Fig. 6

[Continued on next page]

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

(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.
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


BACKGROUND OF THE INVENTION

[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 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

[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 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, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBPI, PTK6, EVAI, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMPRSS2, GDA, MSTIR, ITGB4, ANXA3, CCL15, DPEPI, NOXOI, IFI27, CYP3A43, and PKP2.
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 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, TPKL, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBI, PTK6, PTN6, EVAI, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMPRSS2, GDA, MSTIR, 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, FGFBI, 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, FGFBI, ABCC3, TPMT, IL18, and DPEP1. In some embodiments, one or more of the above-mentioned expression levels 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 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 biomarker genes are compared to the first set and the second set of biomarker gene expression level threshold level values.

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 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, TPKL, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBI, PTK6.
EVAl, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMRSS2, GDA, MSTlR, ...

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, FGFBPl, ABCC3, TPMT, IL18, and DPEPl. 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, FGFBPl, ABCC3, TPMT, IL18, and DPEPl.

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.

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, FGFBPl, ABCC3, TPMT, IL18, or DPEPl are represented in the isolated population of nucleic acids.

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, FGFBPl 1, ABCC3, TPMT, IL18, or DPEPl 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-2478 1. 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, FGFBPl, ABCC3, TPMT, IL18; or DPEPl 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, hi some embodiments, the type of cancer cell determined to be sensitive to an HDAC inhibitor compound HDAC inhibitor compound HDAC inhibitor compound in vitro. In some embodiments, the HDAC inhibitor compound is PCI-2478 1.

[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-2478 1. 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, NOXOI, TM4SF4, PTPN3, EPHA2, FGFBPl, ABC3, TPMT, IL18, or DPEPl 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 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, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXOI, TM4SF4, PTPN3, EPHA2, FGFBPl, ABC3, TPMT, IL18, and DPEPl. 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 µM. In some embodiments, the human cancer cell line is a colon carcinoma cell line. In some 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 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, ABC3, SARG, PPAP2C, NPDC1, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPKI, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBPl, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMPRSS2, GDA, MSTR, ITGB4, ANXA3, CCL15, DPEPl, NOXOI, 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 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 some embodiments, the at least four biomarker genes include DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXOI, TM4SF4, PTPN3, EPHA2, FGFBPl, ABC3, TPMT, IL18, and
In some embodiments, the cancer is colon cancer. In some embodiments, the HDAC inhibitor is PCI-24781.

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 DEF A6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBPI, ABCC3, TPMT, IL18, and DPEPl. 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.

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 DEF A6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBPI, ABCC3, TPMT, IL18, and DPEPl. 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.

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 DEF A6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH, NOXO1, TM4SF4, PTPN3, EPHA2, FGFBPI, ABCC3, TPMT, IL18, and DPEPl; 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, TPKI, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBPI, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMPRSS2, GDA, MSTIR, ITGB4, ANXA3, CCL15, DPEPI, 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, TPKI, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBPI, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMPRSS2, GDA, MSTIR, ITGB4, ANXA3, CCL15, DPEPI, 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.
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.

In a further aspect provided herein is a method for classifying one or more cells, 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, FGFBPI, PTK6, EVAI, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPLD3B, TMRPSS2, GDA, MSTIR, ITGB4, ANXA3, CCL15, DPEP, 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, FGFBPI, 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.

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, FGFBPI, PTK6, EVAI, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPLD3B, TMRPSS2, GDA, MSTIR, ITGB4, ANXA3, CCL15, DPEP, 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, FGFBPI, ABCC3, TPMT, IL18, and DPEP. In some embodiments, the at least four biomarker genes consist of DEFA6, RAB25, TM4SF4, and IL18.
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.

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

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).

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.

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.

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

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.

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 FTDACi 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 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% 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.

15 (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.

20 [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 polynucleotide or polynucleotides which includes unmodified 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 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.

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.

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).

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.

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 % (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.
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.

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.

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.

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**

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.

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.

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.

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.

FIG. 5 is an illustrative bar graph comparing the results of a microarray method versus TaqMan® 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.
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.

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.

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.

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.

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.

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

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.

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.

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.

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)

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.

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.

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.
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.

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.

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.

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.

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTPN3</td>
<td>PTPN3</td>
<td>AK096975</td>
<td>1</td>
</tr>
<tr>
<td>ATP-binding cassette, subfamily C (CFTR/MRP), member 3 specifically androgen-regulated protein</td>
<td>ABCC3</td>
<td>NM_020037</td>
<td>2</td>
</tr>
<tr>
<td>phosphatidic acid phosphatase type 2C</td>
<td>PPAP2C</td>
<td>NM_177526</td>
<td>4</td>
</tr>
<tr>
<td>neural proliferation, differentiation and control, 1</td>
<td>NPDC1</td>
<td>NM_015392</td>
<td>5</td>
</tr>
<tr>
<td>C-terminal tensin-like</td>
<td>CTEN</td>
<td>NM_032865</td>
<td>6</td>
</tr>
<tr>
<td>RAB25, member RAS oncogene family</td>
<td>RAB25</td>
<td>NM_020387</td>
<td>7</td>
</tr>
<tr>
<td>Hephastin</td>
<td>HEPH</td>
<td>NM_138737</td>
<td>8</td>
</tr>
<tr>
<td>thiopurine S-methyltransferase</td>
<td>TPMT</td>
<td>NM_000367</td>
<td>9</td>
</tr>
<tr>
<td>plakophilin 3</td>
<td>PKP3</td>
<td>NM_007183</td>
<td>10</td>
</tr>
<tr>
<td>UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminytransferase 5 (GalNAc-T5)</td>
<td>GALNT5</td>
<td>NM_014568</td>
<td>11</td>
</tr>
<tr>
<td>calmodulin-like 4</td>
<td>CALML4</td>
<td>NM_033429</td>
<td>12</td>
</tr>
<tr>
<td>Gene Name</td>
<td>Gene Symbol</td>
<td>GenBank Accession #</td>
<td>SEQ ID NO</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------</td>
<td>-------------</td>
<td>---------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylglucosaminyltransferase 12 (GalNAc-T12)</td>
<td>GALNT12</td>
<td>AK024865</td>
<td>13</td>
</tr>
<tr>
<td>thiamin pyrophosphokinase 1</td>
<td>TPK1</td>
<td>NM_022445</td>
<td>14</td>
</tr>
<tr>
<td>defensin, alpha 6, Paneth cells-specific</td>
<td>DEFA6</td>
<td>NM_001926</td>
<td>15</td>
</tr>
<tr>
<td>epithelial protein lost in neoplasms beta</td>
<td>EPLIN</td>
<td>NM_016357</td>
<td>16</td>
</tr>
<tr>
<td>chloride intracellular channel 5</td>
<td>CLIC5</td>
<td>NM_016929</td>
<td>17</td>
</tr>
<tr>
<td>PERP, TP53 apoptosis effector</td>
<td>PERP</td>
<td>NM_022121</td>
<td>18</td>
</tr>
<tr>
<td>spleen tyrosine kinase</td>
<td>SYK</td>
<td>NM_003177</td>
<td>19</td>
</tr>
<tr>
<td>solute carrier family 12 (sodium/potassium/chloride transporters), member 2</td>
<td>SLC12A2</td>
<td>NM_001046</td>
<td>20</td>
</tr>
<tr>
<td>guanylate cyclase 2C (heat stable enterotoxin receptor)</td>
<td>GUCY2C</td>
<td>NM_004963</td>
<td>21</td>
</tr>
<tr>
<td>transmembrane 4 superfamily member 4</td>
<td>TM4SF4</td>
<td>NM_004617</td>
<td>22</td>
</tr>
<tr>
<td>transforming growth factor, alpha</td>
<td>TGFA</td>
<td>NM_003236</td>
<td>23</td>
</tr>
<tr>
<td>fibroblast growth factor binding protein 1</td>
<td>FGFBP1</td>
<td>NM_005130</td>
<td>24</td>
</tr>
<tr>
<td>PTK6 protein tyrosine kinase 6</td>
<td>PTK6</td>
<td>NM_005975</td>
<td>25</td>
</tr>
<tr>
<td>epithelial V-like antigen 1</td>
<td>EVA1</td>
<td>NM_005797</td>
<td>26</td>
</tr>
<tr>
<td>EPH receptor A2</td>
<td>EPHA2</td>
<td>NM_004431</td>
<td>27</td>
</tr>
<tr>
<td>integrin, alpha 6</td>
<td>ITGA6</td>
<td>NM_000210</td>
<td>28</td>
</tr>
</tbody>
</table>

Table 1 (continued)

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>tumor necrosis factor receptor superfamily, member 21</td>
<td>TNFRSF21</td>
<td>NM_014452</td>
<td>29</td>
</tr>
<tr>
<td>transmembrane 4 superfamily member 3</td>
<td>TM4SF3</td>
<td>NM_004616</td>
<td>30</td>
</tr>
<tr>
<td>interleukin 18 (interferon-gamma-inducing factor)</td>
<td>IL18</td>
<td>NM_001562</td>
<td>31</td>
</tr>
<tr>
<td>bone morphogenetic protein 4</td>
<td>BMP4</td>
<td>NM_130850</td>
<td>32</td>
</tr>
<tr>
<td>sphingomyelin phosphodiesterase, acid-like 3B</td>
<td>SMPDL3B</td>
<td>NM_014474</td>
<td>33</td>
</tr>
<tr>
<td>transmembrane protease, serine 2</td>
<td>TMPRSS2</td>
<td>NM_005656</td>
<td>34</td>
</tr>
<tr>
<td>guanine deaminase</td>
<td>GDA</td>
<td>NM_004293</td>
<td>35</td>
</tr>
<tr>
<td>macrophage stimulating 1 receptor (c-met-related tyrosine)</td>
<td>MST1R</td>
<td>NM_002447</td>
<td>36</td>
</tr>
<tr>
<td>Kinase</td>
<td>Gene Symbol</td>
<td>Gene ID</td>
<td>HDACiR-BG</td>
</tr>
<tr>
<td>------------------------------------</td>
<td>-------------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>Integrin, beta 4</td>
<td>ITGB4</td>
<td>NM_000213</td>
<td>37</td>
</tr>
<tr>
<td>Annexin A3</td>
<td>ANXA3</td>
<td>NM_005139</td>
<td>38</td>
</tr>
<tr>
<td>Chemokine (C-C motif) ligand 15</td>
<td>CCL15</td>
<td>NM_032965</td>
<td>39</td>
</tr>
<tr>
<td>Dipeptidase 1 (renal)</td>
<td>DPEP1</td>
<td>NM_004413</td>
<td>40</td>
</tr>
<tr>
<td>NADPH oxidase organizer 1</td>
<td>NOXO1</td>
<td>NM_172167</td>
<td>41</td>
</tr>
<tr>
<td>Interferon, alpha-inducible protein 27</td>
<td>IFI27</td>
<td>NM_005532</td>
<td>42</td>
</tr>
<tr>
<td>Cytochrome P450, family 3, subfamily A, polypeptide 43</td>
<td>CYP3A43</td>
<td>NM_057095</td>
<td>43</td>
</tr>
<tr>
<td>Plakophilin 2</td>
<td>PKP2</td>
<td>NM_004572</td>
<td>44</td>
</tr>
</tbody>
</table>

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, TGFalpha, FGFBP1, EPHA2, TNFRSF2, TM4SF3, IL18, TPMRSS2, 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.

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, melanocyte nevus, malignant melanoma, fibroma, fibrosarcoma, an adenocarcinoma, gastrinoma, malignant gastrinoma, an oncocyto, 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, glioblastomultiforme, an oligodendrogloma, 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.

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, orbiological fluids.

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.

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 foury 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, NPDCl, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPKI, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBPI, PTK6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMDPL3B, TPMRSS2, GDA, MSTIR, ITGB4, ANXA3, CCL15, DPEPI, NOXO1, IFI127, 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, FGFBPI, 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

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.

In some embodiments, HDACiR-BG mRNA levels (including cDNA copy or rRNA 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.

[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
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.

[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.
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, NPDCl , CTEN,
RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPKI, DEFA6, EPLIN, CLIC5,
PERP, SYK, SLC1 2A2, GUCY2C, TM4SF4, TGFA, FGFBPI, PTK6, EVAI, EPHA2, ITGA6,
TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMRPRSS2, GDA, MSTIR, ITGB4, ANXA3,
CCL15, DPEPI, NOXOI, IFI27, CYP3A43, and PKP2. In some embodiments, the at least four
biomarker genes are selected from DEFA6, ITGB4, TM4SF4, SYK, PPAP2C, RAB25, HEPH,
NOXOI, TM4SF4, PTPN3, EPHA2, FGFBPI, ABCC3, TPMT, IL18, and DPEP. In some
embodiments, the at least four biomarker genes are DEFA6, RAB25, TM4SF4, and IL18.
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.

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.

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.

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.

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

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

In other embodiments, the internal expression control gene is HDAC3 (NM_003883).

In further embodiments, the internal expression control gene is ZNF217 (NM_006526).

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® from Invitrogen (Carlsbad, CA).

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).


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, electrospary, etc.), or surface plasmon reasonance 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.
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, FGFBPI, EPHA2, TNFRSF2, IL18, CCL15, or TMPRSS2.

Expression Level Reference Samples

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.

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.

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.

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

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

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.

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 µM), 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 IC₅₀ 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.

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).

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, R10781031 14, or R47 1278 1606 described herein.
HDACi inhibitor Compounds.

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):

\[
\begin{array}{c}
\text{O} \\
\text{Q} - \text{L} - \text{C} \\
\text{N} \quad \text{OR}^1 \\
\text{H}
\end{array}
\]

Formula (A)

wherein

- \(Q\) is an optionally substituted \(C_{5-12}\) aryl or an optionally substituted \(C_{5-12}\) heteroaryl;
- \(L\) is a linker having at least 4 atoms;
- \(R^1\) is \(H\) or alkyl;

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

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

\[
\begin{array}{c}
\text{O} \\
\text{Ar}^2 - \text{N} - \text{Y} - \text{X} - \text{Ar}^1 \\
\text{R}^3 \\
\text{O} \quad \text{OR}^1 \\
\text{N} \quad \text{H}
\end{array}
\]

Formula (I)

wherein:

- \(R^1\) is hydrogen or alkyl;
- \(X\) is \(-\text{O}-\), \(-\text{NR}^2\)-, or \(-\text{S(O)}_n\) where \(n\) is 0-2 and \(R^2\) is hydrogen or alkyl;
- \(Y\) is alkylene optionally substituted with cycloalkyl, optionally substituted phenyl, alkylthio, alkylsulfinyl, alkylsulfonyl, optionally substituted phenylalkylthio, optionally substituted phenylalkylsulfinyl, hydroxy, or optionally substituted phenoxy;
- \(Ar^1\) is phenylene or heteroarylene wherein said \(Ar^1\) is optionally substituted with one or two groups independently selected from alkyl, halo, hydroxy, alkoxy, haloalkoxy, or haloalkyl;
- \(R^3\) is hydrogen, alkyl, hydroxyalkyl, or optionally substituted phenyl; and
- \(Ar^2\) 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.

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

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.

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), bryo statin, 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 optionall useful for treating cancer in combination with the compounds described herein.

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

Other anti-cancer agents that are optionally employed in combination with an HDACi compound include Adriamycin, Dactinomycin, Bleomycin, Vinblastine, Cisplatin, acicivirus; aclarcubicin; acodazole hydrochloride; acronine; adozolesin; aldesleukin; altretamine; ambomyacin; amantantrone acetate; aminoglutethimide; amsacrine; anastrozole; anthramycin; asparaginase; asperlin; azacitidine; azetepa; azotomycin; barimastat; benzodepa; bicalutamidine; bisantrene hydrochloride; bisnafide dmesylate; bizelesin; bleomycin sulfate; brequinar sodium; bropirimine; busulfan; caclominycin; calusterone; caracemide; carbetimer; carboplatin; Carmustine; carubicin hydrochloride; carzelesin; cedefingol; chlorambucil; cirolemycin; cladribine; crisnatol mesylate; cyclophosphamide; cytarabine; dacarbazine; daunorubicin hydrochloride; decitabine; dexormaplatin; dezaguanine; dezaganine mesylate; diaziquone; doxorubicin; doxorubicin hydrochloride; drololoxifene; droloxifene citrate; dromostanolone propionate; duazomycin; edatrexate; elfomithine hydrochloride; elsamitracin; 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; fosquizone; fustricin sodium; gemcitabine; gemcitabine hydrochloride; hydroxyurea;
idarubicin hydrochloride; ifosfamide; iimofosine; interleukin II (including recombinant interleukin II,
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; lonustine; losoxantrone hydrochloride;
amasprocol; maytansine; meclorethamine hydrochloride; megestrol acetate; melengestrol acetate;
melphalan; menogaril; mercaptopurine; methotrexate; methotrexate sodium; metoprine; meturedepa;
misotindomide; mitocarcin; mitocromin; mitogillin; mitomalcin; mitomycin; mitosper; mitotane;
mitoxantrone; mycophenolic acid; nocardazoie; nogalamycin; ormaplatin; oxisuran;
pegasparagase; peliomycin; pentamustine; peplomycin sulfate; perfosfamide; pipobroman; piposulfan;
piroxantrone hydrochloride; plicamycin; plomestane; pofimer sodium; porfimycin; prednimustine;
procarbazine hydrochloride; puromycin; puromycin hydrochloride; pyrazofurin; riboprine;
rogletimide; safingol; safingol hydrochloride; semustine; simtrazene; sparfosate sodium;
sparsonmycin; spirogermanium hydrochloride; spiromustine; spiroplatin; streptonigrin; streptozocin;
sulofenur; talisomycin; tecogalan sodium; tegafur; teloxantrone hydrochloride; temoporfin;
teniposide; teroxirone; testolactone; thiamiprine; thioguanine; thiopeta; tiazofurin; tirapazamine;
toremifene citrate; trebolone acetate; triciribine phosphate; trimetrexate; trimetrexate glucuronate;
triptorelin; tubulozole hydrochloride; uracil mustard; uredepa; vaperotide; verteporfin; vinblastine
sulfate; vincristine sulfate; vindesine; vindesine sulfate; vinepidine sulfate; vinglycinate sulfate;
vineurosine 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
compound include: 20-epi-1,25 dihydroxyvitamin D3; 5-ethynyluracil; abiraterone; aclarubicin;
acylfulfene; adecypenol; adezolesin; aldesleukin; ALL-TK antagonists; altretamine; ambamustine;
amidox; amifostine; aminolevulinic acid; amrubicin; amsacrine; anagrelide; anastrozole;
degrangrophilide; angiogenesis inhibitors; antagonist D; antagonist G; antarelix; anti-dorsalizing
morphogenetic protein-1 ; antiandrogen, prostatic carcinoma; antiestrogen; antineoplaston; antisense
oligonucleotides; aphidicolin glycinate; apoptosis gene modulators; apoptosis regulators; apurinic
acid; ara-CDP-DL-PTBA; arginine deaminase; asulcarine; atamestane; axinastatin 1 ;
axinastatin 2; axinastatin 3; azasetron; azatoxin; azatosine; baccatin III derivatives; balanol;
batimastat; BCR/ABL antagonists; benzochlorins; benzoylestaurosporine; beta lactam derivatives;
beta-alethine; betaclamycin B; betulinic acid; bFGF inhibitor; bicalutamide; bisantrene;
bisaziridinylspermine; bisnafide; bistratene A; bizelesin; breflate; bropririmine; budotitane; buthionine
sulfoximine; calciotriol; calphostin C; canpothecin derivatives; canarypox IL-2; capecitabine;
carboxamide-amino-triazole; carboxyamidotriazole; CaRest M3; CARN 700; cartilage derived
inhibitor; carzelesin; casein kinase inhibitors (ICOS); castanospermine; cecropin B; cetrorelix; chlorlns; 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; cyclopentanthaquinones; cycloplatam; cypemycin; cytarabine ofclosfate; cytolytic factor; cytostatin; dacliximab; decitabine; dehydrodideamin B; deslorelin; dexamethasone; dexifosfamide; dexrazoxane; dexverapamil; diaziquone; didemnin B; didox; diethylnorspermine; dihydro-5-azacytidine; 9-dioxamycin; diphenyl spiromustine; docosanol; dolasetron; droxifuridine; droloxefine; dronabinol; duocarmycin SA; ebselen; ecomustine; edelfosine; edrecolomab; eflornithine; elemene; emitefur; epirubicin; epristeride; estramustine analogue; estrogen agonists; estrogen antagonists; etanidazole; etoposide phosphate; exemestane; fadrozole; fazarabine; fenretinide; filgrastim; finasteride; flavopiridol; flezelastine; fludarabine; fluorodaunorunicin hydrochloride; forfenimex; formestane; fotrericin; fotemustine; gadolinium texaphyrin; gallium nitrate; galocitabine; ganirelix; gelatinase inhibitors; gemcitabine; glutathione inhibitors; hepsulfam; heregulin; hexamethylene bisacetamide; hypericin; ibandronic acid; idarubicin; idoxifene; idramantone; ilmofosine; ilomastat; imidazoacridones; imiquimod; immunostimulant peptides; insulin-like growth factor-1 receptor inhibitor; interferon agonists; interferons; interleukins; iobenguane; iododoxorubicin; ipomeanol; 4-; iroplact; irsogladine; isobengazole; isohomohalicdrin B; itasetron; jasplakinolide; kalahalide F; lamellarin-N triacetate; lanreotide; leinamycin; leptomustine; lentinian sulfate; leptomustin; letrozole; leukemia inhibiting factor; leukocyte alpha interferon; leuprolide+estrogen+progesterone; leuprorelin; levamisole; liarozole; linear polyamine analogue; lipophilic disaccharide peptide; lipophilic platinum compounds; lissoclinamide 7; lobarplatin; lombricine; lometrexxol; lonidamine; losoxantrone; lovastatin; loxoribine; lurtotecan; lutetium texaphyrin; lysofylline; lytic peptides; maitansine; mannotin A; marimastat; masoprolcol; maspin; matrysin inhibitors; matrix metalloproteinase inhibitors; menogaril; merbarone; meterelin; methioninase; metcloproamide; 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 chongic gonadotrophin; monophosphoryl lipid A+myobacterium cell wall sk; mopidamol; multiple drug resistance gene inhibitor; multiple 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; naphteo in; nartograstim; nedaplatin; nemorubicin; neridronic acid; neutral endopeptidase; nilutamide; nisamycin; nitric oxide modulators; nitrooxide antioxidant; nitrullyn; O6-benzylguanine; octreotide; okicenone; oligonucleotides; onapristone; ondansetron; ondasetron; oracin; oral cytokine inducer; ormaplatin; osaterone; oxaliplatin; oxaunomycin; palauamine; palmitoylhrizoxin; pamidronic acid; panaxytriol; panomifene; parabactin; pazelliptine; pegaspargase; peldesine; pentosan polysulfate sodium; pentostatin; pentrozole; perflubron;
perofosfamide; perillyl alcohol; phenazinomycin; phenylacetate; phosphatase inhibitors; picibanil; piloca β hydrochloride; pirarubicin; piritekrim; 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; pyrimidine nucleoside phosphorylase inhibitors; purpurins; pyrazoloacridine; pyridoxylated hemoglobin polyoxymethylerie conjugate; raf antagonists; raltitrexed; ramosetron; ras famesyl protein transferase inhibitors; ras inhibitors; ras-GAP inhibitor; retelliptine demethylated; rhenium Re 186 etidronate; rhizoxin; riboymes; Ru retinamide; roglitinde; rohitukine; romurtide; roquinimex; rubiginone B1; ruboxyl; safingol; saintopin; SarCNU; sarcophytol A; sargamostim; Sdi 1 mimesics; semustine; senescence derived inhibitor 1; sense oligonucleotides; signal transduction inhibitors; signal transduction modulators; single chain antigen-binding protein; sizofiran; sobuzoxane; sodium borocapatate; sodium phenylacetate; solventol; 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; taumustine; 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; vitamin; vorozole; zanoteron; zeniplatin; zilasorb; and zinostatin stilalamer.

