttg ctgcggagat ctgaccgcgt tctcccaaag cttctcgatg  
 20 caccactgtt  
 241 gggacgcgtg gggcgcacga gccgcggcct ggcgcgctg cagctggtgg  
 aaacctatc  
 301 tggaggctg ctggcgactg cagagcgcgt ggcacggagc ccgacgatca  
 ctggcttctt  
 25 361 cgcaccgcaa ccctggacc tggagcccgc gctgccaccg ggcagccggg  
 tgatectgcc  
 421 caccocagag gagcagcctc tttctcgcgc tgcgggcccgc ctctccatcc  
 acagtctgga  
 481 ggctcagagc ctgcgctgcc tgcagccctt ctgtaccag gacacgcggg  
 30 ataggccttt  
 541 tcaggcgcag gcccaggaga gcctggacgt gctgctgcgg caccctcag  
 gctggtggct  
 601 ggtggagaac gaagaccggc agaccgcctg gtttccagcg ccctacctgg  
 aggagccggc  
 35 661 cccgggcca ggcggggagg gaggcccgtc cctagggagc agcgggtccc  
 agttctgtgc  
 721 ttcccgcgcc tacgagagca gccgcgcaga tgagctgtcc gtgcccgcgg  
 gggcgcgcgt  
 781 gcgcgtggtg gaaacgtcag accgcggctg gtggctatgc aggtacggcg  
 40 accgggcggg  
 841 cctactcccc gcggtgctgc tgcggccgga agggctgggc gctctcctga  
 gcgggacggg  
 901 gttccgtgga ggagacgacc cggcgggtga ggcccggggc ttcctgaac  
 cctcccaggc  
 45 961 caccgcccct cccccaccg tgcccaccg accttcgccc ggcgccatcc  
 agagccgctg  
 1021 ctgcaccgtc acacgcaggg ccctggagcg gcgcccacgg cgcagggcc  
 gcctcaggg  
 1081 gtgcgtggac tctgtgccgc accccaagac ggagcagtga gcgcgaggat cc

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>interferon, alpha-inducible protein 27</b>	IFI27	NM_005532	42

1 g g g a a c a c a t c c a a g c t t a a g a c g g t g a g g t c a g c t t c a c a t t c t c a g g a  
 5 a c t c t c c t t c  
 61 t t t g g g t c t g g c t g a a g t t g a g g a t c t c t t a c t c t c t a g g c c a c g g a a t t  
 a a c c c g a g c a  
 121 g g c a t g g a g g c c t c t g c t c t c a c c t c a t c a g c a g t g a c c a g t g t g g c c a a  
 a g t g g t c a g g  
 10 181 g t g g c c t c t g c t c t g c c g t a g t t t t g c c c t g g c c a g g a t t g c t a c a g t  
 t g t g a t t g g a  
 241 g g a g t t g t g g c t g t g c c c a t g g t g e t c a g t g c c a t g g g c t t c a c t g c g g c  
 g g g a a t c g c c  
 301 t c g t c c t c c a t a g c a g c c a a g a t g a t g t c c g c g g c g g c c a t t g c c a a t g g  
 15 g g g t g g a g t t  
 361 g c c t c g g g c a g c t t t g t g g c t a c t c t g c a g t c a c t g g g a g c a a c t g g a c t  
 c t c c g g a t t g  
 421 a c c a a g t t c a t c c t g g g c t c a t t g g g t c t g c c a t t g c g g c t g t c a t t g c  
 g a g g t t c t a c  
 20 481 t a g c t c c c t g c c c c t c g c c c t g c a g a g a a g a g a a c c a t g c c a g g g g a g a a  
 g g c a c c c a g c  
 541 c a t c c t g a c c a g c g a g g a g c c a a c t a t a c c c a a a t a t a c c t g g g g t g a a a  
 t a t a c c a a a t  
 601 t c t g c a t c t c a g a g g a a a a t a a g a a t a a a g a t g a a t t g t t g c a a c t c t t c a a a a

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
<b>cytochrome P450, family 3, subfamily A, polypeptide 43</b>	CYP3A43	NM_057095	43

1 a c c t c t g g g c a g a g a a a c a a a g c t c t a t a t g c a c a g c c c a g c a a a g a g a c a  
 g c a c a c a g c t  
 30 61 g a a g a a a a a c t c a g a a g a c a g a g c t g a a a a g a a a a c t g t g a t g g a t c  
 t c a t t c c a a a  
 121 c t t t g c c a t g g a a c a t g g g t t c t t g t g g c t a c c a g c c t g g t a c t c c t c t  
 a t a t t t a t g g  
 181 g a c c c a t t c a c a t a a c t t t t a a g a a g c t g g g a a t t c c t g g g c c a a c c c  
 c t c t g c c t t t  
 35 241 t c t g g g a a c t a t t t t g t t c t a c c t a g g g g t c t t t t g g a a t t t t g a c a g a g  
 a a t g t a a t g a  
 301 a a a a t a c g g a g a a t g t g g g g g c t g a t g a g g g g c a a c a g c c c a t g c t g g  
 t c a t c a t g g a  
 361 t c c c g a c a t g a t c a a a a c a g t g t t a g t g a a a g a a t g t t a c t c t g t c t t c a  
 40 c a a a c c a g a t

421 gccttttaggt ccaatgggat ttctgaaaag tgccttaagt ttgctgaag  
 atgaagaatg  
 481 gaagagaata cgaacattgc tatctccagc tttcaccagt gtaaaattca  
 aggaaatggt  
 5 541 ccccatcatt tcccaatgtg gagatatggt ggtgagaagc ctgaggcagg  
 aagcagagaa  
 601 cagcaagtcc atcaacttga aagatttctt tggggcctac accatggatg  
 taatcactgg  
 661 cacattatth ggagtgaact tggattctct caacaatcca caagatccct  
 10 ttctgaaaaa  
 721 tatgaagaag cttttaaact tggatttttt ggatcccttt ttactcttaa  
 tatcactctt  
 781 tccatttctt accccagttt ttgaagccct aaatatcggg ttgtttccaa  
 aagatggtac  
 15 841 ccatttttta aaaaattcca ttgaaaggat gaaagaaagt cgcctcaaag  
 ataaacaaaa  
 901 gcacgagta gatttctttc aacagatgat cgactcccag aattccaaag  
 aaacaaagtc  
 961 ccataaagct ctgtctgatc tggagcttgt ggcccagtca attatcatca  
 20 tttttgctgc  
 1021 ctatgacaca actagcacca ctctcccctt cattatgtat gaactggcca  
 ctcaaccctga  
 1081 tgtccagcag aaactgcagg aggagattga cgcagtttta cccaataagg  
 cacctgtcac  
 25 1141 ctacgatgcc ctggtacaga tggagtacct tgacatggg gtgaatgaaa  
 cgctcagatt  
 1201 attcccagtt gttagtagag ttacgagagt ctgcaagaaa gatattgaaa  
 tcaatggagt  
 1261 gttcattccc aaagggttag cagtgatggg tccaatctat gctcttcacc  
 30 atgacccaaa  
 1321 gtactggaca gagcctgaga agttctgccc tgaaagggtc agtaagaaga  
 acaaggacag  
 1381 catagatctt tacagataca taccttttgg agctggacc cgaactgca  
 ttggcatgag  
 35 1441 gtttgctctc acaaacataa aacttgctgt cattagagca ctgcagaact  
 tctccttcaa  
 1501 accttgtaaa gagactcaga tcccactgaa attagacaat ctaccaattc  
 ttcaaccaga  
 1561 aaaacctatt gttctaaaag tgeacttaag agatgggatt acaagtggac  
 40 cctgactttc  
 1621 cctaaggact tccactttgt tcaagaaaagc tgtatcccag aacactagac  
 acttcaaatt  
 1681 gttttgtgaa taaaactcag aatgaagat gagcttaatt aacctagat  
 actgggtgaa  
 45 1741 taattagaaa ttctctacat tcattgagct ctcatgtct gggtagagta  
 ttacacgttg  
 1801 catactacaa agcaggtgac aatcaatgc caataagta cagtcattct  
 ctctagttct  
 1861 cataagacta tctccccgcc acctatagtt agtaccctca agtcctcctg  
 50 agctgtgatc  
 1921 agagaataaa catttctcaa caattttacc aacaattttt aatgaaaagg  
 aaaattatac  
 1981 ttgtgattct cgtagtgaca tttatattac atgttcatt tgtgatattc  
 tataataagt

2041 attatattga gaaagtcaac aagcacctct ttacaaaact gttatctgat  
 gtcttctgc  
 2101 atattaagga tgaatctaca gaattagatc aataaggatc aacaaataaa  
 tatttttggg  
 5 2161 catt

Gene Name	Gene Symbol	GenBank Accession #	SEQ ID NO
plakophilin 2	PKP2	NM_004572	44

