



(86) Date de dépôt PCT/PCT Filing Date: 2010/11/19
(87) Date publication PCT/PCT Publication Date: 2011/05/26
(45) Date de délivrance/Issue Date: 2019/07/02
(85) Entrée phase nationale/National Entry: 2012/04/04
(86) N° demande PCT/PCT Application No.: US 2010/057490
(87) N° publication PCT/PCT Publication No.: 2011/063274
(30) Priorité/Priority: 2009/11/23 (US61/263,763)

(51) Cl.Int./Int.Cl. *C12Q 1/6809* (2018.01),
G01N 33/48 (2006.01), *G16B 25/10* (2019.01)

(72) Inventeurs/Inventors:
BAKER, JOFFRE B., US;
CRONIN, MAUREEN T., US;
COLLIN, FRANCOIS, US;
LIU, MEI-LAN, US

(73) Propriétaire/Owner:
GENOMIC HEALTH, INC., US

(74) Agent: SMART & BIGGAR

(54) Titre : PROCÉDES DESTINÉS À PRÉDIRE L'ISSUE CLINIQUE D'UN CANCER
(54) Title: METHODS TO PREDICT CLINICAL OUTCOME OF CANCER

(57) **Abrégé/Abstract:**

The present invention provides methods to determine the prognosis and appropriate treatment for patients diagnosed with cancer, based on the expression levels of one or more biomarkers. More particularly, the invention relates to the identification of genes, or sets of genes, able to distinguish breast cancer patients with a good clinical prognosis from those with a bad clinical prognosis. The invention further provides methods for providing a personalized genomics report for a cancer patient. The inventions also relates to computer systems and software for data analysis using the prognostic and statistical methods disclosed herein.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 May 2011 (26.05.2011)



(10) International Publication Number
WO 2011/063274 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2010/057490

(22) International Filing Date:
19 November 2010 (19.11.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/263,763 23 November 2009 (23.11.2009) US

(71) Applicant (for all designated States except US): **GENOMIC HEALTH, INC.** [US/US]; 301 Penobscot Drive, Redwood City, CA 94063 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **BAKER, Joffre B.** [US/US]; 1400 Avery Street, Montara, CA 94037 (US). **CRONIN, Maureen T.** [US/US]; 771 Anderson Drive, Los Altos, CA 94024 (US). **COLLIN, Francois** [US/US]; 2166 N. Valley, Berkeley, CA 94702 (US). **LIU, Mei-Lan** [US/US]; 760 Pico Avenue, San Mateo, CA 94403 (US).

(74) Agent: **BORDEN, Paula A.**; Bozicevic, Field & Francis LLP, 1900 University Avenue, Suite 200, East Palo Alto, CA 94303 (US).

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

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

Published:

- without international search report and to be republished upon receipt of that report (Rule 48.2(g))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: METHODS TO PREDICT CLINICAL OUTCOME OF CANCER

(57) Abstract: The present invention provides methods to determine the prognosis and appropriate treatment for patients diagnosed with cancer, based on the expression levels of one or more biomarkers. More particularly, the invention relates to the identification of genes, or sets of genes, able to distinguish breast cancer patients with a good clinical prognosis from those with a bad clinical prognosis. The invention further provides methods for providing a personalized genomics report for a cancer patient. The inventions also relates to computer systems and software for data analysis using the prognostic and statistical methods disclosed herein.



WO 2011/063274 A2

METHODS TO PREDICT CLINICAL OUTCOME OF CANCER

INTRODUCTION

[0001] Oncologists have a number of treatment options available to them, including different combinations of therapeutic regimens that are characterized as "standard of care." The absolute benefit from adjuvant treatment is larger for patients with poor prognostic features, and this has resulted in the policy to select only these so-called 'high-risk' patients for adjuvant chemotherapy. See, e.g., S. Paik, et al., J Clin Oncol. 24(23):3726-34 (2006). Therefore, the best likelihood of good treatment outcome requires that patients be assigned to optimal available cancer treatment, and that this assignment be made as quickly as possible following diagnosis.

[0002] Today our healthcare system is riddled with inefficiency and wasteful spending – one example of this is that the efficacy rate of many oncology therapeutics working only about 25% of the time. Many of those cancer patients are experiencing toxic side effects for costly therapies that may not be working. This imbalance between high treatment costs and low therapeutic efficacy is often a result of treating a specific diagnosis one way across a diverse patient population. But with the advent of gene profiling tools, genomic testing, and advanced diagnostics, this is beginning to change.

[0003] In particular, once a patient is diagnosed with breast cancer there is a strong need for methods that allow the physician to predict the expected course of disease, including the likelihood of cancer recurrence, long-term survival of the patient, and the like, and select the most appropriate treatment option accordingly. Accepted prognostic and predictive factors in breast cancer include age, tumor size, axillary lymph node status, histological tumor type, pathological grade and hormone receptor status. Molecular diagnostics, however, have been demonstrated to identify more patients with a low risk of breast cancer than was possible with standard prognostic indicators. S. Paik, The Oncologist 12(6):631-635 (2007).

[0004] Despite recent advances, the challenge of breast cancer treatment remains to target specific treatment regimens to pathogenically distinct tumor types, and ultimately personalize tumor treatment in order to maximize outcome. Accurate prediction of prognosis and clinical outcome would allow the oncologist to tailor the administration of adjuvant chemotherapy such that women with a higher risk of a recurrence or poor prognosis would receive more aggressive treatment. Furthermore, accurately stratifying patients based on risk would greatly advance the understanding of

expected absolute benefit from treatment, thereby increasing success rates for clinical trials for new breast cancer therapies.

[0005] Currently, most diagnostic tests used in clinical practice are frequently not quantitative, relying on immunohistochemistry (IHC). This method often yields different results in different laboratories, in part because the reagents are not standardized, and in part because the interpretations are subjective and cannot be easily quantified. Other RNA-based molecular diagnostics require fresh-frozen tissues, which presents a myriad of challenges including incompatibilities with current clinical practices and sample transport regulations. Fixed paraffin-embedded tissue is more readily available and methods have been established to detect RNA in fixed tissue. However, these methods typically do not allow for the study of large numbers of genes (DNA or RNA) from small amounts of material. Thus, traditionally fixed tissue has been rarely used other than for IHC detection of proteins.

SUMMARY

[0006] The present disclosure provides a set of genes, the expression levels of which are associated with a particular clinical outcome in cancer. For example, the clinical outcome could be a good or bad prognosis assuming the patient receives the standard of care. The clinical outcome may be defined by clinical endpoints, such as disease or recurrence free survival, metastasis free survival, overall survival, etc.

[0007] The present disclosure accommodates the use of archived paraffin-embedded biopsy material for assay of all markers in the set, and therefore is compatible with the most widely available type of biopsy material. It is also compatible with several different methods of tumor tissue harvest, for example, via core biopsy or fine needle aspiration. The tissue sample may comprise cancer cells.

[0008] In one aspect, the present disclosure concerns a method of predicting a clinical outcome of a cancer patient, comprising (a) obtaining an expression level of an expression product (e.g., an RNA transcript) of at least one prognostic gene listed in Tables 1-12 from a tissue sample obtained from a tumor of the patient; (b) normalizing the expression level of the expression product of the at least one prognostic gene, to obtain a normalized expression level; and (c) calculating a risk score based on the normalized expression value, wherein increased expression of prognostic genes in Tables 1, 3, 5, and 7 are positively correlated with good prognosis, and wherein increased expression of prognostic genes in Tables 2, 4, 6, and 8 are negatively associated with good prognosis. In some

embodiments, the tumor is estrogen receptor-positive. In other embodiments, the tumor is estrogen receptor negative.

[0009] In one aspect, the present disclosure provides a method of predicting a clinical outcome of a cancer patient, comprising (a) obtaining an expression level of an expression product (e.g., an RNA transcript) of at least one prognostic gene from a tissue sample obtained from a tumor of the patient, where the at least one prognostic gene is selected from GSTM2, IL6ST, GSTM3, C8orf4, TNFRSF11B, NAT1, RUNX1, CSF1, ACTR2, LMNB1, TFRC, LAPTM4B, ENO1, CDC20, and IDH2; (b) normalizing the expression level of the expression product of the at least one prognostic gene, to obtain a normalized expression level; and (c) calculating a risk score based on the normalized expression value, wherein increased expression of a prognostic gene selected from GSTM2, IL6ST, GSTM3, C8orf4, TNFRSF11B, NAT1, RUNX1, and CSF1 is positively correlated with good prognosis, and wherein increased expression of a prognostic gene selected from ACTR2, LMNB1, TFRC, LAPTM4B, ENO1, CDC20, and IDH2 is negatively associated with good prognosis. In some embodiments, the tumor is estrogen receptor-positive. In other embodiments, the tumor is estrogen receptor negative.

[0010] In various embodiments, the normalized expression level of at least 2, or at least 5, or at least 10, or at least 15, or at least 20, or at least 25 prognostic genes (as determined by assaying a level of an expression product of the gene) is determined. In alternative embodiments, the normalized expression levels of at least one of the genes that co-expresses with prognostic genes in Tables 16-18 is obtained.

[0011] In another embodiment, the risk score is determined using normalized expression levels of at least one a stromal or transferrin receptor group gene, or a gene that co-expresses with a stromal or transferrin receptor group gene.

[0012] In another embodiment, the cancer is breast cancer. In another embodiment, the patient is a human patient.

[0013] In yet another embodiment, the cancer is ER-positive breast cancer.

[0014] In yet another embodiment, the cancer is ER-negative breast cancer.

[0015] In a further embodiment, the expression product comprises RNA. For example, the RNA could be exonic RNA, intronic RNA, or short RNA (e.g., microRNA, siRNA, promoter-associated small RNA, shRNA, etc.). In various embodiments, the RNA is fragmented RNA.

[0016] In a different aspect, the disclosure concerns an array comprising polynucleotides hybridizing to an RNA transcription of at least one of the prognostic genes listed in Tables 1-12.

[0017] In a still further aspect, the disclosure concerns a method of preparing a personalized genomics profile for a patient, comprising (a) obtaining an expression level of an expression product (e.g., an RNA transcript) of at least one prognostic gene listed in Tables 1-12 from a tissue sample obtained from a tumor of the patient; (b) normalizing the expression level of the expression product of the at least one prognostic gene to obtain a normalized expression level; and (c) calculating a risk score based on the normalized expression value, wherein increased expression of prognostic genes in Tables 1, 3, 5, and 7 are positively correlated with good prognosis, and wherein increased expression of prognostic genes in Tables 2, 4, 6, and 8 are negatively associated with good prognosis. In some embodiments, the tumor is estrogen receptor-positive, and in other embodiments the tumor is estrogen receptor negative.

[0018] In various embodiments, a subject method can further include providing a report. The report may include prediction of the likelihood of risk that said patient will have a particular clinical outcome.

[0019] The disclosure further provides a computer-implemented method for classifying a cancer patient based on risk of cancer recurrence, comprising (a) classifying, on a computer, said patient as having a good prognosis or a poor prognosis based on an expression profile comprising measurements of expression levels of expression products of a plurality of prognostic genes in a tumor tissue sample taken from the patient, said plurality of genes comprising at least three different prognostic genes listed in any of Tables 1-12, wherein a good prognosis predicts no recurrence or metastasis within a predetermined period after initial diagnosis, and wherein a poor prognosis predicts recurrence or metastasis within said predetermined period after initial diagnosis; and (b) calculating a risk score based on said expression levels.

[0020] Embodiments of the claimed invention pertain to a method for determining likelihood of good prognosis of a human patient diagnosed with a breast cancer comprising: (a) quantitatively measuring a level of an RNA transcript of IL6ST in a tissue sample obtained from a breast cancer tumor of the patient; (b) normalizing the level of the RNA transcript of IL6ST to obtain a normalized IL6ST expression level; (c) comparing the normalized IL6ST expression level to normalized IL6ST expression level data obtained from a breast cancer reference set; and (d) determining a likelihood of good prognosis for the patient based on the comparison of the normalized IL6ST expression level to the IL6ST expression level data obtained from the breast cancer reference set, wherein increases in normalized IL6ST expression level positively correlate with increased likelihood of good prognosis,

wherein the good prognosis is a reduced likelihood of recurrence or metastasis or an increased likelihood of overall survival.

[0020a] Embodiments of the claimed invention also pertain to a method for predicting likelihood of good prognosis of a human patient diagnosed with breast cancer comprising: (a) extracting RNA from a fixed, wax-embedded breast cancer tissue sample obtained from the patient; (b) reverse transcribing an RNA transcript of IL6ST to produce a cDNA of IL6ST; (c) amplifying the cDNA of IL6ST to produce an amplicon of the RNA transcript of IL6ST; (d) quantitatively assaying a level of the amplicon of the RNA transcript of IL6ST; (e) normalizing the level of the amplicon of the RNA transcript of IL6ST to provide a normalized IL6ST amplicon level; (f) comparing the normalized IL6ST amplicon level to normalized IL6ST amplicon level data obtained from a breast cancer reference set; and (g) determining a likelihood of good prognosis for the patient based on the comparison of the normalized IL6ST amplicon level to the IL6ST amplicon expression level data obtained from the breast cancer reference set, wherein increases in normalized IL6ST amplicon level positively correlate with increased likelihood of good prognosis, wherein the good prognosis is a reduced likelihood of recurrence or metastasis or an increased likelihood of overall survival.

DETAILED DESCRIPTION

DEFINITIONS

[0021] Unless defined otherwise, technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Singleton et al., Dictionary of Microbiology and Molecular Biology 2nd ed., J. Wiley & Sons (New York, NY 1994), and March, Advanced Organic Chemistry

Reactions, Mechanisms and Structure 4th ed., John Wiley & Sons (New York, NY 1992), provide one skilled in the art with a general guide to many of the terms used in the present application.

[0022] One skilled in the art will recognize many methods and materials similar or equivalent to those described herein, which could be used in the practice of the present invention. Indeed, the present invention is in no way limited to the methods and materials described. For purposes of the present invention, the following terms are defined below.

[0023] “Prognostic factors” are those variables related to the natural history of cancer, which influence the recurrence rates and outcome of patients once they have developed cancer. Clinical parameters that have been associated with a worse prognosis include, for example, lymph node involvement, and high grade tumors. Prognostic factors are frequently used to categorize patients into subgroups with different baseline relapse risks.

[0024] The term “prognosis” is used herein to refer to the prediction of the likelihood of cancer-attributable death or progression, including recurrence, metastatic spread, and drug resistance, of a neoplastic disease, such as breast cancer. The term “good prognosis” means a desired or “positive” clinical outcome. For example, in the context of breast cancer, a good prognosis may be an expectation of no recurrences or metastasis within two, three, four, five or more years of the initial diagnosis of breast cancer. The terms “bad prognosis” or “poor prognosis” are used herein interchangeably herein to mean an undesired clinical outcome. For example, in the context of breast cancer, a bad prognosis may be an expectation of a recurrence or metastasis within two, three, four, five or more years of the initial diagnosis of breast cancer.

[0025] The term “prognostic gene” is used herein to refer to a gene, the expression of which is correlated, positively or negatively, with a good prognosis for a cancer patient treated with the standard of care. A gene may be both a prognostic and predictive gene, depending on the correlation of the gene expression level with the corresponding endpoint. For example, using a Cox proportional hazards model, if a gene is only prognostic, its hazard ratio (HR) does not change when measured in patients treated with the standard of care or in patients treated with a new intervention.

[0026] The term “predictive gene” is used herein to refer to a gene, the expression of which is correlated, positively or negatively, with response to a beneficial response to treatment. For example, treatment could include chemotherapy.

[0027] The terms “risk score” or “risk classification” are used interchangeably herein to describe a level of risk (or likelihood) that a patient will experience a particular clinical

outcome. A patient may be classified into a risk group or classified at a level of risk based on the methods of the present disclosure, e.g. high, medium, or low risk. A “risk group” is a group of subjects or individuals with a similar level of risk for a particular clinical outcome.

[0028] A clinical outcome can be defined using different endpoints. The term “long-term” survival is used herein to refer to survival for a particular time period, e.g., for at least 3 years, more preferably for at least 5 years. The term “Recurrence-Free Survival” (RFS) is used herein to refer to survival for a time period (usually in years) from randomization to first cancer recurrence or death due to recurrence of cancer. The term “Overall Survival” (OS) is used herein to refer to the time (in years) from randomization to death from any cause. The term “Disease-Free Survival” (DFS) is used herein to refer to survival for a time period (usually in years) from randomization to first cancer recurrence or death from any cause.

[0029] The calculation of the measures listed above in practice may vary from study to study depending on the definition of events to be either censored or not considered.

[0030] The term “biomarker” as used herein refers to a gene, the expression level of which, as measured using a gene product.

[0031] The term “microarray” refers to an ordered arrangement of hybridizable array elements, preferably polynucleotide probes, on a substrate.

[0032] As used herein, the term “normalized expression level” as applied to a gene refers to the normalized level of a gene product, e.g. the normalized value determined for the RNA expression level of a gene or for the polypeptide expression level of a gene.

[0033] The term “C_t” as used herein refers to threshold cycle, the cycle number in quantitative polymerase chain reaction (qPCR) at which the fluorescence generated within a reaction well exceeds the defined threshold, i.e. the point during the reaction at which a sufficient number of amplicons have accumulated to meet the defined threshold.

[0034] The term “gene product” or “expression product” are used herein to refer to the RNA transcription products (transcripts) of the gene, including mRNA, and the polypeptide translation products of such RNA transcripts. A gene product can be, for example, an unspliced RNA, an mRNA, a splice variant mRNA, a microRNA, a fragmented RNA, a polypeptide, a post-translationally modified polypeptide, a splice variant polypeptide, etc.

[0035] The term “RNA transcript” as used herein refers to the RNA transcription products of a gene, including, for example, mRNA, an unspliced RNA, a splice variant mRNA, a microRNA, and a fragmented RNA. “Fragmented RNA” as used herein refers to

RNA a mixture of intact RNA and RNA that has been degraded as a result of the sample processing (e.g., fixation, slicing tissue blocks, etc.).

[0036] Unless indicated otherwise, each gene name used herein corresponds to the Official Symbol assigned to the gene and provided by Entrez Gene (URL: www.ncbi.nlm.nih.gov/sites/entrez) as of the filing date of this application.

[0037] The terms “correlated” and “associated” are used interchangeably herein to refer to a strength of association between two measurements (or measured entities). The disclosure provides genes and gene subsets, the expression levels of which are associated with a particular outcome measure. For example, the increased expression level of a gene may be positively correlated (positively associated) with an increased likelihood of good clinical outcome for the patient, such as an increased likelihood of long-term survival without recurrence of the cancer and/or metastasis-free survival. Such a positive correlation may be demonstrated statistically in various ways, e.g. by a low hazard ratio (e.g. $HR < 1.0$). In another example, the increased expression level of a gene may be negatively correlated (negatively associated) with an increased likelihood of good clinical outcome for the patient. In that case, for example, the patient may have a decreased likelihood of long-term survival without recurrence of the cancer and/or cancer metastasis, and the like. Such a negative correlation indicates that the patient likely has a poor prognosis, e.g., a high hazard ratio (e.g., $HR > 1.0$). “Correlated” is also used herein to refer to a strength of association between the expression levels of two different genes, such that expression level of a first gene can be substituted with an expression level of a second gene in a given algorithm in view of their correlation of expression. Such “correlated expression” of two genes that are substitutable in an algorithm usually gene expression levels that are positively correlated with one another, e.g., if increased expression of a first gene is positively correlated with an outcome (e.g., increased likelihood of good clinical outcome), then the second gene that is co-expressed and exhibits correlated expression with the first gene is also positively correlated with the same outcome

[0038] The term “recurrence,” as used herein, refers to local or distant (metastasis) recurrence of cancer. For example, breast cancer can come back as a local recurrence (in the treated breast or near the tumor surgical site) or as a distant recurrence in the body. The most common sites of breast cancer recurrence include the lymph nodes, bones, liver, or lungs.

[0039] The term “polynucleotide,” when used in singular or plural, generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or

DNA or modified RNA or DNA. Thus, for instance, polynucleotides 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 may be single-stranded or, more typically, double-stranded or include single- and double-stranded regions. In addition, the term “polynucleotide” as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may 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 “polynucleotide” 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 “polynucleotides” as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritiated bases, are included within the term “polynucleotides” as defined herein. In general, the term “polynucleotide” 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.

[0040] 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 can be made by a variety of other methods, including *in vitro* recombinant DNA-mediated techniques and by expression of DNAs in cells and organisms.

[0041] The phrase “amplification” refers to a process by which multiple copies of a gene or RNA transcript are formed in a particular sample or cell line. The duplicated region (a stretch of amplified polynucleotide) is often referred to as “amplicon.” Usually, the amount of the messenger RNA (mRNA) produced, *i.e.*, the level of gene expression, also increases in the proportion of the number of copies made of the particular gene expressed.

[0042] The term “estrogen receptor (ER)” designates the estrogen receptor status of a cancer patient. A tumor is ER-positive if there is a significant number of estrogen receptors present in the cancer cells, while ER-negative indicates that the cells do not have a

significant number of receptors present. The definition of “significant” varies amongst testing sites and methods (e.g., immunohistochemistry, PCR). The ER status of a cancer patient can be evaluated by various known means. For example, the ER level of breast cancer is determined by measuring an expression level of a gene encoding the estrogen receptor in a breast tumor sample obtained from a patient.

[0043] The term "tumor," as used herein, refers to all neoplastic cell growth and proliferation, whether malignant or benign, and all pre-cancerous and cancerous cells and tissues.

[0044] The terms "cancer" and "cancerous" refer to or describe the physiological condition in mammals that is typically characterized by unregulated cell growth. Examples of cancer include, but are not limited to, breast cancer, ovarian cancer, colon cancer, lung cancer, prostate cancer, hepatocellular cancer, gastric cancer, pancreatic cancer, cervical cancer, liver cancer, bladder cancer, cancer of the urinary tract, thyroid cancer, renal cancer, carcinoma, melanoma, and brain cancer.

[0045] The gene subset identified herein as the “stromal group” includes genes that are synthesized predominantly by stromal cells and are involved in stromal response and genes that co-express with stromal group genes. “Stromal cells” are defined herein as connective tissue cells that make up the support structure of biological tissues. Stromal cells include fibroblasts, immune cells, pericytes, endothelial cells, and inflammatory cells. “Stromal response” refers to a desmoplastic response of the host tissues at the site of a primary tumor or invasion. See, e.g., E. Rubin, J. Farber, Pathology, 985-986 (2nd Ed. 1994). The stromal group includes, for example, CDH11, TAGLN, ITGA4, INHBA, COLIA1, COLIA2, FN1, CXCL14, TNFRSF1, CXCL12, C10ORF116, RUNX1, GSTM2, TGFB3, CAV1, DLC1, TNFRSF10, F3, and DICER1, and co-expressed genes identified in Tables 16-18.

[0046] The gene subset identified herein as the “metabolic group” includes genes that are associated with cellular metabolism, including genes associated with carrying proteins for transferring iron, the cellular iron homeostasis pathway, and homeostatic biochemical metabolic pathways, and genes that co-express with metabolic group genes. The metabolic group includes, for example, TFRC, ENO1, IDH2, ARF1, CLDN4, PRDX1, and GBP1, and co-expressed genes identified in Tables 16-18.

[0047] The gene subset identified herein as the “immune group” includes genes that are involved in cellular immunoregulatory functions, such as T and B cell trafficking, lymphocyte-associated or lymphocyte markers, and interferon regulation genes, and genes

that co-express with immune group genes. The immune group includes, for example, CCL19 and IRF1, and co-expressed genes identified in Tables 16-18.

[0048] The gene subset identified herein as the “proliferation group” includes genes that are associated with cellular development and division, cell cycle and mitotic regulation, angiogenesis, cell replication, nuclear transport/stability, wnt signaling, apoptosis, and genes that co-express with proliferation group genes. The proliferation group includes, for example, PGF, SPC25, AURKA, BIRC5, BUB1, CCNB1, CENPA, KPNA, LMNB1, MCM2, MELK, NDC80, TPX2M, and WISP1, and co-expressed genes identified in Tables 16-18.

[0049] The term “co-expressed”, as used herein, refers to a statistical correlation between the expression level of one gene and the expression level of another gene. Pairwise co-expression may be calculated by various methods known in the art, e.g., by calculating Pearson correlation coefficients or Spearman correlation coefficients. Co-expressed gene cliques may also be identified using a graph theory.

[0050] As used herein, the terms “gene clique” and “clique” refer to a subgraph of a graph in which every vertex is connected by an edge to every other vertex of the subgraph.

[0051] As used herein, a “maximal clique” is a clique in which no other vertex can be added and still be a clique.

[0052] The “pathology” of cancer includes all phenomena that compromise the well-being of the patient. This includes, without limitation, abnormal or uncontrollable cell growth, metastasis, interference with the normal functioning of neighboring cells, release of cytokines or other secretory products at abnormal levels, suppression or aggravation of inflammatory or immunological response, neoplasia, premalignancy, malignancy, invasion of surrounding or distant tissues or organs, such as lymph nodes, etc.

[0053] A “computer-based system” refers to a system of hardware, software, and data storage medium used to analyze information. The minimum hardware of a patient computer-based system comprises a central processing unit (CPU), and hardware for data input, data output (e.g., display), and data storage. An ordinarily skilled artisan can readily appreciate that any currently available computer-based systems and/or components thereof are suitable for use in connection with the methods of the present disclosure. The data storage medium may comprise any manufacture comprising a recording of the present information as described above, or a memory access device that can access such a manufacture.

[0054] To “record” data, programming or other information on a computer readable medium refers to a process for storing information, using any such methods as known in the art. Any convenient data storage structure may be chosen, based on the means used to access the stored information. A variety of data processor programs and formats can be used for storage, e.g. word processing text file, database format, etc.

[0055] A “processor” or “computing means” references any hardware and/or software combination that will perform the functions required of it. For example, a suitable processor may be a programmable digital microprocessor such as available in the form of an electronic controller, mainframe, server or personal computer (desktop or portable). Where the processor is programmable, suitable programming can be communicated from a remote location to the processor, or previously saved in a computer program product (such as a portable or fixed computer readable storage medium, whether magnetic, optical or solid state device based). For example, a magnetic medium or optical disk may carry the programming, and can be read by a suitable reader communicating with each processor at its corresponding station.

[0056] As used herein, “graph theory” refers to a field of study in Computer Science and Mathematics in which situations are represented by a diagram containing a set of points and lines connecting some of those points. The diagram is referred to as a “graph”, and the points and lines referred to as “vertices” and “edges” of the graph. In terms of gene co-expression analysis, a gene (or its equivalent identifier, e.g. an array probe) may be represented as a node or vertex in the graph. If the measures of similarity (e.g., correlation coefficient, mutual information, and alternating conditional expectation) between two genes are higher than a significant threshold, the two genes are said to be co-expressed and an edge will be drawn in the graph. When co-expressed edges for all possible gene pairs for a given study have been drawn, all maximal cliques are computed. The resulting maximal clique is defined as a gene clique. A gene clique is a computed co-expressed gene group that meets predefined criteria.

[0057] “Stringency” of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the

higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

[0058] "Stringent conditions" or "high stringency conditions", as defined herein, typically: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) and 50% formamide at 55°C, followed by a high-stringency wash consisting of 0.1 x SSC containing EDTA at 55°C.

[0059] "Moderately stringent conditions" may be identified as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and %SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

[0060] In the context of the present invention, reference to "at least one," "at least two," "at least five," etc. of the genes listed in any particular gene set means any one or any and all combinations of the genes listed.

[0061] The term "node negative" cancer, such as "node negative" breast cancer, is used herein to refer to cancer that has not spread to the lymph nodes.

[0062] The terms “splicing” and “RNA splicing” are used interchangeably and refer to RNA processing that removes introns and joins exons to produce mature mRNA with continuous coding sequence that moves into the cytoplasm of a eukaryotic cell.

[0063] In theory, the term “exon” refers to any segment of an interrupted gene that is represented in the mature RNA product (B. Lewin. *Genes IV* Cell Press, Cambridge Mass. 1990). In theory the term “intron” refers to any segment of DNA that is transcribed but removed from within the transcript by splicing together the exons on either side of it. Operationally, exon sequences occur in the mRNA sequence of a gene as defined by Ref. SEQ ID numbers. Operationally, intron sequences are the intervening sequences within the genomic DNA of a gene, bracketed by exon sequences and having GT and AG splice consensus sequences at their 5’ and 3’ boundaries.

GENE EXPRESSION ASSAY

[0064] The present disclosure provides methods that employ, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques), microbiology, cell biology, and biochemistry, which are within the skill of the art. Such techniques are explained fully in the literature, such as, “Molecular Cloning: A Laboratory Manual”, 2nd edition (Sambrook et al., 1989); “Oligonucleotide Synthesis” (M.J. Gait, ed., 1984); “Animal Cell Culture” (R.I. Freshney, ed., 1987); “Methods in Enzymology” (Academic Press, Inc.); “Handbook of Experimental Immunology”, 4th edition (D.M. Weir & C.C. Blackwell, eds., Blackwell Science Inc., 1987); “Gene Transfer Vectors for Mammalian Cells” (J.M. Miller & M.P. Calos, eds., 1987); “Current Protocols in Molecular Biology” (F.M. Ausubel et al., eds., 1987); and “PCR: The Polymerase Chain Reaction”, (Mullis et al., eds., 1994).

1. Gene Expression Profiling

[0065] Methods of gene expression profiling include methods based on hybridization analysis of polynucleotides, methods based on sequencing of polynucleotides, and proteomics-based methods. The most commonly used methods known in the art for the quantification of mRNA expression in a sample include northern blotting and in situ hybridization (Parker & Barnes, Methods in Molecular Biology 106:247-283 (1999)); RNase protection assays (Hod, Biotechniques 13:852-854 (1992)); and PCR-based methods, such as reverse transcription polymerase chain reaction (RT-PCR) (Weis et al., Trends in Genetics 8:263-264 (1992)). Alternatively, antibodies may be employed that can recognize

specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes.

2. PCR-based Gene Expression Profiling Methods

a. Reverse Transcriptase PCR (RT-PCR)

[0066] Of the techniques listed above, the most sensitive and most flexible quantitative method is RT-PCR, which can be used to compare mRNA levels in different sample populations, in normal and tumor tissues, with or without drug treatment, to characterize patterns of gene expression, to discriminate between closely related mRNAs, and to analyze RNA structure.

[0067] The first step is the isolation of mRNA from a target sample. The starting material is typically total RNA isolated from human tumors or tumor cell lines, and corresponding normal tissues or cell lines, respectively. Thus RNA can be isolated from a variety of primary tumors, including breast, lung, colon, prostate, brain, liver, kidney, pancreas, spleen, thymus, testis, ovary, uterus, etc., tumor, or tumor cell lines, with pooled DNA from healthy donors. If the source of mRNA is a primary tumor, mRNA can be extracted, for example, from frozen or archived paraffin-embedded and fixed (e.g. formalin-fixed) tissue samples.

[0068] General methods for mRNA extraction are well known in the art and are disclosed in standard textbooks of molecular biology, including Ausubel *et al.*, Current Protocols of Molecular Biology, John Wiley and Sons (1997). Methods for RNA extraction from paraffin embedded tissues are disclosed, for example, in Rupp and Locker, *Lab Invest.* 56:A67 (1987), and De Andrés *et al.*, *BioTechniques* 18:42044 (1995). In particular, RNA isolation can be performed using purification kit, buffer set and protease from commercial manufacturers, such as Qiagen, according to the manufacturer's instructions. For example, total RNA from cells in culture can be isolated using Qiagen RNeasy mini-columns. Other commercially available RNA isolation kits include MasterPure™ Complete DNA and RNA Purification Kit (EPICENTRE®, Madison, WI), and Paraffin Block RNA Isolation Kit (Ambion, Inc.). Total RNA from tissue samples can be isolated using RNA Stat-60 (Tel-Test). RNA prepared from tumor can be isolated, for example, by cesium chloride density gradient centrifugation.

[0069] In some cases, it may be appropriate to amplify RNA prior to initiating expression profiling. It is often the case that only very limited amounts of valuable clinical specimens are available for molecular analysis. This may be due to the fact that the tissues have already be used for other laboratory analyses or may be due to the fact that the original

specimen is very small as in the case of needle biopsy or very small primary tumors. When tissue is limiting in quantity it is generally also the case that only small amounts of total RNA can be recovered from the specimen and as a result only a limited number of genomic markers can be analyzed in the specimen. RNA amplification compensates for this limitation by faithfully reproducing the original RNA sample as a much larger amount of RNA of the same relative composition. Using this amplified copy of the original RNA specimen, unlimited genomic analysis can be done to discovery biomarkers associated with the clinical characteristics of the original biological sample. This effectively immortalizes clinical study specimens for the purposes of genomic analysis and biomarker discovery.

[0070] As RNA cannot serve as a template for PCR, the first step in gene expression profiling by real-time RT-PCR (RT-PCR) is the reverse transcription of the RNA template into cDNA, followed by its exponential amplification in a PCR reaction. The two most commonly used reverse transcriptases are avian myeloblastosis virus reverse transcriptase (AMV-RT) and Moloney murine leukemia virus reverse transcriptase (MMLV-RT). The reverse transcription step is typically primed using specific primers, random hexamers, or oligo-dT primers, depending on the circumstances and the goal of expression profiling. For example, extracted RNA can be reverse-transcribed using a GeneAmp RNA PCR kit (Perkin Elmer, CA, USA), following the manufacturer's instructions. The derived cDNA can then be used as a template in the subsequent PCR reaction. For further details see, e.g. Held *et al.*, *Genome Research* 6:986-994 (1996).

[0071] Although the PCR step can use a variety of thermostable DNA-dependent DNA polymerases, it typically employs the Taq DNA polymerase, which has a 5'-3' nuclease activity but lacks a 3'-5' proofreading endonuclease activity. Thus, TaqMan® PCR typically utilizes the 5'-nuclease activity of Taq or Tth polymerase to hydrolyze a hybridization probe bound to its target amplicon, but any enzyme with equivalent 5' nuclease activity can be used. Two oligonucleotide primers are used to generate an amplicon typical of a PCR reaction. A third oligonucleotide, or probe, is designed to detect nucleotide sequence located between the two PCR primers. The probe is non-extendible by Taq DNA polymerase enzyme, and is labeled with a reporter fluorescent dye and 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

fluorophore. 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.

[0072] TaqMan® RT-PCR can be performed using commercially available equipment, such as, for example, ABI PRISM 7900® Sequence Detection System™ (Perkin-Elmer-Applied Biosystems, Foster City, CA, USA), or LightCycler® 480 Real-Time PCR System (Roche Diagnostics, GmbH, Penzberg, Germany). In a preferred embodiment, the 5' nuclease procedure is run on a real-time quantitative PCR device such as the ABI PRISM 7900® Sequence Detection System™. The system consists of a thermocycler, laser, charge-coupled device (CCD), camera and computer. The system amplifies samples in a 384-well format on a thermocycler. During amplification, laser-induced fluorescent signal is collected in real-time through fiber optics cables for all 384 wells, and detected at the CCD. The system includes software for running the instrument and for analyzing the data.

[0073] 5'-Nuclease assay data are initially expressed as Ct, or the threshold cycle. As discussed above, fluorescence values are recorded during every cycle and represent the amount of product amplified to that point in the amplification reaction. The point when the fluorescent signal is first recorded as statistically significant is the threshold cycle (C_t).

[0074] To minimize errors and the effect of sample-to-sample variation, RT-PCR is usually performed using an internal standard. The ideal internal standard is expressed at a constant level among different tissues, and is unaffected by the experimental treatment. RNAs most frequently used to normalize patterns of gene expression are mRNAs for the housekeeping genes glyceraldehyde-3-phosphate-dehydrogenase (GAPDH) and β -actin.

[0075] The steps of a representative protocol for profiling gene expression using fixed, paraffin-embedded tissues as the RNA source, including mRNA isolation, purification, primer extension and amplification are given in various published journal articles. M. Cronin, Am J Pathol 164(1):35-42 (2004). Briefly, a representative process starts with cutting about 10 μ m thick sections of paraffin-embedded tumor tissue samples. The RNA is then extracted, and protein and DNA are removed. After analysis of the RNA concentration, RNA repair and/or amplification steps may be included, if necessary, and RNA is reverse transcribed using gene specific primers followed by RT-PCR.

b. Design of Intron-Based PCR Primers and Probes

[0076] PCR primers and probes can be designed based upon exon or intron sequences present in the mRNA transcript of the gene of interest. Prior to carrying out primer/probe design, it is necessary to map the target gene sequence to the human genome

assembly in order to identify intron-exon boundaries and overall gene structure. This can be performed using publicly available software, such as Primer3 (Whitehead Inst.) and Primer Express® (Applied Biosystems).

[0077] Where necessary or desired, repetitive sequences of the target sequence can be masked to mitigate non-specific signals. Exemplary tools to accomplish this include the Repeat Masker program available on-line through the Baylor College of Medicine, which screens DNA sequences against a library of repetitive elements and returns a query sequence in which the repetitive elements are masked. The masked intron and exon sequences can then be used to design primer and probe sequences for the desired target sites using any commercially or otherwise publicly available primer/probe design packages, such as Primer Express (Applied Biosystems); MGB assay-by-design (Applied Biosystems); Primer3 (Steve Rozen and Helen J. Skaletsky (2000) Primer3 on the WWW for general users and for biologist programmers. In: Rrawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386).

[0078] Other factors that can influence PCR primer design include primer length, melting temperature (T_m), and G/C content, specificity, complementary primer sequences, and 3'-end sequence. In general, optimal PCR primers are generally 17-30 bases in length, and contain about 20-80%, such as, for example, about 50-60% G+C bases, and exhibit T_m 's between 50 and 80 °C, e.g. about 50 to 70 °C.

[0079] For further guidelines for PCR primer and probe design see, e.g. Dieffenbach, CW. et al, "General Concepts for PCR Primer Design" in: PCR Primer, A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York, 1995, pp. 133-155; Innis and Gelfand, "Optimization of PCRs" in: PCR Protocols, A Guide to Methods and Applications, CRC Press, London, 1994, pp. 5-11; and Plasterer, T.N. Primerselect: Primer and probe design. Methods Mol. Biol. 70:520-527 (1997), the entire disclosures of which are hereby expressly incorporated by reference.

[0080] Table A provides further information concerning the primer, probe, and amplicon sequences associated with the Examples disclosed herein.

c. MassARRAY System

[0081] In the MassARRAY-based gene expression profiling method, developed by Sequenom, Inc. (San Diego, CA) following the isolation of RNA and reverse transcription, the obtained cDNA is spiked with a synthetic DNA molecule (competitor), which matches the targeted cDNA region in all positions, except a single base, and serves as an internal standard. The cDNA/competitor mixture is PCR amplified and is subjected to a post-PCR

shrimp alkaline phosphatase (SAP) enzyme treatment, which results in the dephosphorylation of the remaining nucleotides. After inactivation of the alkaline phosphatase, the PCR products from the competitor and cDNA are subjected to primer extension, which generates distinct mass signals for the competitor- and cDNA-derived PCR products. After purification, these products are dispensed on a chip array, which is pre-loaded with components needed for analysis with matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) analysis. The cDNA present in the reaction is then quantified by analyzing the ratios of the peak areas in the mass spectrum generated. For further details see, e.g. Ding and Cantor, Proc. Natl. Acad. Sci. USA 100:3059-3064 (2003).

d. Other PCR-based Methods

[0082] Further PCR-based techniques include, for example, differential display (Liang and Pardee, Science 257:967-971 (1992)); amplified fragment length polymorphism (iAFLP) (Kawamoto et al., Genome Res. 12:1305-1312 (1999)); BeadArray™ technology (Illumina, San Diego, CA; Oliphant et al., Discovery of Markers for Disease (Supplement to Biotechniques), June 2002; Ferguson et al., Analytical Chemistry 72:5618 (2000)); BeadArray for Detection of Gene Expression (BADGE), using the commercially available Luminex100 LabMAP system and multiple color-coded microspheres (Luminex Corp., Austin, TX) in a rapid assay for gene expression (Yang et al., Genome Res. 11:1888-1898 (2001)); and high coverage expression profiling (HiCEP) analysis (Fukumura et al., Nucl. Acids. Res. 31(16) e94 (2003)).

3. Microarrays

[0083] Differential gene expression can also be identified, or confirmed using the microarray technique. Thus, the expression profile of breast cancer-associated genes can be measured in either fresh or paraffin-embedded tumor tissue, using microarray technology. In this method, polynucleotide sequences of interest (including cDNAs and oligonucleotides) are plated, or arrayed, on a microchip substrate. The arrayed sequences are then hybridized with specific DNA probes from cells or tissues of interest. Just as in the RT-PCR method, the source of mRNA typically is total RNA isolated from human tumors or tumor cell lines, and corresponding normal tissues or cell lines. Thus RNA can be isolated from a variety of primary tumors or tumor cell lines. If the source of mRNA is a primary tumor, mRNA can be extracted, for example, from frozen or archived paraffin-embedded and fixed (e.g. formalin-fixed) tissue samples, which are routinely prepared and preserved in everyday clinical practice.

[0084] In a specific embodiment of the microarray technique, PCR amplified inserts of cDNA clones are applied to a substrate in a dense array. Preferably at least 10,000 nucleotide sequences are applied to the substrate. The microarrayed genes, immobilized on the microchip at 10,000 elements each, are suitable for hybridization under stringent conditions. Fluorescently labeled cDNA probes may be generated through incorporation of fluorescent nucleotides by reverse transcription of RNA extracted from tissues of interest. Labeled cDNA probes applied to the chip hybridize with specificity to each spot of DNA on the array. After stringent washing to remove non-specifically bound probes, the chip is scanned by confocal laser microscopy or by another detection method, such as a CCD camera. Quantitation of hybridization of each arrayed element allows for assessment of corresponding mRNA abundance. With dual color fluorescence, separately labeled cDNA probes generated from two sources of RNA are hybridized pairwise to the array. The relative abundance of the transcripts from the two sources corresponding to each specified gene is thus determined simultaneously. The miniaturized scale of the hybridization affords a convenient and rapid evaluation of the expression pattern for large numbers of genes. Such methods have been shown to have the sensitivity required to detect rare transcripts, which are expressed at a few copies per cell, and to reproducibly detect at least approximately two-fold differences in the expression levels (Schena *et al.*, *Proc. Natl. Acad. Sci. USA* 93(2):106-149 (1996)). Microarray analysis can be performed by commercially available equipment, following manufacturer's protocols, such as by using the Affymetrix GenChip technology, or Agilent's microarray technology.

[0085] The development of microarray methods for large-scale analysis of gene expression makes it possible to search systematically for molecular markers of cancer classification and outcome prediction in a variety of tumor types.

4. Gene Expression Analysis by Nucleic Acid Sequencing

[0086] Nucleic acid sequencing technologies are suitable methods for analysis of gene expression. The principle underlying these methods is that the number of times a cDNA sequence is detected in a sample is directly related to the relative expression of the mRNA corresponding to that sequence. These methods are sometimes referred to by the term Digital Gene Expression (DGE) to reflect the discrete numeric property of the resulting data. Early methods applying this principle were Serial Analysis of Gene Expression (SAGE) and Massively Parallel Signature Sequencing (MPSS). See, e.g., S. Brenner, et al., *Nature Biotechnology* 18(6):630-634 (2000). More recently, the advent of "next-generation" sequencing technologies has made DGE simpler, higher throughput, and more affordable.

As a result, more laboratories are able to utilize DGE to screen the expression of more genes in more individual patient samples than previously possible. See, e.g., J. Marioni, *Genome Research* 18(9):1509-1517 (2008); R. Morin, *Genome Research* 18(4):610-621 (2008); A. Mortazavi, *Nature Methods* 5(7):621-628 (2008); N. Cloonan, *Nature Methods* 5(7):613-619 (2008).

5. *Isolating RNA from Body Fluids*

[0087] Methods of isolating RNA for expression analysis from blood, plasma and serum (See for example, Tsui NB et al. (2002) 48,1647-53 and references cited therein) and from urine (See for example, Boom R et al. (1990) *J Clin Microbiol.* 28, 495-503 and reference cited therein) have been described.