[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., mechloethamine, cyclophosphamide, chlorambucil, etc.), alkyl sulfonates (e.g., busulfan), nitrosoureas (e.g., carmustine, lomusitine, etc.), or triazenes (decarbamazine, 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).

[0014] 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).
Examples of alkylating agents that are optionally employed in combination an HDACi compound include, but are not limited to, nitrogen mustards (e.g., mechlorethamine, cyclophosphamide, chlorambucil, meiphalan, etc.), ethylenimine and methylmelamines (e.g., hexamethylenemelamine, thiotepa), alkyl sulfonates (e.g., busulfan), nitrosoureas (e.g., carmustine, lomusitine, 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, flouxouridine, Cytarabine), purine analogs (e.g., mercaptopurine, thioguanine, pentostatin.

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, carboblatin), anthrancenediole (e.g., mitoxantrone), substituted urea (e.g., hydroxyurea), methyl hydrazine derivative (e.g., procarbazine), adrenocortical suppressant (e.g., miotane, aminoglutethimide).

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), Camudotin 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-1 12378 (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-22365 1 (BASF, also known as 1LX-65 l and LU-22365 L), SAH-49960 (Lilly/Novartis), SDZ-268970 (Lilly/Novartis), AM-97 (Armad/Kyowa Hakko), AM-132 (Armad), AM-138 (Armad/Kyowa Hakko), 1DN-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), Vitileuamidine, Tubulysin A, Canadensol, Centaureidin (also known as NSC-106969), T-138067 (Tularik, also known as T-67, TL-138067 and TI-138067), COBRA-I (Parker Hughes Institute, also known as DDE-261 and WHI-261), HIO (Kansas State University), H16 (Kansas State University), Oncocidin A1 (also known as BTO-956 and DIME), DDE-3 I3 (Parker Hughes Institute), Fijianolide B, Laulimalide, SPA-2 (Parker Hughes Institute), SPA-I (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), Hemisterlin, 3-BAABU (Cytoskeleton/Mt. Sinai School of Medicine, also known as MF-1 91), 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 Desmethylleleutherobin, Desaetylleleutherobin, Isoeleutherobin A, and Z-Eleutherobin), Caribaeside, Caribaecolin, Halichondrin B, D-64131 (Asta Medica), D-68 144 (Asta Medica), Diazonomide 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-3183 15 (Abbott), HTI-286 (also known as SPA-1 10, trifluorooracetate salt) (Wyeth), D-823 17 (Zentaris), D-823 18 (Zentaris), SC-12983 (NCI), Resveratrin 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.

[0019] 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 DEF4A6, ITGB4, TM4SF4, SYK, PAP2C, RAB25, HEPH, NOXOL, TM4SF4, PTPN3, EPHA2, FGFBI, ABCC3, TPMT, IL18, and DPEPL.
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. 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 π, 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 5.83 in HDACi-resistant colon cancer cells and 0.24 in HDACi-sensitive colon cancer cells.

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 cancer.

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 the clinical status of the patient from which the sample was derived. The database is optionally 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 expression profile is associated with resistance to an HDACi compound. Methods for the configuration and construction of databases are widely available, for instance, see U.S. Pat. No. 5,953,727.

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.

Any appropriate computer platform is used to perform the methods for interpreting one or more HDACiR-BG expression profiles by the methods described herein, hi 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 as those available from Silicon Graphics. Client-server environments, 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 vis a 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 trititated 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.

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.

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.

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).
<table>
<thead>
<tr>
<th>Research ID</th>
<th>Cancer Name</th>
<th>Age</th>
<th>Sex</th>
<th>Site</th>
<th>Clinical Diagnosis</th>
<th>Histology</th>
<th>Specimen Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1076103114</td>
<td>Colon Cancer</td>
<td>54</td>
<td>F</td>
<td>R Ovary</td>
<td>Colon Carcinoma</td>
<td>ADENOCARCINOMA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R1065688672</td>
<td>Colon Cancer</td>
<td>72</td>
<td>F</td>
<td>Portion of Terminal Ileum</td>
<td>Colon Carcinoma</td>
<td>NA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R2436509366</td>
<td>Colon Cancer</td>
<td>56</td>
<td>F</td>
<td>Ileocecal Vagi</td>
<td>Rectal Cancer</td>
<td>NA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R2437281936</td>
<td>Colon Cancer</td>
<td>59</td>
<td>M</td>
<td>Colon Resection</td>
<td>Colon Carcinoma</td>
<td>NA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R424782266</td>
<td>Colon Cancer</td>
<td>51</td>
<td>F</td>
<td>Upper Left Lung</td>
<td>Colon Carcinoma</td>
<td>ADENOCARCINOMA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R598915174</td>
<td>Colon Cancer</td>
<td>43</td>
<td>F</td>
<td>Colon</td>
<td>Cecal Carcinoma</td>
<td>NA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R5173297194</td>
<td>Colon Cancer</td>
<td>65</td>
<td>M</td>
<td>Omentum</td>
<td>Colon Carcinoma</td>
<td>ADENOCARCINOMA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R713044976</td>
<td>Colon Cancer</td>
<td>52</td>
<td>F</td>
<td>R Tube &amp; Ovary</td>
<td>Colon Cancer</td>
<td>NA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R9866135153</td>
<td>Colon Cancer</td>
<td>56</td>
<td>F</td>
<td>R Hepatic Lobe</td>
<td>Colon Carcinoma</td>
<td>ADENOCARCINOMA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R2961230689</td>
<td>Colon Cancer</td>
<td>75</td>
<td>F</td>
<td>Colon</td>
<td>Colon Carcinoma</td>
<td>CARCINOMA, PD</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R542724373</td>
<td>Colon Cancer</td>
<td>56</td>
<td>F</td>
<td>Cecum</td>
<td>Colon Carcinoma</td>
<td>COLON CARCINOMA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R862446268</td>
<td>Colon Cancer</td>
<td>47</td>
<td>F</td>
<td>Brain</td>
<td>Colon Carcinoma</td>
<td>NA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R9049310123</td>
<td>Colon Cancer</td>
<td>43</td>
<td>F</td>
<td>L Lower Lung Nodule</td>
<td>Colon Carcinoma</td>
<td>ADENOCARCINOMA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R106526093</td>
<td>Colon Cancer</td>
<td>50</td>
<td>M</td>
<td>Liver</td>
<td>Colon Cancer</td>
<td>ADENOCARCINOMA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R219179233</td>
<td>Colon Cancer</td>
<td>62</td>
<td>F</td>
<td>Ovary</td>
<td>Colon Carcinoma</td>
<td>ADENOCARCINOMA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R449166314</td>
<td>Colon Cancer</td>
<td>40</td>
<td>F</td>
<td>L Ovary</td>
<td>Colon Carcinoma</td>
<td>ADENOCARCINOMA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R4897771011</td>
<td>Colon Cancer</td>
<td>53</td>
<td>F</td>
<td>R Abdominal Sidewall</td>
<td>Colon Carcinoma</td>
<td>ADENOCARCINOMA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R5465691761</td>
<td>Colon Cancer</td>
<td>65</td>
<td>F</td>
<td>Liver Lobes S &amp; E</td>
<td>Met. Colon CA to L</td>
<td>NA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R597610794</td>
<td>Colon Cancer</td>
<td>63</td>
<td>F</td>
<td>Sigmoid Rectum</td>
<td>Colon Carcinoma</td>
<td>NA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R6298195776</td>
<td>Colon Cancer</td>
<td>56</td>
<td>M</td>
<td>Liver</td>
<td>Colon Carcinoma</td>
<td>ADENOCARCINOMA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R632405249</td>
<td>Colon Cancer</td>
<td>55</td>
<td>F</td>
<td>Ovary</td>
<td>Colon Carcinoma</td>
<td>ADENOCARCINOMA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R742410798</td>
<td>Colon Cancer</td>
<td>48</td>
<td>M</td>
<td>Lumbar/Spine Biopsy</td>
<td>Colon Carcinoma</td>
<td>NA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R760142132</td>
<td>Colon Cancer</td>
<td>65</td>
<td>M</td>
<td>Sigmoid Colon</td>
<td>Colon Carcinoma</td>
<td>ADENOCARCINOMA</td>
<td>Solid Tumor Biopsy</td>
</tr>
<tr>
<td>R9418483410</td>
<td>Colon Cancer</td>
<td>56</td>
<td>F</td>
<td>Cecum</td>
<td>Colon Carcinoma</td>
<td>ADENOCARCINOMA</td>
<td>Solid Tumor Biopsy</td>
</tr>
</tbody>
</table>

Table 2
[00131] After determining tumor sensitivity to PCI-2478 1, gene expression profiles were
determined for resistant and sensitive tumors that were treated with PCI-2478 1 (2 µM) 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
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
(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

[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
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
each other and well separated from the treated samples. The major component PCAl 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-2478 1. This suggested that it is easier
to identify patients with the most resistant tumors and exclude them from a clinical trial rather than to
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.
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>z-score</th>
<th>Res./Sens. Fold Expression Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTPN3</td>
<td>PTPN3</td>
<td>AK096975</td>
<td>14.19</td>
<td>2.58</td>
</tr>
<tr>
<td>ATP-binding cassette, sub-family C (CFTR/MRP), member 3</td>
<td>ABCC3</td>
<td>NM_020037</td>
<td>13.24</td>
<td>2.37</td>
</tr>
<tr>
<td>specifically androgen-regulated protein</td>
<td>SARG</td>
<td>NM_023938</td>
<td>13.04</td>
<td>4.00</td>
</tr>
<tr>
<td>phosphatidic acid phosphatase type 2C</td>
<td>PPAP2C</td>
<td>NM_177526</td>
<td>12.95</td>
<td>4.75</td>
</tr>
<tr>
<td>neural proliferation, differentiation and control, 1</td>
<td>NPDC1</td>
<td>NM_015392</td>
<td>11.88</td>
<td>2.45</td>
</tr>
<tr>
<td>C-terminal tensin-like</td>
<td>CTEN</td>
<td>NM_032865</td>
<td>11.32</td>
<td>3.83</td>
</tr>
<tr>
<td>RAB25, member RAS oncogene family</td>
<td>RAB25</td>
<td>NM_020387</td>
<td>10.96</td>
<td>3.51</td>
</tr>
<tr>
<td>hephaestin</td>
<td>HEPH</td>
<td>NM_138737</td>
<td>10.49</td>
<td>3.38</td>
</tr>
<tr>
<td>thiopurine S-methyltransferase</td>
<td>TPMT</td>
<td>NM_000367</td>
<td>9.97</td>
<td>2.56</td>
</tr>
<tr>
<td>plakophilin 3</td>
<td>PKP3</td>
<td>NM_007183</td>
<td>9.31</td>
<td>3.13</td>
</tr>
<tr>
<td>UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylglactosaminyltransferase 5 (GalNAc-T5)</td>
<td>GALNT5</td>
<td>NM_014568</td>
<td>9.31</td>
<td>2.54</td>
</tr>
<tr>
<td>calmodulin-like 4</td>
<td>CALML4</td>
<td>NM_033429</td>
<td>9.14</td>
<td>3.51</td>
</tr>
<tr>
<td>UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylglactosaminyltransferase 12 (GalNAc-T12)</td>
<td>GALNT12</td>
<td>AK024865</td>
<td>8.86</td>
<td>2.51</td>
</tr>
<tr>
<td>thiamin pyrophosphokinase 1</td>
<td>TPK1</td>
<td>NM_022445</td>
<td>8.81</td>
<td>3.55</td>
</tr>
<tr>
<td>defensin, alpha 6, Paneth cell-specific</td>
<td>DEFA6</td>
<td>NM_001926</td>
<td>8.58</td>
<td>12.92</td>
</tr>
<tr>
<td>epithelial protein lost in neoplasm beta</td>
<td>EPLIN</td>
<td>NM_016357</td>
<td>8.49</td>
<td>2.33</td>
</tr>
<tr>
<td>chloride intracellular channel 5</td>
<td>CLIC5</td>
<td>NM_016929</td>
<td>7.20</td>
<td>3.60</td>
</tr>
<tr>
<td>PERP, TP53 apoptosis effector</td>
<td>PERP</td>
<td>NM_022121</td>
<td>6.94</td>
<td>2.60</td>
</tr>
<tr>
<td>spleen tyrosine kinase</td>
<td>SYK</td>
<td>NM_003177</td>
<td>6.90</td>
<td>3.59</td>
</tr>
<tr>
<td>solute carrier family 12 (sodium/potassium/chloride transporters), member 2</td>
<td>SLC12A2</td>
<td>NM_001046</td>
<td>6.75</td>
<td>4.85</td>
</tr>
<tr>
<td>guanylate cyclase 2C (heat stable enterotoxin receptor)</td>
<td>GUCY2C</td>
<td>NM_004963</td>
<td>6.72</td>
<td>3.53</td>
</tr>
<tr>
<td>transmembrane 4 superfamily member 4</td>
<td>TM4SF4</td>
<td>NM_004617</td>
<td>6.54</td>
<td>12.09</td>
</tr>
<tr>
<td>transforming growth factor, alpha</td>
<td>TGFA</td>
<td>NM_003236</td>
<td>6.44</td>
<td>3.11</td>
</tr>
<tr>
<td>Protein Name</td>
<td>Gene Symbol</td>
<td>Accession Number</td>
<td>Fold Change</td>
<td>p-Value</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>-------------</td>
<td>------------------</td>
<td>-------------</td>
<td>---------</td>
</tr>
<tr>
<td>PTK6 protein tyrosine kinase 6</td>
<td>PTK6</td>
<td>NM_005975</td>
<td>6.24</td>
<td>3.10</td>
</tr>
<tr>
<td>epithelial V-like antigen 1</td>
<td>EVA1</td>
<td>NM_005797</td>
<td>5.96</td>
<td>4.55</td>
</tr>
<tr>
<td>EPH receptor A2</td>
<td>EPHA2</td>
<td>NM_004431</td>
<td>5.90</td>
<td>2.18</td>
</tr>
<tr>
<td>integrin, alpha 6</td>
<td>ITGA6</td>
<td>NM_000210</td>
<td>5.53</td>
<td>4.09</td>
</tr>
<tr>
<td>tumor necrosis factor receptor superfamily, member 21</td>
<td>TNFRSF21</td>
<td>NM_014452</td>
<td>5.47</td>
<td>2.16</td>
</tr>
<tr>
<td>transmembrane 4 superfamily member 3</td>
<td>TM4SF3</td>
<td>NM_004616</td>
<td>5.32</td>
<td>3.75</td>
</tr>
<tr>
<td>interleukin 18 (interferon-gamma-inducing factor)</td>
<td>IL18</td>
<td>NM_001562</td>
<td>5.24</td>
<td>5.22</td>
</tr>
<tr>
<td>bone morphogenetic protein 4</td>
<td>BMP4</td>
<td>NM_130850</td>
<td>4.82</td>
<td>3.91</td>
</tr>
<tr>
<td>sphingomyelin phosphodiesterase, acid-like 3B</td>
<td>SMPDL3B</td>
<td>NM_014474</td>
<td>4.62</td>
<td>5.49</td>
</tr>
<tr>
<td>transmembrane protease, serine 2</td>
<td>TMPRSS2</td>
<td>NM_005656</td>
<td>4.62</td>
<td>3.51</td>
</tr>
<tr>
<td>guanine deaminase</td>
<td>GDA</td>
<td>NM_004293</td>
<td>4.56</td>
<td>6.52</td>
</tr>
<tr>
<td>macrophage stimulating 1 receptor (c-met-related tyrosine kinase)</td>
<td>MST1R</td>
<td>NM_002447</td>
<td>4.49</td>
<td>4.52</td>
</tr>
<tr>
<td>integrin, beta 4</td>
<td>ITGB4</td>
<td>NM_000213</td>
<td>4.41</td>
<td>3.98</td>
</tr>
<tr>
<td>annexin A3</td>
<td>ANXA3</td>
<td>NM_005139</td>
<td>4.11</td>
<td>3.34</td>
</tr>
<tr>
<td>chemokine (C-C motif) ligand 15</td>
<td>CCL15</td>
<td>NM_032965</td>
<td>3.87</td>
<td>3.74</td>
</tr>
<tr>
<td>dipeptidase 1 (renal)</td>
<td>DPEP1</td>
<td>NM_004413</td>
<td>3.72</td>
<td>5.53</td>
</tr>
<tr>
<td>NADPH oxidase organizer 1</td>
<td>NOXO1</td>
<td>NM_172167</td>
<td>3.71</td>
<td>8.92</td>
</tr>
<tr>
<td>interferon, alpha-inducible protein 27</td>
<td>IFI27</td>
<td>NM_005532</td>
<td>3.69</td>
<td>3.65</td>
</tr>
<tr>
<td>cytochrome P450, family 3, subfamily A, polypeptide 43</td>
<td>CYP3A43</td>
<td>NM_057095</td>
<td>3.65</td>
<td>3.40</td>
</tr>
<tr>
<td>plakophilin 2</td>
<td>PKP2</td>
<td>NM_004572</td>
<td>3.54</td>
<td>3.45</td>
</tr>
</tbody>
</table>

[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-2478 1.
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® quantitative RT-PCR method as described below.

TaqMan® 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® 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>
<thead>
<tr>
<th>GeneName</th>
<th>GeneCards</th>
<th>Microarrays</th>
<th>Taqman</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Zcut</td>
<td>Resist mean</td>
<td>Sens mean</td>
</tr>
<tr>
<td>defensin, alpha 6, Paneth cell-specific</td>
<td>DEFA6</td>
<td>8.58</td>
<td>8.57</td>
</tr>
<tr>
<td>integrin, beta 4</td>
<td>ITGB4</td>
<td>4.41</td>
<td>0.67</td>
</tr>
<tr>
<td>transmembrane 4 superfamily member 3</td>
<td>TM4SF3</td>
<td>5.32</td>
<td>239.99</td>
</tr>
<tr>
<td>spleen tyrosine kinase</td>
<td>SYK</td>
<td>6.90</td>
<td>5.16</td>
</tr>
<tr>
<td>phosphatidic acid phosphatase type 2C</td>
<td>PPAP2C</td>
<td>12.85</td>
<td>5.35</td>
</tr>
<tr>
<td>RAS25, member RAS oncogene family</td>
<td>RAB25</td>
<td>10.96</td>
<td>55.31</td>
</tr>
<tr>
<td>hphaestin</td>
<td>HEPH</td>
<td>10.49</td>
<td>6.11</td>
</tr>
<tr>
<td>NADPH oxidase organizer 1</td>
<td>NOXO1</td>
<td>3.71</td>
<td>0.98</td>
</tr>
<tr>
<td>transmembrane 4 superfamily member 4</td>
<td>TM4SF4</td>
<td>6.84</td>
<td>2.06</td>
</tr>
<tr>
<td>PTPN3</td>
<td>PTPN3</td>
<td>14.19</td>
<td>5.45</td>
</tr>
<tr>
<td>EPH receptor A2</td>
<td>EPHA2</td>
<td>5.90</td>
<td>29.27</td>
</tr>
<tr>
<td>fibroblast growth factor binding protein 1</td>
<td>FGFBP1</td>
<td>6.27</td>
<td>27.93</td>
</tr>
<tr>
<td>ATP-binding cassette, sub-family C, member 3</td>
<td>ABC3</td>
<td>13.24</td>
<td>4.14</td>
</tr>
<tr>
<td>thiopurine S-methyltransferase</td>
<td>TPMT</td>
<td>9.97</td>
<td>26.21</td>
</tr>
<tr>
<td>interleukin 18 (interferon-gamma-inducing factor)</td>
<td>IL18</td>
<td>5.24</td>
<td>26.57</td>
</tr>
<tr>
<td>diploidase 1 (renal)</td>
<td>OPE1</td>
<td>3.72</td>
<td>2.93</td>
</tr>
<tr>
<td>HDAC3</td>
<td>HDAC3</td>
<td>Not significant</td>
<td>25.66</td>
</tr>
<tr>
<td>Zinc Finger Protein znf217</td>
<td>ZNF217</td>
<td>Not significant</td>
<td>35.07</td>
</tr>
<tr>
<td>TSG101</td>
<td>TSG101</td>
<td>Not significant</td>
<td>40.00</td>
</tr>
</tbody>
</table>
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.