1 gtggcggcctt cgcccgcgag tccagaggca ggcgagcagc tcggtcgccc ccaccggccc  
 61 catggcagcc cccggcgccc cagctgagta cggctacatc cggaccgtcc  
 10 tgggcccagca  
 121 gatcctggga caactggaca gctccagcct ggcgctgccc tccgaggcca  
 agctgaagct  
 181 ggcgggggagc agcggccgag ggcggccagac agtcaagagc ctgaggatcc  
 aggagcaggt  
 15 241 gcagcagacc ctgccccgga agggccgcag ctccgtgggc aacggaaatc  
 ttcaccgaac  
 301 cagcagtgtt cctgagtatg tctacaacct acacttggtt gaaaatgatt  
 ttggtggagg  
 361 cegttcccct gttcctaaaa cctatgacat gctaaaggct ggcacaactg  
 20 ccacttatga  
 421 aggtcgctgg ggaagaggaa cagcacagta cagctcccag aagtcctggtg  
 aagaaaggtc  
 481 cttgaggcat cctctgagga gactggagat ttctcctgac agcagcccgg  
 agagggctca  
 25 541 ctacacgcac agcgattacc agtacagcca gagaagccag gctgggcaca  
 ccctgcacca  
 601 ccaagaaagc aggcgggccc cctcctagt gccaccgaga tatgctcgtt  
 ccgagatcgt  
 661 gggggctcagc cgtgctggca ccacaagcag gcagcgccac tttgacacat  
 30 accacagaca  
 721 gtaccagcat ggctctgtta ggcacaccgt ttttgacagc atccctgcca  
 acccgcccct  
 781 gctcacgtac cccaggccag ggaccagccg cagcatgggc aacctcttgg  
 agaaggagaa  
 35 841 ctacctgacg gcagggctca ctgtcgggca ggtcaggccg ctggtgcccc  
 tgcagcccgt  
 901 cactcagaac agggcttcca ggtcctcctg gcatcagagc tccttcaca  
 gcaccgcac  
 961 gctgagggaa gctgggcca gtgtcgccgt ggattccagc gggaggagag  
 40 cgcacttgac  
 1021 tgtcggccag gcggccgcag ggggaagtgg gaatctgctc actgagagaa  
 gcactttcac  
 1081 tgactcccag ctggggaatg cagacatgga gatgactctg gagcgagcag  
 tgagtatgct  
 45 1141 cgaggcagac cacatgccg catccaggat ttctgctgca gctactttca  
 tacagcacga  
 1201 gtgcttccag aatctgaag ctcggaagag ggttaaccag cttcgtggca  
 tcctcaagct  
 1261 tctgcagctc ctaaaagttc agaatgaaga cgttcagcga gctgtgtgtg  
 50 gggccttgag

1321 aaacttagta tttgaagaca atgacaacaa attggagggtg gctgaactaa  
 atgggggtacc  
 1381 tcggctgctc caggtgctga agcaaaccag agacttggag actaaaaaac  
 aaataacaga  
 5 1441 ccatacagtc aatttaagaa gtaggaatgg ctggccgggc gcggtggctc  
 acgcctgtaa  
 1501 tcccagcact ttgggaggcc aaggcgggcg gatcacgagg tcaggagttc  
 gagaccagcc  
 1561 tgaccaacat ggtttctgtt ggaatttgtc atctaatac aaactcaaga  
 10 atctcatgat  
 1621 aacagaagca ttgcttacgc tgacggagaa tatcatcatc cccttttctg  
 ggtggcctga  
 1681 aggagactac ccaaaagcaa atggtttgtc cgattttgac atattctaca  
 acgtcactgg  
 15 1741 atgcctaaga aacatgagtt ctgctggcgc tgatgggaga aaagcgatga  
 gaagatgtga  
 1801 cggactcatt gactcactgg tccattatgt cagaggaacc attgcagatt  
 accagccaga  
 1861 tgacaaggcc acggagaatt gtgtgtgcat tcttcataac ctctcctacc  
 20 agctggaggc  
 1921 agagctccca gagaaatatt cccagaatat ctatattcaa aaccggaata  
 tccagactga  
 1981 caacaacaaa agtattggat gttttggcag tcgaagcagg aaagtaaaag  
 agcaatacca  
 25 2041 ggacgtgccg atgccggagg aaaagagcaa ccccaagggc gtggagtggc  
 tgtggcattc  
 2101 cattgttata aggatgtatc tgtccttgat cgccaaaagt gtccgcaact  
 acacacaaga  
 2161 agcatcctta ggagctctgc agaacctcac ggccggaagt ggaccaatgc  
 30 cgacatcagt  
 2221 ggctcagaca gttgtccaga aggaaagtgg cctgcagcac acccgaaaga  
 tgctgcatgt  
 2281 tggtgaccca agtgtgaaaa agacagccat ctcgctgctg aggaatctgt  
 cccggaatct  
 35 2341 ttctctgcag aatgaaattg ccaaagaaac tctccctgat ttggtttcca  
 tcattcctga  
 2401 cacagtcccg agtactgacc ttctcattga aactacagcc tctgctgtt  
 acacattgaa  
 2461 caacataatc caaaacagtt accagaatgc acgcgacctt ctaaaccaccg  
 40 ggggcatcca  
 2521 gaaaattatg gccattagtg caggcgatgc ctatgcctcc aacaaagcaa  
 gtaaagctgc  
 2581 ttccgtcctt ctgtattctc tgtgggcaca cacggaactg catcatgctt  
 acaagaaggc  
 45 2641 tcagtttaag aagacagatt ttgtcaacag ccggactgcc aaagcctacc  
 actcccttaa  
 2701 agactgagga aatgacaaa gtattctcgg ctgcaaaaat ccccaaagga  
 aaacacctat  
 2761 ttttctacta cccagcccaa gaaacctcaa aagcatgcct tgtttctatc  
 50 cttctctatt  
 2821 tccgtggctc cctgaatcca gaaaacaaat agaacataat tttatgagtc  
 tccagaaga  
 2881 cctttgcaag tttgccacca gtagataccg gccacaggct cgacaaatag  
 tggctttgt

2941 tattagggct tatggtacat ggcttcctgg aatcaaaatg tgaattcatg  
 tgggaagggac  
 3001 attaatccaa taaataagga aagaagctgt tgcattactg ggattttaaa  
 agtttgattt  
 5 3061 acattttatat tccttttctg gttcccatgt tttgtcactc atgtgcacat  
 tgcttcgcca  
 3121 ttgggctcc agtgtattgt tctgcagtgt tgaaacagaa tggaaatgac  
 aagaaatatac  
 3181 tgcagttatc caggagaaag tataatggca aaattattgg tttctttctt  
 10 tactttgtgc  
 3241 ttgtttttat cccttggggt tgtttttctc tgatttttaa ataaacttaa  
 gaaatttaga  
 3301 ttacagagta tgcattgactg taagaaaaag aaattgagag gaagtgatca  
 tagcaaatta  
 15 3361 aagaagtctt ttctcccag aacttaaagt aaaataaaaa ataaataaat  
 aaataaaatc  
 3421 ttttccacag agaaaggcaa ctgtgatgat aaaatttaac gttcccccaa  
 aactgagtc  
 3481 aatgagattt ttctcaggag atactttacc tataacaacg ccgttaaate  
 20 caaatctctt  
 3541 ctaaactgatg gcattctatg taatgccttt cctggacttt tttggccact  
 gccctggact  
 3601 agtgaaagaa tggactctat ctttatctgc aagaggaact aaggccttct  
 atcagactgc  
 25 3661 ctggccagcc tggggcactg aaaatacggc tcatgttaat gagttacatt  
 atcagccagc  
 3721 ccagccttgc ccaccattta agaaatatca cagagccact agatctcata  
 tgatcttctt  
 3781 caagccatta ttttaactca agaaaactct agagaagaaa agtgaagaag  
 30 tcatgttgaa  
 3841 gaagatgtaa gaatgtgtca agaccatcca gaaatgatat gagaataact  
 gatattttaa  
 3901 atgggtgaca tcatccagcg aatgaatct acattaaatg ttgttttaac  
 tgcgctatga  
 35 3961 ttaaaacat tcatatagag ttagtcttta caactactat tctgttattt  
 ttttttttaa  
 4021 tctgacaaca tttgtcctaa gtaagataag caaaaaaatt cttcaactcc  
 ttttggaag  
 4081 aaaactgtaa cagaaaataa attttgaatg tgtacttaag tctttattat  
 40 atttgaagca  
 4141 attttttttc aattttaaaa gctgaatgaa gacaacttag gttgctaacc  
 tagttcaaaa  
 4201 tgaaattatt tagataccaa tttttaaaat actggagaga atttatatgt  
 ctttttccag  
 45 4261 agttctgatg ataagcattt ggagtgcatt tattcctcca gataataaat  
 gtgtgttcag  
 4321 aactttttgt gttttttaag gcattaataa agccttcgat aatattaaat acaaaatgaa

WHAT IS CLAIMED IS:

1. A method for classifying a cancer in a patient, comprising comparing the expression levels of at least four biomarker genes in the cancer to expression level to a first or second set of expression level threshold values for the biomarker genes, and indicating that the cancer is sensitive to a HDAC inhibitor if the expression levels of the biomarker genes are lower than the first set of expression level threshold values, or indicating that the cancer is resistant to a HDAC inhibitor if the expression levels are greater than the second set of expression level threshold values, wherein the at least four biomarker genes are selected from PTPN3, ABCC3, SARG, PPAP2C, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBP1, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, Tmprss2, GDA, MST1R, ITGB4, ANXA3, CCL15, DPEP1, NOXO1, IFI27, CYP3A43, and PKP2.
2. The method of claim 1, wherein at least one of the at least four marker genes are selected from DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP1.
3. The method of claim 1, wherein the at least four biomarker genes include at least one of DEFA6, RAB25, TM4SF4, or IL18.
4. The method of claim 1, wherein the at least four biomarker genes include DEFA6, ITGB4, TM4SF3, SYK, PPAP2C, and RAB25.
5. The method of claim 1, wherein the at least four biomarker genes include DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP1.
6. The method of claim 1, wherein one or more of the expression levels is an mRNA expression level.
7. The method of claim 1, wherein one or more of the expression levels is a polypeptide expression level.
8. The method of claim 1, wherein the cancer is a colon cancer.
9. The method of claim 1, further comprising determining the level of expression of the at least four biomarker genes in the cancer prior to the step of comparing.