6. *Immunohistochemistry*

[0088] Immunohistochemistry methods are also suitable for detecting the expression levels of the prognostic markers of the present invention. Thus, antibodies or antisera, preferably polyclonal antisera, and most preferably monoclonal antibodies specific for each marker are used to detect expression. The antibodies can be detected by direct labeling of the antibodies themselves, for example, with radioactive labels, fluorescent labels, hapten labels such as, biotin, or an enzyme such as horse radish peroxidase or alkaline phosphatase. Alternatively, unlabeled primary antibody is used in conjunction with a labeled secondary antibody, comprising antisera, polyclonal antisera or a monoclonal antibody specific for the primary antibody. Immunohistochemistry protocols and kits are well known in the art and are commercially available.

7. *Proteomics*

[0089] The term “proteome” is defined as the totality of the proteins present in a sample (e.g. tissue, organism, or cell culture) at a certain point of time. Proteomics includes, among other things, study of the global changes of protein expression in a sample (also referred to as “expression proteomics”). Proteomics typically includes the following steps: (1) separation of individual proteins in a sample by 2-D gel electrophoresis (2-D PAGE); (2) identification of the individual proteins recovered from the gel, e.g. by mass spectrometry or N-terminal sequencing, and (3) analysis of the data using bioinformatics. Proteomics methods are valuable supplements to other methods of gene expression profiling, and can be used, alone or in combination with other methods, to detect the products of the prognostic markers of the present invention.

8. General Description of the mRNA Isolation, Purification, and Amplification

[0090] The steps of a representative protocol for profiling gene expression using fixed, paraffin-embedded tissues as the RNA source, including mRNA isolation, purification, primer extension and amplification are provided in various published journal articles (for example: T.E. Godfrey *et al.*, *J. Molec. Diagnostics* 2: 84-91 [2000]; K. Specht *et al.*, *Am. J. Pathol.* 158: 419-29 [2001]). Briefly, a representative process starts with cutting about 10 μ m thick sections of paraffin-embedded tumor tissue samples. The RNA is then extracted, and protein and DNA are removed. After analysis of the RNA concentration, RNA repair and/or amplification steps may be included, if necessary, and RNA is reverse transcribed using gene specific primers followed by RT-PCR. Finally, the data are analyzed to identify the best treatment option(s) available to the patient on the basis of the characteristic gene expression pattern identified in the tumor sample examined, dependent on the predicted likelihood of cancer recurrence.

9. Normalization

[0091] The expression data used in the methods disclosed herein can be normalized. Normalization refers to a process to correct for (normalize away), for example, differences in the amount of RNA assayed and variability in the quality of the RNA used, to remove unwanted sources of systematic variation in C_t measurements, and the like. With respect to RT-PCR experiments involving archived fixed paraffin embedded tissue samples, sources of systematic variation are known to include the degree of RNA degradation relative to the age of the patient sample and the type of fixative used to preserve the sample. Other sources of systematic variation are attributable to laboratory processing conditions.

[0092] Assays can provide for normalization by incorporating the expression of certain normalizing genes, which genes do not significantly differ in expression levels under the relevant conditions. Exemplary normalization genes include housekeeping genes such as PGK1 and UBB. (See, e.g., E. Eisenberg, et al., *Trends in Genetics* 19(7):362-365 (2003).) Normalization can be based on the mean or median signal (C_T) of all of the assayed genes or a large subset thereof (global normalization approach). In general, the normalizing genes, also referred to as reference genes should be genes that are known not to exhibit significantly different expression in colorectal cancer as compared to non-cancerous colorectal tissue, and are not significantly affected by various sample and process conditions, thus provide for normalizing away extraneous effects.

[0093] Unless noted otherwise, normalized expression levels for each mRNA/tested tumor/patient will be expressed as a percentage of the expression level measured in the

reference set. A reference set of a sufficiently high number (e.g. 40) of tumors yields a distribution of normalized levels of each mRNA species. The level measured in a particular tumor sample to be analyzed falls at some percentile within this range, which can be determined by methods well known in the art.

[0094] In exemplary embodiments, one or more of the following genes are used as references by which the expression data is normalized: AAMP, ARF1, EEF1A1, ESD, GPS1, H3F3A, HNRPC, RPL13A, RPL41, RPS23, RPS27, SDHA, TCEA1, UBB, YWHAZ, B-actin, GUS, GAPDH, RPLPO, and TFRC. For example, the calibrated weighted average C_t measurements for each of the prognostic genes may be normalized relative to the mean of at least three reference genes, at least four reference genes, or at least five reference genes.

[0095] Those skilled in the art will recognize that normalization may be achieved in numerous ways, and the techniques described above are intended only to be exemplary, not exhaustive.

REPORTING RESULTS

[0096] The methods of the present disclosure are suited for the preparation of reports summarizing the expected or predicted clinical outcome resulting from the methods of the present disclosure. A “report,” as described herein, is an electronic or tangible document that includes report elements that provide information of interest relating to a likelihood assessment or a risk assessment and its results. A subject report includes at least a likelihood assessment or a risk assessment, e.g., an indication as to the risk of recurrence of breast cancer, including local recurrence and metastasis of breast cancer. A subject report can include an assessment or estimate of one or more of disease-free survival, recurrence-free survival, metastasis-free survival, and overall survival. A subject report can be completely or partially electronically generated, e.g., presented on an electronic display (e.g., computer monitor). A report can further include one or more of: 1) information regarding the testing facility; 2) service provider information; 3) patient data; 4) sample data; 5) an interpretive report, which can include various information including: a) indication; b) test data, where test data can include a normalized level of one or more genes of interest, and 6) other features.

[0097] The present disclosure thus provides for methods of creating reports and the reports resulting therefrom. The report may include a summary of the expression levels of the RNA transcripts, or the expression products of such RNA transcripts, for certain genes in the cells obtained from the patient’s tumor. The report can include information relating to

prognostic covariates of the patient. The report may include an estimate that the patient has an increased risk of recurrence. That estimate may be in the form of a score or patient stratifier scheme (e.g., low, intermediate, or high risk of recurrence). The report may include information relevant to assist with decisions about the appropriate surgery (e.g., partial or total mastectomy) or treatment for the patient.

[0098] Thus, in some embodiments, the methods of the present disclosure further include generating a report that includes information regarding the patient's likely clinical outcome, e.g. risk of recurrence. For example, the methods disclosed herein can further include a step of generating or outputting a report providing the results of a subject risk assessment, which report can be provided in the form of an electronic medium (e.g., an electronic display on a computer monitor), or in the form of a tangible medium (e.g., a report printed on paper or other tangible medium).

[0099] A report that includes information regarding the patient's likely prognosis (e.g., the likelihood that a patient having breast cancer will have a good prognosis or positive clinical outcome in response to surgery and/or treatment) is provided to a user. An assessment as to the likelihood is referred to below as a "risk report" or, simply, "risk score." A person or entity that prepares a report ("report generator") may also perform the likelihood assessment. The report generator may also perform one or more of sample gathering, sample processing, and data generation, e.g., the report generator may also perform one or more of: a) sample gathering; b) sample processing; c) measuring a level of a risk gene; d) measuring a level of a reference gene; and e) determining a normalized level of a risk gene. Alternatively, an entity other than the report generator can perform one or more sample gathering, sample processing, and data generation.

[00100] For clarity, it should be noted that the term "user," which is used interchangeably with "client," is meant to refer to a person or entity to whom a report is transmitted, and may be the same person or entity who does one or more of the following: a) collects a sample; b) processes a sample; c) provides a sample or a processed sample; and d) generates data (e.g., level of a risk gene; level of a reference gene product(s); normalized level of a risk gene ("prognosis gene")) for use in the likelihood assessment. In some cases, the person(s) or entity(ies) who provides sample collection and/or sample processing and/or data generation, and the person who receives the results and/or report may be different persons, but are both referred to as "users" or "clients" herein to avoid confusion. In certain embodiments, e.g., where the methods are completely executed on a single computer, the user or client provides for data input and review of data output. A "user" can be a health

professional (e.g., a clinician, a laboratory technician, a physician (e.g., an oncologist, surgeon, pathologist), etc.).

[00101] In embodiments where the user only executes a portion of the method, the individual who, after computerized data processing according to the methods of the present disclosure, reviews data output (e.g., results prior to release to provide a complete report, a complete, or reviews an "incomplete" report and provides for manual intervention and completion of an interpretive report) is referred to herein as a "reviewer." The reviewer may be located at a location remote to the user (e.g., at a service provided separate from a healthcare facility where a user may be located).

[00102] Where government regulations or other restrictions apply (e.g., requirements by health, malpractice, or liability insurance), all results, whether generated wholly or partially electronically, are subjected to a quality control routine prior to release to the user.

CLINICAL UTILITY

[00103] The gene expression assay and information provided by the practice of the methods disclosed herein facilitates physicians in making more well-informed treatment decisions, and to customize the treatment of cancer to the needs of individual patients, thereby maximizing the benefit of treatment and minimizing the exposure of patients to unnecessary treatments which may provide little or no significant benefits and often carry serious risks due to toxic side-effects.

[00104] Single or multi-analyte gene expression tests can be used measure the expression level of one or more genes involved in each of several relevant physiologic processes or component cellular characteristics. The expression level(s) may be used to calculate such a quantitative score, and such score may be arranged in subgroups (e.g., tertiles) wherein all patients in a given range are classified as belonging to a risk category (e.g., low, intermediate, or high). The grouping of genes may be performed at least in part based on knowledge of the contribution of the genes according to physiologic functions or component cellular characteristics, such as in the groups discussed above.

[00105] The utility of a gene marker in predicting cancer may not be unique to that marker. An alternative marker having an expression pattern that is parallel to that of a selected marker gene may be substituted for, or used in addition to, a test marker. Due to the co-expression of such genes, substitution of expression level values should have little impact on the overall prognostic utility of the test. The closely similar expression patterns of two genes may result from involvement of both genes in the same process and/or being under common regulatory control in colon tumor cells. The present disclosure thus contemplates

the use of such co-expressed genes or gene sets as substitutes for, or in addition to, prognostic methods of the present disclosure.

[00106] The molecular assay and associated information provided by the methods disclosed herein for predicting the clinical outcome in cancer, e.g. breast cancer, have utility in many areas, including in the development and appropriate use of drugs to treat cancer, to stratify cancer patients for inclusion in (or exclusion from) clinical studies, to assist patients and physicians in making treatment decisions, provide economic benefits by targeting treatment based on personalized genomic profile, and the like. For example, the recurrence score may be used on samples collected from patients in a clinical trial and the results of the test used in conjunction with patient outcomes in order to determine whether subgroups of patients are more or less likely to demonstrate an absolute benefit from a new drug than the whole group or other subgroups. Further, such methods can be used to identify from clinical data subsets of patients who are expected to benefit from adjuvant therapy. Additionally, a patient is more likely to be included in a clinical trial if the results of the test indicate a higher likelihood that the patient will have a poor clinical outcome if treated with surgery alone and a patient is less likely to be included in a clinical trial if the results of the test indicate a lower likelihood that the patient will have a poor clinical outcome if treated with surgery alone.

STATISTICAL ANALYSIS OF GENE EXPRESSION LEVELS

[00107] One skilled in the art will recognize that there are many statistical methods that may be used to determine whether there is a significant relationship between an outcome of interest (e.g., likelihood of survival, likelihood of response to chemotherapy) and expression levels of a marker gene as described here. This relationship can be presented as a continuous recurrence score (RS), or patients may stratified into risk groups (e.g., low, intermediate, high). For example, a Cox proportional hazards regression model may fit to a particular clinical endpoint (e.g., RFS, DFS, OS). One assumption of the Cox proportional hazards regression model is the proportional hazards assumption, i.e. the assumption that effect parameters multiply the underlying hazard.

Coexpression Analysis

[00108] The present disclosure provides genes that co-express with particular prognostic and/or predictive gene that has been identified as having a significant correlation to recurrence and/or treatment benefit. To perform particular biological processes, genes often work together in a concerted way, i.e. they are co-expressed. Co-expressed gene

groups identified for a disease process like cancer can serve as biomarkers for disease progression and response to treatment. Such co-expressed genes can be assayed in lieu of, or in addition to, assaying of the prognostic and/or predictive gene with which they are co-expressed.

[00109] One skilled in the art will recognize that many co-expression analysis methods now known or later developed will fall within the scope and spirit of the present invention. These methods may incorporate, for example, correlation coefficients, co-expression network analysis, clique analysis, etc., and may be based on expression data from RT-PCR, microarrays, sequencing, and other similar technologies. For example, gene expression clusters can be identified using pair-wise analysis of correlation based on Pearson or Spearman correlation coefficients. (See, e.g., Pearson K. and Lee A., *Biometrika* 2, 357 (1902); C. Spearman, *Amer. J. Psychol* 15:72-101 (1904); J. Myers, A. Well, *Research Design and Statistical Analysis*, p. 508 (2nd Ed., 2003).) In general, a correlation coefficient of equal to or greater than 0.3 is considered to be statistically significant in a sample size of at least 20. (See, e.g., G. Norman, D. Streiner, *Biostatistics: The Bare Essentials*, 137-138 (3rd Ed. 2007).) In one embodiment disclosed herein, co-expressed genes were identified using a Spearman correlation value of at least 0.7.

Computer program

[00110] The values from the assays described above, such as expression data, recurrence score, treatment score and/or benefit score, can be calculated and stored manually. Alternatively, the above-described steps can be completely or partially performed by a computer program product. The present invention thus provides a computer program product including a computer readable storage medium having a computer program stored on it. The program can, when read by a computer, execute relevant calculations based on values obtained from analysis of one or more biological sample from an individual (e.g., gene expression levels, normalization, thresholding, and conversion of values from assays to a score and/or graphical depiction of likelihood of recurrence/response to chemotherapy, gene co-expression or clique analysis, and the like). The computer program product has stored therein a computer program for performing the calculation.

[00111] The present disclosure provides systems for executing the program described above, which system generally includes: a) a central computing environment; b) an input device, operatively connected to the computing environment, to receive patient data, wherein the patient data can include, for example, expression level or other value obtained from an assay using a biological sample from the patient, or microarray data, as described in detail

above; c) an output device, connected to the computing environment, to provide information to a user (e.g., medical personnel); and d) an algorithm executed by the central computing environment (e.g., a processor), where the algorithm is executed based on the data received by the input device, and wherein the algorithm calculates a, risk, risk score, or treatment group classification, gene co-expression analysis, thresholding, or other functions described herein. The methods provided by the present invention may also be automated in whole or in part.

Manual and Computer-Assisted Methods and Products

[00112] The methods and systems described herein can be implemented in numerous ways. In one embodiment of particular interest, the methods involve use of a communications infrastructure, for example the Internet. Several embodiments are discussed below. It is also to be understood that the present disclosure may be implemented in various forms of hardware, software, firmware, processors, or a combination thereof. The methods and systems described herein can be implemented as a combination of hardware and software. The software can be implemented as an application program tangibly embodied on a program storage device, or different portions of the software implemented in the user's computing environment (e.g., as an applet) and on the reviewer's computing environment, where the reviewer may be located at a remote site associated (e.g., at a service provider's facility).

[00113] For example, during or after data input by the user, portions of the data processing can be performed in the user-side computing environment. For example, the user-side computing environment can be programmed to provide for defined test codes to denote a likelihood "risk score," where the score is transmitted as processed or partially processed responses to the reviewer's computing environment in the form of test code for subsequent execution of one or more algorithms to provide a results and/or generate a report in the reviewer's computing environment. The risk score can be a numerical score (representative of a numerical value, e.g. likelihood of recurrence based on validation study population) or a non-numerical score representative of a numerical value or range of numerical values (e.g., low, intermediate, or high).

[00114] The application program for executing the algorithms described herein may be uploaded to, and executed by, a machine comprising any suitable architecture. In general, the machine involves a computer platform having hardware such as one or more central processing units (CPU), a random access memory (RAM), and input/output (I/O) interface(s). The computer platform also includes an operating system and microinstruction

code. The various processes and functions described herein may either be part of the microinstruction code or part of the application program (or a combination thereof) that is executed via the operating system. In addition, various other peripheral devices may be connected to the computer platform such as an additional data storage device and a printing device.

[00115] As a computer system, the system generally includes a processor unit. The processor unit operates to receive information, which can include test data (e.g., level of a risk gene, level of a reference gene product(s); normalized level of a gene; and may also include other data such as patient data. This information received can be stored at least temporarily in a database, and data analyzed to generate a report as described above.

[00116] Part or all of the input and output data can also be sent electronically; certain output data (e.g., reports) can be sent electronically or telephonically (e.g., by facsimile, e.g., using devices such as fax back). Exemplary output receiving devices can include a display element, a printer, a facsimile device and the like. Electronic forms of transmission and/or display can include email, interactive television, and the like. In an embodiment of particular interest, all or a portion of the input data and/or all or a portion of the output data (e.g., usually at least the final report) are maintained on a web server for access, preferably confidential access, with typical browsers. The data may be accessed or sent to health professionals as desired. The input and output data, including all or a portion of the final report, can be used to populate a patient's medical record which may exist in a confidential database at the healthcare facility.

[00117] A system for use in the methods described herein generally includes at least one computer processor (e.g., where the method is carried out in its entirety at a single site) or at least two networked computer processors (e.g., where data is to be input by a user (also referred to herein as a "client") and transmitted to a remote site to a second computer processor for analysis, where the first and second computer processors are connected by a network, e.g., via an intranet or internet). The system can also include a user component(s) for input; and a reviewer component(s) for review of data, generated reports, and manual intervention. Additional components of the system can include a server component(s); and a database(s) for storing data (e.g., as in a database of report elements, e.g., interpretive report elements, or a relational database (RDB) which can include data input by the user and data output. The computer processors can be processors that are typically found in personal desktop computers (e.g., IBM, Dell, Macintosh), portable computers, mainframes, minicomputers, or other computing devices.

[00118] The networked client/server architecture can be selected as desired, and can be, for example, a classic two or three tier client server model. A relational database management system (RDMS), either as part of an application server component or as a separate component (RDB machine) provides the interface to the database.

[00119] In one example, the architecture is provided as a database-centric client/server architecture, in which the client application generally requests services from the application server which makes requests to the database (or the database server) to populate the report with the various report elements as required, particularly the interpretive report elements, especially the interpretation text and alerts. The server(s) (e.g., either as part of the application server machine or a separate RDB/relational database machine) responds to the client's requests.

[00120] The input client components can be complete, stand-alone personal computers offering a full range of power and features to run applications. The client component usually operates under any desired operating system and includes a communication element (e.g., a modem or other hardware for connecting to a network), one or more input devices (e.g., a keyboard, mouse, keypad, or other device used to transfer information or commands), a storage element (e.g., a hard drive or other computer-readable, computer-writable storage medium), and a display element (e.g., a monitor, television, LCD, LED, or other display device that conveys information to the user). The user enters input commands into the computer processor through an input device. Generally, the user interface is a graphical user interface (GUI) written for web browser applications.

[00121] The server component(s) can be a personal computer, a minicomputer, or a mainframe and offers data management, information sharing between clients, network administration and security. The application and any databases used can be on the same or different servers.

[00122] Other computing arrangements for the client and server(s), including processing on a single machine such as a mainframe, a collection of machines, or other suitable configuration are contemplated. In general, the client and server machines work together to accomplish the processing of the present disclosure.

[00123] Where used, the database(s) is usually connected to the database server component and can be any device that will hold data. For example, the database can be a any magnetic or optical storing device for a computer (e.g., CDROM, internal hard drive, tape drive). The database can be located remote to the server component (with access via a network, modem, etc.) or locally to the server component.

[00124] Where used in the system and methods, the database can be a relational database that is organized and accessed according to relationships between data items. The relational database is generally composed of a plurality of tables (entities). The rows of a table represent records (collections of information about separate items) and the columns represent fields (particular attributes of a record). In its simplest conception, the relational database is a collection of data entries that "relate" to each other through at least one common field.

[00125] Additional workstations equipped with computers and printers may be used at point of service to enter data and, in some embodiments, generate appropriate reports, if desired. The computer(s) can have a shortcut (e.g., on the desktop) to launch the application to facilitate initiation of data entry, transmission, analysis, report receipt, etc. as desired.

Computer-readable storage media

[00126] The present disclosure also contemplates a computer-readable storage medium (e.g. CD-ROM, memory key, flash memory card, diskette, etc.) having stored thereon a program which, when executed in a computing environment, provides for implementation of algorithms to carry out all or a portion of the results of a response likelihood assessment as described herein. Where the computer-readable medium contains a complete program for carrying out the methods described herein, the program includes program instructions for collecting, analyzing and generating output, and generally includes computer readable code devices for interacting with a user as described herein, processing that data in conjunction with analytical information, and generating unique printed or electronic media for that user.

[00127] Where the storage medium provides a program that provides for implementation of a portion of the methods described herein (e.g., the user-side aspect of the methods (e.g., data input, report receipt capabilities, etc.)), the program provides for transmission of data input by the user (e.g., via the internet, via an intranet, etc.) to a computing environment at a remote site. Processing or completion of processing of the data is carried out at the remote site to generate a report. After review of the report, and completion of any needed manual intervention, to provide a complete report, the complete report is then transmitted back to the user as an electronic document or printed document (e.g., fax or mailed paper report). The storage medium containing a program according to the present disclosure can be packaged with instructions (e.g., for program installation, use, etc.) recorded on a suitable substrate or a web address where such instructions may be obtained. The computer-readable storage medium can also be provided in combination with one or

more reagents for carrying out response likelihood assessment (e.g., primers, probes, arrays, or other such kit components).

[00128] All aspects of the present invention may also be practiced such that a limited number of additional genes that are co-expressed with the disclosed genes, for example as evidenced by statistically meaningful Pearson and/or Spearman correlation coefficients, are included in a prognostic or predictive test in addition to and/or in place of disclosed genes.

[00129] Having described the invention, the same will be more readily understood through reference to the following Examples, which are provided by way of illustration, and are not intended to limit the invention in any way.

EXAMPLE 1:

[00130] The study included breast cancer tumor samples obtained from 136 patients diagnosed with breast cancer (“Providence study”). Biostatistical modeling studies of prototypical data sets demonstrated that amplified RNA is a useful substrate for biomarker identification studies. This was verified in this study by including known breast cancer biomarkers along with candidate prognostic genes in the tissues samples. The known biomarkers were shown to be associated with clinical outcome in amplified RNA based on the criteria outlined in this protocol.

Study design

[00131] Refer to the original Providence Phase II study protocol for biopsy specimen information. The study looked at the statistical association between clinical outcome and 384 candidate biomarkers tested in amplified samples derived from 25 ng of mRNA that was extracted from fixed, paraffin-embedded tissue samples obtained from 136 of the original Providence Phase II study samples. The expression level of the candidate genes was normalized using reference genes. Several reference genes were analyzed in this study: AAMP, ARF1, EEF1A1, ESD, GPS1, H3F3A, HNRPC, RPL13A, RPL41, RPS23, RPS27, SDHA, TCEA1, UBB, YWHAZ, B-actin, GUS, GAPDH, RPLPO, and TFRC.

[00132] The 136 samples were split into 3 automated RT plates each with 2X 48 samples and 40 samples and 3 RT positive and negative controls. Quantitative PCR assays were performed in 384 wells without replicate using the QuantiTect Probe PCR Master Mix® (Qiagen). Plates were analyzed on the Light Cycler® 480 and, after data quality control, all samples from the RT plate 3 were repeated and new RT-PCR data was generated. The data was normalized by subtracting the median crossing point (C_P) (point at which

detection rises above background signal) for five reference genes from the C_p value for each individual candidate gene. This normalization is performed on each sample resulting in final data that has been adjusted for differences in overall sample C_p . This data set was used for the final data analysis.

Data Analysis

[00133] For each gene, a standard z test was run.. (S. Darby, J. Reissland, Journal of the Royal Statistical Society 144(3):298-331 (1981)). This returns a z score (measure of distance in standard deviations of a sample from the mean), p value, and residuals along with other statistics and parameters from the model. If the z score is negative, expression is positively correlated with a good prognosis; if positive, expression is negatively correlated to a good prognosis. Using the p values, a q value was created using a library q value. The poorly correlated and weakly expressed genes were excluded from the calculation of the distribution used for the q values. For each gene, Cox Proportional Hazard Model test was run checking survival time matched with the event vector against gene expression. This returned a hazard ratio (HR) estimating the effect of expression of each gene (individually) on the risk of a cancer-related event. The resulting data is provided in Tables 1-6. A $HR < 1$ indicates that expression of that gene is positively associated with a good prognosis, while a $HR > 1$ indicates that expression of that gene is negatively associated with a good prognosis.

EXAMPLE 2:

Study design

[00134] Amplified samples were derived from 25 ng of mRNA that was extracted from fixed, paraffin-embedded tissue samples obtained from 78 evaluable cases from a Phase II breast cancer study conducted at Rush University Medical Center. Three of the samples failed to provide sufficient amplified RNA at 25 ng, so amplification was repeated a second time with 50 ng of RNA. The study also analyzed several reference genes for use in normalization: AAMP, ARF1, EEF1A1, ESD, GPS1, H3F3A, HNRPC, RPL13A, RPL41, RPS23, RPS27, SDHA, TCEA1, UBB, YWHAZ, Beta-actin, RPLPO, TFRC, GUS, and GAPDH.

[00135] Assays were performed in 384 wells without replicate using the QuantiTect Probe PCR Master Mix. Plates were analyzed on the Light Cycler 480 instruments. This data set was used for the final data analysis. The data was normalized by subtracting the median C_p for five reference genes from the C_p value for each individual candidate gene. This

normalization was performed on each sample resulting in final data that was adjusted for differences in overall sample C_p .

Data analysis

[00136] There were 34 samples with average C_p values above 35. However, none of the samples were excluded from analysis because they were deemed to have sufficient valuable information to remain in the study. Principal Component Analysis (PCA) was used to determine whether there was a plate effect causing variation across the different RT plates. The first principal component correlated well with the median expression values, indicating that expression level accounted for most of the variation between samples. Also, there were no unexpected variations between plates.

Data for Other Variables

[00137] Group – The patients were divided into two groups (cancer/non-cancer). There was little difference between the two in overall gene expression as the difference between median C_p value in each group was minimal (0.7).

[00138] Sample Age – The samples varied widely in their overall gene expression but there was a trend toward lower C_p values as they decreased in age.

[00139] Instrument – The overall sample gene expression from instrument to instrument was consistent. One instrument showed a slightly higher median C_p compared to the other three, but it was well within the acceptable variation.

[00140] RT Plate – The overall sample gene expression between RT plates was also very consistent. The median C_p for each of the 3 RT plates (2 automated RT plates and 1 manual plate containing repeated samples) were all within 1 C_p of each other.

Univariate Analyses for Genes Significantly Different Between Study Groups

[00141] The genes were analyzed using the z-test and Cox Proportional Hazard Model, as described in Example 1. The resulting data can be seen in Tables 7-12.

EXAMPLE 3:

[00142] The statistical correlations between clinical outcome and expression levels of the genes identified in Examples 1 and 2 were validated in breast cancer gene expression datasets maintained by the Swiss Institute of Bioinformatics (SIB). Further information concerning the SIB database, study datasets, and processing methods, is providing in P. Wirapati, et al., Breast Cancer Research 10(4):R65 (2008). Univariate Cox proportional hazards analyses were performed to confirm the relationship between clinical outcome (DFS, MFS, OS) of breast cancer patients and expression levels of the genes identified as

significant in the amplified RNA studies described above. The meta-analysis included both fixed-effect and random-effect models, which are further described in L. Hedges and J. Vevea, *Psychological Methods* 3 (4): 486-504 (1998) and K. Sidik and J. Jonkman, *Statistics in Medicine* 26:1964-1981 (2006) (the contents of which are incorporated herein by reference). The results of the validation for all genes identified as having a statistically significant association with breast cancer clinical outcome are described in Table 13. In those tables, "Est" designates an estimated coefficient of a covariate (gene expression); "SE" is standard error; "t" is the t-score for this estimate (i.e., Est/SE); and "fe" is the fixed estimate of effect from the meta analysis. Several of gene families with significant statistical association with clinical outcome (including metabolic, proliferation, immune, and stromal group genes) in breast cancer were confirmed using the SIB dataset. For example, Table 14 contains analysis of genes included in the metabolic group and Table 15 the stromal group.

EXAMPLE 4:

[00143] A co-expression analysis was conducted using microarray data from six (6) breast cancer data sets. The "processed" expression values are taken from the GEO website, however, further processing was necessary. If the expression values are RMA, they are median normalized on the sample level. If the expression values are MAS5.0, they are: (1) changed to 10 if they are <10; (2) log base e transformed; and (3) median normalized on the sample level.

[00144] Generating Correlation Pairs: A rank matrix was generated by arranging the expression values for each sample in decreasing order. Then a correlation matrix was created by calculating the Spearman correlation values for every pair of probe IDs. Pairs of probes which had a Spearman value ≥ 0.7 were considered co-expressed. Redundant or overlapping correlation pairs in multiple datasets were identified. For each correlation matrix generated from an array dataset, pairs of significant probes that occur in >1 dataset were identified. This served to filter "non-significant" pairs from the analysis as well as provide extra evidence for "significant" pairs with their presence in multiple datasets. Depending on the number of datasets included in each tissue specific analysis, only pairs which occur in a minimum # or % of datasets were included.

[00145] Co-expression cliques were generated using the Bron-Kerbosch algorithm for maximal clique finding in an undirected graph. The algorithm generates three sets of nodes: *compsub*, *candidates*, and *not*. *Compsub* contains the set of nodes to be extended or shrunk by one depending on its traversal direction on the tree search. *Candidates* consists of all the

nodes eligible to be added to *compsub*. *Not* contains the set of nodes that have been added to *compsub* and are now excluded from extension. The algorithm consists of five steps: selection of a candidate; adding the candidate node to *compsub*; creating new sets *candidates* and *not* from the old sets by removing all points not connected to the candidate node; recursively calling the extension operator on the **new** *candidates* and *not* sets; and upon return, remove the candidate node from *compsub* and place in the **old** *not* set.

[00146] There was a depth-first search with pruning, and the selection of candidate nodes had an effect on the run time of the algorithm. By selecting nodes in decreasing order of frequency in the pairs, the run time was optimized. Also, recursive algorithms generally cannot be implemented in a multi-threaded manner, but was multi-threaded the extension operator of the first recursive level. Since the data between the threads were independent because they were at the top-level of the recursive tree, they were run in parallel.

[00147] Clique Mapping and Normalization: Since the members of the co-expression pairs and cliques are at the probe level, one must map the probe IDs to genes (or Refseqs) before they can be analyzed. The Affymetrix gene map information was used to map every probe ID to a gene name. Probes may map to multiple genes, and genes may be represented by multiple probes. The data for each clique is validated by manually calculating the correlation values for each pair from a single clique.

[00148] The results of this co-expression analysis are set forth in Tables 16-18.

TABLE A

Gene	Sequence ID	Official Symbol	F Primer Seq	SEQ ID NO:	R Primer Seq	SEQ ID NO:	Probe Seq	SEQ ID NO:	Target Seq Length	Amplicon Sequence	SEQ ID NO:
A-Catenin	NM_001903.1	CTNNA1	CGTTCCGATCCTCTATAC	1	AGGTCCCTGTGGCCTTATAGG	385	ATGCCTACAGCACCCCTGATGTCGCA	769	78	CGTTCCGATCCTCTATACTGATGCCAGGCATGCCTACAGCACCCCTGATGTCGCA GCCTATTAAGGCCAACAAGGACCT	1153
			GTGTGGCA GTTGGACA CTAA	2	CTCCATCCA CTCCAGGTC TC	386	CGCTTCAAAGGACCAGACCTCTC	770	66	GTGTGGCAGGTGGACACTAAGGAGGAG GTCTGGTCTCTTGAAGCGGGAGACCCTG GAGTGGATGGAG	1154
AAMP	NM_001087.3	AAMP	AAACACCA CTGGAGCAT TGA	3	CAAGCCTGG AACCTATAG CC	387	CTCGCCAAATGATGCTGCTCAAGTT	771	77	AAACACCACTGGAGCAATTGACTACCAG GCTCGCCAAATGATGTGCTCAAGTTAA AGGGCTATAGGTTCCAGGCTTG	1155
			ACCACTGCC ACAATGCA G	4	ATAGCGCTG ACCACTGCC	388	CCATGAGCTGTAGCCGATGTCCA	772	68	ACCAGTCCACAATGACAGTGGCTGGAC ATTCTGGCTACAGCTCATGGGGGGGCA GTGGTCAGCGCTAT	1156
ABCC10	NM_033450.2	ABCC10	TGCAGACTG TACCATGCT GA	5	GGCCAGCAC CATAATCCT AT	389	CTGCACACGGTTCTAGGCTCCG	773	76	TGCAGACTGTACCATGTGACCATTTGC CCATCGCCTGCACACGCTTCTAGGCTCC GATAGGATTATGGTGTGGCC	1157
			ACACGTCTG TCACCATGG AA	6	ACTAGGGTG CTCCGAGTG AC	390	TCCTGCTCTACAAAGCCCATTGACCG	774	67	ACACGTCTGTACCATTGAAGCTCTGC TCTACAAAGCCCATTTGACCGGGTCACTC GGAGCACCCCTAGT	1158
ABR	NM_001092.3	ABR	ATCCGCATT GAAGACCC A	7	ATCCGCTAG AACTGCACC AC	391	CCCCGAGAAAGCACATGGTATTCC	775	66	ATCCGCATTGAAGAACCCACCCCGCAGA AAGCACATGTATTTCTGGGTGCTGCA GTTCTAGCGGAT	1159
			GACTGTCTC GTTTCCCTG GT	8	TGGGCTTAG ATGCTTGAC TC	392	CTCTGTCAACCAATGTGGACCTGCC	776	74	GACTGTCTCGTTTCCCTGGTGACCTCTG TCACCAATGTGGACCTGCCCCCTAAAG AGTCAAGCATCTAAGCCCA	1160
ACVR2B	NM_001106.2	ACVR2B	TCAAAAAGT ACGGACAC CTCCT	9	TGCAAAATGC TTTGATGGA AT	393	TGTAGGTATCTCTTAGTCCGCCATCTGA	777	74	TCAAAAGTACGGACACCTCTCTGTGACA TGGCGGGACTAAGAGATACCTACAAGG ATTCCATCAAAAGCATTTGCA	1161
			GAGCATGC GTCTACTGC CT	10	CTGGTCACG GTCTCCATG T	394	CTGACACTCATCTGAGCCCTCCCA	778	66	GAGCATGCGTCTACTGCCTCACTGACA CTCATCTGAGCCCTCCCATGACATGGA GACCGTGACCAG	1162
ADAM12	NM_021641.2	ADAM12	GAAGTGCC AGGAGGCG ATTA	11	CGGGCACTC ACTGCTATT ACC	395	TGCTACTTGCAAAGGCGTGTCCTACTGC	779	73	GAAGTGCCAGGAGGCGGATTAAATGCTAC TTGCAAAGGCGTGTCTACTGCACAGG TAATAGCAGTGAGTGCCCG	1163
			CAAGGCC CATCTGAAT CA	12	ACCCAGAAT CCAACAGTG CAA	396	CTGCGCTGGATGGACACCGC	780	62	CAAGGCCCATCTGAATCAGCTGCGCT GGATGGACACCGCCTTGCACTGTTGGA TTCTGGGT	1164
ADAM23	NM_003812.1	ADAM23	GCGAGTTCA AAGTGTTCG AG	13	CACAGATGG CCAGTGTIT CT	397	CACACAGGGTGCCATCAATCACCT	781	72	GCGAGTTCAAAAGTGTTCGAGGCCAAAGG TGATTGATGGCACCCCTGTGTGGGCCAG AAACACTGGCCATCTGTG	1165
			TAAGCCAC AAGCACAC GG	14	TGGGCGCCT AAATCCTAA	398	CGAGTGGAAAGTGCTCCC CACTTTC	782	75	TAAGCCACAAGCACACGGGGGCTCCAGC CCCCCGAGTGGAAAGTGCTCCCCACTTT CTTTAGGATTTAGGGCGCCCA	1166

TABLE A

AES	NM_001130.4	AES		ACGAGATG TCCTACGGC TTGA	15	GGGCACAAA TCCCGTTCA G	399	CGATCTCAGCCTGTTTGT GCATCTCGAT	783	78	1167	ACGAGATGTCCTACGGCTTGAAACATCG AGATGCACAAACACAGGCTGAGATCGTCA AAAGGTGAACGGGATTTGTGCC
AGR2	NM_006408.2	AGR2		AGCCAAACA TGTGACTAA TTGGA	16	TCTGATCTC CATCTGCCT CA	400	CAACACGTCACCAACCCT TTGCTCT	784	70	1168	AGCCAAACATGTGACTAAATTGGAAGAAG AGCAAAGGTGTGTGACGTGTGTGATGAG GCAGATGGAGATCAGA
AK055699	NM_194317	LYPD6		CTGCATGTG ATTGAATAA GAAACAAG A	17	TGTGGACCT GATCCCTGT ACAC	401	TGACCACACCAAAAGCCT CCCTGG	785	78	1169	CTGCATGTGATTGAATAAGAAAAAAGA AAGTGACCACACCAAAAGCCTCCCTGGC TGGTGTACAGGGATCAGGTCCACA
AKR7A3	NM_012067.2	AKR7A3		GTGGAAAC GGAGCTCTT CC	18	CCAGAGGGT TGAAGGCAT AG	402	ACCTCAGTCCAAAGTGC CTGAGGC	786	67	1170	GTGGAAACGGGAGCTCTTCCCTGCCTC AGGCACCTTTGGGACTGAGGTTCTATGCCT TCAACCCCTCTGG
AKT3	NM_005465.1	AKT3		TTGTCTCTG CCTTGGACT ATCTACA	19	CCAGCATTA GATTCTCCA ACTTGA	403	TCACGGTACACAATCTTT CCGGA	787	75	1171	TTGTCTCTGCCTTGGACTATCTACATTTC CGGAAAGATTGTGTACCCGTGATCTCAA GTTGGAGAAATCTAATGCTGG
ALCAM	NM_001627.1	ALCAM		GAGGAATA TGGAATCCA AGGG	20	GTGGCGGAG ATCAAGAGG	404	CCAGTTCCTGCCGTCTGC TCTTCT	788	66	1172	GAGGAATATGGAAATCCAAGGGGGCCA GTTCCCTGCCGTCTGTCTTCTTGCCTCTT GATCTCCGCCAC
ALDH4	NM_003748.2	ALDH4A1		GGACAGGG TAAGACCGT GAT	21	AACCGGAAG AAGTCGATG AG	405	CTGCAGCGTCAATCTCC GCTTG	789	68	1173	GGACAGGTTAAGACCCGTGATCCAAGCG GAGATTGACGCTGCAGCGGAACTCATC GACTTCTTCCGGTT
ANGPT2	NM_001147.1	ANGPT2		CCGTGAAA GCTGCTCTG TAA	22	TTGCAGTGG GAAGAACAG TC	406	AAGCTGACACAGCCCCC CCAAGTG	790	69	1174	CCGTGAAAAGCTGCTCTGTAAAAGCTGA CACAGCCCTCCCCAAGTGAGCAGGACTG TTCTTCCCACTGCAA
ANXA2	NM_004039.1	ANXA2		CAAGACAC TAAGGGCG ACTACCA	23	CGTGTGGG CTTCAGTCA T	407	CCACCACACAGGTACAG CAGCGCT	791	71	1175	CAAGACACTAAGGGCGACTACCAGAAA GCGCTGCTGTACCTGTGTGGTGGAGAT GACTGAAGCCCCGACACG
AP-1 (JUN official)	NM_002228.2	JUN		GACTGCAA AGATGGAA ACGA	24	TAGCCATAA GGTCCGCTC TC	408	CTATGACGATGCCCTCA ACGCCTC	792	81	1176	GACTGCAAAGATGGAAACGACCTTCTA TGACGATGCCCTCAACGCCCTCGTCCCTC CCGTCCGAGAGCGGACCTTATGGCTA
APEX-1	NM_001641.2	APEX1		GATGAAGC CTTTCGCA GTT	25	AGGTCTCCA CACAGCAC AG	409	CTTTCGGGAAGCCAGGC CCTT	793	68	1177	GATGAAGCCTTTTCGCAAGTTCCTGAAG GGCCTGGCTTCCCGAAAAGCCCCTTGTG CTGTGTGGAGACCT
APOD	NM_001647.1	APOD		GTTTATGCC ATCGGCACC	26	GGAATACAC GAGGGCATA GTTT	410	ACTGGATCCTGGCCACC GACTATG	794	67	1178	GTTTATGCCATCGGCACCGTACTGGATC CTGGCCACCAGCTATGAGAACTATGCC CTCGTGTATTC
ARF1	NM_001658.2	ARF1		CAGTAGAG ATCCCCGA ACT	27	ACAAGCACA TGGCTATGG AA	411	CTTGTCTTGGGTCACCC TGCA	795	64	1179	CAGTAGAGATCCCCGCAACTCGCTTGT CCTTGGGTACACCTGCAITCCCATAGCCA TGTGCTTGT
ARHI	NM_004675.1	DIRAS3		ATCAGAGA TTACCGCGT CGT	28	ACTTGTGCA GCAGCGTAC TT	412	ACACCAGCGGTGCCGAC TACC	796	67	1180	ATCAGAGATTACCGCGTCTGTGTTAGTC GGCACCGCTGTTGTGGGAAAAGTACG CTGCTGCACAAAGT