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 (NMJM4452), TMF4SF3 (NM_004616), IL18 (NM_001562), TMPRSS2 (NM_005656), and CCL15 (NM_032965).

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-2478 I.

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-2478 I (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-2478 I (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**

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 to identify HDACi responsive mouse tissues. Primary human tumors corresponding to sensitive tissues were then tested ex-vivo with PCI-2478 I, and it was found that tumors from tissues that showed higher levels of activity were sensitive to inhibition by PCI-2478 I, thus validating that this technique does indeed predict sensitive tumor types.

In brief, female BALB/c mice were injected IV with 50 mg/kg PCI-2478 I 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 α-tubulin & histones were evaluated by
immunoblotting.

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

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.

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

<table>
<thead>
<tr>
<th>Common</th>
<th>Description</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slc9a3r1</td>
<td>solute carrier family 9 isoform 3 regulator 1</td>
<td>ION TRANSPORT</td>
</tr>
<tr>
<td>Ing1</td>
<td>inhibitor of growth family, member 1-like</td>
<td>CELL PROLIFERATION AND DIFFERENTIATION</td>
</tr>
<tr>
<td>Gadd45g</td>
<td>growth arrest and DNA-damage-inducible 45 gamma</td>
<td>CELL PROLIFERATION AND DIFFERENTIATION; APOPTOSIS</td>
</tr>
<tr>
<td>Plaur</td>
<td>urokinase plasminogen activator receptor</td>
<td>MULTIPLE</td>
</tr>
<tr>
<td>EST</td>
<td>RIKEN cDNA 281405022 gene</td>
<td>UNKNOWN</td>
</tr>
<tr>
<td>Insi6</td>
<td>insulin-like 6</td>
<td>BIOLOGICAL PROCESS UNKNOWN</td>
</tr>
<tr>
<td>Luc7i</td>
<td>Luc7 homolog (S. cerevisiae)-like</td>
<td>RNA PROCESSING</td>
</tr>
<tr>
<td>Taf9</td>
<td>TAF9 RNA polymerase II</td>
<td>MRNA TRANSCRIPTION</td>
</tr>
<tr>
<td>Gadd45b</td>
<td>growth arrest and DNA-damage-inducible 45 beta</td>
<td>CELL PROLIFERATION AND DIFFERENTIATION</td>
</tr>
<tr>
<td>Syng2</td>
<td>synaptojanin 2</td>
<td>UNKNOWN</td>
</tr>
<tr>
<td>Polr2e</td>
<td>polymerase (RNA) II (DNA directed) polypeptide E</td>
<td>MRNA TRANSCRIPTION</td>
</tr>
<tr>
<td>Kras2</td>
<td>Mouse c-Ki-ras oncogene</td>
<td>ONCOGENE</td>
</tr>
<tr>
<td>Hspa5</td>
<td>heat shock 70kD protein 5</td>
<td>STRESS RESPONSE</td>
</tr>
<tr>
<td>Fgf15</td>
<td>fibroblast growth factor 15</td>
<td>CELL PROLIFERATION AND DIFFERENTIATION</td>
</tr>
<tr>
<td>Tuba4</td>
<td>tubulin, alpha 4</td>
<td>CELL STRUCTURE</td>
</tr>
<tr>
<td>H2afz</td>
<td>H2A histone family, member Z</td>
<td>CHROMATIN PACKAGING</td>
</tr>
</tbody>
</table>
Subsequently, we validated the expression profile of two of HDACi-responsive genes, Fgfl5 and Syngr2, 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.

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 issue distribution pattern was very reproducible across the biomarker set. Ovary showed the highest level of induction, followed by uterus.

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 (EDR) 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.
Table 6: Tumor Resistance to HDAC Inhibitor PCI-24781

<table>
<thead>
<tr>
<th>Tumor</th>
<th>Resistant</th>
<th>Intermediate</th>
<th>Sensitive</th>
<th>Total</th>
<th>%Resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>EDR</td>
<td>IDR</td>
<td>LDR</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>AML</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>10</td>
<td>19</td>
</tr>
<tr>
<td>Multiple Myeloma</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>6</td>
<td>33</td>
</tr>
<tr>
<td>Ovarian</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>12</td>
<td>25</td>
</tr>
<tr>
<td>Glioblastoma</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>29</td>
</tr>
<tr>
<td>Colon</td>
<td>9</td>
<td>3</td>
<td>12</td>
<td>24</td>
<td>38</td>
</tr>
</tbody>
</table>

Note: EDR/LDR status as determined by Onco tech's algorithm from their assay data.
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.
## APPENDIX

Nucleotide Sequences for HDACi Compound Resistance Biomarker Genes

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTPN3</td>
<td>PTPN3</td>
<td>AK096975</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene Name</td>
<td>Gene Symbol</td>
<td>GenBank Accession #</td>
<td>SEQ ID NO</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>---------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>ATP-binding cassette, sub-family C (CFTR/MRP), member 3</td>
<td>ABC3</td>
<td>NM_020037</td>
<td>2</td>
</tr>
</tbody>
</table>

```
1  ctcgggcccc gcgctgtcgc gcgcgttggt cgcaccgccc tcgcctctct tgcagccgcc
gagcaggggg tggccacctt tctaacccca taaccctgcc ccctgccctc ttcctctcct
201 ctcgatagct gcgtctcctc gttcctgctg tgcagccgcc ctgtgccctc ccctgccctc
2161 ctggtgagct ccctctgccg gactgttctt ccctctctct ctcgagcttc ctcggtgttc
gagcagagtg tggccacctt ctcctctctc ctctctctct ctcctctctc ctcctctctc
2211 ctctctctct ctcctctctc ttcctctctc ctcctctctc ctcctctctc ctcctctctc
2221 ctctctctct ctcctctctc ctcctctctc ctcctctctc ctcctctctc ctcctctctc
2281 ctcctctctc ctcctctctc ctcctctctc ctcctctctc ctcctctctc ctcctctctc
5  ctctctctct ctcctctctc ctcctctctc ctcctctctc ctcctctctc ctcctctctc
2341 ctcctctctc ctcctctctc ctcctctctc ctcctctctc ctcctctctc ctcctctctc
2361 ctcctctctc ctcctctctc ctcctctctc ctcctctctc ctcctctctc ctcctctctc
10  ctcctctctc ctcctctctc ctcctctctc ctcctctctc ctcctctctc ctcctctctc
```
2221 caagcgctac cagcagactc tggaggcctg tgccttgcta gctgacctgg agatgctgcc
tggtggggat cagacagaga ttggagagaa gggcattaac ... aggccacccc
taggaactca gtcctgtact ctggggtgct gcctgaatcc attaaaaatg ggagtactga
tgaaataaaa ctacatggtc aacagtaaaa

5

10
tgtggtccag aagcagttta tgagcagact gattcctcctg tctctcgctc tctttcctgc
ttaggaactca gtcctgtact ctggggtgct gcctgaatcc attaaaaatg ggagtactga
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa

15
tgatgccatg cagcagactc tggaggcctg tgccttgcta gctgacctgg agatgctgcc
tggtggggat cagacagaga ttggagagaa gggcattaac ... aggccacccc
taggaactca gtcctgtact ctggggtgct gcctgaatcc attaaaaatg ggagtactga
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa

20

tgcctgctct ctgcctggct cctcctggaa ccctctgatt gcagaggata aggccacccc
tggtggggat cagacagaga ttggagagaa gggcattaac ... aggccacccc
taggaactca gtcctgtact ctggggtgct gcctgaatcc attaaaaatg ggagtactga
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa

25

tgtgccatc cgtgttgact gactcttctt ctgagctgtc gctggtgggc gctggtgggc
tggtggggat cagacagaga ttggagagaa gggcattaac ... aggccacccc
taggaactca gtcctgtact ctggggtgct gcctgaatcc attaaaaatg ggagtactga
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa

30

tgcctgctct ctgcctggct cctcctggaa ccctctgatt gcagaggata aggccacccc
tggtggggat cagacagaga ttggagagaa gggcattaac ... aggccacccc
taggaactca gtcctgtact ctggggtgct gcctgaatcc attaaaaatg ggagtactga
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa

35

tgcctgctct ctgcctggct cctcctggaa ccctctgatt gcagaggata aggccacccc
tggtggggat cagacagaga ttggagagaa gggcattaac ... aggccacccc
taggaactca gtcctgtact ctggggtgct gcctgaatcc attaaaaatg ggagtactga
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa

40

tgcctgctct ctgcctggct cctcctggaa ccctctgatt gcagaggata aggccacccc
tggtggggat cagacagaga ttggagagaa gggcattaac ... aggccacccc
taggaactca gtcctgtact ctggggtgct gcctgaatcc attaaaaatg ggagtactga
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa

45

tgcctgctct ctgcctggct cctcctggaa ccctctgatt gcagaggata aggccacccc
tggtggggat cagacagaga ttggagagaa gggcattaac ... aggccacccc
taggaactca gtcctgtact ctggggtgct gcctgaatcc attaaaaatg ggagtactga
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa

50

tgcctgctct ctgcctggct cctcctggaa ccctctgatt gcagaggata aggccacccc
tggtggggat cagacagaga ttggagagaa gggcattaac ... aggccacccc
taggaactca gtcctgtact ctggggtgct gcctgaatcc attaaaaatg ggagtactga
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa

55

tgcctgctct ctgcctggct cctcctggaa ccctctgatt gcagaggata aggccacccc
tggtggggat cagacagaga ttggagagaa gggcattaac ... aggccacccc
taggaactca gtcctgtact ctggggtgct gcctgaatcc attaaaaatg ggagtactga
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa

60

tgcctgctct ctgcctggct cctcctggaa ccctctgatt gcagaggata aggccacccc
tggtggggat cagacagaga ttggagagaa gggcattaac ... aggccacccc
taggaactca gtcctgtact ctggggtgct gcctgaatcc attaaaaatg ggagtactga
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
tgaaataaaa ctacatggtc aacagtaaaa
<table>
<thead>
<tr>
<th>Gene Name: specifically androgen-regulated protein</th>
<th>Gene Symbol: SARG</th>
<th>GenBank Accession #: NM_023938</th>
<th>SEQ ID NO: 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 gtcgggggccag gcacaggcag atgaaagcatt tacctatcta gttgatgcag caggagctcca</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 1 aaggggagc aacaaagctg gagcagcggg aacagcgcctc tgggctccag tcccgtgcc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>121 gaaacaaaggag tatttcttctg ttcctctcag ccaccccaac ccaggctgc acggcagctgg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2941 cacccaagaa ccgtttcagtt ctcttctgtg cctgatgcag gtagttttaa atttttctca</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2881 ccacactcct ccacagggag cagagcagctc ccacagcgctc ggtgctgtgc ccaccccaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>941 gacactgacc gcagccacgt ccagccacgc ccacgacggg caccaccagt ccaccccaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 1 cccctgggaga cccctgagag cagccttcttc atggtgctgc gcaccccaac ccaccccaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>181 gagaggagcag gacccagactc ccagctgagc cctttctcag ggtgctgtgc ccaccccaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2161 acaccgaggg cccaggctcct caccaggggt ccagccacgg ccaccccaac ccaccccaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25 1 gtagaagaccc ccaggctctct tgcctctctc ccacgctgtg cccaccccaac ccaccccaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2881 gcctggaggg ggtgctgtgc ccaccccaac ccaccccaac ccaccccaac ccaccccaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 1 gtagaagaccc ccaggctctct tgcctctctc ccacgctgtg cccaccccaac ccaccccaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>331 gtgggggggccag gcacaggcag atgaaagcatt tacctatcta gttgatgcag caggagctcca</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>361 gtgggggggccag gcacaggcag atgaaagcatt tacctatcta gttgatgcag caggagctcca</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>421 cagaggctctcc cagactgctcc gttgactgtc gttgatgcag caggagctcca</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>451 cgctgatgcag gatgggagcc caggcttgat aggccagata gaagacagct cccgctgtcc</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>481 cagggctgcc gcagccacgt ccagccacgg ccacgacggg caccaccagt ccaccccaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>511 cagggctgcc gcagccacgt ccagccacgg ccacgacggg caccaccagt ccaccccaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>541 cagggctgcc gcagccacgt ccagccacgg ccacgacggg caccaccagt ccaccccaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>571 cagggctgcc gcagccacgt ccagccacgg ccacgacggg caccaccagt ccaccccaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>601 cagggctgcc gcagccacgt ccagccacgg ccacgacggg caccaccagt ccaccccaac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene Name</td>
<td>Gene Symbol</td>
<td>GenBank Accession #</td>
<td>SEQ ID NO</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>---------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>neural proliferation, differentiation and control, 1</td>
<td>NPDC1</td>
<td>NM_015392</td>
<td>5</td>
</tr>
</tbody>
</table>

20

<table>
<thead>
<tr>
<th>1</th>
<th>gcgcgcctcg</th>
<th>cgccgcctcg</th>
<th>caccccgatc</th>
<th>cctgtgacagc</th>
<th>ctggccaaagt</th>
<th>acatgattgg</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>gcctctgagg</td>
<td>cccacacttc</td>
<td>taqccgcttg</td>
<td>cgaccccgac</td>
<td>tggagccggg</td>
<td>tcaactgcct</td>
</tr>
<tr>
<td>10</td>
<td>gcagtacagag</td>
<td>cactctgaga</td>
<td>aggtgtgcag</td>
<td>ggggaacccc</td>
<td>gctgtatgca</td>
<td>cccagacgcc</td>
</tr>
<tr>
<td>15</td>
<td>gcgtctgagg</td>
<td>cccacacttc</td>
<td>taqccgcttg</td>
<td>cgaccccgac</td>
<td>tggagccggg</td>
<td>tcaactgcct</td>
</tr>
</tbody>
</table>

25

<table>
<thead>
<tr>
<th>1</th>
<th>gcgcgcctcg</th>
<th>cgccgcctcg</th>
<th>caccccgatc</th>
<th>cctgtgacagc</th>
<th>ctggccaaagt</th>
<th>acatgattgg</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>gcctctgagg</td>
<td>cccacacttc</td>
<td>taqccgcttg</td>
<td>cgaccccgac</td>
<td>tggagccggg</td>
<td>tcaactgcct</td>
</tr>
<tr>
<td>10</td>
<td>gcagtacagag</td>
<td>cactctgaga</td>
<td>aggtgtgcag</td>
<td>ggggaacccc</td>
<td>gctgtatgca</td>
<td>cccagacgcc</td>
</tr>
<tr>
<td>15</td>
<td>gcgtctgagg</td>
<td>cccacacttc</td>
<td>taqccgcttg</td>
<td>cgaccccgac</td>
<td>tggagccggg</td>
<td>tcaactgcct</td>
</tr>
</tbody>
</table>

30

<table>
<thead>
<tr>
<th>1</th>
<th>gcgcgcctcg</th>
<th>cgccgcctcg</th>
<th>caccccgatc</th>
<th>cctgtgacagc</th>
<th>ctggccaaagt</th>
<th>acatgattgg</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>gcctctgagg</td>
<td>cccacacttc</td>
<td>taqccgcttg</td>
<td>cgaccccgac</td>
<td>tggagccggg</td>
<td>tcaactgcct</td>
</tr>
<tr>
<td>10</td>
<td>gcagtacagag</td>
<td>cactctgaga</td>
<td>aggtgtgcag</td>
<td>ggggaacccc</td>
<td>gctgtatgca</td>
<td>cccagacgcc</td>
</tr>
<tr>
<td>15</td>
<td>gcgtctgagg</td>
<td>cccacacttc</td>
<td>taqccgcttg</td>
<td>cgaccccgac</td>
<td>tggagccggg</td>
<td>tcaactgcct</td>
</tr>
</tbody>
</table>

35

<table>
<thead>
<tr>
<th>1</th>
<th>gcgcgcctcg</th>
<th>cgccgcctcg</th>
<th>caccccgatc</th>
<th>cctgtgacagc</th>
<th>ctggccaaagt</th>
<th>acatgattgg</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>gcctctgagg</td>
<td>cccacacttc</td>
<td>taqccgcttg</td>
<td>cgaccccgac</td>
<td>tggagccggg</td>
<td>tcaactgcct</td>
</tr>
<tr>
<td>10</td>
<td>gcagtacagag</td>
<td>cactctgaga</td>
<td>aggtgtgcag</td>
<td>ggggaacccc</td>
<td>gctgtatgca</td>
<td>cccagacgcc</td>
</tr>
<tr>
<td>15</td>
<td>gcgtctgagg</td>
<td>cccacacttc</td>
<td>taqccgcttg</td>
<td>cgaccccgac</td>
<td>tggagccggg</td>
<td>tcaactgcct</td>
</tr>
</tbody>
</table>

40

<table>
<thead>
<tr>
<th>1</th>
<th>gcgcgcctcg</th>
<th>cgccgcctcg</th>
<th>caccccgatc</th>
<th>cctgtgacagc</th>
<th>ctggccaaagt</th>
<th>acatgattgg</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>gcctctgagg</td>
<td>cccacacttc</td>
<td>taqccgcttg</td>
<td>cgaccccgac</td>
<td>tggagccggg</td>
<td>tcaactgcct</td>
</tr>
<tr>
<td>10</td>
<td>gcagtacagag</td>
<td>cactctgaga</td>
<td>aggtgtgcag</td>
<td>ggggaacccc</td>
<td>gctgtatgca</td>
<td>cccagacgcc</td>
</tr>
<tr>
<td>15</td>
<td>gcgtctgagg</td>
<td>cccacacttc</td>
<td>taqccgcttg</td>
<td>cgaccccgac</td>
<td>tggagccggg</td>
<td>tcaactgcct</td>
</tr>
</tbody>
</table>

45

<table>
<thead>
<tr>
<th>1</th>
<th>gcgcgcctcg</th>
<th>cgccgcctcg</th>
<th>caccccgatc</th>
<th>cctgtgacagc</th>
<th>ctggccaaagt</th>
<th>acatgattgg</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>gcctctgagg</td>
<td>cccacacttc</td>
<td>taqccgcttg</td>
<td>cgaccccgac</td>
<td>tggagccggg</td>
<td>tcaactgcct</td>
</tr>
<tr>
<td>10</td>
<td>gcagtacagag</td>
<td>cactctgaga</td>
<td>aggtgtgcag</td>
<td>ggggaacccc</td>
<td>gctgtatgca</td>
<td>cccagacgcc</td>
</tr>
<tr>
<td>15</td>
<td>gcgtctgagg</td>
<td>cccacacttc</td>
<td>taqccgcttg</td>
<td>cgaccccgac</td>
<td>tggagccggg</td>
<td>tcaactgcct</td>
</tr>
<tr>
<td>Gene Name</td>
<td>Gene Symbol</td>
<td>GenBank Accession #</td>
<td>SEQ ID NO</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------</td>
<td>-------------</td>
<td>--------------------</td>
<td>-----------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C-terminal tensin-like</td>
<td>CTEN</td>
<td>NM_032865</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 gggcaacagt ctgccacact gttgacacca gactctggtt ctgcagcttg agcaagttag

5 1 ctgccacact 121 cccagtgt gttgacacca gttgacccct gaggacagtc acctcttggtt cttgcagcttg agcaagttag

10 gcagctcagc agggggagact tgtccccatg gttgacacca gttgacccct gaggacagtc acctcttggtt cttgcagcttg agcaagttag

15 gcagctcagc agggggagact tgtccccatg gttgacacca gttgacccct gaggacagtc acctcttggtt cttgcagcttg agcaagttag