10. The method of claim 1, wherein the HDAC inhibitor is PCI-24781.
11. The method of claim 1, wherein the expression levels of the at least four biomarker genes are compared to the first set and the second set of biomarker gene expression level threshold level values.
12. A method for classifying a cancer in a patient, comprising determining the expression levels of at least four biomarker genes in the cancer, comparing the expression levels of the at least four biomarker genes in the cancer to expression level to a first or second set of expression level threshold values for the biomarker genes, and indicating that the cancer is sensitive to a HDAC inhibitor if the expression levels of the biomarker genes are lower than the first set of expression level threshold values, or indicating that the cancer is resistant to a HDAC inhibitor if the expression levels are greater than the second set of expression level threshold values, wherein the at least four biomarker genes are selected from PTPN3, ABCC3, SARG, PPAP2C, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBP1, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, Tmprss2, GDA, MST1R, ITGB4, ANXA3, CCL15, DPEP1, NOXO1, IFI27, CYP3A43, and PKP2.
13. The method of claim 12, wherein at least one of the at least four marker genes are selected from DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP1.
14. The method of claim 12, wherein the at least four biomarker genes include at least one of DEFA6, RAB25, TM4SF4, or IL18.
15. The method of claim 12, wherein the at least four biomarker genes include DEFA6, ITGB4, TM4SF3, SYK, PPAP2C, and RAB25.
16. The method of claim 12, wherein the at least four biomarker genes include DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP1.
17. The method of claim 12, wherein one or more of the expression levels is an mRNA expression level.
18. The method of claim 12, wherein one or more of the expression levels is a polypeptide expression level.
19. The method of claim 12, wherein the cancer is a colon cancer.

20. The method of claim 12, further comprising prescribing or administering an HDAC inhibitor to the patient based on the comparison.
21. The method of claim 12, wherein the HDAC inhibitor is PCI-24781.
22. The method of claim 12, wherein the expression levels of the at least four biomarker genes are compared to the first set and the second set of biomarker gene expression level threshold level values.
23. An isolated population of nucleic acids comprising a plurality of nucleic acids derived from a cancer cell, wherein the cancer cell is a type of cancer cell that is sensitive to an HDAC inhibitor compound.
24. The isolated population of claim 23, wherein the isolated population contains RNAs.
25. The isolated population of claim 23, wherein the isolated population contains cDNAs.
26. The isolated population of claim 23, wherein the HDAC inhibitor is PCI-24781.
27. The isolated population of claim 23, wherein the cell was isolated from a population of cells grown *in vitro*.
28. The isolated population of claim 23, wherein the cancer cell is a colon carcinoma cell.
29. The isolated population of claim 28, wherein the colon carcinoma cell is derived from colon carcinoma R1059261097, R4498160614, R5456781761, R7424107588, or R0948311023.
30. The isolated population of claim 23, wherein the nucleotide sequences of at least four of DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, or DPEP1 are represented in the isolated population.
31. An isolated population of nucleic acids comprising a plurality of nucleic acids derived from a cancer cell, wherein the cancer cell is a type of cancer cell that is resistant to an HDAC inhibitor compound.
32. The isolated population of claim 31, wherein the isolated population contains RNAs.
33. The isolated population of claim 31, wherein the isolated population contains cDNAs.
34. The isolated population of claim 31, wherein the HDAC inhibitor is PCI-24781.

35. The isolated population of claim 31, wherein the cell was isolated from a population of cells grown *in vitro*.
36. The isolated population of claim 31, wherein the cancer cell is a colon carcinoma cell.
37. The isolated population of claim 36, wherein the colon carcinoma cell is derived from colon carcinoma R5247682266, R9866135153, R1078103114, or R4712781606.
38. The isolated population of claim 31, wherein the nucleotide sequences of at least four of DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, or DPEP1 are represented in the isolated population.
39. A kit comprising the isolated population of claim 31 and an insert indicating the ratio of a biomarker gene nucleic acid level in the population to an internal expression control gene nucleic acid level in the population.
40. The kit of claim 31 and an insert indicating the ratio of a biomarker gene nucleic acid level in the population to a nucleic acid level of the biomarker gene in a population of nucleic acids derived from a cancer cell, wherein the cancer cell is a type of cancer cell that is sensitive to the HDAC inhibitor compound.
41. A method for generating an expression level reference population of nucleic acids for expression profiling, comprising deriving an isolated population of nucleic acids from a cancer cell, wherein the cancer cell is a type of cancer cell that is sensitive to an HDAC inhibitor compound.
42. The method of claim 41, wherein the isolated population contains RNAs.
43. The method of claim 41, wherein the isolated population contains cDNAs.
44. The method of claim 41, wherein the HDAC inhibitor compound is PCI-24781.
45. The method of claim 41, wherein the cancer cell is present in a biopsy sample.
46. The method of claim 41, wherein the cancer cell is present in a population of cells grown *in vitro*.
47. The method of claim 41, wherein the cancer cell is a colon carcinoma cell.
48. The method of claim 47, wherein the carcinoma cell is derived from colon carcinoma R1059261097, R4498160614, R5456781761, R7424107588, or R0948311023.

49. The method of claim 41, wherein the nucleotide sequences of at least four of DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, or DPEP1 are represented in the isolated population.
50. The method of claim 41, further comprising determining, prior to the isolating step, that the type of cancer cell is sensitive to an HDAC inhibitor compound.
51. The method of claim 50, wherein the cancer cell determined to be sensitive to an HDAC inhibitor compound is determined to be resistant to an HDAC inhibitor compound *in vitro*.
52. The method of claim 51, wherein the HDAC inhibitor compound is PCI-24781.
53. A method for generating an expression level reference sample for expression profiling, comprising deriving an isolated population of nucleic acids from a cancer cell, wherein the cancer cell is a type of cancer cell that is resistant to an HDAC inhibitor compound.
54. The method of claim 53, wherein the isolated population contains RNAs.
55. The method of claim 53, wherein the isolated population contains cDNAs.
56. The method of claim 53, wherein the HDAC inhibitor compound is PCI-24781.
57. The method of claim 52, wherein the cell is present in a biopsy sample.
58. The method of claim 53, wherein the cell is present in a population of cells grown *in vitro*.
59. The method of claim 53, wherein the cancer cell is a colon carcinoma cell.
60. The method of claim 59, wherein the carcinoma cell is derived from colon carcinoma R5247682266, R9866135153, R1078103114, or R4712781606.
61. The method of claim 53, wherein the nucleotide sequences of at least four of DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, or DPEP1 are represented in the isolated population.
62. The method of claim 53, further comprising, prior to the isolating step, determining that the type of cancer cell is sensitive to an HDAC inhibitor compound.
63. The method of claim 51, wherein the HDAC inhibitor compound is PCI-24781.
64. A human cancer cell line that is resistant to an HDAC inhibitor compound *in vitro*.

65. The human cancer cell line of claim 64, wherein the human cell line expresses DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP1.
66. The human cancer cell line of claim 64, wherein the HDAC inhibitor compound is PCI-24781.
67. The human cancer cell line of claim 66, wherein the human cell line is resistant to a PCI-24781 concentration of at least about 1  $\mu$ M.
68. The human cancer cell line of claim 64, wherein the human cancer cell line is a colon carcinoma cell line.
69. The colon carcinoma cell line of claim 68, wherein the colon carcinoma cell line is R5247682266, R9866135153, R1078103114, or R4712781606.
70. A method for increasing the likelihood of therapeutically effective treatment of a cancer with an HDAC inhibitor, comprising providing an indication that a cancer in a patient is sensitive to treatment with an HDAC inhibitor if expression levels of at least four biomarker genes in a sample from the patient's cancer are lower than expression level threshold values for the four biomarker genes, or providing an indication that the cancer is resistant to treatment with the HDAC inhibitor if the expression levels of the biomarker genes are higher than the expression level threshold values, wherein the at least four biomarker genes are selected from PTPN3, ABCC3, SARG, PPAP2C, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBP1, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, Tmprss2, GDA, MST1R, ITGB4, ANXA3, CCL15, DPEP1, NOXO1, IFI27, CYP3A43, and PKP2, whereby the likelihood of therapeutically effective treatment of the cancer with the HDAC inhibitor is increased.
71. The method of claim 70, wherein the indication is provided in a digital medium.
72. The method of claim 70, wherein the indication is provided in a hardcopy medium.
73. The method of claim 70, wherein the indication is a biomedical publication reference.
74. The method of claim 70, wherein the indication refers to expression levels of at least two of the biomarker genes.
75. The method of claim 70, wherein the at least four biomarker genes include DEFA6, RAB25, TM4SF4, or IL18.

76. The method of claim 70, wherein the at least four biomarker genes include DEFA6, ITGB4, TM4SF3, SYK, PPAP2C, and RAB25.
77. The method of claim 70, wherein the at least four biomarker genes include DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP1.
78. The method of claim 70, wherein the cancer is colon cancer.
79. The method of claim 70, wherein the HDAC inhibitor is PCI-24781.
80. A method for optimizing selection of an anti-cancer agent for treating a cancer in combination with an HDAC inhibitor compound, the method comprising
- (i) comparing a first set of biomarker genes the expression of which is correlated to resistance or sensitivity of the cancer to the anti-cancer agent to a second set of biomarker genes the expression of which is correlated with resistance to the HDAC inhibitor compound; and
  - (ii) selecting the anti-cancer agent for treatment of the cancer in combination with the HDAC inhibitor if the biomarker genes in the first set are different from the biomarker genes in the second set, wherein
- the biomarker genes in the second set are DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP1.
81. The method of claim 80, further comprising comparing the expression level of the second set of biomarker genes in a plurality of cancer cells treated with the HDAC inhibitor together with a second anti-cancer agent.
82. An indication of the likelihood of a therapeutically effective treatment of a cancer with an HDAC inhibitor compound, comprising a means of communicating an interpretation of expression levels of at least four biomarker genes selected from DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP1.
83. The indication of claim 82, further comprising the expression levels of the at least four biomarker genes.
84. The indication of claim 82, wherein the means of communicating is a paper document or an electronic document.