TABLE A

ARNT2	NM_014862.3	ARNT2	GACTGGGTCAGTGATGG	29	GGAGTGACG CATGGACAG A	413	CTAGAGCCATCCTTGGC CATCCTG	797	68	GACTGGGTCAGTGATGGCAACAGGATG GCCAAGGATGGCTCTAGAACACTCTGT CCATGCGTCACTCC	1181
ARSD	NM_001669.1	ARSD	TCCCTGAGA ACGAAACC ACT	30	TGGTGCCAT TTTCCTATG AG	414	CAAGAAATCTTGCAGCAG CATGGCT	798	79	TCCCTGAGAACGAAACCACACTTTTGCAA GAATCTTGCAGCAGCATGGCTATGCAA CCGGCCTCATAGGAAATGGCACCA	1182
AURKB	NM_004217.1	AURKB	AGCTGCAG AAGAGCTG CACAT	31	GCATCTGCC AACTCCTCC AT	415	TGACGAGCAGCGAACAG CCACG	799	67	AGCTGCAGAAAGAGCTGCACATTTGACG AGCAGCGAACAGCCACGATCATGGAGG AGTTGGCAGATGC	1183
B-actin	NM_001101.2	ACTB	CAGCAGAT GTGGATCA GCAAG	32	GCATTTGCG GTGGACGAT	416	AGGAGTATGACGAGTCC GGCCCC	800	66	CAGCAGATGTGGATCAGCAAGCAGGAG TATGACGAGTCCGGCCCCCTCCATCGTCC ACCGCAAAATGC	1184
B-Catenin	NM_001904.1	CTNNB1	GGCTCTTGT GCGTACTGT CCTT	33	TCAGATGAC GAAAGGCAC AGATG	417	AGGCTCAGTGATGTC TTC CCTGTCACCAG	801	80	GGCTCTTGTGCGTACTGTCTTCGGGCT GGTGACAGGGAAAGACATCACTGAGCCT GCCATCTGTGCTCTTCGTCATCTGA	1185
BAD	NM_032989.1	BAD	GGGTCAGG TGCCTCGAG AT	34	CTGCTCACT CGGCTCAAA CTC	418	TGGGCCCCAGAGCATGTT CCAGATC	802	73	GGGTCAGGTGCCTCGAGATCGGGCTTG GGCCAGAGCATGTTCCAGATCCCAGA GTTTGAGCCCGAGTGAGCAG	1186
BAG1	NM_004323.2	BAG1	CGTGTGTCAG CACTTGGAA TACAA	35	GTTCAAACCT CTTCCTGTG GACTGT	419	CCCAATTAAACATGACCC GGCAACCAT	803	81	CGTGTGTCAGCACCTTGGAAATACAAAGATG GTTGCCGGGTTCATGTTAATTGGGAAAA AGAACAGTCCACAGGAAGAGGTTGAAC	1187
BAG4	NM_004874.2	BAG4	CCTACGGCC GCTACTACG	36	GGCGGAAGA GGATATAAG GG	420	AGATGTGCCCGGTACAAC CACCTC	804	76	CCTACGGCCGCTACTACGGGGCTTGGGG GTGGAGATGTGCCGGTACACCCACCTC CACCCCTATATCCTCTTCGCC	1188
BASE	NM_173859.1		GACTCCTCA GGGCAGAC TTTCT	37	CGAAGGCAC TACTCAATG GTTTC	421	CCAGCCTGCAGACAACT GGCCTC	805	72	GACTCCTCAGGGCAGACTTTCCTCCCG CCTGCAGACAACTGGCCTCCAGAAAAACC ATTGAGTAGTGCCCTCG	1189
Bax	NM_004324.1	BAX	CCGCCGTGG ACACAGAC T	38	TTGCCGTCA GAAAACATG TCA	422	TGCCACTCGGAAAAAGA CCTCTCGG	806	70	CCGCCGTGGACACAGACTCCCCCGGAG AGGTCTTTTCCGAGTGGCAGCTGACAT GTTTTCTGACGGCAA	1190
BBC3	NM_014417.1	BBC3	CCTGGAGG GTCCCTGTAC AAT	39	CTAATTGGG CTCCATCTC G	423	CATCATGGGACTCCTGC CCTTACC	807	83	CTTGGAGGTCCTGTACAATCTCATCAT GGGACTCCTGTGCCCTTACCCAGGGGCCA CAGAGCCCCCGAGATGGAGCCCCAATTA G	1191
BCAR1	NM_014567.1	BCAR1	ACTGACAA GACCAGCA GCAT	40	TCCTGGGAG GTGAACCTTA GG	424	AGTCACGACCCCTGCC TCAC	808	65	ACTGACAAAGACCAGCAGCATCCAGTCA CGACCCCTGCCCTCACCCCTTAAGTTCA CCTCCCAGGA	1192
BCAR3	NM_003567.1	BCAR3	TGACTTCCT AGTTCGTGA CTCTCTGT	41	TGAGCGAGG TTCTTCCACT GA	425	CAGCCCTGGGAACTTTG TCCTGACC	809	75	TGACTTCCTAGTTCTGTGACTCTCTGTCC AGCCCTGGGAACCTTGTGCCCTGACCTGIC AGTGGAAAGAACCTCGCTCA	1193
BCAS1	NM_003657.1	BCAS1	CCCCGAGA CAACGGAG ATAA	42	CTCGGGTTT GGCCTCTTT C	426	CTTTCGGTTGGCAJCCGC AACAG	810	73	CCCCGAGACAACGGAGATAAGTGTGT TGC GGATGCCAACGGAAAAAGAAATCTTGG GAAAGAGGCCAAACCCGAG	1194

TABLE A

Bcl2	NM_000633.1	BCL2	CAGATGGA CCTAGTACC CACTGAGA	43	CCTATGATT TAAGGGCAT TTTTCC	427	TTCCACGCCGAAGGACA GCGAT	811	73	CAGATGGACCTAGTACCCACTGAGATT TCCACGCCGAAGGACAGCGAATGGGAAA AATGCCCTTAAATCATAGG	1195
BCL2L12	NM_138639.1	BCL2L12	AACCCACCC CTGTCTTGG	44	CTCAGCTGA CGGAAAAGG	428	TCCGGGTAGCTCTCAA A CTCGAGG	812	73	AACCCACCCCTGTCTTGGAGCTCCGG TAGCTCTCAA A ACTCGAGGCTGCGCACC CCCTTTCCCGTCAGCTGAG	1196
BGN	NM_001711.3	BGN	GAGCTCCGC AAGGATGA C	45	CTTGTGTTC ACCAGGACG A	429	CAAGGGTCTCCAGCACC TCTACGC	813	66	GAGTCCGC AAGGATGACTTCAAGGGT CTCCAGCACCTCTACGCCCTCGTCTCTGG TGAACAACAAG	1197
BIK	NM_001197.3	BIK	ATTCTATG GCTCTGCAA TTGTC	46	GGCAGGAGT GAATGGCTC TTC	430	CCGGTTAACTGTGGCCT GTGCC	814	70	ATTCTATGGCTCTGCAATTTGTACCCG TTAACTGTGGCCTGTGCCAGGAAGAG CCATTCACTCCTGCC	1198
BNIP3	NM_004052.2	BNIP3	CTGGACGG AGTAGCTCC AAG	47	GGTATCTTG TGGTGTCTG CG	431	CTCTCACTGTGACAGCCC ACCTCG	815	68	CTGGACGGAGTAGCTCCCAAGAGCTCTC ACTGTGACAGCCCACCTCGCTCGCAGA CACCACAAGATAACC	1199
BSG	NM_001728.2	BSG	AATTTATG AGGGCCAC GG	48	GTGGCCAAG AGGTCAGAG TC	432	CTGTGTTCCGACTCAGCCT CAGGGA	816	66	AATTTATGAGGGCCACGGGTCTGTGT CGACTCAGCCTCAGGGACGACTCTGAC CTCTTGGCCAC	1200
BTRC	NM_033637.2	BTRC	GTGGGAC ACAGTTGGT CTG	49	TGAAGCAGT CAGTTGTGC TG	433	CAGTCGGCCCCAGGACGG TCTACT	817	63	GTGGGACACAGTGGTCTGCAGTCCG CCCAGGACGGTCTACTCAGCACAACTG ACTGCTTCA	1201
BUB1	NM_004336.1	BUB1	CCGAGGTTA ATCCAGCAC GTA	50	AAGACATGG CGCTCTCAG TTC	434	TGCTGGGAGCCTACACT TGGCCC	818	68	CCGAGGTTAATCCAGCACGTATGGGGC CAAGTGTAGGCTCCCAGCAGGAACCTGA GAGCGCCATGTCTT	1202
BUB1B	NM_001211.3	BUB1B	TCAACAGA AGGCTGAA CCACTAGA	51	CAACAGAGT TTGCCGAGA CACT	435	TACAGTCCCAGCACCCGA CAATTCC	819	82	TCAACAGAAGGCTGAACCACTAGAAAAG ACTACAGTCCCAGCACCCGACAAATCCA AGCTCGAGTGTCTCGGCAAACTCTGTT G	1203
BUB3	NM_004725.1	BUB3	CTGAAGCA GATGGTTCA TCAAT	52	GCTGATTCC CAAGAGTCT AACC	436	CCTCGCTTTGTTTAACAG CCCAGG	820	73	CTGAAGCAGATGGTTTCATCATTTCTCTGG GCTGTTAAACAAAGCGAGGTTAAGGTT AGACTCTTGGGAATCAGC	1204
c-kit	NM_000222.1	KIT	GAGGCAAC TGCTTATGG CTTAATTA	53	GGCACTCGG CTTGAGCAT	437	TTACAGCGACAGTCATG GCCGCAT	821	75	GAGGCAACTGCTTATGGCTTAATTAAG TCAGATCGGGCCATGACTGTGCTGTGA AAGATGCTCAAGCCGAGTGCC	1205
C10orf116	NM_006829.2	C10orf116	CAAGAGCA GAGCCACC GT	54	TGAGACCGT TGGATTGGA TT	438	CCGGAGTCCTAGCCTCC CAAATTC	822	67	CAAGAGCAGAGCCACCCTAGCCCGGAGT CCTAGCCTCCCAAATTCGGAAATCCAA TCCAACGGTCTCA	1206
C17orf37	NM_032339.3	C17orf37	GTGACTGCA CAGGACTCT GG	55	AGGACCAAA GGGAGACCA A	439	CCTGCTCTGTCTGGGGT CCAAAC	823	67	GTGACTGCACAGGACTCTGGGTTCCTG CTCTGTCTGGGGTCCAAACCTTGGTCT CCCTTTGGTCTCT	1207
C20 orf1	NM_012112	TPX2	TCAGCTGTG AGCTGCGG ATA	56	ACGGTCCTA GGTTTGAGG TTAAGA	440	CAGGTCUCCATTCGCGGG CG	824	65	TCAGCTGTGAGCTGCGGATACCGCCCG GCAATGGGACCTGCTCTTTAACCTCAAA CCTAGGACCGT	1208

TABLE A

C6orf66	NM_014165.1	NDUFAF4	GCGGTATCA GGAATTTC ACCT	57	GCGACAGAG GGCTTCATC TT	441	TGATTTCCCGTTCCGCTC GGTTCT	825	70	GCGGTATCAGGAATTTCAACCTAGAGA ACCGAGCGGAACGGGAAATCAGCAAG ATGAAGCCCTCTGTGCG	1209
C8orf4	NM_020130.2	C8orf4	CTACGAGTCA AGCCCATCC AT	58	TGCCCCACGG CTTTCTTAC	442	CATGGCTACCACTTCGA CACAGCC	826	67	CTACGAGTCAAGCCCATCCATCCATGGC TACCACCTTCGACACAGCCTCTCTGTAAG AAAGCCGTGGGCA	1210
CACNA2D 2	NM_006030.1	CACNA2D 2	TGATGCTGC AGAGAACT TCC	59	CACGATGTC TTCTCTCTTG A	443	AAAGCACACCGCTGGCA GGAC	827	67	TGATGCTGCAGAGAACTTCCAGAAAAGC ACACCGCTGGCAGGACAACAATCAAGGA GGAAGACATCGTG	1211
CAT	NM_001752.1	CAT	ATCCATTTCG ATCTCACCA AGGT	60	TCCGGTTTA AGACCAGTT TACCA	444	TGGCCTCACAAAGGACTA CCCTCTCATCC	828	78	ATCCATTTCGATCTCACCAAGGTTTGGCC TCACAAGGACTACCCCTCTCATCCAGTT GGTAAACTGGTCTTAAACCGGA	1212
CAV1	NM_001753.3	CAV1	GTGGCTCAA CATTGTGTT CC	61	CAATGGCCT CCATTTIAC AG	445	ATTTCAGCTGATCAGTIG GGCCTCC	829	74	GTGGCTCAAACATTTGTGTCCCATTTTCAG CTGATCAGTGGGCCTCCAAAGGAGGGGC TGTAATAATGGAGGGCCATTG	1213
CBX5	NM_012117.1	CBX5	AGGGGATG GTCTCTGTC ATT	62	AAAGGGGTG GGTAGAAAAG GA	446	CATAATACATTCACCTCC CTGCCTCCTC	830	78	AGGGGATGGTCTCTGTGTTCATTTCTCTTTG TACATAATAACATTCACCTCCCTGCCTCC TCTCCTTTCTACCAACCCCTTT	1214
CCL19	NM_006274.2	CCL19	GAACGCAT CATCCAGA GACTG	63	CCCTGCAC GGTCATAGG TT	447	CGCTTCATCTTGGGTGAG GTCTC	831	78	GAACGCATCATCCAGAGACTGCAGAGG ACCTCAGCCAAAGATGAAGGCCGCAGC AGTTAACCTATGACCCGTGCAGAGG	1215
CCL3	NM_002983.1	CCL3	AGCAGACA GTGGTCAGT CCTT	64	CTGCATGAT TCTGAGCAG GT	448	CTCTGTGACACTCGAG CCCACAT	832	77	AGCAGACAGTGTCTCAGTCTCTTCTTGG CTCTGTGACACTCGAGCCCCACATTCGG TCACCTGTCTAGAATCATGCAG	1216
CCL5	NM_002985.2	CCL5	AGGTTCTGA GCTCTGGCT TT	65	ATGCTGACT TCCTTCTCG GT	449	ACAGAGCCCTGGCAAAG CCAAG	833	65	AGGTTCTGAGCTCTGGCTTGGCCTTGGC TTTGCCAGGGCTCTGTGACCCAGGAAGG AAGTCAGCAT	1217
CCNB1	NM_031966.1	CCNB1	TTCAGGTTG TTGCAGGA GAC	66	CATCTTCTTG GGCACACAA T	450	TGTTCCATTTATTTGATCG GTTCAIGCA	834	84	TTCAGGTTGTTGCAGGAGACCATGTAC ATGACTGTCTCCATTATTGATCGGTTCA TGCAGAAATAATTGTGTGCCCAAGAAGA TG	1218
CCND3	NM_001760.2	CCND3	CCTCTGTGC TACAGATTA TACCTTTGC	67	CACTGCAGC CCCAATGCT	451	TACCCGCCATCCATGATC GCCA	835	76	CCCTGTGCTACAGATTATACCTTTGCC ATGTACCCGCCATCCATGATCGCCACG GGCAGCATTTGGGGCTGCAGTG	1219
CCNE2 variant 1	NM_057749var 1	CCNE2	GGTCACCA AGAAACAT CAGTATGA A	68	TTCAATGAT AATGCAAGG ACTGATC	452	CCCAGATAATACAGGTG GCCAACAAATTCCT	836	85	GGTCACCAAGAAACATCAGTATGAAAT TAGGAATTGTTGGCCACCTGTATATCT GGGGGATCAGTCCCTTGCAATTATCAIT GAA	1220
CCR5	NM_000579.1	CCR5	CAGACTGA ATGGGGGT GG	69	CTGGTTTGT CTGGAGAAG GC	453	TGGAATAAGTACCTAAG GCGCCCCC	837	67	CAGACTGAATGGGGTGGGGGGGCG CCTTAGGTACTTATTCCAGATGCCTTCT CCAGACAAACCAG	1221
CCR7	NM_001838.2	CCR7	GGATGACA TGCACCTCAG CTC	70	CCTGACATT TCCCTTGTCC T	454	CTCCCATCCCAGTGGAG CCAA	838	64	GGATGACATGCACCTCAGCTCTTGGCTC CACTGGGATGGGAGGAGAGACAAAGG GAAATGTACAG	1222

TABLE A

CD1A	NM_001763.1	CD1A	GGAGTGGA AGGAACTG GAAA	71	TCATGGGCG TAICTACGA AT	455	CGCACCAATTCGGTCATTT GAGG	839	78	GGAGTGGAAGGAACTGGAAACATTATT CCGTATACGCACCATTCGGTCATTTGAG GGAATTCGTAGATACGCCCATGA	1223
CD24	NM_013230.1	CD24	TCCAACATA TGCCACCAC CAA	72	GAGAGAGTG AGACCACGA AGAGACT	456	CTGTTGACTGCAGGGCA CCACCA	840	77	TCCAACATAATGCCACCACCAAGGCGGC TGGTGGTGCCCTGCAGTCAACAGCCAG TCTCTTCGTGGTCTCAGCTCTCTC	1224
CD4	NM_000616.2	CD4	GTGCTGGA GTCGGGACT AAC	73	TCCCTGCAT TCAAGAGGC	457	CAGGTCCCCTTGTCCCAA GTTCCAC	841	67	GTGCTGGAGTCGGGACTAACCCAGGTC CCTTGTCCCAAAGTTCCACTGCTGCCTCT TGAATGCAGGGA	1225
CD44E	X55150		ATCACCGAC AGCACAGA CA	74	ACCTGTGTT TGGATTGTC AG	458	CCCTGCTACCAATATGG ACTCCAGTCA	842	90	ATCACCGACAGCACAGACAGAAATCCCT GCTACCAATATGGACTCCAGTCATAGT ACAACGCTTCAGCCTACTGCAAAATCCA AACACAGT	1226
CD44s	M59040.1		GACGAAGA CAGTCCCTG GAT	75	ACTGGGGTG GAATGTGTC 'T'	459	CACCGACAGCACAGACA GAATCCC	843	78	GACGAAGACAGTCCCTGGATCACCCGAC AGCACAGACAGAAATCCCTGCTACCAGA GACCAAGACACATTCACACCCAGT	1227
CD44v6	AJ251595v6		CTCATACCA GCCATCCAA TG	76	TTGGGTGGA AGAAATCAG TCC	460	CACCAAGCCCCAGAGGAC AGTTCTT	844	78	CTCATACCAAGCCATCCAAATGCAAGGAA GGACAACACCAAGCCCCAGAGGACAGTT CCTGGACTGATTTCTTCAACCCAA	1228
CD68	NM_001251.1	CD68	TGGTTCCCA GCCCCTGTGT	77	CTCCTCCAC CCTGGGTTG T	461	CTCCAAGCCCCAGATTCA GATTCGAGTCA	845	74	TGTTTCCCAGCCCTGTGTCCACCTCCAA GCCCAGATTTCAGATTTCGAGTCATGTAC ACAACCCAGGGTGGAGGAG	1229
CD82	NM_002231.2	CD82	GTGCAGGCT CAGGTGAA G'IG	78	GACCTCAGG GCGATTTCAT GA	462	TCAGCTTCTACAACTGG ACAGACAACGCTG	846	84	GTGCAGGCTCAGGTGAAAGTGCIGCGGC TGGTCAAGCTTCTACAACTGGACAGAC AACGCTGAGCTCATGAATCGCCCTGAG G'IC	1230
CDC20	NM_001255.1	CDC20	TGGATTGGA GTTCTGGGA ATG	79	GCTTGCACT CCACAGGTA CACA	463	ACTGGCCGTGGCACTGG ACAACA	847	68	TGGATTGGAGTTCTGGGAAATGTACTGG CCGTGGCACTGGACAACAGTGTGTACC TGTGGAGTGCAAGC	1231
cdc25A	NM_001789.1	CDC25A	TCTTGCTGG CTACGCCCTC TT	80	CTGCATTGT GGCACAGTT CTG	464	TGTCCCTGTTAGACGTCC TCCGTCCATA	848	71	TCTTGCTGGCTACGCCCTCTTCTGTCCCT GTTAGACGTCTCCTCCGTCCATATCAGAA CTGTGCCACAATGCAG	1232
CDC25C	NM_001790.2	CDC25C	GGTGAGCA GAAGTGGC CTAT	81	CTTCAGTCTT GGCCTGTTC A	465	CTCCCCGTGATGCCAG AGAACT	849	67	GGTGAGCAGAAAGTGGCCTATATCGCTC CCCCGTGATGCCAGAGAACTTGAACAG GCCAAGACTGAAG	1233
CDC4	NM_018315.2	FBXW7	GCAGTCCGC TGTGTTCAA	82	GGATCCCAC ACCTTTACC ATAA	466	TGCTCCACTAACAAACCCT CCTGCC	850	77	GCACTCCGCTGTGTTCAATATGATGGC AGGAGGCTTGTTAGTGGAGCATATGAT TTTATGGTAAAGGTGTGGGATCC	1234
CDC42BP A	NM_003607.2	CDC42BPA	GAGCT'GAA AGACGCAC ACTG	83	GCCGCTCAT TGATCTCCA	467	AATTCTGTGCATGCCAG TTTCCTC	851	67	GAGCTGAAAGACGCACACTGTCAGAGG AAACTGGCCATGCAGGAATTCATGGAG ATCAATGAGCGGC	1235
CDC42EP4	NM_012121.4	CDC42EP4	CGGAGAAG GGCACCCAG TA	84	CCGTCAATTG GCCTTCTTC	468	CTGCCCAAGAGCCTGTGC ATCCAG	852	67	CGGAGAAAGGGCACCCAGTAAAGCTGCCCA AGAGCCTGTTCATCCAGCCCCCGTGAAGA AGGCCAAATGACGG	1236

TABLE A

CDH11	NM_001797.2	CDH11	GTGGCAG AAGCAGGA CT	85	CTACTCATG GGCGGATG	469	CCTTCTGCCCATAGTGAT CAGCGA	853	70	GTGGCAGAAAGCAGGACTTGTACCTTC TGCCCATAGTGATCAGCGATGGCGGCA TCCCGCCCATGAGTAG	1237
CDH3	NM_001793.3	CDH3	ACCCATGTA CCGTCTCTCG	86	CCGCCTTCA GGTTCTCAA T	470	CCAACCCAGATGAAATC GGCAACT	854	71	ACCCATGTACCGTCTCTGGCCAGCCAA CCCAGATGAAATCGGCAACTTTTATAAT TGAGAACTGAAGGCGG	1238
CDK4	NM_000075.2	CDK4	CTTCCCAT CAGCACAG TTC	87	TTGGGATGC TCAAAAGCC	471	CCAGTCGCCCTCAGTAAA GCCACCT	855	66	CCTTCCCATCAGCACAGTTTCGTGAGGT GGCTTTACTGAGGCGACTGGAGGCTTT TGAGCATCCCCAA	1239
CDK5	NM_004935.2	CDK5	AAGCCCTAT CCGATGTAC CC	88	CTGTGGCAT TGAGTTTGG G	472	CACAACATCCCTGGTGA ACGTCTG	856	67	AAGCCCTATCCGATGTACCCGGCCACA ACATCCCTGGTGAACGTCTGTGCCAAA CTCAATGCCACAG	1240
CDKN3	NM_005192.2	CDKN3	TGGATCTCT ACCAGCAA TGTG	89	ATGTCAGGA GTCCCTCCA TC	473	ATCACCCATCATCATCCA ATCGCA	857	70	TGGATCTCTACCCAGCAAATGTGGAATTA TCACCCATCATCATCCAAATCGCAGATG GAGGGACTCCTGACAT	1241
CEACAM1	NM_001712.2	CEACAM1	ACTTGCTG TTCAGAGCA CTCA	90	TGGCAAATC CGAATTAGA GTGA	474	TCCTTCCCCACCCCCAGTC CTGTC	858	71	ACTTGCTGTTCAGAGCACTCATTCCTT CCCCCCCCAGTCTCTGTCTATCACTCT AATTCGGATTTCGCA	1242
CEBPA	NM_004364.2	CEBPA	TGGGTTTG CTCGGATAC TTG	91	GTCTCAGAC CCTTCCCCC	475	AAATGAGACTCTCCCGT CGGCAGC	859	66	TTGGTTTGTCTGGATACTTGCCAAAAAT GAGACTCTCCGTCTGGCAGCTGGGGAA GGGTCTGAGAC	1243
CEGP1	NM_020974.1	SCUBE2	TGACAAATCA GCACACCTG CAT	92	TGTGACTAC AGCCGTGAT CCTTA	476	CAGGCCCTCTTCCGAGC GGT	860	77	TGACAAATCAGCACACCTGCATTCACCG CTCGGAAAGAGGGCTTGAGCTGCATGAA TAAGGATCACGGCTGTAGTCACA	1244
CENPA	NM_001809.2	CENPA	TAAATTAC TCGTGGTGT GGA	93	GCCTCTGT AGGGCCAAT AG	477	CTTCAATTGGCAAGCCC AGGC	861	63	TAAATTCACTCTGTGTGTGGACTTCAAT TGGCAAGCCCCAGGCCCTATTGGCCCTA CAAGAGGC	1245
CGA (CHGA official)	NM_001275.2	CHGA	CTGAAGGA GCTCCAAG ACCT	94	CAAAACCGC TGTGTTCTT C	478	TGCTGATGTGCCCTCTCC TTGG	862	76	CTGAAGGAGCTCCAAGACCTCGCTCTC CAAGGCGCCAAGGAGAGGGCACATCA GCAGAAAGAAACACAGCGGTTTTG	1246
CGalpha	NM_000735.2	CGA	CCAGAATG CACGTACA GGA	95	GCCCATGCA CTGAAGTAT TGG	479	ACCCATTCTTCTCCCCAGC CGGG	863	69	CCAGAATGCACGCTACAGGAAACCCCA TTCTTCTCCCCAGCCGGGTGCCCCCAATAC TTCAGTGCATGGGC	1247
CGB	NM_000737.2	CGB	CCACCATAG GCAGAGGC A	96	AGTCGTGCA GTGCTAGGG AC	480	ACACCCTACTCCCTGTGTC CTCCAG	864	80	CCACCATAGGCAGAGGCAGGCCTTCCT ACACCCTACTCCCTGTGCTCCAGCCTC GACTAGTCCCTAGCACTCGACGACT	1248
CHAF1B	NM_005441.1	CHAF1B	GAGGCCAG TGGTGGAA ACAG	97	TCCGAGGCC ACAGCAAAC	481	AGCTGATGAGTCTGCC TACCGCCTG	865	72	GAGGCCAGTGGTGGAAACAGGTGTGGA GCTGATGAGTCTGCCCTACCCGCTGGT GTTTGCCTGTGCCCTCGGA	1249
CHFR	NM_018223.1	CHFR	AAGGAAGT GGTCCCTCT GTG	98	GACGCAGTC TTTCTGTCTG G	482	TGAAGTCTCCAGCTTTGC CTCAGC	866	76	AAGGAAGTGGTCCCTCTCTGTGGCAAGTG ATGAAGTCTCCAGCTTTGCCCTCAGCTCT CCCAGACAGAAAGACTGCGTC	1250

TABLE A

CHI3L1	NM_001276.1	CHI3L1	AGAATGGG TGTGAAG CG	99	TGCAGAGCA GCACTGGAG	483	CACCAGGACCACAAAGC CTGTTTG	867	66	AGAATGGGTGTGAAGGCGTCTCAAAACA GGCTTTGTGGTCTGTGTCTGCTCCAGT GCTGCTCTGCA	1251
CKS2	NM_001827.1	CKS2	GGCTGGAC GTGGTTTIG TCT	100	CGCTGCAGA AAATGAAAC GA	484		868	62	GGCTGGACGTGGTTTTTGTCTGTGCGCC CGCTCTTCGGGCTCTCGTTTCATTTTCT GCAGCG	1252
Claudin 4	NM_001305.2	CLDN4	GGCTGCTTT GCTGCAACT G	101	CAGAGCGGG CAGCAGAAT A	485	CGCACAGACAAAGCCTTA CTCCGCC	869	72	GGCTGCTTTGTGTGCAACTGTCCACCCCG CACAGACAAGCCTTACTCCGCCAAGTA TTCTGTGCCCCGCTCTG	1253
CLIC1	NM_001288.3	CLIC1	CGGTACTTG AGCAATGC CTA	102	TCGATCTCC TCATCATCT GG	486	CGGGAAGAATTGCGTTT CACCTG	870	68	CGGTACTTGAGCAATGCCCTACGCCCGG GAAGAAATTCGCTTCCACCTGTCCAGAT GATGAGGAGATCGA	1254
CIU	NM_001831.1	CIU	CCCAGGAT ACCTACCAC TACCT	103	TGCGGGACT TGGGAAAGA	487	CCCTTCAGCCTGCCCCAC CG	871	76	CCCAGGATACCTACCACCTACCTGCCCC TCAGCCTGCCCCACCGGAGGCCCTCACT TCTTCTTTCCCAAGTCCCCGCA	1255
CNOT2	NM_014515.3	CNOT2	AAATCGCA GCTTATCAC AAGG	104	TGTTGGTAC CCCTGTGTGTT G	488	ACTCAGTTACCGAGCCA CGTCACG	872	67	AAATCGCAGCTTATCACAAAGGCACCTCA GTTACCGAGCCACGCTCACGCCCAACAAC AGGGGTACCAACA	1256
COL1A1	NM_000088.2	COL1A1	GTGGCCATC CAGCTGACC	105	CAGTGGTAG GTGATGTTT TGGGA	489	TCCTGCGCCTGATGTCCA CCG	873	68	GTGGCCATCCAGCTGACCTTCCCTGCGCC TGATGTCCACCCGAGGCCCTCCCAGAAACA TCACCTACCCTG	1257
COL1A2	NM_000089.2	COL1A2	CAGCCAAG AACTGGTAT AGGAGCT	106	AAACTGGCT GCCAGCAATT G	490	TCCTCCTAGCCAGACGTGT TTCCTGTCTTG	874	80	CAGCCAAGAACTGGTATAGGAGCTCCA AGGACAAAGAAACACGCTCTGGCTAGGAG AAACTATCAATGCTGGCAGCCAGTTT	1258
COMT	NM_000754.2	COMT	CTTTATCGG CTGGAACG AGTT	107	CTCCTTGGT GTCACCCAT GAG	491	CCTGCAGCCCCATCCACA ACCT	875	67	CTTTATCGGCTGGAAACGAGTTCATCCCTG CAGCCCCATCCACAACCTGCTCATGGGT GACACCAAGGAG	1259
Contig 51037			CGACAGTTG CGATGAAA GTTCTAA	108	GGCTGCTAG AGACCATGG ACAT	492	CCTCCTCCTGTGTGCTGCC ACTAATGCT	876	81	CGACAGTTGCGATGAAAGTTCTAATCT CTTCCTCCTCCTGTGTGTGCTGCCACTAAT GCTGATGTCCCATGGTCTCTAGCAGCC	1260
COPS3	NM_003653.2	COPS3	ATGCCCAGT GTTCTTGAC TT	109	CTCCCCATT ACAAGTGCT GA	493	CGAAACGCTATTCTCAC AGGTTACG	877	72	ATGCCCAGTGTTCCTGACTTCGAAAACG CTATTCTCACAGGTTTCAGCTCTTCATCA GCACTTGTAATGGGGAG	1261
CRYAB	NM_001885.1	CRYAB	GATGTGATT GAGGTGCA TGG	110	GAACTCCCT GGAGATGAA AOC	494	TGTTTCATCCTGGCGCTCT TCATGT	878	69	GATGTGATTGAGGTGCATGGAAAACAT GAAGAGCGCCAGGATGAACATGGTTTC ATCTCCAGGAGTTC	1262
CRYZ	NM_001889.2	CRYZ	AAGTCCTGA AATTGCGAT CA	111	CACATGCAT GGACCTTGA TT	495	CCGATTCCAAAAAGACCA TCAGGTTCT	879	78	AAGTCCTGAAATTGCCGATCAGATATTG CAGTACCGGATTCCAAAAGACCATCAGG TTCTAAATCAAGGTCCATGCAITG	1263
CSF1 isoC	NM_172211.1	CSF1	CAGCAAGA ACTGCAAC AACA	112	ATCCCTCGG ACTGCCICT	496	TTTGCTGAATGCTCCAGC CAAGG	880	68	CAGCAAGAAGCTGCAACAACAGCTTTGC TGAATGCTCCAGCCAAGGCCATGAGAG GCAGTCCGAGGGAT	1264

TABLE A

CSF1	NM_000757.3	CSF1	TGCAGCGG CTGATTGAC A	113	CAACTGTTT CTGGTCTAC AAACTCA	497	TCAGATGGAGACCTCGT GCCAAATTACA	881	74	TGCAGCGGCTGATTGACAGTCAGATGG AGACCTCTGTGCCAAATTACATTGAGTT TGTAGACCAGGAACAGTTG	1265
CSF1R	NM_005211.1	CSF1R	GAGCACAA CCAAACCTA CGA	114	CCTGCAGAG ATGGGTATG AA	498	AGCCACTCCCCACGCTG TTGT	882	80	GAGCACAAACCAAACCTACGAGTGCAGG GCCACAAACAGCGTGGGAGTGGCTCC TGGGCTTCATACCCATCTCTGCAGG	1266
CSF2RA	NM_006140.3	CSF2RA	TACCACACC CAGCATTC TC	115	CTAGAGGCT GGTGCCACT GT	499	CGCAGATCCGATTTCTCT GGGATC	883	67	TACCACACCAGCATTCCTCTCTGATCCC AGAGAAATCGGATCTGCGAACAGTGGC ACCAGCCCTCTAG	1267
CSK (SRC)	NM_004383.1	CSK	CCTGAACAT GAAGGAGC TGA	116	CATCACGTC TCCGAACTC C	500	TCCCCGATGGTCTGCAGC AGCT	884	64	CCTGAACATGAAGGAGCTGAAGCTGCT GCAGACCATCGGGAAGGGGAGTTCCGG AGACGTGATG	1268
CTGF	NM_001901.1	CTGF	GAGTTCAA GTGCCCTGA CG	117	AGTTGTAAT GGCAGGCAC AG	501	AACATCATGTTCTTCTTC ATGACCTCGC	885	76	GAGTTCAAAGTGCCCTGACGGCGAGGTC ATGAAGAAACAACATGATGTTTCATCAAG ACCTGTGCTTGCCTATTAACAAT	1269
CTHRC1	NM_138455.2	CTHRC1	GCTCACTTC GGCTAAAA TGC	118	TCAGCTCCA TTGAATGTG AAA	502	ACCAACGCTGACAGCAT GCATTTC	886	67	GCTCACTTCGGCTAAAATGCAGAAAATG CATGCTGTCAAGCGTTGGTATTTACATTT CAATGGAGCTGA	1270
CTSD	NM_001909.1	CTSD	GTACATGAT CCCCGTGTA GAAGGT	119	GGACAGCT TGTAGCCTT TGC	503	ACCCTGCCCGCGATCAC ACTGA	887	80	GTACATGATCCCCCTGTGAGAAAGGTGTC CACCTGCCCGCGATCACACTGAAGCT GGGAGGCAAAAGGCTACAAGCTGTCCC	1271
CTSL2	NM_001333.2	CTSL2	TGTCICACT GAGCGAGC AGAA	120	ACCATTGCA GCCCTGATT G	504	CTTGAGGACGCGAACAG TCCACCA	888	67	TGTCICACTGAGCGAGCAGAAATCTGGT GGAATGTTCGCGTCTCAAGGCAATCA GGGTGCAATGGT	1272
CTSL2int2	NM_001333.2int2		ACCAGGCA ATAACCTAA CAGC	121	CTGTTCTCC AAGCCAAAGA CA	505	AGGTGCAATATGGGCAT ATATCTCCATTG	889	79	ACCAGGCAATAACCTAACAGCACCCCAT TATAGGTGCAATATGGGCATATATCTC CATTGTCTTGGCTTGGAGAACAG	1273
CXCL10	NM_001565.1	CXCL10	GGAGCAAA ATCGATGCA GT	122	TAGGGAAGT GATGGGAGA GG	506	TCGTGTGTCCCATCCTT GGAAGC	890	68	GGAGCAAAATCGATGCAGTGTCTTCCAA GGATGGACCACACAGAGGCTGCCTCTC CCATCACTTCCCTA	1274
CXCL12	NM_000609.3	CXCL12	GAGCTACA GATGCCCAT GC	123	TTTGAGATG CTTGACGTT GG	507	TTCTTCGAAAGCCATGTT GCCAGA	891	67	GAGCTACAGATGCCCATGCCGATTCTT CGAAAGCCATGTTGCCAGAGCCCAACGT CAAGCATCTCAAA	1275
CXCL14	NM_004887.3	CXCL14	TGCGCCCTT TCCTCTGTA	124	CAATGCGGC ATATACTGG G	508	TACCCCTAAGAACGCC CCTCCAC	892	74	TGCGCCCTTCTCTCTGTACATATACCTT TAAGAACGCCCTCTCCACACACTGCC CCCAGTATATGCCGCATTG	1276
CXCR4	NM_003467.1	CXCR4	TGACCGCTT CTACCCCAA TG	125	AGGATAAGG CCAACCATG ATGT	509	CTGAAACTGGAAACACAA CCACCCACAAG	893	72	TGACCGCTTCTACCCCAATGACTTGTGG GTGGTTGTGTTCCAGTTTCAGCATCA TGGTTGGCCCTTATCTT	1277
CYP17A1	NM_000102.2	CYP17A1	CCGGAGTG ACTCTATCA CCA	126	GCCAGCAATT GCCATTATC T	510	TGGACACACTGATGCAA GCCAAGA	894	76	CCGGAGTGACTCTATCACCAACATGCT GGACACACTGATGCAAGCCCAAGATGAA CTCAGATAATGGCAATGCTGTC	1278

TABLE A

CYP19A1	NM_000103.2	CYP19A1	TCCTTATAG GTACTTTCA GCCATTG	127	CACCATGGC GATGTACTT TCC	511	CACAGCCACGGGGCCCA AA	895	70	TCCTTATAGGTA CTTTTGGCCCC TACATCGCCAT GGTG	1279
CYP1B1	NM_000104.2	CYP1B1	CCAGCTTTG TGCCCTGTCA CTAT	128	GGGAATGTG GTAGCCCAA GA	512	CTCATGCCACCACTGGC AACACCTC	896	71	CCAGCTTTGTG CCTGTGCCTGT TGCCACCACTG CCAAACACCT CTCTGTCTTG	1280
CYR61	NM_001554.3	CYR61	TGCTCATTC TTGAGGAG CAT	129	GTGGCTGCA TTAGTGTCC AT	513	CAGCACCCCTTGGCAGTTT CGAAAT	897	76	TGCTCATTTCT TGAGGAGCAT TTTGAAGACT TGCCAAAGGT GTGCTGGCG GATGGACACT AATGCAGCCAC	1281
DAB2	NM_001343.1	DAB2	TGGTGGGTG TAGGTGGTG TA	130	ACCAAAGAT GCTGTGTTC CA	514	CTGTCACACTCCCTCAGG CAGGAC	898	67	TGGTGGGTCTA GGGTGAGGTG GTGTAAGTCA CACTCCCTCAG GCAGGACCAT GGAACA CAGCATCTTTGGT	1282
DCC	NM_005215.1	DCC	AAATGTCC CTCGACTG CT	131	TGAATGCCA TCCTTCTTCC A	515	ATCACTGGAACTCCTCG GTCGGAC	899	75	AAATGTCCCTC TCGACTGCTCG ACTGCTCCGAG GTTCAGGTTC CCAGTATCAA GTGGAAGAAAG ATGGCATTCA	1283
DCC_exons 18-23	X76132_18-23		GGTCACCGT TGGTGTCA CA	132	GAGCGTCGG GTGCAAAATC	516	CAGCCACGATGACCACT ACCAGCACT	900	66	GGTCACCGTGT GGGTGTCATCG TGGCTGTGAT TTTGGC ACCCGACCGCTC	1284
DCC_exons 6-7	X76132_6-7		A'GGAGAT GTGGTCATT CCTAGTG	133	CACCACCCC AAGTATCCG TAAG	517	TGCTTCTCCCACTATCT GAAAATAA	901	74	ATGGAGATGTG GGTCAATCCCT AGTATGAGT ATTTTCAGATA GTGGGAGGAA GCAACT TACGGATACTT GGGGTGGTG	1285
DCK	NM_000788.1	DCK	GCCGCCAC AAGACTAA GGAAT	134	CGATGTTC CTTCGATGG AG	518	AGCTGCCCGTCTTTCTCA GCCAGC	902	110	GCCGCCACAAG ACTAAGGAAT GGCCAC CCCCCCAAAG AGAACTGCTT CTCAGCCAGCT CTGAGGGAC CCCCCAT CAAGAAAATCT CCATCGAAGG GAACAT	1286
DICER1	NM_177438.1	DICER1	TCCAATTCC AGCATCACT GT	135	GGCAGTGAA GGCGATAAA GT	519	AGAAAAGCTGTTGTCT CCCCAGCA	903	68	TCCAATTCCAG CATCACTGTC TCCCCAGCAT ACTTTAT CGCTTCACTGCC	1287
DLC1	NM_006094.3	DLC1	GATTCAGAC GAGGATGA GCC	136	CACCTCTTG CTGTCCCTTT G	520	AAAGTCCATTTTGCCACT GATGGCA	904	68	GATTCAGACG AGGATGAGCC TTGTGCC ATCAGTGGCAA AATGGACTT TCCAAAGG GACAGCAAG AGGTG	1288
DII.4	NM_019074.2	DII.4	CACGGAGG TATAAGGC AGGAG	137	AGAAGGAAG GTCCAGCCG	521	CTACCTGGACATCCCTGC TCAGCC	905	67	CACGGAGGTAT AAGGCAGGAG GCCTACC TGGACATCCC TGTCTCAGCC CCCCCGCGCT GGACCTTCCCTCT	1289
DR5	NM_003842.2	TNFRSF10 B	CTCTGAGAC AGTGCTTCG ATGACT	138	CCATGAGGC CCAACTTCC T	522	CAGACTTGGTGGCCCTTTG ACTCC	906	84	CTCTGAGACAG TGTCTCGATG ACTTTGC AGACTTGTGCT GCCCTTTGACT CTCTGGGA GCCGCTCATG AGGAAAGTT GGGCCCTCAT GG	1290
DSP	NM_004415.1	DSP	TGGCACTAC TGCATGATT GACA	139	CCTGCCGCA TTGTTTTCAG	523	CAGGGCCATGACAATCG CCAA	907	73	TGGCACTACTG CATGATTGAC ATAGAG AAGATCAGGG CCATGACAAAT CGCCAAG CTGAAAAACA ATGCGGCAGG	1291
DTYMK	NM_012145.1	DTYMK	AAATCGCTG GGAACAAG TG	140	AATGCGTAT CTGTCCACG AC	524	CGCCCTGGCTCAACTTTT CCITAA	908	78	AAATCGCTGGG AAACAAGTGC CCGTAAAT TAAGGAAAAG TTGAGCCAGG CGGTGAC CCTCGTCTGTG ACAGATACGCA TT	1292

TABLE A

DUSP1	NM_004417.2	DUSP1	AGACATCA GCTCCTGGT TCA	141	GACAAACAC CCTTCTCC AG	525	CGAGGCCATTGACTTCA TAGACTCCA	909	76	AGACATCAGCTCCTGTTCAACGAGGC CATTGACTTCAATAGACTCCATCAAGAA TGCTGGAGGAAGGTGTTGTG	1293
DUSP4	NM_001394.4	DUSP4	TGGTGACG ATGGAGGA GC	142	CTCGTCCCG GTTTCATCAG	526	TTGAGCACACTGCAGTC CATCTCC	910	68	TGGTACCGATGGAGGAGCTGCGGGAGA TGGACTGCAGTGTGCTCAAAAAGGCTGA TGAACCGGACGAG	1294
E2F1	NM_005225.1	E2F1	ACTCCCTCT ACCCTTGAG CA	143	CAGGCTCA GTTCTCTCA GT	527	CAGAAGAACAGCTCAGG GACCCCT	911	75	ACTCCCTCTACCCCTTGAGCAAGGGCAG GGGTCCCTGAGCTGTTCTTCTGCCCCAT ACTGAAGGAACGTAGGCCTG	1295
EBRP	AF243433.1		CTGCTGGAT GACCTTCCT C	144	CCAACAGTA CAGCCAGTT GC	528	CTCACCAGAAAGCCCCAA CCTCAAC	912	76	CTGCTGGATGACCTTCTCTCCACAGATG GCTCACCAAGAACCCCAACCTCAACAC CAGCAACTGGCTGTACTGTGG	1296
EDN1 endothelin	NM_001955.1	EDN1	TGCCACCTG GACATCAAT TG	145	TGGACCTAG GGCTTCCAA GTC	529	CACTCCCGAGCACGTTG TTCCGT	913	73	TGCCACCTGGACATCATTTTGGGTCAAC ACTCCCGAGCACGTTGTTCCTCGTATGGA CTTGGAAAGCCCTAGGTCCA	1297
EDN2	NM_001956.2	EDN2	CGACAAAG AGTGCGTCT ACTTCT	146	CAGGCCGTA AGGAGCTGT CT	530	CCACTTGGACATCATCTG GGTGAACACTC	914	79	CGACAAAGGAGTGGCTCTACTTCTGCCA CTTGGACATCATCTGGGTGAACACTCCT GAACAGACAGCTCCTTACGGCCTG	1298
EDNRA	NM_001957.1	EDNRA	TTTCCICAA ATTTGCCTC AAG	147	TTACACAATC CAACCAGTG CC	531	CCTTTGCCCTCAGGGCATC CTTTT	915	76	TTTCCICAAATTTGCCCTCAAGATGGAA CCCTTTGCCCTCAGGGCATCCTTTTGGCT GGCACTGTTGGATGTGTAA	1299
EDNRB	NM_000115.1	EDNRB	ACTGTGAAC TGCCTGGTG C	148	ACCACAGCA TGGGTGAGA G	532	TGCTACCTGCCCCCTTGT CATGTG	916	72	ACTGTGAACCTGCCTGGTGCAGTGTCCA CATGACAAAGGGGCAGGTAGCACCCCTC TCTCACCCATGCTGTGGT	1300
EEF1A1	NM_001402.5	EEF1A1	CGAGTGGA GACTGGTGT TCTC	149	CCGTTGTAA CGTTGACTG GA	533	CAAAGGTGACCACCATA CCGGGT	917	67	CGAGTGAGACTGGTGTCTCAAACCC GGTATGGTGGTCACCTTTTGCTCCAGTCA ACGTTACAAACGG	1301
EEF1A2	NM_001958.2	EEF1A2	ATGGACTCC ACAGAGCC G	150	GGCGCTGAC TTCTTGAC T	534	CTCGTCGTAGCGCTTCTC GCTGTA	918	66	ATGGACTCCACAGAGCCGGCCTACAGC GAGAAAGCGCTACGACGAGATCGTCAAG GAAGTCAGCGCC	1302
EEP	NM_005082.2	TRIM25	TTGAACAG AGCCTGACC AAG	151	TGTTGAGAT TCCTCGCAG TT	535	TGATGCTTTCTCCAGAAA CTCGAACTCA	919	74	TTGAACAGAGCCTTGACCAAGAGGGATG AGTTGAGTTTCTGGAGAAAGCATCAA AACTGCGAGGAATCTCAACA	1303
EGR1	NM_001964.2	EGR1	GTCCCCGCT GCAGATCTC T	152	CTCCAGCTT AGGGTAGTT GTCCAT	536	CGGATCCTTTCTCTCACTC GCCCCA	920	76	GTCCCCGCTGCAGATCTCTGACCCGTTC GGATCCTTCTCTCACTCGCCCAACCATGG ACAACTACCCCTAAGCTGGAG	1304
EGR3	NM_004430.2	EGR3	CCATGTGGA TGAATGAG GTG	153	TGCCTGAGA AGAGGTGAG GT	537	ACCCAGTCTCACCTTCTC CCCCAC	921	78	CCATGTGGATGAATGAGGTGTCTCCTTT CCATACCCAGTCTCACCTTCTCCCCACC CTACCCACCTCTTCTCAGGCA	1305
EIF4EBP1	NM_004095.2	EIF4EBP1	GGCGGTGA AGAGTCAC AGT	154	TTGGTAGTG CTCCACACG AT	538	TGAGATGGACATTTAAA GCACCAGCC	922	66	GGCGGTGAAGAGTCAACAGTTTGAGATG GACATTTAAAGCACCCAGCCATCGTGTG GAGCACTACCAA	1306