20 gcagctcagc agggggagact tgtccccatg gttgacacca gttgacccct gaggacagtc acctcttggtt cttgcagcttg agcaagttag
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAB25, member RAS oncogene family</td>
<td>RAB25</td>
<td>NM_020387</td>
<td>7</td>
</tr>
</tbody>
</table>

```
1  cttgcagatg  agtatcacc  cctttttct  ttttatttt  tatttttttt
3001 gagacagagt  ctcactgtca  cccaggctga  actgcagtgg  tgtgatctag  gctcaactgca
gctcgactga  ctgctctagg  tctccgagta  gtctgggatta
3061 acctccacct  cacacctgcc  gcctctgacc  ttggccagca  ccacctgccc
ccactggctt
3121 caggcatgtg  caactcaccc  agctaatttt  gtatttttag  tagagacagg
gtttcaccat
3181 gttggccagg  ctggtcttga  actcctgacc  gcaggtaatc  cacctgcttc
3241 agtgctggga  ttacaggcgc  aagccaccca  gcccagcttc  tttccattcc
ttgataggcg
3301 agtattccaa  agctggtatc  gtagctgccc  taatgttgca  tattaggcgg
cggggcaga
3361 gataagggcc  atctctctgt  gattctgcct  cagctcctgt  cttgctgagc
cctcccccaa
3421 cccacgctcc  aacacacaca  cacacacaca  cacacacaca  cacacacaca
3481 cacgcccctc  tactgctatg  tggcttcaac  cagcctcaca  gccacacggg
ggaagcagag
3541 agtcaagaat  gcaaagaggc  cgcttcccta  agaggcttgg  aggagctggg
cctctatccca
3601 cacccacccc  caccccaccc  ccacccagcc  tccagaagct  ggaaccattt
cctcccgcagg
3661 cctgagttcc  taaggaaacc  accctaccgg  ggtggaaggg  agggtcaggg
aagaaaccca
3721 cattatacctt  acgaggagca  agtgcctgcc  ccctcccagc  agccagccct
gccaaagttg
3781 ctcttgctct  acgaggagca  gtagctgccc  ccctcccagc  agccagccct
gccaaagttg
3841 agtctgagat  cagcccaccc  agccagtggt  cgagcactgc  cccgccgcca
aagtctcag
3901 aatgtgagat  gaggttctca  aggtcacagg  cccccagtccc  agcctggggg
cctggcagagg
3961 ccccccatata  cttcgtctca  gcttcctatca  tgaaaaaaataaatg
```

```
ctctgcttcc  ttacagcacc  cccacctgcc  agagctgatc  ctccctaggc  cctgcctaac
cttgagttgg  cccccaatcc  ctctggctgc  aagaatcccc  ttacccccaa
tgagaggagg
ggcaggacca  gatcttttga  gagctgaggg  ttgagggcat  tgagccaaca
tcgcctctgt  ccccgaaagac  acctgcaccc  tccatgcgga  gccaagatgg
ggaatggaac
tgaggaagat  tataactttg  tcttcaaggt  ggtgctgatc  ggcaatatcag
split into sections
```

```
gaccaatcta  ctctcccgat  tcacgcgcaa  tgagttcagc  cacgacagcc
gcaccaccat
ggggttgag  ttctccaccc  gcactgtgat  tttgggcacc  gctgctgtca
tagctcagat
tggtgggtaa
tttggtgccc  cttgtcccca  cttcagcccc  aggacctttc  cttgcccttt
ggttccagat
```

```
ctctgcttcc  ttacagcacc  cccacctgcc  agagctgatc  ctccctaggc  cctgcctaac
cttgagttgg  cccccaatcc  ctctggctgc  aagaatcccc  ttacccccaa
tgagaggagg
ggcaggacca  gatcttttga  gagctgaggg  ttgagggcat  tgagccaaca
tcgcctctgt  ccccgaaagac  acctgcaccc  tccatgcgga  gccaagatgg
ggaatggaac
tgaggaagat  tataactttg  tcttcaaggt  ggtgctgatc  ggcaatatcag
split into sections
```

```
gaccaatcta  ctctcccgat  tcacgcgcaa  tgagttcagc  cacgacagcc
gcaccaccat
ggggttgag  ttctccaccc  gcactgtgat  tttgggcacc  gctgctgtca
tagctcagat
tggtgggtaa
tttggtgccc  cttgtcccca  cttcagcccc  aggacctttc  cttgcccttt
ggttccagat
```

```
ctctgcttcc  ttacagcacc  cccacctgcc  agagctgatc  ctccctaggc  cctgcctaac
cttgagttgg  cccccaatcc  ctctggctgc  aagaatcccc  ttacccccaa
tgagaggagg
ggcaggacca  gatcttttga  gagctgaggg  ttgagggcat  tgagccaaca
tcgcctctgt  ccccgaaagac  acctgcaccc  tccatgcgga  gccaagatgg
ggaatggaac
tgaggaagat  tataactttg  tcttcaaggt  ggtgctgatc  ggcaatatcag
split into sections
```

```
gaccaatcta  ctctcccgat  tcacgcgcaa  tgagttcagc  cacgacagcc
gcaccaccat
ggggttgag  ttctccaccc  gcactgtgat  tttgggcacc  gctgctgtca
tagctcagat
tggtgggtaa
tttggtgccc  cttgtcccca  cttcagcccc  aggacctttc  cttgcccttt
ggttccagat
```

```
ctctgcttcc  ttacagcacc  cccacctgcc  agagctgatc  ctccctaggc  cctgcctaac
cttgagttgg  cccccaatcc  ctctggctgc  aagaatcccc  ttacccccaa
tgagaggagg
ggcaggacca  gatcttttga  gagctgaggg  ttgagggcat  tgagccaaca
tcgcctctgt  ccccgaaagac  acctgcaccc  tccatgcgga  gccaagatgg
ggaatggaac
tgaggaagat  tataactttg  tcttcaaggt  ggtgctgatc  ggcaatatcag
split into sections
```

```
gaccaatcta  ctctcccgat  tcacgcgcaa  tgagttcagc  cacgacagcc
gcaccaccat
ggggttgag  ttctccaccc  gcactgtgat  tttgggcacc  gctgctgtca
tagctcagat
tggtgggtaa
tttggtgccc  cttgtcccca  cttcagcccc  aggacctttc  cttgcccttt
ggttccagat
```

```
ctctgcttcc  ttacagcacc  cccacctgcc  agagctgatc  ctccctaggc  cctgcctaac
cttgagttgg  cccccaatcc  ctctggctgc  aagaatcccc  ttacccccaa
tgagaggagg
ggcaggacca  gatcttttga  gagctgaggg  ttgagggcat  tgagccaaca
tcgcctctgt  ccccgaaagac  acctgcaccc  tccatgcgga  gccaagatgg
ggaatggaac
tgaggaagat  tataactttg  tcttcaaggt  ggtgctgatc  ggcaatatcag
split into sections
```

```
gaccaatcta  ctctcccgat  tcacgcgcaa  tgagttcagc  cacgacagcc
gcaccaccat
ggggttgag  ttctccaccc  gcactgtgat  tttgggcacc  gctgctgtca
tagctcagat
tggtgggtaa
tttggtgccc  cttgtcccca  cttcagcccc  aggacctttc  cttgcccttt
ggttccagat
```

```
ctctgcttcc  ttacagcacc  cccacctgcc  agagctgatc  ctccctaggc  cctgcctaac
cttgagttgg  cccccaatcc  ctctggctgc  aagaatcccc  ttacccccaa
tgagaggagg
ggcaggacca  gatcttttga  gagctgaggg  ttgagggcat  tgagccaaca
tcgcctctgt  ccccgaaagac  acctgcaccc  tccatgcgga  gccaagatgg
ggaatggaac
tgaggaagat  tataactttg  tcttcaaggt  ggtgctgatc  ggcaatatcag
split into sections
```

```
gaccaatcta  ctctcccgat  tcacgcgcaa  tgagttcagc  cacgacagcc
gcaccaccat
ggggttgag  ttctccaccc  gcactgtgat  tttgggcacc  gctgctgtca
tagctcagat
tggtgggtaa
tttggtgccc  cttgtcccca  cttcagcccc  aggacctttc  cttgcccttt
ggttccagat
```
961 atcagactgt tccctgttca cagcacccctc aggggtcttaa ggtcttcatg
cctatcaca
1021 aataacctcttt tatctgttcc accccctcaca gactaggacc tccaataaa
acctttttat
1081 atcaaaaaaa

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>hephaestin</td>
<td>HEPH</td>
<td>NM_138737</td>
<td>8</td>
</tr>
</tbody>
</table>

1 gcccagcctg ccttgagaaa agtgtctgct cctagcaag atctcctcat
cacaaagatc
gt2ccttggtg
cctcaactgca
gtgcagctg
15 aactatgctc ccaagggaga aatgtcatac acgaaccagc ctctggacag
tgcatagttg
gttccagctg
tcttaaagagc
gagaacctc
tataaagaa atcctagatc
gtctgcgggc
tccctgggc
tttgcccccg
421 tttgcccactg tctctatact catccacacct cattcgtctct tctacgagaa
cagctggggc
ttccttggg
tggaagttccc
tcgactgtcctg
cacacttcat
gctttccctat
cctggagaaaa
ttccctcctc
tcaagtgagaa
tgatggtcctg
ttttgcccga
tggccacttact
tatggagaa
35 781 agctggcact tcaatggaga cattggcact tactgtctcag atcctggctc
gatggcaaaa 841 gaagatgaga catttcaggc gagcaatagg atgcatgcaaa tcaatggtctttttttgg
901 aatatttcgct agctgaacatt gttgacagaa aacggttggg cctggaactt
gttgctgatc
gcgaatgaaa ttgattggtca cagagcatatt tccctgtgacc agatgtgctag
tccgatatgg
tcacacactg atgtggcttaa cacactttttcagcacccttttg tggactggtggtggc
tgccc
45 1081 tgggaaccctg gtacctgtggt aattagttcag caagtgaacca gtcacttttgc
1141 caggcactctt ccaaggtcag gtttggcttcc atgggccccct cttggtgaccctgtccagagc
cagtcggtactcagcag tacactttcat tggagccccat gacattcaat gggacttattg
cccgatgggg
1261 catgatggga gtactgggaa gaatttgaga gagccaggca gtatctcaga
taagtttttc
1321 cagaagagct ccagccgaat tgggggcact tactggaaag... aaagggcatc ctggagcccc atggaggacg
gagtgacatg
2821 gatcgggaat ttgcattgtt gttct 'tgatt tttgatgaaa ataagtcttg
gtatttggag
2881 gaaaatgtgg caacccatgg gtcccaggat ccaggcagta ttaacctaca
ggatgaaact
2941 ttcttggaga gcaataaaat gcatgcaatc aatgggaaac ...
3001 ggatgaaact
tctgaggtgc
3061 gcatgcaatc aatgggaaac tctatgccct atccggaatgg
cggataactc
3121 cggcagatg tgggtgatct gtcccaaggg actttttagg tgtgaggat
ggatgaaact
3181 cggatctacaca ccacccactt ccatccactt tcatgcagag tggcctctg
tggccatggg ccaagatgtg
3241 accctcttca etgttttttc tcgaacagaa acttaagcc ctcccaacgta
catgcagatcc
3301 gagactgaa aagcagtgcc cccccagagac attgaagaag gcaatgtgaa
ggatgaaact
tctgaggtgc
3361 ccttaccatgt accaaggaga acgagtggcc tggtacatgc tggccatggg
tggccatggg ccaagatgtg
3421 accctcttca etgttttttc tcgaacagaa acttaagcc ctcccaacgta
catgcagatcc
3481 aagcactgac gcaataggag gtccatcctg gatgacagct tcaagcttct
ttgccatggg ccaagatgtg
3541 cagtaacatc tggagcctgg agatatcctc atccggaatgg
ggatgaaact
tctgaggtgc
3601 ggccatggac tagtcactaa ccccaacacat ccaggggcat ggggtggtgga
tggccatggg ccaagatgtg
3661 gagcaatcaa tgggtgatct gtcccaaggg actttttagg tgtgaggat
ggatgaaact
tctgaggtgc
3721 ttcaccccttt ttcttacttt ctttgctctta cggcgaacct tggccatgga
tggccatggg ccaagatgtg
3781 attaattaca atcgccacatt tcaacgctg cattatatgt tcctgcacta
tggccatggg ccaagatgtg
3841 atgcagatcc ccataaaaga ttgctgtaag cttgcccttg ttttgggtgc
catgcagatcc
3901 accttcatgct gggtagttttc ctttgccccat gaaattttc ctttcgctatc
tggccatggg ccaagatgtg
3961 gcatgcaatc aatgggaaac tctatgccct atccggaatgg
ttgccatggg ccaagatgtg
4021 aagcagtgcc cccccagagac attgaagaag gcaatgtgaa
ggatgaaact
tctgaggtgc
4081 ttcaccccttt ttcttacttt ctttgctctta cggcgaacct tggccatgga
tggccatggg ccaagatgtg
4141 accttcatgct gggtagttttc ctttgccccat gaaattttc ctttcgctatc
tggccatggg ccaagatgtg
4201 gttaccccttt caagcatgtg ccaggggcat ggggtggtgga
tggccatggg ccaagatgtg
4261 cggccagatg tgggtgatct gtcccaaggg actttttagg tgtgaggat
ggatgaaact
tctgaggtgc
4321 aagcagtgcc cccccagagac attgaagaag gcaatgtgaa
ggatgaaact
tctgaggtgc
4381 ttcaccccttt ttcttacttt ctttgctctta cggcgaacct tggccatgga
tggccatggg ccaagatgtg
4441 accttcatgct gggtagttttc ctttgccccat gaaattttc ctttcgctatc
tggccatggg ccaagatgtg
4501 gttaccccttt caagcatgtg ccaggggcat ggggtggtgga
tggccatggg ccaagatgtg
4561 aagcagtgcc cccccagagac attgaagaag gcaatgtgaa
ggatgaaact
tctgaggtgc
4621 ttcaccccttt ttcttacttt ctttgctctta cggcgaacct tggccatgga
tggccatggg ccaagatgtg
4681 accttcatgct gggtagttttc ctttgccccat gaaattttc ctttcgctatc
tggccatggg ccaagatgtg
4741 gttaccccttt caagcatgtg ccaggggcat ggggtggtgga
tggccatggg ccaagatgtg
4801 aagcagtgcc cccccagagac attgaagaag gcaatgtgaa
ggatgaaact
tctgaggtgc
4861 ttcaccccttt ttcttacttt ctttgctctta cggcgaacct tggccatgga
tggccatggg ccaagatgtg
4921 accttcatgct gggtagttttc ctttgccccat gaaattttc ctttcgctatc
tggccatggg ccaagatgtg
4981 gttaccccttt caagcatgtg ccaggggcat ggggtggtgga
tggccatggg ccaagatgtg
5041 aagcagtgcc cccccagagac attgaagaag gcaatgtgaa
ggatgaaact
tctgaggtgc
5101 ttcaccccttt ttcttacttt ctttgctctta cggcgaacct tggccatgga
tggccatggg ccaagatgtg
5161 accttcatgct gggtagttttc ctttgccccat gaaattttc ctttcgctatc
tggccatggg ccaagatgtg
5221 gttaccccttt caagcatgtg ccaggggcat ggggtggtgga
tggccatggg ccaagatgtg
5281 aagcagtgcc cccccagagac attgaagaag gcaatgtgaa
ggatgaaact
tctgaggtgc
5341 ttcaccccttt ttcttacttt ctttgctctta cggcgaacct tggccatgga
tggccatggg ccaagatgtg
5401 accttcatgct gggtagttttc ctttgccccat gaaattttc ctttcgctatc
tggccatggg ccaagatgtg
5461 gttaccccttt caagcatgtg ccaggggcat ggggtggtgga
tggccatggg ccaagatgtg
5521 aagcagtgcc cccccagagac attgaagaag gcaatgtgaa
ggatgaaact
tctgaggtgc
5581 ttcaccccttt ttcttacttt ctttgctctta cggcgaacct tggccatgga
tggccatggg ccaagatgtg
5641 accttcatgct gggtagttttc ctttgccccat gaaattttc ctttcgctatc
tggccatggg ccaagatgtg
5701 gttaccccttt caagcatgtg ccaggggcat ggggtggtgga
tggccatggg ccaagatgtg
1441 gtagagacag ggtttcacca tgttggtcag gctgatctcg agctcctgac
cctcaggtgat
1501 ctacccacct cggcctccca aagtgctggg attacaggcg
1561 aagtgctggg attacaggcg
ttactcaga
1621 gcattctttt cggcctccca aagtgctggg attacaggcg
ttactcaga
1681 tctctcttaaa
1741 taatgaacaa ctctctcttaaa
1801 ggaacaaaca ttcccataa tctaagaataa atttttcttt
1861 aatgttgttt atatcataatt aatcatcag gcattctttt cggcctccca aagtgctggg attacaggcg
ttactcaga
1921 ctctctcttaaa
1981 ccacatatttgc tctctcttaaa
2041 ttagattgttt tagctcatgt taatcagcgc gcattctttt cggcctccca aagtgctggg attacaggcg
ttactcaga
2101 gcattctttt cggcctccca aagtgctggg attacaggcg
ttactcaga
2161 cccccctttaa cagctgtgaa agctacgctt ccctctctcttaaa
2221 tctctctcttaaa
gggtttcttc ctctctctcttaaa
2281 tatctcttaata
gtctctcttaaa
2341 gcattctttt cggcctccca aagtgctggg attacaggcg
ttactcaga
2401 cctcttttaa cagctgtgaa agctacgctt ccctctctcttaaa
2461 gagagacagctttc gagcgggtctt cagctgtgaa agctacgctt ccctctctcttaaa
2521 tgggttggtt ccctctctcttaaa
2581 attacaggtta gcattctttt cagctgtgaa agctacgctt ccctctctcttaaa
2641 aagcaacccc gtaatcccag ccctctctcttaaa
2701 gtaatcccag ccctctctcttaaa
2761 ttctctcttt cagctgtgaa agctacgctt ccctctctcttaaa
2821 atctctcttaa ccctctctcttaaa
2881 tcctctcttctt cagctgtgaa agctacgctt ccctctctcttaaa
2941 acagcagatga gcagggtcagcagctgtgaa agctacgctt ccctctctcttaaa
3001 ctctctcttaa ccctctctcttaaa
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>plakophilin 3</strong></td>
<td>PKP3</td>
<td>NM_007183</td>
<td>10</td>
</tr>
</tbody>
</table>

```
3061 tagagtcctt tccgcttgca gttgtgcgca aagaatctta aatacaaatg agatatcctt
3121 aggtagttga tcaattatgt aatatgtgtc ttcactgggg aatactgact tccctaaaae
3181 tcaagatgga agatatcctt catgtaaatt attttagagc aataaattgt ttttccaggtat
3241 tttttccaaaa

| 10 | 1 ggcctcgaggg cacagagctg gaagatagtt ggttttgag gcggccgcca ggcctcgaggg |
| 61 | 1 cgggtgcacct cgcggccagt agggacggtaa cttctctggt tccggccggt gcggccgcca |
| 121 | 1 cgggctgtgc tccctggggt ggtggcccctt tacggttggt gcggccgcca |
| 181 | 1 ggcggaggg gagagcagccg gatggggcag ccggttcagc ccggttcagc |
| 241 | 1 gcagctgaggga cagacacgag gcggccgcca agcggcaacc gcggccgcca |
| 301 | 1 tgcagagggg gatcgacatca cccctgtag gcggccgcca |
| 361 | 1 ctagggctgtg agtgggggca cggactggcgc gggccctgggg gcggccgcca |
| 421 | 1 cccagctccc tggctcccgc gtcggggtgt gcgtgcgtgc gcggccgcca |
| 481 | 1 agcccaacac ggggaggagc gcttgctggt gcggccgcca |
| 541 | 1 tcccatgcac accagccggc tggctggggt gcggccgcca |
| 601 | 1 ctatagacaca cctcctctgc gctgcgtgag gcggccgcca |
| 661 | 1 gacgcctgggt ctggagccgc gctggccgcc gcggccgcca |
| 721 | 1 gcagggcccg gacagctgcc gcggccgcca gcggccgcca |
| 781 | 1 tggattcc ccggcgtgcg cccttgggt gcggccgcca |
| 841 | 1 gcagagcccg cggagccgag ggtgtgcgt gcggccgcca |
| 901 | 1 ggtgtgcag gccagcctgg cggactggcgc ggcggctggg gcggccgcca |
| 961 | 1 gcagcttgct gcggccccag cgggttggt gcggccgcca |
| 1021 | 1 ttgtgctgact agtgagtcgg cctgcctgg cggggtgggc gcggccgcca |
| 1081 | 1 gcaggtgcac gcggcgcccag cggggtgggc gcggccgcca |
| 1141 | 1 gcctgctgcc gcggctgccgc ggggtggtgc gcggccgcca |
```
1201 ggaagtgcag cgccatgcca caggtgccat gcgcaacctc atctacgaca
acgctgacaa
1261 caagctggcc ctggtggagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
1321 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
1381 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
1441 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
1501 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
1561 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
1621 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
1681 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
1741 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
1801 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
1861 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
1921 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
1981 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
2041 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
2101 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
2161 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
2221 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
2281 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
2341 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
2401 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
2461 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
2521 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
2581 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
2641 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
2701 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
2761 ctggttgagg agaacgggat cttcgagctg ctgcggacaca
tggagccaga
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 5 (GalNAc-T5)</td>
<td>GALNT5</td>
<td>NM_014568</td>
<td>11</td>
</tr>
</tbody>
</table>

1  agtgttttac  agaacttagc  cagggcagc  caagcaggca  cagatgctct  gctatgaaat

5  gccacgcagg  cagagactga  caagcggtag  gaactgagct  ttccttgcc  acgatgtag  gcaagctcttgg

10  ggaaagcttt  gtaccatgaa  caggatccga  aagtttttc  gaggaagtg

15  ggaggccgata  ggaccctct  gcagagactga  caagcggtag  gaactgagct  ttcctttgg  acatggccgag  aaaaactgag

20  gagagtgtgc  tcaagggcagg  cagagactga  caagcggtag  gaactgagct  ttcctttgg  acatggccgag  aaaaactgag

25  gcaagctctc  ctgcatctgg  ctcctcttttg  acatggccgag  aaaaactgag

30  gacaggaggg  agagtcatag  tcccagcagt  gacacatcaa  aactagcagc  tgcagctctc

35  cccaaactaa  ccatgcttta  actggagggc  tagagccagc  aaaaatcaac

40  ccagagcttc  ctagctgttg  tgcagctctc  gacagcataa  aaaaaactgag  aaaaatcaac

45  acacctttgg  gacacatcaa  aactagcagc  tgcagctctc  gacagcataa  aaaaaactgag  aaaaatcaac
1441 ataactgcca aagccccctc tacagaatac aaccagagtc atataaaagc
ccttttacct
1501 gaagacagtg gaacgcacca ggtgttaaga attgatgtga ... caatcagcaa ttattatgct tggaaggaaa
tttttctcaa
3001 aagatcctga aagtagctgc ctgtgaccca gtgaagccat atcaaaagtg
gaaatttgaa
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>calmodulin-like 4</td>
<td>CALML4</td>
<td>NM_033429</td>
<td>12</td>
</tr>
</tbody>
</table>

```
3061 aaatattatg aagcctgaag tgtaactgat gtttttatat agtaaaccca
ttaaatactg
3121 tgaaaataac a
g
1  ggggctgagg gtggagagag aaggaaggg aaggaaggg aagctttcct
ggccagggta
61 accggcacta agaggctca ctccaagccc cccgaggagcc ttgtgtgggg
cctgacaccc
10 cgtcagccgcccc cctcaagagtc ttcgagaggt gggaagtaa
181 atgggaaacgg gtcgtagcccg gcccactgga ttttgcagcg cgcctcagaa
ttcagcgtgt
241 gaggaagctcg aggcagcagcc tgaggtgaga tgaatctgcg gtcactcagc
aagtacagcg
301 acgggaggg ctgaatccca accgaagctt ggtctacagc
361 cccagccttgc cgcgtgaaggg tggccatcctc tgaagctggg agctttgcctt
421 ccctctcaacag gatggtggtct gggagccctt caatacagcc ccgggtgtaa
481 tctagtccgg gccgctgccgcc agtcccccagag gacacccctc
541 cgtcagcccg cgggccttgcag gagaacacag tggccatgct ccccttgccctt
601 acgcagacag ctgaatccca accgaagctt ggtctacagc
661 ctcggcctcgc ggcggctgctg caggtcgact caggaagctg ggaatctgacc
cggcttctgg
721 cgggccccagctctcaagcccg cccgtgagttc ggggccctcgc aggtgtcctc
gccacgccga
781 acttttaggtg gaaagggapl ctaagcggccctgtctctcttc
gcaggagcc
841 ccggagagctg ccgagccact gagaatttgc ccaagacca
atattaggt
901 gcaagagtgcc ccccttcttg tctgagaacc gcaggagggg gaagatataaa
gccacgccga
961 tctaggtgctg caggtggctg cccgagcgc ggcggagcttg cagcgagcc
ccaagccgc
1021 tggcagaccc gccgtagagcc ggaatggagag cgcgtcttcct cccactcctt
tctgacataattata
1081 tgcacatgca aataaaccag ggaagacacc cagagaaagt aacctcagc
atggtgtaggg
1141 tggagagagc gaaagaggtg taccgctaggt cctccgagcct gcggctaaata
ctcagcgatcc
1201 tggggaggaag ctgcagccagc aaggaaggtg atgacagttt cagggaagcgc
gatagtcgac
1261 ccaatggcag aaggcaatgtgatttacact gccaaatgct cccctcctc
tgcagcggact
```
1321 attgaaggag gagaatgcca gagcctcccc tgggcctgaa aacttggagc
aattaatttt
1381 ttttaaaaag tgttatctttc acttgggaga gatggcaaac ...
1441 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
1501 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tyatcata
1561 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
1621 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
1681 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
1741 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
1801 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
1861 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
1921 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
1981 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
2041 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
2101 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
2161 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
2221 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
2281 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
2341 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
2401 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
2461 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
2521 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
2581 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
2641 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
2701 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
2761 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
2821 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
2881 atctaccaag acactttacct tgtatattgg ggctagaacc actttcattc
tttcaagat
4561 tggttcagat ttttttttt taagatgtat ccaataacac tcacgaagta attaaagcc
4621 actttaaccce tgctaaaaaa