85. The indication of claim 82, wherein the interpretation includes a biomedical publication reference.

86. The indication of claim 82, wherein the interpretation includes a graph.

87. The indication of claim 82, wherein the interpretation includes information that indicates that a cancer in a patient is sensitive to treatment with an HDAC inhibitor if expression levels of the biomarker genes in a sample from the patient's cancer are lower than expression level threshold values for the four biomarker genes, or information that indicates that the cancer is resistant to treatment with the HDAC inhibitor if the expression levels of the biomarker genes are higher than the expression level threshold values.

88. A method for determining the likelihood of effectively treating a cancer in a patient with an HDAC inhibitor compound, comprising

(i) determining in the cancer the expression levels of at least four biomarker genes selected from DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP; and

(ii) comparing the expression levels of that at least four biomarker genes in the cancer to expression levels of the at least four biomarker genes in an expression level reference sample derived from cancer cells previously determined to be resistant to the HDAC inhibitor compound, wherein the likelihood of effectively treating the cancer is higher if the expression level of the at least four biomarkers in the cancer from the patient is lower than the expression levels of the biomarker genes in the expression level reference sample.

89. The method of claim 88, further comprising selecting an anti-cancer agent other than an HDAC inhibitor compound for treating the cancer.

90. A method for classifying a cancer in a patient, comprising comparing the expression levels of at least four biomarker genes in the cancer to a first or second set of expression level values for the biomarker genes, and for each comparison assigning a probability to the biomarker gene expression level that the cancer in the patient is resistant to a histone deacetylase inhibitor compound, wherein

(i) the first set of expression level values were measured in cancer cells determined to be resistant to the histone deacetylase inhibitor compound;

(ii) the second set of expression level values were measured in cancer cells determined to be sensitive to the histone deacetylase inhibitor compound;

(iii) the assigned probability is inversely proportional to a negative deviation of the biomarker gene expression level from the first set of expression level values and directly proportional to a positive deviation of the biomarker gene expression level from the second set of expression level values; and

(iv) the at least four biomarker genes are selected from PTPN3, ABCC3, SARG, PPAP2C, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBP1, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMPRSS2, GDA, MST1R, ITGB4, ANXA3, CCL15, DPEP1, NOXO1, IFI27, CYP3A43, and PKP2.

91. A method for classifying a population of cells, comprising comparing the expression levels of at least four biomarker genes in the population of cells to a first or second set of expression level threshold values for the biomarker genes, and indicating that the population of cells is sensitive to a HDAC inhibitor if the expression levels of the biomarker genes are lower than the first set of expression level threshold values, or indicating that the population of cells is resistant to a HDAC inhibitor if the expression levels are greater than the second set of expression level threshold values, wherein the at least four biomarker genes are selected from PTPN3, ABCC3, SARG, PPAP2C, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBP1, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMPRSS2, GDA, MST1R, ITGB4, ANXA3, CCL15, DPEP1, NOXO1, IFI27, CYP3A43, and PKP2.

92. A method for determining HDAC inhibition *in vivo*, comprising determining the expression level of an HDAC inhibitor-responsive biomarker gene in a biological sample obtained from a subject after the subject had been administered an HDAC inhibitor compound, wherein the HDAC inhibitor-responsive biomarker genes are any of the genes listed in Table 5.

93. A method for determining the most responsive tissues and the tumors derived therefrom to an HDAC inhibitor, comprising

(i) providing a first tissue of the tissue type (including blood) at a first time point and administration of HDAC inhibitor compound to the first tissue by any applicable route at a first time point,

(ii) providing a second tissue of the tissue type (including blood) at a second time point and administration of HDAC inhibitor compound to the second tissue by any applicable route at a second time point,

(iii) determining expression profiles in the first and second tissues for any of the genes listed in Table 5.

94. A method for classifying one or more cells, comprising determining the expression levels of no more than four to fifty biomarker genes in the one or more cells, wherein at least four of the biomarker genes are selected from PTPN3, ABCC3, SARG, PPAP2C, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBP1, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, Tmprss2, GDA, MST1R, ITGB4, ANXA3, CCL15, DPEP1, NOXO1, IFI27, CYP3A43, and PKP2.

95. The method of claim 94, further comprising comparing the expression levels to a first or second set of expression level threshold values for the biomarker genes, and indicating that the one or more cells are sensitive to a HDAC inhibitor if the expression levels of the biomarker genes are lower than the first set of expression level threshold values, or indicating that the one or more cells are resistant to a HDAC inhibitor if the expression levels are greater than the second set of expression level threshold values.

96. The method of claim 94, wherein the one or more cells are cancer cells.

97. The method of claim 94, wherein the four to fifty biomarker comprises one or more genes selected from DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP

98. The method of claim 94, comprising determining the expression levels of no more than four to twenty biomarker genes.

99. The method of claim 98, comprising determining the expression levels of no more than four biomarker genes.

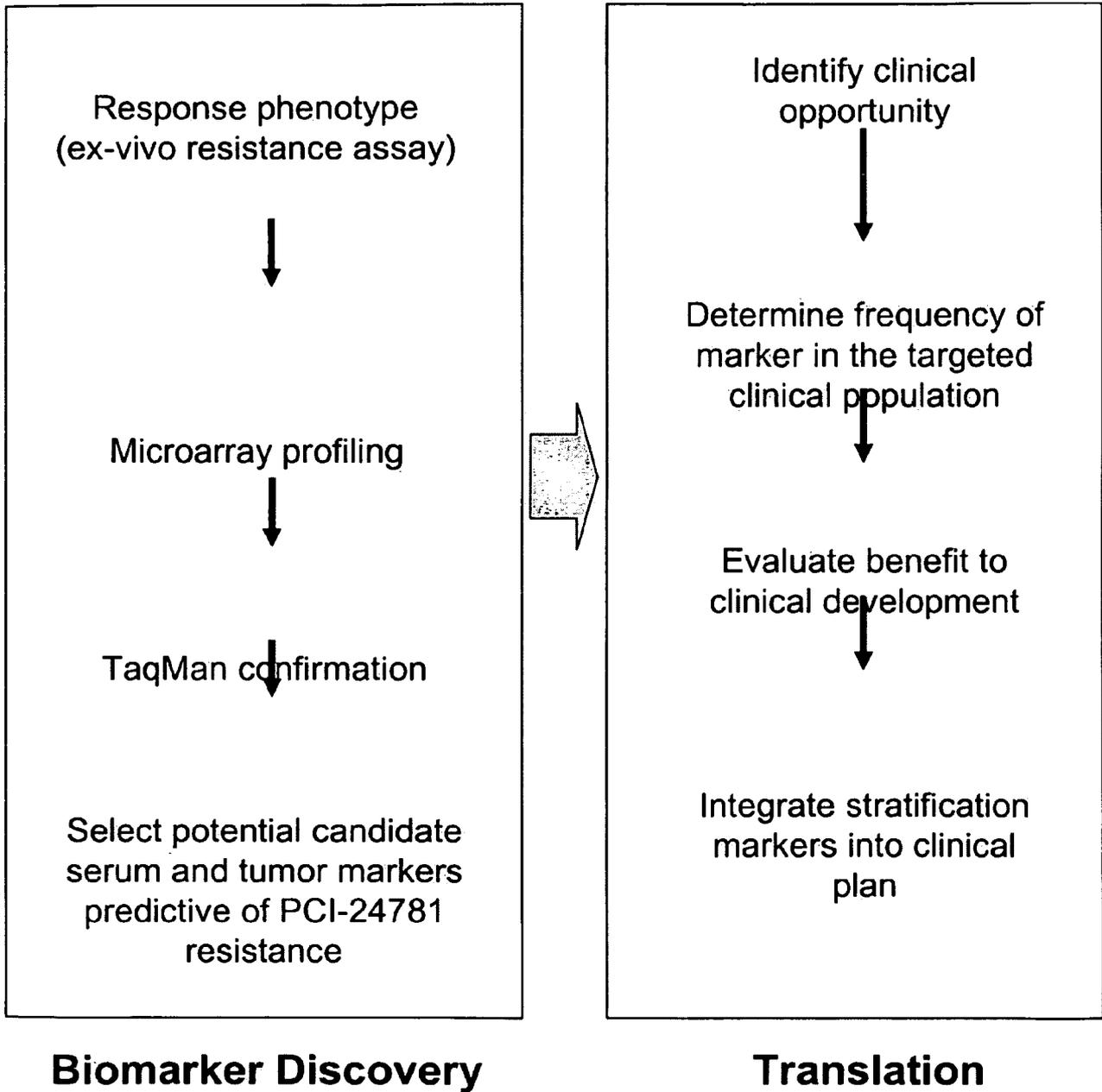
100. The method of claim 99, wherein the four biomarker genes consist of DEFA6, RAB25, TM4SF4, and IL18.

101. A nucleic acid hybridization array, comprising nucleic acid probes that hybridize under high stringency hybridization conditions to nucleic acids of no more than four to fifty biomarker genes, wherein at least four of the biomarker genes are selected from PTPN3, ABCC3, SARG, PPAP2C, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBP1, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, Tmprss2, GDA, MST1R, ITGB4, ANXA3, CCL15, DPEP1, NOXO1, IFI27, CYP3A43, and PKP2.