TABLE A

ELF3	NM_004433.2	ELF3	TCGAGGGC AAGAAAGAG CAA	155	GATGAGGAT GTCCCGGAT GA	539	CGCCAGAGGCACCCAC CTG	923	71	TCGAGGGCAAGAAAGAGCAAGCAGCG CCCAGAGGCACCCACCTGTGGAGTTTC ATCCGGGACATCCTCATC	1307
EMP1	NM_001423.1	EMP1	GCTAGTACT TTGATGCTC CCTTGAT	156	GAAACAGCTG GAGGCCAAG TC	540	CCAGAGAGCCTCCCTGCG AGCCA	924	75	GCTAGTACTTTTGATGCTCCCTTGATGGG GTCCAGAGAGCCTCCCTGCGAGGCCACCA GACTTGGCCTCCAGCTGTTC	1308
ENO1	NM_001428.2	ENO1	CAAGGCCG TGAACGAG AAGT	157	CGGTCAACGG AGCCAATCT	541	CTGCAACTGCTCCTGCT CAAAGTCA	925	68	CAAGGCCGTGAACGAGAAAGTCTCTGCAA CTGCCCTCTGCTCAAAAGTCAACCAGATT GGCTCCGTGACCG	1309
EP300	NM_001429.1	EP300	AGCCCCAG CAACTACA GTCT	158	TGTTCAAAG GTTGACCAT GC	542	CACTGACATCATGGCTG GCCTTG	926	75	AGCCCCAGCAACTACAGTCTGGGATGC CAAGGCCAGCCATGATGTCAGTGGCCC AGCATGGTCAACCTTTGAACA	1310
EP3AM	NM_002354.1	EP3AM	GGGCCCTCC AGAACAAAT GAT	159	TGCACTGCT TGGCCTTAA AGA	543	CCGCTCTCATCGCAGTCA GGATCAT	927	75	GGGCCCTCCAGAACAATGATGGGCTTT ATGATCCTGACTGCGATGAGAGCGGGC TCTTTAAGGCCAAGCAGTGCA	1311
EPHA2	NM_004431.2	EPHA2	CGCCTGTTC ACCAAGATT GAC	160	GTGGCGTGC CTCGAAGTC	544	TGGCCCCGATGAGATCA CCG	928	72	CGCCTGTTCACCAAGATTGACACCAATT GCCCCGATGAGATCACCGTCAGCAGC GACTTCGAGGCACGCCAC	1312
EPHB2	NM_004442.4	EPHB2	CAACCAGG CAGTCCAT C	161	GTAATGCTG TCCACGGTG C	545	CACCTGATGCATGATGG ACACTGC	929	66	CAACCAGGCAGCTCCATCGGCAGTGTC CATCATGCATCAGGTGAGCCGCAACCGT GGACAGCATTAC	1313
EPHB4	NM_004444.3	EPHB4	TGAACGGG GTATCCTCC TTA	162	AGGTACCTC TCGGTCAGT GG	546	CGTCCCATTTGAGCCTGT CAATGT	930	77	TGAACGGGGTATCCTCCTTAGCCACGG GGCCCGTCCCATTGTGAGCCTGTCAATGT CACCCTGACCCGAGAGGTACCT	1314
ER2	NM_001437.1	ESR2	TGGTCCATC GCCAGTTAT CA	163	TGTTCTAGC GATCTTGCT TCACA	547	ATCTGTATGCGGAACCT CAAAAGAGTCCCT	931	76	TGTTCCATCGCCAGTTATCACATCTGTA TGCGGAACCTCAAAAAGAGTCCCTGGTG TGAAGCAAGATCGCTAGAACA	1315
ERBB4	NM_005235.1	ERBB4	TGGCTCTTA ATCAGTTTC GTTACCT	164	CAAGGCATA TCGATCCTC ATAAAGT	548	TGTCCCACGAATAATGC GTAAATTCTCCAG	932	86	TGGCTCTTAATCAGTTTCGTTACCTGCC TCTGGAGAAATTTACGCATTTATTCGTGGG ACAAAACCTTTATGAGGATCGATATGCC TTG	1316
ERCC1	NM_001983.1	ERCC1	GTCCAGGTG GATGTGAA AGA	165	CGGCCAGGA TACACATCT TA	549	CAGCAGGCCCTCAAGGA GCTG	933	67	GTCCAGGTGGATGTGAAAGATCCCCAG CAGGCCCTCAAGGAGCTGGCTAAGATG TGTAATCCTGGCCG	1317
ERG	NM_004449.3	ERG	CCAACACTA GGCTCCCCA	166	CTCCGCCA GGTCTTAG T	550	AGCCATAIGCCTTCTCAT CTGGGC	934	70	CCAACACTAGGCTCCCCACCAGCCATA TGCTTCTCATCTGGGCACTTACTACTA AAGACCTGGCGGAGG	1318
ERRa	NM_004451.3	ESRRA	GGCATTGA GCCTCCTA CATCA	167	TCTCCGAGG AACCCTTIG G	551	AGAGCCGGCCAGCCCCTG ACAG	935	67	GGCATTGAGCCTCTCTACATCAAGGCA GAGCCGGCCAGCCCCTGACAGTCCAAAG GGTTCTCGGAGA	1319
ESD	NM_001984.1	ESD	GTCACCTCG CCACCGTAG	168	CTGTCCAAT TGCTGATTG CTT	552	TGCGCTACCATTTGGTGC AAGCAA	936	66	GTCACTCCGCCACCGTAGAATCGCCTA CCATTTGGTGC AAGCAAAAAGCAATCA GCAATTGGACAG	1320

TABLE A

ESPL1	NM_012291.1	ESPL1	ACCCCCAG ACCGGATC AG	169	TGTAGGGCA GACTTCCTC AAACA	553	CTGGCCCTCATGTCCCCT TCACG	937	70	ACCCCCAGACCGGATCAGGCAAGCTGG CCCTCATGTCCCCCTTCACGGTGTTTGAG GAAGTCTGCCCTACA	1321
ESRRG	NM_001438.1	ESRRG	CCAGCAC ATTGTTGAA GAT	170	AGTCTCTTG GGCATCGAG TT	554	CCCCAGACCAAGTGTGA ATACATGCT	938	67	CCAGCACCATTTGTTGAAAGATCCCCAGA CCAAAGTGTGAATACATGTCTCAACTCGA TGCCCAAGAGACT	1322
ESR1	NM_000125.1	ESR1	CGTGGTGCC CCTCTATGA C	171	GGCTAGTGG GCGCATGTA G	555	CTGGAGATGCTGGACGC CC	939	68	CGTGGTGCCCTCTCTATGACCTGTGCTG GAGATGCTGGACGCGCCACCGCCTACAT GCGCCACTAGCC	1323
ETV5	NM_004454.1	ETV5	ACCATGTAT CGAGAGGG GC	172	TGACCAGGA ACTGCCACA G	556	TTACCAGAGGCGAGGGTT CCCTTCA	940	67	ACCATGTATCGAGAGGGGCCCTTAC CAGAGGCGAGGTTCCCTTCAGCTGTGG CAGTTCTGGTCA	1324
FZH2	NM_004456.3	FZH2	TGGAACA GCGAAGGA TACA	173	CACCGAACA CTCCCTAGT CC	557	TCCTGACTTCTGTGAGCT CATTGCG	941	78	TGGAACAAGCGAAAGGATACAGCCTGTG CACATCCTGACTTCTGTGAGCTCATTGC GCGGACTAGGGAATGTTCTGGTG	1325
F3	NM_001993.2	F3	GTGAAGGA TGTGAAGC AGACGTA	174	AACCGGTGC TCTCCACAT TC	558	TGGCACGGGTCTTCTCCT ACC	942	73	GTGAAGGATGTGAAGCAGACGTACTTG GCACGGTCTTCTCTCTACCCGGCAGGG AATGTGGAGCACCCGGT	1326
FAP	NM_004460.2	FAP	CTGACCAG AACCACGG CT	175	GGAAGTGGG TCATGTGGG	559	CGGCCTGTCCACGAACC ACTTATA	943	66	CTGACCAGAACCAAGGCTTATCCGGCC TGTCCACGAACCACTTATACACCCACA TGACCCACTTCC	1327
FASN	NM_004104.4	FASN	GCCTCTTCC TGTTTCAGC G	176	GCTTTGCC GGTAGCTCT	560	TCGCCCACCTACGTACTG GCCTAC	944	66	GCCTCTCTGTCTGACGGCTCGCCCAC CTACGTACTGGCCTACACCCAGAGCTA CCGGGCAAAAGC	1328
FGFR2 isoform 1	NM_000141.2	FGFR2	GAGGGACT GTTGGCATG CA	177	GAGTGAGAA TTCGATCCA AGTCTTC	561	TCCCAGAGACCAACGTT CAAGCAGTTG	945	80	GAGGGACTGTTGGCATGCAGTGCCCTC CCAGAGACCAACGTTCAAGCAGTTGGT AGAAGACTTGGATCGAATTCTCACTC	1329
FGFR4	NM_002011.3	FGFR4	CTGGCTTAA GGATGGAC AGG	178	ACGAGACTC CAGTGTGA TG	562	CCTTTCATGGGGAGAAC CGCATI	946	81	CTGGCTTAAGGATGGACAGGCCTTTCA TGGGGAGAACCGCATTTGGAGGCATTCG GCTGCGCCATCAGCACCTGGAGTCTCGT	1330
FHIT	NM_002012.1	FHIT	CCAGTGGA GCGCTTCCA T	179	CTCTCTGGG TCGTCTGAA ACAA	563	TCGGCCACTTTCATCAGG ACGCAG	947	67	CCAGTGGAGCGCTTCCATGACCTGCCGT CCTGATGAAGTGGCCGATTGTGTTTCAG ACGACCCAGAGAG	1331
FLOT2	NM_004475.1	FLOT2	GACATCTGC GCTCCATCC	180	CAAACTGGT CCCGGTCCCT	564	AATCTGCTCCACTGTTCAG GGTCCC	948	66	GACATCTGGCTCCATCCTCGGGACCCCT GACAGTGGAGCAGATTATCAGGACCCG GGAOCCAGTTTG	1332
FN1	NM_002026.2	FN1	GGAAGTGA CAGACGTG AAGGT	181	ACACGGTAG CCGGTCACT	565	ACTCTCAGGCGGTCTCC ACATGAT	949	69	GGAAGTGACAGACGTGAAGGTCACCAT CATGTGGACACCGCCTGAGAGTGCAGT GACCGGCTACCGTGT	1333
FOS	NM_005252.2	FOS	CGAGCCCTT TGATGACTT CCT	182	GGAGCGGGC TGTCACAGA	566	TCCCAGCATCATCCAGG CCCAG	950	67	CGAGCCCTTTGATGACTTCCCTGTTCCCA GCATCATCCAGGCCCCAGTGGCTCTGAG ACAGCCCGCTCC	1334

TABLE A

FOXC2	NM_005251.1	FOXC2	GAGAACAA GCAGGGCT GG	183	CTTGACGAA GCACTCGTT GA	567	AGAACAGCATCCGCCAC AACCTCT	951	66	GAGAACAAAGCAGGGCTGGCAGAAACAG CATCCGCCACAACCTCTCTCGCTCAACGA GTGCTTCGTCAAG	1335
FOXO3A	NM_001455.1	FOXO3	TGAAGTCCA GGACGATG ATG	184	ACGGCTTGC TTACTGAAG GT	568	CCTACAGCAGCTCAGC CAGCCTG	952	83	TGAAGTCCAGGACGATGATGGCCTCT CTCGCCCATGCTCTACAGCAGCTCAGC CAGCCTGTCAACCTTCAGTAAAGCAAGCC GT	1336
FOXP1	NM_032682.3	FOXP1	CGACAGAG CTTGTCAC CT	185	GGTCGTCCA TTGGAAATCC T	569	CAGACCAAGCCTTTTGCC CAGAAAT	953	70	CGACAGAGCTTGTGCACCTAAAGCTGCA GACCAAGCCTTTGCCCAGAATTTAAGG ATTCCAATGGACGACC	1337
FOXP3	NM_014009.2	FOXP3	CTGTTTGCT GTCCGGAG G	186	GTGGAGGAA CTCTGGGAA TG	570	TGTTTCCATGGCTACCCC ACAGGT	954	66	CTGTTTGCTGTCCGGAGGCACCTGTGG GGTAGCCATGGAAACAGCACATTTCCCA GAGTTCCTCCAC	1338
FSCN1	NM_003088.1	FSCN1	CCAGCTGCT ACTTTGACA TCGA	187	GGTCACAAA CTTGCCATT GGA	571	TGACCGGCGCATCACAC TGAGG	955	74	CCAGCTGCTACTTTTGACATCGAGTGGC GTGACCGGCGCATCACACTGAGGGCGT CCAAIIGCAAAGTTIGAGCC	1339
FUS	NM_004960.1	FUS	GGATAATT AGACAAAC ACACCATCT	188	TGAAGTAAT CAGCCACAG ACTCAAT	572	TCAATTGTAACTTCTCA CCCAGGCCTTG	956	80	GGATAATTACAGACAACAACACCATCTT TGTGCAAGGCTGGGTGAGAAATGTTAC AATTGAGTCTGTGGCTGATTACTTCA	1340
FYN	NM_002037.3	FYN	GAAGCGCA GATCATGA AGAA	189	CTCCTCAGA CACCACATGC AT	573	CTGAAGCACGACAAAGCT GGTCCAG	957	69	GAAGCGCAGATCATGAAGAAGCTGAA GCACGACAAGCTGGTCCAGCTCTATGC AGTGTGCTTGAGGAG	1341
G-Catenin	NM_002230.1	JUP	TCAGCAGC AAGGGCAT CAT	190	GGTGGTTTT CTTGAGCGT GTACT	574	CGCCCCGAGGCCCTCATC CT	958	68	TCAGCAGCAAGGGCATCATGGAGGAGG ATGAGGCCCTGCGGGCGCCAGTACACGC TCAAGAAAACCCACC	1342
GAB2	NM_012296.2	GAB2	TGTTTGGAG GGAAGGC T	191	GAAGATAGC TGAGGGCTG TGAC	575	TGAGCCAGATTCCACAC CTCACGT	959	74	TGTTTGGAGGGAAGGGCTGGGGCTCTG AGCCAGATTCCACACCTCACGTTCACT CACAGCCCTCAGCTATCTTC	1343
GADD45	NM_001924.2	GADD45A	GTGCTGGTG ACGAATCC A	192	CCCGGCAAA AACAAATAA GT	576	TTCATCTCAATGGAAGG ATCCTGCC	960	73	GTGCTGGTGACGAATCCACATTCATCTC AATGGAAGGATCTTGCCCTTAAGTCAAC TTATTTGTTTTTTGCCGGG	1344
GADD45B	NM_015675.1	GADD45B	ACCTTCGAC AAGACCCAC ACT	193	TGGAGTTC ATGGGTACA GA	577	AAC TTCAGCCCCAGCTC CCAAGTC	961	70	ACCTTCGACAAGACCACACTTTTGGGAC TTGGGAGCTGGGGCTGAAGTTGCTCTG TACCCATGAAC TCCCA	1345
GAPDH	NM_002046.2	GAPDH	ATTCCACCC ATGGCAAA TTC	194	GATGGGATT TCCATTGAT GACA	578	CCGTTCTCAGCCCTTGACG GTGC	962	74	ATTCCACCCATGGCAAAATTCATGGCA CCGTCAAGGCTGAGAACGGGAAGCTTG TCATCAATGGAATCCCATC	1346
GATA3	NM_002051.1	GATA3	CAAAGGAG CTCAGCTGIG GTGTCT	195	GAGTCAGAA TGGCTTAT CACAGATG	579	TGTTCCAACCACTGAATC TGGACC	963	75	CAAAGGAGCTCACTGTGGTGTCTGTGT TCCAACCACTGAATCTGGACCCCACT GTGAATAAGCCATCTTGACTC	1347
GBP1	NM_002053.1	GBP1	TGGGAAAT ATTTGGGCA TT	196	AGAAAGCTAG GGTGGTTGT CC	580	TTGGGACATTTGTAGACTT GGCCAGAC	964	73	TTGGGAAATATTTGGGCATTGGTCTGG CCAAGTCTACAATGTCCCAATATCAAG GACAACCAACCTTAGCTTCT	1348

TABLE A

GBP2	NM_004120.2	GBP2	GCA TGGGA ACCA TCAAC CA	197	TGAGGAGTT TGCCTTGAT TCG	581	CCATGGACCAACTTCAC TATGTGACAGAGC	965	83	GCATGGGAACCATCAACCAAGCAGGCCA TGGACCAAACTTCACTATGTGACAGAGC TGACAGATCGAATCAAGGCAAACTCCT CA	1349
GCLM	NM_002061.1	GCLM	TGTAGAAATC AAACTCTTC ATCATCAAC TAG	198	CACAGAAATC CAGCTGTGC AACT	582	TGCAGTTGACATGGCCT GTTCAGTCC	966	85	TGTAGAAATCAAACTCTTTCATCATCAAACT AGAAGTGCAGTTGACATGGCCGTGTTCA GTCCTTGGAGTTGCACACAGCTGGATTCTG TG	1350
GDF15	NM_004864.1	GDF15	CGCTCCAGA CCTATGATG ACT	199	ACAGTGGAA GGACCAAGGA CT	583	TGTTAGCCAAAGACTGC CACTGCA	967	72	CGTCCAGACCTATGATGACTTGTTAGC CAAGACTGCCACTGCAATATGAGCAGT CCTGGTCTTCCACTGT	1351
GHI	NM_000515.3	GHI	GATCCCAA GGCCCAACT C	200	AGCCATTGC AGCTAGGTG AG	584	TGTCCACAGGACCCCTGA GTGGTTC	968	66	GATCCCAAAGGCCCAAAC TCCCCGAAACCA CTCAGGGTCTGTGGACACAGCTCACCTA GCTGCAATGGCT	1352
GJA1	NM_000165.2	GJA1	GTTCACTGG GGGTGTA TG G	201	AAATACCAA CATGCACCT CTCTT	585	ATCCCC TCCC TCTCCACC CATCTA	969	68	GTTCACTGGGGGTGTATGGGGGTAGATG GGTGGAGAGGAGGGGA TAAAGAGAGG TGCATGTTGTTATTT	1353
GJB2	NM_004004.3	GJB2	TGTCATGTA CGACGGCTT CT	202	AGTCCACAG TGTTGGGAC AA	586	AGGCGTTGCACTTCACC AGCC	970	74	TGTCATGTACGACGGCTTCTCCATGCAG CGGCTGGTGAAGTGC AACGCC TGGCCT TGTCCCAACACTGTGGACT	1354
GMNN	NM_015895.3	GMNN	GTTGCTAC GAGGATTG AGC	203	TGCGTACCC ACTTCTGC	587	CCCTTTGCCCACTTACTG GGTGA	971	67	GTTGCTACGAGGATTGAGCGTCTCCA CCCAGTAA GTGGGCAAGAGCGGCGCAG GAA GTGGGTACGCA	1355
GNAZ	NM_002073.2	GNAZ	TTCTGGACC TGGGACCTT AG	204	AAAGAGCTG TGAGTGGCT GG	588	CCGGGTGACAGCACTAA CCAGACC	972	68	TTCTGGACCTGGGACCTTAGGAGCCGG GTGACAGCACTAACCA GACCTCCAGCC ACTCACAGCTCTT	1356
GPR30	NM_001505.1	GPER	CGTGCCCTCT ACACCATCT TC	205	ATGTTCAAC ACCAGATC AG	589	CTCTTCCCCATCGGCTTT GTGG	973	70	CGTGCCCTACACCATCTTCTCTCTCCC CATCGGCTTTGTGGCAACATCCTGATC CTGGTGGTGAACAT	1357
GPS1	NM_004127.4	GPS1	AGTACAAG CAGGCTGCC AAG	206	GCAGCTCAG GGAAATCAC A	590	CCCTCTGCTGGCTTCCCT TGATCA	974	66	AGTACAAGCAGGCTGCCAAAGTGCCTCC TGCTGGCTTCTTTGATCACTGTGACTT CCCTGAGCTGC	1358
GPX1	NM_000581.2	GPX1	GCTTATGAC CGACCCCA A	207	AAA GTTCCA GGCAACATC GT	591	CTCATCA CCTGGTCTCCG GTGTGT	975	67	GCTTATGACCGACCCCAAGCTCATCAC CTGGTCTCCGGTGTGTGTCGCAACGATGTT GCCTGGAACTTT	1359
GPX2	NM_002083.1	GPX2	CACACAGA TCTCCTACT CCATCCA	208	GGTCCAGCA GTGTCTCCT GAA	592	CATGCTGCATCCTAAGG CTCCTCAGG	976	75	CACACAGATCTCCTACTCCATCCAGTCC TGAGGAGCCTTAGGATGCAGCATGCCT TCAGGAGACACTGCTGGACC	1360
GPX4	NM_002085.1	GPX4	CTGAGTGTG GTTTGCGGA T	209	TACTCCC TG GCTCCTGCT T	593	CTGGCCTTCCC GTGTAAC CAGTTC	977	66	CTGAGTGTGGTTTGGGATCCTTGGCCTT CCC GTTAACCAAGTTCGGGAAGCAGGA GCCAGGGAGTA	1361
GRB7	NM_005310.1	GRB7	CCATCTGCA TCCATCTTG TT	210	GGCCACCAG GGTATTATC TG	594	CTCCCCACCCTTGAGAA GTGCCT	978	67	CCATCTGCATCCATCTTGT TGGGCTCC CCACCCCTTGAGAA GTGCCTCAGATAAT ACCCCTGGTGGCC	1362

TABLE A

GREB1 variant a	NM_014668.2	GREB1	CAGATGAC AATGGCCA CAAT	211	GAAAGCCTTT CTTCCACA GC	595	CACAATTCCAGAGAAA CCAAGAAGAGC	979	71	CAGATGACAAATGGCCACAATGCTCTTC TTGGTTTCTCTGGGAATTGTGTGGCTG TGGAAGAAAAGGCTTC	1363
GREB1 variant b	NM_033090.1	GREB1	TGCTTAGGT GCGGTAAA ACCA	212	CAAGAGCCT GAATGCGTC AGT	596	ACCACGCGAACGGTGCA TCG	980	73	TGCTTAGGTGCGGTAAACCAGCGCTT GTCCGATGCACCGTTCGCGTGGTAAAC TGACGCAATTCAGGCTCTTG	1364
GREB1 variant c	NM_148903.1	GREB1	CCCAGGC ACCAGCTTT A	213	ACTTCGGCT GTGTGTTAT ATGCA	597	TCCCCCGAGCCCAGCAGG ACA	981	64	CCCAGGCACCCAGCTTTACTCCCCGAG CCCAGCAGGACATCTGCATATAACACA CAGCCGAAGT	1365
GRN	NM_002087.1	GRN	TGCCCCCAA GACACTGTG T	214	GAGTCCGT GGTAGCGTT CTC	598	TGACCTGATCCAGAGTA AGTGCCTCTCCA	982	72	TGCCCCAAGACACTGTGTGTGACCTG ATCCAGAGTAAGTGCCTCTCCAAGGAG AACGCTACCACGGACCTC	1366
GS1M1	NM_000561.1	GS1M1	AAGCTATG AGGAAAAG AAGTACAC GAT	215	GGCCCAGCT TGAATTTTC A	599	TCAGCCACTGGCTTCTGT CATAA1CAGGAG	983	86	AAGCTATGAGGAAAAGAAAGTACACGAT GGGGACGCTCCTGTGATTTATGACAGAA CCAGTGGCTGAATGAAAAATTCAAGCT GGGC	1367
GSTM2 gene	NM_000848 ^{gen} _e		CTGGGCTGT GAGGCTGA GA	216	GCGAATCTG CTCCTTTTCT GA	600	CCCGCTACCCCTCGTAA AGCAGATTCA	984	71	CTGGGCTGTGAGGCTGAGAGTGAATCT GCTTTACGAGGGTAGCGGGGGAATCAG AAAAGGAGCAGATTTCG	1368
GSTM2	NM_000848.2	GSTM2	CTGCAGGC ACTCCCTGA AAT	217	CCAAGAAAC CATGGCTGC TT	601	CTGAAGCTCTACTCACA GTTTCTGGG	985	68	CTGCAGGCACCTCCCTGAAATGCTGAAG CTCTACTCACAGTTTCTGGGGAAGCAG CCATGGTTCTTGG	1369
GSTM3	NM_000849.3	GSTM3	CAATGCCAT CTTGCGCTA CAT	218	GTCCACTCG AATCTTTTCT TCTCA	602	CTCGCAAAGCACAAACATG TGTGTGAGA	986	76	CAATGCCATCTTGGCTACATCGTCTGC AAGCACAAACATGTGTGGTGAGACTGAA GAAGAAAAGATTGAGTGGAC	1370
GSTT1	NM_000853.1	GSTT1	CACCATCCC CACCTGTC T	219	GGCCTCAGT GTGCATCAT TCT	603	CACAGCCGCTGAAAGC CACAAAT	987	66	CACCATCCCCACCCTGTCTTCCACAGCC GCCTGAAAGCCACAAATGAGAAATGATGC ACACTGAGGCC	1371
GUS	NM_000181.1	GUSB	CCCACTCAG TAGCCAAGT CA	220	CACGCAGGT GGTATCAGT CT	604	TCAAGTAAACGGGCTGT TTTCCAAACA	988	73	CCCACTCAGTAGCCAAAGTCACAAATGTT TGGAACACAGCCCCGTTTACTTGAGCAA GACTGATACCACCTGCGTG	1372
I13T3A	NM_002107.3	I13T3A	CCAAACGT GTAAACAATT ATGCC	221	TCTTAAAGCA CGTTCTCCA CG	605	AAAGACATCCAGCTAGC ACGCCG	989	70	CCAAACGTGTAAACAATTATGCCAAAAG ACATCCAGCTAGCACGCCGCATACGTG GAGAACGTGCTTAAGA	1373
HDAC1	NM_004964.2	HDAC1	CAAGTACC ACAGCGAT GACTACATT AA	222	GCTTGCTGT ACTCCGACA TGTT	606	TTCTTGCGCTCCATCCGT CCAGA	990	74	CAAGTACCACAGCGATGACTACATTAA ATTCTTGCGCTCCATCCGTCCAGATAAC ATGTCGGAGTACAGCAAGC	1374
HDAC6	NM_006044.2	HDAC6	TCC1GTGCT CTGGAAGC C	223	CTCCACGGT CTCAGTTGA TCT	607	CAAGAACCCTCCAGAAG GGCTCAA	991	66	TCC1GTGCTCTGGAAGCCCTTGAGCCCT TCTGGAGGTTCTTGTGAGATCAACTG AGACCGTGGAG	1375
HER2	NM_004448.1	ERBB2	CGGTGTGA GAAGTGCA GCAA	224	CC1CTCGCA AGTGTCTCCA T	608	CCAGACCATAGCACACT CGGGCAC	992	70	CGGTGTGAGAAAGTGCAGCAAGCCCTGT GCCCGAGTGTGCTATGGTCTGGGCATG GAGCACTTGCGGAGAGG	1376

TABLE A

HES1	NM_005524.2	HES1	GAAAGATA GCTCGCGC A	225	GGAGGTGCT TCACTGTCA TTT	609	CAGAAATGTCCGCTTCTC CAGCTT	993	68	GAAAGATAGCTCGCGGCATTCCAAAGCT GGAGAAAGCGGACATTTCTGGAATGAC AGTGAAGCACCTCC	1377
HGFAC	NM_001528.2	HGFAC	CAGGACAC AAGTGCCA GATT	226	GCAGGGAGC TGGAGTAGC	610	CGCTCACGTTTCTCATCCA AGTGG	994	72	CAGGACACAAGTGCCAGATTGCGGGCT GGGGCCACTTTGGATGAGAACGTGAGCG GCTACTCCAGCTCCCTGC	1378
HLA-DPB1	NM_002121.4	HLA-DPB1	TCCATGATG GTTCTGCAG GTT	227	TGAGCAGCA CCATCAGTA ACG	611	CCCCGGACAGTGGCTCT GACG	995	73	TCCATGATGGTTTCTGCAGGTTTCTGCGG CCCCCGGACAGTGGCTCTGTACGGCGT TACTGATGGTGTCTCTCA	1379
HMGB1	NM_002128.3	HMGB1	TGGCCTGTC CATTGGTGA T	228	GCTTGTCA CTGCAGCAG TGTT	612	TTCCACATCTCTCCCA TTCTTCGCA	996	71	TGGCCTGTCCATTGGTGATGTTGCGAA GAAACTGGGAGAGATGTGGAATAACAC TGCTGCAGATGACAAAGC	1380
HNF3A	NM_004496.1	FOXA1	TCCAGGATG TTAGGAACT GTGAAG	229	GCGTGTCTG CGTAGTAGC TGTT	613	AGTCGCTGGTTTCATGCC CTTCCA	997	73	TCCAGGATGTTAGGAACTGTGAAGATG GAAGGCGATGAAACCCAGCGACTGGAA CAGCTACTACGCAGACACGC	1381
HNRPAB	NM_004499.3	HNRNPAB	AGCAGGAG CGACCAACT GA	230	GTTTGCCAA GTTAAATTT GGTACATAA T	614	CTCCATATCCAAACAAA GCATGTGCG	998	84	AGCAGGAGCGACCAACTGATCGCACAC ATGCTTTGTTGGATATGGAGTGAACA CAATTATGTACCAAAATTTAACTTGGCA AAC	1382
HNRPC	NM_004500.3	HNRNPC	GCAGCAGT CGGCTTCTC T	231	GGGAGGGAG AAGAGATT GAT	615	AGTCTCCTACTCCCGGT TCTGG	999	68	GCAGCAGTCGGCTTCTTACGCAGAAC CCGGGAGTAGGAGACTCAGAATCGAAT CTCTTCTCCCTCCC	1383
HoxA1	NM_005522.3	HOXA1	AGTGACAG ATGGACAA TGCAAGA	232	CCGAGTCGC CACTGCTAA GT	616	TGAACTCCTTCTCCTGGAAT ACCCA	1000	69	AGTGACAGATGGACAATGCAAGAAATGA ACTCCTTCTGGAATACCCATATACTTAG CAGTGGCGACTCGG	1384
HoxA5	NM_019102.2	HOXA5	TCCCCTGTG TTCCCTCTG TGAA	233	GGCAATAAA CAGGCTCAT GATTAA	617	AGCCCTGTCTCTCGTTGCC CTAATTCATC	1001	78	TCCCTTGTGTCTTCTGTGTGAAGAAGCC CTGTTCTCGTTGCCCTAATTCATCTTT AATCATGAGCCTGTATTGCC	1385
HOXB13	NM_006361.2	HOXB13	CGTGCCTTA TGTTACTT TGG	234	CACAGGGTT TCAGCGAGC	618	ACACTCGGCAGGAGTAG TACCCGC	1002	71	CGTGCCTTATGGTTACTTTGGAGGCGG GTACTACTCTCTGCCGAGTGTCCCGGAG CTCGCTGAAACCCCTGTG	1386
HOXB7	NM_004502.2	HOXB7	CAGCCTCAA GTTTCGGTT TC	235	GTGGAAAGC AAACGCACA	619	ACCGGAGCCTTCCCAGA ACAAACT	1003	68	CAGCCTCAAAGTTTCGGTTTTCGCTACCGG AGCCTTCCCAGAACACAACTTCTTGTGC GTTTGTCTTCCAAAC	1387
HSD17B1	NM_000413.1	HSD17B1	CTGGACCGC ACGGACAT C	236	CGCCTCGCG AAAGACTTG	620	ACCGTTCTACCAATACC TCGCCCA	1004	78	CTGGACCGCACGGACATCCACACCTTC CACCGCTTCTACCAATACCTCGCCACA GCAAGCAAAGTCTTTCGCGAGGCG	1388
HSD17B2	NM_002153.1	HSD17B2	GCTTTCCAA GTGGGGAA TTA	237	TGCCTGCGA TATTTGTTA GG	621	AGTTGCTTCCATCCAACC TGGAGG	1005	68	GCTTTCCAAAGTGGGGAATTAAAGTTGC TTCCATCCAAACCTGGAGGCTTCCTAACA AATAJCGCAGGCA	1389
HSN1	NM_017493.3	OTUD4	CAGTCTCGC CATGTTGAA GT	238	ATAAACGCT TCAAAATT TCTCTG	622	CAGAAJGGCTGTATTC ACTATCTTCGAGA	1006	77	CAGTCTCGCCATGTTGAAGTCAGAAATG GCCTGTATTCACTATCTTCGAGAGAAC AGAGAGAAATTTGAAGCGTTTAT	1390

TABLE A

HSPA1A	NM_005345.4	HSPA1A	CTGCTGCGA CAGTCCACT A	239	CAGGTTCGC TCTGGGAAG	623	AGAGTGACTCCCGTTGT CCCAAGG	1007	70	CTGCTGCGACAGTCCACTACCTTTTTCG AGAGTGACTCCCGTTGTGCCCAAGGCTT CCCAGAGCGAACCCTG	1391
HSPA1B	NM_005346.3	HSPA1B	GGTCCGCTT CGTCTTTTCG A	240	GCACAGGTT CGCTCTGGA A	624	TGACTCCCGCGGTCCCA AGG	1008	63	GGTCCGCTTCTGCTTTTCGAGAGTGACTC CCGCGGTCCCAAGGCTTTTCCAGAGCGA ACCTGTGC	1392
HSPA4	NM_002154.3	HSPA4	TTCAGTGTG TCCAGTGCA TC	241	ATCTGTTTC ATTGGCTCC T	625	CATTTTCCTCAGACTTGT GAACCTCCACT	1009	72	TTCAGTGTGTCCAGTGCAATCTTTAGTGG AGGTTCACAAAGTCTGAGGAAAAATGAGG AGCCAATGGAAACAGAT	1393
HSPA5	NM_005347.2	HSPA5	GGCTAGTA GAACTGGA TCCCAACA	242	GGTCTGCC AAATGCTTT TC	626	TAATTAGACCTAGGCCT CAGCTGCACTGCC	1010	84	GGCTAGTAGAACTGGATCCCAACACCA AACTCTTAATTAGACCTAGGCCCTCAGCT GCACTGCCCCGAAAAAGCATTTGGGCAGA CC	1394
HSPA8	NM_006597.3	HSPA8	CCTCCCTCT GGTGGTGGT T	243	GCTACATCT ACACTTGGT TGGCTTAA	627	CTCAGGGCCACCAATTG AAGAGGTG	1011	73	CCTCCCTCTGGTGGTGTCTCTCAGGGC CCACCAATTGAAGAGGTTGATTAAGCCA ACCAAGTGTAGATGTAGC	1395
HSPB1	NM_001540.2	HSPB1	CCGACTGG AGGAGCAT AAA	244	ATGCTGGCT GACTCTGCT C	628	CGCACTTTTCTGAGCAG ACGTCCA	1012	84	CCGACTGGAGGAGCATAAAAGCGCAGC CGAGCCAGCGCCCCCGCACTTTTCTGA GCAGACGTCCAGAGCAGATCAGGCCAG CAT	1396
IBSP	NM_004967.2	IBSP	GAATACCA CACTTTCTG CTACAACAC T	245	GGATTGCAG CTAACCCTG TATACC	629	CCAGGCGTGGCGTCCCTC TCCATA	1013	83	GAATACCACACTTTTCTGCTACAAACACT GGGCTATGGAGAGGACGCCACGCCCTGG CACAGGGTATACAGGGTTAGCTGCAAT CC	1397
ICAM1	NM_000201.1	ICAM1	GCAGACAG TGACCAATCT ACAGCTT	246	CTTCTGAGA CCCTGGCT TCGT	630	CCGGCGCCCAACGTGAT TCT	1014	68	GCAGACAGTGACCATCTACAGCTTTTCC GGCGCCCAACGTGATTCGTGACGAAAGCC AGAGGTCTCAGAAAG	1398
ID1	NM_002165.1	ID1	AGAACCGC AAGGTGAG CAA	247	TCCAACCTGA AGTCCCCTG ATG	631	TGGAGATTCTCCAGCAC GTCAATCGAC	1015	70	AGAACCGCAAGGTGAGCAAGGTGGAG ATTCTCCAGCACGTCAATCGACTACATCA GGGACCTTCAGTTGGA	1399
ID4	NM_001546.2	ID4	TGGCCTGGC TCTTAATT G	248	TGCAATCAT GCAAGACCA C	632	CTTTGTTTTGCCCAGTA TAGACTCGGAAG	1016	83	TGGCTGGCTCTTAATTTGCTTTTGTGTTT GCCCAGTATAGACTCGGAAGTAACAGT TATAGCTAGTGGTCTTGCATGATTGCA	1400
IDH2	NM_002168.2	IDH2	GGTGGAGA GTGGAGCC ATGA	249	GCTCGTTCA GCTTCACAT TGC	633	CCGTGAATGCAGCCCGC CAG	1017	74	GGTGGAGAGTGGAGCCCATGACCAAAG ACCTGGCGGGCTGCATTCACGGCCTCA GCAATGTGAAGCTGAACGAGC	1401
IGF1R	NM_000875.2	IGF1R	GCAATGGTA GCCGAAAGA TTTCA	250	TTTCCGGTA ATAGTCTGT CTCATAGAT ATC	634	CGCGTCATACCAAAATC TCCGATTTTGA	1018	83	GCATGGTAGCCCGAAAGATTTCACAGTCA AAATCGGAGATTTTGGTATGACCGGAG ATATCTATGAGACAGACTATTACCCTCA AA	1402
IGF2	NM_000612.2	IGF2	CCGTGCTTC CGGACAAC TT	251	TGGACTGCT TCCAGGTGT CA	635	TACCCCGTGGGCAAGTT CTTCCAA	1019	72	CCGTGCTTCCGGACAAACTTCCCCAGAT ACCCCGTGGCAAGTTCCTTCCAATAIG ACACCTGGAAGCAGTCCA	1403

TABLE A

IGFBP6	NM_002178.1	IGFBP6	TGAACCGCAGAGACCAACAGAGGAATC CAGGCACCTCTACCAAGCCCTCCCAGC CCAAATTCTGCGGGTGTCCAAGAC	636	ATCCAGGCACCTCTACC ACGCCCTC	1020	77	TGAACCGCAGAGACCAACAGAGGAATC CAGGCACCTCTACCAAGCCCTCCCAGC CCAAATTCTGCGGGTGTCCAAGAC	1404
IGFBP7	NM_001553.1	IGFBP7	GGGTCACTA TGGAGTTCA AAGGA	637	GGGTCTGAA TGGCCAGGT T	1021	68	GGGTCACTATGGAGTTCAAAGGACAGA ACTCCTGCCTGGTGAACCGGGGACAACTT GGCCATTCAAGACCC	1405
IKBKE	NM_014002.2	IKBKE	GCCTCCCAT AGCTCCTTA CC	638	CAGAGCTCT TGCATGTGG AG	1022	66	GCCTCCCATAGCTCCTTACCCAGCCCT ACACGAAAGGACCTGCTTCTCCACATG CAAGAGCTCTG	1406
IL-8	NM_000584.2	IL8	AAGGAACC ATCTCACTG TGTGTAAC	639	ATCAGGAAG GCTGCCAAG AG	1023	70	AAGGAACCATCTCACTGTGTGTAAACA TGACTTCCAAGCTGGCCGTGGCTCTCTTT GGCAGCCTTCCCTGAT	1407
IL10	NM_000572.1	IL10	GGCGCTGTC ATCGATTTC TT	640	TGGAGCTTA TTAAAGGCA TTCTTCA	1024	79	GGCGCTGTCAATCGATTTCCTCCCTGTGA AAACAAGAGCAAGGCCGTGGAGCAGG TGAAGAATGCCCTTTAATAAGCTCCA	1408
IL11	NM_000641.2	IL11	TGGAAGGTT CCACAAGTC AC	641	TCITGACCCTT GCAGCTTTG T	1025	66	TGGAAGGTTCCAACAAGTCACCCCTGTGA TCAACAGTACCCGTATGGGACAAAGCT GCAAGGTCAAGA	1409
IL17RB	NM_018725.2	IL17RB	ACCCCTGG TGGTAAATG GA	642	GGCCCCAAT GAAATAGAC TG	1026	76	ACCCCTGGTGGTAAATGGACATTTTCC TACATCGGCTTCCCTGTAGAGCTGAAC ACAGTCTATTTCATTGGGGCC	1410
IL6ST	NM_002184.2	IL6ST	GGCCTAATG TTCCAGATC CT	643	AAAATTGTG CCTTGGAGG AG	1027	74	GGCCTAATGTTCAGATCCTTCAAGA GTCAATATTGCCAGTGGTCACTCACAC TCTCCAAGGCACAATTT	1411
ING1	NM_005537.2	ING1	ACTTTCCTG CGAGGTCA GTC	644	AACTCCGAG TGGTGATCC A	1028	66	ACTTTCCTGCGAGGTCAGTCAAGGCTTT GGGGGCTCTGTTTTGAATGTGGATCAC CACTCGGAGTT	1412
INHBA	NM_002192.1	INHBA	GTGCCCGA GCCATATAG CA	645	CGGTAGTGG TTGATGACT GTGA	1029	72	GTGCCCGAGCCATATAGCAGGCACGTC CGGTCTCTCACTGTCTCTCCACTCAACA GTCAATCAACCACCTACCG	1413
IRF1	NM_002198.1	IRF1	AGTCCAGCC GAGATGCT AAG	646	AGAAAGTAT CAGGGCTGG AA	1030	69	AGTCCAGCCGAGATGCTAAGAGCAAGG CCAAGAGGAAGTCATGTGGGATTTCCA GCCCTGATACCTTCT	1414
IRS1	NM_005544.1	IRS1	CCACAGCTC ACCTTCTGT CA	647	CCTCAGTGC CAGTCTCTT CC	1031	74	CCACAGCTCACCTTCTGTCAAGGTGTCCA TCCCAGCTCCAGCCAGCTCCCCAGAGAG GAAGAGACTGGCACTGAGG	1415
ITGA3	NM_002204.1	ITGA3	CCATGATCC TCACTCTGC TG	648	GAAGCTTTG TAGCCGGTG AT	1032	77	CCATGATCCTCACTCTGTGTGGTGACTA TACACTCCAGACCTCGCTTAGCATGGT AAATCACCGGCTACAAAGCTTC	1416
ITGA4	NM_000885.2	ITGA4	CAACGCTTC AGTGATCA ATCC	649	GTCTGGCCG GGATTCCTT	1033	66	CAACGCTTCAGTGATCAATCCCCGGGC GATTTACAGATGCAGGATCGGAAAGAA TCCCGGCCAGAC	1417