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 12 (GALNT12)</td>
<td>GALNT12</td>
<td>AK024865</td>
<td>13</td>
</tr>
</tbody>
</table>

1 catttttataa tgaagcctgg tcaacttccc ttcggacagt ttacagtgtc ctggagacat
d6 ccccggtatat ccctgctagaa gaagtgtacc ttgtagatga ctacagtgtagagagaagcc
121 tgaagagcgcc tttgctgcaat gagctttcgg gactgcggcag ggtgagcctg atcccggccca
181 acacgcagaga gggctttggtg cgacgggggc tcggggggg gctgtccggcg agggggccgatg
15 241 ttctgaccttt cctggagcgtc cacttgtagt gcctacagag tctggctgag gcctgctgc
301 agagatccatcg tgaagagggg tcccgagttcg tgtgccccgtg gattgagttg tgtgcacgga
361 acaccccctgta tattctgggg aactccgggg agccccccagct cggcggttctc
gactgcgggc
421 tctggatcag ctggccacaca cttcttgagta gggagaggt gacgatgca aacaagagcccgtgtgtgc
481 atgtcatcag gcctccttaca atggctgtgtg ggctgttttcg tggatgaaat aatctgtgtg
541 aaactcttgga cttcttgagta acaggaatgg aagttggaggag aagagaaac cttctgtgttta
601 ccttttagatg tccgcagtgt ggtgaggttcc tgcaacacag ccatgttccc catgtgtggg
661 atgctttttccc caaacaagct cccctactccc gcaacagcag tctggccaaac agtgtgccgtg
721 cgaccgtttaa atgtctggag tattatccag aagtttgggg ggtaggaag cttcagttttta
781 gctggctgcacc cttttggggtg gtagagagac ggaagcagct cgggacaagcctccttgtgtg
841 aagactctcaaa gttggtctctt gagactgtgtg atccagaact gcattgaccttagagaggc
901 cttgggtcttt cgggattcgtc cagaacaaag gactaacaga cttgctgctttatcatacctg
961 ctctccgatgaa aacccgatttt cttggacacc agttcatctggtaccccggttgatg
1021 gctcagtcttttctagtatacagacagcaagcaagctcattataaactgcctgacaagcctccgctgccc
1081 ctggagggctg cttctggagtt gaagcagggtaaggacacatt tacagctcctggtacccggttgatg
1141 aacgcgttcctg aagagtttccg tgcagctctgtttctgcctggtacccggttgatg
1201 atgccaagaa agtgtccag gctgcgagga aggagtcgag tgacagtttc
gttccactct
1261 tacgagactg caccaactcg gatcatcaga aatggttctt caaagagcgc
1321 gacttagctg agcagtgcacc gaagccacc aaaaactagg ctgcattgtc
ttgaagaggc
1381 atactattttg ccttttgtct aatgccttaa caaatccttc ataatatctc
taatagttga
1441 gacttagtga
ttatccaggt ctgtgtccag gatctgtgag cagcttccct
1501 tacgtgcacct tcgtgccctt tctgtaggta taattcttct
1561 gatgttatgaa
tacgtgctac tacacactgtg cagatgttgca cgatgtgccg
1621 ttgctgccct tcgcaggtct tacagtttcg atacagtttg
1681 tcgacagttc gccacagttc gccacagttc
1741 ttgacagttc
tacaggtct tgcaacctttg cccggtacct cctgtgcatt
1801 actgtgacta ggaacatgtt attgtcatc tcaagttttc
gagagcttgg
tatcagttga
1861 gctggtggag cagagtatct gcaggtcttg ggttgtttct
1921 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
1981 agtccaagaa atgctgtagc gctgcgagga aggagtcgag tgacagtttc
gttccactct
2041 atgactttaa aagagtttta ataggcattt agcagttttc
gtcaggaga
taatagttga
2101 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
2161 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
2221 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
2281 aagagtttta ataggcattt agcagttttc
gtcaggaga
taatagttga
2341 gctggtggag cagagtatct gcaggtcttg ggttgtttct
2401 tacagttttt tacatcgttc ctgagttttt cctgttcacc
2461 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
2521 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
2581 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
2641 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
2701 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
2761 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
2821 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
2881 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
2941 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
3001 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
3061 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
3121 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
3181 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
3241 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
3301 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
3361 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
3421 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
3481 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
3541 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
3601 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
3661 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
3721 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
3781 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
3841 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
3901 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
3961 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
4021 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
4081 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
4141 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
4201 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
4261 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
4321 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
4381 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
4441 tcaagttttt tacatcgttc ctgagttttt cctgttcacc
4501 tcaagttttt tacatcgttc ctgagttttt cctgttcacc

**Gene Name** | **Gene Symbol** | **GenBank Accession #** | **SEQ ID NO**
---|---|---|---
thiamin pyrophosphokinase 1 | TPK1 | NM_022445 | 14
301 attcatcaat ggagacctgg attctattag gcctgaagtc agagaatact atgctactaa
361 gggatgtgag ctcatttcaa ctcctgatca agaccacact gctgttgtgct gctgctgtgct
tccagctccac
421 tgcactctgct gggcctgctttg accagattat ggcattgatgtgctacagc acctggctcaa
481 atgctactaa
541 tccagctgcta ctcctgatca atgctactaa
601 accagagaaac ccaggttgcgc atgctactaa
gtggctccac
661 tctctgctggta cagctctgta atgctactaa
gcctgaagtc
721 tgcactctgct gggcctgctttg accagattat ggcattgatgtgctacagc acctggctcaa
781 tctctgctggta cagctctgta atgctactaa
gcctgaagtc
841 tgcactctgct gggcctgctttg accagattat ggcattgatgtgctacagc acctggctcaa
901 tgaagtggcc atgctactaa
gcctgaagtc
961 tctctgctggta cagctctgta atgctactaa
gcctgaagtc
1021 tgcactctgct gggcctgctttg accagattat ggcattgatgtgctacagc acctggctcaa
1081 tgcactctgct gggcctgctttg accagattat ggcattgatgtgctacagc acctggctcaa
1141 tgcactctgct gggcctgctttg accagattat ggcattgatgtgctacagc acctggctcaa
1201 tgcactctgct gggcctgctttg accagattat ggcattgatgtgctacagc acctggctcaa
1261 tgcactctgct gggcctgctttg accagattat ggcattgatgtgctacagc acctggctcaa
1321 tgcactctgct gggcctgctttg accagattat ggcattgatgtgctacagc acctggctcaa
1381 tgcactctgct gggcctgctttg accagattat ggcattgatgtgctacagc acctggctcaa
1441 tgcactctgct gggcctgctttg accagattat ggcattgatgtgctacagc acctggctcaa
1501 tgcactctgct gggcctgctttg accagattat ggcattgatgtgctacagc acctggctcaa
1561 tgcactctgct gggcctgctttg accagattat ggcattgatgtgctacagc acctggctcaa
1621 tgcactctgct gggcctgctttg accagattat ggcattgatgtgctacagc acctggctcaa
1681 tgcactctgct gggcctgctttg accagattat ggcattgatgtgctacagc acctggctcaa
1741 tgcactctgct gggcctgctttg accagattat ggcattgatgtgctacagc acctggctcaa
1801 tgcactctgct gggcctgctttg accagattat ggcattgatgtgctacagc acctggctcaa
1921 tggtgcatcg aatgtattc acctttcttt tgaagaga ccacctttta
tctttctttcc
1981 acctttctct gtttatgaa accaactgtt gacacacaa ccctgattg
2041 gttccacacgt ttttatgtac acaactctct atggctctt aagaaagtgg
aaagtacgt
2101 tttttctttcttgtttatgaa cagttgtgaa ctcagttgtt ctgtcagaaa
tgaaagaggc
2161 aatgaccggcgtgtctttg gatttaagg gcagttgcccc atgtgactgt
2221 gccaacactt tttctttcttg cttgaccgga agtttaggtt tagcttagtt
tccattttcctta
2281 gtttctgtga actagtttaa tattcttttac tagaaatatata tcataaatatatat
2341 atatatctgtg attttaaaat tttgctacca aagactgtt gttctgtgtg
cctgaaatat
2401 gttaccagtg ttaataatgt gataacattt cccatagaaaa

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>defensin, alpha 6, Paneth cell-specific</td>
<td>DEFA6</td>
<td>NM_001928</td>
<td>15</td>
</tr>
</tbody>
</table>

1 acacatctgct ccctgctctc tctctctcaag gcagccactgc catgagaaccctcaccatccc
61 tcacctgtgc ctctctctgttg cccctcaggg ccagagctgag ccacccatccagcgtgaggtgc
121 atccactgca ggccaaagct tatgaggtgct atgcctcgagg gcagccagggcagagctgaggtgc
181 aggatcttgc cgtctctcttt gcagaggtgct ccagactcaac tctttagagtcttggtctccaa
241 caaggggttt cacttgcactgcgcagaga ggtctgtatcc aacagaatatctcattgatgc
301 cctgctactgtc ctgagagagt aaccacatgc tgtctgtcac ctgaggtgagtt ccagaccccagacaagagcatgacgtgaggtgc
361 cactatatatatt caataaactatttatagct ctttattgacactttaaaactggataactagtttgctctctcctcagagtacagagaagagctgaggtgct ccaccatcactgctgcacgacaagagcatgcgtgaggtgc
421 ggcattcactc tgggtttctc atctcactatatgtggaacagtacagagaagagctgaggtgct ccaccatcactgctgcacgacaagagcatgcgtgaggtgc

...
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>epithelial protein lost in neoplasm beta</td>
<td>EPLIN</td>
<td>NM_016357</td>
<td>16</td>
</tr>
</tbody>
</table>

```
1 gcgctaggta gagcgccggg accttgtgaca gggctggtag cagcgcagag
gaaaggcggc
5 61 ttttagccag gtatttcagt gtctgtagac aagatggaat catctccatt
taatagacgg
121 caatggacct cactatcatt gagggtaaca gccaaagaac tttctcttgt
cacaagaac
181 aagtcagtgc gatttggga aatatctcct aagtgcaaga aagcgtctga
agaacaacaac
241 atggagaga agagaagtaa caccgaaat ctctcccagc actttgaaag
301 acctgtttaa aagaagtg ggaacaagca ggctgggag cagagcttca
cacagctct
361 ctacggaaca gcagcactga gattaggca agagcagacc atcctcttgc
tgaaaggtaca
421 agccacgcttg cttcttgagg caaagctgac caagaagaac aacctaccacc
cagactat
481 ctcaggtcag ctccttgaga ccctctctcag ggtcgtatct cccacatcaca
541 gatcttttaag accactcaca aagaagtaa aaattgaaat attgtctagg
agaactccag
601 catgaagtag aaaaaatcaga aatcagtgaa aacccccca
aatagagaca
661 tataatgttc cgctgaacag gcttaagatg atgttgagaa aagtggaacc
aatccaaact
721 aagattttcc gggccaaag ccgaatgctca agtggaagag aagtctctga
aaacagctat
781 tctcttagatg acctgggaat aggccagttg caggtgctat cttctcttgc
841 aaaaaatgaga gtagagctaa tctggaacct ccacgcctct cagaaacctc
tataaggaat
901 cgaatggccca agt(acc)agag ccgatgtgcc aaacaaagca gtccaaaatcatcaaat
961 gagctgaaga ccagtggttt cgaatctaaa attcataaaa tggagcaaaaa
ggagaaatgg
1021 cccccaggtc ctgaggttgc catcaccctc caggaagggg aaaagatctc
tgaaatgag
1081 aatagcctgg cagtcctggc cccctctgcc gaagatgact cccggtgactc
1141 aaggtggttcc cactgccag gtctgtgtcc cccctctggc
1201 tttactgaaa gttcccttcc caaagcaatg aagaagttgc aggacagctcc
1261 ttcgatcggata cgtcgagcag ccgatctttc ccttggccaa cccagcagggg
1321 tttccatctc cgtctggcctt tggctctctat tgcacactc aactcagctct
1381 gcatctttca atggagagag ctaattgtaag cctctactcct atttatcaga
```
3061  aaataagctt  ttgtatctgc  cagtgaattt  actgtactcc  aaatgattgc
tttcttttct
3121  ggtgatatct  gtcctctctc  taattactga  aagctgaagc  gttaaagga
ggaagaaga
3241  agaatgcgct  tacacacatt  gacacgcagc  ctctaaacc  tggatttccc
ttatgatgc
3301  cccctttttta  gacactaatt  ttttaatct  tacagtctc  gaaatatatt
gattttttac
3421  gtcttaaggt  ttggggacat  tataaacttg  agtacatttg  ttgtcactcat
gttacttctcc
3481  aaattgtatg  gatgggaggg  agaggtgtct  taagctgtag  gcttttcttt
cttcgctcct
181  atggttttgt  gaaaaattaa  acaccgccct  gaagaggagc  ccgctggggc
gacgagcacsggag
241  ggcgcagagtg  ctggcccagg  tgctgcagag  gacgcacttc  cccggcccgg
301  cccggcgcca  acggcatgac  acaactcggcg  acagctaacg  gggacgacag
ccacgacctg
25  1  gacagtcgcg  gatcctgtga  cacctccggg  cagcccgcca  ttttgtgctc
cacgcagcctg
61  ttgctatcctc  cttaaaccggg  cttccccctg  ggcccccggc  cttctccggg
tcttgcctctc
121  ttctatgtga  gcattccggg  cacagcttc  tcagaccccg  ctctgtgggc
tggagaatag
181  atggttttgt  gaaaaattaa  acaccgcctc  gaagaggagc  cccggctgggc
gacgagcacsggag
241  ggcgcagagtg  ctggcccagg  tgctgcagag  gacgcacttc  cccggcccgg
301  cccggcgcca  acggcatgac  acaactcggcg  acagctaacg  gggacgacag
cccgcagccg
35  361  atcagacgctt  tgttgaaggc  tgaaatcgat  ggagaagac  tcggcaactg
tcttttctct
421  caagcgcctct  tctagatcct  ctggcttgaa  ggatcttgt  tcaagtctac
cacttgtgtat
481  ctgaaagaaa  agcccagctga  cctgcacaac  ctagtccccgc  gcagcgaccc
ccacgctcctg
541  accttcaacg  gggacgttiga  acacagcttc  aataagatc  aggagtttctc
ggagagacc
601  ttgaccctctg  aaaagtacc  ccaactggct  gcaaaacacc  gggaatccaa
cacagccggc
45  661  atcagacatct  ttgacatcct  ttctgccctc  atctaaata  ccaagcagcga
gcaatgcct
721  gctttgaaag  gagccttaac  caaggtcttataaattgg  atgactacct
gacacccct

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>chloride intracellular</td>
<td>CLIC5</td>
<td>NM_016929</td>
<td>17</td>
</tr>
<tr>
<td>channel 5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

73
781 ctaccagagg agatgagcgc caacacttgt ggggaagaca aggggtcccg
gcgcaagttc
841 ctggatgggg atgagctgac cctggctgac tgcaatctgt ... gaaaccatgt ggtttgaact tgaagaaaaa
tgtagaccca
2341 tctgggttaa ttttcctaca atctgactca actgccaggt gaaaaaaaaa
aggaaaaatt
2401 tttaagctaa tatttcactc ttttgtcatt ctccttaagt ttcatctcct
aaaaagctta
2461 cccagcctga gcttggggac ctgtgcagag gaaactaaga ... ttgacttgca cagcactatt gggggtgggg
pgaagcagggt
3 961 agtgggagac gaaggcagaa gcaagagtca aactcagaat gactgagttg
aattcactgt
### Gene Name: PERP, TP53 apoptosis effector

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>PERP</td>
<td>NM_022121</td>
<td>18</td>
</tr>
</tbody>
</table>

```plaintext
1 gcttttggctg ccgccccccc gctcgccttc cactccccctg tgcgccccgt
gtcttgattg gtattgactg attcaaataa agttggttta
361 gtttcatcat cctggtgatg ttttcatcccc cctccctctgt
ggcgcaccctg
gctgtgcttct tgcggcctggc ctgccgagcgc tgcctgctgga tcctgcccct
481 tccacccctg
gtgaagctgag atggactcca gaagaagaaa ataaatgctt ttctaatctct
961 attttttaga gatatggcg gtggactgct acacttttata agttttaaat
901 tttttaagt agtttttag ttttcatgct atcttttatt atgtttttgg
gttgctgag
gtttttaagt ctcttctccc tgggactcga tatgtcatct aggaaagtac
ttttcactaa ttactactac ttttcatctct atcctataa
gagttgaga ttcagcagtt ttaacctatac tattttagct
gattttataa actatcagtaa aacaatattta acacctttata aggttaaat
921 gagcttcgaa ctcaatccctg ttttcatctat acatccctac gcacccgacg
gtttttaag gttttgctgat gtcgctcct cccgcccggc ggttttctctt
ttttcatcaaatctc tctctctctagc gcacccgacg
```

This text appears to be a sequence of nucleotide bases, likely part of a genetic or biological study, possibly related to apoptosis or TP53 apoptosis effector gene analysis.
1321 ttgccatagt ttgtaaggct ttctctaaag tgtgaaatat ttagatgaaa
1381 ttttctcttt
1441 taaagttctt tatagggtta gggtgtggga aaatgctata...tgagcaaatt gtcctgatag gatcaaagac
1501 atctcaggtt
1561 ttagtcatata tctctttggt aaacagatgg taataaatct
1621 tgtgtttata tgttcagaac cagagtagac tggattgaaa gatggacttg
1681 tcttgacagt gttgagtgaa aagctaatgt agacagggtc
1741 aacagtctta aatctccttt tgtcttttgtt taataaatct
1801 tatctgggct ccatgactga tagatctgtt aagttgtgta gtaaagcatt aggaggtctg
1861 ctagattat tatttctgta gttgagtgaa aagctaatgt tagttaggttt
1921 tcccaataaa ccaggtattc t

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>spleen tyrosine kinase</td>
<td>SYK</td>
<td>NM_003177</td>
<td>19</td>
</tr>
</tbody>
</table>

1 aggaagagcc gcggcaggcc gcggctgagac caccctgccgc gcggctgagac agagagagag
30 61 agggggtggc cccgcgtgc gcgccccccc gcctccacct gcgaggtgac acacacccacac
50 121 aggtgtgtgc cctccggccc ctgaagcagct gcaagcgcac gcaggtgatg ccagctgccgcac
721 181 caacattctt cctctccttc caactcacc cggaggcagc ccaggtattc gttgagtgaa aagctaatgt
90 241 gggggcagta gttatatggt ctatattgctg cggcagagcc gcaactaacct
108 301 gcctcgctcc tgcggccagcg gcggcagcagc caaccacata ccatcgcagc gcggcgacgt
cagctgcacat
126 361 gcacccacag ccctgcagcg cggcagagcc ccagctgccgc cggccccacat
144 421 caacctcagg agtgtctgatg ccttgctctgc ctcctcaaga gcctctccaa cccgggccccag
162 481 gcggctgacgc ccaagctgct ggcctttggc ctaaagctgc agtttaagag aaaaacctcctacgcggtatat
180 541 gttgagcagca ttgagacac gcagggtcag gcattcgagcc aggccccatcgcgc
198 601 ctgccagag gcggctgacgc ccaagctgct gggcctttggc ctaaagctgc agtttaagag aaaaacctcctacgcggtatat
216 661 ccgcttccgc gggacagctgc cagctgccgc cggccagagc cgggccccacat
234 721 ttctctgctcc gctgcagcagc ccacgctgccgc cggccagagc cgggccccacat
781 aaggtgctgc actatcgcactgacaagagctactctccattccccgaggga
gaggtgctgc actatcgcactgacaagagctactctccattccccgaggga
841 aagaagttcg acacgctctg gcagctagtgtagcattattcttataagcttcattctctccattccccgaggga
841 aagaagttcg acacgctctg gcagctagtgtagcattattcttataagcttcattctctccattccccgaggga
901 ttaagagttccttatcgtccatatgtaaaaaatcggcacactgggaatggaatgc
961 gcccgtccac aacttccagg ttcctcatcct gcaacctttgta cagcgggttggaatttactcatcgc
1021 agaatcagtcctttctttaaccacaagagctactctccattccccgaggga
1081 aacccgcaag aagttaccttgctcatgactcctccattccccgaggga
1141 gcacagatcctgtgctgacagcattttggaggagttg
1201 taggctgacgtgtcatgtcagcagcattttggaggagttg
tagctactaccttactttttttttc
1321 atgaaaaaagatttcaaatatttttccttactttctattttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
2401 ggcccaggga cattgcagag tggcctagag cactctcacc ccaagcggcc ttttccaaat
2461 gcccaaggat gccttagcat gtgactcctg aagggaaggc aaaggcagag gaattttgct
2521 gcttctacgg ccatgagact gatccctggc cactgaaaaag ctttcctgac aataaaatg
2581 tttttaggct ttaaaaaaga aatacaagttt gaccagtgca gttttcaagc atgtacctag
2641 ttaagggaaag aaagaaaaaa