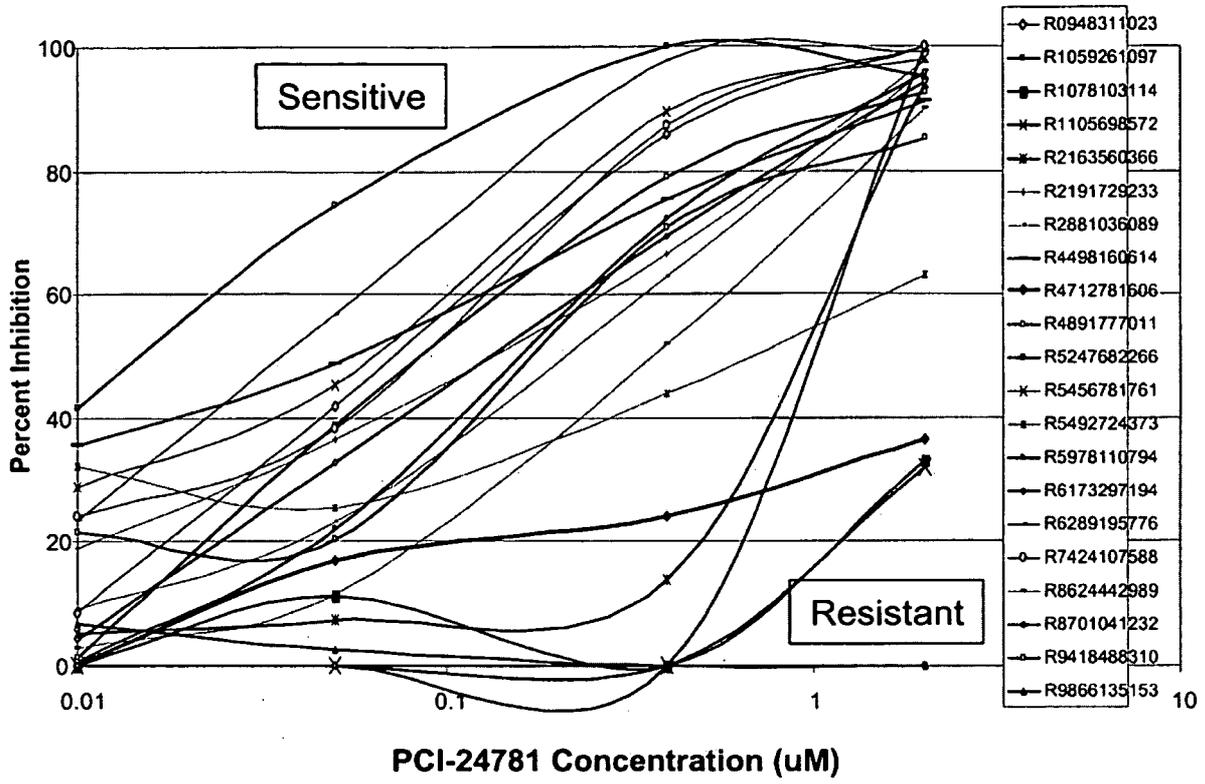
102. The nucleic acid hybridization array of claim 101, wherein the at least four biomarker genes are selected from DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEP.

103. The nucleic acid hybridization array of claim 102, wherein the at least four biomarker genes consist of DEFA6, RAB25, TM4SF4, and IL18.