TABLE A

ITGA5	NM_002205.1	ITGA5	AGGCCAGC CCTACATTA TCA	266	GTCTTCTCC ACAGTCCAG CA	650	TCCTGAGCC TTGTCTCTA TCCGGC	1034	75	AGGCCAGCCCTACATTATCAGAGCAAG AGCCGGATAGAGGACAAGGCTCAGAIC TTGCTGGACTGTGGAGAAAGAC	1418
ITGA6	NM_000210.1	ITGA6	CAGTGACA AACAGCCCT TCC	267	GTTTAGCCT CATGGGCGT C	651	TCGCCATCTTTTGTGTGGGA TTCCCTT	1035	69	CAGTGACA AACAGCCCTTCCAAACCCAA GGAAATCCCA CAAAAGATGGCGATGACG CCCATGAGGCTAAAC	1419
ITGAV	NM_002210.2	ITGAV	ACTCGGACT GCACAAAGC TAAT	268	TGCCATCAC CATTGAAAT CT	652	CCGACAGCCACAGAATA ACCCAAA	1036	79	ACTCGGACTGCACAAAGCTATTTTGTATG ACAGCTATTTGGGTTATTCTGTGGCTGT CGGAGATTTTCAATGGTGATGGCA	1420
ITGB1	NM_002211.2	ITGB1	TCAGAAATTG GATTTGGCT CA	269	CCTGAGCTT AGCTGGTGT TG	653	TGCTAATGTAAAGGCATC ACAGTCTTTTCCA	1037	74	TCAGAAATTGGATTTTGGCTCATTTTGTGGA AAAGACTGTGATGCCTTACATTAGCAC AACACCAGCTAAGCTCAGG	1421
ITGB3	NM_000212.2	ITGB3	ACCGGGGA GCCCTACAT GA	270	CCTTAAGCT CTTTCACTG ACTCAATCT	654	AAATACCTGCAACCGTT ACTGCCGTGAC	1038	78	ACCGGGGAGCCCTACATGACGAAAAATA CCTGCAACCCGTTACTGCCGTGACGAGA TTGAGTCAGTGAAAGAGCTTAAGG	1422
ITGB4	NM_000213.2	ITGB4	CAAGGTGC CCTCAGTGG A	271	GCGCACACC TTCATCTCAT	655	CACCAACCTGTACCCGT ATTGCGA	1039	66	CAAGGTGCCCTCAGTGGAGCTCACCCAA CCTGTACCCGTATTTCGACTATGAGAT GAAGGTGTGCGC	1423
ITGB5	NM_002213.3	ITGB5	TCTGTAAA GATGACCA GGAG	272	GGTGAACAT CATGACGCA GT	656	TGCTATGTTTCTACAAAA CCGCCAAAG	1040	71	TCTGTAAAAGATGACCAAGGAGGCTGTGC TATGTTTCTACAAAACCCGCCAAGGACT GCGTCATGATGTTCCAC	1424
JAG1	NM_000214.1	JAG1	TGGCTTACA CTGGCAATG G	273	GCATAGCTG TGAGATGG G	657	ACTCGATTTCCCAGCCA ACCACAG	1041	69	TGGCTTACACTGGCAATGGTAGTTTCTIG TGGTTGGCTGGGAAATCAGAGTGCCCGCA TCTCACAGCTATGC	1425
JUNB	NM_002229.2	JUNB	CTGTCAGCT GCTGCTTGG	274	AGGGGGTGT CCGTAAAGG	658	CAAGGGACACGCCCTTCT GAACCT	1042	70	CTGTCACTGTCTGCTTGGGGTCAAGGG ACACGCCCTTCTGAACGTCCTCCCTGCCCT TTACGGACACCCCCCT	1426
Ki-67	NM_002417.1	MKI67	CGGACTTTG GGTGGACT T	275	TTACAACTC TTCCACTGG GACGAT	659	CCACTTGTCTGAACCCACC GCTCGT	1043	80	CGGACTTTGGGTGCGACTTGACGAGCG GTGGTTCGACAAAGTGGCCCTGCGGGCC GGATCGTCCCACTGAGTGAAGAGTTGTAA	1427
KIAA0555	NM_014790.3	JAKMIP2	AAGCCCGA GGCACTCAT T	276	TGCTGTGA GCTTGGTCC TG	660	CCCTTCAAGCTGCCAAT GAAGACC	1044	67	AAGCCCGAGGCACATCATTTGTGCCCTTC AAGCTGCCAATGAAGACCTCAGGACCA AGCTCACAGACA	1428
KIAA1199	NM_018689.1	KIAA1199	GCTGGGAG GCAGGACTT C	277	GAAGCAGGT CAGAGTGAG CC	661	CTTCAAGGCCATGCTGA CCATCAG	1045	66	GCTGGGAGGCAGGACTTCTCTTCAAG GCCATGCTGACCATCAGCTGGCTCACT CTGACCTGCTTC	1429
KIF14	NM_014875.1	KIF14	GAGCTCCAT GGCTCATCC	278	TCACACCCA CTGAATCCT ACIG	662	TGCATTCTCTGAGCTCA CTGCTG	1046	69	GAGCTCCATGGCTCATCCCCAGCAGTG AGCTCAGAGGAATGCACACCCAGTAGG ATTCAGTGGGIGTGA	1430
KIF20A	NM_005733.1	KIF20A	TCTCTTGCA GGAAGCCA GA	279	CCGTAGGGC CAATTCAGA C	663	AGTCAGTGGCCCATCAG CAATCAG	1047	67	TCTCTTGCAAGAACCCAGACAAACAGTC AGTGGCCCATCAGCAATCAGGGTCTGA ATTGGCCCTACGG	1431

TABLE A

KIF2C	NM_006845.2	KIF2C	AATTCTCTGC TCCAAAAG AAAGTCTT	280	CGTGATGCG AAGCTCTGA GA	664	AAGCCGCTCCACTCGCA TGTC	1048	73	AATTCTGTCTCCAAAAAGAAAGTCTTCG AAGCCGCTCCACTCGCATGTCCACTGTGTC TCAGAGCTTCGCATCAAG	1432
KLK11	NM_006853.1	KLK11	CACCCCGGC TTCAACAAC	281	CATCTTCAC CAGCATGAT GTCA	665	CCTCCCCAACAAAGACC ACCGCA	1049	66	CACCCCGGCTTCAACAACACAGCCTCCCC AACAAAGACCACCGCAATGACATCATG CTGGTGAAGATG	1433
KLK6	NM_002774.2	KLK6	GACGTGAG GGTCCTGAT TCT	282	TCCTCACTC ATCACGTCC TC	666	TTACCCAGCTCCATCCT TGCATC	1050	78	GACGTGAGGGTCTCTGTATTCTCCCTGGTT TTACCCAGCTCCATCCTTGGCATCACTG GGGAGGACGTGATGAGTGAGGA	1434
KLRC1	NM_002259.3	KLRC1	CATCCTCAT GGATTGGTG TG	283	GCCAAACCA TTCATTGTC AC	667	TTTCGTAACAGCAGTCA TATCCATGG	1051	67	CATCCTCATGGATTGGTGTGTTTCGTAA CAGCAGTCAATCATCCATGGGTGACAAT GAATGGTTTGGC	1435
KNSI.2	BC000712.1		CCACCTCGC CATGATTTT TC	284	GCAATCTCT TCAAAACACT TCATCCT	668	TTTGACCGGGTATTCCCA CCAGGA	1052	77	CCACCTCGCCATGATTTTCCCTTTGACC GGGTATTCCCAACCAGGAAGTGGACAGG ATGAAGTGTTTTGAAGAGATTGC	1436
KNTC2	NM_006101.1	NDC80	ATGTGCCAG TGAGCTTGA GT	285	TGAGCCCT GGTTAACAG TA	669	CCTTGGAGAAACACAAG CACCTGC	1053	71	ATGTGCCAGTGAGCTTGAGTCTCTGGGA GAAACACAAGCACCTGTGCTAGAAAGTAC TGTTAACAGGGGCTCA	1437
KPNA2	NM_002266.1	KPNA2	TGATGGTCC AAATGAAC GAA	286	AAGCTTCAC AAGTTGGGG C	670	ACTCCTGTTTTACCCACC ATGCCA	1054	67	TGATGGTCCAAATGAACGAATTGGCAT GGTGGTGAAACACAGGAGTTGTGCCCCA ACTTGTGAAGCTT	1438
LICAM	NM_000425.2	LICAM	CTTGCTGGC CAATGCCTA	287	TGATTGTCC GCAGTCAGG	671	ATCTACGTTGTCCAGCTG CCAGCC	1055	66	CTTGCTGGCCAATGCCTACATCTACGTT GTCCAGCTGCCAGCCAAAGATCCTGACT GCGGACAATCA	1439
LAMA3	NM_000227.2	LAMA3	CAGATGAG GCACATGG AGAC	288	TTGAAATGG CAGAACGGT AG	672	CTGATTCTCAGGTCTCT GGCCTG	1056	73	CAGATGAGGCACATGGAGACCCAGGCC AAGGACCTGAGGAATCAGTTGCTCAAC TACCGTTCTGCCATTTCAA	1440
LAMA5	NM_005560.3	LAMA5	CTCCTGGCC AACAGCAC T	289	ACACAAGGC CCAGCCTCT	673	CTGTTCTCTGGAGCATGG CCTCTTC	1057	67	CTCCTGGCCAACACAGCACTGCACATAGAA GAGGCCATGTCTCCAGGAACAGCAGAGG CTGGGCCCTGTGT	1441
LAMB1	NM_002291.1	LAMB1	CAAGGAGA CTGGGAGG TGTC	290	CGGCAGAAC TGACAGTGT TC	674	CAAGTGCCTGTACCACA CGGAAGG	1058	66	CAAGGAGACTGGGAGGTGTCTCAAGTG CCTGTACCACACACGGAAGGGGAACACTG TCAGTTCTGCCG	1442
LAMB3	NM_000228.1	LAMB3	ACTGACCA AGCCTGAG ACCT	291	GTCACACTT GCAGCATTT CA	675	CCACTCGCCATACTGGG TGCAGT	1059	67	ACTGACCAAGCCTGAGACCTACTGCAC CCAGTATGGCGAGTGGCAGATGAAATG CTGCAAGTGTGAC	1443
LAMC2	NM_005562.1	LAMC2	ACTCAAAGC GGAAATTG AAGCA	292	ACTCCCTGA AGCCGAGAC ACT	676	AGGTCTTATCAGCACAG TCTCCGCCCTCC	1060	80	ACTCAAAGCGGAAATTTGAAGCAGATAGG TCTTATCAGCACAGTCTCCGCCCTCCTGG ATTCAAGTGTCTCGGCTTCAGGGAGT	1444
LAPTM4B	NM_018407.4	LAPTM4B	AGCGATGA AGATGGTC GC	293	GACATGGCA GCACAAGCA	677	CTGGACGCGGTTCTACTC CAACAG	1061	67	AGCGATGAAGATGTGTCGCGCCCTGGAC GCGGTTCTACTCCACACAGCTGCTGCTTG TGCTGCCATGTC	1445

TABLE A

LGALS3	NM_002306.1	LGALS3	AGCGGAAA ATGGCAGA CAAT	294	CTTGAGGGT TTGGGTTTC CA	678	ACCCAGATAACGCATCA TGGAGCGA	1062	69	AGCGGAAAATGGCAGACAAATTTTTCGC TCCATGATGCGTTATCTGGGICTGGAA ACCCAAAACCCCTCAAG	1446
LIMK1	NM_016735.1		GCTTCAGGT GTTGTGACT GC	295	AAGAGCTGC CCA'TCC'TTCT C	679	TGCCTCCCTGTTCGCACCA GTACTA	1063	67	GCTTCAGGTGTTGTGACTGTCAGTGCCTC CCTGTGCGACACAGTACTATGAGAAAGGA TGGGCAGCTCTT	1447
LIMS1	NM_004987.3	LIMS1	TGAACAGT AATGGGGA GCTG	296	TTCTGGGAA CTGCTGGAA G	680	ACTGAGCGCACACGAAA CACTGCT	1064	71	TGAACAGTAATGGGAGCTGTACCATG AGCAGTGTTCGTGTGCGCTCAGTGCIT CCAGCAGTTCACAGAA	1448
LMNB1	NM_005573.1	LMNB1	TGCAAAAG CTGGTGTCA CA	297	CCCACAGAG TTCTGGTTCT TC	681	CAGCCCCCAACTGACC TCATC	1065	66	TGCAAAAGCTGGTGTACAGCCAGCCCC CCCAACTGACCTCATCTTGGAAGAACCA GAACTCGTGGGG	1449
LOX	NM_002317.3	LOX	CCAATGGG AGAACAAC GG	298	CGCTGAGGC TGGTACTGT G	682	CAGGCTCAGCAAGCTGA ACACCTG	1066	66	CCAATGGGAGAAACAACGGGCAGGTGTT CAGCTTGTGAGCCTGGGCTCACAGTA CCAGCCTCAGCG	1450
LRIG1	NM_015541.1		CTGCAACAC CGAAGTGG AC	299	GTCTCTGGA CACAGGCTG G	683	TTACTCCAGGGGACAAG CCTTCCA	1067	67	CTGCAACACCGAAGTGGACTGTTACTC CAGGGGACAAAGCCTTCCACCCCCAGCC TGTGTCCAGAGAC	1451
LSM1	NM_014462.1	LSM1	AGACCAAG CTGGAAGC AGAG	300	GAGGAATGG AAAGACCTC GG	684	CC TTCAGGGCCTGCAC TT TCAACT	1068	66	AGACCAAGCTGGAAGCAGAGAAAGTTG AAAGTGCAGGCCCTGAAGGACCGAGGT CTTTCACATTCCCTC	1452
LTBP1	NM_206943.1	LTBP1	ACATCCAG GGCTCTGTG G	301	GCAGACACA ATGGAAAGA ACC	685	CTGTGTTAGGCAC TCCC CTTGCG	1069	67	ACATCCAGGGCTCTGTGTCCGC AAGG GGAGTGCC TAAACACAGAGGGTTCTTT CCATTGTGTCTG	1453
LYRIC	NM_178812.2	MTDII	GACCTGGCC TTGCTGAAG	302	CGGACAGTT TCTTCCGGTT	686	TTCTTCTCTGTTCCTCG CTCCGG	1070	67	GACCTGGCCTTGCTGAAGAATCTCCGG AGCGAGGAACAGAAAGAAAGAAACCG GAAGAAACTGTCCG	1454
MAD1L1	NM_003550.1	MAD1L1	AGAAGCTG TCCCTGCAA GAG	303	AGCCGTACC AGCTCAGAC TT	687	CATGTTCTTCACAA TC GC TGCATCC	1071	67	AGAAGCTGTCCCTGCAAGAGCAGGATG CAGCGATTGTGAAGAACATGAAGTCTG AGCTGGTACGGCT	1455
MCM2	NM_004526.1	MCM2	GACTTTTGC CCGCTACCT TTC	304	GCCACTAAC TGCTTCAGT ATGAAGAG	688	ACAGCTCATTGTTGT CAC GCCGGA	1072	75	GACTTTTGCCCGCTACCTTTCATTCCGG CGTGACAAACAATGAGCTGTGTCTCTTC ATACTGAAGCAGTTAGTGGC	1456
MELK	NM_014791.1	MELK	AGGATCGC CTGTCAGAA GAG	305	TGCACATAA GCAACAGCA GA	689	CCCCGGTTGTCTTCCGTC AGATAG	1073	70	AGGATCGCCTGTCTAGAAAGAGAGACCC GGGTGTCTTCCGTCAGATAGTATCTGC TGTTGCTTATGTGCA	1457
MGMT	NM_002412.1	MGMT	GTGAAATG AAACGCAC CACA	306	GACCC TGCT CACAAACCAG AC	690	CAGCCCTTTTGGGGAAGC TGG	1074	69	GTGAAATGAACCGCACCCACACTGGACA GCCCTTTGGGGAAGCTGGAGCTGTCTG GTTGTGAGCAGGGTTC	1458
mGSTI	NM_020300.2	MGSTI	ACGGATCTA CCACACCAT TGC	307	TCCATATCC AACAAAAAA ACTCAAAG	691	TTTGACACCCCTTCCCCA GCCA	1075	79	ACGGATCTACCACACCAATTGCA TATTG ACACCCCTTCCCCAGCCAAATAGAGCT TTGAGTTTTTTTGTGTGATATGGA	1459

TABLE A

MMP1	NM_002421.2	MMP1	GGGAGATC ATCGGGAC AACTC	308	GGGCGTGGT TGAAAAAGCA T	692	AGCAAGATTTCCTCCAG GTCCATCAAAAGG	1076	72	GGGAGATCATCGGACAAACTCTCCCTTT TGATGGACCTGGAGGAAATCTTGCTCA TGCTTTTCAACCAAGGCC	1460
MMP12	NM_002426.1	MMP12	CCAACGCTT GCCAAATCC T	309	ACGGTAGTG ACAGCATCA AAACTC	693	AACCAGCTCTCTGTGAC CCCAATT	1077	78	CCAACGTTGCCAAATCCTGACAAATTC AGAACCAGCTCTCTGTGACCCCAATTT GAGTTTGTGCTGTCTACTACCGT	1461
MMP2	NM_004530.1	MMP2	CCATGATGG AGAGGCAG ACA	310	GGAGTCCGT CCTTACCGT CAA	694	CTGGGAGCATGGCGATG GATACCC	1078	86	CCATGATGGAGAGGCAGACATCATGAT CAACTTTGGCCGCTGGGAGCATGGCGA TGGATACCCCTTTTGACGGTAAGGACGG ACTCC	1462
MMP7	NM_002423.2	MMP7	GGATGGTA GCAGTCTAG GGATTAACT	311	GGAAATGTCC CATACCCAA AGAA	695	CCTGTATGCTGCAACTCA TGAAC TTGGC	1079	79	GGATGGTAGCAGTCTAGGGATTAACTT CCTGTATGCTGCAACTCATGAACTTGGC CATTC TTGGGTATGGGACATTCC	1463
MMP8	NM_002424.1	MMP8	TCACCTCTC ATCTTCACC AGGAT	312	TGTCACCGT GATCTCTTT GGTAA	696	AAGCAATGTTGATATCT GCCTCCTCCCTG	1080	79	TCACCTCTCATCTTTCACCAAGGATCTCAC AGGGAGAGGCAGATATCAACATTGCTT TTTACCAAAAGAGATCACGGTGACA	1464
MMPV-like env	AF346816.1		CCATACGTG CTGCTACCT GT	313	CCTAAAGGT TTGAATGGC AGA	697	TCATCAAAACCATGGTTC ATCACCAATATC	1081	72	CCATACGTGCTGTACTACCTGTAGATATTG GTGATGAACCATGGTTTGATGATTCTGC CATTCAAACCTTTAGG	1465
MNAT1	NM_002431.1	MNAT1	CGAGAGTCT GTAGGAGG GAAACC	314	GGTCCGAT ATTTGGTGG TCCTAC	698	CGAGGGCAACCCCTGATC GTCCA	1082	75	CGAGAGTCTGTAGGAGGGAAACCGCCA TGGACGATCAGGGTTGCCCTCGGTGTA AGACCACCAATATCGGAAACC	1466
MRP1	NM_004996.2	ABCC1	TCATGGTGC CGATCTACT TCCT	315	CGATTGCT TTGCTCTTCA TGIG	699	ACCTGATACGTCTTGGTC TTCATCGCCAT	1083	79	TCATGGTGCCCGTCAATGGCTGTGATGG CGATGAAGACCAAGACGTATCAGGTGG CCCACATGAAGAGCAAAAGACAATCG	1467
MRP3	NM_003786.2	ABCC3	TCATCCTGG CGATCTACT TCCT	316	CCGTTGAGT GGAATCAGC AA	700	TCGTCTCTGGCTGGAGTC GCTTTCAT	1084	91	TCATCCTGGCGATCTACTTCTCTGGCA GAACTAGTCCCTCTGTCTCTGGCTGG AGTCGCTTTCATGGTCTTGCTGATTCCA CTCAACGG	1468
MS4A1	NM_021950.2	MS4A1	TGAGAAAC AAACTGCA CCCA	317	CAAGGCCTC AAATCTCAA GG	701	TGAACTCCGCAGCTAGC ATCCAAA	1085	70	TGAGAAACAAACTGCACCCACTGAAC CCGCAGCTAGCATCCAAATCAGCCCTT GAGATTGAGGCCTTG	1469
MSH2	NM_000251.1	MSH2	GATGCAGA ATTGAGGC AGAC	318	TC TTGGCAA GTCGGTTAA GA	702	CAAGAAAGATTTACTTCG TCGATCCCCAGA	1086	73	GATGCAGAATTGAGGCAGAC TTTACAA GAAGATTTACTTCGTCGATTCCCAGATC TTAACCGACTTGCCAAAGA	1470
MTA3	XM_038567		GCTCGTGGT TCTGTAGTC CA	319	ACAAAGGGA GAGCGTGAA GT	703	TCAGTCAACATCACCCCTC CTAGGATGA	1087	69	GCTCGTGGTTCTGTAGTCCAGTCATCCT AGGAGGCTGATGTTGACTGAGACTTCA CGCTCTCCCTTTGT	1471
MX1	NM_002462.2	MX1	GAAAGGAAT GGGAATCA GTCATGA	320	GTCTATTAG AGTCAGATC CGGGACAT	704	TCACCTGGAGATCAGC TCCCCGA	1088	78	GAAAGGAATGGGAATCAGTCAIGAGCTA ATCACCTGGAGATCAGCTCCCGAGAT GTCCCGGATCTGACTCTAATAGAC	1472
MYBL2	NM_002466.1	MYBL2	GCCGAGAT CGCCAAGA TG	321	CTTTTGATG GTAGAGTTTC CAGTGATTC	705	CAGCATTGCTGTCTCTCC CTGGCA	1089	74	GCCGAGATCGCCAAAGATGTTGCCAGGG AGGACAGACAATGCTGTGAAGAATCAC TGGAAC TCTACCATCAAAAAG	1473

TABLE A

NAT1	NM_000662.4	NAT1	TGGTTTGA GACCACGA TGT	322	TGAATCATG CCAGTGTCTG TA	706	TGGAGTGTCTGTAAACAT ACCCTCCCA	1090	75	TGGTTTTGAGACCACGATGTTGGGAGG GTATGTTTACAGCACTCCAGCCAAAA ATACAGCACTGGCATGATTCA	1474
NAT2	NM_000015.1	NAT2	TAACTGACA TTCTTGAGC ACCAGAT	323	ATGGCTTGC CCACAATGC	707	CGGGCTGTTCCTTTGAG AACCTTAAACA	1091	73	TAACTGACATTCCTTGAGCACCAGATCC GGGCTGTTCCTTTTGAGAACCTTAAACAT GCATTGTGGCAAGCCAT	1475
NRG1	NM_013957.1	NRG1	CGAGACTCT CCTCATAGT GAAAGGTA T	324	CTTGGCGTG TGGAAATCT ACAG	708	ATGACCACCCCGGCTCG TATGTCA	1092	83	CGAGACTCTCCTCATAGTGAAGGTAT GTGTCAGCCCATGACCACCCCGGCTCGT ATGTCACCTGTAGATTTCACACGCCA AG	1476
OPN, osteopontin	NM_000582.1	SPP1	CAACCGAA GTTTTCACT CCAGTT	325	CCTCAGTCC ATAAACCCAC ACTATCA	709	TCCCACACAGTAGACACA TATGATGGCCG	1093	80	CAACCGAAGTTTTCACTCCAGTTGTCC CACAGTAGACACATATGATGGCCGAGG TGATAGTGTGGTTTATGGACTGAGG	1477
p16-INK4	L27211.1		GCGGAAGG TCCCTCAGA CA	326	TGATGATCT AAGTTTCCC GAGGTT	710	CTCAGAGCCTCTCTGGTT CTTTCAAATCGG	1094	76	GCGGAAGTCCCTCAGACATCCCCGAT TGAAAGAACCAAGAGGGCTCTGAGAA ACCTCGGGAACCTTAGATCAACA	1478
PAI1	NM_000602.1	SERPINE1	CCGCAACGT GGTTTCTC A	327	TGCTGGGTT TCTCCTCCTG TT	711	CTCGGTGTGGCCATGCT CCAG	1095	81	CCGCAACGTGGTTTTTCTCACCCCTATGGG GTGGCCTCGGTGTGGCCATGCTCCAG CTGACAACAGGAGGAGAAACCCAGCA	1479
PGF	NM_002632.4	PGF	GTGGTTTTC CCTCGGAGC	328	AGCAAGGGA ACAGCCTCA T	712	ATCTTCTCAGACGTCCCG AGCCAG	1096	71	GTGGTTTTCCCTCGGAGCCCCCTGGCTC GGGACGTCTGAGAAAGATGCCGGTCATG AGGCTGTCCCTTGT	1480
PR	NM_000926.2	PGR	GCATCAGG CTGTCACTA TGG	329	AGTAGTTGT GCTGCCCTT CC	713	TGTCCTTACCTGTGGGAG CTGTAAGTTC	1097	85	GCATCAGGCTGTCAATTATGGTGTCTTA CCTGTGGGAGCTGTAAAGGTCTTCTTTAA GAGGGCAATGGAAGGGCAGCACAACT ACT	1481
PRDX1	NM_002574.2	PRDX1	AGGACTGG GACCCATG AAC	330	CCCATAAATC CTGAGCAAT GG	714	TCCTTTGGTATCAGACCC GAAGCG	1098	67	AGGACTGGGACCCATGAACATTCCTTT GGTATCAGACCCCGAAGCGCACCATTGC TCAGGATTATGGG	1482
PTEN	NM_000314.1	PTEN	TGGCTAAGT GAAGATGA CAATCATG	331	TGCACATAT CATTACACC AGTTCGT	715	CCTTTCCAGCTTTTACAGT GAATTGCTGCA	1099	81	TGGCTAAGTGAAGATGACAATCATGTT GCAGCAATTCACCTGTAAAGCTGGAAAG GGACGAACCTGGTGTAAATGATATGTCA	1483
PTP4A3	NM_007079.2	PTP4A3	AATATTGT GCGGGGTA TGG	332	AACGAGATC CCTGTGCTT GT	716	CCAAGAGAAACGAGATT TAAAAACCCACC	1100	70	AATATTGTGCGGGGTATGGGGGTGGG TTTTTAAATCTCGTTTCTCTTTGGACAAG CACAGGGATCTCGTT	1484
RhoB	NM_004040.2	RHOB	AAGCATGA ACAGGACTT GACC	333	CCTCCCCAA GTCAGTTGC	717	CTTTCCAACCCCTGGGG AAGACAT	1101	67	AAGCATGAACAGGACTTGACCATCTTT CCAAACCCCTGGGGAAGACATTTGCAAC TGACTTGGGGAGG	1485
RPL13A	NM_012423.2	RPL13A	GCAAGGAA AGGGTCTTA GTCAC	334	ACACC TGCA CAATTCTCC G	718	CCTCCCGAAGTTGCTTGA AAGCAC	1102	68	GCAAGGAAAAGGGTCTTAGTACATGCCCT CCCGAAGTTGCTTGAAAGCACTCGGAG AATTGTGCAGGTGT	1486
RPL41	NM_021104.1	RPL41	GAAACCTCT GCGCCATG A	335	TTCTTTTGC CTTCAGCC	719	CATTGCTTCTTCCTCCA CTTGGC	1103	66	GAAACCTCTGCGCCATGAGAGCCAAGT GGAGGAAGAACGGAATGCCGAGGCTG AAGCGCAAAAAAGAA	1487

TABLE A

RPLP0	NM_001002.2	RPLP0	CCATTCTAT CATCAACG GGTACAA	336	TCAGCAAGT GGGAAGGTG TAATC	720	TCGCCACAGACAAGGCC AGGACTCG	1104	75	CCATTCTATCATCAACGGGTACAAACG AGTCTGGCCCTTGCTGTGTGGAGACGGA TTACACCTTCCCACCTTGCTGA	1488
RPS23	NM_001025.1	RPS23	GTTCTGGTT GCTGGATTT GG	337	CCTTAAAGC GGACTCCAG G	721	ATCACCAACAGCATGAC CTTTGCG	1105	67	GTTCTGGTTGCTGGATTGCTGCGCAAAG GTCAATGCTGTGTGGTGATATTCCTGGAGT CCGCTTTAAGG	1489
RPS27	NM_001030.3	RPS27	TCACCACGG TCTTTAGCC A	338	TCCTCCTGT AGGCTGGCA	722	AGGACAGTGGAGCAGCC AACACAC	1106	80	TCACCACGGTCTTTAGCCATGCACAAA CGGTAGTTTTTGTTGTGTGGCTGTCCAC TGTCCTCTGCCAGCCTACAGGAGGA	1490
RRM1	NM_001033.1	RRM1	GGGCTACTG GCAGCTAC ATT	339	CTCTCAGCA TCGGTACAA GG	723	CATTGGAATTGCCATTA GTCCCAGC	1107	66	GGGCTACTGGCAGCTACATTGCTGGGA CTAATGGCAATTCCAAATGGCCTTGTACC GATGCTGAGAG	1491
RRM2	NM_001034.1	RRM2	CAGCGGGA TTAAACAGT CCT	340	ATCTGCGTT GAAGCAGTG AG	724	CCAGCACAGCCAGTTAA AAGATGCA	1108	71	CAGCGGGATTAAACACAGTCCTTTAACCA GCACAGCCAGTTAAAGATGCAGCCTC ACTGCTTCAACGCAGAT	1492
RUNX1	NM_001754.2	RUNX1	AACAGAGA CATTGCCAA CCA	341	GTGATTGTC CCAGGAAAG TTT	725	TTGGATCTGCTTGTGCTC CAAACC	1109	69	AACAGAGACATTGCCAACCATATTGGA TCTGCTTGTGTCTCCAAACCAGCAAACCT CCTGGGCAAAATCAC	1493
S100A10	NM_002966.1	S100A10	ACACCAA ATGCCATCT CAA	342	TTTATCCCC AGCGAATTT GT	726	CACGCCATGGAAACCCAT GATGTTT	1110	77	ACACCAAATGCCAATCICAAAATGGAAC ACGCCATGGAAACCATGATGTTTACAT TTCACAAATTCGCTGGGATAAA	1494
S100A2	NM_005978.2	S100A2	TGGCTGTGC TGGTCACTA CCT	343	TCCCCCTTA CTCAGCTTG AACT	727	CACAAAGTACTCCTGCCA AGAGGGCGAC	1111	73	TGGCTGTGCTGGTCACTACCTTCCACAA GTACTCCTGCCAAGAGGGCGACAAAGTT CAAGCTGAGTAAGGGGGA	1495
S100A4	NM_002961.2	S100A4	GACTGCTGT CATGGCGTG	344	CGAGTACTT GTGGAAGGT GGAC	728	ATCACATCCAGGGCCCTT CTCCAGA	1112	70	GACTGTGTGTCATGGGTGCCCTCTGGA GAAGGCCCTGGATGTGATGGTGTCCAC CTTCCACAAAGTACTCG	1496
S100A7	NM_002963.2	S100A7	CCTGCTGAC GATGATGA AGGA	345	GCGAGGTAA TTTGTGCCCT TT	729	TTCCCCAACTTCCTTAGT GCCTGTGACA	1113	75	CCTGCTGACGATGATGAAGGAGAACTT CCCCAACTTCCTTAGTGCCTGTGACAAA AAGGGCACAAATTACCTCGC	1497
S100A8	NM_002964.3	S100A8	ACTCCCTGA TAAAGGG AATTT	346	TGAGGACAC TCGGTCTCT AGC	730	CATGCCGTCTACAGGGA TGACCTG	1114	76	ACTCCCTGATAAAGGGGAATTTCCATG CCGTCTACAGGGATGACCTGAAGAAAT TGCTAGAGACCGAGTGTCTCA	1498
S100A9	NM_002965.3	S100A9	CACCCTGCC TCTACCCAA C	347	CTAGCCCCA CAGCCAAGA	731	CCCGGGCCTGTATATGTC AAACT	1115	67	CACCCTGCCCTTACCCAAACCAGGGCCC CGGGGCCCTGTATATGTCAAACTGTCTTGG CTGTGGGGCTAG	1499
S100B	NM_006272.1	S100B	CATGGCCGT GTAGACCTT AA	348	AGTTTTAAG GGTGCCCCG	732	CCGGAGGGAACCCCTGAC TACAGAA	1116	70	CATGGCCGTGTAGACCCCTAACCCGGAG GGAACCCCTGACTACAGAAATTACCCCG GGGCACCCCTTAAAACT	1500
S100G	NM_004057.2	S100G	ACCCTGAGC ACTGGAGG AA	349	GAGACTTTG GGGGATTCC A	733	AGGATAAGACCACAGCA CAGGCGC	1117	67	ACCCTGAGCACTGGAGGGAAGAGCGCCT GTGCTGTGCTCTTATCTATCTATGTGGAATC CCCCAAAAGTCTC	1501

TABLE A

S100P	NM_005980.2	S100P	AGACAAAGG ATGCCGTGG ATAA	350	GAAGTCCAC CTGGGCATC TC	734	TTTGCTCAAGGACCTGGA CGCCAA	1118	67	AGACAAAGGATGCCGTGGATAAAATTGCT CAAGGACCTGGACGCCAATGGAGATGC CCAGGTGGACTTC	1502
SDHA	NM_004168.1	SDHA	GCAGAACT GAAGATGG GAAGAT	351	CCCTTTCCA AACTTGAGG C	735	CTGTCCACCAAAATGCAC GCTGATA	1119	67	GCAGAACTGAAGATGGGAAGATTATC AGCGTGCAATTTGGTGGACAGAGCCTCA AGTTTGGAAAGGG	1503
SEMA3F	NM_004186.1	SEMA3F	CGCGAGCC CCTCATTAT ACA	352	CACTCGCCG TTGACATCC T	736	CTCCCCACAGCGCATCG AGGAA	1120	86	CGCGAGCCCTCATTATATACACTGGGCA GCCTCCCCACAGCGCATCGAGGAAATGC GTGCTCTCAGGCAAGGATGTCAACGGC GAGTG	1504
SFRP2	NM_003013.2	SFRP2	CAAGCTGA ACGGTGTGT CC	353	TGCAAGCTG TC1TTGAGC C	737	CAGCACCGATTCTTTCAG GTCCCT	1121	66	CAAGCTGAACGGTGTGTCTCCGAAAGGGA CCTGAAGAAATCGGTGCTGTGGCTCAA AGACAGCTTGCA	1505
SIR2	NM_012238.3	SIR11	AGCTGGGG TG1CTGT1T CAT	354	ACAGCAAGG CGAGCATAA AT	738	CCTGACTTCAGGTCAA GGATGG	1122	72	AGCTGGGGTGTCTGT1TTCATGTGGAAT ACCTGACTTTCAGGTCAAAGGGATGGTAT 11TA1GCT1CGCCT11GCTGT	1506
SKIL	NM_005414.2	SKIL	AGAGGCTG AATATGCA GGACA	355	CTATCGGCC TCAGCATGG	739	CCAATCTCTGCCCTCAGTT CTGCCA	1123	66	AGAGGCTGAATATGCAGGACAGTTGGC AGAACTGAGGCAGAGATTGGACCATGC TGAGGCCGATAG	1507
SKP2	NM_005983.2	SKP2	AGTTGCAG AATCTAAGC CTGGAA	356	TGAGT1TT1T GCGAGAGTA TTGACA	740	CCTGCGGCTTTCGGATCC CA	1124	71	AGTTGCAGAA1TCTAAGCCTGGAAAGGCC TGCGGCTTTCGGATCCCATTTGTCAATAC TCTCGCAAAAAACTCA	1508
SLP1	NM_003064.2	SLP1	ATGGCCAAT GTTTGATGC T	357	ACACTTCAA GTCACGCTT GC	741	TGGCCATCCATCTCACA GAAATTGG	1125	74	ATGGCCAATGTTTGATGCTTAACCCCC CAATTTCTGTGAGATGGATGGCCAGTG CAAGCGTGACTTGAAAGTG	1509
SNAI1	NM_005985.2	SNAI1	CCCAATCGG AAGCCTAA CTA	358	GTAGGGCTG CTGGAAGGT AA	742	TC1GGATTAGAGTCC1GC AGCTCGC	1126	69	CCCAATCGGAAGCCTAACTACAGCGAG CTGCAGGACTCTAATCCAGAGTTTACCT TCCAGCAGCCCTAC	1510
STK15	NM_003600.1	AURKA	CATCTTCCA GGAGGACC ACT	359	TCCGACCTT CAATCATTT CA	743	CTCTGTGGCACCCCTGGA CTACCTG	1127	69	CATCTTCCAGGAGGACCAC1CTCTGTG GCACCTTGGACTACCTGCCCCCCTGAAA TGATTGAAGGTCCGA	1511
STMN1	NM_005563.2	STMN1	AATACCCA ACGCACAA ATGA	360	GGAGACAAT GCNAAACCAC AC	744	CACGT1CTCTGCCCCGTT TCTTG	1128	71	AATACCCAACGCACAAATGACCCGACG TTCTCTGCCCCGTTTCTTGGCCCCAGTGT GGTTTGCATTGTCTCC	1512
STMY3	NM_005940.2	MMP11	CCTGGAGG CTGCAACAT ACC	361	TACAA1TGGC TTTGAGGA TAGCA	745	ATCCTCCTGAAGCCCTTT TCGCAGC	1129	90	CCTGGAGGCTGCAACATACCTCAATCC TG1CCCCAGGCCGGATCCTCCTGAAGCC CTTTTCGCAGCACTGCTATCCTCCCAAAG CCATTGTA	1513
SURV	NM_001168.1	BIRC5	TG11111GAT TCCCCGGCT TA	362	CAAAGCTGT CAGCTCTAG CAAAAAG	746	TGCCTTCTTCCCTCCTCA CTTCTCACCT	1130	80	1G11111GAT1CCC1GGGCT1ACCAGG1GA GAAGTGAGGGAGGAAGAAGGCAGTGT CCCTTTTGCTAGAGCTGACAGCTT1G	1514
SYK	NM_003177.1	SYK	TC1CCAGCA AAAGCGAT GTCT	363	TTCATCCCTC GATATGGCT TCT	747	CCATAGGAGAATGCTTC CCACATCAACACT	1131	85	TC1CCAGCAAAAAGCGATGTCTGGAGCT TTGGAGTGT1GATGTGGGAAGCA1TTCT CCTATGGGCAGAAGCCATATCGAGGGGA TGAA	1515

TABLE A

TAGLN	NM_003186.2		TAGLN	GATGGAGC AGGTGGCTC AGT	364	AGTCTGGAA CATGTCAGT CTTGATG	748	CCCATAGTCCCTCAGCCG CCTTCAG	1132	73	GATGGAGCAGGTGGCTCAGTTCCCTGAA GGCGGCTGAGGACTCTGGGGTCATCAA GACTGACATGTTCCAGACT	1516
TCEA1	NM_201437.1		TCEA1	CAGCCCTGA GGCAAGAG A	365	CGAGCATTT GTCTCATCC TTT	749	CTTCCAGCGGCAATGTA AGCAACA	1133	72	CAGCCCTGAGGCAAGAGAAAGTA CTTCCAGCGGCAATGT'AAGCAACAGAA AGGATGAGACAAATGCTCG	1517
TFRC	NM_003234.1		TFRC	GCCAACTGC TTTCATTTG TG	366	ACTCAGGCC CATTTCCCTTT A	750	AGGGATCTGAACCAATA CAGAGCAGACA	1134	68	GCCAACTGCTTTTCATTTGTGAGGGATCT GAACCAATACAGAGCAGACATAAAGG AAATGGGCTGAGT	1518
TGFB2	NM_003238.1		TGFB2	ACCAGTCCC CCAGAAAGA CTA	367	CCTGGTGCT GTTGTAGAT GG	751	TCCTGAGCCCGAGGAAG TCCC	1135	75	ACCAGTCCCCCAGAAAGACTATCCTGAG CCCGAGGAAGTCCCCCGGAGGTGATT TCCATCTACAAACAGCACCAGG	1519
TGFB3	NM_003239.1		TGFB3	GGATCGAG CTCTTCCAG ATCCT	368	GCCACCAGT ATAGCGCTG TT	752	CGGCCAGATGAGCACAT TGCC	1136	65	GGATCGAGCTCTTCCAGATCCTTCGGGC AGATGAGCACATTTGCCAAACACAGCGCTA TATCGGTGGC	1520
TGFB2	NM_003242.2		TGFB2	AACACCAA TGGGT'TCCA TCT	369	CCTCTTCATC AGGCCAAAC T	753	TTCTGGGCTCCTGATTGC TCAAGC	1137	66	AACACCAATGGGTTCATCTTTCTGGGC TCTTGATTGCTCAAGCACAGTTTGGCCT GATGAAGAGG	1521
TIMP3	NM_000362.2		TIMP3	CTACCTGCC TTGCTTTGT GA	370	ACCGAAAT GGAGAGCAT GT	754	CCAAGAACGAGTGTCTC TGGACCG	1138	67	CTACCTGCCTTTGCITTTGTGACTTCCCAAG AACGAGTGTCTCTGGACCGACATGCTC TCCAAATTTCCGT	1522
TNFRSF11 A	NM_003839.2		TNFRSF11 A	CCAGCCCAC AGACCAGTT A	371	TTCAGAGAA AGGAGGTGT GGA	755	TGTTCCCTCACTGAGCCTG GAAGCA	1139	67	CCAGCCCACAGACCAGTTACTGTTCCTC ACTGAGCCTGGAAAGCAAATCCACACCT CCTTTCCTCTGAA	1523
TNFRSF11 B	NM_002546.2		TNFRSF11 B	TGGCGACC AAGACACC TT	372	GGGAAAGTG GTACGTCTT TGAG	756	AGGGCCTAATGCACGCA CTAAAGC	1140	67	TGGCGACCAAGACACCTTGAAAGGGCCT AATGCACCGCACTAAAGCACTCAAAGAC GTACCACCTTCCC	1524
TNFSF11	NM_003701.2		TNFSF11	CATATCGTT GGATCACA GCAC	373	TTGGCCAGA TCTAACCAT GA	757	TCCACCATCGCTTTCTCT GCTCTG	1141	71	CATATCGTTGGATCACAGCACATCAGA GCAGAGAAAGCGATGTTGGATGGCTCA TGTTTAGATCTGGCCAA	1525
TWIST1	NM_000474.2		TWIST1	GCGCTGCG GAAGATCA TC	374	GCTTGAGGG TCTGAATCT TGCT	758	CCACGCTGCCCTCGGAC AAGC	1142	64	GCGCTGCGGAAGATCATCCCCACGCTG CCCTCGGACAAAGCTGAGCAAGATTTCAG ACCCTCAAAGC	1526