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solute carrier family 12 (sodium/potassium/chloride transporters), member 2</td>
<td>SLC12A2</td>
<td>NM_001046</td>
<td>20</td>
</tr>
</tbody>
</table>

1 ggtggtcctct gttggcgtcc aggtagcgcg cggccccgcag ggcggagggga gaaagacctt
15 61 ctccacctgtg cttgcggcttg tggccaccgc cggccagggg tgtggaggcg gcggccgccgg
121 agacgtcccg cgggctctcg agttccctggg ggggtccggc agctatgga gctctctggtgc
181 cggggtcccc ctccccggtcgc ccgggactgg ccggggtccgg cggacccggcg
cggccggcgg
gccggtcctg
tcaggtcgcc
241 cgtggcgggcg aggcacggtg gaactgcggc gcacccctggt gcctctcggtgc
301 ctggcgccccgg gagccccggg gcggggaggtg gccggagtttg ggcggagtttg
ggcgcggggtgc
cggggctctgg
tggtgacctgg
gttccgagaaa ccgggccggc ggccggtccgg gcggccggccg gcggccggccg
cggggccggggtgc
gccggtcgcgc
cggggtcccgg
gcggccggagc
gccggtcgcgc
gtgccggccgc
ggggggaggtg gacgctggggc cggggtccgg gcggggtccgg
gggggtccgg
361 acgggctcgg cagacccccg gggccccacc cggacccagac cggcgtgggc
421 tttccgagaaa ccggggcgggc ggccggttccgc ccggccggttg gcggccggccg
gcggccggccg
gggggggtgc
ggaggccggggtgc
cggggccggggtgc
ggtctggggtgc
gtcagagccccggc
601 caagcctcgt ctctcggcttg cagacccctg gcggctagatgc gcctctcggct
721 actccggcgg cgggctgcgg agtggccagg tacaggacta ctattatgac accccacaaacgc
781 acacatttct cctcgccaaa tttggggcctgc acaccatggag gctgtgtgcc
agaggctctag
841 actacccggc caccgccccg cagctgggctg agaaacgctgt gcggccctagc
901 tccacgcaggt gtggagaaaaa gaaacctttttg ccggtggttc gcggactggtg
gtaggcctag
tttccacagc agatgcctgtg gtctacgtata cttgcaaaaag aataagaggtc
gctatgtttg
961 ctccacccgagagtgcgtttgt gcggcccagc gtttacacat ttggggtggtg
gtatgttttga
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>guanylate cyclase 2C (heat stable enterotoxin receptor)</td>
<td>GUCY2C</td>
<td>NM_004963</td>
<td>21</td>
</tr>
</tbody>
</table>

1 cgcaaagcaa gtgggcacaa ggagtatgtg tctaaacgtga tggggtcat
gaagacgttg
61 ctgttggact tggcttttgtg gtcactgctc ttccagcccc 99tggtctgc
ccttagtccc
121 caggtgagtc agaactgcac caatggcagg cggtgatcag cgtgacacc gc
40 181 tcagccttgg cagagccccct gaaaaacttg gaagatgcgg tgaatgaggg
gctgaaata
241 gtagaggagc gtcgacaaaa tgctggccta aatgtgacgt tgaagcctcagttc
tttcatgat
301 tcggaggtgtc tgattcctaa tcagagggcag tgcgggagta gcacctgtga
agggctgccac
361 ctactcagga aaatttccaa tgcaaacagc atgggtgtgct cctccatagggc
cctcgcagct
421 acatactcca cctccccagat gtacctggac acagaattga gctaccccat
gatctcagct

84
481  ggaagttttg  gattgctatg  tgactataaa  gaacccctaa  ccaggctgat
                 gtccacagct
541  agaaagttga  tgtacttctt  ggttaacttt  tggaaaacca  acgatctcgg
                 ccaggctgat
541  gaaaccttaa
601  ctctttcttc
       caggaatcta  tcaccaacaa  aacgagactt
                 ggtctccagct
661  tattcctgga  gcacttcgta  tgtttacaag  atggtacag  aaactgagga
                 ctctttcttc
721  tcaccttaatg  ctctggaggc  tagcgtttcc  tatttctccc  acgaactcgg
                 ctgtttctgg
781  gtgttaagac  aagataagga  gtttcaggat  atcctaatgg  accacaacag
                 gtgttaagac
841  gaaagcaat
901  ctgactatat  gaaaaatgtc  ctttgttctga  cgctgtctcc
                 ggtctccagct
961  ccttcctctcc
       caggaatcta  tcaccaacaa  aacgagactt
                 ggtctccagct
1021  tatttgaaatg  gaatccctgtc  ctttggacat  atgcctgaaga  tattttcttg
8aaatggagaa
1141  gagactgtaa  ccccccacat  tcgctcatcg  ttcaagaatc  tcacttttta
                 tgtacttctt
1201  cctctgagca  ccaagaaata  caaggtcttt  tggacctatg  atacccacgt
                 gcagttcctgct
9aattaaagcc
1261  tattcctgtgg  atatgagccc  cacatctact  tggaaagaatc  cttaaacttcc
                 tgtacttctt
1321  acagggccggg  gccctcagat  cctgtatggt  gcagttcctgct  ccctccactgg
                 cctttgagc
1381  ctgctctctgc  tcgctgcctct  cctgatgctc  agaaaatata  gaaagatg
                 cgtacttctt
1441  cagaaaaatg  gtctccacat  tccctctgaa  aatatctttc  cttgttgagac
                 ccaagttcctgct
1501  aatcatgttta  gccctcaagat  cgtatgtgac  aaaaacgag  atacaatcca
                 gcagttcctgct
1561  cagtgcaaat  acgacaaaaa  gcgagtctt  cctcaagatc  tcaagcacaa
                 ggtctccagct
1621  tctctgaaat  acaggtcctgt  ctggaagaatt  aagttgcttc  agattgacta
                 ccaagttcctgct
1681  ttttacccaagtttattct  acggcacagtt  gaaaccttga  accatgtatct  tccctccggt
                 ggtctccagct
1741  gagagagagtgg  aggtagtctgt  ctctgttctt  taacggcata  aagttgcttc
                 gtttacccaagtttattct
1801  atgtagttggg  aagttgcttc  ctctgttctt  taacggcata  aagttgcttc
                 gtttacccaagtttattct
1861  tctctgtggaca  ccaagaaata  caaggtcttt  tggacctatg  atacccacgt
                 gcagttcctgct
1921  acagggccggg  gccctcagat  cctgtatggt  gcagttcctgct  ccctccactgg
                 ccaagttcctgct
1981  ctgctcctgc  tcgctcctgc  cctgatgctc  agaaaatata  gaaagatg
                 ggtctccagct
2041  agctatggga  tcatcgcaca  ggagatcatt  ctgctcctgc  agctatggga
                 ggtctccagct
2101 tgtcgggacc ggaatgagaa gattttcaga gtggaaaatt ccaatggaat
gaacccttc
2161 tgtcgggacc ggaatgagaa gattttcaga gtggaaaatt ccaatggaat
gaaacccttc
2221 tgtcgggacc ggaatgagaa gattttcaga gtggaaaatt ccaatggaat
gaaacccttc
ggacacctca
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga tattcttgga aacagcagag gaaaaagagc
tctcgaagag
tctggtcagtt
tttgtggaact
ttttcagact
ccttgcctaaga aataaaagct aaagg
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>transmembrane 4 superfamily member 4</td>
<td>TM4SF4</td>
<td>NM_004617</td>
<td>22</td>
</tr>
</tbody>
</table>

1 cttcaggtca gggagaatgt ataaatgtcc attgccatcg aggttcctgct atttttgaga
5 61 agctgaagca actccaagga cacagttcac agaaatttgg ttcctagccc caaataactg
10 121 atggaatgg cgagttgcctg cagaaattgg ttcctcttcctg ccgcctggtga
15 181 aggaggcagct gaggagcttt ccttgagaga ataaatgtcc actctaagga
20 241 gggggtctcgt cctctgcatcg cctgtgtggtt ccctttgcagcc
25 301 gctaacatcc ggtttttccgg gaaagctttgt cctccacgac
30 361 caaagagatacg cgtttttccgg gaaagctttgt cctccacgac
35 421 ctgggtttctt gggcccctga gcggtctttcc cccctgcttga
40 481 ggggagccgt cttggcgagt ttttttgggg gaaagctttgt cctccacgac
45 541 ggtcttttcgg gccactcctg ccggtcgttcc cctctgcttga
50 601 aataagactcat ggggctaccc ctttgcgttttg aggttcctgct
55 661 aacaaggtccc gaggccctct ccatgccactc cctgtgtggtt ccctttgcagcc
60 721 ctgggtttctt gggcccctga gcggtctttcc cccctgcttga
65 781 ggggccctcct gccttgcttga ggggccctcct cccctgcttga
70 841 acctccgaga tgacgtctct cagacttacca gcatgacgcct ctaatttcct
75 901 ttttttttttt ccttgcttga gttgcttgtg actgatcttt tgaggctgtc
80 961 gggtttttttt atgtatgtca aattggctatc ggtggtctgt tggctgtaaa
catttgacct
85 1021 tatttttttttt ccttgcttga gttgcttgtg actgatcttt tgaggctgtc
catttgacct
90 1081 tatgtaatatt ccaacagggt ccaacagcta ttcttcgact tttttttttt
catttgacct
95 1141 atgtatagta ctgatgtaag tttgcttttt cttttttttt
100 1201 tggaggtttt accttacatg tttgcttttt cttttttttt
catttgacct
105 1261 ttctggtttttt ctttggcttta ctttggctttt tggggagatgg
110 1321 ttctggtttttt ctttggcttta ctttggctttt tggggagatgg
115 1381 gggttcacgtt tttattgctattttttttt gggaagtttt
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>transforming growth factor, alpha</td>
<td>TGFA</td>
<td>NM_003236</td>
<td>23</td>
</tr>
</tbody>
</table>

```
1  ctggagagcc  tgctgcccgc  ccgcccgtaa  aatggtcccc  tcggctggac
gagctcgccct
5  gttcgctctg  ggtattgtgt  tggctgcgtg  ccaggccttg  gagaacagca
61  gtagtgcagac  ccgcccgtgg  ctgcagcagt  ggtgtcccat  ttaatgact
gccagatcc
tctggtgccgt
121  gttggtctgcc  agccagaaga  agcaggccat  caccccttgc  gtgtgtgtct
ccagtcggtc
136  cctggtctgc  cttatcatca  catgtgtgct  gatacactgc  tgccagtccc
gaaacactg
tcctggtccc
301  tgggtgtgcgtc  cgagatcagat  tattgttttc  tgtcatttgc  cagcaggtcc
agtaaagaag
721  ttaggtgccag  ggtttttttta  gtgggtgctg  tcagcggacc  acgatccttc
agagagataa
781  cctaggttgg  agggttgtgtg  gctgtgtgct  gtgcagcctt  cagaggtacc
agccacaggt
901  attctgcttc  cttctgtcag  ttcttttctt  cagcagggca  ttcacattgg
agagatgacc
ttcagccctc
135  gaaagatgaa  gggagggaaa  cttgatgttttc  gacatccttc  tatctactgc
taaatgcctc
1021  tgaacagatc  cttctgcttc  gttctgtgag  cttcagcctt  ctttcttttc
tgatgttattt  gatgtgcc
1081  ggacagatc  ccagccatc  gccaccctttg  gatgtgatg  tctatctttt
taatgcctc
1141  tggttttttt  ttaatctctc  ttaggtcagat  ttaaatagat  aaaaaaccctt
tcaatccaagctc
1201  ctaatcacat  gtaagtcttg  gttctgagtc  agatgatgtg  ctgactgctc
tgattggtctc
1261  ttagagacagc  taggttccgg  gaaaaacttt  ccaggggttcg  gatgtgacag
tcagagccg
1321  gttactggta  gagaattctt  taggggttctt  gaaatggtgc  cattgaagcc
acaaagccgg
1381  taataatgcctc  aataccttg  gggagaaactt  tttagcaatct  ctcagccag
ctgttgtacct
1441  ctaatgatgat  gtaaatcgtg  gttctgatggc  tcaagagctc  cctttctttc
tctgatgttattt  gatgtgcc
1501  ctaatgatgat  gtaaatcgtg  gttctgatggc  tcaagagctc  cctttctttc
tctgatgttattt  gatgtgcc
```

88
1441 ctctgttggg gagagaggg aagtggtgtg tcttaacag 
atacatattg
1501 tactgctcag gatgtgaatg gctgttcctc tctggaat 
agggacctg
1561 atagaaccat cagccagagc cataggccat tccagccaa 
tccatggaag
1621 cattatattt cagttgggtt gttttgatgt tattttttaa 
aggagatgga
1681 attttttttc tttctctgga gagacagtgg aaaaactttc 
cagtaatttt
1741 tttttttttc ttctaatgca gatgcctggc ttttattaaca 
ctttaataca caagtcacct
1801 cccatttcct tcctcctcaag aagagaggg tacatcatt 
gttcctctga 
1861 cccatatgggac ctcggtttga gacacagcttc gttcccttca 
atatacatgtg 
1921 cctaggaat tttttaaaa atttggaaac agagatactga 
tttataactt 
1981 tatggtgagg aatagctact aatatctgaa aagagaggtg 
tttctatgta
2041 agtgcaacac tctatttaaaa ttcagacact cgggtgaga 
attggctgaa 
2101 aatatttttt cttctcttact acacacac acactttaat 
tactaataa 
2161 atgaatatttt ttcttttataa ttcagacact ctttatttatt 
gttgctgaa 
2221 ttttgccact tttatttaaaa aatggccacac ctgaagcact 
cagagcagac 
2281 cccctagggac ctagttgcag ttccttcctgc gtttggaag 
aagagactt 
2341 tactatcccc gcctccttact ttcagacact cgggtgaga 
attggctgaa 
2401 aatacatgga cactgttttt tagacctctaa tcaatcagaa 
agtgcaacac 
2461 gcagagcggc atggagaagct ggcggacagt ggttccctgc 
ggaggaactt 
2521 gacagtttgtg cttgtctagc cttatttttt cagctggcac 
cgagctgccg 
2581 cccattcccc tgggtctgtga aacagcgcttgc tccagcag 
acacacac ccctgttctct 
2641 aagtgccttc caagctgtttct tttgggcaca tggctggaatt 
tgtgcacatg 
2701 gcaagctggg gctgcttttc cttggtgaca cagctgcaga 
attttgggag 
2761 tttttttttt caggtgcataa gcctgttttc gcctgttttc 
ttttctttact 
2821 ttgtagtttt cttgcctgtttc cttggtgaca cagctgcaga 
atcttcttc 
2881 cctatatgct aatggcgtttg cttggtgaca cagctgcaga 
tttttaaaa 
ttcccagatc 
2941 tagagatcag tagaggtgtttc gaccttcctt taaaactcat 
tttttttttt 
3001 gcgctgtttc gacagctgtttc gcctgttttc gcctgttttc 
ttttttaaaa 
ttggggacta
3061 tttttaaaag gcctccaata ctcttatgga gcctggattt ttcccactgc
tctacaggct
3121 gtgtacttttt ttaagcatcc tgacaggaaa tgtttttcctc tacatggaaa
gatagacgc
3181 agccacacct gatctgggaag acagggcccc ggctggaacac agctgggaacc
aagccaggg
3241 tgggctggcc atttgtctcc cgcaggagag atgggcagaa tggccctaga
gttctcttcc
3301 ctgagaaagg aqaaagaagat gggatgccca ctcaacaccacc cacactggtagagggaggaga
3361 atttgtgctt ctggagcttc tcaagggatt ctcacccacc cacactggtagagggaggaga
3421 acatcttcaagc caggttttcg agggcacatg ggtcaccagat tgctttttca
3481 cccggatgga ctaatgaggc tctacactg ctcaggagac ccctgcccacct
3541 tgtgtggttc tctctcctaa cctgcagcgact cacagaagga ggaatgacctc
3601 aacacaagac acatgcacac agtaagacaca acatgtatat ttttaaatgt
tctacacattaa
3661 gcacgctggct ccctagccat gtatattacca ggttttcatg aagatctagtttgtttggtcttc
3721 gagagagagac gagtactgaa aacagctctgctctctccttct ctttaaataattctacattaa
3781 ggtcaacctt ccattatcttt ctttttaata aacctgatgc ttttttttag
3841 ttctatggtct gtatattttg cactgaaaag gtaatatattt aatgttttaaa
3901 ttttttttga tatahtatgt ttattcttttaa cttcctcc很容易获得
3961 tttcctaaaa cttaataagag gtagtgagac aatataagac gcaatgtgtctttcctattaagag
4021 ttctcactt ctatcacccttattttgaaag gatgagagaatgagtttcatccctacaacc
4081 tttcctaaaa
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTK6 protein tyrosine kinase 6</td>
<td>PTK6</td>
<td>NM_005975</td>
<td>25</td>
</tr>
</tbody>
</table>

```plaintext
361 caaggttgag tgcacctcaat tgagccatctga attttctctgt gtctttgctg gcaatccaac
421 ctcatgtcctta aagctcaagg atgagagagt ctattggaaa tggaccatga attttcctgt gtctttgctg
481 ctccagaa gacatctgta gatattccaa gacagctgtg tgcactcaat gcaatccaac
541 ggttttcgaa gaatccagtc ttaagctagt cagctccact ctattttgga tggaccatga
601 caggaagagg aaaaagagg tggaccatga attttcctgt gtctttgctg gcaatccaac
661 ctctggtcctta gcagtgaccc gagatctggtc cggaatctgcg ctcatgccta aagctcaagg
721 atggagagagt ctattggaaa tggaccatga attttcctgt gtctttgctg gcaatccaac
781 ctctcagcagaaa gacatctggtc cagctccact ctattttgga tggaccatga
841 ggtgcttttgctc aagagatgtc atgagagagt ctattggaaa tggaccatga
901 atggagagagt ctattggaaa tggaccatga attttcctgt gtctttgctg gcaatccaac
961 caggaagagg aaaagagg atgagagagt ctattggaaa tggaccatga
1021 ctctcagcagaaa gacatctggtc cagctccact ctattttgga tggaccatga
```
601 gagggaggag ttcacgctct gcaggaagct ggggtccggc tactttgggg
aggtcttcga
661 ggggctctgg aaagaccggg tccaggtggc cattaaggtg ... attttgaagg tgtctctttc atccatggtt
aagtcataaa
2161 aagcttattg gttttggttt tgactcacct gaaagttttt ttggtttaaa
agaagaatag
2221 gcggggcacg gtggctcatg cctgtaatcc cagcactttg ggaggctgag
gcaggtggat
2281 cacgaggtca ggagatcgac accatcctgg ctaacacggt
gagcttgcag
gcagctaacgt
tgagcgaga
2341 aaatacaaaa aattagctgg gtggctgtgt ggggttggtgcc gcctgtagtct
taatgttacct
2401 ggaggctgta ggaggcagac cctttccgct atagtggagt
gacttttcacc
tgtagtgtaa
2461 tcgcgccact gcactccag ctggccgaga gacgagact ccactctcaaa

1 acaggcacag gtgaggaact caactcaaac tctctctct cggaaacgc
ggtcattgtct
61 cctcccgagag tggccctgagct cagctctctct cttctctgc tggaaacgc
ggtcattgct
121 ggccaaggtgg ggtgattggc aagagctcta catcttcttct
tggaaacgc
ggtcattgtct
181 ctctgccttctc agcctcagcact tctgtggtcct atagcagctg tgtggatcc
tgtagtgtaa
241 gtggtgagag ctggtgctgct gacagatgctct cttctctctc
tgtagtgtaa
301 cctgtgctgct gtcagctgctgct gacagatgctct cttctctctc
tgttagtggct
361 caggggagag tctctctctc agcctcagcact tctgtggtcct atagcagctg tgtggatcc
tgtagtgtaa
421 ccgggagag tctctctctc agcctcagcact tctgtggtcct atagcagctg tgtggatcc
tgtagtgtaa
481 caggggagag tctctctctc agcctcagcact tctgtggtcct atagcagctg tgtggatcc
tgtagtgtaa
541 gtggtgagag tctctctctc agcctcagcact tctgtggtcct atagcagctg tgtggatcc
tgtagtgtaa
601 cctgtgctgct gtcagctgctgct gacagatgctct cttctctctc
tgttagtggct
661 tctgttgtgct gtcagctgctgct gacagatgctct cttctctctc
tgttagtggct
721 aacaggagag tctctctctc agcctcagcact tctgtggtcct atagcagctg tgtggatcc
tgtagtgtaa
781 aacaggagag tctctctctc agcctcagcact tctgtggtcct atagcagctg tgtggatcc
tgtagtgtaa
841 tctcttctagt gtcagctgctgct gacagatgctct cttctctctc
tgttagtggct
901 tcggtcagag tctcttctag ttcagagctgctgct gacagatgctct cttctctctc
tgttagtggct
961 cctagctgctgct gtcagctgctgct gacagatgctct cttctctctc
tgttagtggct
1021 tcggtcagag tctcttctag ttcagagctgctgct gacagatgctct cttctctctc
tgttagtggct
1081 tcggtcagag tctcttctag ttcagagctgctgct gacagatgctct cttctctctc
tgttagtggct