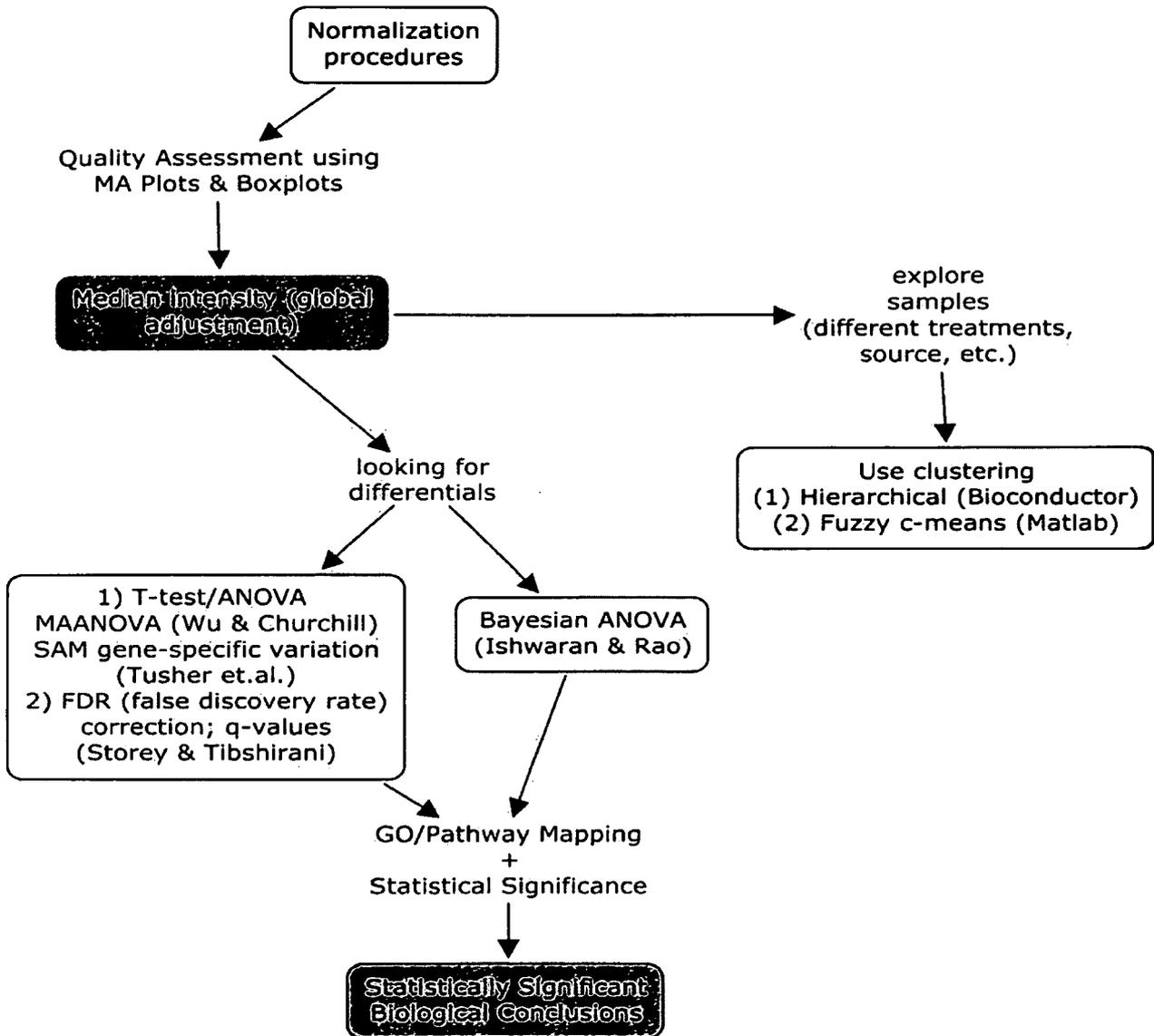
**Fig. 1**



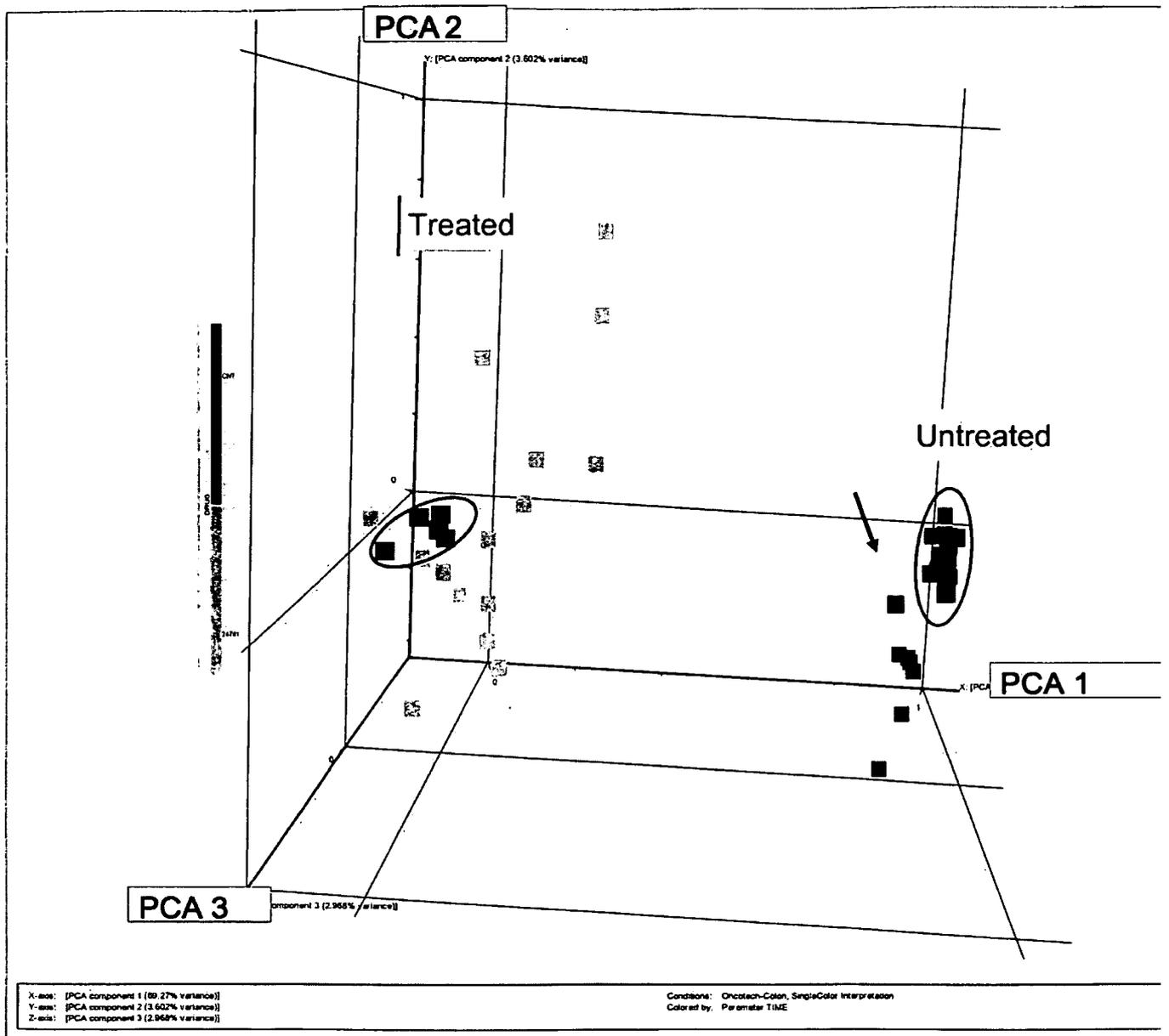
# Fig. 2



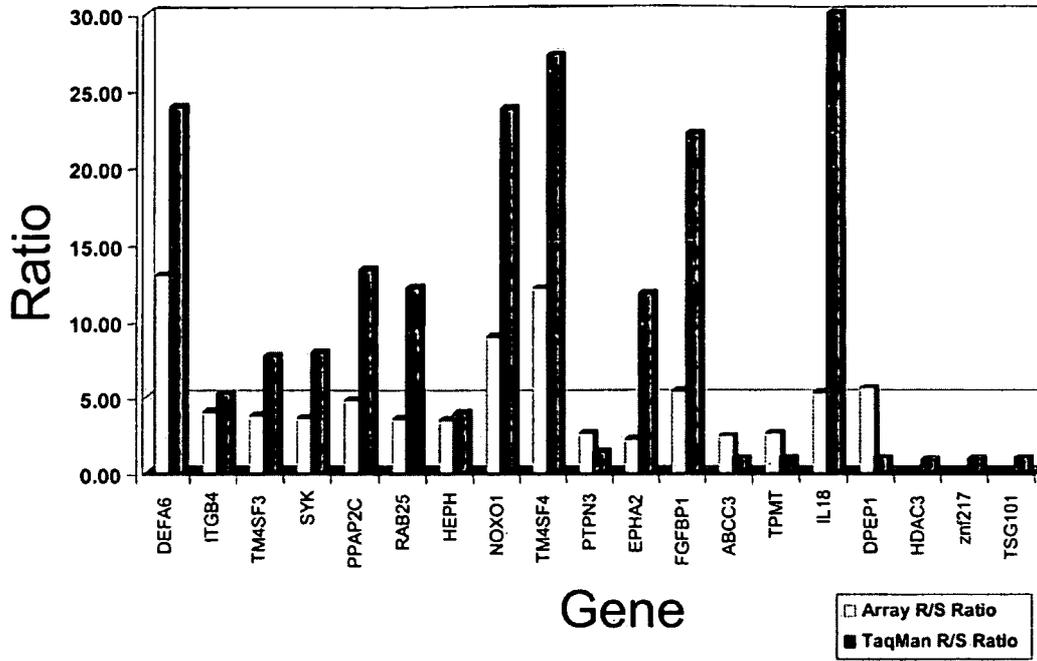
# Fig. 3



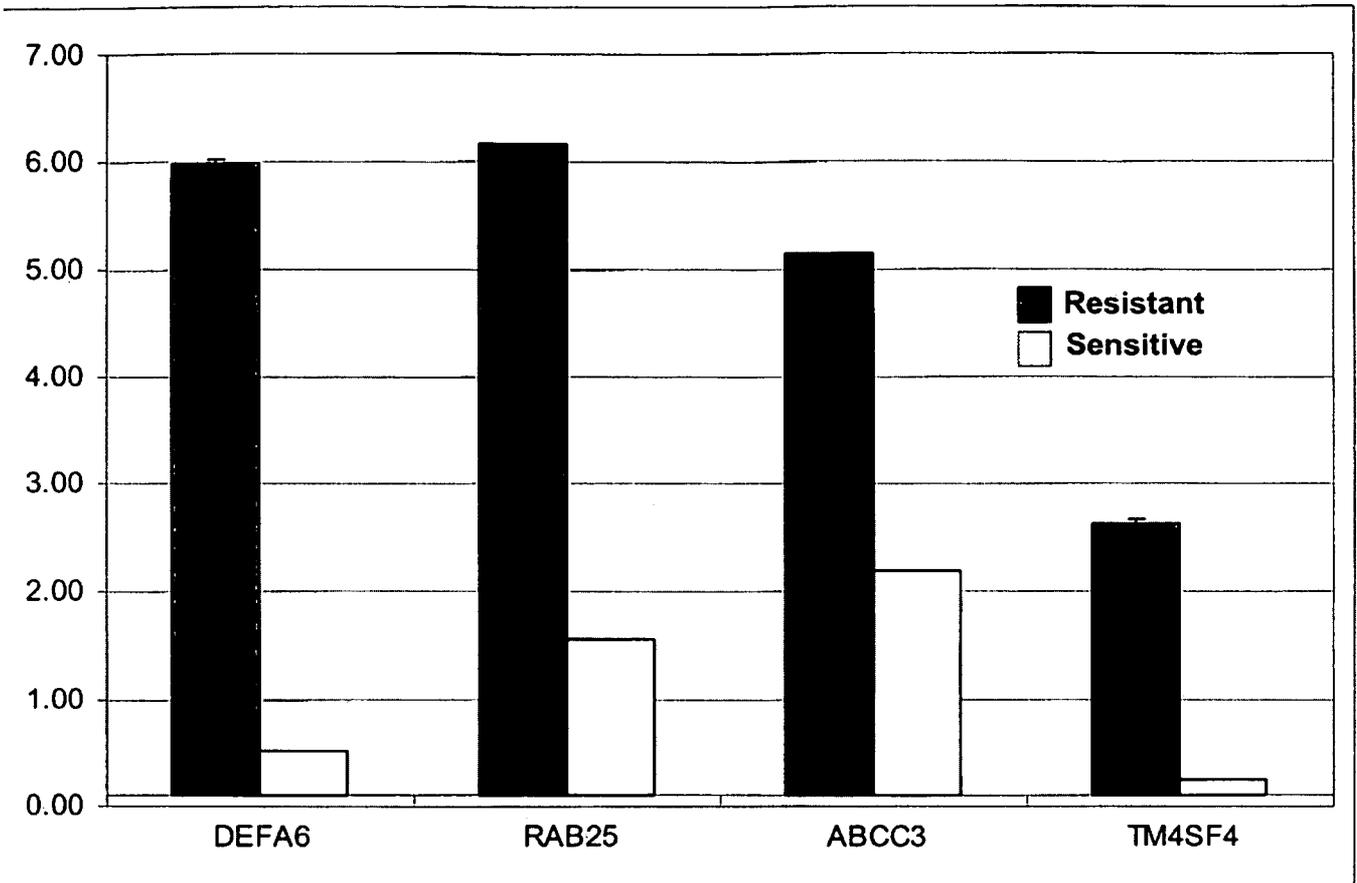
# Fig. 4



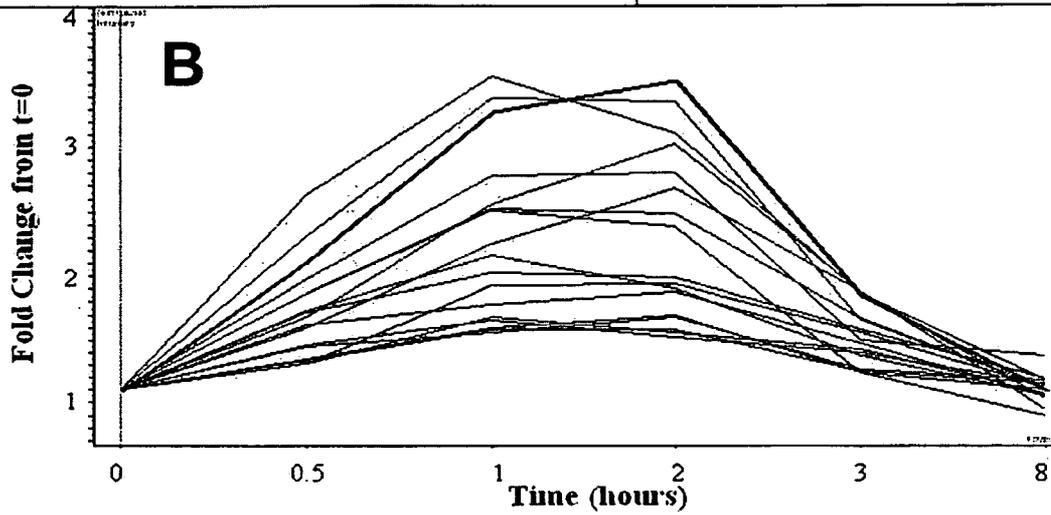
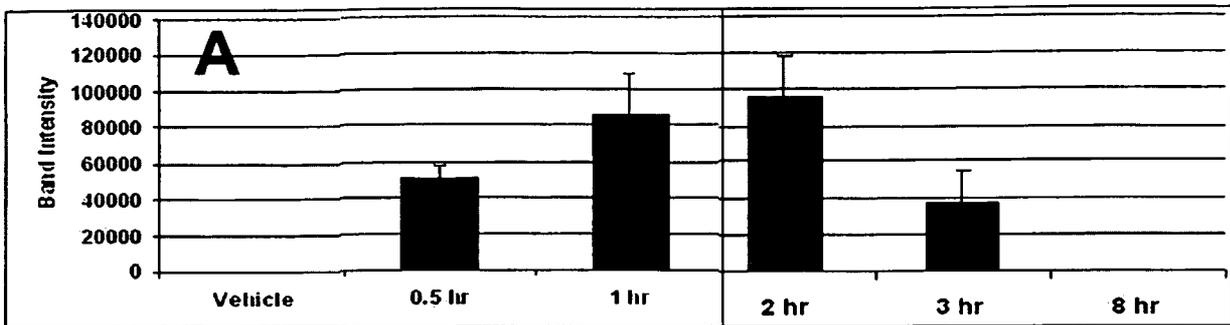
**Fig. 5**