TABLE A

UBB	NM_018955.1	UBB	GAGTCGAC CCTGCACCT G	375	GCGAATGCC ATGACTGAA	759	AATTAACAGCCACCCCT CAGGCG	1143	522	GAGTCGACCCCTGCACCTGGTCCTGCGT CTGAGAGGTGGTATGACAGATCTTCGTG AAGACCCCTGACCGGCAAGACCATCACCC CTGGAAGTGGAGGCCAGTGACACCCATC GAAAATGTGAAGGCCAAGATCCAGGAT AAGAAAGCATCCCTCCCGACACAGCAG AGGCTCATCTTTTGCAAGCAAGCAGCTG GAAGATGGCCGCACCTCTTTCTGACTAC AACATCCAGAAGGAGTCGACCCCTGCAC CTGGTCCTGCTCTGAGAGGTGGTATG CAGATCTTCGTGAAGACCCCTGACCCGC AAGACCATCACTCTGGAAGTGGAGCCC AGTGACACCATCGAAAATGTGAAGGCC AAGATCCAAGATAAAGAAAGCATCCCT CCCGACAGCAGAGGCTCATCTTTGCA GGCAAGCAGCTGGAAGATGGCCGCACT CTTCTGACTACAACATCCAGAAAGAG TCGACCCCTGCACCTGGTCCCTGCGCCTGA GGGTGGCTGTTAATTTCTTCAGTCAATGG CATTCGC	1527
VCAM1	NM_001078.2	VCAM1	TGGCTTCAG GAGCTGAA TACC	376	TGCTGTCTGT GATGAGAAA ATAGTG	760	CAGGCACACACAGGTGG GACACAAAT	1144	89	TGGCTTCAGGAGCTGAATACCCCTCCCA GGCACACACAGGTGGGACACAAAATAA GGGTTTTTGGAAACCACATAATTTTCTCATCA CGACAGCA	1528
VIM	NM_003380.1	VIM	TGCCCTTAA AGGAACCA ATGA	377	GCTTCAAACG GCAAAAGTTC TCTT	761	ATTTACGCACTCTGGCGT TCCA	1145	72	TGCCCTTAAAGGAACCAATGAGTCCCT GGAACGCCAGATGCGGTGAAATGGAAAG AGAACTTTGCCGTTGAAGC	1529
VTN	NM_000638.2	VTN	AGTCAAATCT TCGCACACG G	378	GTA CTGAGC GATGGAGCG T	762	TGGACACTGTGGACCCCT CCCTACC	1146	67	AGTCAAATCTTCGCACACAGCGGAGTGGA CACTGTGGACCCCTCCCTACCCACGCTCC ATCGCTCAGTAC	1530
WAVE3	NM_006646.4	WASF3	CTCTCCAGT GTGGGCAC C	379	GCGGTGTAG CTCCCAGAG T	763	CCAGAAACAGATGCGGAGC AGTCCAT	1147	68	CTCTCCAGTGTGGGCACACAGCCGGCCA GAACAAGATGCGAGCAGTCCAATGACTCT GGGAGCTACACCCG	1531
WISP1	NM_003882.2	WISP1	AGAGGGCAT CCATGAACT TCACA	380	CAAACTCCA CAGTACTTG GGTTGA	764	CGGGCTGCATCAGCACCA CGC	1148	75	AGAGGCATCCATGAACCTTCACACTTGC GGGTGCATCAGCACACGCTCCTATCA ACCCAAGTACTGTGGAGTTTG	1532
Wnt-5a	NM_003392.2	WNT5A	GTATCAGG ACCACATGC AGTACATC	381	TGTCGGAAT TGATACTGG CATT	765	TTGATGCCGTGTCTTCGCG CCTTCT	1149	75	GTATCAGGACCACATGCACTACATCGG AGAAGCGCGAAGACAGGCATCAAAG AATGCCAGTATCAATTCCGACA	1533
Wnt-5b	NM_032642.2	WNT5B	TGTCCTCAG GGTCTTGTC CA	382	GTGCACGTG GATGAAAGA GT	766	TTCCGTAAGAGGCCCTGG TGCTCTC	1150	79	TGTCCTCAGGGTCTTTGTCCAGAATGTAG ATGGGTCCGTAAAGAGGCCCTGGTGCTC TCTTACTCTTTTCATCCACGTGCAC	1534
WWOX	NM_016373.1	WWOX	ATCGCAGCT GGTGGGTGT AC	383	AGCTCCCTG TTGCATGGA CTT	767	CTGCTGTTTACCCTTGGCG AGGCTTTC	1151	74	ATCGCAGCTGGTGGGTGTACACACTGC TGTTTACCTTGGCGAGGCCCTTTCACCAA GTCCATGCAACAGGGAGCT	1535

TABLE A

YWHAZ	NM_003406.2	YWHAZ	GTGGACATC GGATAACC AAG	384	GCAGACAAA AGTTGGAAG GC	768	CCCCCTCCTTCTCCTGCTT CAGCTT	1152	81	GTGGACATCGGATACCCCAAGGAGACGA AGCTGAAAGCAGGAGAAGAGGGGAAA ATTAAACCGGCCTTCCAACTTTTGTCTGC	1536
-------	-------------	-------	------------------------------	-----	------------------------------	-----	-------------------------------	------	----	---	------

Table 1: Cox proportional hazards for Prognostic Genes that are positively associated with good prognosis for breast cancer (Providence study)

Table 1			
Gene_all	z (Coef)	HR	p (Wald)
GSTM2	-4.306	0.525	0.000
IL6ST	-3.730	0.522	0.000
CEGP1	-3.712	0.756	0.000
Bcl2	-3.664	0.555	0.000
GSTM1	-3.573	0.679	0.000
ERBB4	-3.504	0.767	0.000
GADD45	-3.495	0.601	0.000
PR	-3.474	0.759	0.001
GPR30	-3.348	0.660	0.001
CAV1	-3.344	0.649	0.001
C10orf116	-3.194	0.681	0.001
DR5	-3.102	0.543	0.002
DICER1	-3.097	0.296	0.002
FstR1	-2.983	0.825	0.003
BTRC	-2.976	0.639	0.003
GSTM3	-2.931	0.722	0.003
GATA3	-2.874	0.745	0.004
DLC1	-2.858	0.564	0.004
CXCL14	-2.804	0.693	0.005
IL17RB	-2.796	0.744	0.005
C8orf4	-2.786	0.699	0.005
FOXO3A	-2.786	0.617	0.005
TNFRSF11B	-2.690	0.739	0.007
BAG1	-2.675	0.451	0.008
SNAI1	-2.632	0.692	0.009
TGFB3	-2.617	0.623	0.009
NAT1	-2.576	0.820	0.010
FUS	-2.543	0.376	0.011
F3	-2.527	0.705	0.012
GSTM2 gene	-2.461	0.668	0.014
EPHB2	-2.451	0.708	0.014
LAMA3	-2.448	0.778	0.014
BAD	-2.425	0.506	0.015
IGF1R	-2.378	0.712	0.017
RUNX1	-2.356	0.511	0.018
ESRRG	-2.289	0.825	0.022
HSHIN1	-2.275	0.371	0.023
CXCL12	-2.151	0.623	0.031
IGFBP7	-2.137	0.489	0.033
SKIL	-2.121	0.593	0.034
PTEN	-2.110	0.449	0.035
AKT3	-2.104	0.665	0.035
MGMT	-2.060	0.571	0.039
LRIG1	-2.054	0.649	0.040
S100B	-2.024	0.798	0.043
GREB1 variant a	-1.996	0.833	0.046
CSF1	-1.976	0.624	0.048
ABR	-1.973	0.575	0.048
AK055699	-1.972	0.790	0.049

Table 2: Cox proportional hazards for Prognostic Genes that are negatively associated with good prognosis for breast cancer (Providence study)

Table 2			
Gene_all	z (Coef)	HR	p (Wald)
S100A7	1.965	1.100	0.049
MCM2	1.999	1.424	0.046
Contig 51037	2.063	1.185	0.039
S100P	2.066	1.170	0.039
ACTR2	2.119	2.553	0.034
MYBL2	2.158	1.295	0.031
DUSP1	2.166	1.330	0.030
HOXB13	2.192	1.206	0.028
SURV	2.216	1.329	0.027
MELK	2.234	1.336	0.026
HSPA8	2.240	2.651	0.025
cdc25A	2.314	1.478	0.021
C20_orf1	2.336	1.497	0.019
LMNB1	2.387	1.682	0.017
S100A9	2.412	1.185	0.016
CENPA	2.419	1.366	0.016
CDC25C	2.437	1.384	0.015
GAPDH	2.498	1.936	0.012
KNTC2	2.512	1.450	0.012
PRDX1	2.540	2.131	0.011
RRM2	2.547	1.439	0.011
ADM	2.590	1.445	0.010
ARF1	2.634	2.973	0.008
E2F1	2.716	1.486	0.007
TFRC	2.720	1.915	0.007
STK15	2.870	1.860	0.004
LAPTM4B	2.880	1.538	0.004
EpCAM	2.909	1.919	0.004
ENO1	2.958	2.232	0.003
CCNB1	3.003	1.738	0.003
BUB1	3.018	1.590	0.003
Claudin 4	3.034	2.151	0.002
CDC20	3.056	1.555	0.002
Ki-67	3.329	1.717	0.001
KPNA2	3.523	1.722	0.000
IDH2	3.994	1.638	0.000

Table 3: Cox proportional hazards for Prognostic Genes that are positively associated with good prognosis for ER-negative (ER0) breast cancer (Providence study)

Gene_ER0	HR	z (Coef)	p (Wald)
SYK	0.185	-2.991	0.003
Wnt-5a	0.443	-2.842	0.005
WISP1	0.455	-2.659	0.008
CYR61	0.405	-2.484	0.013
GADD45	0.520	-2.474	0.013
TAGLN	0.364	-2.376	0.018
TGFB3	0.465	-2.356	0.018
INHBA	0.610	-2.255	0.024
CDH11	0.584	-2.253	0.024
CHAF1B	0.551	-2.113	0.035
ITGAV	0.192	-2.101	0.036
SNAI1	0.655	-2.077	0.038
IL11	0.624	-2.026	0.043
KIAA1199	0.692	-2.005	0.045
TNFRSF11B	0.659	-1.989	0.047

Table 4: Cox proportional hazards for Prognostic Genes that are negatively associated with good prognosis for ER-negative (ER0) breast cancer (Providence study)

Gene_ER0	HR	z (Coef)	p (Wald)
RPL41	3.547	2.062	0.039
Claudin 4	2.883	2.117	0.034
LYRIC	4.029	2.364	0.018
TFRC	3.223	2.596	0.009
VTN	2.484	3.205	0.001

Table 5: Cox proportional hazards for Prognostic Genes that are positively associated with good prognosis for ER-positive (ER1) breast cancer (Providence study)

Gene_ER1	HR	z (Coef)	p (Wald)
DR5	0.428	-3.478	0.001
GSTM2	0.526	-3.173	0.002
HSHIN1	0.175	-3.031	0.002
ESRRG	0.736	-3.028	0.003
VTN	0.622	-2.935	0.003
Bcl2	0.469	-2.833	0.005
ERBB4	0.705	-2.802	0.005
GPR30	0.625	-2.794	0.005
BAG1	0.339	-2.733	0.006
CAV1	0.635	-2.644	0.008
IL6ST	0.503	-2.551	0.011
C10orf116	0.679	-2.497	0.013
FOXO3A	0.607	-2.473	0.013
DICER1	0.311	-2.354	0.019
GADD45	0.645	-2.338	0.019
CSF1	0.500	-2.312	0.021
F3	0.677	-2.300	0.021
GBP2	0.604	-2.294	0.022
APEX-1	0.234	-2.253	0.024
FUS	0.322	-2.252	0.024
BBC3	0.581	-2.248	0.025
GSTM3	0.737	-2.203	0.028
ITGA4	0.620	-2.161	0.031
EPIB2	0.685	-2.128	0.033
IRF1	0.708	-2.105	0.035
CRYZ	0.593	-2.103	0.035
CCI19	0.773	-2.076	0.038
SKIL	0.540	-2.019	0.043
MRP1	0.515	-1.964	0.050

Table 6: Cox proportional hazards for Prognostic Genes that are negatively associated with good prognosis for ER-positive (ER1) breast cancer (Providence study)

Gene_ER1	HR	z (Coef)	p (Wald)
CTHRC1	2.083	1.958	0.050
RRM2	1.450	1.978	0.048
BUB1	1.467	1.988	0.047
LMNB1	1.764	2.009	0.045
SURV	1.380	2.013	0.044
EpCAM	1.966	2.076	0.038
CDC20	1.504	2.081	0.037
GAPDH	2.405	2.126	0.033
STK15	1.796	2.178	0.029
HSPA8	3.095	2.215	0.027
LAPTM4B	1.503	2.278	0.023
MCM2	1.872	2.370	0.018
CDC25C	1.485	2.423	0.015
ADM	1.695	2.486	0.013
MMP1	1.365	2.522	0.012
CCNB1	1.893	2.646	0.008
Ki-67	1.697	2.649	0.008
E2F1	1.662	2.689	0.007
KPNA2	1.683	2.701	0.007
DUSP1	1.573	2.824	0.005
GDF15	1.440	2.896	0.004

Table 7: Cox proportional hazards for Prognostic Genes that are positively associated with good prognosis for breast cancer (Rush study)

Gene_all	z (Coef)	HR	p (Wald)
GSTM2	-3.275	0.752	0.001
GSTM1	-2.946	0.772	0.003
C8orf4	-2.639	0.793	0.008
ELF3	-2.478	0.769	0.013
RUNX1	-2.388	0.609	0.017
IL6ST	-2.350	0.738	0.019
AAMP	-2.325	0.715	0.020
PR	-2.266	0.887	0.023
FHIT	-2.193	0.790	0.028
CD44v6	-2.191	0.754	0.028
GREB1 variant c	-2.120	0.874	0.034
ADAM17	-2.101	0.686	0.036
EstR1	-2.084	0.919	0.037
NAT1	-2.081	0.878	0.037
TNFRSF11B	-2.074	0.843	0.038
ITGB4	-2.006	0.740	0.045
CSF1	-1.963	0.750	0.050

Table 8: Cox proportional hazards for Prognostic Genes that are negatively associated with good prognosis for breast cancer (Rush study)

Gene_all	z (Coef)	HR	p (Wald)
STK15	1.968	1.298	0.049
TFRC	2.049	1.399	0.040
ITGB1	2.071	1.812	0.038
ITGAV	2.081	1.922	0.037
MYBL2	2.089	1.205	0.037
MRP3	2.092	1.165	0.036
SKP2	2.143	1.379	0.032
LMNB1	2.155	1.357	0.031
ALCAM	2.234	1.282	0.025
COMT	2.271	1.412	0.023
CDC20	2.300	1.253	0.021
GAPDH	2.307	1.572	0.021
GRB7	2.340	1.205	0.019
S100A9	2.374	1.120	0.018
S100A7	2.374	1.092	0.018
HER2	2.425	1.210	0.015
ACTR2	2.499	1.788	0.012
S100A8	2.745	1.144	0.006
ENO1	2.752	1.687	0.006
MMP1	2.758	1.212	0.006
LAPTM4B	2.775	1.375	0.006
FGFR4	3.005	1.215	0.003
C17orf37	3.260	1.387	0.001

Table 9: Cox proportional hazards for Prognostic Genes that are positively associated with good prognosis for ER-negative (ER0) breast cancer (Rush study)

Gene_ER0	z (Coef)	HR	p (Wald)
SEMA3F	-2.465	0.503	0.014
LAMA3	-2.461	0.519	0.014
CD44E	-2.418	0.719	0.016
AD024	-2.256	0.617	0.024
LAMB3	-2.237	0.690	0.025
Ki-67	-2.209	0.650	0.027
MMP7	-2.208	0.768	0.027
GREB1 variant c	-2.019	0.693	0.044
ITGB4	-1.996	0.657	0.046
CRYZ	-1.976	0.662	0.048
CD44s	-1.967	0.650	0.049

Table 10: Cox proportional hazards for Prognostic Genes that are negatively associated with good prognosis for ER-negative (ER0) breast cancer (Rush study)

Gene_ER0	z (Coef)	HR	p (Wald)
S100A8	1.972	1.212	0.049
EEF1A2	2.031	1.195	0.042
TAGLN	2.072	2.027	0.038
GRB7	2.086	1.231	0.037
HER2	2.124	1.232	0.034
ITGAV	2.217	3.258	0.027
CDH11	2.237	2.728	0.025
COL1A1	2.279	2.141	0.023
C17orf37	2.319	1.329	0.020
COL1A2	2.336	2.577	0.020
ITGB5	2.375	3.236	0.018
ITGA5	2.422	2.680	0.015
RPL41	2.428	6.665	0.015
ALCAM	2.470	1.414	0.013
CTHRC1	2.687	3.454	0.007
PTEN	2.692	8.706	0.007
FN1	2.833	2.206	0.005

Table 11: Cox proportional hazards for Prognostic Genes that are positively associated with good prognosis for ER-positive (ER1) breast cancer (Rush study)

Gene_ER1	z (Coef)	HR	p (Wald)
GSTM1	-3.938	0.628	0.000
HNF3A	-3.220	0.500	0.001
EstR1	-3.165	0.643	0.002
Bcl2	-2.964	0.583	0.003
GATA3	-2.641	0.624	0.008
ELF3	-2.579	0.741	0.010
C8orf4	-2.451	0.730	0.014
GSTM2	-2.416	0.774	0.016
PR	-2.416	0.833	0.016
RUNX1	-2.355	0.537	0.019
CSF1	-2.261	0.662	0.024
IL6ST	-2.239	0.627	0.025
AAMP	-2.046	0.704	0.041
TNFRSF11B	-2.028	0.806	0.043
NAT1	-2.025	0.833	0.043
ADAM17	-1.981	0.642	0.048

Table 12: Cox proportional hazards for Prognostic Genes that are negatively associated with good prognosis for ER-positive (ER1) breast cancer (Rush study)

Gene_ER1	z (Coef)	HR	p (Wald)
HSPA1B	1.966	1.382	0.049
AD024	1.967	1.266	0.049
FGFR4	1.991	1.175	0.047
CDK4	2.014	1.576	0.044
ITGB1	2.021	2.163	0.043
EPHB2	2.121	1.342	0.034
LYRIC	2.139	1.583	0.032
MYBL2	2.174	1.273	0.030
PGF	2.176	1.439	0.030
EZH2	2.199	1.390	0.028
HSPA1A	2.209	1.452	0.027
RPLPO	2.273	2.824	0.023
LMNB1	2.322	1.529	0.020
IL-8	2.404	1.166	0.016
C6orf66	2.468	1.803	0.014
GAPDH	2.489	1.950	0.013
P16-INK4	2.490	1.541	0.013
CLIC1	2.557	2.745	0.011
ENO1	2.719	2.455	0.007
ACTR2	2.878	2.543	0.004
CDC20	2.931	1.452	0.003
SKP2	2.952	1.916	0.003
LAPTM4B	3.124	1.558	0.002

WO 2011/063274

PCT/US2010/057490

Table 13: Validation of Prognostic Genes in SIB data sets.

Table 13 Official Symbol	EMC2 ~Est	EMC2~SE	EMC2~t	JRH1~Est	JRH1~SE	JRH1~t	JRH2~Est	JRH2~SE	JRH2~t	MGH~Est	MGH~SE	MGH~t	NCH~Est	NCH~SE	NCH~t	NKI~Est	NKI~SE	NKI~t
AAMP	NA	NA	NA	-0.05212	0.50645	-0.10291	0.105615	1.01216	0.104346	-0.26943	0.620209	-0.43441	0.088826	0.283082	0.313782	0.312939	0.228446	1.36986
ABCC1	NA	NA	NA	NA	NA	NA	2.36153	0.76485	3.087573	0.253516	0.284341	0.891591	0.213191	0.154486	1.380002	0.094607	0.258279	0.366298
ABCC3	NA	NA	NA	0.386945	0.504324	0.767255	0.305901	0.544322	0.561985	0.126882	0.221759	0.572162	-0.00756	0.167393	-0.04517	0.06613	0.096544	0.684974
ABR	NA	NA	NA	0.431151	0.817818	0.527197	0.758422	1.0123	0.749207	NA	NA	NA	NA	NA	NA	-0.06114	0.095839	-0.63795
ACTR2	NA	NA	NA	NA	NA	NA	-0.26297	0.4774	-0.55084	0.071853	0.205648	0.349398	0.131215	0.267434	0.490644	0.539449	0.254409	2.120401
ADAM17	NA	NA	NA	0.078212	0.564555	0.138538	-0.20948	1.06045	-0.19754	0.29698	0.435924	0.681266	-0.18523	0.407965	-0.45402	0.068689	0.12741	0.539115
ADM	NA	NA	NA	NA	NA	NA	0.320052	0.201407	1.589081	0.225324	0.142364	1.582732	0.314064	0.201161	1.561257	0.264131	0.06376	4.142582
LYPD6	NA	NA	NA	NA	NA	NA	NA	NA	NA	-0.38423	0.120883	-3.17855	-0.23802	0.209786	-1.1346	-0.4485	0.106865	-4.19691
AKI3	NA	NA	NA	NA	NA	NA	-2.10931	1.58606	-1.32991	-1.43148	0.576851	-2.48154	0.181912	0.147743	1.231273	0.149731	0.140716	1.064065
ALCAM	NA	NA	NA	-0.17112	0.224449	-0.7624	0.120168	0.212325	0.565963	-0.36428	0.239833	-1.51888	0.002712	0.084499	0.032094	-0.3019	0.094459	-3.19609
APEX1	NA	NA	NA	0.068917	0.410873	0.167732	-0.02247	0.790107	-0.02843	-0.07674	0.181782	-0.42215	-0.00097	0.268651	-0.00361	-0.13398	0.232019	-0.57746
ARF1	NA	NA	NA	0.839013	0.346692	2.420053	0.369609	0.40789	0.906149	2.03958	0.804729	2.534493	-0.15337	0.204529	-0.74984	0.944168	0.204641	4.613777
AURKA	NA	NA	NA	0.488329	0.248241	1.967157	0.285095	0.243026	1.173105	0.270093	0.169472	1.593732	-0.07663	0.213247	-0.35934	0.643963	0.101097	6.369754
BAD	NA	NA	NA	0.027049	0.547028	0.049446	0.121904	0.587599	0.207461	NA	NA	NA	0.38364	0.389915	0.983907	0.149641	0.221188	0.676533
BAG1	NA	NA	NA	0.505074	0.709869	0.711503	-0.13983	0.36181	-0.38648	-0.36295	0.282963	-1.28267	-0.11976	0.203911	-0.58733	-0.41603	0.138093	-3.01265
BBC3	NA	NA	NA	NA	NA	NA	0.182425	0.78708	0.231774	NA	NA	NA	0.056993	0.249671	0.228274	-0.5633	0.158825	-3.54669
BCAR3	NA	NA	NA	NA	NA	NA	-0.29238	0.522706	-0.55935	-0.41595	0.216837	-1.91825	0.072246	0.304443	0.237306	-0.26067	0.114992	-2.26685
BCL2	NA	NA	NA	-1.10678	0.544697	-2.03192	0.124104	0.228026	0.544254	-2.47368	1.23296	-2.00629	NA	NA	NA	-0.30738	0.079518	-3.86357
BIRC5	NA	NA	NA	-0.40529	0.608667	-0.66586	0.319899	0.242736	1.317889	NA	NA	NA	0.268836	0.122325	2.197719	0.390779	0.069127	5.6531
BTRC	NA	NA	NA	NA	NA	NA	0.017988	0.648834	0.027723	NA	NA	NA	-0.63958	0.485936	-1.31618	-0.52394	0.139699	-3.75051
BUB1	NA	NA	NA	0.84036	0.319874	2.627159	0.565139	0.322406	1.75288	0.206656	0.268687	0.769133	0.104644	0.142318	0.735283	0.426611	0.094852	4.49763
C10orf116	NA	NA	NA	-0.1418	0.261554	-0.54216	0.036378	0.182183	0.19968	NA	NA	NA	0.064337	0.14087	0.456713	-0.22589	0.082836	-2.72696
C17orf37	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.1532	0.294177	0.520775	NA	NA	NA
TPX2	NA	NA	NA	NA	NA	NA	0.311175	0.271756	1.145053	NA	NA	NA	-0.01014	0.317222	-0.03198	0.536914	0.116472	4.609812
C8orf4	NA	NA	NA	NA	NA	NA	-0.06402	0.197663	-0.32386	-0.07043	0.106335	-0.66236	-0.03221	0.189009	-0.1704	-0.3396	0.083273	-4.07813
CAV1	NA	NA	NA	-0.20701	0.254401	-0.81372	-0.19588	0.289251	-0.67721	-0.06896	0.2269	-0.30391	0.078825	0.340843	0.231265	-0.30885	0.133788	-2.30848
CCL19	NA	NA	NA	0.101779	0.483649	0.21044	-0.45509	0.26397	-1.71104	0.246585	0.153468	1.606752	0.024132	0.130045	0.185564	-0.08897	0.087102	-1.02143
CCNB1	NA	NA	NA	0.14169	0.276165	0.513063	0.587021	0.249935	2.348695	NA	NA	NA	-0.02016	0.230327	-0.08751	0.495483	0.10424	4.75329
CDC20	NA	NA	NA	-0.82502	0.360648	-2.2876	0.075789	0.208662	0.363213	0.095023	0.198727	0.478159	0.482934	0.216025	2.235547	0.35587	0.125008	2.846778
CDC25A	NA	NA	NA	-0.15046	0.724766	-0.2076	0.358589	0.638958	0.561209	0.257084	0.227966	1.12773	0.078265	0.111013	0.705008	0.48387	0.105238	4.597864
CDC25C	NA	NA	NA	0.047781	0.511454	0.093422	1.07486	0.456637	2.353861	0.340882	0.240266	1.418769	-0.22371	0.269481	-0.83013	0.287063	0.136568	2.101979
CDH11	NA	NA	NA	-0.55211	0.469473	-1.17601	0.072308	0.265898	0.27194	0.028252	0.199053	0.141931	-0.0883	0.124418	-0.70971	-0.13223	0.097541	-1.35364
CDK4	NA	NA	NA	NA	NA	NA	0.759572	0.757398	1.00287	0.18468	0.129757	1.423276	0.304045	0.17456	1.741779	0.267465	0.148641	1.799403

WO 2011/063274

PCT/US2010/057490

Table 13 Official Symbol	EMC2 ~Est	EMC2~SE	EMC2~t	JRH1~Est	JRH1~SE	JRH1~t	JRH2~Est	JRH2~SE	JRH2~t	MGH~Est	MGH~SE	MGH~t	NCH~Est	NCH~SE	NCH~t	NKI~Est	NKI~SE	NKI~t
SCUBE2	NA	NA	NA	NA	NA	NA	-0.0454	0.120869	-0.37564	NA	NA	NA	-0.01783	0.063429	-0.28108	-0.24635	0.048622	-5.0667
CENPA	NA	NA	NA	NA	NA	NA	0.296857	0.253493	1.171066	NA	NA	NA	0.225878	0.249928	0.903772	0.467131	0.081581	5.726013
CHAF1B	NA	NA	NA	0.591417	0.58528	1.010486	0.284056	0.637446	0.445616	0.47534	0.323193	1.470762	0.233081	0.291389	0.799896	0.519868	0.125204	4.152168
CLDN4	NA	NA	NA	-0.54144	0.470758	-1.15014	0.33033	0.351865	0.938798	0.185116	0.314723	0.588187	-0.23129	0.426627	-0.54213	0.564756	0.210595	2.681716
CLIC1	NA	NA	NA	0.678131	0.359483	1.886406	0.764626	0.767633	0.996083	0.171995	0.821392	0.209395	-0.05548	0.414451	-0.13385	0.383134	0.165674	2.312578
COL1A1	NA	NA	NA	NA	NA	NA	0.273073	0.249247	1.095592	NA	NA	NA	0.004033	0.146511	0.027527	NA	NA	NA
COL1A2	NA	NA	NA	NA	NA	NA	0.216939	0.367138	0.590892	0.157848	0.123812	1.274901	0.057815	0.163831	0.352894	-0.00235	0.064353	-0.03653
COMT	NA	NA	NA	0.749278	0.356366	2.101373	-0.05068	0.448567	-0.111298	-2.45771	1.02805	-2.39065	0.526063	0.226489	2.322687	-0.00764	0.129967	-0.05878
CRYZ	NA	NA	NA	NA	NA	NA	-0.31201	0.303615	-1.02766	-0.53751	0.214408	-2.50696	-0.32472	0.253244	-1.28224	-0.25514	0.124909	-2.04264
CSF1	NA	NA	NA	NA	NA	NA	-1.40833	1.21432	-1.15977	NA	NA	NA	-0.14894	0.352724	-0.42226	-0.11194	0.240555	-0.46532
CTHRC1	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.574897	0.535382	1.073807	-0.08389	0.137325	-0.6109	0.024002	0.097692	0.245691
CXCL12	NA	NA	NA	-0.36476	0.372499	-0.97921	-0.4566	0.219587	-2.07935	NA	NA	NA	-0.08863	0.138097	-0.64183	-0.36944	0.138735	-2.66293
CXCL14	NA	NA	NA	-0.23692	0.333761	-0.70985	0.361375	0.159544	2.265049	NA	NA	NA	-0.06392	0.093353	-0.70609	-0.16877	0.054117	-3.11866
CYR61	NA	NA	NA	0.310818	0.515557	0.602878	-0.24435	0.252867	-0.9663	0.571476	0.323144	1.768487	-0.11281	0.164296	-0.68663	0.087147	0.082372	1.057965
DICER1	NA	NA	NA	NA	NA	NA	-0.33943	0.39364	-0.8623	0.038811	0.409835	0.0947	0.086141	0.143687	0.599504	-0.46887	0.150367	-3.11814
DLC1	NA	NA	NA	0.13581	0.37927	0.358083	-0.4102	0.387258	-1.05923	-0.09793	0.247069	-0.39638	-0.03473	0.238947	-0.14533	-0.35001	0.130472	-2.68262
TNFRSF10B	NA	NA	NA	-0.09001	0.619057	-0.1454	0.80742	0.544479	1.482922	0.159018	0.456205	0.348567	-0.19927	0.160381	-1.24248	0.053214	0.164091	0.324294
DUSP1	NA	NA	NA	-0.20229	0.200782	-1.00753	-0.02736	0.224043	-0.12212	NA	NA	NA	-0.03006	0.152909	-0.19657	-0.0472	0.09086	-0.51952
E2F1	NA	NA	NA	NA	NA	NA	0.845576	0.685556	1.233416	-1.06849	0.824212	-1.29638	0.356102	0.38254	0.930888	0.617258	0.121385	5.085126
EEF1A2	0.26278	0.091435	2.873951	NA	NA	NA	0.362569	0.17103	2.119915	NA	NA	NA	-0.0028	0.233293	-0.01199	-0.01585	0.06608	-0.23987
ELF3	NA	NA	NA	1.34589	0.628064	2.142919	0.569231	0.430739	1.321522	0.209853	0.239225	0.87722	0.026264	0.109569	0.2397	0.165848	0.143091	1.159039
ENO1	NA	NA	NA	NA	NA	NA	0.179739	0.312848	0.574525	NA	NA	NA	-0.01727	0.097939	-0.17629	0.3682	0.094778	3.884888
EPHB2	NA	NA	NA	0.155831	0.717587	0.21716	-0.19469	0.90381	-0.21541	1.38257	0.444196	3.112522	-0.46953	0.395102	-1.18837	0.318437	0.123672	2.574851
ERBB2	NA	NA	NA	-0.32795	0.215691	-1.52044	0.065275	0.189094	0.3452	0.314084	0.126321	2.486396	0.23616	0.121533	1.943176	0.08469	0.056744	1.492504
ERBB4	NA	NA	NA	NA	NA	NA	-0.12516	0.182846	-0.68451	-0.13567	0.114364	-1.18626	0.191218	0.114326	1.672568	-0.28508	0.066294	-4.30028
ESRRG	NA	NA	NA	NA	NA	NA	0.122595	0.204322	0.600009	0.356845	0.216506	1.648199	0.023341	0.078378	0.297795	-0.16542	0.093598	-1.76733
ESR1	NA	NA	NA	-0.14448	0.127214	-1.13569	0.009283	0.107091	0.086687	-0.12127	0.111184	-1.09075	0.127143	0.109672	1.159302	-0.16933	0.044665	-3.79121
EZH2	NA	NA	NA	NA	NA	NA	0.36213	0.244107	1.483489	NA	NA	NA	0.008861	0.200897	0.044106	0.478266	0.107424	4.452134
F3	NA	NA	NA	0.719395	0.524742	1.37095	-0.21237	0.363632	-0.58402	-0.00167	0.448211	-0.00372	-0.13187	0.134218	-0.98248	-0.29217	0.093753	-3.11637
FGFR4	NA	NA	NA	0.864262	0.479596	1.802063	0.451249	0.296065	1.524155	0.230309	0.229234	1.00469	-0.15142	0.109674	-1.3806	-0.04922	0.146198	-0.33666
FHIT	NA	NA	NA	1.00058	0.938809	1.065797	-1.58314	0.766553	-2.06527	0.087228	0.322399	0.270559	-0.08366	0.344886	-0.24256	-0.1378	0.121745	-1.13183
FN1	NA	NA	NA	0.056943	0.154068	0.369595	0.282152	0.407361	0.692634	0.417442	0.859619	0.485613	-0.05187	0.111777	-0.46402	0.112875	0.066759	1.690796
FOXAI	NA	NA	NA	NA	NA	NA	0.054619	0.1941	0.281398	NA	NA	NA	-0.04211	0.103534	-0.40677	-0.08953	0.043624	-2.05225
FUS	NA	NA	NA	NA	NA	NA	2.73816	1.95693	1.399212	-0.18397	0.269637	-0.68227	0.119801	0.199389	0.600841	0.115971	0.188545	0.615084

WO 2011/063274

PCT/US2010/057490

Table 13 Official Symbol	EMC2 ~Est	EMC2~SE	EMC2~t	JRH1~Est	JRH1~SE	JRH1~t	JRH2~Est	JRH2~SE	JRH2~t	MGH~Est	MGH~SE	MGH~t	NCH~Est	NCH~SE	NCH~t	NKI~Est	NKI~SE	NKI~t
GADD45A	NA	NA	NA	NA	NA	NA	-0.09194	0.324263	-0.28352	-0.33447	0.236846	-1.41219	-0.43753	0.333292	-1.31276	-0.15889	0.115794	-1.37217
GAPDH	-0.00386	0.125637	-0.03075	0.869317	0.274798	3.163476	0.728889	0.497848	1.464079	NA	NA	NA	0.396067	0.169944	2.330574	0.286211	0.073946	3.870541
GATA3	NA	NA	NA	-0.33431	0.127225	-2.62767	-0.00759	0.145072	-0.05233	0.190453	0.170135	1.119423	0.058244	0.115942	0.502355	-0.13285	0.054984	-2.41625
GBP2	NA	NA	NA	0.120416	0.247997	0.485554	-0.49134	0.289525	-1.69704	0.517501	0.299148	1.729916	0.082647	0.173301	0.4769	-0.19825	0.1358	-1.45985
GDF15	NA	NA	NA	0.219861	0.231613	0.94926	0.317951	0.183188	1.735634	NA	NA	NA	0.200247	0.14325	1.397885	0.052347	0.063101	0.829563
GRB7	NA	NA	NA	-0.46505	0.485227	-0.95842	0.143585	0.218034	0.658544	NA	NA	NA	0.027699	0.459937	0.060224	0.126284	0.054856	2.302117
GSTM1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	-0.18141	0.14912	-1.21652
GSTM2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	-0.15655	0.118111	-1.32547
GSTM3	NA	NA	NA	-1.19919	0.478486	-2.50622	-0.08173	0.176832	-0.46219	NA	NA	NA	-0.09058	0.129247	-0.70086	-0.336	0.086817	-3.87028
HOXB13	NA	NA	NA	NA	NA	NA	0.780988	0.524959	1.487712	0.461343	0.122399	3.769173	0.453876	0.324863	1.39713	0.161713	0.053047	3.048485
OTUD4	NA	NA	NA	NA	NA	NA	-0.54088	1.59038	-0.34009	0.154269	0.633438	0.243542	0.150174	0.149267	1.006076	-0.08847	0.130112	-0.67992
HSPA1A	NA	NA	NA	0.199478	0.304533	0.655029	0.56215	0.592113	0.949396	NA	NA	NA	0.187486	0.231047	0.811463	0.174571	0.117296	1.488295
HSPA1B	NA	NA	NA	NA	NA	NA	0.60089	0.32867	1.828247	NA	NA	NA	NA	NA	NA	0.249602	0.129038	1.934329
IISPA8	NA	NA	NA	0.88406	0.420719	2.101308	1.13504	0.667937	1.699322	0.647034	0.346081	1.869603	0.208652	0.225656	0.924646	0.054243	0.178314	0.304198
IDH2	NA	NA	NA	-0.0525	0.232201	-0.22611	0.151299	0.327466	0.46203	NA	NA	NA	0.265828	0.105592	2.517501	0.284862	0.089145	3.195498
IGF1R	NA	NA	NA	-0.62963	0.509985	-1.23461	-0.05773	0.176259	-0.32753	-0.11077	0.162941	-0.67982	-0.37931	0.371019	-1.02236	-0.13655	0.08362	-1.63299
IGFBP7	NA	NA	NA	NA	NA	NA	0.047112	0.479943	0.098162	NA	NA	NA	0.163138	0.200674	0.81295	0.06541	0.10077	0.649097
IL11	NA	NA	NA	NA	NA	NA	1.19114	1.41017	0.844678	NA	NA	NA	-0.17423	0.144228	-1.20804	-0.048	0.126254	-0.38015
IL17RB	NA	NA	NA	NA	NA	NA	0.143131	0.294647	0.485771	-0.44343	0.132744	-3.3405	NA	NA	NA	-0.01632	0.122679	-0.13305
IL6ST	NA	NA	NA	-0.08851	0.151324	-0.58488	-0.00958	0.287723	-0.03329	-0.76052	0.386504	-1.96769	-0.4336	0.319875	-1.35553	-0.41477	0.111102	-3.73322
IL8	NA	NA	NA	0.222258	0.235694	0.942994	0.262285	0.346572	0.756798	-0.12567	0.154036	-0.81583	-1.28729	0.493461	-2.6087	0.171912	0.07248	2.371858
INHBA	NA	NA	NA	0.095254	0.476446	0.199927	0.342597	0.27142	1.262239	NA	NA	NA	-0.12767	0.132531	-0.96331	0.133895	0.111083	1.20536
IRF1	NA	NA	NA	0.87337	0.941443	0.927693	-0.39282	0.392589	-1.00059	0.474132	0.503423	0.941816	-0.2456	0.294202	-0.8348	-0.08017	0.171067	-0.46864
ITGA4	NA	NA	NA	NA	NA	NA	-0.91318	0.542311	-1.68388	NA	NA	NA	0.034844	0.074049	0.470549	-0.05101	0.133497	-0.38211
ITGA5	NA	NA	NA	1.44044	0.636806	2.261976	0.97846	0.67341	1.452993	0.206218	0.263291	0.783232	0.367111	0.333768	1.099899	0.500604	0.163986	3.052724
ITGAV	NA	NA	NA	0.14845	0.345246	0.429983	0.383127	0.60722	0.630953	-0.23212	0.278464	-0.83358	-0.14166	0.222286	-0.6373	-0.21993	0.158945	-1.38371
ITGB1	NA	NA	NA	1.22836	0.683544	1.797046	-0.0587	1.73799	-0.03378	-0.13651	0.121624	-1.12236	-0.52799	0.346298	-1.52468	0.150333	0.133426	1.126714
ITGB4	NA	NA	NA	0.548277	0.334628	1.638467	0.252015	0.365768	0.689002	-0.12971	0.168517	-0.76973	0.189568	0.163609	1.158665	0.166748	0.175308	0.951172
ITGB5	NA	NA	NA	-0.17231	0.250618	-0.68752	0.037961	0.401861	0.094464	0.682674	0.74847	0.912093	-0.04952	0.16668	-0.29707	0.010302	0.104545	0.098544
MKI67	NA	NA	NA	-0.43304	0.708832	-0.61092	0.482583	0.321739	1.499921	NA	NA	NA	0.128582	0.129422	0.99351	0.397232	0.176102	2.255693
KIAA1199	NA	NA	NA	NA	NA	NA	-0.02195	0.382802	-0.05735	0.081394	0.121221	0.671448	NA	NA	NA	0.238809	0.113748	2.099457
KPNA2	0.301662	0.171052	1.763569	-0.5507	0.55364	-0.99468	0.21269	0.256724	0.828477	-1.6447	1.00101	-1.64304	0.213725	0.196767	1.086183	0.422135	0.089135	4.735922
LAMA3	NA	NA	NA	-0.74591	0.563373	-1.32401	-0.21092	0.29604	-0.71245	NA	NA	NA	-0.03143	0.133752	-0.23497	-0.30023	0.122124	-2.45838
LAMB3	NA	NA	NA	NA	NA	NA	0.345497	0.263827	1.309559	0.03108	0.139904	0.222154	0.106874	0.139587	0.765644	-0.03167	0.069644	-0.45477