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>epithelial V-like antigen</td>
<td>EVA1</td>
<td>NM_005797</td>
<td>26</td>
</tr>
</tbody>
</table>
1 attaaggact cggggcagga ggggcagaag ttgcgcgcag gccggcgggc
gggagcggac
6 1 accgaggccg gcgtgcaggc gtgcgggtgt gcgggagccg ... ctggagcccc acatgaacta caccttcacc
gtggaggccc
1381 gcaatggcgt ctcaggcctg gtaaccagcc gcagcttccg tactgccagt
gtcagcatca
1441 accagacaga gcccccaag gtgaggctgg agggccgcag caccacctcg cttagcgtct
1501 cctggagcat cccccgcgg cagcagagcc gaggtggaa tggctctgctg
gcaggagttt
gaggacgttt
1561 acagccgctg cccccgcgg cagcagagcc gaggtggaa tggctctgctg
gcaggagttt
gaggacgttt
1621 acttcctcca tccacacgct tcaaatgtgc gcccacccca ggggttctcc
cgccggggctc
1681 ccctggagcat cccccgcgg cagcagagcc gaggtggaa tggctctgctg
gcaggagttt
gaggacgttt
1741 cccttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
1801 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
1861 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
1921 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
1981 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
2041 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
2101 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
2161 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
2221 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
2281 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
2341 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
2401 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
2461 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
2521 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
2581 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
2641 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
2701 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
2761 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
2821 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
2881 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
2941 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
3001 accttcctca ttggctctgct gctctggggct tctgtgctgtg
gcaggagttt
gaggacgttt
3061 ccatctgagc ctcgacaggg cctggagccc catcggccaa gaatacttga
agaaacagag
3121 tggcctccct gctgtgccat gctgggccac tggggacttt atttatattct
tgaggagatg acagtggaggg cttggccccag cgccaagtaa acagggtacc tcaagccccca
3181 tccccctgca acttccgctg aggggtctcg gatgacaccc tggcctgaac
tgaggagatg
tttctctaca
3241 aagcaccggcc acgtccccgc acctctgttc cagagagtc ctgcgcacccc
tctgcgcagc
tccctcttag
3301 agggccacgcc acgccttccg atccccctgg tggccttgag cagagagtcct
tctgcgcagc
tccctcttag
3361 atagggatgg atggcctccct gctgtgccat gctgggccac tggggacttt atttattttct
tgaggagatg acagtggaggg cttggccccag cgccaagtaa acagggtacc tcaagccccca
3421 agatggaggg ctggaggccgg cctgcgagacg cacacagctg
tctgcgcagc
tccctcttag
3481 ctaagagggc agactgtgaa tttgcgtcgg tgggacaccc cgggccctgag
tgaggagatg acagtggaggg cttggccccag cgccaagtaa acagggtacc tcaagccccca
3541 ttgcctctttt agacccctcg ccgccatctct acatcctgac tggccaaacc
tgaggagatg acagtggaggg cttggccccag cgccaagtaa acagggtacc tcaagccccca
3601 tgggctctttag caagatgctt ggttgtgttg aggttttttaa atatatatttt
tgaggagatg acagtggaggg cttggccccag cgccaagtaa acagggtacc tcaagccccca
3661 ggagagatgg tgtgtgtgtg gcagggggcc ccgccagggg tgggacaccc cgggccctgag
tgaggagatg acagtggaggg cttggccccag cgccaagtaa acagggtacc tcaagccccca
3721 cattcgtgag ctggggaccc agggacccgt gctgagacagc
tgaggagatg acagtggaggg cttggccccag cgccaagtaa acagggtacc tcaagccccca
3781 ccggccccactt ctctcactcttt tgggataagt ttctatttct tcagtgtaaa aggttttttaa atatatattttt
tgaggagatg acagtggaggg cttggccccag cgccaagtaa acagggtacc tcaagccccca
3841 tgttgtggacca ttttttttca ctctctttat tttttttttt atatatatttttt
tgaggagatg acagtggaggg cttggccccag cgccaagtaa acagggtacc tcaagccccca
3901 tgactttatttt ctgcctctga atataagttgc aagatgatca aaccgaaaaa

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrin, alpha 6</td>
<td>ITGA6</td>
<td>NM_000210</td>
<td>28</td>
</tr>
</tbody>
</table>

1 aacgggctca ttcacgctgcc gcgcagctgg ccggcgggag ccggcgagctgc gcggagctga
gggtgcaaatc
d31 gacgggctccc ggggggtcgg ccggcgcagc gcggcagcgc cggcgagggcc
gggggtgggc
d35 gggtgtcctgta
d41 gcacagccggccc cgcagcccgg cggccgcgggc gcggcgggctgc gcgggcgggctgc
d51 cgggcggggcctttctacatgcgcgggcggggcgggggcggtgcgggcggggcggggcggggcggggggc

d40 cgggcggggcctttctacatgcgcgggcggggcgggggcggtgcgggcggggcggggcggggcggggggc

d45 cgggcggggcctttctacatgcgcgggcggggcgggggcggtgcgggcggggcggggcggggcggggggc

c41 cgggcggggcctttctacatgcgcgggcggggcgggggcggtgcgggcggggcggggcggggcggggggc

d48 cgggcggggcctttctacatgcgcgggcggggcgggggcggtgcgggcggggcggggcggggcggggggc

dataagagcttc

dataagagcttc

dataagagcttc

dataagagcttc

dataagagcttc

dataagagcttc

dataagagcttc

dataagagcttc

dataagagcttc

dataagagcttc

dataagagcttc

481 cgacaggtat cctaggagct gttttcaaat ctaagtctga tcgcattgtg
aatgaaactc
541 tctatgaaaa cacaaagctt ttgagcgcca caggggaaag...acatcatatt ctttcagaga agtgtcccag
gacatgataa
601 taagatgcaa tttgaatctt catcatacga aggatacttt ctagcttgtg
aaaaagagag
721 agacctttttt aaactcatttt tgaaaaaaga ggatgaatttg ggggatagat
cctataatgtt
781 cactgttcaa aacgaagact agctattaaa atttcagcct gggcgagctt
5 841 gtaatcccg ccctttggga ggctgaggcg ggagatcac cagaggtcag
gtttcagact
901 ccagctgcac caccatggtaa aaccctcac cttactaaaa atacaaaaa
attagctgagt
961 gtagtgacgc atgccttcaaa tccccacctc taagagggct gaggcaggag
ggctgaggccaa
1021 actccggagg tagaggttgtg ggtgagccca gattgcctctaa
tgcgctttaa
1081 caaagcgaa ca actccatctc aaaaaataaa ataaataaat aaacaataa
aaatattcata
1141 atgtg

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>bone morphogenetic</td>
<td>BMP4</td>
<td>NM_130850</td>
<td>32</td>
</tr>
<tr>
<td>protein 4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 gagggaggg ccgccccggga agaggaggag gaaggaaga aagaaagcga
20 ggggggggaa
61 gagaggagaag gaagatgcga gaagccagag gaggaggag gggggggga
agcgcctgcag
t25 ccggccccggca agcttggccat tggccatcctc aacactcagc
121 ccggccccggca agcttggccat tggccatcctc aacactcagc
tggtggtctg
181 tcccttgagcct ttccctgcaag tttgtttcaaa gattggctgt caagaatctat
tggtggtctg
241 tatattgtcct gttttctgtct aagacacct gattttcctgt aacactctat
ctgctgactg
301 tttatatgtc agatctctgc taggaggccg gagccatgct aagttgtgatc
tctgagacgg
361 gaagagaaaaa gctgccggaga ttcaggccca cgcgggagaa gcggcctcag
tgctgagacgg
421 tgagctctgt cggtgactgcg aggccacact tctgcagatg tttgggctgc
tgccggcccccc
481 gcagctcgag aagagtgccgc tcattccgga ctacatgcctc gagttccctcc
541 tggggagggag gaggaagagc agatccacag cacctgttctt gagtatcttg
agcgcctgcag
601 ccagcggcccc aacaccgtga gaggcttccca cccacgaaga cactgttgc
acaatccag
40 661 gagcagtgaa aactctgttt ttcgcttccttttaacctc agcagcatcc
tcgtaagcttc
721 ggcgatctccg tctgcagatg tccggctcttt ccggggacag gttggacagg
gcctgattg
45 781 gggaggggct ttccaccgta taacamattta tggaggattt agagccccag
caagaaggtgt
841 gcctgggacac ctcacacag cactactgga cagcagactg gtcgccacaca
tatgtacagcag
421 ccatgaagga gattgagcca gagccagact tcattctctg gactggtgat
gacacgcctc
481 atgtgccccg tgagaacctg ggagagcagg ctgtactgga aatgtggaat
cgcctgcac
5 541 agctcatcag agaggtcttt ccagatatact aagtctagcctg tgccttgtggag
aactcatgtt
601 ttccaccccaaa aaaccagttc ccagctggaa gtaaacaact catcaataccg
atagcagac
661 tatggaaccct ctgtctcttg aatgagtcct tcgctctctt caaaaaagtt
gctctctcact
721 tgtgagaagct gccgggtccc agcggggctg ggcgaattgtg ggtcctcaac
accaatctgt
781 actataccag caatgcgctg acagcagaca tggcggaccc tggccagcag
tcgcctgacca
841 tggagagctgt gctgaccgat gcataaaccag ctggggacat agcagagcag
gacacgcctc
901 tgcccccccggg ttcttttgag aagacgcctaa acaaggtcctgtgctccggag
ggtctctagcg
961 aaaaaattct gaaggttggtcc gcgaagcatact acgctgcatag cagacgcctc
10 20 ttctctgcc
1021 accaccaacac cgacagcttt cggatgtctct atgatgtgctc aggtgtcccc
ataaggccca
1081 tgtctcatcag acctggagtcct aaccctagtaa aaaccacatt acctggagtga
gctagtaagc
1141 ccaacaatcct agccatccagg gttgtagaat atgacggagc cacactgagc
tctgaagagac
1201 tgtgtagaact ttcctgacca ctgagccagg cgaatgctca gggacgcggc
gtgtgagggc
1261 tccagatraacc gcctggagac gccatgggag ggcggagccag cagcgcctccac
ttcagacaca
1321 cagtctgctga cgctgaccag cgcctactgg gcggggagcc gacgctttctgagctctact
1381 cagtctagct ctctgtctggg tctctgcagctg aggctctgacg catgcaacag
1441 tgtggcaggttg gcacattgacc gcttacacca cctgtctgtc tgcctctggcc
accaacccgcc
1501 tgcccccagcct cccgctgctgctg tgctgagctg tgcagctctgctc
tgcgcagactgtgctgac
1561 tgtgccaggtcc cacctctctctc ctgtgaaacc gtaacggggg cagcgcctccac
gatccacccag
1621 agctggaggcct tccaccttttt cctccgccccgc gagaggagtgactgacgataggaatag
gacaccccag
1681 tcaggaagggc agccccccagg agctcagagg atccgctgtc acgcctcagcc
1741 ggccagcacaac gcagactctct cttcacaactgg aaaccagaaa cagaaaaagaa
gagcagcc
1801 agaccccccatacgtatcttttcttgtggtatatgatt tactcacaac
aacaagctc
1861 atcatgctgtt tgaaaaaaa

107
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>transmembrane protease, serine 2</td>
<td>TMPRSS2</td>
<td>NM_005656</td>
<td>34</td>
<td></td>
</tr>
</tbody>
</table>

1 cgcgagctaa gcaggaggcg gaggcggagg cggagggcga ggggcgggga
gccgccgcctg
5 61 gacgcgcggcga ggtcatatgg aacatccca aatactcatca ttactcgatg
cgccgcctg
6 1 gagcgcggca ggtcatattg aacattccag atacctatca ... gggcgggga
10 gcgccgcctg
121 cagcaagatg gctttgactt gctgcacgga atctctactt gacacacttct
teacatcatca
181 tggataacca cccgaaacc cctatccgcgc acagctacact gttgttcgca
241 ggtgccatccg gttcaagtact acccgctccc cgggcctcag taccgggctca
301 gcagacctcc aaccccgctcg ttgcaagca gccaatatcc cacatccgccg
361 cttcaagaact aagaaagcag ttgtaacctt ttggacactt cttgccaacct
tcagatgtac 15 actccctact
421 tgggccggct gttgccctac gctgggactt catgggaga cactggccgtc
481 agagtgccag ttcacctaggta cctgcatcga cccctctaaac tgtgtgtgatg
gctgccaca
541 tgtcggccggc gggagagcg agaatcggtg tgttccgccc tacggacaa
601 tcaagctgatc tcaacctcga ggaagtctcg gcacccctggt tgccaagacg
gtcgccccga
721 gagaatagtg gatgacagcg gatccacag ctttatgaaa ctgaacacaa
gtgccggcga
781 ttgctgatata cttaaaaatc tgtaccacag tgaatggcct tcttacaaag
ccttacctcc
841 ttctcgccttg atagctcgg ggtcaacctt gatccagagc cgccagagca
tcagtttccc
901 cggtagagac gcggctcgggc gggctgagcc ctggcaggtgc agcctgcagc
tcagaacctc
961 ccaagtctgc gcaggtcctca tcatccaccc cgaatggatc gtaaacaggc
cccctctggt
1021 ggaaaaaacct cttaatcattc catggcatttg gacggcatttt gcgggggttt
tgtgacatct
tctaatatga
108 ttctcatgctc tgtagggcgg aacatccaggt agaaaaattg attttctatc
tgacaatct
1141 ctccagagac aagaaacaaag acattgangt gatgaaagcg cagaagcccc
tagtttctca
1201 cqaacactgtg aaacagagtgt ttgctgccccaa ccagccagat atgcgtccgc
cagaaacagt
1261 ctggtggagtt ctcggggtgg gggcaccagg ggagaaaggg aagaccccaag
tcaggtgcaag
1321 cgctgcaagcag tgcgtcttca gttgacacaa gatgtgac ccagatatatc
tctatgacaa

108
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>guanine deaminase</td>
<td>GDA</td>
<td>NM_004293</td>
<td>35</td>
</tr>
</tbody>
</table>

3001 tggcactgta aggtgcctgc tcccaagac acatccataa aggttgttga aatggtgaaaa
3061 cgcttcctct ttttattgcc ccttctatt tatgtgaaca actgttttgtc tttttttgta
3121 tcttttttaa actgtaaagt tcaattgtga aataatgatat catgcaaatatat aattatatgcga
3181 ttttttttcc aaagcaaaaa

<table>
<thead>
<tr>
<th>Residue</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>gtagggagcc agccctctgg gcgcgcctgc aggtaccgg caaccgccgc ggtgaacggg</td>
</tr>
<tr>
<td>15</td>
<td>gacagtgtgc cgcgcctgct gcgcctctcg accagcagac ccgcgctgcgctcctgcagac</td>
</tr>
<tr>
<td>20</td>
<td>gccaataag tgcttttttaga agaagcatct caacagggaa aactgccccaa aagatgtgc</td>
</tr>
<tr>
<td>25</td>
<td>ttaaccgctct gcaccctgtc tcctttgagaag tccttcgataa ctgcttcgcataa</td>
</tr>
<tr>
<td>30</td>
<td>gttcttggctg gcacatgctgac gtagttgattg aagttggcagag ctgcttggcataa</td>
</tr>
<tr>
<td>35</td>
<td>cagattgggt aactagctgc agcgtggctgc atgatttgctg acatataagtg</td>
</tr>
<tr>
<td>40</td>
<td>catataagtg aaaatcgtgta tgaagttgaa gctgtgaaaa acttatacc</td>
</tr>
<tr>
<td>45</td>
<td>cagattgggt aactagctgc agcgtggctgc atgatttgctg acatataagtg</td>
</tr>
<tr>
<td>Gene Name</td>
<td>Gene Symbol</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>macrophage stimulating 1 receptor (c-met-related tyrosine kinase)</td>
<td>MST1R</td>
</tr>
</tbody>
</table>

```
1  ggatccctcta ggggccagc tcggcccgag ggagctcctc cccgcctcgtcctaggtgc
5  ctatggagag cagatcggtc tcggtacctc gcgtgagcgc cttccacccgcaagcctct
4501 tcggctccgc tggctcagct gcttctccgc ctcgtgcgc gagcagcgc gtcgtgcgctc
4506 ctctgctgc tggctctcct gcgtctcgaac gagatcgcct cggccgcgc gcggcccgtcc
```
241 tctagccata cgcaatcgcc tgggcctgac ctgaagtctg
tccagagcct
301 ggccacgggc cctgctggag
361 ccacggccct cccggtgaca cagacacaaa 
421 ctcgacgctg ctggcaagctgaca
tgggcctgac ctgaagtctg
tccagagcct
481 agggacagcc gttgcatcttg cagcagcacg ctgctctctg tcaagcccacc
541 cgatgactgc cccgactgtg
tgggcctgac ctgaagtctg
tccagagcct
601 aggcggagcc tcctatattct acgtggcatc ctcactggac gcagccggtg
ttggtggaag
661 cagcggcagcc tccaggtctct caaggtctac gcctcggggat
tcgccagcct
721 ctcttggtgcc tgggtgctgc tgcacgctgaga
tgttggacc
tggccacgggc cctgctggag accctggctg ccagacgtgt
tccagagcct
781 cagctctccac aagggaggtc tcgtatactctctgctacg tcaagcccacc
841 ggcagcagccg ggcacgctc
901 tgaatatcggct ggcagcagccg ggcacgctc
tccagagctc
tctgtgytcctgc
961 ccaacccgcc cccggtgaca cagacacaaa 
1021 ccaacccgcc cccggtgaca cagacacaaa 
1081 gactgggcaag gatggtgtgc tgggccctgac cagcagcaca
tggggtgagc
tggtggagcc
1141 tggacctgctg gacacactaa ttgatagggg tgtggagggc tgtggtgataa
tggggtgagc
tggtggagcc
1201 tggacctgctg gacacactaa ttgatagggg tgtggagggc tgtggtgataa
tggggtgagc
tggtggagcc
1261 gggtgacagt gggtcagcggc gcaggtggcc tgggtgagc
tggtggagcc
1321 cagcggcagcc tccaggtctct caaggtctac gcctcggggat
tcgccagcct
1381 gatgggtgaca ggttggtggtc ctctgttgtctctgctacgc
tggggtgagc
tggtggagcc
1441 cccgtgcagccc gctgcaaggg ggcatgggctgtg
1501 ggtggtgaag ggcagagccg tgggtgctgc tgggagacat
tggtggtgataa
tggggtgagc
tggtggagcc
1561 tgggtgacagt gggtcagcggc gcaggtggcc tgggtgagc
tggtggagcc
1621 gacgctgtggc gctgccagcc ggcacagcgc tgtggtggtgataa
tggggtgagc
tggtggagcc
1681 ggtggtgacagt gggtcagcggc gcaggtggcc tgggtgagc
tggtggagcc
1741 ctaaggtgtc taacccctacca ctctgtgggctgtg
tgggtgagc
tggtggagcc
1801 cctcttggtg cctgtgggctgtg
tgggtgagc
tggtggagcc
<table>
<thead>
<tr>
<th>Start</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>3481</td>
<td>gccacctgag ggcctgcccc atgtgctgct gccctatatg tgccacggtg acctgcttcca</td>
</tr>
<tr>
<td>3541</td>
<td>gttcatccgc tcacctcagc ggaaccccac cgtgaaggac ctcacacgtctg</td>
</tr>
<tr>
<td>3601</td>
<td>ggtagcccgc ggccttgacgc ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>3661</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>3721</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>3781</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>3841</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>3901</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>3961</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>4021</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>4081</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>4141</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>4201</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>4261</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>4321</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>4381</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>4441</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>4501</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>4551</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>4611</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>4671</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>4731</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>4791</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>4851</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>4911</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>4971</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>5031</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>5091</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>5151</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>5211</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>5271</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>5331</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>5391</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>5451</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>5511</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
<tr>
<td>5571</td>
<td>acctgctcca ggctgctgca ctatgctgct gctgacttttg ccacagcttc</td>
</tr>
</tbody>
</table>

...
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>integrin, beta 4</td>
<td>ITGB4</td>
<td>NM_000213</td>
<td>37</td>
</tr>
</tbody>
</table>

1 gcgctgcccg cctcgtcccc acccccccaa cccccgcgcc cgccctcgga
cagttcccctgc
5 61 tcgccgcgcc gcgccagccgc cagctctccag cggccagccca gcgccggag
gagcaggtccc
121 gccccgaggt aggccgcaca gacagcagcc aggctggcgc
ggcagggag
181 gaagaggagat gcagggccac gcgcagccccc atggggccagg ctgctctcgg
cacgtcctgtc
241 ccgcctgacgc ctcctctggga ccttgccccaa cctgctgaag aaggccccag
tgagaggtgc
301 cagcaggtctg gctgctgtgg atagagactg gcgccctgcga acagacgaga
tgctcagggga
361 ccggcgctgc aacacccagg cggagctgtc ggcgcgcgggc tgccagcgag
agacatcgt
421 ggtctagggag gcagctctcc aaatcacaag ggagacccag atggcacaac
cctgcgggc
481 caagccagatgc tcccccccaa gcctgccccgg cctgctggag cggggcagag
agccgcattt
541 tgtagctggag gtgtttggag cactggagag cccctggaac ctgtcactac
tctggactt
601 ctctcaactcct atgtccgatg atctggaaca cctcaggag aatgggtgaga
cacgtcctgg
661 ggtctctgagct cagctacac ggcagatcctc tattggattt ggcaagtttg
tggcaaacagt
721 cagctctccgc cagccgacaag tgggcttgcag gatgtcccag ggcgagccgc
tgaccactgtc
781 cccccccctcc tctctcaaga agctctcctag cctgcagaga cttggagtgt
841 tgcagcaggt ggagagcggga tccagagcct cctgagactc cctgagggcc
gtctggtgc
901 cagctctcagc aagctctgtg gcacagggga cattggctgg cgcccgaca
gccaccacct
961 gcctgttcttc tccaccgagt cagctttccca ctgtgaggtg atggtcgagt
cgctggtgc
1021 tgccacagctg aggccgccag atgagctggtt ccacctggac accagggcca
cctacacccaca
ttggcactatg aagccgccag atgagctggtt ccacctggac accagggca
ccaccacccaca
1081 gtacagcagc caggtactcc cgtccggtgc cacccctgttg cgcttgcctcg
ccaaggccaca
1141 cagctatccca attttttgctg tcaccaactct tctcatagc tactacgcgga
gctctcacaac
1201 ctatatttcct cttctctccac ggggggtggt gcagggagac tcgctcaaca
tgtgggtgact
1261 gcctgaggag cccttctcagt gatccgcgct ccacctggac accagggcca
taggacccgaga
1321 cggagggcctt cggagagcagtc tcacctgggc gatgtctccag aagagcgagga
tgctggtctct
tcacaatccgg ccggggaagtg ggtttata aaggtgacac cttgccccgccc
tggagcagt
3061 gcagctgctg gtggaggcca tcgacgtgcc cgcaggcact gccaccctcg
gccgccgcct
3121 ggtaaacatc accatcatca aggagcaagc cagagacgtg ... gtgctgctgc ctatggcacc cacctgagcc
cacacgtgcc
4621 ccaccgcgtg ctaagcacat cctccaccct cacacgggac tacaactcac
tgacccgctc
1441 ttatttttaga gcatctcatt tataatgtag cagctcataa atgaaattga
aaatggtatt
1501 aaagatctgc aactactatc caacttatat ttctgctttc aaagtaaga
1561 ttctactcca ttaaatataa agcaagataa taaaaattgt tgcctttttt aaaagtaaaaa