# Fig. 6

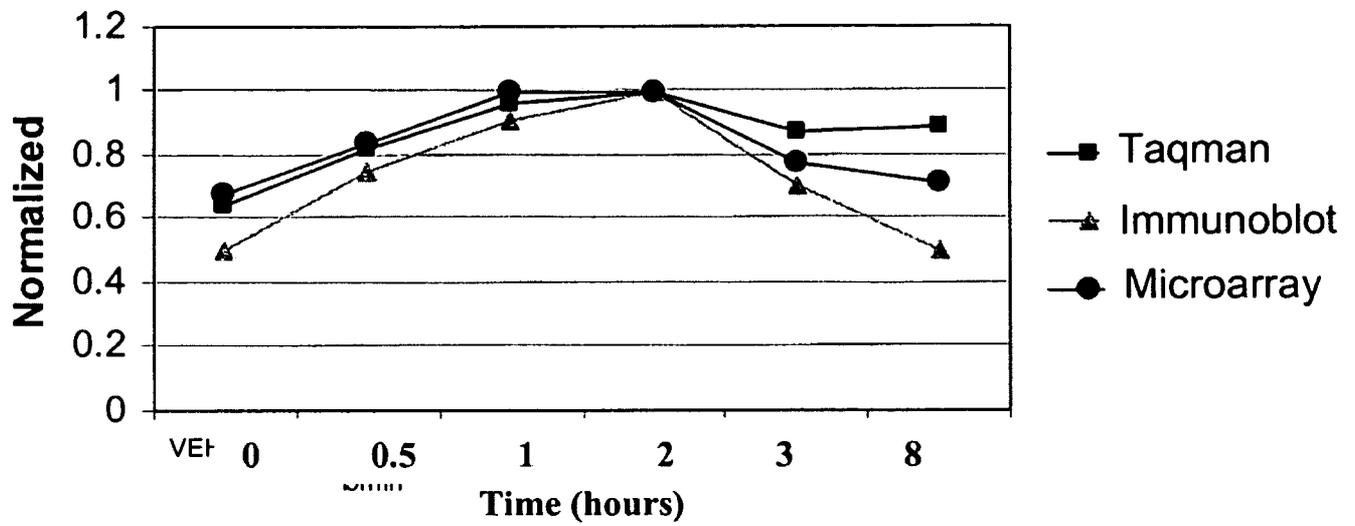


**Fig. 7**

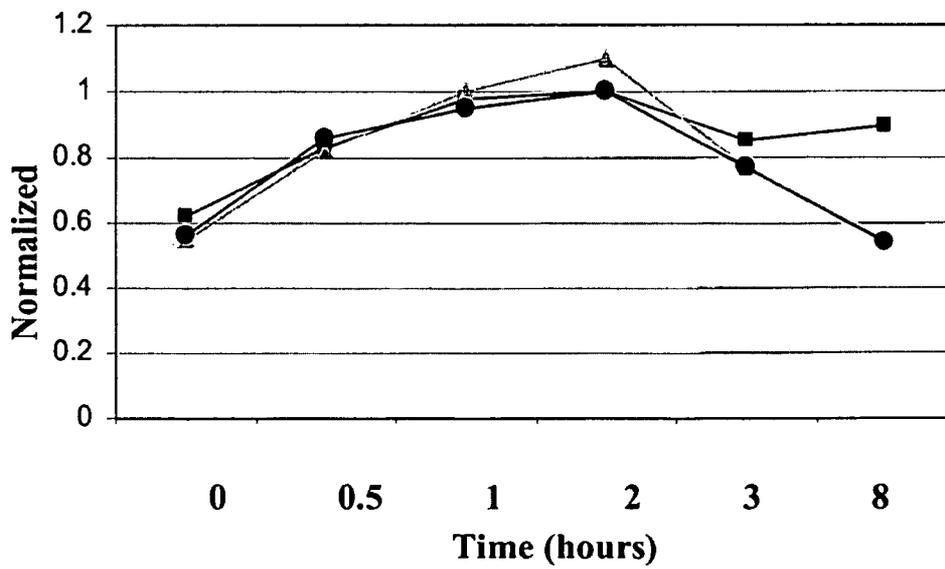


# Fig. 8

## FGF 15

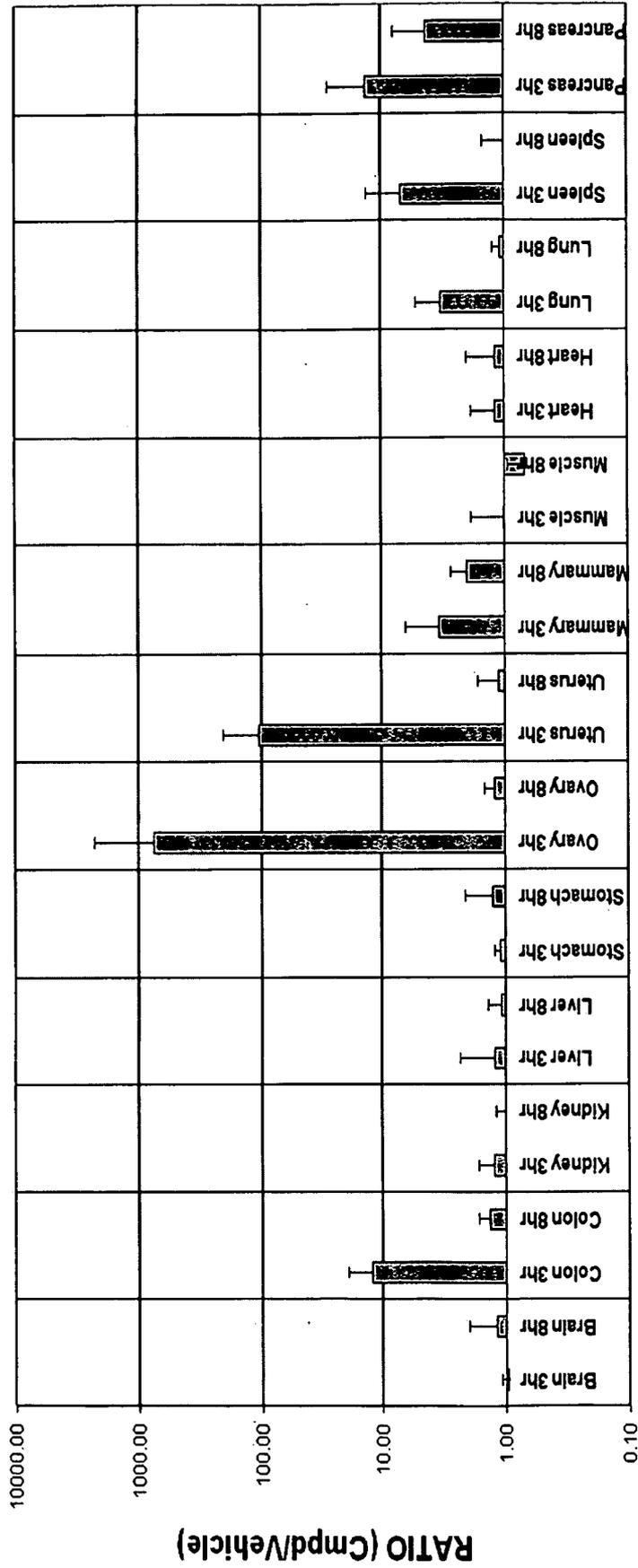


## Synaptogyrin 2



**Fig. 9**

**Marker Set Average**



**Tissues (Time,hrs)**

Brain Colon Kidney Liver Stomach Ovary Uterus Mammary Muscle Heart Lung Spleen Pancreas