WO 2011/063274

PCT/US2010/057490

Table 13 Official Symbol	EMC2 ~Est	EMC2~SE	EMC2~t	JRH1~Est	JRH1~SE	JRH1~t	JRH2~Est	JRH2~SE	JRH2~t	MGH~Est	MGH~SE	MGH~t	NCH~Est	NCH~SE	NCH~t	NKI~Est	NKI~SE	NKI~t
LAPTM4B	NA	NA	NA	NA	NA	NA	-0.04029	0.234986	-0.17148	0.352765	0.40304	0.875261	0.156358	0.140366	1.113931	0.334588	0.083358	4.013853
LMNB1	NA	NA	NA	0.648703	0.285233	2.274292	0.621431	0.389912	1.593772	NA	NA	NA	-0.1517	0.242463	-0.62567	0.461325	0.098382	4.689115
LRIG1	NA	NA	NA	NA	NA	NA	-0.00217	0.260339	-0.00832	-0.61468	0.216033	-2.84532	-0.24368	0.172969	-1.40878	-0.50209	0.1119	-4.48694
MTDH	NA	NA	NA	NA	NA	NA	-0.10827	0.493025	-0.21961	0.084824	0.292285	0.290209	0.039288	0.233351	0.168365	0.430557	0.145357	2.962066
MCM2	NA	NA	NA	0.875004	0.492588	1.77634	0.77667	0.376275	2.064102	0.118904	0.288369	0.412333	0.586577	0.252123	2.326551	0.504911	0.154078	3.276983
MELK	NA	NA	NA	0.850914	0.313784	2.711783	0.16347	0.256575	0.637124	NA	NA	NA	0.216763	0.1352	1.603277	0.471343	0.103644	4.547711
MGMT	NA	NA	NA	NA	NA	NA	0.151967	0.583459	0.260459	0.267185	0.295678	0.903635	-0.37332	0.507157	-0.73611	-0.14716	0.165874	-0.88716
MMP1	NA	NA	NA	0.43277	0.16023	2.70093	-0.02427	0.158939	-0.15272	0.180359	0.078781	2.289386	0.559716	0.331212	1.689903	0.167053	0.064595	2.586172
MMP7	NA	NA	NA	0.198055	0.143	1.385	0.106475	0.193338	0.550719	-1.06791	1.30502	-0.81831	0.012294	0.101346	0.121311	NA	NA	NA
MYBL2	NA	NA	NA	0.731162	0.267911	2.729123	0.098974	0.600361	0.164857	0.612646	0.509356	1.202785	0.396938	0.171503	2.314467	0.751827	0.151477	4.963308
NAT1	NA	NA	NA	-0.57746	15.1186	-0.0382	-0.01397	0.117033	-0.11939	-0.05035	0.105736	-0.47614	-0.15619	0.139368	-1.12073	-0.20435	0.058054	-3.52
PGF	NA	NA	NA	0.901309	0.501058	1.798812	1.43389	1.27617	1.123589	NA	NA	NA	0.05255	0.14245	0.368898	0.055127	0.36118	0.152631
PGR	NA	NA	NA	NA	NA	NA	-0.33243	0.276025	-1.20435	-0.95852	0.593621	-1.61469	-0.01033	0.08386	-0.12312	-0.30421	0.073055	-4.16405
PRDX1	NA	NA	NA	NA	NA	NA	-0.41082	0.47383	-0.86703	NA	NA	NA	0.253047	0.182621	1.38564	0.231612	0.161791	1.431551
PTEN	NA	NA	NA	-0.17429	0.629039	-0.27708	-0.15599	0.541475	-0.28808	-0.10814	0.287261	-0.37645	0.113229	0.228164	0.496261	-0.3204	0.149745	-2.13962
RPL41	NA	NA	NA	NA	NA	NA	1.02038	1.83528	0.555981	0.213155	0.288282	0.739398	0.030854	0.188269	0.163881	-0.08602	0.122667	-0.70126
RPLP0	NA	NA	NA	0.398754	0.282913	1.409458	0.246775	1.2163	0.20289	0.488909	0.174981	2.794069	0.004595	0.198497	0.023148	0.008104	0.079365	0.102105
RRM2	NA	NA	NA	NA	NA	NA	0.196643	0.262745	0.748418	NA	NA	NA	0.229458	0.11665	1.967064	0.434693	0.152104	2.857867
RUNX1	NA	NA	NA	-0.22834	0.318666	-0.71656	0.302803	0.420043	0.720886	0.277566	0.267511	1.037587	0.124568	0.088457	1.408231	-0.18878	0.138365	-1.36435
S100A8	NA	NA	NA	NA	NA	NA	0.066629	0.11857	0.561939	NA	NA	NA	0.142073	0.080349	1.768194	0.094631	0.041656	2.271738
S100A9	NA	NA	NA	NA	NA	NA	0.111103	0.13176	0.843223	NA	NA	NA	0.090314	0.058415	1.546083	0.111093	0.045472	2.443086
S100B	NA	NA	NA	0.097319	0.589664	0.165041	-0.2365	0.349444	-0.67678	NA	NA	NA	0.239753	0.145105	1.652272	0.195383	0.295751	0.660633
S100P	NA	NA	NA	0.378047	0.120687	3.132458	0.302607	0.133752	2.262448	NA	NA	NA	0.202856	0.092114	2.202218	0.103276	0.04811	2.146677
SEMA3F	NA	NA	NA	-0.27556	0.615782	-0.4475	0.498631	0.616195	0.80921	0.107802	0.274191	0.393164	-0.17978	0.185166	-0.97092	NA	NA	NA
SKIL	NA	NA	NA	NA	NA	NA	0.036279	0.587743	0.044712	NA	NA	NA	0.143484	0.103564	1.385462	0.124124	0.120741	1.028019
SKP2	NA	NA	NA	NA	NA	NA	0.2502	0.469372	0.533053	0.470759	0.2802	1.680082	-0.71691	0.354699	-2.02117	0.056728	0.128585	0.441174
SNAIL	NA	NA	NA	NA	NA	NA	0.165897	1.09586	0.151385	0.163855	0.228308	0.717693	-0.04601	0.259767	-0.17711	0.057651	0.124454	0.463235
SYK	NA	NA	NA	-0.26425	0.588491	-0.44903	-0.22515	0.492582	-0.45707	NA	NA	NA	-1.30716	0.591071	-2.21151	0.178238	0.168423	1.058276
TAGLN	NA	NA	NA	NA	NA	NA	0.042223	0.251268	0.168039	0.010727	0.098919	0.108442	0.194543	0.115463	1.684895	0.077881	0.119491	0.651776
TIFRC	NA	NA	NA	-0.91825	0.636275	-1.44317	0.162921	0.352486	0.462206	0.029015	0.193689	0.149803	0.056174	0.166875	0.336622	0.157216	0.10845	1.449663
TGFB3	NA	NA	NA	-1.0219	0.358953	-2.84689	-0.39774	0.470041	-0.84619	0.046498	0.2296	0.202518	-0.30473	0.247338	-1.23202	-0.36531	0.09592	-3.80851
TNFRSF11B	NA	NA	NA	NA	NA	NA	-0.10399	0.440721	-0.23595	-1.15976	0.400921	-2.89274	-0.2492	0.289075	-0.86207	-0.22072	0.10171	-2.17005
VTN	NA	NA	NA	-0.18721	0.475541	-0.39367	-2.39601	1.83129	-1.30837	NA	NA	NA	0.048066	0.34143	0.140779	-0.05675	0.116352	-0.48774
WISPI	NA	NA	NA	NA	NA	NA	0.437936	0.592058	0.739684	-0.03674	0.212861	-0.1726	NA	NA	NA	-0.36317	0.153002	-2.3736

Table 13 Official Symbol	EMC2 ~Est	EMC2~SE	EMC2~t	JRH1~Est	JRH1~SE	JRH1~t	JRH2~Est	JRH2~SE	JRH2~t	MGH~Est	MGH~SE	MGH~t	NCH~Est	NCH~SE	NCH~t	NKI~Est	NKI~SE	NKI~t
WNT5A	NA	NA	NA	NA	NA	NA	0.180255	0.286462	0.629246	0.06984	0.223411	0.312605	-0.14987	0.146576	-1.02248	-0.29433	0.084559	-3.48081
C6orf66	NA	NA	NA	NA	NA	NA	0.35565	0.504627	0.704778	0.179742	0.364806	0.492706	-0.53606	0.448343	-1.19564	0.296686	0.199046	1.49054
FOXO3A	NA	NA	NA	NA	NA	NA	-0.04428	0.39855	-0.1111	0.176454	0.221502	0.796625	0.059822	0.171485	0.348846	-0.2855	0.194121	-1.47074
GPR30	NA	NA	NA	0.01829	0.925976	0.019752	-0.298	0.747388	-0.39872	-0.03208	0.1214	-0.26427	0.157898	0.174583	0.904429	0.080079	0.104254	0.768115
KNITC2	NA	NA	NA	NA	NA	NA	-0.02315	0.289403	-0.07999	-0.14241	0.246904	-0.57677	0.274706	0.14532	1.890352	0.432186	0.120356	3.590897

WO 2011/063274

PCT/US2010/057490

Table 13 Official Symbol	STNO ~Est	STNO ~SE	SINO~t	STOCK ~Est	STOCK ~SE	STOCK ~t	TRANS BIG~Est	TRANS BIG~SE	TRANS BIG~t	UCSF ~Est	UCSF ~SE	UCSF~t	UPP~Est	UPP~SE	UPP~t	fe	sefe
AAMP	0.189376	0.309087	0.612695	0.836415	0.549695	1.521598	0.051406	0.111586	0.460681	0.770516	0.762039	1.011124	1.25423	0.577991	2.169982	0.146929	0.085151
ABCC1	NA	NA	NA	0.640672	0.375725	1.705162	NA	NA	NA	NA	NA	NA	0.274551	0.271361	1.011756	0.281451	0.104466
ABCC3	0.311364	0.100031	3.112675	0.166453	0.159249	1.045237	NA	NA	NA	0.381707	0.250896	1.521375	0.178451	0.097237	1.835219	0.172778	0.048133
ABR	0.095087	0.266216	0.357181	0.08129	0.196104	0.414525	NA	NA	NA	-0.17319	0.728313	-0.23779	-0.16409	0.120793	-1.35847	-0.06034	0.067134
ACTR2	NA	NA	NA	0.302753	0.39656	0.763448	NA	NA	NA	NA	NA	NA	0.21463	0.353554	0.607064	0.199885	0.117995
ADAM17	NA	NA	NA	0.437069	0.276977	1.577997	NA	NA	NA	0.35888	0.433785	0.827322	0.131246	0.194946	0.673243	0.129961	0.090699
ADM	NA	NA	NA	0.555634	0.242705	2.289339	0.025583	0.038218	0.669405	NA	NA	NA	0.361033	0.203349	1.775435	0.119028	0.030564
LYPD6	NA	NA	NA	-0.42358	0.145799	-2.90525	-0.06178	0.031054	-1.98944	NA	NA	NA	-0.1544	0.073668	-2.09587	-0.12675	0.026288
AKT3	NA	NA	NA	0.12232	0.182253	0.671155	NA	NA	NA	NA	NA	NA	-0.06832	0.125172	-0.5458	0.05204	0.071861
ALCAM	-0.14634	0.126842	-1.15369	-0.41301	0.190485	-2.16822	NA	NA	NA	-0.25661	0.251874	-1.01879	-0.1468	0.143998	-1.01942	-0.15502	0.046361
APTX1	0.005151	0.257871	0.019976	0.739037	0.539346	1.370247	NA	NA	NA	-0.96465	0.704753	-1.36878	1.23743	0.466987	2.649817	0.019915	0.10244
ARF1	0	0.107397	0	0.862387	0.279535	3.085077	NA	NA	NA	0.304097	0.58718	0.517894	0.751279	0.361093	2.080569	0.281544	0.07587
AURKA	0.38795	0.127032	3.053955	0.688845	0.210275	3.275924	0.020041	0.064473	0.310835	-0.0146	0.28312	-0.05156	0.427382	0.126638	3.374832	0.262652	0.041246
BAD	-0.30035	0.250277	-1.20006	0.228387	0.543493	0.420221	NA	NA	NA	-0.43933	0.659711	-0.66594	0.351434	0.360157	0.97578	0.059151	0.126378
BAG1	NA	NA	NA	-0.39593	0.380547	-1.04043	NA	NA	NA	0.516764	0.524112	0.98598	0.380154	0.211079	1.801003	-0.16426	0.087173
BBC3	NA	NA	NA	-0.26155	0.219839	-1.18974	-0.04709	0.086372	-0.5452	0.263477	0.606256	0.434597	-0.13039	0.141473	-0.92165	-0.14598	0.061462
BCAR3	NA	NA	NA	-0.49692	0.265837	-1.86927	NA	NA	NA	NA	NA	NA	-0.29435	0.182614	-1.61186	-0.28755	0.080198
BCL2	-0.38181	0.112494	-3.39408	-0.73699	0.228055	-3.23162	NA	NA	NA	-0.3453	0.410691	-0.84078	-0.11988	0.174734	-0.68605	-0.32009	0.056047
BIRC5	0.190534	0.126151	1.510365	0.582957	0.159354	3.658251	0.007906	0.045316	0.174454	0.357332	0.286621	1.246706	0.43455	0.110681	3.926148	0.186649	0.031964
B'IRC	NA	NA	NA	-0.92763	0.307218	-3.01944	NA	NA	NA	NA	NA	NA	-0.0225	0.1807	-0.12451	-0.40405	0.100468
BUB1	0.357653	0.101235	3.532899	1.09451	0.258044	4.241563	0.014276	0.040135	0.355694	0.376719	0.340175	1.107427	0.469009	0.162559	2.885517	0.154368	0.032048
C10orf11 6	-0.09621	0.085948	-1.11936	-0.34745	0.112777	-3.08087	NA	NA	NA	0.013111	156.117	8.40E-05	-0.00923	0.100902	-0.09148	-0.13	0.042521
C17orf37	NA	NA	NA	0.382862	0.185356	2.06555	NA	NA	NA	NA	NA	NA	0.385651	0.113625	3.394068	0.362223	0.092012
TPX2	NA	NA	NA	0.800822	0.195737	4.091316	NA	NA	NA	0.213479	0.284008	0.751665	0.44053	0.139377	3.160708	0.480408	0.073094
C8orf4	NA	NA	NA	-0.36113	0.130038	-2.77713	NA	NA	NA	NA	NA	NA	0.0037	0.109064	0.033921	-0.18346	0.048256
CAV1	0.135002	0.093948	1.436991	-0.65852	0.275751	-2.38811	NA	NA	NA	-0.54391	0.428883	-1.2682	-0.31503	0.150431	-2.09415	-0.11726	0.058989
CCL19	-0.0546	2531.93	-2.16E-05	-0.15743	0.154207	-1.02087	NA	NA	NA	0	0.434462	0	-0.1048	0.106112	-0.98765	-0.05608	0.050769
CCNB1	0.37726	0.156356	2.412827	0.828029	0.223403	3.706436	NA	NA	NA	-0.35808	0.431863	-0.82915	0.611916	0.142007	4.309055	0.456916	0.062513
CDC20	0.059565	1057.7	5.63E-05	0.642601	0.178622	3.597547	NA	NA	NA	-0.65381	0.404188	-1.61759	0.490188	0.130676	3.751171	0.319134	0.064899
CDC25A	0.288245	0.213701	1.348824	0.168571	0.225272	0.7483	NA	NA	NA	-0.31967	0.397525	-0.80414	0.330359	0.191096	1.728759	0.267201	0.060819

WO 2011/063274

PCT/US2010/057490

Table 13 Official Symbol	STNO ~Est	STNO ~SE	SINO~t	STOCK ~Est	STOCK ~SE	STOCK ~t	TRANS BIG~Est	TRANS BIG~SE	TRANS BIG~t	UCSF ~Est	UCSF ~SE	UCSF~t	UPP~Est	UPP~SE	UPP~t	fe	sefe
CDC25C	0.420797	0.155926	2.698697	1.02036	0.337803	3.020577	NA	NA	NA	-0.33774	0.477196	-0.70776	0.827213	0.232669	3.555321	0.382935	0.077595
CDH11	-0.05652	0.1231	-0.45913	-0.21142	0.211537	-0.99942	NA	NA	NA	-0.20367	0.246195	-0.83541	-0.22621	0.164541	-1.37482	-0.11417	0.053045
CDK4	0.279447	0.142472	1.961417	1.40458	0.463254	3.031987	NA	NA	NA	-0.37577	0.674081	-0.55746	0.814832	0.297251	2.741225	0.305255	0.069562
SCUBE2	-0.21559	0.074112	-2.90896	-0.24679	0.122745	-2.01059	0.016505	0.023486	0.702739	NA	NA	NA	-0.14287	0.077009	-1.8552	-0.05439	0.018349
CENPA	NA	NA	NA	0.724539	0.195614	3.703922	0.002888	0.04791	0.060269	0.679912	0.275146	2.471095	0.536476	0.157029	3.416414	0.185486	0.037867
CHAF1B	0.259119	0.162074	1.59877	0.281358	0.148493	1.894756	NA	NA	NA	-0.03447	0.352745	-0.09773	0.209129	0.093425	2.238469	0.300765	0.05807
CLDN4	0.40922	0.128817	3.176755	1.20235	0.33711	3.56664	0.03236	0.053171	0.608591	0	1.8541	0	0.08503	0.258939	0.328378	0.125868	0.045235
CLIC1	0.238723	0.209629	1.138788	2.00024	0.600443	3.331274	-0.26608	0.160756	-1.65519	0.377361	0.532842	0.682584	0.999191	0.414232	2.412153	0.222753	0.088912
COL1A1	0.127256	0.081743	1.556791	0.05098	0.156488	0.325773	0.087944	0.034256	2.567237	NA	NA	NA	-0.05544	0.13355	-0.41509	0.083989	0.029343
COL1A2	-0.01925	0.078156	-0.24625	-0.17504	0.228915	-0.76466	NA	NA	NA	-0.1405	0.184661	-0.76085	-0.15924	0.220113	-0.72346	-0.00069	0.041375
COMT	NA	NA	NA	0.643165	0.360056	1.786292	NA	NA	NA	0.356582	0.628139	0.56768	0.404183	0.257299	1.570869	0.212925	0.092124
CRYZ	-0.38719	0.143353	-2.70092	0.122949	0.340718	0.360853	NA	NA	NA	-0.52792	0.412283	-1.28048	-0.37265	0.225119	-1.65534	-0.33167	0.071579
CSF1	NA	NA	NA	-0.11449	0.197258	-0.58042	-0.09782	0.196881	-0.49684	NA	NA	NA	0.120517	0.148659	0.810694	-0.0334	0.090261
CTHRC1	NA	NA	NA	0.263783	0.247606	1.065334	NA	NA	NA	NA	NA	NA	-0.14789	0.176843	-0.83626	-0.00169	0.069075
CXCL12	0.066487	0.189775	0.350348	-0.65036	0.168426	-3.86137	NA	NA	NA	-0.05795	0.270065	-0.21456	-0.35344	0.150278	-2.35189	-0.28998	0.062826
CXCL14	-0.20969	0.073458	-2.8546	-0.14079	0.096118	-1.46476	NA	NA	NA	NA	NA	NA	-0.1861	0.08384	-2.21976	-0.14219	0.032611
CYR61	NA	NA	NA	-0.38308	0.231645	-1.65372	NA	NA	NA	-0.22327	0.263371	-0.84773	-0.41188	0.174362	-2.36221	-0.04446	0.059831
DICER1	NA	NA	NA	-1.06544	0.322204	-3.30672	NA	NA	NA	0	0.311799	0	0.208326	0.307144	0.678268	-0.19602	0.085879
DLC1	0.519601	0.221066	2.350434	-0.66099	0.298518	-2.21425	NA	NA	NA	-0.31503	0.345828	-0.91094	-0.404	0.200673	-2.01324	-0.19876	0.076441
TNFRSF1 OB	-0.03773	0.174479	-0.21623	-0.03558	0.198203	-0.1795	NA	NA	NA	0.932141	0.524911	1.775808	0.127348	0.157658	0.807748	0.02034	0.072745
DUSP1	0.095682	0.223995	0.42716	-0.14883	0.12682	-1.17351	NA	NA	NA	0.008053	0.779738	0.010327	-0.41475	0.153012	-2.71055	-0.11225	0.054628
E2F1	0.171825	0.110793	1.550865	0.699408	0.207377	3.37264	NA	NA	NA	NA	NA	NA	0.570954	0.172882	3.302565	0.433836	0.067966
EEF1A2	NA	NA	NA	-0.01256	0.130353	-0.09633	NA	NA	NA	0.433528	0.267338	1.621648	-0.04242	0.091692	-0.46259	0.068177	0.041066
ELF3	0.406692	0.148275	2.742822	0.233332	0.357735	0.652248	NA	NA	NA	0.841389	0.55748	1.509272	0.096421	0.256911	0.375307	0.196003	0.066053
ENO1	NA	NA	NA	0.428884	0.194952	2.199947	NA	NA	NA	0.899319	0.369574	2.433394	0.288434	0.179833	1.603899	0.233559	0.058687
EPHB2	NA	NA	NA	0.192999	0.451341	0.427612	NA	NA	NA	0.355634	0.604801	0.588018	0.211632	0.199057	1.063173	0.284709	0.094113
ERBB2	0.268938	0.074504	3.609693	0.092164	0.188964	0.487734	NA	NA	NA	0.301674	0.170749	1.766769	0.349689	0.107646	3.248509	0.181046	0.034939
ERBB4	-0.10396	0.068988	-1.50697	-0.73759	0.209821	-3.51532	NA	NA	NA	NA	NA	NA	-0.1859	0.117619	-1.58055	-0.16266	0.037384
ESRRG	NA	NA	NA	-0.32843	0.127583	-2.57425	NA	NA	NA	NA	NA	NA	-0.04663	0.091723	-0.50839	-0.0602	0.044609
FSR1	-0.14983	0.057346	-2.61275	-0.2159	0.120078	-1.798	-0.0019	0.019747	-0.0963	-0.30054	0.138369	-2.17201	-0.05086	0.082082	-0.6196	-0.04576	0.015905
EZH2	0.293772	0.156133	1.88155	0.79436	0.243012	3.26881	-0.03007	0.04916	-0.61166	0.123884	0.404373	0.306361	0.615257	0.155425	3.958546	0.134411	0.0393
F3	NA	NA	NA	-0.3284	0.132658	-2.47552	NA	NA	NA	-0.08026	0.491948	-0.16315	-0.20405	0.109227	-1.86809	-0.22911	0.055029

Table 13 Official Symbol	STNO ~Est	STNO ~SE	SINO~t	STOCK ~Est	STOCK ~SE	STOCK ~t	TRANS BIG~Est	TRANS BIG~SE	TRANS BIG~t	UCSF ~Est	UCSF ~SE	UCSF~t	UPP~Est	UPP~SE	UPP~t	fe	sefe
FGFR4	0.201581	0.15216	1.324796	-0.06118	0.174787	-0.35001	NA	NA	NA	0.149034	0.333338	0.447096	0.204299	0.102078	2.001401	0.075374	0.053791
FHIT	-0.16819	0.17858	-0.94184	-0.27141	0.367689	-0.73815	NA	NA	NA	0.225378	0.678656	0.332095	0.053025	0.245338	0.216132	-0.11401	0.082797
FN1	0.049279	0.11577	0.425659	0.185381	0.202933	0.913508	NA	NA	NA	0.13258	0.244458	0.542343	-0.15952	0.26761	-0.59607	0.070337	0.045477
FOXAI	NA	NA	NA	-0.18849	0.161048	-1.17039	NA	NA	NA	NA	NA	NA	0.139273	0.160139	0.869701	-0.07105	0.037194
FUS	NA	NA	NA	0.368833	0.437273	0.843485	NA	NA	NA	NA	NA	NA	-0.15247	0.345172	-0.44173	0.063142	0.111165
GADD45 A	0.390085	0.342821	1.137868	-0.24644	0.303688	-0.81148	NA	NA	NA	0.153778	0.296649	0.518384	-0.4297	0.20668	-2.07904	-0.18353	0.077839
GAPDH	NA	NA	NA	0.907441	0.296513	3.060375	NA	NA	NA	NA	NA	NA	0.493907	0.232859	2.121056	0.303991	0.05522
GATA3	-0.20281	0.068842	-2.94607	-0.25592	0.122639	-2.08677	NA	NA	NA	-0.2038	0.135112	-1.50836	0.052882	0.108852	0.485817	-0.12484	0.03218
GBP2	0.104968	0.124764	0.841332	-0.17667	0.338601	-0.52176	NA	NA	NA	0.161775	0.235299	0.687529	0.215873	0.198252	1.088882	0.030811	0.064103
GDF15	-0.02683	0.097032	-0.27646	0.251857	0.169158	1.488886	NA	NA	NA	0.462744	0.465751	0.993544	0.139286	0.128201	1.086466	0.095577	0.04245
GRB7	0.28938	0.08099	3.573025	0.464983	0.21274	2.185687	NA	NA	NA	0.492397	0.361768	1.361085	0.39613	0.142688	2.776197	0.203411	0.041043
GSTM1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	-0.18141	0.14912
GSTM2	NA	NA	NA	NA	NA	NA	NA	NA	NA	-0.12675	0.336406	-0.37676	NA	NA	NA	-0.15328	0.111442
GSTM3	-0.38478	0.15382	-2.50148	-0.43469	0.17404	-2.49766	0.035771	0.038412	0.931246	0.11963	0.323329	0.369995	-0.05308	0.123135	-0.43107	-0.06296	0.030752
HOXB13	NA	NA	NA	0.193	0.369898	0.521765	NA	NA	NA	0.540678	0.49567	1.090802	0.342881	0.212428	1.614105	0.227421	0.046188
OTUD4	0.372577	0.253393	1.470352	-0.19372	0.251083	-0.77155	NA	NA	NA	-0.97971	0.713147	-1.37378	0.231981	0.294286	0.788284	0.034041	0.081167
HSPA1A	NA	NA	NA	0.765501	0.440826	1.736515	NA	NA	NA	NA	NA	NA	0.722677	0.40563	1.781616	0.243271	0.092738
HSPA1B	0.033372	0.19398	0.172039	0.069621	0.248436	0.280237	NA	NA	NA	NA	NA	NA	0.187302	0.176407	1.061761	0.198207	0.083268
HSPA8	0.22166	0.199205	1.112723	0.32649	0.265007	1.232005	NA	NA	NA	-0.30224	0.477926	-0.63239	0.126525	0.166299	0.760828	0.218804	0.082393
IDH2	0.127942	0.255302	0.50114	0.574289	0.193387	2.969636	NA	NA	NA	-0.009	0.554612	-0.01623	0.659908	0.186426	3.539785	0.303626	0.056121
IGF1R	-0.16723	0.112062	-1.49233	-0.35887	0.141569	-2.53498	NA	NA	NA	0.277384	0.391147	0.709155	-0.04996	0.122321	-0.40843	-0.14872	0.0484
IGFBP7	0.121056	0.164973	0.733793	-0.55896	0.34775	-1.60736	NA	NA	NA	-0.50275	0.332753	-1.51087	-0.16594	0.185086	-0.89655	0.005398	0.068861
IL11	NA	NA	NA	0.086327	0.225669	0.38254	NA	NA	NA	NA	NA	NA	0.000507	0.151608	0.003346	-0.05199	0.075711
IL17RB	NA	NA	NA	-0.01403	0.212781	-0.06594	NA	NA	NA	NA	NA	NA	-0.1861	0.139748	-1.33168	-0.16557	0.069337
IL6ST	NA	NA	NA	-0.65682	0.195937	-3.35217	NA	NA	NA	-0.11749	0.19789	-0.5937	-0.26213	0.150485	-1.74192	-0.31568	0.063376
IL8	0.548269	0.238841	2.29554	0.382317	0.203112	1.882296	NA	NA	NA	-0.3673	0.460322	-0.79791	0.076262	0.135635	0.562257	0.136391	0.05243
INHBA	-0.12998	0.113709	-1.14313	0.249729	0.184419	1.354139	NA	NA	NA	0.094476	0.303634	0.311152	0.036575	0.162207	0.225485	0.026824	0.056655
IRF1	0.307333	0.166134	1.84991	0.248132	0.447433	0.554568	NA	NA	NA	0.380822	0.370842	1.026912	-0.01044	0.283877	-0.03676	0.082446	0.091982
ITGA4	0.02688	2341.09	1.15E-05	0.198854	0.302824	0.656665	NA	NA	NA	-0.54938	0.583992	-0.94073	-0.01192	0.18086	-0.0659	0.002027	0.059101
ITGA5	NA	NA	NA	0.025981	0.423908	0.061288	NA	NA	NA	NA	NA	NA	0.406364	0.36399	1.116415	0.431369	0.112958
ITGAV	0	0.216251	0	-0.403	0.45413	-0.88742	NA	NA	NA	-0.59197	0.499066	-1.18615	-0.24399	0.30418	-0.80213	-0.15415	0.089488
ITGB1	0.131284	0.165432	0.793583	0.195878	0.3192	0.613653	NA	NA	NA	0.430257	0.540622	0.795856	-0.18009	0.530248	-0.33962	0.026471	0.072949

WO 2011/063274

PCT/US2010/057490

Table 13 Official Symbol	STNO ~Est	STNO ~SE	SINO~t	STOCK ~Est	STOCK ~SE	STOCK ~t	TRANS BIG~Est	TRANS BIG~SE	TRANS BIG~t	UCSF ~Est	UCSF ~SE	UCSF~t	UPP~Est	UPP~SE	UPP~t	fe	sefe
ITGB4	0.100533	0.106548	0.943547	0.035914	0.241068	0.14898	NA	NA	NA	0.754519	0.285307	2.644586	0.075057	0.181963	0.412483	0.132678	0.060938
ITGB5	-0.19722	0.165947	-1.18843	-0.29946	0.281956	-1.06207	NA	NA	NA	-0.19391	0.378906	-0.51177	-0.21379	0.157719	-1.35549	-0.09296	0.063571
MK167	-0.07823	0.088982	-0.87915	0.96424	0.257398	3.746105	NA	NA	NA	-0.19193	0.462712	-0.4148	0.597931	0.152281	3.926498	0.183915	0.038442
KIAA119 9	NA	NA	NA	0.293164	0.194272	1.509039	NA	NA	NA	NA	NA	NA	0.070065	0.141965	0.493538	0.153718	0.066186
KPNA2	0.328818	0.112579	2.920776	0.857218	0.267225	3.207851	NA	NA	NA	0.32028	0.315031	1.016662	0.615022	0.206117	2.983849	0.374909	0.034897
LAMA3	-0.28334	0.120229	-2.3567	-0.42291	0.12869	-3.28625	NA	NA	NA	-0.14266	0.366741	-0.38899	-0.27285	0.091038	-2.99711	-0.26764	0.050305
LAMB3	NA	NA	NA	-0.15767	0.230936	-0.68274	NA	NA	NA	NA	NA	NA	-0.1353	0.168256	-0.8041	-0.00591	0.051501
LAPTM4 B	0.405684	0.113287	3.581029	0.28652	0.19422	1.475234	NA	NA	NA	NA	NA	NA	0.095487	0.136338	0.700367	0.270104	0.051492
LMNB1	NA	NA	NA	0.755925	0.25541	2.959653	NA	NA	NA	0.121429	0.384263	0.316005	0.805734	0.199208	4.044687	0.481816	0.073226
LRIG1	-0.31422	0.128149	-2.45197	-0.95351	0.258142	-3.69375	NA	NA	NA	NA	NA	NA	-0.05954	0.178366	-0.33383	-0.37679	0.062403
MTDH	0.242242	0.285145	0.84954	0.472647	0.340076	1.389828	0.052038	0.077589	0.670683	NA	NA	NA	0.45556	0.239663	1.900836	0.158361	0.059133
MCM2	0.008185	0.084857	0.096455	0.732134	0.216462	3.382275	NA	NA	NA	0.138969	0.340074	0.408643	0.602555	0.182898	3.294487	0.275153	0.05978
MELK	NA	NA	NA	0.749617	0.195032	3.843559	0.022669	0.036962	0.613293	NA	NA	NA	0.46629	0.128065	3.641042	0.132605	0.031744
MGMT	NA	NA	NA	0.377527	0.48364	0.780595	NA	NA	NA	0.368174	0.453282	0.812241	0.725329	0.346508	2.093253	0.085317	0.117786
MMP1	0.083945	0.055744	1.505895	0.28871	0.081435	3.545299	NA	NA	NA	0.150509	0.33411	0.450477	0.11015	0.051829	2.12525	0.151235	0.027295
MMP7	0.102783	0.072986	1.408258	-0.00343	0.153901	-0.0223	NA	NA	NA	0.166646	0.143301	1.162909	0.039637	0.10332	0.57721	0.08418	0.042799
MYBL2	0.399355	0.118084	3.381957	0.579872	0.192026	3.019758	NA	NA	NA	0.030169	0.282699	0.106717	0.445705	0.102011	4.369186	0.479924	0.057205
NAT1	-0.14333	0.060602	-2.36509	-0.26529	0.117131	-2.26487	NA	NA	NA	-0.1696	0.138069	-1.22836	-0.05668	0.076583	-0.7401	-0.14009	0.030446
PGF	-0.17016	0.153912	-1.10557	-0.08334	0.183966	-0.45304	0.095422	0.145828	0.654349	-1.00442	0.630097	-1.59407	0.038005	0.124883	0.304328	0.009034	0.063633
PGR	NA	NA	NA	-0.18022	0.108941	-1.65427	NA	NA	NA	0.451216	0.527475	0.855426	-0.01652	0.065638	-0.25164	-0.12464	0.038764
PRDX1	NA	NA	NA	1.52553	0.420489	3.62799	NA	NA	NA	0.358079	0.32938	1.08713	0.706059	0.303105	2.32942	0.347764	0.10081
PTEN	0	226.764	0	-0.26976	0.225651	-1.19546	NA	NA	NA	NA	NA	NA	0.110294	0.254356	0.433621	-0.15381	0.092467
RPL41	NA	NA	NA	-0.40807	0.786496	-0.51884	NA	NA	NA	NA	NA	NA	0.24408	0.604521	0.403758	-0.01769	0.094765
RPLP0	NA	NA	NA	0.018324	0.458438	0.039971	NA	NA	NA	NA	NA	NA	0.964584	0.554848	1.738465	0.108162	0.064823
RRM2	0.305217	0.104337	2.9253	0.926244	0.22125	4.186414	0.038487	0.042471	0.906208	-0.03281	0.279791	-0.11727	0.674794	0.149386	4.517117	0.159696	0.03419
RUNX1	-0.17832	0.165636	-1.07657	-0.39722	0.244634	-1.62372	NA	NA	NA	-0.58909	0.385997	-1.52616	-0.2142	0.105479	-2.03071	-0.07498	0.052758
S100A8	0.093477	0.04547	2.055818	0.164366	0.096581	1.701846	NA	NA	NA	0.123771	0.178963	0.691601	0.125784	0.065874	1.909478	0.106936	0.024582
S100A9	NA	NA	NA	0.15514	0.10905	1.42265	NA	NA	NA	NA	NA	NA	0.135096	0.074987	1.801592	0.112811	0.030203
S100B	0.136825	0.163838	0.835124	-0.11862	0.158461	-0.74859	-0.01591	0.034049	-0.46712	-0.05362	0.218098	-0.24584	-0.13315	0.115177	-1.15608	-0.01134	0.030069
S100P	0.19922	0.078236	2.546395	0.201435	0.097583	2.064251	NA	NA	NA	0.416003	0.200351	2.076371	0.174292	0.063687	2.736705	0.179884	0.028697
SEMA3F	0.023257	0.162267	0.143327	0.472655	0.292764	1.614457	NA	NA	NA	NA	NA	NA	0.545294	0.227357	2.398404	0.117569	0.092557

WO 2011/063274

PCT/US2010/057490

Table 13 Official Symbol	STNO ~Est	STNO ~SE	SINO~t	STOCK ~Est	STOCK ~SE	STOCK ~t	TRANS BIG~Est	TRANS BIG~SE	TRANS BIG~t	UCSF ~Est	UCSF ~SE	UCSF~t	UPP~Est	UPP~SE	UPP~t	fe	sefe
SKIL	NA	NA	NA	0.015831	0.262101	0.060402	NA	NA	NA	0.141704	0.348326	0.406814	0.179419	0.152532	1.176271	0.134826	0.065866
SKP2	NA	NA	NA	0.312141	0.339582	0.919192	NA	NA	NA	NA	NA	NA	0.482145	0.194873	2.47415	0.167902	0.091018
SNAI1	NA	NA	NA	0.152799	0.210056	0.72742	NA	NA	NA	NA	NA	NA	0.329059	0.159704	2.060431	0.140674	0.078745
SYK	0.21812	0.150626	1.44809	-0.06882	0.155403	-0.44285	NA	NA	NA	0.159029	0.431388	0.368645	0.066162	0.136668	0.484107	0.063381	0.072639
TAGLN	-0.00434	0.108525	-0.04003	-0.2578	0.197826	-1.30316	NA	NA	NA	NA	NA	NA	-0.06802	0.191196	-0.35574	0.032416	0.049944
TIFRC	0.406546	0.131339	3.095394	0.178145	0.153331	1.161833	-0.03263	0.051129	-0.63826	-0.22576	0.249301	-0.90558	0.545839	0.208978	2.611945	0.062825	0.038345
TGFB3	-0.07166	0.134442	-0.53298	-1.08462	0.322799	-3.36005	0.013681	0.046103	0.296755	-0.25719	0.253264	-1.01551	-0.49773	0.225603	-2.20621	-0.10353	0.03709
TNFRSF1 1B	0	0.08306	0	-0.10987	0.128194	-0.85708	NA	NA	NA	NA	NA	NA	-0.03866	0.087545	-0.44163	-0.09599	0.046815
VTN	-0.01674	0.109545	-0.15278	0.100648	0.186529	0.539584	0.226938	0.091337	2.484623	-0.22804	0.193542	-1.17822	0.167418	0.152274	1.099452	0.063022	0.050706
WISP1	0.03435	0.194412	0.176685	0.236658	0.340736	0.694549	-0.00282	0.068308	-0.04121	NA	NA	NA	-0.29716	0.212939	-1.39552	-0.05687	0.054306
WNT5A	0.121343	0.108022	1.123317	-0.01524	0.172902	-0.08815	NA	NA	NA	-0.96994	0.719267	-1.34851	-0.23507	0.152819	-1.5382	-0.12181	0.051129
C6orf66	NA	NA	NA	0.530409	0.355488	1.492059	NA	NA	NA	NA	NA	NA	-0.04983	0.251179	-0.19837	0.167784	0.123636
FOXO3A	NA	NA	NA	0.087341	0.128833	0.67794	NA	NA	NA	-0.03591	0.49687	-0.07227	-0.00291	0.074227	-0.03914	0.007101	0.054798
GPR30	NA	NA	NA	-0.36866	0.173755	-2.12169	NA	NA	NA	NA	NA	NA	-0.07779	0.125956	-0.61763	-0.02487	0.058543
KNTC2	NA	NA	NA	0.442783	0.170315	2.599789	-0.00276	0.041235	-0.06696	-0.02041	0.366566	-0.05568	0.347484	0.117596	2.954896	0.093083	0.034359

Table 14: Validation of Transferrin Receptor Group genes in SIB data sets.

	Genes						
Study data set	TFRC	ENO1	IDH2	ARF1	CLDN4	PRDX1	GBP1
EMC2~Est	NA	NA	NA	NA	NA	NA	NA
EMC2~SE	NA	NA	NA	NA	NA	NA	NA
EMC2~t	NA	NA	NA	NA	NA	NA	NA
JRH1~Est	-0.91825	NA	-0.0525	0.839013	-0.54144	NA	0.137268
JRH1~SE	0.636275	NA	0.232201	0.346692	0.470758	NA	0.159849
JRH1~t	-1.44317	NA	-0.22611	2.420053	-1.15014	NA	0.858735
JRH2~Est	0.162921	0.179739	0.151299	0.369609	0.33033	-0.41082	-0.07418
JRH2~SE	0.352486	0.312848	0.327466	0.40789	0.351865	0.47383	0.198642
JRII2~t	0.462206	0.574525	0.46203	0.906149	0.938798	-0.86703	-0.37345
MGH~Est	0.029015	NA	NA	2.03958	0.185116	NA	0.15434
MGH~SE	0.193689	NA	NA	0.804729	0.314723	NA	0.188083
MGH~t	0.149803	NA	NA	2.534493	0.588187	NA	0.820595
NCH~Est	0.056174	-0.01727	0.265828	-0.15337	-0.23129	0.253047	0.095457
NCH~SE	0.166875	0.097939	0.105592	0.204529	0.426627	0.182621	0.1323
NCII~t	0.336622	-0.17629	2.517501	-0.74984	-0.54213	1.38564	0.721522
NKI~Est	0.157216	0.3682	0.284862	0.944168	0.564756	0.231612	0.13712
NKI~SE	0.10845	0.094778	0.089145	0.204641	0.210595	0.161791	0.075391
NKI~t	1.449663	3.884888	3.195498	4.613777	2.681716	1.431551	1.818777
STNO~Est	0.406546	NA	0.127942	0	0.40922	NA	0.298139
STNO~SE	0.131339	NA	0.255302	0.107397	0.128817	NA	0.113901
STNO~t	3.095394	NA	0.50114	0	3.176755	NA	2.617528
STOCK~Est	0.178145	0.428884	0.574289	0.862387	1.20235	1.52553	0.068821
STOCK~SE	0.153331	0.194952	0.193387	0.279535	0.33711	0.420489	0.183692
STOCK~t	1.161833	2.199947	2.969636	3.085077	3.56664	3.62799	0.374652
TRANSBIG~Est	-0.03263	NA	NA	NA	0.03236	NA	NA
TRANSBIG~SE	0.051129	NA	NA	NA	0.053171	NA	NA
TRANSBIG~t	-0.63826	NA	NA	NA	0.608591	NA	NA
UCSF~Est	-0.22576	0.899319	-0.009	0.304097	0	0.358079	-0.43879
UCSF~SE	0.249301	0.369574	0.554612	0.58718	1.8541	0.32938	0.874728
UCSF~t	-0.90558	2.433394	-0.01623	0.517894	0	1.08713	-0.50163
UPP~Est	0.545839	0.288434	0.659908	0.751279	0.08503	0.706059	0.119778
UPP~SE	0.208978	0.179833	0.186426	0.361093	0.258939	0.303105	0.117879
UPP~t	2.611945	1.603899	3.539785	2.080569	0.328378	2.32942	1.01611
Fe	0.062825	0.233559	0.303626	0.281544	0.125868	0.347764	0.139381
Sefe	0.038345	0.058687	0.056121	0.07587	0.045235	0.10081	0.044464

Table 15: Validation of Stromal Group genes in SIB data sets.