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>ligand 15</td>
<td>CCL15</td>
<td>NM_032965</td>
<td>39</td>
</tr>
</tbody>
</table>

1 tgcagactga tatggattca ccactgctaa cacctcttg ggagaactac
61 ctggaaagg aaaaaaggca gcattcaca ccaccaactc ctgaatccaa
gagtcttaaa
121 tagtccccc ctcctatccc aggccttagag gattagatta atctctgaga
gggaagctc
181 ttctctgaaa catttttct tattctgctg tagctatgag gataattcgg
gaatccaca
241 ggacagttc aagctcatct ttgtcttctc ttctgtgtgc actctgacc
tttttctctc
301 ttttagaact gcagttgtaa tattatatg taaagaaga gcatctgac
tctgaccctg
361 ggacttccct gatcctcttc ttcttttaaa tacaagggca gagcttgtat
cccggggagc
421 caggaacag cagccctcgg cagccctgcc tgtcctggac
gaggtgaag
481 ctctcctcgg ctgcctctct ctgcttgctg tcttggatc
cagggccccag
541 ttcataaatg atgcagagc agatgtaag atgcaaatgc ttccactgga
aatccgata
601 ttgtctgaca ccctttcact ttgctgtgc ctgctgccctt cttacatcct
acaagcatc
661 ccgtgtctca tcattggtaa tattttggaa acgagcagc agtgctcaca
gcagggtgctc
721 atattctcct ccaagaagg gcggcaagt gcgtccaaac ccagttggtcc
ggagttccag
781 gattgctagta aaagctggaa cccctaactca atataataat aagagcaaa
agaggccag
841 ccacccaccc ccaacacccc ctgtggtttt cttggctcta aataacttaaa
aatatatatat
901 atgggtgcgt gttgaatgca aagtaatgca tctaaataa agtatccaat ttttt
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>dipeptidase 1 (renal)</td>
<td>DPEP1</td>
<td>NM_004413</td>
<td>40</td>
</tr>
</tbody>
</table>

1  cgggggggta          ctgtgcgagc          cctcaaggag          gtggctgttc          tggactgtgc
gagctccctg
5  61 ggtggcagga          ctgaaccttgga          acaccagaaca          caaccctccaa          gctttgtgac
cgggggggta
121 ggaggccggtg          ctgtctccctg          gggacttggg          tggctgagcc          gaggtaactcg
ggaggcagga
181 ccggcccatgg          cagactggct          cctcgcaccc          ctaagctctat          cctctggcac
10  gggccagcca
241 gggccagcaca          gaggccagcag          gggcagctgtg          cacaacaggct          cccggggacc
cctcctggca
301 gaggctggttg          tggctgtggct          ccttcctggct          cctgcgggac          ccggtgtttg
361 ggccagaggg          atcagctggtg          actccccctgt          catttgatggg          cacaatgacc
ttccttcctgga
421 gctgtcgtgat          atgttcaacca          acgggctgcag          ggcagagggg          gccaaccttgga
cacctctgggc
481 gggccacacac          accaacatcacc          ccaagcttgcag          ggccgggtttt          gttgagggcc
taggtctgtgc
541 ctggtcagcc          cctgctgaca          cccagaaaca          agacgcctgtg          cggaggacgcc
tggagctatg
601 ggcggtggttg          caccgcagttg          gcggcaggttg          cccggagacct          ttctctgtatg
tggaggccagc
721 gggccacatcc          attgacagcgac          gttttgggctg          cttgcgggca          ctctctacgc
781 gtaccccgct          ctcacccaca          gctgacaacac          gcccttggtct          gacaatactgc
tgggtgtgagcc
841 ggaggacaggg          gagccccaca          gccaaggctt          gtctccctttt          gggcagctgtg
tgggaagttg
901 gcagctggttc          cttggggctgtg          ggcgtgcagctg          tctgtgagcc          ccctacagggcc
961 cccctccgctg          ctgctcagccag          ccccggtcatc          ttcagcccccct          aacactgctgctgctg
1021 cgcaagccggg          cggcaactgcag          cttgaggcttg          gtagaggctggggccgctg
accioncccag
1081 gtggtaggctgg          aactttctaca          acaattacat          ttctgcacc          aacaaggcaca
tacctctcaca
1141 aqgtggccgac          cactggtgtac          acatcaagggg          ggtggccagga          gcaggagctg
tggttcttttgg
1201 tgggtgctttt          gattggtgttgc          caagggtccctc          tgagggctgtg          gaggacgtctc
tcggttatcc
1261 aagcactgtgtgac        ctcgaggctgac        tcaggagggaa          cttggagccggg          ggcaggatcagcgcctgactg
aggcccgact
1321 ggttgcaacagc         ctcgacggcgg          tgtcgcaggtg          cttccgaggct          gctgggaacagcgcctgactg
acgcacacagcagcag
1381 ttcogcagggg          gagcccctcccc          cggctgagctg          gctggggctgg          tctgcaggggctgggcctgg
cccattacggg
1441 ctactctctt gggctttcaca gctctccatcg ccactggggg ctcctgctgg cctccctcgc
1501 tccctgtgtgt ctcctcttgtg aaacctggga gaccagagtc cccctttaggg
1561 tccgggaaga cccgcccacatc ccaggactcc agatgcaggg agcctctgtg
1621 cccacatgca aggaccagca tctcttgaga ggacgccttg gcttaccttg gggcagagat
1681 gcttggggag cagttcagag acacacacag taggccccgca ataaaaagca cacccttt

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>NADPH oxidase organizer 1</td>
<td>NOXO1</td>
<td>NM_172167</td>
<td>41</td>
</tr>
</tbody>
</table>

1 agccatggca ggccccccagt acccagtttc agtgcaaggg gcagcctcgg tgtgacatca
15 61 gaggcctccaa acggttgcct tctcctgtgg ctctgtcagc gcgcagcgcaca cctccctgtc
121 cagaggttgg gacgaatttc ggccccccagt acacccttac aagagacctt tvtccctcgg
181 ggcggggccgt ctgcggagag tgcagcctgc gcggccagcc ccacgcagctt tctccctgtc
20 caccacgttt 241 gggacgccttg gggccacagc gcgccgctgt gcgcagcttg cagcttgtgg
301 ctcggagcctg ctgcggacgct cagagcgcgt gcgcagccttc cgcagctgtc ctctccatcc
acagctcggga
25 361 gcgcaccgccaa cccctggacc ccagctgttc ccgatgcgtt gcgcagccttc cttccctgtc
tgatcctgcc
421 cacacccagag gacgcacgttc tttctgcggc gcgcagccttc cttccctgtc
acacccagtc
481 gctgctggcc acacgccttc tgcagccttt cgcagccttc gacagccgag ggcgcggttt
541 tccagggcag gctggagcact gcgcagcttt gcgcagccttc tcgcagcttt ggcgcggttt
561 tgggggagag gcgcaggttgc gcgcagccttc ggcgcggttt ggcgcggttt
661 gcgcgcggca ggcgctcgggg gagcgccgtgc cctggccggc cccctgggcc gcgcgcggca
721 tccgctcggc taaccccagag gcgcgagctgc tggggcggtgc tgcggccggc gcgcgcggca
781 gcgggcggtgtg gaaacctcag aacccggttc ggcgttcagc agtgacagtc gcgcgcggca
acccggttc
841 cctacctccc ggctggctgc gcggtgccgga aacccggttc ggcgttcagc gcgcgcggca
901 gtcgctcggc gcggtgccgga aacccggttc ggcgttcagc gcgcgcggca gcgcgcggca
cctccctgcgg
961 cacgcgcctccc tcgccacccag actctgcgcgc gcgccgcacc gcgccgcacc gcgccgcacc
1021 ctgcaccgttc acacgcaggg cccctgcgcgc gcgccgcacc gcgccgcacc gcgccgcacc
1081 ctctgttcgcc accccagcgc ggacgcagtc gcgcgcggca gcgcgcggca gcgcgcggca
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>interferon, alpha-inducible protein 27</td>
<td>IFI27</td>
<td>NM_005532</td>
<td>42</td>
</tr>
</tbody>
</table>

1 gggaacacat ccaagcttaa gacggtgagg tcagcttcac atttcagga
actctccttc
5 61 tttgggtctg gctgaagttg aggatctctt actctctagg ccacggaatt
aacccgagca
121 ggcatggagg cctctgctct cacctcatca gcagtgacca gtgtggccaa
agtgtcaggg
10 181 gtgccccttc gctctgccgt agtctttgcc ctggcagagg ttgctacagt
tgatttggga
241 ggagtgtgg agtgcccat ggtctcaggt gccatgggtctc actgcegpc
gggaatcgc
301 tcgtctccca tagcagccaa gatgatgctc cgggggcagc ttgcaatgg
15 ggctggactt
361 gcctgggcca gctctgctcc tacctgtcag tcactgggag caactggact
ctcggactt
421 accaagtctt ctcctggcct cattggtct gccatgggc agtctattgc
gagctttac
20 481 tagttcctct gcctgctgcc ctgcagagga agaaccatgc caggggagaa
ggccaccgc
541 catctggacc cagcgaggag ccactatcc caaatatacc tggggtgaaa
tatccaaat
601 tcctgcatttc cagagaaaaa taagaataaa agatgaattt ttgcaactct ctaaa

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>cytochrome P450, family 3, subfamily A, polyepitope 43</td>
<td>CYP3A43</td>
<td>NM_057095</td>
<td>43</td>
</tr>
</tbody>
</table>

1 acctctgggc agagaaaaaa agctctatat gcacagccca gcaagagca
gcacagccct
61 gaaagaaaaa ctcagagac agagctgaaa aagaaaaactg gtgtgggcAc
tcctccaaa
121 ctttgccatg gaaacatggg ttttggtgc taccagctgc tacctctctc
atatttatttg
181 gaccacttca cataaacttt ttaagaagct gggaattctt gggccaaccc
tctgctcctt
35 241 tctgggaact aatgttgctct accttagggg tcttgggaat ttggcacagag
aatgaatga
301 aaataacgga gaaatgtgag ggtgtatag ggggaacacag cccatgtcgg
tcatcaggg
361 tcctccgatag atcaaaaagag ttttaggaa agaatgttac ttggttcttca
caaaccagat
2041 attatattga gaaagtcaac aagcacctct ttacaaact gttatctgat
gttctctgct
2101 atattaagga tgaatctaca gaattagatc aataaggatc aacaataaa
tatatttgggt
2161 catt

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Gene Symbol</th>
<th>GenBank Accession #</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>plakophilin 2</td>
<td>PKP2</td>
<td>NM_004572</td>
<td>44</td>
</tr>
</tbody>
</table>

1 gtggccggctt cgccccgcaag tccagaggca ggccagcgcact tcgggtgcac ccacccgcct
22 ggccagccgc cagctgagta cggctcagct cggaccgcct
tggccagca
121 gatcctggga caactgagca gttccagcct ggcgcgtgcac tcgcagcagc
301 cagacaggtt tagtctttat gttcaacacc acctattttcgtg
cagaggttcc
15 ccacttatga
tttgcgttttt gcaaccagcag cagacagcttg ccagttgtttt cagaggttcc
100 cttggagcat ctctctggag gactggagat ttcttctgttc acagccctac
ggaggttcct
24 ggtcggctcg gcaagagggag cgcacgctggag caggtggctgag ctggagagag
cgctgagtttg cagcagtggtc
ttggtgctg
cggccttgac

1321 aaacttagta tttgaagaca atgacaacaa attggaggtg gctgaactaa
atggggtacc
1381 tcggctgctc caggtgctga agcaaaccag agacttggag...gaaaacaaat agaacataat tttatgagtc
ttccagaaga
2881 cctttgcaag tttgccacca gtagataccg gccacaggct cgacaaatag
tggtctttgt
2941 tattagggct tatggtacat ggttctctgg aatcaaaatg tgaattcatg
tgggaaggac
3001 attaatccaa taataaagga aagaagcgtg tgcattaactg ggattttaaa
agttttagtt
3061 acatttatat tccttttcctg gtctcccatgt tttgtcactc atgtgcacat
tgcttcgccca
3121 ttgggctccc agtgtatgtt tctgcagtgt tgaacagaa tgggaatgac
agaatatатc
3181 tgcaattttc caggagaaag tataatggca aatatttgg tttctttctt
3241 ttctttttat cccctttggtt gttttttcctc tgattttttaa ataaacttaa
gaaattaga
3301 ttcatcagatg tgcatactgt taagaaaaag aatattgagag gaagtgatca
tagcaaatta
3361 aagaagcttt tccttccccag aaccttaagtt aaaaataaaa aataaatat
aatattaatc
3421 ttttccacag aagaaagccaa ctgatgatgt aaaaaatttc atttcctgtc
3481 aatgagattt ttctcaggag atacctttacc tataaacagc cctttaaatc
caaatctctt
3541 cttaaagcgatg gcattctatg taatgccttt cctggacatttt tttggccact
3601 aagtgaagaa tggactcttt cttaaatcag aagagcttc aaggtccctct
tcagactgc
3661 ctgccccagcc tgggctcttg aaaaatccgg gcacgatcatt ttgtaaattt
tcagccagc
3721 ccagcttctgc caccacttta agaaatatca cagacccact agatctcata
tgattttctt
3781 caagccatta ttttaactca agaaaaactct agagaagaa aatgaaagaag
tcatgttgaa
3841 gaagatgatgta gaatgtgtca agaccatcca gaaatgata gagaatact
gatatatttt
3901 atggtgtgca taatccagcg aataaatcag acatatatcg aaaaactaac
tgcgtctaga
3961 tttaaaacct tctatcag ttagctctta caactactatct ctgttttaac
ttttctctgt
4021 tctgacacaac ttttgccttaa gtaaataaag caaaaaatt cttcaactcc
tttggcaag
4081 aaaaactgtaa cagaaaaata attttgaatg tgtacttaag tcttttatatt
attggaacag
4141 attttttttc aatattttaa gctgaatgaa gacaacttag gtgtgacacc
tagttttaaaa
4201 tgaatatatt tagataacca aaaaaattt actggagaga atttatatat
ctttctccag
4261 agttctgtag ataaagcattt ggagtgcatt tattctctcc aataaaataat
gttggtccag
4321 aacctttttc gttttaaag gcattaataa agccttccag ataatataaat 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, NPDCl, 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, DPEPl, 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, FITN3, EPHA2, FGFBP1, ABCC3, TPMT, IL18, and DPEPl.

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 DPEPl.

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, NPDCl, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNT12, TPK1, DEFA6, EPLIN, CLIC5, PERP, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBPl, PKT6, EVA1, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMPRSS2, GDA, MIST1R, ITGB4, ANXA3, CCL15, DPEPl, 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, FGFBPl, ABCC3, TPMT, IL18, and DPEPl.

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, FGFBPl, ABCC3, TPMT, IL18, and DPEPl.

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, NOXOL, TM4SF4, PTPN3, EPHA2, FGFBPI, ABCC3, TPMT, IL18, or DPEPl 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, FGFBPI, ABCC3, TPMT, IL1 8, or DPEPI 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 R09483 11023.
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, FGFBPI, ABCC3, TPMT, IL18, or DPEPI 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 RS247682266, R9866 135 153, R 1078 103 114, or R47 1278 1606.

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, FGFBPI, ABCC3, TPMT, IL18, or DPEPI 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, NOXI, TM4SF4, PTPN3, EPHA2, FGFBPI, ABCC3, TPMT, IL18, and DPEPl.

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 µM.

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, NPDCl, CTEN, RAB25, HEPH, TPMT, PKP3, GALNT5, CALML4, GALNTI 2, TPK1, DEFA6, EPLIN, CLIC5, PERP1, SYK, SLC12A2, GUCY2C, TM4SF4, TGFA, FGFBPI, PTK6, EVAI, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPDL3B, TMPRSS2, GDA, MSTIR, ITGB4, ANXA3, CCL15, DPEPl, NOXI, IF127, 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, NOXOl, TM4SF4, PTPN3, EPHA2, FGFBPl, ABCC3, TPMT, IL 18, and DPEPl.

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, NOXOl, TM4SF4, PTPN3, EPHA2, FGFBPl, ABCC3, TPMT, IL18, and DPEPl.

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, NOXOl, TM4SF4, PTPN3, EPHA2, FGFBPl, ABCC3, TPMT, IL18, and DPEP.

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 \( \text{DEF} \), \( \text{ITGB4} \), \( \text{TM4SF} \), \( \text{SYK} \), \( \text{PPAP2C} \), \( \text{RAB25} \), \( \text{HEPH} \), \( \text{NOXO1} \), \( \text{TM4SF} \), \( \text{PTPN3} \), \( \text{EPHA2} \), \( \text{FGFBPI} \), \( \text{ABCC3} \), \( \text{TPMT} \), \( \text{IL1} \), and \( \text{DPEP} \); 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, FGFBPI, PTK6, EVAI, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPD3B, TMRPSS2, GDA, MSTIR, ITGB4, ANXA3, CCL15, DPEPL, NOXO1, EF127, 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, FGFBPI, PTK6, EVAI, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMPD3B, TMRPSS2, GDA, MSTIR, ITGB4, ANXA3, CCL15, DPEPL, NOXO1, EF127, 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, FGFBI, PTK5, EVAI, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMDL3B, TMPRSS2, GDA, MSTIR, ITGB4, ANXA3, CCL15, DPEPI, NOXOI, 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, NOXOI, TM4SF4, PTPN3, EPHA2, FGFBI, ABCC3, TPMT, IL18, and DPEPI.

98. The method of claim 94, comprising determining 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, FGFBI, PTK5, EVAI, EPHA2, ITGA6, TNFRSF21, TM4SF3, IL18, BMP4, SMDL3B, TMPRSS2, GDA, MSTIR, ITGB4, ANXA3, CCL15, DPEPI, NOXOI, 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, NOX0I, TM4SF4, PTPN3, EPHA2, FGFBPI, ABCC3, TPMT, DL18, 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

Biomarker Discovery

Response phenotype (ex-vivo resistance assay)

Microarray profiling

TaqMan confirmation

Select potential candidate serum and tumor markers predictive of PCI-24781 resistance

Translation

Identify clinical opportunity

Determine frequency of marker in the targeted clinical population

Evaluate benefit to clinical development

Integrate stratification markers into clinical plan
Fig. 2

![Graph showing percent inhibition and PCI-24781 concentration (μM) for sensitive and resistant samples.](image)
Fig. 3

Normalization procedures

Quality Assessment using MA Plots & Boxplots

Median intensity (global adjustment)

looking for differentials

1) T-test/ANOVA
   MAANOVA (Wu & Churchill)
   SAM gene-specific variation
   (Tusher et al.)
2) FDR (false discovery rate) correction; q-values
   (Storey & Tibshirani)

Bayesian ANOVA
   (Ishwaran & Rao)

GO/Pathway Mapping + Statistical Significance

explore samples (different treatments, source, etc.)

Use clustering
   (1) Hierarchical (Bioconductor)
   (2) Fuzzy c-means (Matlab)

Statistically Significant Biological Conclusions
Fig. 5

[Graph showing gene expression ratios with labels for each gene and two types of ratios represented by different patterns.]
Fig. 6
Fig. 7

A

Band intensity

Vehicle 0.5 hr 1 hr 2 hr 3 hr 8 hr

B

Fold Change from t=0

Time (hours) 0 0.5 1 2 3 8
Fig. 8

FGF 15

![Graph showing normalized values over time for FGF 15 with Taqman, Immunoblot, and Microarray data points.]

Synaptogyrin 2

![Graph showing normalized values over time for Synaptogyrin 2 with Taqman, Immunoblot, and Microarray data points.]

A. CLASSIFICATION OF SUBJECT MATTER

A61K 49/00(2006.01)i, C12Q 1/68(2006.01)i, GOIN 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, GOIN 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

<table>
<thead>
<tr>
<th>Category*</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>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</td>
<td>101-103</td>
</tr>
</tbody>
</table>

☐ 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
  "&" 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, Seogu, Daejeon 302-701, Republic of Korea
Facsimile No 82-42-472-7140

Authorized officer
PARK, Yeong Gwan
Telephone No 82-42-481-8407
### 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 nuclease 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 64(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.
<table>
<thead>
<tr>
<th>Patent document cited in search report</th>
<th>Publication date</th>
<th>Patent family member(s)</th>
<th>Publication date</th>
</tr>
</thead>
<tbody>
<tr>
<td>US 20030166026 A1</td>
<td>04. 09. 2003</td>
<td>(NONE)</td>
<td></td>
</tr>
</tbody>
</table>