**Fig. 10**

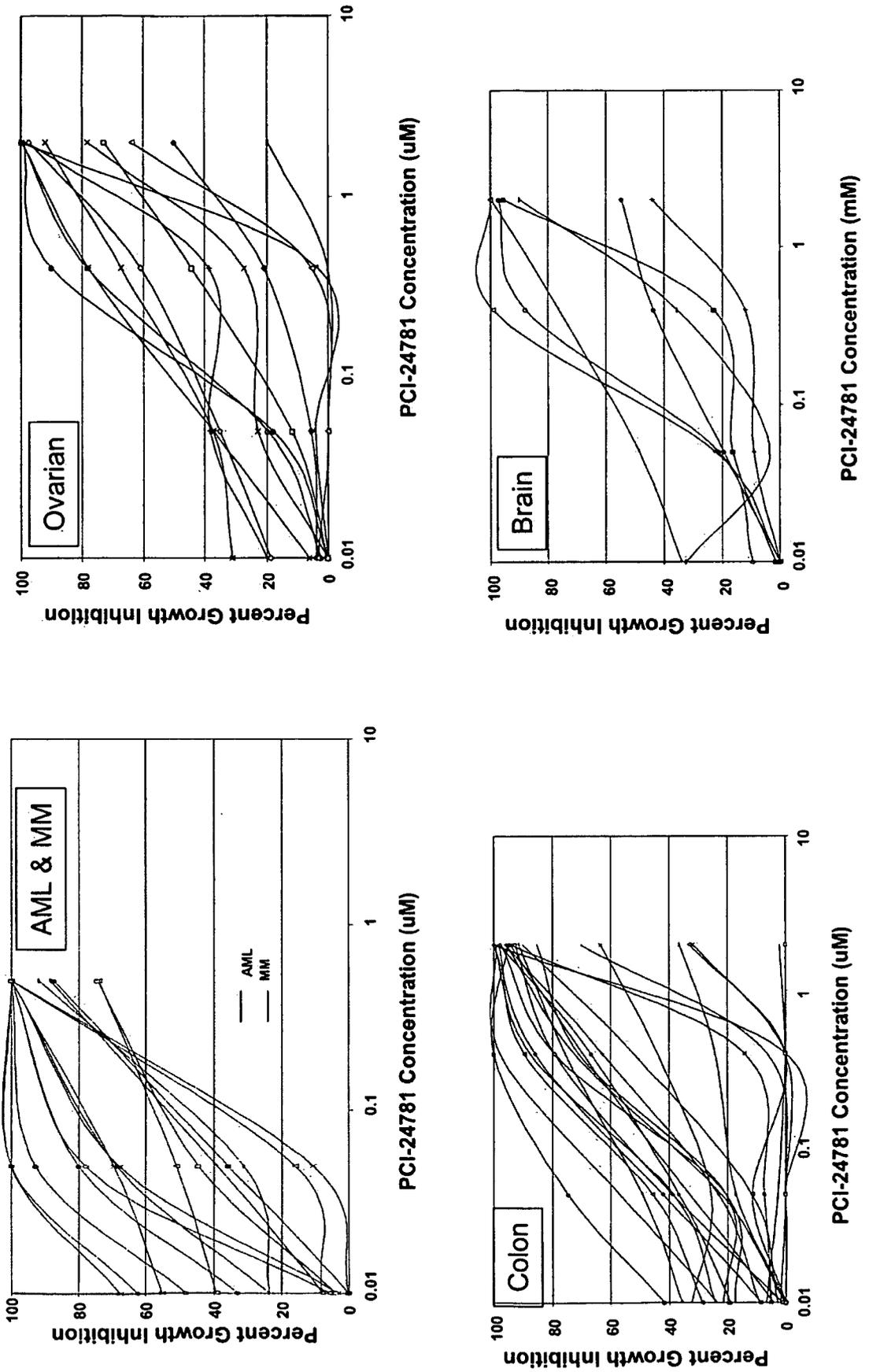
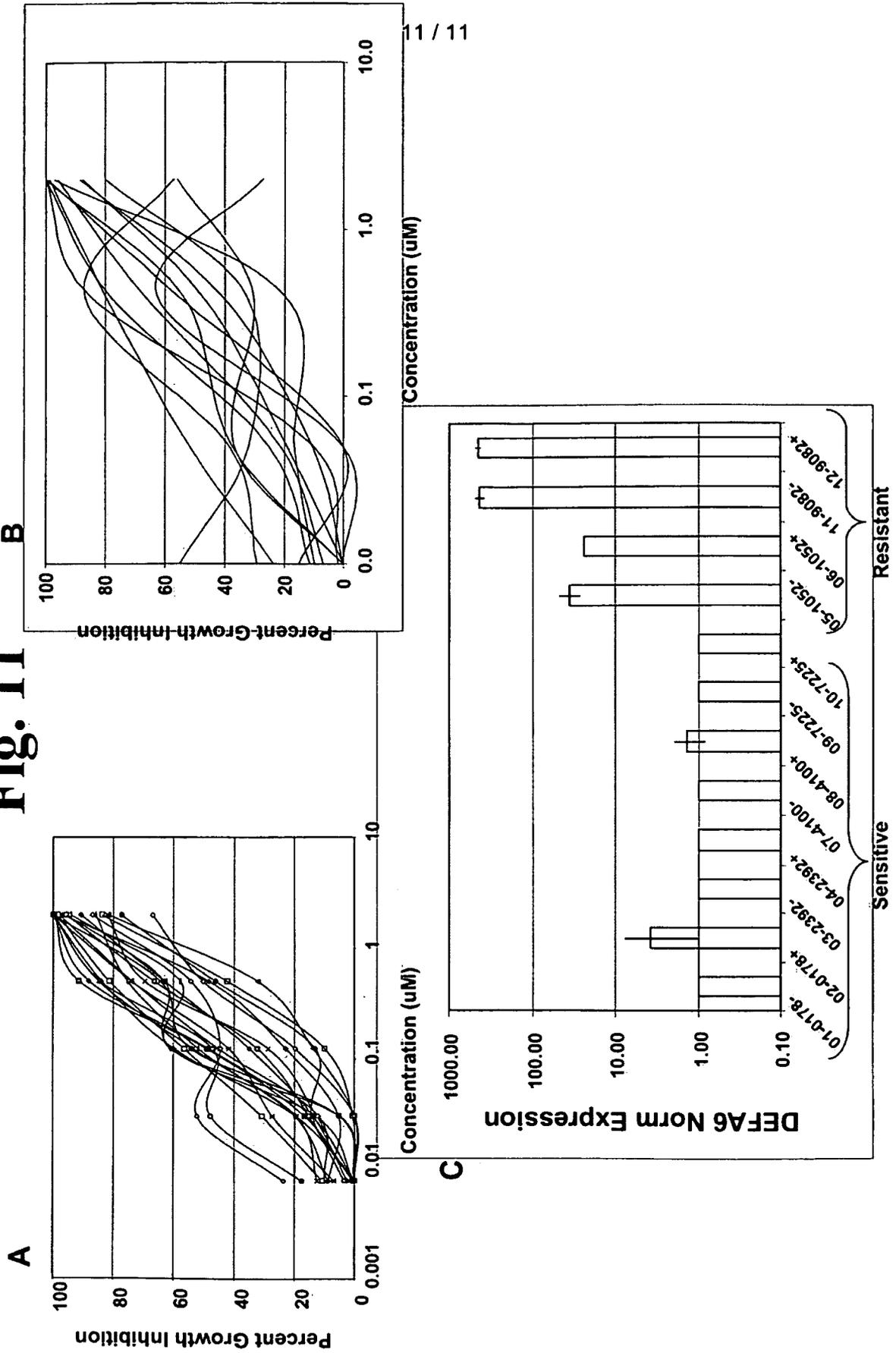


Fig. 11



## INTERNATIONAL SEARCH REPORT

International application No.  
**PCT/US2008/052540****A. CLASSIFICATION OF SUBJECT MATTER***A61K 49/00(2006.01)i, C12Q 1/68(2006.01)i, G01N 33/574(2006.01)i*

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

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

IPC 8 A61K 49/00, C12Q 1/68, G01N 33/574

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

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

eKIPASS(KIPO internal), PubMed, JPO, USPTO, Google

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 20030166026 A1 (LAURIE J. GOODMAN and BENJANIN A. BOWEN) 4 STEMBER 2003 See Abstract, Paragraph 0017-0023, 0235-0236	101-103
Y	Le Page C. et al. 'From gene profiling to diagnostic markers: IL-18 and FGF-2 complement CA125 as serum-based markers in epithelial ovarian cancer.' Int. J. Cancer. 1 April 2006, Vol. 118(7), pp. 1750-1758. See Abstract, Page 1751-1752, Figures 1-2, Table 2	101-103
Y	Natrajan R. et al. 'Array CGH profiling of favourable histology Wilms tumours reveals novel gains and losses associated with relapse.' J. Pathol. September 2006, Vol. 210(1), pp.49-58. See Abstract, Page 50, Figure 3	101-103

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

\* Special categories of cited documents:

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

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search

17 JUNE 2008 (17.06.2008)

Date of mailing of the international search report

**17 JUNE 2008 (17.06.2008)**

Name and mailing address of the ISA/KR

Korean Intellectual Property Office  
Government Complex-Daejeon, 139 Seonsa-ro, Seo-  
gu, Daejeon 302-701, Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

PARK, Yeong Gwan

Telephone No. 82-42-481-8407



**INTERNATIONAL SEARCH REPORT**

International application No.

**PCT/US2008/052540****Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.: 1-22, 41-63, 70-81 and 88-100  
because they relate to subject matter not required to be searched by this Authority, namely:  
Claims 1-22, 41-63, 70-81 and 88-100 pertain to methods for treatment of the human or animal body by therapy, as well as diagnostic methods, and thus relate to a subject matter which this International Searching Authority is not required, under Article 17(2)(a)(i) of the PCT and Rule 39.1(iv) of the Regulations under the PCT, to search.
2.  Claims Nos.: 23-40, 64-69 and 82-87  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:  
The expressions, "an isolated population of nucleic acids" of claims 23-40 and "a human cancer cell line" and/or "HDAC inhibitor" of claims 64-69 and 82-87 are too broad to make meaningful search possible.
3.  Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

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

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No.

**PCT/US2008/052540**

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 20030166026 A1	04.09.2003	(NONE)	