Gene	CXCL14	TNFRSF11B	CXCL12	C10orf116	RUNX1	GSTM2	TGFB3	BCAR3	CAV1	DLC1	TNFRSF10B	F3	DICER1
EMC2~Est	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
EMC2~SE	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
EMC2~t	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
JRH1~Est	-0.23692	NA	-0.36476	-0.1418	-0.22834	NA	-1.0219	NA	-0.20701	0.13581	-0.09001	0.719395	NA
JRH1~SE	0.333761	NA	0.372499	0.261554	0.318666	NA	0.358953	NA	0.254401	0.37927	0.619057	0.524742	NA
JRH1~t	-0.70985	NA	-0.97921	-0.54216	-0.71656	NA	-2.84689	NA	-0.81372	0.358083	-0.1454	1.37095	NA
JRH2~Est	0.361375	-0.10399	-0.4566	0.036378	0.302803	NA	-0.39774	-0.29238	-0.19588	-0.4102	0.80742	-0.21237	-0.33943
JRH2~SE	0.159544	0.440721	0.219587	0.182183	0.420043	NA	0.470041	0.522706	0.289251	0.387258	0.544479	0.363632	0.39364
JRH2~t	2.265049	-0.23595	-2.07935	0.19968	0.720886	NA	-0.84619	-0.55935	-0.67721	-1.05923	1.482922	-0.58402	-0.8623
MGII~Est	NA	-1.15976	NA	NA	0.277566	NA	0.046498	-0.41595	-0.06896	-0.09793	0.159018	-0.00167	0.038811
MGH~SE	NA	0.400921	NA	NA	0.267511	NA	0.2296	0.216837	0.2269	0.247069	0.456205	0.448211	0.409835
MGH~t	NA	-2.89274	NA	NA	1.037587	NA	0.202518	-1.91825	-0.30391	-0.39638	0.348567	-0.00372	0.0947
NCH~Est	-0.06592	-0.2492	-0.08863	0.064337	0.124568	NA	-0.30473	0.072246	0.078825	-0.03473	-0.19927	-0.13187	0.086141
NCH~SE	0.093353	0.289075	0.138097	0.14087	0.088457	NA	0.247338	0.304443	0.340843	0.238947	0.160381	0.134218	0.143687
NCH~t	-0.70609	-0.86207	-0.64183	0.456713	1.408231	NA	-1.23202	0.237306	0.231265	-0.14533	-1.24248	-0.98248	0.599504
NKI~Est	-0.16877	-0.22072	-0.36944	-0.22589	-0.18878	-0.15655	-0.36531	-0.26067	-0.30885	-0.35001	0.053214	-0.29217	-0.46887
NKI~SE	0.054117	0.10171	0.138735	0.082836	0.138365	0.118111	0.09592	0.114992	0.133788	0.130472	0.164091	0.093753	0.150367
NKI~t	-3.11866	-2.17005	-2.66293	-2.72696	-1.36435	-1.32547	-3.80851	-2.26685	-2.30848	-2.68262	0.324294	-3.11637	-3.11814
STNO~Est	-0.20969	0	0.066487	-0.09621	-0.17832	NA	-0.07166	NA	0.135002	0.519601	-0.03773	NA	NA
STNO~SE	0.073458	0.08306	0.189775	0.085948	0.165636	NA	0.134442	NA	0.093948	0.221066	0.174479	NA	NA
STNO~t	-2.8546	0	0.350348	-1.11936	-1.07657	NA	-0.53298	NA	1.436991	2.350434	-0.21623	NA	NA
STOCK~Est	-0.14079	-0.10987	-0.65036	-0.34745	-0.39722	NA	-1.08462	-0.49692	-0.65852	-0.66099	-0.03558	-0.3284	-1.06544
STOCK~SE	0.096118	0.128194	0.168426	0.112777	0.244634	NA	0.322799	0.265837	0.275751	0.298518	0.198203	0.132658	0.322204
STOCK~t	-1.46476	-0.85708	-3.86137	-3.08087	-1.62372	NA	-3.36005	-1.86927	-2.38811	-2.21425	-0.1795	-2.47552	-3.30672

Gene	CXXL14	TNFRSF11B	CXCL12	C10orf116	RUNX1	GSTM2	TGFB3	BCAR3	CAV1	DLC1	TNFRSF10B	F3	DICER1
TRANSBIG~Est	NA	NA	NA	NA	NA	NA	0.013681	NA	NA	NA	NA	NA	N/A
TRANSBIG~SE	NA	NA	NA	NA	NA	NA	0.046103	NA	NA	NA	NA	NA	N/A
TRANSBIG~t	NA	NA	NA	NA	NA	NA	0.296755	NA	NA	NA	NA	NA	N/A
UCSF~Est	NA	NA	-0.05795	0.013111	-0.58909	-0.12675	-0.25719	NA	-0.54391	-0.31503	0.932141	-0.08026	0
UCSF~SE	NA	NA	0.270065	156.117	0.385997	0.336406	0.253264	NA	0.428883	0.345828	0.524911	0.491948	0.311799
UCSF~t	NA	NA	-0.21456	8.40E-05	-1.52616	-0.37676	-1.01551	NA	-1.2682	-0.91094	1.775808	-0.16315	0
UPP~Est	-0.1861	-0.03866	-0.35344	-0.00923	-0.2142	NA	-0.49773	-0.29435	-0.31503	-0.404	0.127348	-0.20405	0.208326
UPP~SE	0.08384	0.087545	0.150278	0.100902	0.105479	NA	0.225603	0.182614	0.150431	0.200673	0.157658	0.109227	0.307144
UPP~t	-2.21976	-0.44163	-2.35189	-0.09148	-2.03071	NA	-2.20621	-1.61186	-2.09415	-2.01324	0.807748	-1.86809	0.678268
Fe	-0.14219	-0.09599	-0.28998	-0.13	-0.07498	-0.15328	-0.10353	-0.28755	-0.11726	-0.19876	0.02034	-0.22911	-0.19602
Sefe	0.032611	0.046815	0.062826	0.042521	0.052758	0.111442	0.03709	0.080198	0.058989	0.076441	0.072745	0.055029	0.085879

Table 16: Genes that co-express with Prognostic genes in ER+ breast cancer tumors (Spearman corr. coef. ≥ 0.7)

Prognostic Gene	Table 16 Co-expressed Genes				
INHBA	AEBP1	CDH11	COL10A1	COL11A1	COL1A2
	COL5A1	COL5A2	COL8A2	ENTPD4	LOXL2
	LRRC15	MMP11	NOX4	PLAU	THBS2
	THY1	VCAN			
CAV1	ANK2	ANXA1	AQP1	C10orf56	CAV2
	CFH	COL14A1	CRYAB	CXCL12	DAB2
	DCN	ECM2	FHL1	FLRT2	GNG11
	GSN	IGF1	JAM2	LDB2	NDN
	NRN1	PCSK5	PLSCR4	PROS1	TGFBR2
NAT1	PSD3				
GSTM1	GSTM2				
GSTM2	GSTM1				
ITGA4	ARHGAP15	ARHGAP25	CCL5	CD3D	CD48
	CD53	CORO1A	EVI2B	FGL2	GIMAP4
	IRF8	LCK	PTPRC	TFEC	TRAC
	TRAF3IP3	TRBC1	EVI2A	FLI1	GPR65
	IL2RB	LCP2	LOC100133233	MNDA	PLAC8
	PLEK	TNFAIP8			
CCL19	ARHGAP15	ARHGAP25	CCL5	CCR2	CCR7
	CD2	CD247	CD3D	CD3E	CD48
	CD53	FLJ78302	GPR171	IL10RA	IL7R
	IRF8	LAMP3	LCK	LTB	PLAC8
	PRKCB1	PTPRC	PTPRCAP	SASH3	SPOCK2
	TRA@	TRBC1	TRD@	PPP1R16B	TRAC
CDH11	TAGLN	ADAM12	AEBP1	ANGPTL2	ASPN
	BGN	BICC1	C10orf56	C1R	C1S
	C20orf39	CALD1	COL10A1	COL11A1	COL1A1
	COL1A2	COL3A1	COL5A1	COL5A2	COL6A1
	COL6A2	COL6A3	COL8A2	COMP	COPZ2
	CRISPLD2	CTSK	DACT1	DCN	DPYSL3
	ECM2	EFEMP2	ENTPD4	FAP	FBLN1
	FBLN2	FBN1	FERMT2	FLRT2	FN1
	FSTL1	GAS1	GLT8D2	HEPH	HTRA1

Prognostic Gene	Table 16 Co-expressed Genes				
	ISLR	ITGBL1	JAM3	KDELC1	LAMA4
	LAMB1	LOC100133502	LOX	LOXL2	LRRC15
	LRRC17	LUM	MFAP2	MFAP5	MMP2
	MRC2	MXRA5	MXRA8	MYL9	NDN
	NID1	NID2	NINJ2	NOX4	OLFML2B
	OMD	PALLD	PCOLCE	PDGFRA	PDGFRB
	PDGFRL	POSTN	PRKCDBP	PRKD1	PTRF
	RARRES2	RCN3	SERPINF1	SERPINH1	SFRP4
	SNAI2	SPARC	SPOCK1	SPON1	SRPX2
	SSPN	TCF4	THBS2	THY1	TNFAIP6
	VCAN	WWTR1	ZEB1	ZFPM2	INHBA
	PLS3	SEC23A	WISP1		
TAGLN	CDH11	ADAM12	AEBP1	ANGPTL2	ASPN
	BGN	BICC1	C10orf56	C1R	C1S
	C20orf39	CALD1	COL10A1	COL11A1	COL1A1
	COL1A2	COL3A1	COL5A1	COL5A2	COL6A1
	COL6A2	COL6A3	COL8A2	COMP	COPZ2
	CRISPLD2	CTSK	DACT1	DCN	DPYSL3
	ECM2	EFEMP2	ENTPD4	FAP	FBLN1
	FBLN2	FBN1	FERMT2	FLRT2	FN1
	FSTL1	GAS1	GLT8D2	HEPH	HTRA1
	ISLR	ITGBL1	JAM3	KDELC1	LAMA4
	LAMB1	LOC100133502	LOX	LOXL2	LRRC15
	LRRC17	LUM	MFAP2	MFAP5	MMP2
	MRC2	MXRA5	MXRA8	MYL9	NDN
	NID1	NID2	NINJ2	NOX4	OLFML2B
	OMD	PALLD	PCOLCE	PDGFRA	PDGFRB
	PDGFRL	POSTN	PRKCDBP	PRKD1	PTRF
	RARRES2	RCN3	SERPINF1	SERPINH1	SFRP4
	SNAI2	SPARC	SPOCK1	SPON1	SRPX2
	SSPN	TCF4	THBS2	THY1	TNFAIP6
	VCAN	WWTR1	ZEB1	ZFPM2	ACTA2
	CNN1	DZIP1	EMILIN1		
ENO1	ATP5J2	C10orf10	CLDN15	CNGB1	DET1
	EIF3CL	HS2ST1	IGHG4	KIAA0195	KIR2DS5
	PARP6	PRH1	RAD1	RIN3	RPL10
	SGCG	SLC16A2	SLC9A3R1	SYNPO2L	THBS1

Prognostic Gene	Table 16 Co-expressed Genes				
	ZNF230				
IDH2	AEBP1	HIST1H2BN	PCDHAC1		
ARF1	CRIM1				
DICER1	ADM	LOC100133583			
AKT3	AKAP12	ECM2	FERMT2	FLRT2	JAM3
	LOC100133502	PROS1	TCF4	WWTR1	ZEB1
CXCL12	ANXA1	C1R	C1S	CAV1	DCN
	FLRT2	SRPX			
CYR61	CTGF				
IGFBP7	VIM				
KIAA1199	COL11A1	PLAU			
SPC25	ASPM	BUB1	BUB1B	CCNA2	CCNE2
	CDC2	CDC25C	CENPA	CEP55	FANCI
	GIN51	HJURP	KIAA0101	KIF11	KIF14
	KIF15	KIF18A	KIF20A	KIF4A	MAD2L1
	MELK	NCAPG	NEK2	NUSAP1	PRC1
	STIL	ZWINT			
WISP1	CDH11	COL5A2			

Table 17: Genes that co-express with Prognostic Genes in ER- breast cancer tumors (Spearman corr. coef. ≥ 0.7)

Prognostic Gene	Table 17 Co-expressed Genes				
IRF1	APOL6	CXCL10	GABBR1	GBP1	HCP5
	HLA-E	HLA-F	HLA-G	HLA-J	INDO
	PSMB8	PSMB9	STAT1	TAP1	UBD
	UBE2L6	WARS	APOBEC3F	APOBEC3G	APOL1
	APOL3	ARHGAP25	BTN3A1	BTN3A2	BTN3A3
	C1QB	CCL5	CD2	CD38	CD40
	CD53	CD74	CD86	CSF2RB	CTSS
	CYBB	FGL2	GIMAP5	GZMA	hCG_1998957
	HCLS1	HLA-C	HLA-DMA	HLA-DMB	HLA-DPA1
	HLA-DQB1	HLA-DQB2	HLA-DRA	HLA-DRB1	HLA-DRB2
	HLA-DRB3	HLA-DRB4	HLA-DRB5	HLA-DRB6	IL10RA
	IL2RB	LAP3	LAPTM5	LOC100133484	LOC100133583
	LOC100133661	LOC100133811	LOC730415	NKG7	PLEK
	PSMB10	PTPRC	RNASE2	SLAMF8	TFEC
	TNFRSF1B	TRA@	TRAC	TRAJ17	TRAV20
	ZNF749				
CDH11	ADAM12	AEBP1	ANGPTL2	ASPN	CFH
	CFHR1	COL10A1	COL11A1	COL1A1	COL1A2
	COL3A1	COL5A1	COL5A2	COL6A3	CRISPLD2
	CTSK	DACT1	DCN	FAP	FBN1
	FN1	HTRA1	LOX	LRRC15	LUM
	NID2	PCOLCE	PDGFRB	POSTN	SERPINF1
	SPARC	THBS2	THY1	VCAN	DAB2
	GLT8D2	ITGB5	JAM3	LOC100133502	MMP2
	PRSS23	TIMP3	ZEB1		
CCL19	ITGA4	ADAM28	AIF1	APOBEC3F	APOBEC3G
	APOL3	ARHGAP15	ARHGAP25	CASP1	CCDC69
	CCR2	CCR7	CD2	CD247	CD27
	CD37	CD3D	CD3G	CD48	CD52
	CD53	CD74	CD86	CD8A	CLEC4A
	CORO1A	CTSS	CXCL13	DOCK10	EVI2A
	EVI2B	FGL2	FLJ78302 (CCR2)	FYB	GIMAP4
	GIMAP5	GIMAP6	GMFG	GPR171	GPR18
	GPR65	GZMA	GZMB	GZMK	hCG_1998957

Prognostic Gene	Table 17				
	Co-expressed Genes				
	HCLS1	HLA-DMA	HLA-DMB	HLA-DPA1	HLA-DQA1
	HLA-DQA2	HLA-DQB1	HLA-DQB2	HLA-DRB1	HLA-DRB2
	HLA-DRB3	HLA-DRB4	HLA-DRB5	HLA-E	IGHM
	IGSF6	IL10RA	IL2RG	IL7R	IRF8
	KLRB1	KLRK1	LAPTM5	LAT2	LCK
	LCP2	LOC100133484	LOC100133583	LOC100133661	LOC100133811
	LOC730415	LPXN	LRMP	LST1	LTB
	LY96	LYZ	MFNG	MNDA	MS4A4A
	NCKAP1L	PLAC8	PLEK	PRKCB1	PSCDBP
	PTPRC	PTPRCAP	RAC2	RNASE2	RNASE6
	SAMHD1	SAMSN1	SASH3	SELL	SELPLG
	SLA	SLAMF1	SLC7A7	SP140	SRGN
	TCL1A	TFEC	TNFAIP8	TNFRSF1B	TRA@
	TRAC	TRAJ17	TRAT1	TRAV20	TRBC1
	TYROBP	ZNF749	ITM2A	LTB	P2RY13
	PRKCB1	PTPRCAP	SELL	TRBC1	
ITGA4	CCL19	ADAM28	AIF1	APOBEC3F	APOBEC3G
	APOL3	ARHGAP15	ARHGAP25	CASP1	CCDC69
	CCR2	CCR7	CD2	CD247	CD27
	CD37	CD3D	CD3G	CD48	CD52
	CD53	CD74	CD86	CD8A	CLEC4A
	CORO1A	CTSS	CXCL13	DOCK10	EVI2A
	EVI2B	FGL2	FLJ78302 (CCR2)	FYB	GIMAP4
	GIMAP5	GIMAP6	GMFG	GPR171	GPR18
	GPR65	GZMA	GZMB	GZMK	hCG_1998957
	HCLS1	HLA-DMA	HLA-DMB	HLA-DPA1	HLA-DQA1
	HLA-DQA2	HLA-DQB1	HLA-DQB2	HLA-DRB1	HLA-DRB2
	HLA-DRB3	HLA-DRB4	HLA-DRB5	HLA-E	IGHM
	IGSF6	IL10RA	IL2RG	IL7R	IRF8
	KLRB1	KLRK1	LAPTM5	LAT2	LCK
	LCP2	LOC100133484	LOC100133583	LOC100133661	LOC100133811
	LOC730415	LPXN	LRMP	LST1	LTB
	LY96	LYZ	MFNG	MNDA	MS4A4A
	NCKAP1L	PLAC8	PLEK	PRKCB1	PSCDBP
	PTPRC	PTPRCAP	RAC2	RNASE2	RNASE6
	SAMHD1	SAMSN1	SASH3	SELL	SELPLG
	SLA	SLAMF1	SLC7A7	SP140	SRGN

Prognostic Gene	Table 17				
	Co-expressed Genes				
	TCL1A	TFEC	TNFAIP8	TNFRSF1B	TRA@
	TRAC	TRAJ17	TRAT1	TRAV20	TRBC1
	TYROBP	ZNF749	MARCH1	C17orf60	CSF1R
	FLJ1	FLJ78302	FYN	IKZF1	INPP5D
	NCF4	NR3C1	P2RY13	PLXNC1	PSCD4
	PTPN22	SERPINB9	SLCO2B1	VAMP3	WIPF1
IDH2	AEBP1	DSG3	HIST1H2BN	PCDHAC1	
ARF1	FABP5L2	FLNB	IL1RN	PAX6	
DICER1	ARS2	IGHA1	VDAC3		
TFRC	RGS20				
ADAM17	TFDP3	GPR107			
CAV1	CAV2	CXCL12	IGF1		
CYR61	CTGF				
ESR1	CBLN1	SLC45A2			
GSTM1	GSTM2				
GSTM2	GSTM1				
IL11	FAM135A				
IL6ST	P2RY5				
IGFBP7	SPARCL1	TMEM204			
INHBA	COL10A1	FN1	SULF1		
SPC25	KIF4A	KIF20A	NCAPG		
TAGLN	ACTA2	MYL9	NNMT	PTRF	
TGFB3	GALNT10	HTRA1	LIMA1		
TNFRSF10B	BIN3				
FOXA1	CLCA2	TFAP2B	AGR2	MLPH	SPDEF
CXCL12	DCN	CAV1	IGF1	CFH	
GBP2	APOL1	APOL3	CD2	CTSS	CXCL9
	CXCR6	GBP1	GZMA	HLA-DMA	HLA-DMB
	IL2RB	PTPRC	TRBC1		

**Table 18: Genes that co-express with Prognostic Genes in all breast cancer tumors
(Spearman corr. coef. ≥ 0.7)**

Prognostic Gene	Table 18 Co-expressed Genes				
S100A8	S100A9				
S100A9	S100A8				
MKI67	BIRC5	KIF20A	MCM10		
MTDH	ARMC1	AZIN1	ENY2	MTERFD1	POLR2K
	PTDSS1	RAD54B	SLC25A32	TMEM70	UBE2V2
GSTM1	GSTM2				
GSTM2	GSTM1				
CXCL12	AKAP12	DCN	F13A1		
TGFB3	C10orf56	JAM3			
TAGLN	ACTA2	CALD1	COPZ2	FERMT2	HEPH
	MYL9	NNMT	PTRF	TPM2	
PGF	ALMS1	ATP8B1	CEP27	DBT	FAM128B
	FBXW12	FGFR1	FLJ12151	FLJ42627	GTF2H3
	HCG2P7	KIAA0894	KLHL24	LOC152719	PDE4C
	PODNL1	POLR1B	PRDX2	PRR11	RIOK3
	RP5-886K2.1	SLC35E1	SPN	USP34	ZC3H7B
	ZNF160	ZNF611			
CCL19	ARHGAP15	ARHGAP25	CCL5	CCR2	CCR7
	CD2	CD37	CD3D	CD48	CD52
	CSF2RB	FLJ78302	GIMAP5	GIMAP6	GPR171
	GZMK	IGHM	IRF8	LCK	LTB
	PLAC8	PRKCB1	PTGDS	PTPRC	PTPRCAP
	SASH3	TNFRSF1B	TRA@	TRAC	TRAJ17
	TRAV20	TRBC1			
IRF1	ITGA4	MARCH1	AIF1	APOBEC3F	APOBEC3G
	APOL1	APOL3	ARHGAP15	ARHGAP25	BTN3A2
	BTN3A3	CASP1	CCL4	CCL5	CD2
	CD37	CD3D	CD48	CD53	CD69
	CD8A	CORO1A	CSF2RB	CST7	CYBB
	EVI2A	EVI2B	FGL2	FLJ1	GBP1
	GIMAP4	GIMAP5	GIMAP6	GMFG	GPR65
	GZMA	GZMK	hCG_1998957	HCLS1	HLA-DMA
	HLA-DMB	HLA-DPA1	HLA-DQB1	HLA-DQB2	HLA-DRA
	HLA-DRB1	HLA-DRB2	HLA-DRB3	HLA-DRB4	HLA-DRB5
	HLA-E	HLA-F	IGSF6	IL10RA	IL2RB

Prognostic Gene	Table 18 Co-expressed Genes				
	IRF8	KLRK1	LCK	LCP2	LOC100133583
	LOC100133661	LOC100133811	LST1	LTB	LY86
	MFNG	MNDA	NKG7	PLEK	PRKCB1
	PSCDBP	PSMB10	PSMB8	PSMB9	PTPRC
	PTPRCAP	RAC2	RNASE2	RNASE6	SAMSN1
	SLA	SRGN	TAP1	TFEC	TNFAIP3
	TNFRSF1B	TRA@	TRAC	TRAJ17	TRAV20
	TRBC1	TRIM22	ZNF749		
ITGA4	IRF1	MARCH1	AIF1	APOBEC3F	APOBEC3G
	APOL1	APOL3	ARHGAP15	ARHGAP25	BTN3A2
	BTN3A3	CASP1	CCL4	CCL5	CD2
	CD37	CD3D	CD48	CD53	CD69
	CD8A	CORO1A	CSF2RB	CST7	CYBB
	EVI2A	EVI2B	FGL2	FLI1	GBP1
	GIMAP4	GIMAP5	GIMAP6	GMFG	GPR65
	GZMA	GZMK	hCG_1998957	HCLS1	HLA-DMA
	HLA-DMB	HLA-DPA1	HLA-DQB1	HLA-DQB2	HLA-DRA
	HLA-DRB1	HLA-DRB2	HLA-DRB3	HLA-DRB4	HLA-DRB5
	HLA-E	HLA-F	IGSF6	IL10RA	IL2RB
	IRF8	KLRK1	LCK	LCP2	LOC100133583
	LOC100133661	LOC100133811	LST1	LTB	LY86
	MFNG	MNDA	NKG7	PLEK	PRKCB1
	PSCDBP	PSMB10	PSMB8	PSMB9	PTPRC
	PTPRCAP	RAC2	RNASE2	RNASE6	SAMSN1
	SLA	SRGN	TAP1	TFEC	TNFAIP3
	TNFRSF1B	TRA@	TRAC	TRAJ17	TRAV20
	TRBC1	TRIM22	ZNF749	CTSS	
SPC25	ASPM	ATAD2	AURKB	BUB1B	C12orf48
	CCNA2	CCNE1	CCNE2	CDC2	CDC45L
	CDC6	CDCA3	CDCA8	CDKN3	CENPE
	CENPF	CENPN	CEP55	CHEK1	CKS1B
	CKS2	DBF4	DEPDC1	DLG7	DNAJC9
	DONSON	E2F8	ECT2	ERCC6L	FAM64A
	FBXO5	FEN1	FOXO1	GINS1	GTSE1
	H2AFZ	HJURP	HMMR	KIF11	KIF14
	KIF15	KIF18A	KIF20A	KIF23	KIF2C
	KIF4A	KIFC1	MAD2L1	MCM10	MCM6

Prognostic Gene	Table 18 Co-expressed Genes				
	NCAPG	NEK2	NUSAP1	OIP5	PBK
	PLK4	PRC1	PTTG1	RACGAP1	RAD51AP1
	RFC4	SMC2	STIL	STMN1	TACC3
	TOP2A	TRIP13	TTK	TYMS	UBE2C
	UBE2S	AURKA	BIRC5	BUB1	CCNB1
	CENPA	KPNA2	LMNB1	MCM2	MELK
	NDC80	TPX2			
AURKA	ASPM	ATAD2	AURKB	BUB1B	C12orf48
	CCNA2	CCNE1	CCNE2	CDC2	CDC45L
	CDC6	CDCA3	CDCA8	CDKN3	CENPE
	CENPF	CENPN	CEP55	CHEK1	CKS1B
	CKS2	DBF4	DEPDC1	DLG7	DNAJC9
	DONSON	E2F8	ECT2	ERCC6L	FAM64A
	FBXO5	FEN1	FOXM1	GIN51	GTSE1
	H2AFZ	HJURP	HMMR	KIF11	KIF14
	KIF15	KIF18A	KIF20A	KIF23	KIF2C
	KIF4A	KIFC1	MAD2L1	MCM10	MCM6
	NCAPG	NEK2	NUSAP1	OIP5	PBK
	PLK4	PRC1	PTTG1	RACGAP1	RAD51AP1
	RFC4	SMC2	STIL	STMN1	TACC3
	TOP2A	TRIP13	TTK	TYMS	UBE2C
	UBE2S	SPC25	BIRC5	BUB1	CCNB1
	CENPA	KPNA2	LMNB1	MCM2	MELK
	NDC80	TPX2	PSMA7	CSE1L	
BIRC5	ASPM	ATAD2	AURKB	BUB1B	C12orf48
	CCNA2	CCNE1	CCNE2	CDC2	CDC45L
	CDC6	CDCA3	CDCA8	CDKN3	CENPE
	CENPF	CENPN	CEP55	CHEK1	CKS1B
	CKS2	DBF4	DEPDC1	DLG7	DNAJC9
	DONSON	E2F8	ECT2	ERCC6L	FAM64A
	FBXO5	FEN1	FOXM1	GIN51	GTSE1
	H2AFZ	HJURP	HMMR	KIF11	KIF14
	KIF15	KIF18A	KIF20A	KIF23	KIF2C
	KIF4A	KIFC1	MAD2L1	MCM10	MCM6
	NCAPG	NEK2	NUSAP1	OIP5	PBK
	PLK4	PRC1	PTTG1	RACGAP1	RAD51AP1
	RFC4	SMC2	STIL	STMN1	TACC3

Prognostic Gene	Table 18 Co-expressed Genes				
	TOP2A	TRIP13	TTK	TYMS	UBE2C
	UBE2S	AURKA	SPC25	BUB1	CCNB1
	CENPA	KPNA2	LMNB1	MCM2	MELK
	NDC80	TPX2	MKI67		
BUB1	ASPM	ATAD2	AURKB	BUB1B	C12orf48
	CCNA2	CCNE1	CCNE2	CDC2	CDC45L
	CDC6	CDCA3	CDCA8	CDKN3	CENPE
	CENPF	CENPN	CEP55	CHEK1	CKS1B
	CKS2	DBF4	DEPDC1	DLG7	DNAJC9
	DONSON	E2F8	ECT2	ERCC6L	FAM64A
	FBXO5	FEN1	FOXO1	GIN5	GTSE1
	H2AFZ	HJURP	HMMR	KIF11	KIF14
	KIF15	KIF18A	KIF20A	KIF23	KIF2C
	KIF4A	KIFC1	MAD2L1	MCM10	MCM6
	NCAPG	NEK2	NUSAP1	OIP5	PBK
	PLK4	PRC1	PTTG1	RACGAP1	RAD51AP1
	RFC4	SMC2	STIL	STMN1	TACC3
	TOP2A	TRIP13	TTK	TYMS	UBE2C
	UBE2S	AURKA	BIRC5	SPC25	CCNB1
	CENPA	KPNA2	LMNB1	MCM2	MELK
	NDC80	TPX2			
CCNB1	ASPM	ATAD2	AURKB	BUB1B	C12orf48
	CCNA2	CCNE1	CCNE2	CDC2	CDC45L
	CDC6	CDCA3	CDCA8	CDKN3	CENPE
	CENPF	CENPN	CEP55	CHEK1	CKS1B
	CKS2	DBF4	DEPDC1	DLG7	DNAJC9
	DONSON	E2F8	ECT2	ERCC6L	FAM64A
	FBXO5	FEN1	FOXO1	GIN5	GTSE1
	H2AFZ	HJURP	HMMR	KIF11	KIF14
	KIF15	KIF18A	KIF20A	KIF23	KIF2C
	KIF4A	KIFC1	MAD2L1	MCM10	MCM6
	NCAPG	NEK2	NUSAP1	OIP5	PBK
	PLK4	PRC1	PTTG1	RACGAP1	RAD51AP1
	RFC4	SMC2	STIL	STMN1	TACC3
	TOP2A	TRIP13	TTK	TYMS	UBE2C
	UBE2S	AURKA	BIRC5	BUB1	SPC25
	CENPA	KPNA2	LMNB1	MCM2	MELK

Prognostic Gene	Table 18 Co-expressed Genes				
	NDC80	TPX2			
CENPA	ASPM	ATAD2	AURKB	BUB1B	C12orf48
	CCNA2	CCNE1	CCNE2	CDC2	CDC45L
	CDC6	CDCA3	CDCA8	CDKN3	CENPE
	CENPF	CENPN	CEP55	CHEK1	CKS1B
	CKS2	DBF4	DEPDC1	DLG7	DNAJC9
	DONSON	E2F8	ECT2	ERCC6L	FAM64A
	FBXO5	FEN1	FOXM1	GIN51	GTSE1
	H2AFZ	HJURP	HMMR	KIF11	KIF14
	KIF15	KIF18A	KIF20A	KIF23	KIF2C
	KIF4A	KIFC1	MAD2L1	MCM10	MCM6
	NCAPG	NEK2	NUSAP1	OIP5	PBK
	PLK4	PRC1	PTTG1	RACGAP1	RAD51AP1
	RFC4	SMC2	STIL	STMN1	TACC3
	TOP2A	TRIP13	TTK	TYMS	UBE2C
	UBE2S	AURKA	BIRC5	BUB1	CCNB1
	SPC25	KPNA2	LMNB1	MCM2	MELK
	NDC80	TPX2			
KPNA2	ASPM	ATAD2	AURKB	BUB1B	C12orf48
	CCNA2	CCNE1	CCNE2	CDC2	CDC45L
	CDC6	CDCA3	CDCA8	CDKN3	CENPE
	CENPF	CENPN	CEP55	CHEK1	CKS1B
	CKS2	DBF4	DEPDC1	DLG7	DNAJC9
	DONSON	E2F8	ECT2	ERCC6L	FAM64A
	FBXO5	FEN1	FOXM1	GIN51	GTSE1
	H2AFZ	HJURP	HMMR	KIF11	KIF14
	KIF15	KIF18A	KIF20A	KIF23	KIF2C
	KIF4A	KIFC1	MAD2L1	MCM10	MCM6
	NCAPG	NEK2	NUSAP1	OIP5	PBK
	PLK4	PRC1	PTTG1	RACGAP1	RAD51AP1
	RFC4	SMC2	STIL	STMN1	TACC3
	TOP2A	TRIP13	TTK	TYMS	UBE2C
	UBE2S	AURKA	BIRC5	BUB1	CCNB1
	CENPA	SPC25	LMNB1	MCM2	MELK
	NDC80	TPX2	NOL11	PSMD12	
LMNB1	ASPM	ATAD2	AURKB	BUB1B	C12orf48
	CCNA2	CCNE1	CCNE2	CDC2	CDC45L

Prognostic Gene	Table 18 Co-expressed Genes				
	CDC6	CDCA3	CDCA8	CDKN3	CENPE
	CENPF	CENPN	CEP55	CHEK1	CKS1B
	CKS2	DBF4	DEPDC1	DLG7	DNAJC9
	DONSON	E2F8	ECT2	ERCC6L	FAM64A
	FBXO5	FEN1	FOXM1	GINS1	GTSE1
	H2AFZ	HJURP	HMMR	KIF11	KIF14
	KIF15	KIF18A	KIF20A	KIF23	KIF2C
	KIF4A	KIFC1	MAD2L1	MCM10	MCM6
	NCAPG	NEK2	NUSAP1	OIP5	PBK
	PLK4	PRC1	PTTG1	RACGAP1	RAD51AP1
	RFC4	SMC2	STIL	STMN1	TACC3
	TOP2A	TRIP13	TTK	TYMS	UBE2C
	UBE2S	AURKA	BIRC5	BUB1	CCNB1
	CENPA	KPNA2	SPC25	MCM2	MELK
	NDC80	TPX2			
MCM2	ASPM	ATAD2	AURKB	BUB1B	C12orf48
	CCNA2	CCNE1	CCNE2	CDC2	CDC45L
	CDC6	CDCA3	CDCA8	CDKN3	CENPE
	CENPF	CENPN	CEP55	CHEK1	CKS1B
	CKS2	DBF4	DEPDC1	DLG7	DNAJC9
	DONSON	E2F8	ECT2	ERCC6L	FAM64A
	FBXO5	FEN1	FOXM1	GINS1	GTSE1
	H2AFZ	HJURP	HMMR	KIF11	KIF14
	KIF15	KIF18A	KIF20A	KIF23	KIF2C
	KIF4A	KIFC1	MAD2L1	MCM10	MCM6
	NCAPG	NEK2	NUSAP1	OIP5	PBK
	PLK4	PRC1	PTTG1	RACGAP1	RAD51AP1
	RFC4	SMC2	STIL	STMN1	TACC3
	TOP2A	TRIP13	TTK	TYMS	UBE2C
	UBE2S	AURKA	BIRC5	BUB1	CCNB1
	CENPA	KPNA2	LMNB1	SPC25	MELK
	NDC80	TPX2			
MELK	ASPM	ATAD2	AURKB	BUB1B	C12orf48
	CCNA2	CCNE1	CCNE2	CDC2	CDC45L
	CDC6	CDCA3	CDCA8	CDKN3	CENPE
	CENPF	CENPN	CEP55	CHEK1	CKS1B
	CKS2	DBF4	DEPDC1	DLG7	DNAJC9

Prognostic Gene	Table 18 Co-expressed Genes				
	DONSON	E2F8	ECT2	ERCC6L	FAM64A
	FBXO5	FEN1	FOXMI	GINSI	GTSE1
	H2AFZ	HJURP	HMMR	KIF11	KIF14
	KIF15	KIF18A	KIF20A	KIF23	KIF2C
	KIF4A	KIFC1	MAD2L1	MCM10	MCM6
	NCAPG	NEK2	NUSAP1	OIP5	PBK
	PLK4	PRC1	PTTG1	RACGAP1	RAD51AP1
	RFC4	SMC2	STIL	STMN1	TACC3
	TOP2A	TRIP13	TTK	TYMS	UBE2C
	UBE2S	AURKA	BIRC5	BUB1	CCNB1
	CENPA	KPNA2	LMNB1	MCM2	SPC25
	NDC80	TPX2			
NDC80	ASPM	ATAD2	AURKB	BUB1B	C12orf48
	CCNA2	CCNE1	CCNE2	CDC2	CDC45L
	CDC6	CDCA3	CDCA8	CDKN3	CENPE
	CENPF	CENPN	CEP55	CHEK1	CKS1B
	CKS2	DBF4	DEPDC1	DLG7	DNAJC9
	DONSON	E2F8	ECT2	ERCC6L	FAM64A
	FBXO5	FEN1	FOXMI	GINSI	GTSE1
	H2AFZ	HJURP	HMMR	KIF11	KIF14
	KIF15	KIF18A	KIF20A	KIF23	KIF2C
	KIF4A	KIFC1	MAD2L1	MCM10	MCM6
	NCAPG	NEK2	NUSAP1	OIP5	PBK
	PLK4	PRC1	PTTG1	RACGAP1	RAD51AP1
	RFC4	SMC2	STIL	STMN1	TACC3
	TOP2A	TRIP13	TTK	TYMS	UBE2C
	UBE2S	AURKA	BIRC5	BUB1	CCNB1
	CENPA	KPNA2	LMNB1	MCM2	MELK
	SPC25	TPX2			
TPX2	ASPM	ATAD2	AURKB	BUB1B	C12orf48
	CCNA2	CCNE1	CCNE2	CDC2	CDC45L
	CDC6	CDCA3	CDCA8	CDKN3	CENPE
	CENPF	CENPN	CEP55	CHEK1	CKS1B
	CKS2	DBF4	DEPDC1	DLG7	DNAJC9
	DONSON	E2F8	ECT2	ERCC6L	FAM64A
	FBXO5	FEN1	FOXMI	GINSI	GTSE1
	H2AFZ	HJURP	HMMR	KIF11	KIF14

Prognostic Gene	Table 18 Co-expressed Genes				
	KIF15	KIF18A	KIF20A	KIF23	KIF2C
	KIF4A	KIFC1	MAD2L1	MCM10	MCM6
	NCAPG	NEK2	NUSAP1	OIP5	PBK
	PIK4	PRC1	PTTG1	RACGAP1	RAD51AP1
	RFC4	SMC2	STIL	STMN1	TACC3
	TOP2A	TRIP13	TTK	TYMS	UBE2C
	UBE2S	AURKA	BIRC5	BUB1	CCNB1
	CENPA	KPNA2	LMNB1	MCM2	MELK
	NDC80	SPC25			
CDH11	INHBA	WISP1	COL1A1	COL1A2	FN1
	ADAM12	AEBP1	ANGPTL2	ASPN	BGN
	BNC2	C1QTNF3	COL10A1	COL11A1	COL3A1
	COL5A1	COL5A2	COL5A3	COL6A3	COMP
	CRISPLD2	CTSK	DACT1	DCN	DKK3
	DPYSL3	EFEMP2	EMILIN1	FAP	FBN1
	FSTL1	GLT8D2	HEG1	HTRA1	ITGBL1
	JAM3	KIAA1462	LAMA4	LOX	LOXL1
	LRP1	LRRC15	LRRC17	LRRC32	LUM
	MFAP5	MICAL2	MMP11	MMP2	MXRA5
	MXRA8	NID2	NOX4	OLFML2B	PCOLCE
	PDGFRB	PLAU	POSTN	SERPINF1	SPARC
	SPOCK1	SPON1	SRPX2	SULF1	TCF4
	THBS2	THY1	VCAN	ZEB1	
INHBA	CDH11	WISP1	COL1A1	COL1A2	FN1
	ADAM12	AEBP1	ANGPTL2	ASPN	BGN
	BNC2	C1QTNF3	COL10A1	COL11A1	COL3A1
	COL5A1	COL5A2	COL5A3	COL6A3	COMP
	CRISPLD2	CTSK	DACT1	DCN	DKK3
	DPYSL3	EFEMP2	EMILIN1	FAP	FBN1
	FSTL1	GLT8D2	HEG1	HTRA1	ITGBL1
	JAM3	KIAA1462	LAMA4	LOX	LOXL1
	LRP1	LRRC15	LRRC17	LRRC32	LUM
	MFAP5	MICAL2	MMP11	MMP2	MXRA5
	MXRA8	NID2	NOX4	OLFML2B	PCOLCE
	PDGFRB	PLAU	POSTN	SERPINF1	SPARC
	SPOCK1	SPON1	SRPX2	SULF1	TCF4
	THBS2	THY1	VCAN	ZEB1	

Prognostic Gene	Table 18				
	Co-expressed Genes				
WISP1	INHBA	CDH11	COL1A1	COL1A2	FN1
	ADAM12	AEBP1	ANGPTL2	ASPN	BGN
	BNC2	C1QTNF3	COL10A1	COL11A1	COL3A1
	COL5A1	COL5A2	COL5A3	COL6A3	COMP
	CRISPLD2	CTSK	DACT1	DCN	DKK3
	DPYSL3	EFEMP2	EMILIN1	FAP	FBN1
	FSTL1	GLT8D2	HEG1	HTRA1	ITGBL1
	JAM3	KIAA1462	LAMA4	LOX	LOXL1
	LRP1	LRRC15	LRRC17	LRRC32	LUM
	MFAP5	MICAL2	MMP11	MMP2	MXRA5
	MXRA8	NID2	NOX4	OLFML2B	PCOLCE
	PDGFRB	PLAU	POSTN	SERPINF1	SPARC
	SPOCK1	SPON1	SRPX2	SULF1	TCF4
	THBS2	THY1	VCAN	ZEB1	
COL1A1	INHBA	WISP1	CDH11	COL1A2	FN1
	ADAM12	AEBP1	ANGPTL2	ASPN	BGN
	BNC2	C1QTNF3	COL10A1	COL11A1	COL3A1
	COL5A1	COL5A2	COL5A3	COL6A3	COMP
	CRISPLD2	CTSK	DACT1	DCN	DKK3
	DPYSL3	EFEMP2	EMILIN1	FAP	FBN1
	FSTL1	GLT8D2	HEG1	HTRA1	ITGBL1
	JAM3	KIAA1462	LAMA4	LOX	LOXL1
	LRP1	LRRC15	LRRC17	LRRC32	LUM
	MFAP5	MICAL2	MMP11	MMP2	MXRA5
	MXRA8	NID2	NOX4	OLFML2B	PCOLCE
	PDGFRB	PLAU	POSTN	SERPINF1	SPARC
	SPOCK1	SPON1	SRPX2	SULF1	TCF4
	THBS2	THY1	VCAN	ZEB1	
COL1A2	INHBA	WISP1	COL1A1	CDH11	FN1
	ADAM12	AEBP1	ANGPTL2	ASPN	BGN
	BNC2	C1QTNF3	COL10A1	COL11A1	COL3A1
	COL5A1	COL5A2	COL5A3	COL6A3	COMP
	CRISPLD2	CTSK	DACT1	DCN	DKK3
	DPYSL3	EFEMP2	EMILIN1	FAP	FBN1
	FSTL1	GLT8D2	HEG1	HTRA1	ITGBL1
	JAM3	KIAA1462	LAMA4	LOX	LOXL1
	LRP1	LRRC15	LRRC17	LRRC32	LUM

Prognostic Gene	Table 18				
	Co-expressed Genes				
	MFAP5	MICAL2	MMP11	MMP2	MXRA5
	MXRA8	NID2	NOX4	OLFML2B	PCOLCE
	PDGFRB	PLAU	POSTN	SERPINF1	SPARC
	SPOCK1	SPON1	SRPX2	SULF1	TCF4
	THBS2	THY1	VCAN	ZEB1	
FN1	INHBA	WISP1	COL1A1	COL1A2	CDH11
	ADAM12	AEBP1	ANGPTL2	ASPN	BGN
	BNC2	C1QTNF3	COL10A1	COL11A1	COL3A1
	COL5A1	COL5A2	COL5A3	COL6A3	COMP
	CRISPLD2	CTSK	DACT1	DCN	DKK3
	DPYSL3	EFEMP2	EMILIN1	FAP	FBN1
	FSTL1	GLT8D2	HEG1	HTRA1	ITGBL1
	JAM3	KIAA1462	LAMA4	LOX	LOXL1
	LRP1	LRRC15	LRRC17	LRRC32	LUM
	MFAP5	MICAL2	MMP11	MMP2	MXRA5
	MXRA8	NID2	NOX4	OLFML2B	PCOLCE
	PDGFRB	PLAU	POSTN	SERPINF1	SPARC
	SPOCK1	SPON1	SRPX2	SULF1	TCF4
	THBS2	THY1	VCAN	ZEB1	

SEQUENCE LISTING IN ELECTRONIC FORM

This description contains a sequence listing in electronic form in ASCII text format. A copy of the sequence listing in electronic form is available from the Canadian Intellectual Property Office.

That which is claimed is:

1. A method for determining likelihood of good prognosis of a human patient diagnosed with a breast cancer comprising:
 - (a) quantitatively measuring a level of an RNA transcript of IL6ST in a tissue sample obtained from a breast cancer tumor of the patient;
 - (b) normalizing the level of the RNA transcript of IL6ST to obtain a normalized IL6ST expression level;
 - (c) comparing the normalized IL6ST expression level to normalized IL6ST expression level data obtained from a breast cancer reference set; and
 - (d) determining a likelihood of good prognosis for the patient based on the comparison of the normalized IL6ST expression level to the IL6ST expression level data obtained from the breast cancer reference set, wherein increases in normalized IL6ST expression level positively correlate with increased likelihood of good prognosis, wherein the good prognosis is a reduced likelihood of recurrence or metastasis or an increased likelihood of overall survival.
2. The method of claim 1, further comprising generating a report based on the normalized IL6ST expression level.
3. The method of claim 1 or 2, wherein the tissue sample is a fixed paraffin-embedded tissue sample.
4. The method of claim 1, 2 or 3, wherein the level of an RNA transcript of IL6ST is measured using a PCR-based method.
5. The method of any one of claims 1 to 4, further comprising measuring a level of an RNA transcript of P2RY5 in the tissue sample.
6. The method of any one of claims 1 to 5, wherein the tissue sample is obtained by core biopsy or fine needle aspiration.

7. The method of any one of claims 1 to 6, wherein the breast cancer is estrogen receptor (ER) positive breast cancer.
8. The method of any one of claims 1 to 7, wherein the level of the RNA transcript of IL6ST is a crossing point (C_p) value and the normalized IL6ST expression level is a normalized C_p value.
9. The method of any one of claims 1 to 7, wherein the level of the RNA transcript of IL6ST is a threshold cycle (C_t) value and the normalized IL6ST expression level is a normalized C_t value.
10. The method of any one of claims 1 to 9, wherein the good prognosis is the reduced likelihood of recurrence or metastasis.
11. A method for predicting likelihood of good prognosis of a human patient diagnosed with breast cancer comprising:
 - (a) extracting RNA from a fixed, wax-embedded breast cancer tissue sample obtained from the patient;
 - (b) reverse transcribing an RNA transcript of IL6ST to produce a cDNA of IL6ST;
 - (c) amplifying the cDNA of IL6ST to produce an amplicon of the RNA transcript of IL6ST;
 - (d) quantitatively assaying a level of the amplicon of the RNA transcript of IL6ST;
 - (e) normalizing the level of the amplicon of the RNA transcript of IL6ST to provide a normalized IL6ST amplicon level;
 - (f) comparing the normalized IL6ST amplicon level to normalized IL6ST amplicon level data obtained from a breast cancer reference set; and
 - (g) determining a likelihood of good prognosis for the patient based on the comparison of the normalized IL6ST amplicon level to the IL6ST amplicon expression level data obtained from the breast cancer reference set, wherein increases in normalized IL6ST amplicon level positively correlate with increased likelihood of good prognosis, wherein the good prognosis is a reduced likelihood of recurrence or metastasis or an increased likelihood of overall survival.
12. The method of claim 11, further comprising generating a report based on the normalized IL6ST amplicon level.

13. The method of claim 11 or 12, wherein the tissue sample is fixed paraffin-embedded tissue.
14. The method of claim 11, 12 or 13, wherein the cDNA of IL6ST is amplified using a PCR-based method.
15. The method of any one of claims 11 to 14, further comprising amplifying an RNA transcript of P2RY5 in the breast cancer tissue sample.
16. The method of any one of claims 11 to 15, wherein the breast cancer is estrogen receptor (ER) positive breast cancer.
17. The method of any one of claims 11 to 15, wherein the level of the amplicon of the RNA transcript of IL6ST is a crossing point (C_p) value and the normalized IL6ST amplicon level is a normalized C_p value.
18. The method of any one of claims 11 to 15, wherein the level of the amplicon of the RNA transcript of IL6ST is a threshold cycle (C_t) value and the normalized IL6ST amplicon level is a normalized C_t value.
19. The method of any one of claims 11 to 18, wherein the good prognosis is the reduced likelihood of recurrence or metastasis.